# Nε- and O-Acetylation in *Mycobacterium tuberculosis* Lineage 7 and Lineage 4 strains: Proteins Involved in Bioenergetics, Virulence and Antimicrobial Resistance are Acetylated

Alemayehu Godana Birhanu[1, 2]*, Solomon Abebe Yimer[1, 3], Carol Holm-Hansen[4], Gunnstein Norheim[4], Abraham Aseffa[5], Markos Abebe[5], Tone Tønjum[1,3]*

[1]Department of Microbiology, University of Oslo, PO Box 4950, Nydalen, NO-0424 Oslo, Norway

[2]Addis Ababa University, Institute of Biotechnology, PO Box 1176, Addis Ababa, Ethiopia

[3]Department of Microbiology, Oslo University Hospital, PO Box 4950, Nydalen, NO-0424 Oslo, Norway

[4]Infection Control and Environmental Health, Norwegian Institute of Public Health, PO Box 4404, Nydalen, NO-0403 Oslo, Norway

[5]Armauer Hansen Research Institute, Jimma Road, PO Box 1005, Addis Ababa, Ethiopia

Email address of all co-authors:

alexbiology97@yahoo.com
s.a.yimer@medisin.uio.no
carol.holm-hansen@fhi.no
gunnstein.norheim@fhi.no
aseffaa@gmail.com
markosabebe@yahoo.com
tone.tonjum@medisin.uio.no

## ABSTRACT

Increasing evidence demonstrates that lysine acetylation is involved in *Mycobacterium tuberculosis* (Mtb) virulence and pathogenesis. However, previous investigations in Mtb have only monitored acetylation at lysine residues using selected reference strains. We analyzed the global $N_\varepsilon$- and O-acetylation of 3 Mtb isolates; 2 lineage 7 clinical isolates and the lineage 4 H37Rv reference strain. Quantitative acetylome analysis resulted in identification of 2490 class-I acetylation sites, among them 2349 O-acetylation and 141 Nε-acetylation sites, derived from 953 unique proteins. Mtb O-acetylation was thereby significantly more abundant than Nε-acetylation. The acetylated proteins were found to be involved in central metabolism, translation, stress responses and antimicrobial drug resistance. Notably, 261 acetylation sites on 165 proteins were differentially regulated between lineage 7 and lineage 4 strains. A total of 257 acetylation sites on 160 proteins were hypoacetylated in lineage 7 strains. These proteins are involved in Mtb growth, virulence, bioenergetics, host-pathogen interaction and stress responses. This study provides the first global analysis of O-acetylated proteins in Mtb. This quantitative acetylome data expand the current understanding regarding the nature and diversity of acetylated proteins in Mtb, and opens a new avenue of research for exploring the role of protein acetylation in Mtb physiology.



**Keywords:** *Mycobacterium tuberculosis*; *lineage 7;* post-translational modifications; *acetylome; $N_\varepsilon$-acetylation; O-acetylation.*

## INTRODUCTION

*Mycobacterium tuberculosis* (Mtb) is the causative agent of tuberculosis (TB) in humans. TB is one of the top ten causes of mortality worldwide and the leading cause of deaths from an infectious disease, leading to 1.8 million deaths and 10.4 million new cases in 2015 [1]. Due to the increasing prevalence of antimicrobial drug resistance (AMR) and challenges in vaccine development, it is crucial to understand fundamental aspects of Mtb biology to achieve future elimination of this disease.

Nε-acetylation, which is acetylation at the ε-amine of lysine (K) residues, is an abundant and evolutionarily conserved post-translational modification (PTM) that regulates a broad range of functions in bacteria, including motility and chemotaxis, transcription, metabolism, DNA metabolism, siderophore biosynthesis, and stress responses [2-5]. Increasing evidence supports the presence and role of lysine acetylation in mycobacteria [6-12]. In Mtb, lysine acetylation presumably confers protein stability and compartmentalization, thereby modulating diverse cellular processes [13-15]. Acetylation of Mtb histone-like nucleoid protein MtHU modulates DNA binding and genome organization [16]. It has also been shown that reversible lysine acetylation regulates fatty acid metabolism in Mtb, and acetate and propionate metabolism in *M. smegmatis* (Msm) [6, 8]. Furthermore, Liu *et al*. have demonstrated the regulatory role of lysine acetylation in the immunogenicity of the secreted protein HspX in Mtb [17].

Lysine acetylation is modulated via both enzymatic and non-enzymatic mechanisms (reviewed in [5]). The enzymatic mechanism is regulated by the opposing actions of acetyltransferases and deacetylases. Transfer of an acetyl group from acetyl-CoA (AcCoA) to an Nε-lysine has been thought to occur enzymatically through lysine acetyltransferases, generating acetylated K. Although a plethora of Mtb acetyltransferases and deacetylases are predicted to be encoded by the Mtb genome [18], only a few of them are characterized to date.

Recently, Lee *et al.* assigned acetyltransferase activity to the Mtb Rv2170 protein that acetylates lysine residues in isocitrate dehydrogenase, leading to a reduction in its enzymatic activity [19]. Other characterized K acetyltransferases in Mtb includeds Rv0998 and Rv3423.1, which acetylate different target proteins [10, 20]. However, studies have demonstrated that lysine acetylation can also occur non-enzymatically in bacteria and mitochondria, with the secondary metabolite acetyl-phosphate (AcP) and AcCoA serving as the acetyl group donor [21-24]. Non-enzymatic acetylation has been shown to occur via direct interaction of the target protein and AcCoA, which is favored by high pH and high AcCoA concentrations such as those in mitochondria [25-26]. Protein deacetylases in Mtb includes Rv1151c, the Mtb homologue of cobB, which has been shown to deacetylate and regulate the activity of Mtb acetyl-CoA synthase (ACS) [2, 7, 17]. Deletion of deacetylases affects Mtb colony morphology and biofilm formation, as well as stress responses [17].

Mukherjee *at al.* discovered that YopJ, a secreted virulence factor from *Yersinia pestis*, acetylates and inhibits kinase activation in the host by blocking phosphorylation [27]. This is the first report regarding the presence and function of O-acetylation, which is acetylation at the –OH group of serine (S) and threonine (T) by a bacterial acetyltransferase. The function of YopJ within *Y. pestis* itself is not known. However, there is no YopJ ortholog in Mtb. Protein acetylation by itself is known to have a regulatory effect. Current evidence suggests that O-acetylation becomes more important when acetylation takes place on kinase substrates, which is known to regulate myriads of signaling pathways in Mtb [28-29]. The transfer of an acetyl group to the –OH group of serine [30] and glycoconjugates (peptidoglycans) [31-32] has also been reported.

Despite phenotypic variability between strains of Mtb, most of the PTM studies performed to date have used the reference laboratory strain H37Rv as a model organism [15, 17, 33]. Furthermore, bacterial acetylome studies have been limited to the analysis of lysine

residues. O-acetylation of proteins has been shown to be involved in regulating key functions in eukaryotes [27]. The delineation of such regulatory mechanisms will not only lead to a better understanding of Mtb basic biology and the discovery of potential new drug targets, but may also facilitate the development of new vaccines and diagnostic tools.

In this study, we analysed the global $N_{\varepsilon}$- and O-acetylome of two Mtb lineage 7 clinical isolates and the lineage 4 reference strain H37Rv using nanoscale liquid chromatography coupled to tandem mass spectrometry (nano LC-MS / MS). The aim of this study is to define the global $N_{\varepsilon}$- and O-acetylome profile of Mtb and to predict its possible contribution to the fitness and survival of Mtb using lineage 7 and lineage 4 strains as relevant model organisms.

## EXPERIMENTAL PROCEDURES

### Mycobacterial Strains and Growth Conditions

Mtb lineage 7 strains (L7-35 and L7-28) and lineage 4 reference strain (H37Rv), were inoculated onto Middlebrook 7H10 plates in triplicates and incubated in a humidified 37 °C, 5% $CO_2$ incubator. After 32 days, the cells were harvested and transferred to 50 mL Falcon® tubes. The cell pellets were resuspended in 30 mL Phosphate-buffered saline (PBS) containing, 10 mM $PO4^{3-}$, 137 mM NaCl, and 2.7 mM KCl, pH 7.4, and centrifuged at 3900 rpm for 20 min at 4 °C. The cell pellets were resuspended in 1 mL PBS, transferred into 2 mL screw capped tubes (Sarstedt, Nümbrecht, Germany) and heat-inactivated at 80 °C for 90 min. Culturing and processing of the Mtb samples prior to heat-inactivation were conducted in a biosafety level 3 facility at Oslo University Hospital, Norway. The heat-inactivated Mtb samples were stored at -20 °C until preparation for MS analysis.

### Preparation of Mtb Cell Lysate

The inactivated cell pellets in lysis buffer containing 2% SDS, 10mM Tris-HCl (pH 7.5), 10 mM DTT, EDTA free protease inhibitor cocktail (Roche) and PhosStop (Roche),

were transferred into Lysing Matrix B tubes (Roche, US) and disrupted mechanically by bead beating with MagNa Lyser for 90 seconds, speed 6.0 (Roche, US). The lysis procedure was repeated six times with 1 min cooling on ice between each bead beating. The lysate was clarified by centrifugation (15,000 × *g* for 15 min) at 21 °C, and the supernatant containing the whole cell lysate proteins was transferred in to new 2 mL screw capped micro tubes (Sarstedt, Germany*)*. Protein concentration was measured by direct detect infrared spectrometer (DirectDetect, Millipore).

**In-gel Trypsin Digestion**

One hundred µg of protein dissolved in NuPAGE LDS sample buffer (1x) and NuPAGE Sample Reducing Agent (1X) (Life Technologies) was incubated for 10 min at 70 °C and pre-fractionated by 1.0 mm, 4%-12% NuPAGE Novex Bis-Tris gel (Life Technologies), at 80 V for 5 min followed by 20 min at 200 V. SDS-PAGE gels were Coomassie stained using a Colloidal Blue Staining kit for NuPAGE according to the manufacturer's instructions. After staining, each gel lane was divided into 6 fractions, and each fraction was subjected to in-gel reduction, alkylation, and tryptic digestion [34]. Proteins were reduced using 10 mM DTT for 1 hour at 56 °C and alkylated with 55 mM iodoacetamide for 1 hour at room temperature. The reduced and alkylated peptides were digested with sequence grade trypsin (Promega, 1:100; w/w) for 16 hours at 37 °C in 50 mM $NH_4HCO_3$. The trypsin digested protein samples were extracted from the gel using acetonitrile (ACN) (50% and 100%), dried by SpeedVac concentrator (Eppendorf, concentrator 5301) and re-suspended using 0.05% trifluoroacetic acid (TFA). The digested protein samples were loaded on to C18 zip-tips activated and equilibrated with 95% ACN/0.1% FA and 0.1 % formic acid (FA), respectively. The loaded samples were washed with 0.05% TFA and eluted with 95% ACN/0.1% FA. The eluent was dried using SpeedVac

concentrator, re-suspended in 0.1% FA, transferred to auto-sampler nano LC vials for LC-MS/MS analysis and stored at -20 °C until injected in to LC-MS/MS.

**LC-MS/MS Analysis**

Peptide characterization and quantitation were performed by nano LC-MS / MS using a Q Exactive hybrid quadropole-orbitrap mass spectrometer interfaced with an EASY1000-nano-electrospray ion source (both from Thermo Scientific).

Peptides were injected in triplicates on to a pre-column (Acclaim PepMap 100, 75 µm x 2 cm, nanoviper, C18, 3 µm, 100Å, Thermo Scientific) and separated on an analytical column (PepMap RSLC, C18, 2µm, 100Å, 50µm x 15cm, Thermo Scientific) at 75 min solvent gradient and flow rate of 0.3 µl/min. Gradients from 2% to 30% solvent B for 30 min followed by 30% to 75% solvent B from 30 to 35 min and 75% to 90% solvent B from 35 to 70 min were used. Thereafter, the gradient was kept at 90% solvent B from 70 to 75 min, using 0.1% FA in 3% ACN as solvent A and 0.1% FA in 97% ACN as solvent B (FA: LC-MS grade, Fluka; ACN: LC-MS grade, Merck). The column was operated at 60 °C. The mass spectrometer was operated in data-dependent acquisition mode with automatic switching between MS and MS/MS scans.

The full MS scans were acquired at 70K resolution, with automatic gain control target of $1 \times 10^6$ ions, maximum injection time of 200 ms and the scan range was 300-1800 m/z for MS scans. Higher energy collision dissociation (HCD) was used for peptide fragmentation with normalized collision energy set to 28. The MS/MS scans were performed using a data-dependent top10 method at a resolution of 17.5K with an automatic gain control target of $5 \times 10^4$ ions at maximum injection time of 100 ms and isolation window of 2.0 m/z units. An underfill ratio of 10% and dynamic exclusion duration of 30 s was applied.

**Protein and PTM identification**

Protein and PTM site identification from the raw MS data was performed by using the MaxQuant software with an integrated Andromeda search engine (version 1.5.7.4) [35-36]. The raw mass spectral data were searched against the Uniprot Mtb protein database (downloaded from http://www.uniprot.org/ on Jan 15, 2017, uniprot ID: UP000001584, Organism/Taxonomy ID: 83332 and with 3993 protein sequences) concatenated to reverse decoy database and protein sequences for common contaminants.

Trypsin/P was specified as a cleavage enzyme allowing up to two missed cleavages. The "re-quantify" and "match between runs" options were utilized with a retention time alignment window of 3 min. Dependent peptide search, second peptide, LFQ and iBAQ were enabled. Carbamidomethylation on cysteine was set as the fixed modification and acetylation on protein N-terminal, conversion of N-terminal glutamine and glutamic acid to pyroglutamic acid and oxidation on methionine were set as the variable modifications. For the PTM analysis, acetyl (KSTY) was set as the variable modification. Unique and razor peptides were used for the quantification of modified peptides (PTM abundance).

Only peptides with a minimum length of seven amino acids and detected in at least one or more of the replicates were considered for identification. For protein identification, a minimum of two peptides, of which at least one was unique, was required per protein group. The threshold of protein identifications were determined by false discovery rate (FDR) of 0.01. All other parameters in MaxQuant were set to default values.

All modified peptide spectra were validated by applying stringent site localization probability of $\geq 0.75$ and PEP of $\leq 0.01$ prior to further analysis. PTM site identifications with localization probability $< 0.75$ and PEP $> 0.01$, protein groups with matches to proteins from the reversed database, and contaminant protein sequences were removed from the analysis. Modified peptides with quantifiable values in at least five of nine biological replicates in the

three strains were considered for label-free relative quantification. Acetylated peptide intensities were used for quantification of PTM abundance.

