The crystal structure of PknI from *Mycobacterium tuberculosis* shows an inactive, pseudokinase-like conformation

María-Natalia Lisa[1,2,3,*,†], Tristan Wagner[1,2,3,*,‡], Matthieu Alexandre[1,2,3], Nathalie Barilone[1,2,3,¶], Bertrand Raynal[2,4], Pedro M. Alzari[1,2,3] and Marco Bellinzoni[1,2,3]

1. Institut Pasteur, Unité de Microbiologie Structurale, 25 rue du Docteur Roux, 75724 Paris Cedex 15, France;

2. CNRS UMR 3528 'Biologie structurale des processus cellulaires et maladies infectieuses', Institut Pasteur, 75724 Paris Cedex 15, France;

3. Université Paris Diderot, Sorbonne Paris Cité, Microbiologie structurale, 75724 Paris Cedex 15, France;

4. Plateforme de Biophysique Moléculaire, Institut Pasteur, 75724 Paris Cedex 15, France.

Article type : Original Article

* Contributed equally to the work.

Present addresses: † Institut Pasteur de Montevideo, Laboratory of Molecular & Structural Microbiology, Mataojo 2020, Montevideo, 11400, Uruguay; ‡ Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Straße 10, 35043 Marburg, Germany; ¶ Institut Pasteur, Unité Récepteurs-Canaux, 75724 Paris Cedex 15, France.

Correspondence: Marco Bellinzoni, Institut Pasteur, Unité de Microbiologie Structurale, 25 rue du Dr. Roux, 75724 Paris Cedex 15, France. Email: marco.bellinzoni@pasteur.fr. Fax: +33-1-45688604

Running title: Structure of *M. tuberculosis* PknI kinase domain

Abbreviations: ePKs: eukaryotic-like Ser/Thr protein kinases; *Mtb: Mycobacterium tuberculosis*; MBP: myelin basic protein

Databases: Atomic coordinates and structure factors for the catalytic domain of *M. tuberculosis* PknI are in the Protein Data Bank under the accession codes **5M06** (wild-type PknI + ADP), **5M07** (PknI_C20A), **5M08** (PknI_C20A_R136A) and **5M09** (PknI_C20A_R136N).

Keywords: X-ray crystallography, Ser/Thr kinase, activation segment, signal transduction.

Abstract

Eukaryotic-like Ser/Thr protein kinases (ePKs) have been identified in many bacterial species, where they are known to mediate signalling mechanisms that share several features with their eukaryotic counterparts. In *Mycobacterium tuberculosis*, PknI is one of the eleven predicted ePKs and it has been related to bacterial virulence. In order to better understand the molecular basis of its role in mycobacterial signalling, we solved the crystal structure of the PknI cytoplasmic domain. We found that even though PknI possesses most conserved elements characteristic of Hanks-type kinases, it is degraded in several motifs that are essential for the ePKs catalytic activity. Most notably, PknI presents a remarkably short activation segment lacking a peptide-substrate binding site. Consistently with this observation and similarly to earlier findings for eukaryotic pseudokinases, no kinase activity was detected for the catalytic domain of PknI, against different

substrates and in various experimental conditions. Based on these results we conclude that PknI may rely on unconventional mechanism(s) for kinase activity and/or it could play alternative role(s) in mycobacterial signalling.

Introduction

The study of signalling elements in *Mycobacterium tuberculosis* (*Mtb*) has been a very active working area during the last fifteen years, since the predicted presence, from the seminal publication of the complete sequence of *M. tuberculosis* H37Rv genome in 1998 [1], of eleven 'eukaryotic-like' Ser/Thr protein kinases (ePKs) in this pathogen. Despite huge efforts to identify their substrates and other interacting partners, leading to a large number of reported enzyme/substrate couples and to exploit some of these kinases as pharmacological targets, still little is known about the signalling pathways concerned, although an increasing bunch of evidence suggests the involvement of kinases in the control of cell division and infection (recently reviewed in [2,3]). However, only two out of eleven ePKs coded by *Mtb*, *i.e.* PknA and PknB, have been proven to be essential for the bacterial survival [4-7], thus raising more general questions about the physiological roles of the ensemble of ePKs, especially in the course of infection, about interactions with other players in signalling like the two-component systems and their possible functional redundancy. Initially among the less studied mycobacterial ePKs, partly due to the non-essentiality of the corresponding gene *in vitro* [8], PknI (Rv2910c) has recently gained interest with the publication of an homology model [9], the identification of two possible interaction partners [10] and the description of the phenotype of a double *pknI*$^-$/*dacB2*$^-$ knockout in *Mtb* [11], confirming initial reports suggesting a role of PknI in infection [8]. In order to structurally characterize and compare PknI with other known ePKs, and eventually investigate the molecular basis of its role in the course of *Mtb* infection, we solved the crystal structure of the PknI catalytic

domain. Here we describe how the structure of PknI, albeit showing most of the conserved elements characteristic of Hanks-type kinases, is not descriptive of an active ePK.

Results

Crystal structure of PknI kinase domain

For a structural characterization, recombinant PknI_1-256 (hereafter PknI) expressed and purified from *Escherichia coli* was subjected to crystallization trials and diffraction-quality crystals grew in the presence of ADP. The final atomic model obtained by X-ray crystallography was refined at 2.0 Å resolution (Table 1) and contains two copies of PknI within the asymmetric unit, encompassing residues 1-254 (plus an additional N-terminal histidine residue from the His_6-tag) and 2-255, respectively (Fig. 1A and 1B). Additionally, *mFo–DFc* sigma-A-weighted electron density maps disclosed intense peaks in the substrate binding cleft of PknI monomer B, where it was possible to refine one ADP molecule and a Ca(II) cation, both present in the crystallization solution (Fig. 1C). The representation of the electrostatic surface potential shows a predominant distribution of negative charges at the crevice between the two lobes, suggesting a role for the divalent ion as a charge neutralizer to allow nucleoside phosphate binding [12] (Fig. 1D).

As shown in Fig. 1A, PknI presents the fold typically found in the family of ePKs. It comprises two lobes, *i.e.* the smaller N-lobe, that includes anti-parallel β-strands and the prominent helix αC (Fig. 1C), and the bigger C-lobe that is mostly α-helical. The two protein chains present in the asymmetric unit are very similar, with an RMSD value of 0.24 Å among 244 Cα, irrespective of the presence/absence of an ADP molecule within the active site. Unexpectedly, the PknI monomers in the asymmetric unit are linked by a disulphide bridge involving the same residue

Cys20 from each monomer (Fig. 1B), part of the conserved 'Gly-rich loop' (or P-loop) on the tip of the β1-β2 turn (Fig. 1B and 1E).

To determine whether the observed interaction between PknI monomers might be relevant in the physiological scenario or if it could constitute an artefact induced by the crystal lattice, we first used analytical ultracentrifugation to investigate the protein species in solution. The distribution of sedimentation coefficients clearly revealed two protein populations, consistent with a PknI monomer ($s^0_{20,w}$ = 2.8 S) and a PknI dimer ($s^0_{20,w}$ = 3.9 S) present in solution. The theoretical sedimentation coefficients calculated for one PknI monomer and the PknI dimer as found in the crystallographic model were 2.8 S and 4.4 S, respectively, showing agreement with the experimental data for the monomeric species only. We therefore decided to replace residue Cys20 by either serine or alanine, as found in the protein orthologs from other *Mycobacterium* species (Fig. 1E), and to solve the crystal structure of PknI_C20S and PknI_C20A. While it was not possible to crystallize PknI_C20S, PknI_C20A crystals, non-isomorphous with respect to crystals of wild-type PknI, readily grew in the presence of high concentrations of NaCl. Crystals of apo PknI_C20A also contained two protein chains in the asymmetric unit, but packing interactions were different from those found in crystals of wild-type PknI and no covalent bonds were observed between monomers. Even though the crystallization assay did not include an added nucleotide, a residual, uninterpretable positive electron density in the mFo-DFc sigmaA-weighted map occupied an area corresponding to the nucleotide pocket of monomer A (Fig. 2), suggesting the presence of either an unidentified ligand or, alternatively, the N-terminal portion of the peptide chain from a neighbouring molecule. Nevertheless, the low RMSD values obtained after superimposing the two protein chains (0.76 Å among 246 Cα) as well as when superimposing wild-type PknI and PknI_C20A chains (RMSD of 0.55 Å over 244 Cα), indicate that residue Cys20 and the disulphide bridge leading to dimerization have no significant impact on the overall conformation of the kinase domain (Fig. 3A). Consistently, wild-type PknI and both mutants

showed virtually identical circular dichroism spectra in the far and near UV (Fig. 3B). Additionally, analytical ultracentrifugation analysis revealed that both PknI_C20A and PknI_C20S behaved as monomers in solution, with no significant amounts of high order species, even at high protein concentrations. Overall, although these results do not allow to rule out a physiological role for the disulphide-linked dimer observed in the crystal structure of wild-type PknI, this notion is especially challenged by the non-conservation of Cys20 among orthologous PknIs from different *Mycobacterium* species (Fig. 1E).

