1 Identification and biochemical characterization of a novel eukaryotic- 2 like Ser/Thr kinase in *E. coli*

3
4 Krithika Rajagopalan[#] and Jonathan Dworkin*

5 Department of Microbiology & Immunology, College of Physicians and Surgeons,
6 Columbia University, New York, NY 10032 USA

7
8 *Mailing address: 701 W. 168th St., Room 1218, New York, NY 10032

9 T: 212.342.3731, email: jonathan.dworkin@columbia.edu

10 [#Present address: Division of Biological Sciences, University of California, San Diego,](#)
11 [San Diego, CA](#)

12
13 **Keywords: Ser/Thr phosphorylation, Yegl**

14
15
16

Abstract

In bacteria, signaling phosphorylation is thought to occur primarily on His and Asp residues. However, phosphoproteomic surveys in phylogenetically diverse bacteria over the past decade have identified numerous proteins that are phosphorylated on Ser and/or Thr residues. Consistently, genes encoding Ser/Thr kinases are present in many bacterial genomes such as *E. coli*, which encodes at least three Ser/Thr kinases. Here we identify a previously uncharacterized ORF, *yegl*, and demonstrate that it encodes a novel Ser/Thr kinase. Yegl lacks several conserved residues including those important for Mg^{2+} binding seen in other bacterial Ser/Thr kinases, suggesting that the consensus may be too stringent. We further find that Yegl is a two-pass membrane protein with both N- and C-termini located intracellularly.

Introduction

Reversible protein phosphorylation is an important regulatory mechanism in eukaryotes and prokaryotes (1). In eukaryotes, signaling phosphorylation typically occurs on serine, threonine or tyrosine residues and is mediated by the combined action of kinases and phosphatases. In prokaryotes, signaling phosphorylation has been thought to occur largely on histidine and aspartate residues mediated by histidine kinases of two-component systems (2). However, mass spectrometry based-phosphoproteomic analyses over past decade have identified numerous Ser/Thr/Tyr phosphorylated proteins in many bacteria, including *Escherichia coli* (3-6). Some of these phosphoproteins and their specific phosphosites are conserved in divergent species (4) suggesting that this regulation is physiologically relevant.

Ser/Thr kinases from phylogenetically diverse bacteria have been described (7). Bacterial Ser/Thr kinases can be divided into two classes based on their structure and sequence homology: eukaryotic-like Ser/Thr kinases (eSTKs) and atypical Ser/Thr kinases. eSTKs bear striking resemblance in structure and function to Hanks family Ser/Thr kinases from eukaryotes (7). Unlike eSTKs, atypical Ser/Thr kinases lack any sequence homology to eukaryotic kinases. So far, three genes (*hipA*, *yeaG*, *yihE*) encoding Ser/Thr kinases have been reported in *E.coli* of which HipA and YihE are atypical Ser/Thr kinases while YeaG is an eSTK. YeaG plays a role in nitrogen starvation (8), YihE is involved in the Cpx stress response (9) and cell death pathways (10) and HipA regulates bacterial persister formation by phosphorylating a tRNA synthetase (11,12). However, the authentic *in vivo* substrates of these kinases and/or their proximal activating stimuli are largely uncharacterized, complicating efforts to understand their precise physiological role.

Here, we report that the previously uncharacterized *E. coli yegI* gene encodes a Ser/Thr kinase. We describe the biochemical characterization of Yegl, which we demonstrate is an integral membrane sSTK. Despite differences in conserved residues observed in other eSTKs, we show that Yegl is an active kinase and is sensitive to the kinase inhibitor staurosporine.

54 Additionally, we show that Yegl kinase activation follows second order kinetics, consistent with
55 the proposed oligomerization-dependent activation of eSTKs.

Results

YegI is a novel eukaryotic-like Ser/Thr kinase

Our lab previously identified a novel PP2C-like phosphatase in the *E. coli* genome encoded by the *pphC* (*yegK*) gene (13). Often, phosphatases of this class are found proximal to eukaryotic-like Ser/Thr kinases (eSTKs) (7). Inspection of the genome of the *E. coli* B strain REL606 revealed the presence of a 1941bp ORF *yegI*, encoding a 646 amino acid protein predicted to be a Ser/Thr kinase, immediately downstream of *pphC*. Of note, BLAST analysis of *E. coli* with previously identified Ser/Thr kinases including YeaG and HipA does not identify *yegI*. In *E. coli* REL606 and in several pathogenic strains such as *E. coli* O157:H7, *yegI* has a four nucleotide overlap (ATGA) at the 5' end with the 3' end of *pphC*, suggesting that the genes are co-transcribed. Interestingly, in *E. coli* K strains (e.g., MG1655, W3110), the ORF *yegJ*, encoding a protein of unknown function, is found between *yegI* and *pphC* (Supp. Fig. 1) suggesting that it may disrupts co-ordinated expression of *pphC* and *yegI*.

Amino acid sequence analysis and alignment of YegI with bacterial eSTKs (Fig. 1) revealed the presence of Hanks signature motifs (I-XI) suggesting that YegI is an authentic eukaryotic-like Ser/Thr kinase (eSTK) (14). However, YegI contains only seven out of the ten absolutely conserved residues commonly found in eSTKs. For example, in motif III, YegI has an aspartic acid residue instead of the conserved glutamic acid (Fig.1). This conserved glutamic acid residue contacts Lys-39 in Motif II and is required for stable binding of α and β phosphates of ATP (15). YegI also lacks the conserved asparagine (replaced by serine) in motif VI that together with catalytic site residues Asp-141 and Asp-159 is important for stability of Mg^{2+} ion binding (15). Lastly, YegI lacks a conserved arginine (replaced by glycine) in motif XI that is important for stabilizing substrate-protein interactions (Fig.1).

A homology model of the YegI kinase domain (aa1-300) based on *M. tuberculosis* eSTK PknB (PDB 1MRU) as a template using Swiss-model (16) demonstrated that despite the absence of these three highly conserved residues, the predicted structure of the YegI kinase domain shows

all the structural features seen in *M. tuberculosis* eSTK PknB (Supplementary Fig. 2A). Specifically, Yegl adopts the canonical bilobed structure seen in eSTKs. The N-terminal lobe of the Yegl kinase domain contains the ATP binding P-loop and the helix C structure (Supplementary Fig. 2B) which is responsible for proper orientation of the active site with the substrate. The C-terminal lobe includes the catalytic loop containing all the residues required for catalysis and an activation loop that is important for kinase activation by autophosphorylation. Thus, although the Yegl sequence is only ~20% identical to PknB (Fig.1), the similarity of the predicted structure of Yegl to that of PknB (Supplementary Fig. 2A), strongly suggests that Yegl is an eSTK.

Yegl is an active kinase

We previously demonstrated that a His₆-tagged version of Yegl is capable of autophosphorylation (13) despite the absence of three of the ten absolutely conserved residues (Fig. 1). To confirm that Yegl belongs to the eSTK family, we examined the consequences of mutations in the conserved lysine at position 39 and a conserved aspartic acid site at position 141, which play a crucial role in ATP binding and catalysis, respectively, in this family (17). We generated Yegl mutant proteins carrying either lysine to aspartic acid (K39D) and aspartic acid to asparagine (D141N) changes. The mutant proteins were expressed and purified in the same manner as the wildtype Yegl. Kinase activity was demonstrated using an *in vitro* radioactive kinase assay using radiolabeled γ -[³²P]ATP. Mutation of either Lys-39 or Asp-141 resulted in complete loss of autophosphorylation activity as compared to the wildtype protein indicating that both residues are essential for kinase activity (Fig. 2A).

