Structure of *Mycobacterium tuberculosis* PknB supports a universal activation mechanism for Ser/Thr protein kinases

Tracy A. Young[1,2], Benedicte Delagoutte[1,2], James A. Endrizzi[1], Arnold M. Falick[3] and Tom Alber[1]

Published online 27 January 2003; doi:10.1038/nsb897

A family of eukaryotic-like Ser/Thr protein kinases occurs in bacteria, but little is known about the structures and functions of these proteins. Here we characterize PknB, a transmembrane signaling kinase from *Mycobacterium tuberculosis*. The intracellular PknB kinase domain is active autonomously, and the active enzyme is phosphorylated on residues homologous to regulatory phospho-acceptors in eukaryotic Ser/Thr kinases. The crystal structure of the PknB kinase domain in complex with an ATP analog reveals the active conformation. The predicted fold of the PknB extracellular domain matches the proposed targeting domain of penicillin-binding protein 2x. The structural and chemical similarities of PknB to metazoan homologs support a universal activation mechanism of Ser/Thr protein kinases in prokaryotes and eukaryotes.

Kinases are ubiquitous regulatory enzymes that provide numerous pharmaceutical targets. The Ser/Thr-Tyr protein kinases comprise the largest family of kinases. In a recent survey[1], 9,600 of >17,000 kinase sequences identified in the nonredundant database were classified in this fold family. Notably, the Ser/Thr-Tyr protein kinases occur almost entirely in eukaryotes. In comparison to >500 members in humans[2] and >100 in yeast, only two members of this family occur in the archaeon *Methanococcus jannaschii* and only one is found in *Escherichia coli*. Thus, the vast divergence of Ser/Thr-Tyr protein kinases correlates with the advent of eukaryotes and the development of animals.

In this context, it is noteworthy that 11 predicted Ser/Thr protein kinases occur in the genome of *Mycobacterium tuberculosis*[3,4]. This pathogen infects nearly one-third of the world's human population and causes over two million deaths annually[5]. Tuberculosis is a frequent immediate cause of death in HIV infections. Eukaryotic kinase inhibitors block growth of *M. tuberculosis*[6,7], but their mechanism of action is unknown. No structures of bacterial Ser/Thr kinases are available to shed light on their functions, origins and mechanistic similarities to eukaryotic homologs.

Nine of the *M. tuberculosis* Ser/Thr protein kinases encode putative membrane proteins with a single predicted transmembrane helix[4]. The kinase domains are thought to be intracellular, and the functions of the predicted extracellular domains, which range from 6 to 276 amino acids in length, have not been assigned. The PknB kinase is prototypical; the kinase homology domain (residues 11–268) is followed by a 50-residue linker to a transmembrane helix and a 274-residue extracellular domain (Fig. 1*a*,*b*). PknB is implicated as a potential regulator of cell growth and division, because it is encoded in an operon that includes genes (*rodA*, *pknA* and *pbpA*) important for these processes[3,4,8]. To compare the prokaryotic Ser/Thr kinases with their eukaryotic homologs and to investigate the basis for function and signaling, we characterized the activity, phosphorylation sites and structure of *M. tuberculosis* PknB.

Activity and phosphorylation sites

As expected for a membrane protein, full-length PknB produced in *E. coli* is insoluble in aqueous buffers[9]. To simplify analysis of the kinase activity, we expressed N-terminal fragments of PknB encompassing the kinase domain alone or including 28 residues of the linker to the transmembrane helix (Fig. 1*a*). We focused on the longer of the two constructs (residues 1–307; PknB-I) because it was more soluble. Similar to many eukaryotic Ser/Thr protein kinases, PknB-I catalyzed phosphorylation of myelin basic protein (MBP) and also catalyzed autophosphorylation (Fig. 1*c*). These results demonstrated that PknB-I folds into the native conformation and catalyzes phosphotransfer in the absence of regulatory cues from the extracellular domain.

Full-length *M. tuberculosis* PknA, PknB, PknG and PknH expressed in *E. coli* also catalyze autophosphorylation[8–10], but the number and location of the phosphorylation sites have not been mapped. Mass spectrometry (MS) showed that His-tagged PknB-I contained a mixture of species whose masses corresponded within 1.5 Da to the expected mass of the protein plus 3–8 phosphates (Fig. 1*d*). The phosphorylated sites (Fig. 1*a*) were mapped by MS and tandem mass spectrometry (MS/MS) analysis of a trypsin digest of PknB-I. Three heterogeneously phosphorylated peptides were found and identified by MS/MS of the corresponding peak. In each of these MS/MS spectra, strong fragment peaks were observed that corresponded to successive losses of 98 Da per phosphate group (usually accompanied by a peak for loss of 80 Da). A single tryptic peptide (residues 162–189) encompassing the sequence homologous to the 'activation loop' in eukaryotic Ser/Thr protein kinases[10] contained 2–4 phosphates (Fig. 1*e*). MS/MS provided evidence for an unphosphorylated peptide corresponding to residues 180–189, suggesting that the four phosphates occur in the 162–179 segment. MS/MS analysis

[1]Department of Molecular and Cell Biology, 229 Stanley Hall 3206, University of California, Berkeley, California 94720-3206, USA. [2]These authors contributed equally to this work. [3]Howard Hughes Medical Institute, 16 Barker Hall 3202, University of California, Berkeley, California 94720-3202, USA.

