doi:10.1016/j.jmb.2012.06.011 J. Mol. Biol. (2012) 422, 488–494

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

COMMUNICATION

Structure of the Sensor Domain of *Mycobacterium tuberculosis* PknH Receptor Kinase Reveals a Conserved Binding Cleft

Alexandra Cavazos, Daniil M. Prigozhin and Tom Alber*

Department of Molecular and Cell Biology, University of California, 374B Stanley Hall, Berkeley, CA 94720, USA

Received 25 April 2012; received in revised form 25 May 2012; accepted 6 June 2012
Available online 20 June 2012

Edited by I. Wilson

Keywords:
orphan receptor; kinase function

Since their discovery over 20 years ago, eukaryotic-like transmembrane receptor Ser/Thr protein kinases (STPKs) have been shown to play critical roles in the virulence, growth, persistence, and reactivation of many bacteria. Information regarding the signals transmitted by these proteins, however, remains scarce. To enhance understanding of the basis for STPK receptor signaling, we determined the 1.7-Å-resolution crystal structure of the extracellular sensor domain of the *Mycobacterium tuberculosis* receptor STPK, PknH (Rv1266c). The PknH sensor domain adopts an unanticipated fold containing two intramolecular disulfide bonds and a large hydrophobic and polar cleft. The residues lining the cleft and those surrounding the disulfide bonds are conserved. These results suggest that PknH binds a small-molecule ligand that signals by changing the location or quaternary structure of the kinase domain.

Introduction

Mycobacterium tuberculosis (*Mtb*) is a persistent human pathogen that currently infects one-third of the world's population and causes over 1.7 million deaths per year.[1] The pathogenicity of *Mtb* stems from its ability to alter its developmental programs and metabolism in different host niches. After phagocytosis by alveolar macrophages, *Mtb* slows growth and alters the composition of cell-wall mycolic and fatty acids to survive the nutrient-poor phagocytic environment and resist microbicides such as nitric oxide and reactive oxygen species.[2] However, little is known about the developmental programs and molecular signals that trigger these adaptive responses. Candidate sensor molecules for transmitting environmental signals into adaptive responses include the 11 eukaryotic-like Ser/Thr protein kinases (STPKs) encoded in the *Mtb* genome, 9 of which have an intracellular N-terminal kinase domain (KD) linked via a single transmembrane helix to an extracellular C-terminal sensor domain.[3] Recent sequencing projects indicate that eukaryotic-like STPKs exist in many prokaryotes, including a wide range of pathogenic bacteria.[4] Since their discovery, STPKs have been shown to regulate diverse cellular functions, such as exit from dormancy,[5,6] protein secretion,[7] cell division,[8] sporulation,[9,10] and cell-wall biosynthesis.[11]

The first bacterial STPK KD structures, which revealed nucleotide complexes of the *Mtb* PknB KD, demonstrated that bacterial and eukaryotic STPKs share close structural similarities and common modes of substrate recognition and regulation.[12,13] Despite advances in understanding the KDs of

*Corresponding author. E-mail address: tom@ucxray.berkeley.edu.

Abbreviations used: STPK, Ser/Thr protein kinase; *Mtb*, *Mycobacterium tuberculosis*; KD, kinase domain; ECD, extracellular domain; SAD, single-wavelength anomalous diffraction; GuHCl, guanidine hydrochloride; PDB, Protein Data Bank; PIM, phosphatidylmyoinositol mannoside.

STPKs, only two of the *Mtb* STPK sensor domains have been structurally characterized. The PknD sensor-domain structure was found to form a rigid, six-bladed β-propeller with a flexible linker to the transmembrane helix,[14] while the PknB sensor domain was found to have four PASTA domains[15] that bind peptidoglycan fragments and localize the kinase to sites of peptidoglycan turnover to regulate cell growth and division.[16] To further understand STPK receptor signaling, we determined the X-ray crystal structure of the extracellular sensor domain of the *Mtb* STPK PknH (Rv1266c).

Protein production and structure determination

To characterize the PknH sensor, we expressed the extracellular domain (ECD; residues 435–626) beginning eight residues after the predicted transmembrane helix. This N-terminally His_6-tagged protein was largely insoluble in *Escherichia coli*. The presence of four conserved Cys residues suggested that the structure may be stabilized by one or two disulfide bonds. To test this idea, we isolated inclusion bodies, denatured the protein in 6 M guanidine hydrochloride (GuHCl), and refolded it on a Ni-NTA column by stepwise dilution of the GuHCl in the presence of a 10:1 mixture of reduced-to-oxidized glutathione. This procedure yielded soluble ECD that was sensitized to proteolysis and precipitation when reduced, suggesting that the disulfides are essential for the stability of the fold. The oxidized protein crystallized after removal of the purification tag. The crystal structure was determined at 1.7-Å resolution by single-wavelength anomalous diffraction (SAD) analysis of a terbium derivative (Table 1). The entire sequence from residue 435 to 626 was ordered.

