

RESEARCH ARTICLE

Open Access

Definition of novel cell envelope associated proteins in Triton X-114 extracts of *Mycobacterium tuberculosis* H37Rv

Hiwa Målen^{1,2}, Sharad Pathak^{1,2}, Tina Søfteland¹, Gustavo A de Souza¹ and Harald G Wiker^{*1,2}

Abstract

Background: Membrane- and membrane-associated proteins are important for the pathogenicity of bacteria. We have analysed the content of these proteins in virulent *Mycobacterium tuberculosis* H37Rv using Triton X-114 detergent-phase separation for extraction of lipophilic proteins, followed by their identification with high resolution mass spectrometry.

Results: In total, 1417 different proteins were identified. *In silico* analysis of the identified proteins revealed that 248 proteins had at least one predicted trans-membrane region. Also, 64 of the identified proteins were predicted lipoproteins, and 54 proteins were predicted as outer membrane proteins. Three-hundred-and-ninety-five of the observed proteins, including 91 integral membrane proteins were described for the first time. Comparison of abundance levels of the identified proteins was performed using the exponentially modified protein abundance index (emPAI) which takes into account the number of the observable peptides to the number of experimentally observed peptide ions for a given protein. The outcome showed that among the membrane- and membrane-associated proteins several proteins are present with high relative abundance. Further, a close examination of the lipoprotein LpqG (Rv3623) which is only detected in the membrane fractions of *M. tuberculosis* but not in *M. bovis*, revealed that the homologous gene in *M. bovis* lack the signal peptide and lipobox motif, suggesting impaired export to the membrane.

Conclusions: Altogether, we have identified a substantial proportion of membrane- and membrane-associated proteins of *M. tuberculosis* H37Rv, compared the relative abundance of the identified proteins and also revealed subtle differences between the different members of the *M. tuberculosis* complex.

Background

Tuberculosis is an airborne infection caused by *Mycobacterium tuberculosis*. It is estimated that one-third of the world's population is latently infected with *M. tuberculosis*, and that each year about three million people die of this disease. The emergence of drug-resistant strains is further escalating the threat to public health (WHO, 2003). In spite of global research efforts, mechanisms underlying pathogenesis, virulence and persistence of *M. tuberculosis* infection remain poorly understood [1].

M. tuberculosis is a facultative intracellular pathogen that resides within the host macrophages [2-4]. When *M. tuberculosis* invades host cells, the interface between the

host and the pathogen includes membrane- and surface proteins likely to be involved in intracellular multiplication and the bacterial response to host microbicidal processes [4]. Recently, the cell wall of *M. tuberculosis* was reported to possess a true outer membrane adding more complexity with regard to bacterial-host interactions and also important information relevant for susceptibility to anti-mycobacterial therapies [5-7]. Revealing the composition of the membrane proteome will have an impact on the design and interpretation of experiments aimed at elucidating the translocation pathways for nutrients, lipids, proteins, and anti-mycobacterial drugs across the cell envelope. According to bioinformatic predictions, 597 genes (~15%) of the *M. tuberculosis* H37Rv genome [8,9], could encode proteins having between 1 and 18 trans-membrane α -helical domains (TMH), which interact with the hydrophobic core of the lipid bilayer. The confirma-

* Correspondence: harald.wiker@gades.uib.no

¹ Section for Microbiology and Immunology, the Gade Institute, University of Bergen, Bergen, Norway

Full list of author information is available at the end of the article

tion of the expression of these genes at the protein level may lead to new therapeutic targets, new vaccine candidates and better serodiagnostic methods.

Membrane proteins resolve poorly in two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and proteomic profiling of mycobacterial membrane proteins remains a major challenge. Their limited solubility in aqueous buffer systems and their relatively low abundance in a background of highly abundant cytoplasmic proteins have yet to be overcome. Several studies have reported extraction of membrane- and membrane-associated proteins using centrifugation to obtain purified cell wall and cell membrane fractions for analysis by sodium-dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE) in combination with liquid chromatography tandem mass spectrometry (LC-MS/MS) [10-13]. Common for these studies is pre-isolation of the membrane and cell wall of the bacteria, and application of different washing techniques prior to protein extraction by detergents. In this study, we separated hydrophobic membrane- and membrane-associated proteins directly from sonicated *M. tuberculosis* H37Rv using phase separation with Triton X-114. The efficacy of this method was shown with *Mycobacterium bovis* BCG in a previous work [14].

Comparison of expressed levels of the identified proteins was performed using the emPAI [15,16]. This approach relates the number of experimentally observed peptide ions in a given protein to the number of theoretically observable peptides. Our results show that among the membrane- and membrane-associated proteins several proteins are present in high relative abundance. Using bioinformatic analysis, we also found that the gene sequence encoding Rv3623 which is annotated as a potential lipoprotein in both *M. tuberculosis* and *M. bovis*, is shorter in *M. bovis* and have lost the N-terminal signal peptide and lipobox that mediate the prelipoprotein translocation and its subsequent lipidation that retains it to the membrane.