The substrate binding cleft of PknI

Displaying the prototypical kinase fold and sequence motifs characteristic of Hanks-type kinases, as already noticed [10,13], the crystal structure of PknI presents some striking features. The activation segment, defined as the stretch running from the DFG to the APE conserved motifs [14], lacks in PknI an activation loop and a P+1 loop and is stabilized within the active site in a 'inward' conformation, incompatible with binding of peptide substrates (Fig. 4A, 4B and 5A). Further, the salt bridge between the catalytic residue Lys41 and Glu60 in helix αC, one of the hallmarks of an ePK active state [15], is hindered in PknI by the side chain of Ile162 in the activation segment (Fig. 4A), similarly to the Src-like autoinhibited conformation observed in another *M. tuberculosis* Ser/Thr kinase, PknA [16]. Thus, the side chains of residues Lys41 and Glu60 are *ca.* 14 Å apart, and Glu60 is found oriented towards the C-lobe, engaged in a salt bridge with residue Arg136 from the HRD motif (Fig. 4A). Also, the amide N of residue Gly161 is not hydrogen bonded by the γ-oxygen in Asp159 side chain, leading the DFG motif to adopt a configuration that has been associated with kinase off states [17]. This conformation is stabilized by residue Arg78 in strand β4, bridging the backbone carbonyl of Phe160 to Phe57 and Thr61 in a kink in helix αC (Fig. 4A, right panel). In line with these observations, the catalytic and regulatory spines [18] are not

properly assembled (Fig. 5B). Overall, our structural data are not descriptive of an active protein kinase [14].

The nucleotide binding site of PknI

The adenine moiety of ADP bound to chain B in wild-type PknI structure is buried in a hydrophobic pocket and makes three direct hydrogen bonds with the enzyme: the amine N at position 6 binds the backbone carbonyl oxygen of Asp90, and the N_1 atom binds both the amide and the carbonyl of Val92 (Fig. 6A, left panel). Besides, van der Waals contacts are established with the side chains of Leu18, Val26, Ala39, Leu73, Met89 and Tyr91. The α-phosphate of ADP is bound by the side chain of Lys41, which also interacts with the carboxylate of Asp159 in the DFG motif, whereas the β-phosphate is bound by residue Asp96 (Fig. 6A, right panel), a position usually found interacting with the ribose moiety in the nucleotide instead [19,20]. Interestingly, the conserved basic residue in the catalytic loop of ePKs (Lys168 in the archetypal kinase PKA), critical for keeping the ATP γ-phosphate and the peptide substrate in the near-attack reactive configuration [21,22], is substituted by asparagine (Asn139) in PknI (Fig. 6A, right panel). Additionally, one calcium ion is coordinated by Asn142 in the catalytic loop, Asp159 in the DFG motif and one water molecule that also contacts both ADP phosphates. Notably, this one cation binding mode is the most common nucleotide binding mode found among pseudokinases structures, first seen in human epidermal growth factor family pseudokinase HER3 (PDB: **4RIW**) [23]. Further, the Gly-rich loop of PknI, which in contrast to most ePKs contains a non-Gly residue in position 21, does not establish hydrogen bonds with the phosphates of ADP, found displaced out of the active site (Fig. 6B, left panel). Finally, Met23 in the Gly-rich loop is involved in a network of hydrophobic interactions comprising residues from several structural motifs, most with key roles in catalysis by ePKs: residues Lys41 and Leu43 in strand β3, Met47 in helix αB,

Phe53 and Phe57 in helix αC, Ile87 in strand β5, Asp159 in the DFG motif and Ile162 in the activation segment (Fig. 6B, right panel).

The kinase domain of PknI has no detectable kinase activity

PknI presents a degraded Gly-rich loop, the activation loop and the P+1 loop are missing from the activation segment, and ADP was found to adopt a binding mode commonly found in eukaryotic pseudokinases. Therefore, to explore PknI ability as a protein kinase, we employed wild-type PknI and Cys20 point mutants to perform kinase assays using different putative substrates (Fig. 7). No autophosphorylation activity was detected in either case in assays containing as much as 25 µM PknI and 10 mM $MnCl_2$ or $MgCl_2$. Besides, we failed to detect phosphorylation of the mycobacterial protein GarA or a 17-mer peptide derived from the N-terminus of GarA and containing the phosphorylatable ETTS motif (Fig. 7) [19].

Active ePKs require a functional activation segment in an open conformation permissive for substrate binding. This lead us to inquire whether modifications in the activation segment, either in its conformation or its nature, may provide PknI with detectable protein kinase activity. As the salt bridge observed between residues Glu60 and Arg136 may be important in stabilizing an inhibitory conformation of PknI activation segment (Fig. 4A), we first substituted Arg136 with either alanine or asparagine, as found in the only known mycobacterial non-RD kinase PknG [19]. The crystal structure of mutants PknI_C20A_R136A and PknI_C20A_R136N revealed the activation segment in the same conformation observed for the wild-type kinase domain (Fig. 8), whereas substitution of residue Arg136 did not provide PknI with detectable kinase activity (Fig. 7). These results showed that neither Arg136 nor the salt bridge Glu60-Arg136 are required for stabilizing the observed conformation of the activation segment or the helix αC of PknI. Second, we replaced the segment 162-165 (IASQ) in PknI by the stretch FAKRVKGRTWDLCGT, a mimic of the phosphorylated activation segment of the archetypal human kinase PKA. Even though this motif

has been found to adopt an open conformation allowing substrate binding in the ePK PKA [20], the resulting mutant PknI_C20A_AS-PKA did not show detectable kinase activity (Fig. 7).

Discussion

One of the key features that account for the success of *Mtb* as a human pathogen is its capability to adapt the physiological processes in response to different environments. Protein kinases provide a platform for the integration of signal transduction networks and structural analyses have greatly facilitated our knowledge about the molecular mechanisms underlying activity and regulation of these enzymes. Here we present a structural, biophysical and biochemical characterization of protein kinase I from *Mtb*.

The overall fold of PknI kinase domain is very similar to that of known ePKs and shows the characteristic bilobal topology with a nucleotide-binding cleft between the two lobes (Fig. 1). Already predicted by homology modelling [9], the ensemble of our crystal structures show that several structural motifs known to be critical for the activity of ePKs are severely degraded in PknI. Strikingly, the activation segment of PknI comprises only eight residues, whereas the same segment contains more than twenty amino acids in the closest mycobacterial orthologues PknA and PknB, and lacks the P+1 loop essential for substrate binding (Fig. 4 and 5A). Indeed, the kinase domain of RNase L, similar to PknI in that it also contains an eight-residue activation segment, is devoid of kinase activity [24]. Besides, comparable to the eukaryotic pseudokinase ILK [25], the activation segment of PknI is well ordered without any phosphorylation, and is stabilized by extensive interactions with the N-lobe in a conformation incompatible with substrate binding (Fig. 4A and 5A). Further, the DFG motif of PknI appears to be locked regardless of the presence of a bound nucleotide, a structural feature that is consistent with the rigidity of the

activation loop. The well-conserved glycine-rich GXGXXG motif in ePKs is replaced by the GXSXXG motif in PknI, that, in our crystal structure, does not form hydrogen bonds with ADP within the nucleotide binding site (Fig. 6B). Notably, ADP binds to PknI by the one cation binding mode commonly found in eukaryotic pseudokinases (Fig. 6A) [23], and would place the ATP γ-phosphate far away from the catalytic loop (Fig. 1C and 6B). Moreover, the catalytic loop of PknI includes residue Asn139 that corresponds to the catalytic lysine Lys168 in PKA (Fig. 6A), critical for phosphotransfer [21,22]. Overall, the deviations observed in the catalytic core show PknI adopting a conformational state that is also incompatible with catalysis as a conventional kinase.

