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

Krithika Rajagopalan[#] and Jonathan Dworkin*

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

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

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

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

Keywords: Ser/Thr phosphorylation, YegI

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, *yegI*, and demonstrate that it encodes a novel Ser/Thr kinase. YegI 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 YegI 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 YegI, which we demonstrate is an integral membrane sSTK. Despite differences in conserved residues observed in other eSTKs, we show that YegI is an active kinase and is sensitive to the kinase inhibitor staurosporine.

54 Additionally, we show that YegI kinase activation follows second order kinetics, consistent with

55 the proposed oligomerization-dependent activation of eSTKs.

56 **Results**

57 **YegI is a novel eukaryotic-like Ser/Thr kinase**

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

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

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

82 all the structural features seen in *M. tuberculosis* eSTK PknB (Supplementary Fig. 2A).
83 Specifically, YegI adopts the canonical bilobed structure seen in eSTKs. The N-terminal lobe of
84 the YegI kinase domain contains the ATP binding P-loop and the helix C structure (Supplementary
85 Fig. 2B) which is responsible for proper orientation of the active site with the substrate. The C-
86 terminal lobe includes the catalytic loop containing all the residues required for catalysis and an
87 activation loop that is important for kinase activation by autophosphorylation. Thus, although the
88 YegI sequence is only ~20% identical to PknB (Fig.1), the similarity of the predicted structure of
89 YegI to that of PknB (Supplementary Fig. 2A), strongly suggests that YegI is an eSTK.
90

91 **YegI is an active kinase**

92 We previously demonstrated that a His_6-tagged version of YegI is capable of autophosphorylation
93 (13) despite the absence of three of the ten absolutely conserved residues (Fig. 1). To confirm
94 that YegI belongs to the eSTK family, we examined the consequences of mutations in the
95 conserved lysine at position 39 and a conserved aspartic acid site at position 141, which play a
96 crucial role in ATP binding and catalysis, respectively, in this family (17). We generated YegI
97 mutant proteins carrying either lysine to aspartic acid (K39D) and aspartic acid to asparagine
98 (D141N) changes. The mutant proteins were expressed and purified in the same manner as the
99 wildtype YegI. Kinase activity was demonstrated using an *in vitro* radioactive kinase assay using
100 radiolabeled γ-[^{32}P]ATP. Mutation of either Lys-39 or Asp-141 resulted in complete loss of
101 autophosphorylation activity as compared to the wildtype protein indicating that both residues are
102 essential for kinase activity (Fig. 2A).

103 We examined how the lack of some of the conserved residues affected YegI function in a
104 series of standard assays of bacterial Ser/Thr kinases. For example, staurosporine is widely used
105 as a small molecule inhibitor for Ser/Thr kinases (18) and bacterial eSTKs are typically sensitive
106 to staurosporine with IC_{50} in the μM range (19). Sensitivity of the YegI kinase to staurosporine
107 was assessed by incubating YegI with increasing concentrations of staurosporine. YegI shows an

108 IC_{50} of ~50 μM (Fig. 2B), comparable to the *Enterococcus faecalis* eSTK IreK (IC_{50} of 150 μM
109 (20)) and the *Staphylococcus epidermis* eSTK Stk (IC_{50} of 320 μM (21)). Bacterial eSTKs require
110 the presence of bivalent cations for activity (22-25). We therefore investigated the effect of
111 different bivalent cations on YegI activity by performing radioactive kinase assays in the presence
112 of either $MgCl_2$/$MnCl_2$/$CaCl_2$/$NiCl_2$. Similar to many eSTKs, YegI catalyzed autophosphorylation
113 only in the presence of $MgCl_2$ and $MnCl_2$, albeit to different levels. YegI displayed much higher
114 autophosphorylation in the presence of Mn^{2+} ion than Mg^{2+} ion suggesting that YegI is a Mn^{2+}
115 dependent kinase (Supplementary Fig. 3). The concentration dependence of YegI kinase activity
116 on Mg^{2+} or Mn^{2+} was measured and the optimal $MnCl_2$ concentration was determined to be
117 between 0.1-1 mM whereas optimal $MgCl_2$ concentration was between 1-10 mM (Fig.3A)
118 indicating that YegI kinase is more active in the presence of $MnCl_2$ than $MgCl_2$.
119 In eSTKs, the highly conserved DFG motif present in domain VII coordinates Mg^{2+} ion
120 binding (26,27). However, amino acid sequence alignment reveals that YegI does not contain the
121 DFG sequence found in bacterial eSTKs but instead the sequence DSD (Fig. 1; Fig. 3B). To
122 examine if addition of the DFG motif to YegI alters preference to $MgCl_2$, we generated a mutant
123 protein carrying changes of Ser-160 to a phenylalanine and Asp-161 to a glycine (S160F, D161G).
124 Additionally, we generated a mutant protein carrying a change of Asp-161 to a glycine (D161G)
125 which would change the DSD sequence to the DSG consensus motif. Mutants were expressed
126 and purified in the same manner as wildtype YegI and kinase activity in the presence of
127 $MgCl_2$/$MnCl_2$ was assessed. Surprisingly, replacement of DSD to DFG or to DSG completely
128 inactivated the kinase (Fig. 3B) indicating that the DSD motif is indispensable for YegI
129 autophosphorylation.
130
131 **Autophosphorylation of YegI**
132 Autophosphorylation is a common model for activation of eukaryotic STKs (27), therefore
133 characterization of the kinetics of autophosphorylation and the sites of autophosphorylation can

134 provide insight into how bacterial homologs of these kinases are activated. We performed
135 standard autophosphorylation assay (Fig. 4A) using a range of YegI concentrations and
136 measured formation of radiolabeled phosphorylated YegI (Fig. 4B). Our results indicate that YegI
137 autophosphorylation is not first order and are consistent with YegI undergoing oligomerization. To
138 map the sites of autophosphorylation, His_6-YegI that had been subjected to an *in vitro* kinase
139 reaction was analyzed by mass spectrometry. Five phosphopeptides that each included at least
140 a single serine residue were identified, leading to the conclusion that Ser-153, Ser-160, and Ser-
141 162 in the predicted activation loop and Ser-602 and Ser-644 at the C-terminus of YegI were
142 phosphorylated (Fig. 5A). To characterize the functional importance of the C-terminal
143 phosphosites, we generated single and double phosphoablative serine to alanine mutants. The
144 mutant proteins were expressed and purified in the same manner as the wildtype protein and
145 assessed for loss of autophosphorylation using a radioactive kinase assay. Since mutation of both
146 Ser-602 and Ser-644 to alanine resulted in only a modest decrease in overall
147 autophosphorylation, phosphorylation of these residues is not necessary for normal YegI activity
148 (Fig.5B).

149 Since eSTKs undergo autophosphorylation on Ser or Thr residues in the activation loop of
150 the kinase domain (28-31), we investigated the phosphosites (Ser-153, Ser-160, Ser-162) in the
151 putative YegI activation loop. All three residues are located between Motif VI and VII in close
152 proximity to the catalytic loop and activation loop suggesting that phosphorylation of these
153 residues could be important for kinase activity. We mutated each Ser to an alanine to examine if
154 phosphorylation of these Ser residues impacted kinase activity. We observed that mutating Ser-
155 162, but not the other phosphosites (Ser-153 or Ser-160), resulted in a complete loss of
156 autophosphorylation (Fig. 5B), demonstrating that autophosphorylation of Ser-162 is important
157 for YegI activation.

158

159 **YegI is an integral membrane protein**

160 We observed that purification of recombinant His_6-YegI required the presence of the anionic
161 detergent N-lauryl sarcosine suggesting that YegI is a membrane protein. Consistently, the
162 membrane topology algorithm Phobius (32) predicts that YegI consists of two transmembrane
163 domains joined by a short periplasmic loop with both the N- and the C-terminus located in the
164 cytosol (Fig. 6A). We used the PhoALacZα dual reporter system (33-36) to experimentally validate
165 this predicted topology. In this system, alkaline phosphatase (PhoA, aa 21-471) lacking the
166 periplasmic localization signal is fused in frame with the α-fragment of *E. coli* β-galactosidase
167 (LacZα, aa 5-60), creating PhoALacZα. Localization of a protein of interest is then analyzed by
168 fusing it to PhoALacZα. Since PhoA is active only in the periplasm and LacZ is only active in the
169 cytoplasm, the PhoALacZα fusion displays high phosphatase activity but no β-galactosidase
170 activity when targeted to the periplasm or to the membrane. Conversely, the fusion displays high
171 β-galactosidase activity but no phosphatase activity when it is targeted to the cytoplasm.
172 PhoAlacZα reporter fusions can be assessed using dual indicator plates with PhoA or LacZ
173 specific chromogenic substrates that allow identification of periplasmic or cytoplasmic fusions
174 (35). Specifically, cells expressing PhoAlacZα protein fusions with high levels of phosphatase
175 activity process the X-phos (5-bromo-4-chloro-3-indolyl phosphate) substrate yielding blue
176 colored colonies, while cells expressing PhoAlacZα protein fusions with high levels of β-
177 galactosidase activity process the Red-gal (6-chloro-3-indolyl-β-D-galactoside) substrate yielding
178 red colored colonies.