Table 1. X-ray data collection, analysis, and refinement statistics for the crystal structure of the PknH ECD

	PknH ECD Tb^{3+} complex
Data collection	
Wavelength (Å)	1.12
Temperature (K)	100
Space group	$P2_1$
Unit cell parameters	
a, *b*, *c* (Å)	47.46, 35.92, 49.31
β (°)	98.36
Resolution (Å) [a]	50.0–1.70 (1.76–1.70)
Number of unique reflections	34,200 (3328)
R_{sym} (%)	6.1 (25.3)
$I/\sigma I$	17.4 (4.9)
Completeness (%)	98.8 (98.1)
Redundancy	3.8 (3.8)
SAD solution	
Proteins per a.u.	1
Terbium sites per a.u.	2
Mean figure of merit	0.424
Refinement	
Resolution (Å)	48.78–1.70
Number of reflections	34,190
R_{work}/R_{free} (%)	16.30/19.74
Number of atoms	
Protein	1467
Solvent	221
Average *B*-factors	
Protein (Å^2)	17
Solvent (Å^2)	25
rmsd	
Bond lengths (Å)	0.013
Bond angles (°)	0.96
Ramachandran plot	
Favored (%)	96
Allowed (%)	4
PDB ID	4ESQ

Data were collected at 100 K at Beamline 8.3.1 at the Lawrence Berkeley National Laboratory Advanced Light Source.[39] Data were reduced and scaled with HKL2000.[40] The structure was determined using PHENIX[41] and the model was adjusted manually using Coot.[42] Phenix.autosol found two terbium sites per asymmetric unit (a.u.), and phenix.autobuild produced a model with 191 residues and an R_{free} of 24% after seven cycles of automatic building and refinement. Building and refinement were completed with phenix.refine and Coot and included addition of a single ordered molecule of 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol buffer that coordinated one of the two terbium ions. The final model was validated using MolProbity.[43]

PknH sensor-domain structure

The PknH sensor domain contains six α-helices and seven antiparallel β-strands with α1–α2–α3–α4–β1–β2–α5–β3–β4–β5–β6–β7–α6 topology (Fig. 1a). Two intramolecular disulfide bonds link α3 to α5 (C482–C545) and β6 to β7 (C587–C604). A 22-residue irregular loop connects α2 and α3. The most prominent feature is a large v-shaped central cleft (Fig. 1b). Five of the seven antiparallel β-strands (β1/β2 and β5–β7) make up one side of this cleft, while α-helices α3 to α5 and β-strands β3 and β4 comprise the other side. The α2–α3 loop forms the rim of the cleft, and residues 486–490 in the α3–α4 loop line the cleft inner wall (Fig. 1). The cleft has a calculated[17] surface area of 1134 Å^2 and a volume of 2768 Å^3.

A BLAST search reveals that PknH orthologs occur only in pathogenic mycobacteria. Homologous sensor domains in STPKs generally show >50% sequence identity. In addition, the PknH sensor domain shows 27–40% sequence identity to mycobacterial LppH proteins, which contain an N-terminal lipid attachment sequence, but no KD. More remote PknH ECD homologs in *Mtb* include LppR, Rv3705c, LpqA, LpqQ, LprH, and the PknJ sensor domain. Plotting the sequence conservation in PknH orthologs on the sensor-domain structure reveals that residues surrounding the disulfide bonds and lining the bottom of the cleft have a high degree of conservation (Fig. 2). In contrast, the residues forming the edges and surface-exposed sides of the cleft are less conserved. These results point to the recent emergence of this fold in mycobacteria and suggest that the sensor-domain homologs are adapted to bind different related ligands in different species.