Our attempts to detect kinase activity by the PknI kinase domain, either in autophosphorylation assays or by using surrogate substrates like the protein GarA or a 17-mer GarA-derived peptide, were unsuccessful (Fig. 7), similar to previous reports [2,26]. In contrast, different groups reported autophosphorylation and phosphorylation of myelin basic protein (MBP) and GarA by recombinant, full-length PknI purified from the soluble fraction of *E. coli* [13,26,27]. Therefore, yet unknown regulation mechanisms possibly to ascribe to the interactions of PknI, embedded in the cell membrane, with other signalling partners cannot be excluded. Besides, known activation mechanisms for mycobacterial kinases, like dimerization, cannot be discarded either. However, even though PknI can form dimers *via* an intermolecular disulfide bridge involving residue Cys20 (Fig. 1B), the fact that this position is not conserved among PknI from different species (Fig. 1E) argues against a conserved role for the observed interaction. On the other hand, mutants PknI_C20A and PknI_C20S behaved as monomers in solution, with no evidence of stable back-to-back dimers as previously found for the mycobacterial kinases PknB [28], PknD [29] and PknE [30]. Thus, the apparently unconventional molecular mechanism mediating PknI kinase activity, reported to be as low as 10^{-5} fmol/min/mg protein at 21 degrees against MBP [27], remains to be elucidated, although mechanisms other than its catalytic activity, like for instance a role as a

signalling scaffold or regulator, might have to be taken into account to explain the physiological role of PknI.

Materials and methods

Chemicals

The synthetic 17-mer peptide SDEVTVETTSVFRADFL, corresponding to residues 14-30 of the protein GarA, was purchased from Thermo Fisher Scientific (Ulm, Germany) with a purity >98%.

Cloning and mutagenesis

Plasmid pET28a-PknI for the expression of the kinase catalytic domain with a hexahistidine tag was constructed by PCR amplification of *pknI* codons 1-336, coding for the whole intracellular region of PknI, from *Mycobacterium tuberculosis* H37Rv genomic DNA, cloning into the *Nde*I and *Hind*III sites in plasmid pET28a (Novagen), then site-directed mutagenesis to replace the codon coding Val257 with a stop codon. Likewise, plasmids pET28a-PknI_C20A and pET28a-PknI_C20S were generated by site directed mutagenesis. Plasmids pET28a-PknI_C20A_R136A, pET28a-PknI_C20A_R136N and pET28a-PknI_C20A_AS-PKA were provided by GenScript (Piscataway, USA) using plasmid pET28a-PknI_C20A as template.

Protein production and purification

Wild-type PknI and mutants were over-produced in *E. coli* BL21(DE3) cells. Proteins were expressed for 18-20 h at 22ºC with 500 µM IPTG and then purified following the same protocol. *E. coli* cells were harvested by centrifugation, re-suspended in lysis buffer (50 mM NaH_2PO_4, 500 mM NaCl, 5% glycerol, 25 mM imidazole, pH 8.0) supplemented with Complete protease inhibitor cocktail (Roche) and sonicated. After clarification by centrifugation, the supernatant was

loaded onto a HisTrap HP column (GE Healthcare) and the His_6-tagged protein was recovered applying a linear imidazole gradient (25–400 mM). The protein was then further purified by size-exclusion chromatography on a Superdex 75 column (GE Healthcare) equilibrated in 25 mM Hepes-NaOH, 150 mM NaCl, 5% glycerol, pH 8.0. Fractions corresponding to PknI, as confirmed by SDS-PAGE, were pooled and concentrated, flash-frozen in liquid nitrogen and stored at -80ºC. GarA was prepared as previously described [31]. Proteins were quantified by using the molar absorption coefficient predicted from the aminoacid sequence by the ProtParam tool (http://web.expasy.org/protparam/).

Protein kinase activity assay

Kinase activity assays were performed in 96-well plates. Each activity measurement was performed in a final volume of 20 µl, containing 50 mM Tris-HCl pH 7.4, 0.1% v/v 2-mercaptoethanol, 10 mM $MnCl_2$ or $MgCl_2$, 100 µM [γ^{32}P]ATP (5-50 cpm/pmol), and 1 mM 17-mer peptide or 25 µM GarA as substrate. PknI concentration in the assays was 6-25 µM. Kinase reactions were started by the addition of 4 µl [γ^{32}P]ATP-Mn^{2+} (or [γ^{32}P]ATP-Mg^{2+}) and were performed at room temperature. Reactions were stopped after 1.5 hours by the addition of phosphoric acid and 4 µl of each reaction were spotted on P81 phosphocellulose papers (Whatman). Papers were washed in 0.01% phosphoric acid, dried, then measured and analyzed using a Phosphorimager apparatus (Storm™, GE Healthcare). Each reaction was performed in duplicates and each assay was performed twice.

Analytical ultracentrifugation

Sedimentation velocity experiments were carried out at 20 °C in an XL-I analytical ultracentrifuge (Beckman Coulter). Samples were spun using an An60Ti rotor and 12-mm double sector epoxy centerpieces. The partial specific volume of PknI (0.729 $ml \cdot g^{-1}$) was estimated from their amino

acid sequences using the software Sednterp. The same software was used to estimate the buffer viscosity (η = 1.036 centipoises) and density (ρ = 1.007 $g \cdot ml^{-1}$). PknI (400 µl at 10 µM, 20 µM and 35 µM) was spun at 42,000 rpm, and absorbance profiles were recorded every 5 minutes. Sedimentation coefficient distributions, c(s), were determined using the software Sedfit 15.01b [32]. Theoretical hydrodynamic radii and sedimentation coefficients of PknI (wild-type and mutants) were calculated from the atomic co-ordinates using Winhydropro [33] with a hydrated radius of 2.84 Å for the atomic elements.

Circular Dichroism

Measurements were performed using a CD 215 spectropolarimeter (Aviv) flushed with N_2. Protein samples were dialyzed against a buffer containing 25 mM Hepes-Na pH 8.0, 150 mM NaCl, 5% glycerol, 1 mM DTT and diluted to 1 mg/ml. Scans were performed in the range 190-260 nm (Far-UV) and 260-330 nm (Near-UV) using a cylindrical 0.2 mm cell (Suprasil) and a 1 nm/s scan rate.

Crystallization and data collection

Crystallization screenings were carried out using the sitting-drop vapor diffusion method and a Mosquito nanolitre-dispensing crystallization robot (TTP Labtech). Following optimization, crystals of PknI + ADP grew after six days from a ~ 35 mg/ml solution of protein supplemented with 5 mM ADP, by mixing 1 µl of protein solution and 1 µl of mother liquor (100 mM Tris-HCl, 18-21% w/v PEG 4000, 240-340 mM $CaCl_2$, 5 mM ADP, pH 8.5) in a hanging drop setup, at 18ºC. Instead, PknI_C20A crystals arose after two weeks, in 80 mM Hepes-Na, NaCl 3.5 M, pH 8.5 as mother liquor, at 4ºC, whereas PknI_C20A_R136A and PknI_C20A_R136N crystals grew within one week, using 100 mM Hepes-Na, NaCl 4.3 M, pH 7.5 as mother liquor, at 18ºC. In each case single crystals were cryprotected in mother liquor supplemented with either 25% methanol (PknI) or 25% glycerol (PknI_C20A_R136A, PknI_C20A_R136N), and flash-frozen in liquid nitrogen.

In the case of PknI_C20A crystals, the cryoprotectant solution was 4.6 M NaCl, 0.1 M Hepes-Na, pH 8.5.