179 We constructed C-terminal PhoALacZα fusions to YegI at amino acid positions 340, 390,
180 414, 432, 439 and 646 (Fig. 6A) and analyzed them on dual indicator plates containing X-phos
181 and Red-gal. DH5α cells carrying empty backbone vector alone served as a negative control
182 ("vector") and were colorless, and DH5α cells expressing PhoALacZα without YegI ("PhoALacZα")
183 served as a positive control and were red/pink indicative of cytoplasmic location (Fig.6B).
184 Consistent with the *in silico* prediction, PhoALacZα fusions to YegI at amino acid positions 340,
185 390, 439 and 646 yielded red/pink colored colonies suggestive of high β-galactosidase activity

186 and cytoplasmic localization (Fig. 6B; Table 1). In contrast, PhoALacZα fusions of YegI at amino
187 acid positions 414 and 432 yielded blue colored colonies on indicator plates suggestive of high
188 alkaline phosphatase activity and periplasmic/TM localization (Fig. 6B). In summary, the
189 PhoAlacZα reporter fusions confirmed the bioinformatic prediction that YegI is an integral inner
190 membrane protein containing two TM domains connected by a short periplasmic loop, with both
191 the N- and C-terminus located in the cytosol.
192
193 **Discussion**
194 Ser/Thr phosphorylation is an important post translational modification that is mediated by Ser/Thr
195 kinases and cognate Ser/Thr protein phosphatases which have been identified in phylogenetically
196 diverse organisms. Although Ser/Thr phosphorylation in the Gram-negative bacterium *E. coli* has
197 been known for >10 years, the enzymes responsible for this reversible modification and their
198 specific substrates are not well understood. In this study, we have described the characterization
199 of YegI, a novel Ser/Thr kinase *E. coli.*
200 *In silico* analysis revealed that YegI belongs to the class of eSTKs that contain all the 11 key
201 structural subdomains that are characteristic of eukaryotic protein kinases (17). Homology search
202 using YegI kinase domain revealed that it was similar to kinase domains of several bacterial
203 eSTKs (Supplementary Fig. 2A). Unlike other bacterial eSTKs, however YegI includes only seven
204 of the ten absolutely conserved residues responsible for ATP binding, catalysis, and substrate
205 recognition (14,17) (Fig. 1). Despite the absence of these conserved residues, YegI encodes an
206 active protein kinase (Fig. 2A) that undergoes autophosphorylation primarily on serine residues
207 (Fig. 5A) Similar to most bacterial eSTKs, autophosphorylation of YegI on serine residues in the
208 activation loop is critical for catalytic activity of YegI (Fig. 5B).
209 Although the predicted structure of YegI closely resembled the structure of the bacterial eSTK
210 PknB (Fig. 4.1D), the YegI sequence shows striking deviations in the conserved regions
211 particularly in the catalytic loop and the Mg^{2+}-binding DFG loop (Fig. 1). Specifically, the YegI

212 sequence contains the catalytic loop VGDXXQXS as opposed to the canonical catalytic loop
213 HRDXXPXN found in almost all other bacterial eSTKs. The conserved histidine residue in the
214 catalytic loop is crucial for binding both the carbonyl group of catalytic aspartate in the catalytic
215 loop and for making a hydrophobic contact with the phenylalanine in the DFG loop. Additionally,
216 YegI lacks the highly conserved DFG (aspartate-phenylalanine-glycine) loop in motif VII (Fig.1,
217 3B) which recognizes the Mg^{2+}-bound ATP (37). Instead, YegI contains a DSD (aspartate-serine-
218 aspartate) sequence suggesting either that YegI might be an inactive kinase or that the DFG loop
219 is not essential for every kinase. Our results suggest the latter, since, despite these differences
220 in highly conserved residues, YegI is an active kinase (Fig. 2A) that requires the bivalent cations
221 Mg^{2+}/Mn^{2+} for activity with preference for Mn^{2+} (Fig. 3A). We speculated that addition of DFG loop
222 could help kinase activation. However, replacing the DSD loop in YegI with the canonical DFG
223 loop failed to increase activity or preference for Mg^{2+} ion and instead completely abolished activity
224 (Fig. 3B).

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

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

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

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

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

264 as ~100 cellular proteins were found to be phosphorylated on Ser/Thr residues in *E. coli* (4). A
265 useful strategy to identify authentic in vivo substrates is comparative phosphoproteomics where
266 the phosphoproteomes of a strain lacking either the eSTK or partner eSTP are compared with the
267 WT strain. This methodology has been utilized previously for identifying phospho substrates of *B.*
268 *subtilis* PrkC using deletion strains of eSTK PrkC and eSTP PrpC (41). This methodology has
269 also been successfully utilized in *M. pneumoniae*, where 4 out of the 63 phosphorylated proteins
270 were specifically enriched in a deletion strain of novel eSTK PrkC (42). Therefore,
271 phosphoproteomic analysis might serve as a useful tool in identifying phosphorylated substrates
272 of YegI by examining Δ*yegI* strains.
273
274 **Figure legends**
275
276 **Figure 1. Amino acid sequence alignment of YegI with bacterial eSTKs**. YegI (*Escherichia*
277 *coli*) was aligned to bacterial eSTKs PrkC (*Bacillus subtilis),* PknB *(Mycobacterium tuberculosis)*
278 and PpkA (*Pseudomonas aeruginosa)* using T-coffee (43). Alignment was visualized using
279 Jalview. Absolutely conserved residues are shaded in dark blue whereas residues conserved in
280 three out of the four kinases are shaded in light blue. The ten absolutely conserved residues found
281 in bacterial eSTKs are depicted with arrows. Hanks conserved domains are depicted on top of
282 the sequence as roman numerals (I-XI). YegI (aa 1- 303), PrkC (aa 1-269), PknB (aa 1-275) and
283 PpkA (aa 1-265) which includes the kinase domain was used for alignment. Structurally and
284 functionally important motifs are indicated below the sequence.
285
286 **Figure 2. YegI is an active kinase**
287 **(A) Kinase activity of YegI**: Autophosphorylation reactions were carried out at 37 ºC with 0.2 µM
288 of YegI (WT or K39D or D141N) in kinase buffer (50 mM Tris pH 7.5, 50 mM KCl, 1 mM DTT, 10
289 mM $MgCl_2$, 10 mM $MnCl_2$, 200 µM cold ATP and 5 µCi γ-[^{32}P]ATP. Reactions were stopped at

290 t=30 mins and run on 12 % SDS-PAGE followed by autoradiography. Molecular weights (kDa)
291 are indicated on the right of the gel.
292 (B) **Sensitivity to staurosporine:** Autophosphorylation reactions were carried out at 37 ºC with
293 1 µM of YegI (WT) in kinase buffer without cold ATP. Different concentrations of staurosporine
294 (µM) was added to the reaction at indicated concentrations. Reactions were stopped at t=30 mins
295 and run on 12% SDS-PAGE followed by autoradiography. Molecular weights are indicated on the
296 right of the gel.
297
298 **Figure 3. YegI is a Mn^{2+} dependent kinase**
299 **(A) Requirement of Mg^{2+}/Mn^{2+} for kinase activity:** Autophosphorylation reactions were carried
300 out at 37 ºC with 1 µM of YegI (WT) in kinase buffer with indicated concentrations of $MgCl_2$/ $MnCl_2$.
301 Reactions were stopped at t=30 mins and run on 12% SDS-PAGE followed by autoradiography.
302 Molecular weights are indicated on the right of the gel.
303 **(B) Requirement of DFG motif**: Autophosphorylation reactions were carried out at 37 ºC with
304 0.2 µM of YegI (WT or S160FD161G or D161G) in kinase buffer with either 10 mM $MgCl_2$ or $MnCl_2$
305 as described in Supp. Figure 3
306
307 **Figure 4**
308 **Kinetics of YegI autophosphorylation**
309 (A) **Autophosphorylation of YegI**: Autophosphorylation reactions were carried out at 37 ºC with
310 indicated concentrations of YegI (WT) in kinase buffer. Reactions were stopped after 30 min and
311 run on 12% SDS-PAGE followed by autoradiography. Molecular weights are indicated on the right
312 of the gel.
313 (B) **Autophosphorylation of YegI follows second order kinetics**: Kinetics of YegI
314 autophosphorylation was assessed using different concentrations of YegI. % P32 incorporation
315 was calculated relative to autophosphorylation in the presence of 2 µM of YegI from Fig 4(A).

316

317 **Figure 5**

318 **YegI undergoes autophosphorylation on serine residues in the kinase domain and the C-**

319 **terminus**

320 (A)**Graphical representation of YegI phosphorylation sites:** Domain structure depicting

321 phosphorylation sites following mass spectrometry analysis of autophosphorylation.

322 Phosphorylated residues are indicated by red circles.

323 (B) **Sites of autophosphorylation on YegI**: Reactions were carried out at 37 ºC with 1 µM of

324 YegI (WT or phosphoablative mutants) in kinase buffer. Reactions were stopped at t=30 mins and

325 run on 12 % SDS-PAGE followed by autoradiography. Molecular weights (kda) are indicated on

326 the right of the gel.

327

328 **Figure 6**

329 **YegI is an integral membrane protein**

330 A) **Predicted topology of YegI using Phobius.** Amino acid positions at which phoA-lacZα

331 reporters are fused are indicated along with associated phenotypes. Phenotypes of phoAlacZα

332 fusions due to LacZ activity are denoted as red ellipses while PhoA activity are denoted as blue

333 ellipses. Predicted transmembrane domains (I and II) are depicted as grey rectangles.

334 (B) **Membrane topology analysis:** DH5α transformed with different phoAlacZα reporters were

335 streaked on an LB plate containing 5-bromo-4-chloro-3-indolyl phosphate disodium salt (X-Phos,

336 80µg/ml), 6-chloro-3-indolyl-β-D-galactoside (Red-Gal, 100 µg/ml, 0.1 % w/v arabinose and 100

337 µg/ml of ampicillin. Control pBAD24: control strain expressing pBAD24 empty vector, control

338 phoAlacZα: strain expressing phoAlacZα in pBAD24 without YegI. Amino acid positions at which

339 phoAlacZα reporters are fused are indicated.

340

341

342 **Table 1** Localization of different phoAlacZα reporter fusions.

343

344

345 **Materials and Methods**

346 **Bacterial strains and growth conditions**

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

353

354 **Cloning and expression of YegI and YegI-NTD**

355 The *yegI* gene was PCR amplified using Phusion polymerase (Thermo Scientific) from *E. coli*
356 REL606 genomic DNA using primers (KR58/ KR59). Sequence for an N-terminal His_6 tag was
357 included in the primer. The PCR product was digested with *NcoI/SphI* and ligated into similarly
358 digested pBAD24 backbone. Ligation products were transformed in DH5α cells and selected on
359 LB/ampicillin plates. The resulting plasmid pKR31 generated an N-terminal His_6-tagged YegI
360 fusion protein. Plasmid cloning was subsequently verified by restriction enzyme digest and DNA
361 sequencing (Operon).

