Phosphoproteomic Approaches to Discover Novel Substrates of Mycobacterial Ser/Thr Protein Kinases

Authors
Seanantha S. Baros, Jonathan M. Blackburn, and Nelson C. Soares

Graphical Abstract

Highlights

- Mapping kinase-substrate relationships is vital in discovering new tuberculosis drug targets.
- LC-MS/MS-based phosphoproteomics expand mycobacterial STPK substrate catalogues.
- We review and integrate MS-generated datasets on novel candidate substrates.
- Validation studies are necessary to confirm true physiological substrates of STPKs.
Phosphoproteomic Approaches to Discover Novel Substrates of Mycobacterial Ser/Thr Protein Kinases*

Seanantha S. Baros‡, Jonathan M. Blackburn‡§, and Nelson C. Soares¶

Mycobacterial Ser/Thr protein kinases (STPKs) play a critical role in signal transduction pathways that ultimately determine mycobacterial growth and metabolic adaptation. Identification of key physiological substrates of these protein kinases is, therefore, crucial to better understand how Ser/Thr phosphorylation contributes to mycobacterial environmental adaptation, including response to stress, cell division, and host-pathogen interactions. Various substrate detection methods have been employed with limited success, with direct targets of STPKs remaining elusive. Recently developed mass spectrometry (MS)-based phosphoproteomics approaches have expanded the list of potential STPK substrate identifications, yet further investigation is required to define the most functionally significant phosphosites and their physiological importance. Prior to the application of MS workflows, for instance, GarA was the only known and validated physiological substrate for protein kinase G (PknG) from pathogenic mycobacteria. A subsequent list of at least 28 candidate PknG substrates has since been reported with the use of MS-based analyses. Herein, we integrate and critically review MS-generated datasets available on novel STPK substrates and report new functional and subcellular localization enrichment analyses on novel candidate protein kinase A (PknA), protein kinase B (PknB) and PknG substrates to deduce the possible physiological roles of these kinases. In addition, we assess substrate specificity patterns across different mycobacterial STPKs by analyzing reported sets of phosphopeptides, in order to determine whether novel motifs or consensus regions exist for mycobacterial Ser/Thr phosphorylation sites. This review focuses on MS-based techniques employed for STPK substrate identification in mycobacteria, while highlighting the advantages and challenges of the various applications. Molecular & Cellular Proteomics 19:233–244, 2020. DOI: 10.1074/mcp.R119.001668.

PTMs are essential for the proper functioning of certain proteins and create heterogeneity and temporal complexity in mycobacterial proteomes (1–3). Recent developments in MS-based proteomics have highlighted the occurrence of numerous types of PTMs in mycobacteria (2, 3). Phosphorylation is a reversible PTM responsible for diversifying protein functions and orchestrating various signaling networks in mycobacteria, mediated by protein kinases and phosphatases (1, 4–6). Other well-studied PTMs in mycobacteria include glycosylation, lipoylation, acetylation, and pupylation, which have a wide array of physiological consequences in mycobacteria, ranging from the regulation of cell functions to modulation of host-pathogen interactions, as reviewed elsewhere (2, 3).

Being the most dynamic of all common PTMs, Ser/Thr protein phosphorylation is a central control mechanism associated with the growth, virulence, and persistence of mycobacterial species, including Mycobacterium tuberculosis, the etiological agent of tuberculosis (7–11). Through phosphorylation, mycobacterial STPKs and two-component systems coordinate signal transduction cascades that regulate physiological states, environmental adaptation, and pathogenesis (10, 12, 13).

Two-component systems consist of a sensor histidine kinase and a cognate response regulator, which are typically genetically linked and transcriptionally coupled (14, 15). The histidine kinase spans the cytoplasmic membrane and activates the response regulator when prompted by specific external stimuli (14). Communication between the proteins occurs via phosphoryl-group transfer from a conserved histidine residue of the sensor kinase to a conserved aspartate residue of the response regulator (16). The response regulator is usually a transcription factor localized in the cytoplasm that mediates environmental adaptive responses (14).

