



Identification of Arginine Phosphorylation in *Mycobacterium smegmatis*

Emmanuel C. Ogbonna,^{a,*} Henry R. Anderson,^b  Karl R. Schmitz^{a,b}

^aDepartment of Biological Sciences, University of Delaware, Newark, Delaware, USA

^bDepartment of Chemistry & Biochemistry, University of Delaware, Newark, Delaware, USA

ABSTRACT Tuberculosis is a leading cause of worldwide infectious mortality. The prevalence of multidrug-resistant *Mycobacterium tuberculosis* infections drives an urgent need to exploit new drug targets. One such target is the ATP-dependent protease ClpC1P1P2, which is strictly essential for viability. However, few proteolytic substrates of mycobacterial ClpC1P1P2 have been identified to date. Recent studies in *Bacillus subtilis* have shown that the orthologous ClpCP protease recognizes proteolytic substrates bearing posttranslational arginine phosphorylation. While several lines of evidence suggest that ClpC1P1P2 is similarly capable of recognizing phosphoarginine-bearing proteins, the existence of phosphoarginine modifications in mycobacteria has remained in question. Here, we confirm the presence of posttranslational phosphoarginine modifications in *Mycobacterium smegmatis*, a nonpathogenic surrogate of *M. tuberculosis*. Using a phosphopeptide enrichment workflow coupled with shotgun phosphoproteomics, we identified arginine phosphosites on several functionally diverse targets within the *M. smegmatis* proteome. Interestingly, phosphoarginine modifications are not upregulated by heat stress, suggesting divergent roles in mycobacteria and *Bacillus*. Our findings provide new evidence supporting the existence of phosphoarginine-mediated proteolysis by ClpC1P1P2 in mycobacteria and other actinobacterial species.

IMPORTANCE Mycobacteria that cause tuberculosis infections employ proteolytic pathways that modulate cellular behavior by destroying specific proteins in a highly regulated manner. Some proteolytic enzymes have emerged as novel antibacterial targets against drug-resistant tuberculosis infections. However, we have only a limited understanding of how these enzymes function in the cell and how they select proteins for destruction. Some proteolytic enzymes are capable of recognizing proteins that carry an unusual chemical modification, arginine phosphorylation. Here, we confirm the existence of arginine phosphorylation in mycobacterial proteins. Our work expands our understanding of a promising drug target in an important global pathogen.

KEYWORDS *Mycobacterium tuberculosis*, mass spectrometry, phosphoarginine, phosphoproteomics, proteolysis, tuberculosis

Tuberculosis is a leading cause of worldwide infectious mortality, ranking above HIV/AIDS and Ebola, and is one of the top 10 leading causes of death overall (1). Advances in diagnosis, vaccinations, and therapeutics have reduced tuberculosis morbidity and mortality. However, the prevalence of multidrug resistance in the causative bacterium, *Mycobacterium tuberculosis*, remains high (1). These statistics underscore the urgent need to discover new drugs and exploit new molecular targets. One promising target is the mycobacterial Clp protease. Several studies show that Clp protease components are strictly essential for *M. tuberculosis* viability (2–6) and are viable targets for anti-*M. tuberculosis* therapeutics (7–12).

Clp proteases mechanically unfold and destroy native cytosolic proteins (13, 14). These large enzymatic complexes consist of a core peptidase and an associated

Editor Emily Weinert, Penn State University

Copyright © 2022 Ogbonna et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Karl R. Schmitz, schmitzk@udel.edu.

*Present address: Emmanuel C. Ogbonna, Harvard Program in Therapeutic Science (HiTS), Harvard Medical School, Boston, Massachusetts, USA.

The authors declare no conflict of interest.

Received 1 June 2022

Accepted 19 September 2022

Published 10 October 2022

hexameric ATP-dependent unfoldase. The mycobacterial Clp peptidase, ClpP1P2, is a heteromer composed of distinct ClpP1 and ClpP2 rings that stack face to face to create a barrel-shaped tetradecamer (15–19). ClpC1 is one of two mycobacterial unfoldases (the other is ClpX) that can assemble with ClpP1P2 to form a functional protease. ClpC1 is an 848-residue protein with a globular N-terminal domain (NTD) and two AAA+ (ATPases associated with various cellular activities) modules. Ring-shaped hexamers of ClpC1 dock coaxially on the surface of the ClpP2 face of the peptidase (20, 21). Protein substrates destined for destruction by ClpC1P1P2 must first be recognized by ClpC1. The unfoldase therefore plays a critical regulatory role in proteolysis.

Clp proteases have become major targets of novel drug development against *M. tuberculosis* and other pathogenic bacteria. Several classes of antimicrobials specifically target the Clp peptidase, including dysregulators (e.g., acyldepsipeptides [22, 23]), activators (e.g., sclerotiamide [24]), and catalytic inhibitors (e.g., β -lactones and boronate compounds [7, 25]). Importantly, multiple cyclic peptides have been discovered that exhibit anti-*M. tuberculosis* activity through dysregulation of ClpC1, including cyclomarin A, lassomycin, ecumicin, rufomycin, and metamarin (8, 9, 26–28). These compounds bind to the ClpC1 NTD and disrupt proteolysis by uncoupling unfoldase activity from proteolysis or by causing uncontrolled degradation of cellular proteins. The antimicrobial activity of ClpC1-targeting compounds underscores the importance of ClpC1P1P2 in mycobacteria. However, the specific proteolytic functions responsible for its essentiality remain obscure, and only a few protein substrates have been identified (29–32).

An expanded understanding of the physiological roles played by ClpC1P1P2 would bolster efforts to develop antimicrobial compounds. In other bacteria, Clp proteases participate in cellular processes ranging from targeted pathway regulation to cell-wide protein quality control (33, 34). Multiple mechanisms of substrate recognition have been described, including direct interactions with the unfoldase (30, 31) and indirect recognition with the aid of adaptors (32, 35). Recent studies with *Bacillus subtilis* and related *Firmicutes* have demonstrated that posttranslational arginine phosphorylation marks some proteins for destruction by ClpCP (34, 36, 37). Dual phosphoarginine (pArg) binding sites on the *B. subtilis* ClpC NTD allow ClpCP to recognize phosphoarginylated proteins as proteolytic substrates (34). Phosphoarginine-mediated proteolysis in these bacteria is upregulated during stress through activation of the arginine kinase McsB (38). Phosphoproteomic studies have uncovered widespread arginine phosphorylation during heat stress, implicating ClpCP in turnover of misfolded proteins (34, 36, 39). Additionally, McsB regulates the global stress response through targeted phosphorylation of the negative transcriptional regulators CtsR and HrcA (36, 38, 40). Proteolysis of these targets by ClpCP allows transcriptional activation of stress response genes (38). Interestingly, ClpC and ClpP are themselves targets of McsB, and specific pArg sites on ClpC are required for its activation by McsB (36, 41). These studies underline the role of arginine phosphorylation in *Firmicutes* as both a degradation signal and a regulatory mechanism.

Several lines of evidence suggest that an analogous pArg-mediated proteolytic pathway exists in mycobacteria. Sequence and structural data reveal overall homology between the NTDs of *M. tuberculosis* ClpC1 and *B. subtilis* ClpC, as well as strong conservation of the residues surrounding the pArg-binding sites (34, 42). Moreover, *in vitro* experiments confirm that the *M. tuberculosis* ClpC1 NTD does indeed interact with both free pArg and with arginine-phosphorylated model substrates (42). However, the existence of pArg-mediated proteolysis in mycobacteria remains in question, as no mycobacterial McsB homologs are known and phosphoarginine modifications have not yet been described for these bacteria.

Here, we confirm the existence of posttranslational phosphoarginine modifications in *Mycobacterium smegmatis*, a nonpathogenic surrogate of *M. tuberculosis*. Using a phosphopeptide enrichment workflow coupled to shotgun phosphoproteomics, we identified arginine phosphosites on several protein targets within the *M. smegmatis* proteome. Our findings suggest that these modifications are widespread among actinobacterial species.