362 For protein expression, plasmid was transformed into *E. coli* LOBSTR (Kerafast) cells and
363 plated on LB ampicillin (100 µg/ml) agar plates. Single colonies were inoculated into 3 ml LB
364 supplemented with ampicillin (100 µg/ml) for overnight cultures. The next day, a dilution of 1:250
365 in 500 ml LB was grown to an OD_{600} of 0.6-0.8. YegI recombinant protein was induced by addition
366 of arabinose (0.2% w/v) for 3 h at 18 ºC. Cells were harvested at 6000 x g for 15 min at 4 ºC.

367 Pellets were washed with ice-cold 50 mM EDTA and centrifuged at 7000 rpm for 15 min at 4 ºC.
368 Washed pellets were saved at -80 ºC until use.
369
370 **Oligonucleotide site directed mutagenesis for YegI point mutants**
371 Point mutation of lysine residue K39 was generated by two step overlap PCR mutagenesis using
372 primer pairs (KR58/KR246) and (KR245/59) with Phusion polymerase. A second PCR was
373 performed with primers (KR58/KR59) using primary PCR products as a template and the
374 subsequent PCR products were digested with *Nco1/Sph1* and ligated into similarly digested
375 pBAD24 to generate an N-terminal His_6 tagged K39D YegI fusion protein. Plasmid cloning was
376 subsequently verified by restriction enzyme digest and DNA sequencing (Operon).
377 Point mutation of aspartic acid residue D141 was generated by two step overlap PCR
378 mutagenesis using primer pairs (KR58/ KR248) and (KR247/KR59) with Phusion polymerase. A
379 second PCR was performed with primers (KR58/KR59) using primary PCR products as a
380 template and the subsequent PCR products were digested with *Nco1/Sph1* and ligated into
381 similarly digested pBAD24 to generate an N-terminal His_6 tagged D141N YegI fusion protein.
382 Plasmid cloning was subsequently verified by restriction enzyme digest and DNA sequencing
383 (Operon).
384 Phosphoablative point mutants used in the study were generated using the same strategy
385 as above and details of plasmid construction are described in Supplementary table S2.
386
387 **Purification of recombinant YegI**
388 Frozen pellets were suspended in lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl, 10 mM β-
389 mercaptoethanol, 1 mM PMSF and 2% w/v sarkosyl) and incubated at room temperature for
390 overnight lysis. Cells were subsequently lysed using sonication. Lysates were cleared by
391 centrifugation at 15000 x g for 30 min at 4 ºC and incubated in Pierce 5 ml columns with Ni-NTA
392 agarose beads (Qiagen) at 4 ºC for 1 h. Lysate was allowed to flow through and beads were

393 washed 10 column volumes of wash buffer (50 mM Tris pH 8.0, 200 mM NaCl, 30 mM imidazole,
394 10 mM β-mercaptoethanol, 0.05% w/v sarkosyl). His tagged protein was eluted using 300 mM
395 imidazole in 50 mM Tris pH 8.0, 200 mM NaCl, 10 mM β-mercaptoethanol and 0.05% w/v
396 sarkosyl. Elution fractions were tested on 12% SDS-PAGE gel and fractions containing protein
397 were pooled and dialyzed overnight at 4 ºC using Snakeskin dialysis tubing 10K MWCO (Thermo
398 Scientific) in kinase storage buffer (20 mM Tris pH 8.0, 125 mM NaCl, 10% glycerol, 1 mM DTT).
399 Dialyzed protein was assessed for purity using 12% SDS-PAGE gel and stored at -80 ºC. YegI
400 variant proteins were purified using the same protocol. Proteins were stored at -80 ºC.
401
402 **Autophosphorylation assays**
403 **Autophosphorylation assay using YegI point mutants**
404 Autophosphorylation of YegI and point mutants was performed by addition of 5 µCi of [γ-^{32}P] ATP
405 (Perkin Elmer) to 0.2 µM of purified YegI or YegI point mutants in 10 µl of reaction buffer containing
406 50 mM Tris pH 7.5, 50 mM KCl, 0.5 mM DTT, 200 µM ATP, 10mM $MgCl_2$ and 10mM $MnCl_2$.
407 Reactions were incubated at 37 ºC for 30 mins and were stopped using 3X Laemmli buffer and
408 boiled for 5 min at 95 ºC. Samples were resolved on a 12% SDS-PAGE gel and visualized by
409 staining with Coomassie dye. Radioactive gels were dried for 30 mins at 80 ºC in a gel dryer,
410 exposed to a phosphoscreen and visualized by autoradiography using a Typhoon Scanner (GE
411 Healthcare).
412 **Inhibition of autophosphorylation by staurosporine**
413 Staurosporine sensitivity experiments were performed using 1 µM of YegI in 10 µl of reaction
414 buffer containing 50 mM Tris pH 7.5, 50 mM KCl, 0.5 mM DTT, 10 mM $MgCl_2$ and 10 mM $MnCl_2$.
415 Staurosporine (Calbiochem) was dissolved in DMSO to a obtain a stock concentration of 10 mM.
416 Staurosporine stock (10 mM) was diluted in DMSO and added to the above kinase reaction to
417 obtain final concentrations in the range of 12.5-100 µM. DMSO was added to the no inhibitor
418 control. Reactions were incubated at 25 ºC for 10mins followed by addition of 5 µCi of [γ-^{32}P] ATP

419 (Perkin Elmer). Reactions were incubated for additional 30 mins at 37 ºC and were stopped using
420 3X Laemmli buffer. Samples were boiled for 5 min at 95 ºC and resolved on a 12% SDS-PAGE
421 gel. Gel was visualized by staining with Coomassie dye. Radioactive dried gel was exposed and
422 visualized by autoradiography using a Typhoon Scanner (GE Healthcare).

423 **Autophosphorylation assay using different bivalent cations**

424 Autophosphorylation of YegI was performed by addition of 5 µCi of [γ-^{32}P] ATP (Perkin Elmer) to
425 1 µM of purified YegI in 10 µl of reaction buffer containing 50 mM Tris pH 7.5, 50 mM KCl, 0.5
426 mM DTT, 200 µM ATP. $MgCl_2$ or $MnCl_2$ was added to a final concentration of 0.01, 0.1,1 and 10
427 mM. Reactions were incubated at 37 ºC for 30 mins and were stopped using 3X Laemmli buffer
428 and boiled for 5 min at 95 ºC. Samples were resolved on a 12% SDS-PAGE gel and visualized
429 by staining with Coomassie dye. Radioactive gels were dried for 30 mins at 80 ºC in a gel dryer,
430 exposed to a phosphoscreen and visualized by autoradiography using a Typhoon Scanner (GE
431 Healthcare). To examine autophosphorylation of YegI mutants (S160FD161G and D161G),
432 reactions were carried out in reaction buffer containing either 10 mM $MgCl_2$ or 10 mM $MnCl_2$.

433 **Autophosphorylation assay to examine kinetics of YegI**

434 Autophosphorylation of YegI was performed by addition of 5 µCi of [γ-^{32}P] ATP (Perkin Elmer) to
435 different concentrations of purified YegI (0.0625,0.125,0.25,0.5,1,2 µM) in 10 µl of reaction buffer
436 containing 50 mM Tris pH 7.5, 50 mM KCl, 0.5 mM DTT, 200 µM ATP, 10mM $MgCl_2$ and 10mM
437 $MnCl_2$. Reactions were incubated at 37 ºC for 30 mins and were stopped using 3X Laemmli buffer
438 and boiled for 5 min at 95 ºC. Samples were resolved on a 12% SDS-PAGE gel and visualized
439 by staining with Coomassie dye. Radioactive gels were dried for 30 mins at 80 ºC in a gel dryer,
440 exposed to a phosphoscreen and visualized by autoradiography using a Typhoon Scanner (GE
441 Healthcare).

442

443 **Membrane topology analysis**

444 **Construction of YegI PhoA-lacZα fusion and derivatives**

445 To study the membrane topology of YegI, plasmid pKR72 was constructed that encodes a dual
446 *pho-lac* reporter in-frame to full length YegI. The YegI-PhoA-lacZα was constructed by a PCR-
447 mediated overlap extension (44). First, the *yegI* gene fragment encoding aa 2-646 was PCR
448 amplified using Phusion polymerase (Thermo Scientific) from *E. coli* REL606 genomic DNA using
449 primers (KR223/KR224). Alkaline phosphatase PhoA fragment (aa 21-471) and the alpha
450 fragment of beta galactosidase LacZ (aa 5-60) was PCR amplified from *E. coli* MG1655 genomic
451 DNA using primers (KR225/KR226) for *phoA* and (KR227/KR228) for *lacZα* respectively.
452 Eighteen nucleotides encoding a GSGSGS linker along with eighteen nucleotides from the *phoA*
453 DNA was introduced in the reverse primer KR224 to generate an overlapping region between
454 *yegI* and *phoA*. Similarly, eighteen nucleotides from the *lacZ* DNA was introduced in the reverse
455 primer KR226 to generate an overlapping region between *phoA* and *lacZ*. The *yegI* and *phoA*
456 PCR products were purified and used as templates for a second PCR to create a fused *yegI-*
457 *linker-pho* DNA that was amplified using primers (KR229/KR226). Sequence for N-terminal His_6
458 tag was included in the primer KR229. Purified PCR products *yegI-linker-phoA* DNA and *lacZ*
459 DNA were used as templates for third and final round of PCR to create a fused *yegI-linker-pho-*
460 *lac* fusion that was amplified using primers (KR229/KR228). The PCR product was digested with
461 *Nhe1/Kpn1* and ligated into similarly digested pBAD24 plasmid backbone. Ligation products were
462 transformed into DH5α cells and selected on LB/ampicillin plates. The resulting plasmid pKR72
463 generated an N-terminal His_6-tagged YegI-PhoA-lacZα fusion protein. Plasmid cloning was
464 subsequently verified by restriction enzyme digest and DNA sequencing (Operon).