## BIOINFORMATICS ANALYSIS

### Statistical Analysis

Bioinformatics analysis was performed using the Perseus software (version 1.5.6.0) as previously described [37]. The protein group output from MaxQuant was used as the basis for all the subsequent statistical and Gene Ontology (GO) enrichment analysis. Following protein identification by a MaxQuant database search, validation for multiple comparisons was corrected using the Benjamini-Hochberg correction [38]. For identification of significantly changed acetylation sites between the two Mtb lineages, a two tailed unpaired student`s T-test with FDR $\leq$ 0.05 and S0=2 was applied.

### GO and Pathway Enrichment Analysis of Acetylated Proteins

The GO annotation of identified modified proteins was derived from the DAVID Bioinformatics Resources 6.7 and Gene Ontology Consortium bioinformatics databases [39-40]**.** The proteins were classified by GO annotation based on three terms; molecular function (MF), biological process (BP) and cellular component (CC). Acetylated proteins were classified based on their functional category using TubercuList Mtb database (http://tuberculist.epfl.ch/) [41-42]. The Kyoto Encyclopedia of Genes and Genomes (KEGG) was utilized to annotate the pathways [43]. The enriched GO terms and KEGG pathways provided corresponding information on p-value, count, percentage and fold enrichment. Any pathway, biological process or molecular function with a p-value $\leq$ 0.05 was considered as significantly enriched. The ScanProsite web-based tool (http://prosite.expasy.org) and GenomeNet Database Resources (http://www.genome.jp/tools/motif/) were used to identify PROSITE signature profiles and active site motifs that match the sequence of identified acetylated peptides [44].

## PTM Motif Analysis

Motif-X software version 1.2 [45] (http://motif-x.med.harvard.edu/motif-x.html) was used to analyze the enrichment of amino acid sequence motifs at specific positions of acetyl-31-mers (15 amino acids upstream and downstream of the site) in all peptide sequences. All protein sequences in the database were used as the background database parameter and other parameters were set at default.

## Protein-Protein Interaction Network Analysis

Protein-protein interaction (PPI) networks were generated and analyzed using STRING database version 10 (http://string-db.org/) [46] with a high confidence threshold of 0.7. Highly connected clusters were identified using MCODE plug-in toolkit and the interaction network was visualized using the Cytoscape software (http://www.cytoscape.org) (version 3.5.0) [47].

## Ethical Approval

The study obtained ethical approval from the Regional Committee for Medical Research Ethics in Eastern Norway (REK Øst) and Ethiopian Science and Technology Ministry in Addis Ababa, Ethiopia. Sample collection was conducted after obtaining written informed consent.

# RESULTS AND DISCUSSION

## The First Combined $N_\varepsilon$- and O- Acetylome Map of Mtb

In this study we analyzed the $N_\varepsilon$- and O-acetylome profile of two Mtb clinical isolates from lineage 7 strain and the lineage 4 reference strain H37Rv. A total of 2490 class-I acetylation sites, 2200 and 2198 acetylation sites in lineage 7 and lineage 4 , respectively(Figure 1A, 1B, Tables 1, SI 1, SI 2 and SI 3). These sites matched to 1568 proteins (953 unique proteins) (Table SI4). A representative spectra of two acetylated peptides from heparin-binding hemagglutinin (HbhA) and conserved hypothetical protein

(Rv2020c) are shown in Figure SI 1 (Figure SI 1). Notably, 2310 (92.77%) of the totally identified acetylome was shared between lineage 7 and H37Rv, highlighting the acetylome conservation and consistency of the acetylome data (Figure 2A). Among the 953 unique proteins, 44.28% of the proteins were singly acetylated while the remaining 55.82% of the proteins were acetylated at two or more sites (Figure 2C). The most heavily acetylated proteins included the 60 kDa chaperonin 2 GroEL2, mycolipanoate synthase Msl3, polyketide synthase PKS13, ATP synthase subunit beta AtpD, chaperone protein DnaK and catalase-peroxidase KatG (Figure 2C, Table SI 4). These heavily acetylated proteins are known to be involved in stress responses, corroborating previous data indicating a role for protein acetylation in stress adaptation in Mtb and other bacteria [17, 48-49].

**O-Acetylation at Serine, Threonine or Tyrosine Represents 94.34 % of the Sites Identified on 88.56 % of the Acetylated Proteins**

Out of a total of 2490 acetylation sites identified, 2349 sites on 844 proteins were found to be O-acetylated. The abundance of acetylation was highest on T residues (39.32%), followed by S (36.39%), Y (18.80%) and K residues (5.66%) (Figure 2B, Table 1). The proportion of O-acetylated residues in Mtb is similar to the phosphorylation profile on the three residues, S, T and Y [50]. The acetylated proteins identified accounted for 23.87% of the total proteins annotated in Mtb, which may indicate that applying an enrichment-based method would provide an even higher number of O-acetylated proteins than what has been reported for the $N_{\varepsilon}$-acetylation [15].

$N_{\varepsilon}$-acetylation is the most commonly studied acetylation both in eukaryotes and prokaryotes [4]. A recent study on the Mtb acetylome, using an anti-acetyllysine antibody enriched sample, reported a total of 1128 lysine acetylation sites on 658 proteins [15]. Although there is no former evidence for the presence of protein O-acetylation in prokaryotes, it is known that O-acetylation is a common modification of bacterial peptidoglycan and other

glycoconjugates [31-32]. Serine O-acetyltransferase in plants and bacteria plays a role in the biosynthesis of cysteine from serine [30, 51]. The O-acetyltransferase, N-hydroxyarylamine O-acetyltransferase (NhoA), has been shown to have an O-acetyltransferase activity in *Salmonella typhimurium* [52]. Based on the evidences discussed so far, including the YopJ O-acetyltransferase, we propose that the mechanism of O-acetylation and deacetylation on S, T and Y residues in Mtb probably follows a similar pattern as the $N_{\varepsilon}$-acetylation involving both enzymatic and non-enzymatic mechanisms and AcCoA as acetyl group donor. Further investigation is needed to elucidate the responsible mechanisms.

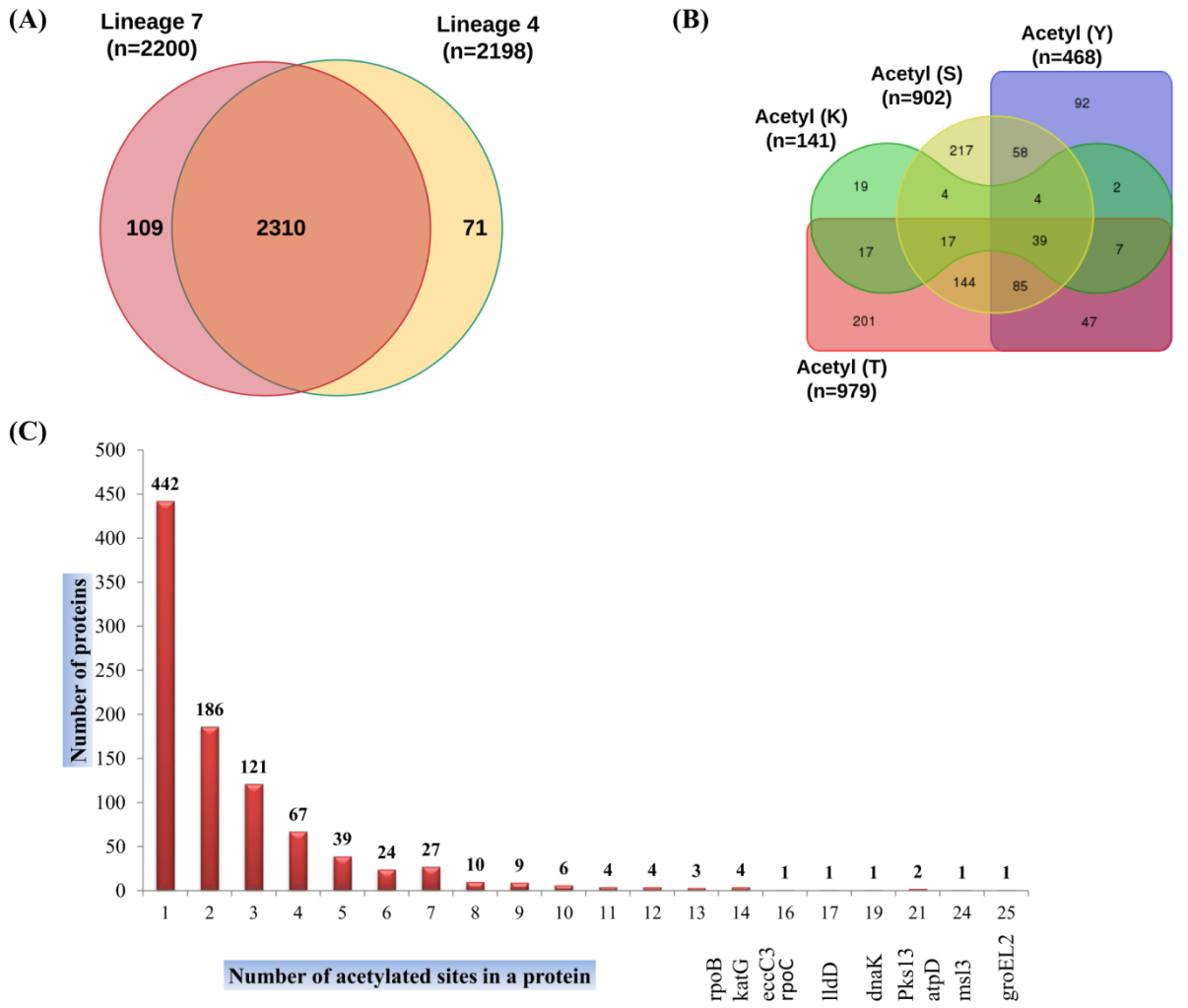



**Figure 1** | (A) Overall difference and overlap in acetylated sites between *Mycobacterium tuberculosis* strains of lineage 7 and lineage 4 (H37Rv). (B) Overall difference and overlap in acetylated proteins between the four residues, S, T, Y and K. "n" in (A) and (B) stand for number of acetylation sites. (C) Distribution of acetylated proteins based on the number of acetylated siteses per protein (Table SI 1,2,3).

**Table 1 |** Number of acetylation sites per residue in *Mycobacterium tuberculosis* proteins and the number and percentage of unique proteins acetylated.

| **Acetylated residues** | **Number of sites** | **Unique proteins** | **Percent (%)** |
|---|---|---|---|
| Acetyl (Lysine, K) | 141 | 109 | **5.66** |
| Acetyl (Serine, S) | 902 | 568 | **36.22** |
| Acetyl (Threonine, T) | 979 | 557 | **39.32** |
| Acetyl ( Tyrosine, Y) | 468 | 334 | **18.80** |
| **Overall unique elements** | **2490** | **953** | **23.87** |

Bioinformatics analysis showed that these proteins share similar functions with $N_{\varepsilon}$-acetylated proteins, as evident from previous reports on lysine acetylation [15, 17]. Even though there was a relatively low coverage of acetylated proteins in this enrichment-free method, a separate enrichment analysis for 109 lysine acetylated proteins and 840 exclusively O-acetylated proteins provided a distinct profile of biological processes and KEGG pathways associated to a particular acetylation. Exclusively O-acetylated proteins were involved in a broad range of KEGG pathway and biological processes, including fatty acid and carbon metabolism, translation, biosynthesis of secondary metabolites and antibiotics, amino acid metabolism, glyoxylate and dicarboxylate metabolism and microbial metabolism in diverse environments (Figure 2A). In contrast, proteins acetylated at K residues were found to be involved in limited KEGG pathways and biological processes, primarily translation and the citrate cycle (TCA cycle) (Figure 2B).

These data may indicate that the stoichiometry of O-acetylation is more plentiful than $N_{\varepsilon}$-acetylation. This may support the assumption that O-acetylation is involved in regulating a myriad of biological processes beyond that of $N_{\varepsilon}$-acetylation. Furthermore, the number of O-acetylated peptides significantly outweighs the $N_{\varepsilon}$-acetylated peptides. These frequent occurrences together with their competitive inhibition of phosphorylation may broaden the role of O-acetylated proteins in regulating bacterial physiology [49, 53]. Further enrichment-

dependent methods for each of the four residues is necessary to complete the catalogue of $N_{\varepsilon}$- and O-acetylated proteins and their respective enriched KEGG pathways.

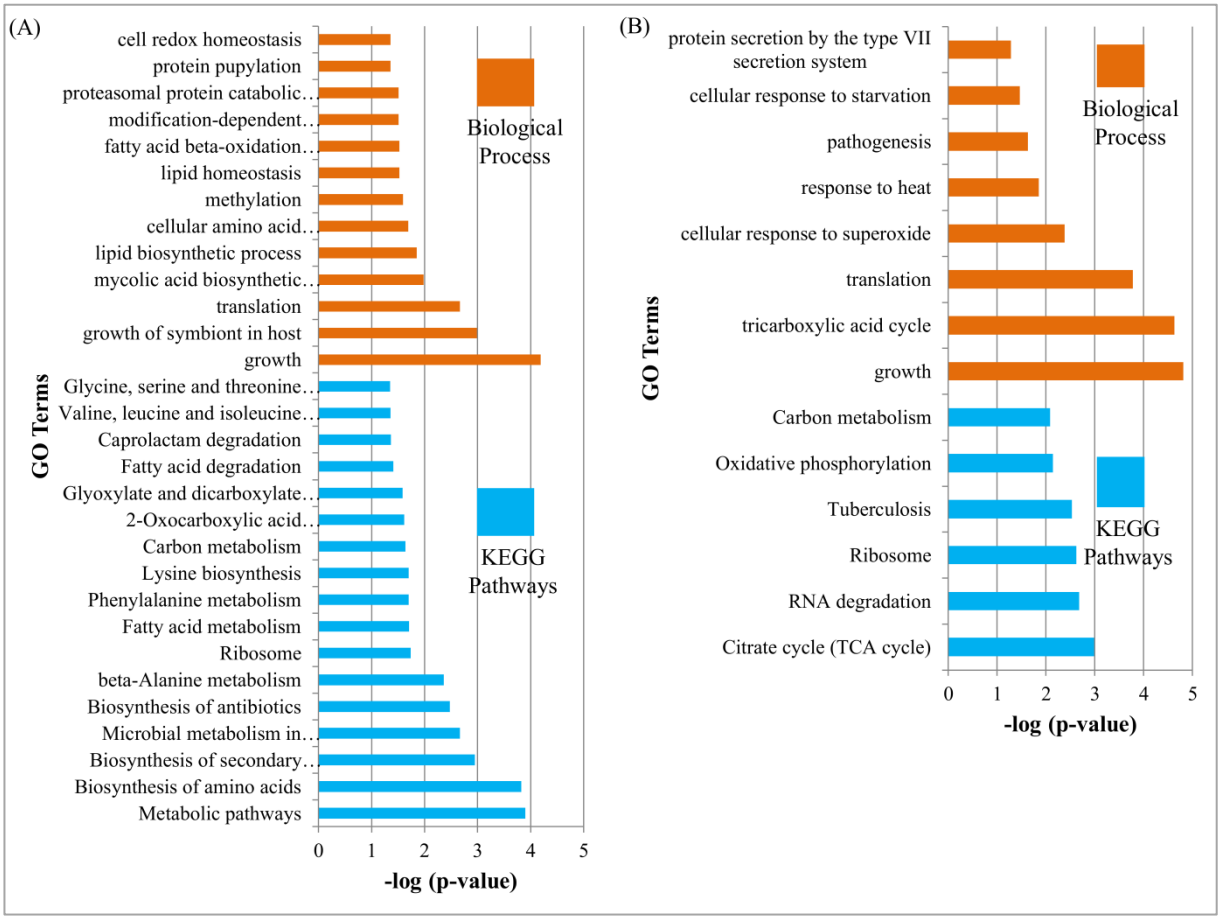



**Figure 2 |** GO enrichment analysis for biological process and KEGG pathways in exclusively O-acetylated *Mycobacterium tuberculosis* proteins (A) and in $N_{\varepsilon}$-acetylated Mtb proteins (B).

