Determinants of Homodimerization Specificity in Histidine Kinases

Published in final edited form as:
J Mol Biol. 2011 October 14; 413(1): 222–235. doi:10.1016/j.jmb.2011.08.011.

Determinants of homodimerization specificity in histidine kinases

Orr Ashenberg[1,2], **Kathryn Rozen-Gagnon**[3], **Michael T. Laub**[1,2,4,5], and **Amy E. Keating**[1,2,5]

[1]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

[2]Computational & Systems Biology Initiative, Massachusetts Institute of Technology, Cambridge, MA 02139

[3]Department of Biology, Barnard College, New York, NY 10027

[4]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139

Abstract

Two-component signal transduction pathways consisting of a histidine kinase and a response regulator are used by prokaryotes to respond to diverse environmental and intracellular stimuli. Most species encode numerous paralogous histidine kinases that exhibit significant structural similarity. Yet in almost all known examples, histidine kinases are thought to function as homodimers. We investigated the molecular basis of dimerization specificity, focusing on the model histidine kinase EnvZ and RstB, its closest paralog in *Escherichia coli*. Direct binding studies showed that the cytoplasmic domains of these proteins each form specific homodimers *in vitro*. Using a series of chimeric proteins, we identified specificity determinants at the base of the four-helix bundle in the dimerization and histidine phosphotransfer domain. Guided by molecular coevolution predictions and EnvZ structural information, we identified sets of residues in this region that are sufficient to establish homospecificity. Mutating these residues in EnvZ to the corresponding residues in RstB produced a functional kinase that preferentially homodimerized over interacting with EnvZ. EnvZ and RstB likely diverged following gene duplication to yield two homodimers that cannot heterodimerize, and the mutants we identified represent possible evolutionary intermediates in this process.

Introduction

Protein-protein interactions are central to most biological processes and typically must be highly selective; that is, proteins must interact preferentially with functionally relevant, or cognate, partners while minimizing interactions with non-cognate proteins. Understanding the basis of interaction specificity is a challenge, especially when the interacting proteins are members of large paralogous families that share similar sequences and structures. Biological strategies for ensuring specificity can be either contextual or intrinsic. In a contextual strategy, proteins that would otherwise interact are kept apart by spatial or temporal localization. In an intrinsic strategy, biophysical properties of the proteins themselves are sufficient to ensure that they interact preferentially with their cognate partners. Intrinsic

[5]corresponding authors: laub@mit.edu, 617-324-0418, keating@mit.edu, 617-452-3398.

specificity has been demonstrated in a growing number of systems, such as eukaryotic and viral bZIP transcription factors, Bcl-2 family proteins, PDZ domains, and bacterial two-component signaling proteins.[1–5]

Two-component signal transduction pathways allow prokaryotes to sense and respond to diverse environmental and intracellular stimuli.[6] These pathways typically pair a sensor histidine kinase with a cognate response regulator. Upon activation, a histidine-kinase dimer undergoes autophosphorylation and then catalyzes transfer of the phosphoryl group to its cognate response regulator, which typically activates a transcriptional response. Two-component signaling pathways are present in most bacteria and have been extensively studied because of their functional importance and because they are a good model for studying fundamental mechanisms of signal transduction. These pathways are also an excellent system in which to explore mechanisms of interaction specificity. Although most bacteria encode dozens of paralogous histidine kinases and response regulators, cross-talk between pathways is limited.[7]

Recent work has shown that histidine kinases harbor a strong preference *in vitro* for phosphorylating their *in vivo* cognate response regulators.[4] Similarly, response regulators show an intrinsic preference for homodimerization *in vitro*.[8] These findings indicate that interaction specificity in these signaling pathways is established largely at the level of molecular recognition and that contextual or cellular strategies are not necessary to insulate different pathways. Because histidine kinases dimerize to autophosphorylate, kinase heterodimerization represents another possible source of cross-talk that presumably must be minimized.[9,10] Many histidine kinases have been shown to homodimerize *in vitro*, but have rarely been tested for heterodimerization.[11–14] There are only a few examples in which heteroassociations between histidine kinases have been reported, including the ethylene receptors ETR1 and ERS2 in *Arabidopsis thaliana*[15,16] and the kinases GacS and RetS in *Pseudomonas aeruginosa*.[17] It thus remains unresolved whether histidine kinases homodimerize specifically and, if so, how specificity is determined.