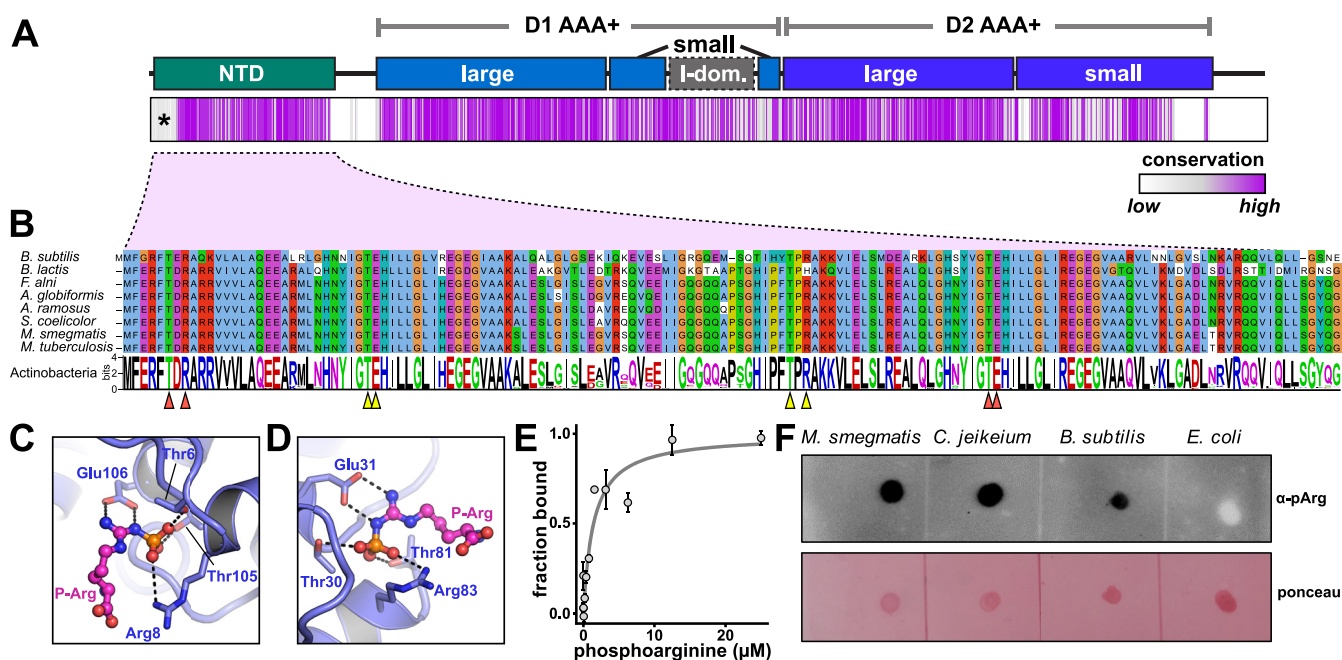


FIG 1 Indirect evidence for phosphoarginine in actinobacteria. (A) The domain organization of ClpC1 is shown above a plot of amino acid conservation among actinobacterial ClpC1 orthologs, where conservation at each alignment position is plotted as a purple vertical strip. The apparent poor conservation at the beginning of the NTD (marked by an asterisk) likely reflects misannotation of the start site in some entries. (B) The sequence of the *Bacillus subtilis* ClpC NTD is shown above equivalent regions from several actinobacterial ClpC orthologs. The sequence logo below shows amino acid conservation in actinobacterial ClpC1 orthologs. Arrows mark positions reported to be important for phosphoarginine binding to site 1 (orange) or site 2 (yellow) (34). Phosphoarginines from the crystal structure of the *B. subtilis* ClpC NTD (PDB code 5HBN) were modeled on the *M. tuberculosis* ClpC1 NTD (PDB code 6PBQ) putative phosphoarginine-binding sites 1 (C) and 2 (D) (47). (E) Binding of phosphoarginine to *M. smegmatis* ClpC1^{NTD} was measured by microscale thermophoresis. Data were fit to a noncooperative binding model (gray curve), yielding a K_D (equilibrium dissociation constant) of $1.4 \pm 0.6 \mu\text{M}$. Values are averages of three technical replicates ± 1 SD. (F) Dot blots of *Mycobacterium smegmatis*, *Corynebacterium jeikeium*, *Bacillus subtilis*, and *Escherichia coli* cell lysates, probed with anti-phosphoarginine antibody (48) or stained with Ponceau S.

RESULTS

Phosphoarginine binding sites in the ClpC1 NTD are conserved across Actinobacteria.

ClpC1 is an essential enzyme in *M. tuberculosis* and *M. smegmatis* (2, 6), yet few specific cellular roles or proteolytic substrates of the ClpC1P1P2 protease are known. To gain insight into its cellular function, we began by examining ClpC1 sequence conservation across the phylum *Actinobacteria* by constructing an alignment of 1,195 actinobacterial orthologs (Fig. 1A). Like most other type-II Clp unfoldases, ClpC1 possesses an ~150-amino-acid (aa) N-terminal domain (NTD), which likely participates in regulation and substrate recognition, but lacks mechanical function or a direct role in catalysis (35, 43). Surprisingly, the NTD is as well conserved as or better conserved than the D1 and D2 AAA+ ATPase rings, suggesting that the NTD has important and conserved functions across *Actinobacteria*. This may help explain why naturally occurring antibiotics that target and dysregulate ClpC1 have evolved to specifically bind the NTD (8, 9, 26, 44–46).

In *Bacillus subtilis* and other *Firmicutes*, ClpCP recognizes phosphorylated arginine residues as degradation signals via twin binding sites on opposite ends of the ClpC NTD (34). Examination of actinobacterial ClpC1 sequences revealed that the residues known to be important for phosphoarginine binding in *B. subtilis* are conserved across actinobacterial orthologs (Fig. 1B). We found that we could readily model bound phosphoarginines in an existing structure of the *M. tuberculosis* ClpC1 NTD (47), with minor side chain rearrangements, based on the observed mode of binding to the *B. subtilis* ClpC NTD (34) (Fig. 1C and D). Finally, we directly assessed binding of pArg to purified *M. smegmatis* ClpC1^{NTD} by microscale thermophoresis (Fig. 1E). Phosphoarginine bound with an affinity of $1.4 \mu\text{M}$, similar to the $\sim 5 \mu\text{M}$ affinity previously reported for the *M. tuberculosis* ClpC1 NTD (42). The ability of the ClpC1^{NTD} to bind pArg and the conservation of pArg binding

modules across *Actinobacteria* provide strong indirect evidence for the existence of this posttranslational modification in this phylum.

Identification of phosphoarginine modifications in *Mycolicibacterium smegmatis*. To test for the physiological existence of arginine phosphorylation, we probed dot blots of several bacterial lysates with a phosphoarginine-specific antibody (Fig. 1F) (48). In line with prior studies, signal was detected in *B. subtilis* lysate (48) but not in lysate from *Escherichia coli*, which lacks arginine phosphorylation (49, 50). We also observed a positive immunoblot reaction in lysates from *M. smegmatis* and a second actinobacterium, *Corynebacterium jeikeium*. These observations suggest that pArg modifications occur at least within the suborder *Corynebacterineae*, which encompasses *Corynebacterium*, *Mycolicibacterium*, and *Mycobacterium*.

To determine which specific cellular proteins carry pArg modifications in *M. smegmatis*, we employed an unbiased shotgun proteomics approach. Phosphoarginine is acid labile and has a short half-life in the acidic trifluoroacetic acid (TFA)-containing solvent systems typically used for proteomic sample preparation and liquid chromatography-tandem mass spectrometry (LC-MS/MS) (36). Optimized protocols for arginine phosphoproteomics have been reported that utilize alternative solvent systems at pHs of ≥ 4 for most steps (34, 36, 39). We adapted these methods to our workflow for sample preparation, phosphopeptide enrichment, and LC-MS/MS, thereby minimizing phosphoarginine hydrolysis (Fig. 2A). Since stress conditions such as heat shock upregulated the occurrence of this modification in *B. subtilis* (36), we analyzed lysates from *M. smegmatis* cultures grown either at normal growth temperature (37°C) or under heat stress. In an initial test of *M. smegmatis* heat tolerance (see Fig. S1 in the supplemental material), we found that cultures grown at 50°C saturated at a lower density but retained substantial viability up to ~ 20 h. Thus, to increase the likelihood of a robust heat shock response, heat-stressed cultures grown at 50°C for 21 h.

A similar number of proteins ($\sim 3,000$) were observed in samples from normal growth (3,147, 3,004, and 3,145 proteins in the three replicates) and heat stress conditions (2,969, 2,724, and 2,899 proteins) (Fig. 2B). This suggests that at this level of stress, compensatory stress responses do not involve a complete shutdown of translational machinery or dramatic upregulation of proteolytic activity. Typically, a major part of the stress response to heat shock is the increased expression of heat shock proteins or chaperones. We evaluated the differential levels of observed proteins based on Sequest HT scores. As shown in Fig. 2C, several proteins were observed at significantly higher levels (P value ≤ 0.05) in heat shock than under normal growth conditions. These included spore protein (Msmeg_5611), which belongs to the Hsp20 small heat shock protein family and was enriched over 2-fold upon heat shock. Other chaperone proteins had slightly higher scores in heat-stressed samples than in normal growth, including DnaJ1, DnaK, and the SecB-like chaperone SmegB. We also performed comparative Gene Ontology (GO) annotation analysis of strongly enriched/depleted proteins, whose levels differed significantly (ratio > 2 ; P value ≤ 0.05) between conditions (Fig. S2; Table S1). Proteins involved in amino acid, lipid, and noncanonical metabolism, along with redox proteins, were strongly depleted during heat stress. Spore protein was the only stress response protein strongly enriched under heat shock; no protein involved in the stress response was strongly depleted. This differential expression suggests the activation of canonical heat shock response in these samples.

Using Proteome Discoverer, we identified arginine-phosphorylated sites in proteins from these samples. Only peptides with a phosphosite localization probability (PhosphoRS/ptmRS score) of $\geq 75\%$ were selected (51). We unambiguously localized six phosphoarginine sites in six different proteins (Fig. 3A; Table S2; Fig. S3). Most pArg sites were observed in multiple replicates, in both heat-stressed and unstressed samples (Fig. 3A). The existence of phosphoarginine modifications in our samples implies the existence of an unidentified *M. smegmatis* pArg kinase. Surprisingly, most pArg sites were observed in both stressed and unstressed samples (Fig. 3A), suggesting that at least these heat stress conditions do not upregulate arginine phosphorylation, in contrast to the paradigm observed in *B. subtilis* (36).