**Acetylated Peptides were Identified Inside or Near PROSITE Signature Motifs**

MS-derived sequence windows data was used to search the PROSITE signature motifs, active sites, and domain profiles. Some of the acetylated peptides identified were found near or inside enzyme active site domains (Table SI 5-1, Table SI 5-2). The polar residues histidine, cysteine, aspartate, glutamate, arginine, lysine, tyrosine, serine, threonine, asparagine, glutamine and tryptophan are the most frequent catalytic amino acid residues [54-55]. S, T and Y are the major substrates for protein kinases that are hubs for complex regulatory networks, and are involved in blocking phagosome-lysosome fusion, the hallmark of Mtb pathogenesis (reviewed in [28-29]). The interplay between acetylation and phosphorylation of the same

residues may have implications in the fine-tuning of certain cellular processes including Mtb pathogenesis. As a result, acetylation of major active site residues, including K, S, T and Y may affect protein activity [27, 56].

Acetylation of K residues found within an enzyme active site neutralizes the positive charge on lysine residues, which results in concomitant alteration in protein activity [57], protein-protein and protein-DNA interactions [58-60], local protein conformation [61], and protein localization [62-63]. Protein acetylation increases the net negative charge [62, 64] on DNA binding proteins that might inhibit their binding capacity of the positively charged lysine residues with the negatively charged phosphate groups on DNA as reviewed by Carabetta *et al.* [5]. The inhibitory effect of acetylation on K residues found in the active sites of enzymes ACS, NhoA, adenosylmethionine synthase (MAT), and MbtA is highlighted in this review.

In our study we found that the pyruvate kinase active site signature (Pyk) was one of the prosite motifs identified to be acetylated at position K221 (Table SI 5-2). The lysine residue in the Pyk active site seems to be the acid/base catalyst responsible for the interconversion of pyruvate and enolpyruvate. In accordance with our finding, mutagenesis of the active site K221 of the pyruvate kinase was shown to reduce the activity of this enzyme by a factor of $10^4$ to $10^5$ in *Bacillus stearothermophilus* [65]. Even though the effect of O-acetylation on bacterial enzyme activity has not been investigated to date, it is possible that direct acetylation of such active site residues may modulate, abolish or induce the enzyme activity [49] or interfere with the phosphorylation event [53].

**The Positively Charged Lysine and Arginine Residues are Enriched Towards the N-Terminus of the Acetylation Sites**

It is highly likely that acetylation events follow conserved linear protein sequence motifs similar to the motifs observed in protein phosphorylation. Two conserved putative

motifs were identified for acetylated peptides on K residues, namely RKac and R*Kac at different abundances (Figure 3A, 3B).

The acetylation motifs identified in both $N_{\varepsilon}$- and O-acetylated residues were consistent, having the positively charged K and/or arginine (R) residues between -1 and -6 positions to the N-terminus of the acetylated residues and L at -1 or -3 positions in some of the acetylated S and T residues (Figure 3A, 3B).

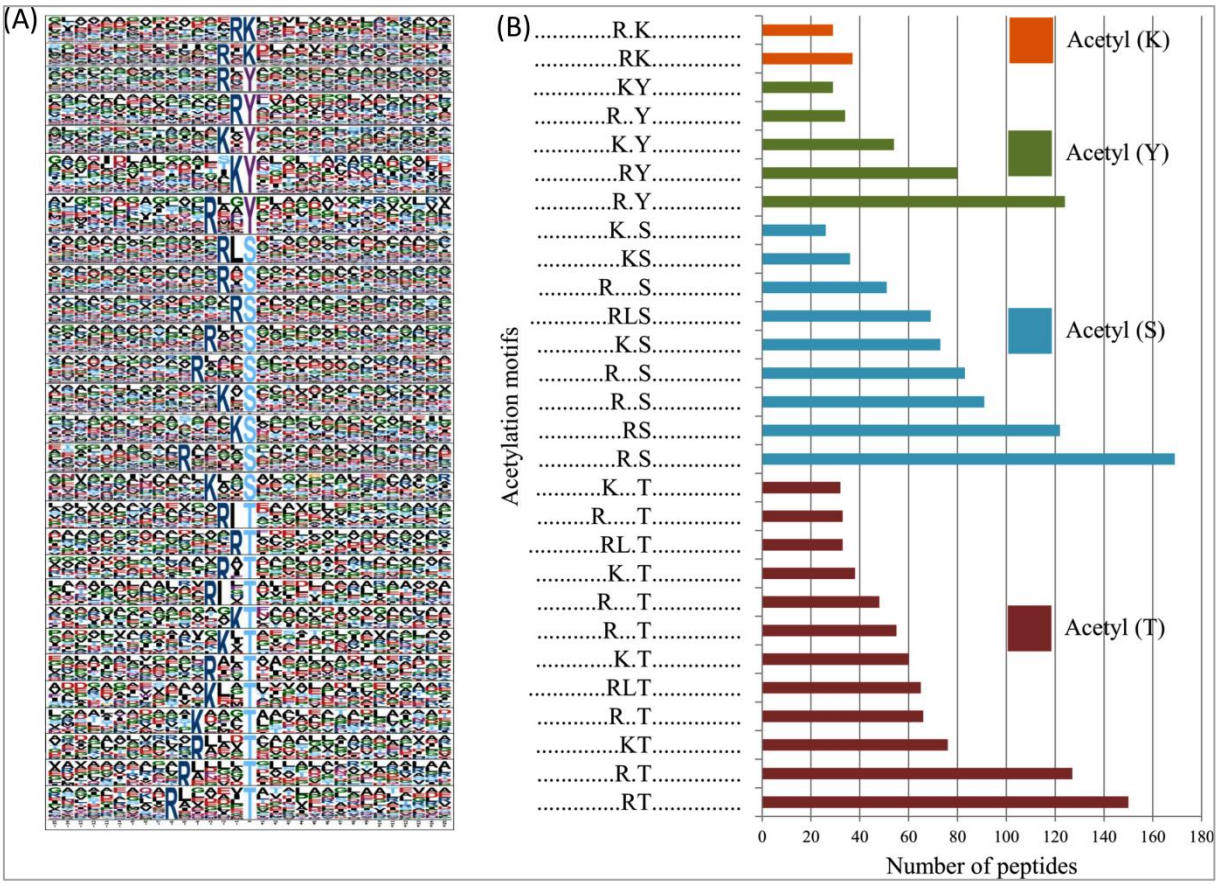



**Figure 3** | (A) *Mycobacterium tuberculosis* protein acetylation motifs and conservation of acetylation sites. (B) Number of identified peptides contained in each conserved motifs for the four residues.

Previous studies on lysine acylation showed that the positively charged K and R residues were significantly enriched residues at the C-terminus [15, 66], N- and C-terminus [67] and N-terminus [49] of the acylation sites. This variation may be attributed to the difference in the methods used in the different studies, for example in antibodies used for immunoaffinity enrichment [5]. These motifs may serve as a recognition signature for putative bacterial

acetyltransferases and deacetylases as observed in eukaryotes, or could be part of an autocatalytic mechanism that facilitates the non-enzymatic acetylation process, thereby regulating substrate specificity, enhancing acetyltransferase activity, and restricting access to non-target proteins [5, 68].

**Acetylated Peptides Identified are Involved in Diverse Cellular Processes**

Among the 953 acetylated Mtb-proteins identified, 78.5% were annotated for KEGG pathways, 92.0% for biological processes and 94.3% for molecular functions, while only 35.1% were annotated for cellular components.

The predicted subcellular location of the acetylated proteins showed that most of the acetylated proteins identified were related to the cytoplasm (60%), while a few proteins were predicted to be membrane associated (20%), in macromolecular complexes (10%), associated to organelles (8%), or located in the extracellular region (2%) (Figure 4A). The GO analysis of biological processes and molecular functions shows that large numbers of acetylated proteins are enzymes (73%) involved in metabolism (60%), respectively (Figure 4B, 4C). Binding proteins are the second largest acetylated protein group in terms of molecular function, accounting for 15% of the total number of acetylated proteins identified (Figure 4B). Proteins associated with cellular processes represented the second largest protein group in terms of biological process covering 24% of all annotated proteins (Figure 4C).

The acetylated proteins identified were grouped based on TubercuList functional categories (http://tuberculist.epfl.ch/). The percentage was calculated by dividing the number of acetylated proteins in each group by the total number proteins known to function in a particular category (Figure 4D) [41]. Proteins involved in information pathways had the most abundant acetylation (105/241, 43.57%), followed by proteins involved in lipid metabolism (106/247, 42.91%) and intermediary metabolism and respiration (373/923, 40.41%).

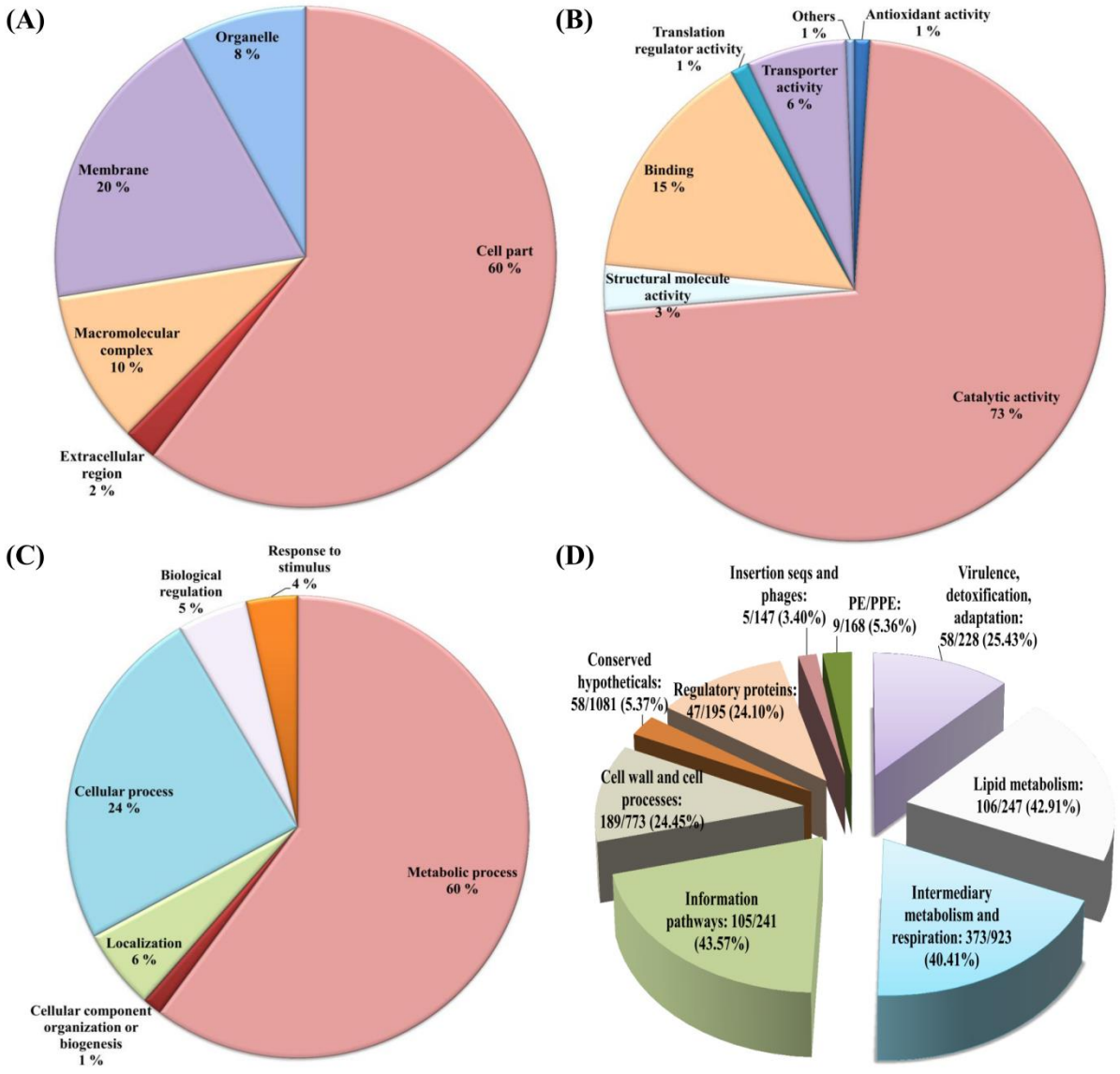



**Figure 4 | Gene Ontology functional classification of the *Mycobacterium tuberculosis* acetylated proteins identified.** (A) Subcellular localization of the acetylated proteins. (B) Classification of the acetylated proteins based on molecular function. (C) Classification of the acetylated proteins based on biological process. (D) Percentage of acetylated proteins within their respective functional category.

