A TPR scaffold couples signal detection to OdhI phosphorylation in metabolic control by the protein kinase PknG

Running title: Regulatory mechanism of PknG from *C. glutamicum*

María-Natalia Lisa [a,b,#], Adrià Sogues [a,*], Nathalie Barilone [a,*], Meike Baumgart [c], Magdalena Gil [d,*], Martín Graña [e], Rosario Durán [d], Ricardo M. Biondi [e,*], Marco Bellinzoni [a], Michael Bott [c] and Pedro M. Alzari [a,#]

[a] Unité de Microbiologie Structurale, Institut Pasteur, CNRS UMR 3528, Université de Paris, 25 Rue du Docteur Roux, 75724 Paris Cedex 15, France.

[b] Instituto de Biología Molecular y Celular de Rosario (IBR, CONICET-UNR), Ocampo y Esmeralda, Rosario 2000, Argentina.

[c] IBG-1: Biotechnology, Institute of Bio- and Geosciences, Forschungszentrum Jülich, Wilhelm-Johnen-Strasse, 52425 Jülich, Germany

[d] Unidad de Bioquímica y Proteómica Analíticas, Instituto de Investigaciones Biológicas Clemente Estable e Institut Pasteur de Montevideo, Mataojo 2020, Montevideo 11400, Uruguay.

[e] Unidad de Bioinformática, Institut Pasteur de Montevideo, Mataojo 2020, Montevideo 11400, Uruguay

[e] Research Group PhosphoSites, Medizinische Klinik 1, Universitätsklinicum Frankfurt, Theodor-Stern-Kai 7, 60590 Frankfurt, Germany

[*] Current addresses: AS: Structural & Molecular Microbiology, VIB-VUB Centre for Structural Biology, Brussels, 1050, Belgium; NB: Unité Récepteurs-Canaux, Institut Pasteur, Paris cedex 15, France; MG: Dynamics of Host-Pathogen Interactions Unit, Institut Pasteur, CNRS UMR3691,

75724 Paris, France; RB: Instituto de Investigación en Biomedicina de Buenos Aires (IBioBA)-CONICET-Partner Institute of the Max Planck Society, Buenos Aires, Argentina.

[#] To whom correspondence should be addressed: MNL (lisa@ibr-conicet.gov.ar) or PMA (pedro.alzari@pasteur.fr).

ABSTRACT

Signal transduction is essential for bacteria to adapt to changing environmental conditions. Among many forms of post-translational modifications, reversible protein phosphorylation has evolved as a ubiquitous molecular mechanism of protein regulation in response to specific stimuli. The Ser/Thr protein kinase PknG modulates the fate of intracellular glutamate by controlling the phosphorylation status of the 2-oxoglutarate dehydrogenase regulator OdhI, a function that is conserved among diverse actinobacteria. PknG has a modular organization characterized by the presence of regulatory domains surrounding the catalytic domain. Here we present an investigation through *in vivo* experiments as well as biochemical and structural methods of the molecular bases of the regulation of PknG from *C. glutamicum* (*Cg*PknG), in the light of previous knowledge available for the kinase from *M. tuberculosis* (*Mtb*PknG). We found that OdhI phosphorylation by *Cg*PknG is regulated by a conserved mechanism that depends on a C-terminal domain composed of tetratricopeptide repeats (TPR) essential for metabolic homeostasis. Furthermore, we identified a conserved structural motif that physically connects the TPR domain and a flexible N-terminal extension of the kinase that is involved in docking interactions with OdhI. Based on our results and previous reports, we propose a model in which the TPR domain of PknG couples signal detection to the specific phosphorylation of OdhI. Overall, the available data indicate that conserved PknG domains in distant actinobacteria retain their roles in kinase regulation in response to nutrient availability.

IMPORTANCE

Bacteria control the metabolic processes by which they obtain nutrients and energy in order to adapt to the environment. In this way, the metabolic characteristics of a microorganism determine its ecological role and its usefulness in industrial processes. Here, we use genetic, biochemical, and structural approaches to study a key component in a system that regulates glutamate production in *C. glutamicum*, a species that is used for the industrial production of amino acids. We elucidated molecular mechanisms involved in metabolic control in *C. glutamicum*, which are conserved in related pathogenic bacteria. The findings have broader significance for diverse actinobacteria, including microorganisms that cause disease as well as environmental species used to produce billions of dollars of amino acids and antibiotics every year.

INTRODUCTION

The large and ancient bacterial phylum *Actinobacteria* comprises species with very diverse lifestyles and physiological adaptations, including soil inhabitants, pathogens as well as plant or animal commensals (1). The eukaryotic-like Ser/Thr protein kinase (STPK) PknG and its FHA (ForkHead-Associated) substrate OdhI (Oxoglutarate dehydrogenase Inhibitor) are at the core of a conserved signal transduction pathway that modulates central metabolism in distant actinobacteria. Both in *Corynebacterium glutamicum*, a soil bacterium used for the industrial production of amino acids, as well as in the pathogen *Mycobacterium tuberculosis*, PknG modulates the 2-oxoglutarate dehydrogenase activity in the Krebs cycle (2–4) by controlling the phosphorylation status of the regulator OdhI (called GarA in the genus *Mycobacterium*) (2–5). Biochemical studies have demonstrated that unphosphorylated OdhI/GarA inhibits the E1 component (OdhA) of the 2-oxoglutarate dehydrogenase complex whereas this inhibition is

relieved by OdhI/GarA phosphorylation by PknG (2–4, 6, 7). Moreover, early studies for the two species revealed that *pknG* disruption leads to an accumulation of intracellular glutamate (2, 8), pointing out that PknG acts by promoting catabolism at the expense of 2-oxoglutarate usage in nitrogen assimilation. On top of this, it was recently found that PknG senses the availability of amino-donor amino acids to control metabolism and virulence in *M. tuberculosis* (9–11). These findings have received much attention (10), as a deeper understanding of PknG regulation can be instrumental for downstream applications in the biotech and pharmaceutical areas.

PknG has a unique modular organization characterized by the ubiquitous presence of a flexible N-terminal segment and a C-terminal domain composed of tetratricopeptide repeats (TPR) flanking the kinase catalytic core (12–14). An additional rubredoxin (Rdx)-like domain occurs immediately adjacent to the catalytic core in PknG from mycobacteria and most other actinobacteria but not in corynebacteria (2). Previous structural studies of PknG have focused on the protein from *M. tuberculosis* (*Mtb*PknG) (12, 13). We have shown that the N-terminal extension and the TPR domain of *Mtb*PknG regulate the selectivity for GarA without significantly affecting the intrinsic kinase activity, whereas the Rdx domain downregulates catalysis by limiting access to a profound substrate-binding site (13). Rdx domains are known to transmit redox stimuli and, consistent with this, evidence has been reported pointing out that perturbations of the metal center in PknG lead to alterations of the kinase activity (15). However, relatively little is known about the regulatory mechanisms of PknG isoforms that lack an Rdx domain.

The gene *pknG* is found within a conserved operon that contains two other genes, *glnX* and *glnH*, which encode a putative transmembrane protein and a putative glutamine-binding lipoprotein, respectively (2, 11). The observation that disruption of any of those genes in *C.*

glutamicum led to a similar phenotype consisting of a growth defect in medium containing glutamine as the sole carbon source (2) suggested a common role of the protein products in metabolic homeostasis. Supporting this early hypothesis, evidence has been recently reported that, in mycobacteria, PknG and GlnX are functionally linked and that GlnH specifically binds amino acids able to stimulate GarA phosphorylation by the kinase (11). This led to the proposal that GlnH senses amino acid availability within the bacterial periplasm and transmits this information across the membrane *via* GlnX to activate PknG by protein-protein interactions (11). Most interesting, a PknG truncation mutant lacking the TPR domain failed to restore the growth defect of a *pknG*-disrupted mycobacterial strain, suggesting that this domain, often involved in protein-protein interactions (16), mediates molecular associations required for the kinase function (11).

To investigate the conservation of mechanisms involved in the regulation of PknG, we studied the kinase isoform from *C. glutamicum* (*Cg*PknG), which is devoid of an Rdx domain. We provide evidence that the C-terminal region of *Cg*PknG, bearing the TPR domain, is crucial for the efficient phosphorylation of OdhI and for the kinase function in metabolic homeostasis. Moreover, our results point out that the recruitment of the FHA substrate is regulated by a conserved phosphorylation-dependent mechanism regardless of the absence of an Rdx domain. Finally, by comparing three high-resolution crystal structures of *Cg*PknG and an available structure of *Mtb*PknG (12), we identified a conserved motif able to link the N-terminal extension and the TPR domain. Interestingly, the evidence suggests that the Rdx domain, absent in corynebacteria, and the TPR domain would constitute independent regulatory mechanisms. Overall, our results indicate that common PknG domains in distant actinobacteria share similar functions in kinase regulation, linking PknG to the control of central metabolism in response to nutrient availability.

RESULTS

The C-terminal region of *Cg*PknG is required for phosphorylation events that modulate metabolism

To investigate the domains required for the function of *Cg*PknG, we employed a previously characterized *C. glutamicum* Δ*pknG* mutant strain able to grow in rich medium but unable to grow in medium containing glutamine as the sole carbon source (2). *Cg*PknG domain boundaries were defined based on a previous characterization of *Mtb*PknG (13) (47% amino acid identity), and plasmids were designed for the expression of *Cg*PknG truncation mutants (Fig. 1A) in *C. glutamicum* Δ*pknG* using the endogenous gene promoter. All strains grew normally in medium containing glucose and all versions of the kinase were detected by Western-blot (Fig. S1).