Results

Identification of Triton X-114 extracted proteins

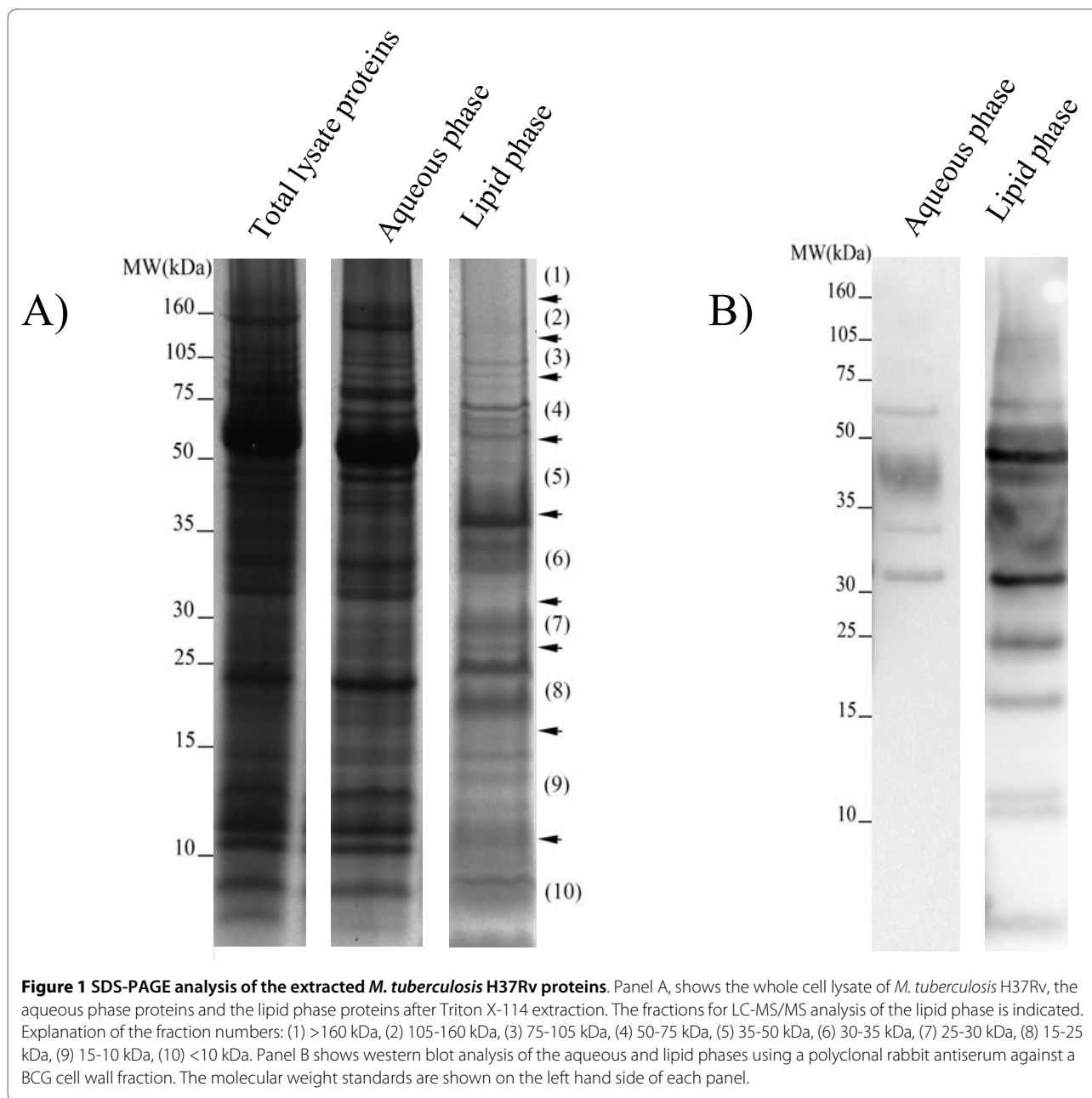
The aim of this study was to enrich and perform a comprehensive proteomic analysis of membrane- and membrane-associated proteins of the virulent reference strain *M. tuberculosis* H37Rv. For this purpose, the hydrophobic proteins were enriched by lysing whole bacilli followed by phase separation with the Triton X-114 detergent. After phase separation, the proteins in the lipid phase were precipitated by acetone and separated by SDS-PAGE. As shown in Figure 1 panel A, the lipid phase was quite complex, but appeared to be enriched for certain proteins as compared to the unfractionated crude lysate. In a parallel experiment, and to validate that the protein content in

the lipid and aqueous phases were different, proteins from both phases were separated and transferred to nitrocellulose membranes which were developed with polyclonal antibodies against a cell wall fraction of *M. bovis* BCG (Figure 1, panel B). Notably, Figure 1 not only demonstrates that the protein content of the aqueous phase and the lipid phase was different, but also clearly shows that the lipid phase was indeed enriched for cell wall proteins. In order to identify the proteins of the Triton X-114 detergent fraction, the protein mixture was separated with SDS-PAGE (Figure 1A), run in duplicate and cut into ten pieces each (twenty fractions in total) and subjected to in-gel digestion by trypsin. The resulting peptides were eluted and analysed by high accuracy mass spectrometry. Additional file 1, Figure S1 illustrates the sequence obtained for ion m/z 1210.62 which was identified by Mascot as peptide CGSPAWDLPTVFGPIAITYNIK from protein Rv0932c with a Mascot score of 79. Such fragmentation data contain a very good coverage of the expected γ - and b-series daughter ions plus the presence of other ions which indicates the correct MS/MS assignment such as two highly abundant γ -ions of proline ($\gamma_{19^{++}}$ and γ_{14}). This is very typical for peptides containing proline.

In total, 1417 proteins extracted with Triton X-114 were identified from the *M. tuberculosis* H37Rv strain out of which 395 are described for the first time. The complete lists of proteins with identified peptides are provided as additional data files (Additional file 2, Table S1 and Additional file 3, Table S2). Information about the criteria for protein identifications, such as number of peptides matching each protein, scores, identification threshold and peak lists are given in Additional file 4, Table S3. Identified proteins were categorized according to functional classification (Table 1). An overview of the number of observed proteins belonging to major groups based on physicochemical properties is shown in Figure 2. These groups are described below:

Membrane proteins

According to TMHMM version 2.0, a bioinformatic algorithm that predicts transmembrane regions in the primary amino acid sequences, 597 genes (~15%) of the *M. tuberculosis* H37Rv genome were found to possess between 1 and 18 TMHs. Each α -helix consists of 10 to 15 amino acid residues which interact with the hydrophobic core of the lipid bilayer. The proteins identified in this study were analysed by the TMHMM algorithm and 248 were predicted to have 1 or more TMH regions (Figure 3), among those, 90 represented novel identifications (Additional file 2, Table S1). Proteins with one TMH were only considered as possible membrane proteins if the TMH region was positioned beyond the first 70 N-terminal amino acids. This was done to avoid confusion with potential secreted proteins.



Lipoproteins

Lipoproteins represent a subgroup of exported proteins characterized by the presence of a lipobox. The lipobox motif is located in the distal C-terminal part of the N-terminal signal peptide [17]. This motif is a recognition signal for lipid modification on the conserved and essential cysteine residue. Precursor lipoproteins are mainly translocated in a Sec-dependent manner across the plasma membrane and are subsequently modified [18]. The proteins identified in this study were analysed by the lipop algorithm <http://www.cbs.dtu.dk/services/LipoP/>, and 63 were predicted as potential lipoproteins (Additional file 2, Table S1) based on the presence of a cleavable signal pep-

ptide and a lipobox motif. Eight lipoproteins are described for the first time. In sum the findings comprises over 56% of all predicted lipoproteins in the genome.

Outer membrane proteins

Outer membrane proteins (OMPs) are a class of proteins residing in the outer membrane of bacterial cells. Identification of OMPs is important as they are exposed on the bacterial surface and so are accessible drug targets. Recently, Song and colleagues analysed the genome of *M. tuberculosis* and predicted 144 proteins as potential OMPs based on the amphiphilicity of the β -strand regions, absence of hydrophobic α -helices and the presence of a signal peptide [19]. In our study, we observed 54 (37.5%)

Table 1: Functional classification of the identified *M. tuberculosis* H37Rv proteins.

Functional group ^a	Functional group no.	Total protein number ^b	Number of observed proteins ^c
Virulence, detoxification, adaptation	0	212	44 (21%)
Lipid metabolism	1	237	84 (35%)
Information pathways	2	232	98 (42%)
Cell wall and cell processes	3	751	313 (42%)
Stable RNAs	4	50	0 (0%)
Insertion sequences and phages	5	147	0 (0%)
PE/PPE	6	168	14 (8%)
Intermediary metabolism and respiration	7	898	412 (46%)
Unknown	8	15	0 (0%)
Regulatory proteins	9	194	54 (28%)
Conserved hypotheticals	10	895	299 (33%)
Conserved hypotheticals with an orthologue in <i>M. bovis</i>	16	262	52 (20%)

^a The functional groups were taken from the Tuberculist database, publicly available at <http://genolist.pasteur.fr/TubercuList/>.

^b Total number of proteins in each group predicted in the genome.