The GO enrichment analysis and protein-protein interaction analysis (PPI) showed that proteins involved in translation and the structural constituents of ribosomes are the most abundant among the significantly enriched acetylated proteins in terms of biological processes and molecular function, respectively (Figure 5, Table SI 6-1, Figure SI 2 ). Despite the

variations in the methods used and the acetylated residues analysed, this finding was in agreement with a previous study on lysine acetylation and succinylation [15, 17, 49].

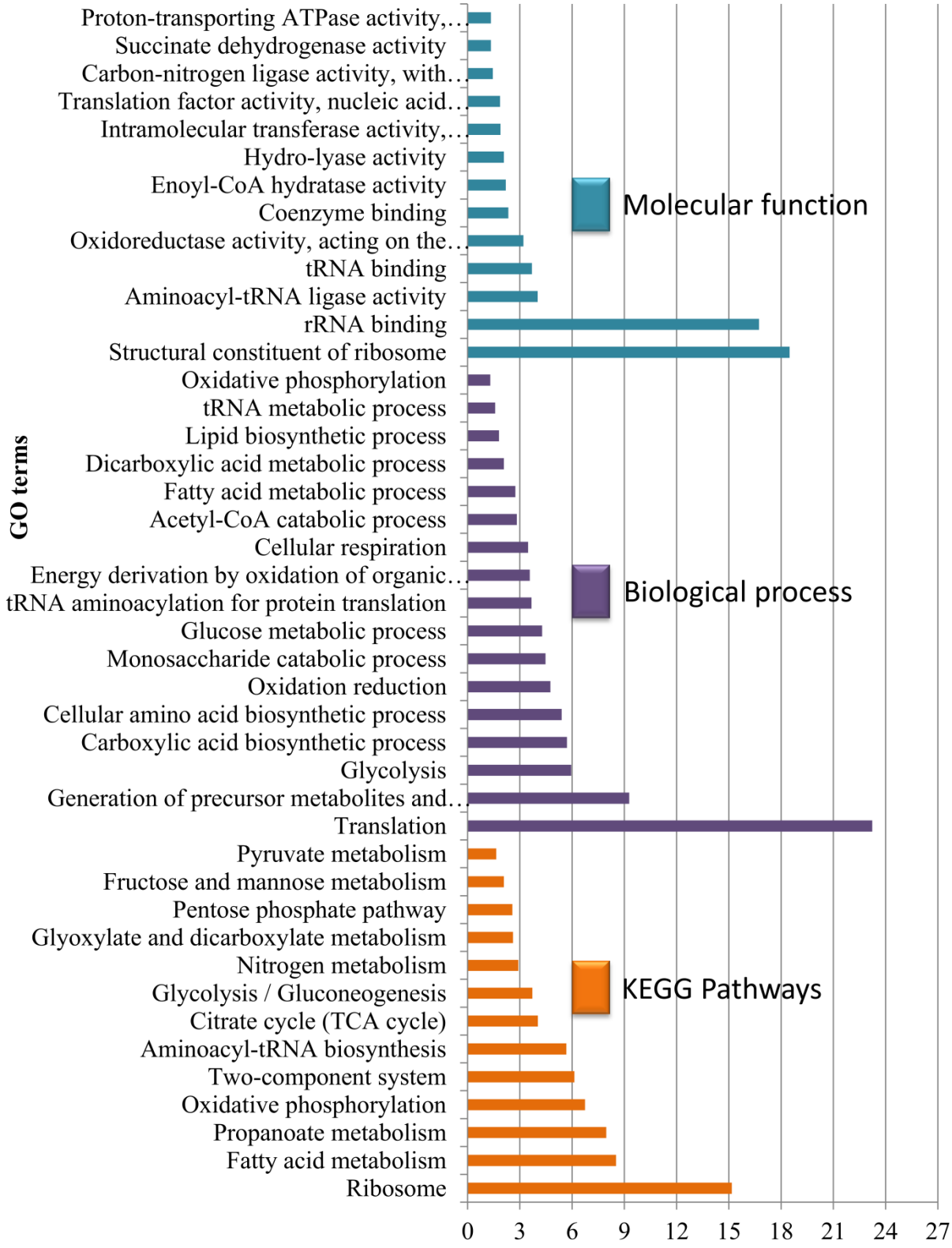



**Figure 5** | Gene Ontology enrichment analysis of identified acetylated *Mycobacterium tuberculosis* proteins based on molecular function, KEGG pathways and biological process.

It has been shown that lysine acetylation regulates cellular metabolism via different mechanisms such as enzymatic activation or inhibition, and by influencing protein stability [17, 69-70]. We identified eight acetylated enzymes involved in the TCA cycle [69, 71] and two enzymes, isocitrate lyase (ICL) and malate synthase G (GlcB), involved in the glyoxylate pathway. Both copies of the ICL genes encoding AceaA and AceaB are essential for survival of Mtb *in vivo* [72-74]. In *E. coli*, it has been shown that His356 is one of the catalytic active site residues in ICL [75]. Furthermore, Wang *et al*. have shown that acetylation of AceA with protein acetyltransferase (Pat) or acetylation-mimicking mutations has led to reduction in AceA activity and its activity was restored by deacetylation via CobB [69]. We found that AceA and Aceb were acetylated at 2 and 5 sites, respectively. AceAb was acetylated at S355, near the catalytic residue H356, indicating that acetylation of this residue may lead to a conformational change in the protein and therefore affect its enzymatic activity. GlcB is also a virulence factor involved in Mtb adherence to lung epithelial cells[76]

ACS is another acetylated protein involved in the synthesis of AcCoA, a key intermediate in energy metabolism and an acetyl group donor in protein acetylation. ACS was the first enzyme in Mtb shown to be regulated by reversible post-translational acetylation via cAMP-dependent protein acetyltransferase [11-12]. The acetylation status of ACS determines the activity that might influence the availability of the acetyl donor, AcCoA, and the metabolic state of the cell [77].

The glyoxylate cycle is another pathway enriched by acetylated proteins. When the TCA cycle is down-regulated upon oxygen and nutrient/glucose depletion, replenishment of TCA cycle intermediates is achieved via the glyoxylate cycle using the AcCoA from fatty acid β-oxidation as a carbon source for subsequent metabolic pathways in the synthesis of biomolecules (glucose, amino acids, DNA, and RNA) [78-80]. Protein acetylation has been shown to regulate the activity of enzymes controlling the direction of glycolysis versus

gluconeogenesis and the branching between TCA cycle and glyoxylate cycle by Pat and deacetylase [69].

Fatty acid metabolism was one of the pathways identified by KEGG pathway enrichment analysis in this study. In addition to being a source of AcCoA, fatty acids are an integral component of the Mtb cell wall and known to be related to Mtb pathogenicity [81]. It has been shown that reversible protein acetylation can regulate the activity of a number of fatty-acid-CoA ligases in Mtb [6, 70]. These findings suggest that acetylation may play a role in the regulation of various cellular processes in Mtb. Additional functional studies are needed to validate these claims.

**Very few Mtb Proteins Involved in Genome Maintenance are Acetylated**

Most components involved in DNA repair, recombination and replication (3R components) were not acetylated even though they are located in the cytosol among core metabolic enzymes. The only 3R components found to be acetylated were DNA gyrases TopA, GyrA and GyrB, single-stranded binding protein SSB, nucleotide excision repair DNA damage sensor UvrA, and recombination factors RecB and RecF. Ku ligase and DnaA were by Nε-lysine acetylation enrichment previously found to be acetylated in Msm and *E. coli,* respectively [82-83]. Thus, along with all surface components, most 3R enzymes were found to be constituents of the non-acetylated complement of Mtb cells (Table SI 4).

**Proteins involved in Antimicrobial Drug Resistance are Acetylated in Mtb**

Several bacterial species, including Mtb, alter their proteins involved in drug resistance or drug targets, which in turn decreases or blocks the affinity for drug binding without affecting normal activity [49]. PTMs may alter the net charge on the protein, conformation, interaction and activity, thereby modulating the bacterial response to drugs. Acetylation of kinase substrates might alter the signaling pathways that lead to drug resistance/sensitivity. Seven Mtb proteins associated with isoniazide (INH) resistance,

including KatG, InhA, NdhA KasA, AhpC, FadE24, and AcpM, were acetylated (Table SI 4) [49, 84]. KatG, a catalase-peroxidase enzyme, is responsible for peroxidative activation of the prodrug INH and acts as a virulence factor to protect against oxidative stress [85]. Mutations at KatG positions S315T and R463L have been shown to diminishe its capacity to activate INH and confer INH resistance to Mtb [86]. An acylation study showed that succinylation of KatG at K310 near the S315T mutation assists the enzyme in retaining its native antioxidant activity, while the INH activating property was reduced by almost 30% and the minimum inhibitory concentration of bacteria increased up to 200-fold [49]. Among the 8 differentially acetylated sites on KatG, S465 was found near the natural mutation R463L. Acetylation of this residue may lead to a conformational change in the protein and therefore affect its activity.

Serine/threonine protein kinases (STPKs) and two-component signal transduction systems are key regulators of metabolic processes, including transcription, cell development, stress response, virulence, host-pathogen interactions and drug resistance [87-88]. Four of the eleven Mtb STPKs, PknD, PknK, PknG and PknH, were acetylated at various residues. OpcA and Wag31, have been shown to be up-regulated in INH-resistant Mtb strains [49, 89]; OpcA and Wag31 are involved in peptidoglycan biosynthesis and oxidative stress responses [90-91]. Another protein, PpiA, is involved in cationic antimicrobial peptide (CAMP) resistance [92]. MurF is a protein involved in cell wall synthesis and implicated in vancomycin resistance in Mtb [93]. Thus, acetylation of a protein in an active site residue or anywhere in the protein sequence may alter the protein function in various modes.

InhA, NADH-dependent enoyl-ACP reductase, is a major enzyme involved in the biosynthesis of mycolic acids. Mutation at InhA I74T has been associated with resistance to ethambutol (EMB), INH, rifampicin (RMP), and streptomycin (SM) [94]. We identified an acetylation site on InhA at position T79 which is only five amino acids away from the natural

mutation I74T. Other acetylated proteins with a role in resistance to first-line anti-TB drugs include RpoB, EmbR, PhoP, FabG1 and RpsL [84, 95].

Mutations in the genes encoding DNA gyrase subunits, *gyrA* and *gyrB*, are the most common mechanisms for acquiring fluoroquinolone (FQ) resistance in Mtb [96]. The most frequent FQ resistance-associated mutations, termed the quinolone resistance-determining region, resides between codons 74 to 113 in *gyrA* and between 461 to 538 in *gyrB* [96-98]. We found two acetylation sites at T500 and S473, which are located within the quinolone resistance-determining region of GyrB, and may play a role in drug resistance, DNA replication, and Mtb survival. Another acetylated proteins involved in drug resistance is enhanced intracellular survival (Eis), an acetyltransferase, that confers resistance to kanamycin by modifying the drug [99-100].

**Proteins Associated to Virulence, Growth and Stress Responses are Differentially Acetylated Between Lineage 7 and Lineage 4 Strains**

From a total of 2490 acetylation sites identified on 953 proteins, 1085 sites on 506 proteins were eligible for quantification. We found that 261 acetylation sites on 165 proteins were differentially acetylated between lineage 7 and lineage 4 at $S_0$=2 and FDR $\leq$ 0.05 (Figure 6A, 6B). Interestingly, 257 sites on 160 proteins involved in Mtb growth and virulence were hypoacetylated in lineage 7 with fold changes between 4.2 and 628.4. Only four sites on four proteins were significantly hyperacetylated in lineage 7 strains. Lineage 7 is a recently identified lineage of Mtb, characterized by slow-growth and reduced virulence phenotypes [101-102]. The GO enrichment analysis of 160 proteins hypoacetylated in lineage 7 revealed that pathogenesis, growth, glycolysis, response to iron ion, response to nitrosative stress and protein folding were among the significantly enriched biological processes (Figure 7A, 7B). Carbon metabolism, TCA cycle, oxidative phosphorylation and microbial

metabolism in diverse environments were some of the significantly enriched pathways from these hypoacetylated proteins (Figure 7B, Table SI 6-2).

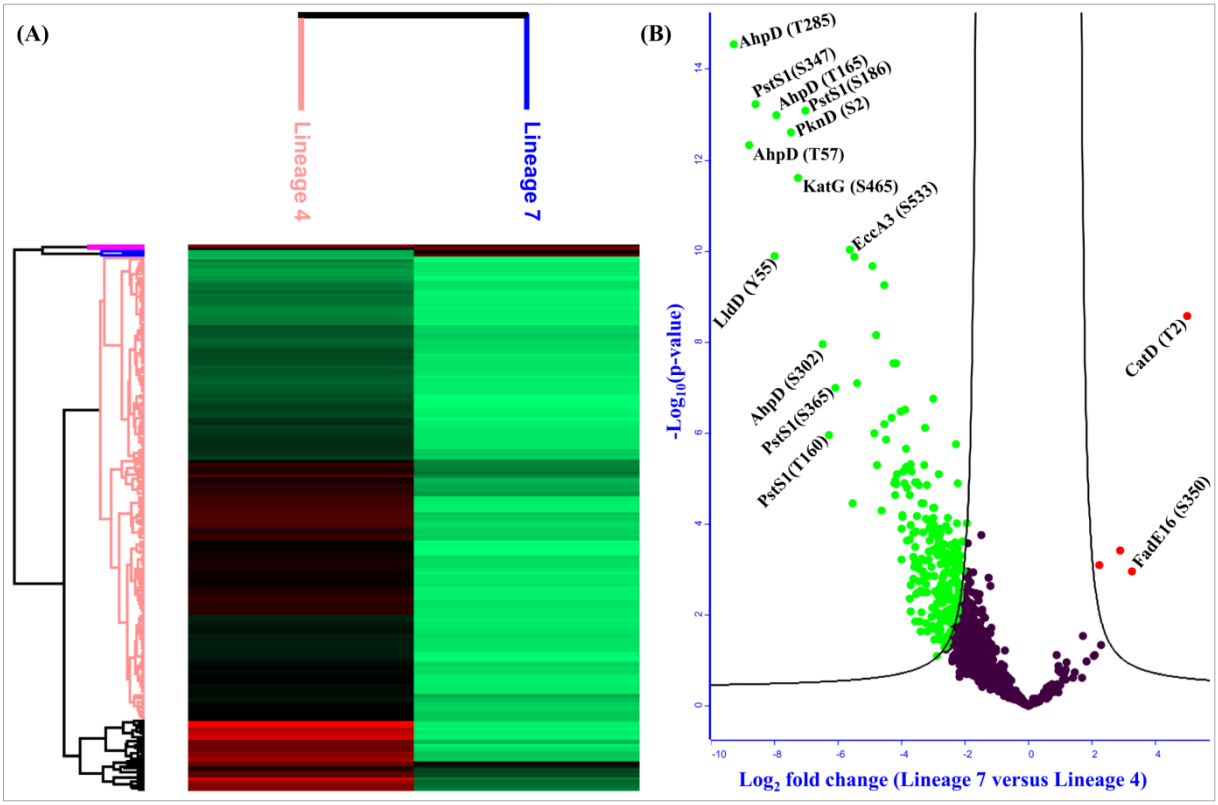



**Figure 6** | (A) Hierarchical clustering of differentially acetylated proteins between lineage 7 and lineage 4 strains of *Mycobacterium tuberculosis* (Mtb). (B) Volcano plot of differentially acetylated proteins between lineage 7 and H37Rv (So=2, FDR ≤ 0.05; -Log Student's T-test p-value in Y-axis and Student's T-test Difference in X-axis ). Red: hyperacetylation and Green: hypoacetylation.