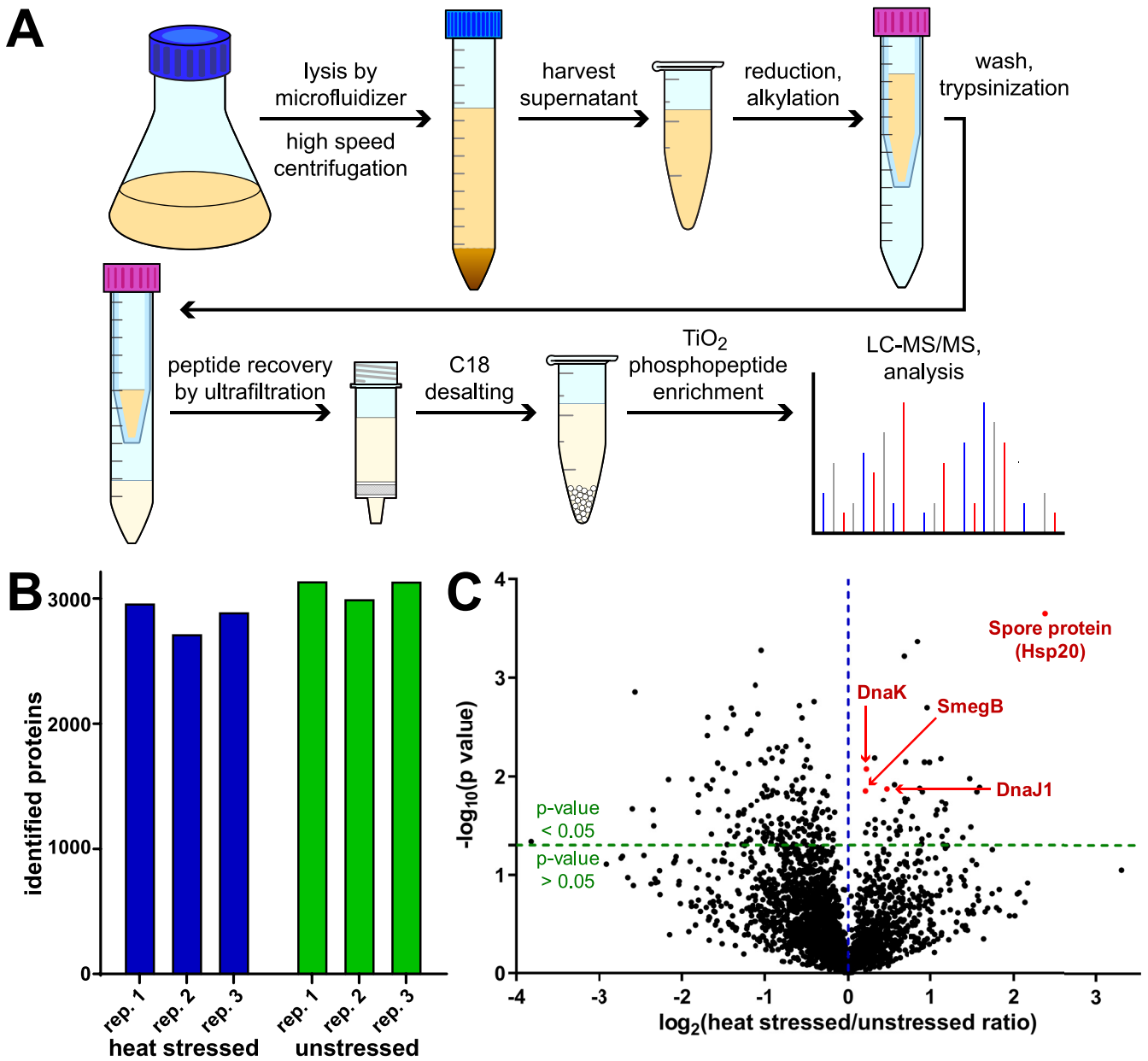


FIG 2 Comparative analysis of mass spectrometry output. (A) Workflow for arginine phosphoproteomics in *Mycobacterium smegmatis*. (B) Number of proteins obtained across biological replicates from cells grown under heat stress (50°C) and normal (37°C) conditions. (C) Volcano plot showing enrichment of proteins observed by LC-MS/MS in heat stress versus normal growth conditions. Plotted on the x axis is the \log_2 of the ratio of average Sequest HT score of stressed to unstressed samples. The y axis shows $-\log_{10}$ of P values obtained by Student's t test. The horizontal green line indicates a cutoff P value = 0.05; the vertical blue line indicates score ratio of 1. Highlighted in red are notable chaperone proteins with stressed/unstressed ratio of >1.

Functional characterization of arginine-phosphorylated proteins. We examined the known or predicted functions of the six proteins observed to carry pArg modifications and found that they have diverse functions, with no obvious common physiological role (Fig. 3A). PhoU1 is a central regulator of the SenX3-RegX3 two-component system responsible for uptake of inorganic phosphate (P_i) during phosphate starvation (52). This raises the possibility that pArg modifications are linked to P_i availability. Two other pArg-bearing targets were metabolic enzymes: the multifunctional 2-oxoglutarate metabolism enzyme Kgd, involved in the tricarboxylic acid cycle (53), and the glycerol dehydratase large subunit (Msmeg_1547), which potentially contributes to the catabolic pathway involving the glycerol dehydration reaction, which yields 3-hydroxypropanal in the presence of adenosylcobalamin coenzyme (Fig. 3A; Fig. S3) (54, 55). pArg may thus play a role in regulating

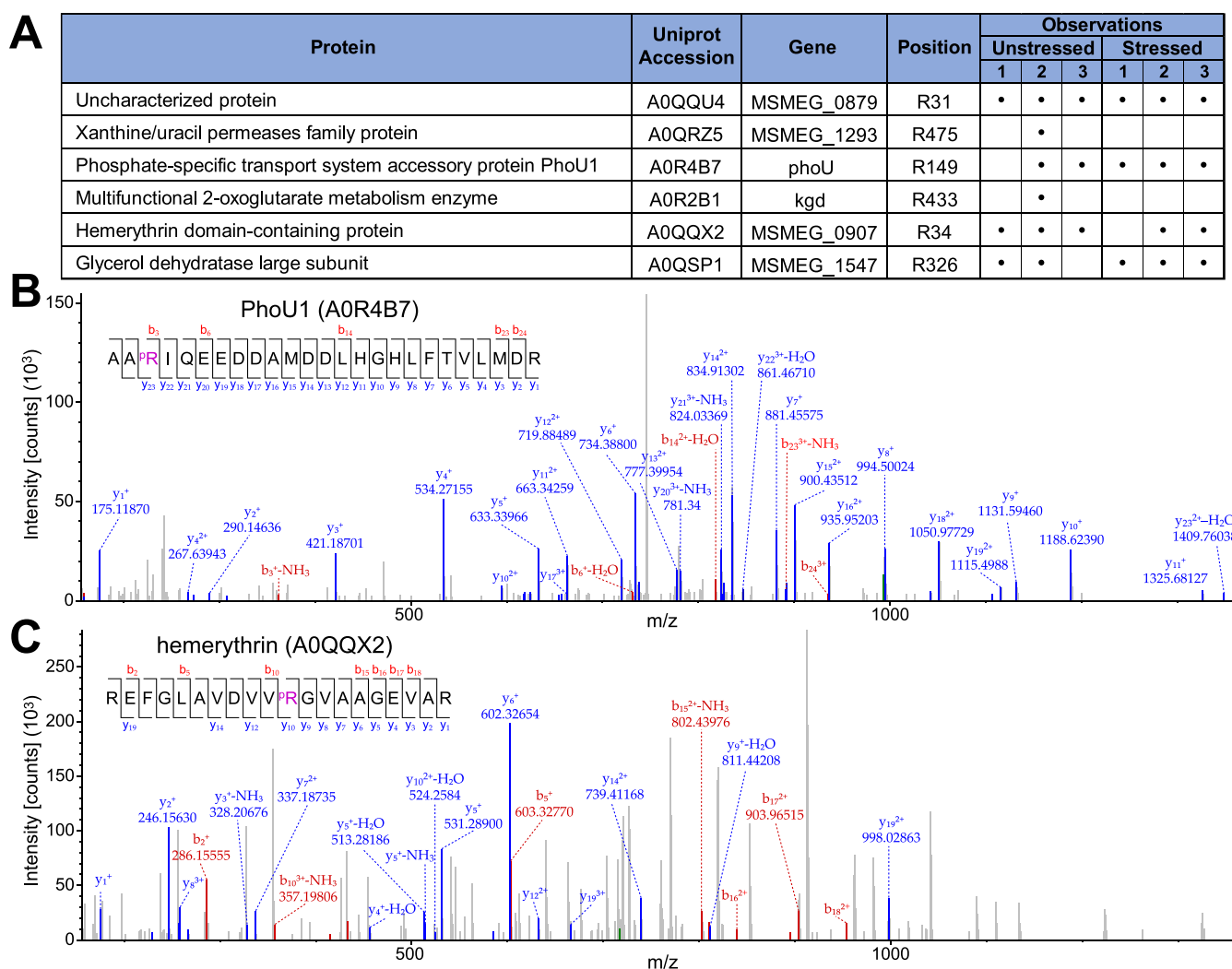


FIG 3 Identification of arginine phosphosites in *M. smegmatis* proteome. (A) Localized arginine phosphosites in six *M. smegmatis* proteins. Localization probability is reported as PhosphoRS/ptmRS score. (B and C) Representative secondary fragmentation spectra show arginine-phosphorylated peptides from PhoU1 (B) and hemerythrin domain-containing protein (C). Additional spectra are shown in Fig. S2.

cellular metabolic pathways. Another phosphosite was detected in the transmembrane xanthine/uracil permease (Msmeg_1293), a nucleobase transporter (Fig. 3A and C). A site was found on the hemerythrin domain-containing protein Msmeg_0907, which belongs to a class of O₂-binding proteins involved in signal transduction, response to H₂O₂, oxygen sensing, and nitric oxide reduction (56–58). Finally, a site was observed on Msmeg_0879, a small (48 aa) uncharacterized protein predicted to be predominantly disordered (Fig. 3A and B). Homologs of Msmeg_0879 were detected in close relatives of *M. smegmatis* but are not widely conserved in actinobacteria.

Structural and physicochemical analysis of arginine phosphorylation sites. We assessed whether arginine phosphorylation occurred at sites with particular physicochemical properties. We aligned 21-residue sequence segments centered on each unique phosphoarginine site (Fig. 4A) but observe no clear consensus motif. Ala, Leu, Val, and Gly appear to be common in flanking positions, although we note that these are the four most abundant amino acids in the *M. smegmatis* proteome (13%, 10%, 9%, and 9% of total, respectively) (Table S3). The sequence diversity surrounding pArg positions suggests that target discrimination is guided by characteristics other than primary sequence.

