2014-12-01

Studies of PhoU in Escherichia coli: Metal Binding, Dimerization,Protein/Protein Interactions, and a Signaling Complex Model

Stewart G. Gardner
Brigham Young University - Provo

Studies of PhoU in *Escherichia coli*: Metal Binding, Dimerization, Protein/Protein Interactions, and a Signaling Complex Model

Stewart G. Gardner

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

William R. McCleary, Chair
Joel Shelby Griffitts
Brent Lynn Nielsen
Julianne House Grose
Zachary Thomas Aanderud

Department of Microbiology and Molecular Biology

Brigham Young University

December 2014

ABSTRACT

Studies of PhoU in *Escherichia coli*: Metal Binding, Dimerization, Protein/Protein Interactions and a Signaling Complex Model

Stewart G. Gardner
Department of Microbiology and Molecular Biology, BYU
Doctor of Philosophy

Phosphate is an essential nutrient for all forms of life. *Escherichia coli* has a PhoR/PhoB two component regulatory system that controls the expression of various genes whose products allow the cell to thrive in low phosphate environments. The signaling mechanism of the PhoR/PhoB system has been studied and the phosphorylation cascade that controls gene expression is well understood. What is still unknown is how PhoR senses the phosphate level of the environment. The PstS, PstC, PstA, PstB, and PhoU proteins play a role in this signal sensing. This work confirms the hypothesis that the PstSCAB complex senses the environmental phosphate and that phosphate signal is passed through PhoU to PhoR. Further, this work characterizes residues important for interaction on PhoU and PhoR and identifies a structural model for interaction. This model points to a potential mechanism for PhoU mediated signaling to PhoR. We tested this model with direct coupling analysis and obtained further confirmation. Further use of these techniques may elucidate more of the interactions necessary for proper phosphate signaling.

Keywords: PhoR, PhoU, PstSCAB, Pho regulon, two component signal transduction, PAS domain, direct coupling analysis, protein structure modeling

ACKNOWLEDGEMENTS

I would like to thank Dr. William R. McCleary for his guidance, encouragement, and mentoring throughout my academic career. I would also like to thank my advisory committee: Drs. Joel Griffitts, Brent Nielsen, Julianne Grose, and Zachary Aanderud. Thanks in addition to Dr. Allen R. Buskirk for serving on my committee. I would also like to thank Dr. Perry Ridge and Dr. Byron Adams for help with bioinformatic and phylogenetic analyses. Thanks to Casey Callison, Kristi Johns, and all the other students who have worked with me in Dr. McCleary's laboratory. Most importantly, thanks to my wife Jamie Gardner, for her support, love, and confidence.

For the research presented in Chapter 2 "The PhoU protein from *Escherichia coli* interacts with PhoR, PstB and metals to form a phosphate-signaling complex at the membrane" (1). I thank my co-authors: Kristine D. Johns, Rebecca Tanner, and William R. McCleary Additionally, I thank Dr. John Bell, Jen Nelson, and Liz Gibbons for their assistance with the fluorescence assays and Lance Moses for help with the ICP-MS analysis. I thank several undergraduate students; Geoffrey Johnston, Rachel Winn, Austin Callison, and Annaliese Gabrielle for help in the initial characterization of the mutant strains and Casey Callison, Kirk Richardson, Michael Barrus, and Gregory Bowden for excellent technical assistance. That work was supported by Public Health Service grant R15GM96222 from the National Institute of General Medical Sciences. R. T. was partially supported by an Undergraduate Mentoring Environment Grant from Brigham Young University.

For the research in Chapter 3 "Genetic analysis, structural modeling, and direct coupling analysis suggest a mechanism for phosphate signaling in *Escherichia coli"*, I thank my co-authors: Justin B. Miller, Tanner Dean, Tanner Robinson , McCall Erickson, Perry Ridge, and

William R. McCleary. Also, I thank Kathryn Hanks, Alex Cummock, Evan Christensen, Bethany Evans, Gregory Bowden, and Michael Barrus for help with sequence collection and mutant characterization. This work was supported by Public Health Service grant R15GM96222 from the National Institute of General Medical Sciences.

Table of Contents

List of Tables

List of Figures

Chapter 1 The Pho Regulon in Escherichia coli

1.1 Introduction

Phosphate is an essential nutrient for all forms of life and is found in many of the basic building blocks of life, such as DNA, RNA, protein modifications, lipid membranes, and more. To survive, organisms need to find and utilize phosphate from their environment. Through the use of a two-component signal transduction system, *Escherichia coli* and other bacteria adjust to the level of phosphate in their environment and adapt to uptake and utilize phosphate more efficiently when it is limiting. This allows for a more flexible life cycle and ability to grow in many different conditions and environments. The mechanisms for regulating this system have been well studied and the deleterious effects of disruptions in proper regulation show how important this system is to the bacterial cell (2-4). Further understanding of how this system functions may shed light on similar two component systems essential for many different organisms, as well as pathogenesis and disease mechanisms of *E. coli* and other related bacteria (5-9).

The coordinated expression control of a collection of genes called the Pho regulon is by a regulatory mechanism in which PhoB, a response regulator, acts as a transcriptional activator. Control of the Pho regulon is essential for proper growth of *E. coli*. Inorganic phosphate (P(i)) is transported into the cell by a transporter made up of PstS, PstC, PstA, and PstB proteins (PstSCAB). These proteins along with PhoR, PhoB, and PhoU are required for proper signaling of the Pho regulon, where gene expression is induced in low phosphate environments and repressed when environmental phosphate is plentiful (3). Studies have shown that in phosphate starvation of *E. coli*, up to 208 proteins show increase in abundance and 205 proteins show a

decrease (10). This study found at least 31 genes whose expression is directly controlled by the PhoB/PhoR two-component signal transduction system (10). In P(i) limiting conditions, PhoR, a histidine kinase, is phosphorylated and phospho-PhoR activates PhoB by phosphorylating it. Phospho-PhoB then acts as a transcriptional activator by interacting with conserved sequences in Pho regulon promoters, called Pho boxes, to increase transcription from those promoters (11)(Fig. 1.1). The exact mechanism by which PhoR receives a signal of the environmental phosphate level has been unknown. This work will discuss the major components of this system, their functions, describe this system's role in pathogenesis, and discuss recent findings that help to identify a potential mechanism for signaling in this system.

1.1.1 Histidine Kinases

Several studies have attempted to understand the mechanism by which histidine kinases receive and pass on environmental signals. Many of these studies base their understanding of signaling mechanisms on structures of various signaling proteins (12-15). In general, histidine kinases form homodimers and are composed of an N-terminal membrane bound periplasmic sensing domain coupled to a C-terminal cytoplasmic kinase domain. The kinase domains of histidine kinases are highly conserved and are responsible for autophosphorylation where a specific histidine residue of the kinase itself is phosphorylated, priming the protein to donate the phosphoryl group to a conserved aspartate residue on the cognate response regulator protein. Additionally, some histidine kinases also display phosphatase activity enabling them to dephosphorylate the phospho-response regulator and change the activation state of the signaling system. This allows for fine tuning of the signal and quick switching from one state to the other (14).

In the process of autophosphorylation, the histidine kinase's formation of the unique high-energy phosphoramidate N-P bond (a higher energy bond than the phosphoester bond formed by serine, threonine, and tyrosine kinases) allows for rapid phosphotransfer to the response regulator in these signaling pathways (12).

1.1.1.1 Conserved domains

Histidine kinases are often composed of several conserved domains. Each domain has a specific activity that when in combination with the other domains allows for proper activation/deactivation in response to various environmental signals. Different combinations and arrangements of the modular domains allows for multiple different activities specific for each of the signals sensed by a wide variety of histidine kinases.

Often, a signal sensing domain is found at the N-terminal end of the histidine kinase that is located outside the cell (in the periplasm of gram negative bacteria). These domains can detect the signal specific to the kinase by binding directly to signal molecules or through protein/protein interaction with other signal binding proteins. Alternatively, some histidine kinases have minimal periplasmic portions and receive signals through other means (16).

Many histidine kinases contain conserved HAMP domains (named for being associated with histidine kinases, adenylyl cylases, methyl-accepting proteins, and other prokaryotic signaling proteins) that are found directly C-terminal to the transmembrane portion of the protein that are implicated in transferring signals from the periplasm to the cytoplasmic domains of the histidine kinase. HAMP domains contain a four helix coiled-coil structure that shows different packing arrangements based on the rotation of the individual helices. This rotation has been linked to activation states of the kinase domains (13, 17). The proposed mechanism is that

binding of a signal molecule to the C-terminal signal sensing domain of the protein causes a conformational shift that is transferred through the membrane by the HAMP domain helix rotation, passing the signal to the cytoplasmic kinase domains.

1.1.1.1.1 PAS domains

Other signaling domains often found in histidine kinases include c-GMP-specific and c-GMP stimulated phosphodiesterases (GAF) domains as well as Per-ARNT-Sim (PAS) domains. These domains are often found in the cytoplasmic region of histidine kinases and have been linked to signal sensing activity. They have a central five (PAS) or six (GAF)-stranded anti-parallel β-sheet cores surrounded by multiple α-helices. PAS domains are found in >33% of all histidine kinases and GAF domains in about 10% of histidine kinases (13).

PAS domains have been shown to play a role in protein/protein interaction (13). Also, PAS domains can bind small ligands and detect changes in light, oxygen, redox potential, and cellular energy level and these signals lead to a conformational change of the PAS domain that causes perturbation of the structural arrangement of the kinase domains leading to altered activities of the histidine kinase (12). Several studies have identified the structures of PAS domains and their mechanism for transferring signals. The core structure of the PAS domain is highly conserved. When signals are sensed, they generate structural changes within the β sheet that are passed onto the adjacent effector domains. The signals often involve changes to the quaternary structure of the protein complex (18).

In one instance the PAS domain of a histidine kinase, WalK from *Streptococcus pneumoniae*, was found to be important for phosphatase activity and virulence in murine pneumonia (19). Another new study found that a PAS domain binds to a cell wall intermediate to

limit the histidine kinase activation (20). These results illustrate specific examples of the PAS domain being essential for proper kinase/phosphatase switch of a histidine kinase.

1.1.1.1.2 Cytoplasmic kinase domains

The actual switch from kinase to phosphatase activity is mediated by the three-dimensional arrangement of the Catalytic and ATP binding (CA, also known as the HATPase_c in the Pfam database) and Dimerization and Histidine phosphotransfer (DHp, also known as the His kinase A domain clan in the Pfam database) domains with the phosphatase activity being found in the DHp domain (12, 17, 21). The CA domain binds to ATP and when in the proper conformation, catalyzes the phosphorylation of the conserved histidine residue of the DHp domain. The DHp domain not only contains the site for phosphorylation, but also is responsible for the dimer formation of the histidine kinase by interacting with the DHp domain of the other histidine kinase of the dimer.

1.1.1.2 Cis or Trans Phosphorylation

Given that histidine kinases function as dimers, one question concerning the mechanism of autophosphorylation has been whether the posphorylation occurs in cis or trans (meaning does the CA domain phosphorylate the conserved His residue of the DHp domain from the same protein or the DHp domain of the other protein in the dimer?). One review addresses this question well (22). Specifically, they mention several histidine kinases that have been shown to catalyze the cis reaction and several others that catalyze the trans. Some evidence points to the loop between the two helices of the DHp domain as determining the type of reaction that a histidine kinase performs. However, they cite an example where two kinases with the same loop

structure undergo different reaction types. The explanation for these results is that there are multiple factors that influence whether a cis or trans reaction occurs for each histidine kinase, the handedness of the loop between the DHp helices and the length of the hinge between the DHp and CA domains both potentially playing roles in determining whether the cis or trans reaction occurs.

1.1.1.3 Phosphatase activity

Many studies have examined the phosphatase activity of histidine kinases (23-25). A review of some of these studies restates that the phosphatase activity is dependent on the DHp domain and that proper three-dimensional conformational arrangement of the CA to DHp domains is essential for phosphatase activity. They also point out that the conserved His residue of the DHp domain does not appear to be essential for phosphatase activity, but may enhance it (25). That finding points to a specific phosphatase reaction for these proteins, not just the reversal of the kinase reaction. Another review confirms the physiological importance of phosphatase activity to prevent cross-talk between different two-component signal transduction systems. They also review a potential mechanism for phosphatase activity similar to known mechanisms of other phosphatase enzymes (24). This same group went on to show that many mutations that affect phosphatase function are likely due to inhibiting proper signal transduction and not necessarily abolishing the activity (23). The molecular mechanism for kinase and phosphatase activity is reviewed in another paper which shows that the switch from kinase to phosphatase activity requires an extra water molecule to be incorporated into the active site. This addition of another molecule to the active site means that the two different activities require unique active site conformations. This prevents the spontaneous dephosphorylation of sites that

were just phosphorylated. The switch between activities of histidine kinases often depends on other domains, like PAS domain (22).

1.1.1.4 Specificity of histidine kinase/response regulator interaction

One potential difficulty of having many two component signal transduction systems (that sense different signals but utilize similar conserved domains for activation) is how to ensure that a specific signal leads to activation of the proper response regulator and not any other response regulators. To prevent phosphorylation/dephosphorylation of response regulators that are not the target for a specific histidine kinase, there is a known interaction surface between the two proteins that allows for specific interaction between histidine kinase and response regulator pairs. This specific protein/protein interface has been identified and shown that if you mutate the residues at the interface you can alter the specificity of which response regulators are activated/deactivated by a histidine kinase (26).

1.1.2 Response Regulators

Histidine kinases act to change the response of the cell to environmental changes by modifying the activity of response regulators via selective phosphorylation/dephosphorylation. There is a modular architecture to the domains of response regulators. The domains consist of a receiver (REC, known as the Response_reg domain in the Pfam database) domain linked to an effector domain (12). Occasionally, proteins contain a REC domain only without an effector domain. These REC domain only proteins may act as phosphorylated intermediates in phospho relays or directly interact with other proteins.

The REC domain of response regulators act to control their cognate effector domain activity through conformational changes based on the REC domain phosphorylation state of a conserved aspartate residue. The phosphotransfer from the histidine kinase depends on catalytic contributions from both proteins, thus the REC acts as an enzyme in addition to its activity on the effector domain. The REC domain also plays a role in catalyzing autodephosphorylation of the response regulator (12).

Effector domains are varied and signal in many different ways. The majority of response regulators identified (63%) contain a DNA binding effector domain, of which the two most common are the OmpR/PhoB winged-helix domain and the NarL/FixJ four-helix helix-turn-helix domain. Additionally, some response regulators have RNA binding domains. Other effector domains include enzymatic domains and protein or ligand interaction domains. With the DNA binding domains, it is difficult to identify all of the DNA targets through bioinformatic methods due to the weak conservation of target sequences. So, experimental methods are more reliable in identifying target genes regulated by a specific response regulator (12).

The mechanism of activating some response regulators is well understood. For OmpR/PhoB response regulators, phosphorylation of the REC domain leads to dimerization coordinated by several salt bridges between highly conserved residues in both proteins. This dimerization brings the DNA binding domains together and allows them to bind to their recognition sites (12).

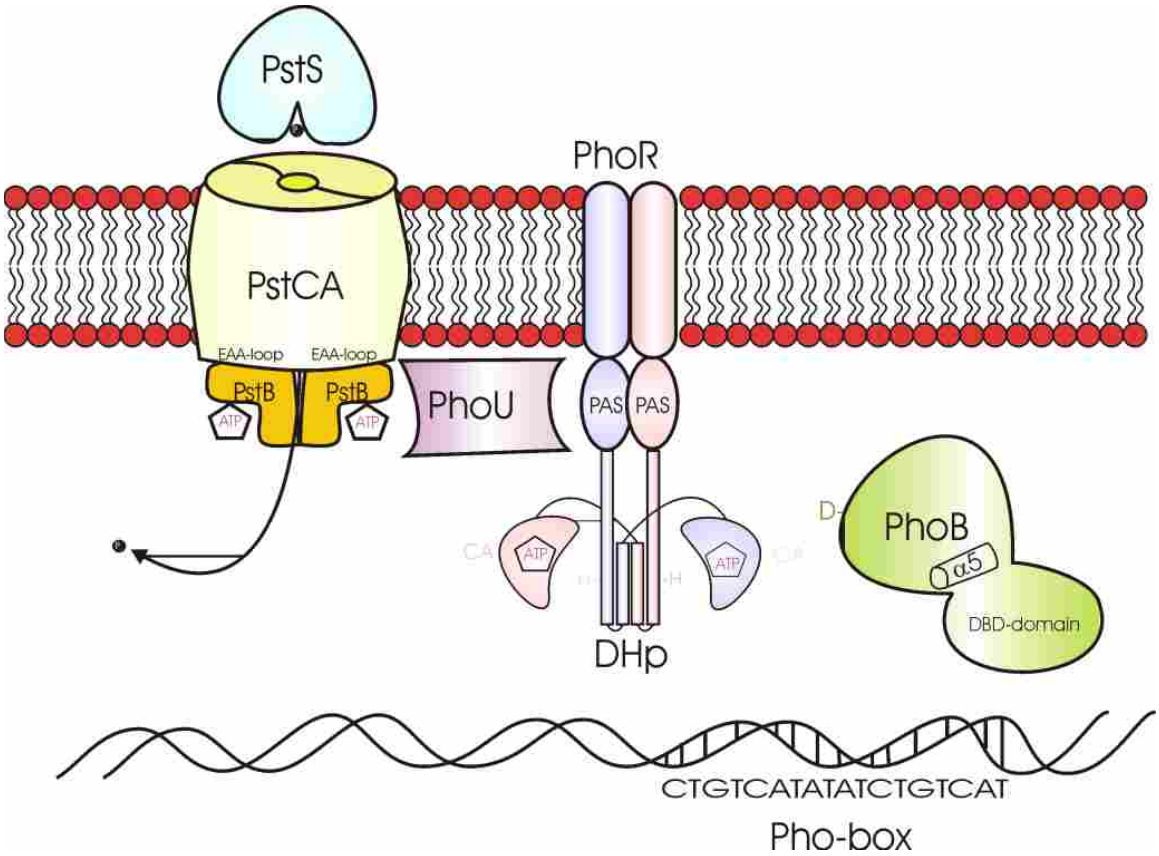


Figure 1.1 Overview of the major regulatory components of the Pho regulon
PhoR is a histidine kinase that receives a signal and catalyzes the autophosphorylation of a conserved histidine residue in an ATP dependent manner. It has a Per-ARNT-Sim (PAS) domain, a dimerization and histidine phosphotransfer (DHp) domain, and a catalytic and ATP binding (CA) domain. PhoB is a response regulator that is activated by phosphorylation on a conserved aspartate (D) residue when it interacts with phospho-PhoR. Once active, PhoB DNA binding domain (DBD) binds DNA that contains the conserved Pho-box sequence to increase transcription from the promoters it binds. The Phosphate specific transporter is made up of PstSCAB proteins and is important in sensing phosphate levels. PhoU is a peripheral membrane protein that acts as negative regulator of the PhoR/PhoB system.

1.1.3 PhoB and PhoR

The control of expression of Pho regulon genes is mediated by a classic two-component signal transduction pair, PhoR and PhoB. There are many examples of two-component signal transduction systems found throughout nature that receive signals and adjust expression of various proteins to respond to changes in the environment. In *E. coli* alone there are at least 62 two component proteins encoded for on the genome (27).

PhoB and PhoR help bacterial cells respond to changes in the environmental phosphate level by regulating the expression of the Pho regulon genes. In low phosphate environments, a signal is sensed and passed through a few proteins, via a phosphorylation cascade that leads to increased gene transcription. PhoR, a histidine kinase, is autophosphorylated in an ATP dependent manner on a conserved histidine residue. In this active state, PhoR phosphorylates the PhoB response regulator on a conserved aspartate. The phosphorylated PhoB then interacts with DNA promoter regions (as a dimer) containing pho boxes and induces transcription. In high phosphate environments PhoR acts as a phosphatase and removes phosphate from PhoB. In this way, the cell can both activate and deactivate expression of this system. Importantly, *phoR* and *phoB* genes are under control of this expression control system. The *phoB* and *phoR* genes form an operon and the *phoB* promoter itself is activated by phospho-PhoB (28). This auto-catalytic loop leads to a control system that has very high levels of expression in low phosphate environments when this system is induced and extremely low expression levels when the system is turned off in high phosphate environments (11).

1.1.3.1 PhoR

PhoR is responsible for the major activation/deactivation of PhoB. The phosphorylation site of PhoR is adjacent to the catalytic region, thus making it a member of the class I family of histidine kinases (29).

It is unclear how PhoR receives a signal that the cell is in a high or low phosphate environment. Many histidine kinases have a periplasmic sensing domain that directly binds signaling molecules to receive an environmental signal. When the signal is sensed, the kinase core is activated to phosphorylate the conserved histidine residue, often as a homodimer(4).

PhoR lacks any significant periplasmic sensing domain and must receive a signal in another manner. PstSCAB and PhoU are important in this signaling but the mechanism is still unknown(3).

Several groups have studied truncations and internal mutations of PhoR to identify what portions of the protein are important for the different functions of PhoR. Initially, one group found that the 158 amino acids on the N-terminus are not essential for kinase function, but do play a role in the phosphatase activity (30). Another group worked to define and characterize any periplasmic portion of PhoR. They found that there is no large domain in the periplasm but that there are two predicted transmembrane helices at the N-terminus (residues 10-60) and propose that the portion of PhoR in the cytoplasm (from residues 52-220), found before the conserved kinase domains, plays a role in regulation by sensing a cytoplasmic signal (31).

Based on studies of EnvZ, another member of the class I family of histidine kinases, the domains of *E. coli* PhoR were identified. Specifically kinase domains identified were: a CA domain from residues 267-431 and a DHp domain from residues 193-267 (32). Additionally, this group identified residues 72-106 as a charged region (CR) consisting of about 35 residues predicted to have a high helical content with six positively charged residues that may aid in stabilizing interaction with the membrane. PhoR also contains a PAS domain from residues 109-196 (33). PAS domains are conserved domains found in all kingdoms of life that play a role as signal sensors and interaction hubs in signal transduction proteins. Often, the PAS domains act to detect physical or chemical signals and stimulate the activity of an output domain for catalysis or DNA binding (18).

Based on an *E. coli* PhoR sequence modeled on the recent structure identified for histidine kinase VicK from *Streptococcus mutans* (34), the amino acid residues that make up each domain appear to be: PAS 109-194, DHp 195-267, and CA 268-431 (Fig 1.2).

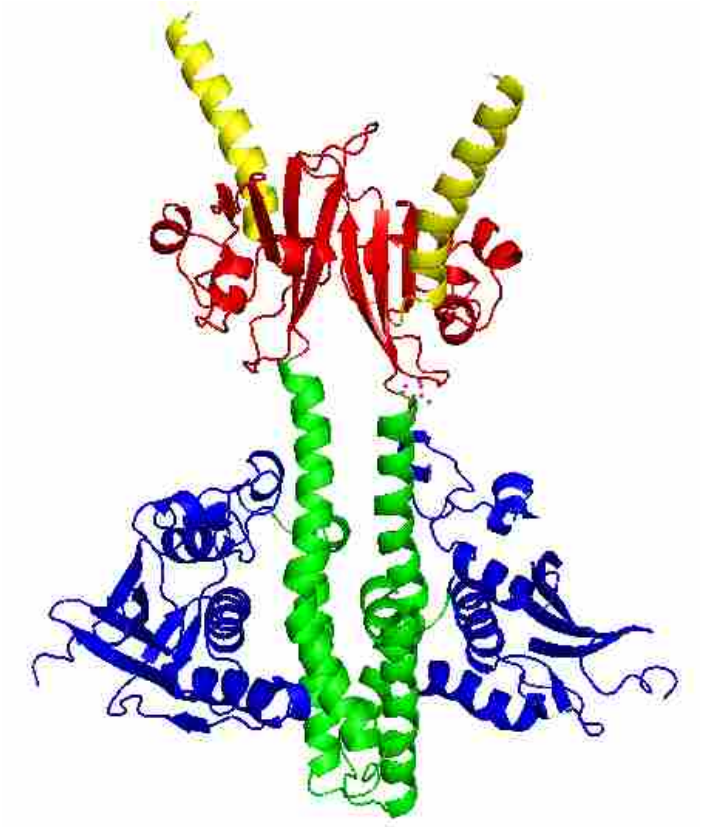


Figure 1.2 Modeled structure of *E. coli* PhoR
A dimer of the cytoplasmic portion of PhoR from *E. coli* modeled on VicK from *Streptococcus mutans* using Protein Homology/analogY Recognition Engine V 2.0 (Phyre2) (35). The PAS domains are in red, the DHp domains are in green, and the CA domains are in blue.

Using precise deletion mutants of PhoR, it was confirmed that the CA and DHp domains are essential for autokinase activity of PhoR and that the phosphorylated DHp domain transfers the phosphate group to PhoB (32). Further, Carmany (2003) showed that the phospho-PhoB phosphatase activity is contained in the DHp domain. However they noted the weak activity and postulated that other domains of PhoR may interact with the DHp domain to trigger phosphatase activity.

Studies of a PhoR PAS domain in *Bacillus subtilis* have added to the knowledge of phosphate based regulation of gene expression (18, 20, 36, 37). PhoR in *E. coli* is similar to *B. subtilis* PhoR, except that it lacks the extracellular PAS domain, but both proteins have a cytoplasmic PAS domain. It appears that while these domains are often found in histidine kinases of two component signal transduction systems, the extracellular PAS domain of *B. subtilis* PhoR is not required for proper signaling and kinase activity. The mechanism for PhoR signaling in *B. subtilis* is still unknown.

1.1.3.2 PhoB

PhoB directly interacts with DNA and RNA polymerase to induce expression of Pho regulon genes. PhoB has two domains, an N-terminal receiver domain and a C-terminal DNA binding domain. The unphosphorylated receiver domain of PhoB acts to silence the DNA binding effector domain (38). Studies have shown that PhoB interacts directly with DNA at specific conserved sequences, called the pho box, and the σ^{70} subunit of RNA polymerase (39, 40).

In addition to PhoR, CreC and acetyl phosphate can transfer a phosphate group to PhoB leading to activation, but the physiological effect of this activation is still unclear (41-43). This non-PhoR mediated phosphorylation has led to speculation on the potential for crosstalk between different regulatory systems through kinases and phosphatases that not only act on their specific response regulator, but also on other response regulators (44). This cross activation may play a role in the expression control of many other genes and allowing the cell to respond to a variety of signals using existing signal transduction systems.

1.1.4 PstSCAB PhoU

One of the major operons in *E. coli* whose expression is controlled by PhoB contains the *pstS, pstC, pstA, pstB,* and *phoU* genes. These genes encode for proteins that play key roles in P(i) uptake as well as expression control of the Pho regulon. The PstSCAB proteins form a phosphate specific ATP Binding Cassette (ABC) transporter and PhoU is a negative regulator of Pho regulon expression.

1.1.4.1 ABC transporters

ABC transporters are a family of proteins that developed anciently and are found in all domains of life. There are two main classes, one where products are transported outside of the cell and the other class that imports substrates. ABC transporters play an important role in the pathogenesis of many organisms (45). These transporters use ATP hydrolysis to power transport of substances across the membrane. Many studies have elucidated the mechanism by which ABC transporters function. Specifically, in bacteria an ABC transporter responsible for transporting maltose into the cell has been well characterized (46).

The maltose transporter complex is composed of two transmembrane subunits, MalF and MalG, and two subunits of MalK, the cytoplasmic ATP binding protein. Additionally, maltose-binding protein (MBP) acts as a periplasmic substrate binding protein that binds maltose and presents it to the transporter complex. Several studies have identified the structure of the transporter at different states of the transport process. Specifically, a resting state where no ATP or MBP are bound where the pore has an inward facing conformation, the maltose binding site is exposed to the cytoplasm. When MBP is present but ATP is not, MBP makes contact with the transporter in a closed conformation where the transmembrane subunits are rotated compared to

the resting state. When ATP is added, the ATP binding domains undergo a conformational change that rotates the transmembrane proteins and opens the MBP. This allows the MBP to transfer maltose into the transmembrane pore and be transported into the cell through ATP hydrolysis. In each state the tertiary structure of the subunits remains the same, but the conformation of the complex changes (46).

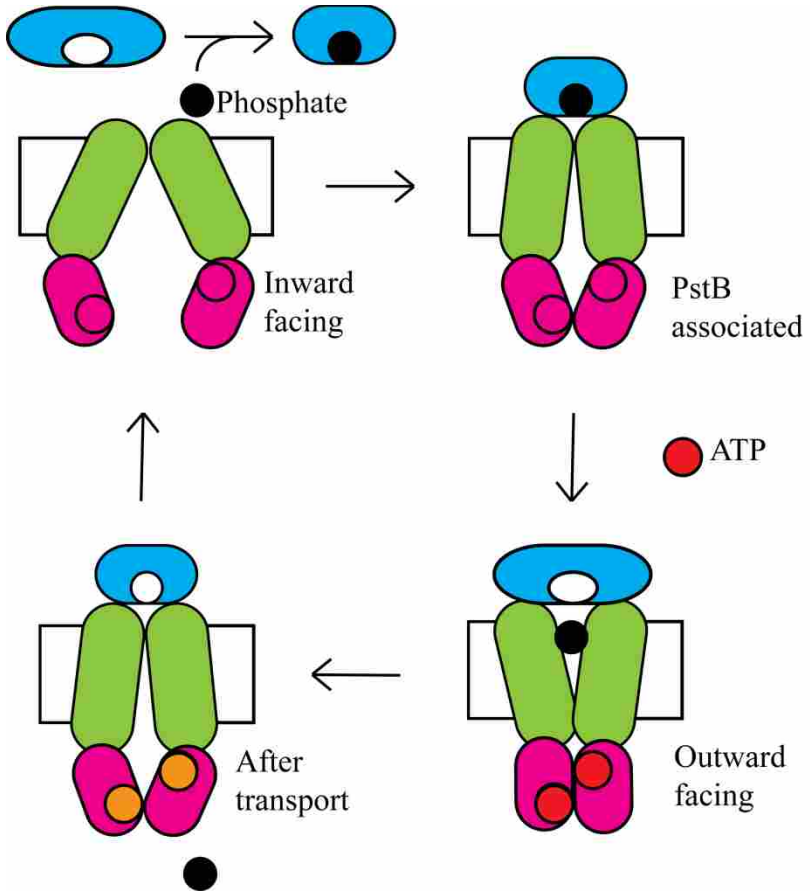


Figure 1.3 ABC transporter structural arrangements throughout transport
In transport, first the periplasmic substrate binding protein binds the substrate (phosphate in the case of PstSCAB). Next, the charged PstS (blue) binds the PstC and PstA (green) pore proteins. When ATP binds the PstB (pink) proteins the conformation of the complex is shifted to allow the substrate to enter the transporter. Hydrolysis of the ATP allows the substrate to complete transport and the transporter to return to its resting state.

Studies have shown that for importers like the maltose transporter, binding of the loaded periplasmic substrate binding protein is necessary to trigger ATP hydrolysis and that substrate

alone does not trigger the hydrolysis of ATP. It is thought ATP binding, when a loaded periplasmic substrate binding protein is attached to the complex, triggers the conformational change to open the pore to accept the substrate. However, it is the hydrolysis of the ATP that powers the complex to reset so that it is ready for another cycle of transport. It is also interesting that a "scoop loop" from the MalG protein is thought to aid in transferring the tightly bound maltose from MBP to the transporter complex by inserting between two lobes of the maltose binding pocket of MBP to not only release the maltose but inhibit its immediate re-association with MBP (46).

Recently, the structures that this model is based on have been analyzed with an elastic network model to better understand the transition between structure conformations. They found that the nucleotide binding domains close first, which leads to the transmembrane domain conformational change. Then the MBP conformation is altered, allowing translocation of the maltose (47).

Another recent study further characterizes the transport mechanism. They questioned why the resting state of the transporter was inward facing. One hypothesis was that the resting state of the transmembrane proteins is because the naturally lowest energy state of these proteins is to be in this conformation. However, when isolated alone in the membrane, the transmembrane proteins formed an intermediate structure between the inward and outward facing conformations. They found that introduction of nucleotide free MalK into the complex allowed for the inward facing conformation. They also showed that mutant MalK, which bound more tightly to the pore complex, inhibited transport (locking the complex into the inward conformation). However, compensatory mutations, that weaken the protein/protein interaction, allowed for proper transport. They propose that it is through ATP hydrolysis that the transition from inward to

outward occurs and allows for release of maltose into the cytosol (48). Another group has studied how ATP binding can direct dimerization of MalK. They found that the entropic force caused by water drove the conformational change and that ATP hydrolysis then leads to a loss of water-entropy and allows for dissociation of the dimer (49).

Additionally, the transport mechanism for a different ABC transporter has been studied using limited proteolysis. For the MisQMP2 transporter (a positively charged amino acid transporter from *Salmonella enterica* serovar Typhimurium) they found that although there are distinct structural differences with the maltose transporter, the conformational changes that occur during transport are similar for both systems (50).

While studying ABC transporters, mechanisms to control the activity of the transporters have been identified. Bacterial cells selectively use a high energy sugar like glucose before maltose when both are available. Enzyme IIA (EIIA) of the glucose specific phosphotransferase system binds to and inhibits a variety of transporters in the presence of glucose. A structure of EIIA interacting with the MalK cytoplasmic portion of the transporter, locking it into the inward facing conformation and preventing ATP hydrolysis and transport through the complex has been characterized (51). There is much more evidence for the common theme that transport proteins are linked to regulation of transcription and metabolism in bacteria (52).

Another interesting finding that may be applicable to the Pho regulon expression control studies is that in a two-component signal transduction systems involved in antibiotic resistance for gram positive bacteria (BceR/BceS), an ABC transporter is the sensor of the regulatory system and passes the signal to the histidine kinase (16). In this example, the ABC transporter is actually an export complex (as opposed to the import function of PstSCAB), but the principle of

a transporter complex acting as the signal sensor for a histidine kinase may still apply to the PhoR/PhoB regulation system and PstSCAB transporter complex.