^c Number of proteins identified and the ratio compared to the total number of proteins assigned to each functional group.

of these proteins, and 9 of them have not been described in previous proteomic works (Additional file 2, Table S1).

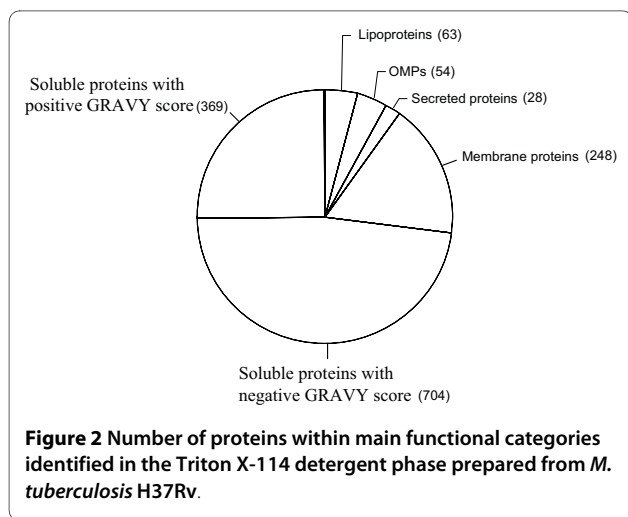
GRAVY

The 'grand mean of hydropathicity' (GRAVY) score is the average hydropathy score for a protein. According to Kyte and Doolittle, integral membrane proteins have a higher GRAVY score than soluble proteins. A positive score > 0.4 suggests increased probability for membrane association; the higher the score, the greater the probability [20]. GRAVY scores were calculated for all the identified proteins using the PROTPARAM tool <http://us.expasy.org/tools/protparam.html>. Three-hundred and sixty nine

proteins without a TMH region had positive GRAVY scores (Additional file 3, Table S2). A substantial proportion of the detected proteins lacked a predicted retention region and had a negative GRAVY score, suggesting that they were soluble proteins. However, it is possible that at least some of them might be functionally membrane-associated through formation of protein complexes with membrane-anchored proteins. In a previous study we showed that several hydrophilic proteins are retained in the lipophilic membrane fraction due to interaction with hydrophobic proteins [21-23].

Relative abundance index

To estimate the relative abundance of the observed proteins, we used the emPAI algorithm, which is based on the calculation of identified peptides per protein and normalized by the theoretical number of peptides for the same protein (PAI). The outcome of the emPAI analysis is given for a selection of membrane proteins and lipoproteins with the highest values in Table 2 and 3, respectively. At the top of the membrane protein list is the possible proline rich antigen *pra* (Rv1078), with 5.66 mol %. This is a small protein with 25 kDa, and has 2 TMHs. When digested with trypsin, it constitutes 6 observable tryptic peptides, where 5 of them were identified. This protein has also been observed in *M. bovis* [14,24]. The membrane proteins Rv1078 and Rv1489 are the most abundant ones, but with no annotated biological functions. In the lipoprotein list only the first three proteins



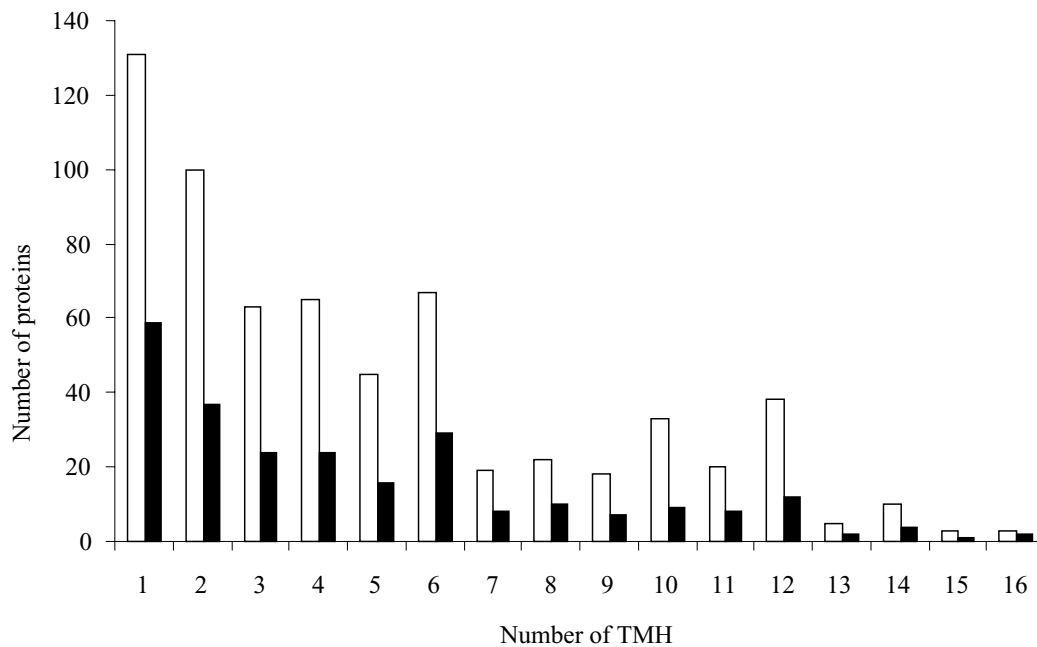


Figure 3 Number of TMH regions in membrane proteins identified in the Triton X-114 lipid phase fraction of *M. tuberculosis* H37Rv. Number of identified proteins compared to the total number of predicted proteins is given. The white bars represent the total number of predicted membrane proteins in the genome based on the TMHMM algorithm version 2.0, while the black bars represent those observed in the present study.

are assigned functions, while the 7 others have unknown biological functions.

Gene sequence analysis

An in-depth analysis of our data indicated that 2 proteins were consistently identified in *M. tuberculosis* and not in *M. bovis* and these were: possible glutamine-transport transmembrane protein ATP binding cassette (ABC) transporter (Rv0072) and possible conserved lipoprotein LpqG (Rv3623). The DNA sequences encoding the two proteins including 100 base pairs (bp) up-stream were obtained from Tuberculist for *M. tuberculosis* and BoviList for *M. bovis* and the sequences were aligned using the Blast 2 algorithm. No differences were found for Rv0072 which had 100% similarity between *M. bovis* and *M. tuberculosis*. However, the conserved lipoprotein LpqG (Rv3623) appeared to be 207 bp shorter in *M. bovis* compared to *M. tuberculosis* with a difference in the N-terminal end of the gene. Consequently, the protein product was 69 amino acids shorter. When the primary sequence of the protein product was analysed by the LipoP algorithm, it appeared that the lipobox was missing in *M. bovis* and the protein cannot be considered as a lipoprotein (Figure 4).

Discussion

Due to the anticipated role of membrane- and membrane-associated proteins of *M. tuberculosis* in virulence,

it is important to characterize these proteins. Therefore, the aim of the present study was to perform a proteomic analysis of these proteins from the virulent reference strain *M. tuberculosis* H37Rv in extracts obtained with the non-ionic detergent Triton X-114. The proteins from the lipid phase of the detergent, which was enriched for membrane proteins as validated by immuno-blotting (Figure 1, panel B), were precipitated, separated, and identified by high accuracy mass spectrometry. In total, 1417 proteins were identified and analysis of the primary amino acid sequences by bioinformatic tools revealed that 31% of the proteins were membrane- or membrane-associated. The list included more than 50% of all predicted integral membrane proteins in the genome.