We examined how the lack of some of the conserved residues affected Yegl function in a series of standard assays of bacterial Ser/Thr kinases. For example, staurosporine is widely used as a small molecule inhibitor for Ser/Thr kinases (18) and bacterial eSTKs are typically sensitive to staurosporine with IC₅₀ in the μ M range (19). Sensitivity of the Yegl kinase to staurosporine was assessed by incubating Yegl with increasing concentrations of staurosporine. Yegl shows an

IC₅₀ of ~50 μ M (Fig. 2B), comparable to the *Enterococcus faecalis* eSTK IreK (IC₅₀ of 150 μ M (20)) and the *Staphylococcus epidermis* eSTK Stk (IC₅₀ of 320 μ M (21)). Bacterial eSTKs require the presence of bivalent cations for activity (22-25). We therefore investigated the effect of different bivalent cations on Yegl activity by performing radioactive kinase assays in the presence of either MgCl₂/MnCl₂/CaCl₂/NiCl₂. Similar to many eSTKs, Yegl catalyzed autophosphorylation only in the presence of MgCl₂ and MnCl₂, albeit to different levels. Yegl displayed much higher autophosphorylation in the presence of Mn²⁺ ion than Mg²⁺ ion suggesting that Yegl is a Mn²⁺ dependent kinase (Supplementary Fig. 3). The concentration dependence of Yegl kinase activity on Mg²⁺ or Mn²⁺ was measured and the optimal MnCl₂ concentration was determined to be between 0.1-1 mM whereas optimal MgCl₂ concentration was between 1-10 mM (Fig.3A) indicating that Yegl kinase is more active in the presence of MnCl₂ than MgCl₂.

In eSTKs, the highly conserved DFG motif present in domain VII coordinates Mg²⁺ ion binding (26,27). However, amino acid sequence alignment reveals that Yegl does not contain the DFG sequence found in bacterial eSTKs but instead the sequence DSD (Fig. 1; Fig. 3B). To examine if addition of the DFG motif to Yegl alters preference to MgCl₂, we generated a mutant protein carrying changes of Ser-160 to a phenylalanine and Asp-161 to a glycine (S160F, D161G). Additionally, we generated a mutant protein carrying a change of Asp-161 to a glycine (D161G) which would change the DSD sequence to the DSG consensus motif. Mutants were expressed and purified in the same manner as wildtype Yegl and kinase activity in the presence of MgCl₂/MnCl₂ was assessed. Surprisingly, replacement of DSD to DFG or to DSG completely inactivated the kinase (Fig. 3B) indicating that the DSD motif is indispensable for Yegl autophosphorylation.

Autophosphorylation of Yegl

Autophosphorylation is a common model for activation of eukaryotic STKs (27), therefore characterization of the kinetics of autophosphorylation and the sites of autophosphorylation can

provide insight into how bacterial homologs of these kinases are activated. We performed standard autophosphorylation assay (Fig. 4A) using a range of Yegl concentrations and measured formation of radiolabeled phosphorylated Yegl (Fig. 4B). Our results indicate that Yegl autophosphorylation is not first order and are consistent with Yegl undergoing oligomerization. To map the sites of autophosphorylation, His₆-Yegl that had been subjected to an *in vitro* kinase reaction was analyzed by mass spectrometry. Five phosphopeptides that each included at least a single serine residue were identified, leading to the conclusion that Ser-153, Ser-160, and Ser-162 in the predicted activation loop and Ser-602 and Ser-644 at the C-terminus of Yegl were phosphorylated (Fig. 5A). To characterize the functional importance of the C-terminal phosphosites, we generated single and double phosphoablative serine to alanine mutants. The mutant proteins were expressed and purified in the same manner as the wildtype protein and assessed for loss of autophosphorylation using a radioactive kinase assay. Since mutation of both Ser-602 and Ser-644 to alanine resulted in only a modest decrease in overall autophosphorylation, phosphorylation of these residues is not necessary for normal Yegl activity (Fig.5B).

Since eSTKs undergo autophosphorylation on Ser or Thr residues in the activation loop of the kinase domain (28-31), we investigated the phosphosites (Ser-153, Ser-160, Ser-162) in the putative Yegl activation loop. All three residues are located between Motif VI and VII in close proximity to the catalytic loop and activation loop suggesting that phosphorylation of these residues could be important for kinase activity. We mutated each Ser to an alanine to examine if phosphorylation of these Ser residues impacted kinase activity. We observed that mutating Ser-162, but not the other phosphosites (Ser-153 or Ser-160), resulted in a complete loss of autophosphorylation (Fig. 5B), demonstrating that autophosphorylation of Ser-162 is important for Yegl activation.

Yegl is an integral membrane protein

We observed that purification of recombinant His₆-YegI required the presence of the anionic detergent N-lauryl sarcosine suggesting that YegI is a membrane protein. Consistently, the membrane topology algorithm Phobius (32) predicts that YegI consists of two transmembrane domains joined by a short periplasmic loop with both the N- and the C-terminus located in the cytosol (Fig. 6A). We used the PhoALacZ α dual reporter system (33-36) to experimentally validate this predicted topology. In this system, alkaline phosphatase (PhoA, aa 21-471) lacking the periplasmic localization signal is fused in frame with the α -fragment of *E. coli* β -galactosidase (LacZ α , aa 5-60), creating PhoALacZ α . Localization of a protein of interest is then analyzed by fusing it to PhoALacZ α . Since PhoA is active only in the periplasm and LacZ is only active in the cytoplasm, the PhoALacZ α fusion displays high phosphatase activity but no β -galactosidase activity when targeted to the periplasm or to the membrane. Conversely, the fusion displays high β -galactosidase activity but no phosphatase activity when it is targeted to the cytoplasm. PhoALacZ α reporter fusions can be assessed using dual indicator plates with PhoA or LacZ specific chromogenic substrates that allow identification of periplasmic or cytoplasmic fusions (35). Specifically, cells expressing PhoALacZ α protein fusions with high levels of phosphatase activity process the X-phos (5-bromo-4-chloro-3-indolyl phosphate) substrate yielding blue colored colonies, while cells expressing PhoALacZ α protein fusions with high levels of β -galactosidase activity process the Red-gal (6-chloro-3-indolyl- β -D-galactoside) substrate yielding red colored colonies.

We constructed C-terminal PhoALacZ α fusions to YegI at amino acid positions 340, 390, 414, 432, 439 and 646 (Fig. 6A) and analyzed them on dual indicator plates containing X-phos and Red-gal. DH5 α cells carrying empty backbone vector alone served as a negative control ("vector") and were colorless, and DH5 α cells expressing PhoALacZ α without YegI ("PhoALacZ α ") served as a positive control and were red/pink indicative of cytoplasmic location (Fig.6B). Consistent with the *in silico* prediction, PhoALacZ α fusions to YegI at amino acid positions 340, 390, 439 and 646 yielded red/pink colored colonies suggestive of high β -galactosidase activity

and cytoplasmic localization (Fig. 6B; Table 1). In contrast, PhoALacZα fusions of Yegl at amino acid positions 414 and 432 yielded blue colored colonies on indicator plates suggestive of high alkaline phosphatase activity and periplasmic/TM localization (Fig. 6B). In summary, the PhoALacZα reporter fusions confirmed the bioinformatic prediction that Yegl is an integral inner membrane protein containing two TM domains connected by a short periplasmic loop, with both the N- and C-terminus located in the cytosol.

Discussion

Ser/Thr phosphorylation is an important post translational modification that is mediated by Ser/Thr kinases and cognate Ser/Thr protein phosphatases which have been identified in phylogenetically diverse organisms. Although Ser/Thr phosphorylation in the Gram-negative bacterium *E. coli* has been known for >10 years, the enzymes responsible for this reversible modification and their specific substrates are not well understood. In this study, we have described the characterization of Yegl, a novel Ser/Thr kinase *E. coli*.