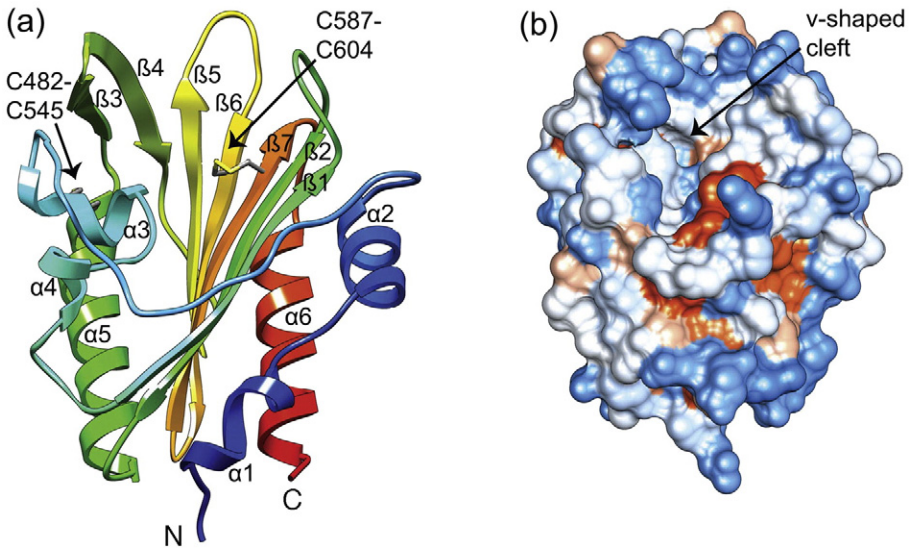

Fig. 1. (a) Crystal structure of the PknH extracellular sensor domain. The ribbon diagram was color coded from the N-terminus (blue) to the C-terminus (red). The v-shaped cleft is surrounded by the β-sheet, α3–α5, and the long α2–α3 loop. PknH residues 435–626 were PCR amplified from *Mtb* H37Rv genomic DNA and cloned into the pET28-based destination vector pHGWA using Gateway enzymes (Invitrogen). The pHGWA vector has an N-terminal His_6 tag followed by a tobacco etch virus (TEV) protease cleavage site[44] that leaves Gly–His at the N-terminus of the protein after TEV cleavage. The protein was expressed in BL21 (DE3) CodonPlus cells (Stratagene) using auto-induction.[45] Cells were grown at 37 °C for 6 h and then at 30 °C for 18 h. Cell paste was resuspended in Buffer A {200 mM NaCl, 5% v/v glycerol, and 20 mM 1,3-Bis[tris(hydroxymethyl)methylamino]propane at pH 8.5} plus protease inhibitors E-64 and leupeptin and lysed by sonication on ice. The lysate was centrifuged for 40 min at 16,000 rpm and the resulting supernatant was discarded. The pellet was resuspended in Buffer A plus 0.1% Triton X-100, and this mixture was sonicated and centrifuged at 9000 rpm for 20 min. The supernatant was discarded and the detergent extraction step was repeated twice. The pellet was resuspended at 4 °C overnight in Buffer A plus 6 M GuHCl, 1 mM reduced glutathione, and 0.1 mM oxidized glutathione. The GuHCl solution was sonicated and centrifuged at 17,000 rpm for 30 min. The supernatant was incubated with Ni-NTA agarose beads (Qiagen) for 4 h, and stepwise washes that reduced the GuHCl concentration in 1-M increments were used to refold the protein in the presence of the 10:1 ratio of reduced-to-oxidized glutathione. The protein was eluted with Buffer A plus 300 mM imidazole. After dialysis into Buffer A to remove the imidazole, the protein was incubated with TEV protease overnight at 4 °C. Cleaved PknH protein was purified by a second IMAC step and further purified by size-exclusion chromatography using a HiLoad 26/60 Superdex 75 column in Buffer B (100 mM NaCl, 5% v/v glycerol, and 20 mM Pipes, pH 6.5). A single peak corresponding to the PknH sensor-domain monomer was concentrated to 5 mg/ml and was crystallized at 4 °C by vapor diffusion in a 1:1.5 ratio of protein-to-initial crystallization buffer {0.1 M 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol (pH 5.5), 0.2 M ammonium acetate, and 25% polyethylene glycol 3350}. Additive screening of the initial crystal hit led to improved crystals with $Tb(NO_3)_3$, which were used for SAD phasing. Optimal crystals were grown overnight at 18 °C using 250 nL of the initial crystallization buffer, 500 nL of 5 mg/ml PknH in Buffer B, and 100 nL of 100 mM $Tb(NO_3)_3$. Crystals were cryoprotected with mother liquor containing 15% v/v glycerol. Images of the structure were created using Chimera v1.5.3.[46] (b). PknH sensor-domain surface was colored according to the Kyte and Doolittle hydrophobicity scale.[47] The most hydrophobic areas are in orange and the most polar areas are in blue. The cleft (arrow) has a surface area of 1134 $Å^2$ and a volume of 2768 $Å^3$ calculated using CASTp.[17]

Structural comparisons

No structures of proteins with similar sequences have been reported, making the structure of PknH ECD the first for its Pfam[18] family (PF 14032). A search for similar structures using the DALI server[19] found DIP2269, a hypothetical protein from *Corynebacterium diphtheriae* [Protein Data Bank (PDB) ID: 3V7B, Z = 9.6, rmsd = 3.6 Å over 118 aligned residues], TM1622, a GTP-binding regulator from *Thermotoga maritima* (PDB ID: 1VR8, Z = 8.8, rmsd = 3.5 Å over 120 aligned residues), and BT1490, an uncharacterized protein from *Bacteroides thetaiotaomicron* (PDB ID: 3HLZ, Z = 8.8, rmsd = 3.2 Å over 119 aligned residues). Structural alignment of PknH ECD with these three proteins using POSA[20] revealed a shared core of two α-helices and six antiparallel β-strands (β1–β2–α5–β4 extension–β5–β6–β7–α6 in PknH, Fig. 2c) that comprises 60% of the structure. The longer, unique N-terminal segment and an extra β-hairpin (β3–β4) between the core elements α5 and β5 in PknH provide evidence for a new fold. Nonetheless, the presence of the common core also is consistent with the idea that the