In contrast to two-component systems, STPKs are pleiotropic signal transducers that can phosphorylate multiple substrates to elicit branched signaling (14). The majority of mycobacterial STPKs are transmembrane proteins with an extracellular sensor domain and an intracellular kinase do-

From the ‡Division of Chemical & Systems Biology, Department of Integrative Biomedical Sciences, Faculty of Health Sciences, University of Cape Town, South Africa; §Institute of Infectious Disease & Molecular Medicine, Faculty of Health Sciences, University of Cape Town, South Africa; ¶Sharjah Institute for Medical Research, University of Sharjah, Sharjah, United Arab Emirates; College of Pharmacy, University of Sharjah, Sharjah, United Arab Emirates

Received July 9, 2019, and in revised form, November 15, 2019

Published, MCP Papers in Press, December 15, 2019, DOI 10.1074/mcp.R119.001668
Classification of the 11 mycobacterial STPKs according to their sequence similarity

<table>
<thead>
<tr>
<th>Clade</th>
<th>Mycobacterial STPK</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>PknA, PknB, PknL</td>
</tr>
<tr>
<td>II</td>
<td>PknD, PknE, PknH</td>
</tr>
<tr>
<td>III</td>
<td>PknF, PknI, PknJ</td>
</tr>
<tr>
<td>IV</td>
<td>PknK</td>
</tr>
<tr>
<td>V</td>
<td>PknG</td>
</tr>
</tbody>
</table>

The extracellular sensor domain transduces extracytoplasmic signals to the intracellular kinase domain, resulting in kinase activation and phosphorylation of Ser/Thr residues on substrate proteins. This phosphorylation may directly affect protein function or influence protein-protein interactions. Notably, Ser/Thr phosphorylation rarely results in direct regulation of transcription, which is often facilitated by two-component signal transduction (10).

Genomic studies have revealed that *M. tuberculosis* strains are equipped with 11 conserved STPKs (PknA, PknB, PknD, PknE, PknF, PknG, PknH, PknI, PknJ, and PknK) (17, 18). Furthermore, in mycobacteria with larger genomes, such as *M. marinum* or *M. smegmatis*, the number of STPKs outnumbered that of two-component system kinases. This suggests that the majority of the signal transduction pathways in these species is via Ser/Thr phosphorylation/dephosphorylation events (19, 20). Hence, mycobacterial STPKs form the focus of this review.

A phylogenetic analysis clustered the 11 mycobacterial STPKs into five groups, Clades I-V (Table I) (20). The group of transmembrane sensor kinases, PknA, PknB, and PknL belong to Clade I. PknD, PknE, and PknH belong to Clade II, representing the group of integral membrane receptor and cytoplasmic kinases. Clade III STPKs comprise PknF, PknI, and PknJ. Finally, Clades IV and V represent the group of soluble kinases that cluster along with PknK and PknG, respectively (20).

PknA and PknB are essential for sustaining the growth and survival in *in vitro* and *in vivo* (11, 19, 21) while modulating cell morphology (22). PknG, a well-studied mycobacterial STPK, is involved in metabolism regulation (7, 23) and promotes the intracellular survival of mycobacteria within macrophages by blocking phagosome-lysosome fusion (24). The remaining kinases have proposed roles in various biological processes, including the regulation of growth and cell division, cell wall synthesis, environmental sensing, and adaptive responses to different abiotic stresses (25–29).

These STPKs, along with their cognate phosphatases, interact functionally to generate a complex signaling architecture (6, 13, 30). Notably, kinases are distinguished from one another based on their substrate repertoires and their modes of regulation (31). Mapping kinase-substrate relationships is, therefore, key to understanding fundamental biochemical pathways and signaling networks in mycobacteria and identifying new, plausible pharmaceutical targets for therapeutic intervention (1).