In silico analysis revealed that Yegl belongs to the class of eSTKs that contain all the 11 key structural subdomains that are characteristic of eukaryotic protein kinases (17). Homology search using Yegl kinase domain revealed that it was similar to kinase domains of several bacterial eSTKs (Supplementary Fig. 2A). Unlike other bacterial eSTKs, however Yegl includes only seven of the ten absolutely conserved residues responsible for ATP binding, catalysis, and substrate recognition (14,17) (Fig. 1). Despite the absence of these conserved residues, Yegl encodes an active protein kinase (Fig. 2A) that undergoes autophosphorylation primarily on serine residues (Fig. 5A) Similar to most bacterial eSTKs, autophosphorylation of Yegl on serine residues in the activation loop is critical for catalytic activity of Yegl (Fig. 5B).

Although the predicted structure of Yegl closely resembled the structure of the bacterial eSTK PknB (Fig. 4.1D), the Yegl sequence shows striking deviations in the conserved regions particularly in the catalytic loop and the Mg²⁺-binding DFG loop (Fig. 1). Specifically, the Yegl

sequence contains the catalytic loop VGDXXQXS as opposed to the canonical catalytic loop HRDXXPXN found in almost all other bacterial eSTKs. The conserved histidine residue in the catalytic loop is crucial for binding both the carbonyl group of catalytic aspartate in the catalytic loop and for making a hydrophobic contact with the phenylalanine in the DFG loop. Additionally, Yegl lacks the highly conserved DFG (aspartate-phenylalanine-glycine) loop in motif VII (Fig.1, 3B) which recognizes the Mg^{2+} -bound ATP (37). Instead, Yegl contains a DSD (aspartate-serine-aspartate) sequence suggesting either that Yegl might be an inactive kinase or that the DFG loop is not essential for every kinase. Our results suggest the latter, since, despite these differences in highly conserved residues, Yegl is an active kinase (Fig. 2A) that requires the bivalent cations Mg^{2+}/Mn^{2+} for activity with preference for Mn^{2+} (Fig. 3A). We speculated that addition of DFG loop could help kinase activation. However, replacing the DSD loop in Yegl with the canonical DFG loop failed to increase activity or preference for Mg^{2+} ion and instead completely abolished activity (Fig. 3B).

This result suggests that the presence of a serine and an aspartic acid residue in the DSD loop might be critical for activity and could affect Yegl autophosphorylation. Consistently, we observed that the serine residue in the DSD loop is one of the three serine residues that undergoes autophosphorylation and is required for activity of the kinase (Fig. 5A,B). Substitution of all three serine residues with alanine residues (Ser153Ala, Ser160Ala, Ser162Ala) decreased Yegl kinase activity to different extents indicating that these residues are important for autophosphorylation of Yegl (Fig.5B). Particularly, mutation of Ser-162 adjacent to Asp-161 completely abolished Yegl kinase activity further validating the importance of DSD loop in Yegl autophosphorylation.

Two possible mechanisms of kinase activation have been described: direct activation by ligand-promoted dimerization or indirect activation. The first evidence of ligand-promoted dimerization and activation of the kinase came from the structure of *M. tuberculosis* PknB (15,38) which forms back-to-back dimers of the N-terminal lobe of the catalytic domain that facilitates

autophosphorylation. This mechanism of ligand-induced dimerization and activation was also observed for two other eSTKs, *M. tuberculosis* PknD (39), and *Pseudomonas aeruginosa* PpkA (40). In cases where the kinase lacks a ligand binding domain, an indirect mechanism has been proposed as a model for activation. For example, the structure of the PknB kinase domain alone with an ATP competitive inhibitor revealed formation of asymmetric front-to-front dimers via interaction between the α G helix and an ordered activation loop on one domain and α G helix and a disordered activation loop of the second domain. Substitutions in the α G helix interface resulted in decreased autophosphorylation suggesting that α G helix in the C-terminal lobe of the kinase is important for activation. The conformation of the two proteins in the dimer is such that one monomer acts like an activator while the other monomer acts like a substrate very similar to a transphosphorylation model (28). This provides an alternate mechanism of activation especially for soluble kinases that lack a membrane bound domain. Unlike most eSTKs, Yegl lacks an extra cytoplasmic sensor domain and instead is a double pass membrane protein with both the N and C terminus in the cytoplasm (Fig. 6B). Additionally, our studies show that Yegl follows second order kinetics which is consistent with dimer formation (Fig. 4A,B). Future studies will be aimed at understanding the mechanism of activation of Yegl.

An interesting feature of the genomic organization of *yegl* is that it differs in closely related *E. coli* strains (Supplementary Fig.1). In most pathogenic strains and lab B strains of *E. coli*, *yegl* occurs immediately downstream of *pphC*, a PP2C-like phosphatase with a four nucleotide overlap while in K strains this locus is disrupted by a gene, *yegJ*, encoding an ORF without similarity previously identified proteins, suggesting potential regulatory differences between strains. In REL606 strains, *yegl* expression was detected only in stationary phase but at very low levels (data not shown). However, since the identity of the physiological stimulus that initiates *yegl* expression and activation remains unknown, it is difficult to interpret the functional implications of the altered gene locus.

An additional open question is the identity of Yegl substrate(s). There are many candidates

as ~100 cellular proteins were found to be phosphorylated on Ser/Thr residues in *E. coli* (4). A useful strategy to identify authentic in vivo substrates is comparative phosphoproteomics where the phosphoproteomes of a strain lacking either the eSTK or partner eSTP are compared with the WT strain. This methodology has been utilized previously for identifying phospho substrates of *B. subtilis* PrkC using deletion strains of eSTK PrkC and eSTP PrpC (41). This methodology has also been successfully utilized in *M. pneumoniae*, where 4 out of the 63 phosphorylated proteins were specifically enriched in a deletion strain of novel eSTK PrkC (42). Therefore, phosphoproteomic analysis might serve as a useful tool in identifying phosphorylated substrates of Yegl by examining $\Delta yegl$ strains.

Figure legends

Figure 1. Amino acid sequence alignment of Yegl with bacterial eSTKs. Yegl (*Escherichia coli*) was aligned to bacterial eSTKs PrkC (*Bacillus subtilis*), PknB (*Mycobacterium tuberculosis*) and PpkA (*Pseudomonas aeruginosa*) using T-coffee (43). Alignment was visualized using Jalview. Absolutely conserved residues are shaded in dark blue whereas residues conserved in three out of the four kinases are shaded in light blue. The ten absolutely conserved residues found in bacterial eSTKs are depicted with arrows. Hanks conserved domains are depicted on top of the sequence as roman numerals (I-XI). Yegl (aa 1- 303), PrkC (aa 1-269), PknB (aa 1-275) and PpkA (aa 1-265) which includes the kinase domain was used for alignment. Structurally and functionally important motifs are indicated below the sequence.

Figure 2. Yegl is an active kinase

(A) Kinase activity of Yegl: Autophosphorylation reactions were carried out at 37 °C with 0.2 μ M of Yegl (WT or K39D or D141N) in kinase buffer (50 mM Tris pH 7.5, 50 mM KCl, 1 mM DTT, 10 mM MgCl₂, 10 mM MnCl₂, 200 μ M cold ATP and 5 μ Ci γ -[³²P]ATP. Reactions were stopped at

t=30 mins and run on 12 % SDS-PAGE followed by autoradiography. Molecular weights (kDa) are indicated on the right of the gel.

(B) Sensitivity to staurosporine: Autophosphorylation reactions were carried out at 37 °C with 1 µM of Yegl (WT) in kinase buffer without cold ATP. Different concentrations of staurosporine (µM) was added to the reaction at indicated concentrations. Reactions were stopped at t=30 mins and run on 12% SDS-PAGE followed by autoradiography. Molecular weights are indicated on the right of the gel.