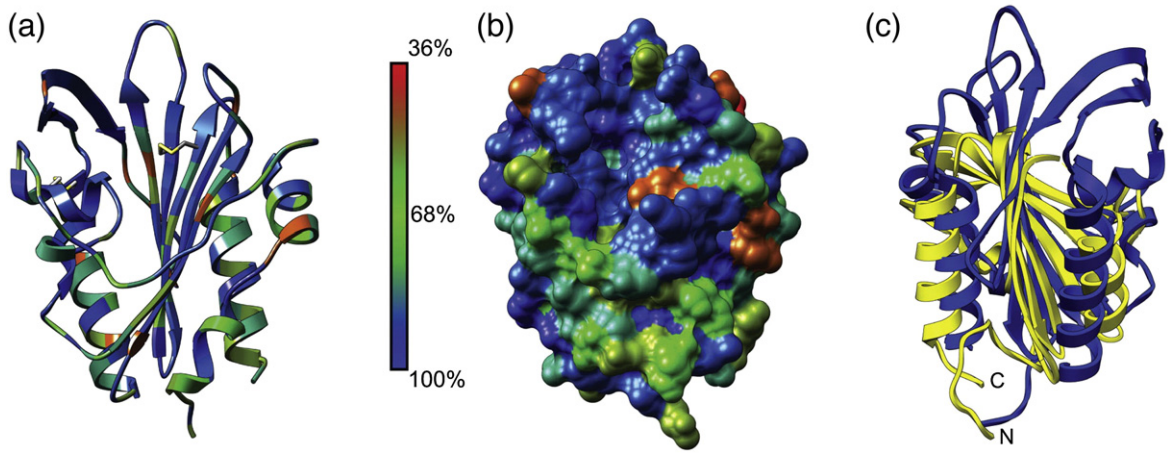

Fig. 2. Sequence conservation among PknH sensor-domain orthologs identifies the disulfide bonds and the cleft as critical functional elements while structural alignment suggests similarities to the Mog1p fold. (a) Ribbon representation of the PknH sensor domain with residues colored according to conservation among 11 unique orthologous sequences from different mycobacterial species. The orthologs are as follows: *M. tuberculosis* H37Rv (GI: 15608406), *M. marinum* DL240490 (GI: 169245959), *M. tuberculosis C* (GI: 254231523), *M. bovis* BCG Tokyo 172 (GI: 224989669), *M. bovis* AF2122/97 (GI: 31792458), *M. tuberculosis* K85 (GI: 289573926), *M. marinum* M (GI: 183984142), *M. parascrofulaceum* ATCC BAA-614 (GI: 296170175), *M. avium* 104 (GI: 118465657), *M. colombiense* CECT 3035 (GI: 342862145), and *M. intracellulare* ATCC 13950 (GI: 254820140). The alignment for the first nine sequences was downloaded from PhyloFacts[48] family bpg0185066 (redundant sequences removed), and the two additional sequences were identified using BLAST and added to the alignment using Chimera Multalign Viewer. (b) Surface representation of residue conservation. Conservation was mapped onto the PknH sensor domain and imaged using version 1.5.3 of Chimera.[46] (c) Ribbon representation of the PknH ECD (blue) and DIP2269 (yellow), a Mog1p-fold protein from *C. diphtheriae* (PDB 3V7B). The PknH ECD shares a central antiparallel β-sheet and two α-helices with DIP2269, but PknH has an extra β3–β4 hairpin and a significantly different N-terminus.

PknH ECD may be a distant variant of the SCOP Mog1p fold.[21] In any case, the modest *Z*-scores, partial alignments, distinct secondary structures, high C^{α} rmsds, and uncharacterized activities led us to conclude that direct structural comparisons between the PknH sensor domain and the DALI search results do not provide insights into the possible functions of this STPK.