1.1.4.2 PstSCAB

Since phosphate sources do not readily diffuse through cell membranes, it is essential for all forms of life to have a way to import phosphate into the cell. In *E. coli* there are two main systems for transport of phosphate: phosphate inorganic transport (Pit) systems that have low affinity and, the phosphate specific transport (Pst) system that displays high affinity for phosphate. The Pst system is encoded by genes in the *pstSCAB phoU* operon (53-55). PstSCAB proteins compose an ATP-binding cassette (ABC) transporter, a common family of transporters that use ATP hydrolysis to power transport of P(i) across the cell membrane. PstS is a periplasmic protein that binds P(i), PstC and PstA form a pore to transport the P(i), and PstB is an ATPase that probably exists as a dimer in this ABC transporter and uses the energy from ATP hydrolysis to transport P(i) into the bacterial cell (11). The transport of P(i) by this complex is very specific. Interestingly, in *Mycobacterium smegmatis* a second ABC transporter (Phn) also appears to transport P(i) directly in addition to the Pst transporter complex (56).

PstSCAB is essential for sensing environmental P(i). It seems that PhoR cannot directly detect P(i) on its own. Proper negative regulation of this signal cascade requires PstSCAB, PhoU, and PhoR. The membrane localization of all these components has led to a theory that they interact at the membrane to form an inhibitory complex when environmental P(i) is abundant (3, 11). The conformational state of the PstSCAB transporter, that signals the abundance of P(i), may be recognized by PhoU and that signal passed along to PhoR to control the phosphate response.

There are other phosphate transporters found in *E. coli*. PitA is a low-affinity P(i) transporter that uses proton motive force to import metal phosphates into the cell and is expressed constitutively. PitB is another low affinity P(i) transporter that is also found in *E. coli* (57, 58). One study found that in a *pstS* mutant bacteria, over expression of *pitA* or *pitB* restored both P(i) transport and allowed for some phosphate signaling of the PhoR/PhoB system, but that this regulation still required the presence of the *pstC*, *pstA*, and *pstB* genes (59). Recent studies imply that the PstSCAB complex acts as the main transporter of P(i) in *E. coli* given: PitA is primarily a transporter of divalent metal cations that are in complex with P(i) and its expression is controlled by the environmental divalent cation concentration; and PitB is poorly expressed under normal growth conditions (3). Another group has found that PitA has a role in copper and arsenic resistance in the archaeal species *Metallosphaera sedula*, further confirming its primary role in metal transport (60).

1.1.4.3 PhoU

PhoU is a member of the *pstSCABphoU* operon and does not have significant homology to proteins with known activity. PhoU has a role in the regulation of the Pho regulon. Studies of PhoU has revealed its role in regulation, proteins that PhoU interacts with, and a potential protein/protein interaction model that implies a mechanism for signaling the switch of PhoR activity.

1.1.4.3.1 Role in Regulation

PhoU was found early on in the study of the phosphate response to play a role in negatively regulating the Pho system. In *E. coli* cells where PhoU is mutated, Pho regulon genes

are constitutively expressed and cells grow poorly. Often cultures accumulate compensatory mutations in the PstSCAB and/or the PhoB and PhoR encoding genes. It seems that constitutive expression of PstSCAB in the absence of PhoU is toxic to the cells. The reason for this is not well known, but it appears that PhoU plays a negative regulatory role in transport by the PstSCAB complex. This may also explain the high frequency of compensatory mutations in *phoU* knockout strains (61-63).

PhoU has been implicated in aiding in the regulation of other systems in addition to the phosphate response. In one study, they identified a link between regulation of the phosphate response and iron response to control virulence genes in *Edwardsiella tarda* (64). PhoU is implicated in bridging the regulation of the phosphate response and a ferric uptake regulator (Fur) that coordinates pathogenesis with low phosphate and iron concentrations as would be found inside of host of this pathogen.

1.1.4.3.2 Structures

Multiple crystal structures have been solved for PhoU proteins from various organisms (65-67). In general, the structures consist of two fairly symmetric, three alpha-helix bundles. Often, metal ions have been found associated with these structures that are coordinated by highly conserved amino acids. It is still not clear whether these ions play a role in PhoU's function or if they are an artifact from the purification and crystallization of the proteins. The work described in Chapter 2 addresses metal binding of PhoU from *E. coli* and finds that metal binding is important for proper function when *phoU* is expressed from a single gene on the chromosome, but the phenotype can be ameliorated by expression *phoU* from a medium copy number plasmid (1).

One of these structure studies suggests that PhoU has folds similar to Bag domains (a class of cofactors of the eukaryotic chaperone Hsp70 family). Given this similarity, they propose that PhoU may associate with the ATPase domain of PhoR and cause it to release PhoB, thus turning off the signaling cascade in a manner similar to Bag domain's association with Hsp70 (65). PhoU/PhoR interaction has since been mapped to be between PhoU and the PAS domain of PhoR in addition to the ATPase domain (1)(Chapter 3 of this work). A recent study of PhoU structure from *Pseudomonas aeruginosa* confirmed the same fold and a dimer quaternary structure. However, they did not observe metal binding (67).

1.2 Regulation of the Pho Regulon

In addition to PhoB, recent studies have identified other proteins that interact with PhoR. The work presented in Chapter 2 describes results that show PhoR interacts with both PhoU and PstB (1). Another study found that PhoR interacts with the EIIA protein of the nitrogen-related phosphotransferase system (68). That study shows that $EIIA^{Ntr}$ plays a role in enhancing the activity of PhoR in the phosphate response. These results demonstrate that the regulation of the phosphate response has multiple signal inputs that may help cells adapt specifically to various environments and combinations of nutrients and environmental signals. Better understanding about these different signal inputs and their affect on Pho regulon expression will more fully elucidate the phosphate response and the complex regulation control that is needed for optimal growth in a variety of environments and may be pertinent to other regulatory systems with similar control mechanisms and networks of signal inputs.

1.2.1 Crosstalk

One unique theme in studies of Pho regulon control and in studies of other two-component signal transduction systems is the potential for crosstalk between different regulatory systems. PhoR is the major activator of PhoB, but other histidine kinases and acetyl phosphate molecules have been shown to phosphorylate PhoB as well. Some have hypothesized that these interactions allow for multiple signals to play a role in expression of the Pho regulon genes. Acetyl phosphate activity to phosphorylate PhoB may indicate a role of fatty acid availability in the activation of PhoB. CreC is a predicted catabolite repression sensor kinase that has been shown to activate PhoB by phosphorylation (11, 69). This interaction may imply a cross regulation between the Pho regulon and central carbon metabolism (70). Additionally, the interactions of PhoU and PhoR with Fur and $EIIA^{Ntr}$ respectively may point to roles in regulation of other response genes independent of the Pho regulon genes as another mode of crosstalk between bacterial regulatory networks (64, 68).

1.2.2 Stochastic activation

One study of the activation of Pho regulon looked at the activation of this system on a single cell level (44). This study found that the Pho system is often expressed in a stochastic manner in different bacterial cells of the same population. In this manner, the overall response of the population of bacteria can be observed as a degree of activation, but each cell is either fully active or inactive. Some potential causes of this activation profile include activation of PhoB by non-cognate histidine kinases, for example, the CreC and acetyl phosphate molecules that phosphorylate PhoB in the absence of a phospho-PhoR signal. Given that *phoR* background expression is low, this stochastic activation by PhoR independent mechanisms may be necessary

for quick adaptation to environmental changes. This may be a reason why PhoB is sensitive to these non-PhoR signal sources. So that at any given time, there can be a few cells in every population that are primed to induce expression of the Pho regulon, thus making the population more flexible in its response to phosphate level changes. PhoR translation is slow, so there may be some cells in the population that have no PhoR proteins synthesized to initiate the signal (3).

Another study observed the transcript and protein levels of several components during phosphate starvation (70). They followed interactions of various proteins by using FRET analysis. They found that CreC interacts with PhoB faster than the PhoR/PhoB interaction.

1.2.3 Other Pho Regulon Genes

There are 31 genes shown to be under direct control of PhoB mediated transcriptional control and several other genes that are believed to be but direct evidence is still lacking (3). Many of these genes have roles in increasing the cell's efficiency in P(i) uptake and scavenging P(i) from other sources of phosphate. The following section will discuss some of the major genes found in the Pho regulon.

1.2.3.1.1 PhoE

It is essential that phosphate can enter the periplasmic space of gram negative organisms so that is available for transport into the cell across the inner membrane. One member of the Pho regulon that aids in meeting this requirement is *phoE*. PhoE protein forms a porin in the outer membrane. It is considered a nonspecific porin (similar to OmpF and OmpC). However, it does show some selectivity toward anions such as organophosphates and P(i) due to several lysine

residues. The pore is formed by three monomers that each contain a central channel forming a trimer that has tight association with membrane lipid head groups of the outer membrane (11).

1.2.3.1.2 Phosphonate Pathways

Phosphonates, molecules that have a direct carbon phosphorous bond, are found in several organisms and are an environmental source of phosphorous that *E. coli* has developed a method to transport and metabolize. In particular, there are two pathways for utilization of phosphonates under Pho regulon expression control: the Phosphonatase and the C-P lyase pathways. These pathways differ in their mechanisms for breaking down phosphonates, but they both offer a way for the cell to utilize and incorporate phosphate from phosphonate molecules. Each of these pathways contain enzymes and transport mechanisms specific for the type of phosphonates that are broken down (11).

1.2.3.1.3 Bap

Cells can increase the local availability of free P(i) by cleaving phosphate from molecules outside the cell. Bacterial alkaline phosphatase (Bap), a nonspecific phosphomonoesterase is an enzyme encoded by the *phoA* gene that is part of the Pho regulon that can cleave phosphate from other molecules. When P(i) is limiting, Bap synthesis can increase over 1,000-fold due to phopho-PhoB activation. This protein is secreted into the periplasm and is post-translationally modified by cleaving an N-terminal signal peptide and the removal of Arg-22 by an aminopeptidase. The monomers then form disulfide bonds and dimerize to make an active protein (11). Various innate properties of Bap have made is useful for following Pho regulon expression control as well as for use in many other techniques (1, 43, 63, 71-74).

1.2.3.1.4 Ugp

Another source of phosphate available to *E. coli* cells is *sn*-glycerol-3-phosphate (G3P). This organophosphate is transported into the cell with a specific transporter made up of the Ugp proteins and processed by UgpQ, a glycerophosphoryl phosphodiesterase. These proteins are encoded by the *ugpBAECQ* operon which is under Pho regulon expression control (11). The structure of this transporter complex has been identified and the transporter activity characterized to show that transport activity did not appear to be sensitive to phosphate or PhoU (75). Interestingly, there is another G3P transport system in *E. coli*, GlpT. However, the GlpT transporter only transports G3P for use as a carbon source for the bacterial cell. It is independent of the Ugp system and not a part of the Pho regulon (11).

1.3 Pho Regulon and Pathogenesis

There is strong evidence of a link between the Pho regulon and pathogenesis (6, 9, 76). There are many studies that find the pathogenicity of various organisms is affected by Pho regulon members (5-8, 76-79). These studies show that the disruption of the Pho regulon leads to a variety of effects, such as: bacterial resistance to host responses, reduced ability to form plaques, and slower growth rates. All of these functions may be impacted by the Pho regulon and can lead to poor pathogenicity when the Pho regulon expression control is lost.

These results are not unexpected and show that uncontrolled expression of the Pho regulon has a deleterious effect on bacterial cells. This is seen in the phenotype of PhoR/B and PstSCAB PhoU mutants where cell surface modification and response to oxidative stress are linked to Pho Regulon control. Mutations in the PstSCAB complex proteins also affect certain

bacterial cells' ability to adhere to host's intestinal cells, weakening their pathogenesis. All of these studies demonstrate that not only is the regulation of the phosphate response important for adapting to changing levels of environmental phosphate, but it is also important in many cell pathways involved in virulence. Continued study of these interactions may lead to new insight about disease as well as hew to prevent and treat these diseases.

1.3.1 PhoU and Pathogenesis

PhoU plays a role in some bacterial cells' ability to successfully colonize a host and cause disease. Specifically, a few studies have looked that the effect that PhoU has in allowing the development of persisters in *Mycobacterium tuberculosis* and *E. coli* (5, 7). A persister is a pathogen that develops increased resistance to treatment. The mechanisms for persistence are not well understood and are still being studied. PhoU and the Pho regulon may play an important role in these pathways and further study could lead to advances in treatment.

Another study linked PhoU to toluene resistance in *Pseudomonas putida* (79). They found that a transposon insertion into the *phoU* gene of *P. putida* led to susceptibility to toluene, in an otherwise resistant strain. They hypothesize that this is due to an increase in a porin protein that makes the membrane more permeable to toluene than the wild-type cells.

The direct link between PhoU and the ferric uptake regulator (Fur) in *Edwardsiella tarda* is an example where PhoU may directly affect control of the expression of virulence genes that code for a type III and VI secretion systems (64).

The many systems that link Pho regulon expression control mechanisms to pathogenesis present several interesting questions. Is the environmental level of nutrients important for a successful bacterial colonization of a host? Are there other nutrient sensing systems that affect

pathogenesis? Are these phenotypes due solely to a poor growth phenotype caused by unregulated expression of the Pho regulon or is there a direct mechanistic link between the Pho regulon and control of virulence factors? While current data address some of these questions, further study of these issues is essential to fully understand the mechanisms of how pathogenesis is affected by the Pho regulon and expression control.

1.4 Conclusions

The study of the Pho regulon and its control has added understanding to how *E. coli* cells adapt to changes in the environment. Studies of the phosphate response show how complex and intertwined the regulatory mechanisms for adapting to environmental changes are in bacteria. The effects of the Pho regulon and its expression control are seen throughout the cell and lead to changes in cell growth, equilibrium, and pathogenesis. Continued study of this system will no doubt bring further understanding of how these cells grow and adapt to changes. In particular, the mechanism for transferring the signal may be deduced with further study and lead to greater understanding of how this and similar two component signal transduction systems work.

While many parts of this system have been studied and are now well understood, there is still much to learn. How does PhoR receive signals correlated to the phosphate level of the environment? What is PstSCAB and PhoU's role in this signaling? Is there a small molecule that is produced to pass on the signal or is direct protein/protein interaction essential? How do PhoU and other Pho regulon genes cause the effects seen in pathogenesis and could these properties then become targets for therapies?

Knowledge gained about this system may apply to the many other two-component regulatory systems found in other forms of life. This may lead to advances in our understanding

of molecular interactions of signaling proteins, transcriptional control, signal transduction, and many more areas that have potential for application in molecular biology, medicine, ecology, bioremediation, and a multitude of other fields.

Chapter 2 The PhoU protein from *Escherichia coli* interacts with PhoR, PstB and metals to form a phosphate-signaling complex at the membrane

2.1 Abstract

Robust growth in many bacteria is dependent upon proper regulation of the adaptive response to phosphate (P_i) limitation. This response enables cells to acquire P_i with high affinity and utilize alternate phosphorous sources. The molecular mechanisms of P_i signal transduction are not completely understood. PhoU along with the high affinity, P_i-specific ATP-binding cassette transporter PstSCAB and the two-component proteins PhoR and PhoB are absolutely required for P_i signaling in *Escherichia coli*. Little is known about the role of PhoU and its function in regulation. We demonstrate using bacterial two-hybrid analysis and confirmatory co-elution experiments that PhoU interacts with PhoR through its PAS domain and that it also interacts with PstB, the cytoplasmic component of the transporter. We also show that the soluble form of PhoU is a dimer that binds manganese and magnesium. Alteration of highly conserved residues in PhoU by site-directed mutagenesis shows that these sites play a role in binding metals. Analysis of these *phoU* mutants suggests that metal binding may be important for PhoU membrane interactions. Taken together, these results support the hypothesis that PhoU is involved in the formation of a signaling complex at the cytoplasmic membrane that responds to environmental P_i levels.

2.2 Introduction

Escherichia coli employ seven genes whose products sense environmental phosphate (P_i) and control the expression of the Pho regulon (3, 11). These genes include *phoB*, *phoR*, *pstS*, *pstC*, *pstA*, *pstB* and *phoU*. Together, these genes are necessary and sufficient for P_i-signal transduction. While the identity of these signaling proteins has been known for some time(2, 80, 81), the mechanisms by which they function to transduce the P_i signal have not yet been fully elucidated.

The hub of the signaling pathway consists of the two-component signaling proteins PhoB and PhoR (82). PhoB is a typical winged-helix response regulator that upon aspartyl phosphorylation forms a dimer, which binds to DNA sequences upstream of Pho regulon genes to recruit RNA polymerase and initiate transcription (42, 83-85). PhoR is the bifunctional histidine autokinase/phospho-PhoB phosphatase that donates a phosphoryl group to PhoB when environmental P_i is limiting and removes the phosphoryl group from PhoB~P when environmental P_i is abundant (32, 84). PhoR is an integral membrane protein that does not contain a significant periplasmic domain, but does contain a membrane spanning region (Mem), a cytoplasmic charged region (CR), a Per-ARNT-Sim (PAS) domain (32, 86), with prototypical Dimerization/Histidine phosphorylation (DHp) and Catalytic ATP binding (CA) domains at its C-terminus (29). Since PhoR does not contain a significant periplasmic sensory domain, it is assumed that its PAS domain senses a cytoplasmic signal, but the nature of the signal is not known. Changes in intracellular P_i concentration are not the signal because these levels remain relatively constant in the presence of various extracellular P_i concentrations and different P_i-signaling states (87, 88).

It has been suggested that the PhoR/PhoB proteins assess P_i availability by monitoring the activity of the Pst transporter (89). The Pst proteins form a type I ATP-binding cassette (ABC) importer that exhibits high affinity, P_i-specific transport (90-92). PstS is the periplasmic P_i-binding protein, PstC and PstA are integral membrane proteins that form the pore through which P_i passes and PstB is the dimeric, cytoplasmic ATPase. Based upon data from other well-described ABC transporters (93), P_i-bound PstS likely stimulates the ATPase activity of PstB to facilitate the significant conformational changes that are involved in P_i transport (91). When external P_i levels are above ~ 4 μM the transporter exists in a high activity state that stabilizes the phospho-PhoB phosphatase state of PhoR (2). The low activity state transporter signals for the high autokinase activity of PhoR that leads to elevated levels of phospho-PhoB.

In addition to the four transporter proteins, PhoU is also required for P_i-signal transduction, but not for transport through the PstSCAB complex (62). When *phoU* is mutated or deleted, PhoR is a constitutive PhoB kinase leading to high expression of the Pho regulon genes. This is accompanied by poor growth and frequently leads to the accumulation of compensatory mutations in *phoR, phoB,* or the *pstSCAB* genes. It has been reported that PhoU is a peripheral membrane protein that is a negative regulator of the signaling pathway and that it modulates P_i transport through the PstSCAB proteins (62, 63, 94). Multiple crystal structures have been reported for PhoU proteins in various organisms (65, 66, 95). Each of the structures shows that PhoU consists of two fairly symmetric, three alpha-helix bundles (See Fig. 2.1). These structures show several different quaternary structures: monomer, dimer, trimer, and even hexamer. Metal ions are found associated with two of these structures and are coordinated by highly conserved amino acid residues that are found in each three-helix bundle (66, 95). One metal group is bound

by residues that correspond to *E. coli* PhoU D58, N62, E100 and D104 and another bound by D161, D165, E200, and D204 (66) (Fig. 2.1).

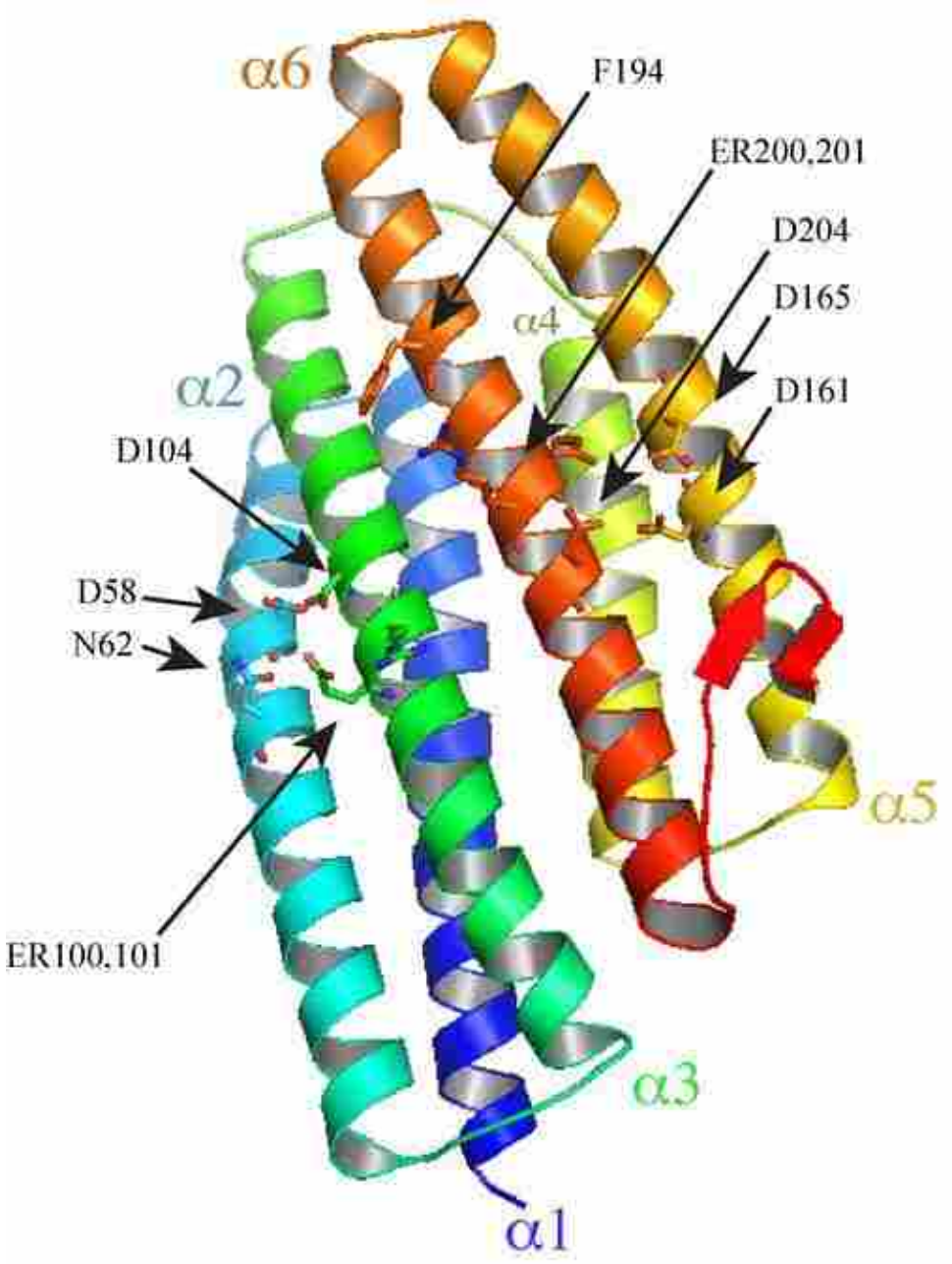


Figure 2.1 Modeled structural fold of *E. coli* PhoU
The structure of the PhoU protein from *Escherichia coli* was modeled based upon known structures from the Protein Data Bank using the Phyre2 server. The protein is colored blue to red going from the amino terminus to the carboxyl terminus. The six helices are labeled from to α1 to α6. The side chains of several highly conserved amino acids are shown. In particular, D58, N62, E100, and D104 form a putative metal binding site between helix 2 and 3 and D161, D165, E200, and D204 form another metal binding site between helix 5 and 6. Also shown is the F194 residue used in the fluorescence assays.

It is not known how PhoU functions in the signaling pathway. However, two general classes of models have been suggested; PhoU may mediate the formation of a signaling complex between the PstSCAB transporter and PhoR (11, 65) or it may produce a soluble messenger that is recognized by the cytoplasmic domains of PhoR (consistent with observations reported by Hoffer and Tommassen (59) and by Rao and Torriani (96)). No experimental evidence has yet been presented supporting physical interactions between signaling proteins.

We employed bacterial adenylate cyclase two-hybrid (BACTH) analysis as well as co-elution experiments to show that PhoU interacts with both PhoR and PstB. Our results show that the PhoU/PhoR interaction involves the PAS domain of PhoR. This is consistent with the presence of a P_i-signaling complex consisting of the PstSCAB transporter, PhoU, and PhoR. We also show that *E. coli* that soluble PhoU is a dimer that binds manganese and magnesium at conserved sites and that metal binding may be important for localizing PhoU to the membrane.

2.3 Materials and methods

2.3.1 Bacterial stains, growth media, and growth conditions.

The bacterial strains and plasmids used in this study are shown in Table 2.1. P1*clr* transductions were carried out as previously described (97). Strains were grown in LB medium (98) at 37°C or in MOPS defined media containing either 0.06% glucose and 2.0 mM P_i (MOPS HiP_i) or 0.4 % glucose and 0.1 mM Pi (MOPS LoP_i) (99). When indicated, kanamycin, ampicillin or chloramphenicol were used at 50 μg/ml, 50 μg/ml, and 40 μg/ml respectively. BACTH strains were grown on MOPS LoP_i plates with 0.2% maltose as a carbon source with ampicillin (50 μg/ml), kanamycin (30 μg/ml), isopropyl-β-D-thiogalactopyranoside (IPTG) at

0.5 mM, and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) at 40 μg/ml. The His-tagged versions of PhoR, PhoU, and PstB, and the other site-directed mutants were created using the QuikChange site-directed mutagenesis kit from Agilent Technologies and verified by DNA sequence analysis (see Table S1 for primer sequences).

Table.2.1 Strains and plasmids

***E. coli* strains**	**Description**	**Source**
BW25113	wild-type strain	(100)
CSH126	*recA Tn*10, tet^R	(97)
BM261	P_{pstS}::P_{tac} Δ*pitA*::frt Δ*pitB*::frt pRR48	(63)
BM263	P_{pstS}::P_{tac} Δ*pitA*::frt Δ*pitB*::frt Δ*phoU* pRR48	(63)
BM265	BM263 *pstB*::pBU3 *recA*1 encodes wild-type *phoU*	this study
BM266	BM263 *pstB*::pBQ3 *recA*1 encodes *phoU*Quad	this study
BTH101	F^- *cya-99, araD139, galE15, galK16, rpsL1 (*Str^R*), hsdR2, mcrA1, mcrB1*	EuroMedex
BW26337	BW25113 Δ*pstSCAB-phoU*::frt	(100)
BW26390	BW25113 Δ*pstB*::frt	Yale CGSC
JW0390-2	BW25113 Δ*phoR::kan*	Yale CGSC
ANCH1	Δ*phoBR::kan*	(101)
SG1	BW26337, Δ*phoBR::kan* from ANCH1 by P1*clr* transduction	this study
Plasmids		
pIB307	pMAK705-based vector with a temperature-sensitive replicon	(102)
pKG116	pACYC184-based replicon cam^R *nahR*	(103)
p116U2	cam^R *phoU* expression plasmid, salicylate inducible	(63)
p116U2His	p116U2 with a C-terminal 6X-His tag	this study
p116U2D58A	p116U2 with D58A mutation	this study
p116U2E100A	p116U2 with E100A mutation	this study
p116U2R101A	p116U2 with R101A mutation	this study
p116U2E200A	p116U2 with E200A mutation	this study
p116U2R201A	p116U2 with R201A mutation	this study
p116U2D204A	p116U2 with D204A mutation	this study
p116U2D58A,N62A	p116U2 with D58A and N62A mutations	this study
p116U2E100A,R101A	p116U2 with E100A and R101A mutations	this study
p116U2E200A,R201A	p116U2 with E200A and R201A mutations	this study
p116U2Quad	p116U2 with E100A, R101A, E200A, and R201A mutations	this study
pBU3	cam^R temperature-sensitive encodes *phoU*	this study
pBQ3	cam^R temperature-sensitive encodes *phoU*-Quad	this study
pRR48	pBR322-based replicon, Amp^r $lacI^q$	(104)
p48phoR	pRR48 based *phoR* expression plasmid, lactose inducible	this study
p48phoRNHis	p48phoR with an amino terminal 6X-His tag	this study
p48pstB	pRR48 with *pstB* cloned into the NdeI and KpnI sites, lactose inducible	this study
p48pstB-His	p48pstB with a carboxyl-terminal 6X-His tag	ths study
pKT25	kanR, BACTH plasmid for T25 fragment of AC	EuroMedex
pUT18C	ampR, BACTH plasmid for T18 fragment of AC	EuroMedex
pUT18zip	ampR, BACTH plasmid for T18 fragment fused to leucine zipper	EuroMedex
pKT25zip	kanR, BACTH plasmid for T25 fragment fused to leucine zipper	EuroMedex
pKT25phoU	kanR, BACTH plasmid for *phoU*-T25 fusion	this study
pKT25quad	kanR, BACTH plasmid for *phoUQuad*-T25 fusion	this study
pKT25pstB	kanR, BACTH plasmid for *pstB*-T25 fusion	this study
pUT18phoU	ampR, BACTH plasmid for *phoU*-T18 fusion	this study
pUT18Cquad	ampR, BACTH plasmid for *phoUQuad*-T18 fusion	this study
pUT18CD85A	ampR, BACTH plasmid for *phoUD85A*-T18 fusion	this study
pUT18CE100A	ampR, BACTH plasmid for *phoUE100A*-T18 fusion	this study
pUT18CA147E	ampR, BACTH plasmid for *phoUA147E*-T18 fusion	this study
pUT18CE200A	ampR, BACTH plasmid for *phoUE200A*-T18 fusion	this study
pUT18CE200A, R201A	ampR, BACTH plasmid for *phoUE200A, R201A*-T18 fusion	this study
pUT18CphoRN-C	ampR, BACTH plasmid for *phoR*-T18 fusion	this study
pUT18CRN-DHp	ampR, BACTH plasmid for N-DHp portion of PhoR-T18 fusion	this study
pUT18CRN-PAS	ampR, BACTH plasmid for N-PAS portion of PhoR-T18 fusion	this study
pUT18CRN-CR	ampR, BACTH plasmid for N-CR portion of PhoR-T18 fusion	this study
pUT18CRPAS	ampR, BACTH plasmid for PhoR PAS-T18 fusion	this study
pUT18CRCR-C	ampR, BACTH plasmid for CR-C portion of PhoR-T18 fusion	this study
pUT18CRPAS-C	ampR, BACTH plasmid for PAS-C portion of PhoR-T18 fusion	this study
pUT18CRDHp-C	ampR, BACTH plasmid for DHp-C portion of PhoR-T18 fusion	this study

2.3.1.1 *E. coli* PhoU structure prediction

The three-dimensional structure of the PhoU protein from *E. coli* was predicted by using the Phyre2 web site (http://www.sbg.bio.ic.ac.uk/phyre2) (35). The PhoU sequence was submitted as a query and was analyzed in intensive mode. The server indicated that 216 residues were modeled at >90% accuracy. The predicted structures were displayed using MacPyMol (Schrödinger, LLC).

2.3.1.2 Bacterial adenylate cyclase two-hybrid analysis

Two-hybrid screens were carried out using the plasmids pKT25 and pUT18C from the BACTH kit and grown in the BTH101 indicator strain (EuroMedex). Control plasmids, coding for T18 and T25 fragments that were fused to a leucine zipper (GCN4), were also provided by the BACTH System Kit. Plasmid constructs encoding *phoR*, *phoR* truncations, *pstB*, and *phoU* fusions to the T18 or T25 fragments were generated by cloning PCR fragments that incorporated *Xba*I and *Kpn*I restriction sites by using the primers listed in Table S1. About 1 µl of overnight cultures were spotted on MOPS Maltose, ampicillin, kanamycin, IPTG, X-Gal plates and grown at 30°C until color developed.

2.3.2 β-galactosidase activity assays.

β-galactosidase has long been utilized as an easily assayable enzyme to measure gene expression and in the BACTH system gene expression is linked to protein-protein interactions (97). Assays of β-galactosidase activity were based on methods published previously (105). Specifically, quadruplicate cultures in LB with ampicillin, kanamycin, and IPTG were grown at 30°C with shaking overnight. Then 50 µl of overnight culture were added to 150 µl LB in a 96

well flat bottom plate (Greiner Bio-One) and the OD_{600} values were read. In addition, in a 1.7 ml centrifuge tube, 200 µl of overnight culture were added to 800 µl of Z buffer (16 g $Na_2HPO_4*12H_2O$, 6.25 g $NaH_2PO_4*H_2O$, 0.75 g KCl, 0.246 g $MgSO_4*7H_2O$, and 2.7 ml of β-mercaptoethanol added to 1 liter of water and pH adjusted to 7.0 (105)). 1 drop of 0.1% SDS and 2 drops of chloroform were added to the cells and samples were vortexed vigorously for 15 sec. The tubes were spun for 1 min. at 16,000 X g in a benchtop centrifuge to pellet cell debris and chloroform. 200 µl of the cell lysates were then loaded into wells of a 96 well flat bottom plate. 40 µl of 0.4% ONPG (***o***-nitrophenyl-β-d-galactopyranoside) in Z buffer were added to each sample and the OD_{420} values were read once a minute for 30 min in a plate reader that was maintained at 28°C (BioTek Synergy HT). Various studies have reported different units of activity for β-galactosidase activity assays in 96 well plates (106, 107). Units of activity were calculated as follows: Units = (1000*Slope of a line fit to OD_{420} in mOD/min) / (4*OD_{600} of 1:4 diluted overnight sample), based on previously described tests for the BACTH system (105).

2.3.3 PstB/PhoU co-elution experiments

Cell cultures were grown overnight in 250 ml MOPS LoP_i medium (except the mixed culture which combined the Δ*pstB* strain containing pRR48 in 125 ml MOPS HiP_i with the Δ*pstB* strain containing p48pstB-His in 125 ml MOPS LoP_i). Samples were collected by centrifugation at 5,000 X g and stored at -20°C. Cell pellets were thawed on ice and resuspended in 5 ml PstB lysis buffer (50 mM Tris-HCl pH 7.2, 300 mM NaCl, 20 mM Imidazole) with Protease Inhibitor Cocktail for use in purification of Histidine-tagged proteins (Sigma). Cells were lysed by one passage at 4000 psi and three additional passages at 18,000 psi using a Microfluidics LV1 cell disruptor. The crude lysate was then cleared by centrifugation in a

Sorval SLA-600 TC rotor at 10,000 X g at 4°C for 10 min. An aliquot of the cleared, crude lysate (0.5 ml) was collected and stored at -20°C for later analysis. The remaining cleared crude lysate was loaded onto a fresh 1 ml HisTrap FF nickel column (GE Healthcare). The HisTrap column was then loaded onto a GE Healthcare AKTA Prime Plus liquid chromatography system. The column was washed with 20 ml of PstB lysis buffer followed by elution with 5 ml of PstB elution buffer (Tris-HCl pH 7.2, 300 mM NaCl, 250 mM Imidazole). Wash and elution samples were collected in 1 ml fractions and stored at -20°C before analysis.

