New substrates and interactors of the mycobacterial Serine/Threonine protein kinase PknG identified by a tailored interactomic approach

Magdalena Gil, Analía Lima, Bernardina Rivera, Jéssica Rossello, Estefanía Urdániz, Alessandro Cascioferro, Federico Carrión, Annemarie Wehenkel, Marco Bellinzoni, Carlos Batthyany, et al.

New substrates and interactors of the mycobacterial Serine/Threonine protein kinase PknG identified by a tailored interactomic approach

Magdalena Gil[a § #], Analía Lima[a #], Bernardina Rivera[a], Jessica Rossello[a], Estefanía Urdániz[b], Alessandro Cascioferro[c], Federico Carrión[d], Annemarie Wehenkel[e], Marco Bellinzoni[e], Carlos Batthyány[a], Otto Pritsch[d], Ana Denicola[f], María N. Alvarez[g], Paulo C. Carvalho[h], María-Natalia Lisa[e,i], Roland Brosch[c], Mariana Piuri[b], Pedro M. Alzari[e], Rosario Durán[a *]

[a] Unidad de Bioquímica y Proteómica Analíticas, Institut Pasteur de Montevideo & Instituto de Investigaciones Biológicas Clemente Estable, Uruguay.

[b] Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Argentina.

[c] Integrated Mycobacterial Pathogenomics Unit, Institut Pasteur, Paris, France.

[d] Unidad de Biofísica de Proteínas, Institut Pasteur de Montevideo, Uruguay.

[e] Unité de Microbiologie Structurale & CNRS URA 2185, Institut Pasteur, Paris, France.

[f] Laboratorio de Fisicoquímica Biológica, Facultad de Ciencias, Universidad de la República, Uruguay.

[g] Departamento de Bioquímica, Facultad de Medicina, CEINBIO, Universidad de la República, Uruguay.

[h] Laboratory for Proteomics and Protein Engineering, Carlos Chagas Institute, Fiocruz-Paraná, Brazil.

[§] Magdalena Gil's current address is Unit of Dynamics of Host-Pathogen Interactions, Institut Pasteur, Paris, France.

[i] María-Natalia Lisa's current address is Instituto de Biología Molecular y Celular de Rosario (IBR, CONICET-UNR), Ocampo y Esmeralda, S2002LRK, Rosario, Argentina.

[#] These authors contributed equally to this work

* **Corresponding author**: Rosario Durán, Unidad de Bioquímica y Proteómica Analíticas, Institut Pasteur de Montevideo, Mataojo 2020, Montevideo 11400, Uruguay, Tel: +598 2 5220910, FAX: +598 2 5224185, e-mail: duran@pasteur.edu.uy

Keywords: PknG; Serine/Threonine protein kinase; glutamine synthetase; FhaA; Affinity purification-mass spectrometry; *Mycobacterium tuberculosis*.

Abbreviations: AP-MS, affinity purification-mass spectrometry; DIA, Differential In gel Analysis; FHA, forkhead associated domain; GS, glutamine synthetase; STD, internal standard; STPK, Serine/Threonine protein kinase; TPR, tetratricopeptide repeats; WT, wild type.

Abstract

PknG from *Mycobacterium tuberculosis* is a multidomain Serine/Threonine protein kinase that regulates bacterial metabolism as well as the pathogen's ability to survive inside the host by still uncertain mechanisms. To uncover PknG interactome we developed an affinity purification-mass spectrometry strategy to stepwise recover PknG substrates and interactors; and to identify those involving PknG autophosphorylated docking sites. We report a confident list of 7 new putative substrates and 66 direct or indirect partners indicating that PknG regulates many physiological processes, such as nitrogen and energy metabolism, cell wall synthesis and protein translation. GarA and the 50S ribosomal protein L13, two previously reported substrates of PknG, were recovered in our interactome. Comparative proteome analyses of wild type and *pknG* null mutant *M. tuberculosis* strains provided evidence that two kinase interactors, the FHA-domain containing protein GarA and the enzyme glutamine synthetase, are indeed endogenous substrates of PknG, stressing the role of this kinase in the regulation of nitrogen metabolism. Interestingly, a second FHA protein was identified as a PknG substrate. Our results show that PknG phosphorylates specific residues in both glutamine synthetase and FhaA *in vitro*, and suggest that these proteins are phosphorylated by PknG in living mycobacteria.

1. Introduction

Mycobacterium tuberculosis, the etiological agent of tuberculosis, is a major health problem and the main cause of death due to a single infectious agent. According to the World Health Organization, this pathogen has caused 10.4 million new cases and 1.7 million deaths worldwide during 2016 [1]. One crucial feature of *M. tuberculosis* pathogenesis is its ability to respond to environmental signals switching from dormant to replicating bacilli in different disease stages [2,3]. In particular, signal transduction pathways involving protein phosphorylation play key roles in the adaptive response of *M. tuberculosis* [4,5].

Mass spectrometry-based phosphoproteomic approaches allowed the identification of more than 500 Ser- and Thr-phosphorylated residues in *M. tuberculosis*, and showed that phosphorylation patterns in mycobacteria changed dramatically in response to environmental stimuli [5–7]. However, the identification of the specific enzyme responsible for every phosphorylation event and the characterization of the function of each phosphoprotein still lags behind [8,9].

Genomic analysis of *M. tuberculosis* revealed the presence of 11 Serine/Threonine protein kinases (STPKs) [10]. Nine of them are receptor-like proteins with an intracellular kinase domain and an extracellular sensor domain, whereas the two other members of the family (PknG and PknK) are soluble proteins. These mycobacterial STPKs have been related to the regulation of several processes, including transcription, cell division and host-pathogen interactions, and eight of them are expressed during infection [11–14]. Among them, PknG became of special interest as it was found to play dual roles in mycobacterial metabolism and pathogenicity through mechanisms still not completely understood. On one hand, disruption of *pknG* gene reduced *M. tuberculosis* viability *in vitro* and in infection models, and caused retarded mortality in highly susceptible infected mice [12,15]. Furthermore, a *M. bovis* BCG *pknG* null mutant strain was found to be unable to block phagosome maturation in infected macrophages, and it has been proposed that PknG secretion and interference with the host cell signalling could be the underlying mechanism of this effect [16]. Very recently, other reports also support a role for PknG in facilitating bacterial growth in conditions mimicking host environment, like hypoxia or acidic

78 environments [17,18]. In addition, the deletion of *pknG* caused a multidrug sensitive phenotype in
79 contraposition to the intrinsic antibiotic resistance of pathogenic mycobacteria [19].
80 Other functions have been reported for PknG in bacterial metabolism. We have shown that PknG
81 participates in the control of glutamate metabolism *via* the phosphorylation of the endogenous regulator
82 GarA [20]. In addition, it was shown that PknG can regulate the activity of the Nudix hydrolase RenU
83 through the phosphorylation of the ribosomal protein L13 [21]. Besides these better characterized
84 substrates, very recent phosphoproteomics and protein microarray analysis have expanded the list of
85 putative PknG substrates and interactors [22–24], and few candidates were further selected to test its
86 interaction with, or phosphorylation by PknG. These studies suggest that RmlA and MurC activities are
87 also regulated by PknG mediated phosphorylation [22,24].
88 PknG presents a unique modular domain organization. Flanking the conserved catalytic kinase
89 domain, PknG has a N-terminal rubredoxin-like domain and a C-terminal domain composed of
90 tetratricopeptide repeats (TPR). The Rbx domain of PknG has been the only protein motif found to
91 regulate the intrinsic kinase activity [25,26], while the TPR domain can lead to PknG dimerization
92 without evident effects on the kinase activity [26,27]. Besides, in its N-terminal end PknG has a possibly
93 unstructured extension with up to four autophosphorylation sites that act as essential anchoring points for
94 the recruitment of the forkhead-associated (FHA) domain-containing regulator GarA [20]. FHA domains
95 specifically recognize phosphorylated Thr residues and play important roles in phosphorylation dependent
96 signal transduction [28]. Thus, PknG phosphorylates GarA within a conserved N-terminal motif,
97 triggering the self-recognition of the phosphorylated residue by the FHA domain in the C-terminus of the
98 molecule [20,29,30]. This interaction serves as a switch to activate/inhibit GarA control of downstream
99 metabolic enzymes that use alpha-ketoglutarate as substrate [20,31].
100 The central role of PknG in mycobacterial physiology and virulence is well documented, but the
101 molecular mechanisms underlying these effects, as well as the protein partners involved, are still poorly
102 understood. Even when the list of putative substrates and interactors is rapidly expanding, there is very
103 little overlap between these high-throughput studies and only few PknG substrates have been validated as

physiologically relevant. This evidences that the identification of *bona-fide* kinase substrates and the processes they regulate still represents a challenge from a methodological point of view and requires the use of multiple experimental approaches.

With the aim of contributing to a better understanding of the biological processes regulated by PknG, we developed a tailored interactomic approach. We combined the use of different constructions of PknG with specific sequential elution steps to identify PknG mediated protein complexes and, in particular, to discriminate those interactions relying on PknG's autophosphorylated docking sites. First, phosphorylation conditions were used to elute PknG substrates that, similar to the model substrate GarA, are released after phosphorylation. Second, a phosphatase treatment was employed to disrupt interactions mediated by phosphoresidues, including PknG autophosphorylated sites in the case of the whole length construction. A third and final elution step was performed to recover the remaining interacting partners. Using this experimental approach, we recovered 66 direct or indirect PknG partners, including the two previously reported substrates GarA and the 50S ribosomal protein L13. Further, two other proteins identified in the interactome of PknG, the enzyme glutamine synthetase (GS) and the protein FhaA, were validated as PknG substrates *in vitro* and evidence that both proteins are endogenous substrates of this kinase is provided. Altogether, our results suggest that PknG regulates a wide range of cellular processes including protein translation, nitrogen assimilation and cell wall biosynthesis.

2. Materials and methods

2.1. Preparation of mycobacterial lysates

M. tuberculosis Δ*pknG* was kindly provided by Dr. J. Av-Gay [12]. Wild type *M. tuberculosis* H37Rv (WT) and a *pkn*G null mutant strain (Δ*pkn*G) were grown in Middlebrook 7H9 supplemented with 0.05% Tween® 80 containing ADC supplement (BD Biosciences) until early-logarithmic phase. Cells were washed and then resuspended in minimum medium supplemented with 10 mM asparagine and cultured for 4-5 additional days. *Mycobacterium smegmatis* MC^2155 was grown in Sauton's medium

supplemented with 0.05% Tween® 80 until logarithmic phase. Mycobacterial cells were harvested and resuspended in PBS (*M. tuberculosis*) or 25 mM HEPES, 150 mM NaCl, 1% glycerol, 1 mM EDTA, pH 7.4 (protein interaction buffer for *M. smegmatis*) plus Complete EDTA-free Protease Inhibitor Cocktail (Roche). An equal amount of acid-washed glass beads (≤ 106 μm, Sigma) was added to the cell pellet and lysis was achieved by vortexing at top speed for 10 min. Cell debris and beads were removed by centrifugation and protein quantification in the supernatant was performed by densitometry analysis in gel. Protein extracts of each strain were prepared from three independent biological replicates.

2.2. Affinity purification and sequential elution of PknG interactors

Full-length PknG and $PknG_{\Delta 73}$ were produced as described before [20] and immobilized on NHS-Activated Sepharose 4 Fast Flow (GE Healthcare) following the supplier's instructions. In mock experiments, a control resin prepared by blocking active groups with ethanolamine was used.

PknG kinase activity was confirmed by using autophosphorylation and GarA phosphorylation assays as previously described [20,25]. Briefly, for autophosphorylation assay, immobilized kinase was incubated with 1 mM $MnCl_2$ in 50 mM HEPES, pH 7.4, for 30 min at 37 °C with or without (control) 500 μM ATP. Proteins immobilized in the resins were further digested overnight with trypsin (sequence grade, Promega) and the peptides recovered were analyzed by MALDI-TOF MS (4800 MALDI-TOF/TOF Analyzer, Abi Sciex) using linear mode and *m/z* range from 2000 to 6000. Phosphorylation of the recombinant substrate was performed by incubation of immobilized PknG with GarA (final concentration 70 μM) under the same conditions used for autophosphorylation assay, with or without ATP. GarA phosphorylation status was verified by whole molecular mass measurements using a MALDI-TOF mass spectrometer (4800 MALDI-TOF/TOF Analyzer, Abi Sciex) operated in linear mode.

Immobilized PknG, $PknG_{\Delta 73}$ or control resin pre-equilibrated in the interaction buffer were incubated with 800 μg of *M. smegmatis* protein extract (three biological replicates) supplemented with 2 mM EDTA, overnight at 4 °C, with gentle agitation. The matrix was packed into Pierce™ Micro-Spin

Columns (Thermo Scientific) and the retained proteins were eluted using the following protocol: (a) for the elution of interactors under phosphorylation conditions (E1), the PknG, $PknG_{\Delta 73}$ and control matrices were incubated with 1 mM $MnCl_2$ and 500 µM ATP in 50 mM HEPES, pH 7.0, for 30 min at 37 °C and the eluted proteins were recovered; (b) for the elution of proteins under dephosphorylation conditions (E2), the resins were further incubated with 0.18 U/µL of calf intestine alkaline phosphatase (Roche) for 40 min at 37 °C and the eluted proteins were recuperated; (c) for the elution of proteins under unspecific conditions (E3), the matrices containing the remaining interactors were incubated for 10 min at 25 °C with 70% ACN, 0.1% formic acid and the eluted proteins were collected. The whole experiment was run in triplicates, using independent biological samples, and each eluate was analyzed by nano-LC MS/MS.

