Supporting Information

Prisic et al. 10.1073/pnas.0913482107

SI Materials and Methods

Cell Cultures, Protein Extraction, and Digestion. *Mycobacterium tuberculosis* H37Rv was grown under several conditions in Middlebrook 7H9 liquid culture medium (Difco) supplemented with 0.5% albumin, 0.2% glucose, 0.085% NaCl, and 0.05% Tween 80, and harvested at different growth stages. A total of 10 culture extracts were obtained as follows: (*i*) log phase (200 mL), (*ii*) log phase (400 mL), (*iii*) late stationary phase (400 mL), (*iv*) induced hypoxia (Wayne model) (1) (200 mL), (*v*) early stationary phase (400 mL), (*vi*) log phase treated with 300 μM diethylenetriamine (DETA)-NO overnight (200 mL), (*vii*) log phase treated with 10 μM cumene hydroperoxide overnight (200 mL), (*viii*) log phase grown on 0.1% sodium acetate in place of glucose (400 mL), (*ix*) log phase treated with 300 μM DETA-NO 90 min (200 mL), and (*x*) log phase treated with 10 μM cumene hydroperoxide 90 min (200 mL).

Cell pellets were washed with 40 mM Tris (pH 7.4), and proteins were isolated using TRIzol (Invitrogen) according to the manufacturer's manual.

The ethanol protein precipitate was solubilized in Laemmli reducing sample buffer (BioRad), and proteins were separated by SDS/PAGE. The gels were stained (GelCode Blue; Pierce Biotechnology), cut horizontally into eight strips, and each strip was further cut into small cubes. Before digestion, the pieces were destained with 5 mL 50% acetonitrile (ACN) in 0.1 M ammonium bicarbonate (ABC) and incubated at least 1 h in 5 mL 0.1 M ABC. The solution was replaced with fresh 1.5 mL 0.1 M ABC, and proteins were reduced for 30 min at 60 °C after adding 100 μL 45 mM DTT and then alkylated with 100 μL 0.1 M iodoacetamide at room temperature for 30 min in the dark. The reaction mixture was replaced with 5 mL 50% ACN in 0.1 M ABC, and slices were shaken for 1 h. After removing the wash solution, 0.5 mL of 100% ACN was added to shrink the gel pieces. ACN was then removed, and gel pieces were dried under vacuum. The gel was reswollen with 20 μL trypsin solution [50 μg/mL of Trypsin Gold MS grade (Promega) in 25 mM ABC], and 25 mM of ABC was added to completely cover the pieces (≈700 μL total). Proteins were digested overnight at 37 °C and extracted with 25 mM ABC and 2 × 2 mL 60% ACN/0.1% trifluoroacetic acid (TFA). All digests were pooled, split in three equal volumes, and dried under vacuum with medium heat.

Phosphopeptide Enrichment. To enrich for phosphopeptides, we used three different commercial kits: ProteoExtract Phosphopeptide TiO2 Enrichment (Calbiochem), ProteoExtract Phosphopeptide Enrichment SCIMAC (Calbiochem), and TopTip TiO_2 (Glygen). For the first two kits manufacturer's instructions were followed, and for TopTip TiO_2 we used 40% ACN 5% TFA solution for dissolving the sample, column conditioning, and washing. Peptides bound to the column were washed once more with 0.2% TFA 40% ACN solution and eluted with 40 μL 0.1 M ABC and 40 μL 0.4 M ammonium hydroxide 40% ACN directly into 50 μL 40% ACN 5% TFA buffer.

Phosphopeptide Identification. All samples were dried under vacuum and run on a Thermo Finnigan LTQ mass spectrometer. Briefly, samples were resuspended in 2.5% formic acid 4% ACN, loaded on to the column (Magic C18 resin), and washed with aqueous solvent (buffer A: 0.2% formic acid and 99.8% H_2O) for 5 min, followed by 30 min gradient elution (buffer B: 0.2% formic acid and 99.8% ACN), and a 5-min buffer B wash. The column was equilibrated with aqueous solvent for 5 min after each run. Liquid chromatography/mass spectrometry/electrospray ionization runs were programmed to generate MS2 spectra for the top 6 MS1 peaks and conditional MS3 spectra from any MS2 containing a neutral loss of 98 Da.