These results show a significant improvement compared to the two studies of mycobacterial plasma membrane proteins by Gu *et al.* [25] and Xiong *et al.*, [26]. In these studies, membrane proteins were enriched by differential centrifugation and alkaline treatment of crude membranes with sodium carbonate and urea and separated by SDS-PAGE followed by protein identification with LC-MS/MS. The study by Gu *et al.* revealed 739 *M. tuberculosis* H37Rv proteins including 85 membrane proteins (11.5%), while Xiong *et al.* identified 349 proteins, of which 100 were predicted membrane proteins (28.7%). The low percentage of integral plasma membrane proteins among the proteins identified in these studies was

Table 2: List of the 14 most frequently observed membrane proteins.

<i>Sanger ID</i>	<i>Gene name</i>	<i>Protein identity</i>	<i>No. of TMH^a</i>	<i>No. of observed peptides^b</i>	<i>emPAI (Mol %)^c</i>	<i>References</i>
Rv1078	<i>pra</i>	Possible proline rich antigen	2	5	5.66	[14,24]
Rv1489	-	Conserved hypothetical protein	2	5	1.30	[26]
Rv1306	<i>atpF</i>	Possible ATP synthase b chain	1	7	0.36	[14,24-26]
Rv2563	-	Possible glutamine-transport transmembrane protein	4	13	0.35	[14,25,26,32]
Rv1234	-	Possible transmembrane protein	2	7	0.26	[25,26]
Rv0072	-	Possible glutamine-transport transmembrane protein	4	11	0.23	[25,26]
Rv0479c	-	Possible conserved membrane protein	1	11	0.23	[24-26]
Rv2969c	-	Possible conserved membrane or secreted protein	1	11	0.19	[14,24-26,40]
Rv2200c	<i>ctaC</i>	Possible transmembrane cytochrome C oxidase	3	13	0.17	[14,24-26,32]
Rv2195	<i>qcrA</i>	Possible rieske iron-sulfur protein	3	15	0.16	[14,24-26,40,54]
Rv1223	<i>htrA</i>	Possible serine protease	1	19	0.15	[24,26,54]
Rv1822	-	Phosphatidylglycerophosphate synthase	4	5	0.14	[14]
Rv2721c	-	Possible conserved transmembrane protein	2	12	0.13	[14,24-26,32]
Rv3273	-	Possible transmembrane carbonic anhydrase	10	11	0.11	[24-26,54]

^a Number of TMH regions predicted by TMHMM version 2.0 publicly available at <http://www.cbs.dtu.dk/services/TMHMM/>.

^b Number of observed unique peptides from each protein.

^c Relative protein abundance provided in mol % concentration.

probably based in the membrane enrichment methods. We reduced the soluble protein contamination by phase separation of whole bacterial sonicates, and also applied state-of-the-art mass spectrometry analysis for identification of peptides.

More than 50% of all predicted lipoproteins in the genome were found. These are proteins translocated across the cell membrane and retained in the cell envelope by post-translational lipid modification. They are functionally diverse, and are suggested to be involved in host-pathogen interactions [27,28]. They are also of interest with respect to development of serodiagnostic tests for tuberculosis due to their strong immunogenicity [29,30].

We also found 37% of all predicted OMPs [19], which is an essential group of proteins involved in import of nutrients, secretion processes and host-pathogen interactions

in gram-negative bacteria [31], and this is also likely to be of great importance in mycobacteria because it is now firmly established that they have a true outer membrane [5-7].

Even though a considerable number of observed proteins were predicted as integral membrane- or membrane-associated proteins, a substantial proportion of the detected proteins lacked a predicted retention region. For those proteins we measured the GRAVY score which express the total hydrophobicity of a protein as an indicator for membrane association. However, this is just a measure of increased probability for membrane association based on the fact that most integral membrane proteins have a positive GRAVY value. If a protein has a positive value, even though it lacks a retention signal, it is probably associated with the membrane. On the other hand, some of the hydrophilic proteins with a negative

Table 3: List of the 10 most frequently observed lipoproteins.

Sanger ID	Gene name	Protein identity	No. of observed peptides ^a	emPAI (Mol %) ^b	References
Rv0432	<i>sodC</i>	Possible periplasmic superoxide dismutase	6	2.36	[14,24-26,40]
Rv3763	<i>lpqH</i>	19 kda lipoprotein antigen precursor	3	1.05	[14,24-26,40,55]
Rv0932c	<i>pstS2</i>	Periplasmic phosphate-binding lipoprotein	9	0.39	[14,24-26,45]
Rv2945c	<i>lppX</i>	Possible conserved lipoprotein	6	0.21	[14,24-26,45,54]
Rv1411c	<i>lprG</i>	Possible conserved lipoprotein	6	0.19	[14,24-26,40,54]
Rv0928	<i>pstS3</i>	Periplasmic phosphate-binding lipoprotein	7	0.16	[14,24,26,45]
Rv0583c	<i>lpqN</i>	Probable conserved lipoprotein	3	0.12	[14,25,26,32]
Rv1275	<i>lprC</i>	Possible lipoprotein	6	0.12	[14,24,25,54]
Rv2116	<i>lppK</i>	Probable conserved lipoprotein	4	0.12	[14,25,26]
Rv3623	<i>lpqG</i>	Possible conserved lipoprotein	7	0.11	[25,26,40]

^a Number of observed unique peptides from each protein.

^b Relative protein abundance provided in mol % concentration.

GRAVY value might still be retained in the membrane through formation of protein complexes with membrane-anchored proteins [21-23]. Several proteins in this group are encoded in operons of well known integral enzyme complexes [14].

Using state-of-the-art proteomic instrumentation and techniques, subtle details could be revealed at the individual protein level, such as experimental identification of signal peptide cleavage sites of predicted secreted proteins [32], or confirmation of the start codon, or identification of peptides from regions predicted to be non-coding thus indicating a more up-stream start codon [33,34], or even detection of novel genes [35]. Therefore, the data obtained in this study was examined both in detail and in the context of what have been reported in the literature. To examine the amounts of individual proteins in the membrane fraction we applied the emPAI algorithm. The emPAI calculation gives an approximate

estimate of the abundance of a certain protein, and it calculates the protein concentration (in mol %) [15,16]. An advantage of this method is that it gives a more realistic picture of the protein profile compared to the mRNA levels, which could be difficult to relate to the actual protein amount. The membrane proteins (14 proteins) and the lipoproteins (10 proteins), with the highest relative abundance values are listed in Tables 2 and 3, respectively.