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

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

479 **Membrane topology assay**

480 To assess membrane topology *in vivo*, *E. coli* DH5α cells were transformed with the resulting pho-
481 lac fusion plasmids and streaked on dual indicator plates containing LB agar with 5-bromo-4-
482 chloro-3-indolyl phosphate disodium salt (X-Phos) (Sigma) at a final concentration of 80μg/ml and
483 6-chloro-3-indolyl-β-D-galactoside (Red-Gal) (Sigma) at a final concentration of 100 μg/ml as
484 indicators. L-arabinose at a final concentration of 0.1% was added as an inducer for reporter
485 expression. 100μg/ml of ampicillin was added for plasmid selection. Plates were incubated at 25°
486 C for 24 h and scanned using EPSON image scanner.

487

488 **Acknowledgements**

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

493

494 **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

539 14. Hanks, S. K., and Hunter, T. (1995) Protein kinases 6. The eukaryotic protein kinase
540 superfamily: kinase (catalytic) domain structure and classification. FASEB J 9, 576-596
541 15. Young, T. A., Delagoutte, B., Endrizzi, J. A., Falick, A. M., and Alber, T. (2003) Structure
542 of Mycobacterium tuberculosis PknB supports a universal activation mechanism for
543 Ser/Thr protein kinases. Nat Struct Biol 10, 168-174
544 16. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.
545 T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018) SWISS-
546 MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46,
547 W296-W303
548 17. Hanks, S. K., Quinn, A. M., and Hunter, T. (1988) The protein kinase family: conserved
549 features and deduced phylogeny of the catalytic domains. Science 241, 42-52
550 18. Ruegg, U. T., and Burgess, G. M. (1989) Staurosporine, K-252 and UCN-01: potent but
551 nonspecific inhibitors of protein kinases. Trends Pharmacol Sci 10, 218-220
552 19. Meggio, F., Donella Deana, A., Ruzzene, M., Brunati, A. M., Cesaro, L., Guerra, B., Meyer,
553 T., Mett, H., Fabbro, D., Furet, P., and et al. (1995) Different susceptibility of protein
554 kinases to staurosporine inhibition. Kinetic studies and molecular bases for the resistance
555 of protein kinase CK2. Eur J Biochem 234, 317-322
556 20. Hall, C. L., Tschannen, M., Worthey, E. A., and Kristich, C. J. (2013) IreB, a Ser/Thr kinase
557 substrate, influences antimicrobial resistance in Enterococcus faecalis. Antimicrob Agents
558 Chemother 57, 6179-6186
559 21. Liu, Q., Fan, J., Niu, C., Wang, D., Wang, J., Wang, X., Villaruz, A. E., Li, M., Otto, M.,
560 and Gao, Q. (2011) The eukaryotic-type serine/threonine protein kinase Stk is required for
561 biofilm formation and virulence in Staphylococcus epidermidis. PLoS One 6, e25380
562 22. Arora, G., Sajid, A., Arulanandh, M. D., Misra, R., Singhal, A., Kumar, S., Singh, L. K.,
563 Mattoo, A. R., Raj, R., Maiti, S., Basu-Modak, S., and Singh, Y. (2013) Zinc regulates the
564 activity of kinase-phosphatase pair (BasPrkC/BasPrpC) in Bacillus anthracis. Biometals :
565 an international journal on the role of metal ions in biology, biochemistry, and medicine
566 26, 715-730
567 23. Boitel, B., Ortiz-Lombardia, M., Duran, R., Pompeo, F., Cole, S. T., Cervenansky, C., and
568 Alzari, P. M. (2003) PknB kinase activity is regulated by phosphorylation in two Thr
569 residues and dephosphorylation by PstP, the cognate phospho-Ser/Thr phosphatase, in
570 Mycobacterium tuberculosis. Molecular microbiology 49, 1493-1508
571 24. Rajagopal, L., Clancy, A., and Rubens, C. E. (2003) A eukaryotic type serine/threonine
572 kinase and phosphatase in Streptococcus agalactiae reversibly phosphorylate an
573 inorganic pyrophosphatase and affect growth, cell segregation, and virulence. The Journal
574 of biological chemistry 278, 14429-14441
575 25. Srinivasan, V. B., Vaidyanathan, V., Mondal, A., Venkataramaiah, M., and Rajamohan, G.
576 (2012) Functional characterization of a novel Mn2+ dependent protein serine/threonine
577 kinase KpnK, produced by Klebsiella pneumoniae strain MGH78578. FEBS Lett 586,
578 3778-3786
579 26. Knighton, D. R., Zheng, J. H., Ten Eyck, L. F., Ashford, V. A., Xuong, N. H., Taylor, S. S.,
580 and Sowadski, J. M. (1991) Crystal structure of the catalytic subunit of cyclic adenosine
581 monophosphate-dependent protein kinase. Science 253, 407-414
582 27. Nolen, B., Taylor, S., and Ghosh, G. (2004) Regulation of protein kinases; controlling
583 activity through activation segment conformation. Mol Cell 15, 661-675
584 28. Mieczkowski, C., Iavarone, A. T., and Alber, T. (2008) Auto-activation mechanism of the
585 Mycobacterium tuberculosis PknB receptor Ser/Thr kinase. EMBO J 27, 3186-3197
586 29. Shakir, S. M., Bryant, K. M., Larabee, J. L., Hamm, E. E., Lovchik, J., Lyons, C. R., and
587 Ballard, J. D. (2010) Regulatory interactions of a virulence-associated serine/threonine
588 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

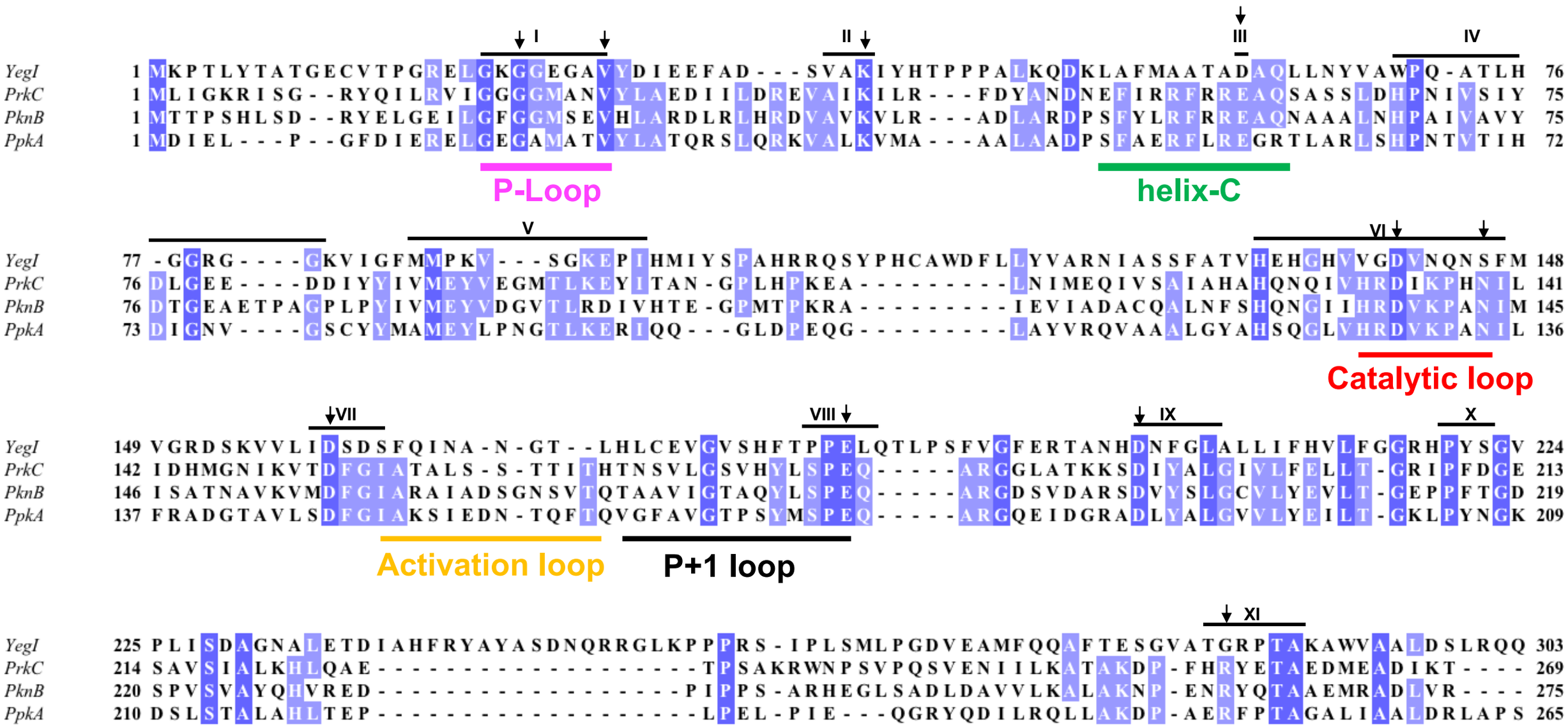
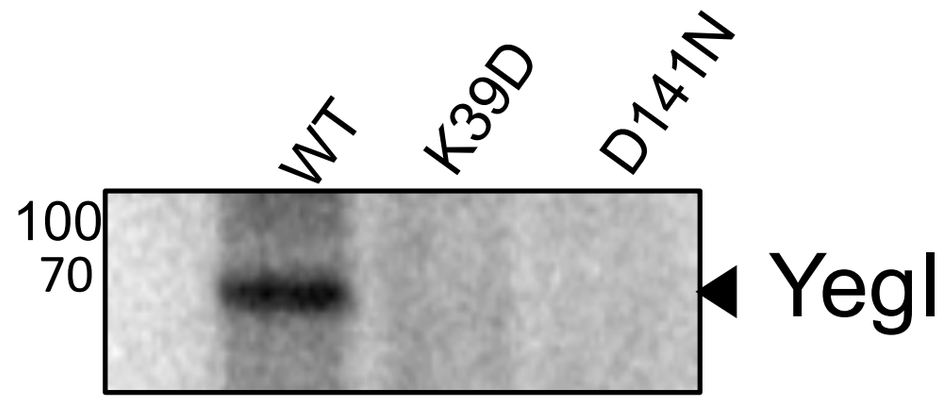