Data were searched with the Paragon algorithm using ProteinPilot 2.0 (Applied Biosystems) with emphasis on phosphorylation. The following parameters were used:

Cys alkylation: iodoacetamide
Digestion: trypsin
Instrument: LTQ
Special factors: phosphorylation emphasis; gel-based ID
ID focus: biologic modifications
Search effort: rapid

The data were also analyzed using Mascot (Matrix Science). Parameters were as follows:

Type of search: MS/MS ion search
Enzyme: trypsin
Mass values: average
Protein mass: unrestricted
Peptide mass tolerance: ±1.5 Da
Fragment mass tolerance: ±0.8 Da
Max missed cleavages: 1

Posttranslational and chemical modifications, including phosphorylation on S, T, Y (+80 Da in MS2 or −18 in MS3), oxidation on M (+16), deamidation of N, Q (+1), and pyro-Glu at the N terminus from Q (−17) or from E (−18), were conditionally searched. Cysteine was unconditionally modified using iodoacetamide (+57).

Custom scripts in the Perl programming language were written to perform a number of downstream processing steps as described below.

Determination of Spectral Identifications. The outputs from all Mascot runs were harvested to collect all peptide identifications with a score ≥40, and all ProteinPilot (PP) searches were harvested to collect all peptide identifications ≥99. Decoy analysis was performed in Mascot and indicated that a score of 35.4 had a calculated 1% false discover rate (2, 3). In cases in which multiple identifications existed for a given spectrum, only peptide sequences (without modifications) agreeing with the top-scoring peptide sequence (also without modifications) were retained for further analysis. In no case did the top-scoring PP sequence (without modifications) disagree with the top-scoring Mascot hit (without modifications) for a given spectrum. Results from two sequences did occasionally disagree with respect to modifications and position of modifications including phosphates, a situation that was addressed as described below.

Phosphate Localization Analysis. In a number of spectra, multiple possible positions were identified for any given phosphorylation. The Ascore server (4) was used to determine the location of phosphates with greater confidence. To use the Ascore server, all runs were first converted to mzXML format from their native RAW file format, and each spectrum with a significant hit (i.e., ≥40 Mascot score or ≥99% confidence in PP) was extracted into a .dta format spectral file. Next, to create the Sequest output files expected by the Ascore program, each spectrum was rapidly "researched" using a database containing solely the previously identified Mascot or PP peptide and any of the aforementioned

modifications. The purpose of this step was not to independently search the spectra with a different search engine, but rather to create a file format expected by the Ascore program. Because the Ascore source code is not available, it is not possible to know the importance of modified residues other than phosphorylation. Therefore, we took a conservative approach whereby the top Sequest hit was required to match at least one of the Mascot or PP hits for a given spectrum in all respects except phosphorylation location before sending it to the Ascore server. The spectral file (.dta) and cognate output (.out) files were sent to the Ascore server for localization scoring.

In a few instances, Mascot or PP searches found hits assuming that a spectrum was for a +2 charged peptide or a +3 charged peptide—one of which was likely incorrect. In these few instances, the proper charge state was confirmed by these Sequest runs, and only results based on the correct charge state were retained.

If the Ascore program was able to produce a score of ≥15 for a particular position, then the phosphorylation was deemed to be localized. Scores of <15 for a given site denote a *nonlocalized* phosphate. Because of their uncertainty, such sites that cannot be localized with high confidence should not be used to annotate proteomic phosphorylation events.

Although recent evidence points to the existence of a tyrosine kinase in *M. tuberculosis* (5), at the time of our analyses Tyr phosphorylation was not known to occur and we therefore did not include the option of phosphotyrosine in our Mascot searches, although Ascore did identify one possible Tyr phosphorylation. PP, however, which was run in a mode to search for this modification, did not find any phospho-Tyr containing peptides.

Ascore Independent Localization. Only MS2 spectra were sent to Ascore for processing because it could not localize MS3 spectra arising from a neutral phosphate loss during CID fragmentation. Despite this, phosphorylation localizations in either MS2 or MS3 spectra were generated for any hits in which such assignments could be made unambiguously because the number of phosphates exactly matched the number of Ser or Thr residues in the peptide. Such assignments overrode any Ascore assessments.

Protein Identification and Classification. All peptides were searched against all proteins in the *M. tuberculosis* H37Rv strain protein database (obtained from tbdb.org in September 2009). All Rv numbers were assigned a functional classification obtained from Tuberculist (http://genolist.pasteur.fr/TubercuList/). All peptides, with the exception of QAQVDAGARPGTTTEESAELKR, which is present in each of the 16 copies of IS6110 present in the H37Rv strain, were found exactly once in the proteome, and thus were used to unambiguously determine the protein identification and its Rv accession number. We identified 381 distinct tryptic fragments with phosphorylations. These mapped to 301 different proteins.