2.3.4 PhoR/PhoU co-elution experiments

SG1 cells (a Δ*phoBphoR*, Δ*pstSCABphoU* strain) were transformed with p48phoRNHis or pRR48 (as a negative control) and p116U2 (Table 2.1). 50 ml cultures were grown overnight in LB with ampicillin, kanamycin, and 100 μM IPTG. These cells were collected by centrifugation, resuspended in 5 ml PhoU Lysis Buffer (50 mM Tris, 0.5 M NaCl, 2 mM Imidazole, 14 mM β-Mercaptoethanol, pH 7.2) and 25 μl of Protease Inhibitor Cocktail for use in purification of Histidine-tagged proteins (Sigma). Then cells were lysed by 1 passage at 5,000 psi and 3 passages at 18,000 psi through a Microfluidics LV1 microfluidizer. Lysates were cleared by centrifugation and a lysate sample was taken.

Cleared lysates were mixed with 1 ml of Ni-NTA Agarose slurry (Qiagen) and shaken for 1 hour on ice. Samples were then loaded onto a 1 ml disposable polypropylene column (Qiagen); flow through was collected and reapplied to the column. 10 ml of PhoU Lysis buffer was applied for the first wash, followed by 6 ml PhoU Wash buffer (same as PhoU Lysis buffer, but with 35 mM Imidazole) for the second wash. Finally, 2 ml of PhoU Elution buffer (PhoU Lysis buffer with 0.5 M Imidazole) was applied to elute off the bound proteins. We added equal volumes of

Laemmli Sample buffer (Bio-Rad) to each sample, boiled for 5 min., and 5 µl of each was used for Western blot analysis (the lysate sample was diluted 1:20 to run on the gel).

2.3.5 Alkaline phosphatase (AP) and Western immunoblot assays.

The AP assays were carried out as described previously (108). The immunoblot assays were performed as described (109, 110) using a polyclonal rabbit α-PhoU antibody and a mouse α-Penta-His antibody (Qiagen). Immunoblots were visualized using the WesternBreeze chemiluminescent Western blot immunodetection kit (Invitrogen).

2.3.6 PhoU purification and gel filtration.

The DH5α *E. coli* strain harboring p116U2His was grown overnight in LB medium containing 40 µg/ml chloramphenicol and 100 µM sodium salicylate. Cells were harvested by centrifugation and resuspended in lysis buffer (20 mM $NaPO_4$, 0.5 M NaCl, 20 mM imidazole, 0.5 mM DTT, 10% glycerol, and 0.1 mM PMSF, pH 7.4) after which they were lysed by1 passage at 5,000 psi and 3 passages at 18,000 psi through a Microfluidics LV1 microfluidizer. The crude lysate was centrifuged at 12,000 X g for 15 min and the supernatant fraction was used for purification on an AKTA Prime-plus chromatography system (GE Healthcare) using a 1 ml HisTrap column as directed by the manufacturer.

Specifically, we washed with 20 ml of a mixture of 80% lysis buffer and 10% elution buffer (20 mM $NaPO_4$, 0.5 M NaCl, 0.5 M imidazole, 0.5 mM DTT, 10% glycerol). Then we eluted with 100% elution buffer collecting 1 ml fractions. All steps were carried out at 1 ml/min flow rate. It should be noted that the PhoU-His protein remained soluble only at concentrations below ~ 0.7 mg/ml, whereas mutant derivatives fell out of solution at concentrations above ~ 0.1

mg/ml. Gel filtration chromatography was also performed on this system using a HiPrep 16/60 Sephacryl S-200 HR column (GE Healthcare and a 0.5 ml sample loop run at 0.7 ml/min at 4°C. Standards of 450, 158, 45, 25, 12.5 kDa were run under the same conditions to create a calibration curve used to predict the size of soluble PhoU.

2.3.7 Inductively coupled plasma mass spectrometry (ICP-MS)

Samples of purified PhoU-His were collected from the HisTrap column in 1ml fractions. For a buffer only control, the same purification protocol was followed using these same buffers without loading any cell lysate onto the HisTrap column. PhoU-His and the corresponding buffer only 1ml fractions were collected during elution and then diluted to 12 ml of 5% nitric acid. These samples were incubated for 3 hours at 55°C, centrifuged 10 min at 2250 X g, filtered through a 0.2 µm Nalgene syringe filter (Thermo Scientific) and analyzed for metal content with ICP-MS on Elan6000 using the TotalQuant method. We scanned over the whole range using distilled deionized water as the blank and an optimization solution containing 10 ppb of various metals as the standard. Five replicates were performed of each and the metal content of the PhoU-His fraction was compared to the buffer only control.

2.3.8 Fluorescence assays for metal binding.

To confirm metal binding, we used PhoU-His F194W protein purified by metal affinity chromatography, dialyzed extensively in dialysis buffer (20 mM Sodium phosphate, 0.5 M NaCl, 0.1 mM DTT, 10% glycerol, pH7.4), and assayed in a FlouroMax-3 spectrofluorometer with excitation at 280 nm (the excitation wavelength for Tryptophan), at 21°C in a temperature-controlled continuously stirred chamber, and scanned for fluorescence from 300-400 nm.

For manganese interactions, we assayed purified protein (70 μg F194W PhoU and 26.4 μg E100A, R201A, F194W, E200A, R201A PhoU (QuadW)) in dialysis buffer. To the protein samples, we added 10 mM $MnCl_2$ stepwise to get 5 μM, 25 μM, 75 μM, and 150 μM final concentrations of manganese and scanned the sample for fluorescence. These scans were performed in triplicate. We averaged the fluorescence values for the 345-355 nm wavelengths, subtracted the corresponding buffer plus $MnCl_2$ control values, corrected for dilution, and then fit curves to the data to determine the K_d, and derived the peak fluorescence of each scan. For $MgCl_2$, we followed the same protocol as for $MnCl_2$ but added stocks of $MgCl_2$ stepwise to get 25 μM, 75 μM, 150 μM, 250 μM, 500 μM, 1.5 mM, and 2.75 mM final concentrations.

2.3.9 Integrated strain construction.

Temperature-sensitive plasmids encoding a 300-bp fragment of *pstB* upstream of *phoU* were constructed in order to introduce *phoU* (or its mutant derivative) back into the chromosome at its native location. To create the plasmid insert, the 3' end of the *pstB* gene extending 67 bp into the *phoU* gene was first amplified from chromosomal DNA obtained from BW25113 using primers PstBfor and PstBrev (See Table S1 for primer sequences). The *phoU* gene was then amplified from p116U2 or p116U2Quad using the PhoUfor and PhoUrev primers. These two PCR fragments were purified, diluted 1:50, combined and amplified again using the PstBfor and PhoUrev primers to generate a PCR fragment encoding the 3' end of the *pstB* gene upstream of *phoU*, exactly how it is found in the chromosome. This PCR product was subsequently digested with XbaI and HindIII (sites that were engineered into the primers) and ligated into the temperature-sensitive plasmid pIB307, which had been similarly digested. The resulting

plasmids were called pBU3 and pBQ3, and encoded *phoU* and *phoU* E100A R101A E200A R201A (*phoU*Quad), respectively.

The plasmids were then transformed into BM263, which contains a chromosomal deletion of *phoU* and subsequently grown at 30°C. To select for cells in which the whole plasmid integrated into the chromosome via RecA-mediated recombination into the *pstB* region, several colonies from each transformation were streaked onto fresh LB chloramphenicol plates and incubated at 42°C. To prevent excision of the integrated plasmid, isolated colonies were grown in LB medium at 37°C and then transduced with P1 phage carrying a marked *recA* allele from *E. coli* strain CSH126. The resulting strain with an integrated *phoU* gene was called BM265 and the strain with the integrated quadruple mutant was called BM266. Strains were confirmed by PCR analysis.

2.3.10 Cell fractionation

Membranes were prepared from the indicated strains by harvesting cells from 300 ml of liquid cultures grown to stationary phase in LB. Cells were harvested by centrifugation and resuspended in 14 ml of cold Buffer A (PBS, 1 mM DTT, 0.1 mM PMSF, pH 7.2). The cells were then lysed by three to four passages through a Microfluidics M-110L microfluidizer at 4°C. The crude lysates were centrifuged for 10 minutes at 13,000 X g to remove unbroken cells and aggregated proteins and 5 ml of the supernatant fractions were then subjected to ultracentrifugation at 100,000 X g for 1 hour at 4°C in a L100XP Beckman Coulter ultracentrifuge using a Type 90 Ti rotor. The soluble fractions were saved and the pelleted fractions were resuspended in 5 ml of 0.1% SDS. 50 µl of each fraction were mixed with 50 µl Laemmli Sample Buffer (Bio-Rad) prepared with β-mercaptoethanol and boiled for 10 min.

2.4 Results

2.4.1 Bacterial two-hybrid analysis shows interactions between P_i-signaling proteins

Since preliminary experiments in our laboratory to isolate a stable signaling complex from membranes using affinity chromatography with non-ionic detergents were unsuccessful, we employed alternate approaches to address this important question. If a signaling complex exists, then we should be able to detect physical interactions between multiple proteins. Since PhoU, PstB and PhoR are predicted to be found on the cytoplasmic surface of the cytoplasmic membrane, we probed for interactions between these three proteins. We first conducted bacterial two-hybrid experiments using the BACTH system. When the T18 and T25 domains of adenylate cyclase from *Bordetella pertussis* are in close proximity, they create an active enzyme that catalyzes the production of cAMP. By attaching different proteins to the T18 and T25 fragments, cAMP production is a measure of whether the fused proteins interact. Since cAMP binds to the CRP protein, a functional adenylate cyclase can be measured indirectly by assaying the expression of certain catabolic gene products, such as β-galactosidase. We prepared plasmids that express T18-PhoU, T25-PhoR and T25-PstB and introduced various combinations into an *E. coli* strain that was deleted for adenylate cyclase (BTH101). The strains were plated onto MOPS-maltose plates containing X-gal to provide a qualitative measure of β–galactosidase activity. For a more quantitative evaluation of protein interactions, we performed β–galactosidase assays in 96-well plates. We found that PhoU interacts with both PhoR and PstB, with the interaction between PhoU and PhoR being stronger than the PhoU/PstB interaction (Fig. 2.2A). We did not observe interactions between the T18-PhoU construct and the T25 domain

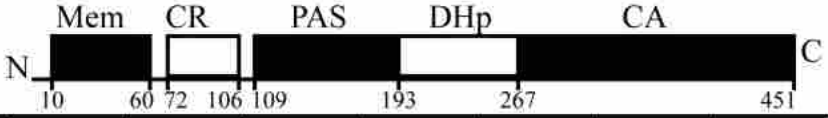
alone or the T25 domain fused to a leucine zipper domain. Interactions were also detected between T25-PhoU and T18-PhoR and T18-PstB (data not shown).

A

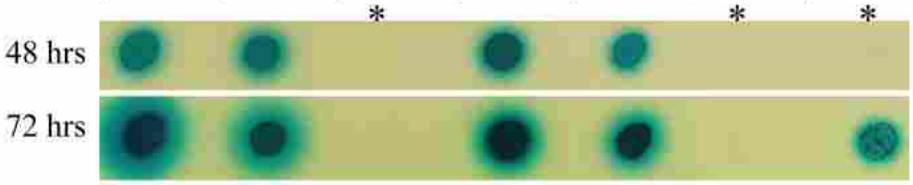
pUT18C	Zip	PhoU	PhoU	PhoU	PhoU
pKT25	Zip	PhoR	PstB	Zip	Empty
β-Gal Units	2712	1071	690	90.7	66.7
SD	1201	147	285	10.6	34.6

* = P value <0.01 for a two tailed unpaired t test compaired to PhoU/PhoR

B PhoR Domains



pUT18C	PhoU	PhoU	PhoU	PhoU	PhoU	PhoU	PhoU
pKT25	N-DHp	N-PAS	N-CR	CR-C	PAS-C	DHP-C	PAS
β-Gal Units	1296	1432	134	3051	2583	72.9	153
SD	360	514	5.4	285	661	34.6	48.5



C

pUT18C	E100A	E200A	ER200,201AA	Quad
pKT25	PhoR	PhoR	PhoR	PhoR
β-Gal Units	2072	1956	3607	3441
SD	499	422	296	598

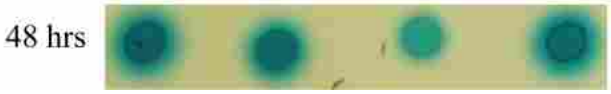


Figure 2.2 PhoU interacts with PhoR and PstB.
A BACTH system was used to identify PhoU interactions.Cultures were spotted onto MOPS-HiP_i-maltose plates containing X-Gal and incubated at 30°C for the indicated time. All samples shown were from the same plate. The average β-galactosidase activities (β-Gal Units) are reported with the standard deviation (SD)(n = 4). A. We used leucine zippers (Zip) fused to the adenylate cyclase domains as a positive control. We observed interaction between PhoU and both PhoR and PstB. As negative controls, we combined PhoU with the Zip plasmid as well as a plasmid that did not have any protein fused to the T25 domain. B. We engineered several PhoR truncations to identify where the interaction occurs. We see that any construct without the PAS domain has significantly less interaction than the full length PhoR (two-tailed t test gave P values of < 0.01). The PAS domain alone shows weak interactions with PhoU (as seen with color development after 72 hrs as well as significantly higher β-galactosidase units than either of the negative controls (two-tailed t test gave P values of < 0.05). C. We tested various PhoU mutant proteins and found that none of the mutations prevented the PhoU/PhoR interaction.

Several nested deletions of *phoR* were constructed to identify the region(s) of the protein that interacts with PhoU (see Fig 2.2B). The deletions were named PhoRN-'domain' to indicate a protein that extends from the amino terminus of PhoR through a particular domain. Constructs were called PhoR'domain'-C when deletions removed residues from the amino terminus of PhoR. The PhoRN-DHp and PhoRN-PAS constructs showed that the CA and DHp domains are not required for the PhoU interaction. Importantly, when the PAS domain was removed (PhoRN-CR), the interaction was lost. When PhoR was truncated from the amino terminus, we found that neither the membrane-spanning region nor the CR were required for interaction (PhoRCR-C and PhoRPAS-C), but loss of the PAS domain again disrupted the interaction (PhoRDHp-C). The PAS domain alone showed a weak interaction with PhoU. PhoU fused with T25 and PhoR truncations with T18 gave similar results (data not shown).

2.4.2 PstB and PhoR co-elute with PhoU during affinity chromatography

As a complementary method we employed protein co-elution experiments using His-tagged versions of either PstB or PhoR. The *pstB* gene was cloned into the pRR48 plasmid in such a manner that it expressed versions of PstB with and without a C-terminal 6XHis tag and both constructs were tested for P_i-signaling. Functional P_i-signaling is most readily apparent when cells are grown under P_i-replete conditions because cells produce very low amounts of alkaline phosphatase (AP). In contrast, elevated AP levels are observed when signaling is defective, such as when mutations occur in any of the *pstSCAB* genes, *phoU*, or even *phoR*. Wild-type cells also display high AP levels when cells are grown in low P_i media.

Under high P_i growth conditions, Δ*pstB* strains expressing either *pstB* or *pstB*-His exhibited very low AP levels, similar to the wild-type control strain, indicating that a functional

signaling pathway was reconstituted for both of these strains (see Fig. 2.3A). The empty vector control showed elevated AP levels, as expected for a strain that is deficient in P_i-signaling. However, in low P_i conditions, while the reconstituted BW26390 p48*pstB* strain demonstrated an expected increase in AP activity, the strain carrying the p48*pstB*-His plasmid still exhibited very low AP levels. This indicates that the *pstB*-His construct constitutively signals a high-P_i environment regardless of the P_i levels in the growth medium.

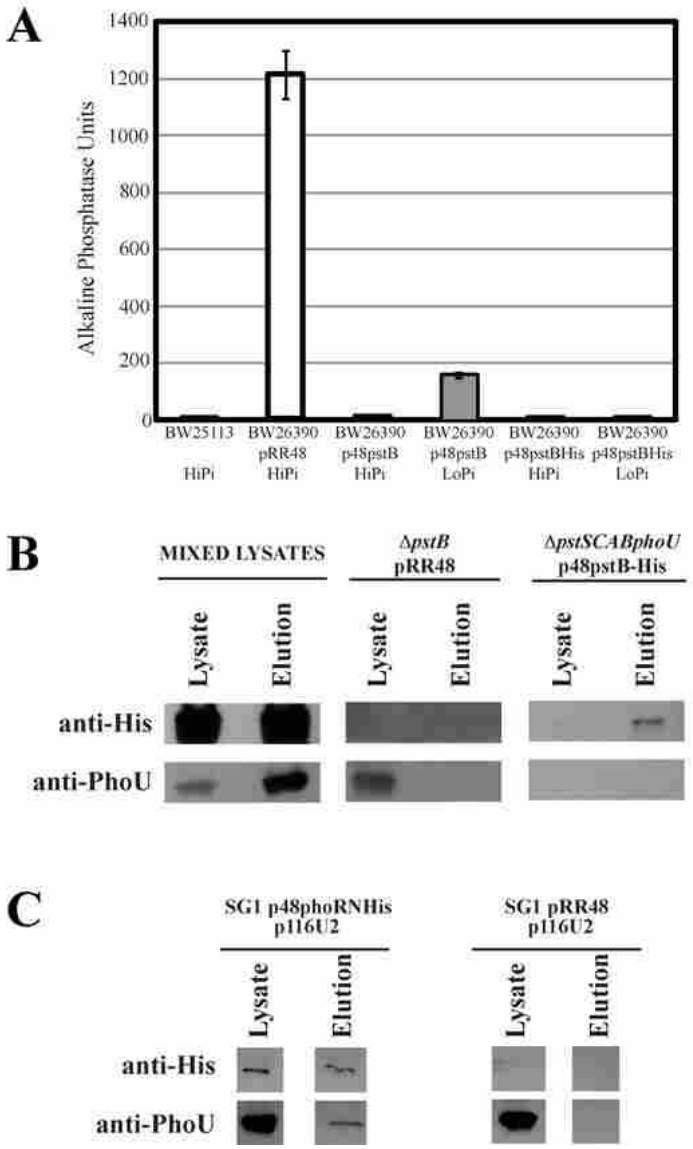


Figure 2.3 PstB and PhoR co-elute with PhoU.
A. Genetic complementation of a Δ*pstB* mutation with p48*pstB* and p48*pstB*-His. The Δ*pstB E. coli* strain BW26390 was transformed with the indicated plasmids and AP assays were performed. Cells were grown in MOPS HiP_i or MOPS LoP_i media, where indicated. B. Binding of PhoU to PstB-His. Cells were lysed and the soluble extracts were subjected to nickel affinity purification by passage over a HisTrap column. Identical samples from the lysate and elution fraction were separated by SDS-PAGE, transferred to nitrocellulose membranes and analyzed by immunodetection. The mixed lysate sample contained extracts from a strain expressing high levels of PstS, PstC, PstA, and PhoU with a sample expressing high levels of PstB-His. The use of the mixed lysate was necessary because the expression of PstB-His decreased the expression of the Pho regulon in the Δ*pstB* genetic background. C. Plasmids expressing PhoR-His (PhoR with an amino terminal 6X His tag) and PhoU were both transformed into a Δ*phoBR* Δ*pstSCAB-phoU* strain (SG1). Nickel affinity chromatography was used to retain the PhoR-His on the affinity column. Cell lysate and elusion fractions were analyzed by Western blotting to see that PhoU co-elutes with PhoR-His.

We were well aware of the possibility that potential interactions between PstB and PhoU could be transient and difficult to detect due to the large conformational changes that occur in the PstSCAB protein as a consequence of P_i transport. We also recognized that the PstB-His construct may have the fortuitous effect of trapping a complex between PstB and PhoU because it is locked into a "high-P_i" signaling conformation making the PstB/PhoU interaction easier to detect.

Once the signaling characteristics of the p48*pstB* and p48*pstB*-His plasmids were determined, we tested whether protein-protein interactions between PstB and PhoU occurred. This was performed by retaining PstB-His on a HisTrap nickel column along with any interacting proteins and then performing an immunoblot on the subsequent lysate and elution fractions assaying for the presence of PhoU with the eluted PstB-His.

Since we desired elevated levels of expression of the Pst transporter and the strains expressing PstB-His were repressed for the Pho regulon, we chose to mix extracts from two Δ*pstB* strains to detect protein-protein interactions. The first strain harbored the pRR48 plasmid and highly expressed the *pstSCA-phoU* operon (including *phoU*), but was missing PstB. The second Δ*pstB* strain harbored the p48*pstB*-His plasmid and expressed elevated levels of PstB-His, but did not produce much of the other Pst proteins or PhoU. Cells were grown in MOPS HiP_i media, pelleted, and then mixed together before cell disruption. The cleared, lysed sample was incubated for several minutes on ice, subjected to affinity chromatography and then processed for Western immunoblotting with anti-His antibody to detect the PstB-His protein and anti-PhoU antibody to detect PhoU. As shown in Fig. 2.3B, the mixed sample showed bands for PstB-His and PhoU in the lysate and elution fractions. These results demonstrate that PstB-His retained PhoU on the column through the washing procedure. As a control experiment, a Δ*pstB*

strain harboring the empty vector pRR48 was grown in MOPS LoP_i medium and subjected to the same protocol. As shown in Fig. 2.3A, this strain constitutively expresses high levels of AP, and as expected, also produced elevated amounts of PhoU (Fig. 2.3B, central panel). No PstB-His was observed in any fraction, but a strong PhoU band was observed only in the crude fraction, indicating that PhoU does not bind to the HisTrap nickel column in the absence of PstB-His. When a Δ*pstSCAB-phoU* strain containing p48*pstB*-His was analyzed it showed low level expression of PstB-His in the crude fraction but accumulation in the elution fraction. This strain also showed the absence of PhoU in every fraction, indicating that the antibody was specific for PhoU.

We confirmed the PhoU/PhoR interaction with co-elution experiments. We constructed a plasmid that expressed PhoR with an amino-terminal 6XHis tag (p48PhoRNHis) and confirmed its signaling activity by introducing it into the Δ*phoR* deletion strain JW0390-2 (111) and performing AP assays (data no shown). Because of difficulties with protein expression levels in this genetic background, we assayed interactions between PhoR and PhoU in SG1, a strain in which genes for all seven of the P_i-signaling proteins were deleted. Either pRR48 or p48PhoRNHis were introduced into the SG1 strain containing p116U2. The new strains were grown in LB medium, pelleted, lysed, and subjected to affinity chromatography. The crude and elution fractions were then probed by Western blot analysis for both PhoR and PhoU. Fig. 2.3C shows that PhoU was observed in the lysate fractions of both strains, but was only found in the elution fraction when PhoR-His was also present. The anti-His antibodies showed that PhoR-His was only produced when SG1 harbored p48PhoRNHis.

The results presented above show that PhoU interacts with both PhoR and PstB and are consistent with the presence of a P_i-signaling complex. To further understand the role of PhoU

in P_i-signaling we have sought to characterize its native structure in *E. coli*, to explore the functions of the most highly conserved residues in the PhoU family by mutational analysis, and to determine whether PhoU binds metal ions.

2.4.3 *E. coli* PhoU forms a dimer

We cloned *phoU,* to encode a native protein or a version with a C-terminal 6XHis-tag (PhoU-His) into pKG116 and confirmed that the constructs were functional by complementation analysis (data not shown). We then purified PhoU-His using immobilized metal ion affinity chromatography. To investigate its native structure, we analyzed this purified form of PhoU-His by gel filtration chromatography. Under nonreducing conditions, PhoU-His (a predicted molecular weight of 28,240 Da) eluted in several peaks corresponding to high molecular weight oligomers (data not shown). The PhoU oligomers were resolved when β-mercaptoethanol was included in the sample buffer. We think that these oligomers represent nonfunctional forms of PhoU because the side chains of each of its five Cys residues are predicted to extend towards the interior of the protein. Any covalent bonds involving these Cys residues would require PhoU to be partially or completely unfolded. When PhoU-His was analyzed with β-mercaptoethanol in the column buffer, it eluted as a single peak with an apparent molecular weight of 52.1 kDa which is comparable to the predicted molecular weight of a PhoU dimer (Fig. 2.4A).

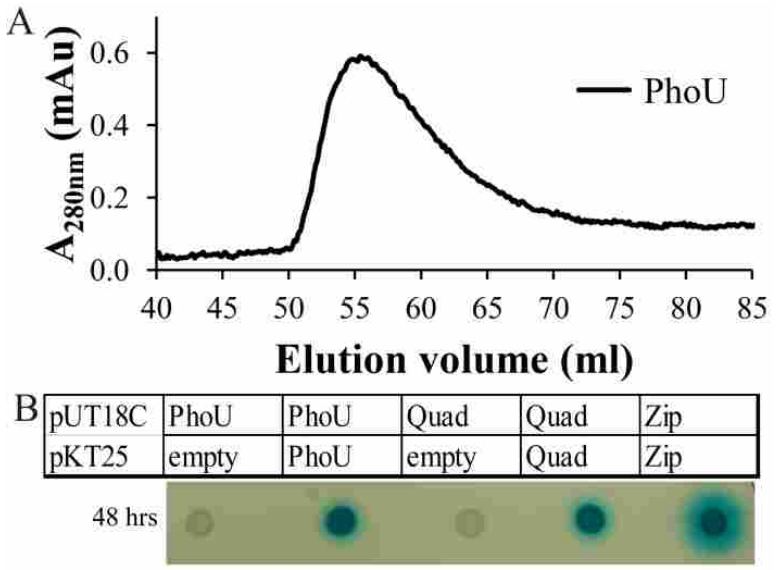


Figure 2.4 PhoU forms a dimer.
A. *E. coli* PhoU-His was purified and run on a gel filtration column. The PhoU-His protein eluted as a single peak between the 45 kDa and 68 kDa standards. PhoU-His has a predicted molecular weight of 28.2 kDa, so this peak corresponds with the expected size of a PhoU dimer. B. BACTH analysis Strains grown in LB medium with different combinations of plasmids were spotted onto MOPS-HiP_i-maltose plates with X-Gal and incubated at 30°C. These results show that PhoU interacts with itself as well as the PhoU E100A, R101A, E200A, R201A mutant (Quad). As a negative control, each construct is shown in combination with an empty vector. As a positive control, constructs with leucine zipper domains (Zip) fused to T18 and T25 were combined.

We used the BACTH system to confirm our gel filtration experiments. Using compatible plasmids that express T18-PhoU and T25-PhoU, we observed that PhoU interacts with itself (Fig. 2.4B), which is consistent with PhoU gel filtration results. We did not observe interactions when T18-PhoU was co-expressed with the T25 fragment alone or when T25-PhoU was co-expressed with the T18 fragment.

2.4.4 *E. coli* PhoU binds manganese and magnesium

X-ray crystal structures show that PhoU from *Thermotoga maritima* coordinates trinuclear and tetranuclear iron clusters at conserved sites (66), while PhoU from *Streptococcus pneumoniae* shows zinc ions bound to these same sites(95). It is not known whether these metals

are important for physiological function or if they are artifacts of crystallization conditions. We wanted to determine if *E. coli* PhoU binds metal, and if it does, which metal it binds. To screen for potential bound metals, the PhoU-His protein was purified from *E. coli* and analyzed by inductively coupled plasma mass spectrometry (ICP-MS). We identified manganese as a potential bound metal because significantly more manganese was found with the PhoU-His sample than with a buffer only control (t test with a P value of $2.4x10^{-7}$). Nickel and zinc were also identified as potentially binding to PhoU; however these metals may be an artifact from the 6XHis tag on PhoU and the nickel column protein purification. We did not detect any significant increase of iron in the sample containing PhoU compared with the control.

We used a fluorescence assay to further investigate interaction of PhoU with these metals. Tryptophan fluorescence intensity and maximum wavelength change when polarity of its local environment changes. Assays that follow changes in intrinsic tryptophan fluorescence are frequently used to detect changes in local protein environments (112). We engineered a PhoU-His F194W mutant and confirmed that the construct still functioned similar to the wild type protein in P_i-signaling (data not shown). This engineered tryptophan is located 6 residues upstream of the conserved putative metal binding residue, E200, and should be on the same face of the α-helix as one of the predicted metal binding sites (Fig. 2.1). We tested this purified protein with excitation at 280 nm and followed the fluorescence from 345 to 355 nm. When manganese was added to the sample, fluorescence was enhanced in a pattern that fit a predicted binding curve (Fig. 2.5A). Fitting a curve to the fluorescence data with various concentrations of manganese, we found that PhoU-His F194W bound manganese with an apparent K_d (dissociation constant) of 18.3 μM with a standard error of 6.2, based on three replicates (Fig. 2.5B). This value is similar to intracellular levels of manganese in wild-type *E. coli* reported between 15 μM

(113) and 21.1 μM (114), although it is proposed that the free manganese levels in *E. coli* are much lower than this (115). Magnesium and manganese often overlap in functions for cellular processes and magnesium is in much higher concentration in the cell than manganese (116). So, we also tested whether our purified PhoU-His F194W protein bound magnesium and found that it bound this metal with a K_d of ~1.5 ± 0.68 mM (Fig. S3). The high K_d for magnesium may explain why we did not observe this metal bound to PhoU in the ICP-MS experiment. Since it has been reported that the amount of free magnesium in *E. coli* cell is between 1 and 2 mM (117), binding this metal may be important for its physiological function. We saw a shift in the wavelength of peak fluorescence with metal binding. We did not see any significant fluorescence change with zinc or nickel (data not shown). Zinc and nickel are known to bind to 6XHis tags (118). Apparently, binding at the C-terminal 6XHis tag on PhoU is far enough away that it does not alter the F194W fluorescence. A change in fluorescence with magnesium and manganese implies that the metal binds at or near the conserved metal binding site or that metal binding significantly alters protein structure around the F194W residue.

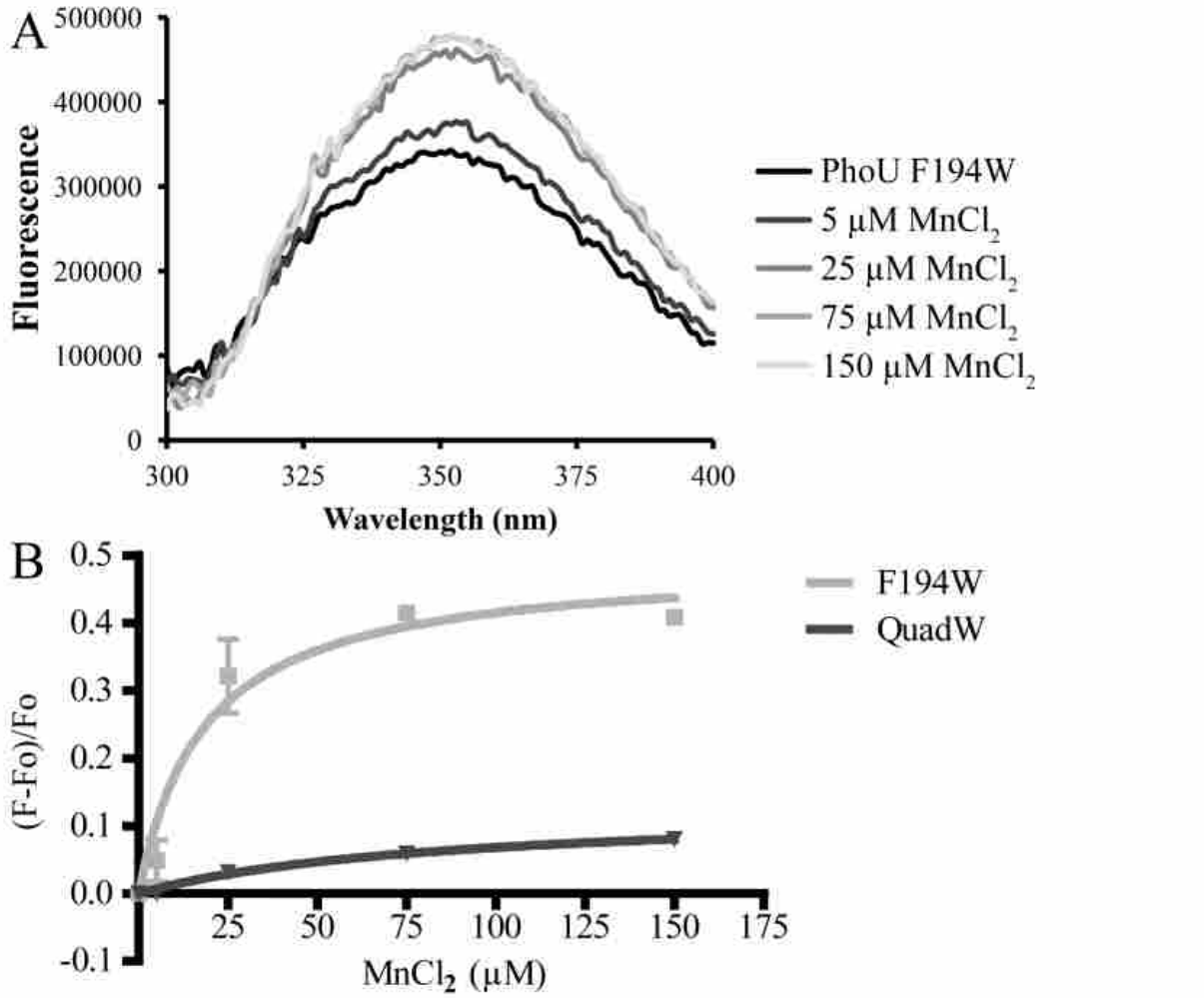


Figure 2.5 PhoU binds manganese.
A. A representative emission scan of tryptophan fluorescence when purified *E. coli* PhoU-His F194W (F194) is mixed with increasing levels of manganese and excited at 280 nm. The addition of manganese leads to an increase in the fluorescence and a shift in peak fluorescence. B. With the addition of manganese, the mean change in fluorescence between 345 nm to 355 nm was plotted and a binding curve was fit to the data (error bars represent ± standard error, n = 3). We see that PhoU-His F149W binds manganese to a greater degree than the PhoU-His E100A, R101A, F194W, E200A, R201A, F194W (QuadW) mutant protein.