Despite the lack of close structural homologs, visual comparison between the PknH sensor domain and the glycolipid-binding *Mtb* lipoproteins LprG (Rv1411c)[22,23] and LppX (Rv2945c)[24] shows that all three have a large central cleft bordered on one side by a β-sheet and on the other by two to three α-helices (Fig. 3). The cleft of PknH has a surface area of 1134 $Å^2$ and a volume of 2768 $Å^3$, compared to the clefts for LprG (1549 $Å^2$ and 2679 $Å^3$) and LppX (1375 $Å^2$ and 2835 $Å^3$).[17]

LprG is a widely distributed and conserved *Mtb* lipoprotein with TLR2 agonist activity that has been shown to bind the cell-wall precursor molecule Ac_1PIM_2.[22] Phosphatidylmyoinositol mannosides (PIMs) are glycolipids found in the inner and outer membrane of the cell envelopes of all mycobacteria. They consist of a phosphatidylmyoinositol lipid anchor that carries one to six mannose residues with up to four acyl chains.[25] In addition to being critical structural components of the cell envelope, PIMs such as Ac_1PIM_2 are precursors for lipomannan and lipoarabinomannan, which modulate host–pathogen interactions over the course of a tuberculosis infection.[26] LppX is another conserved *Mtb* lipoprotein that shares 31% sequence identity with LprG and is predicted to bind to phthiocerol dimycocerosates and transport them to the outer layer of the mycobacterial cell envelope.[24]

PknH function

The PknH STPK sensor domain adopts a distinctive fold with two intramolecular disulfide bonds and a large v-shaped cleft. Possible functions for the PknH sensor domain can be inferred from its genomic location and reported substrates. The *pknH* gene (Rv1266c) is adjacent to the gene for the EmbR transcriptional regulator (Rv1267) on the *Mtb* chromosome. *In vitro* phosphorylation by PknH enhances EmbR binding to the promoter of the *embCAB* arabinosyltransferase genes and leads to increased transcription of these enzymes.[27] EmbA and EmbB are glycosyltransferases that create the terminal hexaarabinoside motif in *Mtb* cell-wall arabinogalactan,[28] while EmbC synthesizes the arabinan portion of lipoarabinomannan.[29]

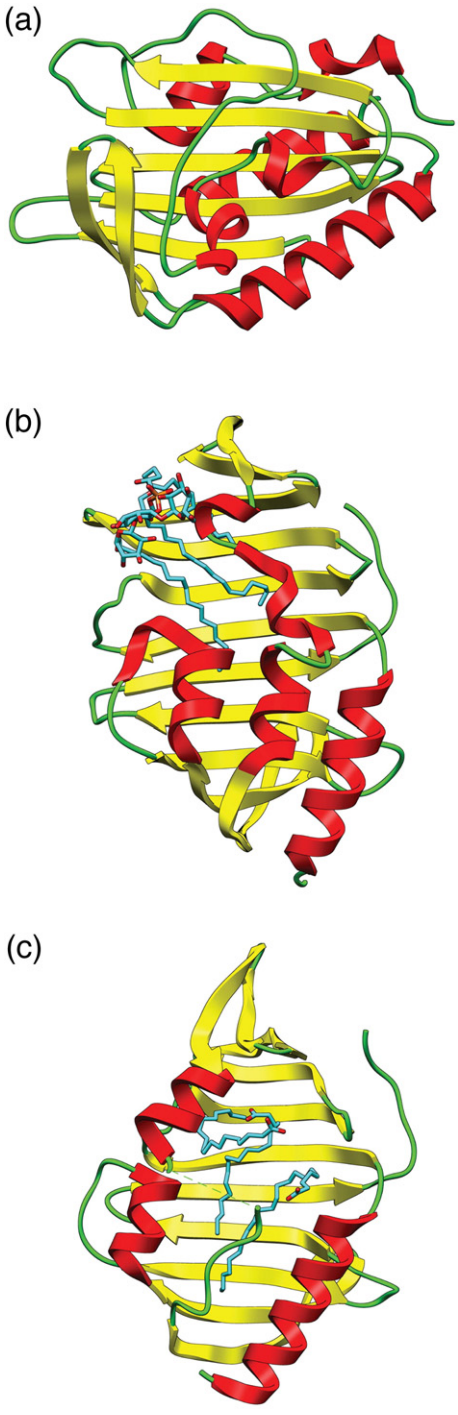

Fig. 3. Structures qualitatively resembling the PknH sensor domain. Ribbon representation of (a) the PknH sensor domain, (b) *Mtb* lipoprotein LprG bound to the glycolipid cell-wall precursor molecule AC_1PIM_2 (PDB 3MHA), and (c) *Mtb* lipoprotein LppX bound to two molecules of *cis*-vaccenic acid (C18:1) and one docosanoic acid (C22:0) molecule (PDB 2BYO).

Deletion of PknH in *Mtb* results in a hypervirulent phenotype in mice and decreased transcription of *embB* and *embC* in cultures treated with sublethal concentrations of ethambutol.[30] By phosphorylating EmbR, PknH may control the ratio of lipoarabinomannan to lipomannan, a critical determinant of *Mtb* virulence.