In contrast to wild type *Cg*PknG, the mutant *Cg*PknG$_{K205A}$, which harbors a substitution of the invariant catalytic lysine, did not complement the growth defect of *C. glutamicum* Δ*pknG* on glutamine (Fig. 1B), indicating that the kinase activity is required for protein function. Additionally, a *Cg*PknG truncation mutant lacking residues 433-822 was unable to restore bacterial growth on glutamine, pointing out, in agreement with previous results for *Mtb*PknG (11), that the region of *Cg*PknG located C-terminally to the catalytic core is necessary for the kinase role in the control of metabolism. Moreover, *Mtb*PknG did complement the growth defect of *C. glutamicum* Δ*pknG*, stressing the functional conservation between distant kinase isoforms. A *Cg*PknG deletion mutant devoid of residues 1-130 failed to restore the growth of *C. glutamicum* Δ*pknG* on glutamine, however the low amount detected of this kinase version precludes drawing conclusions from this

observation. Together, these results support a conserved requirement of the C-terminal region of PknG for phosphorylation events that modulate metabolism in response to amino acid availability.

A conserved phosphorylation-dependent mechanism for substrate recruitment

To investigate the molecular mechanisms of metabolic control by the kinase activity of *Cg*PknG, we first tested the ability of recombinant *Cg*PknG to phosphorylate OdhI and GarA *in vitro*. *Cg*PknG phosphorylated OdhI and GarA to a similar extent (Fig. 2A), confirming the ability of *Cg*PknG to phosphorylate the FHA substrate and evidencing that structural differences between OdhI and GarA (4, 17), either in the FHA domain or in the N-terminal phosphorylatable region, do not influence the kinase activity. Moreover, *Cg*PknG phosphorylated GarA in the same peptide as *Mtb*PknG (3) (Fig. S2), equivalent to the OdhI peptide phosphorylated by *Cg*PknG (2).

The N-terminal segment of *Mtb*PknG contains auto-phosphorylation sites (Thr23, Thr32, Thr63 and Thr64) (3) (Fig. 1A and Fig. S3) that act as essential anchoring points for the recruitment of GarA by interacting with the pThr-binding FHA domain of the regulator (3, 13). Despite the crucial role of the N-terminal extension of the kinase in substrate selectivity, its primary structure is poorly conserved. Therefore, to determine whether or not the role of the kinase N-terminal segment in the recruitment of the FHA substrate is conserved in spite of sequence divergence, we first investigated the auto-phosphorylation of *Cg*PknG. Despite no phosphorylation was detected in the purified recombinant protein, four phosphorylation sites (Thr14, Thr68, Thr92 and Thr93) were identified by mass spectrometry within the N-terminal extension of *Cg*PknG after incubating the kinase with ATP and Mn(II) (Fig. S3 and Fig. S4).

Next, we studied the ability of *Cg*PknG to phosphorylate a substrate lacking an FHA domain, using for this the previously reported 17-mer SDEVTV**ETTS**VFRADFL peptide (13) centered around the phosphorylatable ETTS motif that is conserved among OdhI/GarA homologs (2). The kinase activity of *Cg*PknG varied linearly with the concentration of the 17-mer peptide up to 1 mM, indicating a high K_M (> 1 mM) and the slope providing a measure of the catalytic efficiency (k_{cat}/K_M) of (9.0 ± 0.4) 10^{-3} pmol μM^{-2} min^{-1} for this substrate (Fig. 2B). By comparison, the phosphorylation of OdhI by *Cg*PknG was approximately 3-fold higher than for the 17-mer peptide even though a *ca.* 15-fold lower concentration of OdhI was used (Fig. 2A and Fig. 2C), indicating a *ca.* 45-fold higher activity towards OdhI due to the FHA domain acting as a kinase docking site.

Finally, we tested the kinase activity of a *Cg*PknG deletion mutant lacking residues 1-129 and 434-822. *Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$ displayed a *ca.* 7-fold lower activity against OdhI compared to the full-length enzyme, whereas phosphorylation of the 17-mer substrate was unaffected (Fig. 2A and Fig. 2C). These results indicate that neither residues 1-129 within the N-terminal extension, nor the TPR domain of *Cg*PknG had an effect on the intrinsic kinase activity, supporting previous evidence for *Mtb*PknG (13) that both regions contribute to stabilize the enzyme-FHA substrate complex.

Overall, our results indicate that diverse PknG isoforms recruit the FHA substrate OdhI (or GarA) *via* a conserved phosphorylation-dependent mechanism.

A conserved overall topology

To investigate the structural basis of the regulation of a PknG isoform lacking an Rdx domain, we solved a high-resolution crystal structure of *Cg*PknG$_{\Delta N\text{-}t}$ (see below) in complex with the non-

hydrolysable ATP analog AMP-PNP (Table 1). The final atomic model contains two copies of *Cg*PknG within the asymmetric unit, encompassing residues 123-799 and 125-798, respectively, including a short fragment of the N-terminal segment (hence the name *Cg*PknG$_{\Delta N\text{-}t}$), the kinase catalytic core and the TPR domain (Fig. 3 and Fig. S5A). The protein is monomeric, consistent with analytical ultracentrifugation that did not provide evidence in favor of *Cg*PknG dimerization (Fig. S5B), similar to previous results for *Mtb*PknG (13). Additionally, *mFo–DFc* sigma-A-weighted electron density maps clearly revealed the bound nucleotide and two Mg(II) atoms at the active site of each *Cg*PknG molecule. Notably, even though we used full-length *Cg*PknG in our crystallization assays, we found no evidence for residues 1-122 in electron density maps. Edman degradation experiments revealed that the N-terminal residue of crystallized *Cg*PknG was Val123, suggesting that the kinase N-terminal segment was partially degraded during crystal growth and that, as similarly reported for *Mtb*PknG (12), it is probably unstructured in most of its length.

*Cg*PknG and *Mtb*PknG (12) share the same overall fold and topology, except for the absence of a regulatory Rdx domain in *Cg*PknG that leads to a more accessible active site (Fig. 3A). As expected, kinase domain residues or motifs involved in contacts with the Rdx domain in *Mtb*PknG (12, 13) adopt distinct conformations in *Cg*PknG (Fig. 3B). Residue Trp188 in *Cg*PknG (equivalent to Trp164 in *Mtb*PknG), located in the β_2 strand and adjacent to the G-rich loop, interacts with the N-terminal segment. The loop connecting strands β4 and β5 (loop β4-β5) is found in *Cg*PknG in close association with the kinase N-lobe, with residue Val246 (His223 in *Mtb*PknG) buried within a pocket and residues Asp243 and Arg245 in contact with the strand β0. Besides, the helix αC does not interact with strands β4 and β5 and its C-terminal tip is displaced, in *Cg*PknG compared to *Mtb*PknG, towards the kinase activation loop.

Regardless of these differences, nucleotide binding within the active site of *Cg*PknG parallels the previous description for *Mtb*PknG (13) (Fig. 3C), consistent with a conserved set of residues within the ATP binding site region of the kinase. Also similar to *Mtb*PknG (12, 13), most functionally important and conserved motifs in the active site of *Cg*PknG exhibit conformations compatible with a standard eukaryotic protein kinase active state, and the activation loop is stabilized in an open and extended conformation, permissive for substrate binding in the absence of phosphorylation (Fig. 3D). Nevertheless, *Cg*PknG residue Glu222 is found away from the catalytic Lys205, pointing out of the active site due to an outward conformation of the helix αC, as previously reported for *Mtb*PknG (12, 13).

Compared to *Mtb*PknG, *Cg*PknG contains an additional motif (residues 604-661) in the TPR domain, adjacent to the catalytic core (Fig. 1 and Fig. 3A). However, a *Cg*PknG truncation mutant lacking residues 604-661 did complement the growth defect of *C. glutamicum* Δ*pknG* on glutamine, suggesting that this motif is not crucial for the kinase function.

A conserved motif connects the N-terminal segment and the TPR domain

The TPR domain of *Mtb*PknG influences the FHA substrate selectivity and we have previously proposed that this depends on the stabilization of a β-hairpin in the N-terminal extension of the kinase (13). In spite of sequence divergence, this secondary structure motif is conserved in *Cg*PknG (Fig. 4A). Both in *Cg*PknG and *Mtb*PknG the N-terminal β-hairpin is stabilized by interactions with the catalytic core and the linker between this and the TPR domain (linker C-T, see also Figs. 1A and 3A). Notably, the linker C-T simultaneously contacts the N-terminal segment, the catalytic core and the TPR domain of the kinase. To explore the significance of such interactions, we solved the

high-resolution crystal structures of the truncation mutant *Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$ in two different isoforms (Table 1). According to the electron density maps, the N-terminal β-hairpin was not stabilized in any of the structures of *Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$ (Fig. 4B), suggesting that this motif is responsive to the C-terminal region of the kinase. These results indicate that the linker C-T physically connects the conserved N- and C-terminal regions flanking the kinase catalytic core.