Identified peptides are presented in Table S1, which contains Mascot and PP search results, a localized version of the peptide (if it exists), protein identification, functional classification, and experimental conditions. Table S2 provides additional information for each identified peptide.

Peptides Not Matching Proteins in H37Rv Strain. Initially, peptides were searched in multiple *M. tuberculosis* strains, and several peptides were identified that did not map to the H37Rv strain. All these peptides were either partially overlapping an ORF at its amino terminus or were entirely upstream of an annotated H37Rv gene. After analysis of these peptides and comparison with other strains, we decided to report them because at least some of them are likely to be expressed, and their absence in the H37Rv genome may be the result of differences in annotation methods. Four such peptides from four proteins were identified and are shown in Table S1b and Table S2b, using the Rv number of the adjacent protein. These sequences include five additional phosphorylation sites, all localized with high confidence by Ascore.

These peptides and phosphorylation sites were not used for bioinformatic analysis or included in final proteome list, but five peptides derived from these sites were tested in vitro for phosphorylation by recombinant protein kinases.

Recombinant Protein Kinase Cloning and Expression. The kinase domains of all 11 STPKs were cloned from *M. tuberculosis* H37Rv genomic DNA using either restriction enzyme digestion/ligation or Gateway (Invitrogen) technology and expressed as GST-fusion proteins in *E. coli* [pDEST15 (Invitrogen) or pGEX-4T-3 (GE Healthcare)], with the exception of PknK, which was expressed as a His-tagged protein (pDEST17; Invitrogen). Three kinases, PknG, I, and J, were also expressed as full-length proteins in *Mycobacterium smegmatis* as His-tagged fusion proteins using pAL5000-based vector pMV261 (Medimmune) in which the hsp60 promoter was replaced with acetamidase promoter.

For expression in *E. coli*, C41(DE3) cells (Lucigen) were used. Overnight cultures grown at 37 °C in LB with 100 μg/mL ampicillin were diluted in 500 mL of fresh medium and grown at 37 °C until mid log (OD600 0.5–0.6). The temperature was lowered to 16 °C for 1 h before 1 mM IPTG (isopropyl β-D-1-thiogalactopyranoside) was added to induce expression overnight. Cell pellets were lysed, and proteins were purified using either B-PER GST (Pierce Biotechnology) or His Bind purification kits (Novagen) with the exception of PknA-GST, for which we used MagneGST purification kit (Promega).

For expression in *M. smegmatis*, cultures were grown in Middlebrook 7H9-ADC with 25 μg/mL kanamycin at 37 °C with shaking at 150–180 rpm until saturation. Fresh medium was inoculated with an aliquot-saturated culture grown at 37 °C with shaking until OD600 of 0.6 and induced overnight with 0.2% acetamide. Pelleted cells were lysed by sonication, and His-tagged proteins were affinity purified according to the supplier's procedures (Novagen).

Protein preparations were dialyzed against 40% glycerol in Tris-buffered saline (TBS) and analyzed for purity using SDS/PAGE. Protein concentrations were determined by A (280 nm), and preparations were stored at −20 °C until use.

Mutagenesis of PknB and GarA. The WT GST-tagged PknB kinase domain in pGEX-4T-3 was mutated using QuikChange (Stratagene) according to the manufacturer's instructions, expressed in *E.coli*, and purified as described above for the WT. GarA was cloned into pENTR vector (Invitrogen), mutated using QuikChange (Stratagene), and transferred to the pDEST17 expression vector (Invitrogen) for expression as a His-tagged protein. Tagged mutated and WT GarA proteins were expressed in C41(DE3) cells (Lucigen) after induction at mid log phase with 1 mM IPTG for 2 h at 37 °C. Cell pellets were lysed in BugBuster (Novagen) and purified using HisPur (Pierce Biotechnology). All mutants were expressed and purified at levels comparable to their corresponding WT proteins. The BioRad protein assay was used to determine protein concentrations.

Two PknB mutants (V176D and V176R) were analyzed for phosphorylation of an ideal peptide (peptide 1 in Fig. 3 in main text) using FlashPlates, in the same manner as the peptide library screening described in more detail below. The major difference was that a 300-μL reaction mixture containing 25 ng/μL of each kinase was prepared in kinase buffer with ATP [10 μM ATP, 0.5 μCi [γ-^{33}P] ATP, 50 mM Mops (pH 7.4), 10 mM $MgCl_2$, and 10mM $MnCl_2$] and 50-μL aliquots were transferred at 0, 15, 45, 60, 120, 240 min to 100 μL 0.1 M EDTA/4 M NaCl solution. Phosphorylation was not linear beyond 120 min, so that the linear regression shown in Fig. 5 (main text) was done omitting the 240-min data point. The assay was performed at room temperature.

