Toward understanding the essence of post-translational modifications for the *Mycobacterium tuberculosis* immunoproteome

Cécile A. C. M. van Els¹*, Véronique Corbière², Kaat Smits³, Jacqueline A. M. van Gaans-van den Brink¹, Martien C. M. Poelen¹, Françoise Mascart²,³, Hugo D. Meiring⁴† and Camille Locht⁵,⁶,⁷,⁸†

¹ Centre for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, Netherlands
² Laboratory for Vaccinology and Mucosal Immunity, Université Libre de Bruxelles (UL.B.), Brussels, Belgium
³ Immunobiology Clinic, Hôpital Erasme, Université Libre de Bruxelles (UL.B.), Brussels, Belgium
⁴ Institute for Translational Vaccinology, Bilthoven, Netherlands
⁵ Institut Pasteur de Lille, Center for Infection and Immunity of Lille, Lille, France
⁶ INSERM U1019, Lille, France
⁷ CNRS UMR8204, Lille, France
⁸ Université Lille Nord de France, Lille, France

*Correspondence:*
Cécile A. C. M. van Els, Centre for Infectious Disease Control, National Institute for Public Health and the Environment, Antonie van Leeuwenhoeklaan 9, 3721 MA Bilthoven, Netherlands
e-mail: cecile.van.els@rivm.nl

†Hugo D. Meiring and Camille Locht have contributed equally to this work.

CD₄⁺ T cells are prominent effector cells in controlling *Mycobacterium tuberculosis* (*Mtbc*) infection but may also contribute to immunopathology. Studies probing the CD₄⁺ T cell response from individuals latently infected with Mtbc or patients with active tuberculosis using either small or proteome-wide antigen screens so far revealed a multi-antigenic, yet mostly invariable repertoire of immunogenic Mtbc proteins. Recent developments in mass spectrometry-based proteomics have highlighted the occurrence of numerous types of post-translational modifications (PTMs) in proteomes of prokaryotes, including Mtbc. The well-known PTMs in Mtbc are glycosylation, lipidation, or phosphorylation, known regulators of protein function or compartmentalization. Other PTMs include methylation, acetylation, and pupylation, involved in protein stability. While all PTMs add variability to the Mtbc proteome, relatively little is understood about their role in the anti-Mtbc immune responses.

Here, we review Mtbc protein PTMs and methods to assess their role in protective immunity against Mtbc. Hyperconservation of human Mtbc T cell peptide epitopes has been described, perhaps arguing for a beneficial effect of recognition by the host for the pathogen (26, 27), yet epitope sequence variability has also been reported (3, 28, 29).

Several genome-wide screens and bioinformatics-guided approaches further added to the identification of novel protein antigens and immunodominant epitopes for a number of antigen presentation platforms (5, 13, 24, 29–33). Altogether, the picture emerging from these studies is consistent with a multi-epitopic, multi-antigenic IFNγ response during Mtbc infection. To investigate whether different protein classes have the same or diverse functional characteristics, Lindestam Arlehamn et al. combined genome-wide HLA class II binding predictions with high-throughput cellular screens of peptides to interrogate CD4⁺ T cell responses from latently infected individuals. A significant clustering was seen of the majority of targeted proteins, representing 42% of the total response to three broadly immunodominant antigenic islands, to only 0.55% of the total open reading frames (ORFs) (3). However, no quantitative, functional, or phenotypical distinction was observed between T cells elicited by the various protein classes involved, such as those assigned to be secreted or others belonging to secretion systems themselves, or to cell wall or...
cellular processes. Hence, because of equal functionality, no anti-
gen class could be implied in a more protective (or non-protective) profile over others.

Even though greatly informative, preselecting epitope candidates from the full Mtb proteome of approximately 4,000 ORFs based on bioinformatics has limitations. Binding algorithms may not be 100% effective and certain protective Mtb epitopes with weaker binding properties could perhaps rank too low in the assignment to be selected.

Moreover, the assumption that the immunoproteome is merely a direct translation of the coding genome is an oversimplification. As an additional level of proteome complexity, primary protein structures can be modified after translation. Multiple post-translational modifications (PTMs) occur in higher and lower organisms, involving proteolytic events or transfer of modifying groups to one or more amino acids of the proteins. These PTMs may influence the protein’s active state, compartmentalization, turnover, and/or interactions with other proteins. The rich nature of PTMs of prokaryotic proteomes has started to become unraveled only recently (34), essentially through advances in mass spectrometry (MS) (35). However, their presence in the Mtb proteome and their role in virulence and immunity have not received sufficient attention yet. Here, we review PTMs currently known to occur in the Mtb proteome and discuss whether they modify the Mtb immunoproteome indirectly, by engaging eukaryotic innate receptor signaling or antigen-processing pathways, or directly by persisting as structural moieties in the immunogenic epitopes. In addition, we highlight technologies enabling the unbiased detection and identification of the Mtb T cell epitope repertoire, modified or unmodified.

**POST-TRANSLATIONAL MODIFICATIONS OF Mtb PROTEINS**

Current advances in MS-based proteomics have revealed that, like in eukaryotes, PTMs can create an enormous diversity and complexity of gene products in prokaryotes, as was reviewed recently elsewhere (34). PTMs are covalent-processing events chemically changing protein structure, often catalyzed by substrate-specific enzymes. Hundreds of types of PTMs are known, some of which can occur in parallel to create even more heterogeneity in the protein arsenal (36, 37). There are several technical obstacles still to overcome in PTM analysis. In proteome measurements, each protein can be identified based on combined mass and fragmentation patterns from various cleaved peptides. In PTM measurements, each modification site is only represented by a single peptide. Modified peptides can be of low abundance and furthermore may have chemical properties requiring optimization of liquid chromatography (LC) separation techniques or fragmentation modules, used in MS identification. As a solution, robust MS-based proteomic workflows have been designed, including affinity-based enrichment strategies, identified novel glycosylation sites directly from culture filtrates (38), whereas others, comparing several fragmentation strategies, identified novel glycosylation sites directly from culture filtrate proteins (45, 71). These localizations corroborate with data suggesting that O-glycosylation and Sec-translocation, a process shuttling proteins across the bacterial cell envelope, are linked (39). As the number of bona fide identified Mtb glycoproteins is increasing, a glycosylation site motif is emerging, frequently observed in Gram-positive bacteria or in pathogenic mycobacterial species.