Correspondence should be addressed to T.A. *e-mail: tom@ucxray.berkeley.edu*

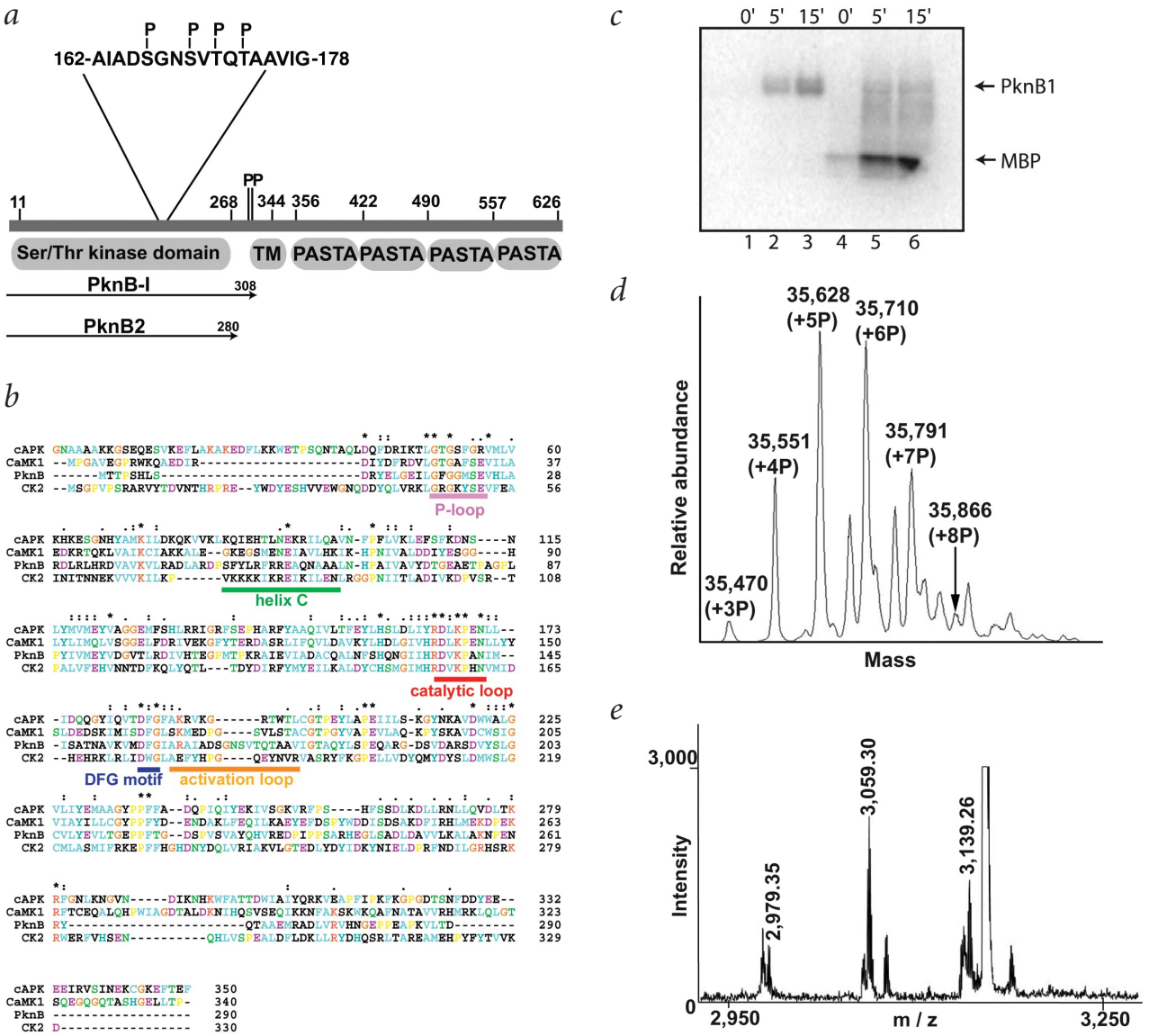


Fig. 1 Domain structure of PknB and phosphorylation of the intracellular domain. ***a***, Schematic drawing of the domains of PknB. *M. tuberculosis* PknB encodes a predicted membrane protein with a single transmembrane helix. A low-complexity linker joins the kinase domain to the transmembrane helix. The extracellular domain contains four predicted PASTA domains. Arrows indicate the boundaries of the expressed constructs of the kinase domain. Tryptic fragments corresponding to 86% of the PknB-I sequence were identified by MS, and likely autophosphorylation sites are designated with a 'P'. ***b***, Sequence alignment of the kinase domains of PknB, cyclic AMP–dependent protein kinase (cAPK, mouse), Ca^{2+}-calmodulin–dependent protein kinase (CaMK1, rat) and casein kinase 2 (CK2, human). Several conserved and functionally important motifs are indicated. No significant similarity was detected by BLAST beyond PknB residue 234. The levels of sequence identity with the PknB kinase domain (1–286) are 24% for cAPK, 21% for CK2 and 27% for CaMK1. Stars denote invariant residues; double and single dots indicate higher and lower similarity, respectively. ***c***, Catalytic activity of the PknB intracellular domain. Autophosphorylation of PknB-I and phosphorylation of MBP were visualized by SDS-PAGE and autoradiography. Lanes 1–3 include PknB-I was mixed with [γ-^{32}P]ATP in assay buffer, and aliquots were quenched at indicated times. The appearance of radioactivity in PknB-I indicates that additional phosphates were added to the purified phospho-protein. In lanes 4–6, PknB-I was mixed with MBP and [γ-^{32}P-ATP] in assay buffer, and aliquots were quenched at indicated times. Multiple bands in the autoradiogram indicate phosphorylation of MBP. ***d***, Deconvoluted electrospray mass spectrum of intact PknB-I showing peaks corresponding to the protein plus 3–8 phosphoryl groups. The peaks are labeled with their experimental masses and the number of phosphoryl groups. ***e***, Phosphoryl-group mapping. Representative partial mass spectrum of a tryptic digest of PknB-I showing peaks corresponding to the activation loop peptide, residues 162–189. The labeled peaks correspond to the monoisotopic MH^+ masses of the peptide modified with two (2,979), three (3,059) or four (3,139) phosphates. No evidence was observed for other states of phosphorylation of this peptide. The large (off-scale) peak at m/z 3,150.538 (MH^+, monoisotopic) corresponds to a fragment comprising residues 60–90.

also indicated that in the linker segment joining the kinase homology domain to the transmembrane helix, Thr294 was fully phosphorylated and Ser295 was predominantly phosphorylated (data not shown). In addition, peptides corresponding to the His-tag fragment containing zero, one or two phosphoserines were detected (data not shown). These results accounted for all eight phosphorylation sites, with two sites occurring in the His-tag and six in PknB-I (Fig. 1*a*).

PknB-I crystal structure

To explore the basis of the kinase activity, we used MAD to determine the 3.0 Å resolution X-ray crystal structure of the SeMet-

