Mechanism and specificity of bacterial two-component signaling systems

by

Emma A. Lubin

A.B. Biochemistry and Molecular Biology
Dartmouth College, 2006

Submitted to the Department of Biology in partial fulfillment of the requirements for the degree of

Master of Science in Biology

at the

Massachusetts Institute of Technology

June 2010

Signature of Author: _____

Certified by: _____
Michael T. Laub
Assistant Professor of Biology
Thesis Supervisor

Accepted by: _____
Stephen P. Bell
Professor of Biology
Chair, Biology Graduate Committee

Mechanism and specificity of bacterial two-component signaling systems

by

Emma A. Lubin

Submitted to the Department of Biology in partial fulfillment
of the requirements for the degree of Master of Science in Biology

Abstract

Bacterial two component signaling (TCS) systems are the predominant means by which bacteria sense and respond to external signals. These systems represent a large family of paralogous proteins; often hundreds of the histidine kinase (HK) and response regulator (RR) pairs that make up a TCS system can be found in a single cell. How do these systems maintain faithful signal transmission and avoid cross-talk? To understand how specificity is determined, we examined co-evolving residues between HKs and RRs, and guided by this, aimed to rewire specificity of several activities of TCS systems. Previous work in the lab has successfully rewired specificity of histidine kinases for response regulators in the phosphotransfer reaction. By mutating different subsets of these co-evolving residues, we were able to rewire specificity of RRs in the phosphotransfer reaction, and partially rewire specificity of HKs and RRs in the phosphatase reaction. Additionally, we identified residues that may dictate specificity between two domains of the histidine kinase, and found that mutating them altered the rate of autophosphorylation. These analyses will allow rational rewiring of two component systems *in vivo*, and permit us to examine the fitness consequences of this altered specificity, providing insight into the evolutionary pressures on TCS systems.

Thesis advisor: Michael T. Laub, Assistant Professor of Biology

Acknowledgements

This work could not have been completed without the constant help and support of the Laub lab, particularly that of Jeff Skerker, Barrett Perchuk, Erin Chen, and Christos Tsokos. I would also like to thank my family, and Steve Bell and Graham Walker, for their kindness and advice as mentors, and their inspiration as scientific role models.

I am very grateful to Victor Ambros, Candy Lee, Allison Abbott, and Jim DiRenzo for their encouragement and guidance, and to Betsey Walsh and the Biology education office, for everything.

I would like to especially thank Mike, for his unwavering support, and for many exciting scientific discussions.

Table of contents

Introduction

Motivation: signaling specificity in paralogous gene families

In many biological systems, a single module can be duplicated and altered to be used in different contexts. This process often involves gene duplication and divergence, which results in the formation of large, paralogous gene families (Zhang, 2003). In mammals, the olfactory system comprises one such family: it employs a set of as many as a thousand G-protein coupled receptor-like proteins to sense different odors (Firestein, 2001). In bacteria, large sets of two-component signal transduction systems are used to sense and respond to many different environmental stimuli. How are so many similar systems able to interact specifically and evolve without incurring a detrimental level of crosstalk? This is an important and widespread problem in understanding biological systems. In higher organisms, factors such as scaffolding proteins or differential sub-cellular localization can help to prevent cross-talk between different members of the same signaling families. However, in other cases, specificity is dictated at the level of molecular recognition (Newman and Keating, 2003; Zarrinpar et al., 2003; Stiffler et al., 2007). Bacteria can harbor tens to hundreds of different two-component systems in a single cell, with similar localization patterns and without the aid of scaffolding proteins. Specificity in this case would appear to be dependent on the signaling molecules themselves. Because of this, two-component systems represent an ideal model for studying how specificity is determined and how it evolves within paralogous gene families.

Two-component signal transduction systems in bacteria

Bacterial two-component signaling (TCS) systems are the predominant signaling modality in bacteria. Several have also been indentified in eukaryotes, where signaling is dominated by cascading Ser/Thr/Tyr phosphorylation systems, rather than stoichiometric His/Asp systems such as TCS systems (Stock et al., 2000). TCS systems are comprised of a sensor histidine kinase and cognate response regulator. In response to an environmental stimulus, the dimerized histidine kinase (HK) autophosphorylates on a conserved histidine residue, and then transfers these phosphoryl groups to a conserved aspartate residue on the response regulator (RR). Phosphorylation activates the response regulator to affect an output within the cell, often by regulating transcription (Stock, 2000). The response regulator can then carry out an autodephosphorylation reaction to terminate the signal. Though the regulator contains all the residues necessary for catalysis of dephosphorylation, this reaction is most commonly stimulated by a second interaction with the histidine kinase to remove the signal (Fig. 1A). Histidine kinases are thus bifunctional, as they can promote both the phosphorylation and dephosphorylation of response regulators, and the conserved histidine, required for autokinase and phosphotransfer, as important in dephosphorylation as well (Hsing and Silhavy, 1997).

E. coli has 30 known histidine kinases and 32 response regulators (Mizuno, 1997). Histidine kinases range from about 40 to 200 kDa, and most are membrane-bound. They have been divided into eleven subfamilies, based on their features in multiple sequence alignments (Grebe and Stock, 1999). EnvZ, an *E. coli* kinase and one of the best-studied HKs, belongs to histidine kinase subfamily 2. Members of this family have a

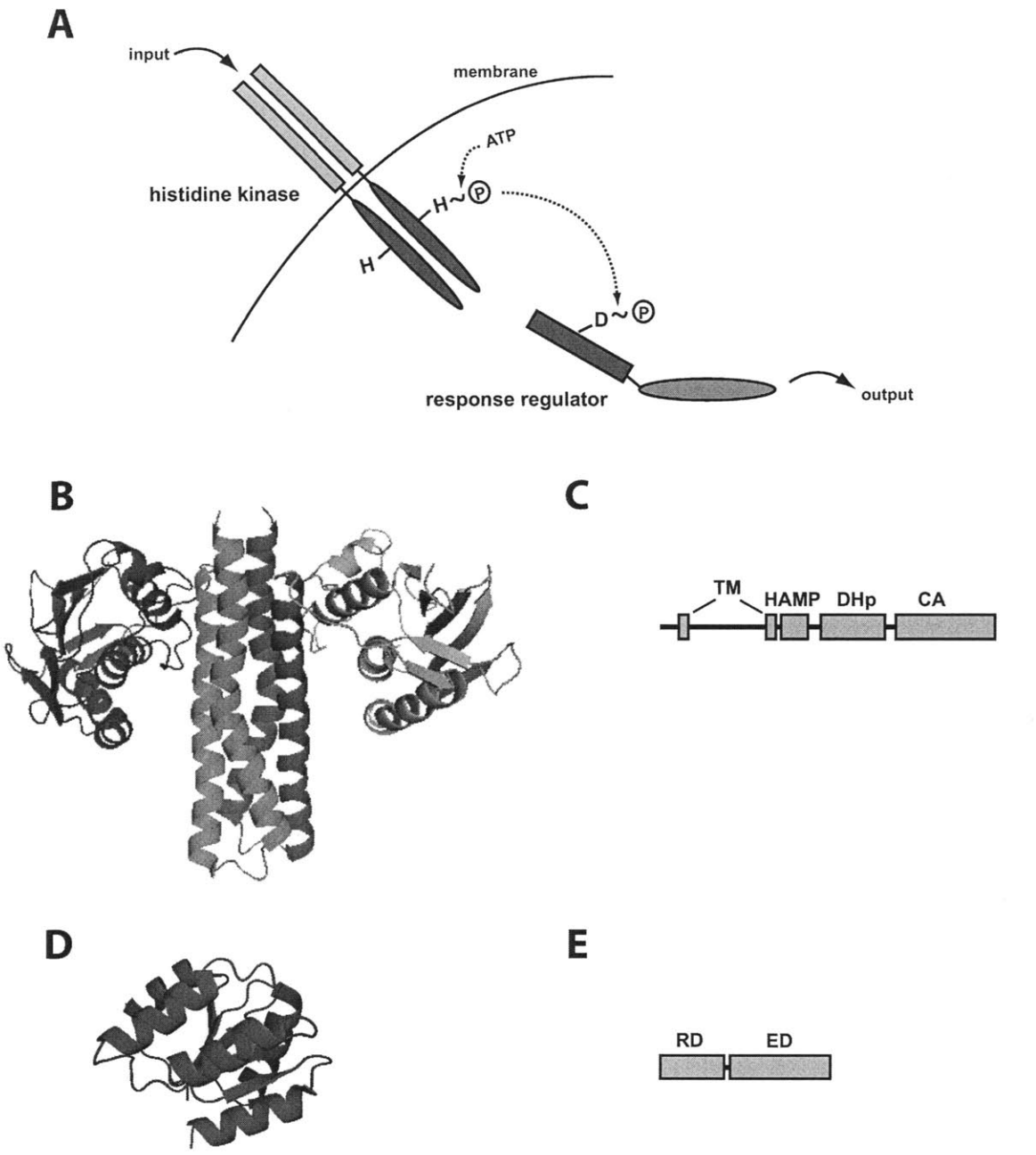

Figure 1: Organization of a prototypical two-component signal transduction system
(A) Schematic of the two-component signaling paradigm. An external signal stimulates the histidine kinase (HK) to autophosphorylate. The phosphoryl group is then transferred to a cognate response regulator (RR), which can effect changes in gene expression or other processes. The HK can interact with its cognate RR a second time, to dephosphorylate it. RRs can also, at a slow rate autodephosphorylate themselves.
(B) HK crystal structure (HK853)
(C) Schematic of the domain organization of a prototypical histidine kinase.
(D) RR crystal structure (RR486). Only receiver domain, not effector domain, is shown.
(E) Schematic of the domain organization of a response regulator.

phenylalanine, leucine or methionine six residues proximal to the conserved histidine (Grebe and Stock, 1999).

EnvZ, which is a prototypical HK, has two transmembrane domains, while others can have more (Stock et al, 2000). Evidence suggests that autophosphorylation of EnvZ takes place in *trans* (Cai and Inouye, 2003), but other kinases autophosphorylate in *cis* (Casino et al., 2009). This autophosphorylation reaction produces a phosphoramidate bond, in contrast to the phosphoester bonds of serine/threonine/tyrosine protein kinases. The cellular portion of EnvZ is linked to its transmembrane domains by a HAMP linker domain, a helix-turn-helix motif that is common to many prokaryotic signaling proteins (Khorchid and Ikura, 2006; Aravind and Ponting, 2006). The HAMP linker is thought to transmit the extracellular signal to the intracellular HK domains through conformational rotation (Hulko et al., 2006). EnvZ increases autophosphorylation in response to signaling (Slauch et al., 1988; Kenney, 2010), but in some other systems, ligand binding stimulates phosphatase activity of the kinase (Freeman et al., 2000; de Wulf et al., 2000; Zhang et al., 2010). Like all HKs, EnvZ have two intracellular domains, a dimerization and histidine phosphotransferase (DHp) domain that contains the conserved histidine, and an ATP-binding catalytic (CA) domain (Fig. 1C). In addition to maintaining interaction specificity between HKs and RRs, HKs must maintain the interaction between its DHp and CA domains.

After autophosphorylation, the high-energy phosphoryl group on the HK is transferred to a conserved aspartate residue on the receiver domain of a response regulator protein. This induces a conformational change that allows the effector domain of the response regulator to carry out the response to the signal (Lee et al., 2001; Stock

and Guhaniyogi, 2006; Stock et al., 2000). The half-life of the phosphorylated aspartate of a response regulator varies, from seconds to hours, depending on the autodephosphorylation rate of the regulator, and the phosphatase activity of the kinase (Stock et al., 2000). The three interactions that dictate the kinetics of signal transfer are summarized in (1) below:

$$\begin{aligned} &\text{Autophosphorylation: HK-His + ATP} \Leftrightarrow \text{HK-His}\sim\text{P + ADP} \\ &\text{Phosphotransfer: HK-His}\sim\text{P + RR-Asp} \Leftrightarrow \text{HK-His + RR-Asp}\sim\text{P} \\ &\text{Phosphatase: RR-Asp}\sim\text{P + H2O} \Leftrightarrow \text{RR-Asp + Pi} \end{aligned} \qquad (1)$$

Though most TCS systems can be described by this simple paradigm, others have slightly different arrangements of these modular domains. In such systems, hybrid histidine kinases can have phosphodonor and phosphoacceptor sites in the same protein. The domains themselves remain remarkably similar to those in free HKs and RRs (Stock et al., 2000). Hybrid kinases represent a scenario in which the constraints on interaction specificity may be lessened because the interacting domains are tethered together.

Though TCS systems are found predominantly in bacteria, a few TCS systems, and others with homology to TCS systems, are found in higher organisms. In yeast, a modified two-component system helps to regulate the HOG1 osmosensing pathway: a sensor histidine kinase, SLN1, phosphorylates two response regulators, SSK1 and SKN7, though an intermediate YPD1 (Brown et al., 1994; Ketela et al., 1998; Liet al., 1998; Posas et al., 1996). In plants, one example of how two-component systems have been adapted for use in eukaryotes is that of phytochrome proteins, which transmit photosensory information. These proteins show homology to histidine protein kinases,

but only a few have the canonical H-boxes of HKs and some are thought to function as Ser/Thr kinases (Muller et al., 2009). Direct targets of the kinase portion of phytochromes have not been identified, so it is not known whether they transfer to response regulator-like proteins. However, mutation of the same conserved region that, when mutated disrupts bacterial histidine kinase function (Hsing et al., 1998) also effects phenotypes downstream of phytochrome signaling. A cyanobacterial phytochrome, Cph1_ssp, does phosphotransfer to a response regulator, Rph1 (Grebe and Stock, 1999), and this may represent an evolutionary intermediate between bacterial and eukaryotic uses of the TCS signaling modality.

Structural data offer insight into TCS system mechanism and specificity

Despite their diverse cellular roles, TCS proteins, particularly in bacteria, retain remarkable sequence homology. Structural approaches have traditionally been employed to understand how these proteins interact and function. However, crystallization of TCS proteins has been notoriously difficult, and structural data has accumulated from a variety of sources – only within the last two years has a more complete structural picture of TCS systems emerged. While response regulator receiver domains have been largely amenable to crystallization, and a number of structures exist (Stock et al., 1989; Baikalov et al., 1996; Lewis et al., 1999; Bachhawat and Stock, 2007) the first high-resolution crystal structure of a histidine kinase was not published until 2005 (Marina et al., 2005) and the first high-resolution co-crystal was produced only in 2009 (Casino et al., 2009).

Prior to the availability of an HK-RR co-crystal, an understanding of possible interactions between HKs and RRs was inferred from a co-crystal between the *Bacillus*

subtilis phosphorelay proteins Spo0B and Spo0F (Zapf et al, 2000). Spo0F is a response regulator while Spo0B is a histidine phosphotransferase; this protein can accept and donate a phosphoryl group on a conserved histidine, and dimerizes and forms a four-helix bundle as HK DHp domains do, but lacks an N-terminal transmembrane domain and C-terminal CA domain.

Structural understanding of HK-RR interactions has improved with the publication of an HK crystal structure and HK-RR co-crystal structure. Both structures are from the HK853-RR468 TCS system from the thermophilic bacterium *Thermatoga maritima.* The 1.9 Å resolution structure of the entire cytoplasmic domain of HK853 (Fig. 1B) contains an asymmetric subunit of the dimerized kinase complexed with one sulfate ion, water molecules, and one molecule of ADP-β-N, the hydrolysis product of the AMPPNP provided in the crystallization buffer (Marina et al, 2005). This structure presents one possible conformation of the histidine kinase in the absence of the response regulator, and identifies several points of interaction between the DHp and CA domains in this orientation. Though AMPPNP is an ATP analog, it was hydrolyzed in the buffer to ADP-β-N, more analogous to ADP since it lacks the γ phosphate. Likely because of this, the ATP lid is disordered in the structure. The sulfate ion is positioned such that it may be mimicking the phosphate on the histidine. The presence of the ADP analog in combination with the positioning of the sulfate ion indicates that this structure is representative of a histidine kinase that has just completed autophosphorylation, rather than one prepared to bind and dephosphorylate a response regulator (Marina et al., 2005).

Complementary to these data, the co-crystal of HK853 with its cognate response regulator, RR486 (Fig. 3B), and structures of RR486 alone (Fig. 1D) provide a fuller, but

still incomplete, picture of possible interactions. In the co-crystal of HK853 and RR486, HK853 has a different orientation of its DHp and CA domains compared to the structure of the kinase alone. RR486 is a single-domain response regulator. RR486 was crystallized both in the presence and absence of BeF_3-, which, if present, can bind noncovalently at the conserved aspartate to mimic phosphorylation (Casino et al., 2009). The combined data from these structures suggest that the response regulator in the complex is in the conformation associated with the phosphorylated state, since it is most similar to the structure of RR468 in the presence of BeF_3- (Casino et al., 2009). The conformation of the histidine kinase in the co-crystal appears to be representative of the phosphatase interaction.