2.3. Production of recombinant FhaA

Plasmid pLAM12-fhaA, for *Strep*-tag® II-FhaA overexpression in mycobacteria, was constructed by PCR amplification of Rv0020c (*fha*A) from genomic DNA of *M. tuberculosis* H37Rv using the oligonucleotides AAATTCATATG**TGGAGCCACCCGCAGTTCGAAAAAAGCGCT**GGTAGCCAGAAAAGGCTGGTTC (FhaA-Fw) and ATATTGAATTCTCAGTGCATGCGGACGATGATC-3' (FhaA Rv), and cloning the insert between the sites NdeI and EcoRI (underlined) in the extrachromosomally-replicating parent vector pLAM12 [32], under the control of the *M. smegmatis* acetamidase promoter. FhaA-Fw contains a 5' extension coding for the STAG II sequence (bold). Plasmid pLAM12-fhaA was used to transform electrocompetent *M. smegmatis* MC^2155 cells. Transformants were grown in Middlebrook 7H9 broth supplemented with ADC, 0.05% Tween® 80 and kanamycine, and the expression of *Strep*-tag® II-FhaA was induced by the addition of 0.2% acetamide during exponential growth. *Strep*-tag® II-FhaA (herein after named FhaA) was purified using Strep-Tactin® Sepharose® (IBA) and eluted in a competitive manner with D-desthiobiotin.

2.4. Production of recombinant GS

pCRT7::Rv2220 plasmid was kindly provided by Dr. S. Mowbray and protein production was performed as previously reported [33]. Briefly, *Escherichia coli* BL21 pLysS were transformed with plasmid pCRT7::Rv2220 and cells were grown in autoinduction medium for 4 h at 37 °C and later overnight at 20°C. GS was first purified by metal-affinity chromatography on a His-Trap FF crude column equilibrated in 50 mM HEPES pH 8.0, 500 mM NaCl, 5% glycerol and 5 mM imidazol, using a linear imidazole gradient from 5 to 500 mM. Size exclusion chromatography was performed on fractions containing GS, using a Sephacryl S400 16/60 column (GE Healthcare) equilibrated in 20 mM HEPES pH 7.7, 150 mM NaCl, 3% glycerol and 1 mM $MgCl_2$. Oligomeric state of GS was determined using either Dynamic Light Scattering or native gels (NativePAGE™ 3-12% Bis-Tris Protein Gels, Thermo). Native Mark Unstained Molecular Weight Marker (Thermo) was used as standard.

2.5. Mapping FhaA or GS phosphorylation sites

To assay FhaA as a substrate of PknG *in vitro*, a cell lysate of *M. smegmatis* MC^2155 overexpressing FhaA was initially loaded onto a Strep-Tactin® Sepharose® column (IBA). Immobilized FhaA was later dephosphorylated with 0.07 U/µL calf intestine alkaline phosphatase (Roche) for 40 min at 37 °C. Dephosphorylated FhaA was then eluted and further incubated with 1 mM $MnCl_2$, 250 µM ATP and 10 nM PknG, for 30 min at 37 °C. As a control, we performed the same experiment in the absence of PknG. Additionally, to identify the sites of FhaA phosphorylated *in vivo*, FhaA was purified from *M. smegmatis* cells transformed with plasmid pLAM12-fhaA and induced with acetamide.

To evaluate GS as a substrate of PknG *in vitro*, recombinant GS was incubated with 1 mM $MnCl_2$, 250 µM ATP and PknG (molar ratio PknG:GS 1:150), for 30 min at 37 °C.

Protein samples were digested overnight with trypsin and analysed by nano-LC MS/MS using an ion trap instrument (LTQ Velos, Thermo) for the identification of phosphorylation sites.

2.6. Surface Plasmon Resonance analysis

For surface plasmon resonance analysis, FhaA was diluted in 10 mM sodium acetate pH 4.5 at a concentration of 5 µg/mL and immobilized on a CM5 sensorchip by standard amine coupling. A BIACORE 3000 (Biacore AB, Uppsala, Sweden) was used, achieving a final density of 25 RU. Phosphorylated PknG and $PknG_{\Delta 73}$ were diluted in 10 mM HEPES pH 7.4, 0.15 M NaCl, 3 mM EDTA, 0.005% v/v Surfactant P20 to a final concentration of 30 nM and injected during 3 min at a flow rate of 80 µL/min over the immobilized and a reference surfaces. Regeneration was achieved by extensively washing with running buffer. All injections were done at 25 °C and double referenced by subtracting the reference cell signal and a buffer injection.

2.7. Differential Gel Electrophoresis (DIGE)

Comparative proteome analyses between wild type *M. tuberculosis* and Δ*pknG* strains were performed on three biological replicates, using the Ettan DIGE System (GE Healthcare) and following the manufacturer's instructions. Samples were purified using the 2D Clean-up kit (GE Healthcare) and protein pellets were solubilized in 30 mM Tris pH 8.5, 7 M urea, 2 M thiourea, 4% CHAPS.

Equal amounts (25 µg) of samples WT and Δ*pknG* were mixed to set up an internal standard (STD) for multiplex matching of DIGE images, spot normalization and calculation of spots abundance changes. Then, the STD and the protein samples WT and Δ*pknG* (50 µg each) were differentially labelled with the N-hydroxysuccinimidyl ester derivatives of the cyanine dyes Cy2, Cy3, and Cy5 following the manufacturer's instructions for minimal labelling (GE Healthcare, Munich, Germany). The differentially labelled WT and Δ*pknG* samples together with the STD were mixed and rehydration solution (7 M urea, 2 M thiourea, 4% CHAPS, 0.5% IPG Buffer 4-7 (GE Healthcare)) was added. This mixture was later used to rehydrate IPG strips (13 cm pH 4-7) overnight. Isoelectric focusing was performed in an IPGphor Unit (Pharmacia Biotech) applying the following voltage profile: constant phase of 500 V for 1 h, linear increase from to 5000V in 1 h, followed by another linear increase to 8000 V in 2 h 30 min, and a final constant phase of 8000 V in 35 min, reaching a total of 17.5 kVh. Prior to running the second dimension,

IPG-strips were reduced for 15 min in equilibration buffer (6 M urea, 75 mM Tris–HCl pH 8.8, 29.3% glycerol, 2% SDS, 0.002% bromophenol blue) supplemented with DTT (10 mg/mL) and subsequently alkylated for 15 min with iodoacetamide (25 mg/mL). The separation in the second-dimension was performed on 12.5% SDS-PAGE, at 20 °C, in a SE 600 Ruby Standard Dual Cooled Vertical Unit (GE Healthcare). Gels were scanned with a Typhoon FLA 9500 laser scanner (GE Healthcare), at a resolution of 100 µm, using laser wavelengths and filters recommended for each dye.

Images were analyzed using the DeCyder™ 2D software (v7.2) (GE Healthcare). The module for Differential In-gel Analysis (DIA) was used for spots co-detection, spot quantification by normalization and calculation of the ratio between different samples in the same gel. Biological Variation Analysis software module was used to perform inter-gel matching and statistical analyses. An unpaired Student's t-test was employed to determine if the standardized spot volume abundances from the triplicate samples showed significant changes between the two conditions. Spots displaying a pI shift profile consistent with the presence of post-translational modification such as phosphorylation (spots in trains with increased abundance in WT strain), significant volume abundances differences ($p \leq 0.05$) and a fold change of at least 25% were filtered and selected for further analyses.

To confirm the phosphorylation status of the differential spots, we compared wild type *M. tuberculosis* protein extracts pre-labelled with CyDyes with and without treatment with alkaline phosphatase (Cy3 and Cy2 respectively), using DIA analysis (DeCyder™ 2D software). The same strategy was use to compare Δ*pknG* protein extracts with and without previous incubation with recombinant PknG under phosphorylation conditions (Cy5 and Cy2 respectively).

2.8. Sample preparation for MS analysis

Protein mixtures obtained from affinity purification (AP) experiments were reduced and alkylated with iodoacetamide prior to treatment with trypsin. Peptide quantification was performed at 280 nm using a NanoDrop 1000 spectrophotometer (Thermo Scientific) prior to its analysis by nano-LC MS/MS (LTQ Velos, Thermo).

For identification of differential spots from DIGE experiments, the master gel was silver stained [34] and selected spots were manually picked from this gel and processed for protein identification using MALDI-TOF MS. Briefly, spots were *in-gel* digested with trypsin and the resultant peptides were extracted using aqueous 60% ACN/0.1% trifluoroacetic acid, and concentrated by vacuum drying. Samples were desalted using C18 micro-columns (C_{18} OMIX pipette tips, Agilent) and eluted directly onto the sample plate for MALDI-MS with α-cyano-4-hydroxycinnamic acid matrix solution in aqueous 60% ACN/0.1% trifluoroacetic acid. Additionally, some spots were analyzed by nano-LC MS/MS, in this case elution was performed in 70% ACN/0.1%f formic acid.

2.9. Protein identification by MALDI-TOF/TOF MS

Tryptic peptides obtained from the spots of the silver stained master gel were analysed using a 4800 MALDI-TOF/TOF Analyzer (Abi Sciex, Framingham, MA) equipped with 355mn Nd:YAG laser. Spectra were acquired in reflector mode using the following parameters: detector voltage: 1.75 kV; laser ranging from 50 and 70% of maximum intensity; 160 shots per subspectra and 25 subspectra acquired. MS/MS analyses of selected peptides were performed for all the analyzed spots in DIGE experiments. For MS/MS spectra acquisition source 1 and source 2 voltages were 8 kV and 15 kV, respectively.

Proteins were identified by database searching of measured *m/z* values using the MASCOT search engine (Matrix Science http://www.matrixscience.com) in the sequence Query mode, using the database from NCBI (20160821) and restricting the taxonomy to the *Mycobacterium tuberculosis* complex. The following search parameters were set: monoisotopic mass tolerance, 0.03 Da; fragment mass tolerance, 0.5 Da; methionine oxidation and phosphorylation of Ser, Thr and Tyr as variable modification, cysteine carbamidomethylation as fixed modification and one missed tryptic cleavage allowed. Significant protein scores ($p < 0.05$), at least three peptides per protein and at least one fragmented peptide with ion significant score ($p < 0.05$) per protein were used as criteria for positive identification.

2.10. LC MS/MS data acquisition

Samples from AP experiments, selected DIGE spots and phosphorylated proteins were analyzed by nano-LC MS/MS. Each sample was injected into a nano-HPLC system (EASY-nLC 1000, Thermo Scientific) fitted with a reverse-phase column (EASY-Spray column, 50 cm × 75 µm ID, PepMap RSLC C18, 2 µm, Thermo Scientific). For AP experiments, peptides were separated on a linear gradient of solvent B (ACN 0.1% formic acid (v/v)) from 5% to 55% in 75 min. In the case of DIGE and DIA spots, as well as analysis of phosphorylation status of recombinant proteins, the gradient was modified as follow: linear gradient of solvent B from 5% to 55% in 60 min. In both cases, the flow rate was set at 250 nL/min at 45 °C.

Peptide analysis was carried out in a LTQ Velos nano-ESI linear ion trap instrument (Thermo Scientific) set in a data-dependent acquisition mode using a dynamic exclusion list. NSI-source parameters were set as follows: spray voltage (kV): 2.3 and 260 °C capillary temperature. Mass analysis was performed with Xcalibur 2.1 in two steps: (1) acquisition of full MS scan in the positive ion mode with *m/z* between 400 and 1200 Da, (2) CID fragmentation of the ten most intense ions with the following parameters; normalized collision energy: 35, activation Q: 0.25; activation time: 15 ms.

2.11. Proteomic data analysis

A target-decoy database including sequences from *M. smegmatis* strain ATCC 700084/MC^{2}155, *M. tuberculosis* strain ATCC 25618/H37Rv downloaded from Uniprot consortium in November 2014, and 127 most common mass spectrometry contaminants was generated using PatternLab for Proteomics (version 3.2.0.3) [35,36], giving rise a database with 10.712 entries. For analyses of AP samples and protein identification of selected spots from DIGE gel, the Comet search engine was set as follows: tryptic peptides; oxidation of methionine and phosphorylation of Ser, Thr or Tyr as variable modifications and carbamidomethylation as fixed modification; and 700 ppm of tolerance from the measured precursor *m/z*. XCorr and Z-Score were used as the primary and secondary search engine scores, respectively. Peptide spectrum matches were filtered using the Search Engine Processor (SEPro) using the following parameters; acceptable FDR: 1% at the protein level; a minimum of two peptides per protein. All reported

proteins have at least one unique peptide. Comparisons between proteins present in each elution of PknG, $PknG_{\Delta 73}$ and mock experiments was performed using PatternLab's Approximately Area Proportional Venn Diagram module, and verification of statistical validity of the proteins was performed according to the Bayesian model integrated into PatternLab for proteomics [36,37]. Briefly, the model considered quantitative data and number of appearances in different biological replicates to assign p-values and ultimately shortlist proteins that are likely to be real interactors.