2.4.5 Mutational analysis of conserved residues

To examine the role of PhoU metal binding on P_i-signaling, we mutated the *phoU* gene on the p116U2 plasmid to encode proteins with several of the highly conserved charged residues

converted to alanine residues. The residues that were initially altered were D58, E100, R101, E200, R201 and D204. These plasmids were introduced into a Δ*phoU* strain that replaced the *pstSCAB-phoU* promoter with a lactose inducible promoter (BM263) to prevent the poor growth phenotype seen in *phoU* mutants that overexpress a functional PstSCAB transporter (63). Strains were grown overnight in a high-P_i medium and AP assays were performed to assess signaling (Fig. 2.6A). Each of the mutants displayed nearly wild-type signaling activity, repressing the expression of alkaline phosphatase. The strongest phenotype was associated with the R201A mutation, which still left PhoU with nearly 80% of its wild-type activity. Immunoblot analysis showed that each of the mutant proteins was expressed at a similar level to each other (data not shown).

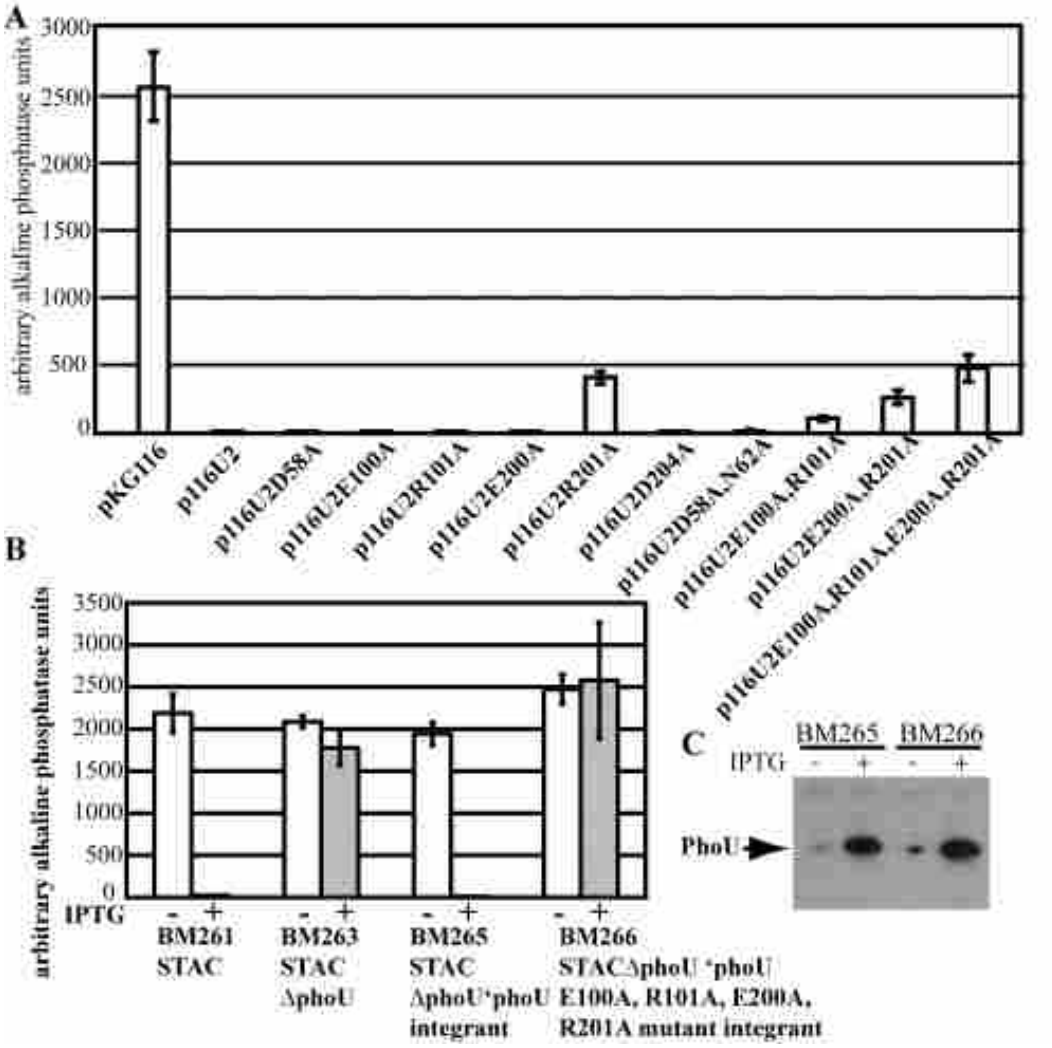


Figure 2.6 Signaling phenotypes of PhoU mutants.
A. Alkaline phosphatase expression was used as a reporter of P_i signaling. Plasmids with the mutant *phoU* constructs were introduced into a strain for which the native P_{pst} promoter was replaced by the P_{tac} promoter. For ease in referring to this strain, it is called STAC, for *PstS* operon, P_{tac} promoter. A Δ*phoU* STAC strain (BM263) was used to test PhoU mutants for signaling. Assays were performed in triplicate and the error bars represent ± the standard deviations of the measurements. B. Signaling phenotypes of the integrated *phoU* constructs. Cells were grown overnight on a roller drum at 37°C in MOPS HiP_i medium in the presence or absence of 50 μM IPTG and AP assays were performed in triplicate. The error bars represent ± the standard deviations. C. BM265 (*phoU*) and BM266 (*phoU*Quad) were grown in the presence or absence of 50 μM IPTG in MOPS HiP_i, then harvested and processed for visualizing protein expression by immunoblot analysis.

We created three mutants that carried two mutations in a single site: D58A/N62A, E100A/R101A, and E200A/R201A to test if it was necessary to alter more than a single residue within a putative metal binding pocket to block function.. Each of these double mutants also

retained significant PhoU signaling activity. Since these mutations only affected a single three-helix bundle, we wanted to know if there is a functional redundancy between the two PhoU domains. We created a PhoU quadruple mutation, E100A R101A E200A R201A (PhoUQuad), which had mutations in each three-helix bundle putative metal binding sites (E100A and E200A) as well as mutations of two other highly conserved residues (R101A and R201A). This version of PhoU also retained ~80% of its signaling activity.

2.4.6 Analysis of *phoU* integrants

Since the mutations were carried on a medium-copy-number plasmid, we postulated that the lack of a prominent phenotype was due to elevated expression levels rather than to those residues not being important for function. To address this possibility, we engineered two new strains, BM265 and BM266, which inserted the wild-type *phoU gene* or the *phoU*Quad variation into the chromosome of the *ΔphoU* strain BM263 at its normal location. AP assays were performed on strains that were grown overnight in MOPS HiP_i medium ± 50 μM IPTG. This amount of IPTG was found to be the minimal amount needed to achieve full repression of the Pho regulon in BM261 (a strain for which the native P_{pst} promoter was replaced by the P_{tac} promoter) (data not shown). Analysis of the parent strain BM261 showed that repression of the Pho regulon required expression of the *pstSCAB-phoU* operon as AP was derepressed in the absence of IPTG (Fig.2. 6B). Because of its inherent defect in signaling, the Pho regulon was induced both in the presence and absence of IPTG in the Δ*phoU* strain BM263. The strain containing the integrated *phoU* gene showed a response that was indistinguishable from the BM261 parent strain, indicating that the reconstruction of the *pstSCAB-phoU* operon was functional and was regulated in a normal pattern. However, the *phoU*Quad mutant showed a

pattern that was like the Δ*phoU* strain BM263, indicating that when expressed at lower levels from the chromosome, the PhoUQuad mutant was not functional in signaling a high P_i environment. We performed a Western blot to ensure that the signaling differences between the *phoU*$^+$ construct and the *phoU*Quad mutant were not due to differences in protein expression or protein stability (Fig. 2.6C). These results indicate that PhoU and PhoUQuad were equally expressed from their chromosomal locations and that at this expression level the mutant version of PhoU was nonfunctional whereas the wild-type version was fully functional. This supports the conclusion that elevated expression of *phoU* from a multicopy-number plasmid suppresses the phenotype of the site-directed *phoU* mutations described above.

What type of signaling function may be alleviated by over-expression? One possibility is that these conserved sites may be involved in targeting PhoU to its proper cellular location. When expressed from a plasmid, there may be sufficient PhoU throughout the cell so that targeted membrane localization is not required for signal transmission. However, proper cellular targeting would be essential for signaling when the *phoU* gene is present in single copy and expression levels are low. To test this membrane affinity hypothesis for these conserved residues, plasmids expressing several *phoU* mutants were introduced into the Δ*phoU* strain BM263 and membrane localization experiments were performed. The amount of PhoU in the membrane fraction progressively decreased as the number of mutations in PhoU increased from zero to four (Fig. 2.7).

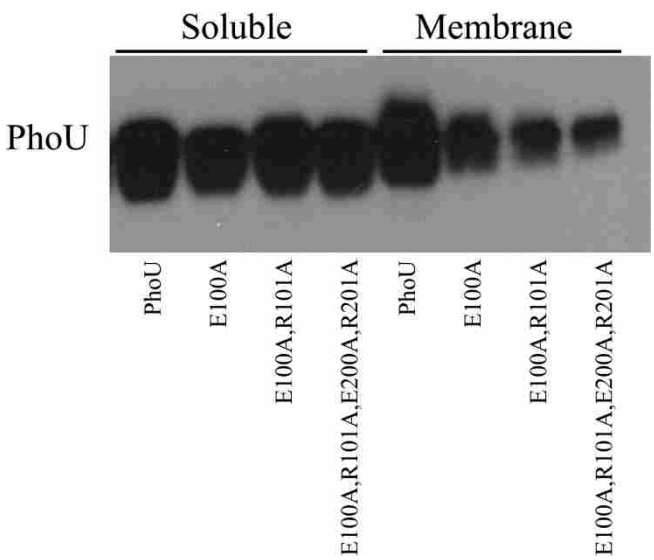


Figure 2.7 Membrane localization of PhoU and its mutant derivatives.
BM263 cells harboring p116U2 or several derived plasmids carrying mutations that alter conserved residues were grown at 37°C in LB medium. Cells were disrupted and the cell components were separated into soluble and membrane fractions by ultracentrifugation. A total of 1.5 μg of proteins from the soluble fractions and 5 μg from the membrane fractions were separated by SDS-PAGE, transferred to nitrocellulose, and probed with polyclonal rabbit antiserum to PhoU.

2.4.7 Biochemical analysis of the PhoUQuad protein

To further understand the role of these conserved residues, we analyzed the PhoUQuad mutant protein for dimer formation and metal binding. When examined by gel filtration chromatography, the PhoUQuad protein eluted from the column at the same volume as the PhoU-His protein (Fig. S1). We also tested the PhoUQuad mutant with the BACTH system and found that it also interacts with itself (see Fig. 2.4D). We also repeated the PhoU-His gel filtration experiments with buffers that contained either EDTA to chelate metal or excess manganese to saturate metal binding and did not see any changes in the elution profiles (Fig. S2). From these observations we conclude that PhoU forms a dimer and that the E100, R101, E200, and R201 residues are not essential for dimer formation.

We also tested the purified PhoU-His Quad F194W protein using the fluorescence assay to measure metal binding and found that it did not bind manganese as well as PhoU-His F194W (K_d of 85.2 μM ± 26.3 versus a K_d of 18.3 μM ± 6.2 with no overlap of 94% confidence intervals of the calculated K_d's based on triplicate tests). The fluorescence peak shift observed with the PhoU-His Quad F194W was significantly less than the shift caused by PhoU-His F194W (4.2 nm for F194W versus 1.7 nm for Quad F194W, P value of 0.020 from a two-tailed t test). We did not observe fluorescence enhancement when the PhoUQuadF194W protein was reacted with magnesium. Since the wild-type and mutant proteins both form dimers and the mutant protein does not appear to bind metals as well, we conclude that PhoU dimerization is independent of metal binding.

To further analyze the role of these residues in PhoU function, we employed the BACTH system to test whether the conserved charged residues that are involved in metal binding are important for the interactions between PhoU and PhoR. The E100A, E200A, E200A/R201A, and *phoUQuad* mutations were introduced into pUT18C*phoU* and interactions with PhoU were determined as described above. As shown in Fig. 2.2C, none of the mutations disrupted the interactions between PhoU and PhoR.

2.5 Discussion

This work provides evidence that the PhoU, PhoR, and PstB proteins physically interact. The demonstration of this interaction is important because it provides a framework for understanding P_i-signal transduction. Our current model of this signaling pathway is that PhoU is required for the formation of a signaling complex that is comprised of PstSCAB and PhoR

(See Fig. 2.8). We propose that PhoU contains multiple non-overlapping binding sites where it interacts with the membrane, PhoR and PstB.

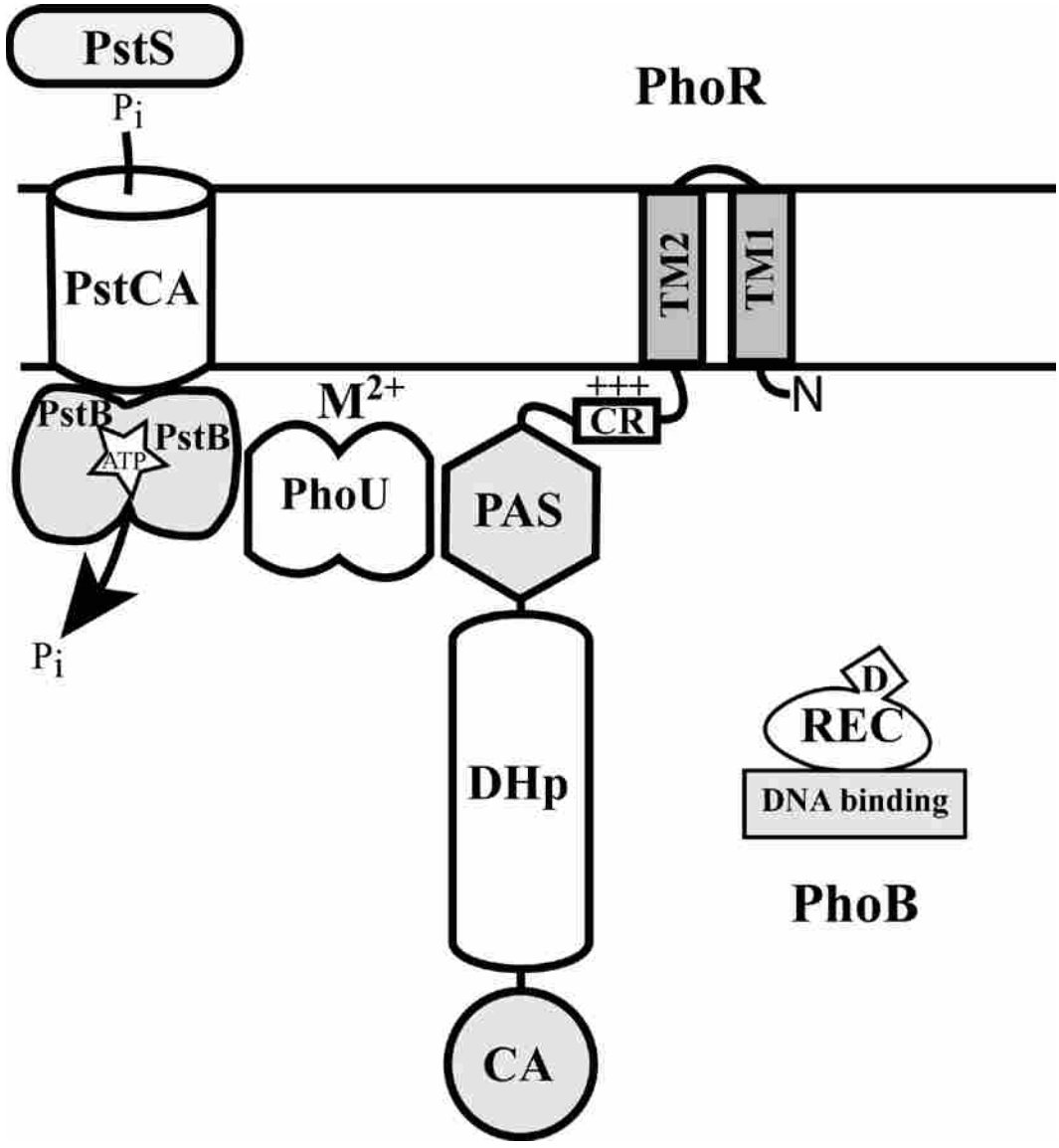


Figure 2.8 Model of Pho regulon expression control.
PhoR contains two transmembrane segments (TM1 and TM2), a charged region (CR), PAS, DHp, and CA domains. PhoR either phosphorylates or dephosphorylates the response regulator PhoB on a conserved aspartic acid residue (D) of its receiver domain (REC), Upon phosphorylation PhoB binds DNA and activates expression of Pho regulon genes. The PstSCAB transporter may signal P_i levels through the alternating conformations that are inherent to its transport process. Thus, PhoU may respond to mechanical forces in its interaction with the PstSCAB transporter and transmit that information to PhoR through its PAS domain. Proper signaling complex formation requires PhoU to localize to the cytoplasmic face of the inner membrane. Metal binding (M^{2+}) by PhoU is important for its interaction with the membrane, especially when expressed at low levels, as is the case when cells are grown under phosphate-replete conditions.

We found that PhoU's interaction with PhoR requires the PAS domain of PhoR (Fig. 2.2B). PAS domains often function in signaling through sensing physical or chemical stimuli (18). Structural studies of histidine kinases have shown that the correct spatial positioning of the DHp and CA domains is essential for their kinase and phosphatase activities (12, 34, 119, 120). In some cases, it is thought that physical interactions between the CA and DHp domains lead to phosphatase activity and that altering these interactions upon receiving a signal allows for autokinase activity (12). There have also been several studies of PAS domain structures in histidine kinases (13, 18, 119, 121). In one study, the authors proposed that the PAS domain interacts with the CA domain to inhibit the CA/DHp domain interaction that is required for kinase activity, and that signal binding to the PAS domain frees up the CA domain and allows for kinase activity (121). Disruption of the interactions between the PAS domain and other domains may lead to changes in activity (34). One example reported that the PAS domain is essential for kinase function (122) while another found that the PAS domain is essential for phosphatase function (19). Clearly, PAS domains play a role in regulating the activity of many histidine kinases. Our data supports a model in which the PAS domain of PhoR receives its input through direct interactions with PhoU. PAS domains frequently bind small molecules, our data do not address whether the PAS domain of PhoR binds a small molecule that allows it to interact with PhoU.

According to our signaling complex model, PhoR is able to sense through PhoU the conformational states of PstSCAB as a consequence of P_i transport and then modulate its kinase/phosphatase equilibrium towards the appropriate response. This model could accommodate variations in which the signaling complex is stable regardless of the activity of the

Pst transporter or the possibility that the complex is transient and is only formed under a subset of conditions. For example, in high P_i conditions, PstB may interact with PhoU and favor the binding of PhoR at the PAS domain. This may allow the PAS domain to interact with the CA domain and shift the activity toward phospho-PhoB phosphatase activity. Conversely, in low P_i conditions, PhoU may not have a stable interaction with the PAS domain of PhoR, which stabilizes the autokinase activity. This hypothesis is consistent with the observed unregulated kinase activity of PhoR in the absence of PhoU or a functional PstSCAB transporter (62).

In support of a signaling complex involving PhoU, we have previously shown that PhoU modulates P_i transport through the Pst transporter (63). It seems likely that such modulation would involve direct protein interactions between PhoU and PstSCAB. Moreover, Organesyan et al. reported that PhoU contains folds similar to Bag domains (a class of cofactors of the eukaryotic chaperone Hsp70 family) and proposed that PhoU may associate with the ATPase domain of PhoR (CA domain) and cause it to release PhoB, thus turning off the signaling cascade in a manner similar to Bag domain's association with Hsp70 (65). While supporting an interaction between PhoU and PhoR, our results show that unlike the Bag/Hsp70 paradigm, that PhoU interacts with the PAS domain of PhoR and not its ATP-binding CA domain. The dimeric nature of PhoU may also be important as it interacts with the dimeric PstB and PhoR proteins.

Our data suggest that the roles of the most highly conserved residues within the PhoU protein family are to bind metal ions, which function to target PhoU to the membrane. Membrane targeting may be especially important when PhoU is present at low concentrations, such as when P_i is abundant. There are several examples where metal binding by proteins allows them to localize to the membrane. For example, annexins play a crucial role in Ca^{2+} signaling by binding Ca^{2+} that forms a bridge between the protein and the phospholipid head groups (123).

Metal is bound by the protein through interactions with carbonyl and carboxyl groups of the protein and bound to the membrane through interaction with the phosphoryl moieties of the phospholipids (124). A similar mechanism is employed by the α-toxin protein from *Clostridium perfringens*, where binding of Ca^{2+} is linked to membrane binding and triggers the opening of the active site (125).

The highly conserved metal-binding residues of PhoU are important for signaling only when the *phoU* gene is expressed in single copy from the chromosome. They are not essential for signaling when expressed from a plasmid. It therefore seems unlikely that these residues are part of an enzymatic active site that produces a soluble message that is part of the signaling process.

By assaying for co-elution of PhoU following the retention of PstB-His on a HisTrap nickel column we were able to detect protein-protein interactions between these two proteins. We assume that this method was effective because the PstB-His protein was incorporated into a complete transporter consisting of PstC and PstA where the proper protein conformations required for physical interactions were maintained. Coomassie stained gels of the elution fractions showed many bands (not shown), including those of the predicted sizes for PstC and PstB, indicating that the affinity chromatography step enriched for the PstB-His protein, but did not produce a purified complex. It seems likely that the complexity of the eluate was due to the nature of the sample applied to the column, which consisted of soluble proteins as well as proteins imbedded in membrane vesicles.

A few studies have investigated PhoU's interaction with other proteins. One group used fluorescence resonance energy transfer (FRET) analysis to determine protein-protein interactions of various proteins involved in the P_i starvation response and failed to find an interaction between PhoU and either PhoR or PhoB (70). Another study used a Yeast two-hybrid assay to

show that in *Edwardsiella tarda* PhoU interacts with PhoB and Fur, but they did not see any interaction with PhoR (64). However, the conflicting results between these studies imply that more evidence is necessary to fully understand all of the proteins that interact with PhoU. It is possible that fusing PhoU to enhanced cyan fluorescent protein in the FRET analysis and using a truncated PhoR expressed in yeast cells prevented the PhoU/PhoR interaction or the detection methods were not sensitive enough to identify the PhoU/PhoR interaction.

We used a BACTH system to identify and characterize PhoU/PhoR interactions. Others have had success using these systems to identify domains from bacterial proteins that interact and have found the BACTH system is especially useful for testing membrane bound proteins (105, 126). Our results confirm that PhoU does interact with PhoR. Given the multiple PhoR and PhoU constructs that show interaction (Fig. 2.2) it is unlikely that all of the different constructs are false positives. These results are also confirmed with PhoU co-eluting with PhoR-His (Fig. 2.3).

Chapter 3 Genetic analysis, structural modeling, and direct coupling analysis suggest a mechanism for phosphate signaling in *Escherichia coli*

3.1 Abstract

Background: Proper phosphate signaling is essential for robust growth of *Escherichia coli* and many other bacteria. The phosphate signal is mediated by a classic two component signal system composed of PhoR and PhoB. The PhoR histidine kinase is responsible for phosphorylating/dephosphorylating the response regulator, PhoB, which controls the expression of genes that aid growth in low phosphate conditions. The mechanism by which PhoR receives a signal of environmental phosphate levels has remained elusive. A transporter complex composed of the PstS, PstC, PstA, and PstB proteins as well as a negative regulator, PhoU, have been implicated in signaling environmental phosphate to PhoR.

Results: This work confirms that PhoU and the PstSCAB complex are necessary for proper signaling of high environmental phosphate. Also, we identify residues important in PhoU/PhoR interaction with genetic analysis. Using protein modeling and docking methods, we show an interaction model that points to a potential mechanism for PhoU mediated signaling to PhoR to modify its activity. This model is tested with direct coupling analysis.

Conclusions: These bioinformatics tools, in combination with genetic and biochemical analysis, help to identify and test a model for phosphate signaling and may be applicable to several other systems.

3.2 Background

Adapting to changes in the environment is one of the hallmarks of life. For all life, phosphate is an essential nutrient. Bacteria have several mechanisms to scavenge phosphate that are only expressed when the level of available environmental phosphate is limited: including a phosphate specific ABC transporter complex (PstSCAB) and a periplasmic phosphate scavenging enzyme (alkaline phosphatase (AP)); the product of the *phoA* gene) (11). Expression control of these genes is essential for optimal growth and has been implicated in the regulation of pathogenesis in several organisms (6, 9).

In *Escherichia coli*, a classic two-component signal transduction system, composed of the PhoR histidine kinase and the PhoB response regulator, is responsible for expression control of a group of genes called the Pho regulon. PhoR consists of an N-terminal membrane-spanning region, as well as cytoplasmic PAS, DHp and CA domains. The PAS domain was named for the *Drosophila* Per, Arnt, and Sim proteins, in which this domain was originally described and has been found in many signaling proteins. Many PAS domains bind cofactors such as heme (33). The DHp domain is conserved in histidine kinases and functions in dimerization and contains the site of histidine phosphorylation. The CA domain is the catalytic, ATP-binding part of the protein. PhoB consists of an N-terminal phosphorylation domain that receives a phosphoryl group from PhoR and a C-terminal DNA binding domain. In low phosphate conditions, PhoR acts as a PhoB kinase. Upon phosphorylation, PhoB recruits RNA polymerase to promoters of the Pho regulon that contain a Pho box. In high phosphate conditions, PhoR acts as a phospho-PhoB phosphatase and removes the phosphate from PhoB to keep the expression level of Pho regulon genes very low. One unanswered question with this system is how PhoR perceives external phosphate concentrations. PhoR lacks a significant periplasmic domain that could

detect phosphate abundance outside the cell. Past work has shown that the PstSCAB transporter and the PhoU protein play important roles in phosphate signaling to PhoR (11). The mechanism of this signal has not been fully elucidated.

A deletion mutation of the *phoU* gene leads to poor growth and the frequent development of compensatory mutations in the other Pho regulon expression control genes, *pstSCAB*, *phoR*, and *phoB* (62). The poor growth phenotype is likely due to overexpression and under-regulation of a functional PstSCAB transporter (61), which leads to phosphate poisoning when cells are grown in high phosphate environments (63). Reference (63) proposed that PhoU modulates phosphate transport through the PstSCAB complex by inhibiting transport when internal phosphate levels are too high.

Recently, PhoU was shown to directly interact with PstB and PhoR (1). This observation suggested a model that PhoU interacts with PstB to sense environmental phosphate levels and that it passes that signal along to PhoR to modulate its kinase/phosphatase activities (Fig 3.1). To further characterize these interactions, this work analyzes several *phoU* mutants for signaling activity, interactions with PhoR, and interactions with PstB. A scanning mutagenesis screen of the PAS domain of PhoR identified potential residues important for interaction with PhoU. We modeled potential docking interactions between PhoR and PhoU. One of the docking models showed a close proximity of identified residues in PhoU with a predicted interaction loop of the PAS domain of PhoR. To validate this model, we performed a Direct Coupling Analysis (DCA) between PhoR and PhoU using sequences from the gammaproteobacteria. The DCA results are consistent with the proposed docking model and may point to a mechanism of action for PhoU in controlling the opposing kinase and phosphatase activities of PhoR.

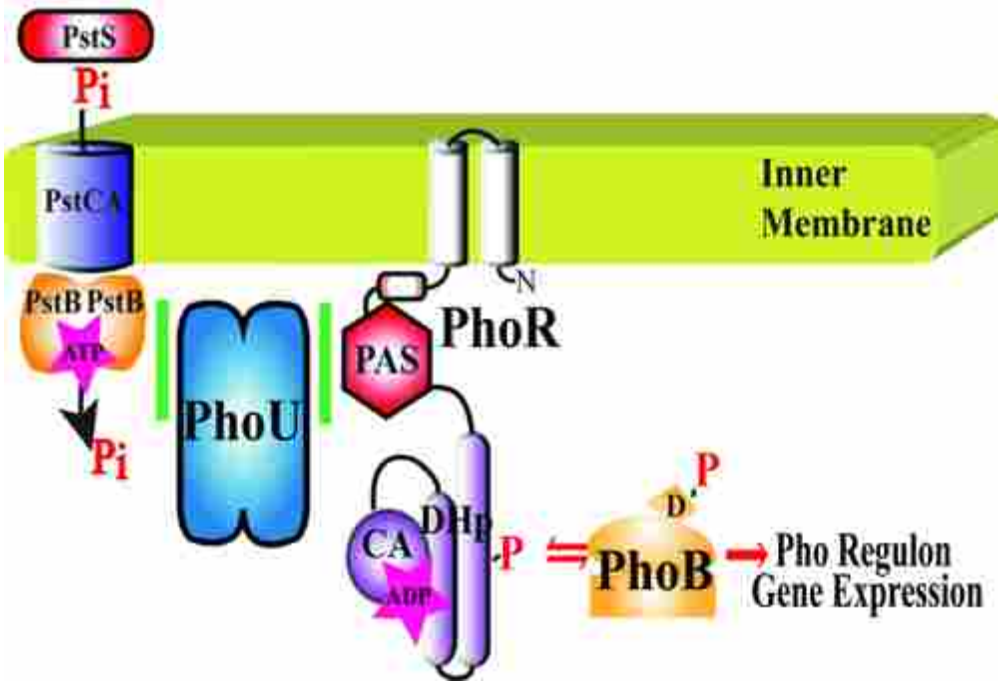


Figure 3.1 Pho regulon expression control.
Proper control of Pho regulon expression involves the histidine kinase, PhoR that has a Per-Arnt-Sim (PAS), dimerization/histidine phosphorylation (DHp), and catalytic (CA) domains; the response regulator PhoB (with a conserved aspartate (D) residue that is phosphorylated); the phosphate specific ABC transporter composed of PstS (Periplasmic phosphate binding protein), PstC and PstA (proteins that form the pore in the inner membrane), and PstB (the ATP binding portion of the transporter); and PhoU (a negative regulator of PhoR).

3.3 Results and Discussion

3.3.1 PhoU and PstSCAB necessity for phosphate signaling

Mutations in any of the *pstSCAB* transporter genes or *phoU* lead to loss of phosphate signaling. To determine whether PhoU could act independently of the PstSCAB transporter, we expressed these two genetic entities from separate plasmids in *E. coli* strain BW26337, which contains a deletion of the *pstSCABphoU* operon and tested strains for control of the Pho regulon. We cloned the *pstSCAB* genes into the pRR48 plasmid (Amp^R) (104) and the *phoU* gene into the compatible pKG116 plasmid (Cam^R) (103). Control of the Pho regulon was analyzed by

determining Alkaline Phosphatase (AP) expression. Fig 3.2 shows that AP expression was unregulated when cells were grown under high phosphate conditions in strains expressing neither protein (containing two empty vectors), or expressing each protein individually. However, when both proteins were expressed in the same cell, AP expression levels were significantly reduced. These results show that *phoU* expression is necessary but not sufficient for proper phosphate signaling and are consistent with a model that these proteins act together in signal transduction.

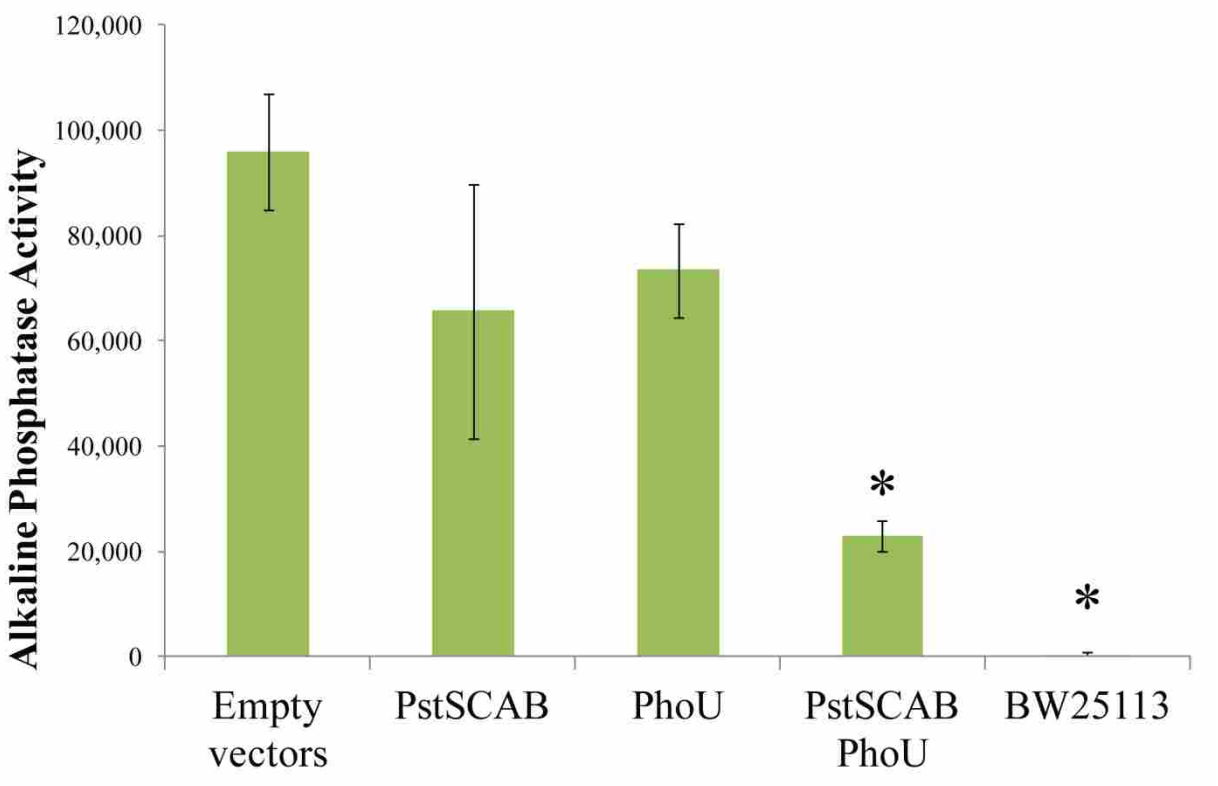


Figure 3.2 Signaling necessity and sufficiency of PhoU and PstSCAB.
Triplicate cultures of BW26337 (a Δ*pstSCABphoU* strain) cells with pKG116 and pRR48 (Empty vectors), pKG116 and pRR48SCAB (PstSCAB), p116U2 and pRR48 (PhoU), p116U2 and pRR48SCAB (PstSCAB PhoU) and BW25113 (a wild-type strain) were grown in LB and Bacterial Alkaline Phosphatase activity was assayed. Error bars represent the standard deviation. * = p-value of < 0.05 compared to the Empty vectors with a two-tailed T-test.