However, despite the increase in the definition of the functional roles of these individual STPKs, the physiologically relevant substrates and networks responsible for integrating cellular signals remain obscure. Various techniques, including genetic screening, *in vitro* kinase assays, protein microarrays, two hybrid-based studies and proteomics, have been described for the screening of kinase substrates (32). These approaches often attempt to mimic *in vivo* contexts to provide useful insights, however, validation is still required. Consequently, improved methods and novel strategies for mycobacterial STPK substrate detection are essential.

The sensitivity and high-throughput capacity of MS have improved significantly in recent years and proved to be of considerable benefit to the detection of phosphorylation events. MS-based methods now surpass classical approaches, especially since MS is able to confidently assign the localization of phosphosites (7–9). Coupled with the use of kinase knock-out mutants or chemical inhibitors, MS also allows for the quantitative analysis of phosphorylation changes in the presence or absence of a specific kinase.

Quantitative phosphoproteomic strategies have been successfully applied to identify novel substrates of protein kinases in bacteria such as *Bacillus subtilis* (33, 34) and *Escherichia coli* (35), but also in higher organisms, including yeast (36) and human cells (37, 38). More recently, high-resolution MS has been applied for the identification of STPK physiological substrates and interacting partners in mycobacteria. Accordingly, this review will hereafter focus on quantitative MS-based mycobacterial STPK substrate identification.

**MS-Based Ser/Thr Phosphoproteomic Analysis—Classical Methods for STPK Substrate Identification**—Technical developments in MS, together with its increasing accessibility, has enabled the refinement of proteomic strategies for the identification of novel physiological substrates of mycobacterial STPKs (4, 7–9, 39). Commonly utilized methods combine *in vitro* kinase assays with MS to detect kinase substrates. This often entails the use of purified, active protein kinases to phosphorylate proteins or cell lysates in *in vitro*, followed by protein purification and MS analysis for phosphopeptide identification. Applying this strategy, the forkhead-associated (FHA) domain-containing protein, GarA, was identified as a substrate of PknG in *in vitro* (40). Employing a similar proteomic approach, in combination with 2D-PAGE and autoradiography, Rv2175c and GarA were found to be substrates of PknL and PknB, respectively (41, 42). This method allowed for the differentiation of several protein species and highlighted the complementarity of 2D-PAGE and MS for increasing proteome coverage.

An alternative approach, termed KESTREL, uses fractions of cell lysates to separate endogenous protein kinases from their various substrates, thereby minimizing background phosphorylation and sample complexity for subsequent MS
Mueller and Pieters (23) presented a KESTREL-based approach for mycobacterial STPK substrate identification combining proteome fractionation, in vitro kinase assays and reversed-phase chromatography with 2D-PAGE and tandem MS protein identification (23). Using this method, GarA was, again, identified as a substrate for mycobacterial PknG. However, as demonstrated in previous studies, high enzyme-to-substrate ratios are required to detect phosphorylation using in vitro assays (1, 23, 40). This limitation suggests that these assays may not capture and reflect the true biological selectivity of protein kinases, which are also known to display promiscuity in vitro. The physiologically relevant specificity of these kinases is often lost because of the elevated, nonphysiological kinase and/or substrate concentrations used in kinase assays. Thus, mycobacterial phosphoproteomics research should ideally aim toward in vivo phosphorylation network reconstruction in order to link STPKs with their physiological substrates and binding partners.