DISCUSSION

The phosphorylation-dependent stabilization of enzyme-substrate complexes is a widespread mechanism among STPKs that enables the efficient phosphorylation of specific cellular targets (18). PknG controls metabolism in corynebacteria and mycobacteria by modulating the phosphorylation status of the FHA regulator OdhI (or GarA) (2, 9), a task that requires the N-terminal extension of the kinase. Despite the relatively high sequence divergence of this segment, it has a roughly conserved distribution of charged amino acids, Pro and Gly residues in diverse species (Fig. S3), and comprises auto-phosphorylation sites both in *Cg*PknG and in *Mtb*PknG (3) (Fig. S3 and Fig. S4). The N-terminal extension of PknG is dispensable for the phosphorylation of a surrogate peptide lacking an FHA domain (Fig. 2C and (13)) and, conversely, the presence of the FHA domain in OdhI or GarA enables a much more efficient phosphorylation by full-length PknG (Fig. 2A, Fig. 2C and (13)). Overall, our results support a conserved phosphorylation-dependent mechanism for the recruitment of the FHA substrate *via* the kinase N-terminal extension.

Kinase domain motifs that play regulatory roles in eukaryotic protein kinases (ePKs) adopt different conformations in PknG isoforms depending on the presence or the absence of an Rdx domain. In *Cg*PknG the loop β4-β5 fills the pocket formed by the β-sheet in the kinase N-lobe,

246 whereas this loop is exposed to the solvent in *Mtb*PknG (12, 13) (Fig. 3B). The pocket and the
247 motifs that may fill it (*i.e.*, the N-lobe cap) lay on top of the catalytic Lys and are features
248 associated with the regulation of ePKs (19). Besides, the helix αC, an important regulatory motif in
249 ePKs (20, 21), is displaced in *Cg*PknG towards the kinase activation loop when compared to
250 *Mtb*PknG (12, 13) (Fig. 3B). Consistent with previous findings for ePKs (22), the crystal structures
251 of both *Cg*PknG and *Mtb*PknG (12, 13) exhibit relatively high B-factors for the loop β3-αC and the
252 N-terminal end of the helix αC, indicating that this motif is highly dynamic. Interestingly, while the
253 Rdx domain in *Mtb*PknG restraints the position of the helix αC by interacting with the loop β3-αC
254 (12, 13), the position adopted by the helix αC in *Cg*PknG generates a pocket that is reminiscent of
255 the PIF-pocket in AGC kinases (22, 23) (Fig. S6). However, irrespective of the structural differences
256 noted between *Cg*PknG and *Mtb*PknG (13), in both kinase isoforms the ATP phosphates are
257 properly positioned in the active site despite the absence of a salt bridge between the conserved
258 Glu in the helix αC and the catalytic Lys, while other conserved catalytically relevant motifs exhibit
259 conformations compatible with an ePK active state (20) (Fig. 3C and Fig. 3D). Thus far, there is no
260 evidence revealing regulatory mechanisms that depend exclusively on motifs within the kinase
261 catalytic domain. The Rdx module of *Mtb*PknG (absent in *Cg*PknG) remains the sole regulatory
262 element known to modulate the intrinsic activity of PknG (13, 15). It is worth noting that Rdx-
263 mediated regulation appears to act independently of the modulation of substrate specificity by
264 FHA-mediated docking interactions.

265 As the assembly of new domain combinations into complex proteins is linked to speciation
266 and segregation into distinct phylogenetic groups (24, 25), we performed a phylogenetic analysis
267 of PknG orthologs to seek for hints about the PknG-Rdx association (Fig. S7). In line with such

notion, PknG orthologs, distinguished by their unique domain organization, are broadly distributed within *Actinobacteria* and also mostly restricted to this bacterial phylum. A homologue of *Mtb*PknG is, however, found in *Ktedonobacter racemifer*. This Gram-positive spore-forming bacterium belongs to *Chloroflexi* and grows in filamentous colonies similarly to a number of actinobacteria (26). *Chloroflexi* is an ancient phylum proposed to be at or very close to the root of the bacterial phylogenetic tree (27). Besides, a readily detectable homologue of PknG from *K. racemifer* is that from *Calothrix* sp. from the ancient phylum *Cyanobacteria*. The fact that both of these PknG homologues harbor an Rdx domain (defined by the presence of a PknG_rubred Pfam PF16919 domain or two CxxCG motifs) suggests that such domain architecture either preceded the evolution of *Actinobacteria*, being then differentially lost in some lineages, or that the gene of an Rdx-containing PknG homolog was horizontally transferred to *Chloroflexi* and *Cyanobacteria*. We favor the former, more parsimonious hypothesis because several non-actinobacterial ancient sequences include an Rdx domain whereas the genus *Corynebacterium* lacks the module. It remains enigmatic why the Rdx domain was lost in evolution in this genus.

The overall topology of PknG is conserved irrespective of the presence or the absence of an Rdx domain (Fig. 3A). The relative position of the TPR and the catalytic domain of *Cg*PknG is similar to that of *Mtb*PknG (12). Compared to the mycobacterial isoform, *Cg*PknG contains an additional, intriguing motif (residues 604-661) in the TPR domain (Fig. 1A and Fig. 3A) that increases its interface with the catalytic core. However, according to our *in vivo* tests, such motif is not essential for the role of *Cg*PknG in metabolic homeostasis (Fig. 1B). In contrast, the C-terminal region of *Cg*PknG (residues 433-822) was required for complementing the *C. glutamicum* Δ*pknG* mutant strain (Fig. 1B), replicating previous results for *Mtb*PknG (11) and pointing to a conserved

role of the TPR domain in signal transduction. Notably, in *Cg*PknG as in *Mtb*PknG the linker C-T bridges the N-terminal segment and the TPR domain (Fig. 4A and (12)), both regions involved in the regulation of the kinase selectivity for the FHA substrate (13). The linker C-T is stabilized by conserved interactions with residues along the concave surface of the TPR domain (Fig. 3A and (12)). According to a recent proposal (11), this surface might constitute a binding site for GlnX, so that the transduction of extracellular stimuli would imply a conformational change of the linker C-T from its position in the free form of the kinase.

Taking together the available evidence, we propose that the TPR domain of PknG functions as a localization scaffold that, by mediating an interaction between the kinase and the transmembrane protein GlnX, transduces a signal about amino acid availability detected by GlnH (Fig. 5). The PknG-GlnX interaction likely produces a conformational change in the linker C-T, which couples the detection of the signal to the specific recruitment of the FHA substrate *via* the N-terminal segment of the kinase. Given that the specific set of multidomain proteins in genomes sets constraint on the topology of pathways and networks that carry out regulatory processes (28), the co-occurrence of *pknG*, *glnX*, *glnH* and *odhI* in actinobacteria (2, 11), together with the functional links found among the respective proteins, therefore suggests the conservation of the associated molecular mechanism that evolved in this phylum to control metabolism in response to nutrient availability.

MATERIALS AND METHODS

Complementation assays

All plasmids used in this study are listed in Table 2. Plasmids for complementation assays were generated by Genscript (Leiden, The Netherlands) from the previously described pEKEx2-*pknG*$_{St}$ template plasmid (2). The *C. glutamicum* Δ*pknG* strain (2) was transformed with each of the plasmids carrying the relevant *pknG* variants, or with the pEKEx2 vector lacking an insert, as previously described (29). Then, strains were first streaked on BHI medium (BD BBL). In each case, single colonies were subsequently plated both on CGXII-glucose (30) and CGXII-glutamine. The CGXII-glutamine broth is a modified version of medium CGXII that is devoid of $(NH_4)_2SO_4$, urea and glucose and is supplemented with 100 mM glutamine. Plates were cultivated for 3 days at 30°C.

Detection of PknG versions by Western-blot

Transformed *C. glutamicum* Δ*pknG* (2) cells were grown at 30°C in BHI broth (BD BBL) with agitation until reaching 3 units of optical density at 600 nm. Protein expression was then induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM, and the incubation was continued for 20 hours at 30°C. Cells were then harvested by centrifugation. Cell pellets were suspended in lysis buffer (50 mM Bis-Tris, 75 mM 6-aminocaproic acid, 1 mM $MgSO_4$, 1 U/ml benzonase, cOmplete EDTA-free protease inhibitor cocktail (Roche) in the amount specified by the manufacturer, pH 7.4) and disrupted by using 0.1 mm glass beads and a homogenizer (Precellys 24) operated at 4°C. 120-250 µg of crude extracts were run in a pre-cast 4-12% SDS-PAGE gradient gel (Biorad) and then electro-transferred onto a 0.2 µm nitrocellulose membrane (Biorad). Blocking was performed with PBS buffer supplemented with 3% w/v BSA and 0.05% v/v Tween 20. The membrane was subsequently incubated with an anti-Strep antibody (StrepMAB-Classic, IBA Lifesciences) at 4°C overnight. After 3 washes with TBS-Tween buffer (10

mM Tris-HCl, 150 mM NaCl, 0.05 v/v Tween 20, pH 8.0) for 5 minutes each, the membrane was incubated with a secondary anti-Rabbit horseradish peroxidase conjugated antibody (GE Healthcare) for 45 minutes at room temperature. Finally, the membrane was washed 3 times with TBS-Tween buffer for 5 min each, revealed with the horseradish peroxidase substrate (Immobilon Forte, Millipore) and imaged using the ChemiDoc MP Imaging System (Biorad).