Early studies of phosphate homeostasis in *E. coli* isolated mutants that constitutively expressed AP (73). One mutant, named C4 (127) was characterized (128) and was later named *phoU35* because it was independent of the phosphate transport genes (129). When the genes of the *pstSCABphoU* operon were sequenced and the *phoU35* mutant was analyzed , they found this mutation encoded a change from alanine at position 147 to glutamic acid (A147E) (71). Since a *phoU* deletion mutation results in loss of phosphate signaling and causes a severe growth phenotype, but the *phoU*35 allele was only reported to cause a loss of signaling without the accompanying poor growth, we hypothesized that the *phoU*35 mutation may disrupt PhoU's interaction with PhoR, preventing the signal for the switch to PhoR phosphatase activity, but may maintain the interaction with PstB, limiting excess transport of phosphate into the cell during phosphate replete conditions.

3.3.2 PhoU and PhoU A1417E growth phenotypes in a Δ*phoU* strain

We wanted to confirm that strains expressing the *phoU*35 allele do not show a severe growth phenotype. We employed a *phoU* deletion strain in which the normal *pstS* promoter was replaced by the *tac* promoter to uncouple expression of this operon from mutations in any of its genes (STACΔphoU) (1). By growing this strain in the absence of IPTG the accumulation of compensatory mutations was avoided. Fig 3.3 shows the results of characterizing the growth yield of various cultures grown in a high-phosphate medium with IPTG overnight with shaking. The STAC strain with a wildtype copy of *phoU* still in the chromosome grew normally. We observed a significant growth defect when the *phoU* knockout strain containing an empty vector, pKG116, was grown in identical conditions. When the p116U2 plasmid was introduced into this strain, expressing wild-type *phoU*, and grown under identical conditions, no growth defect was

observed. Moreover, when the *phoUA147E* allele was introduced into the strain, the growth phenotype was similar to that of wild-type *phoU*. These results confirm the lack of growth defect in the *phoU35* mutant and are consistent with the hypothesis that the A147 residue of PhoU is important for interactions with PhoR.

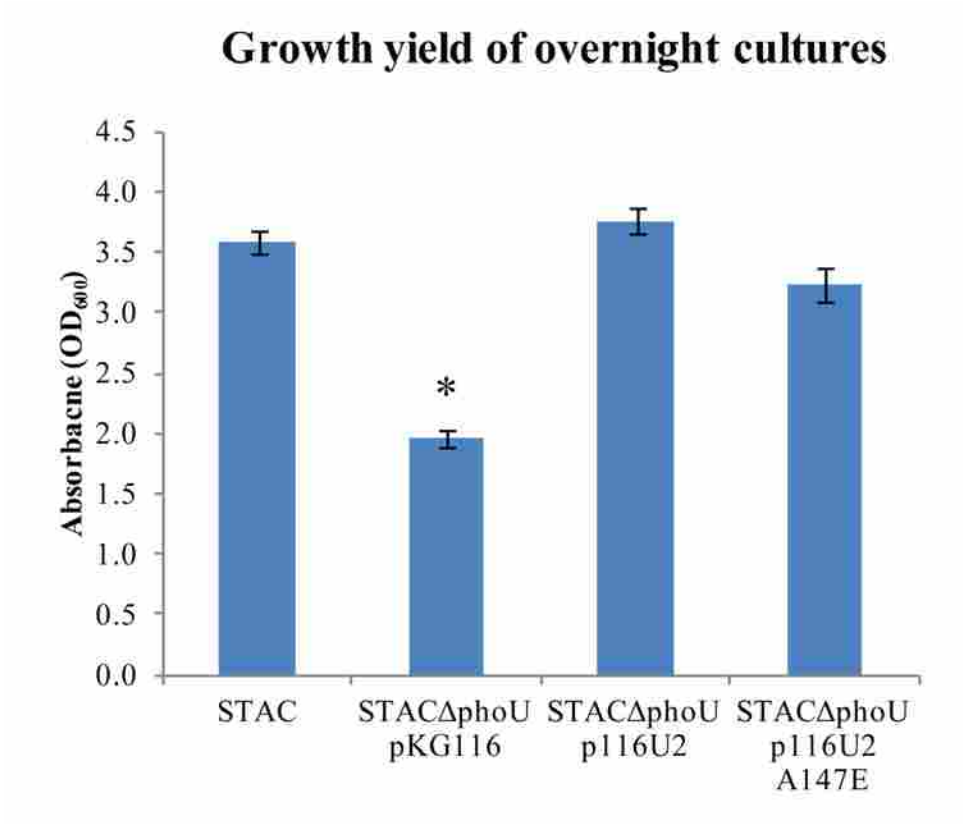


Figure 3.3 Growth yield of overnight cultures.
Triplicate cultures were grown overnight in LB and the absorbance of 600nm wavelength of light for each culture was recorded. Error bars represent standard deviations. * = p-value of < 0.01 compared to STAC with a two-tailed T-test.

3.3.3 PhoU mutant's signaling phenotypes

To verify that the *phoU*35 allele caused a signaling defect when expressed from a multicopy plasmid, we tested AP expression in the strains constructed for the previous experiments. Fig 3.4 shows that high AP levels were observed in the STACΔphoU pKG116 strain and the STACΔphoU p116U2 A47E strain grown under high phosphate conditions, but

that reduced AP levels were observed in the STACΔphoU p116U2 strain. This showed that PhoUA147E could not regulate PhoR and was consistent with the model presented. We wondered if other mutations that caused changes in the *phoU* protein in the vicinity of A147 would also display a signaling defect. We therefore introduced additional mutations by site-directed mutagenesis in the region near A147 (A147, R148, and D150) to create versions of PhoU with different charges and sizes of amino acid side chains in this region. The mutations that we created were A147K, R148E, R148A, and D150A. The rightmost four bars in Fig 3.4 show that the A147K, R148E and R148A mutants lost phosphate signaling activity as they expressed elevated AP levels, but that the D150A mutation retained signaling activity. These results suggest that both A147 and R148 are important for interactions with PhoR.

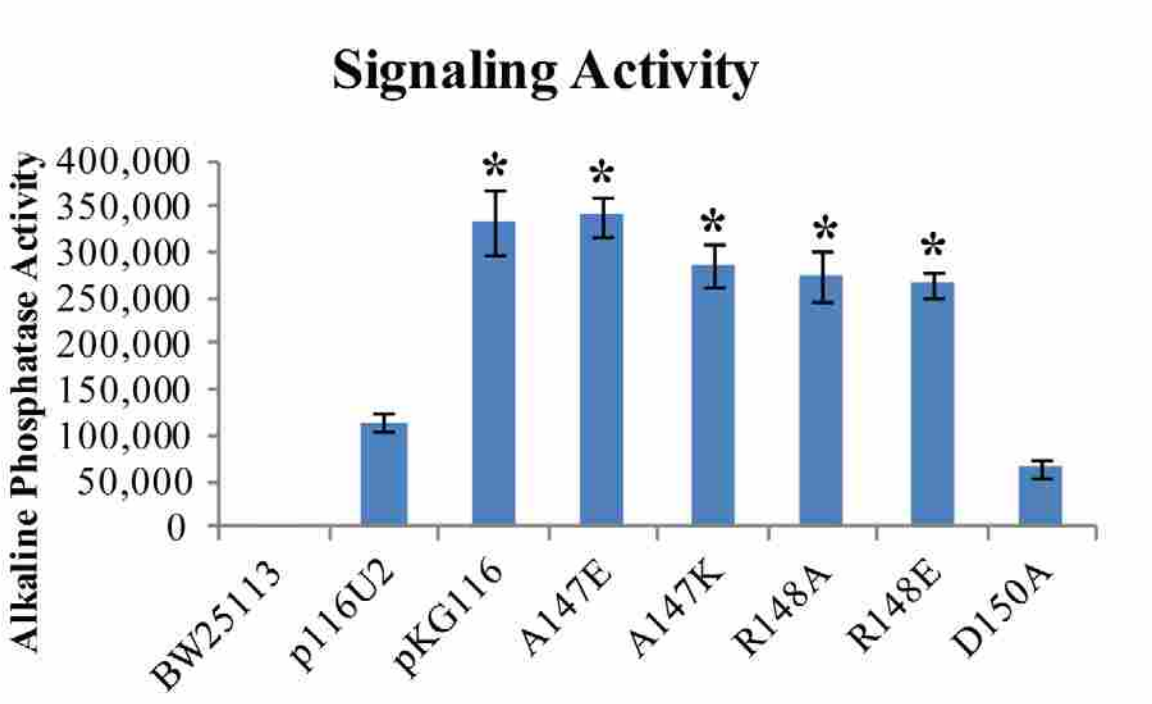


Figure 3.4 Signaling activity of various PhoU mutants.
Triplicate cultures were grown in LB and assayed for Bacterial Alkaline Phosphatase activity. BW25113 is a wildtype control. Other strains were STACΔphoU with p116U2 (wildtype *phoU*), pKG116 (empty parent plasmid), and various p116U2 plasmids with specific mutations in *phoU* (A147E, A147K, R148A, R148E, and D150A). Error bars represent standard deviations. Values significantly greater than p116U2 are labeled with * = p-value of < 0.001 compared to p116U2 with a two-tailed T-test.

3.3.4 BACTH of PhoU mutants interaction with PhoR and PstB

The combination of no growth inhibition and disrupted signaling are the expected phenotypes of a *phoU* mutant that can interact with PstB to limit phosphate transport but can no longer interact with PhoR to signal the phosphate level. We further examined this hypothesis by using a bacterial adenylate cyclase two-hybrid (BACTH) system to test for interactions between mutant versions of PhoU and PstB and PhoR (Fig 3.5A and 3.5B).

The BACTH system employs the separable T18 and T25 domains of adenylate cyclase from *Bordetella pertussis*. When these two domains are in close proximity they create an active enzyme that produces cAMP. By creating gene fusions in which protein domains are connected to the T18 and T25 fragments, cAMP production is a measure of whether the fused proteins interact. Since cAMP binds to the CRP protein, cAMP can be measured indirectly by assaying β-galactosidase. We previously used this method to show interactions between the wild-type version of PhoU and PstB and various domains of PhoR.

Fig 3.5A shows the results of β-galactosidase activity assays of various BACTH strains. Each sample is BTH101 strain containing one plasmid that expresses the T18 domain fused to PhoR and another plasmid that expresses either the T25 domain alone (Empty vector), T25 fuzed to PhoU (PhoU/PhoR), or various PhoU mutant proteins (A147E, A147K, A148A, R148E, and D150A). The A147 and R148 mutants of *phoU* had a significantly weaker interaction with PhoR (as represented by low β-galactosidase activities) and the D150A mutant retained PhoR interaction. We also tested these plasmids expressing the T25-PhoU fusions for interaction with a T18-PstB fusion protein. Interestingly, the A147 and R148 mutant proteins maintained

interactions with PstB (Fig 3.5B). This implies that the loss of interaction with PhoR is not due to protein instability or radical protein misfolding.

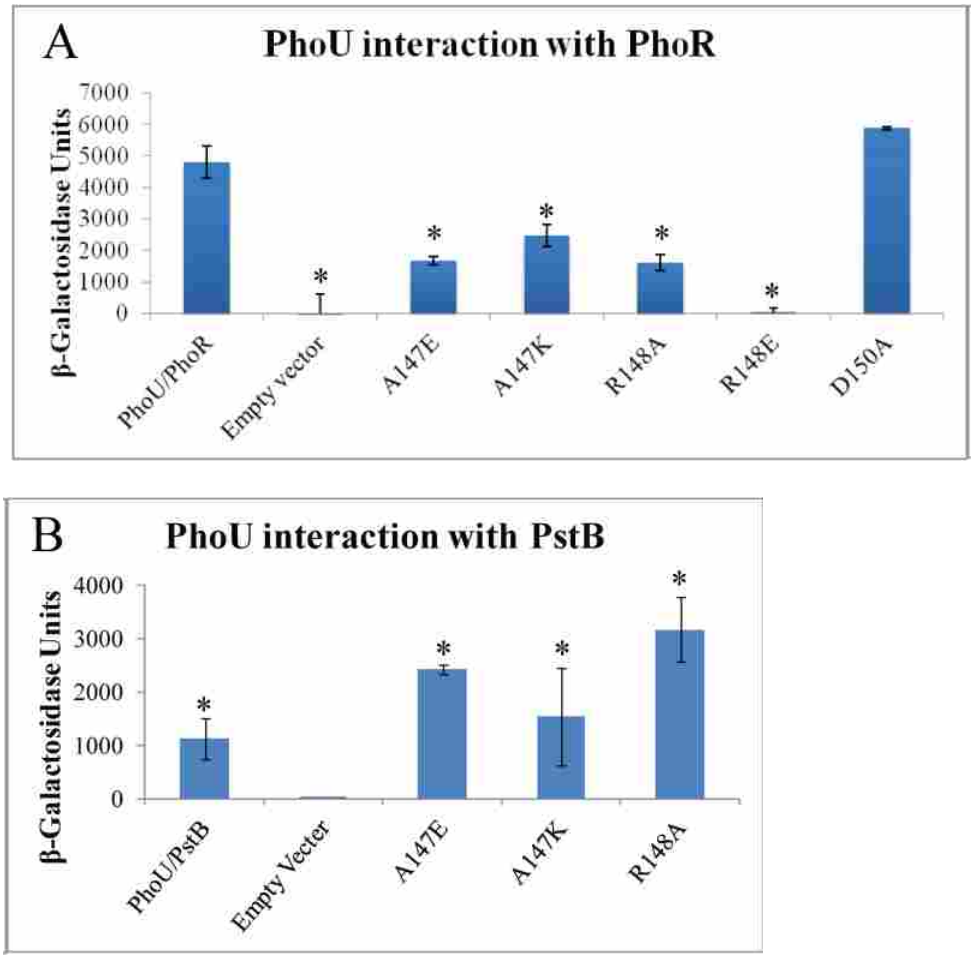


Figure 3.5 PhoU mutant interaction s with PhoR and PstB.
Values shown are the average of triplicate samples with the error bars representing standard deviation. A. BACTH analysis of PhoU/PhoR interaction. Various combinations of PhoU and mutants or an empty vector negative control were tested for interaction with PhoR. Activity of β-Galactosidase correlates with protein/protein interaction. * = p-value < 0.01 compared to PhoU/PhoR with a two-tailed T-test. B. This chart shows the interaction of PstB with various PhoU mutants and an empty vector negative control. * = $p < 0.05$ compared to Empty Vector with a two-tailed T-test.

3.3.5 Scanning mutagenesis of the PAS domain of PhoR

Previous work showed that the PhoR/PhoU interaction was dependent on the PAS domain of PhoR (1). We used scanning mutagenesis of a plasmid expressing a T18-PhoR N-PAS (the portion of the PhoR protein from the C terminus through the PAS domain, but without the

CA and DHp domains) domain fusion protein to identify the residues within this domain that are important for the interaction. Every two amino acids of the PhoR PAS domain were mutated to alanine and cysteine residues by sequentially replacing the six bases encoding adjacent codons with a *SphI* restriction site. The mutant versions were then introduced into the appropriate tester strain expressing wild-type *phoU* fused to the T25 fragment of adenylate cyclase and β-galactosidase assays were performed. Many mutants lost the ability to interact with PhoU as indicated by significantly reduced β-galactosidase levels (shown as blue bars in Fig. 3.6B). Among the mutations there were several regions where neighboring residues showed loss of interaction (for example, residues 141-146, 157-162 and 169-176). There were also several single mutations (111, 115, 189) that reduced protein interactions.

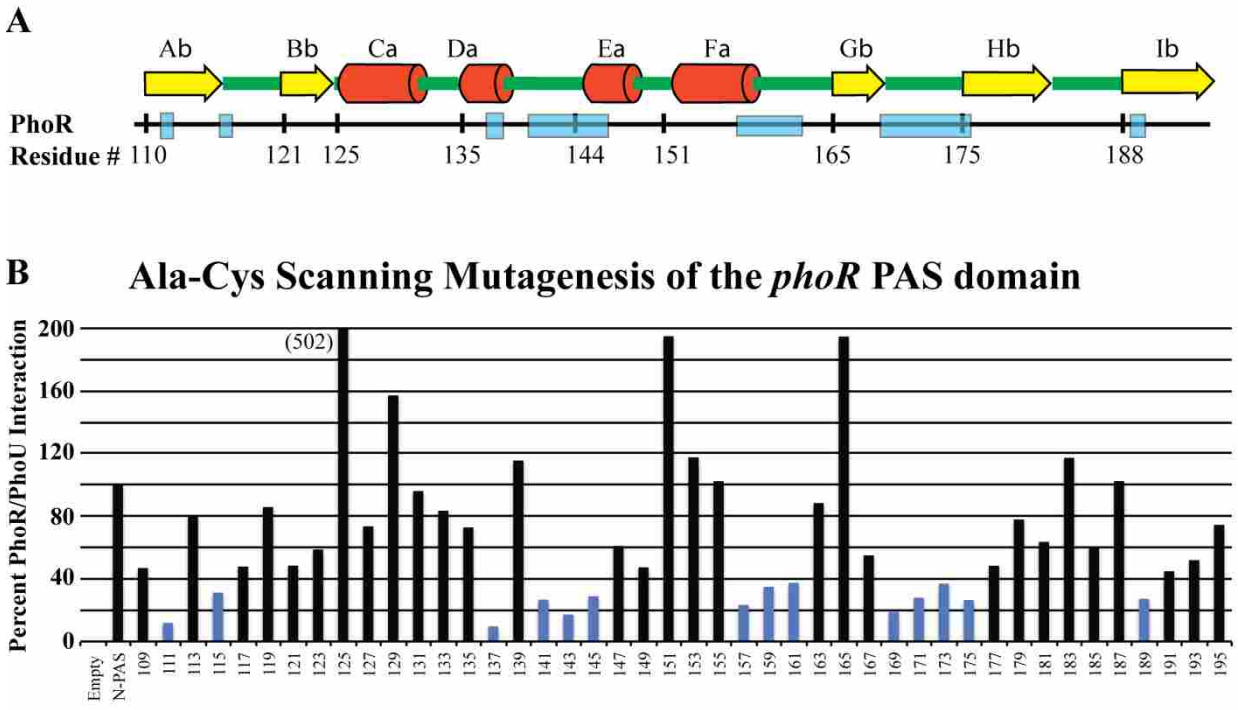


Figure 3.6 PAS domain scanning mutagenesis.
A. Predicted secondary structure of the PhoR PAS domain with alpha helixes labeled with yellow arrows and β-sheets labeled with red cylinders. B. Scanning mutagenesis of the PhoR PAS domain used BACTH to identify regions of the protein that are essential for interaction with PhoU. Every two amino acids were changed to code for alanine and cysteine. Each construct was tested in triplicate for interaction with PhoU. Blue bars represent samples that had less than 40% of the activity of unmutated PhoR.

3.3.6 Protein structure models and a docking model of PhoR/PhoU interaction

We modeled the cytoplasmic portion of the *E. coli* PhoR structure based upon the structure of the VicK protein from *Streptococcus mutans* using the Phyre2 webserver (Fig 3.7A). This modeled structure predicts that the 141-146 and the 157-162 regions of PhoR have residues that are exposed to the surface and face outwards, consistent with an interaction with another protein (Fig 3.7A, the residues colored red and purple respectively.) The 111, 115, 189, and the 165-176 mutations all map to regions of PhoR that are predicted to lie along a β-sheet that has been shown in other histidine kinases to be important for PAS domain dimerization contacts. For this reason, these mutations may not be involved in PhoU/PhoR interactions, but for maintaining a proper protein conformation.

We docked a modeled *E. coli* PhoU protein onto the modeled PhoR dimer. The server returned ten potential PhoR/PhoU models with most of the models being unreasonable because they docked PhoU onto the part of PhoR that interacts with the cytoplasmic membrane. One of the reasonable models showed an interesting interaction between PhoR and PhoU (Fig 3.7B). In this model the PhoU residues that were identified through genetic analysis as important for interactions with PhoR were in close proximity to regions of the PhoR PAS domain that were identified by scanning mutagenesis. Fig 3.7C shows that some of the PhoR 157-162 residues (shown in green) appear to interact with the A147 and R148 residues of PhoU (shown in blue).

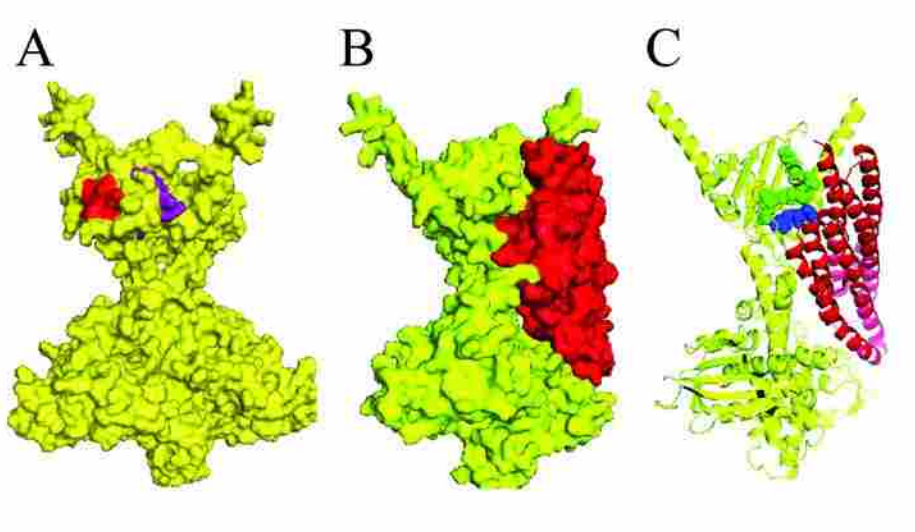


Figure 3.7 Structures of modeled PhoR and PhoU.
A. Interaction sites identified from scanning mutagenesis that are found on the surface of a modeled PhoR structure, 141-146 are red and the 157-162 residues are purple. B. Shows the protein surface predictions of the PhoU/PhoR docking model. PhoR is colored yellow and PhoU is colored red. C. A cartoon of protein backbones with PhoU A147 and R148 shown as spheres in blue and PhoR K157-S162 shown as green spheres.

3.3.7 Direct Coupling Analysis

Support for this interaction model was obtained through a bioinformatic method, called direct coupling analysis (DCA) that identifies covariance between residues. DCA identifies which residues from a sequence tend to co-evolve with any other residue by measuring the predictive power of one residue on another. This can identify direct interactions between residues, such as residues that are directly involved in protein/protein interactions. Also, residues that are involved in indirect interaction (for example residues that play a role in proper structural conformation) and mechanistic interactions (as would be found in residues associated with an enzyme active site) are identified with DCA (130). Mutual information (MI) quantifies the direct and indirect coupling of residue pairs. MI has been used to identify both intradomain and interdomain interactions (131-137). Direct information (DI) is a value that attempts to eliminate the indirect coupling of MI due to neighboring residues. Biologically, MI may identify residues

that play a role in interaction through indirect means that may not be identified with DI. We used MI to sort our DCA results to identify the residues that play an important role in PhoU/PhoR interaction,

We analyzed sequences of PhoU and PhoR from 196 gammaproteobacteria. We concatenated the *phoU* and *phoR* sequences from each species with 20 ala residues in between the two genes to aid in alignment. We aligned the sequences and then compared all the residues with DCA. The alignment had some gaps, mostly at the beginning and end of the two genes (Fig. S.4)

Our previous work identified PhoU A147 and R148 residues as being involved in interaction with PhoR. To isolate residues from the PhoR PAS domain that show covariance with PhoU A147 and R148, we sorted the top co-varying residues of the PhoR PAS domain by MI for both A147 and R148 residues of PhoU. The top ten predicted pairs are shown in Table 3.1 and Table 3.2. Several of these co-varying residues fall in the same region of the PAS domain of PhoR identified by the BACTH analysis. When the top five residues for A147 and R148 are highlighted in the docking model, we see that many cluster near the predicted interaction surface. Given that some of the residues for A147 overlap with the R148 residues we colored residues that co-vary with A147 green, residues that co-vary with R148 cyan, and residues that co-vary with both purple (Fig 3.8A).

Table 3.1 DCA of PhoU A147 vs. PhoR PAS domain

PhoR	PhoU	Mutual information	Direct Information	Distance (Å)
Y149	A147	0.5876	0.0016	12.1
N166	A147	0.5475	0.0006	14.7
F161	A147	0.5189	0.0005	12.2
P164	A147	0.5045	0.0009	12.9
R163	A147	0.4823	0.0005	12.1
G118	A147	0.4796	0.0011	19.0
Q186	A147	0.4565	0.0009	16.6
L187	A147	0.4459	0.0012	13.7
D160	A147	0.4442	0.0006	6.0
Q154	A147	0.4438	0.0008	10.9

Table 3.2 DCA of PhoU R148 vs. PhoR PAS domain

PhoR	PhoU	Mutual information	Direct Information	Distance (Å)
R148	R148	0.4980	0.0045	11.4
Y149	R148	0.4927	0.0006	7.2
G118	R148	0.4121	0.0011	16.3
N145	R148	0.4082	0.0011	12.9
N166	R148	0.3901	0.0007	3.1
P164	R148	0.3861	0.0004	2.0
Q154	R148	0.3831	0.0006	12.1
N139	R148	0.3795	0.0004	21.9
P150	R148	0.3781	0.0017	12.0
R163	R148	0.3767	0.0003	3.0

Table 3.3 DCA of PhoU E121 vs. PhoR CA domain

PhoR	PhoU	Mutual information	Direct Information	Distance (Å)
K428	E121	0.512	0.003	17.8
N429	E121	0.489	0.001	19.2
D316	E121	0.475	0.001	10.6
T340	E121	0.461	0.004	23.8
I426	E121	0.460	0.003	10.2
L274	E121	0.453	0.001	6.4
T271	E121	0.453	0.001	3.4
E375	E121	0.451	0.003	22.3
S430	E121	0.443	0.001	20.1
K277	E121	0.437	0.001	14.3

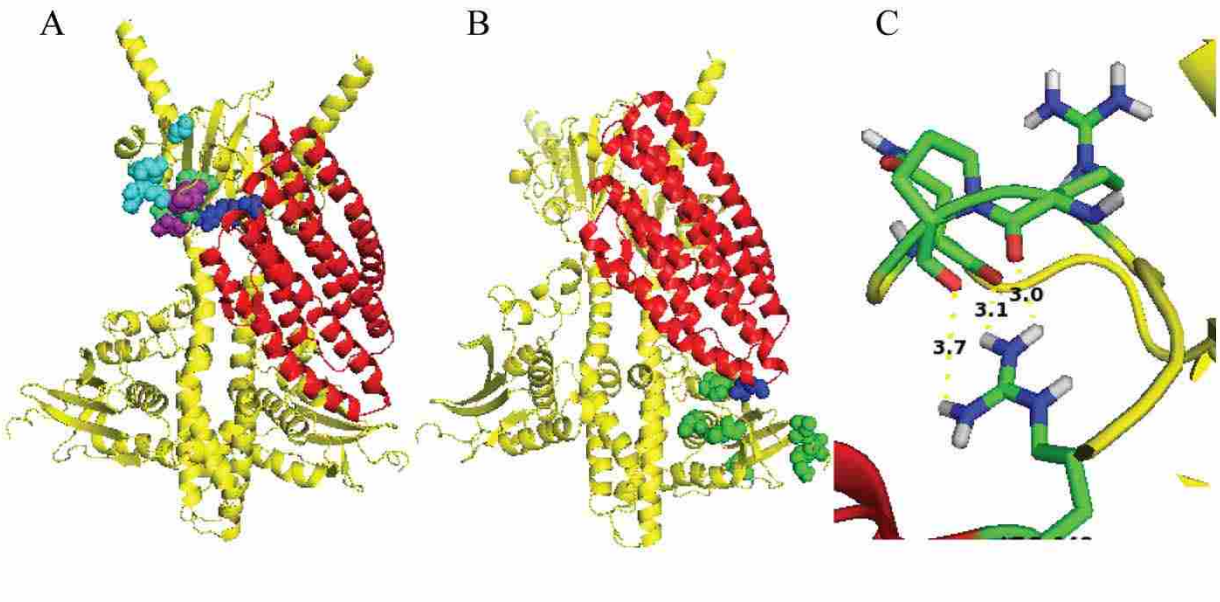


Figure 3.8 Residues from DCA analysis.
A. PhoR in yellow and PhoU in red with A147 and R148 highlighted in blue and the top five DCA residues of PhoR PAS domain highlighted as spheres (residues that co-vary with A147 are green, R148 are cyan, and residues that co-vary with both are purple). B. PhoU E121 highlighted in blue and the top seven residues in the DCA analysis of the PhoR CA domain highlighted in green spheres. C. PhoU R148 potential interactions with the side chain of PhoR PAS R163, P164, and N166 residues (distances labeled are Å).

For a more complete analysis, we also sorted the potential interacting pairs by DI and found that several of the same residues were found in the top pairs (Tables 3.4 and 3.5). In the tables, residues that were also found in the top ten residues of MI sorted data are underlined. Several other residues sorted by MI show relatively large DI values (Tables 3.1 and 3.2). Also, it is interesting that many of the interacting pairs sorted by DI fall into the loop regions identified by the PAS domain scanning mutagenesis (Fig. 3.6).

Table 3.4 DCA of PhoU A147 sorted by Direct Information.
Underlined PhoR residues are also in the top ten residues when sorted by Mutual Information

PhoR	PhoU	Mutual information	Direct Information	Distance (Å)
R148	A147	0.4267	0.0028	16.9
T195	A147	0.1738	0.0019	18.6
Y149	A147	0.5876	0.0016	9.3
L133	A147	0.2515	0.0016	20.9
I120	A147	0.1997	0.0015	20.0
E151	A147	0.4273	0.0014	8.0
L146	A147	0.3957	0.0014	15.8
L126	A147	0.2256	0.0014	22.2
V168	A147	0.3870	0.0014	15.0
F121	A147	0.2997	0.0013	22.1

Table 3.5 DCA of PhoU R148 sorted by Direct Information
Underlined PhoR residues are also in the top ten residues when sorted by Mutual Information

PhoR	PhoU	Mutual information	Direct Information	Distance (Å)
R148	R148	0.4980	0.0045	11.4
I143	R148	0.2236	0.0019	11.2
L167	R148	0.2752	0.0017	5.4
P150	R148	0.3781	0.0017	11.6
L126	R148	0.1135	0.0016	14.5
E176	R148	0.3762	0.0014	13.5
I120	R148	0.2126	0.0013	15.2
F152	R148	0.2972	0.0012	13.4
N145	R148	0.4082	0.0011	12.9
G118	R148	0.4121	0.0011	16.3

With the docking model it appears that the loops on the opposite side of PhoU from the PAS interacting surface are in close proximity to the CA domain of the opposite PhoR of the dimer. To observe any covariance between PhoU and PhoR at these surfaces, we compared the PhoU E121 residue to the CA domain with DCA. We see that the top PhoR CA domain residues that co-vary with E121 (shown as green spheres) cluster at the surface of PhoR nearest to PhoU E121 (shown as blue spheres) (Table 3.3, Fig 3.8B).

We noticed that there are few residues that appear to interact based on the docking model and DCA that did not show loss of function in our scanning mutagenesis (~PhoR163-166). To better understand these results we looked closer at the residues in our docking model structures. When we highlight the side chain of the PhoU R148 residue and the side chains of residues in positions 163, 164, and 166 of the PhoR PAS domain, it appears that the R148 may interact with the amino acid backbone for these sites and not the side chain residues as they appear to point away from the PhoU structure (Fig 3.8C). Changing the side chains of these amino acids in our scanning mutagenesis of the PAS domain may not disrupt PhoU/PhoR interactions, which explains our scanning mutagenesis results.

Additionally, when distances between predicted interacting pairs is measured, several residues are found near R148 (within ~8Å) but few are found that near to A147. One explanation for this is that PhoU A147 itself does not directly interact with PhoR, but mutations of PhoU A147 may affect the ability of R148 to interact with PhoR. For example, the A147E mutation places a large acidic side chain right next to the R148 basic side chain and may disrupt proper PhoU/PhoR interaction and phosphate signaling. Using MI to sort our DCA results allows for identifying residues like this that are important for proper interaction but may not be directly involved. This phenomenon may also explain some of the PhoR PAS residues identified with DCA, like PhoR G118. The PhoR G118 residue is not found very near to PhoU R148 or A147 in our docking model (Table 3.1 and Table 3.2). However, G118 is on the surface of the PhoR PAS domain on a loop between α-helices that appears to form the top of the PhoU binding pocket. One would expect that mutations of a highly flexible glycine at this position could be associated with compensating mutations of PhoU in the R148 region.