Recent Advances in Mycobacterial STPK Substrate Identification—Various quantitative methods for ex vivo LC-MS/MS analysis have been successfully applied to generate phosphoproteomic datasets for mycobacterial species (12, 39, 44–46). For example, Prisic et al. reported global Ser/Thr phosphorylation profiles of the M. tuberculosis laboratory strain H37Rv grown under conditions of hypoxia, nitric oxide stress, oxidative stress, and using glucose or acetate as a carbon source (39). MS analysis of these samples enabled the detection of 506 phosphosites on 301 proteins involved in a vast array of functions. Similar numbers of phosphoproteins were reported in other M. tuberculosis strains (12, 47) and other mycobacterial species (45, 46). However, the high-throughput capability of MS and phosphoproteome profiling is often unable to determine an exact relationship between protein kinases and their substrates due to the presence of multiple different active kinases in a given biological system. Consequently, more laborious methods are required to associate these substrates with their cognate protein kinases (4).

Recently, Nakedi et al. employed label-free quantitative phosphoproteomics to screen for novel physiological substrates of M. bovis BCG PknG in vivo (7). The study compared the phosphoproteomic dynamics of the batch culture growth site; PTM, post-translational modification; Ser/Thr, serine/threonine; STPK, serine/threonine protein kinase; TB, tuberculosis; TiO2, titanium dioxide.
of *M. bovis* BCG wild-type against the respective PknG knock-out mutant strain. Titanium dioxide (TiO2) beads were used for phosphopeptide enrichment, followed by LC-MS/MS (Fig. 1). This ex vivo workflow facilitated the identification of 55 differentially phosphorylated phosphopeptides, of which 21 phosphopeptides were phosphorylated only in the *M. bovis* BCG wild-type and not in the PknG knock-out mutant. Phosphopeptides were identified by mapping acquired peptide MS/MS spectra against the *M. bovis* BCG Pasteur 1172 reference proteome using the MaxQuant suite, Andromeda (48). Modified peptides underwent score-based filtering, including an Andromeda search score >40, delta score >8, false discovery rate <0.05 and a localization probability ≥0.75 (Fig. 2).

A selected number of these novel candidate PknG substrates were further validated through targeted parallel reaction monitoring MS assays. Notably, this study exemplified the utility of parallel reaction monitoring as a new means to validate phosphorylation events in phosphoproteomic experiments (49). Following this publication, several other studies employed similar approaches in the search of mycobacterial protein kinase substrates (6, 8, 9, 50, 51). Gil et al. developed an affinity purification-MS strategy to stepwise recover PknG substrates and interactors (8). Using this tailored interactome approach, seven novel candidate PknG substrates were identified, of which six substrates overlapped with those previously reported by Nakedi et al. (7). In addition, this study recovered the only two previously well-characterized PknG substrates, namely GarA and the 50 ribosomal protein L13, thereby substantiating the notion that these proteins represent plausible physiological substrates or interactors of PknG.

In order to identify potential PknA and PknB targets, Carette et al. opted for a phosphoproteomic approach using a small-molecule inhibitor of PknA and PknB (52). They observed differential levels of protein phosphorylation and identified 46 candidate substrates of PknA and PknB. Kaur et al. identified 73 potential targets of PknB using a tandem mass tag-based...
Fig. 3. **Top 10 significantly enriched biological process GO terms of novel candidate PknA, PknB, and PknG substrates.** (A) Bar chart ranking of the significantly enriched biological process GO terms of novel candidate PknA and PknB substrates identified by Carette et al. (52) and Kaur et al. (9). (B) Bar chart ranking of the significantly enriched biological process GO terms of novel candidate PknG substrates identified by Nakedi et al. (7) and Gil et al. (8). Details of the biological process GO enrichments are presented in Table S3.
quantitative phosphoproteomics approach with a PknB tetra mutant, wherein all four putative ligand interacting residues were modified. Simultaneous mutation of these penicillin-binding proteins and Ser/Thr kinase-associated 3–4 linker region residues appeared to abrogate ligand binding (9). Nine of these 73 targets overlapped with those reported previously by Carette et al. (52), providing partial independent verification of this significantly expanded list of candidate PknB substrates. However, the general lack of overlap between the candidate PknB substrates reported by these two studies is notable—even when accepting that the data presented by Carette et al. did not distinguish between PknA and PknB.