A computational and experimental approach to identify the determinants of TCS system specificity

Even with these new structural data, experimental work is needed to understand the molecular basis of specificity in TCS systems. Previous work in our lab has demonstrated that HKs and RRS do in fact interact specifically, and we have used computational and experimental methods to identify the determinants of HK specificity in the phosphotransfer reaction. These studies have enabled the rationally rewiring of specificity both *in vitro* and *in vivo*. By purifying large numbers of HK and RR proteins, and using a robust assay for effective phosphotransfer between the two, the lab has demonstrated that HKs exhibit a kinetic preference *in vitro* for their *in vivo* cognate response regulators (Laub et al., 2006). To identify the amino acid residues responsible for this preference, we first employed a computational approach, using mutual

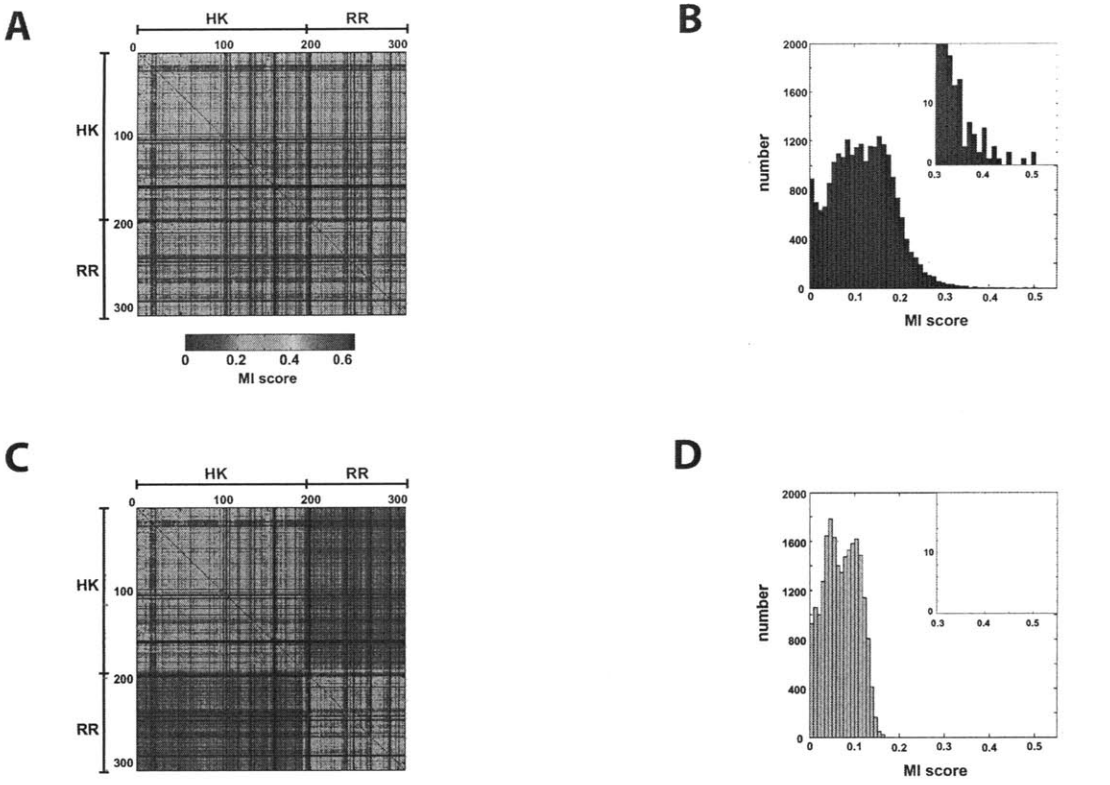

Figure 2: Mutual information analysis of two-component signal transduction systems (A) Covarying residues in histidine kinases and response regulators were identified by mutual information analysis. Operon pairs of cognate HKs and RRs were concatenated and aligned. Each pair of positions in this multiple sequence alignment was analyzed for covariation using mutual information analysis. The heat map in (A) is a plot of the matrix of scores for each pair of positions. Positions corresponding to the histidine kinase and response regulator are indicated. (B) Histogram of interprotein (HK-RR) mutual information scores from (A). (C) Pairings of HK and RR sequences were then randomized, and mutual information analysis repeated. The heatmap of scores from this randomized alignment reduces interprotein, but not intraprotein covariation scores. (D) Histogram of interprotein (HK-RR) mutual information scores from (C). (E) For the MI analysis in (A) of cognate HK-RR pairs, the graph in (E) plots pairs of residues with scores > 0.35. Regions corresponding to HK DHp and CA domains, and RR receiver domains are indicated.

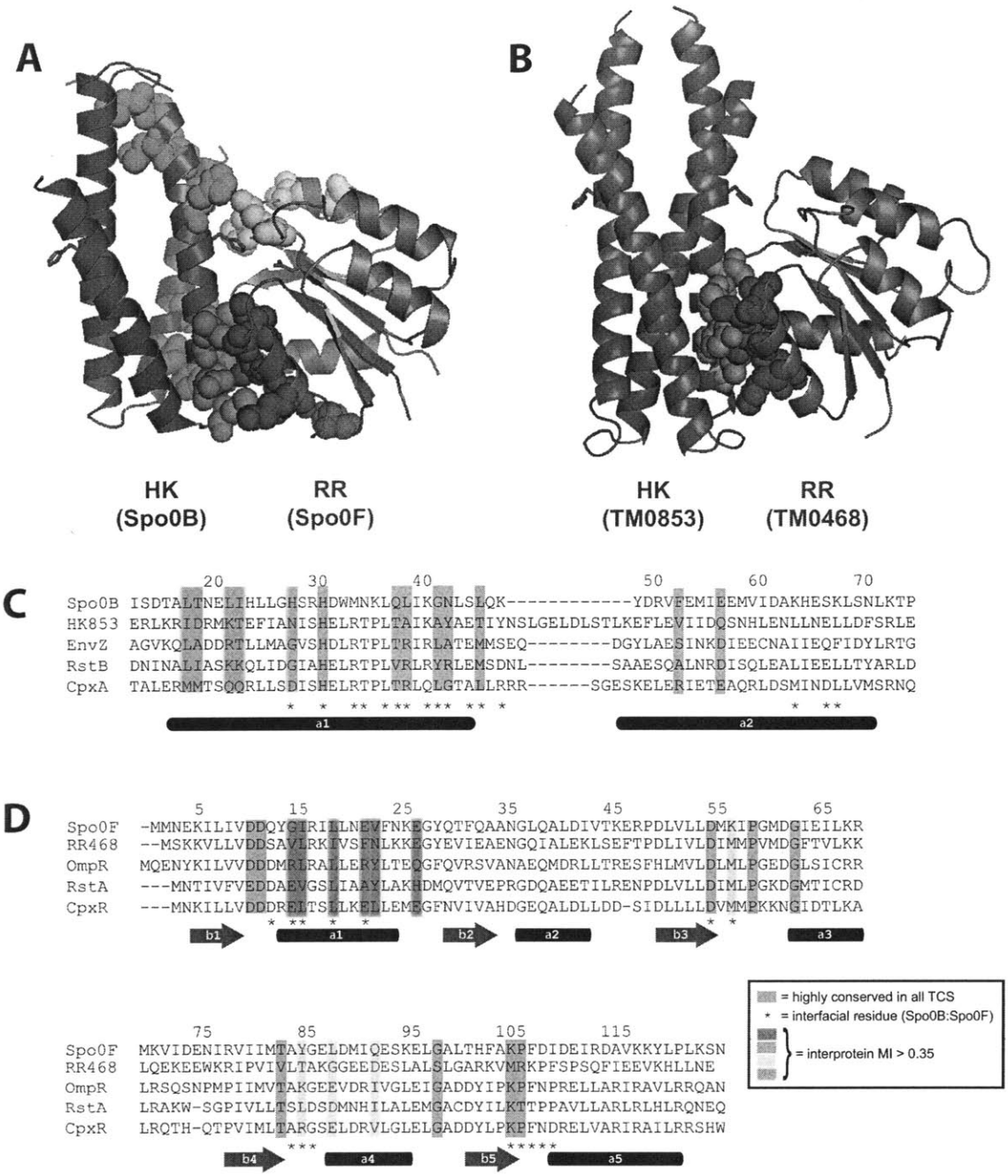

Figure 3:

(A) The Spo0B/Spo0F co-crystal. (PDB IF51). Predicted interacting residues shown in orange and green (kinase) and red and yellow (regulator).

(B) The HK853/RR486 co-crystal. Only one set of interacting residues is shown; the second set does not appear proximal enough in the crystal structure to represent a potential interaction.

(C) Primary sequence alignment of *Bacillus subtilis* phosphotransferase Spo0B, and histidine kinases *Thermatoga maritima* HK853, *E. coli* EnvZ, RstB, and CpxA. As in (A) and (B), highly co-varying residues are shown in green and orange.

(D) Primary sequence alignment of the cognate response regulators to the proteins in (C): Spo0F, RR468, OmpR, RstA, and CpxR. As in (A) and (B), highly co-varying residues are shown in red and yellow. Conserved residues are grey, and Spo0B/0F contacts have asterisks.

information (MI) algorithms. Using large multiple sequence alignments of HK and RR sequences, Skerker et al. searched for residues that co-vary with each other, with the hypothesis that residues that co-vary between HKs and RRs are likely to be the specificity-determining residues (Skerker et al., 2008).

This coevolution analysis identified two patches of potential specificity residues in the HK and RR (Fig. 3). One, clustered around the conserved histidine in the kinases, appeared best positioned to interact with the response regulator, and so these residues were tested for their ability to determine specificity. Using the *E. coli* histidine kinase EnvZ as a backbone, we replaced the residues identified by mutual information (MI) analysis in EnvZ with the corresponding residues in a number of other *E. coli* kinases. Doing so resulted in a partial specificity switch, but did not completely alter specificity for all kinases. A flexible loop is positioned in the midst of these MI residues, in between $\alpha 1$ and $\alpha 2$ of the histidine kinase; because this loop varies so widely in the number and types of residues across HK sequences, it did not align well enough to be accessible to MI analysis. By replacing this loop in the MI alone mutants with the corresponding loop from the target kinase, we were able to completely switch specificity in all cases tested. These so-called MI+loop mutants no longer phosphotransferred to OmpR, the cognate regulator of EnvZ, and each phosphorylated the cognate regulator of its target kinase with kinetics similar to the wild-type target kinase (Skerker et al., 2008). When several of these mutant kinases were introduced into *E. coli* strains containing a reporter for the output of the signaling system, the mutant kinases functioned effectively *in vivo* with the specificity of their target kinase (Skerker et al., 2008).

In this thesis I have aimed to further characterize and re-engineer the determinants of HK and RR interactions in order to understand how specificity is dictated in two-component systems, and to ask how it evolves. By re-wiring specificity first *in vitro* we can demonstrate how in two-component systems it is prescribed by molecular interaction alone. Future work will be required to demonstrate, as has already been done with HK rewiring, that molecular recognition is sufficient to achieve specificity *in vivo* as well. By understanding the interprotein and interdomain interactions between the two components of the system, we can more fully understand how to alter the determinants of phosphate flow through a TCS system, and rewire specificity without changing other activities of the proteins. By using all this knowledge to rewire interactions, we will in the future be able to ask what the constraints are on these interactions as they evolve.

Results

Mutual information identifies co-varying residues

We based this work on the same mutual information analysis that was previously successful in identifying specificity residues in histidine kinases (Skerker et al, 2008). Mutual information is one of several computational approaches researchers have developed to the problem of identifying interacting residues. One other successful method is statistical coupling analysis (SCA), which is, like mutual information, based on the idea that if residues are functionally coupled, they will co-evolve. SCA assumes that given a large and diverse sequence alignment and conservation of functional coupling at given positions, there should be statistical correlations between the amino acid distributions at these sites (Socolich, 2005). This analysis has proven successful in the analysis of WW domain folding; however, we chose to employ mutual information, both because of the need to validate it as a technique, and because software was freely available.

We used the same mutual information analysis as in Skerker et al. (2008); briefly, cognate pairs of HKs and RRs were identified by the fact that such pairs are generally found in the same operon. From nearly 200 bacterial genomes, roughly 1300 pairs were identified. Cognate HK and RR sequences were concatenated and then aligned. Using available software (Fodor and Aldrich, 2004), we calculated the mutual information between each pair of sites (Fig 2). The mutual information between two positions X,Y in an alignment is defined as:

$$\sum_{j=1}^{n} \sum_{k=1}^{m} p_{jk} \log \frac{p_{jk}}{p_j q_k} \qquad (2)$$

Where column X has n different residues, column Y has m different residues, p_j is the probability of finding residue j in column X, q_k is the probability of finding residue k in column Y, and p_{jk} is the number of sequences with residue j in column X and residue k in column Y divided by the total number of sequences (Skerker et al., 2008; Atchley et al., 2000; Fodor and Aldrich, 2004) (Fig. 2A,B). In order to distinguish significant mutual information scores from background noise due to sampling bias, we calculated the MI scores of randomized HK-RR pairs (Fig. 2C,D). These scores between intermolecular residue pairs in the randomized alignment all fell well below 0.35, while in the alignment of cognate pairs, 43 intermolecular residue pairs had MI scores above 0.35 (Fig. 2E). This 43 pairs were composed of 16 histidine kinase residues and 12 response regulator residues. We took these pairs to be significant; of them, we focused on the subset of 36 pairs between residues in the RR and HK DHp domain, while the remaining 7 were between residues in the RR and HK CA domain.

When the co-varying pairs identified by mutual information are mapped onto the Spo0B/Spo0F structure, they cluster into two main regions of interacting residues (Fig. 3A). The residues required to switch specificity in histidine kinases fall into one cluster, while those in the second cluster were found to have no effect on specificity (Skerker et al, 2008). This first region co-varies with residues clustered in $\propto$ helix 1 of the response regulator receiver domain, and appears to contact them in the Spo0B/0F co-crystal. The

second region co-varies with a set of residues in the RR spread over several regions in the receiver domain. These residues overlap with some implicated in response regulator autodephosphorylation rate (Pazy et al., 2009; Thomas et al., 2006), while others have not been assigned a function.

Re-wiring response regulator specificity in vitro

Previous work in the lab has identified and re-wired specificity determinants in histidine kinases. We aimed to determine the corresponding specificity determinants in the response regulator and to further validate mutual information as a strategy for predicting co-varying residues. Based on our mutual information data, we focused our initial mutational analyses on $\propto 1$ of the receiver domain (Fig. 3D).

Using *E. coli* OmpR, the cognate regulator to EnvZ, as a backbone, we mutated putative specificity residues to the corresponding amino acids found in RstA and CpxR, the cognate regulators of RstB and CpxA, and tested the resulting purified proteins for their ability to be phosphorylated by EnvZ, RstB, and CpxA. Though RstB and CpxA have higher homology to EnvZ relative to other E. coli RRs, each shows a strong kinetic preference for its cognate response regulator (Skerker et al., 2008). Since the putative specificity residues occupied most of α helix 1 in the response regulator, we first replaced the entire α helix of OmpR with those of RstA and CpxR, reasoning that this would grant OmpR the specificity of RstA or CpxR if mutual information residues are in fact the sole specificity determinants. However, one of these constructs, OmpR with the α helix 1 of CpxA, did not produce a soluble protein, and the other, with the α helix of RstB, was not phosphorylated by RstA, and only minimally by EnvZ (data not shown). These mutant

proteins likely do not fold properly as the changes introduced included both solvent-exposed and buried residues.

We then tested whether the residues identified by MI alone, all of which are solvent-exposed, could switch specificity. We replaced putative specificity residues in α helix 1 of OmpR with the corresponding residues in RstA, terming the mutants OmpR[RstA], and CpxR. Substitution of a single residue produced little effect on regulator specificity; these OmpR mutants were still recognized by EnvZ in the phosphotransfer reaction, and were not phosphorylated by RstB or CpxR (data not shown). However, mutating multiple residues simultaneously produced a more significant switch in specificity. OmpR mutants in which three MI residues were replaced with the corresponding residues in RstA or CpxA (Fig. 4A,C) were now only weakly phosphorylated by EnvZ (Fig. 4B,D), although RstB did not recognize the OmpR[RstA] mutant, and CpxA only partially phosphorylated the OmpR[CpxR] mutant. Because these assays were carried out at only one timepoint, we tested whether CpxA displayed a kinetic preference for wild-type CpxR over the OmpR[CpxR] specificity mutant (Fig 4E). Indeed, CpxA recognizes the mutant better than OmpR, but not as well as CpxR, indicating that residues outside of this region are required to determine specificity. We tested whether the second set of residues identified by mutual information contributed to specificity. These residues cluster near the loop between β-sheet 4 and α-helix 4. This region is implicated in modulating the autodephosphorylation rate of RRs, and mutating these residues in one regulator to the corresponding amino acids of a second regulator can confer the autodephosphorylation rate of the second regulator on the first (Pazy et al., 2009, Thomas et al., 2006). However, Spo0B/0F

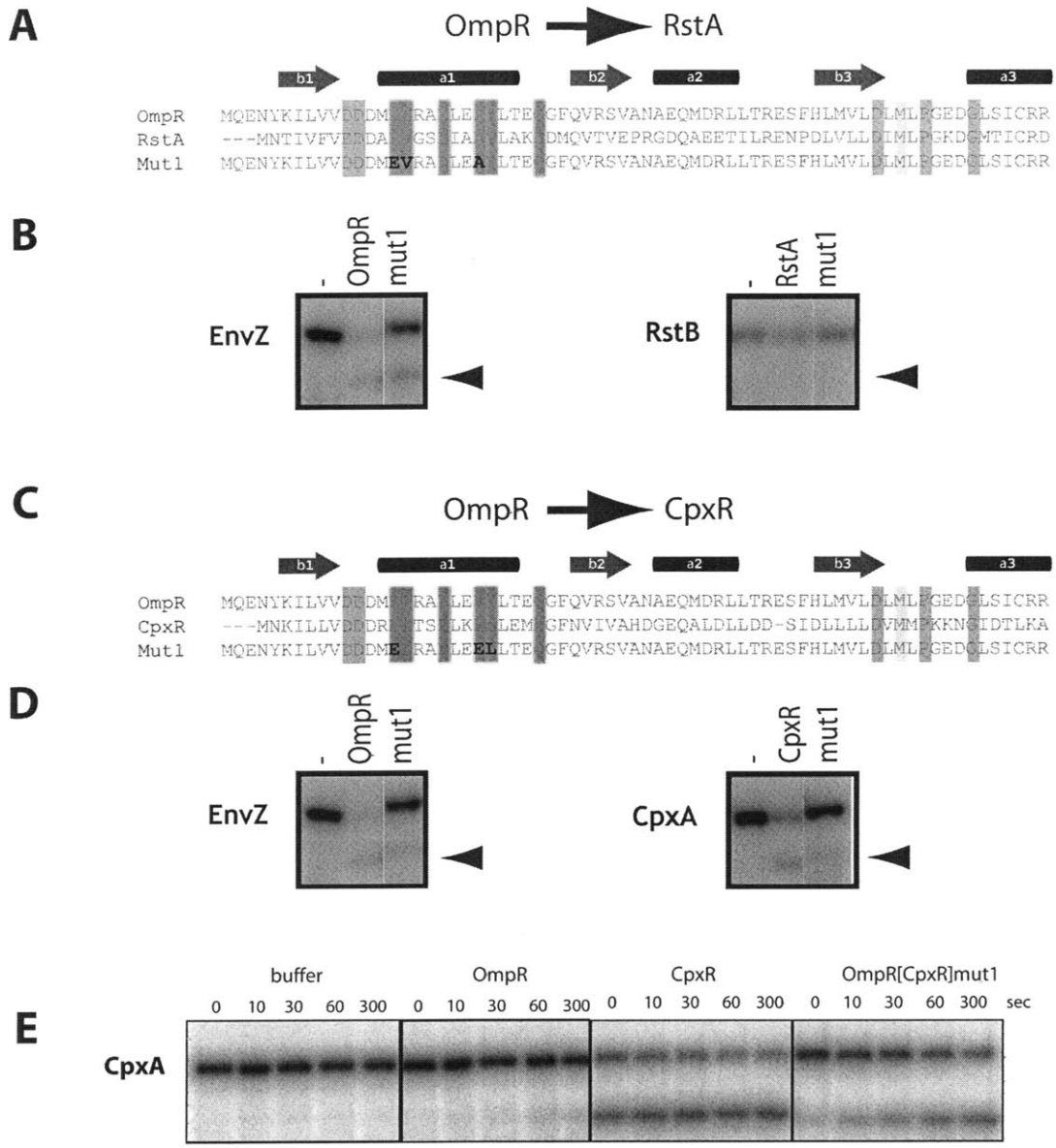

Figure 5: Mutual information predicts RR specificity determinants
(A) Amino acid sequences of the receiver domains of RR specificity point mutant from OmpR to RstA. In OmpR, co-varying residues identified by MI analysis were replaced with corresponding residues in RstA as determined by their positions in an MSA. Mutated residues are shown in bold. Specificity residues predicted by MI are highlighted. For complete protein sequences see Appendix.
(B) Phosphotransfer specificity of OmpR[RstA]MI point-mutant response regulator. Each wild-type and mutant regulator was incubated with buffer or with autophosphorylated cognate kinase (EnvZ) or specificity target kinase (RstB). Incubations were for 10 seconds. Black arrows denote the phosphorylated band corresponding to the positionof the regulator.
(C) Amino acid sequences of the receiver domains of RR specificity point mutant from OmpR to CpxR.
(D) Phosphotransfer specificity of OmpR[CpxR] MI point-mutant response regulators. for cognate (EnvZ) and target (CpxA) kinases. Phosphotransfer reactions ran for 10 seconds.
(E) Timecourse of phosphotransfer specificity of OmpR[CpxR MI] mut1 for target kinase CpxA. CpxA was autophosphorylated and free phosphate and ATP purified away. Purified, autophosphorylated kinase was incubated with each response regulator (wild-type cognate RR, wild-type target RR, and specificity mutant RR) or with buffer alone for 0, 10, 30, 60, or 300 seconds before reaction was stopped by the addition of SBB.