Figure 3. Yegl is a Mn²⁺ dependent kinase

(A) Requirement of Mg²⁺/Mn²⁺ for kinase activity: Autophosphorylation reactions were carried out at 37 °C with 1 µM of Yegl (WT) in kinase buffer with indicated concentrations of MgCl₂/ MnCl₂. Reactions were stopped at t=30 mins and run on 12% SDS-PAGE followed by autoradiography. Molecular weights are indicated on the right of the gel.

(B) Requirement of DFG motif: Autophosphorylation reactions were carried out at 37 °C with 0.2 µM of Yegl (WT or S160FD161G or D161G) in kinase buffer with either 10 mM MgCl₂ or MnCl₂ as described in Supp. Figure 3

Figure 4

Kinetics of Yegl autophosphorylation

(A) Autophosphorylation of Yegl: Autophosphorylation reactions were carried out at 37 °C with indicated concentrations of Yegl (WT) in kinase buffer. Reactions were stopped after 30 min and run on 12% SDS-PAGE followed by autoradiography. Molecular weights are indicated on the right of the gel.

(B) Autophosphorylation of Yegl follows second order kinetics: Kinetics of Yegl autophosphorylation was assessed using different concentrations of Yegl. % P32 incorporation was calculated relative to autophosphorylation in the presence of 2 µM of Yegl from Fig 4(A).

Figure 5

Yegl undergoes autophosphorylation on serine residues in the kinase domain and the C-terminus

(A) **Graphical representation of Yegl phosphorylation sites:** Domain structure depicting phosphorylation sites following mass spectrometry analysis of autophosphorylation. Phosphorylated residues are indicated by red circles.

(B) **Sites of autophosphorylation on Yegl:** Reactions were carried out at 37 °C with 1 µM of Yegl (WT or phosphoablative mutants) in kinase buffer. Reactions were stopped at t=30 mins and run on 12 % SDS-PAGE followed by autoradiography. Molecular weights (kda) are indicated on the right of the gel.

Figure 6

Yegl is an integral membrane protein

A) **Predicted topology of Yegl using Phobius.** Amino acid positions at which phoA-lacZα reporters are fused are indicated along with associated phenotypes. Phenotypes of phoAlacZα fusions due to LacZ activity are denoted as red ellipses while PhoA activity are denoted as blue ellipses. Predicted transmembrane domains (I and II) are depicted as grey rectangles.

(B) **Membrane topology analysis:** DH5α transformed with different phoAlacZα reporters were streaked on an LB plate containing 5-bromo-4-chloro-3-indolyl phosphate disodium salt (X-Phos, 80µg/ml), 6-chloro-3-indolyl-β-D-galactoside (Red-Gal, 100 µg/ml, 0.1 % w/v arabinose and 100 µg/ml of ampicillin. Control pBAD24: control strain expressing pBAD24 empty vector, control phoAlacZα: strain expressing phoAlacZα in pBAD24 without Yegl. Amino acid positions at which phoAlacZα reporters are fused are indicated.

Table 1 Localization of different phoAlacZα reporter fusions.

Materials and Methods

Bacterial strains and growth conditions

E. coli DH5α cells were used for regular cloning and *E. coli* LOBSTR (BL21-DE3) strains were used for expression of recombinant proteins. *E. coli* cells for expression were grown in LB Lennox broth supplemented with ampicillin (100 µg/ml) at 37 °C with shaking (220 rpm) unless otherwise specified. Genomic DNA from *E. coli* REL606 was isolated using a Wizard genomic DNA purification kit (Promega) following manufacturer's instructions. Details of strains, plasmids and primers used in the study are described in Supplementary Tables S1, S2, and S3, respectively.

Cloning and expression of Yegl and Yegl-NTD

The *yegl* gene was PCR amplified using Phusion polymerase (Thermo Scientific) from *E. coli* REL606 genomic DNA using primers (KR58/ KR59). Sequence for an N-terminal His₆ tag was included in the primer. The PCR product was digested with *Nco*I/*Sph*I and ligated into similarly digested pBAD24 backbone. Ligation products were transformed in DH5α cells and selected on LB/ampicillin plates. The resulting plasmid pKR31 generated an N-terminal His₆-tagged Yegl fusion protein. Plasmid cloning was subsequently verified by restriction enzyme digest and DNA sequencing (Operon).

For protein expression, plasmid was transformed into *E. coli* LOBSTR (Kerafast) cells and plated on LB ampicillin (100 µg/ml) agar plates. Single colonies were inoculated into 3 ml LB supplemented with ampicillin (100 µg/ml) for overnight cultures. The next day, a dilution of 1:250 in 500 ml LB was grown to an OD₆₀₀ of 0.6-0.8. Yegl recombinant protein was induced by addition of arabinose (0.2% w/v) for 3 h at 18 °C. Cells were harvested at 6000 x g for 15 min at 4 °C.

Pellets were washed with ice-cold 50 mM EDTA and centrifuged at 7000 rpm for 15 min at 4 °C.
Washed pellets were saved at -80 °C until use.

Oligonucleotide site directed mutagenesis for Yegl point mutants

Point mutation of lysine residue K39 was generated by two step overlap PCR mutagenesis using primer pairs (KR58/KR246) and (KR245/59) with Phusion polymerase. A second PCR was performed with primers (KR58/KR59) using primary PCR products as a template and the subsequent PCR products were digested with *Nco1/Sph1* and ligated into similarly digested pBAD24 to generate an N-terminal His₆ tagged K39D Yegl fusion protein. Plasmid cloning was subsequently verified by restriction enzyme digest and DNA sequencing (Operon).

Point mutation of aspartic acid residue D141 was generated by two step overlap PCR mutagenesis using primer pairs (KR58/ KR248) and (KR247/KR59) with Phusion polymerase. A second PCR was performed with primers (KR58/KR59) using primary PCR products as a template and the subsequent PCR products were digested with *Nco1/Sph1* and ligated into similarly digested pBAD24 to generate an N-terminal His₆ tagged D141N Yegl fusion protein. Plasmid cloning was subsequently verified by restriction enzyme digest and DNA sequencing (Operon).

Phosphoablative point mutants used in the study were generated using the same strategy as above and details of plasmid construction are described in Supplementary table S2.

Purification of recombinant Yegl

Frozen pellets were suspended in lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl, 10 mM β-mercaptoethanol, 1 mM PMSF and 2% w/v sarkosyl) and incubated at room temperature for overnight lysis. Cells were subsequently lysed using sonication. Lysates were cleared by centrifugation at 15000 x g for 30 min at 4 °C and incubated in Pierce 5 ml columns with Ni-NTA agarose beads (Qiagen) at 4 °C for 1 h. Lysate was allowed to flow through and beads were

washed 10 column volumes of wash buffer (50 mM Tris pH 8.0, 200 mM NaCl, 30 mM imidazole, 10 mM β -mercaptoethanol, 0.05% w/v sarkosyl). His tagged protein was eluted using 300 mM imidazole in 50 mM Tris pH 8.0, 200 mM NaCl, 10 mM β -mercaptoethanol and 0.05% w/v sarkosyl. Elution fractions were tested on 12% SDS-PAGE gel and fractions containing protein were pooled and dialyzed overnight at 4 °C using Snakeskin dialysis tubing 10K MWCO (Thermo Scientific) in kinase storage buffer (20 mM Tris pH 8.0, 125 mM NaCl, 10% glycerol, 1 mM DTT). Dialyzed protein was assessed for purity using 12% SDS-PAGE gel and stored at -80 °C. Yegl variant proteins were purified using the same protocol. Proteins were stored at -80 °C.