Fig. 4. Top 10 significantly enriched cellular component GO terms of novel candidate PknA, PknB, and PknG substrates. (A) Bar chart ranking of the significantly enriched cellular component GO terms of novel candidate PknA and PknB substrates identified by Carette et al. (52) and Kaur et al. (9). (B) Bar chart ranking of the significantly enriched cellular component GO terms of novel candidate PknG substrates identified by Nakedi et al. (7) and Gil et al. (8). Details of the cellular component GO enrichments are presented in Table S4.
substrates. This suggests that further research is required to verify the full sets of candidate substrates to minimize overreporting and to identify the true physiological PknB substrates.

Notwithstanding caveats concerning the need for robust validation of new candidate substrates for individual kinases, LC-MS/MS phosphoproteomic-based workflows have proven to be the most promising strategy for the ex vivo identification of novel STPK substrates in mycobacteria thus far. Moreover, the cost demand of these workflows can be minimized using label-free quantitation, allowing for the investigation of multiple time points and/or conditions. Above all, these approaches allow for substrate discovery under more biologically relevant conditions. In contrast, in vitro strategies ideally need to factor in subcellular compartmentalization of proteins, as well as physiologically relevant enzyme/substrate concentrations, which ultimately determine whether a protein kinase can interact with, and subsequently phosphorylate, a putative substrate. In vitro kinase assays can, however, serve as a validation technique for novel putative substrates recovered during in vivo phosphoproteomic studies. As part of this review, we have created a manually curated list of all currently reported mycobacterial kinase substrates and their specific phosphosites (Table S1). Table I presents a truncated table of known mycobacterial PknG substrates.

**Integration of Novel Substrates Identified by LC-MS/MS:**

Functional and Localization Enrichment Analyses—The roles of PknA, PknB, and PknG in the cell—Novel candidate PknA, PknB and PknG substrates were analyzed for GO functional and localization enrichment. The putative PknG substrates reported by Nakedi et al. (7) and Gil et al. (8) were converted to *M. tuberculosis* (strain ATCC 25618/H37Rv) orthologs using Orthologous Matrix Browser (https://omabrowser.org) (53). Potential targets of PknA and PknB identified in Carette et al. (52) and Kaur et al. (9) were classified using Mycobrowser (https://mycobrowser.epfl.ch/) (54). The corresponding protein identifiers (Table S2) were used to perform statistical overrepresentation tests of GO biological processes and GO cellular components using the PANTHER Classification System (http://pantherdb.org) version 14.1 (55) and GO database version 1.2 (released 2019-07–03). The *M. tuberculosis* (strain ATCC 25618/H37Rv) proteome was used as a reference list.

Fig. 3A shows the functions that were enriched in the PknA and PknB candidate substrate list, with a clear enrichment of protein folding, cellular macromolecule biosynthetic process, regulation of molecular function, and gene expression. Moreover, the cellular component enrichment suggests that these substrates are localized in the cell wall, external encapsulation structure, cell periphery, and plasma membrane (Fig. 4A). Both of these results reaffirm the role of PknA and PknB in modulating cell division, cell wall synthesis, cell morphology, and metabolic processes (19, 21, 22, 56–58), while corroborating the localization of these kinases to the cell wall and cell membrane (21, 59).

The biological processes that were highly enriched in the PknG candidate substrate list were protein folding, response to temperature stimulus, and the regulation of various processes, including protein maturation, plasminogen activation, and protein processing (Fig. 3B). These findings agree with
the conclusions drawn in previous studies that suggest PknG to be involved in the regulation of protein folding and translation (7, 8). The cellular component enrichment shows these substrates to be considerably localized in the cytosol, cell wall, and capsule (Fig. 4B). This agrees with PknG being a soluble STPK and cytosolic protein, which is hypothesized to be translocated under specific environmental conditions (60, 61). Furthermore, these results also support the localization of PknG to the cell wall (61–63) and its proposed regulatory role in cell wall biogenesis and integrity (62, 64).