Construction of plasmids for the production of recombinant proteins

Plasmids pET28a-*Cg*PknG and pET28a-*Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$ (Table 2) were constructed by PCR amplification of *pknG* regions 1-822 and 130-433, respectively, from *C. glutamicum* ATCC 13032 genomic DNA, followed by digestion and ligation of the amplification products into the *Nde*I and *Sac*I sites in plasmid pET28a (Novagen). The oligonucleotides employed were the following (the TEV protease cleavage sites are underlined):

*Cg*PknG-F: ATTATCATATGGAGAATCTTTATTTTCAGGGCATGAAGGATAATGAAGATTTCGATCC

*Cg*PknG-R: ATATTGAGCTCTCACTAGAACCAACTCAGTGGCCGCACGGC

Δ1-129,Δ434-822-F:

TATATTATCATATGGAGAATCTTTATTTTCAGGGCGTTGCTGATGGCATGGTGGAATTG

Δ1-129,Δ434-822-R: TATATATTGAGCTCTCATTTGCCGTCGCGGACTGCCAAAATTTC

Protein production and purification

Wild type *Cg*PknG and the truncation mutant *Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$ were both overproduced in *E. coli* BL21(DE3) cells cultivated in LB broth. Wild type *Cg*PknG was produced for 18 h at 15°C with 500 µM IPTG, whereas *Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$ was expressed after 3 h of induction at 30°C with

250 µM IPTG. Both of these proteins were then purified following the same protocol. *E. coli* cells were harvested by centrifugation, re-suspended in lysis buffer (25 mM Hepes, 500 mM NaCl, 20% v/v glycerol, 20 mM imidazole, pH 8.0), supplemented with cOmplete EDTA-free protease inhibitor cocktail (Roche) as specified by the manufacturer and sonicated. After clarification by centrifugation, the supernatant was loaded onto a HisTrap HP column (GE Healthcare) and the His-tagged protein was purified applying a linear imidazole gradient (20–500 mM) in lysis buffer. The His6-tag was later removed by over-night incubation at 4°C with 0.2 equivalents of His6-tagged TEV protease, followed by separation on a Ni-NTA agarose column (Qiagen). The protein was then further purified by size-exclusion chromatography on a 16/600 Superdex 200 column (GE Healthcare) equilibrated in either 50 mM Tris-HCl, 250 mM NaCl, 5% glycerol, pH 8.0 (wild type *Cg*PknG) or 25 mM Hepes, 150 mM NaCl, 5% glycerol, pH 7.5 (*Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$), using a flow rate of 0.5-1 ml/min. Fractions corresponding to *Cg*PknG or *Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$, as confirmed by SDS-PAGE, were pooled and concentrated, flash-frozen in liquid nitrogen and stored at -80°C.

GarA and OdhI were prepared as previously described (17, 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 assays

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 $MnCl_2$, 100 µM [γ-^{32}P]ATP (5-50 cpm/pmol), and 330 µM 17-mer peptide or 25 µM OdhI (or GarA) as substrate. The enzyme concentration in the assays was 0.7-3 µM and

0.15-0.9 µM when using the 17-mer peptide or OdhI (or GarA) as substrates, respectively. The kinase reactions were started by the addition of 4 µl [γ-^{32}P]ATP-Mn^{+2} and were performed at room temperature. The reactions were stopped by the addition of phosphoric acid and 4 µl of each reaction were spotted on P81 phosphocellulose papers (Whatman) using the epMotion 5070 (Eppendorf) workstation. The papers were washed in 0.01% phosphoric acid, dried, then measured and analyzed using the PhosphorImager (FLA-9000 Starion, Fujifilm). Each reaction was performed in duplicates (<5% variation). In all cases, specific activity values were derived from reactions performed employing three different enzyme concentrations within the indicated ranges (<10% variation), verifying a linear dependence of activity with the enzyme concentration. Each assay was performed at least twice. The proportion of 17-mer peptide or OdhI (or GarA) phosphorylated in the reactions was lower than 10% and 30%, respectively. OdhI (or GarA) phosphorylation was verified to be linear in time up to 50% of its initial concentration. Under the experimental conditions employed to test phosphorylation of the 17-mer peptide or OdhI (or GarA), *Cg*PknG auto-phosphorylation represented less than 5% of the total signal. The measured signal was at least five times higher than the measure on the background.

The 17-mer peptide SDEVTV**ETTS**VFRADFL was produced with a purity >98% by Thermo Fisher Scientific.

Mass spectrometry analysis

The kinase activity of *Cg*PknG was assayed using GarA as substrate and the molecular mass of unphosphorylated and phosphorylated GarA was then determined as previously described (15).

398 *Cg*PknG was incubated with ATP and $MnCl_2$ and then sequentially digested with trypsin and
399 endoproteinase GluC for 3 h at 37°C. The resulting peptides were separated using a nano-HPLC
400 system (Proxeon EasynLC, Thermo) with a reverse-phase column (easy C18 column, 3 µm; 75 µm
401 ID×10 cm; Proxeon, Thermo) and eluted with a 0.1% v/v formic acid (in water) to acetonitrile
402 gradient (0–40% acetonitrile in 50 min; flow 300 nl/min). Online MS analysis was carried out in a
403 linear ion trap instrument (LTQ Velos,Thermo) in data dependent acquisition mode (full scan
404 followed by MS/MS of the top 5 peaks in each segment, using a dynamic exclusion list). Raw
405 MS/MS spectra were extracted by the Proteome Discoverer software package (v.1.3.0.339,
406 Thermo) and submitted to Sequest for database searching against sequences from *E. coli* (strain
407 K12) downloaded from Uniprot consortium (April, 2021) to which the sequence of PknG from *C.*
408 *glutamicum* was added. Search parameters were set as follows: peptide tolerance: 0.8 Da; MS/MS
409 tolerance: 0.8 Da; methionine oxidation and Ser/Thr/Tyr phosphorylation as the allowed variable
410 modifications. PhosphoRS was used as phospho-site localization tool (32). We considered a
411 positive phospho-site identification when more than one spectrum for the phospho-peptide was
412 obtained, pRS probability was >95% and manual inspection of the MS/MS spectra showed at least
413 two confirmatory fragment ions.
414
415 **Crystallization and data collection**
416 Crystallization screenings were carried out using the sitting-drop vapor diffusion method and
417 a Mosquito nanolitre-dispensing crystallization robot (TTP Labtech). Crystals of *Cg*PknG$_{\Delta N\text{-}t}$ + AMP-
418 PNP and *Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$ + AMP-PNP grew after 20-30 and 7-10 days, respectively, from 10
419 mg/ml protein solutions supplemented with 5 mM AMP-PNP, by mixing 200 nl of protein solution

and 200 nl of mother liquor (100 mM Tris-HCl, 17% w/v PEG 20 k, 100 mM $MgCl_2$, pH 8.5; and 100 mM Tris-HCl, 27-30% w/v PEG 4 k, 200 mM $MgCl_2$, pH 8.8, respectively), at 18°C. Single crystals reaching a size of $(100\ \mu m)^3$ were cryprotected in mother liquor containing 25% glycerol and flash-frozen in liquid nitrogen. X-ray diffraction data were collected at the synchrotron beamlines Proxima 2 (Synchrotron Soleil, Saint-Aubin, France) and ID29 (ESRF, Grenoble, France) at 100 K. Employed wavelengths were 0.9801 Å and 0.97625 Å for *Cg*PknG$_{\Delta N\text{-}t}$ + AMP-PNP and *Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$ + AMP-PNP crystals, respectively. The diffraction data were processed using XDS (33) and scaled with Aimless (34) from the CCP4 program suite.

Structure determination and refinement

The crystal structure of *Cg*PknG$_{\Delta N\text{-}t}$ + AMP-PNP was solved by molecular replacement using the program Phaser (35) and the atomic coordinates of *Mtb*PknG residues 138-405 from PDB 4Y0X (13) and residues 406-750 from PDB 2PZI (12) as search probes. The structures of *Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$ + AMP-PNP were solved similarly by using the atomic coordinates of *Cg*PknG$_{\Delta N\text{-}t}$ residues 165-425. Ligand molecules were manually placed in *mFo–DFc* sigma-A-weighted electron density maps employing *Coot* (36). Models were refined through iterative cycles of manual model building with *Coot* and reciprocal space refinement with phenix.refine (37). The final models were validated through the MolProbity server (38). In each case, the final model contained more than 97% of residues within favored regions of the Ramachandran plot, with no outliers. Figures were generated and rendered with Pymol 1.8.x. (Schrödinger, LLC).