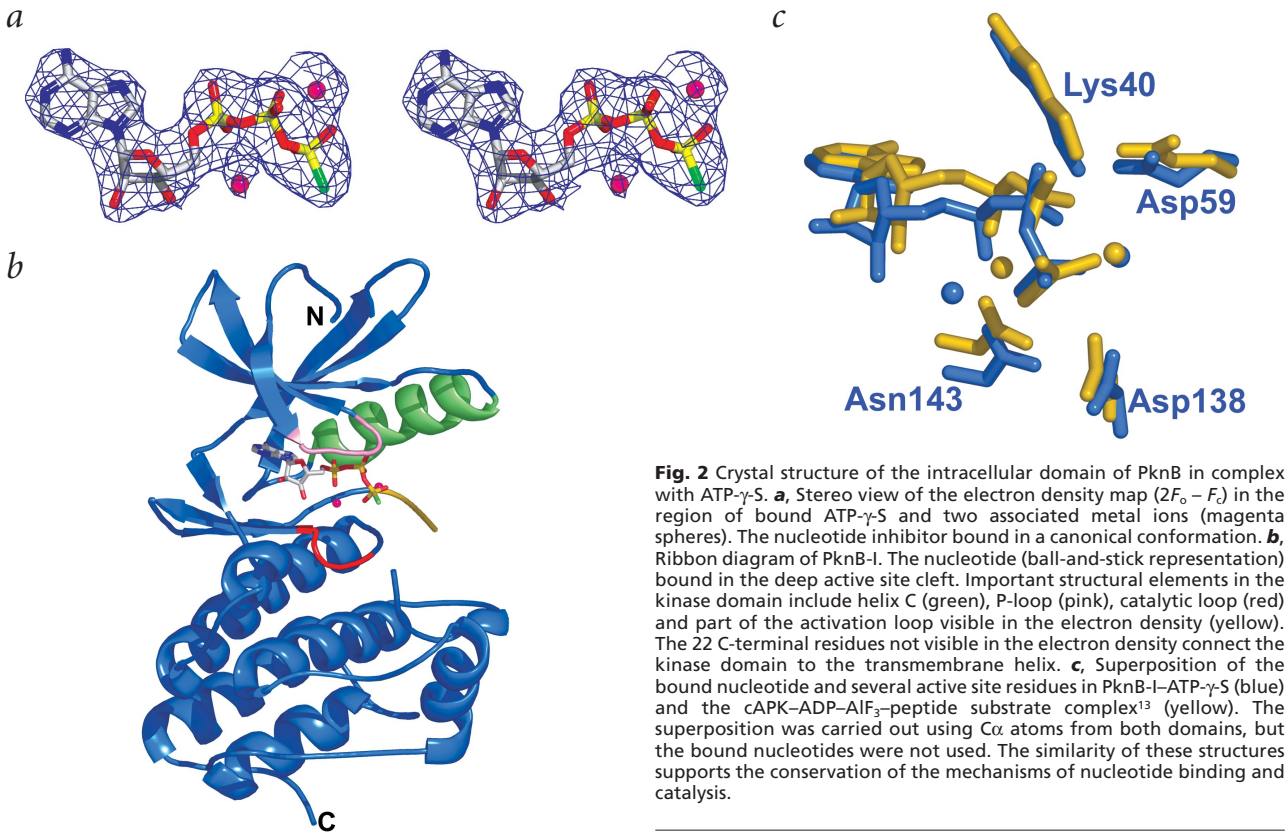


Fig. 2 Crystal structure of the intracellular domain of PknB in complex with ATP-γ-S. ***a***, Stereo view of the electron density map ($2F_o - F_c$) in the region of bound ATP-γ-S and two associated metal ions (magenta spheres). The nucleotide inhibitor bound in a canonical conformation. ***b***, Ribbon diagram of PknB-I. The nucleotide (ball-and-stick representation) bound in the deep active site cleft. Important structural elements in the kinase domain include helix C (green), P-loop (pink), catalytic loop (red) and part of the activation loop visible in the electron density (yellow). The 22 C-terminal residues not visible in the electron density connect the kinase domain to the transmembrane helix. ***c***, Superposition of the bound nucleotide and several active site residues in PknB-I–ATP-γ-S (blue) and the cAPK–ADP–AlF_3–peptide substrate complex[13] (yellow). The superposition was carried out using Cα atoms from both domains, but the bound nucleotides were not used. The similarity of these structures supports the conservation of the mechanisms of nucleotide binding and catalysis.

labeled, PknB-I–ATP-γ-S complex (Table 1; Fig. 2). The two PknB-I molecules in the asymmetric unit were modeled in two segments comprising residues 1–163 (or 3–163) and 179–286. No electron density was apparent for residues 164–178 or the C-terminal 21 residues. Except for the N-terminal seven residues, which mediate crystal contacts in molecule B, the two independent PknB-I molecules had similar structures. The r.m.s. deviation of equivalent Cα atoms was 0.7 Å. Consistent with the locations of the phosphorylated residues determined by MS, no phosphorylated residues were seen in the ordered parts of the structure.

The PknB kinase domain in complex with ATP-γ-S adopts the characteristic two-lobed structure of eukaryotic Ser/Thr kinases[11] (Fig. 2*b*). The topology of the kinase matches that of the eukaryotic homologs, and the nucleotide is sandwiched in the deep cleft between the lobes. Most active site signature sequences (Fig. 1*b*) adopt conformations that define the activated states of eukaryotic Ser/Thr kinases[11] (Fig. 2*c*). Lys40 contacts the α- and β-phosphates, as well as the conserved Glu59 in helix C. The hexapeptide 18-GFGGMS-23 forms a standard P-loop in which the main chain amides of the glycines coordinate the phosphates. Asp156 of the DFG motif and Asn143, which makes a canonical contact to Asp138 in the active-site loop (137–143), coordinate the Mg^{2+} ions. Residues in the adenine-binding pocket also are structurally equivalent to eukaryotic enzymes.

The structural similarities to the activated states of eukaryotic Ser/Thr protein kinases encompass nearly the entire PknB kinase domain (Fig. 3). The PknB-I–ATP-γ-S complex shows the characteristic closed conformation of the hinge between the N- and C-terminal domains associated with nucleotide binding. In addition, the conformation of the N-terminal β-hairpin resembles that seen in activated Ser/Thr kinases (such as casein kinase-2 (CK2)–ATP-γ-S[12] and cyclic AMP–dependent protein kinase (cAPK)–ADP–AlF_3–peptide substrate complex[13]) but not in Ser/Thr kinases in the 'off' state[11,14] (Fig. 3). In contrast to activated eukaryotic homologs[11,13], however, residues 164–178 in the phosphorylated activation loop are disordered. In addition, unlike the cAPK–ADP–AlF_3–peptide substrate complex[13], helix C is tipped away from the active site to a position similar to that found in inhibited Ser/Thr kinases (Fig. 3).

PknB extracellular domain

To explore the signaling mechanism of PknB, we analyzed computationally the sequence of the extracellular domain. A BLAST search revealed the presence of four sequence repeats of ~66 amino acids with up to 43% identity with each other (Fig. 1*a*). Clues about function were obtained using the 3D-PSSM fold recognition method[15], which revealed structural compatibility (e-value = 6.9×10^{-4}, >95% confidence) of the PknB repeats with the proposed targeting domain of *Streptococcus pneumoniae* penicillin-binding protein 2x (Pbp2x)[16], a cell wall biosynthetic-enzyme. The Pbp2x targeting domain contains two homologous modules in tandem, and each repeat adopts a unique $\beta_3\alpha$ structural topology. This predicted structural homology was confirmed using the Conserved Domain Architecture Retrieval Tool[17] (CDART). CDART classified the repeats as PASTA (Pbp and Ser/Thr kinase attached) domains[18] and retrieved 127 sequences containing PASTA-domain homologs. These proteins, including 29 predicted Ser/Thr kinases, 82 penicillin-binding proteins and 13 hypothetical proteins, were all from bacteria.