We next examined structural characteristics of phosphorylated positions, based on predicted structures generated by AlphaFold2 (59), except for the small protein Msmeg_0879,

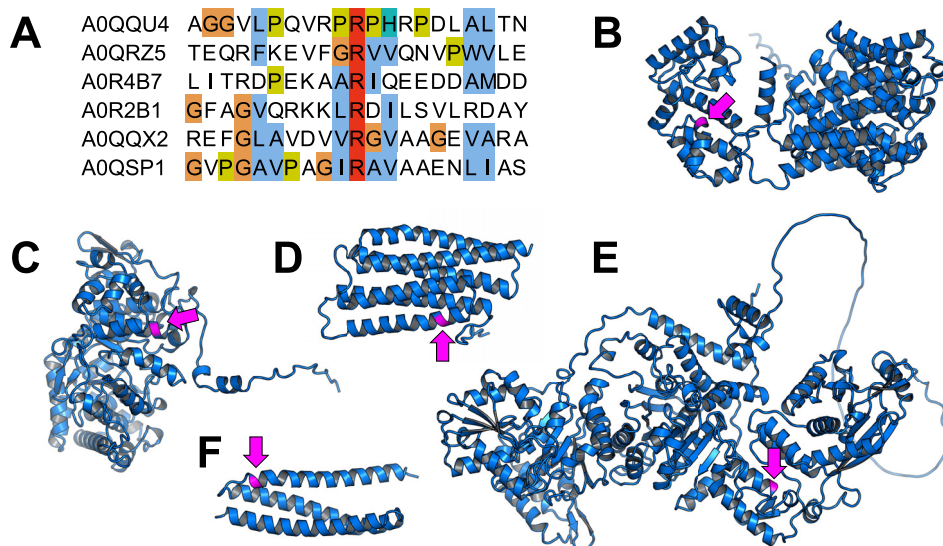


FIG 4 Structural characteristics of phosphoarginine sites. (A) Sequence alignment of 20 residues flanking the arginine phosphosites. (B to F) Structural models of Msmeg_1293 (B), Msmeg_1547 (C), PhoU1 (D), Kgd (E), and Msmeg_0907 (F) are shown with the phosphorylated arginine colored magenta and marked with an arrow. All models were generated by AlphaFold2 (59).

for which structure prediction failed. (Notably, the AlphaFold2 prediction for Kgd was virtually identical to its reported X-ray crystal structure [PDB code 2XT6], with a root mean square deviation [RMSD] of ~ 0.5 Å in the region of the phosphosite [53].) As expected for this charged residue, arginine phosphosites were solvent exposed (Fig. S4) but were located on structured elements rather than loops. Interestingly, in all structures, arginine phosphosites occurred proximal to the beginning or the end of an alpha helix (Fig. 4B to F), which may reflect a structural constraint important for recognition by an arginine kinase.

Finally, we assessed the sequence conservation of arginine at these positions. We aligned the arginine-phosphorylated *M. smegmatis* proteins with homologs within the *Corynebacterineae* suborder and analyzed the positional conservation of arginine at the respective positions. As shown in Fig. 5, conservation varies. The phosphosite arginine was well conserved in Kgd (99.8%), Msmeg_0907 (79%), and Msmeg_0879 (62%) (Fig. 5A to C). Conservation was lower in Msmeg_1547 (34.1%) and PhoU1 (32.63%) (Fig. 5D and E). Arg was rarely present in this position in homologs of Msmeg_1293 (0.23%) (Fig. 5F). In cases where the phosphorylated arginine was well conserved, it may indicate that phosphorylation at this position is a conserved phenomenon with functional or regulatory significance.

DISCUSSION

Protein phosphorylation is a ubiquitous mechanism of signal propagation and pathway regulation in bacteria (60). While examples of His, Asp, Ser, Thr, and Tyr phosphorylation are widespread (61–64), the importance of arginine phosphorylation has become apparent only recently. A growing body of data links pArg modifications to protein quality control pathways and spore germination in *Firmicutes*, including *B. subtilis* (34, 36–38, 41, 42, 48, 65). However, it has remained unclear how prevalent pArg modifications are in other bacterial phyla. Here, we provide direct evidence for the existence of arginine phosphorylation in *Mycobacterium smegmatis*, an actinobacterium.

In *B. subtilis* and *Staphylococcus aureus*, this modification is upregulated by stress (34, 36–38, 41, 42, 48, 65). Hence, we performed our study by examining pArg levels in cells grown under both normal conditions and long-term heat stress. While we saw expected increases in levels of some heat shock proteins overall, arginine phosphorylation abundance was not altered by heat stress, as occurs in *Firmicutes*. It is possible that pArg modifications play no role in the mycobacterial stress response and instead occur in the

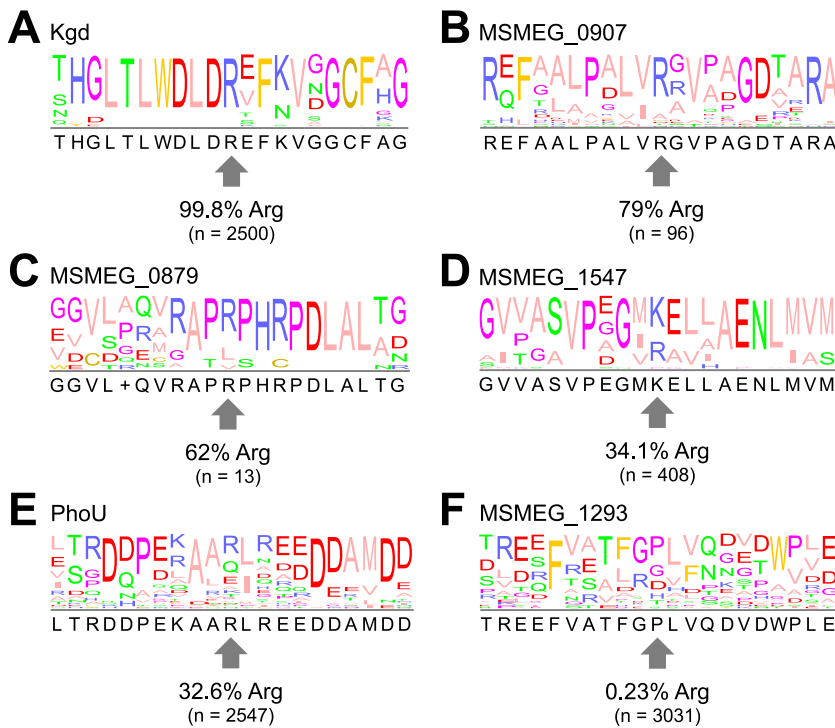


FIG 5 Sequence conservation of phosphorylated arginine residues. Sequence logos show conservation near phosphorylated arginines in Kgd R433 (A), Msmeg_0907 R34 (B), Msmeg_0879 R31 (C), Msmeg_1547 R326 (D), PhoU1 R149 (E), and Msmeg_1293 R475 (F). Percentages indicate positional conservation of the phosphorylated arginine; “n” indicates the number of aligned orthologous sequences used to generate the sequence logo.

context of targeted pathway regulation. On the other hand, we cannot rule out a role for pArg in other stress contexts that were not tested in this study, including milder heat shock conditions. Notably, the sustained heat stress conditions in this study may have impaired the activity of the yet-unidentified mycobacterial arginine kinase.

The six arginine phosphosites localized in this work were far fewer than those identified in other bacterial studies. Two independent studies in *B. subtilis* observed 121 and 217 sites in 87 and 134 proteins, respectively (36, 41), while in *S. aureus* 207 sites were identified in 126 proteins (37). On its face, the small number of pArg modifications found in *M. smegmatis* argues against a proteome-wide quality control role and appears more consistent with targeted regulation of selected proteins and processes. However, we note that phosphoproteomic studies with *Firmicutes* utilized deletion strains lacking known arginine phosphatases (*B. subtilis* YwlE and *S. aureus* PtpB), which elevate levels of pArg-bearing proteins. No arginine phosphatase has yet been identified in *M. smegmatis* or *M. tuberculosis*; thus, our approach utilized wild-type *M. smegmatis*. If a mycobacterial arginine phosphatase is eventually identified, it would be interesting to observe the effect of its knockout or knockdown on pArg levels.

For arginine phosphorylation to be useful to the cell, it must presumably be applied selectively through the regulated activity of arginine kinases. Many protein kinases recognize substrates through characteristic sequence motifs surrounding the phosphosite. In contrast, our examination of mycobacterial arginine phosphosites revealed no clearly enriched consensus sequence around pArg. Prior analysis of Arg phosphosites in *B. subtilis* similarly revealed no evident consensus sequence (36). This degeneracy suggests either that distal interactions guide *M. smegmatis* arginine kinases to phosphosites or that other characteristics of target proteins guide substrate selection. Supporting the latter possibility, we noted that all phosphorylated arginine residues (for which structural models could be obtained) were found near one end of an alpha helix, which may indicate a structural feature recognized by an arginine kinase.

What role does arginine phosphorylation play in mycobacterial cells? In *B. subtilis*, pArg tagging helps enforce protein quality control by marking misfolded proteins for destruction by ClpCP (34, 36, 41). Prior studies have demonstrated that mycobacterial ClpC1P1P2 can degrade model substrates *in vitro* bearing pArg modifications (42). The pArg-bearing proteins identified in this work may therefore be recognized as proteolytic substrates by ClpC1P1P2. Future studies will be required to test whether the pArg-bearing population of these proteins experiences increased ClpC1-dependent turnover. Alternatively, pArg may modulate some aspect of the targets' function or interactions with other proteins. We observed pArg on PhoU1, which helps regulate P_i uptake by inhibiting the activity of the SenX3-RegX3 two-component system when P_i is readily available (52). Additionally, the signal transduction protein Msmeg_0907, metabolic proteins (Kgd and Msmeg_1547), a transmembrane xanthine/uracil transport protein (Msmeg_1293), and an uncharacterized protein (Msmeg_0879) were found to be phosphorylated. In terms of essentiality, a prior study showed that the *M. tuberculosis* homolog of Kgd is essential for viability (2); another report showed that PhoU1 is jointly essential with the PhoU2 protein (Msmeg_1605) for *in vitro* *M. smegmatis* growth (52). More work is required to determine whether pArg modifications regulate these pathways and whether such modification plays an essential role in mycobacterial physiology. Along other lines, recent studies point to a role for the ClpC1 NTD in regulating unfoldase activity by modulating the formation of higher-order ClpC1 oligomers (66). Binding of pArg to the NTD may thus influence the overall activity state of the enzyme toward other targets.