A number of enzymes involved in carbon metabolism, fatty acid metabolism, stress response, growth, virulence and the Esx-3 secretion system were hypoacetylated in lineage 7 (Figure 7A, 7B, Table 2, Table SI 7). The dihydrolipoyllysine-residue acetyltransferase component of the pyruvate dehydrogenase complex (DlaT) and aconitate hydratase A (AcnA) are two proteins involved in the TCA cycle. Both enzymes were found to be hyperacetylated at lysine residues in lineage 4 strains. AcnA was acetylated at K273 position with fold change of 7.83, whereas DlaT was acetylated at two positions, K273 and K287, with fold changes of 8.27 and 6.66, respectively. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and enolase (Eno) are two hypoacetylated enzymes involved in glycolytic pathway. Eno was

592 acetylated at 5 sites with fold changes between -8.14 to 15.59 , while GAPDH was acetylated
593 at position S259 with a fold change of 7.03 (Table 2, Table SI 7).

594 Energy metabolism is associated to growth and virulence in intracellular bacteria,
595 including Mtb [103-104]. Central carbon metabolism uses different carbon sources to generate the
596 building blocks, cofactors, and energy for cell growth. Wang *et al*. have shown that enzymes
597 involved in *Salmonella enterica* central carbon metabolism are regulated by reversible lysine
598 acetylation, involving protein acetyltransferase (Pat) and deacetylases (CobB) [69]. In *S.*
599 *enterica*, acetylation of GAPDH has been shown to favor the glycolytic pathway while
600 inhibiting gluconeogenesis by more than 30% while deacetylation of GAPDH by the sirtiun
601 CobB stimulates gluconeogenesis and inhibits glycolytic pathway [69]. Furthermore, it has been
602 shown that *S. enterica* deficient in *cobB* (with high acetylation) grew faster than the wild-type
603 cells in minimal glucose medium but grew slower than the wild type in minimal citrate
604 medium. This indicates the importance of protein acetylation in regulating bacterial growth [69].
605 The detailed mechanism of the regulation needs to be investigated. The activity of pyruvate
606 dehydrogenase, one of the hypoacetylated enzymes involved in the TCA cycle, is reduced by
607 lysine acetylation in eukaryotes [105-106], and this might also be true in bacteria. ACS was found
608 to be hypoacetylated in lineage 7 strains with a fold change of 5.44. AcCoA is the substrate
609 for the glyoxylate cycle [107]. The glyoxylate cycle has been shown to be up-regulated during
610 Mtb growth arrest and involved in Mtb persistence [108-109]. Moreover, previous works indicated
611 the possible role of ACS in controlling the expression and/or activity of the glyoxylate cycle
612 in *E. coli* [110-111]. Therefore, the acetylation status of Mtb ACS might have a role in modulating
613 Mtb physiology and persistence [11, 107]. Mtb utilizes lactate as a potential carbon and energy
614 source inside macrophages [112]. L-lactate dehydrogenases (LldD), an enzyme involved in the
615 oxidation of lactate, was hypoacetylated in lineage 7 with fold change of -258.83.

AcCoA can be generated from different substrates, including glucose, fatty acids, amino acids and citrate. Once generated, it is used in ATP synthesis via TCA cycle, synthesis of fatty acids, amino acids and other metabolites (reviewed in [77]). In addition, AcCoA is an acetyl group donor for protein acetylation. Thus, acetylation of enzymes involved in AcCoA synthesis, ACS, may also modulate all the processes involving AcCoA. These processes also consume AcCoA and thereby alter the availability of AcCoA in the cell and metabolic state of the cell.

In addition to the enzymes involved in TCA and glyoxylate cycles, a number of long-chain-fatty-acid-CoA ligases associated with fatty acid metabolism were hypoacetylated in Mtb lineage 7 strains with fold changes ranging from 4.54 to 14.35 (Figure 7A, Table 2, Table SI 5). Enzymes involved in fatty acid β-oxidation were hypoacetylated in lineage 7 (Tables 2 and SI 7). Fatty acid metabolism, both synthesis and catabolism, is an important cellular process for Mtb fitness and survival. More importantly, fatty acids and their derivatives are an integral component of the cell wall complex and implicated in Mtb pathogenicity, fitness and survival [17, 81, 113]. Mtb utilizes fatty acids as principal source of energy during dormancy and reactivation[104, 114]. Survival of Mtb inside the phagolysosome depends on the pathogen's ability to synthesize virulence factors (proteins) and other biomolecules in a glucose-limited stress environment [80]. Reversible acetylation has been shown to modulate the activity of several fatty-acid-CoA ligases [70]. Polyketide synthase (PKS13) was hypoacetylated in lineage 7 at 6 sites with fold changes ranging from -3.89 to -5.98. PKS13 is an enzyme involved in the final steps of mycolic acid biosynthesis. Mycolic acids are an integral component of Mtb cell wall and known to be related to its pathogenicity [81].

ESX-3 secretion system components were found to be highly acetylated in H37Rv. Esx-3 is implicated in essential physiologic processes and metal homeostasis and crucial for Mtb growth *in vivo* and *in vitro* [115]. The phosphate-binding protein PstS 1 is involved in

641 inorganic phosphate uptake and its disruption has been shown to be associated to decreased
642 virulence and attenuated growth [116]. PstS 1 was hypoacetylated at 6 sites in lineage 7 with
643 fold changes between -388.45 and -27.53.

**Table 2 |** List of proteins hypoacetylated in *Mycobacterium tuberculosis* lineage 7 strains.

| Protein name | Gene ID | Acetylation | Fold change |
|---|---|---|---|
| Alkyl hydroperoxide reductase AhpD | Rv2159c | T285 | -628.43 |
| Alkyl hydroperoxide reductase AhpD | Rv2159c | T57 | -450.62 |
| Phosphate-binding protein PstS 1 | *pstS1* | S347 | -388.45 |
| Putative L-lactate dehydrogenase | *lldD* | Y55 | -258.83 |
| Alkyl hydroperoxide reductase AhpD | Rv2159c | T165 | -246.35 |
| Serine/threonine-protein kinase PknD | *pknD* | S2 | -179.07 |
| Catalase-peroxidase | *katG* | S465 | -153.46 |
| Phosphate-binding protein PstS 1 | *pstS1* | S186 | -130.58 |
| Alkyl hydroperoxide reductase AhpD | Rv2159c | S302 | -89.98 |
| Phosphate-binding protein PstS 1 | *pstS1* | T160 | -78.49 |
| Phosphate-binding protein PstS 1 | *pstS1* | S365 | -68.62 |
| ESX-3 secretion system protein EccA3 | *eccA3* | S533 | -49.83 |
| ESAT-6-like protein EsxO | *esxO* | T2 | -46.69 |
| ESX-3 secretion system protein EccC3 | *eccC3* | T891 | -45.42 |
| Isoniazid-induced protein IniB | *iniB* | S20 | -42.15 |
| Mycolipanoate synthase | *msl3* | Y1279 | -30.14 |
| Phosphate-binding protein PstS 1 | *pstS1* | T328 | -28.94 |
| Probable CDP-diacylglycerol pyrophosphatase | *cdh* | Y84 | -28.23 |
| Phosphate-binding protein PstS 1 | *pstS1* | K324 | -27.53 |
| Chaperone protein DnaK | *dnaK* | Y106 | -24.68 |
| Nitrate reductase alpha subunit | *narG* | T2 | -23.58 |
| Ferritin BfrB | *bfrB* | Y49 | -18.81 |
| Probable thiol peroxidase | *tpx* | T13 | -18.61 |
| Cytochrome BD ubiquinol oxidase subunit I | *cydA* | S307 | -18.27 |
| ESAT-6-like protein EsxO | *esxO* | Y65 | -18.09 |
| Alpha-crystallin | *hspX* | T101 | -18.03 |
| ESX-3 secretion system protein EccC3 | *eccC3* | T726 | -17.80 |
| Mycolipanoate synthase | *msl3* | S2016 | -16.50 |
| ESX-3 secretion system protein EccC3 | *eccC3* | T176 | -16.15 |
| Enolase | *eno* | K335 | -15.59 |
| ESX-3 secretion system protein EccA3 | *eccA3* | T433 | -15.40 |
| Mycolipanoate synthase | *msl3* | T1259 | -14.74 |
| ESX-3 secretion system protein EccC3 | *eccC3* | S125 | -14.54 |
| Long-chain-fatty-acid--CoA ligase FadD15 | *fadD15* | T160 | -14.35 |
| Enolase | *eno* | S39 | -14.15 |
| Probable thiol peroxidase | *tpx* | T98 | -13.50 |
| Enolase | *eno* | S198 | -13.30 |
| Alpha-crystallin | *hspX* | S91 | -13.27 |

644 Other groups of differentially acetylated proteins involved in stress response, virulence
645 and pathogenesis includes chaperon proteins (DnaK, HspX), oxidoreductases (AhpD, Tpx and

KatG), ESX-1 secretion-associated proteins (EsxB, EspI, EspR and EspF), mammalian cell entry proteins (Mce1B, Mce1F and Mce1C) and STPKs (PknD and PknH), drug resistance-associated proteins (RpoB, RpoC and IniB) (Table 2, Table SI 7). These proteins are indispensable for Mtb in signal transduction mechanisms that lead to bacterial adaptation to its environment [88], detoxification and drug resistance [86, 117], or involved in entry and survival of the pathogen inside macrophages [118]. Although the exact mechanism remains unknown, acetylation of enzymes involved in Mtb fitness and survival may lead to a change in the net charge of the protein, alter stability and compartmentalization, brings a conformational change and/or block kinase substrates and may thus modulate activity.

The protein-protein interaction (PPI) network was established to investigate the physical and functional interactions among the 261 differentially acetylated proteins. Eight interconnected protein clusters were identified using the MCODE plug-in toolkit, and the PPI network was visualized using the Cytoscape software. The clusters include proteins involved in ESX-3 secretion (cluster 1), host-pathogen interactions (cluster 2), glycolysis (cluster 3), TCA cycle (cluster 4), fatty acid metabolism (cluster 5), stress response (cluster 6), transcription and translation (cluster 7) and ATP synthase (cluster 8) (Figure 7A).

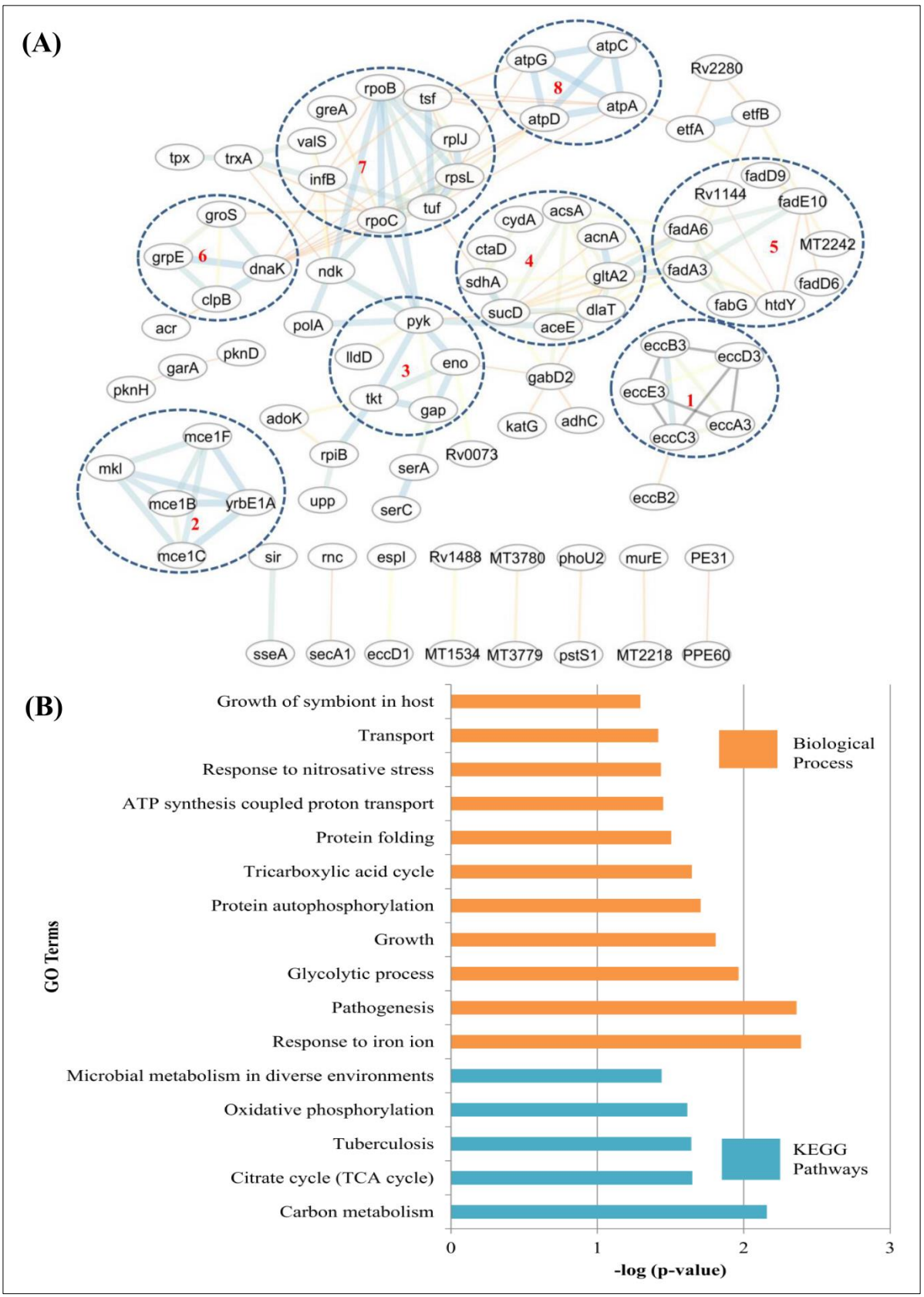



**Figure 7 |** (A) Protein-protein interaction network of differentially acetylated protein groups involved in; 1: ESX-3 secretion, 2: host-pathogen interaction, 3: glycolysis, 4: TCA cycle,5: fatty acid metabolism, 6: stress response, 7: transcription and translation and, 8: ATP synthase. (B) GO enrichment analysis of proteins hypoacetylated in lineage 7.