Comparison with eukaryotic Ser/Thr kinases

The crystal structure of the PknB-I–ATP-γ-S complex establishes the close conformational similarity of the prokaryotic and eukary-

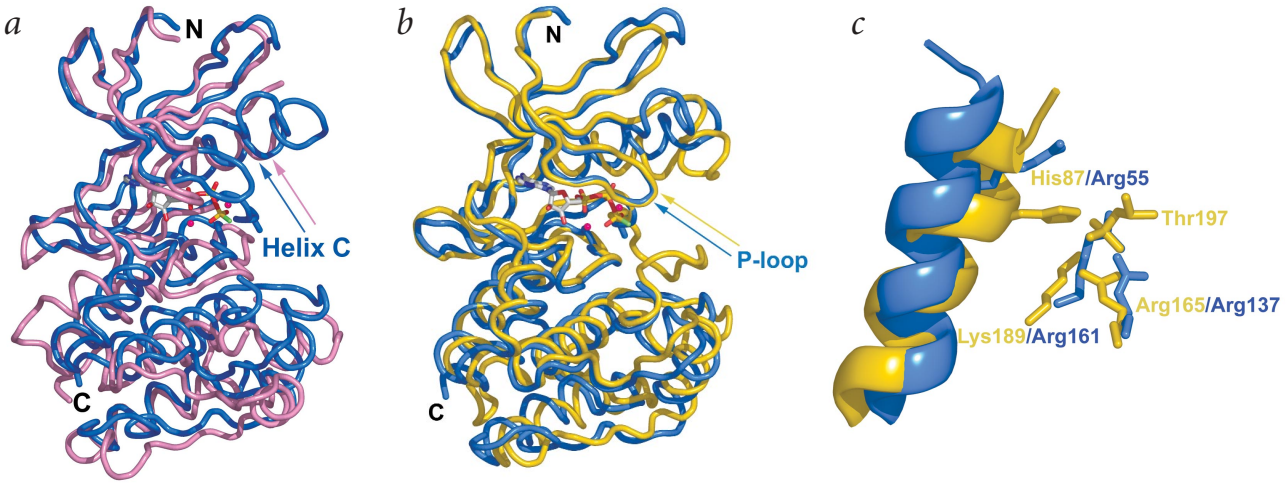


Fig. 3 Similarity of the PknB kinase domain and activated eukaryotic Ser/Thr kinases. Superposition of PknB-I–ATP-γ-S (blue) with ***a***, rat CaMK1 (magenta, PDB entry 1A06)[14] and ***b***, mouse cAPK–ADP–AlF_3–peptide substrate complex[13] (yellow, PDB entry 1L3R). The superpositions involved the N-terminal lobe of the kinase domain (residues 8–99 for PknB-I, residues 17–101 for CaMK1 and residues 40–127 for cAPK). The CaMK1 structure represents an inactive conformation, and the cAPK complex structure represents the active, closed conformation. The PknB main chain generally resembles the 'on' state of cAPK. ***c***, Incomplete binding site for the phosphorylated activation loop in PknB-I–ATP-γ-S (blue) compared to cAPK–ADP–AlF_3–peptide substrate[13] (yellow). The shift in PknB-I helix C positions Arg55 3.0 Å from the homologous His87 in the cAPK pThr-binding site.

otic Ser/Thr kinases. Despite sequence identity of <27% (Fig. 1*b*), the r.m.s. deviation for equivalent Cα atoms in PknB-I and seven different structures of eukaryotic Ser/Thr kinases[12–14,19–22] ranges between 1.36 and 1.72 Å. The closest match was to activated cAPK in complex with a transition-state analog[13]. The eukaryotic Ser/Thr kinases switch between activated ('on') and inactivated ('off') states, and the switch is mediated by a set of coupled conformational transitions[11]. The PknB-I–ATP-γ-S crystal structure displays extensive local and global similarities to the structure of cAPK in the activated state (Figs. 2*c*, 3).

The activating conformational changes in eukaryotic Ser/Thr kinases are linked to changes at regulatory sites that include phospho-acceptors in the activation loop. Among the six phosphorylated residues in PknB-I (Ser166, Ser169, Thr171, Thr173, Thr294 and Ser295), four occurred in the predicted activation loop (Fig. 1*a*). Ser166 is homologous to the phospho-acceptor in the activation loop of cAPK, and Thr173 is homologous to the phospho-acceptor in CaMK1 (Fig. 1*b*). It remains to be shown directly that phosphorylation of PknB at one or more of these residues controls kinase activity. Nonetheless, the PknB-I structure contains basic residues that correspond to the phospho-Thr (pThr) binding site in cAPK (Fig 3*c*). The presence of these conserved basic residues in PknB supports the idea that phosphorylation of the activation loop may play a regulatory role.

For a structure so similar to the 'on' state, the disorder of the PknB-I activation loop and the position of helix C (Fig. 3*b*) were unusual. Hyperphosphorylation or heterogeneous phosphorylation of the PknB-I activation loop may block formation of a unique, docked structure. In addition, two of the three putative phosphoryl-binding residues in the PknB-I–ATP-γ-S structure are positioned at the binding site (Fig. 3*c*), because helix C is tipped away from the active site. Because of this shift, the binding sites for the phosphorylated activation loop and the protein substrate[11,13] are not fully formed.

Conserved surfaces of PknB-I

To explore the basis for regulation and protein substrate recognition, we mapped, on the surface of PknB-I, the sequence variation in the kinase domains of 22 bacterial PknB orthologs (Fig. 4*a*). The greater divergence of the bacterial Ser/Thr kinases[23,24] compared to those in eukaryotes facilitates the interpretation of conservation patterns. Notably, the conserved sites in PknB orthologs map precisely to the ATP-binding cleft, the tips of the loops adjacent to the γ-phosphate and the surface analogous to the peptide substrate–binding groove identified in cAPK[13] (Fig 4*a*). The conserved surface extends beyond the peptide contact surface identified in the complex of cAPK with the transition-state analog[13], supporting the prediction that the protein substrate–binding site is larger than that envisioned previously. Consistent with roles in conferring protein-substrate specificity[11], intermediate conservation occurs in the activation loop and in the putative substrate-binding groove of the 11 *M. tuberculosis* Ser/Thr kinases (Fig 4*b*). The residues most conserved in these paralogs surround the ATP-binding cleft, consistent with common nucleotide-binding and phospho-acceptor recognition functions.