Interestingly, two of the proteins (Rv0072 and Rv2563) among those with the highest relative abundance values were "possible glutamine-transport transmembrane ABC transporter protein", with sequence motifs that belong to the ABC transport system. Glutamine is a major cell wall component of pathogenic mycobacteria only [36]. Its production is mainly catalyzed extracellularly by glutamine synthetase *GlnA1* (Rv2220) [37]. Tullius et al., 2003 showed that a *M. tuberculosis glnA1* mutant requires a relatively high level of exogenous L-glutamine for growth

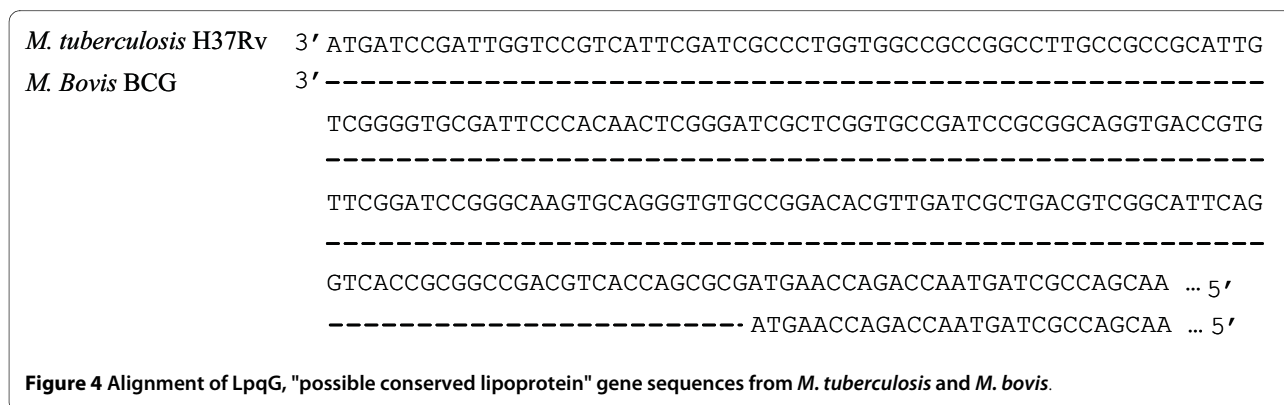


Figure 4 Alignment of LpqG, "possible conserved lipoprotein" gene sequences from *M. tuberculosis* and *M. bovis*.

in vitro, and the mutant was attenuated for intracellular growth in differentiated THP-1 cells, and it was also avirulent in infected guinea pigs [38]. Identification of two related proteins among the most abundant membrane proteins in *M. tuberculosis*, underlines the importance of production and transport of glutamine for the pathogen and its virulence.

The Rv0072 protein is only reported in studies conducted on *M. tuberculosis* [25,26] and not on *M. bovis* BCG (11, 17). It was identified by 11 different peptides giving sequence coverage of 44%, and the high emPAI value observed for this membrane protein suggests that it is abundantly present in the membrane of the virulent *M. tuberculosis* H37Rv strain. The open reading frames and sequences 100 bp up-stream to the start codon from *M. tuberculosis* H37Rv and *M. bovis* BCG 1173P2 and AF2122/97 were aligned, but the DNA sequences were identical and could not explain why Rv0072 has not been observed in *M. bovis* (data not shown).

Among the 10 most abundant lipoproteins 7 were not assigned any biological function, reflecting a fundamental lack of knowledge about these proteins. A careful examination revealed that the possible conserved lipoprotein LpqG (Rv3623) lies on the border of region of difference 9 (RD9) [39]. RD9 is deleted from all *M. bovis* lineages and consequently this protein has only been identified in proteomic studies performed on *M. tuberculosis* H37Rv [25,40], but not been reported in previous proteomic works on *M. bovis* BCG [14,24,41]. This RD region is also missing in other mycobacterial strains such as *Mycobacterium microti* or *Mycobacterium pinnipedii*. This region was first described by Gordon *et al.*, 1999 [42] as RD8 and later put in an evolutionary context by Brosch *et al.*, 2002 [43], which now corresponds to the region RD09 described by Behr *et al.*, [39]. A close examination of the gene encoding Rv3623 revealed that it is 207 bp shorter with a deletion in the N-terminal region that includes the signal peptide and the predicted lipo-box in the genomic sequences of *M. bovis* AF2122/97 and *M. bovis* BCG Pasteur 1173P2. The gene is annotated to encode a lipoprotein in the *M. bovis* strains even though the lipo-box is missing and it is therefore questionable whether it should be considered as a lipoprotein in *M. bovis*. The identification of this protein with 7 peptides covering 34% of its sequence in *M. tuberculosis* H37Rv suggests that it is a major lipoprotein.

The two lipoproteins listed in Table 3, annotated as "periplasmic phosphate-binding lipoprotein" (Rv0932c) is a known antigen [44] that also induces antibody responses in tuberculosis patients [45]. The 19 kDa lipoprotein antigen precursor (Rv3763) have been extensively studied due to its immunogenic properties [46-49]. Enrichment and analysis of lipoproteins with respect to humoral and cell-mediated immunity in infected individ-

uals might ultimately lead to the identification of additional antigens that can serve as biomarkers for *M. tuberculosis* infection.

Conclusion

In summary, we have enriched and extracted membrane- and membrane-associated proteins from *M. tuberculosis* H37Rv using Triton X-114, and identified the largest number of this subset of proteins reported so far. Further analysis of the data obtained in this study with bioinformatic tools suggests that several of these proteins are major membrane proteins. We have described one major lipoprotein of *M. tuberculosis* which has become a pseudogene by the RD9 deletion in *M. bovis*.

Methods

Preparation of crude bacterial extracts

The mycobacterial reference strain *M. tuberculosis* H37Rv (ATCC 27294), used in this study was kindly provided by Dr Harleen Grewal, The Gade Institute, University of Bergen, Bergen, Norway. The bacilli were cultured on Middelbrook 7H10 agar plates with OADC enrichment (BD Difco) at 37°C and 5% CO₂ for 3-4 weeks. Bacterial colonies were harvested by using an extraction buffer consisting of phosphate-buffered saline (PBS), pH 7.4 with freshly added Roche Protease Inhibitor Cocktail (Complete, EDTA-free, Roche GmbH, Germany). Six hundred µl of this extraction buffer was added to each agar plate and the mycobacterial colonies were gently scraped off the agar surface using a cell scraper. Aliquots of the resulting pasty bacterial mass was transferred into 2 ml cryo-tubes with O-rings (Sarstedt, Norway) containing 250 µl of acid washed glass beads (≤ 106 µm; Sigma-Aldrich, Norway) and an additional 600 µl of extraction buffer, and stored at -80°C until protein extraction was performed. For protein extraction, the mycobacteria were disrupted mechanically by bead-beating in a Ribolyser (Hybaid, UK) at max speed (6.5) for 45 seconds.

Triton X-114 extraction of exported proteins from whole bacteria

Triton X-114 phase-separation was used to isolate lipophilic proteins following the method of Bordier [50]. In brief, 3-4 week old bacilli were lysed by bead beating and centrifuged, initially at 2300 g to remove unbroken cells and cell-wall debris. Triton X-114 was added to the supernatant (final detergent concentration 2%, v/v) and the suspension was stirred at 4°C for 20 minutes to obtain the protein extract in a single phase. Residual insoluble matter was removed by centrifugation at 15700 g for 10 min, and the solution separated into two phases, an upper (aqueous) and lower (detergent) phase after 10 minutes incubation at 37°C. The detergent phase was collected and proteins were precipitated by acetone.