Data collection, structure determination and refinement

X-ray diffraction data were collected at 100 K at the beamlines Proxima 1 (Synchrotron Soleil, Saint-Aubin, France) and ID29 (ESRF, Grenoble, France). The diffraction data were processed using XDS [34] and scaled with Aimless or Scala from the CCP4 program suite [35]. The crystal structure of PknI was solved by multi-wavelength anomalous diffraction (MAD) on a 2.9 Å resolution dataset from a SeMet-labelled crystal, collected at Proxima 1 (Table 1). Thirteen selenium sites were located with the program SHELXD [36] processing only the Se K-edge peak dataset; subsequent MAD phasing and density modification by solvent flattening were carried out with autoSHARP [37]. A preliminary model was obtained by placing the C-terminal domain of *M. tuberculosis* PknB [38] (PDB: **1O6Y**) in the resulting solvent flattened electron density map with Molrep [39]; the model was then extended through manual tracing with COOT [40], alternated with refinement with refmac5 [41] and further tracing in the resulting Fourier difference maps, and eventually used as the search model for molecular replacement on a native dataset (Table 1). The structures of PknI_C20A, PknI_C20A_R136A and PknI_C20A_R136N were then solved by using the refined atomic coordinates of wild-type PknI (Table 1). Ligand molecules were manually placed in *mFo–DFc* sigma-A-weighted electron density maps employing COOT [40]. Structures were refined through iterative cycles of manual model building with COOT [40] and reciprocal space refinement with BUSTER (Global Phasing Ltd.), applying non crystallographic symmetry restraints [42] and validated through the Molprobity server (http://molprobity.biochem.duke.edu) [43]. Figures were generated and rendered with Pymol 1.5.0.2. (Schrödinger, LLC). Atomic coordinates and structure factors have been deposited in the Protein Data Bank under the accession

codes **5M06** (wild-type PknI + ADP), **5M07** (PknI_C20A), **5M08** (PknI_C20A_R136A) and **5M09** (PknI_C20A_R136N).

Acknowledgements

We thank Ahmed Haouz and Patrick Weber (Institut Pasteur) for their help with robot-driven crystallization screenings, Bruno Baron for the assistance during circular dichroism measurements, and Clément Moroldo for his work in optimizing cryoprotection of PknI crystals. We acknowledge the synchrotron sources Soleil (Saint-Aubin, France) and ESRF (Grenoble, France) for granting access to their facilities, and their staff for the helpful assistance. MNL received postdoctoral fellowships from EMBO (European Molecular Biology Organization) and FRM (Fondation pour la Recherche Médicale, France). This work was partially supported by grants from the Institut Pasteur and the CNRS.

Author contributions

MNL, TW, PMA and MB designed research. MNL, TW, MA, NB and MB performed research. MNL, TW, MA, NB, BR, PMA and MB analysed data. MNL and MB wrote the paper. All the authors revised and approved the paper.

References

1. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S & Barrell BG (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome

sequence. *Nature* **393**, 537–544.

2. Prisic S & Husson RN (2014) *Mycobacterium tuberculosis* Serine/Threonine Protein Kinases. *Microbiol Spectr* **2** MGM2-0006-2013.

3. Cousin C, Derouiche A, Shi L, Pagot Y, Poncet S & Mijakovic I (2013) Protein-serine/threonine/tyrosine kinases in bacterial signaling and regulation. *FEMS Microbiol. Lett.* **346**, 11–19.

4. Nagarajan SN, Upadhyay S, Chawla Y, Khan S, Naz S, Subramanian J, Gandotra S & Nandicoori VK (2015) Protein kinase A (PknA) of *Mycobacterium tuberculosis* is independently activated and is critical for growth in vitro and survival of the pathogen in the host. *J Biol Chem* **290**, 9626–9645.

5. Chawla Y, Upadhyay S, Khan S, Nagarajan SN, Forti F & Nandicoori VK (2014) Protein kinase B (PknB) of *Mycobacterium tuberculosis* is essential for growth of the pathogen in vitro as well as for survival within the host. *J Biol Chem* **289**, 13858–13875.

6. Forti F, Crosta A & Ghisotti D (2009) Pristinamycin-inducible gene regulation in mycobacteria. *J Biotechnol* **140**, 270–277.

7. Fernandez P, Saint-Joanis B, Barilone N, Jackson M, Gicquel B, Cole ST & Alzari PM (2006) The Ser/Thr protein kinase PknB is essential for sustaining mycobacterial growth. *J Bacteriol* **188**, 7778–7784.

8. Gopalaswamy R, Narayanan S, Chen B & Av-Gay Y (2009) The serine/threonine protein kinase PknI controls the growth of *Mycobacterium tuberculosis* upon infection. *FEMS Microbiol. Lett.* **295**, 23–29.

9. Kandasamy S, Hassan S, Gopalaswamy R & Narayanan S (2014) Homology modelling, docking, pharmacophore and site directed mutagenesis analysis to identify the critical amino acid residue of PknI from *Mycobacterium tuberculosis*. *J Mol Graph Model* **52**, 11–19.

10. Venkatesan A, Hassan S, Palaniyandi K & Narayanan S (2015) In silico and experimental

validation of protein-protein interactions between PknI and Rv2159c from Mycobacterium tuberculosis. *J Mol Graph Model* **62**, 283–293.

11. Kandasamy S & Narayanan S (2015) Phenotypic characterization of a novel double knockout PknI/DacB2 from *Mycobacterium tuberculosis*. *Microbiol. Res.* **170**, 255–262.

12. Kahraman A, Morris RJ, Laskowski RA, Favia AD & Thornton JM (2010) On the diversity of physicochemical environments experienced by identical ligands in binding pockets of unrelated proteins. *Proteins* **78**, 1120–1136.

13. Gopalaswamy R, Narayanan PR & Narayanan S (2004) Cloning, overexpression, and characterization of a serine/threonine protein kinase *pknI* from *Mycobacterium tuberculosis* H37Rv. *Protein Expr. Purif.* **36**, 82–89.

14. Nolen B, Taylor S & Ghosh G (2004) Regulation of protein kinases; controlling activity through activation segment conformation. *Mol. Cell* **15**, 661–675.

15. Huse M & Kuriyan J (2002) The conformational plasticity of protein kinases. *Cell* **109**, 275–282.

16. Wagner T, Alexandre M, Duran R, Barilone N, Wehenkel A, Alzari PM & Bellinzoni M (2015) The crystal structure of the catalytic domain of the ser/thr kinase PknA from *M. tuberculosis* shows an Src-like autoinhibited conformation. *Proteins* **83**, 982–988.

17. Kornev AP, Haste NM, Taylor SS & Eyck Ten LF (2006) Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc Natl Acad Sci U S A* **103**, 17783–17788.

18. Kornev AP & Taylor SS (2010) Defining the conserved internal architecture of a protein kinase. *Biochim Biophys Acta* **1804**, 440–444.

19. Lisa M-N, Gil M, Andre-Leroux G, Barilone N, Duran R, Biondi RM & Alzari PM (2015) Molecular Basis of the Activity and the Regulation of the Eukaryotic-like S/T Protein Kinase PknG from *Mycobacterium tuberculosis*. *Structure* **23**, 1039–1048.

20. Zheng J, Trafny EA, Knighton DR, Xuong NH, Taylor SS, Eyck Ten LF & Sowadski JM (1993) 2.2 A refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor. *Acta Crystallogr D Biol Crystallogr* **49**, 362–365.

21. Cheng YH, Zhang YK & McCammon JA (2005) How does the cAMP-dependent protein kinase catalyze the phosphorylation reaction: An ab initio QM/MM study. *J Am Chem Soc* **127**, 1553–1562.

22. Taylor SS, Yang J, Wu J, Haste NM, Radzio-Andzelm E & Anand G (2004) PKA: a portrait of protein kinase dynamics. *Biochim Biophys Acta* **1697**, 259–269.

23. Hammaren HM, Virtanen AT & Silvennoinen O (2016) Nucleotide-binding mechanisms in pseudokinases. *Bioscience Reports* **36**, e00282–e00282.

24. Han Y, Donovan J, Rath S, Whitney G, Chitrakar A & Korennykh A (2014) Structure of human RNase L reveals the basis for regulated RNA decay in the IFN response. *Science* **343**, 1244–1248.

25. Fukuda K, Gupta S, Chen K, Wu C & Qin J (2009) The pseudoactive site of ILK is essential for its binding to alpha-Parvin and localization to focal adhesions. *Mol. Cell* **36**, 819–830.

26. Baer CE, Iavarone AT, Alber T & Sassetti CM (2014) Biochemical and spatial coincidence in the provisional Ser/Thr protein kinase interaction network of *Mycobacterium tuberculosis*. *J Biol Chem* **289**, 20422–20433.