In an effort to expand our DCA we completed an analysis using 415 sequences from various species belonging to the proteobacteria group. This analysis further confirmed our findings that PhoU interacts with the PAS domain of PhoR. When we compared the entire PhoR protein with either A147 or R148 residues of PhoU, we found that 14 of the top 20 pairs sorted by MI for both A147 and R148 fall within the PAS domain of PhoR. With aligning so many diverse PhoR sequences, there were several gaps in the alignment (Fig. S5). Often these gaps correspond to the loop regions between the α helices of the PAS domain. Where the DCA residue for PhoR fell into a gap in the *E. coli* PhoR aligned sequence, we report the residues on either side of the gap (Table 3.6 and Fig 3.10). Interestingly, many of the top ten residues fall in the loops between predicted α helices (Fig 3.5A) that form the same predicted PhoU binding pocket of our docking model (Fig3.10). There is also a residue that maps to the predicted PhoU/PhoR CA domain interaction surface that was identified (H272-L273). This again may be due to our use of MI and the fact that even though the residue is far from the direct interaction site, it plays a role in docking and any mutations there may need to be compensated for at other interaction sites. This also appears to support the hypothesis that PhoU interacts with the CA domain of PhoR.

```
> Escherichia_coli_K-12_MG1655
--------------MDS--LNLNK-HISGQF--NAELESIRTQVMTMGGMVEQQLSDAITAMHNQD--SDLA
KRVIEGDKNVNMMEVAIDEACVRIIAKRQPTASDLRLVMVISKTIAELERIGDVADKICRTALE--KFSQQH
------QPLLVS----LESLGRHTIQMLHDVLDAFARMDIDEAVRIYREDKKVDQEYEGI--VRQLMTYMM-
EDSRTIPSVLTALF----CARSIERIGDRCQNICEFIFYYVKGQDFRHVGGDE-LDKLL----------A--
-----GKDSDK--------------------AAAAAAAAAAAAAAAAAAAA--------------------
------------------------------M-----------------------------------LERLSW
KR--------------------------------------------------------------------
-----------------------------LVLELLLCCLPAFIL----GA-----FFGYL----------
-----------PWFLLASVTG-LLIWHFWNLLRLSWWLWVDRS-------MT-PPPG-RGS-WEPLLYGLHQ
---MQLRNKKRRRELGNLIK-RFRSGAE---------SLPDAVVLTT-E-------EGGIFWCNGLAQQIL-
GLRWPEDNGQ--NILN--LLRYPEFTQYL------KTR-DFSRPLNL---V--LNT----------------
----GR----HLEIRVMPY--------------------------TH--KQLLMVARDVTQMHQLEGARRN
FFANVSHEL------RTP----LTVLQGYLEMMN-------------EQPL----EGA---VREKALHTMRE
QTQRMEGLVKQLLTLSKIEA-APTH-LLNEKVDVPM-MLRVVEREAQTLS--------QK----K-QTFTF-
--E-I---DN------------GLKVSGNEDQLRSAISNLVYNAVNH-----T----PEG------------
------------TH----ITVRWQR-V----PHG--------------------------A----EFSVED
NGPGIAPEHIPRLTERFYRVDKAR-SRQT-GGSGLGLAIVKHAVNHHESRLNIESTV----G-KGT------
------------------RFSFVI--PERLIAKN-SD-----------------------------------
-------
```

Figure 3.9 Alignment of *phoU* and *phoR* from *E. coli* used for DCA analysis
For DCA, we concatenated the *phoU* sequence with the corresponding *phoR* sequence from the same bacterial species. We added 20 alanine residues between the two gene sequences to aid in correct multiple sequence alignment. After aligning sequences from the different species, we used the FASTA formatted files for DCA.

Table 3.6 DCA of PhoU R148 compared to all of the PhoR residues using 415 sequences from bacterial species across the proteobacteria group

PhoR	PhoU	Mutual Information	Direct Information
T171-G172	R148	0.832	0.002
D138	R148	0.825	0.005
H272-L273	R148	0.820	0.002
W56	R148	0.802	0.003
N145	R148	0.755	0.003
H184-K185	R148	0.751	0.002
T171-G172	R148	0.750	0.002
L156-K157	R148	0.739	0.003
L156-K157	R148	0.737	0.007
L156-K157	R148	0.733	0.002

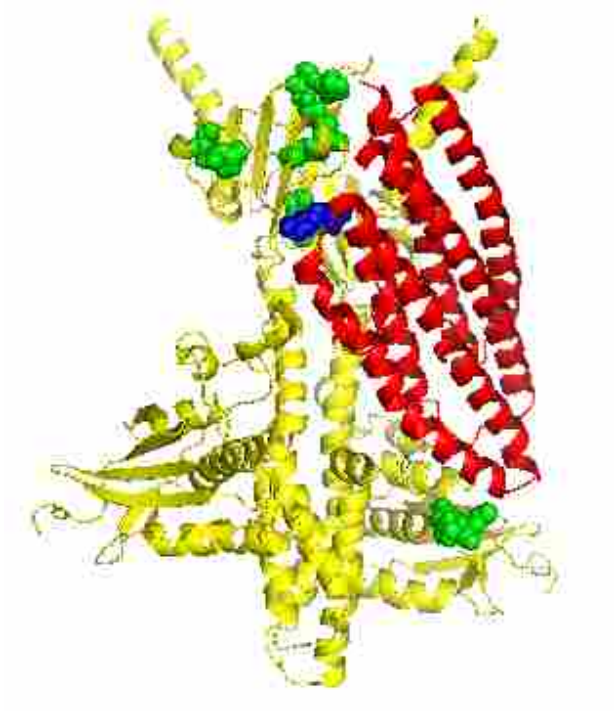


Figure 3.10 DCA analysis of PhoU R148 compared to all of the PhoR residues using sequences from bacterial species across the proteobacteria group
PhoR is shown as yellow and PhoU is shown as red. The R148A residue of PhoU is highlighted as blue spheres and green spheres represent the top ten residues identified through DCA using 415 different sequences of various species from the proteobacteria group.

3.3.8 Potential model for PhoU mediated signaling to PhoR

Our results confirm that PhoU is necessary for proper signaling in high phosphate, but that the PstSCAB transporter is also necessary. We have identified the A148 and R149 residues of PhoU as being important for interaction with PhoR. Using scanning mutagenesis, we identified residues of the PhoR PAS domain essential for interaction with PhoU. A docking model was identified and tested with DCA. The docking model shows PhoU interacting with both the PAS domain and the CA domain of PhoR, pointing to a potential mechanism for PhoU mediated switching of PhoR's kinase to phosphatase activity.

The switch from kinase to phosphatase activity for histidine kinases is dependent upon the conformation between the CA and DHp domains (21, 138). The CA domain must remain

flexible for proper kinase function. Also, the DHp domain was previously shown biochemically to be the portion of PhoR that has phospho-PhoB phosphatase activity (32). It is possible that when PhoU interacts with PhoR, that interaction may constrain the CA domain to inhibit kinase activity and expose the DHp domain to allow phosphatase activity of PhoR. These results are the first to point to a specific molecular mechanism for PhoU mediated modification of PhoR activity.

3.4 Conclusions

These studies have combined genetic, mutagenesis, computer modeling, and DCA analyses in studying the molecular interaction of PhoU and PhoR. It will be interesting to apply these techniques to identify residues involved in other signaling pathways and develop interaction models. For example, PhoU interaction with PstB appears to be weaker than PhoU interaction with PhoR based on our BACTH (Fig 3.6A and 3.6B), making it more difficult to directly identify residues essential for interaction. Using modeling and DCA, potential interacting residues may be identified. Efforts to isolate a complete signaling complex of PhoR, PhoU, and PstSCAB have been unsuccessful in the past. However, if the sites of interaction were identified, perhaps an entire signaling complex could be isolated and characterized using methods directed toward the specific interaction sites.

3.5 Methods

3.5.1 Strains, plasmids, and reconstructing of signaling system

Plasmids that were used include pKG116(103), p116U2(63) (a pKG116 with the *phoU* gene under salycilate expression control), p116A147E, p116U2 A147K, p116U2 R148A, p116U2 R148E, p116U2 D150A (all the mutant p116U2 plasmids were constructed with QuikChange site-directed mutagenesis kit from Agilent Technologies and verified by DNA sequence analysis as described in (1) (primers listed in Table S2)), pRR48(104), and p48SCAB (a pRR48 with the *pstSCAB* genes inserted also as described in (1) (primers are listed in Table S2)). For signaling analysis, combinations of p48SCAB and a pKG116 derived plasmid were introduced into the *E. coli* strain BW26337 (having a chromosomal deletion of the *pstSCABphoU* genes constructed from the wildtype strain BW25113(100)). The cultures were grown at 37°C with shaking overnight in morpholinepropanesulfonic acid (MOPS) defined medium with 0.06% glucose and 2.0 mM phosphate (a high phosphate minimal media)(1).

3.5.2 Growth of PhoU A147E mutant

STAC and STACΔphoU (1) were used with pKG116, p116U2, or p116U2 A147E (a p116U2 plasmid mutated with QuikChange site-directed mutagenesis kit from Agilent Technologies and verified by DNA sequence analysis (primers listed in Table S2). Triplicate overnight cultures were grown in LB media with 100 μM IPTG and the OD_{600} of a 1:5 dilution of culture was measured. Values were adjusted for dilution and averages are reported with error bars representing the standard deviation.

3.5.3 Alkaline Phosphatase Assays

We based our Alkaline Phosphatase (AP) Assays on a 96 well plate β-Galactosidase Activity assay previously described (1). Cultures were grown overnight at 37°C with shaking.

The OD_{600} was read on a 1:4 dilution of the culture. 1 ml of culture was collected and resuspended in 900 µl of 1 M Tris-HCl pH 8.2. 1 drop of 0.1% sodium dodecyl sulfate and 2 drops of chloroform were added to each tube and tubes were vortexed vigorously for 15 sec. Tubes were then spun for 1 min at 16,000 x g and 200 µl of each sample was loaded into a well of a 96 well flat bottomed plate. 40 µl of 20 mM p-Nitrophenyl phosphate in 1 M Tris-HCl pH 8.2 was added. The OD_{420} values were read once a minute for 20 min at 37°C. Units of activity were calculated as (1000 x slope of a line fit to OD_{420} in mOD/min)/(4 x OD_{600} of 1:4 dilution of the overnight culture). Cultures were grown in triplicate and the average is reported with error bars representing the standard deviation.

3.5.4 BACTH and β-Galactosidase Assays for scanning mutagenesis

BACTH analysis and β-Galactosidase Assays were performed as described in (1). Briefly, using the QuikChange Lightning Site-Directed Mutagenesis Kit from Agilent Technologies, the p18CRN-PAS plasmid was mutated to change every two amino acids of the PhoR PAS domain to code for alanine-cysteine (primers listed in Table S2). For example, p18CRN-PAS 109 had PhoR D109A and A110C changes. These residues were chosen because they are not predicted to cause major secondary structural changes. Each p18CRN-PAS plasmid of the library of alanine-cysteine mutants was transformed with pKT25phoU into the tester strain, BTH101 from EuroMedex. Cultures were grown in triplicate in LB and assayed as described in (1). The percent of PhoU/PhoR Interaction was found by dividing the average activity of each sample with the activity of an unmutated p18CRN-PAS control and multiplied by 100.

3.5.5 Protein structure modeling and protein docking modeling

We used the ClusPro webserver (139-142) to dock structures of a dimer of the cytoplasmic portion from *E. coli* PhoR (structure modeled from the structure for VicK from *Streptococcus mutans* (34)), and PhoU (modeled from *Thermatoga maratima* PhoU (95)) using Protein Homology/analogY Recognition Engine V 2.0 (Phyre2) (35).

3.5.6 Direct Coupling Analysis

We used PhoU and PhoR sequences from 196 species of bacteria from the gammaproteobacteria group. Sequences were concatenated starting with PhoU, followed by PhoR from the same species. 20 alanine residues were placed between the two protein sequences to aid in differentiating the PhoU from the PhoR after the alignment. Using MAFFT version 7 (website is http://mafft.cbrc.jp/alignment/server/); we were able to create an alignment file to use in DCA analysis. A MATLAB implementation of direct coupling analysis reported in (143) distinguished differences between the direct information and mutual information between all protein residues. The command line argument used was matlab -nodisplay -nojvm -nosplash -r "dca $INPUT_ALIGNED_FASTA $OUT_FILE" . The DCA implementation used the input aligned fasta file in its calculations, outputting a file with N(N-1)/2 (N = length of the sequences) rows and four columns: residue i (column 1), residue j (column 2), MI(i,j) (Mutual Information between i and j), and DI(i,j) (Direct Information between i and j). All inserts columns were removed from the alignment during the analysis. Concluding this analysis, a simple python script was used to remove all rows in which residues i and j were within five residue numbers of each other. This was done to eliminate false positive residue interactions caused by their proximity to

other amino acids. Mutual information was later screened to signify relatedness between two residues.

Chapter 4 Additional research findings involving PhoU

4.1 Metal Binding

As we analyzed the metal binding curves of PhoU with various metals, a pattern appeared that may be significant. As seen in Fig.S3, it appears that at lower concentrations of magnesium the fluorescence is initially quenched, but as the concentration increases fluorescence is enhanced. This trend appeared consistently in our experiments. We did not see the same trend with manganese. However, it may be that the increased binding affinity for manganese made it so that we did not test concentrations low enough to see an initial quenching trend.

One possible explanation of the different effect on tryptophan fluorescence at different metal concentrations may be that when PhoU initially binds metal, the environment around the tryptophan is less hydrophobic (based on the properties of tryptophan fluorescence as discussed in (112)), but as another metal is bound by the same PhoU, perhaps at the metal binding site proximal to the tryptophan, the local environment is then more hydrophobic or more exposed to the metal ion. This may imply a two-step mechanism for PhoU binding metal where the PhoU dimer "breaths" to allow metal binding where binding of a single metal ion allows the PhoU dimer to exclude water, but when the second metal is bound the dimer is less tight and exposes the tryptophan to the metal ion and water. This metal mediated more open dimer structure may aid in allowing a sufficient portion of the PhoU in the cell to be in a monomer conformation to interact with the membrane and be sequestered to the cellular localization of PhoR and PstSCAB transporter.

One model for PhoU function is that soluble PhoU is present as a dimer, but a fraction of PhoU may be found as monomers where the metal bound allows it to interact with the

phospholipid head groups of the membrane and localize near PhoR and PstB. Then, PhoU monomer interaction with PstB in the open conformation (associated with phosphate transport indicating a high phosphate environment) allows PhoU to specifically interact with PhoR and trigger phosphatase activity. If PhoU interacts with the membrane through the metal binding surface, perhaps binding metal allows PhoU to localize to the membrane to be in a higher effective concentration near membrane bound PhoR, allowing for proper interaction. Loss of metal binding activity can be made up for by over expressing PhoU, thus having a sufficient number so that the small fraction of PhoU in monomeric states can interact with PhoR independent of metal mediated membrane interaction.

4.2 Phylogeny of bacteria and gene organization of *phoB, phoR, pstS, pstC, pstA, pstB,* and *phoU*

As we collected sequences for DCA analysis, several different gene organizations of the Pho regulon expression control genes were identified. We created a phylogenetic tree based on the various gene orders to identify branches where various genes may have been altered or lost using various species of proteobacteria (Figure 4.1). The phylogenetic tree was generated using iTOL Interactive Tree Of Life webserver (144, 145). The general trend appears to follow known phylogenetic branching with genes being rearranged and lost with different speciation events.

It is interesting to note that species of the epsilonproteobacteria do not appear to have any genomic *phoU* genes. Given that many epsilonproteobacteria live in the digestive tract of various hosts, or in the deep ocean, it may be that the fairly constant phosphate concentration of these environments did not provide sufficient evolutionary selection to maintain this signaling system

and so it has been lost. This may also explain the lack of annotated phosphate regulon control genes in *Neisseria meningitidis*.

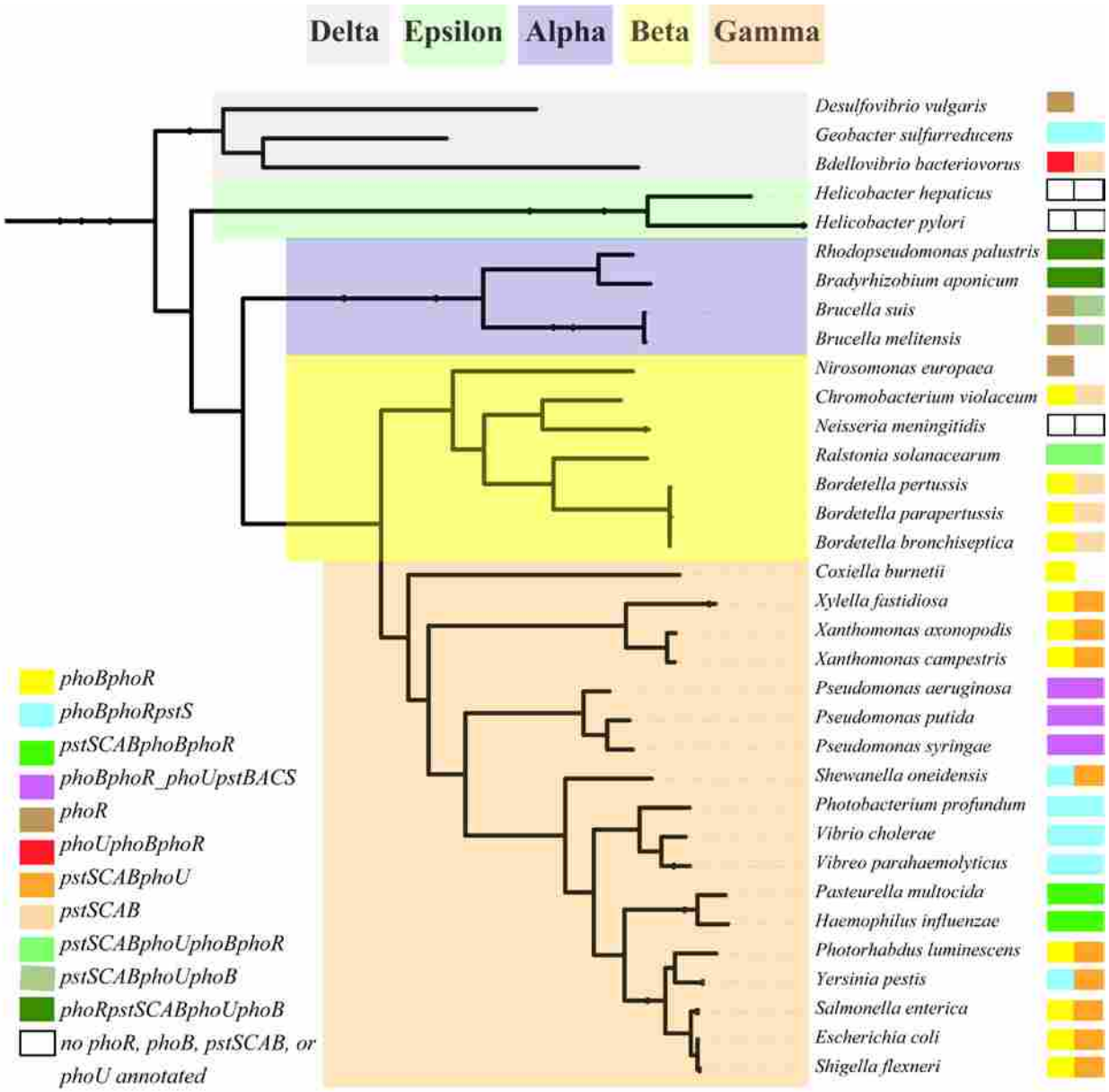


Figure 4.1 Phylogenetic Tree of Pho regulon control gene arrangements.
This tree was constructed with iTOL Interactive Tree Of Life webserver. Bacterial genomes were searched using KEGG webserver for annotated genes. Species that did not have an annotated *phoU* gene were checked by running a BLAST search against *E. coli* MG1655 PhoU protein and no clear PhoU protein sequences were found.

The alphaproteobacteria species highlighted in our trees do not appear to have *phoR* and *phoB* as a single operon, but maintain the *pstSCABphoU* order of genes. With *Bradyrhizobium japonicum* and *Rhodopseudomonas palustris*, it appears that the *pstSCABphoU* operon has inserted between the *phoB* and *phoR* genes. However, with *Brucella melitensis* and *Brucella suis* it appears that the *phoR* gene from this group has relocated to another area of the chromosome.

Also, the *Coxiella burnetii* genome analyzed did not show any *pstSCABphoU* genes. This is not unexpected given the obligate intracellular nature of *C. burnetii*. These genes may have been lost due to the environment that this bacteria inhabits maintains a fairly constant concentration of phosphate as part of the normal homeostasis of the host cells. Perhaps the remaining *phoBphoR* operon in this species has been co-opted for some other purpose than the phosphate response.

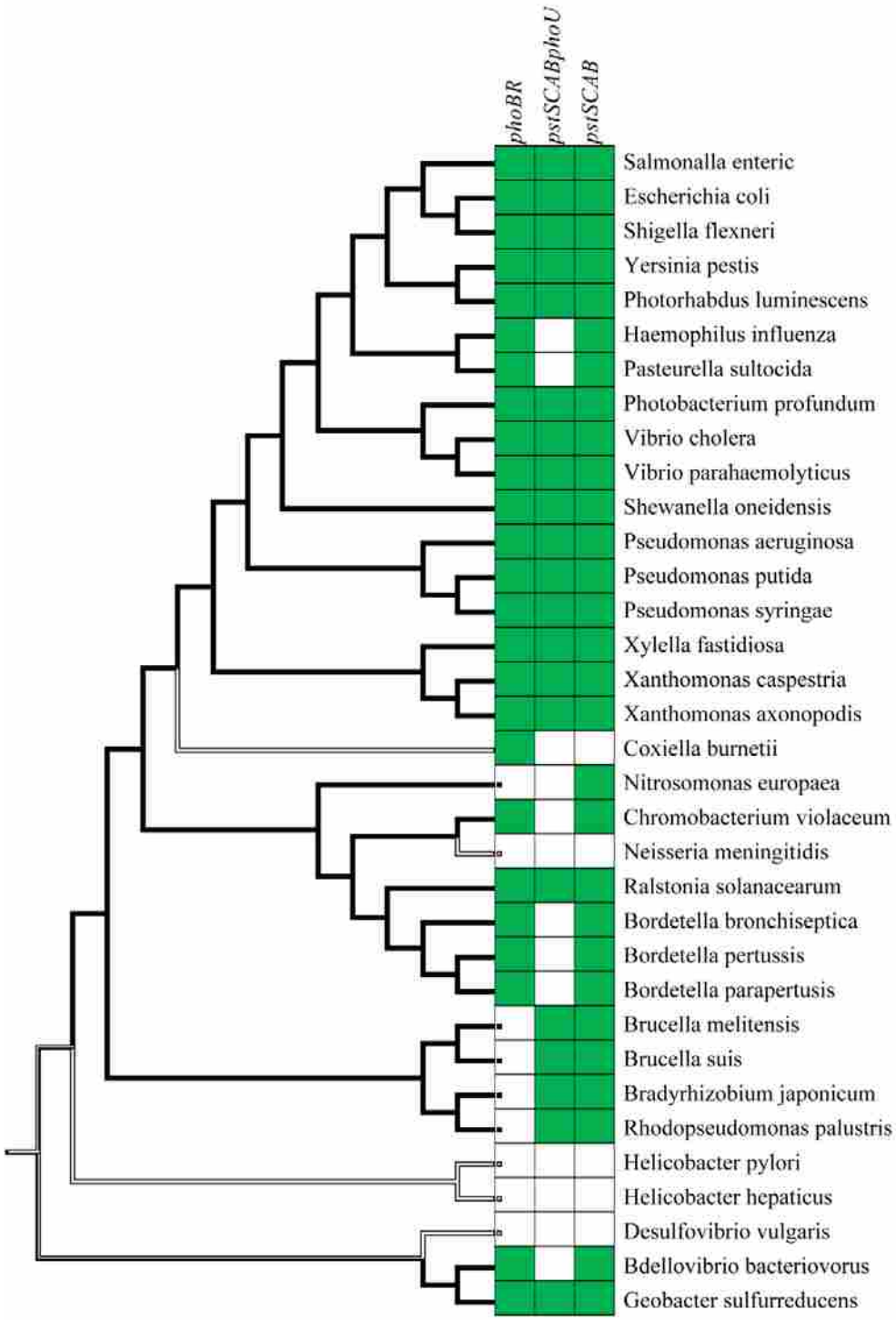


Figure 4.2 A phylogenetic tree with various gene states highlighted
This tree was created using the program Mesquite: A modular system for evolutionary analysis version: 3.01 (146). This tree shows black branches as where the *pstSCAB* genes are found as a single operon on the chromosome. White branches are where the *pstSCAB* genes are not found as a single operon in the species shown. On the right of the tree, green boxes represent the presence and white boxes represent the absence of *phoBphoR* genes as an operon, *pstSCABphoU* genes as an operon, and the *pstSCAB* genes as an operon respectively.

4.3 Attempts to confirm protein/protein interactions

Initially, to further characterize residues involved in PhoU/PhoR interaction we utilized the BACTH system. With this system, protein interaction is linked to the reporter cells ability to grow with maltose as the sole carbon source. We used this phenotype as a selection for a screen of potential compensatory mutations in PhoU or PhoR mutant proteins, that we identified significant loss of interaction, that restored interaction. For example, we created a mutant library of *phoR* containing plasmids by growing them in XL1 Red (Agilent Technologies) cells for several generations. This library of plasmids was used to transform BTH101 cells containing a corresponding *phoU35* containing plasmid that showed weak interaction with wildtype PhoR. The cells were then grown on maltose indicator plates to select for cells where interaction was restored. Potential compensatory mutants were isolated, re-transformed to confirm phenotype, and sequenced to identify mutations that restored interaction. Unfortunately, none of the potential compensatory mutants examined showed mutations in the genes we were studying.

One possible explanation for this finding may be that with the initial mutant proteins, whose interaction was significantly weaker than the wildtype but still existent, significant interaction was restored by mutations in promoter sequences or in regulator proteins that allowed increased expression of the mutant protein from the plasmid, thus giving the appearance of restored interaction.

With our predicted docking model identified, we wanted to confirm the sites of interaction using protein crosslinking and/or co-elution studies. Specifically, we cloned mutants of PhoU that would code for a cysteine residue at the A147 position similar to our scanning mutagenesis of PhoR's PAS domain. We then attempted to identify specific crosslinking of the

mutant PhoU with various PhoR mutants that contained cysteine residues at predicted interaction sites (PhoU AR147,148AC; PhoU MD149,150AC; PhoR RD159, 160AC; PhoR FS161,162A; and PhoR DA109,110AC (as a negative control because the 109,110 sites do not appear to interact with PhoU in our docking model)). Unfortunately, every condition that was tested showed crosslinking of all samples, even the negative controls. It appears that the native cysteine residues of wildtype PhoU can crosslink with the cysteine containing PhoR proteins non-specifically under all the conditions we tested.

For co-elution studies, we based our conditions on the conditions used to confirm the initial PhoU/PhoR interaction (Figure 2.3C). We used various PhoU mutants in an attempt to confirm the weaker interaction of PhoU with PhoR seen in our BACTH analysis of PhoU A147E (Figure 3.5A). However, we could not find conditions that specifically allowed native PhoU to co-elute with PhoR but prevented the co-elution of PhoU A147E. Given that the BACTH showed the PhoU A147E maintained a significant amount of interaction with PhoR, it may be that co-elution is not sensitive enough to resolve the small but significant difference between the binding affinity of PhoU versus PhoU A147E binding to PhoR.

4.4 Attempts to isolate a PhoU mutant that signals independently of PstSCAB

One question that we had as we have studied the role of PhoU in Pho regulon expression control was: can PhoU cause the shift to PhoR phosphatase activity in the absence of PstSCAB proteins. We attempted to address this question with the reconstructed signaling system in BW26337 cells (Fig 3.2). We saw that wildtype PhoU could not function to signal in the absence of the PstSCAB proteins in our initial studies. However, we wondered if there were any possible mutations in PhoU that could allow for signaling to PhoR in the absence of the PstSCAB

complex. To test this theory, we mutated the p116U2 plasmid (a plasmid that contains a wildtype copy of *phoU*) using XL1 cells (Agilent Technologies) to create a library of mutated *phoU* genes encoded on the pKG116 plasmid. We transformed this mutated library of plasmids into BW26337 cells and plated the cells onto LB plates. After colonies formed, an agar overlay containing 5-Bromo 4-Cloro 3-Indolyl Phosphate, p-Toluidine Salt (X-Phos) allowed for the differentiation of cells where regulation was dysfunctional (blue colonies like the BW26337 parent phenotype) and cells that no longer expressed AP in the high phosphate growth conditions (white colonies, as would be expected if the plasmid encoded PhoU restored negative regulation in the absence of the PstSCAB complex).

We screened >164,000 colonies and picked several white colonies. We found a frequency of ~0.54% of colonies with a white phenotype. However, with subsequent rounds of growth, most of the white colonies isolated returned to the blue phenotype. A few plasmids that maintained the white phenotype after repeated retransformation into BW26337 were sequenced, but no mutation in *phoU* was identified. Based on our work, we believe it is unlikely that any minor mutation will allow PhoU to signal to PhoR in the absence of the PstSCAB complex. This correlates with the finding that even in the PstS mutant background, the PstC, PstA, and PstB proteins are necessary for a PitA or PitB mediated suppression of the mutant phenotype (59). The high frequency of white colonies may be due to an increased mutation rate linked to the absence of functional repression of the PhoR/PhoB expression control system.

Chapter 5 Future work and conclusions

5.1 Future work

The methods used in these studies have potential to help better understand the signaling mechanism of the Pho regulon and have potential application to a variety of other signal transduction systems. Specifically, studies of the PhoU/PstB interaction have been difficult due to the relatively weak interaction (note the absolute values of the BACTH analysis of PhoU/PstB interactions versus PhoU/PhoR in Figure 3.5).

Interestingly, the EIIA proteins of various phosphotransferase systems have come up a few times in relation to various Pho regulon expression control genes. Specifically, the EIIA from the glucose phosphotransferase system is known to bind and inhibit transport of various transporters (51). However, the EIIA protein of the nitrogen-related phosphotransferase system has been shown to interact with PhoR and alter activity (68). In addressing the function of PhoU, it also appears to interact with an ABC transporter to limit transport and bind to PhoR to alter its activity. While the sequences and secondary structures of the EIIA protein and PhoU do not share high identity, perhaps they still function in similar manners. It would be interesting to observe the effect that EIIA and PhoU have on each other in binding and modulating PhoR activity (68). These study methods may be used to identify additional proteins that interact in this system, like an EIIA protein, and characterize their effect on signaling phosphate and identify potential cross-regulation mechanisms of various systems in *E. coli*.

5.1.1 PhoU and PstB

Using DCA to compare PstB sequences with PhoU sequences within the same species of bacteria may identify potential sites of interaction on each protein. Once potential sites are identified, these sites could then be mutated to test for loss of interaction. Additionally, sites may be mutated to cysteine residues and interaction may be confirmed through crosslinking experiments. Knowing the sites of interaction may allow a study where specific mutations disrupt interaction and further mutation of the predicted interaction interface may restore interaction (similar to work done with histidine kinases and response regulators in (26)).

If specific surface residues are identified that are responsible for PhoU/PstB interaction, docking models based on PhoU and PstB modeled structures can be identified where the sites of interaction of both proteins are in close proximity. With a PhoU/PstB docking model, perhaps a full signaling complex structural model may be proposed by docking the PstSCAB, PhoU, and PhoR protein structures as a complete phosphate signaling complex.

5.1.2 Complex isolation

Furthermore, using combinations of cysteine mutant PstB, PhoU, and PhoR perhaps an entire signaling complex could be crosslinked and isolated. This would confirm the signaling complex hypothesis that has been proposed for years but not yet confirmed with experimental data. Isolating the complex would be significant evidence for our proposed signaling mechanism for the Pho regulon's response to environmental phosphate levels. Additionally, with an entire complex isolated, perhaps co-elution experiments could identify if other proteins also interact with the complex. While there is no evidence that *E. coli* requires any proteins in addition to the PstSCAB, PhoU, PhoR, and PhoB proteins, perhaps other proteins may interact and modify the phosphate response in *E. coli*, like the $EIIA^{Ntr}$. Identifying all of the players that have an effect

on PhoR/PhoB mediated activation may elucidate how the Pho regulon is involved in virulence in many species and other potential cross-regulation of bacterial systems with the phosphate response.

5.2 Conclusions

This work has led to significant progress toward understanding PhoU's role in phosphate signaling in *E. coli*. Our work shows that the highly conserved residues of PhoU may play a role in metal binding and affect signaling function of PhoU, but the loss of signaling can be overcome with over expression of mutant PhoU. These results imply that metal binding may be essential for proper cellular localization of PhoU but not essential for direct signaling activity. This tends to preclude the hypothesis that PhoU acts as an enzyme. If metals were essential cofactors for enzymatic activity, overexpressed non-metal binding mutant proteins should not show any signaling activity.

We have shown that PhoU does in fact interact with PhoR at its PAS domain and also with PstB. This interaction adds evidence for the hypothesis that signaling occurs through the formation of a signaling complex between the PstSCAB transporter, PhoU, and PhoR. Additionally, we have identified residues in PhoU and PhoR that are important for protein/protein interaction. A potential docking model of PhoU and PhoR where these residues are found in close proximity has been proposed and tested by comparing the model with DCA results of co-varying residues in PhoU and PhoR. This model implies a potential molecular mechanism for PhoU mediated signaling to PhoR to alter the kinase/phosphatase activity through direct interaction with the CA domain of PhoR.

Future work characterizing the PhoU/PstB interaction as well as identifying the whole signaling complex may explain more fully how interaction with PstB can allow PhoU to signal for PhoR phosphatase activity. Identifying a structure of PstB associated with phosphate transport may aid in more fully understanding the conformational arrangement of proteins in ABC transporters during active transport. These results would answer remaining questions about this signaling system that has been studied for many years.