**GLYCOSYLATION**

Prokaryotes possess conserved N- and O-linked glycosylation pathways, capable of enzyme-catalyzed covalently coupling glycans (oligosaccharides) to proteins (65–67). N-linked glycosylation, in which oligosaccharide precursors are first assembled on a cytoplasmic carrier molecule before being transferred en bloc to the amide nitrogen of an Asn in the acceptor protein, has not been observed in Gram-positive bacteria or in pathogenic mycobacterial species. O-glycosylation in bacteria can proceed en bloc or stepwise, but for Mtb it is thought to be the latter. A model was proposed in which the initial glycosyl molecule is transferred to the hydroxy oxygen of the acceptor Thr or Ser residue, a process catalyzed by the protein O-mannosyltransferase (PMT) (Rv1002c) (39). Hereafter, further sugars are added one at a time, but the enzymes involved in this elongation are unknown. While the precise role of O-glycosylation of Mtb proteins is still elusive (68), this PTM appears essential for Mtb virulence, since Rv1002c deficient strains are highly attenuated in immunocompromised mice (69).

Initially, glycoproteins of Mtb were reported to contain glycan moieties based on their ability to bind the lectin concanavalin A (ConA), e.g., 38 kDa (PstS1) protein (40). MS then enabled assessment of glycosylation patterns of Mtb proteins, first the alanine-proline-rich 45–47 kDa antigen Apa (41, 70), followed by others, e.g., the lipoproteins (19 kDa) LpqH (42, 43) and SodC (44). Using ConA affinity capture or other sugar-based partitioning methods, and dedicated proteomics, Bell et al. reported a wealth of candidate Mtb glycoproteins, associated with membrane fractions and with culture filtrates (38), whereas others, comparing several fragmentation strategies, identified novel glycosylation sites directly from culture filtrate proteins (45, 71). Some O-glycosylated Mtb proteins constitute B cell antigens for serodiagnostics, such as the 38 kDa protein (72). Furthermore, they might contribute to the virulence of Mtb by binding as adhesins to innate immune receptors, promoting invasion of the host cells. The 19-kDa glycolipoprotein was shown to bind to the macrophage mannose receptor (MR) of monocytic THP-1 cells, hereby promoting the uptake of bacteria (73). Apa, secreted, as well as cell wall associated, binds to human pulmonary Surfactant Protein A (SP-A), an important lung C-type lectin (74). These two glycoproteins were also reported to be involved in Mtb binding to DC-SIGN on dendritic cells, although this needs further investigation (75).

**PHOSPHORYLATION**

Since Mtb can exist under various physiological states in the host, including dormancy and active replication, it makes use of a versatile mechanism to sense signals from the host and regulate cellular processes. Signal transduction through reversible
| PTM | Structure (example) | Function and notes | Mtb proteins exhibiting this PTM | Reference |
|-----|---------------------|--------------------|---------------------------------|-----------|
| O-glycosylation | O-(6-D-Mannopyranosyl)-L-Threonine | Pathogenesis | Apa/Rv1880; Mpt83/Rv2873; 19 kDa LpqH/Rv3763; 38 kDa PstS1/Rv0934; SodC/Rv0432; WGA enriched candidate glycoproteins | (38–45) |
| Phosphorylation | Phosphoserine | Regulation | 301 proteins | (46, 47) |
| Methylation | N,N-Dimethyllysine | Protease resistance | HBHA/Rv0475; LBP/Rv2986c | (49) |
| Acetylation | N-Acetylaspartate | Stability | Esat-6 (N-terminal threonine) | (49) |
| Lipidation | Lipidated Cysteine | Compartmentalization | 99 Putative lipoproteins; 42 lipoproteins | (38, 42, 44, 50–55) |
| Deamidation | Deamidated glutamine | Regulator of protein-ligand interaction | Pup/Rv2111c | (56) |
| N-formylation | N-Formylmethionine | Start bacterial protein synthesis (fMet) | Rv0476, Rv0277C, Rv0749, Rv1688C | (67, 58) |
| Pupylation | N-Lysine pupylation | Degradation signal (reversible) | 1,305 proteins | (56, 59–64) |

*a amino acid.

b Mass increment of modified aa (Da).
c Formally not a PTM but a modified aa.

d protein phosphorylation participates in this function. The Mtb genome encodes multiple serine/threonine protein kinases, and Ser/Thr/Tyr protein phosphorylation occurs extensively. In addition, Mtb makes extensive use of two-component signal transduction systems, which rely on a phosphorylation cascade involving His kinases. Using TiO$_2$-phosphopeptide enrichment, Prisic et al. assigned 301 phosphoproteins in Mtb grown under six different conditions and identified corresponding phosphorylation site motifs. These likely represent only a part of the Mtb phosphoproteome. However, little is known on the role of this PTM in...
the function or pathogenicity of these proteins, with exception of the His kinases in two-component systems (46).