Unexpected conservation also occurs in a surface site on the 'backside' of the N-terminal lobe (Fig. 4*a*, right). This site is buried in a broad intermolecular interface (1,045 $Å^2$ buried surface area per molecule) that includes the invariant Leu33 and the contact site for Leu33 at the C terminus of helix C in the adjacent molecule (related by noncrystallographic, two-fold symmetry) in the crystal. A conserved, dyad-symmetric network in the interface encompasses an intermolecular ion pair between Arg10 and Asp76, which makes an intermolecular hydrophobic contact with Leu33 and an intramolecular hydrogen bond with Tyr11. Such strict surface conservation would be unusual in the absence of functional selection. A possible function is suggested by the observation that analogous, allosteric, autoinhibitory interactions control the position of helix C in several eukaryotic protein kinases[11,22,25]. The analogous intermolecular interaction in the PknB-I crystals may account for the unexpected conformation of helix C and recapitulate a functional regulatory contact.

Structural conservation across kingdoms

Our results provide both chemical and structural evidence suggesting that the activated state of Ser/Thr kinases is universal in the prokaryotic and eukaryotic kingdoms. The activated

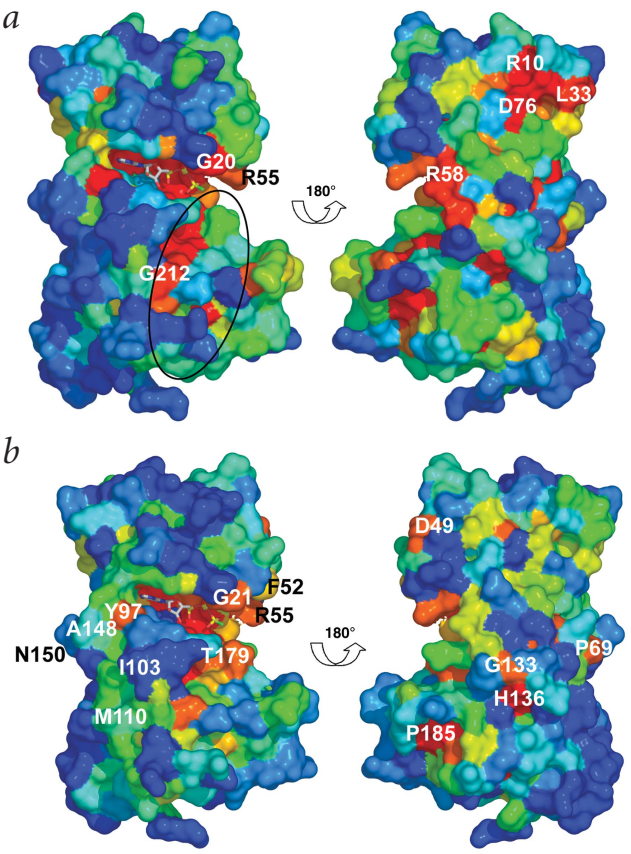
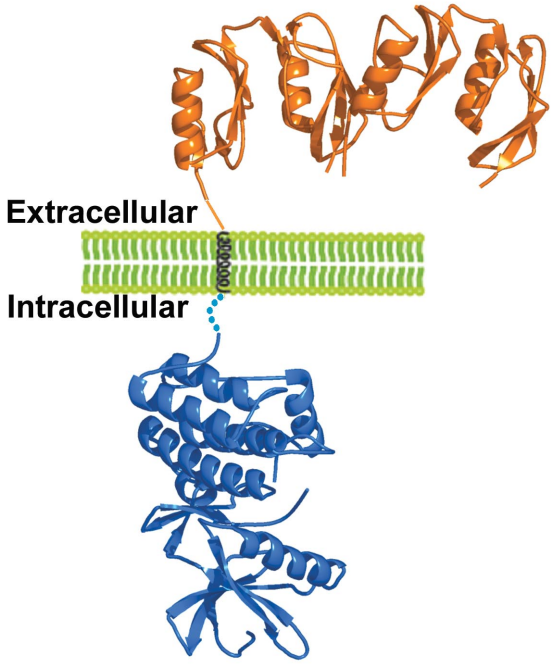


Fig. 4 Conserved surface sites in PknB. ***a***, Sequence variation in 22 PknB orthologs mapped onto surface of PknB-I. Orthologs were predicted to contain at least one PASTA domain fused through a single transmembrane helix to a kinase domain. The solvent-accessible surface is colored according to sequence conservation, from invariant (red) to variable (dark blue). A 180° rotation relates the left and right images. The most conserved residues include the ATP-binding site and Met20, Phe52, Arg55 and Thr179 neighboring the γ-phosphate in position to contact the Thr or Ser phospho-acceptor. Highly conserved residues also occur in a diagonal groove on the C-terminal domain (ellipse, left) analogous to the peptide substrate–binding cleft of cAPK[13] and in the continuation of this groove beyond the γ-phosphate (right). These residues presumably contact the protein substrate. The PknB residues (Val98, Ile103, Glu107, Met110, Arg114, Glu117, Ile146, Ala148 and Asn150) analogous to the residues of MAP kinase p38 that contact substrate docking sites[37] form a poorly conserved surface patch distinct from the putative peptide-binding cleft. The lack of conservation of this surface suggests that the peptide-binding groove plays a primary role in substrate selection. Conserved residues Leu33, Arg10 and Asp76 in the N-terminal lobe (top right) mediate intermolecular contacts in the crystals and correspond to a surface that transduces regulatory signals in eukaryotic protein kinases, such as c-Src and TGF-β receptor[11,22,25]. ***b***, The PknB-I surface colored from red (high) to blue (low) corresponding to the level of sequence similarity of PknB to the mycobacterial Ser/Thr protein kinases. These kinase domains show 27% (PknG) to 45% (PknA) sequence identity with PknB. Conserved surfaces include the ATP-binding site. Intermediate conservation occurs in the putative substrate-binding cleft and the potential regulatory site on the back of the N-terminal domain (top right). Conserved patches on the side opposite the active site (right) include buried ion pairs — for example, Arg262–Glu186 — that play apparent structural roles in the C-terminal domain.