In addition, PknH and several other *Mtb* STPKs have been shown to phosphorylate KasA, KasB, and MtFabH. KasA and KasB are β-ketoacyl-ACP synthases that elongate mycolic acid precursors.[31] Mycolic acids are long-chain (C60–C90) α-alkyl-β-hydroxy fatty acids that promote bacterial resistance to antibiotics and environmental stress[32] and thus increase *Mtb* virulence[33] and persistence.[32] Phosphorylation-induced inhibition of KasA activity presumably leads to immature mycolic acids while phosphorylation-induced stimulation of KasB activity is thought to ensure production of the full-length mycolates required for bacterial survival and virulence.[34,35] MtFabH is the β-ketoacyl-ACP synthase III enzyme that catalyzes the condensation of FAS-I-derived acyl-CoAs with malonyl-AcpM, thus linking the FAS-I and FAS-II systems in *Mtb*.[36] MtFabH is phosphorylated *in vitro* by PknH, PknA, and PknF.[37] If PknH functions as a feedback regulator, the sensor domain may be responsive to signals generated in the complex *Mtb* cell wall. Alternatively, PknH may regulate cell-wall production in response to environmental cues, including compounds that are unrelated to the mycobacterial cell wall.

The novel structure of the sensor domain affords few clues about the signaling ligand(s). The putative recognition cleft is narrow, deep, and conserved, as expected for small-molecule binding sites in proteins.[38] Visual comparison between the cleft in PknH and the glycolipid-binding clefts of the *Mtb* lipoproteins LprG and LppX indicates that the PknH binding site is less hydrophobic, making it unlikely that cell-wall glycolipids are the signals. The mixed hydrophobic and polar character of the PknH cleft is consistent with a more polar signal. The stabilizing disulfides impart rigidity to the fold, and the N-terminus, which connects the domain to the TM helix, extends away from the structure. These general characteristics of structural stiffness and loose tethering to the TM helix are shared by the PknD and PknB sensor domains, implying a common signaling mechanism in which ligands may modulate the localization or oligomerization of the kinase.

Coordinates

The coordinates and structure factors were deposited in the PDB under the accession number 4ESQ.

Acknowledgements

We thank James Holton, George Meigs, and Jane Tanamachi at Beamline 8.3.1 at Lawrence Berkeley National Laboratory for crystallographic assistance. This work was supported by National Institutes of Health grants R01 GM70962 and P01 AI068135 to T.A.