co-crystal identifies this region as having potential contacts with the histidine kinase (Zapf et al., 2000), and so it seemed possible that this region has dual roles, in autodephosphorylation and specificity. We mutated MI residues in this region alone and in tandem with α1 mutual information residues. In both cases, these mutations did not alter specificity (data not shown). The finding that residues identified by MI have no effect on specificity appears to lessen the power of mutual information as an approach to determining interacting residues. One possible explanation for the identification of these positions by mutual information is two-component systems may in general evolve with the constraint that they must maintain a certain rate of phosphate flow through the system. If this is the case, phosphotransfer and other steps in signal transfer, such as autodephosphorylation, will co-evolve to some degree. This is a matter for further study, and our work will provide the tools to elucidate questions such as these.

Mutual information residues dictate specificity in combination with loop regions

Since HK specificity had proven to be determined both by MI residues and by a flexible loop region that was inaccessible to computational methods, we asked whether RR sequences might have a similar, poorly alignable loop region. In the background of several α helix 1 specificity mutants, we mutated a loop region that was not identified by mutual information, but was implicated as a potential HK contact site in the Spo0B/Spo0F co-crystal. There is only a single amino acid difference in this loop between OmpR and CpxR (Fig 5C) and a three amino acid difference between OmpR and RstA (Fig 5A). We tested these mutants for their ability to accept a phosphoryl group both from OmpR's cognate kinase, EnvZ, and from the new specificity target kinase,

CpxA or RstB. We first autophosphorylated each kinase and then assayed for phosphotransfer to mutant and wild-type regulators. These mutants are no longer recognized by EnvZ, and are phosphorylated by their specificity target kinase, CpxA or RstB (Fig. 5D,B). In both sets of mutants, the target kinase appears to better recognize its own cognate regulator than the specificity mutants, but the switch in specificity is significant. Further, adding the loop region improves on the specificity of mutations in α helix 1 alone (Fig. 4). We constructed several different α helix 1 mutations, both adding an additional specificity residue, and returning one of the initially mutated residues (Fig. 4) to that of OmpR. We found that the additional specificity residue, in Mut3 and Mut5, had a significant impact in both OmpR[RstA] mutants and OmpR[CpxR] mutants – it increased the ability of the mutant regulator to receive a phosphoryl group from its target kinase. Replacing one of the three initial α1 specificity residues (Mut2) with the corresponding OmpR residue (Mut4) had a more ambiguous effect, but it appears that this residue does not have a significant effect on phosphotransfer preference in these cases. Indeed, subsequent mutual information analysis has found that this residue does not co-vary with HK residues as highly as others (Michael Laub, unpublished). These findings further demonstrate the power of mutual information as a means of identifying interacting residues, but also indicate the necessity of combining computational with experimental approaches. Though we have found β5-α5 loop to be important, the bulk of specificity appears to be determined by the residues in α1, indicating that the loop region may serve to stabilize this interaction by forming supplementary contacts with the kinase. Mutating the loop region alone produces a regulator that is still minimally recognized by

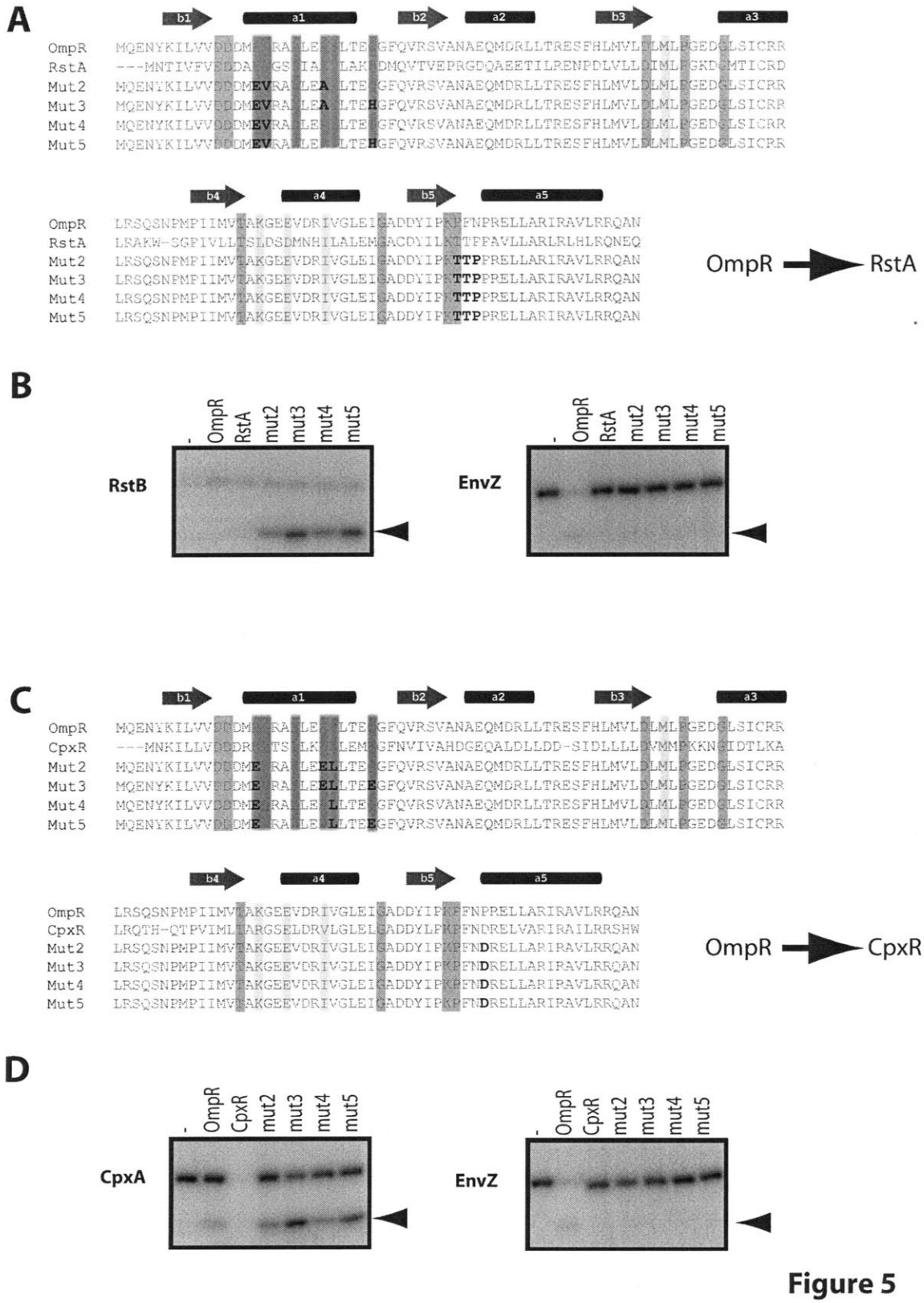
A
b1
a1
b2
a2
b3
a3
OmpR
RstA
Mut2
Mut3
Mut4
Mut5
b4
a4
b5
a5
OmpR → RstA
B
-
OmpR
RstA
mut2
mut3
mut4
mut5
RstB
EnvZ
C
CpxR
OmpR → CpxR
D
CpxR
CpxA
EnvZ

Figure 5

Figure 5: Mutation of a loop region is needed to switch response regulator specificity
(A) Amino acid sequences of the receiver domains of OmpR → RstA RR specificity MI+loop mutants. Residues of OmpR were replaced with corresponding residues of RstA as determined by a multiple sequence alignment. Mutated residues are shown in bold. Specificity residues predicted by mutual information are highlighted. MI residues in alpha helix 1 were mutated, in conjunction with residues in one loop region. For complete protein sequences see Appendix.
(B) Phosphotransfer specificity of OmpR → RstA MI + loop response regulator mutants for cognate and target kinase. Autophosphorylated cognate (EnvZ) or target (RstB) kinase was incubated with buffer or with an equimolar amount of response regulator. Incubations were for 10 seconds with EnvZ and for 15 minutes with RstB due to the differential rate of phosphotransfer activities of these kinases (data not shown).
(C) Amino acid sequences of the receiver domains of OmpR → CpxR RR specificity MI+loop mutants. Residues of OmpR were replaced with corresponding residues of RstA as determined by a multiple sequence alignment. Mutated residues are shown in bold. Specificity residues predicted by mutual information are highlighted. MI residues in alpha helix 1 were mutated, in conjunction with residues in one loop region. For complete protein sequences see Appendix.
(D) Phosphotransfer specificity of OmpR → CpxR MI + loop response regulator mutants for cognate and target kinase. Autophosphorylated cognate (EnvZ) or target (CpxA) kinase was incubated with buffer or with an equimolar amount of response regulator. Incubations were for 10 seconds with EnvZ and for 15 minutes with CpxA due to the differential rate of phosphotransfer activities of these kinases (data not shown).

EnvZ in the phosphotransfer region, and only partially recognized (OmpR[RstA]) or not recognized (OmpR[CpxR]) by the target kinase (data not shown).

Histidine kinases have highly specific phosphatase activity

Histidine kinases have been shown to engage in two reactions with their cognate response regulator: phosphotransfer and dephosphorylation. Though histidine kinases can be referred to as phosphatases, they formally do not catalyze this reaction. Rather, the catalytic residues reside in the response regulator; it has been hypothesized that HKs increase the rate of the dephosphorylation reaction, possibly by stabilizing the transition state, since mutation of the conserved histidine largely eliminates phosphatase activity (Tomomori et al., 1999; Stock, 2001). Having both activities is common, but many HKs have been described in the literature as having only kinase activity, termed monofunctional (Alves and Savageau, 2003). Differentiating monofunctional from

bifunctional kinases in sequence or structure has proven a difficult problem (Alves and Savageau, 2003). Another understanding of monofunctional and bifunctional kinases seems more plausible: the dual function of kinases can be examined as a balance of two competing activities, which can be swayed by molecular changes that make one interaction more favorable than the other. In some kinases, this balance lies heavily to the kinase side, so that there is little or no detectible phosphatase activity – in this case the histidine kinase appears monofunctional. If a kinase were to exist in a monofunctional state over a long period through evolution, it is possible that additional changes to the system may occur – for example, the kinase no longer recognizing its cognate regulator in a phosphatase interaction – but such changes may be difficult to detect and categorize on a systems level. In this explanation, in which HKs in general have both phosphatase and kinase activities, one should be able to identify determinants in HKs of both phosphatase specificity and of the phosphotransfer/phosphatase balance. These changes should affect interaction with the regulator and interaction between the DHp and CA domains, respectively. This is supported by the identification of EnvZ mutants with hyperactive kinase activity but low phosphatase activity (K+ P-), and low kinase activity but high phosphatase activity (K- P+) (Hsing et al., 1998; data not shown).

Though our lab has previously shown that HK phosphotransfer activity is highly specific, the specificity of the phosphatase reaction has never been studied. The importance of HK phosphatase activity has been noted in the regulation of several particular TCS systems (Atkinson et al., 1994), but it is not known to what degree phosphatase activity functions to regulate TCS systems in general, and whether HKs show a kinetic preference for their cognate regulators in dephosphorylation as well as in

phosphotransfer. Further, if not all kinases have high levels of phosphatase activity, and if the determinants of the phosphatase interaction are different from the phosphotransfer interaction, it is possible that there is less selective pressure to maintain specific phosphatase activity.

To determine whether HKs have specific phosphatase activity, we first examined the phosphatase specificity of EnvZ and several other *E. coli* kinases for the cognate partner of EnvZ, phosphorylated OmpR. To obtain purified phosphorylated OmpR we incubated OmpR with radioactive ATP and a small amount of a K+P- EnvZ mutant in which the phosphatase activity had been decreased so that it was able to autophosphorylate and phosphotransfer, but act only as a very weak phosphatase (Fig. 6B). Since the half-life of OmpR~P is approximately 90 minutes (Zhu et al., 2000), the absence of phosphatase activity in the reaction allowed us to load OmpR with phosphate for an extended time. After removing free phosphate, we then added a stoichiometric amount of OmpR's cognate HK EnvZ or another E. coli kinase. OmpR was readily dephosphorylated by EnvZ, but not by other kinases (Fig. 6A); these kinases have phosphatase activity on their own cognate regulators (data not shown). These data indicate that, as with phosphotransfer, histidine kinases have a strong preference as phosphatases *in vitro* for their *in vivo* cognate substrates.

HK phosphatase specificity is separable from phosphotransfer specificity

We next asked whether the molecular determinants of phosphatase specificity were identical to those of kinase specificity. Previous structural studies had not differentiated between a phosphotransfer and phosphatase interaction between kinases

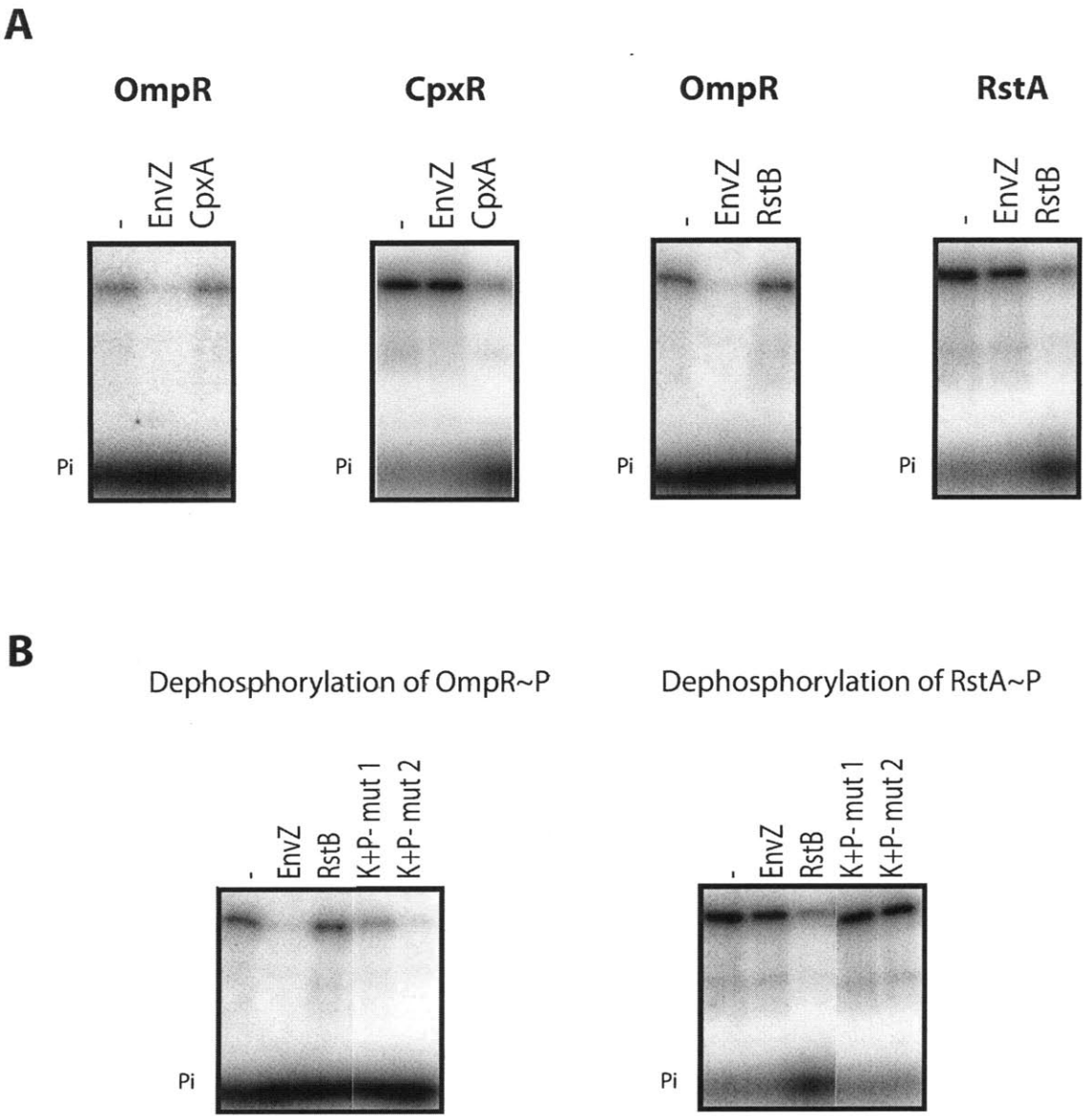

Figure 8: Histidine kinases have specific phosphatase activity
(A) Hisitidine kinases were tested for their specificity for cognate and non-cognate regulators in the phosphatase reaction. First, OmpR, RstA, and CpxR were each incubated with a small amount of a K+P- version of EnvZ to phosphorylate the regulator. Free phosphate was then purified away, and phosphorylated regulator was added either to buffer or to an approximately equimolar amount of cognate or non-cognate kinase.
(B) We made single point mutations in EnvZ to produce two mutants that phosphotransfer effectively (data not shown) but that have impaired phosphatase activity (K+ P-). K+P- mut1 is more impaired phosphatase than K+P- mut2, and so this was chosen as the loading kinase to produce phosphorylated regulator in A. Diminishing phosphatase activity did not alter the specificity of the kinase. Purified, phosphorylated cognate (OmpR) or non-cognate (RstA) regulator was incubated with buffer or wild-type or mutant kinase.

and regulators, but we hypothesized that the two interactions would differ for several reasons. The position of the RR relative to the HK most likely needs to be different in the phosphatase reaction for the conserved histidine to stabilize the transition state, and the phosphatase interaction is complicated by the fact that, though most regulators are thought to be monomers when unphosphorylated, they form dimers or oligomers once phopshorylated (Stock et al., 2000). Previously identified phosphatase mutations have been mapped to regions within the DHp domain (Hsing et al., 1998); though it is not clear whether these mutants altered specificity or simply phosphatase activity of the kinase.