conformations of Ser/Thr protein kinases may be preserved for historical or functional reasons. The hypothesis that the bacterial Ser/Thr kinases arose by a process that included horizontal gene transfer from ancient, ancestral eukaryotes[23,24] is supported at the structural level by the conformational similarity of PknB-I–ATP-γ-S with eukaryotic homologs. In addition, three of the *M. tuberculosis* Ser/Thr protein kinases contain predicted motifs, β-propeller (PknD), TPR (PknG) or PDZ (PknK) domains[4], that also are largely restricted to eukaryotes. The juxtaposition of two predominantly eukaryotic motifs within individual bacterial proteins may reflect a eukaryotic origin of the bacterial Ser/Thr kinases. The activated state of PknB also may be conserved as a result of functional interactions with eukaryotic proteins. Such interactions would require, however, that the kinase domain function outside the bacterial cell (like the *Yersinia pestis* protein kinase, YpkA/YopO[26]). Extracellular function of PknB would require a mechanism (such as bacterial lysis) to expose the intracellular kinase domain. Although intracellular proteins of *M. tuberculosis* are known to be antigenic[27], such a 'Trojan horse' function for the Ser/Thr kinases has not been tested. In contrast, the occurrence within *M. tuberculosis* of an intracellular Ser/Thr protein phosphatase[3] and proteins containing predicted pThr–binding motifs[23,28] supports the idea that the Ser/Thr kinases signal within the bacterium.



Functional implications

The combined crystallographic and computational analysis supports a model of full-length PknB (Fig. 5) containing an intracellular kinase domain joined by a segment of ~70 residues (including the transmembrane helix) to four extracellular PASTA domains. Tandem PASTA domains have been proposed to target the *S. pneumoniae* Pbp2x to sites of cell growth[16]. Strikingly, *pbpA*, encoding a homolog of Pbp2x (21% sequence identity and 36% similarity over 395 residues), occurs in the *pknB* operon in the *M. tuberculosis* genome[3]. PbpA contains a predicted transmembrane helix followed by a transpeptidase domain, but it lacks C-terminal PASTA domains. Protein fusions

Fig. 5 Model of full-length PknB. The extracellular domain of PknB is modeled as four linked PASTA domains taken from the structure of *S. pneumoniae* Pbp2x[16]. The phosphorylated linker between the kinase domain and the transmembrane helix is apparently flexible, as judged by the lack of corresponding electron density in the PknB-I crystal structure, the sensitivity of this region to limited proteolysis (data not shown) and the enrichment of proline and polar residues compared to well-folded sequences. The fusion of an extracellular targeting domain to an intracellular kinase through a disordered linker suggests that oligomerization or localization control the activity and substrate specificity of PknB.

Table 1 Data collection, phasing and refinement statistics for the PknB-I–ATP-γ-S complex

Data collection and phasing		
	λ_1 (1.1000 Å)	λ_2 (0.9798 Å)
Resolution (Å)	110–3.0	110–3.0
Completeness (%)	97.4	98.2
R_{merge} (%)[1,2]	4.3 (17.1)	6.5 (32.4)
$<I / \sigma I>$[2]	14.3 (7.4)	15.1 (7.0)
Phasing power (acentric/centric)[3]	1.39/1.41	
Mean figure of merit[4]		
(after solvent flattening)	0.420 (0.804)	
Refinement		
Resolution (Å)	30–3.0	
Reflections	13,890	
R-factor / R_{free} (%)	23.65/29.59	
R.m.s. deviations from ideal		
Bonds (Å)	0.0098	
Angles (°)	1.66	
Average B-factor (Å^2)		
A molecule	53.43	
B molecule	51.23	
Main chain dihedral angles (%)		
Most favored	81.7	
Allowed	17.3	

[1]$R_{merge} = \Sigma| I - <I> | / \Sigma I$, where I is intensity.
[2]Parentheses denote value for highest resolution shell.
[3]Phasing power = $[\Sigma_n|F_H|^2 / \Sigma_n|E|^2]^{1/2}$, where F_H is the calculated heavy atom scattering factor and E is lack of closure error.
[4]Mean figure of merit = $<|\Sigma_\alpha P(\alpha)e^{i\alpha} / \Sigma_\alpha P(\alpha)|>$, where α is the phase and $P(\alpha)$ is the phase probability distribution.
[5]$R_{cryst} = \Sigma|F_o - F_c| / \Sigma F_o$, where F_o is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude.

in one organism can sometimes imply functional or physical interactions between separate proteins in other organisms[29]. The occurrence of PASTA domains in Pbp2x provides a 'Rosetta stone'[29] that predicts a functional linkage between the coexpressed mycobacterial PknB and PbpA.

How does the fusion of a targeting domain to a kinase domain mediate a regulatory signal? A striking feature of the PknB-I–ATP-γ-S crystal structure is the apparent disorder of the C-terminal 21 residues in the linker to the transmembrane helix. The phosphorylated linker apparently does not form a rigid domain that could transmit conformational signals from the extracellular domain. Instead, the activity of the isolated intracellular kinase domain (Fig. 1*c*) and the possible targeting function of the extracellular PASTA domains[18] support a signaling mechanism in which localization or oligomerization controls PknB kinase activity. Such a role for assembly in determining substrate specificity resembles the mechanism of activation of several classes of eukaryotic transmembrane signaling protein kinases.

Methods

Amplification and cloning. Expression vectors for PknB fragments were created by sticky-end PCR[30] from the genomic DNA of *M. tuberculosis* HRv37. The annealed PCR products were ligated with pET28b (Novagen) digested with *Nde*I and *Xho*I, and transformed into *E. coli* XL1-Blue (Stratagene). Constructs were confirmed by DNA sequencing.

Purification of PknB-I for activity assays and phosphate mapping. BL21-CodonPlus(DE3) cells (Stratagene) harboring the PknB-I expression plasmid were cultured in Terrific Broth containing 100 μg ml^{-1} kanamycin and 34 μg ml^{-1} chloramphenicol, and grown at 37 °C until the OD_{600} was 0.3. The culture was shaken at 25 °C for 30 min, 250 μM IPTG (final concentration) was added and growth was continued for 4 h. The cells were harvested by centrifugation and resuspended in 0.3 M NaCl, 50 mM NaH_2PO_4, pH 8.0, 10 mM imidazole and 0.5 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP). After sonication on ice and centrifugation, the clarified cell lysate was loaded onto a 5 ml immobilized metal affinity chromatography (IMAC) column (Amersham-Pharmacia) equilibrated with 0.1 M $NiSO_4$. The protein was eluted with a 10–250 mM imidazole gradient in 100 ml and dialyzed into 75 mM NaCl, 20 mM Tris-HCl, pH 7.0, and 0.5 mM TCEP.