For phosphosite identification of FhaA and GS recombinant proteins, Proteome Discoverer software package (v.1.3.0.339, Thermo) with Sequest as search engine was used. Database was compiled from *M. tuberculosis* strain ATCC 25618/H37Rv (April 2015) and *M. smegmatis* strain ATCC 700084/$MC^2$155 (November 2014) Uniprot protein sequences. The precursor mass tolerance and fragment mass tolerance were set as 1 Da and 1.5 Da respectively, oxidation of methionine and phosphorylation of Ser, Thr or Tyr were set as variable modifications. Peptide validator node was used to perform a decoy database search, and the target FDR (strict) was set to 0.01. Only high-confident peptides with strict target false discovery rate were considered. To localize phosphosites in validated sequences phospho*RS* node was used. In addition, we considered positive phosphosite identification when more than one spectrum for the phosphopeptide was obtained. Manual inspection of the MS/MS spectra was performed to corroborate peptide and phosphorylation site assignments.

2.12. Bioinformatic analyses

The list of *M. smegmatis* identified proteins was converted to *M. tuberculosis* orthologs and classified using TubercuList (*M. tuberculosis* H37Rv database, March 2013 release 27 [38]). If a protein ortholog was absent in TubercuList, a blastp sequence and similarity search was performed using *M. smegmatis* sequence as query and *M. tuberculosis* complex (taxid: 77643) as database and a statistical significance threshold of 10 (default value) [39]. Protein-protein interactions analyses were performed with STRING database (v 10.0) (http://string-db.org/), using the entire set of putative PknG's interactors (MSMEG identifiers) as an input gene list, and the following parameters: "known interactions" (from curated

databases and experimentally determined) and a minimum interaction score of 0.7 (high confidence) [40]. A statistical overrepresentation test of GO biological processes (release 20160715) was performed using the Panther Server (http://pantherdb.org) version 11.1. The release date of GO ontology dataset was 2017-01-26. *Mycobacterium tuberculosis* database as reference list and the Tuberculist gene identifiers were used [41].

2.13 Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD005950 [42]. Reviewer account details to access to proteomics data: Username: reviewer87925@ebi.ac.uk; Password: whLcyv5a.

3. Results

3.1. An affinity chromatography-sequential elution strategy to find new PknG protein interactors

To identify PknG protein partners we developed an affinity chromatography-sequential elution strategy to stepwise recover PknG interactors with different binding modes and to specifically discriminate those recruited through PknG autophosphorylated docking sites. PknG was initially immobilized on NHS-Activated Sepharose and incubated with ATP/Mn^{2+} to allow autophosphorylation, which was later verified by MALDI-TOF MS analysis of tryptic digestion mixtures. As shown in **Supplementary Figure 1A**, after the autophosphorylation reaction we observed an intense signal corresponding to the diphosphorylated ion of sequence 10-60 (*m/z* 5556.3), in agreement with previous results [20]. We also verified that the immobilized kinase behaved as an active enzyme towards the substrate GarA **(Supplementary Figure 1B**). In parallel, we immobilized an active deletion mutant of PknG lacking the phosphorylatable sequence 1-73 ($PknG_{\Delta 73}$) that participates in the recruitment of the FHA-containing substrate GarA [20]. Thus, immobilized PknG, $PknG_{\Delta 73}$ and control resins were used to fish out interacting proteins from total *M. smegmatis* protein extracts. First, elution was performed under phosphorylation conditions (ATP/Mn^{2+}) to recover substrates released after its phosphorylation by the

immobilized kinase (E1). A subsequent elution step was performed under dephosphorylation conditions, to specifically disrupt interactions relying on phosphoresidues (E2), including kinase autophosphorylated sites in the case of the whole length PknG. Finally, the remaining interacting partners were recovered in a third elution step (E3). For each elution step, positive interactors were identified by comparison with mock experiments, using quantitative data and number of appearances in different biological replicates to assign p-values (filtering options, $p<0.05$) using Patternlab for Proteomics software.

3.2. Identification of new PknG substrates and protein interactors

Eight proteins were identified in at least two of three biological replicates of E1 when using full-length PknG as bait, whereas six of them were also recovered with $PknG_{\Delta 73}$ (**Tables 1 and S1**). Interestingly, we could not identify proteins in any of the three replicates of mock resin. Therefore, the kinase N-terminal segment was critical for the recruitment of two putative phosphorylatable interactors. As expected, one of these proteins is the regulator GarA, validating our experimental approach. The other protein in this group was the enzyme glutamine synthetase 1 (GS, MSMEG_4290). While GarA contains an FHA domain that specifically recognizes pThr residues in PknG, no phospho-recognition domain has been predicted for GS, and its possible interaction through the N-terminal extension of PknG deserves further investigation. No protein was identified exclusively in the interactome of $PknG_{\Delta 73}$ under the same conditions. However, six proteins were systematically recovered using both full-length PknG and $PknG_{\Delta 73}$, namely chaperone protein DnaK, alcohol dehydrogenase, 60 kDa chaperonin 1, 4Fe-4S ferredoxin, iron-sulfur binding protein, inorganic pyrophosphatase and acetyl-/propionyl-coenzyme A carboxylase alpha chain. Interestingly, six out of the eight proteins recovered in E1 have been previously reported as phosphorylated proteins (**Table S2**).

Under dephosphorylation conditions (E2), the FHA protein FhaA (MSMEG_0035) was the only protein systematically recovered when using full-length PknG as bait, but not $PknG_{\Delta 73}$ or the mock resin (**Tables 1 and S3**). The presence of a pThr recognition domain in FhaA possibly explains its specific

interaction with the N-terminal segment of the kinase. Additionally, 45 proteins were statistically significantly identified as interactors in E2 when compared to mock experiments using both baits, including PknG itself (**Table S3**). Among these proteins we recovered the previously reported substrate of PknG, the 50S ribosomal protein L13 [21], highlighting the usefulness of our interactomic approach to identify *bona-fide* kinase substrates. GS was also detected in all replicates of E2 elution with very high numbers of spectra assigned to its sequence, strongly supporting that this protein is indeed a PknG interactor. The fact that GS was recovered using both PknG constructs suggests that, besides the N-terminal segment of PknG, the kinase core and/or the TPR domain also play a role in its direct or indirect binding. Notably, 23 out of 46 proteins recovered in E2 are ribosomal proteins, including 7 known interactors of the ribosomal protein L13 (RplR, RplF, RplX, RplV, RplC, RplB and RplA), reflecting the recovery of physiologically relevant protein complexes (https://string-db.org/).

In the third elution step (E3) we employed astringent conditions to recover the remaining interactors of PknG. Using the statistical model of Venn Diagram analysis for comparison with mock experiments, 24 proteins were identified in this fraction using whole length PknG and 22 of them were also recovered using $PknG_{\Delta 73}$, evidencing a confident list of partners (**Table S4**).

Overall, we identify 66 proteins in the PknG interactome (**Table S5**). Bioinformatics analysis showed that our set of putative PknG partners presented a significant enrichment of previously reported protein-protein interactions (PPI enrichment p-value < 1.0e-16). The biological processes that were statistically enriched in the PknG interactome were: ATP synthesis coupled proton transport; proton transmembrane transport; translation and ribosome biogenesis (**Table S6**). Moreover, the differential analysis of the interactome of PknG and $PknG_{\Delta 73}$ indicates that GarA and FhaA are recruited exclusively by full-length PknG, pointing to a specific interaction with pThr residues of the kinase N-terminal extension.

3.3. Comparative analysis of the proteomes of wild type M. tuberculosis and a PknG-null mutant

To evaluate if the newly identified interactors could represent physiological substrates of PknG, we compared the proteomes of wild type *M. tuberculosis* and a *pknG* deletion mutant (Δ*pknG*) by difference

in gel electrophoresis (DIGE). This approach allows an estimation of the relative abundance of proteoforms with different isoelectric points (pI), representing therefore a valuable tool for the analysis of kinase substrates. Image analysis led to the detection of around 2100 spots in each gel. Spots exhibiting normalized abundance fold changes greater than 25% and p-values lower than 0.05 (Student's t-test) were considered as differential spots. Not surprisingly, the differential spots that presented the largest overall fold changes could be assigned by LC-MS/MS analysis to GarA (**Fig. 1A and B**). Four spots corresponding to GarA proteoforms with different apparent pI values (spots 1 to 4) presented decreased normalized abundance in the Δ*pknG* strain, with fold changes ranging from 1.52 to 4.38 (**Fig. 1B and Table S7**). At the same time, the more basic GarA spot (spot 5 in **Figure 1**) was overrepresented in Δ*pknG* strain, with an increased normalized abundance of 2.30 (**Fig. 1B and Table S7**). The pattern of phosphorylated proteoforms of GarA was previously reported [30], and our results are consistent with diminished levels of GarA phosphorylation in the Δ*pknG* strain. In fact, Thr_{21}, the specific residue modified by PknG, was unequivocally detected as phosphorylated in the most acidic proteoform of GarA overrepresented in wild type *M. tuberculosis* (**Table S7**). This finding confirms previous reports indicating that GarA is a physiological substrate of PknG [15,20] and demonstrates the utility of the DIGE approach to identify protein substrates of this kinase.

The analysis by DIGE allowed the identification of an additional area with a differential pattern of spots between wild type *M. tuberculosis* and the Δ*pknG* mutant strain, possibly reflecting differences in protein phosphorylation. Interestingly, these spots were identified as GS (Rv2220) (spots 6 to 13 in **Fig. 1B and Table S7**). However, we failed to detect phosphopeptides in these very faint spots by MS, possibly due to a combination of the low amount of available protein and the analytical difficulties associated with the ionization and fragmentation of the phosphorylated peptides. To evaluate if phosphorylation by PknG could be responsible for these differential spots, we performed two complementary experiments. First, we treated protein extracts from wild type *M. tuberculosis* with alkaline phosphatase and analyzed changes in the relative abundance of spots by differential in gel

analysis (DIA). Then, the same approach was used to compare protein extracts from the *M. tuberculosis* Δ*pknG* mutant strain before and after incubation with PknG. Spots 6-8 were not completely resolved in this DIA gel, therefore they were analyzed as a single spot. Interestingly, the relative intensity of GS acidic spots, that were differential between strains, decreased after the treatment with phosphatase (fold change of 6.5 for spot 6-8 and fold change of 5.3 for GS spot 9) (**Fig. 2**). Additionally, the relative intensity of GS spot 9, underrepresented in Δ*pknG* strain, increased upon incubation with PknG (fold change of 2.16), while spot 6-8 didn't show difference in its abundance.

However, incubations with PknG did not change the relative abundance of the more basic GS spots (spots 10-13). Overall, these results confirm that GS is modified by phosphorylation *in vivo* and show that the incubation of protein extracts of *M. tuberculosis* Δ*pknG* with PknG can partially recapitulate the pattern of GS spots observed in the wild type strain.

3.4. PknG phosphorylates specific residues of GS

The results described above point to GS as a new PknG substrate. To further validate these results, and to identify the involved phosphoresidues, recombinant GS was incubated with PknG under phosphorylation conditions using different enzyme:substrate ratios. Phosphorylated peptides were subsequently identified by tryptic digestion and LC-MS/MS. After incubation with PknG, the N-terminal region of GS was systematically phosphorylated, even with an enzyme:substrate molar ratio as low as 1:150. Two peptides of *m/z* 833.4 and 1168.2, corresponding to doubly charged ions of phosphorylated sequences 45-59 and 60-79 respectively, were detected. Thr_{77} were unequivocally identified as the phosphorylated residue within sequence 60-79 by phosphoRS analysis (pRS probability 100%) (**Fig. 3B** and **Table S8**). In addition, manual inspection of the spectra indicated the presence of several y ions (y_5^+, y_5^+-H_3PO_4, y_7^+, y_7^+- H_3PO_4, y_8^+, y_8^+-H_3PO_4) that supports the phosphosite assignment. In the case of sequence 45-59, Ser_{57} was identified as the most probably phosphorylated residue by phosphoRS (pRS probability S_{57}:98.9%; $Ser_{56:}$ 0.2%), and this was confirmed by the presence of a signal corresponding to y_3^+ ion in MS/MS spectra (**Fig. 3A** and **Table S8**). While Thr_{77} is exposed to the solvent and relatively

distant from the active site, Ser_{57} is located in close proximity to an oligomerization interface and points to the active site of an adjacent monomer in the dodecameric structure of GS (**Fig. 4**). Thus, it is tempting to speculate that phosphorylation of Ser_{57} could interfere with substrate binding and/or GS oligomerization. Additionally, a third phosphopeptide (residues 404-429) was systematically identified in phosphorylation assays (triply charged ion of *m/z* 933.1), and although unambiguous phosphosite identification is not possible from the spectra obtained, Thr_{421} is the most probably phosphorylated residue within this sequence (pRS site probability 90.5%) (**Fig. 3C** and **Table S8**). Further support of Thr_{421} phosphorylation by PknG is provided by the detection of sequence 404-429 with two modifications: phosphorylation (Thr_{421}) and AMPylation (Tyr_{406}). In this case, the presence in the MS/MS spectrum of the fragment ions y_{10}^+, y_{10}^+-H_3PO_4; and a small signal corresponding to y_9^+-H_3PO_4 point to Thr_{421} and not Thr_{420} as the phosphorylated residue (**Fig. 3D** and **Table S8)**. Interestingly, Thr_{421} is close in the primary and the tertiary structure of the protein to the conserved residue Tyr_{406}, whose AMPylation is a well-known mechanism for the modulation of bacterial GS activity [43,44]. The fact that the phosphorylation of this sequence can occur alone or in combination with the AMPylation of Tyr_{406} suggests that both residues could have related regulatory roles.