Supplementary Tables and Figures

Table S.1 Oligonucleotides

Primer name	Sequence 5'-3'
PstBfor	GACGCAGGCC**CATATG**AGTATGGTTGAAACTGCCCCGAGTAAAATTCAGG
PstBrev	GACGCAGGCC**GGTACC**TCAACCGTAACGACCGGTGATGTAG
PhoUfor	AAACATATTTCCGGCCAGTTCAACGCCGAACTGG
PhoUrev	GCG**AAGCTT**TTATTTGTCGCTATCTTTCCCCGCCAGCAGTTTATCCAGC
PhoUCHis for	AAGATAGCGACAAACACCACCACCACCACCACTAAGGTACCGAGCTCCTCGA
PhoUCHis rev	TCGAGGAGCTCGGTACCTTAGTGGTGGTGGTGGTGGTGTTTGTCGCTATCTT
D58A forward	CGCGTCATCGAAGGCGCAAAGAACGTCAACATGATGG
D58A reverse	CCATCATGTTGACGTTCTTTGCGCCTTCGATGACGCG
D85A forward	CCGACGGCGAGCGCACTGCGACTGGTTA
D85A reverse	TAACCAGTCGCAGTGCGCTCGCCGTCGG
E100A forward	CCATTGCCGAGCTGGCTCGTATTGGCGACG
E100A reverse	CGTCGCCAATACGAGCCAGCTCGGCAATGG
R101A forward	ATTGCCGAGCTGGAGGCAATTGGCGACGTGGCGG
R101A reverse	CCGCCACGTCGCCAATTGCCTCCAGCTCGGCAAT
E200A forward	GCGCGTTCTATCGCACGTATTGGCGACC
E200A reverse	GGTCGCCAATACGTGCGATAGAACGCGC
R201A forward	GCGCGTTCTATCGAAGCAATTGGCGACCGCTGCC
R201A reverse	GGCAGCGGTCGCCAATTGCTTCGATAGAACGCGC
D204A forward	CTATCGAACGTATTGGCGCACGCTGCCAGAATATTTGTGA
D204A reverse	TCACAAATATTCTGGCAGCGTGCGCCAATACGTTCGATAG
D58A, N62A for	GCGTCATCGAAGGCGCCAAGAACGTCGCCATGATGGAAGTGGC
D58A, N62A rev	GCCACTTCCATCATGGCGACGTTCTTGGCGCCTTCGATGACGC
E100A,R101A for	CCATTGCCGAGCTGGCGGCTATTGGCGACGTGGC
E100A R101A rev	GCCACGTCGCCAATAGCCGCCAGCTCGGCAATGG
E200A R201A for	GCGCGCGTTCTATCGCAGCTATTGGCGACCGCTG
E200A R202A rev	CAGCGGTCGCCAATAGCTGCGATAGAACGCGCGC
PhoRfor	CCAGGT**CATATG**CTGGAACGGCTGTCGTGGAAAAG
PhoR rev	CCAGGT**GGTACC**TTAATCGCTGTTTTTGGCAATTAAACGTTCCGGG
PhoRNHis for	TAAAGGAGTGTCATATGCACCATCACCATCACCATCTGGAACGGCTGTCGTG
PhoRNHis rev	CACGACAGCCGTTCCAGATGGTGATGGTGATGGTGCATATGACACTCCTTTA
PstBHis rev	GTA**GGTACC**TCAGTGGTGGTGGTGGTGGTGACCGTAACGACCG
PhoU BACTH for	GACTGACTGA**TCTAGA**CGGTGGTATGGACAGTCTCAATCTTAATAAACATATTTCCG
PhoU BACTH rev	GACTGACTGA**GGTACC**TTATTTGTCGCTATCTTTCCCCGCC
PstB BATCHfor	GACTGACTGA**TCTAGA**CGGTGGTATGAGTATGGTTGAAACTGCCCCG
PstB BATCHrev	GACTGACTGA**GGTACC**TCAACCGTAACGACCGGTGATGTA
PhoR BACTH for	GACTGACTGA**TCTAGA**CGGTGGTATGGTGCTGGAACGGCTGTCGTGGAAAAGGCTGG
PhoR CR BACTH for	GACTGACTGA**TCTAGA**CGGTGGTATGCGCAGTATGACCCCGCCACCGGGGCGTGGTA
PhoR PAS BACTH for	GACTGACTGA**TCTAGA**CGGTGGTATGCTGCCCGACGCGGTGGTGCTGACCACGGAAG
PhoR DHp BACTH for	GACTGACTGA**TCTAGA**CGGTGGTATGACGCAATATCTGAAAACGCGTGATTTTTCTCGC
PhoR CA BACTH for	GACTGACTGA**TCTAGA**CGGTGGTATGGAAGCCGCACCGACGCATTTGCTCAATGAA
PhoR BACTH rev	GACTGACTGA**GGTACC**TTAATCGCTGTTTTTGGCAATTAA
PhoR DHp BACTH rev	GACTGACTGA**GGTACC**TTATTCTATTTTCGACAGCGTCAGCAA
PhoR PAS BACTH rev	GACTGACTGA**GGTACC**TTATTGCGTGACATCACGCGCCACCAT
PhoR CR BACTH rev	GACTGACTGA**GGTACC**TTACGACTCCGCGCCGCTACGAAAGCG

Bold indicates restriction sites

Table S.2 Primers

Primer name	Sequence 5'-3'
pstSFor	GCATGATCGTCATATGAAAGTTATGCGTACCACCGTC
pstBRev	GCATGACGTGGTACCTCAACCGTAACGACCGGTGAT
A147E	GCTGGACGCGTTCGAGCGGATGGACATTG
A147E antisense	CAATGTCCATCCGCTCGAACGCGTCCAGC
A147K	CGTGCTGGACGCGTTCAAGCGGATGGACATTGAC
A147K antisense	GTCAATGTCCATCCGCTTGAACGCGTCCAGCACG
R148A	TGGACGCGTTCGCGGCGATGGACATTGACG
R148A antisense	CGTCAATGTCCATCGCCGCGAACGCGTCCA
R148E	CTGGACGCGTTCGCGGAGATGGACATTGACGA
R148E antisense	TCGTCAATGTCCATCTCCGCGAACGCGTCCAG
D150A	GTTCGCGCGGATGGCCATTGACGAAGCGG
D150A antisense	CCGCTTCGTCAATGGCCATCCGCGCGAAC
109 SphI	CGCGGAGTCGCTGCCCGCATGCGTGGTGCTGACCACGG
109 SphI antisense	CCGTGGTCAGCACCACGCATGCGGGCAGCGACTCCGCG
111 SphI	AGTCGCTGCCCGACGCGGCATGCCTGACCACGGAAGAGGG
111 SphI antisense	CCCTCTTCCGTGGTCAGGCATGCCGCGTCGGGCAGCGACT
113 SphI	GCCCGACGCGGTGGTGGCATGCACGGAAGAGGGCGGTA
113 SphI antisense	TACCGCCCTCTTCCGTGCATGCCACCACCGCGTCGGGC
115 SphI	CGACGCGGTGGTGCTGACCGCATGCGAGGGCGGTATTTTCTGGT
115 SphI antisense	ACCAGAAAATACCGCCCTCGCATGCGGTCAGCACCACCGCGTCG
117 SphI	CGGTGGTGCTGACCACGGAAGCATGCGGTATTTTCTGGT
117 SphI antisense	ACCAGAAAATACCGCATGCTTCCGTGGTCAGCACCACCG
119 SphI	GCTGACCACGGAAGAGGGCGCATGCTTCTGGTGTAACGGTCTGG
119 SphI antisense	CCAGACCGTTACACCAGAAGCATGCGCCCTCTTCCGTGGTCAGC
121 SphI	CCACGGAAGAGGGCGGTATTGCATGCTGTAACGGTCTGGCGC
121 SphI antisense	GCGCCAGACCGTTACAGCATGCAATACCGCCCTCTTCCGTGG
123 SphI	CAAGAATTTGTTGCGCCAGACCGCATGCCCAGAAAATACCGCCCTCTTCCG
123 SphI antisense	CGGAAGAGGGCGGTATTTTCTGGGCATGCGGTCTGGCGCAACAAATTCTTG
125 SphI	CGCAAACCAAGAATTTGTTGCGCGCATGCGTTACACCAGAAAATACCGCCCTC
125 SphI antisense	GAGGGCGGTATTTTCTGGTGTAACGCATGCGCGCAACAAATTCTTGGTTTGCG
127 SphI	TATTTTCTGGTGTAACGGTCTGGCATGCCAAATTCTTGGTTTGCGCTGGC
127 SphI antisense	GCCAGCGCAAACCAAGAATTTGGCATGCCAGACCGTTACACCAGAAAATA
129 SphI	GTGTAACGGTCTGGCGCAAGCATGCCTTGGTTTGCGCTGGCCGG
129 SphI antisense	CCGGCCAGCGCAAACCAAGGCATGCTTGCGCCAGACCGTTACAC
131 SphI	GGTGTAACGGTCTGGCGCAACAAATTGCATGCTTGCGCTGGCCGGA
131 SphI antisense	TCCGGCCAGCGCAAGCATGCAATTTGTTGCGCCAGACCGTTACACC
133 SphI	GTAACGGTCTGGCGCAACAAATTCTTGGTGCATGCTGGCCGGAAGATAAC
133 SphI antisense	GTTATCTTCCGGCCAGCATGCACCAAGAATTTGTTGCGCCAGACCGTTAC
135 SphI	GGCGCAACAAATTCTTGGTTTGCGCGCATGCGAAGATAACGGGCAGAACATCCTTA
135 SphI antisense	TAAGGATGTTCTGCCCGTTATCTTCGCATGCGCGCAAACCAAGAATTTGTTGCGCC
137 SphI	ATTCTTGGTTTGCGCTGGCCGGCATGCAACGGGCAGAACATCCTTAAC
137 SphI antisense	GTTAAGGATGTTCTGCCCGTTGCATGCCGGCCAGCGCAAACCAAGAAT
139 SphI	CTTGGTTTGCGCTGGCCGGAAGATGCATGCCAGAACATCCTTAACCTACTG

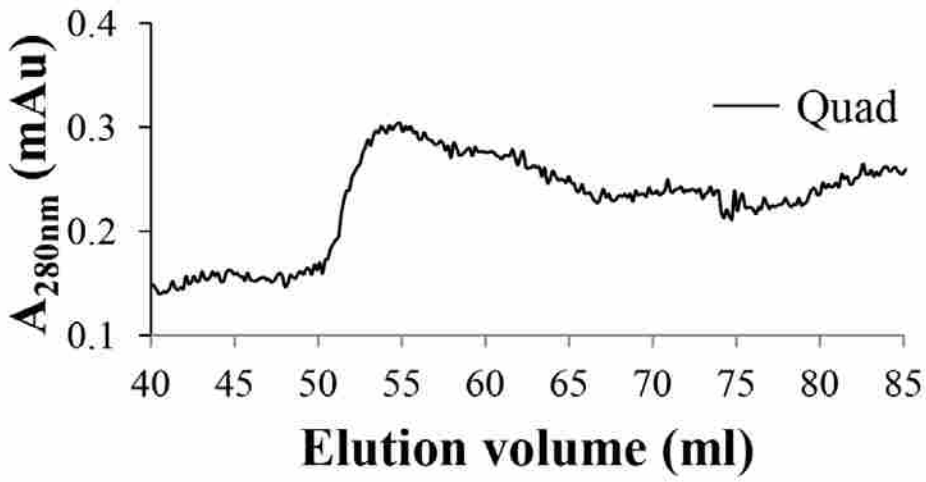


Figure +S.1 PhoU-His Quad is a dimer.
E. coli PhoU-His E100A, R101A, E200A, R101A (Quad) was purified and run on a gel filtration column similar to Fig. 4A and elutes as at a similar size.

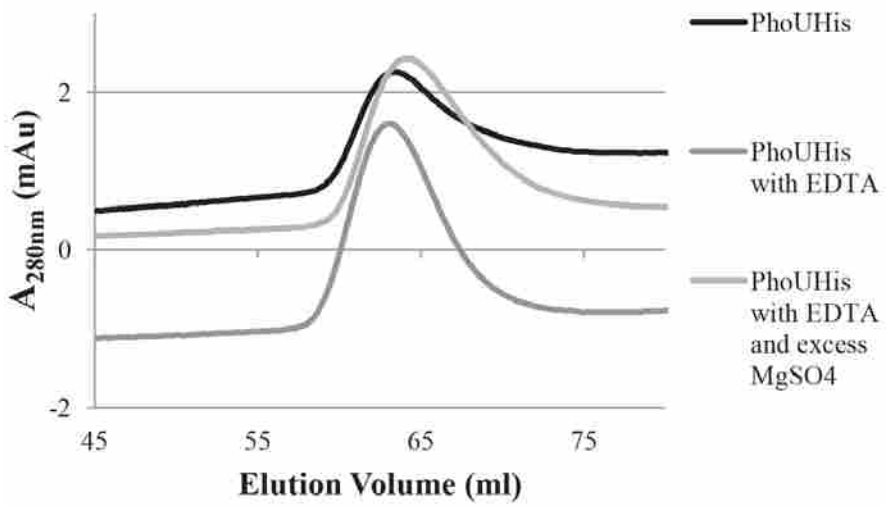


Figure S.2 Metal binding does not affect PhoU dimerization.
For analysis, 0.6 ml of purified *E. coli* PhoU-His (0.65 mg/ml) was treated with 2 µl of 0.5 M EDTA, or EDTA and 20 µl of 1 M $MgSO_4$ Then, untreated PhoU-His and the two treated samples were run on a gel filtration column. We see that all of the samples eluted at similar volumes.

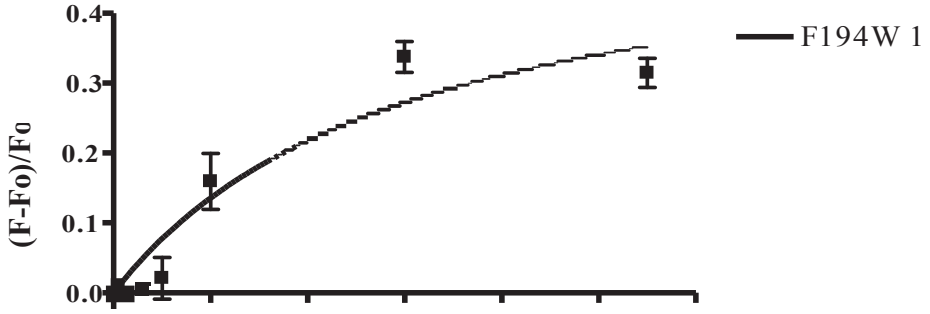
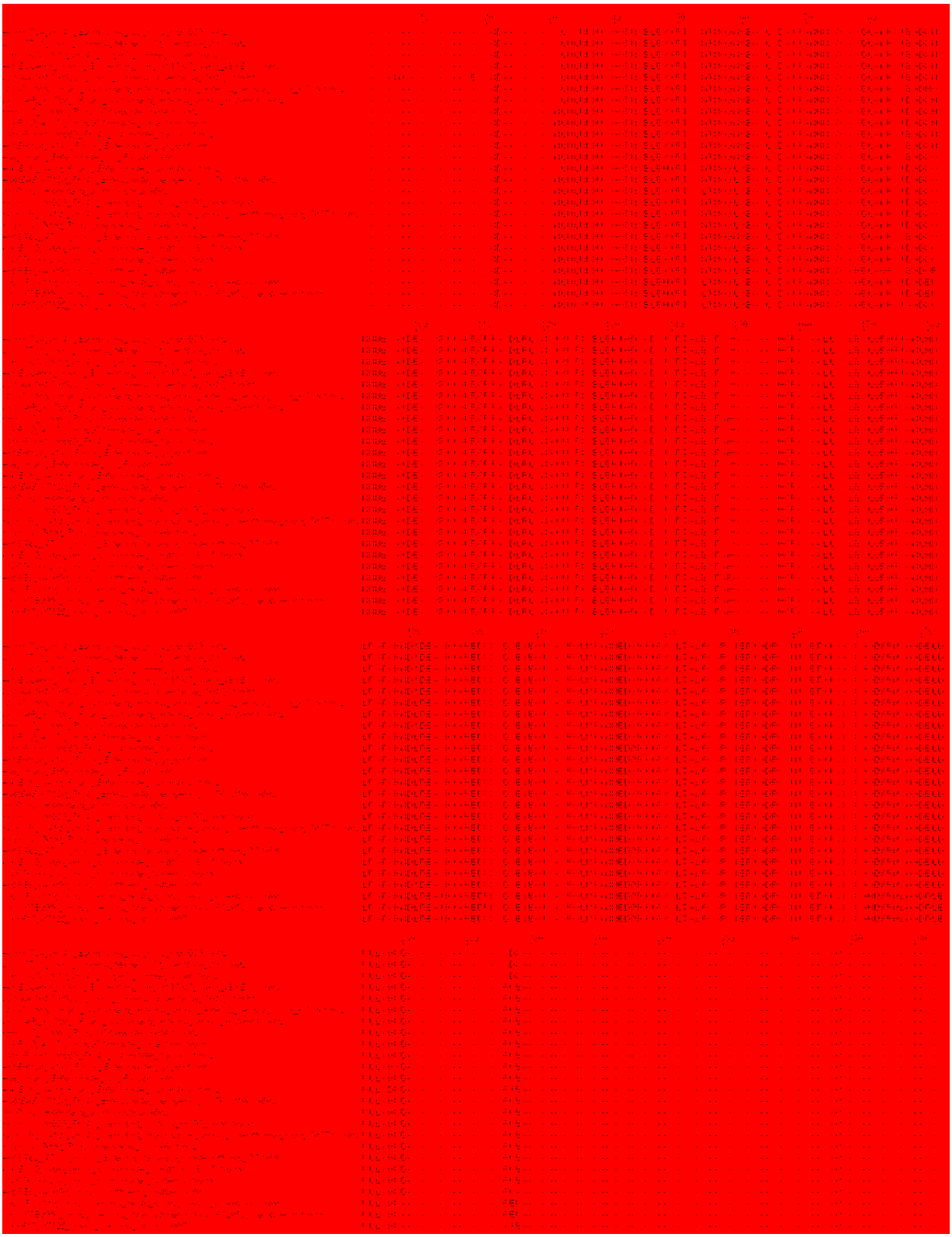
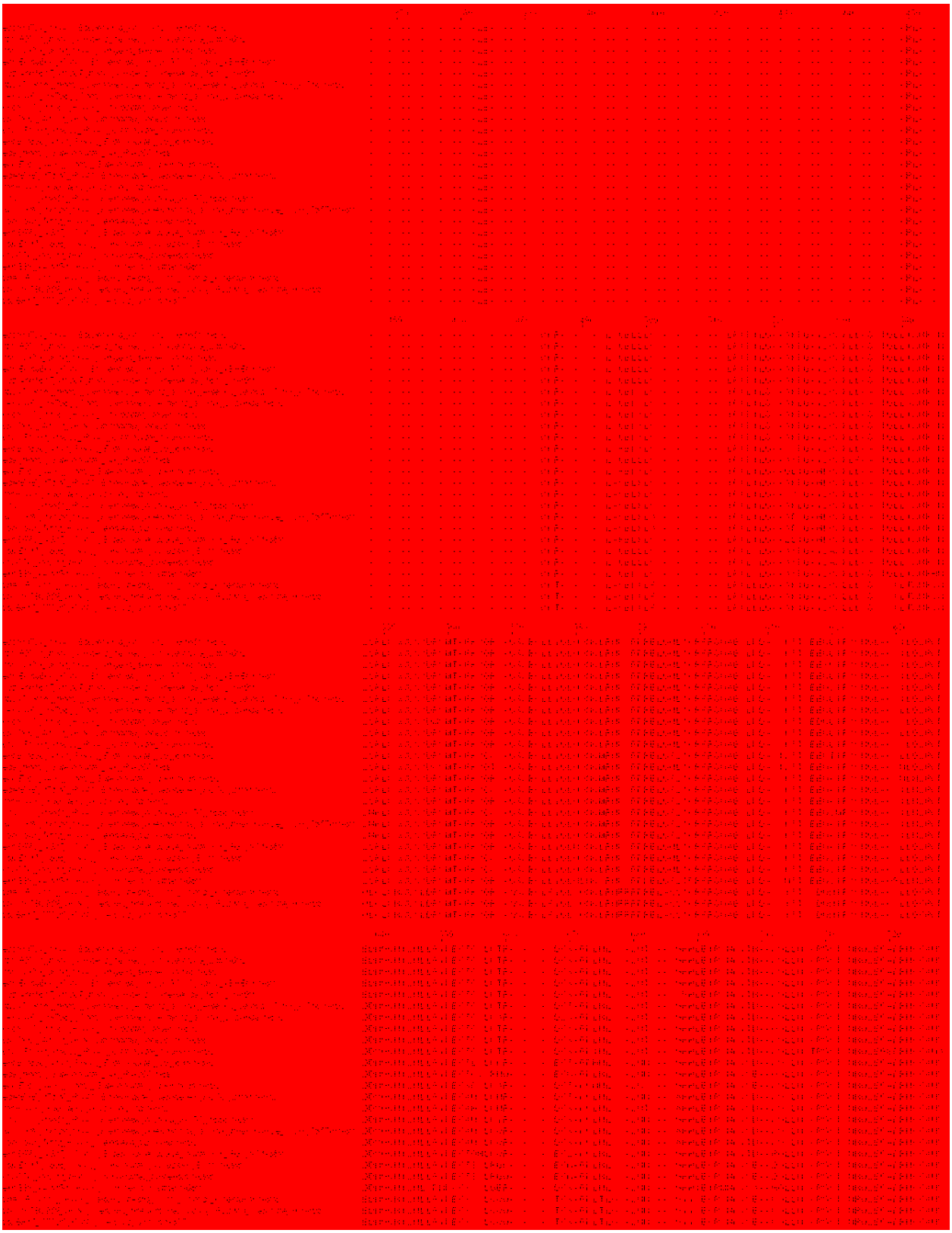


Figure S.3 PhoU F194W binds magnesium.
Scans of PhoU F194W with addition of $MgCl_2$ were performed similar to those shown in Figure 2.5. The mean change in fluorescence between 345 nm to 355 nm was plotted and a binding curve was fit to the data (error bars represent +/- standard error n = 3).





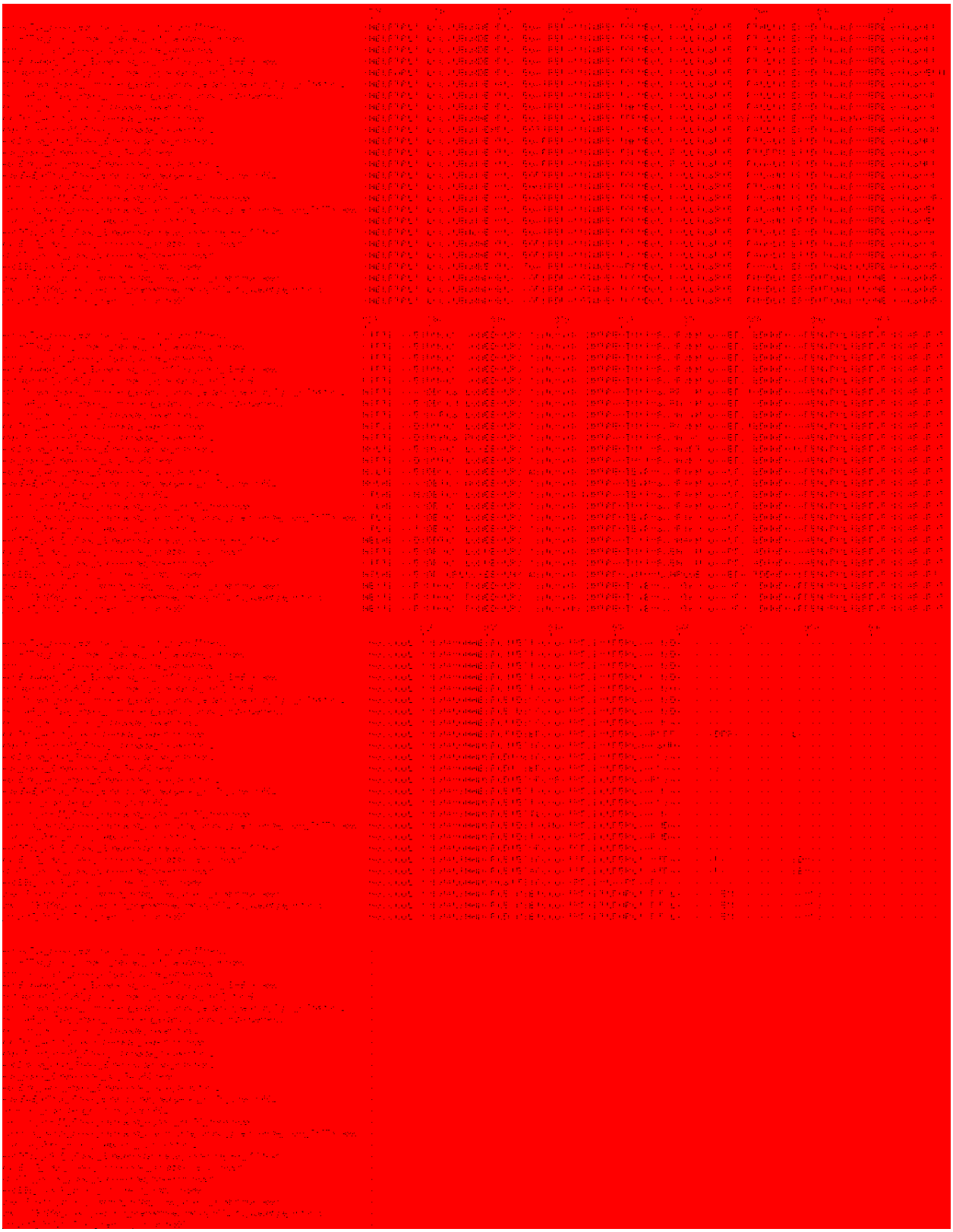
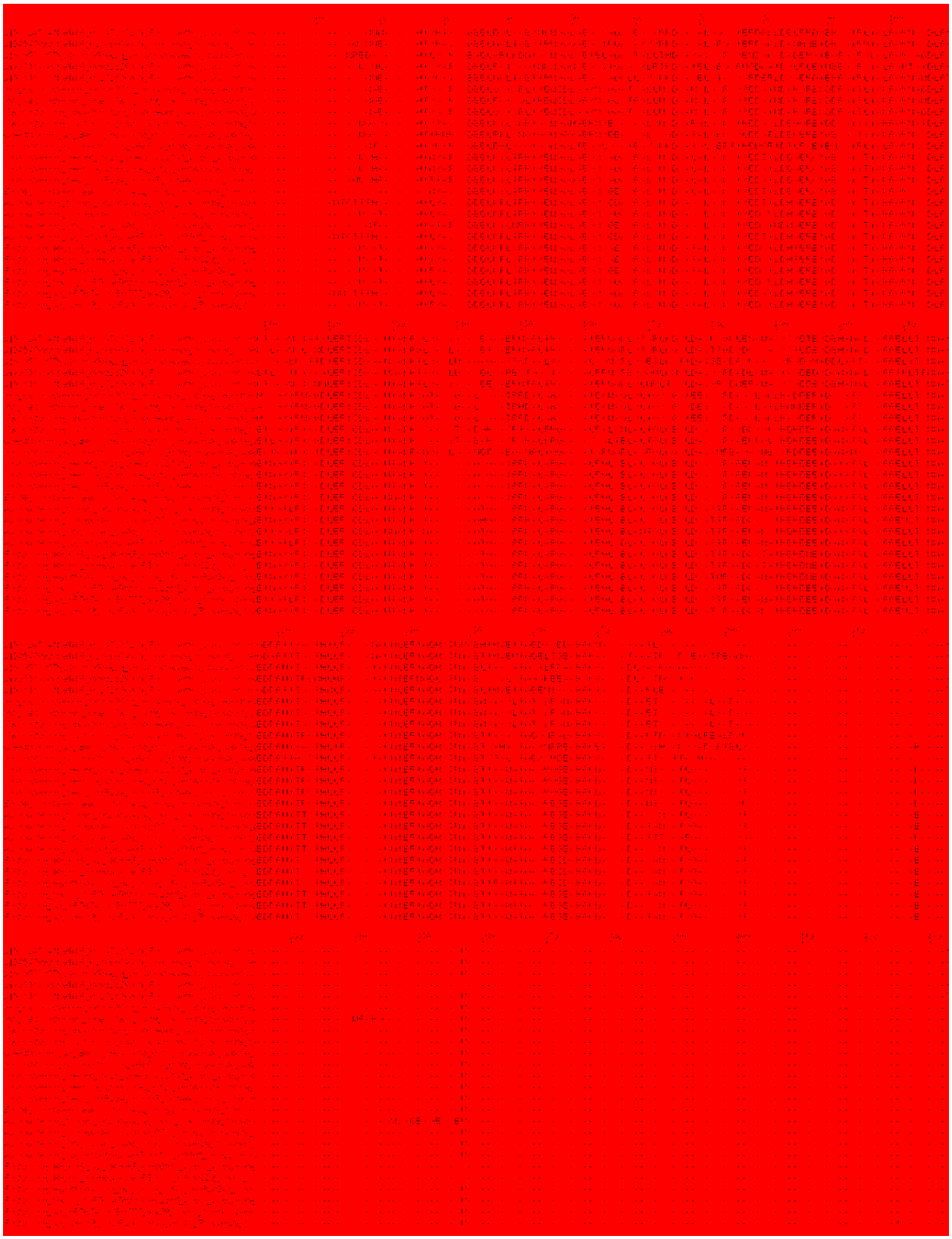
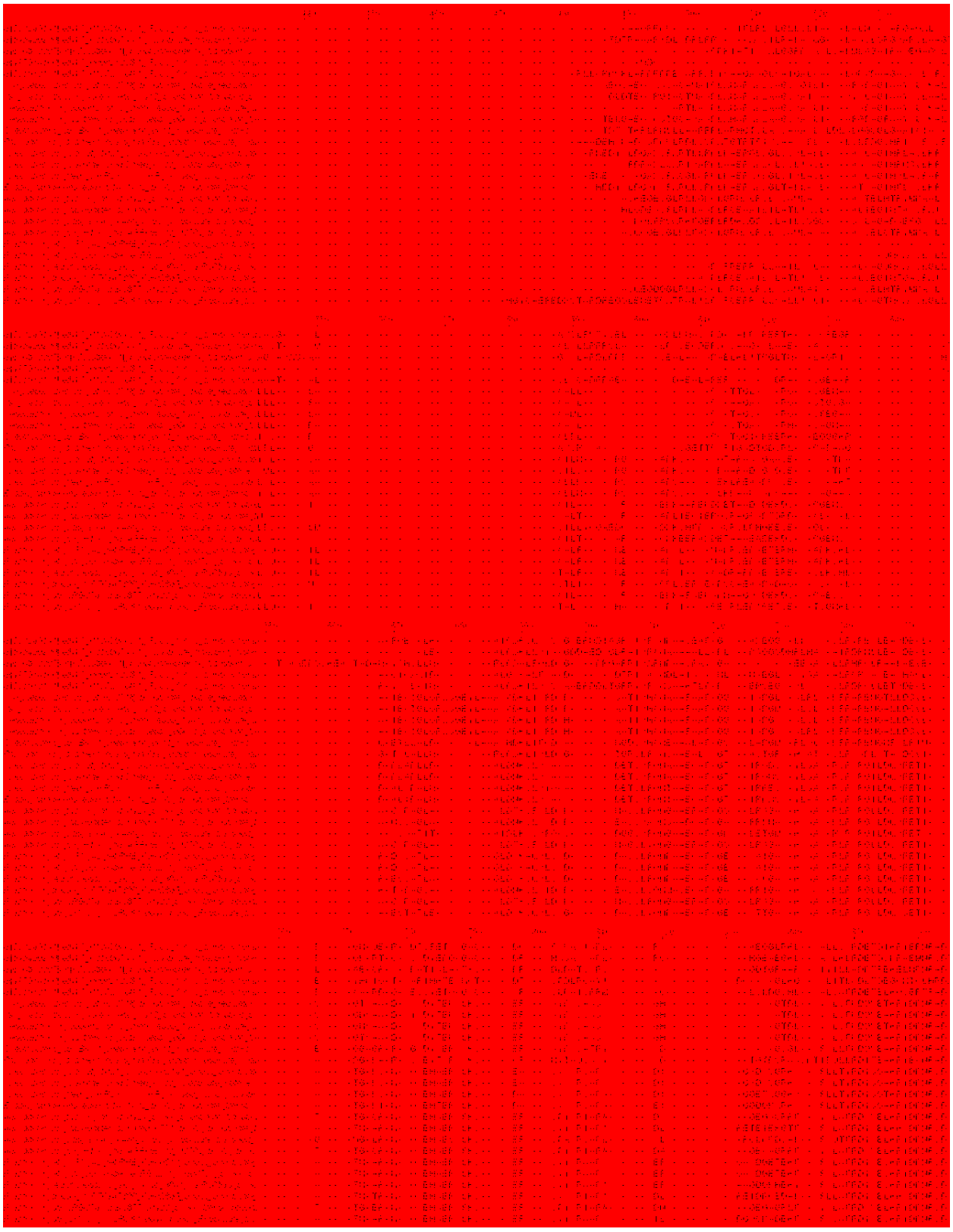
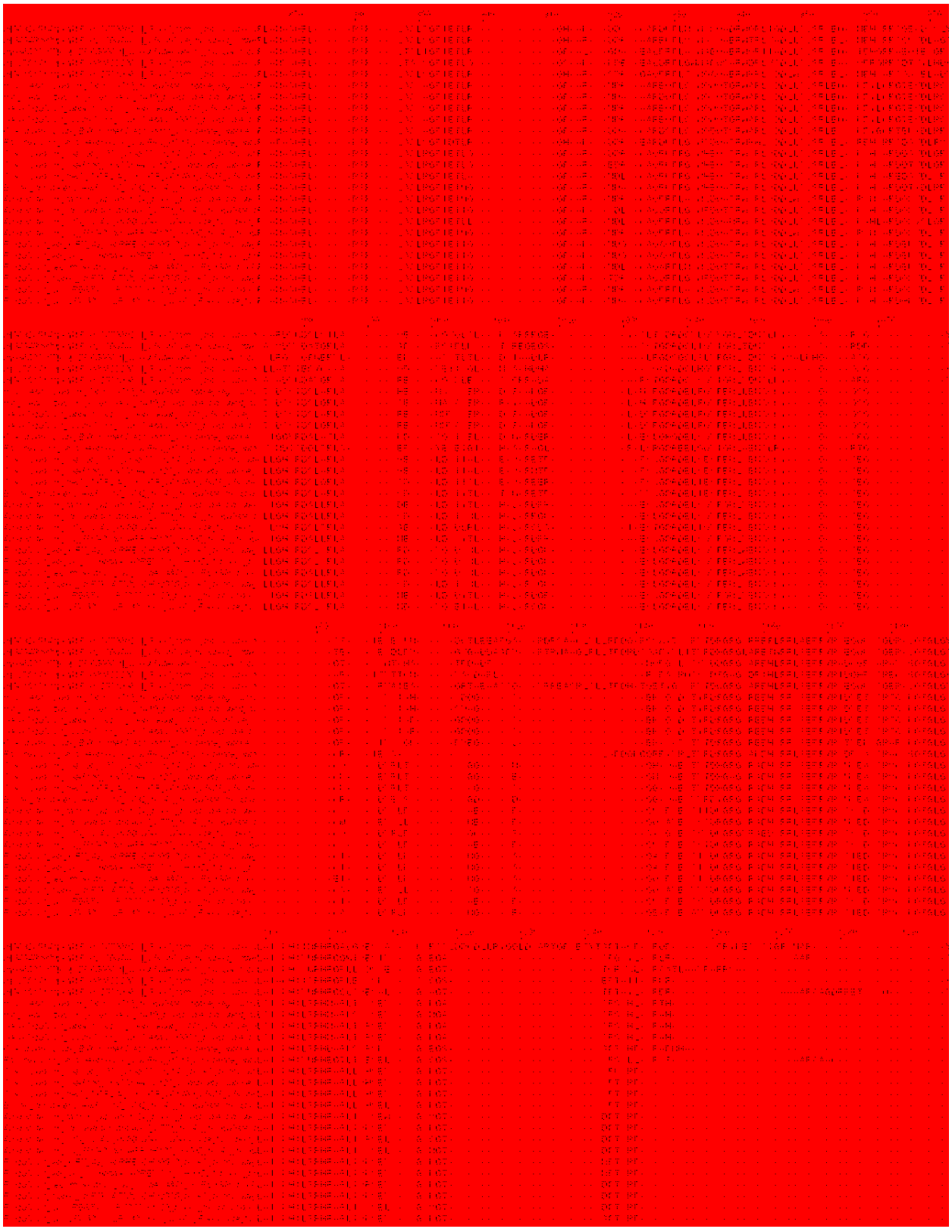


Figure S.4 Alignment of *phoU* and *phoR* genes from gammaproteobacterial species
This shows the alignment of the gammaproteobacteria gene sequences that we used for our initial DCA. The first 25 species' genes are shown.