Here, we investigate the homodimerization specificity of the model histidine kinase EnvZ from *E. coli*.[18] The closest paralog of EnvZ is RstB, which shares a common domain organization and has ~30% sequence similarity to EnvZ in the cytoplasmic dimerization domain. If there were cross-talk in *E. coli*, RstB is the histidine kinase most likely to heterodimerize with EnvZ. Both EnvZ and RstB function as integral membrane homodimers. Periplasmic sensory domains in each protein are linked to highly conserved cytoplasmic domains through a transmembrane alpha helix.[19] The cytoplasmic portion of each kinase consists of a parallel, dimeric four-helix bundle HAMP (*h*istidine kinases, *a*denylyl cyclases, *m*ethyl-accepting chemotaxis proteins, and *p*hosphatases) domain[20,21], a parallel, dimeric four-helix bundle DHp (*d*imerization and *h*istidine *p*hosphotransfer) domain containing a conserved histidine that is the site of autophosphorylation[12,22–25], and an α/β-sandwich CA (*c*atalytic and *A*TP-binding) domain that catalyzes autophosphorylation.[26] An NMR structure of the DHp domain of EnvZ[23] shows both similarities and interesting differences compared to a crystal structure of the DHp domain of TM0853, a histidine kinase from *Thermotoga maritima*.[12] The two DHp domains share the same overall four-helix bundle architecture but differ in the ordering of the helices. The HAMP domains in histidine kinases appear to be linked to DHp domains through a continuous alpha helix that comprises the last alpha helix of the HAMP domain and the first alpha helix of the DHp domain. Both the DHp and HAMP domains likely dimerize in EnvZ, but whether either or both domains contribute to homodimerization specificity is unclear.

Computational analysis of amino-acid coevolution, combined with knowledge of protein structures, has proven very useful for elucidating the specificity determinants of histidine

kinase-response regulator interactions.[27,28] This approach looks for amino-acid covariation at pairs of positions within large multiple sequence alignments, and is particularly powerful for examining prokaryotic protein families as there are often thousands, or even tens of thousands, of homologous sequences in genome databases. Highly covarying pairs can indicate that two residues interact to maintain protein structure or function. If histidine kinases must homodimerize specifically to suppress cross-talk, then specificity determining residues could show statistically significant covariation. This was shown to be the case for residues at the histidine kinase-response regulator interface and for residues mediating polyketide synthase protein interactions.[28,29] For the former, subsequent analyses have demonstrated how kinase-regulator specificity can be rewired with minimal sequence changes localized to regions originally identified using covariation approaches.

Here, we took a biochemical approach to show that the cytoplasmic domains of EnvZ and RstB specifically self-associate *in vitro*. Guided by an analysis of amino-acid coevolution within histidine kinases, we identified a small number of residues in the DHp domain of EnvZ that can be substituted with the corresponding residues from RstB to create a functional homodimer that preferentially self-associates rather than associating with EnvZ. Our results suggest that histidine-kinase dimerization specificity is, like kinase-regulator interaction and response-regulator dimerization, hardwired at the level of molecular recognition.

Results

EnvZ and RstB specifically homodimerize

To assess whether the cytoplasmic domains of EnvZ and RstB homodimerize specifically, we used fluorescence resonance energy transfer (FRET) to examine interactions between $EnvZ_{HDC}$ and $RstB_{HDC}$ fused to CFP and YFP. The domains present in each kinase construct are denoted with subscripts where H, D, and C indicate HAMP, DHp, and CA, respectively (for details, see Methods). If each kinase specifically homodimerizes, we expected that a mixture of CFP- and YFP-labeled subunits would produce a FRET signal that could be inhibited by increasing concentrations of an unlabeled copy of the same kinase, but not the other kinase (Fig. 1A). All FRET experiments described below were performed in this competition format.