To test whether the residues that dictate phosphotransfer specificity were sufficient to determine phosphatase specificity as well, we assayed the phosphatase activity of six sets of MI and MI+loop mutants that were shown to partially and fully switch phosphotransfer specificity, respectively, from EnvZ to a target kinase (Skerker et al., 2008). When these constructs were incubated with purified, phosphorylated OmpR, they no longer dephosphorylated OmpR~P at a timepoint at which EnvZ fully dephosphorylates it (data not shown), indicating that specificity is partly switched by these mutations. We then tested three sets of these MI and MI+loop mutants (in which EnvZ had been given the specificity residues of RstB, CpxA, or PhoR) for dephosphorylation of a new, targeted RR. None of the mutants tested fully dephosphorylated their target regulators (Fig. 7B), although, with one exception (EnvZ[CpxA MI]), each mutant was a better phosphatase for the target regulators than was EnvZ. These results indicate that the phosphatase and phosphotransfer specificity determinants overlap, but are not identical.

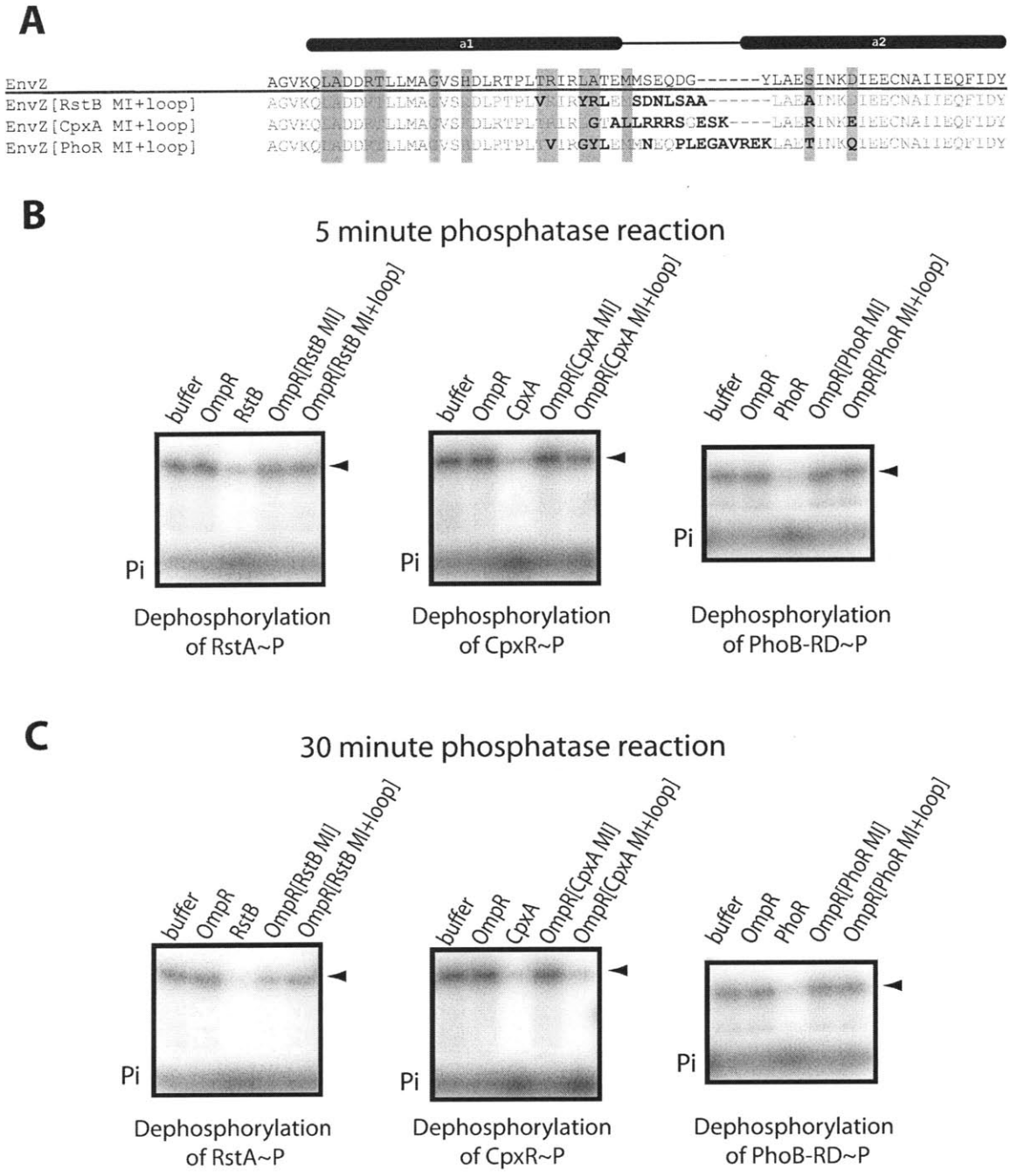

Figure 7: MI+loop mutants sufficient to switch HK phosphotransfer specificity are not sufficientto switch phosphatase specificity.

(A) Amino acid sequences of the DHp domains of HK MI + loop specificity mutants in EnvZ. These mutations are sufficient to fully switch phosphotransfer specificity of EnvZ so that it no longer phosphorylates OmpR but instead phosphorylates target non-cognate regulator (Skerker et al., 2008). For complete protein sequences see Appendix.

(B) Phosphatase specificity of histidine kinase MI + loop mutants. Each kinase was incubated with buffer or with purified, phosphorylated OmpR or specificity target regulator, RstA, CpxR, or the receiver domain alone of PhoB. Incubations were for 5 minutes.

(C) The same phosphatase reactions as in (B) were carried out for 30 minutes.

Additional determinants may be found elsewhere in the DHp domain, or, alternatively, in the CA domain. The second region of mutual information residues identified in HKs (Fig. 3C) represents one possibility. However, initial mutations in these residues show no effect on phosphatase specificity (data not shown). We also tested domain-level chimeric proteins which contained the DHp domain of a target kinase, in this case RstB, fused to the CA domain of EnvZ (Fig. 8A). We reasoned that if specificity resides solely in the DHp domain, these proteins would have the same phosphatase specificity as that of the DHp donor kinase. These chimeric kinases have been shown to autophosphorylate and phosphotransfer, with the specificity of phosphotransfer matching that of the DHp donor kinase (Skerker et al., 2008). Chimeric RstB-EnvZ did not dephosphorylate OmpR~P, but did act as a weak phosphatase for RstA~P (Fig. 8B). The chimeric is not as efficient a phosphatase for RstA~P as full-length RstB is. This could indicate two things: either that some phosphatase specificity determinants reside in the CA domain, or that specificity is in fact determined by the DHp domain alone, but that the interactions between the DHp and CA domain needed for phosphatase activity have been disrupted by incompatible regions in the DHp of one kinase and the CA domain of the other.

To differentiate between these possibilities, we examined of the activity of isolated DHp domains. The isolated EnvZ DHp domain has previously been shown to function as a phosphatase *in vitro* (Zhu et al., 2000), albeit more weakly. If specificity is dictated solely by the DHp domain, and the chimeric proteins did not have fully switched specificity due to disrupted inter-domain interactions, isolated DHp domains should

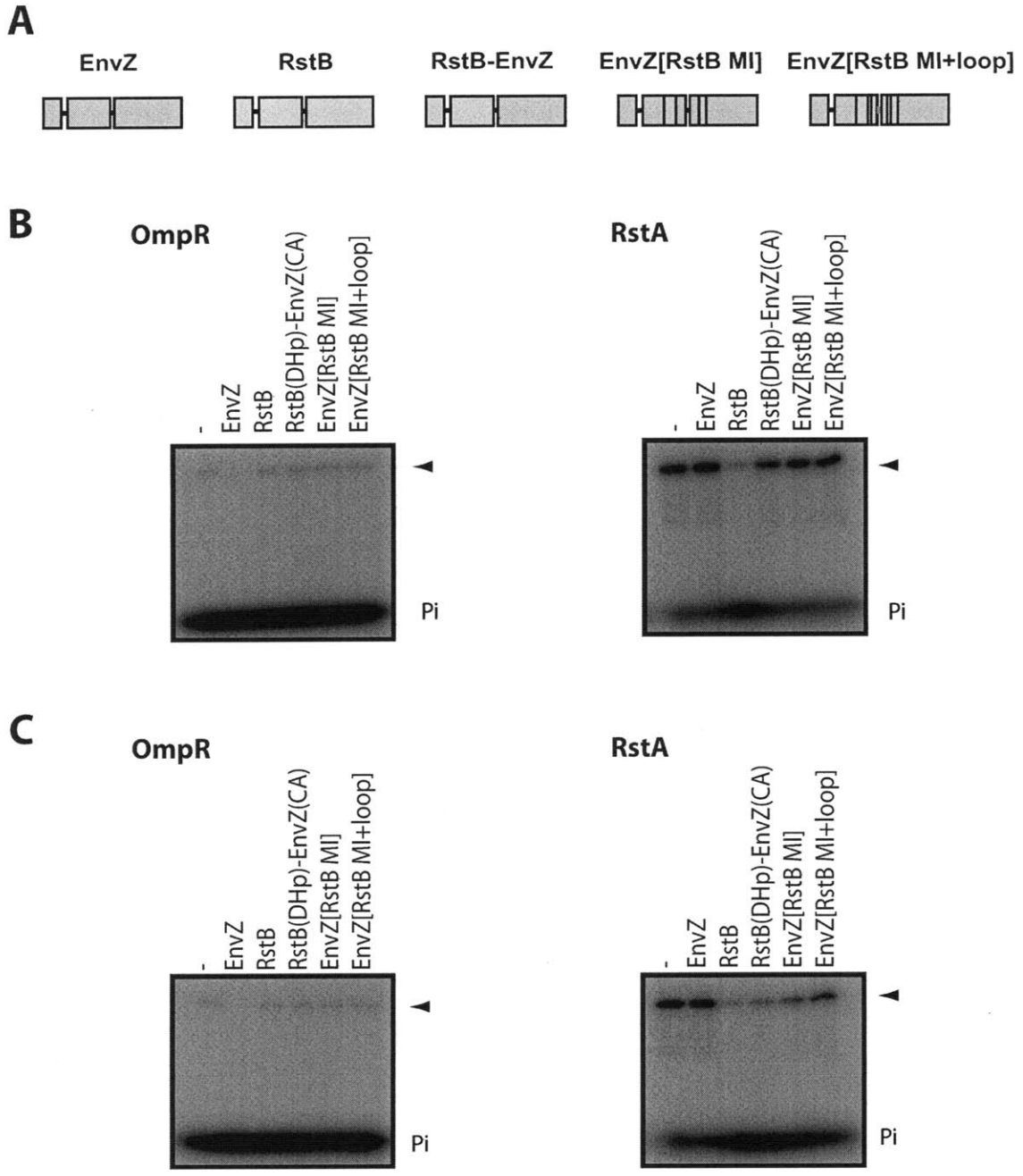

Figure 8: Full domain HK chimerics and MI+loop specificity mutants that fully switch phosphotransfer specificity incompletely switch phosphatase specificity

(A) Schematic of constructs used. See Appendix for full protein sequences.

(B) Phosphatase specificity of full-domain chimerics compared to that of MI and MI+loop mutants. Cognate regulator OmpR and specificity target regulator RstA were phosphorylated, purified, and added either to buffer, wild-type EnvZ, wild-type RstB, or one of three mutant kinases. Phosphatase reaction incubations were for 5 minutes. Black arrowheads denote the position of the phosphorylated band corresponding to the response regulator.

(C) The reactions described in (B) were carried out, with 30 minute phosphatase reaction incubations.

function specifically. Not all DHp domains are capable of acting as phosphatases alone (data not shown), but we found that in addition to EnvZ, the DHp domain of its ortholog in *C. crescentus*, CC1181, does act as a weak phosphatase. We therefore purified the isolated DHp domains of EnvZ and CC1181. We also purified EnvZ DHp domains in which the MI and MI+loop specificity residues identified in Skerker et al., 2008 had been replaced with those of CC1181, and compared the ability of these constructs to dephosphorylate OmpR~P and CC1182~P with wild-type EnvZ and CC1181 DHp domains. While EnvZ DHp specifically dephosphorylates OmpR and CC1181 DHp specifically dephosphorylates CC1182, the mutant constructs do not dephosphorylate OmpR, even at long timepoints. Instead, they now dephosphorylate CC1182, but with slower kinetics than does CC1181 (Fig. 9B). The full-length proteins display similar patterns of phosphatase activity to their isolated DHp domains (Fig. 9C). This result indicates that although the residues dictating phosphotransfer specificity contribute to phosphatase specificity, additional residues in the DHp domain are required. This experiment does not rule out that additional residues in the CA domain may also be required, a hypothesis which is supported by the co-crystal of the subcellular domains of *Thermatoga* HK853 with its cognate response regulator. This crystal structure indicates that there are points of contact between the CA domain and RR receiver domain in the phosphatase interaction (Casino et al., 2009).

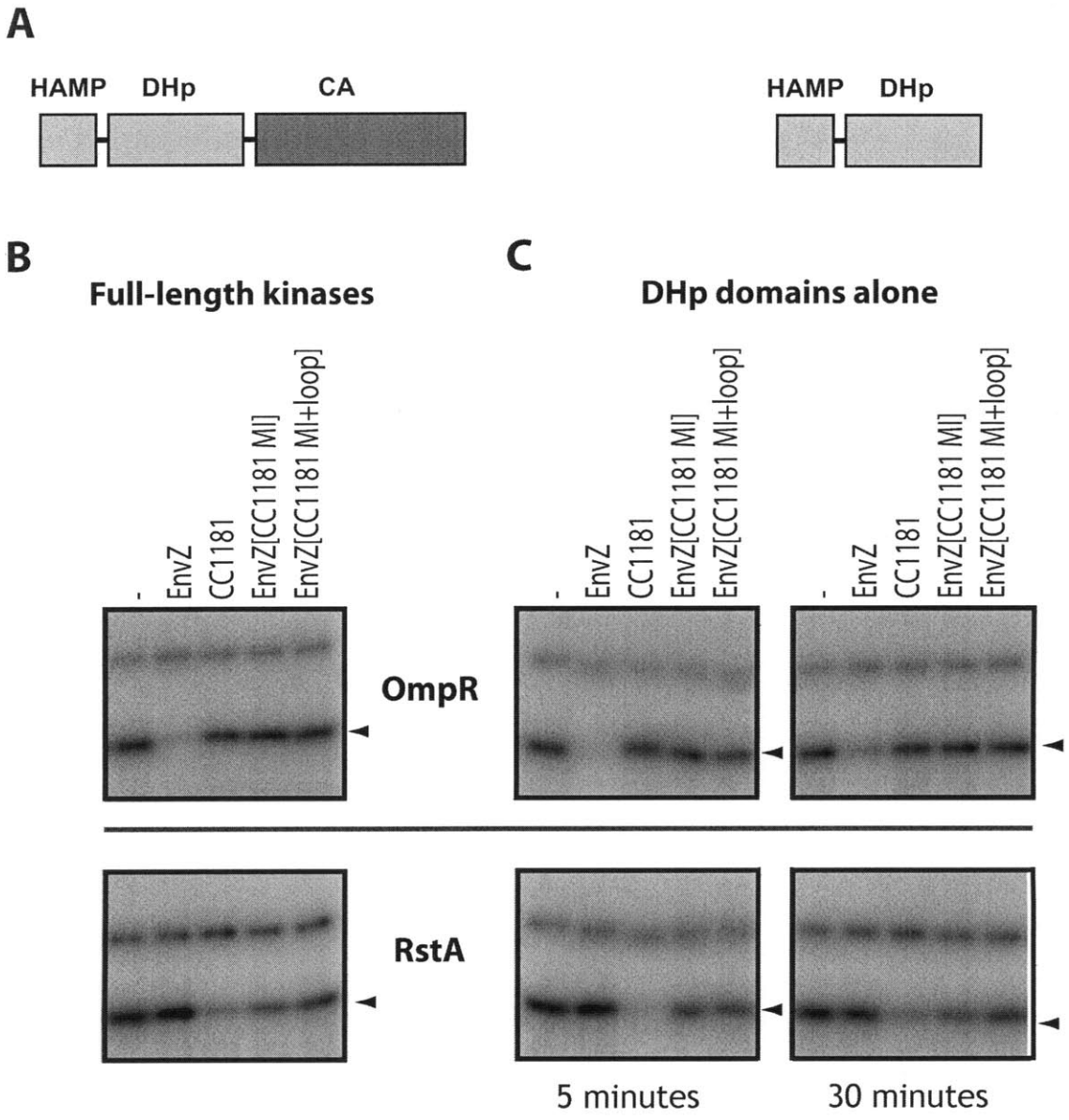

Figure 9: HK MI+loop mutations do not have improved specificity in DHp domains alone compared to full-length kinases

(A) Schematic of the types of construct used: full length kinase and DHp domain alone.