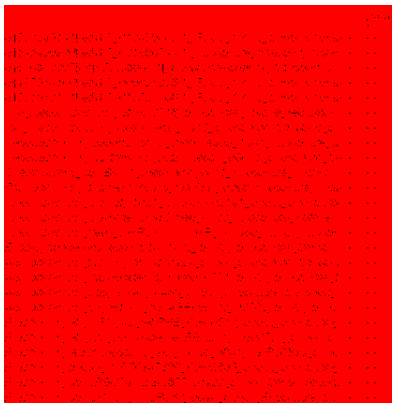


Figure S.5 Alignment of *phoU* and *phoR* genes from species across the proteobacteria phylogenetic group This shows the alignment of the proteobacteria gene sequences that we used for DCA. The first 25 species' genes are shown

REFERENCES

1. **Gardner SG, Johns KD, Tanner R, McCleary WR.** 2014. The PhoU protein from *Escherichia coli* interacts with PhoR, PstB, and metals to form a phosphate-signaling complex at the membrane. J Bacteriol **196:**1741-1752.
2. **Wanner BL.** 1993. Gene regulation by phosphate in enteric bacteria. J. Cell. Biochem. **51:**47-54.
3. **Hsieh YJ, Wanner BL.** 2010. Global regulation by the seven-component Pi signaling system. Curr Opin Microbiol **13:**198-203.
4. **Stock AM, Robinson VL, Goudreau PN.** 2000. Two-component signal transduction. Annu. Rev. Biochem. **69:**183-215.
5. **Li Y, Zhang Y.** 2007. PhoU is a persistence switch involved in persister formation and tolerance to multiple antibiotics and stresses in *Escherichia coli*. Antimicrob Agents Chemother **51:**2092-2099.
6. **Lamarche MG, Wanner BL, Crepin S, Harel J.** 2008. The phosphate regulon and bacterial virulence: a regulatory network connecting phosphate homeostasis and pathogenesis. FEMS Microbiol Rev **32:**461-473.
7. **Shi W, Zhang Y.** 2010. PhoY2 but not PhoY1 is the PhoU homologue involved in persisters in *Mycobacterium tuberculosis*. J Antimicrob Chemother **65:**1237-1242.
8. **Buckles EL, Wang X, Lockatell CV, Johnson DE, Donnenberg MS.** 2006. PhoU enhances the ability of extraintestinal pathogenic *Escherichia coli* strain CFT073 to colonize the murine urinary tract. Microbiology **152:**153-160.

9. **Chekabab SM, Harel J, Dozois CM.** 2014. Interplay between genetic regulation of phosphate homeostasis and bacterial virulence. Virulence **5**.

10. **VanBougelen RA, Olson, E. R., Wanner, B. L., and Neidhardt, F. C.** 1996. Global analysis of proteins synthesized during phosphorus restriction in *Escherichia coli*. J. Bacteriol. **178:**4344-4366.

11. **Wanner BL.** 1996. Phosphorus assimilation and control of the phosphate regulon, p. 1357-1381. *In* Neidhardt FC, III RC, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umbarger HE (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology. American Society for Microbiology, Washington D. C.

12. **Gao R, Stock AM.** 2009. Biological insights from structures of two-component proteins. Annual review of microbiology **63:**133-154.

13. **Szurmant H, White RA, Hoch JA.** 2007. Sensor complexes regulating two-component signal transduction. Current opinion in structural biology **17:**706-715.

14. **West AH, Stock AM.** 2001. Histidine kinases and response regulator proteins in two-component signaling systems. Trends in biochemical sciences **26:**369-376.

15. **Ferris HU, Coles M, Lupas AN, Hartmann MD.** 2014. Crystallographic snapshot of the Escherichia coli EnvZ histidine kinase in an active conformation. Journal of structural biology **186:**376-379.

16. **Mascher T.** 2014. Bacterial (intramembrane-sensing) histidine kinases: signal transfer rather than stimulus perception. Trends in microbiology **22:**559-565.

17. **Ferris HU, Dunin-Horkawicz S, Hornig N, Hulko M, Martin J, Schultz JE, Zeth K, Lupas AN, Coles M.** 2012. Mechanism of regulation of receptor histidine kinases. Structure **20:**56-66.

18. **Moglich A, Ayers RA, Moffat K.** 2009. Structure and signaling mechanism of Per-ARNT-Sim domains. Structure **17:**1282-1294.

19. **Gutu AD, Wayne KJ, Sham LT, Winkler ME.** 2010. Kinetic characterization of the WalRKSpn (VicRK) two-component system of *Streptococcus pneumoniae*: dependence of WalKSpn (VicK) phosphatase activity on its PAS domain. J Bacteriol **192:**2346-2358.

20. **Botella E, Devine S, Hubner S, Salzberg L, Gale R, Brown E, Link H, Sauer U, Codee J, Noone D, Devine K.** 2014. PhoR autokinase activity is controlled by an intermediate in wall teichoic acid metabolism that is sensed by the intracellular PAS domain during the PhoPR-mediated phosphate limitation response of *Bacillus subtilis*. Molecular microbiology **doi: 10.1111/mmi.12833. [Epub ahead of print]**.

21. **Dago AE, Schug A, Procaccini A, Hoch JA, Weigt M, Szurmant H.** 2012. Structural basis of histidine kinase autophosphorylation deduced by integrating genomics, molecular dynamics, and mutagenesis. Proceedings of the National Academy of Sciences of the United States of America **109:**E1733-1742.

22. **Casino P, Rubio V, Marina A.** 2010. The mechanism of signal transduction by two-component systems. Current opinion in structural biology **20:**763-771.

23. **Huynh TN, Noriega CE, Stewart V.** 2013. Missense substitutions reflecting regulatory control of transmitter phosphatase activity in two-component signalling. Molecular microbiology **88:**459-472.

24. **Huynh TN, Stewart V.** 2011. Negative control in two-component signal transduction by transmitter phosphatase activity. Molecular microbiology **82:**275-286.

25. **Kenney LJ.** 2010. How important is the phosphatase activity of sensor kinases? Curr Opin Microbiol **13:**168-176.

26. **Podgornaia AI, Casino P, Marina A, Laub MT.** 2013. Structural basis of a rationally rewired protein-protein interface critical to bacterial signaling. Structure **21:**1636-1647.

27. **Mizuno T.** 1997. Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of Escherichia coli. DNA research : an international journal for rapid publication of reports on genes and genomes **4:**161-168.

28. **Nakata A, Amemura M, Shinagawa H.** 1984. Regulation of the phosphate regulon in *Escherichia coli* K-12: regulation of the negative regulatory gene *phoU* and identification of the gene product. J. Bacteriol. **159:**979-985.

29. **Dutta R, L. Qin, and M. Inouye.** 1999. Histidine kinases: diversity of domain organization. Mol. Microbiol. **34:**633-640.

30. **Yamada M, Makino K, Shinagawa H, Nakata A.** 1990. Regulation of the phosphate regulon of *Escherichia coli*: properties of *phoR* deletion mutants and subcellular localization of PhoR protein. Mol. Gen. Genet. **220:**366-372.

31. **Scholten M, Tommassen J.** 1993. Topology of the PhoR protein of *Escherichia coli* and functional analysis of internal deletion mutants. Molecular microbiology **8:**269-275.

32. **Carmany DO, Hollingsworth, K., and McCleary, W. R.** 2003. Genetic and biochemical studies of phosphate activity of PhoR. J. Bacteriol. **185:**1112-1115.

33. **Taylor BL, Zhulin IB.** 1999. PAS domains: internal sensors of oxygen, redox potential, and light. Microbiology and molecular biology reviews : MMBR **63:**479-506.

34. **Wang C, Sang J, Wang J, Su M, Downey JS, Wu Q, Wang S, Cai Y, Xu X, Wu J, Senadheera DB, Cvitkovitch DG, Chen L, Goodman SD, Han A.** 2013. Mechanistic insights revealed by the crystal structure of a histidine kinase with signal transducer and sensor domains. PLoS biology **11:**e1001493.

35. **Kelley LA, Sternberg MJ.** 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nature protocols **4:**363-371.

36. **Shi L, Hulett FM.** 1999. The cytoplasmic kinase domain of PhoR is sufficient for the low phosphate-inducible expression of pho regulon genes in *Bacillus subtilis*. Molecular microbiology **31:**211-222.

37. **Chang C, Tesar C, Gu M, Babnigg G, Joachimiak A, Pokkuluri PR, Szurmant H, Schiffer M.** 2010. Extracytoplasmic PAS-like domains are common in signal transduction proteins. J Bacteriol **192:**1156-1159.

38. **Ellison DW, and McCleary, W. R.** 2000. The unphosphorylated receiver domain of PhoB silences the activity of its output domain. J. Bacteriol. **182:**6592-6597.

39. **Kim SK, Makino, K., Amemura, M., Nakata, A., and Shinagawa, H.** 1995. Mutational analysis of the role of the first helix of region 4.2 of the sigma 70 subunit of *Escherichia coli* RNA polymerase in transcriptional activation by activator protein PhoB. Mol. Gen. Genet. **248:**1-8.

40. **Blanco AG, Sola M, Gomis-Ruth FX, Coll M.** 2002. Tandem DNA recognition by PhoB, a two-component signal transduction transcriptional activator. Structure. **10:**701-713.

41. **Wanner BL.** 1995. Signal transduction and cross regulation in the *Escherichia coli* phosphate regulon by PhoR, CreC, and Acetyl Phosphate. ASM Press, Washington, D. C.

42. **McCleary WR.** 1996. The activation of PhoB by acetylphosphate. Molecular microbiology **20:**1155-1163.

43. **Wanner BL.** 1987. Control of *phoR*-dependent bacterial alkaline phosphatase clonal variation by the *phoM* region. J. Bacteriol. **169:**900-903.

44. **Zhou L, Gregori, G., Blackman, J. M., Robinson, J. P., and Wanner, B. L.** 2005. Stochastic activation of the response regulator PhoB by noncognate histidine kinases. Journal of Integrative Bioinformatics**:**11-24.

45. **Lewis VG, Ween MP, McDevitt CA.** 2012. The role of ATP-binding cassette transporters in bacterial pathogenicity. Protoplasma **249:**919-942.

46. **Chen J.** 2013. Molecular mechanism of the Escherichia coli maltose transporter. Current opinion in structural biology **23:**492-498.

47. **Li CH, Yang YX, Su JG, Liu B, Tan JJ, Zhang XY, Wang CX.** 2014. Allosteric transitions of the maltose transporter studied by an elastic network model. Biopolymers **101:**758-768.

48. **Bao H, Duong F.** 2014. Nucleotide-free MalK drives the transition of the maltose transporter to the inward-facing conformation. J Biol Chem **289:**9844-9851.

49. **Hayashi T, Chiba S, Kaneta Y, Furuta T, Sakurai M.** 2014. ATP-Induced Conformational Changes of Nucleotide-Binding Domains in an ABC Transporter. Importance of the Water-Mediated Entropic Force. The journal of physical chemistry. B.

50. **Heuveling J, Frochaux V, Ziomkowska J, Wawrzinek R, Wessig P, Herrmann A, Schneider E.** 2014. Conformational changes of the bacterial type I ATP-binding cassette importer HisQMP2 at distinct steps of the catalytic cycle. Biochimica et biophysica acta **1838:**106-116.

51. **Chen S, Oldham ML, Davidson AL, Chen J.** 2013. Carbon catabolite repression of the maltose transporter revealed by X-ray crystallography. Nature **499:**364-368.

52. **Vastermark A, Saier MH, Jr.** 2014. The involvement of transport proteins in transcriptional and metabolic regulation. Curr Opin Microbiol **18:**8-15.

53. **Surin BP, Rosenberg H, Cox GB.** 1985. Phosphate-specific transport system of *Escherichia coli*: nucleotide sequence and gene-polypeptide relationships. J. Bacteriol. **161:**189-198.

54. **Surin BP, Jans DA, Fimmel AL, Shaw DC, Cox GB, Rosenberg H.** 1984. Structural gene for the phosphate-repressible phosphate-binding protein of Escherichia coli has its own promoter: complete nucleotide sequence of the phoS gene. J Bacteriol **157:**772-778.

55. **Magota K, Otsuji N, Miki T, Horiuchi T, Tsunasawa S, Kondo J, Sakiyama F, Amemura M, Morita T, Shinagawa H, et al.** 1984. Nucleotide sequence of the phoS gene, the structural gene for the phosphate-binding protein of Escherichia coli. J Bacteriol **157:**909-917.

56. **Gebhard S, Tran SL, Cook GM.** 2006. The Phn system of Mycobacterium smegmatis: a second high-affinity ABC-transporter for phosphate. Microbiology **152:**3453-3465.

57. **Harris RM, Webb DC, Howitt SM, Cox GB.** 2001. Characterization of PitA and PitB from *Escherichia coli*. J. Bacteriol. **183:**5008-5014.

58. **Hoffer SM, Schoondermark P, van Veen HW, Tommassen J.** 2001. Activation by gene amplification of *pitB*, encoding a third phosphate transporter of *Escherichia coli* K-12. J. Bacteriol. **183:**4659-4663.

59. **Hoffer SM, Tommassen J.** 2001. The phosphate-binding protein of *Escherichia coli* is not essential for P(i)-regulated expression of the pho regulon. J. Bacteriol. **183:**5768-5771.

60. **McCarthy S, Ai C, Wheaton G, Tevatia R, Eckrich V, Kelly R, Blum P.** 2014. Role of an Archaeal PitA Transporter in the Copper and Arsenic Resistance of Metallosphaera sedula, an Extreme Thermoacidophile. J Bacteriol **196:**3562-3570.

61. **Haldimann A, Daniels LL, Wanner BL.** 1998. Use of new methods for construction of tightly regulated arabinose and rhamnose promoter fusions in studies of the *Escherichia coli* phosphate regulon. J. Bacteriol. **180:**1277-1286.

62. **Steed PM, Wanner BL.** 1993. Use of the rep technique for allele replacement to construct mutants with deletions of the *pstSCAB-phoU* operon: evidence of a new role for the PhoU protein in the phosphate regulon. J. Bacteriol. **175:**6797-6809.

63. **Rice CD, Pollard JE, Lewis ZT, McCleary WR.** 2009. Employment of a promoter-swapping technique shows that PhoU modulates the activity of the PstSCAB2 ABC transporter in *Escherichia coli*. Appl Environ Microbiol **75:**573-582.

64. **Chakraborty S, Sivaraman J, Leung KY, Mok YK.** 2011. Two-component PhoB-PhoR regulatory system and ferric uptake regulator sense phosphate and iron to control virulence genes in type III and VI secretion systems of *Edwardsiella tarda*. J Biol Chem **286:**39417-39430.

65. **Oganesyan V, Oganesyan N, Adams PD, Jancarik J, Yokota HA, Kim R, Kim SH.** 2005. Crystal structure of the "PhoU-like" phosphate uptake regulator from *Aquifex aeolicus*. J Bacteriol **187:**4238-4244.

66. **Liu J, Lou Y, Yokota H, Adams PD, Kim R, Kim SH.** 2005. Crystal structure of a PhoU protein homologue: A new class of metalloprotein containing multinuclear iron clusters. J. Biol. Chem. **280:**15960-15966.

67. **Lee SJ, Park YS, Kim SJ, Lee BJ, Suh SW.** 2014. Crystal structure of PhoU from *Pseudomonas aeruginosa*, a negative regulator of the Pho regulon. Journal of structural biology **188:**22-29.

68. **Luttmann D, Gopel Y, Gorke B.** 2012. The phosphotransferase protein EIIA(Ntr) modulates the phosphate starvation response through interaction with histidine kinase PhoR in Escherichia coli. Molecular microbiology **86:**96-110.

69. **Baek JH, Lee SY.** 2007. Transcriptome analysis of phosphate starvation response in Escherichia coli. J Microbiol Biotechnol **17:**244-252.

70. **Baek JH, Kang YJ, Lee SY.** 2007. Transcript and protein level analyses of the interactions among PhoB, PhoR, PhoU and CreC in response to phosphate starvation in *Escherichia coli*. FEMS Microbiol Lett **277:**254-259.

71. **Muda M, Rao NN, Torriani A.** 1992. Role of PhoU in phosphate transport and alkaline phosphatase regulation. J. Bacteriol. **174:**8057-8064.

72. **Wanner BL.** 1986. Bacterial alkaline phosphatase clonal variation in some *Escherichia col*i K-12 *phoR* mutant strains. J. Bacteriol. **168:**1366-1371.

73. **Torriani A, Rothman F.** 1961. Mutants of *Escherichia coli* constitutive for alkaline phosphatase. J Bacteriol **81:**835-836.

74. **Horiuchi T, Horiuchi S, Mizuno D.** 1959. A possible negative feedback phenomenon controlling formation of alkaline phosphomonoesterase in *Escherichia coli*. Nature **183:**1529-1530.

75. **Wuttge S, Bommer M, Jager F, Martins BM, Jacob S, Licht A, Scheffel F, Dobbek H, Schneider E.** 2012. Determinants of substrate specificity and biochemical properties of the sn-glycerol-3-phosphate ATP binding cassette transporter (UgpB-AEC2) of Escherichia coli. Molecular microbiology **86:**908-920.

76. **Crepin S, Chekabab SM, Le Bihan G, Bertrand N, Dozois CM, Harel J.** 2011. The Pho regulon and the pathogenesis of *Escherichia coli*. Vet Microbiol **153:**82-88.

77. **Daigle F, Fairbrother JM, Harel J.** 1995. Identification of a mutation in the *pst-phoU* operon that reduces pathogenicity of an *Escherichia coli* strain causing septicemia in pigs. Infect. Immun. **63:**4924-4927.

78. **Cheng C, Tennant SM, Azzopardi KI, Bennett-Wood V, Hartland EL, Robins-Browne RM, Tauschek M.** 2009. Contribution of the *pst-phoU* operon to cell adherence by atypical enteropathogenic *Escherichia coli* and virulence of *Citrobacter rodentium*. Infect Immun **77:**1936-1944.

79. **Lee K, Jung J, Kim K, Bae D, Lim D.** 2009. Overexpression of outer membrane protein OprT and increase of membrane permeability in phoU mutant of toluene-tolerant bacterium Pseudomonas putida GM730. J Microbiol **47:**557-562.

80. **Rao NN, Torriani A.** 1990. Molecular aspects of phosphate transport in Escherichia coli. Molecular microbiology **4:**1083-1090.

81. **Torriani-Gorini A.** 1987. The birth of the Pho regulon, p. 3-11. *In* Torriani-Gorini Aea (ed.), Phosphate metabolism and cellular regulation in mircroorganisms. Am. Soc. Microbiol.

82. **Makino K, Shinagawa H, Amemura M, Kawamoto T, Yamada M, Nakata A.** 1989. Signal transduction in the phosphate regulon of *Escherichia coli* involves phosphotransfer between PhoR and PhoB proteins. J. Mol. Biol. **210:**551-559.

83. **Bachhawat P, Swapna GV, Montelione GT, Stock AM.** 2005. Mechanism of activation for transcription factor PhoB suggested by different modes of dimerization in the inactive and active states. Structure **13:**1353-1363.

84. **Makino K, Shinagawa H, Amemura M, Nakata A.** 1986. Nucleotide sequence of the *phoR* gene, a regulatory gene for the phosphate regulon of *Escherichia coli*. J. Mol. Biol. **192:**549-556.

85. **Makino K, Amemura M, Kim SK, Nakata A, Shinagawa H.** 1993. Role of the sigma 70 subunit of RNA polymerase in transcriptional activation by activator protein PhoB in Escherichia coli. Genes & development **7:**149-160.

86. **Etzkorn M, Kneuper H, Dunnwald P, Vijayan V, Kramer J, Griesinger C, Becker S, Unden G, Baldus M.** 2008. Plasticity of the PAS domain and a potential role for signal transduction in the histidine kinase DcuS. Nature structural & molecular biology **15:**1031-1039.

87. **Rao NN, Roberts MF, Torriani A, Yashphe J.** 1993. Effect of glpT and glpD mutations on expression of the phoA gene in Escherichia coli. J Bacteriol **175:**74-79.

88. **Shulman RG, Brown TR, Ugurbil K, Ogawa S, Cohen SM, den Hollander JA.** 1979. Cellular applications of 31P and 13C nuclear magnetic resonance. Science **205:**160-166.

89. **Peterson CN, Mandel MJ, Silhavy TJ.** 2005. Escherichia coli starvation diets: essential nutrients weigh in distinctly. J Bacteriol **187:**7549-7553.

90. **Rees DC, Johnson E, Lewinson O.** 2009. ABC transporters: the power to change. Nature reviews. Molecular cell biology **10:**218-227.

91. **Webb DC, Rosenberg H, Cox GB.** 1992. Mutational analysis of the *Escherichia coli* phosphate-specific transport system, a member of the traffic ATPase (or ABC) family of membrane transporters. A role for proline residues in transmembrane helices. J. Biol. Chem. **267:**24661-24668.

92. **Amemura M, Makino K, Shinagawa H, Kobayashi A, Nakata A.** 1985. Nucleotide sequence of the genes involved in phosphate transport and regulation of the phosphate regulon in Escherichia coli. J Mol Biol **184:**241-250.

93. **Davidson AL, Dassa E, Orelle C, Chen J.** 2008. Structure, function, and evolution of bacterial ATP-binding cassette systems. Microbiology and molecular biology reviews : MMBR **72:**317-364, table of contents.

94. **Surin BP, Dixon NE, Rosenberg H.** 1986. Purification of the PhoU protein, a negative regulator of the pho regulon of *Escherichia coli* K-12. J. Bacteriol. **168:**631-635.

95. **Madej T, Addess KJ, Fong JH, Geer LY, Geer RC, Lanczycki CJ, Liu C, Lu S, Marchler-Bauer A, Panchenko AR, Chen J, Thiessen PA, Wang Y, Zhang D, Bryant SH.** 2012. MMDB: 3D structures and macromolecular interactions. Nucleic acids research **40:**D461-464.

96. **Rao NN, Wang E, Yashphe J, Torriani A.** 1986. Nucleotide pool in pho regulon mutants and alkaline phosphatase synthesis in *Escherichia coli*. J Bacteriol **166:**205-211.

97. **Miller J.** 1992. A short course in bacterial genetics: a laboratory manual and handbook for Escherichia coli and related bacteria. Cold Springs Harbor Laboratory Press, Cold Spring Harbor.

98. **Sambrook J, Fritsch RF, Mantiatis T.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.

99. **Neidhardt FC, P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. J. Bacteriol. **119:**736-747.

100. **Datsenko KA, Wanner. BL.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. **USA 97:**6640-6645.

101. **Yamada M, Makino K, Amemura M, Shinagawa H, Nakata A.** 1989. Regulation of the phosphate regulon of *Escherichia coli*: analysis of mutant *phoB* and *phoR* genes causing different phenotypes. J. Bacteriol. **171:**5601-5606.

102. **Blomfield IC, Vaughn V, Rest RF, Eisenstein BI.** 1991. Allelic exchange in Escherichia coli using the Bacillus subtilis sacB gene and a temperature-sensitive pSC101 replicon. Molecular microbiology **5:**1447-1457.

103. **Buron-Barral MC, Gosink KK, Parkinson JS.** 2006. Loss- and gain-of-function mutations in the F1-HAMP region of the *Escherichia coli* aerotaxis transducer Aer. J Bacteriol **188:**3477-3486.

104. **Studdert CA, Parkinson JS.** 2005. Insights into the organization and dynamics of bacterial chemoreceptor clusters through in vivo crosslinking studies. Proceedings of the National Academy of Sciences of the United States of America **102:**15623-15628.

105. **Battesti A, Bouveret E.** 2012. The bacterial two-hybrid system based on adenylate cyclase reconstitution in *Escherichia coli*. Methods **58:**325-334.

106. **Thibodeau SA, Fang R, Joung JK.** 2004. High-throughput beta-galactosidase assay for bacterial cell-based reporter systems. BioTechniques **36:**410-415.

107. **Griffith KL, Wolf RE, Jr.** 2002. Measuring beta-galactosidase activity in bacteria: cell growth, permeabilization, and enzyme assays in 96-well arrays. Biochemical and biophysical research communications **290:**397-402.

108. **Zundel CJ, D. C. Capenter, and W. R. McCleary.** 1998. Analysis of the conserved acidic residues in the regulatory domain of PhoB. FEBS Lett. **441:**242-246.

109. **Schurdell MS, Woodbury GM, McCleary WR.** 2007. Genetic evidence suggests that the intergenic region between *pstA* and *pstB* plays a role in the regulation of *rpoS* translation during phosphate limitation. J Bacteriol **189:**1150-1153.

110. **Ruiz N, Silhavy TJ.** 2003. Constitutive activation of the *Escherichia coli* Pho regulon upregulates *rpoS* translation in an Hfq-dependent fashion. J Bacteriol **185:**5984-5992.

111. **Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H.** 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Molecular systems biology **2:**2006 0008.

112. **Vivian JT, Callis PR.** 2001. Mechanisms of tryptophan fluorescence shifts in proteins. Biophysical journal **80:**2093-2109.

113. **Anjem A, Varghese S, Imlay JA.** 2009. Manganese import is a key element of the OxyR response to hydrogen peroxide in *Escherichia coli*. Molecular microbiology **72:**844-858.

114. **Waters LS, Sandoval M, Storz G.** 2011. The *Escherichia coli* MntR miniregulon includes genes encoding a small protein and an efflux pump required for manganese homeostasis. J Bacteriol **193:**5887-5897.

115. **Golynskiy MV, Gunderson WA, Hendrich MP, Cohen SM.** 2006. Metal binding studies and EPR spectroscopy of the manganese transport regulator MntR. Biochemistry **45:**15359-15372.

116. **Silver S, Johnseine P, King K.** 1970. Manganese Active Transport in Escherichia coli. J Bacteriol **104:**1299-1306.

117. **Alatossava T, Jutte H, Kuhn A, Kellenberger E.** 1985. Manipulation of intracellular magnesium content in polymyxin B nonapeptide-sensitized Escherichia coli by ionophore A23187. J Bacteriol **162:**413-419.

118. **Lopez C, Sanchez J, Hermida L, Zulueta A, Marquez G.** 2004. Cysteine mediated multimerization of a recombinant dengue E fragment fused to the P64k protein following immobilized metal ion affinity chromatography. Protein expression and purification **34:**176-182.

119. **Winnen B, Anderson E, Cole JL, King GF, Rowland SL.** 2013. Role of the PAS sensor domains in the Bacillus subtilis sporulation kinase KinA. J Bacteriol **195:**2349-2358.

120. **Sousa EH, Tuckerman JR, Gondim AC, Gonzalez G, Gilles-Gonzalez MA.** 2013. Signal transduction and phosphoryl transfer by a FixL hybrid kinase with low oxygen affinity: importance of the vicinal PAS domain and receiver aspartate. Biochemistry **52:**456-465.

121. **Yamada S, Sugimoto H, Kobayashi M, Ohno A, Nakamura H, Shiro Y.** 2009. Structure of PAS-linked histidine kinase and the response regulator complex. Structure **17:**1333-1344.

122. **Echenique JR, Trombe MC.** 2001. Competence repression under oxygen limitation through the two-component MicAB signal-transducing system in Streptococcus pneumoniae and involvement of the PAS domain of MicB. J Bacteriol **183:**4599-4608.

123. **Gerke V, Creutz CE, Moss SE.** 2005. Annexins: linking Ca2+ signalling to membrane dynamics. Nature reviews. Molecular cell biology **6:**449-461.

124. **Swairjo MA, Concha NO, Kaetzel MA, Dedman JR, Seaton BA.** 1995. Ca(2+)-bridging mechanism and phospholipid head group recognition in the membrane-binding protein annexin V. Nature structural biology **2:**968-974.

125. **Titball RW, Naylor CE, Miller J, Moss DS, Basak AK.** 2000. Opening of the active site of Clostridium perfringens alpha-toxin may be triggered by membrane binding. International journal of medical microbiology : IJMM **290:**357-361.

126. **Mikkelsen H, Hui K, Barraud N, Filloux A.** 2013. The pathogenicity island encoded PvrSR/RcsCB regulatory network controls biofilm formation and dispersal in *Pseudomonas aeruginosa* PA14. Molecular microbiology **89:**450-463.

127. **Garen A, Otsuji N.** 1964. Isolation of a Protein Specified by a Regulator Gene. J Mol Biol **8:**841-852.

128. **Zuckier G, Torriani A.** 1981. Genetic and physiological tests of three phosphate-specific transport mutants of *Escherichia coli*. J. Bacteriol. **145:**1249-1256.

129. **Amemura M, Shinagawa H, Makino K, Otsuji N, Nakata A.** 1982. Cloning of and complementation tests with alkaline phosphatase regulatory genes (*phoS* and *phoT*) of *Escherichia coli*. J. Bacteriol. **152:**692-701.

130. **Dwyer RS, Ricci DP, Colwell LJ, Silhavy TJ, Wingreen NS.** 2013. Predicting functionally informative mutations in Escherichia coli BamA using evolutionary covariance analysis. Genetics **195:**443-455.

131. **Liu Y, Eyal E, Bahar I.** 2008. Analysis of correlated mutations in HIV-1 protease using spectral clustering. Bioinformatics **24:**1243-1250.

132. **Liu Y, Bahar I.** 2012. Sequence evolution correlates with structural dynamics. Mol Biol Evol **29:**2253-2263.

133. **Clark GW, Ackerman SH, Tillier ER, Gatti DL.** 2014. Multidimensional mutual information methods for the analysis of covariation in multiple sequence alignments. BMC bioinformatics **15:**157.

134. **Simonetti FL, Teppa E, Chernomoretz A, Nielsen M, Marino Buslje C.** 2013. MISTIC: Mutual information server to infer coevolution. Nucleic acids research **41:**W8-14.

135. **Mahony S, Auron PE, Benos PV.** 2007. Inferring protein-DNA dependencies using motif alignments and mutual information. Bioinformatics **23:**i297-304.

136. **Weil P, Hoffgaard F, Hamacher K.** 2009. Estimating sufficient statistics in co-evolutionary analysis by mutual information. Computational biology and chemistry **33:**440-444.

137. **Zhang S, McDonald PW, Thompson TA, Dennis AF, Akopov A, Kirkness EF, Patton JT, McDonald SM.** 2014. Analysis of human rotaviruses from a single location over an 18-year time span suggests that protein coadaption influences gene constellations. Journal of virology **88:**9842-9863.

138. **Albanesi D, Martin M, Trajtenberg F, Mansilla MC, Haouz A, Alzari PM, de Mendoza D, Buschiazzo A.** 2009. Structural plasticity and catalysis regulation of a thermosensor histidine kinase. Proceedings of the National Academy of Sciences of the United States of America **106:**16185-16190.

139. **Kozakov D, Brenke R, Comeau SR, Vajda S.** 2006. PIPER: an FFT-based protein docking program with pairwise potentials. Proteins **65:**392-406.

140. **Comeau SR, Gatchell DW, Vajda S, Camacho CJ.** 2004. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics **20:**45-50.

141. **Comeau SR, Gatchell DW, Vajda S, Camacho CJ.** 2004. ClusPro: a fully automated algorithm for protein-protein docking. Nucleic acids research **32:**W96-99.

142. **Kozakov D, Beglov D, Bohnuud T, Mottarella S, Xia B, Hall DR, Vajda S.** 2013. How good is automated protein docking? Proteins: Structure, Function, and Bioinformatics **81:**2159-2166.

143. **Morcos F, Pagnani A, Lunt B, Bertolino A, Marks DS, Sander C, Zecchina R, Onuchic JN, Hwa T, Weigt M.** 2011. Direct-coupling analysis of residue coevolution captures native contacts across many protein families. Proceedings of the National Academy of Sciences of the United States of America **108:**E1293-1301.

144. **Letunic I, Bork P.** 2011. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic acids research **39:**W475-478.

145. **Letunic I, Bork P.** 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics **23:**127-128.

146. **Maddison WP, Maddison DR.** 2014. Mesquite: a modular system for evolutionary analysis. Version 3.01 http://mesquiteproject.org.