Mass spectrometry. The molecular masses of His-tagged PknB were measured on a Bruker/Agilent Esquire-LC ion trap mass spectrometer. The spectra were deconvoluted using the Protein Reconstruct algorithm from BioAnalyst (Applied Biosystems). For phosphate mapping, 1 mg purified protein was digested 1 h with 0.1 mg bovine trypsin (Sigma) in 100 mM ammonium bicarbonate, pH 8.5, at room temperature. The tryptic digest was diluted 50-fold with 0.1% (v/v) trifluoroacetic acid (TFA), and 0.5 μl of this solution was combined with an equal volume of matrix solution (10 g l^{-1} solution of α-cyano-4-hydroxycinnamic acid in 50% acetonitrile/50% 0.1% aqueous TFA)) and allowed to dry on the MALDI target. Reflector mass spectra were measured with an Applied Biosystems 4700 Proteomics Analyzer (a TOF/TOF instrument with a MALDI ion source)[31]. Some spectra were acquired with a Bruker Reflex III instrument. Reflector spectra were calibrated with trypsin autolysis peaks and predicted tryptic-fragment masses of PknB, typically giving mass accuracy exceeding 5–10 p.p.m. on the TOF/TOF. MS/MS CID spectra were acquired manually with the TOF/TOF on selected peptides, using air as the collision gas. Default calibration of the mass scale for MS/MS spectra provided fragment masses accurate to <0.1 Da.

Kinase activity assays. His-tagged PknB-I (9 μg) was incubated with or without 50 μg of myelin basic protein (MBP; Sigma) in 20 mM PIPES, pH 7.2, 1 mM $MgCl_2$ and 1 mM $MnCl_2$. The reactions were started by adding 1 μCi of [γ-^{32}P]ATP (800 Ci mmol^{-1} and ~10 mCi ml^{-1}; ICN) at room temperature. The reactions were stopped with SDS-PAGE loading buffer and separated on a 12% Tris-glycine polyacrylamide gel (Fisher Scientific). Proteins were electroblotted onto a PVDF membrane (Fisher Scientific) and radioactivity was quantified with a Molecular Dynamics Typhoon 8600 phosphoimager.

Expression and purification of SeMet PknB-I. SeMet PknB-I was produced as described[32] in BL21-CodonPlus(DE3) cells harboring the expression plasmid. After induction with 250 μM IPTG for 6 h at 25 °C, the cells were harvested and lysed in 50 mM Na-cacodylate, pH 7.5, 0.3 M NaCl, 10 mM imidazole, 5 mM L-Met, 0.2 mM EDTA and 0.5 mM TCEP, and the protein was purified by IMAC, as described above. Thrombin (200 U l^{-1} of culture, Sigma) was added to the eluted protein, which was dialyzed at 4 °C into lysis buffer. The thrombin-cleaved products were separated by IMAC, and the flow-through was purified by size-exclusion chromatography using a HiLoad 26/60 Superdex 75 column (Amersham Pharmacia) equilibrated with 40 mM Tris-HCl, pH 7, 100 mM NaCl, 5 mM L-Met, 0.2 mM EDTA and 0.5 mM TCEP. Peaks corresponding to monomer and dimer were observed. The PknB-I monomer peak was loaded onto a 5 ml HiTrap Q Sepharose ion-exchange column (Amersham Pharmacia) and eluted with a 0.1–1 M NaCl gradient in 100 ml. The purified protein was dialyzed into 20 mM Tris, pH 7.0, 75 mM NaCl, 5 mM L-Met, 0.2 mM EDTA and 0.5 mM TCEP.

Crystallization. PknB-I was crystallized by the vapor diffusion method. Hanging drops containing a 1:1 mixture of 10 mg ml^{-1} protein and 12.5% (w/v) PEG 4000, 0.1 M Tris-HCl (w/v) , pH 8.0, 0.2 M sodium acetate, 10 mM spermine, 3 mM $MgCl_2$ and 3 mM ATP-γ-S were equilibrated at 4 °C against 12.5% (w/v) PEG 4000, 0.1 M Tris-HCl (w/v), pH 8.0, and 0.2 M sodium acetate. For data collection, crystals were equilibrated for three days against a well solution containing 50% (w/v) PEG 6000, mounted on a loop and frozen in liquid N_2.

Structure determination. X-ray data were collected at the Lawrence Berkeley National Laboratory Advanced Light Source beamline 8.3.1 at 100 K at the Se peak and low-remote energies. The crystals had the symmetry of space group $P2_1$, with unit cell parameters a = 63.10 Å, b = 50.17 Å, c = 111.07 Å and β = 96.17°. Data were processed to 3 Å resolution using the Elves automation program (J.M. Holton and T.A., unpublished results), which found 11 of the expected 14 Se sites and produced a clear electron density map after solvent flattening. The initial model was built with O[33] using the structures of CK2 (PDB entry 1JWH)[12] and CaMK1 (1A06)[14] as guides. After locating the methionines in the map, an additional Se site was added and refined with Elves, and model building continued with the resulting improved electron density map.

Structure refinement. The model was refined initially at 3.0 Å resolution with REFMAC[34], using a single overall *B*-value and tight non-crystallographic symmetry restraints between the two monomers in the asymmetric unit. The R_{free} was calculated with a random 5% of the data. Rebuilding was carried out with O[33]. Subsequent models were refined with CNS[35]. Initial stages used torsional refinement. Refinement of restrained, individual *B*-values improved the *R*-factor and R_{free} values. The final model contains two PknB molecules (residues 1–163 or 3–163 and 179–286), two ATP-γ-S molecules and four Mg^{2+} ions per asymmetric unit. The stereochemistry of the model inspected using PROCHECK[36] scored better overall (*G*-factor = 0.16) than most structures refined at 3.0 Å resolution. No residues were in the disallowed region of the Ramachandran plot. Figures were drawn with PyMOL (*http://www.pymol.org*).

Coordinates. The coordinates and structure factors were deposited in the Protein Data Bank (accession number 1MRU).

Acknowledgments

We thank D.S. King for protein molecular weight measurements, K.F. Medzihradszky for LC/MS/MS measurements, J. Holton and E. Skordalakes for help with structure determination, C.A. Settineri for help with the deconvoluted protein mass spectrum, and L. Gay and M. Good for stimulating discussions. We are indebted to J. Kuriyan for encouragement and for pointing out the importance of the regulatory interaction sites in PknB-I. K. Strenge and S. Shieh provided technical support. We are indebted to T. Terwilliger and the TB Structural Genomics Consortium for support. The work was supported by a grant from the NIH. The Advanced Light Source Beamline 8.3.1 was funded by the NSF, the University of California and Henry Wheeler.

Competing interests statement

The authors declare that they have no competing financial interests.