WT and mutated PknB (1.25 μg each) were tested for autophosphorylation in the presence of 1 μCi [γ-^{33}P] ATP (EasyTide; PerkinElmer) in the same kinase buffer at 30 °C for 40 min. The reactions were run on SDS/PAGE, and radioactive signal was detected using Storm PhosphorImager (GE Healthcare).

A similar kinase assay was performed to test phosphorylation of GarA and its mutants by PknB. Identical amounts of WT or mutated recombinant GarA (0.4 μg or 23 pmol) were mixed with a recombinant PknB (0.6 μg or 10 pmol) in the presence of 1 μCi [γ-^{33}P] ATP and kinase buffer. The reaction was stopped by adding reducing sample buffer and run on SDS/PAGE. The radioactive signal was read using a Storm PhosphorImager.

In Vitro Peptide Phosphorylation. A set of peptides corresponding to in vivo phosphorylations was synthesized on the basis of initial identifications of phosphorylated tryptic peptides by Mascot (score >44) or PP (>99% confidence). From the initial set, several peptides were excluded on the basis of manual review identifying low-quality spectra. We randomly checked PP identifications and all Mascot identifications in this group. The final list of 336 synthetic peptides that were subject to in vitro phosphorylation and analyzed in this work includes 148 phosphorylations that were localized with high confidence by Ascore and 188 that could not be definitively localized. Table S3 contains a list of these synthetic 13-mer peptides, the proteins from which they are derived, and results of in vitro kinase assays with nine recombinant proteins (see below).

The peptides were synthesized as crude product containing a minimum 50 nmol peptide in 96-well plate format (BioTides; JPT Peptide Technologies). During synthesis, an additional Gly was added to the C terminus and biotin at the N terminus. Stocks were made in DMSO to ≈0.5-mM concentration and stored at −80 °C. Working solutions were made by diluting stock 10-fold in 10 mM Mops (pH 7.4) and were stored up to several weeks at −80 °C.

Recombinant protein kinases (≈10–20 ng/μL for most kinases, except for PknL, which required higher concentration, i.e., ≈200 ng/μL) were mixed with 10 μM ATP in 50 mM Mops (pH 7.4), 10 mM $MgCl_2$, and 10 mM $MnCl_2$. After 15 min preincubation, 0.5 μCi [γ-^{33}P] ATP (EasyTide; PerkinElmer) was added, and 45 μL of the mixture was aliquoted to wells in OptiPlates (Perkin-Elmer). Peptides were added (5 μL) to make 5-μM final concentration, and the 50-μL reaction was incubated at 30 °C for 4–7 h, depending on estimated kinase activity. Phosphorylation was stopped by adding 100 μL 0.1 M EDTA 4 M NaCl and transferring the entire reaction volume to streptavidin-coated FlashPlates (Perkin-Elmer) containing 50 μL water. Plates were incubated overnight at room temperature, washed five times with TBS 0.1% Tween-20, and counted for 2 min in a TopCount instrument (Perkin-Elmer). Table S3 shows the ratio of phosphorylation of each peptide relative to background, by each kinase except for PknI and PknJ, which did not show significant activity in this assay.

For motif confirmation, 32 peptides were synthesized and purified to at least 70% purity JPT (Peptide Technologies) and assayed as described above, except that incubation times were significantly shorter—1.5 h for PknB, PknD, and PknF, and 2 h for PknA, PknE, and PknH. PknG, PknK, and PknL were also tested for phosphorylation of a 32-peptide set, with 2.5 h incubation. However, for these three enzymes no significant phosphorylation was observed.

Bioinformatic Analysis. ***Motif-x.*** **Motif determination of in vivo substrates.** All localized peptides from the localized column of Table S1 were used as input to an internal working version of *motif-x* (6) to determine motifs for these phosphorylated substrates. Parameters for the *motif-x* analyses were as follows: foreground central residue = S or T (dependent on analysis), width = 13, occurrences = 5, significance = 0.000041687, subtract foreground from background option selected, and background = *M. tuberculosis* proteome. Additionally, the analyses were carried using a breadth-first search implementation of *motif-x*. To visualize sequence differences between the preferred phosphorylation sites, *motif-x* probability logos (pLOGos) were generated. These probability log-based logos depict the statistical significance of residues under a user-defined background, with larger, more statistically significant, residues stacked closer to the midline. Red horizontal lines represent the 0.01 significance level (after Bonferroni correction). Residues above the midline are overrepresented, whereas those below the midline are underrepresented. Fixed residues within motifs are drawn at 100% height and can be distinguished from other residues by having no underrepresented residues beneath them. The pLOGos resulting from this analysis are shown in Fig. S1.