Autophosphorylation assays

Autophosphorylation assay using Yegl point mutants

Autophosphorylation of Yegl and point mutants was performed by addition of 5 μ Ci of [γ -³²P] ATP (Perkin Elmer) to 0.2 μ M of purified Yegl or Yegl point mutants in 10 μ l of reaction buffer containing 50 mM Tris pH 7.5, 50 mM KCl, 0.5 mM DTT, 200 μ M ATP, 10mM MgCl₂ and 10mM MnCl₂. Reactions were incubated at 37 °C for 30 mins and were stopped using 3X Laemmli buffer and boiled for 5 min at 95 °C. Samples were resolved on a 12% SDS-PAGE gel and visualized by staining with Coomassie dye. Radioactive gels were dried for 30 mins at 80 °C in a gel dryer, exposed to a phosphoscreen and visualized by autoradiography using a Typhoon Scanner (GE Healthcare).

Inhibition of autophosphorylation by staurosporine

Staurosporine sensitivity experiments were performed using 1 μ M of Yegl in 10 μ l of reaction buffer containing 50 mM Tris pH 7.5, 50 mM KCl, 0.5 mM DTT, 10 mM MgCl₂ and 10 mM MnCl₂. Staurosporine (Calbiochem) was dissolved in DMSO to obtain a stock concentration of 10 mM. Staurosporine stock (10 mM) was diluted in DMSO and added to the above kinase reaction to obtain final concentrations in the range of 12.5-100 μ M. DMSO was added to the no inhibitor control. Reactions were incubated at 25 °C for 10mins followed by addition of 5 μ Ci of [γ -³²P] ATP

(Perkin Elmer). Reactions were incubated for additional 30 mins at 37 °C and were stopped using 3X Laemmli buffer. Samples were boiled for 5 min at 95 °C and resolved on a 12% SDS-PAGE gel. Gel was visualized by staining with Coomassie dye. Radioactive dried gel was exposed and visualized by autoradiography using a Typhoon Scanner (GE Healthcare).

Autophosphorylation assay using different bivalent cations

Autophosphorylation of Yegl was performed by addition of 5 µCi of [γ -³²P] ATP (Perkin Elmer) to 1 µM of purified Yegl in 10 µl of reaction buffer containing 50 mM Tris pH 7.5, 50 mM KCl, 0.5 mM DTT, 200 µM ATP. MgCl₂ or MnCl₂ was added to a final concentration of 0.01, 0.1, 1 and 10 mM. Reactions were incubated at 37 °C for 30 mins and were stopped using 3X Laemmli buffer and boiled for 5 min at 95 °C. Samples were resolved on a 12% SDS-PAGE gel and visualized by staining with Coomassie dye. Radioactive gels were dried for 30 mins at 80 °C in a gel dryer, exposed to a phosphoscreen and visualized by autoradiography using a Typhoon Scanner (GE Healthcare). To examine autophosphorylation of Yegl mutants (S160FD161G and D161G), reactions were carried out in reaction buffer containing either 10 mM MgCl₂ or 10 mM MnCl₂.

Autophosphorylation assay to examine kinetics of Yegl

Autophosphorylation of Yegl was performed by addition of 5 µCi of [γ -³²P] ATP (Perkin Elmer) to different concentrations of purified Yegl (0.0625, 0.125, 0.25, 0.5, 1, 2 µM) in 10 µl of reaction buffer containing 50 mM Tris pH 7.5, 50 mM KCl, 0.5 mM DTT, 200 µM ATP, 10mM MgCl₂ and 10mM MnCl₂. Reactions were incubated at 37 °C for 30 mins and were stopped using 3X Laemmli buffer and boiled for 5 min at 95 °C. Samples were resolved on a 12% SDS-PAGE gel and visualized by staining with Coomassie dye. Radioactive gels were dried for 30 mins at 80 °C in a gel dryer, exposed to a phosphoscreen and visualized by autoradiography using a Typhoon Scanner (GE Healthcare).

Membrane topology analysis

Construction of Yegl PhoA-lacZα fusion and derivatives

To study the membrane topology of Yegl, plasmid pKR72 was constructed that encodes a dual *pho-lac* reporter in-frame to full length Yegl. The Yegl-PhoA-lacZα was constructed by a PCR-mediated overlap extension (44). First, the *yegl* gene fragment encoding aa 2-646 was PCR amplified using Phusion polymerase (Thermo Scientific) from *E. coli* REL606 genomic DNA using primers (KR223/KR224). Alkaline phosphatase PhoA fragment (aa 21-471) and the alpha fragment of beta galactosidase LacZ (aa 5-60) was PCR amplified from *E. coli* MG1655 genomic DNA using primers (KR225/KR226) for *phoA* and (KR227/KR228) for *lacZα* respectively. Eighteen nucleotides encoding a GSGSGS linker along with eighteen nucleotides from the *phoA* DNA was introduced in the reverse primer KR224 to generate an overlapping region between *yegl* and *phoA*. Similarly, eighteen nucleotides from the *lacZ* DNA was introduced in the reverse primer KR226 to generate an overlapping region between *phoA* and *lacZ*. The *yegl* and *phoA* PCR products were purified and used as templates for a second PCR to create a fused *yegl-linker-pho* DNA that was amplified using primers (KR229/KR226). Sequence for N-terminal His₆ tag was included in the primer KR229. Purified PCR products *yegl-linker-phoA* DNA and *lacZ* DNA were used as templates for third and final round of PCR to create a fused *yegl-linker-pho-lac* fusion that was amplified using primers (KR229/KR228). The PCR product was digested with *Nhe1/Kpn1* and ligated into similarly digested pBAD24 plasmid backbone. Ligation products were transformed into DH5α cells and selected on LB/ampicillin plates. The resulting plasmid pKR72 generated an N-terminal His₆-tagged Yegl-PhoA-lacZα fusion protein. Plasmid cloning was subsequently verified by restriction enzyme digest and DNA sequencing (Operon).

Plasmid derivatives of pKR72 with *pho-lac* fusions were constructed in frame with different codons of Yegl using the Q5 Site-Directed mutagenesis kit (NEB). Primers used to generate *pho-lac* fusions in frame to different codons E439 (KR230/KR231), T432 (KR230/KR232), R414 (KR230/KR233), E390 (KR230/KR234), E340 (KR230/KR235) were designed using NEBaseChanger. Primers were designed such that the 5' ends were annealing back-to-back

following manufacturer's instructions. PCR amplification was performed using Q5 hotstart high fidelity master mix (NEB) with respective primers and 2 ng of Plasmid pKR72 as template DNA. The PCR product was subsequently treated with Kinase/Ligase/Dpn1 mix for 5 mins and transformed in NEB5alpha competent *E. coli* cells (NEB C2987). This method generated plasmids with in-frame *pho-lac* fusions to different codons of YegI at positions E439 (pKR73), T432 (pKR74), R414 (pKR75), E390 (pKR76), E340 (pKR77). pKR78, a derivative of pKR72 lacking *yegI* was used as a control plasmid and was constructed using the same method using primers KR230/KR240. Plasmid cloning was subsequently verified by colony PCR and DNA sequencing (Operon).

Membrane topology assay

To assess membrane topology *in vivo*, *E. coli* DH5α cells were transformed with the resulting *pho-lac* fusion plasmids and streaked on dual indicator plates containing LB agar with 5-bromo-4-chloro-3-indolyl phosphate disodium salt (X-Phos) (Sigma) at a final concentration of 80μg/ml and 6-chloro-3-indolyl-β-D-galactoside (Red-Gal) (Sigma) at a final concentration of 100 μg/ml as indicators. L-arabinose at a final concentration of 0.1% was added as an inducer for reporter expression. 100μg/ml of ampicillin was added for plasmid selection. Plates were incubated at 25°C for 24 h and scanned using EPSON image scanner.