A total of 71 acetylated sites on 69 proteins and 109 acetylated sites on 93 proteins were exclusively identified in Mtb lineage 7 and lineage 4, respectively. GO enrichment analysis of the exclusively identified acetylated proteins showed a strain-specific enrichment of biological processes. Carbon and fatty acid metabolism were enriched in lineage 7 (Figure SI 3), while translation and amino acid metabolism were enriched in lineage 4 (H37Rv) (Figure SI 3). Enrichment of carbon and fatty acid metabolism in proteins exclusively identified in lineage 7 might be considered as a compensatory mechanism for energy generation.

## CONCLUSIONS

This study generated the first extensive data on O-acetylation in Mtb. We identified a total of 2490 acetylation sites on 953 Mtb unique proteins involved in a wide variety of fundamental cellular processes. These acetylated proteins are involved in Mtb core metabolic processes, bioenergetics, virulence and drug resistance. Our findings provide novel insight into a range of functions predicted to be regulated by $N_{\varepsilon}$-and O-acetylation in Mtb. The stoichiometry of Mtb O-acetylation is significantly higher than that of Nε-acetylation. This fact may in turn support the assumption that O-acetylation is involved in a broader range of processes than the Nε-acetylated counterparts. Due to the limited number of functional studies on the impact of acetylation conducted to date, the contribution of lysine acetylation in Mtb drug resistance is not fully understood. This study provides new knowledge on acetylated proteins and acetylation sites found near naturally occurring mutations that are known to be involved in Mtb drug resistance, and may be used as a basis for further functional studies. Differentially acetylated proteins involved in Mtb virulence, energy metabolism/growth and stress responses were uniformly biased towards being hypoacetylated in lineage 7 strains. This may lead to a metabolic state that makes one strain better fitted to a certain environment

compared to the other. The real impact (inhibition/activation) of acetylation on protein activity needs to be further investigated through functional studies.

This MS-based prediction focusing on the nature and role of protein acetylation in Mtb should be verified by using both *in vivo* and *in vitro* functional studies, which can provide the proof of concept for the role of acetylation in Mtb physiology. This investigation relied on two methods, deletion of acetyltransferases/deacetylases and substitution. Deletion of either the acetyltransferases or deacetylases can be used to unveil the role of these enzymes in the acetylation process, while point mutations are used to assess the functional role of a specific acetylation site. Finally, the use of immunoaffinity-enriched samples for nanoLC-MS/MS analysis may enable the identification of large numbers of acetylated proteins to establish the complete Mtb O-acetylome.

## ASSOCIATED CONTENT

### Supporting Information

- Figure SI 1. Representative spectra of $N_{\varepsilon}$- and O-acetylated peptides in Mtb (PDF).
- Figure SI 2. Protein-Protein interaction network analysis of $N_{\varepsilon}$- and O-acetylated proteins in Mtb (PDF).
- Figure SI 3. GO Enrichment analysis of acetylated proteins exclusively identified in Mycobacterium tuberculosis strains of lineage 7 (A) and lineage 4 (B) (PDF).
- Table SI 1. List of class-I acetylation sites (n=2198) identified in H37Rv; localization probability >0.75 and PEF <0.05 (xlsx).
- Table SI 2. List of class-I acetylation sites (n=2200) identified in Lineage 7; localization probability >0.75 and PEF<0.05 (xlsx).
- Table SI 3. List of total identified class-I acetylation sites (n=2490); localization probability >0.75 and PEF<0.05 (xlsx).

- Table SI 4. List of 953 identified acetylated proteins and number of acetylation sites per protein (xlsx).
- Table SI 5-1. PROSITE patterns of acetylated proteins. In red the PROSITE signature , in green the identified acetylated site in *M. tuberculosis*, and the identified peptides are between brackets (xlsx).
- Table SI 5-2. List of identified PROSITE profiles of acetylated peptides (xlsx).
- Table SI 6-1. Gene Ontology enrichment analysis for Nε- and O-acetylated proteins using DAVID classification system, P-value < 0.05 (xlsx).
- Table SI 6-2. Gene Ontology enrichment analysis for differentially acetylated proteins between lineage 7 and lineage 4 (H37Rv) using DAVID classification system (xlsx).
- Table SI 7. List of differentially acetylated proteins (n=261) in lineage 7 versus lineage 4 (xlsx).

## AUTHOR INFORMATION

***Corresponding Author**

Alemayehu Godana Birhanu

E-mail: alexbiology97@yahoo.com. Tel: (+47) 94294333.

**ORCID ID:** 0000-0001-5302-0919

Tone Tønjum

E-mail: tone.tonjum@medisin.uio.no. Tel: (+47) 90152936.

**ORCID ID:** 0000-0002-1709-6921

**Author Contributions**

TT, AGB, and SAY conceived the study and study design. SAY collected the lineage 7 isolates. AGB performed the MS sample preparation. AGB performed bioinformatics and statistical analyses. AGB and TT evaluated and interpreted the data and drafted the paper. All authors edited and approved the final manuscript.

## Notes

The authors declare no competing financial interest.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [119] partner repository with the dataset identifier PXD006630.


## FUNDING

Funding was received from the Research Council of Norway (RCN) FRIMEDBIO project # 204747 and RCN GLOBVAC projects #234506 to TT and 192468 to CH, and Norwegian South-Eastern Health Authority project 2013080 to SAY and TT.

## ACKNOWLEDGMENTS

We thank Tahira Riaz for acquiring the MS data. We also thank the patients for consenting to participate in the study and selected health care facilities in the Amhara Region, Ethiopia, for facilitating the study. We are grateful to the Armauer Hansen Research Institute (AHRI), Addis Ababa, Ethiopia, and the Norwegian Institute of Public Health for facilitating the transfer of lineage 7 strains for WGS at Oslo University Hospital.


## ABBREVIATIONS

Mtb, *Mycobacterium tuberculosis*; PTM, post-translational modification; MS, mass spectrometry; GO, Gene Ontology; FDR, false discovery rate; PEP, posterior error probability; KEGG, Kyoto Encyclopedia of Genes and Genomes.

## REFERENCES


1. WHO *Global Tuberculosis Report 2016. World Health Organization, Geneva; 2016.* *http://www.who.int/tb/publications/global_report/en/. (accessed May 15, 2017).*
2. Cain, J. A.; Solis, N.; Cordwell, S. J., Beyond gene expression: the impact of protein post-translational modifications in bacteria. *J Proteomics* **2014,** *97*, 265-86.
3. Vergnolle, O.; Xu, H.; Tufariello, J. M.; Favrot, L.; Malek, A. A.; Jacobs, W. R., Jr.; Blanchard, J. S., Post-translational Acetylation of MbtA Modulates Mycobacterial Siderophore Biosynthesis. *J Biol Chem* **2016,** *291* (42), 22315-22326.
4. Ouidir, T.; Kentache, T.; Hardouin, J., Protein lysine acetylation in bacteria: Current state of the art. *Proteomics* **2016,** *16* (2), 301-9.
5. Carabetta, V. J.; Cristea, I. M., The regulation, function, and detection of protein acetylation in bacteria. *J Bacteriol* **2017**, JB. 00107-17.

6. Nambi, S.; Gupta, K.; Bhattacharyya, M.; Ramakrishnan, P.; Ravikumar, V.; Siddiqui, N.; Thomas, A. T.; Visweswariah, S. S., Cyclic AMP-dependent protein lysine acylation in mycobacteria regulates fatty acid and propionate metabolism. *J Biol Chem* **2013,** *288* (20), 14114-24.
7. Gu, J.; Deng, J. Y.; Li, R.; Wei, H.; Zhang, Z.; Zhou, Y.; Zhang, Y.; Zhang, X. E., Cloning and characterization of NAD-dependent protein deacetylase (Rv1151c) from Mycobacterium tuberculosis. *Biochemistry (Mosc)* **2009,** *74* (7), 743-8.
8. Hayden, J. D.; Brown, L. R.; Gunawardena, H. P.; Perkowski, E. F.; Chen, X.; Braunstein, M., Reversible acetylation regulates acetate and propionate metabolism in Mycobacterium smegmatis. *Microbiology* **2013,** *159* (Pt 9), 1986-99.
9. Lee, H. J.; Lang, P. T.; Fortune, S. M.; Sassetti, C. M.; Alber, T., Cyclic AMP regulation of protein lysine acetylation in Mycobacterium tuberculosis. *Nat Struct Mol Biol* **2012,** *19* (8), 811-8.
10. Nambi, S.; Basu, N.; Visweswariah, S. S., cAMP-regulated protein lysine acetylases in mycobacteria. *J Biol Chem* **2010,** *285* (32), 24313-23.
11. Xu, H.; Hegde, S. S.; Blanchard, J. S., Reversible Acetylation and Inactivation of Mycobacterium tuberculosis Acetyl-CoA Synthetase Is Dependent on cAMP. *Biochemistry* **2011,** *50* (26), 5883-5892.
12. Li, R.; Gu, J.; Chen, P.; Zhang, Z.; Deng, J.; Zhang, X., Purification and characterization of the acetyl-CoA synthetase from Mycobacterium tuberculosis. *Acta Biochim Biophys Sin* **2011,** *43* (11), 891-899.
13. van Els, C. A.; Corbiere, V.; Smits, K.; van Gaans-van den Brink, J. A.; Poelen, M. C.; Mascart, F.; Meiring, H. D.; Locht, C., Toward Understanding the Essence of Post-Translational Modifications for the Mycobacterium tuberculosis Immunoproteome. *Front Immunol* **2014,** *5*, 361.
14. Okkels, L. M.; Müller, E. C.; Schmid, M.; Rosenkrands, I.; Kaufmann, S. H.; Andersen, P.; Jungblut, P. R., CFP10 discriminates between nonacetylated and acetylated ESAT‐6 of Mycobacterium tuberculosis by differential interaction. *Proteomics* **2004,** *4* (10), 2954-2960.
15. Xie, L.; Wang, X.; Zeng, J.; Zhou, M.; Duan, X.; Li, Q.; Zhang, Z.; Luo, H.; Pang, L.; Li, W.; Liao, G.; Yu, X.; Li, Y.; Huang, H.; Xie, J., Proteome-wide lysine acetylation profiling of the human pathogen Mycobacterium tuberculosis. *Int J Biochem Cell Biol* **2015,** *59*, 193-202.
16. Ghosh, S.; Padmanabhan, B.; Anand, C.; Nagaraja, V., Lysine acetylation of the Mycobacterium tuberculosis HU protein modulates its DNA binding and genome organization. *Mol Microbiol* **2016,** *100* (4), 577-88.
17. Liu, F.; Yang, M.; Wang, X.; Yang, S.; Gu, J.; Zhou, J.; Zhang, X. E.; Deng, J.; Ge, F., Acetylome analysis reveals diverse functions of lysine acetylation in Mycobacterium tuberculosis. *Mol Cell Proteomics* **2014,** *13* (12), 3352-66.
18. Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. V.; Eiglmeier, K.; Gas, S.; Barry, C. E., 3rd; 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, M. A.; Rajandream, M. A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G., Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* **1998,** *393* (6685), 537-44.
19. Lee, W.; VanderVen, B. C.; Walker, S.; Russell, D. G., Novel protein acetyltransferase, Rv2170, modulates carbon and energy metabolism in Mycobacterium tuberculosis. *Sci Rep* **2017,** *7* (1), 72.