(B) Phosphatase specificity of full-length specificity mutants from E. coli EnvZ to its ortholog in C. crescentus, CC1181. OmpR and and CC1182, the cognate regulator of CC1181, were phosphorylated and purified. Phosphorylated regulator was then incubated with buffer, wild-type EnvZ, wild-type CC1181, EnvZ[CC1181 MI] mutant or EnvZ[CC1181 MI+loop] mutant. Phosphatase reactions were incubated for 11 minutes. Band corresponding tophosphorylated regulator is denoted with a black arrowhead (lower band). Higher band corresponds to phosphorylated loading kinase. Top row shows phosphorylation of OmpR; bottom row shows phosphorylation of RstA.

(C) Phosphatase specificity of DHp domain equivalents of full-length kinases tested in (B). A portion of the N-terminal HAMP domain remains attached, as with the full-length constructs, while the entire CA domain has been removed. Phosphatase reactions were incubated for 5 minutes and 30 minutes.

Response regulators have specific phosphatase activity

To further demonstrate that phosphotransfer determinants are insufficient to dictate phosphatase specificity, we examined a response regulator mutant in which phosphotransfer specificity had been rewired significantly. This approach helped to remove some concerns about maintaining interdomain interactions, since the response regulator effector domain has not been implicated in the catalysis of the dephosphorylation reaction, and some regulators exist without an effector domain (Casino et al., 2009). We assayed the phosphatase specificity of OmpR[CpxR]mut5, a mutant of OmpR in which specificity residues had been replaced with those of CpxR against EnvZ and CpxA (Fig. 5C). These mutations effect only a partial switch in phosphatase specificity – the protein is not dephosphorylated by either EnvZ or CpxA (Fig. 10), while in the phosphotransfer reaction it is not recognized by EnvZ but transferred to by CpxA (Fig. 5D). These data are incomplete – this regulator was not tested for an alteration in its autodephosphorylation rate – however, it is unlikely that this is changed as none of the mutated residues were those identified by Pazy et al., 2009 as controlling autodephosphorylation rate. This experiment further demonstrates that the phosphatase interaction is specific, and that its specificity is determined by residues that overlap with, but are not identical to, the phosphotransfer specificity determinants.

Determinants of histidine kinase autophosphorylation rate

In order to more fully understand how TCS systems maintain specificity, we aimed to understand additional potential pressures on specificity residues. In particular, residues face pressures to maintain interdomain interactions, not just interprotein ones,

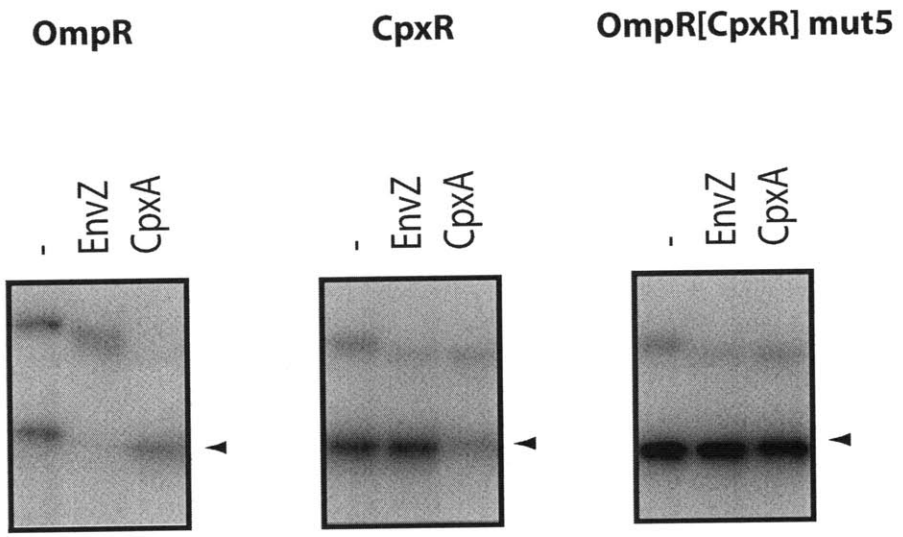

Figure 10: The determinants of RR phosphotransfer and phosphatase specificity overlap but are not identical.
Mutations that confer CpxR phosphotransfer specificity on OmpR do not confer CpxR phosphatase specificity on OmpR.; rather, it partially switched phosphatase specificity so that OmpR[CpxR]mut5 is not dephosphorylated by EnvZ or by CpxA. The three different response regulators have different initial levels of phosphorylation, due to differential recognition by the loading kinase, and possibly different phosphoryl stability. Black arrowheads indicate the position of the phosphorylated band corresponding to the RR (lower band). The upper band represents the loading kinase.

and in HKs and RRs there may be overlap between residues responsible for DHp-CA interactions and those responsible for HK-RR ones. To study this, we asked what residues were responsible for modulating intraprotein interactions in HKs. Both HKs and RRs are composed of two main interacting domains, but residues involved in RR receiver domain-effector domain interactions may be more variable, since effector domains carry out different functions, and are even absent in some RRs. The amino acids responsible for the other major RR activity, autodephosphorylation, have already been studied extensively (Pazy et al., 2009; Thomas et al., 2006).

In addition to coordinating two interactions with the response regulator, the kinase must maintain specific interactions between its DHp and CA domain to autophosphorylate, and to dephosphorylate its response regulator. We aimed to characterize the elements controlling HK autophosphorylation rate in order to provide the basis for understanding and modeling of phosphate flow through TCS systems, and to better allow for directed rewiring of one kinase function without accidentally affecting another. This is a concern because some HK mutants that affect specificity have inadvertently altered autophosphorylation rate as well (Jeffrey Skerker, unpublished). By aiming to alter autophosphorylation rate without damaging the integrity of specific transfer we may eventually parse apart and differentially alter the elements affecting the various conformations and reactions of two-component signaling systems, as much as is possible.

We focused our mutational analyses on the second set of co-varying residues identified by mutual information, which map to α helix 1 of the HK DHp domain. Residues proximal to this region have been implicated in DHp-CA domain interactions

by the structure of *Thermatoga* HK853 (Marina et al., 2005), and, further, if this region has a role in regulating phosphate flow, it might explain why it is identified by MI as co-varying with the response regulator. We made point mutations in EnvZ, both to the corresponding residues in RstB, and to alanine. Though no time-course data were obtained for these mutants, from two data points all of these mutations appear to alter autophosphorylation rate; most decrease it, but one, A231I, increases rate (Fig. 11). These mutations have no effect on the specificity of the kinase for OmpR (data not shown). It is difficult to conclude whether or not the mutations from EnvZ to RstB grant EnvZ an autophosphorylation rate more similar to that of RstB because although *in vitro* RstB has a slower autophosphorylation rate than EnvZ (data not shown), we have found *in vitro* autophosphorylation function to vary with the length of the HAMP linker included in the construct (Jeffrey Skerker, unpublished), though in general we have found EnvZ constructs to be better autokinases than RstB constructs. However, it is notable that all mutations in this region alter autophosphorylation rate.

We then tested the effect of switching specificity in these mutants. A set of three point mutations converts the specificity of EnvZ to that of RstB (Skerker et al., 2008), but these alone also give it a slightly decreased autophosphorylation rate (Jeffrey Skerker, unpublished). A decreased autophosphorylation rate could have confounding effects if this mutant were introduced *in vivo*. Therefore, we aimed to test whether combining autophosphorylation rate mutants with this specificity mutant could have an additive or compensatory effect on its existing autophosphorylation defect. We added one rate-increasing mutation (A231I) and one rate-decreasing mutation (A231I R234K T235K) to the specificity mutant. The resulting two mutants had switched kinase specificity (data

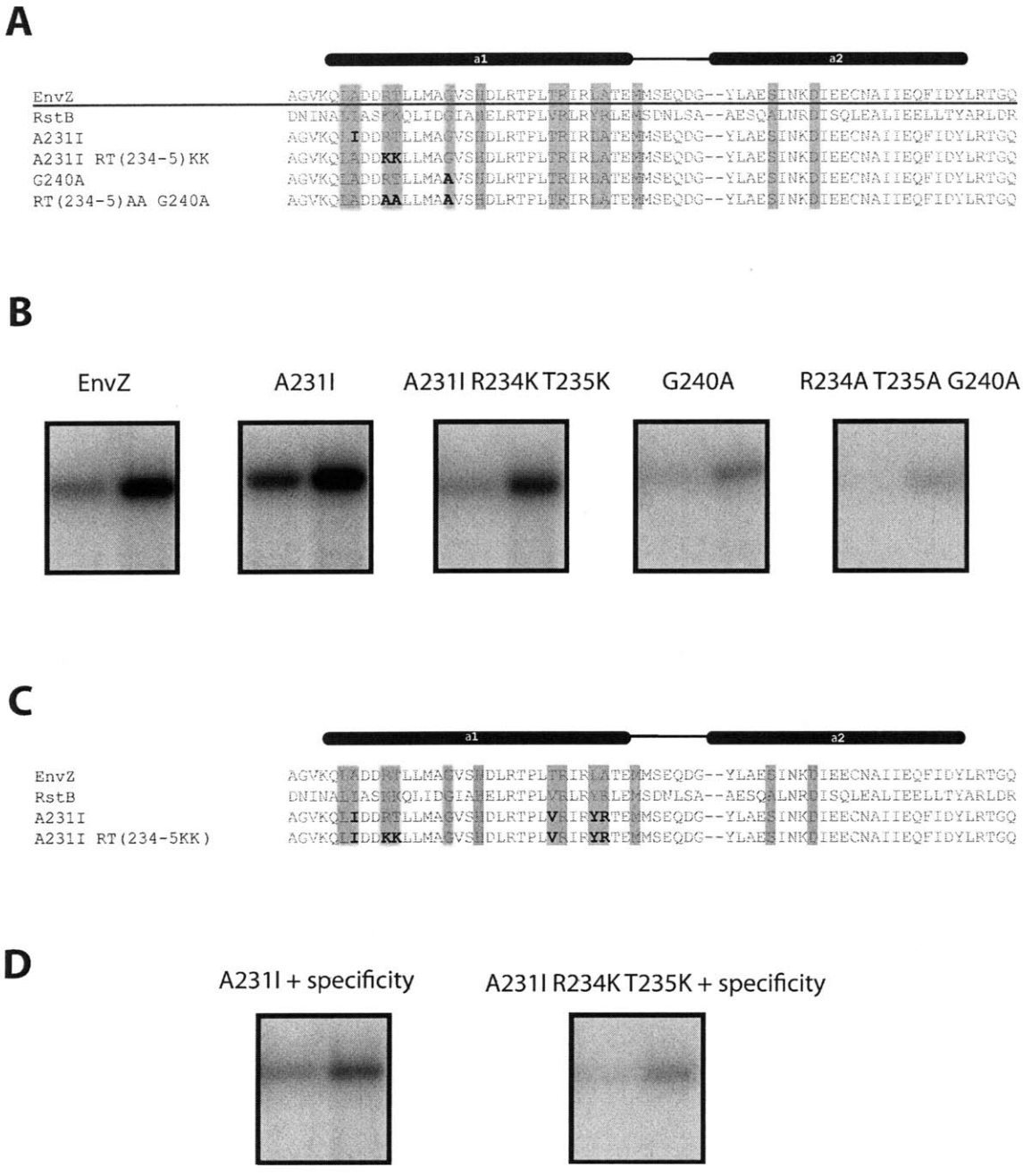

Figure 11: Mutation of mutual information residues affects HK autophosphorylation rate
(A) Amino acid sequences of EnvZ autophosphorylation mutants. See Appendix for complete protein sequences.
(B) Each kinase was incubated with ATP for 10 (left lane) and 60 (right lane) seconds. Black arrowheads indicate the position of the kinase.
(C) Amino acid sequences of EnvZ autophosphorylation and specificity combined mutants. See Appendix for complete protein sequences.
(D) Each kinase was incubated with ATP for 10 (left lane) and 60 (right lane) seconds. Black arrowheads indicate the position of the kinase.

not shown), as well as diminished autophosphorylation rates compared to A231I and A231I R234K T235K alone (Fig. 11). In the context of the rate increasing mutation A231I, however, the autophosphorylation rate of the kinase is now closer to that of wild type than was the autophosphorylation rate of the specificity mutant alone (Fig. 11 and data not shown). In the mutant with an already decreased autophosphorylation rate, A231I R234K T235K, autophosphorylation was further decreased when three EnvZ specificity residues were swapped for those of RstB.

These data offer not only individual insights into the mechanism of two-component system interactions, but, combined, provide the tools for directed rewiring of them. With a more complete understanding of the determinants of signal transfer, we can rationally redirect phosphate flow through a TCS system to understand its regulation, and probe the constraints on two-component proteins as they evolve.

Discussion

Conclusions and future experiments

The maintenance of specificity within paralogous gene families is a fundamental and little-understood problem in biology. In bacterial two-component signaling systems, a histidine kinase must interact specifically with its cognate regulator, avoiding interaction with the many non-cognate response regulators present in the same cell. Paralogous gene families are thought to arise through a process of gene duplication and divergence (Zhang et al., 2003), but it is not understood what pressures drive this divergence and maintain pathway insulation.

In this thesis I identified in a response regulator the majority of the molecular determinants of its phosphotransfer interaction with histidine kinases. I also identified a portion of the molecular determinants of the histidine kinase and response regulator phosphatase interaction, and those responsible for the rates of the histidine kinase autophosphorylation reaction. By understanding the sites that control the rate and direction of phosphate flow, we have provided the tools to rationally rewire TCS systems. Rewiring two-component systems can allow us to specifically alter aspects of a system to understand and model points of regulation, and to probe the constraints laid on TCS systems as they evolve.

More cases are required to demonstrate the generalizability of these findings; I have begun this process by partially switching *E. coli* PhoB phosphotransfer specificity to that of *C. crescentus* PhoB, and *C. crescentus* PhoB specificity to *E. coli* PhoB

specificity, as well as altered the autophosphorylation rate of *C. crescentus* PhoR to more closely resemble that of *S. meliloti* PhoR (data not shown). In addition, further work remains to completely rewire these activities. While we have almost entirely switched RR phosphotransfer specificity in two cases, it appears from these data that another residue or two may be important in determining molecular recognition of the cognate HK; as more two-component systems are identified and added to our multiple sequence alignment, this residue may be identified by mutual information.

More significant molecular determinants of phosphatase specificity remain undiscovered. Our results indicate that, in addition to DHp-RR contacts, interaction between the histidine kinase CA domain and the response regulator may be required to fully switch phosphatase specificity. This is supported by the recent co-crystal structure of an HK and RR, in what appears to be the phosphatase interaction (Casino et al., 2009).

One promising target for determinants of response regulator phosphatase specificity is a patch of residues near those involved in autodephosphorylation rate, the loop region between β4 and α4 (Thomas et al., 2006). We have previously ruled out these residues as being involved in phosphotransfer specificity. However, three of these positions were identified as potential contact residues in the Spo0B/Spo0F co-crystal, and three were found by MI analysis to co-vary highly with HK residues. More tellingly, two positions in this loop are identified as contacting the HK linker region between the DHp and CA domains in the co-crystal structure (Casino et al., 2009). We have taken OmpR[RstA MI+loop] and OmpR[CpxR MI+loop] mutants similar to those tested in Figure 5 and in addition to these mutations exchanged the β4-α4 loop for the corresponding one in RstA or CpxR. The addition of this second loop does not alter

phosphotransfer specificity compared to the background mutations alone in either case; however, it does appear to alter the specificity of the OmpR[RstA] mutant in the phosphatase reaction so that it is better dephosphorylated by RstB (data not shown). These data are preliminary, and inferred from an phosphotransfer experiment in which both phosphotransfer and dephosphorylation were permitted to occur. Controls to ensure that simply that autodephosphorylation rate has not been increased have not been done; however, an increase in autodephosphorylation rate would be unlikely to have an effect, since in both OmpR and RstA this rate is slow relative to the timescale of the experiment. The HK-RR co-crystal identifies another loop region, between β3 and α3, as contacting the CA domain, and this represents another potential site of partner recognition in the phosphatase interaction. The HK-RR co-crystal structure provides insight into potential determinants of phosphatase specificity in the histidine kinase as well, identifying several residues in the linker region between the DHp and CA domains, and in the CA domain that may interact with these residues in the regulator.

Finally, these mutants must be re-introduced into bacteria to demonstrate that molecular recognition is sufficient to dictate specificity *in vivo* as well. This has already been shown in several cases for HK phosphotransfer specificity (Skerker et al., 2008), indicating that this may also be the case for RR phosphotransfer, and HK and RR phosphatase specificity.

Understanding the evolution of specificity in paralogous gene families

This research provides tools with which we can better understand evolution in paralogous gene families. In particular, it has been widely hypothesized, but never

measured, that loss of pathway insulation is detrimental to a cell. The hypothesis implicit in this idea is that sequence space is, in general, small, and that negative selection against cross-talk is the driving force in the divergence of paralogous systems, rather than positive selection or neutral drift.

Such hypotheses are difficult to test, but a range of investigations have approached this problem by offering evidence in support of negative selection. A small-scale study examined a pseudogene in *Yersinia pestis*. *Y. pestis*, which forms biofilms inside fleas, evolved from *Y. pseudotuberculosis*, which forms biofilms in other contexts but fails to do so inside insects. *Y. pseudotuberculosis* contains a gene, *rscA*, whose ortholog in *Y. pestis* is a pseudogene . Replacement of the pseudogene in *Y. pestis* with *rscA* now produces a mutant unable to form biofilms in insects, indicating that the active form of this gene had been selected against (Sun et al, 2008).

Several systems approaches to testing negative selection have also been employed. Zarrinpar et al (2003) studied the specificity of yeast SH3 domains. In *S. cerevisiae* there are 27 identified SH3 domains. Zarrinpar et al examined the interaction of one such domain, that of Sho1 with its target, the kinase Pbs2 and reported that the Sho1 SH3, but none of the 26 other SH3 domains can interact with Pbs2, while SH3 domains from non-yeast proteins can interact with Pbs2. This suggests that negative selection has had a role in forming the specificity of yeast SH3 domains; Sho1 has evolved in the absence of non-yeast SH3 domains, and so overlap in their specificities has not been selected against. A microarray study of the binding specificity of mouse PDZ domains found that their specificities are spread out, rather than falling into clusters,

indicating that cross-reactivity is selected against across the proteome (Stiffler et al, 2007).