**LIPIDATION**

Lipidation of proteins is predicted for a small percentage (0.9–2.5%) of ORFs in mycobacterial genomes, and is required for their anchoring and sorting to the cell surface [reviewed in Ref. (50, 76)]. The first step in Mtb lipoprotein biogenesis occurs in the N-terminal leader of preprocipropoproteins having a so-called lipobox motif, involving the attachment of diacylglycerol to the thiol group of a Cys, by Lgt (phosphatidylglycerol-pre-prolipoprotein diacylglycerol transferase). Second, the signal peptide directly upstream of the modified Cys is cleaved off by LspA (prolipoprotein signal peptidase/signal peptidase II). Only recently, proof was found that slow-growing Gram-positive mycobacteria also share the third step in lipoprotein biosynthesis with Gram-negative bacteria, i.e., adding a third acyl residue to the free amino group of the modified Cys by Lnt (phospholipid-apolipoprotein N-acyltransferase) (51). Brulé et al. described the BCG 2070c as the major ORF in BCG to encode a functional Lnt using a mycobacteria-specific acyl substrate, tuberculostearic acid (52). Lipoprotein genesis is essential for Mtb. Deletion of lgt was not possible (77), while an lspA deletion mutant was viable but had an attenuated phenotype (78, 79). For Mtb, multiple (candidate) lipoproteins have been identified, and classified as components of transport systems, enzymes, or as molecules involved in cell adhesion or in signaling (38, 50), several of which were not only lipidated but also glycosylated (42, 44, 52). In line with the dogma that lipoproteins are pathogen associated molecular patterns (PAMPs) sensed by TLR2 (80), Sanchez et al. showed that the glycolipoprotein 38 kDa PstS1 triggers a TLR2 and caspase-dependent apoptotic pathway in human macrophages (53). Besides this mechanism, the 19-kDa glycolipoprotein LpqH was shown also to induce a caspase independent apoptotic mechanism, involving mitochondrial apoptosis-inducing factor (AIF), killing macrophages (54). Furthermore, TLR2-dependent inhibition of MHC class II function was observed for LpqH (81). The cumulative data on LpqH suggest that through its PTMs, this glycolipoprotein exploits multiple innate immune receptors and mechanisms to enter (73), incapacitate, and kill mononuclear phagocytes. Notably, Lopez et al. reported that the lipid moiety of LpqH was not required for the TLR2-dependent apoptosis of macrophages (82). As another innate feature, LpqH and the lipoprotein LprG were found to directly stimulate TLR2/TLR1 on memory CD4+ T cells (55), presumably via engaging TLR2 and TLR1 pockets by their thioether-linked diacylglycerol and amide-linked third acyl chain, respectively (83).

**FORMYLATION**

Formylation/de-formylation of proteins is a typical hallmark of bacterial proteomes. Protein synthesis in bacteria is initiated with a formylated methionine (fMet) residue, which is then enzymatically cleaved by peptide deformylase (PDF) and methionine aminopeptidase to generate mature proteins. The human immune system can benefit from this unique formylation pathway to distinguish self from non-self proteins. Although formylation is not strictly a PTM, but comes with the first “modified” building block of protein synthesis, the presence of the formyl group can be considered a variation of plain translation of the genetic code. What might be the life span of the formylated state of proteins is unknown so far. However, short formylated Mtb protein fragments have been identified that can be presented as epitopes via non-classical murine MHC class Ib molecules of infected macrophages and appear to be protective in a Mtb challenge model (57, 58). This suggests that in vivo-formylated proteins can enter antigen-processing pathways before the enzymatic removal of the N-terminal fMet residue has occurred. Recently, N-formylated peptides of ESAT-6 and glutamine synthetase were found to have immunotherapeutic potential in a Mtb mouse infection model. A role for formyl peptide-receptor recognition in activation of innate immune cells was implied (84), but presentation via non-classical MHC molecules may also play a role.

**PUPYLATION**

Pupylation is a protein-to-protein modification, first identified in Mtb. It covalently attaches the C-terminal Glu of the 6.9-kDa “Protein Ubiquitin-like Protein” (Pup) to the ε-amine of Lys side chains of an interacting protein partner (59). Although the full purpose of the pupylation pathway in Actinobacteria remains to be elucidated, it is assumed that in Mtb, disposing of a proteasomal system, tagging proteins with Pup renders them susceptible for proteasomal degradation (60–62), similar to the well-known ubiquitin-initiated protein degradation pathway. The C-terminal Gpu of LpqH itself is generated by another PTM, i.e., deamidation of the C-terminal Gln (56). From various large-scale proteomic studies, a database of the mycobacterial “pupylome,” containing >150 verified pupylated proteins and >1,000 candidate pupylated proteins, was annotated (63). Depupylation activity also occurs (64), hence the modification can be reversed.

**ACETYLATION AND ACETYL-LIKE MODIFICATIONS**

Transferring an acetyl, propionyl, maloyl, or succinyl group to the ε-amine of lysines (Nε-modification) or to the α-amine of protein N-termini (Nα-modification) are widely occurring PTMs in prokaryotes (34). Mtb encodes multiple proteins annotated as putative acetyltransferases acting on protein substrates (85). A well-studied Nε-acetylated Mtb protein is the virulence factor and immunodominant antigen, early secretory antigenic target 6 (Esat-6) (49). Acetylation presumably confers protein stability and compartmentalization, and occurs at Thr2, becoming the N-terminus after removal of the fMet residue at position 1.

**METHYLATION**

This PTM involves the addition of one or several methyl groups to either the ε-amine of lysines or to the side chain carboxyl of Glu. Although this PTM occurs in Mtb, genes encoding Mtb protein-methyltransferases have not been identified yet. Two Mtb adhesins, heparin-binding hemagglutinin (HBHA, Rv0475) and laminin-binding protein (LPB, Rv2986c) were shown to be methylated (48). HBHA is a 28-kDa multifunctional protein found on the surface and in culture filtrates of mycobacteria. Automated Edman degradation and mass spectrometric analysis indicate that at least 13 out of 16 Lys residues in the Lys-Ala-Pro rich C-terminal region of HBHA can be mono- or dimethylated, generating a
spectral envelop of isoforms (Figure 1A) (48). HBHA mediates mycobacterial adherence to epithelial cells via the interactions of this C-terminus with sulfated glycoconjugates on the surface of epithelial cells and methylation was implied to play a role in resistance to proteases present in bronchoalveolar lavage fluids (86–88). Recently, Sohn et al. showed that HBHA from Mtb also targeted murine macrophages and induced apoptosis via a mechanism involving mitochondria (89). Interestingly, HBHA purified from Mycobacterium avium subsp. paratuberculosis contains an N-terminal acetylated alanine residue in addition to the methylated lysines (90), whereas there is no evidence for acetylation of the N-terminal residue of Mtb HBHA (88).