Kinase-specific phosphorylation site motifs. Motifs for each of the *M. tuberculosis* kinases that were used in the in vitro kinase asays were extracted using the *motif-x* algorithm as described above (6). Phosphorylation (scintillation counts) of each peptide was divided by the median value from all peptides phosphorylated in the same experiment. For each of the kinases, only those peptides that were at least 3-fold above the median in duplicate experiments were used for the motif analysis. Peptides containing predicted N- or C-terminal phosphoacceptors (i.e., Ser and Thr that were off centered) were excluded from this analysis. The five peptides that did not map to the annotated H37Rv proteome were also excluded from this analysis, reducing the total number of peptides to 318. Motifs of peptides that were phosphorylated in vitro vs. the ones that failed to phosphorylate in vitro, but were phosphorylated in vivo, are shown in Figs. S3 and S4.

The pLOGos for the peptides phosphorylated by each kinase were created using *M. tuberculosis* proteomic background. pLOGos for each kinase are shown in Fig. S5 and in Fig. 2 in the main text. Note, for these pLOGos, *motif-x* was run in a mode in which no significant residues were allowed to become fixed. Therefore, these pLOGos represent the overall motif for each kinase's data set.

Threading algorithm. From the in vitro kinase assays, phosphorylated peptides were chosen whose scintillation counts were at least 5-fold higher than the median value for each experiment. The selected peptides were then used to construct a frequency matrix with elements f_{ij} consisting of rows $(i \in [1,13])$ corresponding to each position in the peptide sequence and columns corresponding to the frequency of occurrence of each amino acid residue $(j \in [1,20])$.

To correct for background, the frequency of amino acids (matrix with elements d_{ij}) was computed from the entire set of peptides. The resulting matrix with elements $\tilde{f}_{ij}$ was obtained by subtracting the background matrix from the frequency matrix $\tilde{f}_{ij} = f_{ij} - d_{ij}$. $\tilde{f}_{ij}$ was used as the starting point for the motif search.

Each peptide p_i (a vector whose entries contain one of 20 amino acids) was then compared against the $\tilde{f}$ matrix, and a score S was obtained by summing the entries of the peptide sequence that match the elements of the $\tilde{f}$ matrix; that is, $S = \sum_i \tilde{f}_{ip_i}$. To reduce the effects of potential false positives obtained from the assay, only those peptides with a score sufficiently higher than a threshold value $S^T (S > S^T)$ were used to compute the motif. Therefore, the set of peptides with $S > S^T$ were then used as input into the weblogo application (http://weblogo.berkeley.edu) for computing motifs using sequence entropies (7, 8). Weblogos for each kinase are shown in Fig. S6.

Modeling of PknB structure in complex with an optimal peptide substrate. The x-ray crystal structures of the *M. tuberculosis* PknB kinase domain [Protein Data Bank (PDB) ID: 1o6y] and a phosphorylase kinase–peptide substrate complex (PDB ID: 2phk) were aligned using the align function in PyMOL. The activation loop (residues 168–179) and substrate peptide (all residues) from

phosphorylase kinase were added to PknB, and the remaining phosphorylase kinase residues were removed. Activation loop residues were mutated to their PknB counterparts (sequence from gi:15607156, DPGEKLREVCGT → nsvtqtaavigt), and rotamers were manually selected. Rather than attempting to model the extra residues found in PknB, the end of the loop was left disconnected. The N-epsilon atom of Lys140 was added in an extended conformation. Peptide substrate residues were mutated to those favored for PknB substrates (RQMSFRL → aeltgei). Additional N-terminal residues for the substrate were added in a mostly extended conformation, and the C and N termini were converted to N-methylamide and N-acetyl, respectively, to give the initial model for AELTGEIPII complexed to PknB. The Maestro molecular modeling package (version 8.5; Schrodinger) was used to first minimize protein and substrate side chain conformations and next to minimize the complete peptide substrate with protein/substrate H-bond constraints similar to those observed in the 2phK x-ray structure and with PknB atoms frozen. The terminal atoms where the activation loop residues were disconnected (I163 and D168) were frozen during all of the minimizations. Possible contacts between PknB active site residues and the substrate peptide were calculated using a 4-Å radius around individual peptide residues in PyMol session and are shown in Fig. 4*B* in the main text.