Baer et al. suggested that mycobacterial STPKs should phosphorylate substrates localized in the same subcellular compartment in vivo (30). Therefore, we classified the candidate novel substrates of PknA, PknB (9, 52), and PknG (1, 2) according to UniProt subcellular locations. In cases where no annotation was recorded or ambiguity existed, the tuberculosis-pred online tool (www.tbpred.com) was used. Fig. 5 illustrates the subcellular localization of the MS identified novel candidate PknA, PknB (9, 52), and PknG (7, 8) substrates.

The subcellular localization of a substrate influences its ability to access protein kinases. PknA and PknB are both transmembrane proteins comprising an intracellular kinase domain (18), thereby accounting for the higher percentage of cytosolic protein substrates than membrane protein substrates. Nevertheless, it remains important to determine the circumstances by which a substrate becomes available to interact with a kinase anchored to the cell membrane.

PknG demonstrated preferential phosphorylation of cytoplasmic proteins, as anticipated. However, to understand the recovery of membrane and membrane-associated proteins by Nakedi et al. (7), we analyzed these substrates using the transmembrane topology predictor, Phobius (http://phobius.cgb.ki.se) (65). Intriguingly, none of the detected phosphorylation site motifs were analyzed using a foreground composed of available sequence windows extracted from the studies cited in Table II. The M. tuberculosis (strain ATCC 25618/H37Rv) proteome was used as the background database. The red horizontal lines (± 4.06) illustrate the relative statistical significance (p value ≤ 0.05, after Bonferroni correction) of residues flanking the central phosphorylation site. Overrepresented residues are above the midline, whereas underrepresented residues are below the midline. No distinct motifs were observed for phosphorylation on either Thr or Ser sites.

Characterizing Substrate Specificity Patterns in Mycobacteria—Mycobacterial STPKs have long been suspected of promiscuity toward various substrates, a notion supported in part by the aggregated data on reported mycobacterial substrates (Table S1), which suggests an unexpectedly large number of substrates for each mycobacterial kinase. This
contrasts with the tight specificity typically observed for eu-
karyotic STPKs that underpins the integrity of eukaryotic sig-
naling networks. The apparent promiscuity of mycobacterial
STPKs is further supported by the absence of any obvious
sequence motifs surrounding preferential Ser/Thr phosphory-
lation sites for individual STPKs (Fig. 6). However, recent
phosphoproteomic-based studies show that a subset of sub-
strates is specifically phosphorylated in vivo only when in the
presence of STPK activity (7–9, 52), arguing that these STPKs
demonstrate a preference for certain substrates. Furthermore,
it is equally clear from the GO enrichment analyses previously
presented that the substrates for the individual mycobacterial
STPKs are nonrandom at the functional level. Future attempts
to determine possible specificity across mycobacterial STPK
substrates may, therefore, require the use of more sophisti-
cated Hidden Markov model-type approaches to identify po-
tentially distal structural motifs that can account for the ob-
served substrate selectivity of the mycobacterial STPKs.