Gel electrophoresis and in-gel digestion of proteins

Extracted proteins (50 µg) were mixed with 25 µl SDS loading buffer and boiled for 5 minutes before separation on a 10 cm long 1 mm thick 12% SDS polyacrylamide gel (Invitrogen, Carlsbad, CA, U.S.A.). The protein migration was allowed to proceed until the bromophenol dye had migrated to the bottom of the gel. The protein bands were visualized with Coomassie Brilliant Blue R-250 staining (Invitrogen). Protein lanes were excised and divided in fractions according to the bands of the protein standard, ranging from ~3 kDa to ~188 kDa. The gel pieces were washed twice with 50% acetonitrile (ACN) in 25 mM ammonium bicarbonate (NH₄HCO₃) for 15 minutes at room temperature (RT), and subsequently dehydrated by incubating them with 50 µl 100% ACN for 20 minutes at RT. The proteins were reduced using 10 mM dithiothreitol and alkylated with 55 mM iodoacetamide; both in 100 mM NH₄HCO₃. The gel pieces were dehydrated by 100% ACN as described above, and rehydrated in 25 mmol/l NH₄HCO₃ followed by in-gel protein digestion with trypsin (Promega, Madison, U.S.A.) for 16-20 h at 37°C. The digested peptides were eluted by incubating the gel pieces with 50 µl 1% formic acid (FA) for 20 minutes at RT. The supernatant containing the peptides were collected after centrifugation at 15700 g for 10 minutes. Then, the gel pieces were incubated with 50 µl 0.1% FA in 50% ACN for 20 minutes at RT, followed by centrifugation at 15700 g. The supernatant was collected and combined with the previous one. Finally, the gel pieces were dehydrated with 50 µl 100% ACN for 20 minutes at RT, and the supernatant was collected after centrifugation as described above and added to the pool.

Mass spectrometry

Experiments were performed on a Dionex Ultimate 3000 nano-LC system (Sunnyvale CA, USA) connected to a linear quadrupole ion trap-Orbitrap (LTQ-Orbitrap) mass spectrometer (Thermo Electron, Bremen, Germany) equipped with a nano-electrospray ion source. The mass spectrometer was operated in the data-dependent mode to automatically switch between Orbitrap-MS and LTQ-MS/MS acquisition. Survey full scan MS spectra (from m/z 400 to 2,000) were acquired in the Orbitrap with resolution R = 60,000 at m/z 400 (after accumulation to a target of 1,000,000 charges in the LTQ). The method allowed sequential isolation of up to five of the most intense ions for fragmentation on the linear ion trap using collision induced dissociation at a target value of 100,000 charges.

For accurate mass measurements the lock mass option was enabled in MS mode and the polydimethylcyclsiloxane (PCM) ions generated in the electrospray process from ambient air (protonated (Si(CH₃)₂O)₆; m/z 445.120025) were used for internal recalibration during

the analysis [51]. Target ions already selected for MS/MS were dynamically excluded for 30 seconds. General mass spectrometry conditions were: electrospray voltage, 1.9 kV Ion selection threshold was 500 counts for MS/MS, an activation Q-value of 0.25 and activation time of 30 ms was also applied for MS/MS.

The obtained data was searched against the publicly available Tuberculist database version R10 <http://genolist.pasteur.fr/TubercuList/> using MASCOT software version 2.1 (Matrix Science, UK). The database was in-house modified to include reversed sequences of the original ORFs in order to determine false-positive thresholds of the Mascot identification engine [52]. Tuberculist was preferred over secondary annotations performed by independent institutes because previous data from our group demonstrated that the Tuberculist annotation appear to be more reliable [33]. The criteria for the Mascot search were as follows: Cysteine carbamidomethylation was set as fixed modification, methionine oxidation and N-acetylation (protein) as variable modifications. Up to 3 missed cleavages were allowed. Peptide (precursor) ion mass tolerance was 15 ppm, and the fragment ion tolerance was 0.5 Da. Mascot scoring showed that p > 0.01 was equivalent to a score of 24. The criterion for a positive identification of proteins identified with at least 2 peptides was a minimal score of 24 for each peptide which represents a 1:10,000 false positive rate at protein level. The maximal score for a peptide from a reversed entry of the annotated *M. tuberculosis* H37Rv database was found to be 31 (data not shown). This was considered as a threshold for false-positive identifications, and all proteins identified in this study with only one peptide were based on a score higher than 37 (25:10,000). No false positive identifications were observed from the reversed database using these criteria. For visualization and validation of spectra, MSQuant version +1.4.2 was used. MSQuant is an open source tool available at <http://msquant.sourceforge.net> and is widely used for LC-MS/MS data analysis [51].

Western blot

Proteins from both lipid and aqueous phase were separated by SDS-PAGE, electroblotted to nitrocellulose membranes (Amersham Biosciences) and blocked with 5% non-fat milk in PBS containing 0.5% Tween 20 (PBST) for 1 hour at RT. The membranes were then washed with PBST for 10 min. This was repeated three times. After the last wash, the membranes were incubated overnight at 4°C with rabbit antisera raised against 1) a cell wall fraction and 2) a crude whole cell lysate of *M. bovis* BCG. Sera were diluted 1:500 in PBS with 1% non-fat milk and 0.1% Tween 20. The blots were washed thoroughly with PBST as described above, and probed with Horse Radish Peroxidase (HRP) conjugated anti-rabbit IgG (1:2000 dilution) (Amersham Biosciences) for 1 hour at RT. Antigen-antibody complexes were visualized by a chemilumi-

nescent reaction (Pierce, Rockford, IL, U.S.A.) using Chemidoc XRS (Bio-Rad, Hercules, CA, USA).

Gene and protein sequence analysis

Gene and protein sequences were obtained from Tuberculist <http://genolist.pasteur.fr/TubercuList/> and BoviList <http://genolist.pasteur.fr/BoviList/>. Sequences alignments were done using the Blast 2 algorithm <http://blast.ncbi.nlm.nih.gov/Blast.cgi>. For prediction of lipoproteins, the LipoP algorithm was used <http://www.cbs.dtu.dk/services/LipoP/>. For detection of potential secreted proteins SignalP version 3.0 was used <http://www.cbs.dtu.dk/services/SignalP/>.

Estimation of protein abundance

The abundance of each protein was estimated by calculating the protein abundance index (PAI) [53], and the emPAI [15]. The estimation is based on the calculation of identified peptides per protein normalized by the theoretical number of peptides for the same protein. This is considered to be a good method for quantitative estimation because it takes into account that larger proteins are expected to generate more observable peptides in the mass spectrometry analysis, compared to smaller ones [15,16]. The final peptide list obtained from the MS analysis was submitted to a publicly available tool <http://empai.iab.keio.ac.jp/>, and emPAI values were calculated using the following parameters: *M. tuberculosis* H37Rv Tuberculist version R10 database; trypsin enzyme, carbamidomethyl (C) modification; peptide MW range from 300 to 6000 Da; no retention time filtering; peptide score higher than 24 as filtered by Mascot.

Additional material

Additional file 1 Figure S1: Collision induced dissociation fragmentation pattern of ion M+2H 1210.62. The sequence identified by the Mascot engine was CGSPAWDLPTVFGPIAITYNIK₁₁₉₋₁₄₀ from protein Rv0932c.