POST-TRANSLATIONAL Mtb PROTEIN MODIFICATIONS IN PROTECTIVE IMMUNITY AND VACCINE CANDIDATES

The rich variety of PTMs to a large proportion of the Mtb proteome is likely to play a major role in the successful intracellular lifestyle of Mtb during chronic and sometimes lifelong infections. In the quest of novel vaccines, urgently needed to improve the limited protective capacity of BCG, it may be useful to understand the role of these PTMs in the host response to Mtb infection. Over thousands of years, a balance has been reached in which Mtb avoids excessive immunity allowing it to survive in the host, and in which a certain level of immunity allows the host not to succumb to the infection.

While the primary Mtb proteome shows features of hyperconservation, suggesting an evolutionary advantage to ensure stable epitope recognition by CD4+ T cells (26), PTMs superimpose a high level of complexity. This may complicate the identification of protective protein antigens based on in silico analyses and recombinant DNA technologies. Once protective protein antigens have been identified, the exact structural features need to be known for optimization and process development of the antigen. Furthermore, it will be important to know whether a particular PTM acts as an immune modulator, or/and whether it is part of the structural antigen moiety targeted by the adaptive immune system. Illustrative in this respect are three examples of Mtb protein antigens with PTMs, currently considered as vaccine candidates because of their immunodominance in humans and/or protective effect in animal models.

The 45–47 kDa secretory and cell-surface adhesin Apa is a major mycobacterial antigen with different O-mannosylation patterns in pathogenic versus non-pathogenic mycobacterial species that are critical for its T cell antigenicity in vivo and in vitro (70, 91). T cells from BCG-vaccinated PPD-responsive individuals recognize either both native mannosylated Apa (nApa) and recombinant non-mannosylated Apa (rApa), or nApa only. These latter T cells did, in contrast to the former, not recognize synthetic peptides corresponding to the Apa protein sequence. Together with the finding that recognition of nApa required active antigen processing, these data suggest that mannosylation does not induce alternate processing of nApa but rather that the carbohydrate moiety is an intrinsic part of the T cell epitope(s) (92). Protection by Apa was shown in guinea pig and mouse models in the context of various vaccine platforms (protein, DNA, and poxvirus boost) and routes (intanasal and subcutaneous), as a subunit or as a BCG-booster vaccine (70, 92–94). In a mouse model, adjuvanted nApa was found to induce higher frequencies of CD4+ T cells, producing more cytokines, compared to adjuvanted rApa. However, both antigens were equally protective against virulent Mtb infection when used as a subunit vaccine or as a BCG-booster vaccine (92). This indicates that O-mannosylation is not required for the protective effect in this model. However, understanding of the impact of the different immune responses evoked by nApa and rApa, as well as the nature of the putative naturally processed glycopeptide(s), need further investigation.

In contrast to Apa, the natural PTM of HBHA, methylation, is essential for providing high levels of protection against Mtb challenge in mice, in addition to its antigenicity in Mtb-infected human individuals (95, 96). However, immunization of mice with purified non-methylated HBHA induces antibodies and Th1 cytokines at levels similar to those induced by immunization with
methylated HBHA. Also, the antibody isotype profiles are similar in both instances. Interestingly, however, only splenocytes isolated from mice immunized with methylated HBHA, and not with non-methylated HBHA, induce IFNγ secretion upon incubation with Mtb-pulsed macrophages. Methylated HBHA-specific T cell responses are likely to participate in protection against disease in humans, since T cells from patients with active TB secrete significantly lower amounts of IFNγ after stimulation with methylated HBHA than subjects with latent Mtb infection (Figure 1B) (23, 97, 98). HBHA is being considered as a BCG-booster vaccine (99), as responses to methylated HBHA were found to be primed in BCG-vaccinated infants (100). It is not yet known whether the PTM affects the presentation of non-modified protective T cell epitopes via modulation of antigen uptake or processing, or whether methylation is part of the protective T cell epitope(s) involved.

The N-terminal-Thr acetylated antigen ESAT-6 is known as an immunological hotspot in humans (6). During natural infection or after subunit vaccination in mice, vigorous Th1 type CD4⁺ T cell responses are directed to the N-terminal immunodominant epitope ESAT-6₁₋₁₅, whereas other epitopes are masked (101). These can be revealed by redesigning ESAT-6 analogs in which the dominant epitope is removed, resulting in the engagement of protective CD4⁺ T cell responses that resist infection-driven terminal differentiation (102). To our knowledge, the role of the N-acetylation at Thr2 in generating the ESAT-6 peptide repertoire has not been interrogated, yet in view of ESAT-6’s current status as a vaccine candidate in clinical testing (99, 103), such assessment has not been interrogated, yet in view of ESAT-6’s current status as a vaccine candidate in clinical testing (99, 103), such assessment may be important.

In order to fully characterize these candidate vaccine antigens, it will be important to elucidate the exact roles of the added glyco-, methyl-, or N-acetyl moieties, respectively. Does their presence modulate effective antigen processing, perhaps by steering proteolysis and immunodominance through masking certain enzyme cleavage sites as was shown for O-linked glycans (104), or are they part of the protective immunoproteome itself? Clearly more studies are needed, including epitope identification approaches to unravel, in these and other candidate vaccine candidates, the role of PTMs in the Mtb immunoproteome. Knowledge on the precise role of the PTM of Mtb vaccine candidates may be of great help to optimize vaccine candidates and potentially to simplify vaccine design and process development.

TOWARD UNBIASED ASSESSMENT OF THE Mtb IMMUNOPROTEOME

Protein antigens, modified or not, are translated for T cell surveillance into immunogens in antigen-processing pathways of antigen-presenting cells. This translation consists of enzymatic cleavage and rescue of protein fragments onto the molecules of a relevant antigen-presenting platform, such as classical class I or II MHC molecules (105), non-classical MHC molecules, including class Ib MHC molecules (16), or CD1 isoforms (17). The identification of the exact nature of the naturally processed and presented Mtb immunoproteome would require dedicated technologies such as LC-MS, first pioneered MHC class I ligands by Hunt et al. more than two decades ago (106, 107). Typically, cell lines would be grown at large scale (>1 × 10⁹ cells) and, after detergent solubilization and immunoaffinity purification of MHC-ligand complexes, bound peptide epitopes would be eluted. The purified endogenous MHC class I ligands were characterized by dedicated LC-MS and MS/MS sequencing.