Edman degradation

442 The crystal employed to solve the structure of *Cg*PknG$_{\Delta N\text{-}t}$ was dissolved in water and Edman
443 degradation was performed by the Functional Genomics Center of Zurich
444 (https://fgcz.ch/omics_areas/prot/applications/protein-characterization.html). As a control, an
445 aliquot of recombinant *Cg*PknG as used in crystallization screenings was also analyzed, and the
446 sequence of the protein N-terminus resulted GMKDN, as expected.
447
448 **Analytical ultracentrifugation**
449 Sedimentation velocity experiments were carried out at 20°C in an XL-I analytical
450 ultracentrifuge (Beckman Coulter). Samples were spun using an An60Ti rotor and 12-mm double
451 sector epoxy centerpieces. The partial specific volume of *Cg*PknG (0.734 ml g^{-1}) was estimated
452 from their amino acid sequences using the software Sednterp. The same software was used to
453 estimate the buffer viscosity (η = 1.040 centipoises) and density (ρ = 1.010 g·ml^{-1}). *Cg*PknG (400 µl
454 at 1 mg/ml) was spun at 42,000 rpm, and absorbance profiles were recorded every five minutes.
455 Sedimentation coefficient distributions, c(s), were determined using the software Sedfit 14.1 (39).
456
457 **Database searches, alignments and phylogenetic analyses**
458 BLASTp searches (40) were conducted against complete protein sequences available at the
459 Integrated Microbial Genome (IMG; http://img.jgi.doe.gov) (41), performing a taxon sampling on
460 finished assembled genomes within the phyla *Cyanobacteria*, *Chloroflexi*, *Chlorobi*, *Fusobacteria*,
461 *Sinergistetes*, *Firmicutes*, *Tenericutes*, *Acidobacteria*, *Nitrospirae*, *Spirochaetes*, *Aquificae* and
462 *Thermotogae*, all in the vicinity of *Actinobacteria* in an updated tree of life (27). The sequence of
463 *Mtb*PknG was used as queries for searches to identify homologues in such genomes using an

464 expected inclusion threshold e-value < 1 e^{-20}. Once the existence of the domain combinations was
465 confirmed, we focused on 91 complete *Actinobacteria* genomes available from IMG (April 2021).
466 The final selection was preprocessed using PREQUAL (42) to mask non-homologous sequence
467 stretches. A CD-HIT (43) cut-off value of 90% pairwise identity was applied for the entire set of
468 sequences retrieved as described. The final set of 40 sequences was aligned with MAFFT (version
469 7.467) using the L-INS-I strategy (44) and columns with more than 90% gaps were removed with
470 TrimAL. The phylogenetic tree displayed in Fig. S7 was computed with IQ-TREE (version 1.6.12,
471 (45)) using ModelFinder (46) to select the evolutionary model and the ultrafast bootstrap method
472 (47) (options "-bb 1000 -alrt 1000"). The model selected with the Bayesian Information Criterion
473 was the evolutionary matrix EX_EHO (48) with empirical frequencies and four categories of free
474 rate (EX_EHO+F+R4).
475
476 **Data availability**
477 Atomic coordinates and structure factors have been deposited in the Protein Data Bank
478 under the accession codes 7mxb (*Cg*PknG$_{\Delta N\text{-}t}$ + AMP-PNP), 7mxj (*Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$ + AMP-PNP_1)
479 and 7mxk (*Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$ + AMP-PNP_2).
480
481 **ACKNOWLEDGEMENTS**
482 MNL received fellowships from the EMBO (European Molecular Biology Organization) and
483 the Fondation pour la Recherche Médicale (FRM, France). This work was funded by grants from
484 the Institut Pasteur, the CNRS (France), the Agence Nationale de la Recherche (ANR, France,
485 contract number ANR-09-BLAN-0400) and the European Commission Seventh Framework

Programme (contract HEALTH-F3-2011-260872). MNL acknowledges support from the Agencia I+D+I (grant PICT 2017-1932). RMB acknowledges support from DFG (grant BI 1044/12-1).

We thank Ahmed Haouz and Patrick Weber for their help with robot-driven crystallization screenings and Bertrand Raynal for his help with analytical ultracentrifugation experiments. We acknowledge William Shepard for his assistance in the collection of diffraction data at beamline Proxima 2 and Martin Cohen-Gonsaud for providing plasmid pET-15b-Tev-OdhI for OdhI production.

Author contributions: MNL designed experiments, prepared proteins, performed kinase activity assays, carried out crystallographic studies and structural analysis, analyzed data and wrote the paper; ASC performed complementation assays, analyzed data and wrote the paper; NB prepared plasmids pET28a-*Cg*PknG and pET28a-*Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$, optimized the production of recombinant proteins and performed analytical ultracentrifugation experiments; MGi carried out mass spectrometry analyses; MGa performed phylogenetic analysis; RD designed and performed mass spectrometry studies; RMB designed kinase activity assays and analyzed data; MBe, MBo and PMA designed research and analyzed data. All authors copy edited the paper.

REFERENCES

1. Barka EA, Vatsa P, Sanchez L, Gaveau-vaillant N, Jacquard C, Klenk H-P, Clément C, Ouhdouch Y, van Wezel GP. 2016. Taxonomy, Physiology, and Natural Products of Actinobacteria. Microbiol Mol Biol Rev 80:1–43.
2. Niebisch A, Kabus A, Schultz C, Weil B, Bott M. 2006. Corynebacterial protein kinase G controls 2-oxoglutarate dehydrogenase activity via the phosphorylation status of the OdhI

protein. J Biol Chem 281:12300–12307.

3. O'Hare HM, Durán R, Cerveñansky C, Bellinzoni M, Wehenkel AM, Pritsch O, Obal G, Baumgartner J, Vialaret J, Johnsson K, Alzari PM. 2008. Regulation of glutamate metabolism by protein kinases in mycobacteria. Mol Microbiol 70:1408–1423.

4. Nott TJ, Kelly G, Stach L, Li J, Westcott S, Patel D, Hunt DM, Howell S, Buxton RS, O'Hare HM, Smerdon SJ. 2009. An Intramolecular Switch Regulates Phosphoindependent FHA Domain Interactions in Mycobacterium tuberculosis. Sci Signal 2:ra12.

5. Ventura M, Rieck B, Boldrin F, Degiacomi G, Bellinzoni M, Barilone N, Alzaidi F, Alzari PM, Manganelli R, O'Hare HM. 2013. GarA is an essential regulator of metabolism in mycobacterium tuberculosis. Mol Microbiol 90:356–366.

6. Krawczyk S, Raasch K, Schultz C, Hoffelder M, Eggeling L, Bott M. 2010. The FHA domain of OdhI interacts with the carboxyterminal 2-oxoglutarate dehydrogenase domain of OdhA in Corynebacterium glutamicum. FEBS Lett 584:1463–1468.

7. Wagner T, André-Leroux G, Hindie V, Barilone N, Lisa M, Hoos S, Raynal B, Vulliez-Le Normand B, O'Hare HM, Bellinzoni M, Alzari PM. 2019. Structural insights into the functional versatility of an FHA domain protein in mycobacterial signaling. Sci Signal 12:eaav9504.

8. Cowley S, Ko M, Pick N, Chow R, Downing KJ, Gordhan BG, Betts JC, Mizrahi V, Smith DA, Stokes RW, Av-Gay Y. 2004. The Mycobacterium tuberculosis protein serine/threonine kinase PknG is linked to cellular glutamate/glutamine levels and is important for growth in vivo. Mol Microbiol 52:1691–1702.

9. Rieck B, Degiacomi G, Zimmermann M, Cascioferro A, Boldrin F, Lazar-Adler NR, Bottrill AR,

le Chevalier F, Frigui W, Bellinzoni M, Lisa M-N, Alzari PM, Nguyen L, Brosch R, Sauer U, Manganelli R, O'Hare HM. 2017. PknG senses amino acid availability to control metabolism and virulence of Mycobacterium tuberculosis. PLoS Pathog 13:e1006399.

10. York A. 2017. Bacterial physiology: An inside job on metabolism. Nat Rev Microbiol 15:383–383.

11. Bhattacharyya N, Nkumama IN, Newland-Smith Z, Lin LY, Yin W, Cullen RE, Griffiths JS, Jarvis AR, Price MJ, Chong PY, Wallis R, O'Hare HM. 2018. An aspartate-specific solute-binding protein regulates protein kinase G activity to control glutamate metabolism in mycobacteria. MBio 9:1–13.

12. Scherr N, Honnappa S, Kunz G, Mueller P, Jayachandran R, Winkler F, Pieters J, Steinmetz MO. 2007. Structural basis for the specific inhibition of protein kinase G, a virulence factor of Mycobacterium tuberculosis. Proc Natl Acad Sci U S A 104:12151–12156.

13. Lisa M-N, Gil M, André-Leroux G, Barilone N, Durán 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.

14. Reckel S, Hantschel O. 2015. Kinase Regulation in Mycobacterium tuberculosis: Variations on a Theme. Structure 23:975–976.

15. Gil M, Graña M, Schopfer FJ, Wagner T, Denicola A, Freeman BA, Alzari PM, Batthyány C, Durán R. 2013. Inhibition of Mycobacterium tuberculosis PknG by non-catalytic rubredoxin domain specific modification: Reaction of an electrophilic nitro-fatty acid with the Fe-S center. Free Radic Biol Med 65:150–161.

16. Zeytuni N, Zarivach R. 2012. Structural and functional discussion of the tetra-trico-peptide

repeat, a protein interaction module. Structure 20:397–405.

17. Barthe P, Roumestand C, Canova MJ, Kremer L, Hurard C, Molle V, Cohen-Gonsaud M. 2009. Dynamic and Structural Characterization of a Bacterial FHA Protein Reveals a New Autoinhibition Mechanism. Structure 17:568–578.

18. Jin J, Pawson T. 2012. Modular evolution of phosphorylation-based signalling systems. Philos Trans R Soc B Biol Sci 367:2540–2555.

19. Thompson EE, Kornev AP, Kannan N, Kim C, Ten Eyck LF, Taylor SS. 2009. Comparative surface geometry of the protein kinase family. Protein Sci 18:2016–2026.