3.5. FhaA as a new substrate of M. tuberculosis PknG

M. tuberculosis encodes several FHA-containing proteins and all of them have been identified as substrates of multiple STPKs *in vitro* [30,45–48]. With our interactomic approach we systematically recovered one of these proteins, FhaA, as the only PknG interactor recovered under dephosphorylation conditions when using the full-length kinase but not $PknG_{\Delta 73}$, suggesting that the interaction between FhaA and PknG strongly relies on the N-terminal pThr residues of PknG. A recent work reported the interaction of FhaA with several mycobacterial kinases, but failed to detect an interaction with PknG [24]. We employed surface plasmon resonance to determine whether FhaA is able to establish a direct interaction with PknG. Phosphorylated PknG formed a stable complex with immobilized FhaA whereas truncated $PknG_{\Delta 73}$ lacking all phosphosites did not (**Supplementary Fig. 2**). Overall, our results show

that PknG and FhaA can interact in the absence of additional factors and that the phosphorylated N-terminal extension of PknG is essential for this binding.

Then, we decided to assess FhaA phosphorylation by PknG *in vitro*. While no phosphopeptides were identified in LC-MS/MS analysis of recombinant phosphatase-treated FhaA, two FhaA phosphopeptides were confidently identified after incubation of the protein with PknG, and the phosphorylated residues were unequivocally localized. A peptide of *m/z* 472.3 (corresponding to the triply charged ion of phosphorylated sequence 14-25) and a peptide of *m/z* 816.2 (corresponding to the doubly charged ion of phosphorylated sequence 108-120) were systematically detected in the phosphorylation assays. These two sequences mapped to the N-terminal domain of FhaA and MS/MS spectra allowed the unequivocal identification of Thr_{18} and Thr_{116} (pRs scores of 79 and 110 respectively; 100% pRS probability in both cases) as the phosphorylated residues (**Fig. 5A** and **Fig. 5B,** respectively; and T**able S8**). Manual inspection of these spectra allowed the detection of numerous signals that support the specific modification of these residues (**Fig. 5A** and **B**, respectively).

To determine if these sites were also phosphorylated in living mycobacteria, recombinant *M. tuberculosis* FhaA produced in *M. smegmatis* was digested and analyzed by nano-LC MS/MS. Three FhaA phosphopeptides were found, two of them matching those previously identified in the *in vitro* phosphorylation assay (Thr_{18} in sequence 14-25 and Thr_{116} in sequence 108-120) (**Fig. 6A** and **Fig. 6B**, respectively; and **Table S8**). An additional phosphopeptide was detected corresponding to sequence 368-378, with Thr_{377} identified as the phosphorylation site (**Fig. 6C** and **table S8**). Previous studies reported that the kinase PknB is able to phosphorylate the FhaA residue Thr_{116} *in vitro* and several phosphoproteomics reports showed that this residue is also phosphorylated *in vivo* [49,50]. Even when it is difficult to exclude the possibility of physiologically non-relevant modifications due to protein overexpression, our results suggest that PknG might also contribute to the phosphorylation of Thr_{116} and is likely responsible for the specific phosphorylation of Thr_{18} in living mycobacteria.

4. Discussion

In this work we develop an AP-MS protocol for the identification of specific substrates and interactors of the mycobacterial protein kinase PknG by combining sequential elution steps and the use of different PknG constructions as baits. The proposed strategy presents several advantages that allowed us to report a confident list of PknG's interactors. On one hand the use of specific elution conditions (E1 and E2) allowed us to fractionate PknG interactome, facilitating the detection of less abundant proteins and at the same time decreasing the number of unspecific proteins eluted. On the other hand, the use PknG constructions with and without the autophosphorylated sites allowed us to discern which interactions relies on this docking site. Altogether, we report 66 proteins that participate in PknG-mediated protein complexes in mycobacteria. One striking feature of the interactome of PknG is the presence of a high number of ribosomal proteins and proteins related to translation. It is interesting to note that many of the identified PknG interactors are predicted to be part of the RelA regulon, which mediates the stringent response under nutrient limitation conditions [51] (**Table S5**). The presence of a large number of ribosomal proteins among PknG interactors might reflect the recovery of multiprotein complexes, possibly due to the simultaneous interaction of PknG with several specific proteins partners. Prisic *et al*. suggested that translation was regulated by Serine/Threonine protein kinase(s), as they identified phosphorylation sites in several ribosomal and ribosome-associated proteins in *M. tuberculosis* [5]. The occurrence of phosphorylated ribosomal proteins was further confirmed by phosphoproteomic analysis in related mycobacteria [6,52] . However, the impact of these modifications on protein biosynthesis is still not fully understood. Ribosomal proteins are among the most common contaminants in interactomes, but they are also part of physiologically relevant macromolecular complexes [53]. However, the analysis of proteins interacting with a mock resin, using three biological replicates, indicates that these proteins are not unspecific interactors. Thus, based on the employed experimental approach, the high enrichment of proteins related to translation and the available data reporting ribosomal proteins as kinases substrates (including PknG), we predict that ribosomal proteins are genuine partners of PknG. This conclusion is also supported by a recent report that identified protein translation as one of the main processes regulated by PknG [23]. Nakedi *et al*. used a label free phosphoproteomic approach to identify proteins that were

differentially phosphorylated in wild type *M. bovis* BCG and a PknG knockout mutant strain, and their list of 22 candidate PknG substrates include two ribosomal proteins (the 50S ribosomal protein L2 and the 30S ribosomal protein S16), both of which were also identified in the present work (**Table S5**) Overall, we recovered 6 out of the 22 PknG substrates reported by Nakedi *et al*, namely DnaK, RNA polymerase-binding protein RbpA, the 50S ribosomal protein L2, the 30S ribosomal protein S16, the ATP synthase beta subunit and the protein FhaA, thus strengthen the idea that these proteins represent physiological substrates/interactors of the kinase. Protein microarrays have also been used to identify PknG direct interactors [22,24]. Deng *et al* reported 59 protein interactors of PknG, while Wu *et al* reported 122 interactors among which 31 were exclusive of PknG while the remaining 91 were also interactors of at least one more protein kinase. Using the same microarray and an optimized protocol the authors increased from 59 to 122 the number of interactors found, but only 21 hits were common to the lists in both papers. This strategy seeks a different objective than ours; in one case direct interactors are identified while in the other protein complexes that include direct and indirect interactors are analyzed. However, there is strikingly little overlap between the lists of interactors found by protein microarrays and those obtained by using other approaches. In particular, none of the interactors reported by Wu *et al* was identified in our interactomic approach, while only one was identified by Nakedi *et al* (uncharacterized protein A0A0H3M4P0). Very interestingly, the strategy designed here allowed us to recover the only two previously well-characterized substrates of PknG: GarA and the 50 ribosomal protein L13, pointing to a confident list of biologically relevant interactors. None of the other reports on substrates or direct interactors of PknG succeeded in the identification of these validated substrates [22–24].

To cluster PknG protein partners we used a sequential elution scheme. In a first step using phosphorylation conditions, we recovered 8 proteins that represent either putative PknG substrates or components of substrate-mediated complexes. However, we were not able to identify phosphopeptides to support phosphorylation by PknG during elution. Several factors preclude phosphopeptide identification in protein mixtures. On one hand, phosphopeptides have lower ionization efficiencies and ionization is frequently suppressed by the presence of non-phosphorylated peptides. Additionally, phosphopeptides

generate low quality MS/MS spectra dominated by the neutral loss of the phosphate group, resulting in lower confidence of spectral matching [54,55]. In spite of this, several pieces of evidence support the relevance of this list of putative PknG substrates. First, GarA was systematically identified in all replicates of this fraction. Additionally, 6 out of 8 proteins recovered in E1 (GarA, GS, DnaK, 60kDa chaperonin, inorganic pyrophosphatase, and Acetyl-/propionyl-coenzyme A carboxylase alpha chain) have been previously reported as kinases substrates in phosphoproteomics and/or *in vitro* phosphorylation assays [7,20,30,49,56]. Remarkably, differences in DnaK phosphorylation have been recently reported for a strain lacking PknG [23]. Finally, comparative proteomics between wild type *M. tuberculosis* and a Δ*pknG* mutant strain using DIGE strongly suggest that two of the proteins recovered in E1, GarA and GS, are indeed differentially phosphorylated in the absence of PknG (**Fig. 1**).

The enzyme GS catalyzes the condensation of glutamate and ammonium to produce glutamine in the first step of nitrogen assimilation [57,58]. Overall, our results indicate that GS might represent a new substrate of PknG. First, we corroborated that GS is phosphorylated by PknG *in vitro* and we identified three residues as specific phosphorylation sites (**Fig. 3**). In addition, we provide evidence that GS is indeed phosphorylated in *M. tuberculosis,* and that phosphorylated proteoforms are underrepresented in a PknG null mutant (**Fig. 2**). Furthermore, the abundance of one of these GS proteoforms increased upon incubation of *M. tuberculosis* Δ*pknG* protein extracts with PknG, pointing to a direct effect of PknG on GS phosphorylation. Unfortunately, due to the low abundance of GS phosphorylated spots we could not confirm its phosphorylation site *in vivo*. Interestingly, a recent report supports the phosphorylation of GS *in vivo* within its N-terminal sequence, including the specific phosphorylation on Ser_{57} [59]. Thus, phosphorylation sites were assigned to Ser_{56} or Ser_{57} in the case of the sequence $_{45}$SVFDDGLAFDGSSIR$_{59}$ and to Ser_{63} or Ser_{67} in the case of the sequence $_{60}$GFQSIHESDMLLLPDPETAR$_{79}$. Moreover, these authors did not find any difference in the phosphorylation status of these peptides when inhibiting PknA and PknB, indicating that other kinase(s) must be responsible for these phosphorylation events. GS plays a key role in bacterial survival inside the host, regulating not only ammonium assimilation but also the synthesis of the poly L-glutamate/glutamine

cell wall structure [58,60]. GS synthesis and activity are tightly regulated at different levels and, in particular, the reversible AMPylation of the conserved residue Tyr_{406} is a well-established regulatory mechanism [61–63]. Our results open the possibility that protein phosphorylation also participates in the modulation of GS activity; however, more experiments will be required to verify this hypothesis. Notably, GS is involved in the bacterial nitrogen metabolism, similar to three other metabolic enzymes indirectly regulated by PknG. It has been shown previously that the unphosphorylated form of GarA is an allosteric inhibitor of alpha-ketoglutarate dehydrogenase and glutamate dehydrogenase, and an activator of glutamate synthase [20,31]. PknG switches off these modulatory activities by phosphorylating GarA at Thr_{21} [20,31]. Thus, GarA regulates in a very coordinated manner three of the four enzymes responsible for the balance between ammonium assimilation and glutamate oxidative deamination. Strikingly, the fourth enzyme that participates in this process, GS, was identified in the current study as a putative PknG substrate/interactor. In this way, PknG could regulate the four key enzymes for ammonium assimilation in mycobacteria through the concerted action on GarA and GS. Altogether, our results strongly suggest that GS is a new substrate of PknG. However, the specific effect of GS phosphorylation on its activity, oligomerization state and/or modulation by AMPylation remains to be established.

One of the aims of our interactomic approach was to identify PknG partners specifically interacting with the kinase N-terminal docking site(s). Two proteins consistently recruited only by full-length PknG contained an FHA domain: GarA and FhaA (**Table 1**). We demonstrated that FhaA is a substrate of PknG *in vitro* (**Fig. 5**) and that the phosphorylated N-terminus of PknG is required for the interaction (**Supp. Fig. 2**). The fact that FhaA was recovered in fraction E2 but not in E1 suggest that, opposite to GarA, the intramolecular recognition of the pThr residues by the FHA domain in FhaA is not a favoured and therefore, phosphorylation does not trigger disruption of the PknG-FhaA complex. FhaA is a two-domain protein with a C-terminal FHA domain and a N-terminal domain of unknown function connected *via* a long unstructured linker region (300 residues in *M. tuberculosis*) [50]. The *fhaA* gene is encoded by a highly conserved mycobacterial operon involved in the control of cell shape and cell division [64]. A previous study described the recruitment of FhaA by a phosphorylated Thr residue in the pseudokinase

MviN and the involvement of this complex in mycobacterial cell wall biosynthesis [64]. Global phosphoproteomic studies in mycobacteria showed that FhaA is phosphorylated *in vivo* both on Thr and on Tyr, however the kinase(s) responsible(s) for each phosphorylation event are still poorly characterized. [49]. Here we identified three phosphorylation sites in *M. tuberculosis* FhaA overproduced in *M. smegmatis* (Thr_{18}, Thr_{116} and Thr_{377}) and two of them (Thr_{116} and Thr_{377}) have been previously detected *in vivo*. In particular, previous studies reported that the kinase PknB is able to phosphorylate Thr_{116} [65] while it has been proposed that PknG can phosphorylate the conserved Thr_{377} (Thr_{371} in *M. bovis* BCG) [23]. Thus, Thr_{18} represents a putative new FhaA phosphorylation site. In addition, our results also suggest that PknG might contribute to the phosphorylation of Thr_{116} in FhaA *in vivo*. Indeed, the phosphorylation of the same substrate by different kinases, possibly in response to different stimuli, has already been reported in *M. tuberculosis* [20,30]. We have tried to compare FhaA phosphorylation in wild type *M. tuberculosis* and the Δ*pknG* mutant strain by DIGE and shotgun approaches; however we failed to obtain reasonable read-outs, possibly due to the low abundance of these proteins in mycobacterial extracts. In spite of this, our results suggest that PknG might play a role in cell wall synthesis through FhaA phosphorylation. Interestingly, Nakedi *et al* recently identified the FhaA orthologue in *M. bovis* as a physiological substrate of PknG [23].