Nowadays, ever evolving LC-MS/MS systems have greatly added to our understanding of the endogenous peptide repertoire and binding motifs of many MHC class I and II molecules (108–111), as well as of class1b MHC molecules (112). For the classical MHC pathways, the notion has emerged that antigen-presenting cells express approximately 100,000 MHC class I and II molecules at their surface, presenting thousands of different endogenous peptides, at widely divergent abundances (113). LC-MS/MS sequencing can unambiguously identify the epitopes as they are eluted from their antigen-presenting molecules in a qualitative and quantitative manner, revealing both primary epitope sequences, as well as any modifications to them (114). LC-MS/MS analyses have shown that processing inside antigen-presenting cells can generate modified or unpredictable MHC epitopes, such as deamidated (115), citullinated (116), or cysteinylated (117) ligands, as well as ligands arising from protein splicing (118–120) or from alternative reading frames or read-throughs of protein-encoding genes (121–123).

Pathogen-encoded immunoproteomes, including PTMs, generated from the proteome inside infected or antigen endocytosing antigen-presenting cells, should be detectable through LC-MS/MS sequencing approaches as well, although pathogen-derived ligands will be needles in the haystack of eluted self epitopes. To facilitate the identification of these non-self pathogen-derived antigens, targeted LC-MS/MS approaches have been developed (124–127). Foreign epitopes that originate from proteins synthesized during infection inside antigen-presenting cells, such as viral MHC class I epitopes during infection, can be traced using algorithms detecting isotopic patterns in the mass chromatograms of MHC immunoproteomes from carefully mixed infected and non-infected cell cultures that were metabolically labeled during growth (128). Alternatively, epitopes that arise from exogenous proteins endocytoosed by antigen-presenting cells during infection, such as bacterial MHC class II epitopes, can be traced back in the MHC-bound peptide repertoire after metabolic labeling of antigen during the prokaryotic cell growth (126, 129). However, if PTMs are suspected in the foreign MHC immunoproteome, chromatography, ion fragmentation strategy, and even affinity enrichment strategies will have to be considered accordingly. Until now, only a single study has reported the identification of several Mtb epitopes presented by MHC class I via LC-MS (130). More approaches are underway to extend our knowledge on the naturally processed and MHC-presented Mtb epitopes, including those derived from methylated HBHA, using dedicated LC-MS. These studies include large-scale human monocyte or dendritic cell cultures and either in vitro Mtb infection or targeted antigen pulsing. Inhibition of MHC class II presentation upon incubation with live Mtb, mycobacterial lysates, or purified antigens may frustrate these attempts (131, 132). Dedicated isolation and analytical discovery procedures should then help to identify the Mtb epitope “needles” in the self“haystack,” and increase our knowledge on the role of PTM in the Mtb immunoproteome.
CONCLUDING REMARKS

Fast developments in LC-MS/MS-based proteomics have enabled the detection of many types of PTMs in proteomes of prokaryotes, including Mtbc. Elucidating the role of PTMs in the immunoproteome of protective Mtb protein antigens is important for the molecular optimization of vaccine candidates, and will also greatly benefit from technical advancements in LC-MS/MS.

ACKNOWLEDGMENTS

We thank Ad de Jong for stimulating discussions. Cécile A. C. M. van Els, Véronique Corbière, Kaat Smits, Hugo D. Meiring, François Mascart, and Camille Lochte received funding on Mtbc epitime discovery by the EU FP7 Collaborative Project NEWTBVAC (grant no. HEALTH-F32009-241745).