20. Jose, L.; Ramachandran, R.; Bhagavat, R.; Gomez, R. L.; Chandran, A.; Raghunandanan, S.; Omkumar, R. V.; Chandra, N.; Mundayoor, S.; Kumar, R. A., Hypothetical protein Rv3423.1 of Mycobacterium tuberculosis is a histone acetyltransferase. *Febs Journal* **2016,** *283* (2), 265-281.
21. Weinert, B. T.; Iesmantavicius, V.; Wagner, S. A.; Scholz, C.; Gummesson, B.; Beli, P.; Nystrom, T.; Choudhary, C., Acetyl-Phosphate Is a Critical Determinant of Lysine Acetylation in E. coli. *Molecular Cell* **2013,** *51* (2), 265-272.
22. Kuhn, M. L.; Zemaitaitis, B.; Hu, L. I.; Sahu, A.; Sorensen, D.; Minasov, G.; Lima, B. P.; Scholle, M.; Mrksich, M.; Anderson, W. F.; Gibson, B. W.; Schilling, B.; Wolfe, A. J., Structural, kinetic and proteomic characterization of acetyl phosphate-dependent bacterial protein acetylation. *PLoS One* **2014,** *9* (4), e94816.
23. Kosono, S.; Tamura, M.; Suzuki, S.; Kawamura, Y.; Yoshida, A.; Nishiyama, M.; Yoshida, M., Changes in the Acetylome and Succinylome of Bacillus subtilis in Response to Carbon Source. *PLoS One* **2015,** *10* (6), e0131169.
24. Paik, W. K.; Pearson, D.; Lee, H. W.; Kim, S., Nonenzymatic acetylation of histones with acetyl-CoA. *Biochim Biophys Acta* **1970,** *213* (2), 513-22.
25. Wagner, G. R.; Payne, R. M., Widespread and enzyme-independent Nϵ-acetylation and Nϵ-succinylation of proteins in the chemical conditions of the mitochondrial matrix. *Journal of Biological Chemistry* **2013,** *288* (40), 29036-29045.
26. Baeza, J.; Smallegan, M. J.; Denu, J. M., Site-specific reactivity of nonenzymatic lysine acetylation. *ACS Chem Biol* **2015,** *10* (1), 122-8.
27. Mukherjee, S.; Keitany, G.; Li, Y.; Wang, Y.; Ball, H. L.; Goldsmith, E. J.; Orth, K., Yersinia YopJ acetylates and inhibits kinase activation by blocking phosphorylation. *Science* **2006,** *312* (5777), 1211-4.
28. Cousin, C.; Derouiche, A.; Shi, L.; Pagot, Y.; Poncet, S.; Mijakovic, I., Protein-serine/threonine/tyrosine kinases in bacterial signaling and regulation. *FEMS Microbiol Lett* **2013,** *346* (1), 11-9.
29. Chao, J.; Wong, D.; Zheng, X.; Poirier, V.; Bach, H.; Hmama, Z.; Av-Gay, Y., Protein kinase and phosphatase signaling in Mycobacterium tuberculosis physiology and pathogenesis. *Biochim Biophys Acta* **2010,** *1804* (3), 620-7.
30. Denk, D.; Bock, A., L-cysteine biosynthesis in Escherichia coli: nucleotide sequence and expression of the serine acetyltransferase (cysE) gene from the wild-type and a cysteine-excreting mutant. *J Gen Microbiol* **1987,** *133* (3), 515-25.
31. Anonsen, J. H.; Borud, B.; Vik, A.; Viburiene, R.; Koomey, M., Structural and genetic analyses of glycan O-acetylation in a bacterial protein glycosylation system: evidence for differential effects on glycan chain length. *Glycobiology* **2017**, 1-12.
32. Moynihan, P. J.; Clarke, A. J., O-Acetylated peptidoglycan: controlling the activity of bacterial autolysins and lytic enzymes of innate immune systems. *Int J Biochem Cell Biol* **2011,** *43* (12), 1655-9.
33. Yang, M.; Wang, Y.; Chen, Y.; Cheng, Z.; Gu, J.; Deng, J.; Bi, L.; Chen, C.; Mo, R.; Wang, X.; Ge, F., Succinylome analysis reveals the involvement of lysine succinylation in metabolism in pathogenic Mycobacterium tuberculosis. *Mol Cell Proteomics* **2015,** *14* (4), 796-811.
34. Shevchenko, A.; Tomas, H.; Havlis, J.; Olsen, J. V.; Mann, M., In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* **2006,** *1* (6), 2856-60.
35. Cox, J.; Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **2008,** *26* (12), 1367-72.

36. Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M., Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* **2011,** *10* (4), 1794-805.
37. Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M. Y.; Geiger, T.; Mann, M.; Cox, J., The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **2016,** *13* (9), 731-40.
38. Benjamini, Y.; Drai, D.; Elmer, G.; Kafkafi, N.; Golani, I., Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* **2001,** *125* (1-2), 279-84.
39. Ashburner, M.; Ball, C. A.; Blake, J. A.; Botstein, D.; Butler, H.; Cherry, J. M.; Davis, A. P.; Dolinski, K.; Dwight, S. S.; Eppig, J. T.; Harris, M. A.; Hill, D. P.; Issel-Tarver, L.; Kasarskis, A.; Lewis, S.; Matese, J. C.; Richardson, J. E.; Ringwald, M.; Rubin, G. M.; Sherlock, G., Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **2000,** *25* (1), 25-9.
40. Huang da, W.; Sherman, B. T.; Lempicki, R. A., Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **2009,** *4* (1), 44-57.
41. Lew, J. M.; Kapopoulou, A.; Jones, L. M.; Cole, S. T., TubercuList--10 years after. *Tuberculosis (Edinb)* **2011,** *91* (1), 1-7.
42. Camus, J. C.; Pryor, M. J.; Medigue, C.; Cole, S. T., Re-annotation of the genome sequence of Mycobacterium tuberculosis H37Rv. *Microbiology* **2002,** *148* (Pt 10), 2967-73.
43. Kanehisa, M.; Goto, S., KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **2000,** *28* (1), 27-30.
44. de Castro, E.; Sigrist, C. J.; Gattiker, A.; Bulliard, V.; Langendijk-Genevaux, P. S.; Gasteiger, E.; Bairoch, A.; Hulo, N., ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res* **2006,** *34* (Web Server issue), W362-5.
45. Chou, M. F.; Schwartz, D., Biological sequence motif discovery using motif‐x. *Current Protocols in Bioinformatics* **2011**, 13.15. 1-13.15. 24.
46. Jensen, L. J.; Kuhn, M.; Stark, M.; Chaffron, S.; Creevey, C.; Muller, J.; Doerks, T.; Julien, P.; Roth, A.; Simonovic, M.; Bork, P.; von Mering, C., STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res* **2009,** *37* (Database issue), D412-6.
47. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T., Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **2003,** *13* (11), 2498-504.
48. Ma, Q.; Wood, T. K., Protein acetylation in prokaryotes increases stress resistance. *Biochem Biophys Res Commun* **2011,** *410* (4), 846-51.
49. Xie, L.; Liu, W.; Li, Q.; Chen, S.; Xu, M.; Huang, Q.; Zeng, J.; Zhou, M.; Xie, J., First succinyl-proteome profiling of extensively drug-resistant Mycobacterium tuberculosis revealed involvement of succinylation in cellular physiology. *J Proteome Res* **2015,** *14* (1), 107-19.
50. Fortuin, S.; Tomazella, G. G.; Nagaraj, N.; Sampson, S. L.; Gey van Pittius, N. C.; Soares, N. C.; Wiker, H. G.; de Souza, G. A.; Warren, R. M., Phosphoproteomics analysis of a clinical Mycobacterium tuberculosis Beijing isolate: expanding the mycobacterial phosphoproteome catalog. *Front Microbiol* **2015,** *6*, 6.
51. Saito, K.; Yokoyama, H.; Noji, M.; Murakoshi, I., Molecular cloning and characterization of a plant serine acetyltransferase playing a regulatory role in cysteine biosynthesis from watermelon. *J Biol Chem* **1995,** *270* (27), 16321-6.
52. Saito, K.; Yamazoe, Y.; Kamataki, T.; Kato, R., Mechanism of activation of proximate mutagens in Ames' tester strains: The acetyl-CoA dependent enzyme in Salmonella

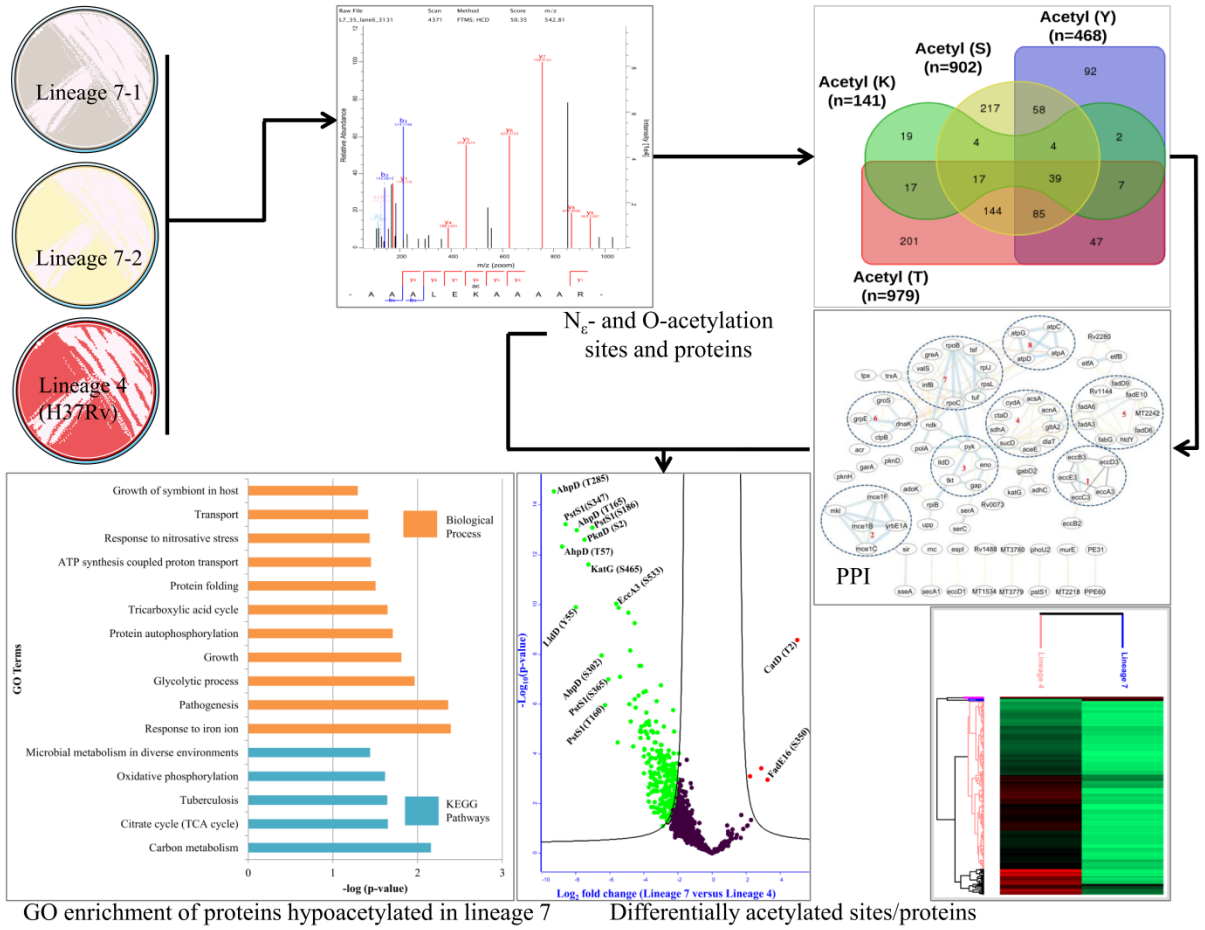
typhimurium TA98 deficient in TA981, 8-DNP6 catalyzes DNA-binding as the cause of mutagenicity. *Biochemical and biophysical research communications* **1983,** *116* (1), 141-147.
53. Mukherjee, S.; Hao, Y. H.; Orth, K., A newly discovered post-translational modification--the acetylation of serine and threonine residues. *Trends Biochem Sci* **2007,** *32* (5), 210-6.
54. Holliday, G. L.; Mitchell, J. B.; Thornton, J. M., Understanding the functional roles of amino acid residues in enzyme catalysis. *J Mol Biol* **2009,** *390* (3), 560-77.
55. Bartlett, G. J.; Porter, C. T.; Borkakoti, N.; Thornton, J. M., Analysis of catalytic residues in enzyme active sites. *J Mol Biol* **2002,** *324* (1), 105-21.
56. Thao, S.; Escalante-Semerena, J. C., Control of protein function by reversible Nvarepsilon-lysine acetylation in bacteria. *Curr Opin Microbiol* **2011,** *14* (2), 200-4.
57. Zhang, T.; Wang, S.; Lin, Y.; Xu, W.; Ye, D.; Xiong, Y.; Zhao, S.; Guan, K. L., Acetylation negatively regulates glycogen phosphorylase by recruiting protein phosphatase 1. *Cell Metab* **2012,** *15* (1), 75-87.
58. Glozak, M. A.; Sengupta, N.; Zhang, X.; Seto, E., Acetylation and deacetylation of non-histone proteins. *Gene* **2005,** *363*, 15-23.
59. Gu, W.; Roeder, R. G., Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell* **1997,** *90* (4), 595-606.
60. Thao, S.; Chen, C.-S.; Zhu, H.; Escalante-Semerena, J. C., Nε− lysine acetylation of a bacterial transcription factor inhibits its DNA-binding activity. *PloS one* **2010,** *5* (12), e15123.
61. Howe, F. S.; Boubriak, I.; Sale, M. J.; Nair, A.; Clynes, D.; Grijzenhout, A.; Murray, S. C.; Woloszczuk, R.; Mellor, J., Lysine acetylation controls local protein conformation by influencing proline isomerization. *Mol Cell* **2014,** *55* (5), 733-44.
62. Li, T.; Diner, B. A.; Chen, J.; Cristea, I. M., Acetylation modulates cellular distribution and DNA sensing ability of interferon-inducible protein IFI16. *Proc Natl Acad Sci U S A* **2012,** *109* (26), 10558-63.
63. Ishfaq, M.; Maeta, K.; Maeda, S.; Natsume, T.; Ito, A.; Yoshida, M., Acetylation regulates subcellular localization of eukaryotic translation initiation factor 5A (eIF5A). *FEBS Lett* **2012,** *586* (19), 3236-41.
64. Shaw, B. F.; Schneider, G. F.; Bilgicer, B.; Kaufman, G. K.; Neveu, J. M.; Lane, W. S.; Whitelegge, J. P.; Whitesides, G. M., Lysine acetylation can generate highly charged enzymes with increased resistance toward irreversible inactivation. *Protein Science* **2008,** *17* (8), 1446-1455.
65. Sakai, H., Mutagenesis of the active site lysine 221 of the pyruvate kinase from Bacillus stearothermophilus. *J Biochem* **2005,** *137* (2), 141-5.
66. Pan, J.; Ye, Z.; Cheng, Z.; Peng, X.; Wen, L.; Zhao, F., Systematic analysis of the lysine acetylome in Vibrio parahemolyticus. *J Proteome Res* **2014,** *13* (7), 3294-302.
67. Pan, J.; Chen, R.; Li, C.; Li, W.; Ye, Z., Global Analysis of Protein Lysine Succinylation Profiles and Their Overlap with Lysine Acetylation in the Marine Bacterium Vibrio parahemolyticus. *J Proteome Res* **2015,** *14* (10), 4309-18.
68. Berndsen, C. E.; Denu, J. M., Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Curr Opin Struct Biol* **2008,** *18* (6), 682-9.
69. Wang, Q.; Zhang, Y.; Yang, C.; Xiong, H.; Lin, Y.; Yao, J.; Li, H.; Xie, L.; Zhao, W.; Yao, Y.; Ning, Z. B.; Zeng, R.; Xiong, Y.; Guan, K. L.; Zhao, S.; Zhao, G. P., Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* **2010,** *327* (5968), 1004-7.
70. Guan, K. L.; Xiong, Y., Regulation of intermediary metabolism by protein acetylation. *Trends Biochem Sci* **2011,** *36* (2), 108-16.
71. Zhao, S.; Xu, W.; Jiang, W.; Yu, W.; Lin, Y.; Zhang, T.; Yao, J.; Zhou, L.; Zeng, Y.; Li, H.; Li, Y.; Shi, J.; An, W.; Hancock, S. M.; He, F.; Qin, L.; Chin, J.; Yang, P.; Chen, X.;