References

1. World Health Organization. (2011). Global tuberculosis control report 2011. (WHO/HTM/TB/2011.16). Available: http://www.who.int/tb/publications/global_report/2011/gtbr11_full.pdf. Accessed 2012 April 4.
2. Schnappinger, D., Ehrt, S., Voskuil, M. I., Liu, Y., Mangan, J. A., Monahan, I. M. *et al.* (2003). Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: insights into the phagosomal environment. *J. Exp. Med.* **198**, 693–704.
3. Av-Gay, Y. & Everett, M. (2000). The eukaryotic-like Ser/Thr protein kinases of *Mycobacterium tuberculosis*. *Trends Microbiol.* **8**, 238–244.
4. Galperin, M. Y., Higdon, R. & Kolker, E. (2010). Interplay of heritage and habitat in the distribution of bacterial signal transduction systems. *Mol. Biosyst.* **6**, 721–728.
5. Shah, I. M., Laaberki, M. H., Popham, D. L. & Dworkin, J. (2008). A eukaryotic-like Ser/Thr kinase signals bacteria to exit dormancy in response to peptidoglycan fragments. *Cell*, **135**, 486–496.
6. Shah, I. M. & Dworkin, J. (2010). Induction and regulation of a secreted peptidoglycan hydrolase by a membrane Ser/Thr kinase that detects muropeptides. *Mol. Microbiol.* **75**, 1232–1243.
7. Mougous, J. D., Gifford, C. A., Ramsdell, T. L. & Mekalanos, J. J. (2007). Threonine phosphorylation post-translationally regulates protein secretion in *Pseudomonas aeruginosa*. *Nat. Cell Biol.* **9**, 797–803.
8. Fiuza, M., Canova, M. J., Zanella-Cleon, I., Becchi, M., Cozzone, A. J., Mateos, L. M. *et al.* (2008). From the characterization of the four serine/threonine protein kinases (PknA/B/G/L) of *Corynebacterium glutamicum* toward the role of PknA and PknB in cell division. *J. Biol. Chem.* **283**, 18099–18112.
9. Madec, E., Laszkiewicz, A., Iwanicki, A., Obuchowski, M. & Seror, S. (2002). Characterization of a membrane-linked Ser/Thr protein kinase in *Bacillus subtilis*, implicated in developmental processes. *Mol. Microbiol.* **46**, 571–586.
10. Madec, E., Stensballe, A., Kjellstrom, S., Cladiere, L., Obuchowski, M., Jensen, O. N. & Seror, S. J. (2003). Mass spectrometry and site-directed mutagenesis identify several autophosphorylated residues required for the activity of PrkC, a Ser/Thr kinase from *Bacillus subtilis*. *J. Mol. Biol.* **330**, 459–472.
11. Fiuza, M., Canova, M. J., Patin, D., Letek, M., Zanella-Cleon, I., Becchi, M. *et al.* (2008). The MurC ligase essential for peptidoglycan biosynthesis is regulated by the serine/threonine protein kinase PknA in *Corynebacterium glutamicum*. *J. Biol. Chem.* **283**, 36553–36563.
12. Young, T. A., Delagoutte, B., Endrizzi, J. A., Falick, A. M. & Alber, T. (2003). Structure of *Mycobacterium tuberculosis* PknB supports a universal activation mechanism for Ser/Thr protein kinases. *Nat. Struct. Biol.* **10**, 168–174.
13. Ortiz-Lombardia, M., Pompeo, F., Boitel, B. & Alzari, P. M. (2003). Crystal structure of the catalytic domain of the PknB serine/threonine kinase from *Mycobacterium tuberculosis*. *J. Biol. Chem.* **278**, 13094–13100.
14. Good, M. C., Greenstein, A. E., Young, T. A., Ng, H. L. & Alber, T. (2004). Sensor domain of the *Mycobacterium tuberculosis* receptor Ser/Thr protein kinase, PknD, forms a highly symmetric beta propeller. *J. Mol. Biol.* **339**, 459–469.
15. Barthe, P., Mukamolova, G. V., Roumestand, C. & Cohen-Gonsaud, M. (2010). The structure of PknB extracellular PASTA domain from *Mycobacterium tuberculosis* suggests a ligand-dependent kinase activation. *Structure*, **18**, 606–615.
16. Mir, M., Asong, J., Li, X., Cardot, J., Boons, G. J. & Husson, R. N. (2011). The extracytoplasmic domain of the *Mycobacterium tuberculosis* Ser/Thr kinase PknB binds specific muropeptides and is required for PknB localization. *PLoS Pathog.* **7**, e1002182.
17. Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y. & Liang, J. (2006). CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* **34**, W116–W118.
18. Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C. *et al.* (2012). The Pfam protein families database. *Nucleic Acids Res.* **40**, D290–D301.
19. Holm, L. & Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549.
20. Ye, Y. & Godzik, A. (2005). Multiple flexible structure alignment using partial order graphs. *Bioinformatics*, **21**, 2362–2369.
21. Xu, Q., Krishna, S. S., McMullan, D., Schwarzenbacher, R., Miller, M. D., Abdubek, P. *et al.* (2006). Crystal structure of an ORFan protein (TM1622) from *Thermotoga maritima* at 1.75 Å resolution reveals a fold similar to the Ran-binding protein Mog1p. *Proteins*, **65**, 777–782.
22. Drage, M. G., Tsai, H. C., Pecora, N. D., Cheng, T. Y., Arida, A. R., Shukla, S. *et al.* (2010). *Mycobacterium tuberculosis* lipoprotein LprG (Rv1411c) binds triacylated glycolipid agonists of Toll-like receptor 2. *Nat. Struct. Mol. Biol.* **17**, 1088–1095.
23. Sieling, P. A., Hill, P. J., Dobos, K. M., Brookman, K., Kuhlman, A. M., Fabri, M. *et al.* (2008). Conserved mycobacterial lipoglycoproteins activate TLR2 but also require glycosylation for MHC class II-restricted T cell activation. *J. Immunol.* **180**, 5833–5842.
24. Sulzenbacher, G., Canaan, S., Bordat, Y., Neyrolles, O., Stadthagen, G., Roig-Zamboni, V. *et al.* (2006). LppX is a lipoprotein required for the translocation of phthiocerol dimycocerosates to the surface of *Mycobacterium tuberculosis*. *EMBO J.* **25**, 1436–1444.
25. Guerin, M. E., Kordulakova, J., Alzari, P. M., Brennan, P. J. & Jackson, M. (2010). Molecular basis of phosphatidyl-myo-inositol mannoside biosynthesis and regulation in mycobacteria. *J. Biol. Chem.* **285**, 33577–33583.