I
II
III
IV
YegI 1 MKPTLYTATGECVTPGRELGKGGEGAVYDIEEFAD---SVAKIYHTPPPALKQDKLAFMAATADAQLLNYVAWPQ-ATLH 76
PrkC 1 MLIGKRISG--RYQILRVIGGGGMANVYLAEDIILDREVAIKILR---FDYANDNEFIRRFRREAQSASSLDHPNIVSIY 75
PknB 1 MTTPSHLSD--RYELGEILGFGGMSEVHLARDLRLHRDVAVKVLR---ADLARDPSFYLRFRREAQNAAALNHPAIVAVY 75
PppA 1 MDIEL---P--GFDIERELGEGAMATVYLATQRSLQRKVALKVMA---AALAADPSFAERFLREGRTLARLSHPNTVTIH 72
P-Loop
helix-C
V
VI
YegI 77 -GGRG----GKVIGFMMPKV---SGKEPIHMIYSPAHRRQSYPHCAWDFLLYVARNIASSFATVHEHGHVVGDVNQNSFM 148
PrkC 76 DLGEE----DDIYYIVMEYVEGMTLKEYITAN-GPLHPKEA---------LNIMEQIVSAIAHAHQNQIVHRDIKPHNIL 141
PknB 76 DTGEAETPAGPLPYIVMEYVDGVTLRDIVHTE-GPMTPKRA---------IEVIADACQALNFSHQNGIIHRDVKPANIM 145
PppA 73 DIGNV----GSCYYMAMEYLPNGTLKERIQQ---GLDPEQG---------LAYVRQVAAALGYAHSQGLVHRDVKPANIL 136
Catalytic loop
VII
VIII
IX
X
YegI 149 VGRDSKVVLIDSDSFQINA-N-GT--LHLCEVGVSHFTPPELQTLPSFVGFERTANHDNFGLALLIFHVLFGGRHPYSGV 224
PrkC 142 IDHMGNIKVTDFGIATALS-S-TTITHTNSVLGSVHYLSPEQ-----ARGGLATKKSDIYALGIVLFELLT-GRIPFDGE 213
PknB 146 ISATNAVKVMDFGIARAIADSGNSVTQTAAVIGTAQYLSPEQ-----ARGDSVDARSDVYSLGCVLYEVLT-GEPPFTGD 219
PppA 137 FRADGTAVLSDFGIAKSIEDN-TQFTQVGFAVGTPSYMSPEQ-----ARGQEIDGRADLYALGVVLYEILT-GKLPYNGK 209
Activation loop
P+1 loop
XI
YegI 225 PLISDAGNALETDIAHFRYAYASDNQRRGLKPPPRS-IPLSMLPGDVEAMFQQAFTESGVATGRPTAKAWVAALDSLRQQ 303
PrkC 214 SAVSIALKHLQAE----------TPSAKRWNPSVPQSVENIILKATAKDP-FHRYETAEDMEADIKT---- 269
PknB 220 SPVSVAYQHVRED----------PIPPS-ARHEGLSADLDAVVLKALAKNP-ENRYQTAAEMRADLVR---- 275
PppA 210 DSLSTALAHLTEP----------LPEL-PIE---QGRYQDILRQLLAKDP-AERFPTAGALIAALDRLAPS 265

Figure 2

A


WT
K39D
D141N
100
70
YegI

B

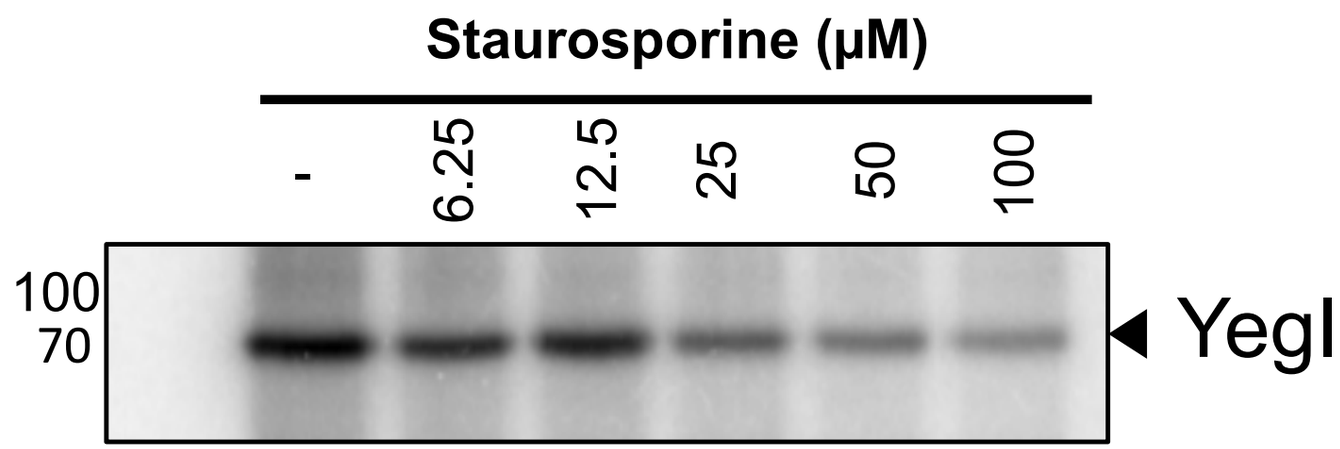
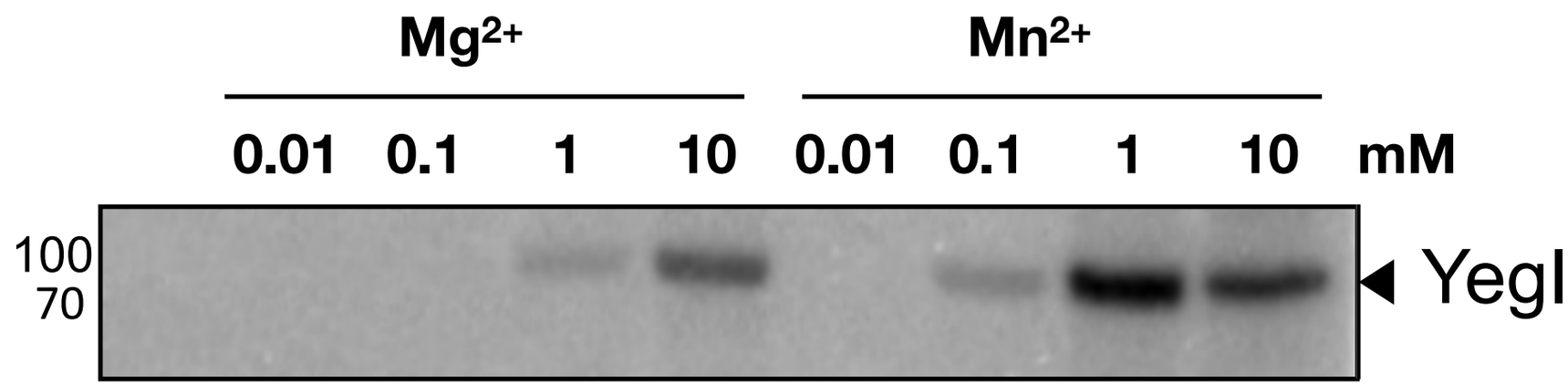
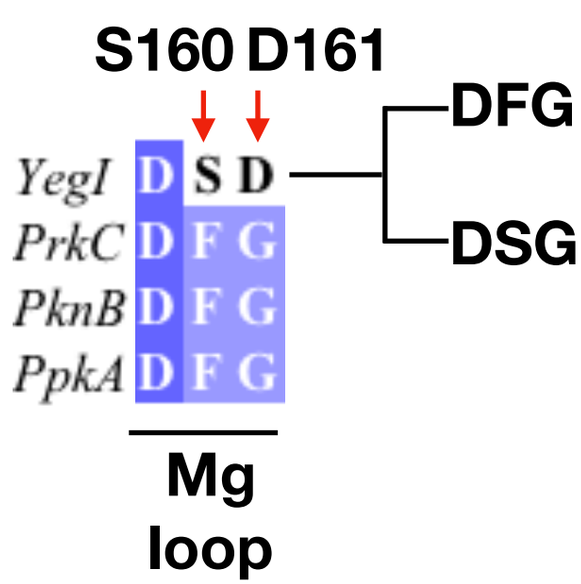