All of these studies adopt a retrospective approach, rather than attempting to recreate a process of evolution. On the other end of the spectrum, researchers have used experimental evolution to map fitness landscapes (Poelwijk et al., 2007; Romero and Arnold, 2009). In such experiments, an understanding of the accessible evolutionary trajectories of a protein is achieved, which offers insight into the ease or difficulty with which a protein sequence can reach a neighboring sequence through neutral drift. If a single change occurs in an amino acid sequence, the new sequence is separated from its parent sequence by a distance of one amino acid. Though these individual differences are small, the space of possible proteins is very large. Even with mutations occurring in a step-wise fashion, depending on which mutations occur, and in which order, vastly different proteins can be reached by different evolutionary trajectories. These different proteins give different fitnesses to the organism. By trying to evolve proteins in a directed fashion, sequencing mutations that arise and measuring fitness levels, an understanding of a fitness landscape emerges (Poelwijk et al., 2007; Romero and Arnold, 2009).

Future directions: a novel approach to examining the fitness consequence of cross-talk between signaling systems

These studies offer important contributions to our understanding of evolution, but leave unanswered questions. In particular, the consequence of cross-talk as it might occur during the course of evolution has not been directly measured. What is the resulting degree of cross-talk, and the fitness defect incurred, on the scale of one or two amino acid

changes? With our knowledge of TCS systems, we can make small, rational changes in a TCS protein, and measure both the consequences of these changes to the affected systems, and to cellular fitness.

In future work, we aim to quantify the fitness cost of small changes in a two-component signaling system. If a point mutation is introduced into, for example, a histidine kinase, this mutation can have several possible consequences for signal transfer. Given our new understanding of the residues controlling the various activities and interactions of two-component proteins, we can control and prevent alterations to activities other than specific interaction with the response regulator have occurred by this mutation. A specificity mutation in an HK may have no effect on the specificity of the kinase (Fig. 12A); it may introduce new specificity for a non-cognate regulator (Fig. 12B; it may decrease the kinase's specificity for its cognate regulator (Fig. 12C); or both (B) and (C) may occur simultaneously (Fig. 12D). In the case of (A) we in general anticipate that no fitness defect to be incurred. In (B) and (C), depending on the cellular importance of the signaling systems involved and the degree of change, we expect some fitness defect to be incurred. This experiments will explore to what degree negative selection against cross-talk is biologically relevant – perhaps no fitness defect will be incurred by a slight introduction of cross-talk, or perhaps loss of signaling integrity is often more detrimental than the introduction of cross-talk.

Figure 12: There are different biological outcomes to mutation in a specificity residue
(A) When a mutation is introduced into the specificity residues of a two-component protein, there are four possible effects on specificity. In (A), both pathway insulation and integrity of phosphotransfer are preserved.
(B) Phosphotransfer with the cognate protein may be maintained, but unwanted crosstalk to a non-cognate protein may be introduced.
(C) Phosphotransfer to the cognate protein may be decreased, but pathway insulation may be maintained.
(D) Both phosphotransfer with the cognate protein and pathway insulation may be compromised.

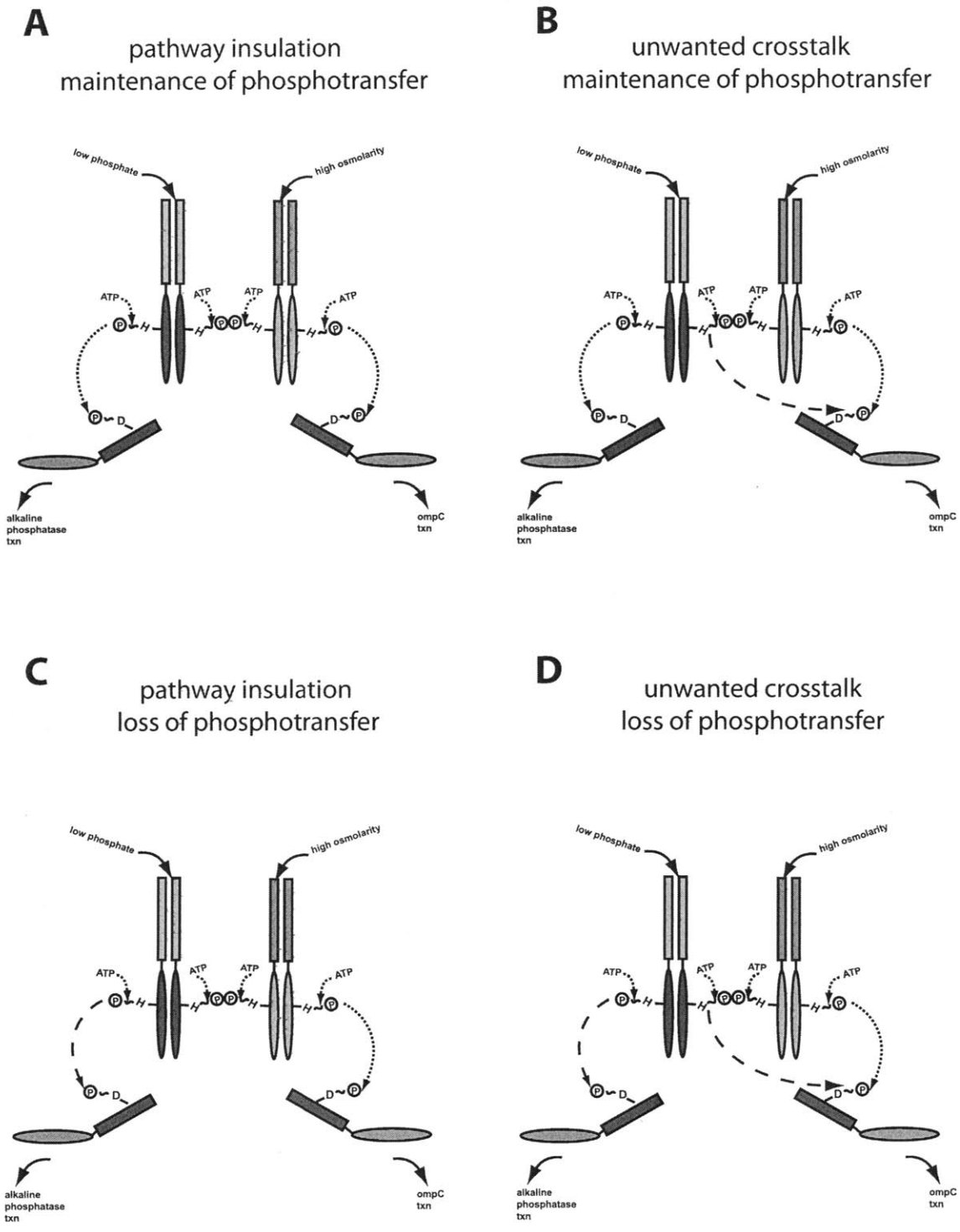
A
pathway insulation
maintenance of phosphotransfer
low phosphate
high osmolarity
ATP
ATP
ATP
ATP
alkaline phosphatase txn
ompC txn
B
unwanted crosstalk
maintenance of phosphotransfer
low phosphate
high osmolarity
ATP
ATP
ATP
ATP
alkaline phosphatase txn
ompC txn
C
pathway insulation
loss of phosphotransfer
low phosphate
high osmolarity
ATP
ATP
ATP
ATP
alkaline phosphatase txn
ompC txn
D
unwanted crosstalk
loss of phosphotransfer
low phosphate
high osmolarity
ATP
ATP
ATP
ATP
alkaline phosphatase txn
ompC txn

Figure 12

Experimental design: measuring in vivo fitness costs associated with introduction of cross-talk and loss of signaling integrity

Several components are required to carry out this experiment, but each has been tested previously in a different context. Competition experiments have widely been used to measure cellular fitness (Hegreness et al., 2008). Two fluorescent markers can be integrated into the genomes of two bacterial strains, wild type and a mutant. As the two strains are co-cultured, the relative fitness of the mutant strain compared to wild type can be measured by assaying the fluorescent marker (Fig. 13).

Using our understanding of TCS systems we can engineer mutations to alter the specificity of a kinase. We have, *in vitro*, analyzed and rewiring the relevant elements that constrain a two-component system interaction. Some constraints on TCS systems are not directly relevant for understanding specificity because they depend on the modular domains that vary widely between TCS systems, such as changes in system input and output. There are also system-specific constraints involving additional proteins; several TCS inhibiting or activating proteins have been found to bind to the intracellular domains of the kinase or to the regulator (Paul et al., 2008; Wang et al., 1997; Atkinson et al., 1994; Stephenson and Perego, 2002). These regulatory proteins themselves often appear to be derived from TCS proteins.

Others are more relevant, though still not directly involved in specificity, such as the balance between kinase and phosphatase activities in the HK, and the autodephosphorylation rate of the RR. We have partially characterized these, in order to be better able to alter specificity separately from other HK and RR functions. Primarily, however, we have focused on the constraints on the molecular interactions between the

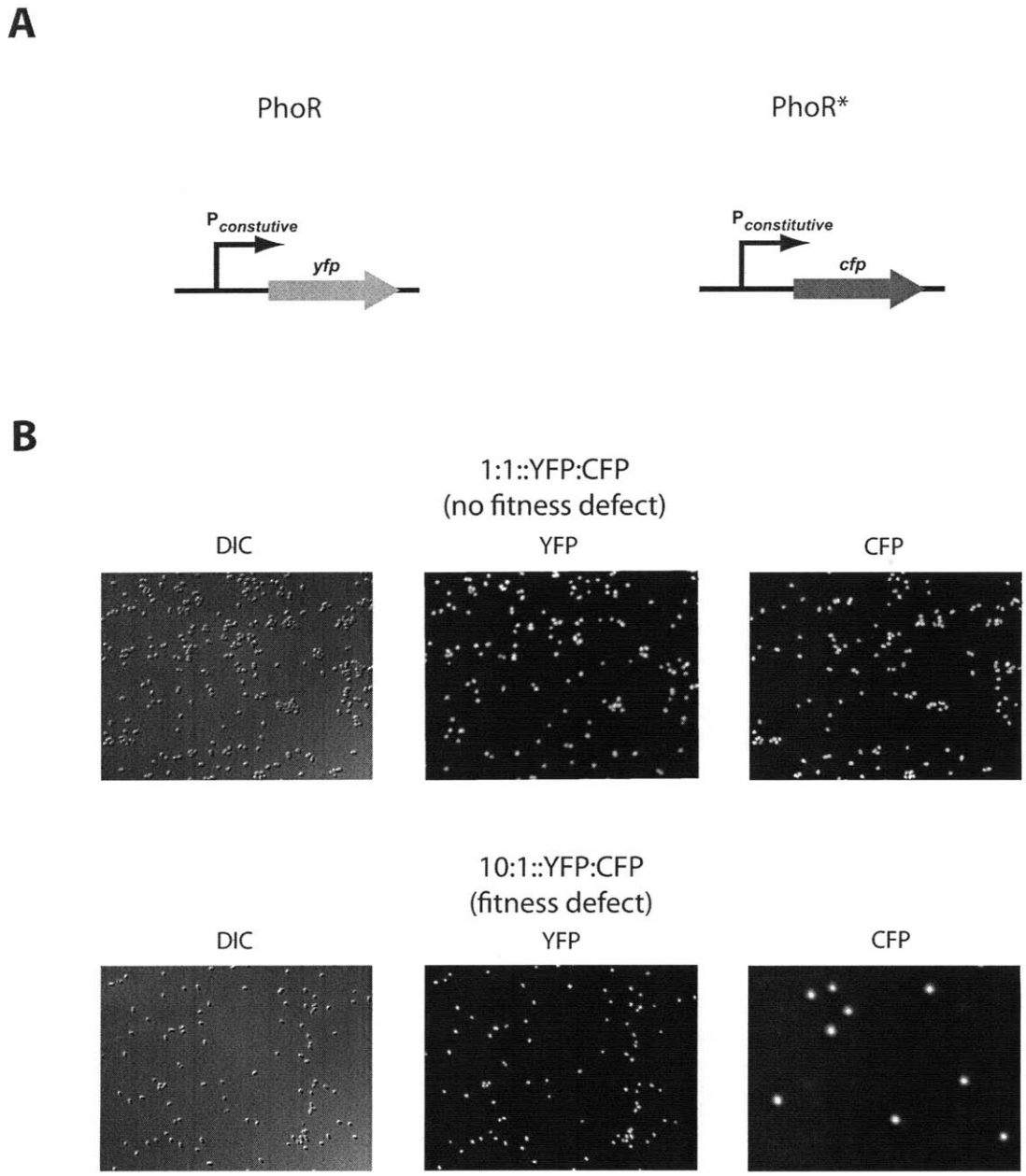

Figure 13: Proportions of different strains in co-culture are readily measurable
(A) Two strains with different fluorescent markers, YFP and CFP, integrated into the genome have been used in competition experiments to report the ratio of the two strains over time; this provides a measure of the relative fitness of the strains (Hegreness et al., 2008).
(B) Ratios between two strains can be accurately reported. For each of the ratios, the same microscope view is shown in DIC and with CFP and YFP filters.

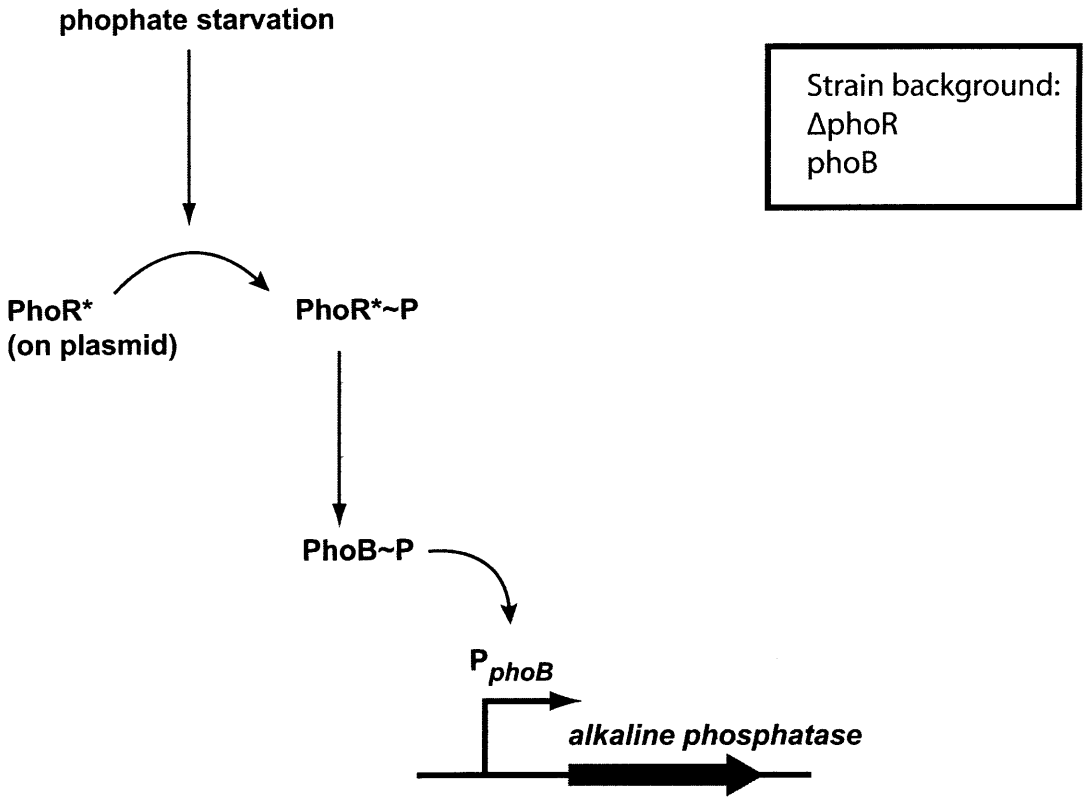

Figure 14: Integrity of the PhoR-PhoB signaling pathway can be reported by alkaline phosphatase acitivity.
Reporting systems can be used to quantify phosphotransfer to cognate regulator. Several TCS systems are especially amenable to reporters because of their biological relevance. In particular, output of the PhoR/B system can be measured by assaying for alkaline phosphatase activity as in (Spira and Ferenci, 2007; Fisher et al., 1995; Chaffin et al., 1999).

two proteins themselves. As these interacting residues evolve, they may be constrained by the need to maintain structure, the need to maintain signaling rate within that system, and the need to prevent signaling between paralogous systems. Though it may not be wholly generalizable, the lab has demonstrated in numerous cases that specificity can be drastically altered without compromising HK or RR structure.

We can also, with a careful choice of TCS system, assay the degree of signal transfer of the mutated system *in vivo*. In this work, we have focused particular attention on the EnvZ/OmpR system and PhoR/B systems in several different bacteria. The PhoR/B system directly regulates the transcription of alkaline phosphatase, commonly used as a reporter, and readily assayed and quantified (Fisher et al., 1995; Chaffin and Rubens, 1999) (Fig. 14).

This set-up will enable us to ask what the fitness consequence is of loss of signaling integrity. We can examine how great a loss of specificity a TCS system can tolerate, and how much unwanted cross-talk a set of TCS systems can tolerate. Using this basic design, we can introduce different types and degrees of specificity mutations, in addition to mutations that simply decrease phosphate flow or introduce promiscuity into a kinase. A chart detailing an example of how in one such an experiment fitness will be computed is given in Figure 15. In this example, a mutation is introduced into PhoR to change its specificity slightly so that while it will still transfer to PhoB, it now also transfers to OmpR. In such a scenario, we can ask, if a fitness defect is incurred by this mutation, whether it is due to the loss of signaling integrity (Fig. 12C), introduction of cross-talk (Fig. 12B) or both (Fig. 12D). Loss of signaling integrity is a concern not just because introducing new specificity may remove part of the old specificity, but also

A

defect	relative fitness
phoR	1.0
*phoR**	0.8
ΔompR	0.7
ΔphoB	0.5

B

where x is the quantity of fitness defect due to loss of pathway insulation and y is the quantity of fitness defect due to loss of phosphotransfer, x+y=0.2

C

phoR* insulation	+	−	−
phoR* phosphotransfer	−	+	−
genotype		*relative fitness*	
ΔphoR + phoR	1	1	1
*ΔphoR + phoR**	0.8	0.8	0.8
ΔphoR ΔphoB + phoR	0.5	0.5	0.5
*ΔphoR ΔphoB + phoR**	0.5	0.8*0.5	(1-x)*0.5
ΔphoR +phoR	1	1	1
*ΔphoR + phoR**	0.8	0.8	0.8
ΔphoR ΔompR + phoR	0.7	0.7	0.7
*ΔphoR ΔompR + phoR**	0.8*0.7	0.7	(1-y)*0.7

Figure 15: Proposed experiment: dissection of non-cognate cross-talk and loss of cognate phosphotransfer
(A) Hypothetical relative fitnesses of mutations. *phoR** contains a specificity mutation that may introduce cross-talk to OmpR *in vivo.*
(B) The resulting fitness defect may have two components.
(C) Genetic dissection of the different contributions to a fitness defect.
When a specificity mutation, *phoR**, is introduced in combination with the removal of either its cognate partner, PhoB, or its non-cognate target, OmpR, different fitness levels, dependent on the fitnesses of the individual mutations, are expected.

because the kinase must now divide its signal in some proportion between its cognate and new target regulator. In an *E. coli* strain in which wild type *phoR* has been knocked out, either wild type or mutant *phoR* can be reintroduced. By measuring the fitness of these mutants in combination with those in which *phoB*, the cognate of *phoR,* or *ompR*, the target of specificity mutant *phoR,* have been knocked out, we may be able to differentiate between the scenarios in Figure 12 B, C, and D. We can then ask further questions. For example, if loss of signaling integrity causes a great fitness defect, we can then probe, through rational mutation, what the effect of increasing PhoR autophosphorylation rate in this context is on the fitness of that bacterium.