26. Mishra, A. K., Driessen, N. N., Appelmelk, B. J. & Besra, G. S. (2011). Lipoarabinomannan and related glycoconjugates: structure, biogenesis and role in *Mycobacterium tuberculosis* physiology and host–pathogen interaction. *FEMS Microbiol. Rev.* **35**, 1126–1157.
27. Sharma, K., Gupta, M., Pathak, M., Gupta, N., Koul, A., Sarangi, S. *et al.* (2006). Transcriptional control of the mycobacterial embCAB operon by *PknH* through a regulatory protein, EmbR, in vivo. *J. Bacteriol.* **188**, 2936–2944.
28. Escuyer, V. E., Lety, M. A., Torrelles, J. B., Khoo, K. H., Tang, J. B., Rithner, C. D. *et al.* (2001). The role of the embA and embB gene products in the biosynthesis of the terminal hexaarabinofuranosyl motif of *Mycobacterium smegmatis* arabinogalactan. *J. Biol. Chem.* **276**, 48854–48862.
29. Zhang, N., Torrelles, J. B., McNeil, M. R., Escuyer, V. E., Khoo, K. H., Brennan, P. J. & Chatterjee, D. (2003). The Emb proteins of mycobacteria direct arabinosylation of lipoarabinomannan and arabinogalactan via an N-terminal recognition region and a C-terminal synthetic region. *Mol. Microbiol.* **50**, 69–76.
30. Papavinasasundaram, K. G., Chan, B., Chung, J. H., Colston, M. J., Davis, E. O. & Av-Gay, Y. (2005). Deletion of the *Mycobacterium tuberculosis* pknH gene confers a higher bacillary load during the chronic phase of infection in BALB/c mice. *J. Bacteriol.* **187**, 5751–5760.
31. Molle, V., Brown, A. K., Besra, G. S., Cozzone, A. J. & Kremer, L. (2006). The condensing activities of the *Mycobacterium tuberculosis* type II fatty acid synthase are differentially regulated by phosphorylation. *J. Biol. Chem.* **281**, 30094–30103.
32. Daffe, M. & Draper, P. (1998). The envelope layers of mycobacteria with reference to their pathogenicity. *Adv. Microb. Physiol.* **39**, 131–203.
33. Dubnau, E., Chan, J., Raynaud, C., Mohan, V. P., Laneelle, M. A., Yu, K. *et al.* (2000). Oxygenated mycolic acids are necessary for virulence of *Mycobacterium tuberculosis* in mice. *Mol. Microbiol.* **36**, 630–637.
34. Bhatt, A., Fujiwara, N., Bhatt, K., Gurcha, S. S., Kremer, L., Chen, B. *et al.* (2007). Deletion of kasB in *Mycobacterium tuberculosis* causes loss of acid-fastness and subclinical latent tuberculosis in immunocompetent mice. *Proc. Natl Acad. Sci. USA*, **104**, 5157–5162.
35. Bhatt, A., Molle, V., Besra, G. S., Jacobs, W. R., Jr & Kremer, L. (2007). The *Mycobacterium tuberculosis* FAS-II condensing enzymes: their role in mycolic acid biosynthesis, acid-fastness, pathogenesis and in future drug development. *Mol. Microbiol.* **64**, 1442–1454.
36. Brown, A. K., Sridharan, S., Kremer, L., Lindenberg, S., Dover, L. G., Sacchettini, J. C. & Besra, G. S. (2005). Probing the mechanism of the *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III mtFabH: factors influencing catalysis and substrate specificity. *J. Biol. Chem.* **280**, 32539–32547.
37. Veyron-Churlet, R., Molle, V., Taylor, R. C., Brown, A. K., Besra, G. S., Zanella-Cleon, I. *et al.* (2009). The *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III activity is inhibited by phosphorylation on a single threonine residue. *J. Biol. Chem.* **284**, 6414–6424.
38. Liang, J., Edelsbrunner, H. & Woodward, C. (1998). Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein Sci.* **7**, 1884–1897.
39. MacDowell, A. A., Celestre, R. S., Howells, M., McKinney, W., Krupnick, J., Cambie, D. *et al.* (2004). Suite of three protein crystallography beamlines with single superconducting bend magnet as the source. *J. Synchrotron Radiat.* **11**, 447–455.
40. Otwinowski, Z. & Minor, W. (1997). Processing of x-ray diffraction data collected in oscillation mode. Methods in Enzymology. In *Macromolecular Crystallography, Part A* (Carter, C. W. & Sweet, R. M., eds), Volume 276, pp. 307–326. Academic Press, New York.
41. Adams, P. D., Alfonin, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N. *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 213–221.
42. Emsley, P. & Cowtan, K. (2004). Coot. model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
43. Chen, V. B., Arendall, W. B., III, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J. *et al.* (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 12–21.
44. Busso, D., Delagoutte-Busso, B. & Moras, D. (2005). Construction of a set Gateway-based destination vectors for high-throughput cloning and expression screening in *Escherichia coli*. *Anal. Biochem.* **343**, 313–321.
45. Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234.
46. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
47. Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132.
48. Datta, R. S., Meacham, C., Samad, B., Neyer, C. & Sjolander, K. (2009). Berkeley PHOG: PhyloFacts orthology group prediction web server. *Nucleic Acids Res.* **37**, W84–W89.