Additional file 2 Table S1: List of observed membrane- and membrane-associated proteins from *M. tuberculosis* H37Rv.

Additional file 3 Table S2: List of all observed *M. tuberculosis* H37Rv proteins in the lipid phase of Triton X-114 detergent, sorted by their Sanger IDs.

Additional file 4 Table S3: Information about the criteria for protein identifications, such as number of peptides matching each protein, scores, identification threshold and peak lists.

Authors' contributions

HM contributed to overall conception and design, analysis and interpretation of data, and manuscript drafting. SP cultured *M. tuberculosis* and extracted proteins. TS contributed with protein separation and mass spectrometry analysis. GAdS contributed with LTQ-Orbitrap expertise, data acquisition and critical revision of the data. HGW contributed with design, project coordination, manuscript drafting and critical revision. All authors have read and approved the final manuscript.

Acknowledgements

This work was supported by grants from the Regional Health Authorities of Western Norway (Projects 911077, 911117 and 911239) and by the National Programme for Research in Functional Genomics in Norway (FUGE) funded by the Norwegian Research Council (Project 175141/S10). We thank Dr. Benjamin

Thomas and the Proteomic Facility at the Dunn School of Pathology, Oxford University, for providing time at the LTQ-Orbitrap used on this work. We thank the Proteomic unit, PROBE, University of Bergen for analytical services. We are indebted to Professor Lars Haarr for critical comments to the manuscript.

Author Details

¹Section for Microbiology and Immunology, the Gade Institute, University of Bergen, Bergen, Norway and ²Department of Microbiology and Immunology, Haukeland University Hospital, Bergen, Norway

Received: 18 September 2009 Accepted: 29 April 2010

Published: 29 April 2010

References

1. Kaufmann SH: Tuberculosis: back on the immunologists' agenda. *Immunity* 2006, **24**:351-357.
2. Camacho LR, Ensergueix D, Perez E, Gicquel B, Guilhot C: Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. *Mol Microbiol* 1999, **34**:257-267.
3. Russell RB, Eggleston DS: New roles for structure in biology and drug discovery. *Nat Struct Biol* 2000, **7**(Suppl):928-930.
4. Daffe M, Etienne G: The capsule of *Mycobacterium tuberculosis* and its implications for pathogenicity. *Tuber Lung Dis* 1999, **79**:153-169.
5. Zuber B, Chami M, Houssin C, Dubochet J, Griffiths G, Daffe M: Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. *J Bacteriol* 2008, **190**:5672-5680.
6. Hoffmann C, Leis A, Niederweis M, Plitzko JM, Engelhardt H: Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proc Natl Acad Sci USA* 2008, **105**:3963-3967.
7. Velayati AA, Farnia P, Ibrahim TA, Haroun RZ, Kuan HO, Ghanavi J, Farnia P, Kabarei AN, Tabarsi P, Omar AR, Varahram M, Masjedi MR: Differences in Cell Wall Thickness between Resistant and Nonresistant Strains of *Mycobacterium tuberculosis*: Using Transmission Electron Microscopy. *Chemotherapy* 2009, **55**:303-307.
8. Camus JC, Pryor MJ, Medigue C, Cole ST: Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology* 2002, **148**:2967-2973.
9. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE III, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998, **393**:537-544.
10. Gu S, Chen J, Dobos KM, Bradbury EM, Belisle JT, Chen X: Comprehensive Proteomic Profiling of the Membrane Constituents of a *Mycobacterium tuberculosis* Strain. *Mol Cell Proteomics* 2003, **2**:1284-1296.
11. Mawuenyega KG, Forst CV, Dobos KM, Belisle JT, Chen J, Bradbury EM, Bradbury AR, Chen X: *Mycobacterium tuberculosis* functional network analysis by global subcellular protein profiling. *Mol Biol Cell* 2005, **16**:396-404.
12. Sinha S, Kosalal K, Arora S, Namane A, Sharma P, Gaikwad AN, Brodin P, Cole ST: Immunogenic membrane-associated proteins of *Mycobacterium tuberculosis* revealed by proteomics. *Microbiology* 2005, **151**:2411-2419.
13. Xiong Y, Chalmers MJ, Gao FP, Cross TA, Marshall AG: 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**:855-861.
14. Målen H, Berven FS, Sjøteland T, Arntzen MØ, D'Santos CS, De Souza GA, Wiker HG: 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**:1859-1870.
15. Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M: Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* 2005, **4**:1265-1272.