20. Huse M, Kuriyan J. 2002. The conformational plasticity of protein kinases. Cell 109:275–282.

21. Kornev AP, Taylor SS. 2010. Defining the conserved internal architecture of a protein kinase. Biochim Biophys Acta - Proteins Proteomics 1804:440–444.

22. Hindie V, Stroba A, Zhang H, Lopez-Garcia LA, Idrissova L, Zeuzem S, Hirschberg D, Schaeffer F, Jørgensen TJD, Engel M, Alzari PM, Biondi RM. 2009. Structure and allosteric effects of low-molecular-weight activators on the protein kinase PDK1. Nat Chem Biol 5:758–764.

23. Leroux AE, Biondi RM. 2020. Renaissance of Allostery to Disrupt Protein Kinase Interactions. Trends Biochem Sci 45:27–41.

24. Vogel C, Bashton M, Kerrison ND, Chothia C, Teichmann SA. 2004. Structure, function and evolution of multidomain proteins. Curr Opin Struct Biol 14:208–216.

25. Han JH, Batey S, Nickson AA, Teichmann SA, Clarke J. 2007. The folding and evolution of multidomain proteins. Nat Rev Mol Cell Biol 8:319–330.

26. Cavaletti L, Monciardini P, Bamonte R, Schumann P, Ronde M, Sosio M, Donadio S. 2006. New lineage of filamentous, spore-forming, gram-positive bacteria from soil. Appl Environ

574 Microbiol 72:4360–4369.

575 27. Castelle CJ, Banfield JF. 2018. Major New Microbial Groups Expand Diversity and Alter our

576 Understanding of the Tree of Life. Cell 172:1181–1197.

577 28. Zmasek CM, Godzik A. 2012. This Déjà Vu Feeling-Analysis of Multidomain Protein Evolution

578 in Eukaryotic Genomes. PLoS Comput Biol 8.

579 29. van der Rest ME, Lange C, Molenaar D. 1999. A heat shock following electroporation

580 induces highly efficient transformation of Corynebacterium glutamicum with xenogeneic

581 plasmid DNA. Appl Microbiol Biotechnol 52:541–545.

582 30. Frunzke J, Engels V, Hasenbein S, Gätgens C, Bott M. 2008. Co-ordinated regulation of

583 gluconate catabolism and glucose uptake in Corynebacterium glutamicum by two

584 functionally equivalent transcriptional regulators, GntR1 and GntR2. Mol Microbiol 67:305–

585 322.

586 31. England P, Wehenkel A, Martins S, Hoos S, André-Leroux G, Villarino A, Alzari PM. 2009. The

587 FHA-containing protein GarA acts as a phosphorylation-dependent molecular switch in

588 mycobacterial signaling. FEBS Lett 583:301–307.

589 32. Taus T, Köcher T, Pichler P, Paschke C, Schmidt A, Henrich C, Mechtler K. 2011. Universal

590 and Confident Phosphorylation Site Localization Using phosphoRS. J Proteome Res 10:5354–

591 5362.

592 33. Kabsch W. 2010. XDS. Acta Crystallogr Sect D 66:125–132.

593 34. Evans PR, Murshudov GN. 2013. How good are my data and what is the resolution? Acta

594 Crystallogr Sect D Biol Crystallogr 69:1204–1214.

595 35. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser

crystallographic software. J Appl Crystallogr 40:658–674.

36. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. Acta Crystallogr Sect D Biol Crystallogr 66:486–501.

37. Afonine P V., Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr Sect D Biol Crystallogr 68:352–367.

38. Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall WB, Snoeyink J, Adams PD, Lovell SC, Richardson JS, Richardson DC. 2018. MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci.

39. Schuck P. 2000. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. Biophys J 78:1606–1619.

40. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res 25:3389–3402.

41. Markowitz VM, Chen IMA, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang J, Williams P, Huntemann M, Anderson I, Mavromatis K, Ivanova NN, Kyrpides NC. 2012. IMG: The integrated microbial genomes database and comparative analysis system. Nucleic Acids Res 40:D115–D122.

42. Whelan S, Irisarri I, Burki F. 2018. PREQUAL: detecting non-homologous characters in sets of unaligned homologous sequences. Bioinformatics 34:3929–3930.

618 43. Li W, Godzik A. 2006. Cd-hit: A fast program for clustering and comparing large sets of
619 protein or nucleotide sequences. Bioinformatics 22:1658–1659.
620 44. Katoh K, Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment
621 program. Brief Bioinform 9:286–298.
622 45. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear
623 R. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the
624 Genomic Era. Mol Biol Evol 37:1530–1534.
625 46. Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. 2017. ModelFinder:
626 Fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589.
627 47. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: Improving the
628 Ultrafast Bootstrap Approximation. Mol Biol Evol 35:518–522.
629 48. Le SQ, Gascuel O. 2010. Accounting for Solvent Accessibility and Secondary Structure in
630 Protein Phylogenetics Is Clearly Beneficial. Syst Biol 59:277–287.
631 49. Eikmanns BJ, Kleinertz E, Liebl W, Sahm H. 1991. A family of Corynebacterium
632 glutamicum/Escherichia coli shuttle vectors for cloning, controlled gene expression, and
633 promoter probing. Gene. 102:93-98.
634

635 **TABLES**

636 **Table 1. Data collection and refinement statistics.**

	*Cg*PknG$_{\Delta N\text{-}t}$	*Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$_1	*Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$_2
Data collection			
Space group	$P2_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	104.66 42.74 175.33	37.62 55.94 123.94	37.81 54.59 146.49
α, β, γ (°)	90.00 95.31 90.00	90.00 90.00 90.00	90.00 90.00 90.00
Resolution (Å)	46.68-2.20 (2.24-2.20)*	41.52-1.92 (1.97- 1.92)	48.83-1.99 (2.04-1.99)
R_{merge}	0.086 (0.514)	0.070 (0.746)	0.067 (0.654)
$I/\sigma I$	9.9 (2.0)	13.6 (1.9)	17.1 (2.3)
CC (1/2)	0.996 (0.675)	0.999 (0.805)	0.999 (0.751)
Completeness (%)	99.1 (90.8)	99.9 (100.0)	98.4 (84.9)
Redundancy	3.4 (2.6)	5.9 (5.6)	6.0 (5.2)
Refinement			
Resolution (Å)	43.02-2.20	41.52-1.92	43.77- 1.99
No. reflections	78,937	20,701	21,183
R_{work}/R_{free}	0.210/0.246	0.199/0.225	0.209/0.237
No. atoms			
Protein	10,369	2,190	2,255
Ligands	68	33	33
Solvent	910	237	157
Average *B*-factors			
Protein	37.82	32.46	47.67
Ligands	24.54	30.19	33.21
Solvent	38.11	36.72	42.30
R.m.s. deviations			
Bond lengths (Å)	0.002	0.003	0.003
Bond angles (°)	0.54	0.63	0.59
Ramachandran			
Favored (%)	98.17	97.85	97.57
Allowed (%)	1.83	2.15	2.43
Outliers (%)	0	0	0
PDB code	7mxb	7mxj	7mxk

637 *One protein crystal was employed for structure determination in each case. Values in

638 parentheses are for highest-resolution shell.