Staurosporine (μM)
-
6.25
12.5
25
50
100
100
70
YegI

Figure 3

A


Mg2+
Mn2+
0.01 0.1 1 10 0.01 0.1 1 10 mM
100
70
YegI

B


S160 D161
YegI D S D
PrkC D F G
PknB D F G
PpkA D F G
DFG
DSG
Mg loop

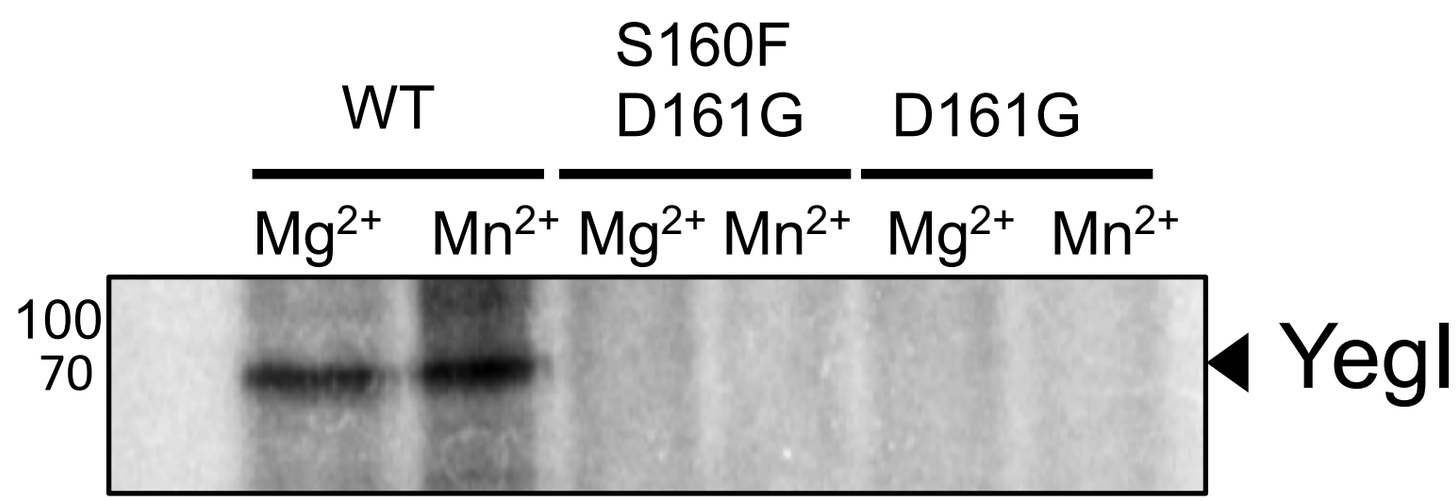

WT
S160F D161G
D161G
Mg2+ Mn2+ Mg2+ Mn2+ Mg2+ Mn2+
100
70
YegI

Figure 4

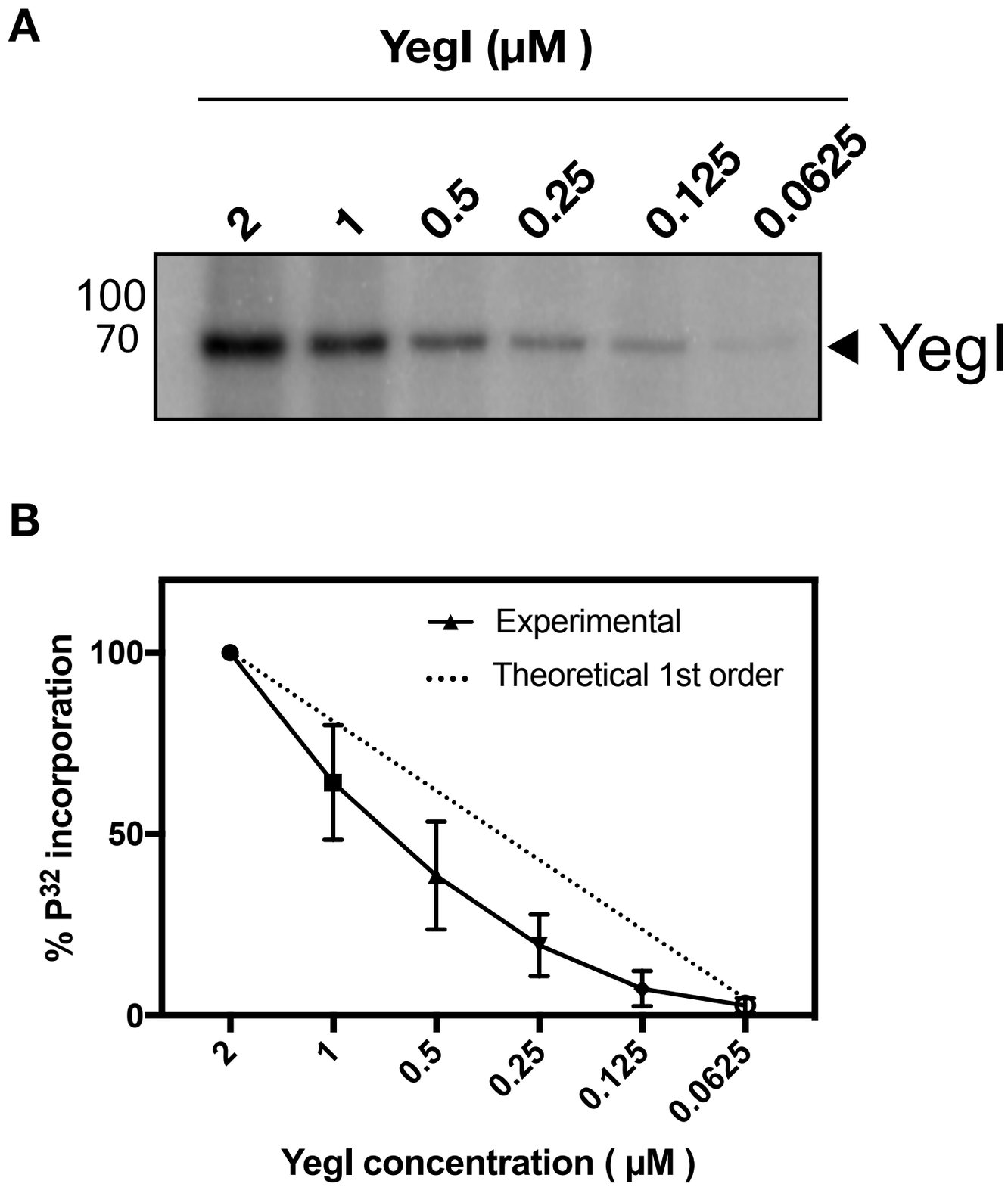

A
YegI (μM)
2
1
0.5
0.25
0.125
0.0625
100
70
YegI
B
Experimental
Theoretical 1st order
% P32 incorporation
100
50
0
2
1
0.5
0.25
0.125
0.0625
YegI concentration (μM)

Figure 5

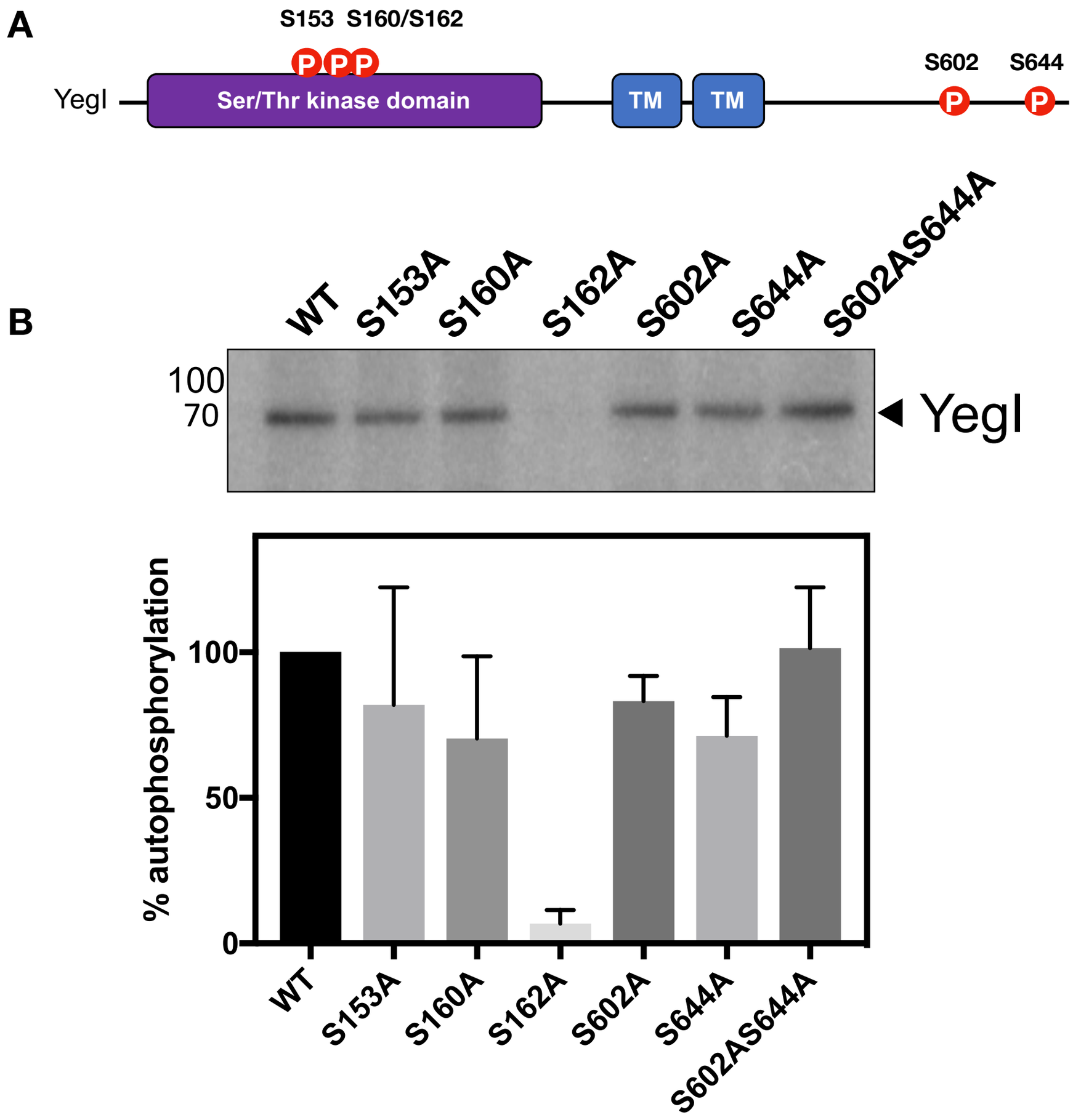
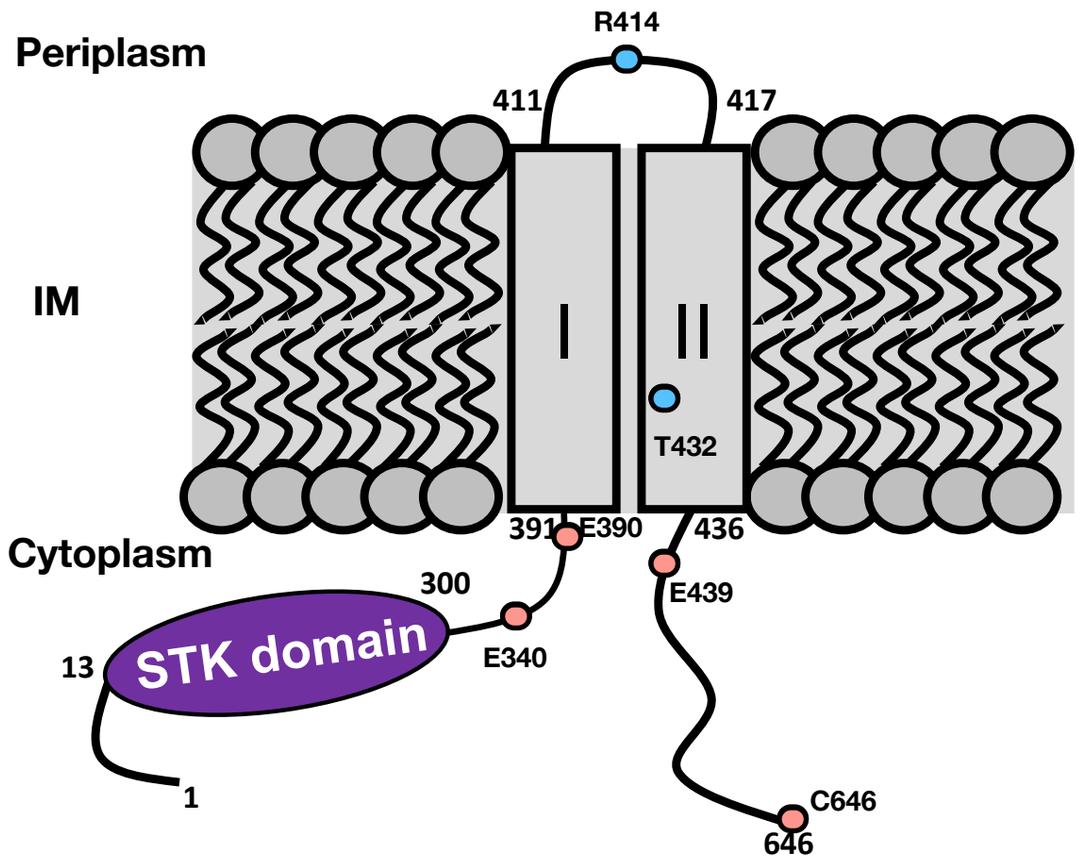