Altogether, we used a reliable AP-MS protocol to study the interactome of PknG and identify new substrates, interactors and processes regulated by this kinase. Our results point to nitrogen and energy metabolism, cell wall biosynthesis and protein translation as processes potentially modulated by PknG.

M. tuberculosis must reprogram its metabolism and gene expression profile in response to external stimuli to survive inside the host. The accumulating evidence indicates that PknG is involved in the regulation of core processes in the bacterial physiology, essential for the adaptation to the host environment. Further studies on the newly reported substrates/interactors are required to test this hypothesis, which warrants interesting insights about new drug targets to fight tuberculosis.

5. Conclusion

PknG is a Serine/Threonine protein kinase recognized as a key player in mycobacterial physiology and pathogenesis. The complete set of substrates and interactors that participate in signal transduction together with PknG needs to be identified to understand the role of the kinase at the molecular level.

Using a tailored approach to study the interactome of PknG, we have identified two new substrates: the enzyme glutamine synthetase and the protein FhaA. Our results show that PknG phosphorylates *in vitro* specific residues in both glutamine synthetase and FhaA, and we provide evidence that these proteins are likely phosphorylated by PknG in living mycobacteria. Our work complements previous reports on PknG substrates and interactors, and reinforces a central role of PknG in the control of the bacterial nitrogen metabolism. In addition, we provide evidence that PknG might regulate other processes in mycobacterial physiology, including protein translation and cell wall synthesis.

Acknowledgments: This work was funded by grants from the Agencia Nacional de Investigación e Innovación, Uruguay (ANII) [FCE_3_2013_1_100358 and FCE_1_2014_1_104045] and FOCEM (MERCOSUR Structural Convergence Fund, COF 03/11). MG, JR and BR were supported by a fellowship from ANII [POS_NAC_2012_1_8824, POS_NAC_2015_1_109755, POS_FCE_2015_1_1005186] and AC was supported by Agence Nationale pour la Recherche (France) [grant 09 BLAN 0400 01]. We thank M. Portela for her excellent technical support. We also thank Dr. Av-Gay and Dr. S. Mowbray for kindly providing *M. tuberculosis ΔpknG* strain and Rv2220 plasmid respectively.

Conflict of interest: The authors declare that they have no conflicts of interest with the contents of this article.

References:

[1] World Health Organization. Annual TB Report, (2017). http://www.searo.who.int/tb/documents/annual_tb_repot_2017/en/.

[2] M. Gengenbacher, S.H.E. Kaufmann, *Mycobacterium tuberculosis* : success through dormancy, FEMS Microbiol. Rev. 36 (2012) 514–532. doi:10.1111/j.1574-6976.2012.00331.x.

[3] D.G. Russell, P.-J. Cardona, M.-J. Kim, S. Allain, F. Altare, Foamy macrophages and the progression of the human tuberculosis granuloma, Nat. Immunol. 10 (2009) 943–948. doi:10.1038/ni.1781.

[4] C. Ortega, R. Liao, L.N. Anderson, T. Rustad, A.R. Ollodart, A.T. Wright, D.R. Sherman, C. Grundner, Mycobacterium tuberculosis Ser/Thr Protein Kinase B Mediates an Oxygen-Dependent Replication Switch, PLoS Biol. 12 (2014). doi:10.1371/journal.pbio.1001746.

[5] S. Prisic, S. Dankwa, D. Schwartz, M.F. Chou, J.W. Locasale, C.-M. Kang, G. Bemis, G.M. Church, H. Steen, R.N. Husson, Extensive phosphorylation with overlapping specificity by Mycobacterium tuberculosis serine/threonine protein kinases., Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 7521–6. doi:10.1073/pnas.0913482107.

[6] S. Fortuin, G.G. Tomazella, N. Nagaraj, S.L. Sampson, N.C. Gey van Pittius, N.C. Soares, H.G. Wiker, G.A. de Souza, R.M. Warren, Phosphoproteomics analysis of a clinical Mycobacterium tuberculosis Beijing isolate: Expanding the mycobacterial phosphoproteome catalog, Front. Microbiol. 6 (2015) 1–12. doi:10.3389/fmicb.2015.00006.

[7] U. Kusebauch, C. Ortega, A. Ollodart, R.S. Rogers, D.R. Sherman, R.L. Moritz, C. Grundner, Mycobacterium tuberculosis supports protein tyrosine phosphorylation, Proc. Natl. Acad. Sci. 111 (2014) 9265–9270. doi:10.1073/pnas.1323894111.

[8] S. Prisic, R.N. Husson, Mycobacterium tuberculosis Serine/Threonine Protein Kinases, Microbiol. Spectr. 2 (2014). doi:10.1128/microbiolspec.MGM2-0006-2013.

[9] D.R. Sherman, C. Grundner, Agents of change - concepts in Mycobacterium tuberculosis Ser/Thr/Tyr phosphosignalling., Mol. Microbiol. 94 (2014) 231–41. doi:10.1111/mmi.12747.

694 [10] S.T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier,
695 S. Gas, C.E. Barry, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R.
696 Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A.
697 Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M.A. Quail, M.-A. Rajandream, J.
698 Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J.E. Sulston, K. Taylor, S.
699 Whitehead, B.G. Barrell, Deciphering the biology of Mycobacterium tuberculosis from the
700 complete genome sequence, Nature. 393 (1998) 537–544. doi:10.1038/31159.
701 [11] Y. Av-Gay, M. Everett, The eukaryotic-like Ser/Thr protein kinases of Mycobacterium
702 tuberculosis., Trends Microbiol. 8 (2000) 238–44.
703 http://www.ncbi.nlm.nih.gov/pubmed/10785641.
704 [12] S. Cowley, M. Ko, N. Pick, R. Chow, K.J. Downing, B.G. Gordhan, J.C. Betts, V. Mizrahi, D. a
705 Smith, R.W. Stokes, Y. Av-Gay, The Mycobacterium tuberculosis protein serine/threonine kinase
706 PknG is linked to cellular glutamate/glutamine levels and is important for growth in vivo., Mol.
707 Microbiol. 52 (2004) 1691–702. doi:10.1111/j.1365-2958.2004.04085.x.
708 [13] N.A. Kruh, J. Troudt, A. Izzo, J. Prenni, K.M. Dobos, Portrait of a pathogen: the Mycobacterium
709 tuberculosis proteome in vivo., PLoS One. 5 (2010) e13938. doi:10.1371/journal.pone.0013938.
710 [14] A. Wehenkel, M. Bellinzoni, M. Graña, R. Duran, A. Villarino, P. Fernandez, G. Andre-Leroux, P.
711 England, H. Takiff, C. Cerveñansky, S.T. Cole, P.M. Alzari, Mycobacterial Ser/Thr protein
712 kinases and phosphatases: physiological roles and therapeutic potential., Biochim. Biophys. Acta.
713 1784 (2008) 193–202. doi:10.1016/j.bbapap.2007.08.006.
714 [15] B. Rieck, G. Degiacomi, M. Zimmermann, A. Cascioferro, F. Boldrin, N.R. Lazar-Adler, A.R.
715 Bottrill, F. le Chevalier, W. Frigui, M. Bellinzoni, M.N. Lisa, P.M. Alzari, L. Nguyen, R. Brosch,
716 U. Sauer, R. Manganelli, H.M. O'Hare, PknG senses amino acid availability to control metabolism
717 and virulence of Mycobacterium tuberculosis, 2017. doi:10.1371/journal.ppat.1006399.
718 [16] J. Walburger, A., Koul, A., Nguyen, L., Prescianotto-Baschong, C., Huygen, K., Klebl, B.,
719 Thompson, C., Bacher, C., Pieters, Protein Kinase G from Pathogenic Mycobacteria Promotes

Survival Within Macrophages, Science (80-.). 304 (2004) 1800–1804. doi:10.1126/science.1099384.

[17] M.Z. Khan, A. Bhaskar, S. Upadhyay, P. Kumari, R.S. Rajmani, P. Jain, A. Singh, D. Kumar, N.S. Bhavesh, V.K. Nandicoori, Protein kinase G confers survival advantage to Mycobacterium tuberculosis during latency-like conditions., J. Biol. Chem. 292 (2017) 16093–16108. doi:10.1074/jbc.M117.797563.

[18] R. Paroha, R. Chourasia, R. Mondal, S.K. Chaurasiya, PknG supports mycobacterial adaptation in acidic environment, Mol. Cell. Biochem. (2017) 1–12. doi:10.1007/s11010-017-3211-x.

[19] K. a Wolff, H.T. Nguyen, R.H. Cartabuke, A. Singh, S. Ogwang, L. Nguyen, Protein kinase G is required for intrinsic antibiotic resistance in mycobacteria., Antimicrob. Agents Chemother. 53 (2009) 3515–9. doi:10.1128/AAC.00012-09.

[20] H.M. O'Hare, R. Durán, C. Cerveñansky, M. Bellinzoni, A.M. Wehenkel, O. Pritsch, G. Obal, J. Baumgartner, J. Vialaret, K. Johnsson, P.M. Alzari, Regulation of glutamate metabolism by protein kinases in mycobacteria., Mol. Microbiol. 70 (2008) 1408–23. doi:10.1111/j.1365-2958.2008.06489.x.

[21] K.A. Wolff, A.H. de la Peña, H.T. Nguyen, T.H. Pham, L.M. Amzel, S.B. Gabelli, L. Nguyen, A Redox Regulatory System Critical for Mycobacterial Survival in Macrophages and Biofilm Development, PLoS Pathog. 11 (2015) 1–20. doi:10.1371/journal.ppat.1004839.

[22] J. Deng, L. Bi, L. Zhou, S. Guo, J. Fleming, H. Jiang, Y. Zhou, J. Gu, Q. Zhong, Z. Wang, Z. Liu, R. Deng, J. Gao, T. Chen, W. Li, J. Wang, X. Wang, H. Li, F. Ge, G. Zhu, H. Zhang, J. Gu, F. Wu, Z. Zhang, D. Wang, H. Hang, Y. Li, L. Cheng, X. He, S. Tao, X. Zhang, Mycobacterium tuberculosis proteome microarray for global studies of protein function and immunogenicity., Cell Rep. 9 (2014) 2317–29. doi:10.1016/j.celrep.2014.11.023.

[23] K.C. Nakedi, B. Calder, M. Barnejee, A. Giddey, A.J. Nel, S. Garnett, J.M. Blackburn, N.A. Da Cruz Soares, Identification of novel physiological substrates of Mycobacterium Bovis BCG Protein Kinase G (PknG) by label-free quantitative phosphoproteomics., Mol. Cell. Proteomics.

(2018) mcp.RA118.000705. doi:10.1074/mcp.RA118.000705.

[24] F.-L. Wu, Y. Liu, H.-W. Jiang, Y. Luan, H. Zhang, X. He, Z.-W. Xu, J.-L. Hou, L.-Y. Ji, Z. Xie, D.M. Czajkowsky, W. Yan, J.-Y. Deng, L.-J. Bi, X.-E. Zhang, S.-C. Tao, The Ser / Thr Protein Kinase Protein-Protein Interaction Map of M . tuberculosis, Mol. Cell. Proteomics. 16 (2017) 1–42. doi:10.1074/mcp.M116.065771.

[25] M. Gil, M. Graña, F.J. Schopfer, T. Wagner, A. Denicola, B.A. Freeman, P.M. Alzari, C. Batthyány, R. Durán, R.D. M. Gil, M. Graña, F. J. Schopfer, T. Wagner, A. Denicola, B. A. Freeman, P. M. Alzari, C. Batthyány, Inhibition of Mycobacterium tuberculosis PknG by non-catalytic rubredoxin domain specific modification: reaction of an electrophilic nitro-fatty acid with the Fe-S center, Free Radic. Biol. Med. 65 (2013) 150–161. doi:10.1038/jid.2014.371.

[26] M.N. Lisa, M. Gil, G. André-Leroux, N. Barilone, R. Durán, R.M. Biondi, P.M. Alzari, Molecular Basis of the Activity and the Regulation of the Eukaryotic-like S/T Protein Kinase PknG from Mycobacterium tuberculosis, Structure. 23 (2015) 1039–1048. doi:10.1016/j.str.2015.04.001.

[27] N. Scherr, S. Honnappa, G. Kunz, P. Mueller, R. Jayachandran, F. Winkler, J. Pieters, M.O. Steinmetz, Structural basis for the specific inhibition of protein kinase G, a virulence factor of Mycobacterium tuberculosis., Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 12151–6. doi:10.1073/pnas.0702842104.

[28] D. Durocher, S.P. Jackson, The FHA domain, FEBS Lett. 513 (2002) 58–66. doi:10.1016/S0014-5793(01)03294-X.

[29] P. England, A. Wehenkel, S. Martins, S. Hoos, G. André-Leroux, A. Villarino, P.M. Alzari, The FHA-containing protein GarA acts as a phosphorylation-dependent molecular switch in mycobacterial signaling, FEBS Lett. 583 (2009) 301–307. doi:10.1016/j.febslet.2008.12.036.

[30] A. Villarino, R. Duran, A. Wehenkel, P. Fernandez, P. England, P. Brodin, S.T. Cole, U. Zimny-Arndt, P.R. Jungblut, C. Cerveñansky, P.M. Alzari, Proteomic identification of M. tuberculosis protein kinase substrates: PknB recruits GarA, a FHA domain-containing protein, through activation loop-mediated interactions., J. Mol. Biol. 350 (2005) 953–63.

doi:10.1016/j.jmb.2005.05.049.