639 **Table 2. Plasmids used in this study.**

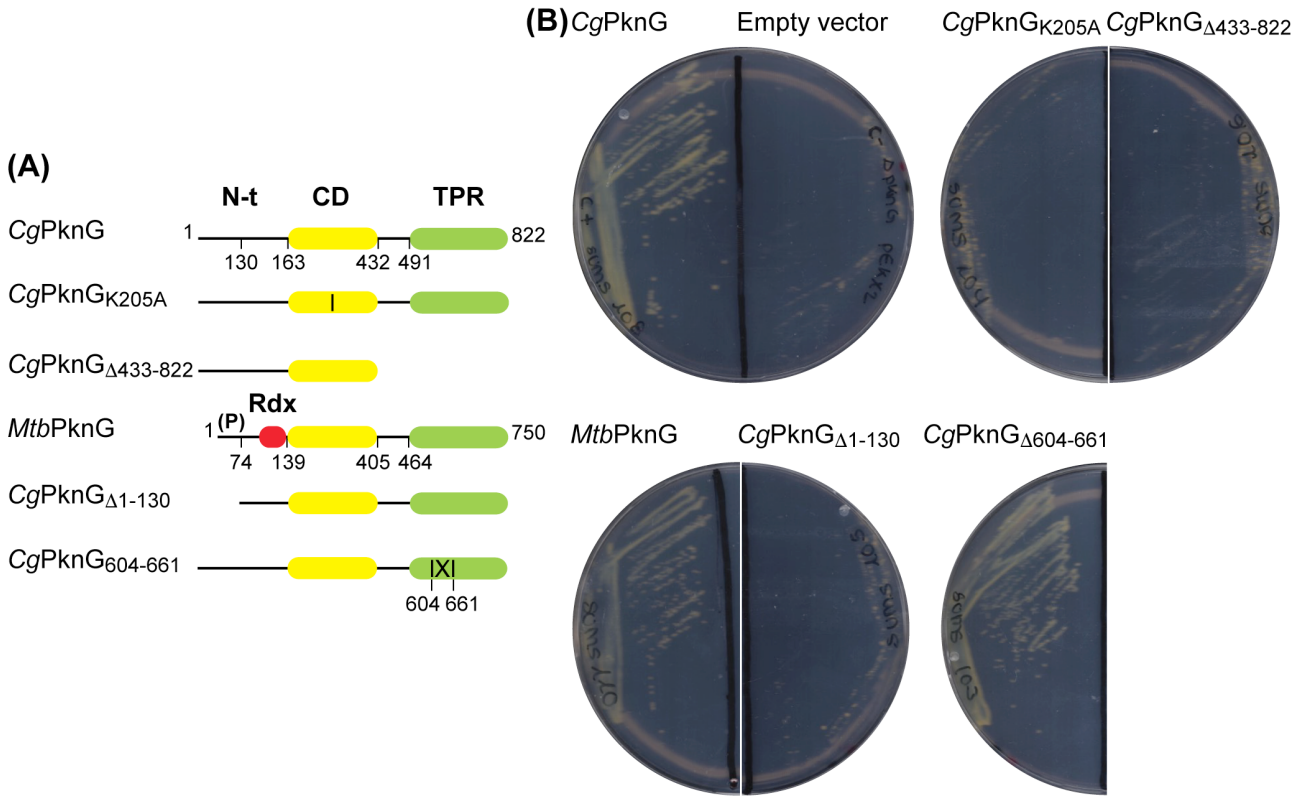
Plasmid	Description	Reference
pEKEx2	Kan^R. Allows the IPTG-inducible production of proteins in *C. glutamicum.*	(49)
pEKEx2-*pknG*$_{St}$	Kan^R, derived from pEKEx2. Designed for the production in *C. glutamicum* of C-terminally Strep-tagged full-length *Cg*PknG from the endogenous gene promoter.	(2)
pEKEx2-*Cg*PknG$_{K205A}$	Kan^R, derived from pEKEx2-*pknG*$_{St}$. Used for the production in *C. glutamicum* of C-terminally Strep-tagged full-length *Cg*PknG carrying substitution K205.	This work
pEKEx2-*Cg*PknG$_{\Delta 433-822}$	Kan^R, derived from pEKEx2-*pknG*$_{St}$. Used for the production in *C. glutamicum* of C-terminally Strep-tagged *Cg*PknG lacking residues 433-822.	This work
pEKEx2-*Mtb*PknG	Kan^R, derived from pEKEx2-*pknG*$_{St}$. Used for the production in *C. glutamicum* of C-terminally Strep-tagged full-length *Mtb*PknG.	This work
pEKEx2-*Cg*PknG$_{\Delta 1-130}$	Kan^R, derived from pEKEx2-*pknG*$_{St}$. Used for the production in *C. glutamicum* of C-terminally Strep-tagged *Cg*PknG lacking residues 1-130.	This work
pEKEx2-*Cg*PknG$_{\Delta 604-661}$	Kan^R, derived from pEKEx2-*pknG*$_{St}$. Used for the production in *C. glutamicum* of C-terminally Strep-tagged *Cg*PknG lacking residues 604-661.	This work
pET28a-*Cg*PknG	Kan^R, derived from pET28a. Used for the IPTG-inducible production in *E. coli* of N-terminally His6-tagged full-length *Cg*PknG.	This work
pET28a-*Cg*PknG$_{\Delta 1-129,\Delta 434-822}$	Kan^R, derived from pET28a. Used for the IPTG-inducible production in *E. coli* of N-terminally His6-tagged full-length *Cg*PknG lacking residues 1-129 and 434-822.	This work

640

641 **FIGURES**

642 **Figure 1**

643



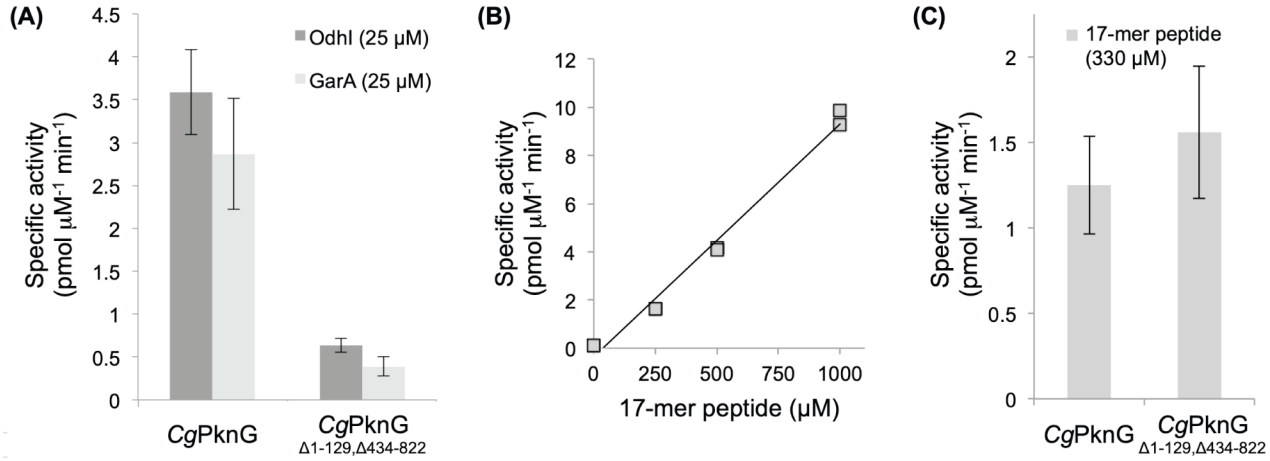
644

645

646 **Complementation of *C. glutamicum* Δ*pknG*** (2) **with different PknG versions.** (A) Schematic
647 representation of the kinase variants tested in complementation assays in this study. The
648 structured domains of the protein are shown as colored rectangles: the Rdx domain in red, the
649 catalytic domain (CD) in yellow, and the TPR domain in green. The vertical line in the CD of mutant
650 *Cg*PknG$_{K205A}$ represents the amino acid substitution. The (P) symbol indicates the cluster of
651 autophosphorylation sites in the N-terminal region (N-t) of *Mtb*PknG (13). The IXI symbol in the
652 TPR domain of *Cg*PknG$_{604-661}$ represents an internal segment of deleted amino acids. (B)
653 Complementation of the *ΔpknG* strain with different *pknG* versions. Complementation was

assessed by growth on CGXII plates with 100 mM glutamine as sole carbon source after 3 days at 30°C. PknG variants capable to complement the *ΔpknG* strain were *Cg*PknG, *Mtb*PknG and *Cg*PknG$_{\Delta 604\text{-}661}$. The empty pEKEx2 vector was used as a negative control.