A
S153
S160/S162
P
P
P
YegI
Ser/Thr kinase domain
TM
TM
S602
S644
P
P
B
WT
S153A
S160A
S162A
S602A
S644A
S602AS644A
100
70
YegI
% autophosphorylation
100
50
0
WT
S153A
S160A
S162A
S602A
S644A
S602AS644A

Figure 6

A



B

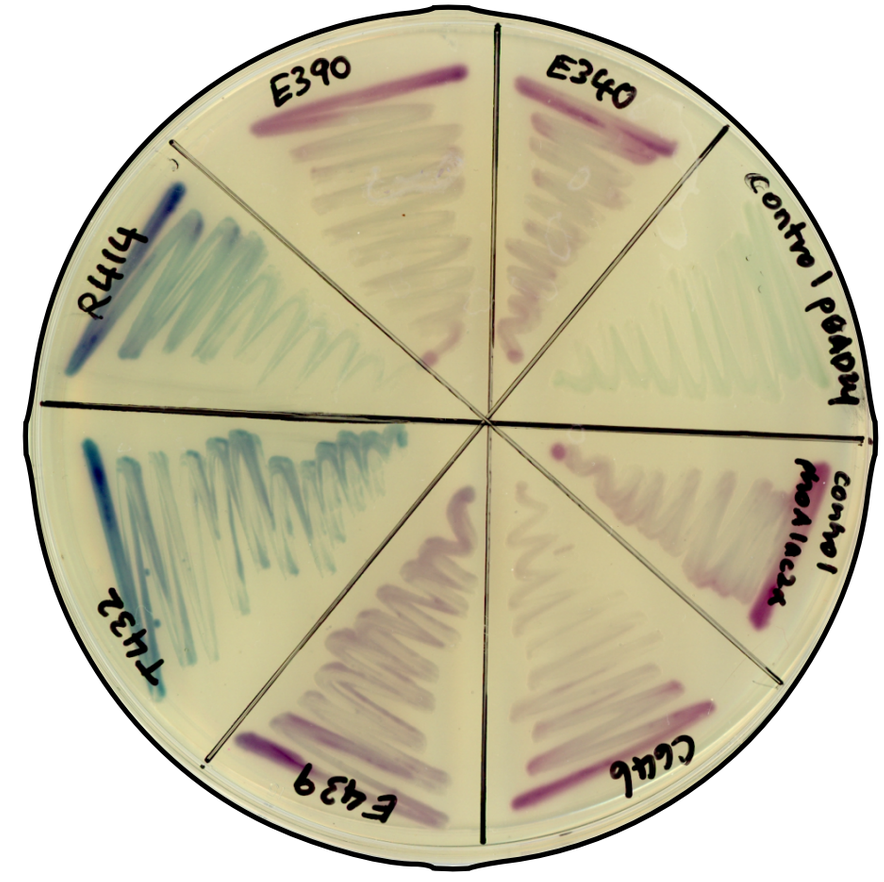
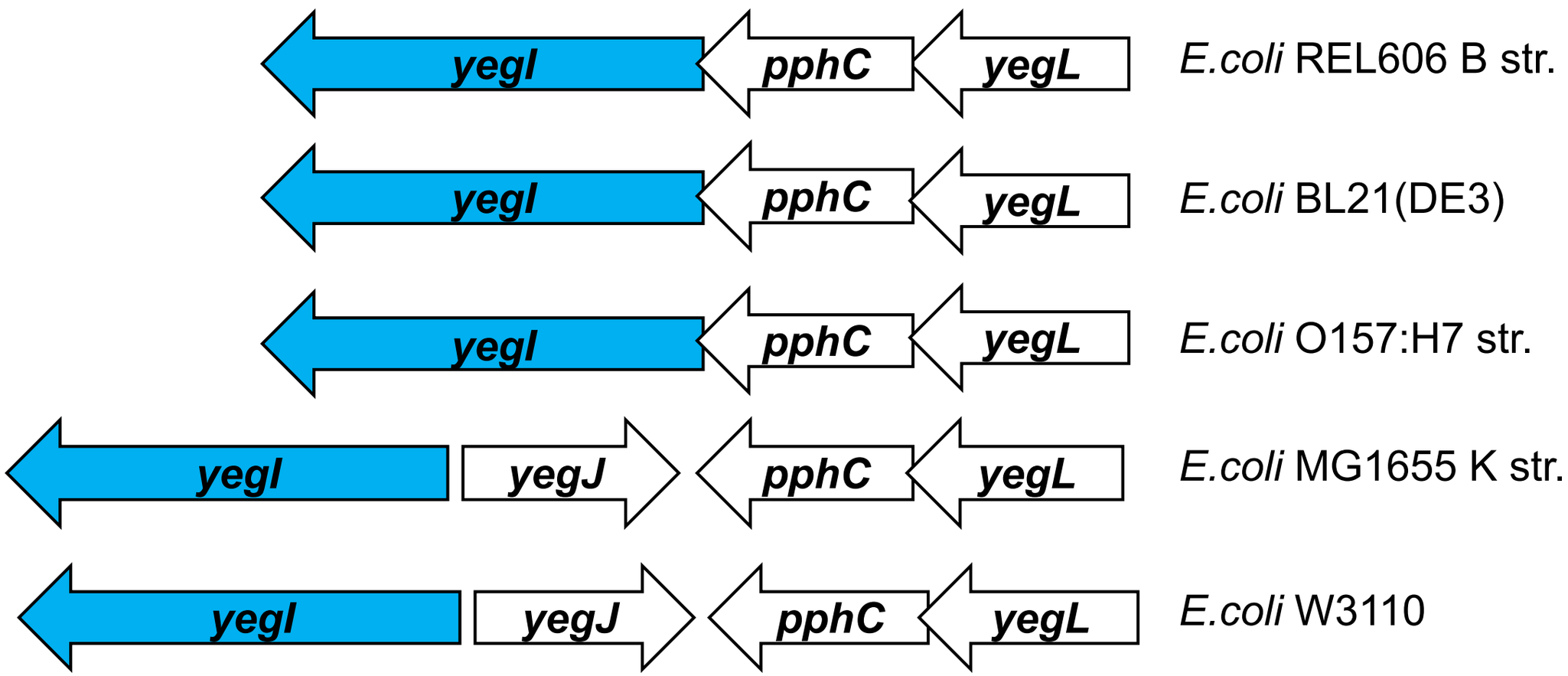