[31] T.J. Nott, G. Kelly, L. Stach, J. Li, S. Westcott, D. Patel, D.M. Hunt, S. Howell, R.S. Buxton, H.M. O'Hare, S.J. Smerdon, An Intramolecular Switch Regulates Phosphoindependent FHA Domain Interactions in Mycobacterium tuberculosis, Sci. Signal. 2 (2009) ra12-ra12. doi:10.1126/scisignal.2000212.

[32] J.C. van Kessel, L.J. Marinelli, G.F. Hatfull, Recombineering mycobacteria and their phages, Nat. Rev. Microbiol. 6 (2008) 851–857. doi:10.1038/nrmicro2014.

[33] W.W. Krajewski, T.A. Jones, S.L. Mowbray, Structure of Mycobacterium tuberculosis glutamine synthetase in complex with a transition-state mimic provides functional insights, Proc. Natl. Acad. Sci. 102 (2005) 10499–10504. doi:10.1073/pnas.0502248102.

[34] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels, Anal. Chem. 68 (1996) 850–858. doi:10.1021/ac950914h.

[35] P.C. Carvalho, J.S.G. Fischer, T. Xu, J.R. Yates, V.C. Barbosa, PatternLab: From mass spectra to label-free differential shotgun proteomics, Curr. Protoc. Bioinforma. (2012) 1–18. doi:10.1002/0471250953.bi1319s40.

[36] P.C. Carvalho, D.B. Lima, F. V Leprevost, M.D.M. Santos, J.S.G. Fischer, P.F. Aquino, J.J. Moresco, J.R. Yates, V.C. Barbosa, J.R.Y. Iii, V.C. Barbosa, Integrated analysis of shotgun proteomic data with PatternLab for proteomics 4.0, Nat. Protoc. 11 (2016) 102–117. doi:10.1038/nprot.2015.133.

[37] P.C. Carvalho, J.S.G. Fischer, J. Perales, J.R. Yates, V.C. Barbosa, E. Bareinboim, Analyzing marginal cases in differential shotgun proteomics., Bioinformatics. 27 (2011) 275–6. doi:10.1093/bioinformatics/btq632.

[38] A. Kapopoulou, J.M. Lew, S.T. Cole, The MycoBrowser portal: A comprehensive and manually annotated resource for mycobacterial genomes, Tuberculosis. 91 (2011) 8–13. doi:10.1016/j.tube.2010.09.006.

[39] W. Gish, D.J. States, Identification of protein coding regions by database similarity search, Nat Genet. 3 (1993) 266–272. doi:10.1038/ng0393-266.

[40] D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K.P. Tsafou, M. Kuhn, P. Bork, L.J. Jensen, C. Von Mering, STRING v10: Protein-protein interaction networks, integrated over the tree of life, Nucleic Acids Res. 43 (2015) D447–D452. doi:10.1093/nar/gku1003.

[41] H. Mi, S. Poudel, A. Muruganujan, J.T. Casagrande, P.D. Thomas, PANTHER version 10: Expanded protein families and functions, and analysis tools, Nucleic Acids Res. 44 (2016) D336–D342. doi:10.1093/nar/gkv1194.

[42] J.A. Vizcaino, A. Csordas, N. Del-Toro, J.A. Dianes, J. Griss, I. Lavidas, G. Mayer, Y. Perez-Riverol, F. Reisinger, T. Ternent, Q.W. Xu, R. Wang, H. Hermjakob, 2016 update of the PRIDE database and its related tools, Nucleic Acids Res. 44 (2016) D447–D456. doi:10.1093/nar/gkv1145.

[43] R. Mehta, J.T. Pearson, S. Mahajan, A. Nath, M.J. Hickey, D.R. Sherman, W.M. Atkins, Adenylylation and catalytic properties of Mycobacterium tuberculosis glutamine synthetase expressed in Escherichia coli versus mycobacteria., J. Biol. Chem. 279 (2004) 22477–82. doi:10.1074/jbc.M401652200.

[44] B.M. Shapiro, H.S. Kingdon, E.R. Stadtman, Regulation of glutamine synthetase. VII. Adenylyl glutamine synthetase: a new form of the enzyme with altered regulatory and kinetic properties., Proc. Natl. Acad. Sci. U. S. A. 58 (1967) 642–9. http://www.ncbi.nlm.nih.gov/pubmed/4860756 (accessed April 25, 2018).

[45] C. Grundner, L.M. Gay, T. Alber, Mycobacterium tuberculosis serine/threonine kinases PknB, PknD, PknE, and PknF phosphorylate multiple FHA domains., Protein Sci. 14 (2005) 1918–1921. doi:10.1110/ps.051413405.

[46] M. Gupta, A. Sajid, G. Arora, V. Tandon, Y. Singh, Forkhead-associated domain-containing protein Rv0019c and polyketide-associated protein PapA5, from substrates of serine/threonine

protein kinase PknB to interacting proteins of mycobacterium tuberculosis, J. Biol. Chem. 284 (2009) 34723–34734. doi:10.1074/jbc.M109.058834.

[47] V. Molle, D. Soulat, J.M. Jault, C. Grangeasse, A.J. Cozzone, J.F. Prost, Two FHA domains on an ABC transporter, Rv1747, mediate its phosphorylation by PknF, a Ser/Thr protein kinase from Mycobacterium tuberculosis, FEMS Microbiol. Lett. 234 (2004) 215–223. doi:10.1016/j.femsle.2004.03.033.

[48] K. Sharma, M. Gupta, A. Krupa, N. Srinivasan, Y. Singh, EmbR, a regulatory protein with ATPase activity, is a substrate of multiple serine/threonine kinases and phosphatase in Mycobacterium tuberculosis., FEBS J. 273 (2006) 2711–21. doi:10.1111/j.1742-4658.2006.05289.x.

[49] B. Calder, C. Albeldas, J.M. Blackburn, N.C. Soares, Mass spectrometry offers insight into the role of ser/thr/tyr phosphorylation in the mycobacteria, Front. Microbiol. 7 (2016) 1–8. doi:10.3389/fmicb.2016.00141.

[50] C. Roumestand, J. Leiba, N. Galophe, E. Margeat, A. Padilla, Y. Bessin, P. Barthe, V. Molle, M. Cohen-Gonsaud, Structural insight into the Mycobacterium tuberculosis Rv0020c protein and its interaction with the PknB kinase., Structure. 19 (2011) 1525–34. doi:10.1016/j.str.2011.07.011.

[51] J.L. Dahl, C.N. Kraus, H.I.M. Boshoff, B. Doan, K. Foley, D. Avarbock, G. Kaplan, V. Mizrahi, H. Rubin, C.E. Barry, The role of RelMtb-mediated adaptation to stationary phase in long-term persistence of Mycobacterium tuberculosis in mice., Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 10026–31. doi:10.1073/pnas.1631248100.

[52] R. Verma, S.M. Pinto, A.H. Patil, J. Advani, P. Subba, M. Kumar, J. Sharma, G. Dey, R. Ravikumar, S. Buggi, P. Satishchandra, K. Sharma, M. Suar, S.P. Tripathy, D.S. Chauhan, H. Gowda, A. Pandey, S. Gandotra, T.S.K. Prasad, Quantitative Proteomic and Phosphoproteomic Analysis of H37Ra and H37Rv Strains of Mycobacterium tuberculosis, J. Proteome Res. 16 (2017) 1632–1645. doi:10.1021/acs.jproteome.6b00983.

[53] J.S. Rees, N. Lowe, I.M. Armean, J. Roote, G. Johnson, E. Drummond, H. Spriggs, E. Ryder, S.

Russell, D.S. Johnston, K.S. Lilley, In Vivo Analysis of Proteomes and Interactomes Using Parallel Affinity Capture (iPAC) Coupled to Mass Spectrometry, Mol. Cell. Proteomics. 10 (2011) M110.002386-M110.002386. doi:10.1074/mcp.M110.002386.

[54] N. Dephoure, K.L. Gould, S.P. Gygi, D.R. Kellogg, Mapping and analysis of phosphorylation sites: a quick guide for cell biologists, Mol. Biol. Cell. 24 (2013) 535–542. doi:10.1091/mbc.E12-09-0677.

[55] M. Mann, S. Ong, M. Gr, H. Steen, O.N. Jensen, A. Pandey, Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome., Trends Biotechnol. 20 (2002) 261–8. doi:10.1016/S0167-7799(02)01944-3.

[56] M.J. Canova, R. Veyron-Churlet, I. Zanella-Cleon, M. Cohen-Gonsaud, A.J. Cozzone, M. Becchi, L. Kremer, V. Molle, 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 (2008) 521–533. doi:10.1002/pmic.200700442.

[57] A. Gouzy, Y. Poquet, O. Neyrolles, Nitrogen metabolism in Mycobacterium tuberculosis physiology and virulence, Nat. Rev. Microbiol. 12 (2014) 729–737. doi:10.1038/nrmicro3349.

[58] M. V Tullius, G. Harth, M. a Horwitz, Glutamine Synthetase GlnA1 Is Essential for Growth of Mycobacterium tuberculosis in Human THP-1 Macrophages and Guinea Pigs Glutamine Synthetase GlnA1 Is Essential for Growth of Mycobacterium tuberculosis in Human THP-1 Macrophages and Guinea Pigs, Infect. Immun. 71 (2003) 3927–3936. doi:10.1128/IAI.71.7.3927.

[59] X. Carette, J. Platig, D.C. Young, M. Helmel, A.T. Young, Z. Wang, L.-P. Potluri, C.S. Moody, J. Zeng, S. Prisic, J.N. Paulson, J. Muntel, A.V.R. Madduri, J. Velarde, J.A. Mayfield, C. Locher, T. Wang, J. Quackenbush, K.Y. Rhee, D.B. Moody, H. Steen, R.N. Husson, Multisystem Analysis of Mycobacterium tuberculosis Reveals Kinase-Dependent Remodeling of the Pathogen-Environment Interface., MBio. 9 (2018) e02333-17. doi:10.1128/mBio.02333-17.

[60] G. Harth, M. a Horwitz, An inhibitor of exported Mycobacterium tuberculosis glutamine

synthetase selectively blocks the growth of pathogenic mycobacteria in axenic culture and in human monocytes: extracellular proteins as potential novel drug targets., J. Exp. Med. 189 (1999) 1425–1436. doi:10.1084/jem.189.9.1425.

[61] P. Carroll, C.A. Pashley, T. Parish, Functional analysis of GlnE, an essential adenylyl transferase in Mycobacterium tuberculosis, J. Bacteriol. 190 (2008) 4894–4902. doi:10.1128/JB.00166-08.

[62] J.A. Leigh, J.A. Dodsworth, Nitrogen Regulation in Bacteria and Archaea, Annu. Rev. Microbiol. 61 (2007) 349–377. doi:10.1146/annurev.micro.61.080706.093409.

[63] E.R. Stadtman, The Story of Glutamine Synthetase Regulation, J. Biol. Chem. 276 (2001) 44357–44364. doi:10.1074/jbc.R100055200.

[64] C.C.L. Gee, K.G.K. Papavinasasundaram, S.R. Blair, C.E. Baer, A.M. Falick, D.S. King, J.E. Griffin, H. Venghatakrishnan, A. Zukauskas, J.-R. Wei, R.K. Dhiman, D.C. Crick, E.J. Rubin, C.M. Sassetti, T. Alber, A phosphorylated pseudokinase complex controls cell wall synthesis in mycobacteria., Sci. Signal. 5 (2012) 1–24. doi:10.1126/scisignal.2002525.

[65] C. Roumestand, J. Leiba, N. Galophe, E. Margeat, A. Padilla, Y. Bessin, P. Barthe, V. Molle, M. Cohen-Gonsaud, Structural insight into the Mycobacterium tuberculosis Rv0020c protein and its interaction with the PknB kinase, Structure. 19 (2011) 1525–1534. doi:10.1016/j.str.2011.07.011.

Table 1. Proteins identified in E1 and E2 fractions from the AP-MS experiment.

Proteins identified as exclusive interactors of full length PknG in E1		
UniProt identifier	**Description**	**Gene name**
A0QYG2	Glycogen accumulation regulator GarA	MSMEG_3647
A0R079	Glutamine synthetase	MSMEG_4290
Proteins identified as common interactors of PknG and $PknG_{\Delta 73}$ in E1		
UniProt identifier	**Description**	**Gene name**
A0QQC8	Chaperone protein DnaK	MSMEG_0709
A0R5M3	Alcohol dehydrogenase, iron containing	MSMEG_6242
A0QQU5	60 kDa chaperonin 1	MSMEG_0880
A0R2I1	4Fe-4S ferredoxin, iron-sulfur binding protein	MSMEG_5122
A0R597	Inorganic pyrophosphatase	MSMEG_6114
A0QTE1	Acetyl-/propionyl-coenzyme A carboxylase alpha chain	MSMEG_1807
Proteins identified as exclusive interactors of full length PknG in E2		
UniProt identifier	**Description**	**Gene name**
A0QNG7	FHA domain protein	MSMEG_0035

FIGURE LEGENDS

Figure 1. Analysis of the differential proteoform-patterns present in wild type *M. tuberculosis* and a PknG null mutant by DIGE.

A. Representative DIGE in pseudocolours. Protein spots over-represented in WT and Δ*pknG* strains are shown in green and red, respectively. Differential spots with patterns consistent with changes in protein phosphorylation are labelled. Protein spots in these areas were assigned to glutamine synthetase (GS,) and GarA. Average standardized abundance fold changes and protein identification details for each spot are shown in Table S7. **B.** Bar graphs showing the average standardized abundance fold changes of GarA spots (spots 1-5, upper panel) and GS spots (spots 6-13, lower panel). The estimated pI range of each group of spots is indicated below each graph. Fold changes for spots 1-12 have $p<0.05$ values.