27. Singh A, Singh Y, Pine R, Shi LB, Chandra R & Drlica K (2006) Protein kinase I of *Mycobacterium tuberculosis*: Cellular localization and expression during infection of macrophage-like cells. *Tuberculosis* **86**, 28–33.

28. Mieczkowski C, Iavarone AT & Alber T (2008) Auto-activation mechanism of the *Mycobacterium tuberculosis* PknB receptor Ser/Thr kinase. *EMBO J* **27**, 3186–3197.

29. Greenstein AE, Echols N, Lombana TN, King DS & Alber T (2007) Allosteric activation by dimerization of the PknD receptor Ser/Thr protein kinase from *Mycobacterium tuberculosis*. *J Biol*

Chem **282**, 11427–11435.

30. Gay LM, Ng H-L & Alber T (2006) A conserved dimer and global conformational changes in the structure of apo-PknE Ser/Thr protein kinase from *Mycobacterium tuberculosis*. *J Mol Biol* **360**, 409–420.

31. England P, Wehenkel A, Martins S, Hoos S, Andre-Leroux G, Villarino A & Alzari PM (2009) The FHA-containing protein GarA acts as a phosphorylation-dependent molecular switch in mycobacterial signaling. *FEBS Lett* **583**, 301–307.

32. Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78**, 1606–1619.

33. García De La Torre J, Huertas ML & Carrasco B (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys J* **78**, 719–730.

34. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132.

35. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A & Wilson KS (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235–242.

36. Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A Found Adv* **64**, 112–122.

37. Vonrhein C, Blanc E, Roversi P & Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* **364**, 215–230.

38. Ortiz-Lombardía M, Pompeo F, Boitel B & Alzari PM (2003) Crystal structure of the catalytic domain of the PknB serine/threonine kinase from Mycobacterium tuberculosis. *J Biol Chem* **278**, 13094–13100.

39. Vagin A & Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**, 22–25.

40. Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.

41. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–367.

42. Smart OS, Womack TO, Flensburg C, Keller P, Paciorek W, Sharff A, Vonrhein C & Bricogne G (2012) Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr D Biol Crystallogr* **68**, 368–380.

43. Chen VB, Arendall WB3, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12–21.

44. Baker NA, Sept D, Joseph S, Holst MJ & McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* **98**, 10037–10041.

45. Dolinsky TJ, Nielsen JE, McCammon JA & Baker NA (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* **32**, W665–7.

46. Di Tommaso P, Moretti S, Xenarios I, Orobitg M, Montanyola A, Chang J-M, Taly J-F & Notredame C (2011) T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res* **39**, W13–7.

47. Gouet P, Courcelle E, Stuart DI & Metoz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305–308.

48. Karplus PA & Diederichs K (2012) Linking Crystallographic Model and Data Quality. *Science* **336**, 1030–1033.

Table 1. *Crystallographic data collection and refinement statistics*. Values in parentheses refer to the highest-resolution shell.

Data collection	SeMet_PknI Peak	SeMet_PknI Infl. point	SeMet_PknI High remote	PknI + ADP	PknI _C20A	PknI_C20A_R136A	PknI_C20A_R136N
Beamline	Soleil Proxima 1	Soleil Proxima 1	Soleil Proxima 1	Soleil Proxima 1	Soleil Proxima 1	ESRF ID29	ESRF ID29
Space group	C 2	C 2	C 2	C 2 2 2_1	P 4_3 2 2	P 4_3 2 2	P 4_3 2 2
Cell dimensions							
a, *b*, *c* (Å)	61.63, 68.57, 130.50	61.67, 68.59, 130.56	61.71, 68.57, 130.59	60.00, 69.54, 257.42	108.30, 108.30, 185.47	109.31, 109.31, 184.12	110.87, 110.87, 181.76
α, β, γ (°)	90, 94.27, 90	90, 94.18, 90	90, 94.07, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	33.15 – 2.80 (2.95 – 2.80)	33.16 – 2.90 (3.08 – 2.90)	33.16 – 2.90 (3.08 – 2.90)	45.43 – 2.00 (2.05 – 2.00)	48.44 – 2.50 (2.60 – 2.50)	48.88 – 3.03 (3.21- 3.03)	47.94 – 2.98 (3.16 – 2.98)
Unique reflections	13405 (1933)	12117 (1958)	12119 (1965)	36891	38990	22374	23530
R_{merge}	0.050 (0.423)	0.051 (0.484)	0.061 (0.856)	0.052 (0.840)	0.051 (1.836)	0.061 (0.444)	0.068 (0.406)
$<I / \sigma I>$	15.8 (2.9)	16.1 (2.6)	13.6 (1.4)	17.0 (2.0)	29.2 (1.6)	9.7 (2.0)	10.5 (2.0)
Completeness (%)	99.3 (99.3)	99.5 (99.6)	99.4 (99.7)	99.8 (97.6)	100.0 (100.0)	99.9 (100)	98.9 (99.7)
Redundancy	3.7 (3.5)	3.7 (3.7)	3.7 (3.7)	5.9 (4.9)	12.7 (13.0)	5.4 (5.4)	5.5 (5.5)
$CC_{1/2}$ *	0.996 (0.812)	0.997 (0.776)	0.998 (0.547)	0.998 (0.668)	1.000 (0.794)	0.999 (0.905)	0.998 (0.945)
Phasing							
Phasing power†	1.237	0.667	0.523	-	-	-	-
Refinement							
Resolution (Å)	-	-	-	28.09 – 2.00	36.1 – 2.50	48.88 – 3.03	47.94 – 2.98
No. reflections	-	-	-	36813	38927	22374	23530
R_{work} / R_{free}	-			0.212 / 0.239	0.209 / 0.235	0.184 / 0.209	0.190 / 0.231
No. atoms							
Protein	-	-	-	3697	3714	3710	3 682
Ions‡	-	-	-	1 (Ca^{2+})	2 (Na^{+})	-	2 (Na^{+})
Water	-	-	-	127	154	22	58
B-factors							
Wilson B (Å^2)	-	-	-	48.1	89.0	96.2	82.5
Mean B value (Å^2)	-	-	-	62.2	85.3	87.2	64.6
R.m.s. deviations							
Bond lengths (Å)	-	-	-	0.010	0.010	0.010	0.010
Bond angles (°)	-	-	-	0.92	1.01	1.03	1.01
PDB ID	-	-	-	5M06	5M07	5M08	5M09

* $CC_{1/2}$ according to [48].

† Anomalous phasing power: $< |F_h(calc)| /$ phase-integrated lack of closure $>$. Calculated figures of merit (FOM) after MAD phasing were 0.354/0.136 (acentric/centric reflections).

‡ Ion species modelled as either Ca^{2+} or Na^{+} according both to their chemical environment and considering the components of protein buffers and crystallization solutions.