By providing concrete examples of the effect of small changes on cellular fitness, we can understand the cost of loss of signaling integrity and introduction of cross-talk. The work presented here provides the basis for these studies. By identifying the specificity determinants of the interactions between histidine kinases and response regulators, we can rationally rewire their specificities, and perhaps in the future, differentially rewire kinase and phosphatase activities, as well as autophosphorylation rate and autodephosphorylation rates, with the work of Pazy et al., 2009 and Thomas et al., 2006. The ability to rewire different aspects of a two-component system will not only allow us to study systems-level evolution in a paralogous gene family, but will aid in the modeling and study of phosphate flow through individual signal transduction systems.

Materials and Methods

Computational analyses

Computational analyses were performed as described previously in Skerker et al, 2008. Putative cognate two-component proteins were identified in sequenced bacterial genomes by selecting adjacent genes predicted to encode a histidine kinase (HK) and a response regulator (RR), by using custom PERL scripts. The sensor and transmembrane domains of histidine kinases were omitted, as were the effector domains of response regulators. This retained the dimerization and histidine phosphotransfer (DHp) domain and catalytic and ATP-binding (CA) domain of the histidine kinases and the receiver domain (RD) of the response regulators. The domains of each cognate pair were concatenated into a single sequence and aligned with PCMA (Pei et al., 2003) with some manual adjustment. Analysis of mutual information was performed using published software (Fodor and Aldrich, 2004). Columns in the alignment containing more than 10% gaps were eliminated from consideration. We also ensured that no two sequences in the alignment had greater that 90% identity to one another. This step helped to minimize the detection of amino acids that co-vary due to phylogenetic relationships rather than functional relationships.

Specificity-determining residues were mapped onto the Spo0B:Spo0F crystal structure (PDB: 1F51) using PyMOL (DeLano, 2002). The asymmetric unit contains four Spo0B and four Spo0F molecules. For clarity, only the four-helix bundle of one Spo0B dimmer in complex with one Spo0F molecule is shown in (Fig. 3A). Distances between residues were measured as the shortest distance between any nonhydrogen atoms.

Cloning and protein purification

Cloning and protein purification were carried out as previously described (Skerker et al., 2005; Skerker et al., 2008). All protein constructs contained only subcellular domains. We have found that, in some cases, removing the full HAMP linker domain prevents HKs from phosphorylating *in vitro* (Jeffrey Skerker, unpublished). In these cases, the subcellular domains of the kinase have been purified with part or all of the HAMP domain attached. All proteins are derived from the genome of *E. coli* K12, except for CC1181 and CC1182 which were derived from *Caulobacter crescentus* CB15N

Phosphorylation and phosphotransfer assays

In vitro analyses of phosphorylation and phosphotransfer were performed as previously described (Skerker et al. 2005). Briefly, histidine kinases in 10 mM HEPES-KOH (pH 8.0), 50 mM KCl, 10% glycerol, 0.1 mM EDTA, 2 mM DTT, 5 mM MgCl2 were autophosphorylated with 500 μM ATP and 0.5 μCi/μl [γ^{32}P]-ATP (from a stock at ~6000 Ci/mmol, Amersham Biosciences), and then subsequently incubated with a response regulator. Kinase and regulator were present at 2.5 μM each. Reactions were incubated at room temperature, and products were then separated by 10% SDS-PAGE, exposed to a phosphor screen, and quantified by using a Typhoon 9400 Scanner (GE Healthcare) with ImageQuant 5.2.

Phosphatase assays

In vitro analyses of response regulator dephosphorylation were performed as follows. Response regulators in 10 mM HEPES-KOH (pH 8.0), 50 mM KCl, 10% glycerol, 0.1 mM EDTA, 2 mM DTT, 5 mM MgCl2 (HKEDG) were incubated with a small amount of loading kinase and 0.5 [γ^{32}P]-ATP (from a stock at ~6000 Ci/mmol, Amersham Biosciences). Total reaction volume was 50 µl. Final concentrations in the loading reaction were 5 µM response regulator, 0.1 µM loading kinase, and 0.05 µCi/ul [γ^{32}P]-ATP. Reactions were allowed to proceed for between 1 and 2 hours, at 30° C. After incubation, remaining ATP was hydrolyzed by addition of 7 µl hexokinase and its substrate, 4.7 mM (D)-glucose and incubated at room temperature for 7 minutes. EDTA was generally added to a final concentration of 12.4 mM. Free phosphate was removed by washing four times in 10k Nanosep columns with 450 µl of HKEDG.

Reaction volume was brought to 50 µl with HKEDG and 5 µl of the phosphorylated regulator reaction was incubated with 5 µl of histidine kinase in HKEDG and ADP. Final concentrations in the phosphates reaction were <2.5 µM RR~P (approximate), 2.5 µM histidine kinase, <0.1 µM loading kinase (approximate), and 0.5 mM ADP. Phosphatase reactions were incubated at room temperature for the time specified and reaction was stopped with 3.5 µl of 4X sample buffer (500 mM Tris [pH 6.8], 8% SDS, 40% glycerol, 400 mM β-mercaptoethanol) and stored on ice until loaded. Products were separated by 10% SDS-PAGE, exposed to a phosphor screen, and quantified using a Typhoon 9400 Scanner (GE Healthcare) with ImageQuant 5.2.

References

Alves, R. and Savageau, MA. "Comparative analysis of prototype two-component systems with either bifunctional or monofunctional sensors: differences in molecular structure and physiological function." *Molecular Microbiology* (2003). 48(1): 25-51.

Aravind, L. and Ponting, C.P. (2006). The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins. FEMS Microbiology Letters. *176(1)*, 111-116.

Atkinson, M.R., Kamberov, E.S., Weiss, R.L., and Ninfa, A.J. (1994). Reversible uridylylation of the *Escherichia coli* PII signal transduction protein regulates its ability to stimulate the dephosphorylation of the transcription factor nitrogen regulator I (NRI or NtrC). J. Biol. Chem. *269*, 28288-28293.

Bachhawat, P. and Stock, A.M. (2007). Crystal structures of the receiver domain of the response regulator PhoP from *Escherichia coli* in the absence and presence of the phosphoryl analog beryllofluoride. Journal of Bacteriology. *189(16)*, 5987-5995.

Baikalov, I., Schroder, I., Kaczor-Greskowiak, M., Grzeskowiak, K., Gunsalus, R.P., and Dickerson, R.E. (1996). Structure of the *Escherichia coli* response regulator NarL. Biochemistry. *35(34)*, 11053-11061.

Brown, J.L., Bussey, H., and Stewart, R.C. (1994). Yeast Skn7p functions in a eukaryotic two-component regulatory pathway. EMBO J. *13*, 5186-5194.

Cai, S.J., and Inouye, M. (2003) Spontaneous subunit exchange and biochemical evidence for trans-autophosphorylation in a dimmer of Escherichia coli histidine kinase (EnvZ). J. Mol. Biol. *329(3)*, 495-503.

Chaffin, D.O. and Rubens, C.E. (1999). Blue/white screening of recombinant plasmids in Gram-positive bacteria by interruption of alkaline phosphatase gene (*phoZ*) expression. Gene *219(1-2)*, 91-99.

De Wulf, P. and Lin, E.C. (2000). Cpx two-component signal transduction in *Escherichia coli*: excessive CpxR~P levels underlie CpxA* phenotypes. J. Bacteriol. *182*, 1423-1426.

Fisher, S.L., Jiang, W., Wanner, B.L., and Walsh, C.T. (1995). Cross-talk between the histidine protein kinase VanS and the response regulator PhoB: Characterization and identification of a VanS domain that inhibits activation of PhoB. J. Bact. Chem. *270*, 23143-23149.

Freeman, J.A., Lilley, B.N., and Bassler, B.L. (2000). A genetic analysis of the functions of LuxN: a two-component hybrid sensor kinase that regulates quorum sensing in *Vibrio harveyi*. Mol. Microbiol. *35(1)*, 139-149.

Grebe, TW and Stock, JB. "The Histidine Protein Kinase Superfamily," *Advances in Microbial Physiology* 1999;41:139-227.

Hegreness, M.*, Shoresh, N.*, Damian, D., Hartl, D., and Kishony, R. (2008). Accelerated evolution of resistance in multi-drug environments. PNAS. *105*, 13977-13981.

Hsing, W. and Silhavy, T.J. (1997). Function of conserved histidine-243 in phosphatase activity of EnvZ, the sensor for porin osmoregulation in *Escherichia coli*. J. Bacteriol. *179(11)*, 3729-3735.

Hsing, W., Russo, F.D., Bernd, K.K., and Silhavy, T.J. (1998). Mutations that alter the kinase and phosphatase activities of the two-component sensor EnvZ. J. Bacteriol. *180(17)*, 4538-4546.

Hulko, M., Berndt, F., Gruber, M., Linder, J.U., Truffault, V., Schultz, A., Martin, J., Schultz, J.E., Lupas, A.N., and Coles, M. (2006). The HAMP domain structure implies helix rotation in transmembrane signaling. Cell. *126(5)*, 929-940.

Kenney, L.J. (2010). How important is the phosphatase activity of sensor kinases? Current Opinion in Microbiology. *13(2)*, 168-176.

Ketela, T., Brown, J.L., Stewart, R.C., and Bussey, H. (1998). Yeast Skn7p activity is modulated by the Sln1p-Ydp1p osmosensor and contributes to regulation of the HOG1 pathway. Mol. Gen. Genet. *259*, 372-378.

Khorchid, A. and Ikura, M. (2006). Bacterial histidine kinase as signal sensor and transducer. The Int. J. of Biochem. & Cell Biol. *38(3)*, 307-312.

Laub, M.T. and Goulian, M. (2007). Specificity in two-component signal transduction pathways. Annual Review of Genetics. *41*, 121-145.

Laub, M.T., Biondi, E.G., Skerker, J.M. (2006). Phosphotransfer profiling: systematic mapping of two-component signal transduction pathways and phosphorelays. Molecular Microbiology. *59*, 386-401.

Lee, S.Y., Cho, H.S., Pelton, J.G., Yan, D., Henderson, R.K., King, D.S., Huang, L.-S., Kustu, S., Berry, E.A., and Wemmer, D.E. (2001). Crystal structure of an activated response regulator bound to its target. Nat. Struct. Biol. *8*, 52-56.

Lewis, R.J., Brannigan, J.A., Muchova, K., Barak, I., and Wilkinson, A.J. (1999). J. Mol. Biol. *294*, 9-15.

Li, S., Ault, A., Malone, C.. Raitt, D., Dean, S., Johnston, L.H., Deschenes, R.J., and Fassler, J.S. (1998). The yeast histidine protein kinase, Sln1p, mediates phosphotransfer to two response regulators, Ssk1p and Skn7p. EMBO J. *17*, 6952-6962.

Mizuno, T. (1997). Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of *Escherichia coli*. DNA Research. *4*, 161-168.

Paul, R., Jaeger, T., Abel, S., Wiederkehr, I., Folcher, M., Biondi, E.G., Laub, M.T., and Jenal, U. (2008). Allosteric regulation of histidine kinases by their cognate response regulator determines cell fate. Cell. *133(3)*, 452-461.

Pazy, Y., Wollish, A.C., Thomas, S.A., Miller, P.J., Collins, E.J., Bourret, R.B., and Silversmith, R.E. (2009). Matching biochemical reaction kinetics to the timescales of life: Structural determinants that influence the autodephosphorylation rate of response regulator proteins. Journal of Molecular Biology. *392(5)*, 1205-1220.

Posas, F., Wurgler-Murphy, S.M., Maeda, T., Witten, E.A., Thai, T.C., and Saito, H. (1996). Yeast HOG1 MAP kinase cascade is regulated by a multistep phosphorelay mechanism in the SLN1-YPD1-SSK1 "two-component" osmosensor. Cell *86*, 865-875.

Skerker, J.M., Prasol, M.S., Perchuk, B.S., Biondi, E.G., and Laub, M.T. (2005). Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. PLOS Biology. *3(10)*, e334, 0001-0019.

Slauch, J.M., Garrett, S., Jackson, D.E., and Silhavy, T.J. (1988). EnvZ functions through OmpR to control porin gene expression in *Escherichia coli* K-12. J. Bacteriol. *170*, 439-441.

Socolich, M; Lockless, SW; Russ, WP; Lee, H; Gardner, KH, and Ranganathan, R. "Evolutionary information for specifying a protein fold," *Nature*. Vol 437/22 September 2005. 512-518

Spira, B. and Ferenci, T. (2007). Alkaline phosphatase as a reporter of σ^S levels and *rpoS* polymorphisms in different *E. coli* strains. Archives of Microbiology. *189(1)*, 43-47.

Stephenson, S.J. and Perego, M. (2002). Interaction surface of the Spo0A response regulator with the Spo0E phosphatase. Mol. Micro. *44(6)*, 1455-1467.

Stock, A.M. and Guhaniyogi, J. (2006). A new perspective on response regulator activation. J. Bacteriol. *188*, 7328-7330.

Stock, A.M., Mottonen, J.M., Stock, J.B., and Schutt, C.E. (1989). Three-dimensional structure of CheY, the response regulator of bacterial chemotaxis. Nature. *337*, 745-749.

Sun, Y; Hinnebusch, BJ; Darby, C. "Experimental evidence for negative selection in the evolution of a *Yersinia pestis* pseudogene. PNAS vol 105, no 23: 8097-8101. June 10, 2008.

Thomas, S.A., Brewster, J.A., and Bourret, R.B. (2008). Two variable active site residues modulate response regulator phosphoryl group stability. Mol. Microbiol. *69*, 453-465.

Tomomori, C., Tanaka, T., Dutta, R., Park, H., Saha, S.K., Zhu, Y., Ishima, R., Liu, D., Tong, K.I., Kurokawa, H., Qian, H., Inouye, M., and Ikura, M. (1999). Solution structure of the homodimeric core domain of *Escherichia coli* histidine kinase EnvZ. Nature Struct. Biol. *6(8)*, 729-734.

Wang, L., Grau, R., Perego, M., and Hoch, J.A. (1997). A novel histidine kinase inhibitor regulating development in *Bacillus subtilis*. Genes & Development. *11*, 2569-2579.

Zhang, J. (2003) Evolution by gene duplication: an update. TRENDS in Ecology and Evolution. *18(6)*, 292-298.

Zhu, Y., Qin, L., Yoshida, T., and Inouye, M. (2000). Phosphatase activity of histidine kinase EnvZ without kinase catalytic domain. PNAS. *97(14)*, 7808-7813.