Figure 2. Confirmation of GS phosphorylation status. A. DIA analysis of wild type *M. tuberculosis* protein extracts with or without treatment with protein phosphatase. **B.** DIA analysis of *M. tuberculosis* Δ*pknG* protein extracts with or without pre-incubation with PknG under phosphorylation conditions. Spots labels refers to spots in DIGE gel in Fig. 4, all identified as GS by MS; spots 6-8 were not completely resolved in this DIA gel, therefore, they were analyzed as a single spot. Spots 6-8 $\log_{10}$ volume ratio (WT/WT+phosphatase) = 1.65, Mascot protein score: 267; spot 9 $\log_{10}$ volume ratio (WT/WT+phosphatase) = 1.53, Mascot protein score: 149; spot 9 volume ratio (Δ*pknG*/Δ*pknG*+PknG) = 1.32, Mascot protein score: 111. Volume ratios are expressed as $10.\log_{10}$ (spot volume condition 1/spot volume condition 2). **C-D.** 3D view of the GS spot 9 showing the spot volume differences between each pair of conditions.

Figure 3. Identification of GS phosphorylation sites. In each case, C-terminal (blue) and N-terminal (red) fragment ions assigned to the sequence are indicated. Ions presenting the neutral loss characteristic of phosphorylation are indicated. **A.** Representative MS/MS spectrum of the doubly charged peptide

SVFDDGLAFDGS**p**S_{57}IR (*m/z* observed 833.4; Xcorr value: 4.81). PhoshpoRS indicates that Ser_{57} is the phosphorylated residue within the sequence (pRS score of 114; pRS site probablitiy S_{57}:99.8%; $Ser_{56:}$ 0.2%). The presence of y_3^+ fragment ion points to S_{57} and not S_{56} as the phosphoresidue. **B.** Representative MS/MS spectrum of doubly charged ion of sequence GFQSIHESDMLLLPDPE**p**T_{77}AR (*m/z* observed 1168.2; Xcorr: 4.34). PhoshpoRS indicates that Thr_{77} is the phosphorylated residue within the sequence (pRS score of 100; pRS probablitiy T_{77} 100%). The presence of many y ion fragments from $y3^+$ to $y8^+$, either containing the +80 modification and/or the phosphate neutral loss corroborates the assignment of Thr_{77} and not S_{67} or S_{63} as the posphoresidue. **C**. Representative MS/MS spectrum of triply charged ion of sequence DLYELPPEEAASIPQTP**p**T_{421}QLSDVIDR (*m/z* observed: 993.1; Xcorr: 4.06) and pRS score 58. PhosphoRS indicates that the most probably modified residue is Thr_{421} (pRS probability 90.5%). However the site is difficult to identify by manual inspection. **D.** MS/MS spectrum of the triply charged ion of sequence DL$_{\mathbf{AMP}}$YELPPEEAASIPQTP**p**T_{421}QLSDVIDR (*m/z* observed 1102.9; Xcorr 5.05) showing that both the AMPylation of Tyr_{406} [43] residue and the phosphorylation of a residue within this sequence, most probably Thr_{421}, can take place in the same peptide (pRS site probability for Thr421 92.8%). The presence of the fragment ions y_{10}^+, y_{10}^+-H_3PO_4; and a small signal corresponding to y_9^+-H_3PO_4 point to Thr_{421} and not Thr_{420} as the phosphorylated residue. Additional information regarding each phosphosite identification is depicted in Supplementary Table S8.

Figure 4. Location of phosphorylatable residues on GS structure. Crystal structure of the GS dodecameric form (PDB code 2BVC [33]) highlighting the three phosphorylatable sites (Ser_{57}, Thr_{77} and Thr_{421}) identified in this work and residue Tyr_{406}, found AMPylated in previous work [43]. Protomers are shown in cartoon representation in different colours, highlighted protein residues are shown in spheres and molecules in the substrate-binding pocket are shown in sticks. Right: view of the interfacial plane between the two GS 6-mer rings perpendicular to the protein 6-fold axis; center: lateral view of the GS dodecamer; left: amplification of the lateral view of the GS dodecamer.

Figure 5. Identification of *in vitro* phosphorylation sites of FhaA. *Strep*-tag® II-FhaA was purified from *M. smegmatis*, dephosphorylated with alkaline phosphatase, submitted to phosphorylation experiments using PknG as kinase and digested previous to MS analysis. y- (blue) and b- (red) ions assigned to the sequence are indicated. Ions presenting the neutral loss characteristic of phosphorylation are shown. **A**. Representative MS/MS spectrum of doubly charged ion of sequence FEQSSNLH**p**T_{116}GQFR (observed m/z 816.2, Xcorr 3.90). pRS score of 110 and pRS probability 100% for T_{116} allowed the unequivocal identification of the phosphorylation site. The presence of several y ions in the spectrum confirms this assignment (y_5^+; y_6^+; y_6^+-H_3PO_4; y_7^+; y_8^+). **B**. Representative MS/MS spectrum of triply charged ion of sequence KLEQ**p**T_{18}VGDAFAR (*m/z* 472.3, Xcorr value of 4.31). pRS score of 79 and pRS probability 100% for T_{18} allowed to unequivocally identify the phosphoresidue, which is confirmed by the presence of b ions from b_5 to b_{11}.

Figure 6. Identification of FhaA phosphorylation sites *in vivo*. *Strep*-tag® II-FhaA was purified from *M. smegmatis* and digested previous to mass spectrometry analysis. y- (blue) and b- (red) fragment ions assigned to the sequence are indicated. Ions presenting the neutral loss characteristic of phosphorylation are also shown. **A.** Representative MS/MS spectrum of doubly charged ion of sequence FEQSSNLH**p**T_{116}GQFR (*m/z* observed: 816.1; Xcorr: 4.35). Analysis using PhosphoRS indicates that T_{116} is the phosphorylated residue within the sequence (pRS score of 106; pRS probablitiy T_{116} 100%). Manual inspection of the spectra revealed the presence of fragment ions that allow confirming the phosphorylation of T_{116} and not S_{111} or S_{112}: y_5^+; y_6; $^+$ y_6^+-H_3PO_4; y_7^+; y_8^+). **B.** Representative MS/MS spectrum of triply charged ion of sequence KLEQ**p**T_{18}VGDAFAR (observed *m/z* 472.1; Xcorr value of 4.25). Both phosphoRS and manual inspection of the spectra fully supports the phosphorylation of T_{18} within the sequence (pRS score of 84; pRS probability 100%). **C.** Representative MS/MS spectrum of the doubly charged ion of sequence QDYGGGADY**p**T_{377}R (observed *m/z* 641.9; Xcorr value of 3.54). PhosphoRS analysis point to T_{377} as the *in vivo* phosphorylation site (pRS score of 91; pRS probability

99.5% for T_{377}). The presence of y_2^+ and y_2^+-H_3PO_4 fragment ions confirms the phosphorylation of the Thr_{377}, and not the Tyr_{376} residue.

Supporting information:

Supplementary Figures Legends

Supplementary Figure 1. Kinase activity of immobilized PknG. A. PknG autophosphorylation. Linear MALDI-TOF spectra of tryptic digestion of immobilized PknG (upper panel) and immobilized PknG previously incubated with ATP under phosphorylation conditions (bottom panel). Non-phosphorylated sequence 10-60 (theoretical *m/z* 5395.7), and its mono and diphosphorylated forms are indicated (theoretical *m/z* 5475.7 and 5555.7, respectively). While the predominant signal in recombinant PknG spectrum corresponds to the unphosphorylated peptide, the diphosphorylated specie is the most intense ion of sequence 10-60 after autophosphorylation. **B**. GarA phosphorylation by PknG. Linear MALDI-TOF mass spectra of recombinant GarA (*m/z* 20446, black line) and GarA phosphorylated by PknG immobilized on NHS-Activated Sepharose 4 Fast Flow (*m/z* 20526, grey line). The mass shift corresponds to the incorporation of one phosphate group (80 Da).

Supplementary Figure 2. PknG interacts with FhaA. FhaA was immobilized on a CM5 sensorchip by standard amine coupling. Phosphorylated PknG and $PknG_{\Delta 73}$ were diluted to a final concentration of 30 nM and injected during 180 s over the immobilized and a reference surfaces.

Supplementary Tables

Supplementary Table S1: Proteins identified in E1.

Supplementary Table S2: Proteins identified in E1: reported phosphosites in mycobacterial phosphoproteomic analysis.

Supplementary Table S3: Proteins identified in E2.

Supplementary Table S4: Proteins identified in E3.

Supplementary Table S5: List of proteins identified in PknG interactome.

Supplementary Table S6: Analysis of protein-protein interaction enrichment (STRING, http://string-db.org/cgi/network.pl). Overrepresentation biological processes enrichment test (PANTHER, http://www.pantherdb.org/).

Supplementary Table S7: Identification of protein spots from DIGE analysis.

Supplementary Table S8. Phosphorylation sites reported in this work.