**TABLE II**

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Gene</th>
<th>UniProt ID</th>
<th>P-site</th>
<th>Species</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>30S ribosomal protein S16</td>
<td>rpsP</td>
<td>A1KM03</td>
<td>S162</td>
<td>M. bovis BCG</td>
<td>(7)</td>
</tr>
<tr>
<td>4HBT domain-containing protein</td>
<td>BCG_1584c</td>
<td>A0A0H3M6A1</td>
<td>S2</td>
<td>M. bovis BCG</td>
<td>(7)</td>
</tr>
<tr>
<td>4-hydroxy-3-methylbut-2-en-1-yl diposphate synthase (flavodoxin)</td>
<td>ispG</td>
<td>A1KML3</td>
<td>S387</td>
<td>M. bovis BCG</td>
<td>(7)</td>
</tr>
<tr>
<td>50S ribosomal protein L2</td>
<td>rplB</td>
<td>A1KGI5</td>
<td>S32</td>
<td>M. bovis BCG</td>
<td>(7)</td>
</tr>
</tbody>
</table>
| 5-methyltetrahydroperoyltriglutamate–
homocysteine methyltransferase                     | metE       | A1KHS4     | S713   | M. bovis BCG     | (7)       |
| 60 kDa chaperonin 1 (GroL1)                       | groL1/MSMEG_0880 | A0QQ05 | –      | M. smegmatis     | (8)       |
| AAA domain-containing protein                     | BCG_3931   | A0A0H3M7C9 | S277   | M. bovis BCG     | (7)       |
| Acetyl-/propionyl-coenzyme A carboxylase alpha chain | MSMEG_1807 | A0QTE1    | –      | M. smegmatis     | (8)       |
| Alcohol dehydrogenase, iron-containing            | MSMEG_6242 | A0R5M3     | –      | M. smegmatis     | (8)       |
| Antitoxin                                         | BCG_00306c | A0A0H3M1S6 | T152   | M. bovis BCG     | (7)       |
| Antitoxin                                         | BCG_3207c  | A0A0H3M8A1 | T109   | M. bovis BCG     | (7)       |
| ATP synthase subunit beta                         | atpD       | A1KI98     | S16    | M. bovis BCG     | (7)       |
| CbiA domain-containing protein                    | BCG_0573   | A0A0H3M2H1 | T25    | M. bovis BCG     | (7)       |
| Chaperone protein CipB                            | clpB       | A0A0H3M7W9 | T79    | M. bovis BCG     | (7)       |
| Chaperone protein DnaK                            | dnaK       | A1KHF2     | T391   | M. bovis BCG     | (7)       |
| Chaperone protein DnaK                            | MSMEG_0709 | A0QQC8     | –      | M. smegmatis     | (8)       |
| Conserved hypothetical proline and threonine-rich protein | BCG_0352 | A0A0H3M3A7 | S403   | M. bovis BCG     | (7)       |
| DNA gyrase subunit A                               | gyrA1      | A0A0G2Q9F8 | S263   | M. bovis BCG     | (7)       |
| Ferredoxin                                        | MSMEG_5122 | A0R211     | –      | M. smegmatis     | (8)       |
| FHA domain protein                                | MSMEG_0035 | A0QNG7     | T18, 116, 137, 277 | M. tuberculosis, M. smegmatis | (8) |
| FHA domain-containing protein                     | BCG_0573   | A0A0H3M2H1 | T25    | M. bovis BCG     | (7)       |
| Glutamine synthetase                              | MSMEG_4290 | A0R079     | T421, 277, 57    | M. bovis BCG, M. smegmatis | (8) |
| Glycogen accumulation regulator                   | BQ2027_MB1858 | P64889, | T21    | M. bovis BCG, M. tuberculosis, M. smegmatis | (8, 23, 40) |
| GarA                                             | garA       | A0A0H3M1S6 | S16    | M. bovis BCG     | (7)       |
| HpcH, HpaI domain-containing protein              | BCG_3100c  | A0A0H3M8P9 | T12    | M. bovis BCG     | (7)       |
| Inorganic pyrophosphatase                         | MSMEG_6114 | A0R597     | –      | M. smegmatis     | (8)       |
| Malate dehydrogenase                              | mdh        | A1KI28     | S238, S239 | M. bovis BCG, M. tuberculosis | (7, 67) |
| Probable conserved membrane protein               | BCG_0236   | A0A0H3M0X2 | S19    | M. bovis BCG     | (7)       |
| Protein kinase domain-containing protein           | BCG_0696c  | A0A0H3M4V4 | S177   | M. bovis BCG     | (7)       |
| RNA polymerase-binding protein                    | rbpA       | A0A0H3M6B6 | T18    | M. bovis BCG     | (7)       |
| UDP-N-acetylmuramate–L-alanine ligase (MurC)       | murC       | P9WJL7     | –      | M. tuberculosis   | (64)      |
| Uncharacterized protein                           | BCG_1803c  | A0A0H3M4P0 | S11    | M. bovis BCG     | (7)       |
| Uncharacterized protein                           | BCG_2052c  | A0A0H3M7J9 | T7     | M. bovis BCG     | (7)       |
**Future Perspectives**—Quantitative phosphoproteomics is now established as a powerful and reliable strategy for mapping connections between phosphorylated substrates and their respective kinases or phosphatases. As discussed in the preceding, the employment of LC-MS/MS-based strategies has dramatically increased the number of potential substrates for mycobacterial STPKs, specifically for PknB and PknG. While this represents a unique opportunity to better understand the regulatory role of STPKs, it is important to note potential pitfalls associated with such approaches. For instance, it could be argued that differentially phosphorylated proteins may not be direct substrates of the studied kinase but, rather, phosphorylated by another kinase whose activity was affected by the absence of the kinase and/or the presence of a kinase-specific inhibitor. It is, therefore, imperative to further validate these novel substrates. This may include the use of *in vitro* assays to confirm that the newly identified phosphosites are, in fact, phosphorylated by the corresponding STPK. Additionally, in cases where a new role and/or phenotype is associated with the activity of a protein kinase, it is important to determine the phosphorylation site occupancy of the novel substrate(s). Researchers should then, ideally, verify whether the phenotype is altered by the replacement of the substrate phosphorylation site with an alanine or glutamate residue via site-directed mutagenesis (*i.e.*, knock-out or knock-in of a constitutive phosphorylation phenotype).