REFERENCES

1. Flynn JL, Chan J. Immunoology of tuberculosis. *Annu Rev Immunol* (2001) 19:93–128. doi:10.1146/annurev.immunol.19.1.93
2. Kaufmann SH. How can immunology contribute to the control of tuberculosis? *Nat Rev Immunol* (2001) 1:20–30. doi:10.1038/35095558
3. Blythe MJ, Zhang Q, Vaughan K, de Castro R Jr, Salimi N, Bui HH, et al. An analysis of the epitope knowledge related to mycobacteria. *Immunome Res* (2007) 3:10. doi:10.1186/1674-7580-3-10
4. Kunnath-Velayudhan S, Porcelli SA. Recent advances in defining the immunoproteome of *Mycobacterium tuberculosis*. *Front Immunol* (2013) 4:335. doi:10. 3388/fimmu.2013.00335
5. Lindestam Arlehamn CS, Gerasimova A, Mele F, Henderson R, Swann J, Greenbaum JA, et al. Memory T cells in latent Mycobacterium tuberculosis infection are directed against three antigenic islands and largely contained in a C5cR3+CCR6+ Th1 subset. *PLoS Pathog* (2013) 9:e1003130. doi:10.1371/journal.ppat.1003130
6. Axelson-Robertson R, Magalhães I, Parida SK, Zulma A, Maurer M. The immunological footprint of *Mycobacterium tuberculosis* T-cell epitope recognition. *J Infect Dis* (2012) 205(Suppl 2):S301–15. doi:10.1093/infdis/jis198
7. Cala-de Paepe D, Layre E, Giacometti G, Garcia-Alles LF, Mori L, Hanau D, et al. Comparative analysis of *Mycobacterium tuberculosis* pe and ppe genes reveals a novel mycobacterial antigen processing pathway directing CD8+ T-cell response in natural hosts. *PLoS One* (2012) 7:e38090. doi:10.1371/journal.pone.0038090
8. Bastian M, Braun T, Bruns H, Rollinghoff M, Stenger S. Mycobacterial lipopolysaccharides elicit CD4+ CD8− T cells in *Mycobacterium tuberculosis*-infected humans. *J Immunol* (2008) 180:3436–46. doi:10.4049/jimmunol.180.5.3436
9. Seshadri C, Turner MT, Lewinsohn DM, Mycobacterium tuberculosis RD1, 2013. doi:10.1038/nmeth.2632
10. Skjoer RL, Brock I, Arend SM, Munk ME, Theisen M, Ottenhoff TH, et al. Epitope mapping of the immunodominant antigen TNFα and the two homologous proteins TNFα and TNFα2, which constitute a subfamily of the est-6 gene family. *Infect Immun* (2002) 70:4546–43. doi:10.1128/IAI.70.12.4546–4546.2002
11. Vordermeier HM, Levinson RW, Wilkinson BI, Wilkinson KA, Gideon HP, Young DB, et al. Conserved immune recognition hierarchy of mycobacterial PE/PPE proteins during infection in natural hosts. *PLoS One* (2012) 7:e40890. doi:10.1371/journal.pone.0040890
12. Cosma I, Chakravarti J, Small PM, Galagan J, Niemann S, Kremer K, et al. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat Genet* (2010) 42:498–503. doi:10.1038/ng.950
13. McEvoy CR, Cloete R, Muller B, Schuch AC, van Helden PD, Gagneux S, et al. Comparative analysis of *Mycobacterium tuberculosis* pe and ppe genes reveals high sequence variation and an apparent absence of selective constraints. *PLoS One* (2012) 7:e30593. doi:10.1371/journal.pone.0030593
14. Uplekar S, Heym B, Rougemont J, Cole ST. Comparative genomics of *Mycobacterium tuberculosis* provides evidence for gene conversion and epitope variation. *Infect Immun* (2011) 79:4042–9. doi:10.1128/IAI.05344-11
15. Axelson-Robertson R, Loxton AG, Walzl G, Ehlers MM, Kock MM, Zumla A, et al. A broad profile of co-dominant epitopes shapes the peripheral *Mycobacterium tuberculosis* specific CD8+ T-cell immune response in South African patients with active tuberculosis. *PLoS One* (2013) 8:e58309. doi:10.1371/journal. plosgen.0058309
16. Cho S, Mehra V, Thoma-Uszynski S, Stenger S, Serbina N, Mazzaccaro RJ, et al. Antimicrobial activity of MHC class I-restricted CD8+ T cells in human tuberculosis. *Proc Natl Acad Sci U S A* (2000) 97:12210–5. doi:10.1073/pnas.210391479
17. Hammond AS, Klein MR, Corrath T, Fox A, Jaye A, McDaniel KP, et al. *Mycobacterium tuberculosis* genome-wide screen exposes multiple CD8 T cell epitopes. *Clin Exp Immunol* (2005) 140:109–16. doi:10.1111/j.1365-2249.2005.02751.x
18. van Els et al. The role of PTM in the immunoproteome of Mtbc.
van Els et al. The role of PTM in the immunoproteome of Mtb

34. Cain JA, Solis N, Cordwell SJ. Beyond gene expression: the impact of protein post-translational modifications in bacteria. J Proteomics (2014) 97:265–86. doi:10.1016/j.jprot.2013.08.012
35. Olsen JV, Mann M. Status of large-scale analysis of post-translational modifications by mass spectrometry. Mol Cell Proteomics (2013) 12:3444–52. doi:10.1074/mcp.O113.034181
36. Jensen ON. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. Curr Opin Chem Biol (2004) 8:33–41. doi:10.1016/j.copcb.2003.12.009
37. Mann M, Jensen ON. Proteomic analysis of post-translational modifications. Nat Biotechnol (2003) 21:235–61. doi:10.1038/nbt0303-255
38. Bell C, Smith GT, Sweredoski MJ, Brennan PJ, Belisle JT. Definition of the N-terminal clustering of the O-glycosylation sites in the Mycobacterium tuberculosis lipoprotein SodC. Glycobiology (2009) 19:58–51. doi:10.1093/glycob/cwp102
39. Vander Ven BC, Harder JD, Crick DC, Belisle JT. Export-mediated assembly of mycobacterial glycoproteins parallels eukaryotic pathways. Science (2005) 309:941–3. doi:10.1126/science.1114347
40. Espitia C, Mancilla R. Identification, isolation and partial characterization of Mycobacterium tuberculosis glycoprotein antigens. Clin Exp Immunol (1989) 77:378–83
41. Dobos KM, Khoo KH, Swiderek KM, Brennan PJ, Belisle JT. Definition of the full extent of glycosylation of the 45-kilodalton glycoprotein of Mycobacterium tuberculosis. J Bacteriol (1996) 178:2498–506.
42. Herrmann JL, O’Gaorain P, Gallagher A, Thole JE, Young D. Bacterial glycoproteins: a link between glycosylation and proteolytic cleavage of a 19 kDa antigen from Mycobacterium tuberculosis. EMBO J (1996) 15:3547–54.
43. Herrmann JL, Denlay R, Gallagher A, Robertson B, Young D. Analysis of post-translational modification of mycobacterial proteins using a cassette expression system. FEBS Lett (2000) 473:358–62. doi:10.1016/S0014-5793(00)01553-2
44. Sartain MJ, Belisle JT. N-terminal clustering of the O-glycosylation sites in the Mycobacterium tuberculosis lipoprotein SodC. Glycobiology (2009) 19:58–51. doi:10.1093/glycob/cwp102
45. Smith GT, Sweredoski MJ, Hess S. O-linked glycosylation sites profiling in Mycobacterium tuberculosis culture filtrate proteins. J Proteomics (2014) 97:296–306. doi:10.1016/j.jprot.2013.05.011
46. Kusebauch U, Ortega C, Loirard A, Rogers RS, Sherman DR, Moritz RL, et al. Mycobacterium tuberculosis supports protein tyrosine phosphorylation. Proc Natl Acad Sci U S A (2014) 111:9265–70. doi:10.1073/pnas.1323894111
47. Prisci S, Dankwa S, Schwartz D, Chou MF, Locasale JW, Kang CM, et al. Extensive phosphorylation with overlapping specificity by Mycobacterium tuberculosis serine/threonine protein kinases. Proc Natl Acad Sci U S A (2010) 107:7521–6. doi:10.1073/pnas.0913487107
48. Peiter K, Bilani P, Proebec H, Serhergaut C, Debs E, Locht C, et al. Mycobacterial heparin-binding hemagglutinin and laminin-binding protein share anti-idiotype antibodies that confer resistance to proteinase. Proc Natl Acad Sci U S A (2002) 99:10779–84. doi:10.1073/pnas.162248698
49. Okkels LM, Muller EC, Schmid M, Rosenkrands I, Kaufmann SH, Andersen P, et al. CEP101 discriminates between nonacylated and acetylated ESAT-6 of Mycobacterium tuberculosis by differential interaction. Proteomics (2004) 4:2954–60. doi:10.1002/pmc.200404906
50. Sutcliffe IC, Harrington DJ. Lipoproteins of Mycobacterium tuberculosis: an abundant and functionally diverse class of cell envelope components. FEMS Microbiol Rev (2004) 28:645–59. doi:10.1016/j.femsre.2004.06.002
51. Okuda S, Tokuda H. Lipoprotein sorting in bacteria. Anna Rev Microbiol (2011) 65:239–59. doi:10.1146/annurev-micro-090110-102859
52. Brulle JK, Tschumi A, Sander P. Lipoproteins of slow-growing Mycobacteria carry three fatty acids and are N-acetylated by alkaloprotein N-acetytransferase BCG 2070c. BMC Microbiol (2013) 13:223. doi:10.1186/1471-2180-13-223
53. Sanchez A, Espinosa P, Esparza MA, Colon M, Bernal G, Mancilla R. Mycobacterium tuberculosis 38-kDa lipoprotein is apoptogenic for human monocyte-derived macrophages. Scand J Immunol (2009) 69:260–8. doi:10.1111/j.1365-3038.2008.02193.x
54. Sanchez A, Espinosa P, Garcia T, Mancilla R. The 19 kDa Mycobacterium tuberculosis lipoprotein (LpqH) induces macrophage apoptosis through extrinsic and intrinsic pathways: a role for the mitochondrial apoptosis-inducing factor. Clin Dev Immunol (2012) 2012:950503. doi:10.1155/2012/950503
55. Lanciai CL, Li Q, Thomas JI, Ding X, Thiel B, Drage MG, et al. Mycobacterium tuberculosis lipoproteins directly regulate human memory CD4+ T cell activation via Toll-like receptors 1 and 2. Infect Immun (2011) 79:663–73. doi:10.1128/IAI.00806-10
56. Striebel F, Imkamp F, Sutter M, Steiner M, Mann M, et al. Induction of M3-restricted cytotoxic T lymphocyte responses by N-formylated peptides derived from Mycobacterium tuberculosis. J Exp Med (2001) 193:2123–20. doi:10.1084/jem.193.6.8306
57. Doi T, Yamada Y, Yajima T, Wajiwalku W, Hayata T, Yoshihaki Y. H2-M3-restricted CD8+ T cells induced by peptide-pulsed dendritic cells confer protection against Mycobacterium tuberculosis. J Immunol (2007) 178:3806–13. doi:10.4049/jimmunol.178.6.8306
58. Sanchez A, Espinosa P, Garcia T, Mancilla R. The 19 kDa lipoprotein of Mycobacterium tuberculosis is a major adhesin that binds the mannose receptor of THP-1 monocytic cells and promotes phagocytosis of mycobacteria. Microb Path (2009) 39:97–107. doi:10.1016/j.micpath.2005.06.002
59. Ragas A, Roussel L, Puzo G, Riviere M, Belisle JT. Mycobacterium tuberculosis cell-surface glycoprotein apo as a potential adhesin to colonize target cells via the...
innate immune system pulmonary C-type lectin surfactant protein A. J Biol Chem (2007) 282:5313–4. doi:10.1074/jbc.M106183200