Protein domain identification and functional classification of phosphoproteins vs. total protein. Kinase domains, transmembrane segments, and other protein features were identified using SMART (Simple Modular Architecture Research Tool) (9). The functional classification use in the *M. tuberculosis H37Rv* genome annotation (10) was used to assign phosphoproteins to functional categories. Distribution of phosphoproteins vs. total proteins was compared using the χ^2 test with Bonferroni correction for multiple comparisons.

1. Wayne LG, Hayes LG (1996) An in vitro model for sequential study of shiftdown of *Mycobacterium tuberculosis* through two stages of nonreplicating persistence. *Infect Immun* 64:2062–2069.
2. Elias JE, Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* 4:207–214.
3. Moore RE, Young MK, Lee TD (2002) Qscore: An algorithm for evaluating SEQUEST database search results. *J Am Soc Mass Spectrom* 13:378–386.
4. Beausoleil SA, Villén J, Gerber SA, Rush J, Gygi SP (2006) A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat Biotechnol* 24:1285–1292.
5. Bach H, Wong D, Av-Gay Y (2009) *Mycobacterium tuberculosis* PtkA is a novel protein tyrosine kinase whose substrate is PtpA. *Biochem J* 420:155–160.
6. Schwartz D, Gygi SP (2005) An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat Biotechnol* 23: 1391–1398.
7. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: A sequence logo generator. *Genome Res* 14:1188–1190.
8. Schneider TD, Stephens RM (1990) Sequence logos: A new way to display consensus sequences. *Nucleic Acids Res* 18:6097–6100.
9. Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc Natl Acad Sci USA* 95:5857–5864.
10. Cole ST, et al. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393:537–544.

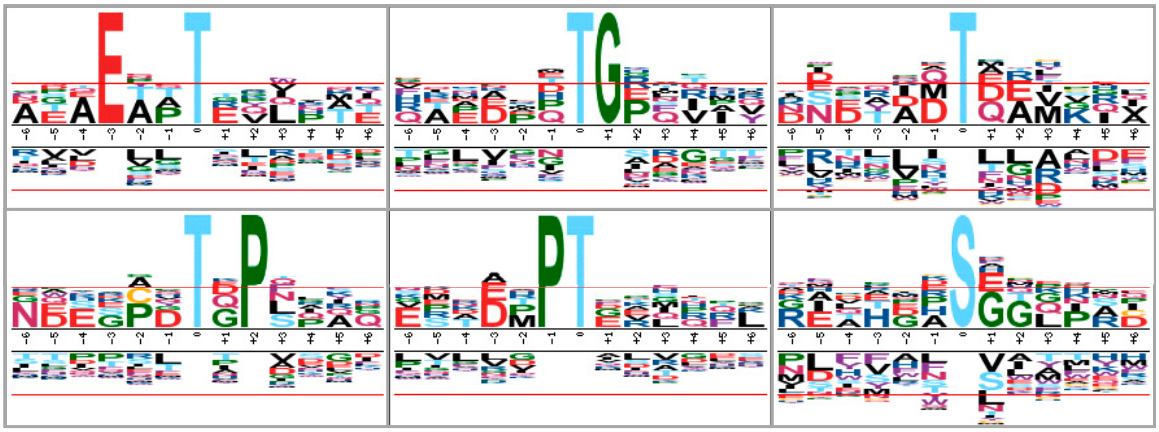

Fig. S1. *Motif-x* analysis of localized in vivo phosphorylation sites from all experimental conditions against the *M. tuberculosis* proteomic background.

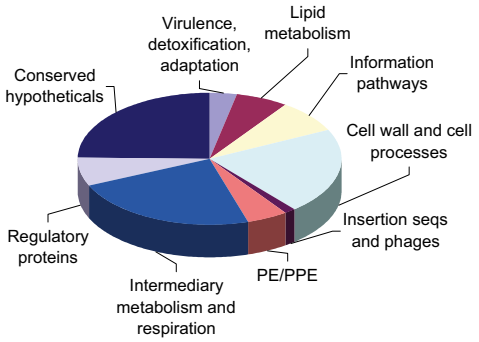

Fig. S2. Classification of *M. tuberculosis* phosphoproteins. Functional categories are those used in the *M. tuberculosis* H37Rv genome sequence annotation [Camus JC, Pryor MJ, Médigue C, Cole ST (2002) Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology* 148:2967–2973]. PE/PPE are families of *M. tuberculosis* proteins containing Pro-Glu or Pro-Pro-Glu motifs.