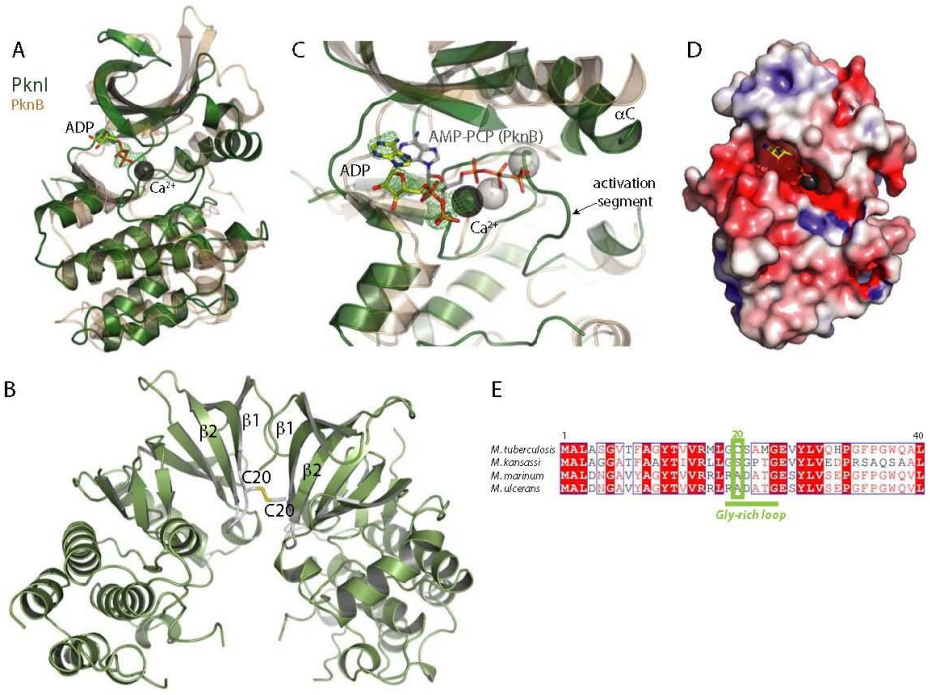


Figure 1. *Crystal structure of PknI kinase domain.* (A) Cartoon representation of the catalytic domain of wild-type PknI (green) superimposed to the equivalent domain from *M. tuberculosis* PknB (translucent light brown; PDB: **1O6Y**) as reference [38]. (B) The crystal structure of wild-type PknI contains two protein chains in the asymmetric unit, which are linked by a disulfide bridge involving the same residue Cys20 from each monomer, part of the conserved 'Gly-rich loop' (in gray) on the tip of the β1-β2 turn. (C) Close view of the active site, showing the bound ADP molecule (sticks with yellow carbon atoms) and the relative mFo-DFc electron density map (computed before the ligand was added to the model) and contoured at the 3σ level. Only one over two monomers in the asymmetric unit shows density for ADP. The bound Ca^{2+} ion is shown as a dark grey sphere. In translucent grey, AMP-PCP/Mg^{2+} bound to PknB (PDB: **1O6Y**) as in (A). The arrow highlights the conformation of the activation segment. (D) Poisson-Boltzmann electrostatic surface potential calculated with the APBS [44] plugin in pymol from atom charges and radii provided by the pdb2pqr server, applying the CHARMM force field [45]. The bound ADP/Ca^{2+} are shown as in (A). (E) Sequence alignment of PknI from different *Mycobacterium* species, showing the non-conservation of residue Cys20.

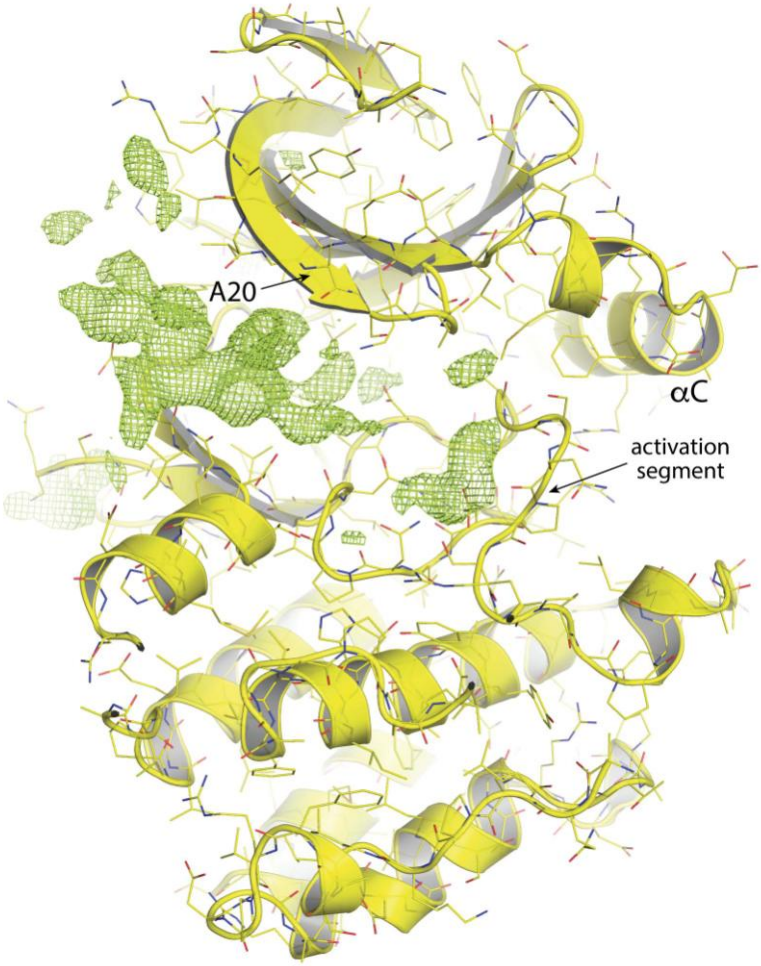


Figure 2. *Electron density in the nucleotide binding pocket of PknI_C20A (monomer A only).* Cartoon diagram of PknI_C20A, limited to the monomer A, showing residual, uninterpretable positive Fourier difference electron density in the nucleotide binding pocket of monomer A. Side chains are represented as thin sticks. Electron density is here shown as a green mesh, corresponding to the mFo-DFc sigmaA-weighted electron density map as provided by BUSTER (version 2.10.2; Global Phasing Ltd., Cambridge, UK) after the last refinement step, and contoured at the 3σ level (~0.122 $e^-/Å^3$).

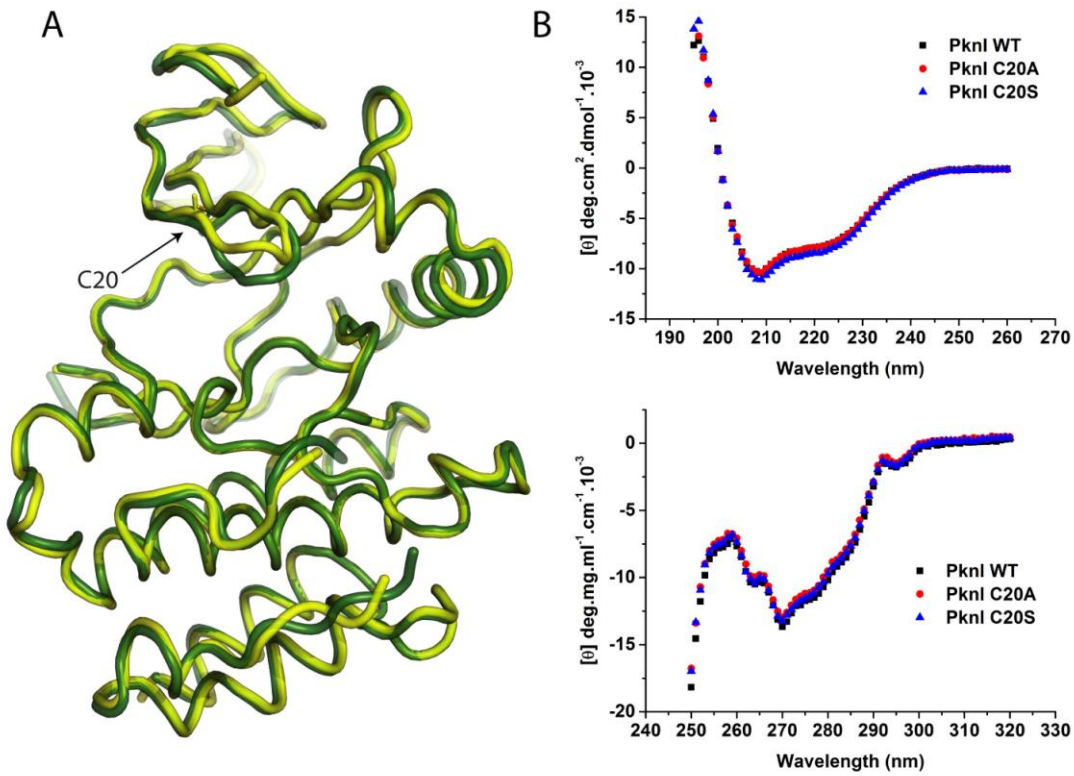
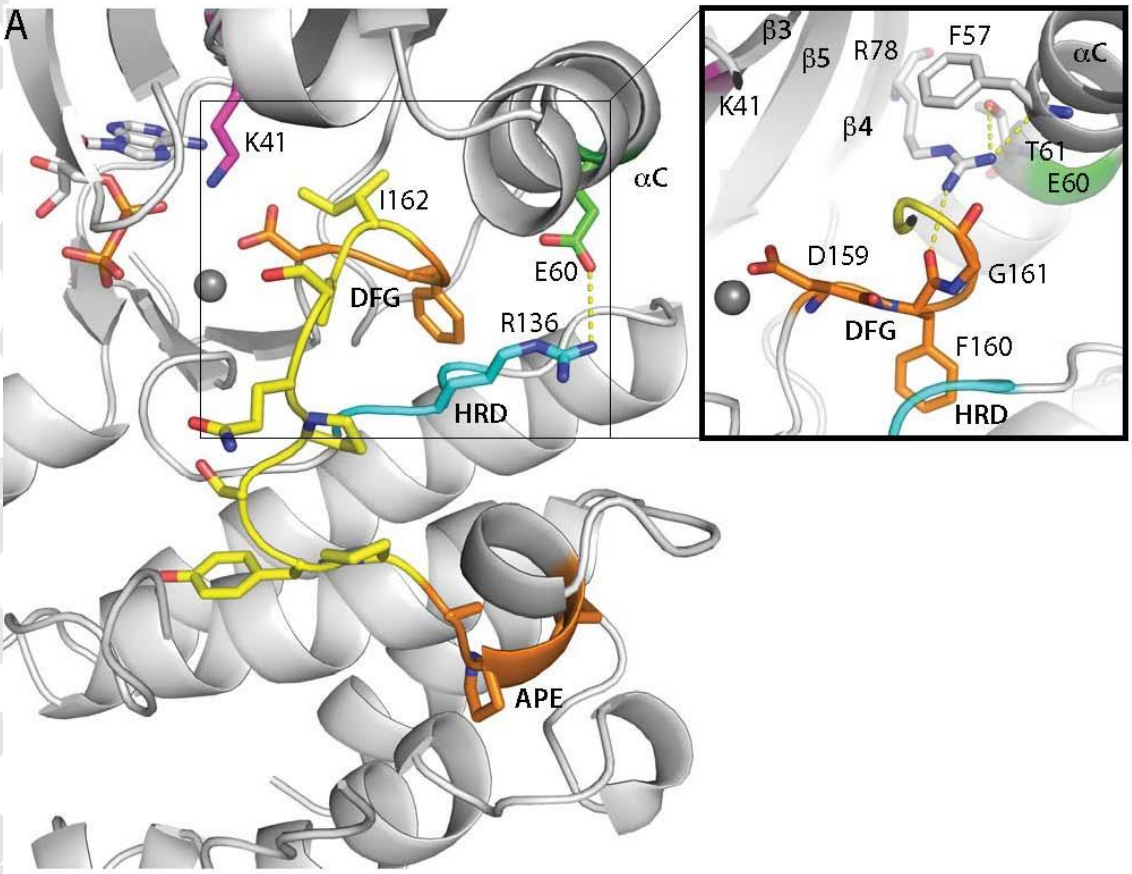