This study lays the groundwork for future efforts to expound on the roles of phosphoarginine modifications in mycobacteria. The immediate impediment to further understanding this system is the absence of identifiable orthologs of known arginine kinases or phosphatases. Nevertheless, the lack of consensus motif in the identified phosphosites predicts a promiscuous kinase that targets cellular proteins, in a manner independent of a specific primary sequence. Future work will be required to decipher how arginine phosphorylation is regulated and how it contributes to the physiology of mycobacteria and other actinobacterial species.

MATERIALS AND METHODS

ClpC1 sequence analysis. Actinobacterial homologs of *M. smegmatis* ClpC1 were identified using HMMER (67) and aligned using Clustal Omega (68). Fragmentary sequences were omitted from analysis. To reduce overrepresentation of similar taxa (e.g., multiple *M. tuberculosis* strains), alignments were pruned such that no two sequences had greater than 90% sequence identity. Positional conservation scores were exported from Jalview (69) and plotted as a heatmap in Prism (GraphPad). The sequence logo of the NTD region was generated using WebLogo (70).

Protein purification and binding assays. The nucleotide sequence encoding the NTD of ClpC1 (codons 1 to 147; referred to as ClpC1^{NTD}) was amplified from *M. smegmatis* MC²155 genomic DNA (gDNA) (ATCC) and cloned into a pET22b-derived vector (EMD Millipore) in frame with a C-terminal LPETGG sortase recognition sequence (71) and 6×His tag. ClpC1^{NTD} was overexpressed in *E. coli* strain ER2566 (New England BioLabs [NEB]) by induction with 0.5 mM isopropyl β -D-1-thiogalactopyranoside at 30°C for 4 h. Cells were harvested by centrifugation, resuspended in lysis buffer (25 mM HEPES, 500 mM NaCl, 10 mM imidazole, 10% glycerol [pH 7.5]), and lysed by sonication. ClpC1^{NTD} was purified from clarified lysate by nickel-nitrilotriacetic acid (Ni-NTA; Marvelgent Biosciences) and anion exchange (Source Q; Cytiva) chromatography. ClpC1^{NTD} was fluorescently labeled via sortase transpeptidation (71) by 2 h of incubation with sortase A (~1:30 molar ratio) and a 2-fold molar excess of a Gly-Gly-Asn-Lys-(fluorescein isothiocyanate) peptide (Biomatik) in PBS buffer (10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, 137 mM NaCl, 2.7 mM KCl, 10% glycerol [pH 7.4]). Excess peptide was removed by size exclusion chromatography (Superdex 75; Cytiva). Purified ClpC1^{NTD-FITC} was concentrated and stored in CPD buffer (25 mM HEPES, 200 mM NaCl, 10 mM MgCl₂, 0.1 mM EDTA [pH 7.0]). Binding of phosphoarginine (Millipore Sigma) to 0.1 μ M purified ClpC1^{NTD-FITC} was assayed in CPD buffer supplemented with 0.05% Tween 20 and 8 mg/mL of bovine serum albumin (BSA) by microscale thermophoresis using a Monolith NT.115 (NanoTemper) (72). Thermophoretic data were fit to a quadratic single site binding equation in Prism (GraphPad).

Cell culture conditions. Liquid starter cultures of *M. smegmatis* MC²155 were prepared in 20 mL of Middlebrook broth base (HiMedia) containing 0.2% (vol/vol) glycerol (Fisher Scientific) and 0.05% (vol/vol) Tween 80 and grown for 48 h at 37°C with orbital shaking at 250 rpm. Saturated starter cultures were subcultured into 900 mL of fresh medium at a starting A_{600} of ~0.02 and grown at 37°C until reaching an A_{600} of ~1.0. A total of 500 mL of culture was collected and added to 400 mL of fresh medium. Unstressed control samples were further grown at 37°C, while samples for heat stress were grown at 50°C for 21 h. Cells were harvested by centrifugation at 9,000 \times g for 20 min at 4°C, and pellets were resuspended in 5 mL of ice-cold lysis

buffer (25 mM HEPES, 200 mM KCl, 10 mM MgCl₂, 0.1 mM EDTA [pH 7.5]) containing 10 mM ATP (Fisher), 200 μ L of phosphatase inhibitor cocktail (Millipore Sigma), 30 mM sodium pervanadate (Acros Organics), 3 mM NaF, and 3 mM sodium pyrophosphate (both from Fisher). Cells were lysed at high pressure using a microfluidizer (Microfluidics). Lysates were clarified at 15,000 \times *g* for 30 min at 4°C, and supernatant was stored at –80°C prior to further processing. Total protein content was estimated by Bradford assay (Bio-Rad).

Filter-aided sample preparation for mass spectrometry. Replicate samples of clarified cell lysates (approximately 10 mg total protein) were prepared for mass spectrometry using filter-aided sample preparation (FASP) (36, 73). Reduction of disulfide bonds was achieved by the addition of 200 mM dithiothreitol (DTT; Fisher), followed by incubation at 56°C for 50 min. Afterwards, samples were diluted in 7 mL of 0.1 M Tris (pH 7.5; Gold Biotechnology) containing 8 M urea and 25 mM 2-iodoacetamide (both from Acros Organics) in 15-mL tubes and incubated in the dark at room temperature for 45 min to carbamidomethylate cysteines. After alkylation, samples were transferred to an Amicon filter (10,000-molecular-weight cutoff [MWCO]; Millipore Sigma), washed twice with 5 mL of 0.1 M Tris (pH 7.5) containing 8 M urea by centrifugation at 4,000 \times *g*, then washed twice with 5 mL of 0.1 M Tris (pH 7.5) containing 4 M urea, and finally washed twice with 5 mL of 100 mM ammonium bicarbonate (Honeywell). After the last wash step, retentate was reduced to less than 1 mL by spin concentration.

In-solution trypsin proteolytic cleavage. MS-grade trypsin protease (Pierce) was dissolved to obtain a 20- μ g/mL stock in 100 mM ammonium bicarbonate. Trypsin digestion of processed *M. smegmatis* lysates was performed in an Amicon filter unit at a protein/trypsin ratio of approximately 1,500:1 overnight at 37°C. Upon digestion, peptides were recovered by centrifugation at 4,000 \times *g* for 10 min, followed by the addition of 0.5 M NaCl. Eluted samples were dried under vacuum in a SpeedVac centrifuge (Thermo Fisher).

TiO₂ enrichment of phosphopeptides. Phosphopeptides were enriched from tryptic peptide samples using Titansphere TiO₂ beads (5 μ m; GL Sciences) under buffer conditions that limited acid hydrolysis of phosphoarginine (34, 36, 74). Five milligrams of TiO₂ resin was resuspended in 1 mL of binding buffer (300 mg/mL lactic acid, 12.5% acetic acid, 60% acetonitrile, 0.2% heptafluorobutyric acid [pH 4] with NH₄OH). Lyophilized peptide samples were redissolved in the TiO₂ suspension, incubated for 35 min at 20°C with gentle agitation, and then transferred to graphite spin columns (Thermo Fisher). Unbound peptides were removed by a wash step with 150 μ L of binding buffer and spun at 2,000 \times *g* for 1 min, followed by three wash steps using 400 μ L of wash solution A (200 mg/mL of lactic acid, 75% acetonitrile, 2% trifluoroacetic acid, 2% heptafluorobutyric acid), wash solution B (200 mg/mL of lactic acid, 75% acetonitrile, 10% acetic acid, 0.1% heptafluorobutyric acid [pH 4] with NH₄OH), and wash solution C (80% acetonitrile, 10% acetic acid). The resin was then incubated with 100 μ L of elution solution 1 (1% NH₄OH, 30 mM ammonium phosphate) and elution solution 2 (1.25% NH₄OH in 50% acetonitrile) for 15 min each. The eluate containing phosphopeptides was collected by centrifugation after each incubation. To remove salts, samples were desalted using a HyperSep C₁₈ column (Thermo Scientific). Samples were lyophilized and stored at –80°C prior to mass spectrometry.

LC-MS/MS. Lyophilized and desalted tryptic digests were resuspended in 20 μ L of 0.5% acetic acid (pH 4.5). An Orbitrap Eclipse mass spectrometer (MS; Thermo Scientific) coupled with an Ultimate 3000 nano-liquid chromatography (nano-LC) system and a FAIMS Pro Interface (Thermo Scientific) was used for the LC-tandem mass spectrometry (MS/MS) analysis. Peptide samples were first loaded onto a trap column (PepMap C₁₈; 2 cm by 100 μ m [inside diameter]) and afterwards separated at a flow rate of 300 nL/min on an analytical column (PepMap C₁₈; 3.0 μ m; 10 cm by 75 mm [inside diameter]; Thermo Scientific). A binary buffer system (buffer A, 0.1% formic acid in water; buffer B, 0.1% formic acid in acetonitrile) with a 165-min gradient (1% to 25% buffer B over 125 min, 25% to 32% buffer B in 10 min, then 95% buffer B over 3 min, back to 1% B in 5 min, and stay equilibration at 1% buffer B for 20 min) was utilized. To achieve field asymmetric ion mobility spectrometry (FAIMS) separation, multiple compensation voltages [CVs] (–45, –60, and –80) were applied. For all experiments, the survey scans (MS1) were acquired over a mass range of 375 to 1,500 *m/z* at a resolution of 60,000 in the Orbitrap. Isolation of precursors was done with a width of 1.6 *m/z* for MS/MS acquisition. The precursors were subsequently fragmented with higher-energy collisional dissociation (HCD) using 30% collision energy with a maximum injection time of 100 ms and collected in Orbitrap at 15,000 resolution. The dynamic exclusion was set to 60 s and was shared across different FAIMS experiments. LS-MS/MS data were collected in independent biological triplicates.