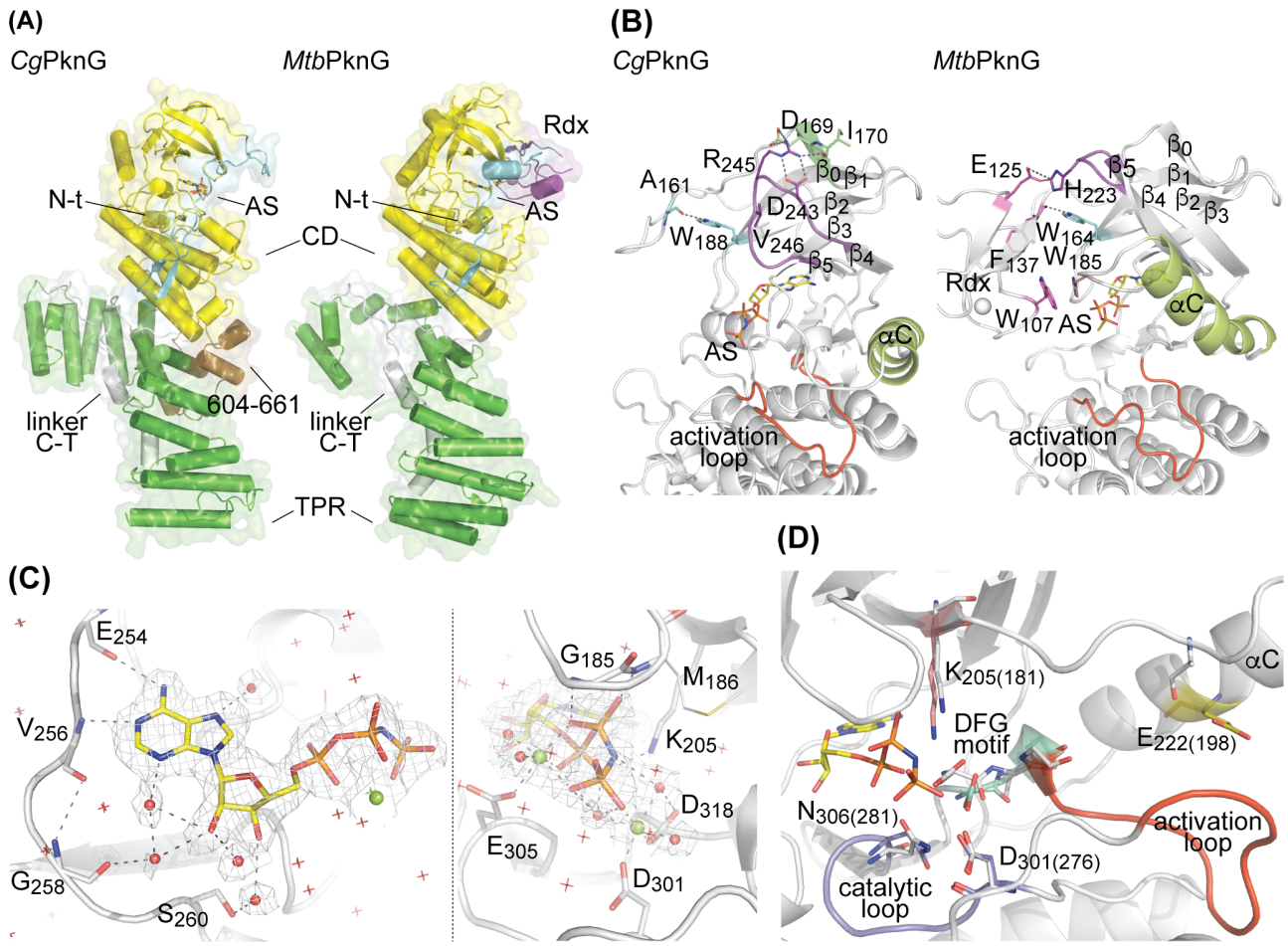
Figure 2



***Cg*PknG and CgPknG$_{Δ1-129, Δ434-822}$ and their relative kinase activities.** (A) Relative kinase activities of *Cg*PknG and CgPknG$_{Δ1-129, Δ434-822}$ against OdhI and GarA. (B) Kinase activity of *Cg*PknG for different concentrations of the 17-mer peptide substrate SDEVTV**ETTS**VFRADFL. (C) Relative kinase activity of *Cg*PknG and CgPknG$_{Δ1-129, Δ434-822}$ against the 17-mer peptide. Measurements were performed at least twice; error bars represent the scattering among average values obtained in independent determinations.

668 **Figure 3**

669



670

671

672 **The crystal structure of *Cg*PknG$_{\Delta N\text{-}t}$.** (A) Comparison of *Cg*PknG$_{\Delta N\text{-}t}$ and *Mtb*PknG$_{\Delta 1\text{-}73}$ (12) (PDB

673 code 2PZI). The chain A in each crystal structure is shown (RMSD of 2.35 Å for 532 aligned

674 residues). The non-hydrolysable ATP analog AMP-PNP bound to the active site (AS) of *Cg*PknG$_{\Delta N\text{-}t}$

675 is depicted in sticks. N-t: N-terminal region; CD: catalytic domain; linker C-T: linker between the

676 catalytic domain and the TPR domain. (B) Comparison of *Cg*PknG$_{\Delta N\text{-}t}$ and *Mtb*PknG$_{\Delta 1\text{-}73,\ \Delta 406\text{-}750}$ (13)

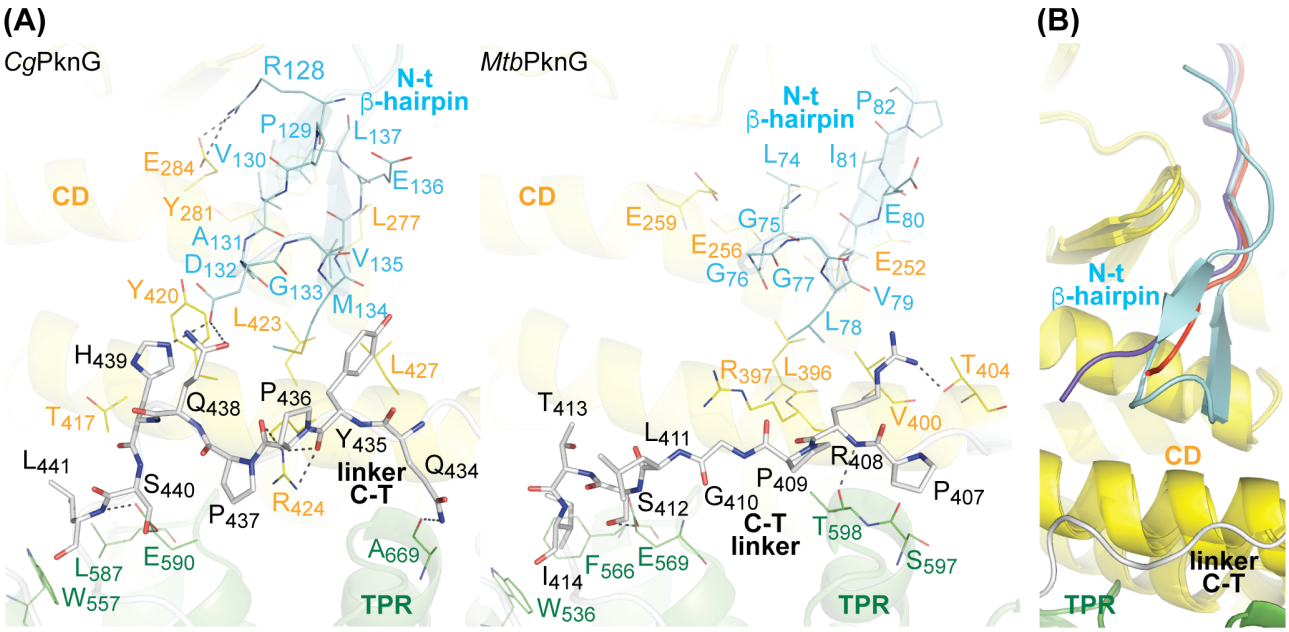
677 (PDB code 4Y12). The highlighted kinase domain residues or motifs adopt distinct conformations

678 in the absence or in the presence of an Rdx domain. (C) The ATP binding site of *Cg*PknG $_{\Delta N\text{-}t}$ with a

679 bound AMP-PNP molecule. The AMP-PNP molecule and the protein residues interacting with it are
680 shown in sticks. Water molecules are depicted as red spheres or stars and Mg(II) atoms are shown
681 as green spheres. The *2mFo–DFc* electron density is contoured to 1.0 σ and presented as a mesh.
682 Dashed lines represent atomic interactions. (D) Functionally important and conserved residues
683 within the kinase active site are shown for *Cg*PknG$_{\Delta N\text{-}t}$. Gray sticks correspond to residues in
684 *Mtb*PknG$_{\Delta 1\text{-}73,\ \Delta 406\text{-}750}$ (13) (PDB code 4Y12), numbered between brackets.
685

686 **Figure 4**

687



688

689

690 **The linker C-T simultaneously interacts with an N-terminal β-hairpin, the catalytic core and the**

691 **TPR domain of PknG.** (A) Comparison of the crystal structures of *Cg*PknG$_{\Delta N\text{-}t}$ (this work) and

692 *Mtb*PknG$_{\Delta 1\text{-}73}$ (12) (PDB code 2PZI). The chain A in each crystal structure is shown. Selected

693 residues within the linker C-T are shown in sticks. Residues conforming the N-terminal β-hairpin

694 are depicted as lines. Residues of the catalytic core or the TPR domain involved in polar or

695 hydrophobic interactions with the N-terminal β-hairpin or the linker C-T are also shown as lines.

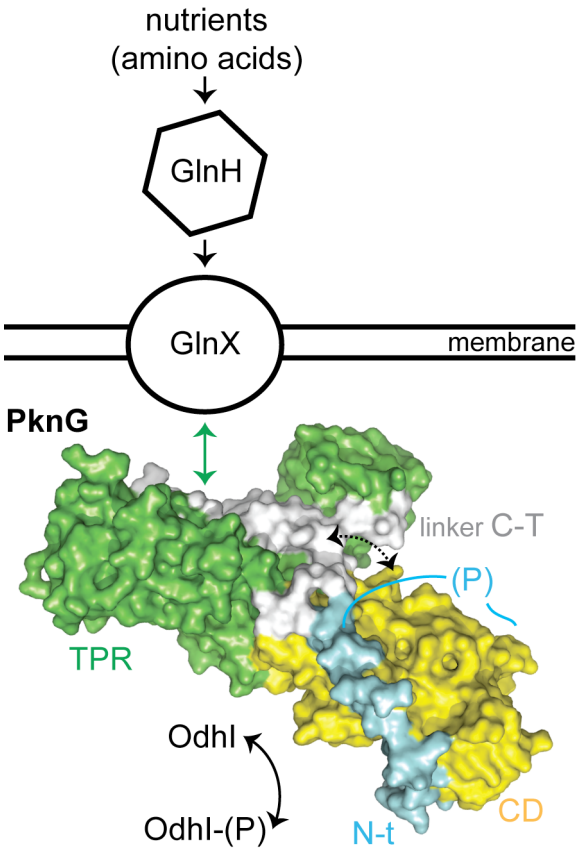
696 Dashed lines represent polar interactions. (B) The crystal structures of *Cg*PknG$_{\Delta N\text{-}t}$ and *Cg*PknG$_{\Delta 1\text{-}}$

697 $_{129,\Delta 434\text{-}822}$ are superimposed. The RMSD values between the chain A in the structure of *Cg*PknG$_{\Delta N\text{-}t}$

698 and the structures of *Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$ are 1.04 Å and 0.79 Å for 282 and 288 aligned residues,

699 respectively. The N-terminal extension of *Cg*PknG$_{\Delta 1\text{-}129,\Delta 434\text{-}822}$ is colored in blue or red.

Figure 5



Proposed model for the role of the TPR domain in the *Cg*PknG function. The available genetic, biochemical, and structural evidence suggests that the TPR domain might act as a localization scaffold that, providing a surface for the interaction between the kinase and the transmembrane protein GlnX, would couple signal detection to OdhI phosphorylation by modulating the conformation of the linker C-T.