Received 17 September, 2002; accepted 6 January, 2002.

1. Cheek, S., Zhang, H., & Grishin, N.V. Sequence and structure classification of kinases. *J. Mol. Biol.* **320**, 855–881 (2002).
2. Manning, G., Whyte, D.B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912–1934 (2002).
3. Cole, S.T. *et al.* Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544 (2000).
4. Av-Gay, Y. & Everett, M. The eukaryotic-like Ser/Thr protein kinases of *Mycobacterium tuberculosis*. *Trends Microbiol.* **8**, 238–244 (2000).
5. Andersen, P. TB vaccines: progress and problems. *Trends Immunol.* **22**, 160–168 (2001).
6. Ono-Saito, N., Niki, I. & Hidaka, H. H-series protein kinase inhibitors and potential clinical applications. *Pharmacol. Ther.* **82**, 123–131 (1999).
7. Drews, S.J., Hung, F. & Av-Gay, Y. A protein kinase inhibitor as an antimicrobial agent. *FEMS Microbiol. Lett.* **205**, 369–374 (2001).
8. Chaba, R., Raje, M. & Chakrobarti, P.K. Evidence that a eukaryotic-type serine/threonine protein kinase from *Mycobacterium tuberculosis* regulates morphological changes associated with cell division. *Eur. J. Biochem.* **269**, 1078–1085 (2002).
9. Av-Gay, Y., Jamil, S. & Drews, S.J. Expression and characterization of the *Mycobacterium tuberculosis* serine/threonine protein kinase PknB. *Infection Immunol.* **67**, 5676–5682 (1999).
10. Koul. A. *et al.* Serine/threonine protein kinases PknF and PknG of *Mycobacterium tuberculosis*: characterization and localization. *Microbiology* **147**, 2307–2314 (2001).
11. Huse, M. & J. Kuriyan. The conformational plasticity of protein kinases. *Cell* **109**, 275–282 (2002).
12. Niefind, K., Guerra, B., Ermakowa, I. & Issinger, O.G. Crystal structure of human protein kinase Ck2: Insights into basic properties of the Ck2 holoenzyme. *EMBO J.* **20**, 5320–5331 (2001).
13. Madhusudan, Akamine, P., Xuong, N.-H. & Taylor, S.S. Crystal structure of a transition state mimic of the catalytic subunit of cAMP-dependent protein kinase. *Nat. Struct. Biol.* **9**, 273–277 (2002).
14. Goldberg J., Nairn, A.C., & Kuriyan, J. Structural basis for the autoinhibition of calcium/calmodulin-dependent protein kinase I. *Cell* **84**, 875–887 (1996).
15. Kelley, L.A., MacCallum, R.M. & Sternberg, M.J. Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.* **299**, 499–520 (2000).
16. Gordon, E., Mouz, N., Duee, E. & Dideberg, O. The crystal structure of the penicillin-binding protein 2x from *Streptococcus pneumoniae* and its acyl-enzyme form: implication in drug resistance. *J. Mol. Biol.* **299**, 477–485 (2000).
17. Marchler-Bauer, A. *et al.* CDD: a database of conserved domain alignments with links to domain three-dimensional structure. *Nucleic Acids Res.* **30**, 281–283 (2002).
18. Yeats, C. Finn, R.D. & Bateman, A. The PASTA domain: a β-lactam binding domain. *Trends Biochem. Sci.* **9**, 438–40 (2002).
19 Schulze-Gahmen, U., *et al.* Multiple modes of ligand recognition: crystal structures of cyclin-dependent protein kinase 2 in complex with ATP and two inhibitors, olomoucine and isopentenyladenine. *Proteins* **22**, 378–391 (1995).
20. Xie, X. *et al.* Crystal structure of JNK3: a kinase implicated in neuronal apoptosis. *Structure* **6**, 983–991 (1998).
21. Zhang F., Strand A., Robbins D., Cobb M.H., & Goldsmith E.J. Atomic structure of the MAP kinase ERK2 at 2.3 Å resolution. *Nature* **367**, 704–711 (1994).
22. Xu, W., Doshi, A., Lei, M., Eck, M.J. & Harrison, S.C. Crystal structures of c-Src reveal new features of its autoinhibitory mechanism. *Mol. Cell* **3**, 629–638 (1999).
23. Leonard, C.J., Aravind, L. & Koonin, E.V. Novel families of putative protein kinases in bacteria and archaea: evolution of the "eukaryotic" protein kinase superfamily. *Genome Res.* **8**, 1038–1047 (1998).
24. Han, G. & Zhang, C.-C. On the origin of Ser/Thr kinases in a prokaryote. *FEMS Microbiol. Lett.* **200**, 79–84 (2001).
25. Young, M.A., Gonfloni, S., Superti-Furga, G., Roux, B. & Kuriyan, J. Dynamic coupling between the SH2 and SH3 domains of c-Src and Hck underlies their inactivation by C-terminal tyrosine phosphorylation. *Cell* **105**, 115–126 (2001).
26. Juris S.J., Rudolph, A.E., Huddler, D., Orth, K. & Dixon, J.E. A distinctive role for the Yersinia protein kinase: actin binding, kinase activation, and cytoskeleton disruption. *Proc. Natl. Acad. Sci. USA* **15**, 9431–9436 (2000).
27. Havlir, D.V. *et al.* Human immune response to *Mycobacterium tuberculosis* antigens. *Infection Immunol.* **59**, 665–670 (1991).
28. Durocher, D. & Jackson, S.P. The FHA domain. *FEBS Let.* **513**, 58–66 (2002).
29. Marcotte, E.M. *et al.* Detecting protein function and protein–protein interactions from genome sequences. *Science* **285**, 751–753 (1999).
30. Pham, K., LaForge, K.S. & Kreek, M.J. Sticky-end PCR: new method for subcloning. *Biotechniques* **25**, 206–208 (1998).
31. Bienvenut, W.V. *et al.* Matrix-assisted laser desorption/ionization-tandem mass spectrometry with high resolution and sensitivity for identification and characterization of proteins. *Proteomics* **2**, 868–876 (2002).
32. Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L. & Clardy, J. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* **229**, 105–124 (1993).
33. Jones T.A., Zou J.Y., Cowan S.W. & Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
34. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
35. Brunger A.T., *et al.* Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
36. Laskowski, R.A, MacArthur, M.W., Moss, D.S. & Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structure. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
37. Chang, C.-I., Xu, B., Akella, R. Cobb, M.H. & Goldsmith, E.J. Crystal structures of MAP kinase p38 complexed to the docking sites on its nuclear substrate MEF2A and activator MKK3b. *Mol. Cell* **9**, 1241–1249 (2002).