Mass spectrometry data analysis. Proteomic analysis was performed in the Proteome Discoverer software suite (version 2.2; Thermo Fisher). Raw data were searched against the *M. smegmatis* (strain ATCC 700084 [MC²155]) UniProt Reference Proteome (Proteome identifier [UP000000757](#); 6,602 entries in total) using Sequest HT (University of Washington and Thermo Fisher) (75). Iodoacetamide-mediated cysteine carbamidomethylation was set as a static modification, while methionine oxidation and phosphorylation of arginine, serine, threonine, tyrosine, and histidine residues were entered as dynamic modifications. Complete trypsinization with a maximum of two missed cleavages was allowed. Precursor mass tolerance was set at 10 ppm while allowing fragment ions to have a mass deviation of 0.02 Da for the HCD data. Validation of peptide-spectrum matches (PSM) based on *q* value was done using Percolator, with target false-discovery rates (FDR) of 1% and 5% for stringent and relaxed validations, respectively. The false-discovery rate of high-confidence protein and peptide identification was 1%. Localization probability of phosphopeptide hits was analyzed using the PhosphoRS (ptmRS) node of the Proteome Discoverer software (76). Only modifications with a PhosphoRS (ptmRS) score of \geq 75% were selected.

Gene Ontology annotation analysis. Gene Ontology (GO) annotation was performed using the OmicsBox software suite (77). Homologs were identified by BLAST (Basic Local Alignment Search Tool) search against the Swiss-Prot/UniProt database (78), with a minimum expectation value 10^{–3}. Homolog annotations

were compiled from the InterPro database (79). BLAST and InterPro results were used to generate GO terms in terms of biological process. For proteins still unannotated, direct UniProt BLAST was performed, and tentative assignment of functional group was based on those of obtained homologs.

Sequence and structural analysis of arginine phosphosites. For alignment analysis, a 21-mer sequence was obtained containing 10 residues flanking each side of phosphoarginine. All 21-mers were then aligned in BioEdit (80). Structural information, if available, was obtained from Protein Data Bank (PDB) (81). Predicted structures were retrieved using a neural network-based machine learning model, AlphaFold (version 2 [59]), and molecular images were prepared in PyMOL (version 2.5.2; Schrödinger). Protein BLAST of arginine-phosphorylated proteins was performed in the NCBI suite (82). Analysis was restricted to the *Corynebacterineae* suborder. Sequences were aligned using Clustal Omega (68), and conservation and consensus values were obtained in Jalview (69).

Data availability. The mass spectrometry data from this publication have been submitted to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository (83) and assigned the identifier PXD032083.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, XLSX file, 7.3 MB.

SUPPLEMENTAL FILE 2, XLSX file, 0.02 MB.

SUPPLEMENTAL FILE 3, XLSX file, 0.01 MB.

SUPPLEMENTAL FILE 4, XLSX file, 0.01 MB.

SUPPLEMENTAL FILE 5, PDF file, 4.3 MB.

ACKNOWLEDGMENTS

We thank members of the Schmitz lab for advice, V.P., A.J.L., P.C.B., M.P., A.D., P.A.O., and O.U. for comments on the manuscript, and the Parashar lab for use of instrumentation. For technical help with mass spectrometry, we thank Yanbao Yu and the University of Delaware Mass Spectrometry Facility. Sortase A was a kind gift from S. Novo and J. M. Fox. The anti-pArg antibody was a kind gift from Paul Thompson, UMass Medical School.

K.R.S. was supported by NIH NIGMS award P20GM104316. The UD Department of Chemistry & Biochemistry Mass Spectrometry core was additionally supported by NIH NIGMS award P30GM110758-02. This content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

E.C.O. and K.R.S. designed research, E.C.O. and H.R.A. performed research, E.C.O. and K.R.S. analyzed data, and E.C.O. and K.R.S. wrote the paper.

REFERENCES

1. WHO. 2021. Global tuberculosis report 2021. World Health Organization, Geneva, Switzerland.
2. Sasseti CM, Boyd DH, Rubin EJ. 2003. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* 48:77–84. <https://doi.org/10.1046/j.1365-2958.2003.03425.x>.
3. Griffin JE, Gawronski JD, DeJesus MA, Ioerger TR, Akerley BJ, Sasseti CM. 2011. High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog* 7:e1002251. <https://doi.org/10.1371/journal.ppat.1002251>.
4. Raju RM, Jedrychowski MP, Wei JR, Pinkham JT, Park AS, O'Brien K, Rehren G, Schnappinger D, Gygi SP, Rubin EJ. 2014. Post-translational regulation via Clp protease is critical for survival of *Mycobacterium tuberculosis*. *PLoS Pathog* 10:e1003994. <https://doi.org/10.1371/journal.ppat.1003994>.
5. Raju RM, Unnikrishnan M, Rubin DH, Krishnamoorthy V, Kandror O, Akopian TN, Goldberg AL, Rubin EJ. 2012. *Mycobacterium tuberculosis* ClpP1 and ClpP2 function together in protein degradation and are required for viability in vitro and during infection. *PLoS Pathog* 8:e1002511. <https://doi.org/10.1371/journal.ppat.1002511>.
6. DeJesus MA, Gerrick ER, Xu W, Park SW, Long JE, Boutte CC, Rubin EJ, Schnappinger D, Ehart S, Fortune SM, Sasseti CM, Ioerger TR. 2017. Comprehensive essentiality analysis of the *Mycobacterium tuberculosis* genome via saturating transposon mutagenesis. *mBio* 8:e02133-16. <https://doi.org/10.1128/mBio.02133-16>.
7. Compton CL, Schmitz KR, Sauer RT, Sello JK. 2013. Antibacterial activity of and resistance to small molecule inhibitors of the ClpP peptidase. *ACS Chem Biol* 8:2669–2677. <https://doi.org/10.1021/cb400577b>.
8. Gavriš E, Sit CS, Cao S, Kandror O, Spoering A, Peoples A, Ling L, Fetterman A, Hughes D, Bissell A, Torrey H, Akopian T, Mueller A, Epstein S, Goldberg A, Clardy J, Lewis K. 2014. Lassomycin, a ribosomally synthesized cyclic peptide, kills mycobacterium tuberculosis by targeting the ATP-dependent protease ClpC1P2. *Chem Biol* 21:509–518. <https://doi.org/10.1016/j.chembiol.2014.01.014>.
9. Gao W, Kim JY, Anderson JR, Akopian T, Hong S, Jin YY, Kandror O, Kim JW, Lee IA, Lee SY, McAlpine JB, Mulugeta S, Sunoqrot S, Wang Y, Yang SH, Yoon TM, Goldberg AL, Pauli GF, Suh JW, Franzblau SG, Cho S. 2015. The cyclic peptide ecumicin targeting ClpC1 is active against *Mycobacterium tuberculosis* in vivo. *Antimicrob Agents Chemother* 59:880–889. <https://doi.org/10.1128/AAC.04054-14>.
10. Moreira W, Ngan GJ, Low JL, Poulsen A, Chia BC, Ang MJ, Yap A, Fulwood J, Lakshmanan U, Lim J, Khoo AY, Flotow H, Hill J, Raju RM, Rubin EJ, Dick T. 2015. Target mechanism-based whole-cell screening identifies bortezomib as an inhibitor of caseinolytic protease in mycobacteria. *mBio* 6:e00253-15. <https://doi.org/10.1128/mBio.00253-15>.
11. Famulla K, Sass P, Malik I, Akopian T, Kandror O, Alber M, Hinzen B, Ruebsamen-Schaeff H, Kalscheuer R, Goldberg AL, Brotz-Oesterhelt H. 2016. Acyldepsipeptide antibiotics kill mycobacteria by preventing the physiological functions of the ClpP1P2 protease. *Mol Microbiol* 101:194–209. <https://doi.org/10.1111/mmi.13362>.
12. Moreno-Cinos C, Goossens K, Salado IG, Van Der Veken P, De Winter H, Augustyns K. 2019. ClpP protease, a promising antimicrobial target. *Int J Mol Sci* 20:2232. <https://doi.org/10.3390/ijms20092232>.