Acknowledgements

We would like to thank present and former members of our lab for helpful discussions. We would like to thank Elizabeth Nagle for performing initial experiments characterizing YegI. This work was supported by an HHMI International Student Fellowship to KR and by NIH grant GM114213 to JD.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

KR performed all of the experiments and KR and JD wrote the manuscript.

References

1. Hunter, T. (1995) Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* 80, 225-236
2. Stock, J. B., Ninfa, A. J., and Stock, A. M. (1989) Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol Rev* 53, 450-490
3. Lim, S., Marcellin, E., Jacob, S., and Nielsen, L. K. (2015) Global dynamics of *Escherichia coli* phosphoproteome in central carbon metabolism under changing culture conditions. *J Proteomics* 126, 24-33
4. Macek, B., Gnad, F., Soufi, B., Kumar, C., Olsen, J. V., Mijakovic, I., and Mann, M. (2008) Phosphoproteome analysis of *E. coli* reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. *Molecular & cellular proteomics : MCP* 7, 299-307
5. Qu, Y., Wu, S., Zhao, R., Zink, E., Orton, D. J., Moore, R. J., Meng, D., Clauss, T. R., Aldrich, J. T., Lipton, M. S., and Pasa-Tolic, L. (2013) Automated immobilized metal affinity chromatography system for enrichment of *Escherichia coli* phosphoproteome. *Electrophoresis* 34, 1619-1626
6. Soares, N. C., Spat, P., Krug, K., and Macek, B. (2013) Global dynamics of the *Escherichia coli* proteome and phosphoproteome during growth in minimal medium. *Journal of proteome research* 12, 2611-2621
7. Pereira, S. F., Goss, L., and Dworkin, J. (2011) Eukaryote-like serine/threonine kinases and phosphatases in bacteria. *Microbiol Mol Biol Rev* 75, 192-212
8. Figueira, R., Brown, D. R., Ferreira, D., Eldridge, M. J., Burchell, L., Pan, Z., Helaine, S., and Wigneshweraraj, S. (2015) Adaptation to sustained nitrogen starvation by *Escherichia coli* requires the eukaryote-like serine/threonine kinase YeaG. *Sci Rep* 5, 17524
9. Zheng, J., He, C., Singh, V. K., Martin, N. L., and Jia, Z. (2007) Crystal structure of a novel prokaryotic Ser/Thr kinase and its implication in the Cpx stress response pathway. *Molecular microbiology* 63, 1360-1371
10. Dorsey-Oresto, A., Lu, T., Mosel, M., Wang, X., Salz, T., Drlica, K., and Zhao, X. (2013) YihE kinase is a central regulator of programmed cell death in bacteria. *Cell Rep* 3, 528-537
11. Kaspy, I., Rotem, E., Weiss, N., Ronin, I., Balaban, N. Q., and Glaser, G. (2013) HipA-mediated antibiotic persistence via phosphorylation of the glutamyl-tRNA-synthetase. *Nat Commun* 4, 3001
12. Germain, E., Castro-Roa, D., Zenkin, N., and Gerdes, K. (2013) Molecular mechanism of bacterial persistence by HipA. *Mol Cell* 52, 248-254
13. Rajagopalan, K., and Dworkin, J. (2018) Identification and biochemical characterization of a novel PP2C-like Ser/Thr phosphatase in *E. coli*. *Journal of bacteriology* 200, e00225-00218

14. Hanks, S. K., and Hunter, T. (1995) Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J* 9, 576-596
15. Young, T. A., Delagoutte, B., Endrizzi, J. A., Falick, A. M., and Alber, T. (2003) Structure of *Mycobacterium tuberculosis* PknB supports a universal activation mechanism for Ser/Thr protein kinases. *Nat Struct Biol* 10, 168-174
16. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46, W296-W303
17. Hanks, S. K., Quinn, A. M., and Hunter, T. (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* 241, 42-52
18. Ruegg, U. T., and Burgess, G. M. (1989) Staurosporine, K-252 and UCN-01: potent but nonspecific inhibitors of protein kinases. *Trends Pharmacol Sci* 10, 218-220
19. Meggio, F., Donella Deana, A., Ruzzene, M., Brunati, A. M., Cesaro, L., Guerra, B., Meyer, T., Mett, H., Fabbro, D., Furet, P., and et al. (1995) Different susceptibility of protein kinases to staurosporine inhibition. Kinetic studies and molecular bases for the resistance of protein kinase CK2. *Eur J Biochem* 234, 317-322
20. Hall, C. L., Tschannen, M., Worthey, E. A., and Kristich, C. J. (2013) IreB, a Ser/Thr kinase substrate, influences antimicrobial resistance in *Enterococcus faecalis*. *Antimicrob Agents Chemother* 57, 6179-6186
21. Liu, Q., Fan, J., Niu, C., Wang, D., Wang, J., Wang, X., Villaruz, A. E., Li, M., Otto, M., and Gao, Q. (2011) The eukaryotic-type serine/threonine protein kinase Stk is required for biofilm formation and virulence in *Staphylococcus epidermidis*. *PLoS One* 6, e25380
22. Arora, G., Sajid, A., Arulanandh, M. D., Misra, R., Singhal, A., Kumar, S., Singh, L. K., Mattoo, A. R., Raj, R., Maiti, S., Basu-Modak, S., and Singh, Y. (2013) Zinc regulates the activity of kinase-phosphatase pair (BasPrkC/BasPrpC) in *Bacillus anthracis*. *Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine* 26, 715-730
23. Boitel, B., Ortiz-Lombardia, M., Duran, R., Pompeo, F., Cole, S. T., Cervenansky, C., and Alzari, P. M. (2003) PknB kinase activity is regulated by phosphorylation in two Thr residues and dephosphorylation by PstP, the cognate phospho-Ser/Thr phosphatase, in *Mycobacterium tuberculosis*. *Molecular microbiology* 49, 1493-1508
24. Rajagopal, L., Clancy, A., and Rubens, C. E. (2003) A eukaryotic type serine/threonine kinase and phosphatase in *Streptococcus agalactiae* reversibly phosphorylate an inorganic pyrophosphatase and affect growth, cell segregation, and virulence. *The Journal of biological chemistry* 278, 14429-14441
25. Srinivasan, V. B., Vaidyanathan, V., Mondal, A., Venkataramaiah, M., and Rajamohan, G. (2012) Functional characterization of a novel Mn²⁺ dependent protein serine/threonine kinase KpnK, produced by *Klebsiella pneumoniae* strain MGH78578. *FEBS Lett* 586, 3778-3786
26. Knighton, D. R., Zheng, J. H., Ten Eyck, L. F., Ashford, V. A., Xuong, N. H., Taylor, S. S., and Sowadski, J. M. (1991) Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* 253, 407-414
27. Nolen, B., Taylor, S., and Ghosh, G. (2004) Regulation of protein kinases; controlling activity through activation segment conformation. *Mol Cell* 15, 661-675
28. Mieczkowski, C., Iavarone, A. T., and Alber, T. (2008) Auto-activation mechanism of the *Mycobacterium tuberculosis* PknB receptor Ser/Thr kinase. *EMBO J* 27, 3186-3197
29. Shakir, S. M., Bryant, K. M., Larabee, J. L., Hamm, E. E., Lovchik, J., Lyons, C. R., and Ballard, J. D. (2010) Regulatory interactions of a virulence-associated serine/threonine phosphatase-kinase pair in *Bacillus anthracis*. *Journal of bacteriology* 192, 400-409