Figure 1

Click here to download high resolution image

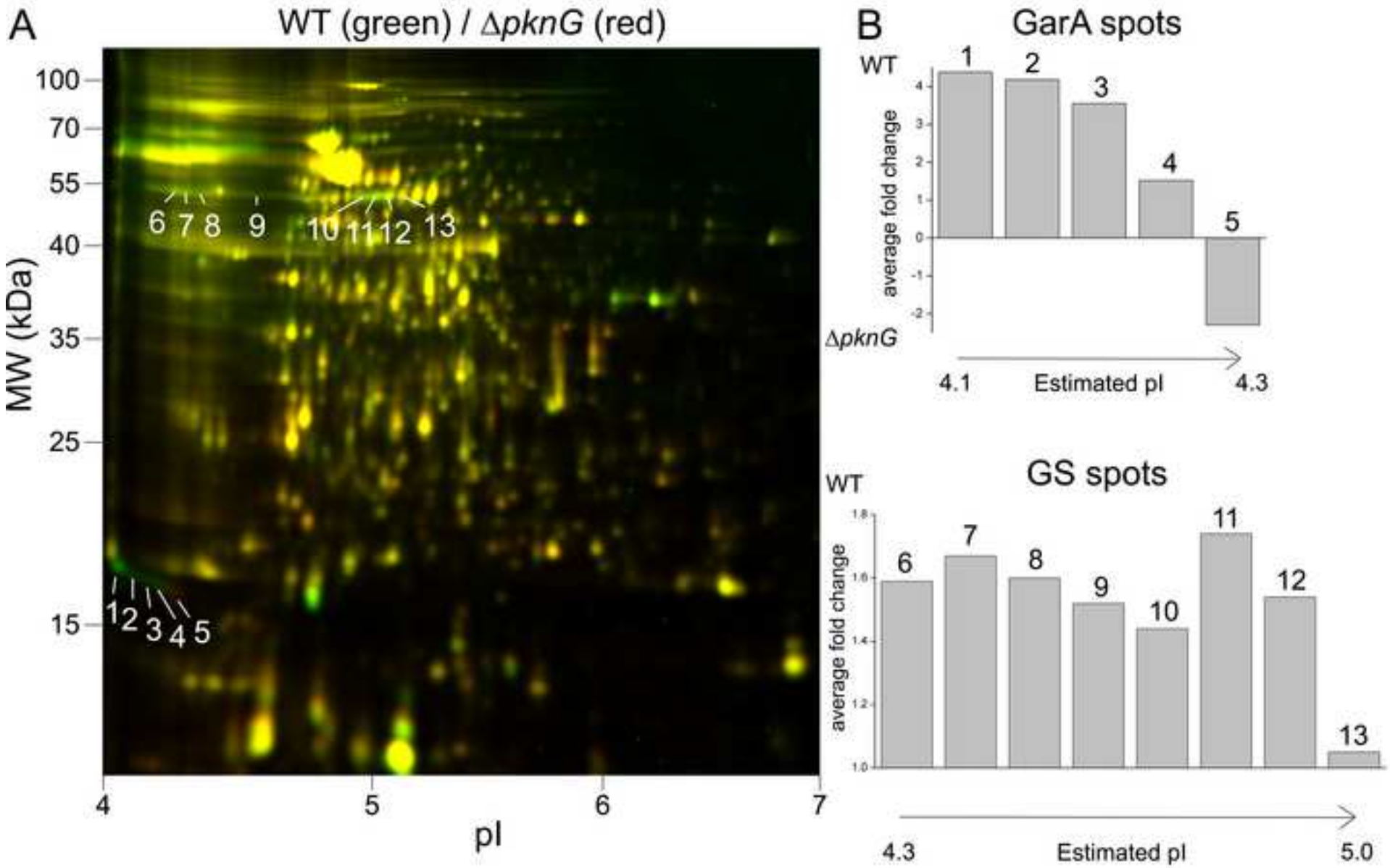

Figure 2

Click here to download high resolution image

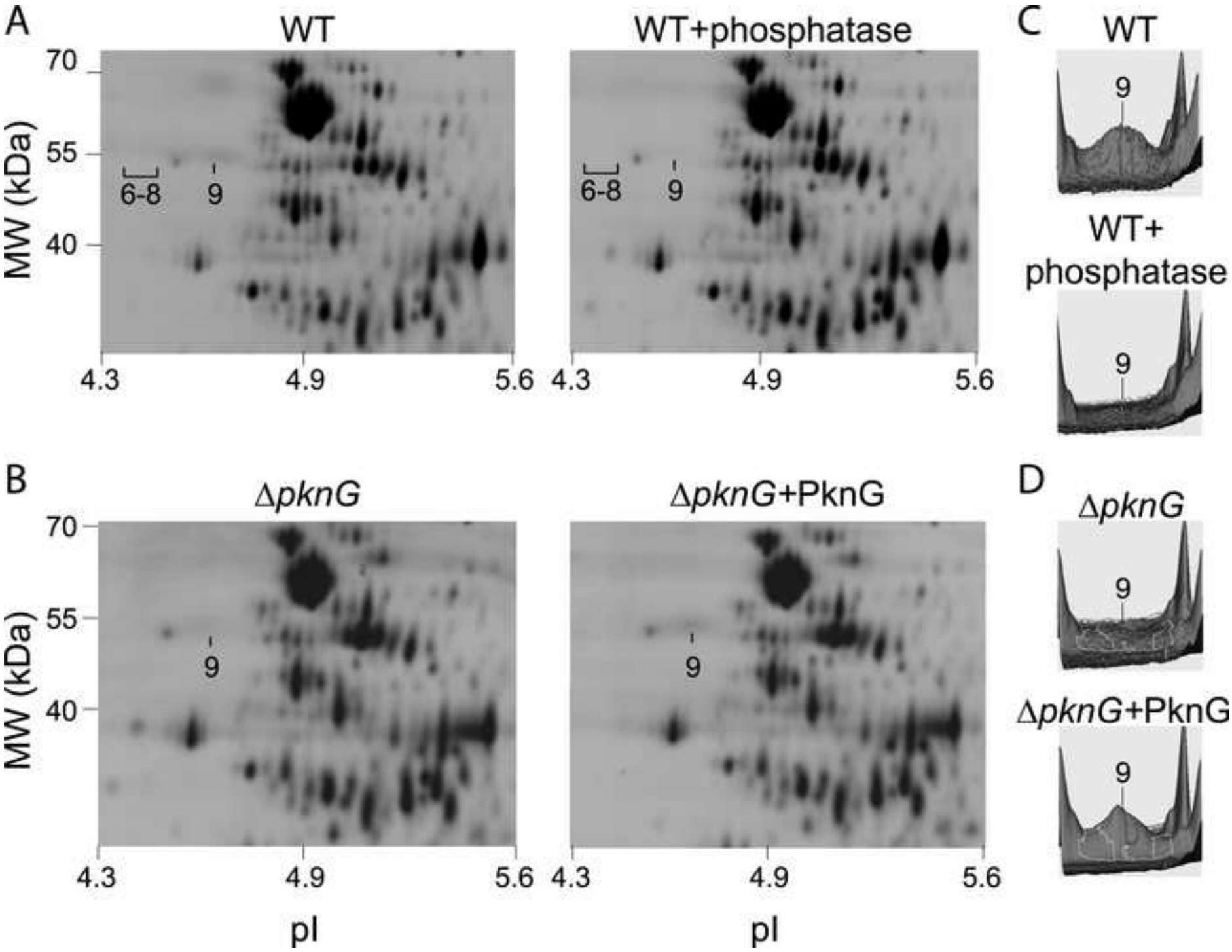

Figure 3

Click here to download high resolution image

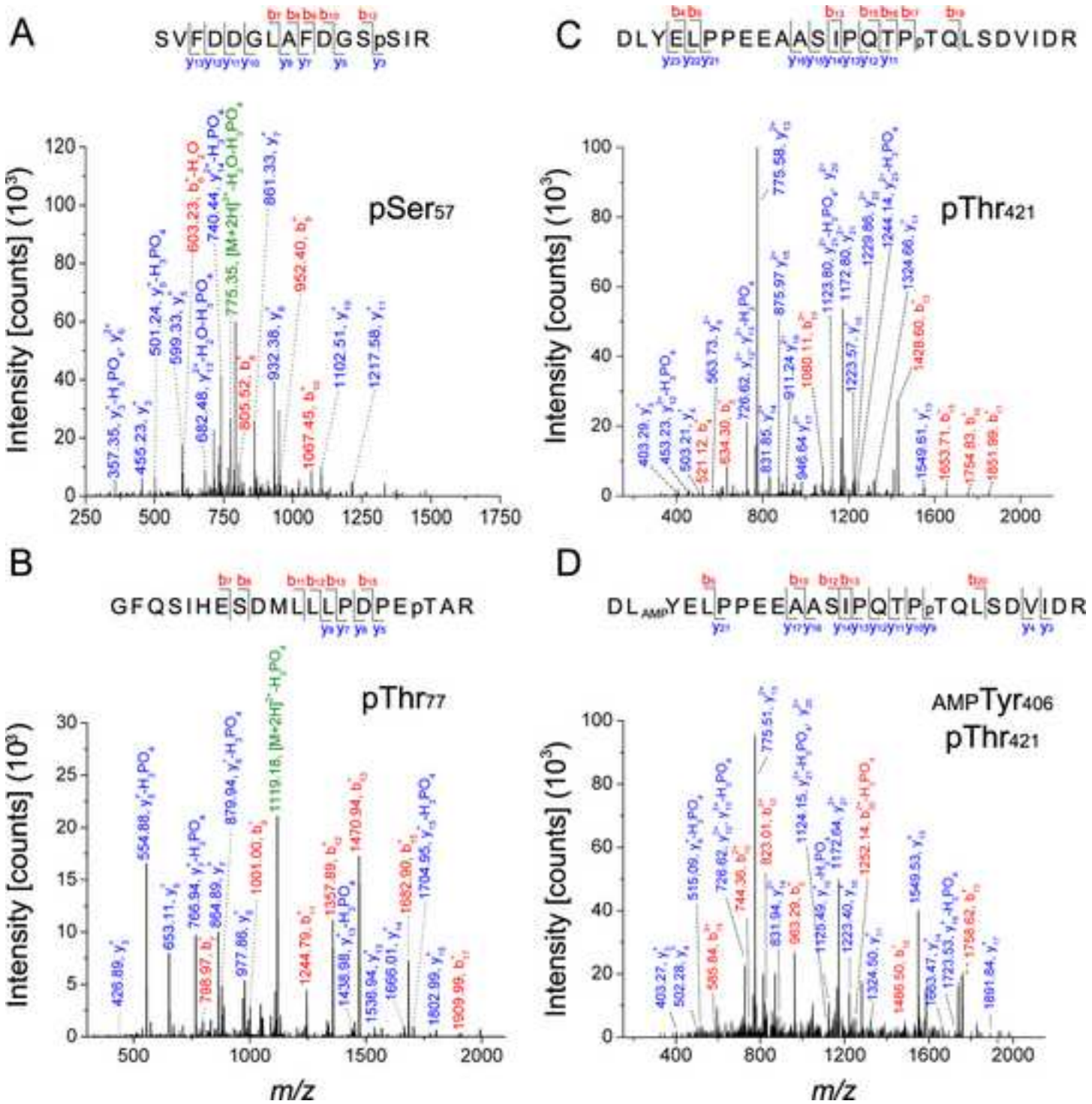

Figure 4

Click here to download high resolution image

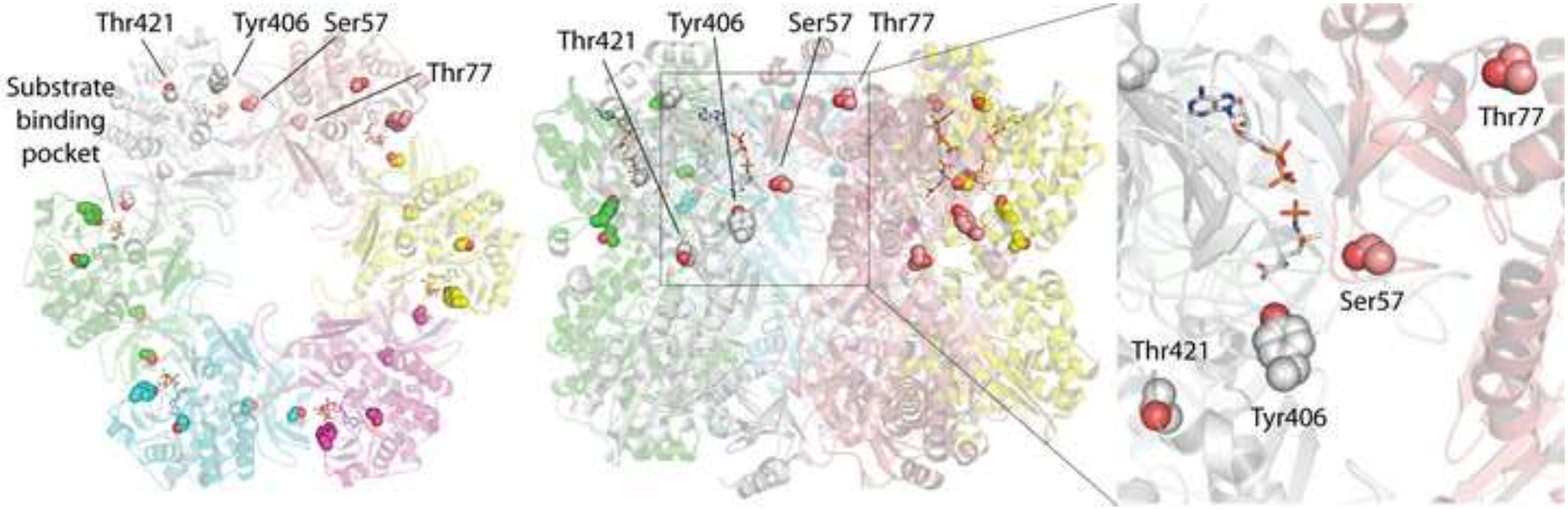

Figure 5

Click here to download high resolution image

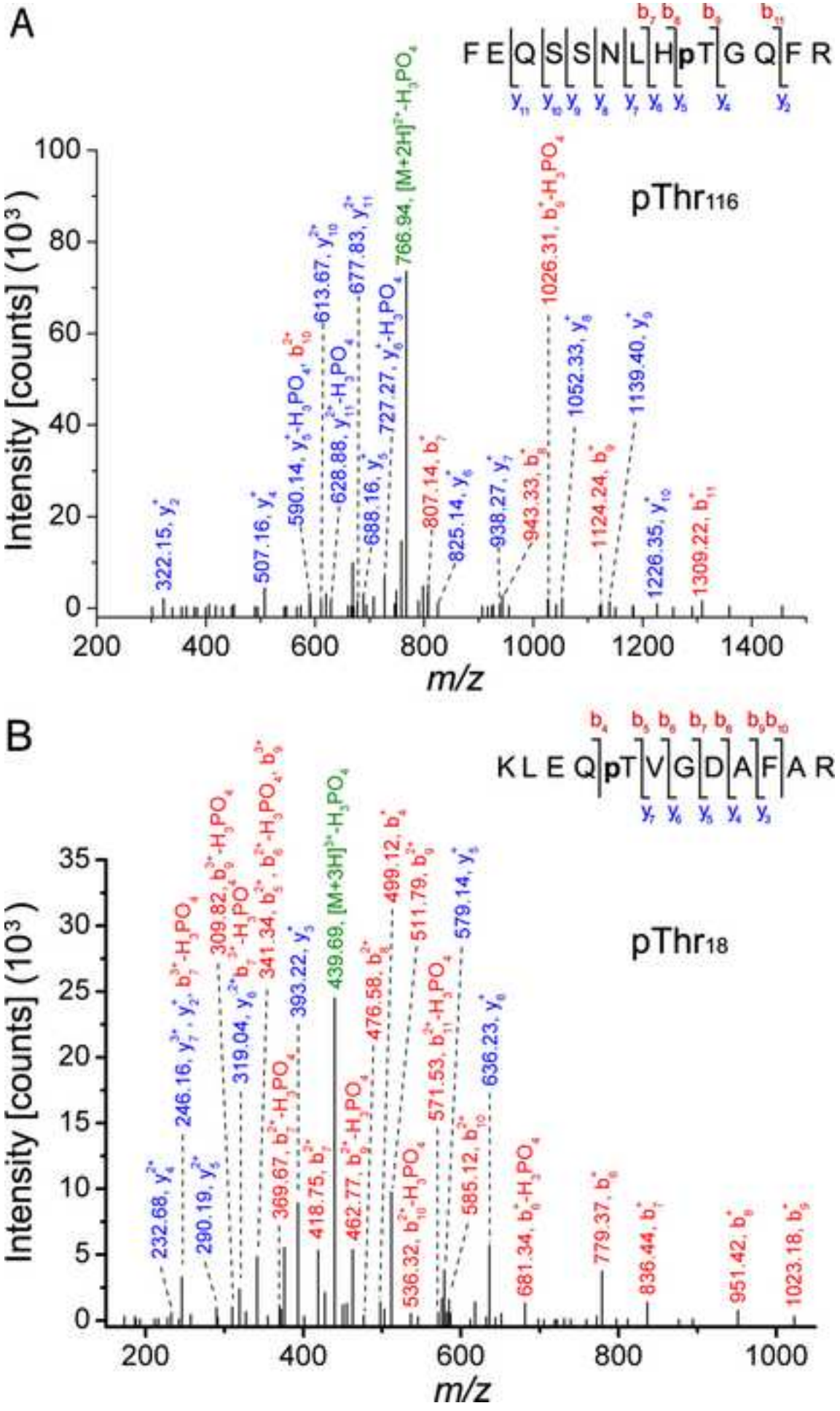

Figure 6

Click here to download high resolution image