16. Ishihama Y, Schmidt T, Rappsilber J, Mann M, Hartl FU, Kerner MJ, Frishman D: **Protein abundance profiling of the *Escherichia coli* cytosol.** *BMC Genomics* 2008, **9**:102.
17. Babu MM, Priya ML, Selvan AT, Madera M, Gough J, Aravind L, Sankaran K: **A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins.** *J Bacteriol* 2006, **188**:2761-2773.
18. Rezwani M, Grau T, Tschumi A, Sander P: **Lipoprotein synthesis in mycobacteria.** *Microbiology* 2007, **153**:652-658.
19. Song H, Sandie R, Wang Y, Andrade-Navarro MA, Niederweis M: **Identification of outer membrane proteins of *Mycobacterium tuberculosis*.** *Tuberculosis (Edinb)* 2008, **88**:526-544.
20. Kyte J, Doolittle RF: **A simple method for displaying the hydrophobic character of a protein.** *J Mol Biol* 1982, **157**:105-132.
21. Althage M, Bizouan T, Kindlund B, Mullins J, Alander J, Rydstrom J: **Cross-linking of transmembrane helices in proton-translocating nicotinamide nucleotide transhydrogenase from *Escherichia coli*: implications for the structure and function of the membrane domain.** *Biochim Biophys Acta* 2004, **1659**:73-82.
22. Guenebaut V, Vincentelli R, Mills D, Weiss H, Leonard KR: **Three-dimensional structure of NADH-dehydrogenase from *Neurospora crassa* by electron microscopy and conical tilt reconstruction.** *J Mol Biol* 1997, **265**:409-418.
23. Guenebaut V, Schlitt A, Weiss H, Leonard K, Friedrich T: **Consistent structure between bacterial and mitochondrial NADH:ubiquinone oxidoreductase (complex I).** *J Mol Biol* 1998, **276**:105-112.
24. Mattow J, Siejak F, Hagens K, Schmidt F, Koehler C, Treumann A, Schaible UE, Kaufmann SH: **An improved strategy for selective and efficient enrichment of integral plasma membrane proteins of mycobacteria.** *Proteomics* 2007, **7**:1687-1701.
25. Gu S, Chen J, Dobos KM, Bradbury EM, Belisle JT, Chen X: **Comprehensive proteomic profiling of the membrane constituents of a *Mycobacterium tuberculosis* strain.** *Mol Cell Proteomics* 2003, **2**:1284-1296.
26. Xiong Y, Chalmers MJ, Gao FP, Cross TA, Marshall AG: **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**:855-861.
27. Sander P, Rezwani M, Walker B, Rampini SK, Kroppenstedt RM, Ehlers S, Keller C, Keeble JR, Hagemeyer M, Colston MJ, Springer B, Bottger EC: **Lipoprotein processing is required for virulence of *Mycobacterium tuberculosis*.** *Mol Microbiol* 2004, **52**:1543-1552.
28. Pennini ME, Pai RK, Schultz DC, Boom WH, Harding CV: ***Mycobacterium tuberculosis* 19-kDa lipoprotein inhibits IFN-gamma-induced chromatin remodeling of MHC2TA by TLR2 and MAPK signaling.** *J Immunol* 2006, **176**:4323-4330.
29. Young DB, Garbe TR: **Lipoprotein antigens of *Mycobacterium tuberculosis*.** *Res Microbiol* 1991, **142**:55-65.
30. Abebe F, Holm-Hansen C, Wiker HG, Bjune G: **Progress in serodiagnosis of *Mycobacterium tuberculosis* infection.** *Scand J Immunol* 2007, **66**:176-191.
31. Nikaïdo H: **Molecular basis of bacterial outer membrane permeability revisited.** *Microbiol Mol Biol Rev* 2003, **67**:593-656.
32. Målen H, Berven FS, Fladmark KE, Wiker HG: **Comprehensive analysis of exported proteins from *Mycobacterium tuberculosis* H37Rv.** *Proteomics* 2007, **7**:1702-1718.
33. De Souza GA, Målen H, Søfteland T, Saelensminde G, Prasad S, Jonassen I, Wiker HG: **High accuracy mass spectrometry analysis as a tool to verify and improve gene annotation using *Mycobacterium tuberculosis* as an example.** *BMC Genomics* 2008, **9**:316.
34. Jungblut PR, Müller EC, Mattow J, Kaufmann SH: **Proteomics reveals open reading frames in *Mycobacterium tuberculosis* H37Rv not predicted by genomics.** *Infect Immun* 2001, **69**:5905-5907.
35. De Souza GA, Søfteland T, Koehler CJ, Thiede B, Wiker HG: **Validating divergent ORF annotation of the *Mycobacterium leprae* genome through a full translation data set and peptide identification by tandem mass spectrometry.** *Proteomics* 2009, **9**:3233-3243.
36. Harth G, Horwitz MA: **An inhibitor of exported *Mycobacterium tuberculosis* glutamine synthetase selectively blocks the growth of pathogenic mycobacteria in axenic culture and in human monocytes: extracellular proteins as potential novel drug targets.** *J Exp Med* 1999, **189**:1425-1436.
37. Harth G, Clemens DL, Horwitz MA: **Glutamine synthetase of *Mycobacterium tuberculosis*: extracellular release and characterization of its enzymatic activity.** *Proc Natl Acad Sci USA* 1994, **91**:9342-9346.
38. Tullius MV, Harth G, Horwitz MA: **Glutamine synthetase GlnA1 is essential for growth of *Mycobacterium tuberculosis* in human THP-1 macrophages and guinea pigs.** *Infect Immun* 2003, **71**:3927-3936.
39. Behr MA, Wilson MA, Gill WP, Salamon H, Schoolnik GK, Rane S, Small PM: **Comparative genomics of BCG vaccines by whole-genome DNA microarray.** *Science* 1999, **284**:1520-1523.
40. Sinha S, Kosalai K, Arora S, Namane A, Sharma P, Gaikwad AN, Brodin P, Cole ST: **Immunogenic membrane-associated proteins of *Mycobacterium tuberculosis* revealed by proteomics.** *Microbiology* 2005, **151**:2411-2419.
41. Zheng J, Wei C, Leng W, Dong J, Li R, Li W, Wang J, Zhang Z, Jin Q: **Membrane subproteomic analysis of *Mycobacterium bovis* bacillus Calmette-Guérin.** *Proteomics* 2007, **7**:3919-3931.
42. Gordon SV, Brosch R, Billault A, Garnier T, Eiglmeier K, Cole ST: **Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays.** *Mol Microbiol* 1999, **32**:643-655.
43. Brosch R, Philipp WJ, Stavropoulos E, Colston MJ, Cole ST, Gordon SV: **Genomic analysis reveals variation between *Mycobacterium tuberculosis* H37Rv and the attenuated *M. tuberculosis* H37Ra strain.** *Infect Immun* 1999, **67**:5768-5774.
44. Tanghe A, Lefevre P, Denis O, D'Souza S, Braibant M, Lozes E, Singh M, Montgomery D, Content J, Huygen K: **Immunogenicity and protective efficacy of tuberculosis DNA vaccines encoding putative phosphate transport receptors.** *J Immunol* 1999, **162**:1113-1119.
45. Målen H, Søfteland T, Wiker HG: **Antigen analysis of *Mycobacterium tuberculosis* H37Rv culture filtrate proteins.** *Scand J Immunol* 2008, **67**:245-252.
46. Greenaway C, Lienhardt C, Adegbola R, Brusasca P, McAdam K, Menzies D: **Humoral response to *Mycobacterium tuberculosis* antigens in patients with tuberculosis in the Gambia.** *Int J Tuberc Lung Dis* 2005, **9**:1112-1119.
47. Bothamley GH: **Epitope-specific antibody levels demonstrate recognition of new epitopes and changes in titer but not affinity during treatment of tuberculosis.** *Clin Diagn Lab Immunol* 2004, **11**:942-951.
48. Bothamley GH, Rudd R, Festenstein F, Ivanyi J: **Clinical value of the measurement of *Mycobacterium tuberculosis* specific antibody in pulmonary tuberculosis.** *Thorax* 1992, **47**:270-275.
49. Bothamley GH, Beck JS, Potts RC, Grange JM, Kardjito T, Ivanyi J: **Specificity of antibodies and tuberculin response after occupational exposure to tuberculosis.** *J Infect Dis* 1992, **166**:182-186.
50. Bordier C: **Phase separation of integral membrane proteins in Triton X-114 solution.** *J Biol Chem* 1981, **256**:1604-1607.
51. Olsen JV, de Godoy LM, Li G, Macek B, Mortensen P, Pesch R, Makarov A, Lange O, Horning S, Mann M: **Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap.** *Mol Cell Proteomics* 2005, **4**:2010-2021.
52. Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP: **Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome.** *J Proteome Res* 2003, **2**:43-50.
53. Rappsilber J, Ryder U, Lamond AI, Mann M: **Large-scale proteomic analysis of the human spliceosome.** *Genome Res* 2002, **12**:1231-1245.
54. Mawuenyega KG, Forst CV, Dobos KM, Belisle JT, Chen J, Bradbury EM, Bradbury AR, Chen X: ***Mycobacterium tuberculosis* functional network analysis by global subcellular protein profiling.** *Mol Biol Cell* 2005, **16**:396-404.
55. Rosenkrands I, King A, Weldingh K, Moniatte M, Moertz E, Andersen P: **Towards the proteome of *Mycobacterium tuberculosis*.** *Electrophoresis* 2000, **21**:3740-3756.

doi: 10.1186/1471-2180-10-132

Cite this article as: Målen et al., Definition of novel cell envelope associated proteins in Triton X-114 extracts of *Mycobacterium tuberculosis* H37Rv *BMC Microbiology* 2010, **10**:132