13. Sauer RT, Baker TA. 2011. AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem* 80:587–612. <https://doi.org/10.1146/annurev-biochem-060408-172623>.
14. Baker TA, Sauer RT. 2012. ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochim Biophys Acta* 1823:15–28. <https://doi.org/10.1016/j.bbamer.2011.06.007>.
15. Akopian T, Kandror O, Raju RM, Unnikrishnan M, Rubin EJ, Goldberg AL. 2012. The active ClpP protease from *M. tuberculosis* is a complex composed of a heptameric ClpP1 and a ClpP2 ring. *EMBO J* 31:1529–1541. <https://doi.org/10.1038/emboj.2012.5>.
16. Schmitz KR, Sauer RT. 2014. Substrate delivery by the AAA+ ClpX and ClpC1 unfoldases activates the mycobacterial ClpP1P2 peptidase. *Mol Microbiol* 93:617–628. <https://doi.org/10.1111/mmi.12694>.
17. Schmitz KR, Carney DW, Sello JK, Sauer RT. 2014. Crystal structure of Mycobacterium tuberculosis ClpP1P2 suggests a model for peptidase activation by AAA+ partner binding and substrate delivery. *Proc Natl Acad Sci U S A* 111:E4587–E4595. <https://doi.org/10.1073/pnas.1417120111>.
18. Li M, Kandror O, Akopian T, Dharkar P, Wlodawer A, Maurizi MR, Goldberg AL. 2016. Structure and functional properties of the active form of the proteolytic complex, ClpP1P2, from Mycobacterium tuberculosis. *J Biol Chem* 291:7465–7476. <https://doi.org/10.1074/jbc.M115.700344>.
19. Vahidi S, Ripstein ZA, Juravsky JB, Rennella E, Goldberg AL, Mittermaier AK, Rubinstein JL, Kay LE. 2020. An allosteric switch regulates Mycobacterium tuberculosis ClpP1P2 protease function as established by cryo-EM and methyl-TROSY NMR. *Proc Natl Acad Sci U S A* 117:5895–5906. <https://doi.org/10.1073/pnas.1921630117>.
20. Leodolter J, Warweg J, Weber-Ban E. 2015. The Mycobacterium tuberculosis ClpP1P2 protease interacts asymmetrically with its ATPase partners ClpX and ClpC1. *PLoS One* 10:e0125345. <https://doi.org/10.1371/journal.pone.0125345>.
21. Nagpal J, Paxman JJ, Zammit JE, Alhuwaider A, Truscott KN, Heras B, Dougan DA. 2019. Molecular and structural insights into an asymmetric proteolytic complex (ClpP1P2) from Mycobacterium smegmatis. *Sci Rep* 9:18019. <https://doi.org/10.1038/s41598-019-53736-8>.
22. Brotz-Oesterhelt H, Beyer D, Kroll HP, Endermann R, Ladel C, Schroeder W, Hinzen B, Raddatz S, Paulsen H, Henninger K, Bandow JE, Sahl HG, Labischinski H. 2005. Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat Med* 11:1082–1087. <https://doi.org/10.1038/nm1306>.
23. Kirstein J, Hoffmann A, Lilie H, Schmidt R, Rubsamen-Waigmann H, Brotz-Oesterhelt H, Mogk A, Turgay K. 2009. The antibiotic ADEP reprogrammes ClpP, switching it from a regulated to an uncontrolled protease. *EMBO Mol Med* 1:37–49. <https://doi.org/10.1002/emmm.200900002>.
24. Lavey NP, Coker JA, Ruben EA, Duerfeldt AS. 2016. Sclerotiamide: the first non-peptide-based natural product activator of bacterial caseinolytic protease P. *J Nat Prod* 79:1193–1197. <https://doi.org/10.1021/acs.jnatprod.5b01091>.
25. Akopian T, Kandror O, Tsu C, Lai JH, Wu W, Liu Y, Zhao P, Park A, Wolf L, Dick LR, Rubin EJ, Bachovchin W, Goldberg AL. 2015. Cleavage specificity of Mycobacterium tuberculosis ClpP1P2 protease and identification of novel peptide substrates and boronate inhibitors with anti-bacterial activity. *J Biol Chem* 290:11008–11020. <https://doi.org/10.1074/jbc.M114.625640>.
26. Schmitt EK, Riawanto M, Sambandamurthy V, Roggo S, Miault C, Zwillingstein C, Krastel P, Noble C, Beer D, Rao SP, Au M, Niyomrattanakit P, Lim V, Zheng J, Jeffery D, Pethe K, Camacho LR. 2011. The natural product cyclomarin kills Mycobacterium tuberculosis by targeting the ClpC1 subunit of the caseinolytic protease. *Angew Chem Int Ed Engl* 50:5889–5891. <https://doi.org/10.1002/anie.201101740>.
27. Wolf NM, Lee H, Choules MP, Pauli GF, Phansalkar R, Anderson JR, Gao W, Ren J, Santarsiero BD, Lee H, Cheng J, Jin YY, Ho NA, Duc NM, Suh JW, Abad-Zapatero C, Cho S. 2019. High-resolution structure of ClpC1-rufomycin and ligand binding studies provide a framework to design and optimize anti-tuberculosis leads. *ACS Infect Dis* 5:829–840. <https://doi.org/10.1021/acsinfecdis.8b00276>.
28. Li L, MacIntyre LW, Ali T, Russo R, Koirala B, Hernandez Y, Brady SF. 2021. Bio-synthetic interrogation of soil metagenomes reveals metamarin, an uncommon cyclomarin congener with activity against Mycobacterium tuberculosis. *J Nat Prod* 84:1056–1066. <https://doi.org/10.1021/acs.jnatprod.0c01104>.
29. Barik S, Sureka K, Mukherjee P, Basu J, Kundu M. 2010. RseA, the SigE specific anti-sigma factor of Mycobacterium tuberculosis, is inactivated by phosphorylation-dependent ClpC1P2 proteolysis. *Mol Microbiol* 75:592–606. <https://doi.org/10.1111/j.1365-2958.2009.07008.x>.
30. Gopal P, Sarathy JP, Yee M, Ragnathan P, Shin J, Bhushan S, Zhu J, Akopian T, Kandror O, Lim TK, Gengenbacher M, Lin Q, Rubin EJ, Grüber G, Dick T. 2020. Pyrazinamide triggers degradation of its target aspartate decarboxylase. *Nat Commun* 11:1661. <https://doi.org/10.1038/s41467-020-15516-1>.
31. Lunge A, Gupta R, Choudhary E, Agarwal N. 2020. The unfoldase ClpC1 of Mycobacterium tuberculosis regulates the expression of a distinct subset of proteins having intrinsically disordered termini. *J Biol Chem* 295:9455–9473. <https://doi.org/10.1074/jbc.RA120.013456>.
32. Ziemski M, Leodolter J, Taylor G, Kerschenmeyer A, Weber-Ban E. 2021. Genome-wide interaction screen for Mycobacterium tuberculosis ClpCP protease reveals toxin-antitoxin systems as a major substrate class. *FEBS J* 288:99–114. <https://doi.org/10.1111/febs.15335>.
33. Gerth U, Kock H, Kusters I, Michalik S, Switzer RL, Hecker M. 2008. Clp-dependent proteolysis down-regulates central metabolic pathways in glucose-starved Bacillus subtilis. *J Bacteriol* 190:321–331. <https://doi.org/10.1128/JB.01233-07>.
34. Trentini DB, Suskiewicz MJ, Heuck A, Kurzbauer R, Deszcz L, Mechtler K, Clausen T. 2016. Arginine phosphorylation marks proteins for degradation by a Clp protease. *Nature* 539:48–53. <https://doi.org/10.1038/nature20122>.
35. Marsee JD, Ridings A, Yu T, Miller JM. 2018. Mycobacterium tuberculosis ClpC1 N-terminal domain is dispensable for adaptor protein-dependent allosteric regulation. *Int J Mol Sci* 19:3651. <https://doi.org/10.3390/ijms19113651>.
36. Schmidt A, Trentini DB, Spiess S, Fuhrmann J, Ammerer G, Mechtler K, Clausen T. 2014. Quantitative phosphoproteomics reveals the role of protein arginine phosphorylation in the bacterial stress response. *Mol Cell Proteomics* 13:537–550. <https://doi.org/10.1074/mcp.M113.032292>.
37. Junker S, Maaß S, Otto A, Michalik S, Morgenroth F, Gerth U, Hecker M, Becher D. 2018. Spectral library based analysis of arginine phosphorylations in Staphylococcus aureus. *Mol Cell Proteomics* 17:335–348. <https://doi.org/10.1074/mcp.RA117.000378>.
38. Fuhrmann J, Schmidt A, Spiess S, Lehner A, Turgay K, Mechtler K, Charpentier E, Clausen T. 2009. McsB is a protein arginine kinase that phosphorylates and inhibits the heat-shock regulator CtsR. *Science* 324:1323–1327. <https://doi.org/10.1126/science.1170088>.
39. Trentini DB, Fuhrmann J, Mechtler K, Clausen T. 2014. Chasing phosphoarginine proteins: development of a selective enrichment method using a phosphatase trap. *Mol Cell Proteomics* 13:1953–1964. <https://doi.org/10.1074/mcp.O113.035790>.
40. Suskiewicz MJ, Hajdusits B, Beveridge R, Heuck A, Vu LD, Kurzbauer R, Hauer K, Thoeny V, Rumpel K, Mechtler K, Meinhardt A, Clausen T. 2019. Structure of McsB, a protein kinase for regulated arginine phosphorylation. *Nat Chem Biol* 15:510–518. <https://doi.org/10.1038/s41589-019-0265-y>.
41. Elsholz AK, Turgay K, Michalik S, Hessling B, Gronau K, Oertel D, Mäder U, Bernhardt J, Becher D, Hecker M, Gerth U. 2012. Global impact of protein arginine phosphorylation on the physiology of Bacillus subtilis. *Proc Natl Acad Sci U S A* 109:7451–7456. <https://doi.org/10.1073/pnas.1117483109>.
42. Weinhäupl K, Brennich M, Kazmaier U, Lelievre J, Ballell L, Goldberg A, Schanda P, Fraga H. 2018. The antibiotic cyclomarin blocks arginine-phosphate-induced millisecond dynamics in the N-terminal domain of ClpC1 from Mycobacterium tuberculosis. *J Biol Chem* 293:8379–8393. <https://doi.org/10.1074/jbc.RA118.002251>.
43. Schlothauer T, Mogk A, Dougan DA, Bukau B, Turgay K. 2003. MecA, an adaptor protein necessary for ClpC chaperone activity. *Proc Natl Acad Sci U S A* 100:2306–2311. <https://doi.org/10.1073/pnas.0535717100>.
44. Vasudevan D, Rao SP, Noble CG. 2013. Structural basis of mycobacterial inhibition by cyclomarin A. *J Biol Chem* 288:30883–30891. <https://doi.org/10.1074/jbc.M113.493767>.
45. Bürstner N, Roggo S, Ostermann N, Blank J, Delmas C, Freuler F, Gerhartz B, Hinniger A, Hoepfner D, Liechty B, Mihalic M, Murphy J, Pistorius D, Rottmann M, Thomas JR, Schirle M, Schmitt EK. 2015. Gift from nature: cyclomarin A kills mycobacteria and malaria parasites by distinct modes of action. *ChemBiochem* 16:2433–2436. <https://doi.org/10.1002/cbic.201500472>.
46. Jung IP, Ha NR, Kim AR, Kim SH, Yoon MY. 2017. Mutation analysis of the interactions between Mycobacterium tuberculosis caseinolytic protease C1 (ClpC1) and ecumicin. *Int J Biol Macromol* 101:348–357. <https://doi.org/10.1016/j.ijbiomac.2017.03.126>.
47. Wolf NM, Lee H, Zagal D, Nam JW, Oh DC, Lee H, Suh JW, Pauli GF, Cho S, Abad-Zapatero C. 2020. Structure of the N-terminal domain of ClpC1 in complex with the antituberculosis natural product ecumicin reveals unique binding interactions. *Acta Crystallogr D Struct Biol* 76:458–471. <https://doi.org/10.1107/S2059798320004027>.
48. Fuhrmann J, Subramanian V, Thompson PR. 2015. Synthesis and use of a phosphonate amidine to generate an anti-phosphoarginine-specific antibody. *Angew Chem Int Ed Engl* 54:14715–14718. <https://doi.org/10.1002/anie.201506737>.