Lei, Q.; Xiong, Y.; Guan, K. L., Regulation of cellular metabolism by protein lysine acetylation. *Science* **2010,** *327* (5968), 1000-4.
72. McKinney, J. D.; Honer zu Bentrup, K.; Munoz-Elias, E. J.; Miczak, A.; Chen, B.; Chan, W. T.; Swenson, D.; Sacchettini, J. C.; Jacobs, W. R., Jr.; Russell, D. G., Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature* **2000,** *406* (6797), 735-8.
73. Zu Bentrup, K. H.; Miczak, A.; Swenson, D. L.; Russell, D. G., Characterization of Activity and Expression of Isocitrate Lyase in Mycobacterium avium andMycobacterium tuberculosis. *Journal of bacteriology* **1999,** *181* (23), 7161-7167.
74. Munoz-Elias, E. J.; McKinney, J. D., Mycobacterium tuberculosis isocitrate lyases 1 and 2 are jointly required for in vivo growth and virulence. *Nat Med* **2005,** *11* (6), 638-44.
75. Rehman, A.; Mcfadden, B. A., The Consequences of Replacing Histidine 356 in Isocitrate Lyase fromEscherichia coli. *Archives of biochemistry and biophysics* **1996,** *336* (2), 309-315.
76. Kinhikar, A. G.; Vargas, D.; Li, H.; Mahaffey, S. B.; Hinds, L.; Belisle, J. T.; Laal, S., Mycobacterium tuberculosis malate synthase is a laminin‐binding adhesin. *Molecular microbiology* **2006,** *60* (4), 999-1013.
77. Shi, L.; Tu, B. P., Protein acetylation as a means to regulate protein function in tune with metabolic state. Portland Press Limited: 2014.
78. Wayne, L. G.; Lin, K. Y., Glyoxylate metabolism and adaptation of Mycobacterium tuberculosis to survival under anaerobic conditions. *Infect Immun* **1982,** *37* (3), 1042-9.
79. Dunn, M. F.; Ramirez-Trujillo, J. A.; Hernandez-Lucas, I., Major roles of isocitrate lyase and malate synthase in bacterial and fungal pathogenesis. *Microbiology* **2009,** *155* (Pt 10), 3166-75.
80. Lorenz, M. C.; Fink, G. R., Life and death in a macrophage: role of the glyoxylate cycle in virulence. *Eukaryot Cell* **2002,** *1* (5), 657-62.
81. Guenin-Mace, L.; Simeone, R.; Demangel, C., Lipids of pathogenic Mycobacteria: contributions to virulence and host immune suppression. *Transbound Emerg Dis* **2009,** *56* (6-7), 255-68.
82. Zhou, Y.; Chen, T.; Zhou, L.; Fleming, J.; Deng, J.; Wang, X.; Wang, L.; Wang, Y.; Zhang, X.; Wei, W.; Bi, L., Discovery and characterization of Ku acetylation in Mycobacterium smegmatis. *FEMS Microbiol Lett* **2015,** *362* (6).
83. Zhang, Q.; Zhou, A.; Li, S.; Ni, J.; Tao, J.; Lu, J.; Wan, B.; Li, S.; Zhang, J.; Zhao, S., Reversible lysine acetylation is involved in DNA replication initiation by regulating activities of initiator DnaA in Escherichia coli. *Scientific reports* **2016,** *6*, 30837.
84. Smith, T.; Wolff, K. A.; Nguyen, L., Molecular biology of drug resistance in Mycobacterium tuberculosis. In *Pathogenesis of Mycobacterium tuberculosis and its Interaction with the Host Organism*, Springer: 2012; pp 53-80.
85. Ng, V. H.; Cox, J. S.; Sousa, A. O.; MacMicking, J. D.; McKinney, J. D., Role of KatG catalase-peroxidase in mycobacterial pathogenesis: countering the phagocyte oxidative burst. *Mol Microbiol* **2004,** *52* (5), 1291-302.
86. Cohen, T.; Becerra, M. C.; Murray, M. B., Isoniazid resistance and the future of drug-resistant tuberculosis. *Microb Drug Resist* **2004,** *10* (4), 280-5.
87. Gotoh, Y.; Eguchi, Y.; Watanabe, T.; Okamoto, S.; Doi, A.; Utsumi, R., Two-component signal transduction as potential drug targets in pathogenic bacteria. *Curr Opin Microbiol* **2010,** *13* (2), 232-9.
88. Prisic, S.; Husson, R. N., Mycobacterium tuberculosis Serine/Threonine Protein Kinases. *Microbiol Spectr* **2014,** *2* (5).
89. Jhingan, G. D.; Kumari, S.; Jamwal, S. V.; Kalam, H.; Arora, D.; Jain, N.; Kumaar, L. K.; Samal, A.; Rao, K. V.; Kumar, D.; Nandicoori, V. K., Comparative Proteomic Analyses

of Avirulent, Virulent, and Clinical Strains of Mycobacterium tuberculosis Identify Strain-specific Patterns. *J Biol Chem* **2016,** *291* (27), 14257-73.
90. Mukherjee, P.; Sureka, K.; Datta, P.; Hossain, T.; Barik, S.; Das, K. P.; Kundu, M.; Basu, J., Novel role of Wag31 in protection of mycobacteria under oxidative stress. *Mol Microbiol* **2009,** *73* (1), 103-19.
91. Kang, C. M.; Nyayapathy, S.; Lee, J. Y.; Suh, J. W.; Husson, R. N., Wag31, a homologue of the cell division protein DivIVA, regulates growth, morphology and polar cell wall synthesis in mycobacteria. *Microbiology* **2008,** *154* (Pt 3), 725-35.
92. Unal, C. M.; Steinert, M., Microbial peptidyl-prolyl cis/trans isomerases (PPIases): virulence factors and potential alternative drug targets. *Microbiol Mol Biol Rev* **2014,** *78* (3), 544-71.
93. Munshi, T.; Gupta, A.; Evangelopoulos, D.; Guzman, J. D.; Gibbons, S.; Keep, N. H.; Bhakta, S., Characterisation of ATP-dependent Mur ligases involved in the biogenesis of cell wall peptidoglycan in Mycobacterium tuberculosis. *PLoS One* **2013,** *8* (3), e60143.
94. Tseng, S. T.; Tai, C. H.; Li, C. R.; Lin, C. F.; Shi, Z. Y., The mutations of katG and inhA genes of isoniazid-resistant Mycobacterium tuberculosis isolates in Taiwan. *J Microbiol Immunol Infect* **2015,** *48* (3), 249-55.
95. Ramaswamy, S. V.; Reich, R.; Dou, S. J.; Jasperse, L.; Pan, X.; Wanger, A.; Quitugua, T.; Graviss, E. A., Single nucleotide polymorphisms in genes associated with isoniazid resistance in Mycobacterium tuberculosis. *Antimicrob Agents Chemother* **2003,** *47* (4), 1241-50.
96. Chien, J. Y.; Chiu, W. Y.; Chien, S. T.; Chiang, C. J.; Yu, C. J.; Hsueh, P. R., Mutations in gyrA and gyrB among Fluoroquinolone- and Multidrug-Resistant Mycobacterium tuberculosis Isolates. *Antimicrob Agents Chemother* **2016,** *60* (4), 2090-6.
97. Wang, J. Y.; Lee, L. N.; Lai, H. C.; Wang, S. K.; Jan, I. S.; Yu, C. J.; Hsueh, P. R.; Yang, P. C., Fluoroquinolone resistance in Mycobacterium tuberculosis isolates: associated genetic mutations and relationship to antimicrobial exposure. *J Antimicrob Chemother* **2007,** *59* (5), 860-5.
98. Pitaksajjakul, P.; Wongwit, W.; Punprasit, W.; Eampokalap, B.; Peacock, S.; Ramasoota, P., Mutations in the gyrA and gyrB genes of fluoroquinolone-resistant Mycobacterium tuberculosis from TB patients in Thailand. *Southeast Asian J Trop Med Public Health* **2005,** *36 Suppl 4*, 228-37.
99. Chen, W.; Biswas, T.; Porter, V. R.; Tsodikov, O. V.; Garneau-Tsodikova, S., Unusual regioversatility of acetyltransferase Eis, a cause of drug resistance in XDR-TB. *Proc Natl Acad Sci U S A* **2011,** *108* (24), 9804-8.
100. Tsodikov, O. V.; Green, K. D.; Garneau-Tsodikova, S., A random sequential mechanism of aminoglycoside acetylation by Mycobacterium tuberculosis Eis protein. *PLoS One* **2014,** *9* (4), e92370.
101. Yimer, S. A.; Norheim, G.; Namouchi, A.; Zegeye, E. D.; Kinander, W.; Tonjum, T.; Bekele, S.; Mannsaker, T.; Bjune, G.; Aseffa, A.; Holm-Hansen, C., Mycobacterium tuberculosis lineage 7 strains are associated with prolonged patient delay in seeking treatment for pulmonary tuberculosis in Amhara Region, Ethiopia. *J Clin Microbiol* **2015,** *53* (4), 1301-9.
102. Yimer, S. A.; Birhanu, A. G.; Kalayou, S.; Riaz, T.; Zegeye, E. D.; Beyene, G. T.; Holm-Hansen, C.; Norheim, G.; Abebe, M.; Aseffa, A.; Tonjum, T., Comparative Proteomic Analysis of Mycobacterium tuberculosis Lineage 7 and Lineage 4 Strains Reveals Differentially Abundant Proteins Linked to Slow Growth and Virulence. *Front Microbiol* **2017,** *8*, 795.

103. Eisenreich, W.; Dandekar, T.; Heesemann, J.; Goebel, W., Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. *Nat Rev Microbiol* **2010,** *8* (6), 401-12.
104. Munoz-Elias, E. J.; McKinney, J. D., Carbon metabolism of intracellular bacteria. *Cell Microbiol* **2006,** *8* (1), 10-22.
105. Zhang, X.; Ji, R.; Liao, X.; Deng, L. Y.; Castillero, E.; Givens, R.; George, I.; Schulze, P. C., Lysine Acetylation of Pyruvate Dehydrogenase Reduces Enzymatic Activity and Contributes to Impaired Substrate Metabolism in the Failing Myocardium. Am Heart Assoc: 2014.
106. Ozden, O.; Park, S. H.; Wagner, B. A.; Song, H. Y.; Zhu, Y.; Vassilopoulos, A.; Jung, B.; Buettner, G. R.; Gius, D., SIRT3 deacetylates and increases pyruvate dehydrogenase activity in cancer cells. *Free Radic Biol Med* **2014,** *76*, 163-172.
107. Li, R.; Gu, J.; Chen, P.; Zhang, Z.; Deng, J.; Zhang, X., Purification and characterization of the acetyl-CoA synthetase from Mycobacterium tuberculosis. *Acta Biochim Biophys Sin (Shanghai)* **2011,** *43* (11), 891-9.
108. McKinney, J. D.; zu Bentrup, K. H.; Muñoz-Elías, E. J.; Miczak, A., Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature* **2000,** *406* (6797), 735.
109. Shi, L.; Sohaskey, C. D.; Pheiffer, C.; Datta, P.; Parks, M.; McFadden, J.; North, R. J.; Gennaro, M. L., Carbon flux rerouting during Mycobacterium tuberculosis growth arrest. *Mol Microbiol* **2010,** *78* (5), 1199-215.
110. Wolfe, A. J., The acetate switch. *Microbiol Mol Biol Rev* **2005,** *69* (1), 12-50.
111. Contiero, J.; Beatty, C.; Kumari, S.; DeSanti, C. L.; Strohl, W. R.; Wolfe, A., Effects of mutations in acetate metabolism on high-cell-density growth of Escherichia coli. *Journal of Industrial Microbiology & Biotechnology* **2000,** *24* (6), 421-430.
112. Billig, S.; Schneefeld, M.; Huber, C.; Grassl, G. A.; Eisenreich, W.; Bange, F.-C., Lactate oxidation facilitates growth of Mycobacterium tuberculosis in human macrophages. *Scientific reports* **2017,** *7* (1), 6484.
113. Banerjee, R.; Vats, P.; Dahale, S.; Kasibhatla, S. M.; Joshi, R., Comparative genomics of cell envelope components in mycobacteria. *PLoS One* **2011,** *6* (5), e19280.
114. Daniel, J.; Sirakova, T.; Kolattukudy, P., An acyl-CoA synthetase in Mycobacterium tuberculosis involved in triacylglycerol accumulation during dormancy. *PLoS One* **2014,** *9* (12), e114877.
115. Chen, J. M., Mycosins of the Mycobacterial Type VII ESX Secretion System: the Glue That Holds the Party Together. *MBio* **2016,** *7* (6), e02062-16.
116. Peirs, P.; Lefevre, P.; Boarbi, S.; Wang, X.-M.; Denis, O.; Braibant, M.; Pethe, K.; Locht, C.; Huygen, K., Mycobacterium tuberculosis with disruption in genes encoding the phosphate binding proteins PstS1 and PstS2 is deficient in phosphate uptake and demonstrates reduced in vivo virulence. *Infection and immunity* **2005,** *73* (3), 1898-1902.
117. Colangeli, R.; Helb, D.; Sridharan, S.; Sun, J.; Varma-Basil, M.; Hazbon, M. H.; Harbacheuski, R.; Megjugorac, N. J.; Jacobs, W. R., Jr.; Holzenburg, A.; Sacchettini, J. C.; Alland, D., The Mycobacterium tuberculosis iniA gene is essential for activity of an efflux pump that confers drug tolerance to both isoniazid and ethambutol. *Mol Microbiol* **2005,** *55* (6), 1829-40.
118. Zhang, F.; Xie, J. P., Mammalian cell entry gene family of Mycobacterium tuberculosis. *Mol Cell Biochem* **2011,** *352* (1-2), 1-10.
119. Vizcaíno, J. A.; Csordas, A.; del-Toro, N.; Dianes, J. A.; Griss, J.; Lavidas, I.; Mayer, G.; Perez-Riverol, Y.; Reisinger, F.; Ternent, T., 2016 update of the PRIDE database and its related tools. *Nucleic acids research* **2016,** *44* (D1), D447-D456.

**For Table of content (TOC) abstract only**




GO enrichment of proteins hypoacetylated in lineage 7 Differentially acetylated sites/proteins