Figure 3. *Comparison of wild-type PknI and C20 muteins.* (A) Superimposition of the crystal structures of wild-type PknI (green) and PknI_C20A (yellow), showing the same overall conformation. (B) Circular dichroism spectra in the far-UV (upper panel) and near-UV (lower) of wild-type PknI and Cys20 muteins.



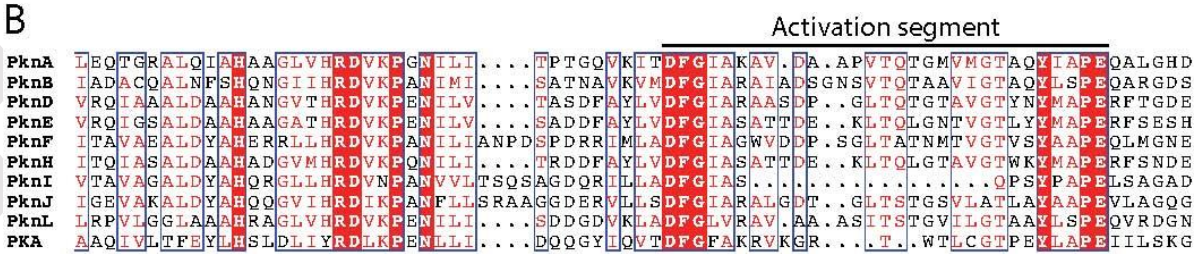


Figure 4. *The substrate binding cleft of PknI.* (A) View of PknI showing the activation segment (highlighted with yellow and orange C atoms), *i.e.* the segment running from the DFG (Asp159 to Gly161) to the APE (Ala170 to Glu172) motif. The inset at the right shows in detail the conformation of the DFG motif. The bound Ca^{2+} ion is shown as a grey sphere. (B) Multiple

sequence alignment of *M. tuberculosis* transmembrane kinases (PknA, PknB, PknD, PknF, PknH, PknI, PknJ, PknL) and human PKA [19] as reference (PDB: **1ATP**), realized with T-Coffee [46] and limited to the activation loop region, showing the apparent deletion of most of the activation segment in PknI. Alignment visualized through Espript [47].

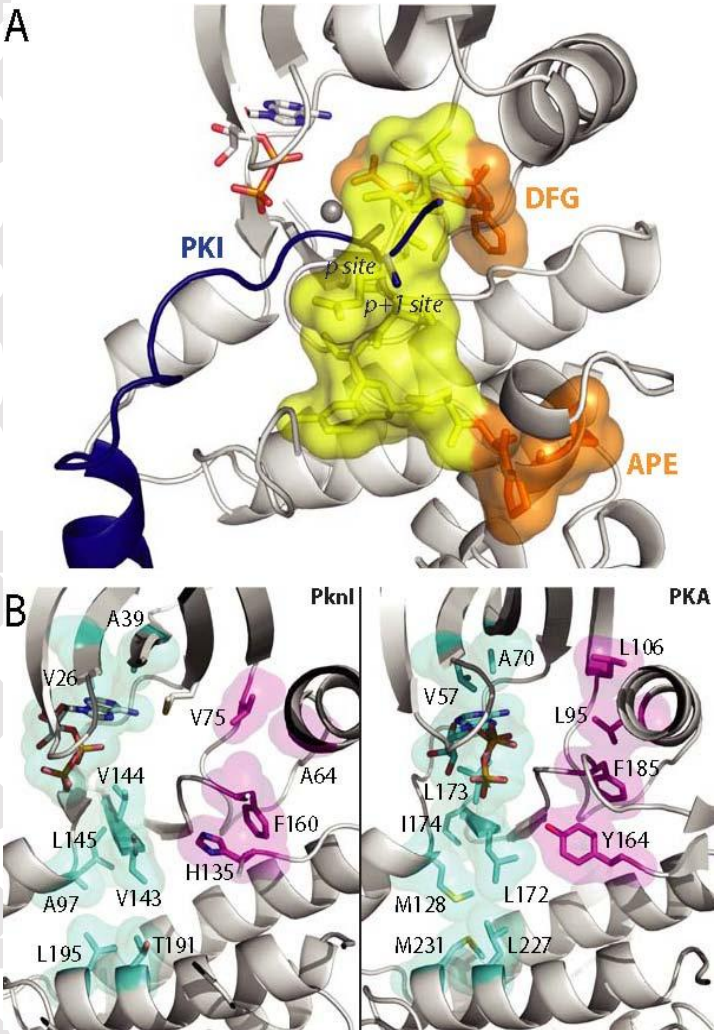


Figure 5. *The crystal structure of PknI is not descriptive of an active protein kinase.* (A) View of PknI showing the activation segment (surface highlighted), *i.e.* the segment included between the DFG (Asp159 to Gly161) and the APE (Ala170 to Glu172) motifs. Peptide PKI (in blue) was positioned by superimposition of the eukaryotic kinase PKA (not shown) in complex with PKI (PDB: **1ATP**) onto PknI. (B) View of PknI (left) and PKA (right), highlighting residues in the catalytic (cyan) and regulatory (pink) spines, respectively.