75. Piteaque S, Herrmann JL, Duteyrat JL, Jackson M, Stewart GR, Leceinte F, et al. Deciphering the molecular bases of Mycobacterium tuberculosis binding to the lectin DC-SIGN reveals an underestimated complexity. Biochem J (2005) 389:261–4. doi:10.1042/BJ20050709

76. Rezwan M, Grau T, Tschumi A, Sander P. Lipoprotein synthesis in mycobacteria. Microbiology (2007) 153:652–8. doi:10.1099/mic.0.2006/000216-0

77. Tschumi A, Grau T, Albrecht D, Rezwan M, Antelmann H, Sander P. Functional analyses of mycobacterial lipoprotein diacylglycerol transferase and companion secretome analysis of a mycobacterial lgt mutant. J Bacteriol (2012) 194:3938–49. doi:10.1128/JB.00127-12

78. Rampini SK, Selchow P, Ehlers S, Biet F, Delogu G, Brennan MJ, Locht C, et al. The heparin-binding hemagglutinin of Mycobacterium tuberculosis. Microbiology (2007) 153:5133–42. doi:10.1099/mic.0.2006/000216-0

79. Sander P, Rezwan M, Walker B, Rampini SK, Kroppenstedt RM, Ehlers S, et al. Lipoprotein processing is required for virulence of Mycobacterium tuberculosis. Mol Microbiol (2004) 52:1543–52. doi:10.1111/j.1365-2958.2004.04041.x

80. Zahringer U, Lindner B, Inamura S, Heine H, Alexander C. TLR2 – promiscuous or specific? A critical re-evaluation of a receptor expressing apparent broad specificity. Immunobiology (2008) 213:205–24. doi:10.1016/j.imbio.2008.02.005

81. Noss EH, Pai RK, Sellati TJ, Radolf JD, Belsle I, Golenbock DT, et al. Toll-like receptor 2-dependent inhibition of macrophage class II MHC expression and antigen processing by 19-kDa lipoprotein of Mycobacterium tuberculosis. J Immunol (2001) 167:910–8. doi:10.4049/jimmunol.167.2.910

82. Lopez M, Sily LM, Loo J, Young D, Cooper H, Reiner NE. The 19-kDa Mycobacterium tuberculosis protein induces macrophage apoptosis through Toll-like receptor-2. J Immunol (2003) 170:2409–16. doi:10.4049/jimmunol.170.7.2409

83. Jin MS, Kim SE, Heo JY, Lee ME, Kim HM, Paik SG, et al. Crystal structure of the TL1R–TL2 heterodimer induced by binding of a triacylated lipopeptide. Cell (2007) 130:97–108. doi:10.1016/j.cell.2007.09.008