It is equally important that future studies investigate and validate the biological significance of these findings in order to increase understanding of the regulatory function(s) of newly identified substrate phosphorylation sites. For example, until very recently, the physiological role of PknG was largely associated with the regulation of mycobacterial metabolism, but newer datasets now suggest that PknG phosphorylation activity likely plays an important role in protein processing, translation, and folding machinery (7, 8). Moreover, it is also important to develop phosphoproteomic strategies to investigate the role of mycobacterial STPKs in host cells in order to further reveal the complexity of host-pathogen interactions. In a recent study, 31 proteins were reported as exclusive interactors of PknG using a human proteome microarray (66). This study provides a valuable foundation for further research on host-pathogen protein-protein interactions to reveal the role of PknG during pathogenic conditions. However, the biological relevance of these findings remains to be verified by phosphoproteomic studies that account for the subcellular compartmentalization of proteins and the accessibility of a host substrate to a mycobacterial protein kinase *in vivo*. We, therefore, propose that the next goal in MS-based mycobacterial research should be the identification of host substrates for mycobacterial STPKs during live infections. Work toward this objective is currently underway in our laboratory.

**Concluding Remarks**—We anticipate that a comprehensive integration of the large phosphoproteomic datasets, aimed at complete identification of novel STPK substrates, will provide the scientific community with significant new insight into the network mechanisms by which STPKs regulate mycobacterial environmental responses. In addition, this should allow researchers to pinpoint potential sites for new pharmacological interventions. The phosphorylation status of novel substrates can be utilized as biomarkers of the corresponding STPK activity. This, in turn, may aid in determining the efficiency and mode of action of potential inhibitors targeting specific mycobacterial STPKs. Adopting such strategies could provide a new platform to guide further development of next-generation drugs for clinical applications.

**Acknowledgments**—S.S.B. thanks the NRF for a Master’s bursary. JMB thanks the NRF for a South African Research Chair grant. NCS thanks the South African Medical Research Council for a Junior Research Fellowship.

*This work is based on the research supported in part by the National Research Foundation (NRF) of South Africa (Grant Numbers: 467126 and 95984). The authors declare that they have no conflicts of interest with the contents of this article.*

**REFERENCES**