30. Zheng, W., Cai, X., Li, S., and Li, Z. (2018) Autophosphorylation Mechanism of the Ser/Thr Kinase Stk1 From *Staphylococcus aureus*. *Front Microbiol* 9, 758
31. Ravala, S. K., Singh, S., Yadav, G. S., Kumar, S., Karthikeyan, S., and Chakraborti, P. K. (2015) Evidence that phosphorylation of threonine in the GT motif triggers activation of PknA, a eukaryotic-type serine/threonine kinase from *Mycobacterium tuberculosis*. *The FEBS journal* 282, 1419-1431
32. Kall, L., Krogh, A., and Sonnhammer, E. L. (2007) Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. *Nucleic Acids Res* 35, W429-432
33. Alexeyev, M. F., and Winkler, H. H. (1999) Membrane topology of the *Rickettsia prowazekii* ATP/ADP translocase revealed by novel dual pho-lac reporters. *J Mol Biol* 285, 1503-1513
34. Islam, S. T., Taylor, V. L., Qi, M., and Lam, J. S. (2010) Membrane topology mapping of the O-antigen flippase (Wzx), polymerase (Wzy), and ligase (WaaL) from *Pseudomonas aeruginosa* PAO1 reveals novel domain architectures. *MBio* 1
35. Karimova, G., and Ladant, D. (2017) Defining Membrane Protein Topology Using pho-lac Reporter Fusions. *Methods Mol Biol* 1615, 129-142
36. Karimova, G., Robichon, C., and Ladant, D. (2009) Characterization of YmgF, a 72-residue inner membrane protein that associates with the *Escherichia coli* cell division machinery. *Journal of bacteriology* 191, 333-346
37. Fabbro, D., Cowan-Jacob, S. W., and Moebitz, H. (2015) Ten things you should know about protein kinases: IUPHAR Review 14. *Br J Pharmacol* 172, 2675-2700
38. Ortiz-Lombardia, M., Pompeo, F., Boitel, B., and Alzari, P. M. (2003) Crystal structure of the catalytic domain of the PknB serine/threonine kinase from *Mycobacterium tuberculosis*. *The Journal of biological chemistry* 278, 13094-13100
39. Greenstein, A. E., Echols, N., Lombana, T. N., King, D. S., and Alber, T. (2007) Allosteric activation by dimerization of the PknD receptor Ser/Thr protein kinase from *Mycobacterium tuberculosis*. *The Journal of biological chemistry* 282, 11427-11435
40. Hsu, F., Schwarz, S., and Mougous, J. D. (2009) TagR promotes PpkA-catalysed type VI secretion activation in *Pseudomonas aeruginosa*. *Molecular microbiology* 72, 1111-1125
41. Ravikumar, V., Shi, L., Krug, K., Derouiche, A., Jers, C., Cousin, C., Kobir, A., Mijakovic, I., and Macek, B. (2014) Quantitative phosphoproteome analysis of *Bacillus subtilis* reveals novel substrates of the kinase PrkC and phosphatase PrpC. *Molecular & cellular proteomics : MCP* 13, 1965-1978
42. Schmidl, S. R., Gronau, K., Pietack, N., Hecker, M., Becher, D., and Stulke, J. (2010) The phosphoproteome of the minimal bacterium *Mycoplasma pneumoniae*: Analysis of the complete known Ser/ Thr kinome suggests the existence of novel kinases. *Molecular & cellular proteomics : MCP*
43. Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302, 205-217
44. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K., and Pease, L. R. (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77, 61-68

Figure 1

Figure 2

A

B

Figure 3

A

B

Figure 4

Figure 5

A

B

Figure 6

A

B

Table 1

YegJ Fusion	LacZα	PhoA	Localization
R414	-	+	Periplasm
T432	-	+	TM
E340	+	-	Cytoplasm
E390	+	-	Cytoplasm
E439	+	-	Cytoplasm
C646	+	-	Cytoplasm

Supp fig 1: Genome architecture of *yeg* operon

Supp fig 2: Predicted structure of YegI kinase domain resembles a bacterial eSTK

A

B

Escherichia coli eSTK YegI kinase domain (Predicted structure)

Mycobacterium tuberculosis eSTK PknB (PDB:1MRU)

Supp fig 3: Requirement of bivalent cations

Supplementary figure legends

Figure S1: Genome organization of *yeg* operon in different *E. coli* strain backgrounds.

Thick arrows denote orfs. Genes *yegL*, *pphC*, *yegI* encode a predicted Von Willebrand factor, a PP2C-like phosphatase and an eukaryotic-like Ser/Thr kinase respectively. The *yegJ* gene encodes a protein of unknown function. *E. coli* strain backgrounds are indicated on the right.

Figure S2: Predicted structure of Yegl kinase domain (aa 1-302).

(A) Structural overlay of predicted structure of Yegl kinase domain (blue) and *Mycobacterium tuberculosis* eSTK PknB (pink) (Protein data bank PDB:1MRU). Mg^{2+} ion is represented as a grey sphere.

(B) SWISS MODEL algorithm (1) was used to predict structural model of Yegl kinase domain. Structurally important motifs such as P-loop (pink), Helix C (green), Catalytic loop (red), activation loop (yellow) and P+1 loop (black) are indicated in structure. Corresponding amino acid sequences are depicted in Fig 1.

Figure S3: Requirement of bivalent cations.

Autophosphorylation reactions were carried out at 37 °C with 1 μ M of Yegl (WT) in kinase buffer with either 10 mM of $MgCl_2$ / $MnCl_2$, $CaCl_2$ / $NiCl_2$. Reactions were stopped at t=30 mins and run on 12% SDS-PAGE followed by autoradiography

Table S1: Bacterial Strains

Strain	Relevant Genotype	Description	Reference
DH5a			Invitrogen
C43DE3			(2)
LOBSTR-BL21(DE3)			Kerafast (3)
JDE2735	REL606		(4)
JDE3049	pKR31 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged WT YegI from <i>E.coli</i> REL606	This study
JDE3050	pKR81 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged K39D YegI from <i>E.coli</i> REL606	This study
JDE3051	pKR83 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged D141N YegI from <i>E.coli</i> REL606	This study
JDE3052	pKR97 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S160FD161G YegI from <i>E.coli</i> REL606	This study
JDE3053	pKR98 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged D161G YegI from <i>E.coli</i> REL606	This study
JDE3054	pKR110 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S153A YegI from <i>E.coli</i> REL606	This study
JDE3055	pKR111 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S160A YegI from <i>E.coli</i> REL606	This study
JDE3056	pKR112 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S162A YegI from <i>E.coli</i> REL606	This study
JDE3057	pKR62 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S602A YegI from <i>E.coli</i> REL606	This study
JDE3058	pKR63 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S644A YegI from <i>E.coli</i> REL606	This study
JDE3059	pKR64 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S602AS644A YegI from <i>E.coli</i> REL606	This study
JDE3042	pKR73 amp	DH5a expressing N terminal 6X-His tagged YegI (aa 1-439) fused with PhoA lacZα from <i>E.coli</i> REL606	This study
JDE3043	pKR74 amp	DH5a expressing N terminal 6X-His tagged YegI (aa 1-432) fused with	This study

		PhoA lacZ α from E.coli REL606	
JDE3044	pKR75 amp	DH5a expressing N terminal 6X-His tagged YegI (aa 1-414) fused with PhoA lacZ α from E.coli REL606	This study
JDE3045	pKR77 amp	DH5a expressing N terminal 6X-His tagged YegI (aa 1-340) fused with PhoA lacZ α from E.coli REL606	This study
JDE3046	pKR76 amp	DH5a expressing N terminal 6X-His tagged YegI (aa 1-390) fused with PhoA lacZ α from E.coli REL606	This study
JDE3047	pKR77 amp	DH5a expressing N terminal 6X-His tagged YegI (aa 1-646) fused with PhoA lacZ α from E.coli REL606	This study
JDE3048	pKR78 amp	DH5a expressing N terminal 6X-His tagged PhoA (aa 21-471) in frame with lacZ α (aa 5-60) from E.coli REL606	This study
JDE659	pBAD24 amp	DH5a expressing pBAD24	This study