84. Mir SA, Sharma S. Immunotherapeutic potential of N-formylated peptides of ESAT-6 and glutamine synthetase in experimental tuberculosis. Int Immunopharmacol (2014) 18:298–303. doi:10.1016/j.intimp.2013.09.010

85. Vetting MW, Errey JC, Blanchard JS. Rv0802c from Mycobacterium tuberculosis affects phagosome maturation arrest. J Bacteriol (2008) 190:2659–68. doi:10.1128/JB.00127-12

86. Pethe K, Alonso S, Biet F, Delogu G, Brennan MJ, Locht C, et al. Identification of a heparin-binding hemagglutinin present in mycobacteria. J Exp Med (1994) 180:993–1001. doi:10.1084/jem.184.3.993

87. Sohn H, Kim JS, Shin SJ, Kim K, Won CJ, Kim WS, et al. Targeting of Mycobacterium tuberculosis heparin-binding hemagglutinin to mitochondria in macrophages. PLoS Pathog (2011) 7:e1002435. doi:10.1371/journal.ppat.1002435

88. Menozzi FD, Rouse JJ, Alavi M, Laude-Sharp M, Muller J, Bischoff R, et al. Identification of a heparin-binding hemagglutinin present in mycobacteria. J Exp Med (1996) 184:1929–37. doi:10.1084/jem.184.10.1929

89. Hancock REW, Lehrer RI, Thornsberry C. Antimicrobial peptides and proteins of the innate immune system. Clin Microbiol Rev (2001) 14:133–352. doi:10.1128/CMR.14.1.133-352.2001

90. Normark S. The role of PTM in the immunoproteome of Mtb

www.frontiersin.org August 2014 | Volume 5 | Article 361 | 9
112. Lampen MH, Hassan C, Sluiter M, Geluk A, Dijkman K, Tjon JM, et al. Alternative peptide repertoire of HLA-E reveals a binding motif that is strikingly similar to HLA-A2. *Mol Immunol* (2013) 53:126–31. doi:10.1016/molimm.2012.07.009

113. van Els CA, Herberts CA, van der Heeft E, Poelen MC, van Gaans-van den Brink JA, van der Kooi A, et al. A naturally processed measles virus peptide fully dominates the HLA-A*0201-associated peptide display and is mutated at its anchor position in persistent viral strains. *Eur J Immunol* (2000) 30:1172–81. doi:10.1002/(SICI)1521-4141(200004)30:4<1172::AID-IMMU1172>3.0.CO;2-J

114. Hillen N, Stevanovic S. Contribution of mass spectrometry-based proteomics to immunology. *Expert Rev Proteomics* (2006) 3:653–64. doi:10.1586/14789450.3.6.653

115. Skipper JC, Hendrickson RC, Gulden PH, Brichard V, Van PeL A, Chen Y, et al. An HLA-A2-restricted tyrosinase antigen on melanoma cells results from posttranslational modification and suggests a novel pathway for processing of membrane proteins. *J Exp Med* (1996) 183:527–34. doi:10.1084/jem.183.2.527

116. Ireland JM, Unanue ER. Autophagy in antigen-presenting cells results in presentation of citrullinated peptides to CD4 T cells. *J Exp Med* (2011) 208:2625–32. doi:10.1084/jem.20110640

117. Pierce RA, Field ED, den Haan JM, Caldwell JA, White FM, Marto JA, et al. Cutting edge: the HLA-A*0101-restricted HY minor histocompatibility antigen originates from DFFRY and contains a cysteinylated cysteine residue as identified by a novel mass spectrometric technique. *J Immunol* (1999) 163:6360–4.

118. Hanada K, Yewdell JW, Yang JC. Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature* (2004) 427:252–6. doi:10.1038/nature02240

119. Vigneron N, Stroobant V, Chapiron J, Ooms A, Degiovanni G, Morel S, et al. An antigenic peptide produced by peptide splicing in the proteasome. *Science* (2004) 304:587–90. doi:10.1126/science.1095522

120. Engelhard VH. Creating new peptide antigens by slicing and splicing proteins. *Nat Immunol* (2004) 5:128–9. doi:10.1038/nimi0204-128

121. Bullock TN, Patterson AE, Franlin LL, Notidis E, Eisenlohr LC. Initiation mechanism. *Eur J Immunol* (1999) 29:681–90. doi:10.1002/(SICI)1521-4002(199903)29:3<681::AID-IEIJ6>3.0.CO;2-X

122. Malarkannan S, Horng T, Shih PP, Schwab S, Shastri N. Presentation of out-of-frame peptide/MHC class I complexes by a novel translation initiation mechanism. *Immunity* (1999) 10:681–90. doi:10.1016/S1074-7613(00)80067-9

123. Mandic M, Almunia C, Vicedo S, Gillet D, Janjic B, Coval K, et al. The alternative open reading frame of LAGE-1 gives rise to multiple promiscuous HLA-DR-restricted epitopes recognized by T-helper 1-type tumor-reactive CD4+ T cells. *Cancer Res* (2003) 63:5506–15.

124. Dengjel J, Decker P, Schoor O, Altenberend F, Weinschenk T, Rammensee HG, et al. Identification of a naturally processed cyclin D1 T-helper epitope by a novel combination of HLA class II targeting and differential mass spectrometry. *Eur J Immunol* (2004) 34:3644–51. doi:10.1002/eji.200425510

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 02 April 2014; accepted: 14 July 2014; published online: 11 August 2014. Citation: van Els CACM, Corbière V, Smits K, van Gaans-van den Brink JA, Poelen MCM, Mascart F, Meiring HD and Locht C (2014) Toward understanding the essence of post-translational modifications for the Mycobacterium tuberculosis immunoproteome. *Front. Immunol.* 5:361. doi:10.3389/fimmu.2014.00361 This article was submitted to Microbial Immunology, a section of the journal Frontiers in Immunology. Copyright © 2014 van Els, Corbière, Smits, van Gaans-van den Brink, Poelen, Mascart, Meiring and Locht. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.