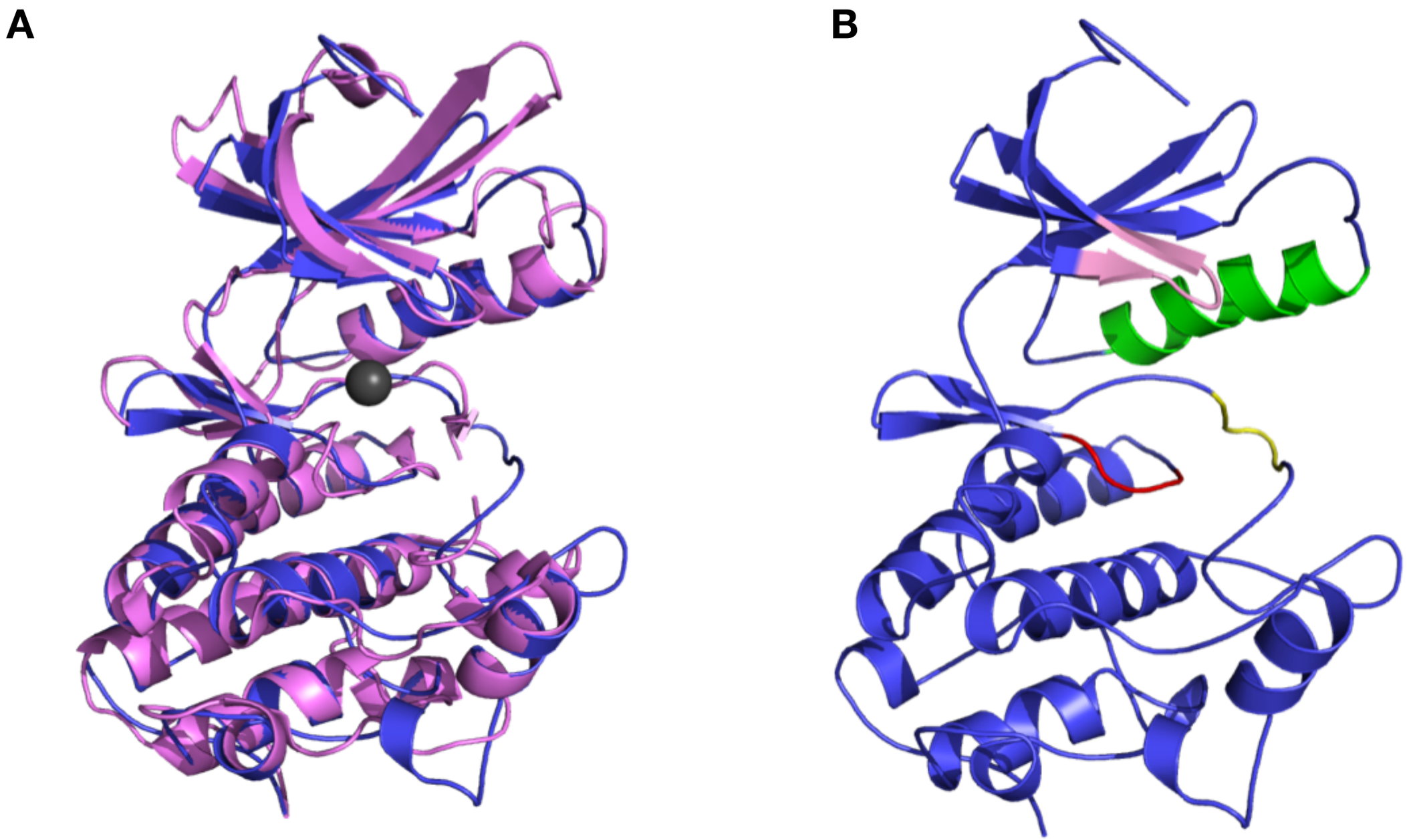
Table 1

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

Supp fig 1: Genome architecture of *yeg* operon



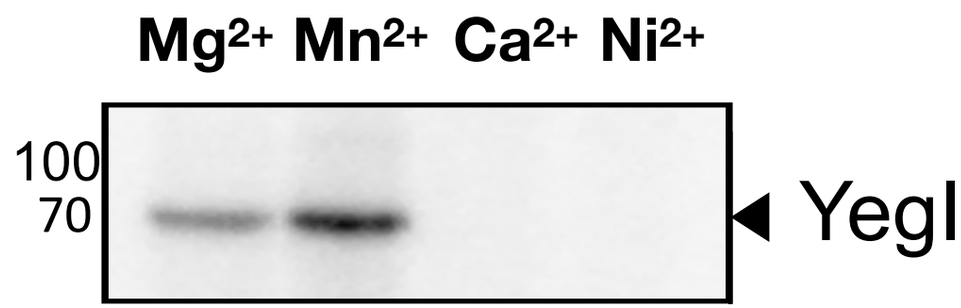
Supp fig 2: Predicted structure of YegI kinase domain resembles a bacterial eSTK



***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 YegI kinase domain (aa 1-302)**.

(A) Structural overlay of predicted structure of YegI 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 YegI 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 ºC with 1 µM of YegI (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 *E.coli* REL606	This study
JDE3050	pKR81 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged K39D YegI from *E.coli* REL606	This study
JDE3051	pKR83 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged D141N YegI from *E.coli* REL606	This study
JDE3052	pKR97 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S160FD161G YegI from *E.coli* REL606	This study
JDE3053	pKR98 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged D161G YegI from *E.coli* REL606	This study
JDE3054	pKR110 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S153A YegI from *E.coli* REL606	This study
JDE3055	pKR111 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S160A YegI from *E.coli* REL606	This study
JDE3056	pKR112 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S162A YegI from *E.coli* REL606	This study
JDE3057	pKR62 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S602A YegI from *E.coli* REL606	This study
JDE3058	pKR63 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S644A YegI from *E.coli* REL606	This study
JDE3059	pKR64 amp	LOBSTR-BL21 (DE3) expressing N terminal 6X-His tagged S602AS644A YegI from *E.coli* REL606	This study
JDE3042	pKR73 amp	DH5a expressing N terminal 6X-His tagged YegI (aa 1-439) fused with PhoA lacZα from E.coli REL606	This study
JDE3043	pKR74 amp	DH5a expressing N terminal 6X-His tagged YegI (aa 1-432) fused with	This study

		PhoA lacZα from E.coli REL606	
JDE3044	pKR75 amp	DH5a expressing N terminal 6X-His tagged YegI (aa 1-414) fused with PhoA lacZα from E.coli REL606	This study
JDE3045	pKR77 amp	DH5a expressing N terminal 6X-His tagged YegI (aa 1-340) fused with PhoA lacZα from E.coli REL606	This study
JDE3046	pKR76 amp	DH5a expressing N terminal 6X-His tagged YegI (aa 1-390) fused with PhoA lacZα from E.coli REL606	This study
JDE3047	pKR77 amp	DH5a expressing N terminal 6X-His tagged YegI (aa 1-646) fused with PhoA lacZα from E.coli REL606	This study
JDE3048	pKR78 amp	DH5a expressing N terminal 6X-His tagged PhoA (aa 21-471) in frame with lacZα (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 yegI 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 *yegI-phoA-lacZ* 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 YegI E439-phoA lacZ was generated by Q5 site directed mutagenesis using KR230/KR231 and pKR72 as a template.	This study
pKR74	REL606 YegI T432-phoA lacZ was generated by Q5 site directed mutagenesis using KR230/KR232 and pKR72 as a template.	This study
pKR75	REL606 YegI R414-phoA lacZ was generated by Q5 site directed mutagenesis using KR230/KR233 and pKR72 as a template.	This study
pKR76	REL606 YegI E390-phoA lacZ was generated by Q5 site directed mutagenesis using KR230/KR234 and pKR72 as a template.	This study
pKR77	REL606 YegI E340-phoA lacZ was generated by Q5 site directed mutagenesis using KR230/KR235 and pKR72 as a template.	This study
pKR78	Control phoA lacZ 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	ggaaaaccatggca*catcatcatcatcatcat*AAACCCACTTTATATACTGCTAC	NcoI; *6X-His tag*	This study
KR59	cgtccagcatgcTCAGCAGCGGCTTAAATCTG	SphI; STOP	This study
KR152	GGCTGGAA**GCG**GCGTTGAC	S602A	This study
KR153	GTCAACGC**CGC**TTCCAGCC	S602A	This study
KR154	GCAGATTTA**GCC**CGCTGCTGA	S644A	This study
KR155	TCAGCAGCG**GGC**TAAATCTGC	S644A	This study
KR157	GTCGCGAC**GCC**AAAGTGGT	S153A	This study
KR158	ACCACTTT**GGC**GTCGCGAC	S153A	This study
KR223	ccatggca*catcatcatcatcatcat*AAACCCACTTT ATATACTGCTACT	NcoI; *6X-His tag*	This study
KR224	catttctggtgtccgggcgctaccgctgccgctaccGCAG CGGCTTAAATCTGCCTG	PhoA; Linker	This study
KR225	GCCCGGACACCAGAAATG	PhoA	This study
KR226	gacggccagtgaatccgtTTTCAGCCCCAGAGC GGC	lacZ;PhoA	This study
KR227	ACGGATTCACTGGCCGTCGTT	lacZ	This study
KR228	cgtccaggtaccttaGCGCCATTCGCCATTCAG GC	Kpn1; STOP	This study
KR229	ggaaaagctagcaggaggaattCCATGGCACATC ATCATCATCAT	Nhe1; 6X-His-tag	This study
KR230	GGTAGCGGCAGCGGTAGC	Q5 SDM_Linker	This study
KR231	TTCTGCTTTATAAGCTTTGCTTGTCAGA CTG	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	TTCGCCGAGATCAATAAAATAGATAAC GCC	Q5_SDM_E340	This study
KR240	ATGATGATGATGATGATGTGCCATGGT G	Q5_SDM_phoAlacZ	This study
KR245	AGCGTCGCC**GAC**ATTTATCAC	K39D	This study

KR246	GTGATAAAT**GTC**GGCGACGCT	K39D	This study
KR247	GTCGTGGGG**AAC**GTAAACCA	D141N	This study
KR248	TGGTTTAC**GTT**CCCCACGAC	D141N	This study
KR285	GTTGATCGAT**TTTGGC**TCCTTTCAGAT TAACGCCAATG	S160FD161G	This study
KR286	TGGTTTAC**GTT**CCCCACGAC	S160FD161G	This study
KR305	GTTGATCGAT**GCT**GACTCCTTT	S160A	This study
KR306	AAAGGAGTC**AGC**ATCGATCAAC	S160A	This study
KR307	CGATAGTGAC**GCC**TTTCAGAT	S162A	This study
KR308	ATCTGAAA**GGC**GTCACTATCG	S162A	This study

Homology sequences to gene are shown in uppercase. Non homology sequences are indicated in lower case including His_6-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