49. Potel CM, Lin MH, Heck AJR, Lemeer S. 2018. Widespread bacterial protein histidine phosphorylation revealed by mass spectrometry-based proteomics. *Nat Methods* 15:187–190. <https://doi.org/10.1038/nmeth.4580>.
50. Fu S, Fu C, Zhou Q, Lin R, Ouyang H, Wang M, Sun Y, Liu Y, Zhao Y. 2020. Widespread arginine phosphorylation in human cells—a novel protein PTM revealed by mass spectrometry. *Sci China Chem* 63:341–346. <https://doi.org/10.1007/s11426-019-9656-7>.
51. Bäsell K, Otto A, Junker S, Zühlke D, Rappen GM, Schmidt S, Hentschker C, Macek B, Ohlsen K, Hecker M, Becher D. 2014. The phosphoproteome and its physiological dynamics in *Staphylococcus aureus*. *Int J Med Microbiol* 304:121–132. <https://doi.org/10.1016/j.ijmm.2013.11.020>.
52. Brokaw AM, Eide BJ, Muradian M, Boster JMM, Tischler AD. 2017. Mycobacterium smegmatis PhoU proteins have overlapping functions in phosphate signaling and are essential. *Front Microbiol* 8:2523. <https://doi.org/10.3389/fmicb.2017.02523>.
53. Wagner T, Bellinzoni M, Wehenkel A, O'Hare HM, Alzari PM. 2011. Functional plasticity and allosteric regulation of α -ketoglutarate decarboxylase in central mycobacterial metabolism. *Chem Biol* 18:1011–1020. <https://doi.org/10.1016/j.chembiol.2011.06.004>.
54. Liu Y, Gallo AA, Bajpai RK, Chistoserodov A, Nelson A, Segura L, Xu W. 2010. The diversity and molecular modelling analysis of B₁₂ and B₁₂-independent glycerol dehydratases. *Int J Bioinform Res Appl* 6:484–507. <https://doi.org/10.1504/IJBRA.2010.037988>.
55. Nasir A, Ashok S, Shim JY, Park S, Yoo TH. 2020. Recent progress in the understanding and engineering of coenzyme B(12)-dependent glycerol dehydratase. *Front Bioeng Biotechnol* 8:500867. <https://doi.org/10.3389/fbioe.2020.500867>.
56. Isaza CE, Silaghi-Dumitrescu R, Iyer RB, Kurtz DM, Jr, Chan MK. 2006. Structural basis for O₂ sensing by the hemerythrin-like domain of a bacterial chemotaxis protein: substrate tunnel and fluxional N terminus. *Biochemistry* 45:9023–9031. <https://doi.org/10.1021/bi0607812>.
57. Li X, Li J, Hu X, Huang L, Xiao J, Chan J, Mi K. 2015. Differential roles of the hemerythrin-like proteins of *Mycobacterium smegmatis* in hydrogen peroxide and erythromycin susceptibility. *Sci Rep* 5:16130. <https://doi.org/10.1038/srep16130>.
58. Ma Z, Abendroth J, Buchko GW, Rohde KH, Davidson VL. 2020. Crystal structure of a hemerythrin-like protein from *Mycobacterium kansasii* and homology model of the orthologous Rv2633c protein of *M. tuberculosis*. *Biochem J* 477:567–581. <https://doi.org/10.1042/BCJ20190827>.
59. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Židek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596:583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
60. Macek B, Forchhammer K, Hardouin J, Weber-Ban E, Grangeasse C, Mijakovic I. 2019. Protein post-translational modifications in bacteria. *Nat Rev Microbiol* 17:651–664. <https://doi.org/10.1038/s41579-019-0243-0>.
61. Chow K, Ng D, Stokes R, Johnson P. 1994. Protein tyrosine phosphorylation in *Mycobacterium tuberculosis*. *FEMS Microbiol Lett* 124:203–207. <https://doi.org/10.1111/j.1574-6968.1994.tb07285.x>.
62. Canova MJ, Veyron-Churllet R, Zanella-Cleon I, Cohen-Gonsaud M, Cozzzone AJ, Becchi M, Kremer L, Molle V. 2008. The *Mycobacterium tuberculosis* serine/threonine kinase PknL phosphorylates Rv2175c: mass spectrometric profiling of the activation loop phosphorylation sites and their role in the recruitment of Rv2175c. *Proteomics* 8:521–533. <https://doi.org/10.1002/pmic.200700442>.
63. Prisc S, Dankwa S, Schwartz D, Chou MF, Locasale JW, Kang CM, Bemis G, Church GM, Steen H, Hussion RN. 2010. Extensive phosphorylation with overlapping specificity by *Mycobacterium tuberculosis* serine/threonine protein kinases. *Proc Natl Acad Sci U S A* 107:7521–7526. <https://doi.org/10.1073/pnas.0913482107>.
64. Nakedi KC, Nel AJ, Garnett S, Blackburn JM, Soares NC. 2015. Comparative Ser/Thr/Tyr phosphoproteomics between two mycobacterial species: the fast growing *Mycobacterium smegmatis* and the slow growing *Mycobacterium bovis* BCG. *Front Microbiol* 6:237. <https://doi.org/10.3389/fmicb.2015.00237>.
65. Zhou B, Semanjski M, Orlovetskie N, Bhattacharya S, Alon S, Argaman L, Jarrous N, Zhang Y, Macek B, Sinai L, Ben-Yehuda S. 2019. Arginine dephosphorylation propels spore germination in bacteria. *Proc Natl Acad Sci U S A* 116:14228–14237. <https://doi.org/10.1073/pnas.1817742116>.
66. Taylor G, Frommherz Y, Katikaridis P, Layer D, Sinning I, Carroni M, Weber-Ban E, Mogk A. 2022. Antibacterial peptide CyclomarinA [sic] creates toxicity by deregulating the *Mycobacterium tuberculosis* ClpC1/ClpP1P2 protease. *J Biol Chem* 298:102202. <https://doi.org/10.1016/j.jbc.2022.102202>.
67. Eddy SR, Mitchison G, Durbin R. 1995. Maximum discrimination hidden Markov models of sequence consensus. *J Comput Biol* 2:9–23. <https://doi.org/10.1089/cmb.1995.2.9>.
68. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539. <https://doi.org/10.1038/msb.2011.75>.
69. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189–1191. <https://doi.org/10.1093/bioinformatics/btp033>.
70. Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res* 14:1188–1190. <https://doi.org/10.1101/gr.849004>.
71. Glasgow JE, Salit ML, Cochran JR. 2016. In vivo site-specific protein tagging with diverse amines using an engineered sortase variant. *J Am Chem Soc* 138:7496–7499. <https://doi.org/10.1021/jacs.6b03836>.
72. Jerabek-Willemsen M, André T, Wanner R, Roth HM, Duhr S, Baaske P, Breitsprecher D. 2014. MicroScale thermophoresis: interaction analysis and beyond. *J Mol Struct* 1077:101–113. <https://doi.org/10.1016/j.molstruc.2014.03.009>.
73. Wiśniewski JR, Zougman A, Nagaraj N, Mann M. 2009. Universal sample preparation method for proteome analysis. *Nat Methods* 6:359–362. <https://doi.org/10.1038/nmeth.1322>.
74. Thingholm TE, Jørgensen TJ, Jensen ON, Larsen MR. 2006. Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nat Protoc* 1:1929–1935. <https://doi.org/10.1038/nprot.2006.185>.
75. Eng JK, McCormack AL, Yates JR. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* 5:976–989. [https://doi.org/10.1016/1044-0305\(94\)80016-2](https://doi.org/10.1016/1044-0305(94)80016-2).
76. Taus T, Köcher T, Pichler P, Paschke C, Schmidt A, Henrich C, Mechtler K. 2011. Universal and confident phosphorylation site localization using phosphoRS. *J Proteome Res* 10:5354–5362. <https://doi.org/10.1021/pr200611n>.
77. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674–3676. <https://doi.org/10.1093/bioinformatics/bti610>.
78. Bairoch A, Bucher P, Hofmann K. 1997. The PROSITE database, its status in 1997. *Nucleic Acids Res* 25:217–221. <https://doi.org/10.1093/nar/25.1.217>.
79. Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L, Finn RD, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Laugraud A, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, Mistry J, Mitchell A, Mulder N, Natale D, Orengo C, Quinn AF, Selengut JD, Sigrist CJ, Thimma M, Thomas PD, Valentin F, Wilson D, Wu CH, Yeats C. 2009. InterPro: the integrative protein signature database. *Nucleic Acids Res* 37:D211–D215. <https://doi.org/10.1093/nar/gkn785>.
80. Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98.
81. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. *Nucleic Acids Res* 28:235–242. <https://doi.org/10.1093/nar/28.1.235>.
82. States DJ, Gish W. 1994. Combined use of sequence similarity and codon bias for coding region identification. *J Comput Biol* 1:39–50. <https://doi.org/10.1089/cmb.1994.1.39>.
83. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, Pérez E, Uszkoreit J, Pfeuffer J, Sachsenberg T, Yilmaz S, Tiwary S, Cox J, Audain E, Walzer M, Jarnuczak AF, Ternent T, Brazma A, Vizcaino JA. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* 47:D442–D450. <https://doi.org/10.1093/nar/gky1106>.