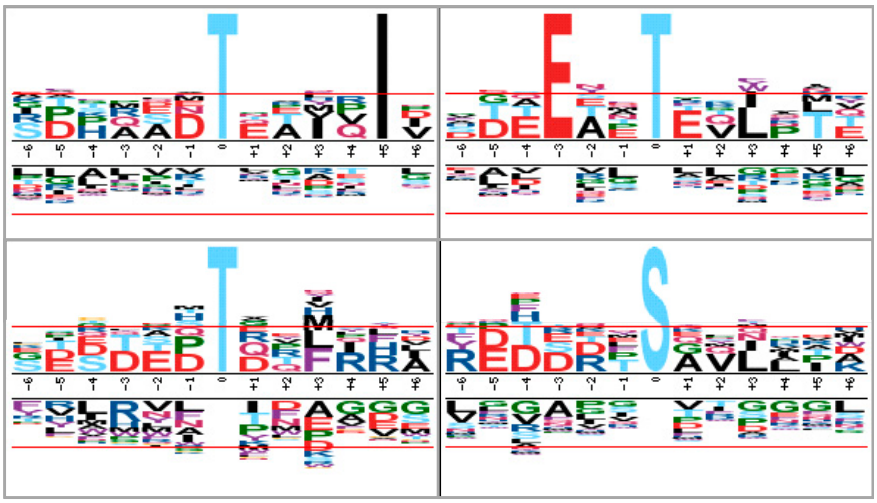

Fig. S3. pLOGos derived from peptides that are phosphorylated by recombinant protein kinases in vitro. *Motif-x* analysis was done against *M. tuberculosis* proteomic background.

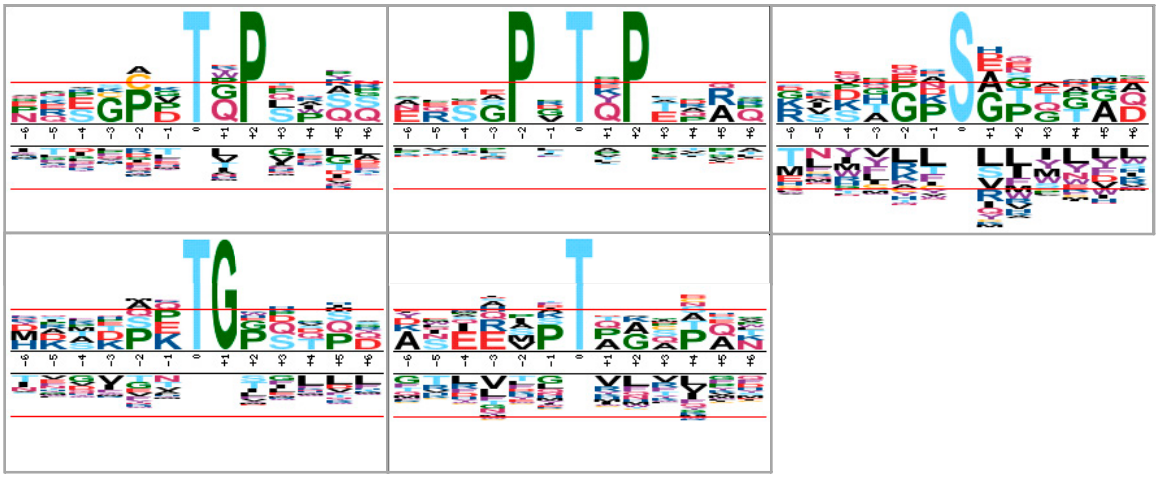

Fig. S4. pLOGos derived from peptides that are not phosphorylated by recombinant protein kinases in vitro. *Motif-x* analysis was done against *M. tuberculosis* proteomic background.