Appendix

Complete protein sequences of mutants used

>OmpR
LQENYKILVVDDDMRLRALLERYLTEQGFQVRSVANAEQMDRLLTRESFHLMV
LDLMLPGEDGLSICRRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNP
RELLARIRAVLRRQANELPGAPSQEEAVIAFGKFKLNLGTREMFREDEPMPLTSG
EFAVLKALVSHPREPLSRDKLMNLARGREYSAMERSIDVQISRLRRMVEEDPAH
PRYIQTVWGLGYVFVPDGSKA*

>OmpR[RstA]mut1
LQENYKILVVDDDMEVRALLEAYLTEQGFQVRSVANAEQMDRLLTRESFHLMV
LDLMLPGEDGLSICRRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNP
RELLARIRAVLRRQANELPGAPSQEEAVIAFGKFKLNLGTREMFREDEPMPLTSG
EFAVLKALVSHPREPLSRDKLMNLARGREYSAMERSIDVQISRLRRMVEEDPAH
PRYIQTVWGLGYVFVPDGSKA*

>OmpR[RstA]mut2
LQENYKILVVDDDMEVRALLEAYLTEQGFQVRSVANAEQMDRLLTRESFHLMV
LDLMLPGEDGLSICRRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKTTPP
RELLARIRAVLRRQANELPGAPSQEEAVIAFGKFKLNLGTREMFREDEPMPLTSG
EFAVLKALVSHPREPLSRDKLMNLARGREYSAMERSIDVQISRLRRMVEEDPAH
PRYIQTVWGLGYVFVPDGSKA*

>OmpR[RstA]mut3
LQENYKILVVDDDMEVRALLEAYLTEHGFQVRSVANAEQMDRLLTRESFHLMV
LDLMLPGEDGLSICRRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKTTPP
RELLARIRAVLRRQANELPGAPSQEEAVIAFGKFKLNLGTREMFREDEPMPLTSG
EFAVLKALVSHPREPLSRDKLMNLARGREYSAMERSIDVQISRLRRMVEEDPAH
PRYIQTVWGLGYVFVPDGSKA*

>OmpR[RstA]mut4
LQENYKILVVDDDMEVRALLERYLTEQGFQVRSVANAEQMDRLLTRESFHLMV
LDLMLPGEDGLSICRRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKTTPP
RELLARIRAVLRRQANELPGAPSQEEAVIAFGKFKLNLGTREMFREDEPMPLTSG
EFAVLKALVSHPREPLSRDKLMNLARGREYSAMERSIDVQISRLRRMVEEDPAH
PRYIQTVWGLGYVFVPDGSKA*

>OmpR[RstA]mut5
LQENYKILVVDDDMEVRALLERYLTEHGFQVRSVANAEQMDRLLTRESFHLMV
LDLMLPGEDGLSICRRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKTTPP
RELLARIRAVLRRQANELPGAPSQEEAVIAFGKFKLNLGTREMFREDEPMPLTSG
EFAVLKALVSHPREPLSRDKLMNLARGREYSAMERSIDVQISRLRRMVEEDPAH
PRYIQTVWGLGYVFVPDGSKA*

>OmpR[CpxR]mut1
LQENYKILVVDDDMELRALLEELLTEQGFQVRSVANAEQMDRLLTRESFHLMVL
DLMLPGEDGLSICRRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNPR
ELLARIRAVLRRQANELPGAPSQEEAVIAFGKFKLNLGTREMFREDEPMPLTSGE
FAVLKALVSHPREPLSRDKLMNLARGREYSAMERSIDVQISRLRRMVEEDPAHP
RYIQTVWGLGYVFVPDGSKA*

>OmpR[CpxR]mut2
LQENYKILVVDDDMELRALLEELLTEQGFQVRSVANAEQMDRLLTRESFHLMVL
DLMLPGEDGLSICRRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNDR
ELLARIRAVLRRQANELPGAPSQEEAVIAFGKFKLNLGTREMFREDEPMPLTSGE
FAVLKALVSHPREPLSRDKLMNLARGREYSAMERSIDVQISRLRRMVEEDPAHP
RYIQTVWGLGYVFVPDGSKA*

>OmpR[CpxR]mut3
LQENYKILVVDDDMELRALLEELLTEEGFQVRSVANAEQMDRLLTRESFHLMVL
DLMLPGEDGLSICRRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNDR
ELLARIRAVLRRQANELPGAPSQEEAVIAFGKFKLNLGTREMFREDEPMPLTSGE
FAVLKALVSHPREPLSRDKLMNLARGREYSAMERSIDVQISRLRRMVEEDPAHP
RYIQTVWGLGYVFVPDGSKA*

>OmpR[CpxR]mut4
LQENYKILVVDDDMELRALLERLLTEQGFQVRSVANAEQMDRLLTRESFHLMV
LDLMLPGEDGLSICRRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFND
RELLARIRAVLRRQANELPGAPSQEEAVIAFGKFKLNLGTREMFREDEPMPLTSG
EFAVLKALVSHPREPLSRDKLMNLARGREYSAMERSIDVQISRLRRMVEEDPAH
PRYIQTVWGLGYVFVPDGSKA*

>OmpR[CpxR]mut5
LQENYKILVVDDDMELRALLERLLTEEGFQVRSVANAEQMDRLLTRESFHLMVL
DLMLPGEDGLSICRRLRSQSNPMPIIMVTAKGEEVDRIVGLEIGADDYIPKPFNDR
ELLARIRAVLRRQANELPGAPSQEEAVIAFGKFKLNLGTREMFREDEPMPLTSGE
FAVLKALVSHPREPLSRDKLMNLARGREYSAMERSIDVQISRLRRMVEEDPAHP
RYIQTVWGLGYVFVPDGSKA*

>PhoB
LARRILVVEDEAPIREMVCFVLEQNGFQPVEAEDYDSAVNQLNEPWPDLILLDW
MLPGGSGIQFIKHLKRESMTRDIPVVMLTARGEEEDRVRGLETGADDYITKPFSP
KELVARIKAVMRRISPMAVEEVIEMQGLSLDPTSHRVMAGEEPLEMGPTEFKLL
HFFMTHPERVYSREQLLNHVWGTNVYVEDRTVDVHIRRLRKALEPGGHDRMV
QTVRGTGYRFSTRF*

>RstA
LNVMNTIVFVEDDAEVGSLIAAYLAKHDMQVTVEPRGDQAEETILRENPDLVLL
DIMLPGKDGMTICRDLRAKWSGPIVLLTSLDSDMNHILALEMGACDYILKTTPPA

VLLARLRLHLRQNEQATLTKGLQETSLTPYKALHFGTLTIDPINRVVTLANTEISL
STADFELLWELATHAGQIMDRDALLKNLRGVSYDGLDRSVDVAISRLRKKLLDN
AAEPYRIKTVRNKGYLFAPHAWE*

>CpxR
LNKILLVDDDRELTSLLKELLEMEGFNVIVAHDGEQALDLLDDSIDLLLLDVMMP
KKNGIDTLKALRQTHQTPVIMLTARGSELDRVLGLELGADDYLPKPFNDRELVA
RIRAILRRSHWSEQQQNNDNGSPTLEVDALVLNPGRQEASFDGQTLELTGTEFTL
LYLLAQHLGQVVSREHLSQEVLGKRLTPFDRAIDMHISNLRRKLPDRKDGHPWF
KTLRGRGYLMVSAS*

>CC1182
LENVQNAAQSELEAVRGAPSRILIVDDDPGIRDVVSDFLAKHGYVVETAQDGRT
MEQVLARGPIDLIVLDVMLPGEDGLAICRRLSATPEAPAIIMLSAMGEETDRIVGL
ELGADDYLPKPCNPRELLARVRAVLRRRQEPRAVDDAMGAACEFAGWRLDLV
RRELRSPQSIVVNLSSGEFSLLRAFVERPQRVLTRDQLLDLARGRDSDAYDRAID
VQISRLRRKLDDGGGSELIRTIRSEGYMFTAKVVRTP*

>EnvZ
LAAGVKQLADDRTLLMAGVSHDLRTPLTRIRLATEMMSEQDGYLAESINKDIEE
CNAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVK
MHPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRK
HLFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVP
VTRAQGTTKEG*

>EnvZ_T247A
LAAGVKQLADDRTLLMAGVSHDLRAPLTRIRLATEMMSEQDGYLAESINKDIEE
CNAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVK
MHPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRK
HLFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVP
VTRAQGTTKEG*

>EnvZ_T247R
LAAGVKQLADDRTLLMAGVSHDLRRPLTRIRLATEMMSEQDGYLAESINKDIEE
CNAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVK
MHPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRK
HLFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVP
VTRAQGTTKEG*

>RstB-EnvZ_full_domain_chimeric
PHWQDMLKLEAAAQRFGDGHLNERIHFDEGSSFERLGVAFNQMADNINALIASK
KQLIDGIAHELRTPLVRLRYRLEMSDNLSAAESQALNRDISQLEALIEELLTYART
GQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVKMHPLSIKRAVANM
VVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRKHLFQPFVRGDSAR
TISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVPVTRAQGTTKEG*

>RstB
MRPHWQDMLKLEAAAQRFGDGHLNERIHFDEGSSFERLGVAFNQMADNINALI
ASKKQLIDGIAHELRTPLVRLRYRLEMSDNLSAAESQALNRDISQLEALIEELLTY
ARLDRPQNELHLSEPDLPLWLSTHLADIQAVTPDKTVRIKTLVQGHYAALDMRL
MERVLDNLLNNALRYCHSTVETSLLLSGNRATLIVEDDGPGIAPENREHIFEPFVR
LDPSRDRSTGGCGLGLAIVHSIALAMGGTVNCDTSELGGARFSFSWPLWHNIPQF
TSA*

>EnvZ[RstB_MI]
LAAGVKQLADDRTLLMAGVSHDLRTPLVRIRYRTEMMSEQDGYLAEAINKDIEE
CNAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVK
MHPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRK
HLFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVP
VTRAQGTTKEG*

>EnvZ[RstB_MI+loop]
LAAGVKQLADDRTLLMAGVSHDLRTPLVRIRYRLEMSDNLSAALAEAINKDIEE
CNAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVK
MHPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRK
HLFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVP
VTRAQGTTKEG*

>PhoR
DVTQMHQLEGARRNFFANVSHELRTPLTVLQGYLEMMNEQPLEGAVREKALHT
MREQTQRMEGLVKQLLTLSKIEAAPTHLLNEKVDVPMMLRVVEREAQTLSQKK
QTFTFEIDNGLKVSGNEDQLRSAISNLVYNAVNHTPEGTHITVRWQRVPHGAEFS
VEDNGPGIAPEHIPRLTERFYRVDKARSRQTGGSGLGLAIVKHAVNHHESRLNIE
STVGKGTRFSFVIPERLIAKNSD*

>EnvZ[PhoR_MI]
LAAGVKQLADDRTLLMAGVSHDLRTPLTVIRGYTEMMSEQDGYLAETINKQIEE
CNAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVK
MHPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRK
HLFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVP
VTRAQGTTKEG*

>EnvZ[PhoR_MI+loop]
LAAGVKQLADDRTLLMAGVSHDLRTPLTVIRGYLEMMNEQPLEGAVREKLAETI
NKQIEECNAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPG
SIEVKMHPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIA
PEQRKHLFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRA
WLPVPVTRAQGTTKEG*

>CpxA
FLAAGASFNQMVTALERMMTSQQRLLSDISHELRTPLTRLQLGTALLRRRSGESK
ELERIETEAQRLDSMINDLLVMSRNQQKNALVSETIKANQLWSEVLDNAAFEAE

QMGKSLTVNFPPGPWPLYGNPNALESALENIVRNALRYSHTKIEVGFAVDKDGIT
ITVDDDGPGVSPEDREQIFRPFYRTDEARDRESGGTGLGLAIVETAIQQHRGWVK
AEDSPLGGLRLVIWLPLYKRS*

>EnvZ[CpxA_MI]
LAAGVKQLADDRTLLMAGVSHDLRTPLTRIRLGTELMSEQDGYLAERINKEIEEC
NAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVKM
HPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRKH
LFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVPV
TRAQGTTKEG*

>EnvZ[CpxA_MI+loop]
LAAGVKQLADDRTLLMAGVSHDLRTPLTRIRLGTALLRRRSGESKLAERINKEIE
ECNAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVK
MHPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRK
HLFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVP
VTRAQGTTKEG*

>CC1181
ARRLAQPITAFADAAERLGKDPRTPPLNMTGSGEVVAAASAFNMMQERLRRYV
EDRTAMVGAIAHDLRTPLTRLKFRIEAAPEDIRPKLAADIDQMEAMISATLGFVR
DTNRPAERTKLELSSLLESVMDEAAETGGDATVERSEKTVIEGDPVALKRLVSNL
VENALKYGGRARGRVFSEDGMAIIEIDDDGPGVPPAELERVFEPFYRGEPSRNRE
TGGIGLGLAVVRSLARAHGGDVVLANRLGGGLRATVKLPA*

>EnvZ[CC1181_MI]
LAAGVKQLADDRTLLMAGVSHDLRTPLTRIRFRTEAMSEQDGYLAEKINKDIEE
CNAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVK
MHPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRK
HLFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVP
VTRAQGTTKEG*

>EnvZ[CC1181_MI+loop]
LAAGVKQLADDRTLLMAGVSHDLRTPLTRIRFRIEAAPEDLAEKINKDIEECNAII
EQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVKMHPLS
IKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRKHLFQP
FVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVPVTRAQ
GTTKEG*

>EnvZ_DHp
LAAGVKQLADDRTLLMAGVSHDLRTPLTRIRLATEMMSEQDGYLAESINKDIEE
CNAIIEQFIDYLR*

>CC1181_DHp
ARRLAQPITAFADAAERLGKDPRTPPLNMTGSGEVVAAASAFNMMQERLRRYV
EDRTAMVGAIAHDLRTPLTRLKFRIEAAPEDIRPKLAADIDQMEAMISATLGFVR*

>EnvZ[CC1181_MI]_DHp
LAAGVKQLADDRTLLMAGVSHDLRTPLTRIRFRTEAMSEQDGYLAEKINKDIEE
CNAIIEQFIDYLR*

>EnvZ[CC1181_MI+loop]_DHp
LAAGVKQLADDRTLLMAGVSHDLRTPLTRIRFRIEAAPEDLAEKINKDIEECNAII
EQFIDYLR

>EnvZ_G240A
LAAGVKQLADDRTLLMAAVSHDLRTPLTRIRLATEMMSEQDGYLAESINKDIEE
CNAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVK
MHPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRK
HLFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVP
VTRAQGTTKEG*

>EnvZ_G240A_R234A_T235A
LAAGVKQLADDAALLMAAVSHDLRTPLTRIRLATEMMSEQDGYLAESINKDIEE
CNAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVK
MHPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRK
HLFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVP
VTRAQGTTKEG*

>EnvZ_A231I
LAAGVKQLIDDRTLLMAGVSHDLRTPLTRIRLATEMMSEQDGYLAESINKDIEEC
NAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVKM
HPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRKH
LFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVPV
TRAQGTTKEG*

>EnvZ_A231I_R234K_T235K
LAAGVKQLIDDKKLLMAGVSHDLRTPLTRIRLATEMMSEQDGYLAESINKDIEEC
NAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVKM
HPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRKH
LFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVPV
TRAQGTTKEG*

>EnvZ_T250V_L254Y_A255R
LAAGVKQLADDRTLLMAGVSHDLRTPLVRIRYRTEMMSEQDGYLAESINKDIEE
CNAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVK
MHPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRK
HLFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVP
VTRAQGTTKEG*

>EnvZ_A231I_T250V_L254Y_A255R

LAAGVKQLIDDRTLLMAGVSHDLRTPLVRIRYRTEMMSEQDGYLAESINKDIEEC
NAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVKM
HPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRKH
LFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVPV
TRAQGTTKEG*

>EnvZ_A231I_R234K_T235K_T250V_L254Y_A255R
LAAGVKQLIDDKKLLMAGVSHDLRTPLVRIRYRTEMMSEQDGYLAESINKDIEE
CNAIIEQFIDYLRTGQEMPMEMADLNAVLGEVIAAESGYEREIETALYPGSIEVK
MHPLSIKRAVANMVVNAARYGNGWIKVSSGTEPNRAWFQVEDDGPGIAPEQRK
HLFQPFVRGDSARTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAWLPVP
VTRAQGTTKEG*

Novel primers used in constructing mutants

Mutant	**Template**	**Mutation**
OmpR[RstA]mut1	*E. coli K12 ompR*	R15E L16V GTCGATGACGACATGGAAGTGCGTGCGCTGCTGGAA TTCCAGCAGCGCACGCACTTCCATGTCGTCATCGAC R22A CGTGCGCTGCTGGAAGAGTATCTCACCGAACAA TTGTTCGGTGAGATACGCTTCCAGCAGCGCACG
OmpR[RstA]mut2	OmpR[RstA]mut1	R15E L16V R22A P106T F107T N108P GACTACATTCCAAAAACGACGCCGCCGCGTGAACTGCTG CAGCAGTTCACGCGGCGGCGTCGTTTTTGGAATGTAGTC
OmpR[RstA]mut3	OmpR[RstA]mut2	R15E L16V R22A Q27H P106T F107T N108P GCGTATCTCACCGAACATGGCTTCCAGGTTCGA TCGAACCTGGAAGCCATGTTCGGTGAGATACGC
OmpR[RstA]mut4	OmpR[RstA]mut2	R15E L16V P106T F107T N108P CGTTATCTCACCGAACATGGCTTCCAGGTTCGA TCGAACCTGGAAGCCATGTTCGGTGAGATAACG
OmpR[RstA]mut5	OmpR[RstA]mut3	R15E L16V Q27H P106T F107T N108P CGTTATCTCACCGAACATGGCTTCCAGGTTCGA TCGAACCTGGAAGCCATGTTCGGTGAGATAACG
OmpR[CpxR]mut1	*E. coli K12 ompR*	R15E CGTGCGCTGCTGGAAGAATATCTCACCGAACAA TTGTTCGGTGAGATATTCTTCCAGCAGCGCACG R22E Y23L CGTGCGCTGCTGGAAGAACTGCTCACCGAACAAGGC GCCTTGTTCGGTGAGCAGTTCTTCCAGCAGCGCACG
OmpR[CpxR]mut2	OmpR[CpxR]mut1	R15E R22E Y23L P109D CCAAAACCGTTTAACGACCGTGAACTGCTGGCC GGCCAGCAGTTCACGGTCGTTAAACGGTTTTGG

OmpR[CpxR]mut3	OmpR[CpxR]mut2	R15E R22E Y23L Q27E P109D GAACTGCTCACCGAAGAAGGCTTCCAGGTTCGA TCGAACCTGGAAGCCTTCTTCGGTGAGCAGTTC
OmpR[CpxR]mut4	OmpR[CpxR]mut2	R15E Y23L P109D CGTCTGCTCACCGAAGAAGGCTTCCAGGTTCGA TCGAACCTGGAAGCCTTCTTCGGTGAGCAGACG
OmpR[CpxR]mut5	OmpR[CpxR]mut3	R15E Y23L Q27E P109D CGTCTGCTCACCGAAGAAGGCTTCCAGGTTCGA TCGAACCTGGAAGCCTTCTTCGGTGAGCAGACG
EnvZ A231I	*E. coli K12 envZ*	A231I GGTGTTAAGCAACTGATCGATGACCGCACGCTG CAGCGTGCGGTCATCGATCAGTTGCTTAACACC
EnvZ A231I T250V L254Y A255R	T250V L254Y A255R (Jeffrey Skerker)	A231I T250V L254Y A255R GGTGTTAAGCAACTGATCGATGACCGCACGCTG CAGCGTGCGGTCATCGATCAGTTGCTTAACACC
EnvZ G240A	*E. coli K12 envZ*	G240A ACGCTGCTGATGGCGGCAGTAAGTCACGACTTG CAAGTCGTGACTTACTGCCGCCATCAGCAGCGT
EnvZ A231I R234K T235K	A231I	A231I R234K T235K CAACTGATCGATGACAAGAAACTGCTGATGGCGGGG CCCCGCCATCAGCAGTTTCTTGTCATCGATCAGTTG
EnvZ A231I R234K T235K T250V L254Y A255R	A231I T250V L254Y A255R	A231I R234K T235K T250V L254Y A255R CAACTGATCGATGACAAGAAACTGCTGATGGCGGGG CCCCGCCATCAGCAGTTTCTTGTCATCGATCAGTTG
EnvZ R234A T235A G240A	G240A	R234A T235A G240A CAACTGGCGGATGACGCGGCACTGCTGATGGCGGCA TGCCGCCATCAGCAGTGCCGCGTCATCCGCCAGTTG
EnvZ T247R	*E. coli K12 envZ*	T247R AGTCACGACTTGCGCCGTCCGCTGACGCGTATT AATACGCGTCAGCGGACGGCGCAAGTCGTGACT

Mutants for which no primers are listed were either constructed by Biobasic gene synthesis, or primers were obtained from Jeffrey Skerker and Barrett Perchuk in Michael Laub's laboratory.