Table S2: Plasmids

Plasmid Name	Description	Reference
pBAD24	Cloning vector, ampR	(5)
pKR31	yegI was amplified from REL606 genomic DNA using KR58/KR59 primers, digested with NcoI/SphI and ligated into pBAD24	This study
pKR81	REL606 K39D yegI was generated by two step PCR using KR58/KR246 and KR245/KR59 primer pairs followed by amplification with KR58/KR59. Resulting product was digested with NcoI/SphI and ligated into pBAD24	This study
pKR83	REL606 D141N yegI was generated by two step PCR using KR58/KR248 and KR247/KR59 primer pairs followed by amplification with KR58/KR59. Resulting product was digested with NcoI/SphI and ligated into pBAD24	This study
pKR97	REL606 S160FD161G yegI was generated by Q5 site directed mutagenesis using KR285/KR286 and pKR31 as a template.	This study
pKR98	REL606 D161G yegI was generated by Q5 site directed mutagenesis using KR285/KR286 and pKR31 as a template.	This study
pKR110	REL606 S153A yegI was generated by two step PCR using KR58/KR158 and KR157/KR59 primer pairs followed by amplification with KR58/KR59. Resulting product was digested with NcoI/SphI and ligated into pBAD24	This study
pKR111	REL606 S160A yegI was generated by two step PCR using KR58/KR306 and KR305/KR59 primer pairs followed by amplification with KR58/KR59. Resulting product was digested with NcoI/SphI and ligated into pBAD24	This study
pKR112	REL606 S162A yegI was generated by two step PCR using KR58/KR308 and KR307/KR59 primer pairs followed by amplification with KR58/KR59. Resulting product was digested with NcoI/SphI and ligated into pBAD24	This study
pKR62	REL606 S602A yegI was generated by two step PCR using KR58/KR153 and KR152/KR59 primer pairs followed by amplification with KR58/KR59. Resulting product was digested with NcoI/SphI and ligated into pBAD24	This study
pKR63	REL606 S644A yegI was generated by two	This study

	step PCR using KR58/KR155 and KR154/KR59 primer pairs followed by amplification with KR58/KR59. Resulting product was digested with NcoI/SphI and ligated into pBAD24	
pKR64	REL606 S602AS644A <i>yegI</i> was generated by two step PCR using pKR62 as template and KR58/KR155 and KR154/59 primer pairs followed by amplification with KR58/KR59. Resulting product was digested with NcoI/SphI and ligated into pBAD24	This study
pKR72	REL606 <i>yegI-phoA-lacZ</i> fusion was generated by overlap extension PCR using KR223/KR224 , KR225/KR226, KR227/KR228 primer pairs followed by amplification with KR229/KR228. Resulting product was digested with NheI/KpnI and ligated into pBAD24	This study
pKR73	REL606 <i>YegI E439-phoA lacZ</i> was generated by Q5 site directed mutagenesis using KR230/KR231 and pKR72 as a template.	This study
pKR74	REL606 <i>YegI T432-phoA lacZ</i> was generated by Q5 site directed mutagenesis using KR230/KR232 and pKR72 as a template.	This study
pKR75	REL606 <i>YegI R414-phoA lacZ</i> was generated by Q5 site directed mutagenesis using KR230/KR233 and pKR72 as a template.	This study
pKR76	REL606 <i>YegI E390-phoA lacZ</i> was generated by Q5 site directed mutagenesis using KR230/KR234 and pKR72 as a template.	This study
pKR77	REL606 <i>YegI E340-phoA lacZ</i> was generated by Q5 site directed mutagenesis using KR230/KR235 and pKR72 as a template.	This study
pKR78	Control <i>phoA lacZ</i> was generated by Q5 site directed mutagenesis using KR230/KR240 and pKR72 as a template.	This study

Table S3: Primers

Oligo	Sequence (5'-3')	Features	Origin
KR58	ggaaaac <u>catggc</u> acatcatcatcatcatcatAAACCCACTTTATATACTGCTAC	<u>NcoI</u> ; 6X-His tag	This study
KR59	cgtccag <u>catgc</u> TCAGCAGCGGCTTAAATCTG	<u>SphI</u> ; STOP	This study
KR152	GGCTGGAAG GCG GCGTTGAC	S602A	This study
KR153	GTCAACG CGCTT CCAGCC	S602A	This study
KR154	GCAGATTTAG CCCG CTGCTGA	S644A	This study
KR155	TCAGCAGCG GGCT AAATCTGC	S644A	This study
KR157	GTCGCGAC GCC AAAGTGGT	S153A	This study
KR158	ACCACTTT GCG TCGCGAC	S153A	This study
KR223	<u>ccatggc</u> acatcatcatcatcatcatAAACCCACTTTATATACTGCTACT	<u>NcoI</u> ; 6X-His tag	This study
KR224	catttctggtgtccgggcgctaccgctgccgtaccGCAGCGGCTTAAATCTGCCTG	PhoA; Linker	This study
KR225	GCCCGGACACCAGAAATG	PhoA	This study
KR226	gacggccagtgaatccgtTTTCAGCCCCAGAGC GGC	lacZ;PhoA	This study
KR227	ACGGATTCACTGGCCGTCGTT	lacZ	This study
KR228	cgtccaggtac <u>ctta</u> GCGCCATTCGCCATTCAGGC	<u>KpnI</u> ; STOP	This study
KR229	ggaaaagctagcaggaggaattCCATGGCACATCATCATCATCAT	Nhe1; 6X-His-tag	This study
KR230	GGTAGCGGCAGCGGTAGC	Q5 SDM Linker	This study
KR231	TTCTGCTTTATAAGCTTTGCTTGTCAGACTG	Q5 SDM_E439	This study
KR232	TGTCAGACTGCCAATAATCCAGATAGC	Q5_SDM_E432	This study
KR233	ACGCGGTTCTGCCTGAAGG	Q5_SDM_E414	This study
KR234	TTCGCGCCGTAACAGGCC	Q5_SDM_E390	This study
KR235	TTCGCCGAGATCAATAAAATAGATAACGCC	Q5_SDM_E340	This study
KR240	ATGATGATGATGATGATGTGCCATGGTG	Q5_SDM_phoAlacZ	This study
KR245	AGCGTCGCC GAC ATTTATCAC	K39D	This study

KR246	GTGATAAAT GT CGGCGACGCT	K39D	This study
KR247	GTCGTGGGGA AC GTAACCA	D141N	This study
KR248	TGGTTTAC GT TCCCCACGAC	D141N	This study
KR285	GTTGATCGATTTT GG CTCCTTTCAGAT TAACGCCAATG	S160FD161G	This study
KR286	TGGTTTAC GT TCCCCACGAC	S160FD161G	This study
KR305	GTTGATCGAT GCT GACTCCTTT	S160A	This study
KR306	AAAGGAGTC AG CATCGATCAAC	S160A	This study
KR307	CGATAGTGAC GC CTTTCAGAT	S162A	This study
KR308	ATCTGAA AGGC GTCACATATCG	S162A	This study

Homology sequences to gene are shown in uppercase. Non homology sequences are indicated in lower case including His₆-tag (*italics*), restriction enzyme sites (*underlined*). Point mutations are highlighted in bold.

References

1. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-W303
2. Miroux, B., and Walker, J. E. (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* **260**, 289-298
3. Andersen, K. R., Leksa, N. C., and Schwartz, T. U. (2013) Optimized *E. coli* expression strain LOBSTR eliminates common contaminants from His-tag purification. *Proteins* **81**, 1857-1861
4. Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., Lenski, R. E., and Kim, J. F. (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243-1247
5. Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of bacteriology* **177**, 4121-4130