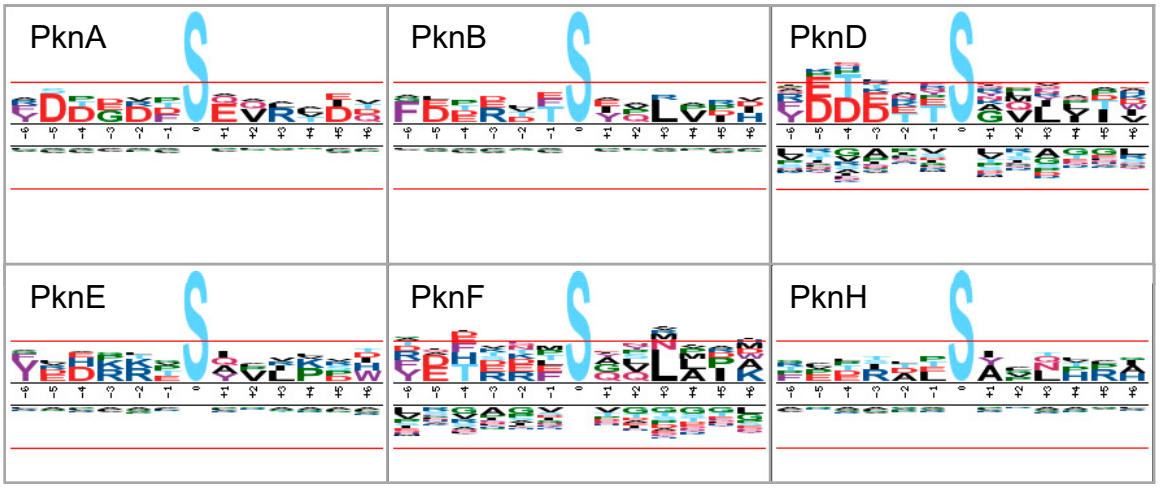

Fig. S5. pLOGos derived from peptides that are phosphorylated by six protein kinases in vitro with Ser as phosphoacceptor. Analysis was done against *M. tuberculosis* proteomic background. Note that only PknD shows significantly overrepresented acidic resides at -5 position when Ser is phosphorylated.

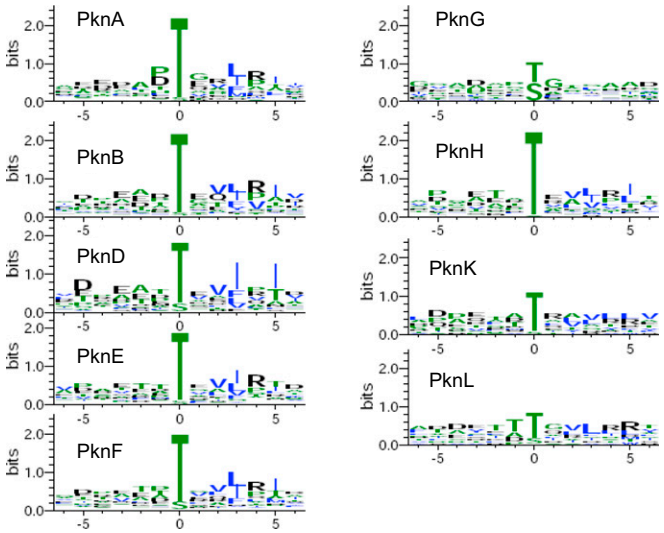

Fig. S6. Threading algorithm analysis of synthetic peptides phosphorylated in vitro.

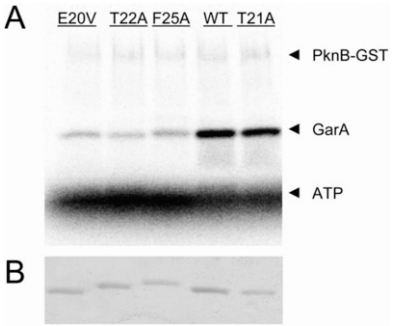

Fig. S7. Phosphorylation of GarA and GarA mutants with PknB. (*A*) Autoradiography of GarA and various point mutations after incubation with recombinant PknB in presence of -P33-ATP and running on SDS/PAGE. (*B*) GelCode Blue (Pierce Biotechnology) stained SDS/PAGE gel showing the amount of total protein loaded.

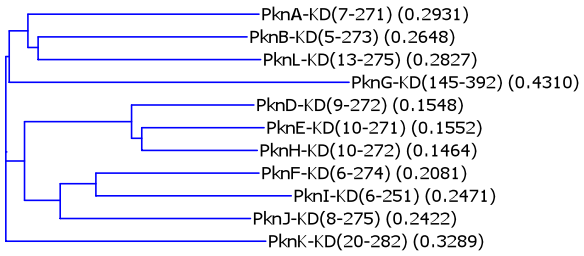

Fig. S8. Dendrogram of kinase domains of *M. tuberculosis* STPKs.

Other Supporting Information Files

Table S1 (XLS)
Table S2 (XLS)
Table S3 (XLS)