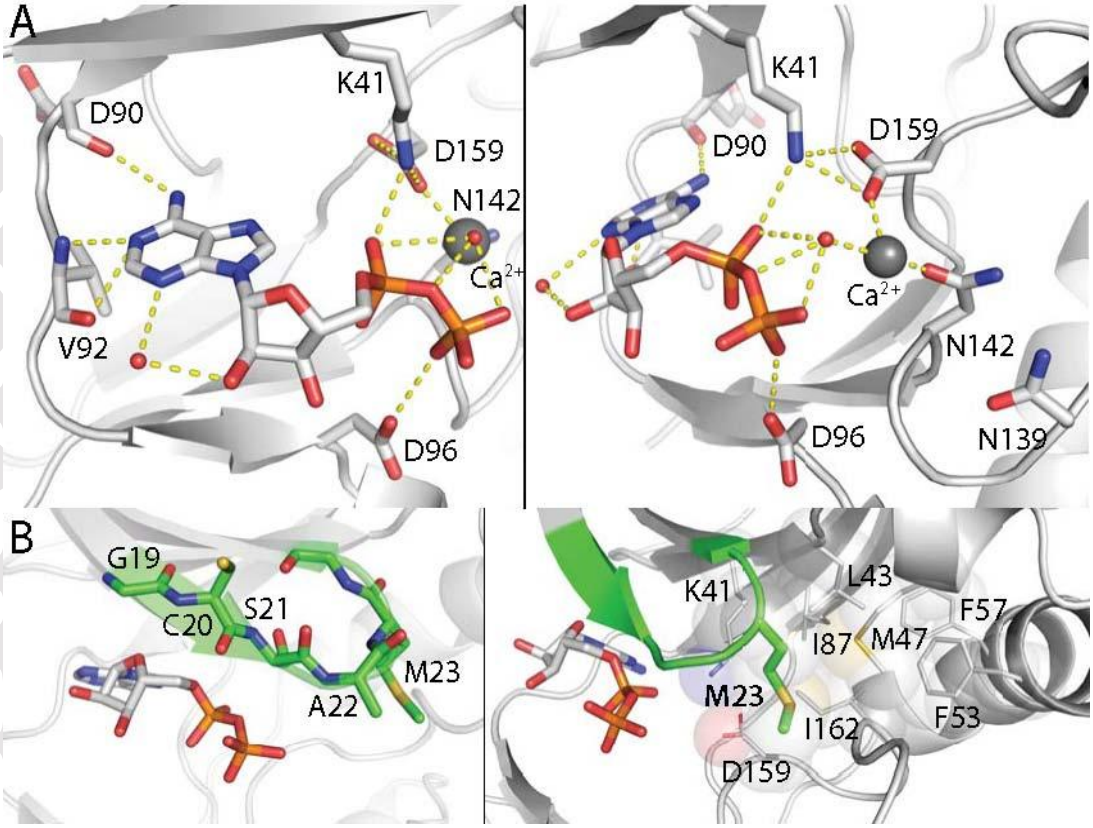


Figure 6. *The nucleotide binding site of PknI.* (A) Chain B is depicted as grey ribbons. The ADP molecule and the protein residues interacting with it are shown in sticks with atoms coloured by type (C, grey; N, blue; O, red; P, orange). Water molecules are presented as red spheres and the calcium atom is shown as a dark grey sphere. Dashed lines represent atomic interactions. (B) Left panel: highlight of the Gly-rich loop in PknI (sticks with green C atoms). Right panel: network of hydrophobic interactions involving residue Met23 in the Gly-rich loop of PknI.

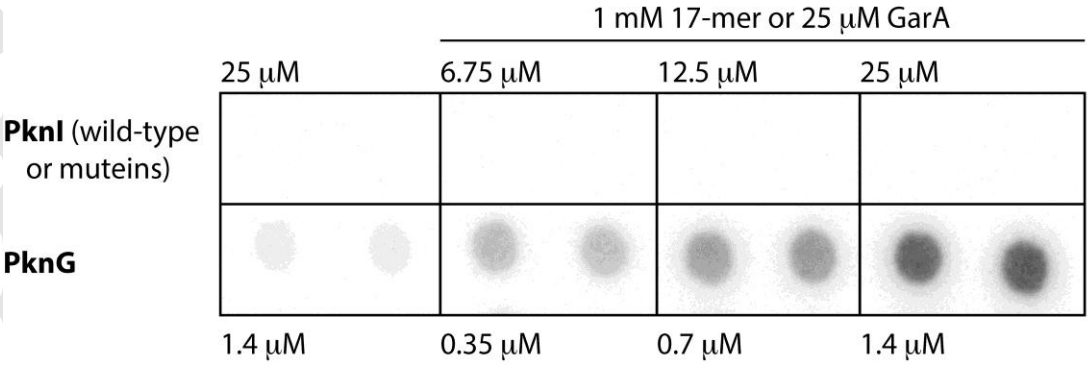


Figure 7. *The kinase domain of PknI is devoid of kinase activity.* The kinase activity of wild-type PknI or mutants was assayed in medium containing 50 mM Tris-HCl pH 7.4, 0.1% v/v 2-mercaptoethanol, 10 mM $MnCl_2$ and 100 μM [γ^{32}P]ATP (5-50 cpm/pmol), with the protein alone

or using 1 mM 17-mer SDEVTVETTSVFRADFL or 25 μM GarA as substrate. Reactions were performed at room temperature and stopped after 1.5 hours. Identical results were obtained when replacing Mn(II) by Mg(II) in the reaction medium. The protein kinase PknG from *M. tuberculosis* was used as a positive control [19].

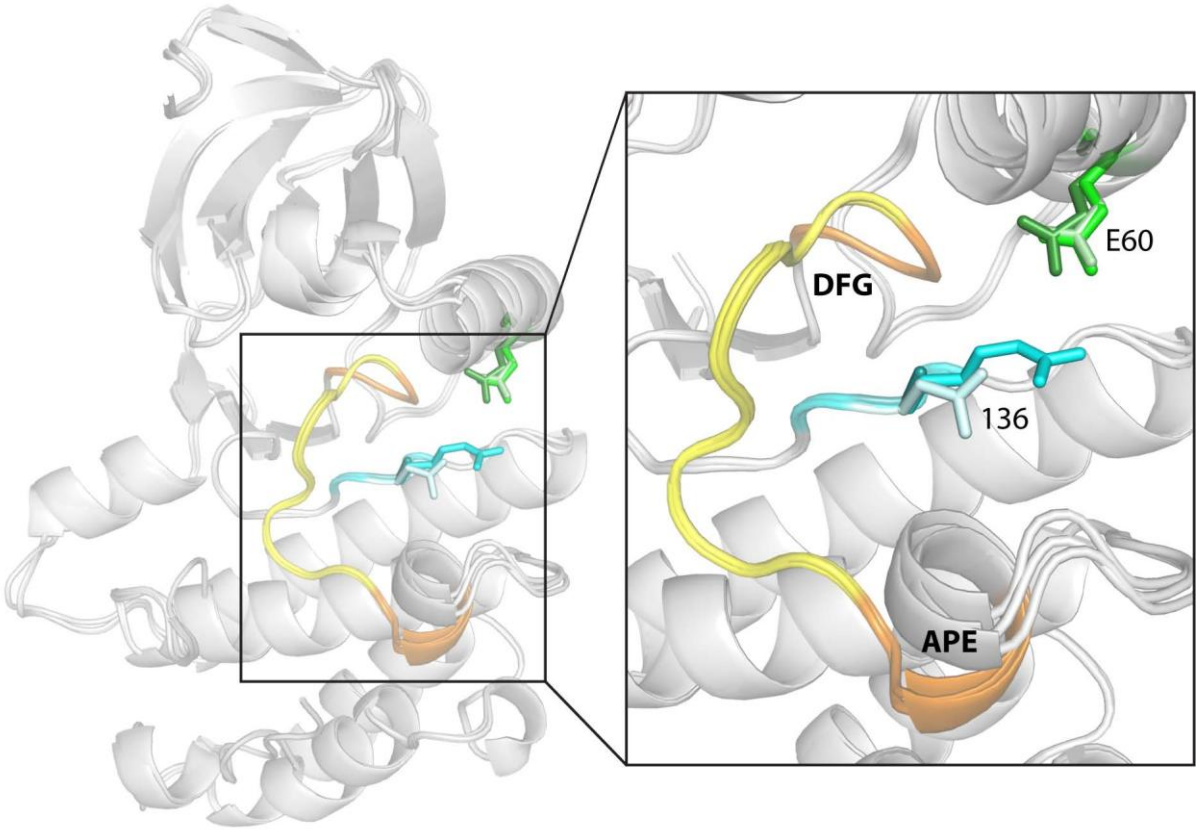


Figure 8. *Substitution of PknI residue Arg136.* Mutations R136A and R136N do not impact on the conformation of the PknI kinase domain, as shown by the superimposition of each crystal structure (Table 1) to wild-type PknI. The DFG and APE motifs, as well as the side chain of Glu60 are indicated.