A mixture of CFP-$EnvZ_{HDC}$ and YFP-$EnvZ_{HDC}$ produced a FRET signal that was reduced in a concentration-dependent manner by the addition of unlabeled $EnvZ_{HDC}$ (Fig. 1B). Similarly, the FRET signal from a mixture of CFP-$RstB_{HDC}$ and YFP-$RstB_{HDC}$ decreased with the addition of unlabeled $RstB_{HDC}$ (Fig. 1B). These data confirmed the expected homodimerization of our $EnvZ_{HDC}$ and $RstB_{HDC}$ constructs. We next tested whether EnvZ and RstB could heteroassociate by adding unlabeled $RstB_{HDC}$ to the mixture of CFP-$EnvZ_{HDC}$ and YFP-$EnvZ_{HDC}$. Even at concentrations up to 100 times higher than the labeled EnvZ concentration, unlabeled $RstB_{HDC}$ did not significantly inhibit homodimerization. Similarly, unlabeled $EnvZ_{HDC}$ did not disrupt the $RstB_{HDC}$ homodimer (Fig. 1B). To quantify the relative stabilities of interactions between kinases, we fit homodimer dissociation constants and heterodimer dissociation constants, as detailed in the methods (Table 1, Supp. Table 1, Supp. Fig. 5, 6). The homodimer K_d for $EnvZ_{HDC}$ was 0.4 μM, less than the previously reported value of 10 μM measured using a pull-down assay with His_6-labeled $EnvZ_{HDC}$.[11] Similarly, the homodimer K_d of $RstB_{HDC}$ was 0.3 μM, whereas the $EnvZ_{HDC}$-$RstB_{HDC}$ heterodimer K_d was > 50 μM. Note that when competition in the FRET assay is very weak, the sensitivity of this competition assay is limited and any K_d weaker than 50 μM fits the $EnvZ_{HDC}$-$RstB_{HDC}$ data equally well (Supp. Fig. 4). Overall, these experiments indicate $EnvZ_{HDC}$ and $RstB_{HDC}$ homodimers are each more stable than the $EnvZ_{HDC}$-$RstB_{HDC}$ heterodimer.

We also assayed dimerization of $EnvZ_{HDC}$ and $RstB_{HDC}$ using a pull-down assay. $EnvZ_{HDC}$ or $RstB_{HDC}$, N-terminally labeled with a FLAG epitope, was mixed with MBP-tagged $EnvZ_{HDC}$ or $RstB_{HDC}$. The protein complexes formed were isolated using anti-FLAG beads and examined by SDS-PAGE (Fig. 1C). Homospecific kinase complexes were clearly identified at a level above that due to non-specific bead binding (compare lanes 2 vs. 10 and 4 vs. 12) but heteromeric complexes were much weaker (compare lanes 6 vs. 12 and 8 vs. 10). These data corroborated our FRET data indicating that $EnvZ_{HDC}$ and $RstB_{HDC}$ specifically homodimerize.

A recent crystal structure of a HAMP-DHp fusion protein shows the HAMP and DHp domains make up the cytoplasmic dimerization interface (PDB id: 3zrx). Furthermore, in EnvZ the CA domain is a monomer and is not required for dimerization.[30] To experimentally test contributions of the HAMP and DHp domains to dimerization specificity, we made constructs in which most of the HAMP domain was removed. The constructs we made had the same N-terminus as the EnvZ construct for which an NMR structure was solved, such that only 10 residues from the HAMP domain remained; in the NMR structure this region is unfolded and non-interacting.[23] Interestingly, removing the HAMP domain significantly destabilized RstB but not EnvZ. We estimated a K_d for the wild-type $EnvZ_{DC}$ homodimer of 0.1 μM and a K_d for the wild-type $RstB_{DC}$ homodimer of ~300 μM (Table 1). To explore the dimerization specificity of these constructs, FRET competition experiments were performed at higher concentrations (20 μM) where both $EnvZ_{DC}$ and $RstB_{DC}$ dimerize. These higher concentrations allowed measurement of a range of weaker K_d's. Under these conditions, $EnvZ_{DC}$ homodimerization was inhibited by unlabeled $EnvZ_{DC}$, but not by $RstB_{DC}$, even at high concentrations (Fig. 2). Similarly, $RstB_{DC}$ homodimerization was inhibited by unlabeled $RstB_{DC}$, but not by $EnvZ_{DC}$ (Fig. 2). We estimated the $EnvZ_{DC}$-$RstB_{DC}$ heterodimer K_d as >100 μM; any K_d >100 μM fits the data equally well (Supp. Fig. 4). We conclude that $EnvZ_{DC}$, like $EnvZ_{HDC}$, retains a high degree of homodimerization specificity. Whether $RstB_{DC}$ specifically homodimerizes is less clear because we cannot determine the relative stabilities of $RstB_{DC}$ homodimer and $EnvZ_{DC}$-$RstB_{DC}$ heterodimer.

Identification of putative specificity determining residues within the DHp domain

The specific homodimerization of $EnvZ_{DC}$ motivated us to look for specificity-determining residues within the DHp domain. To guide this search, we performed a computational analysis of amino-acid covariation in histidine kinases. The assumption underlying this approach is that homodimerization specificity-determining residues must coevolve to maintain the self-association of a kinase while disfavoring competing heterodimer states. We built a multiple sequence alignment of 4272 histidine-kinase sequences that each contain a HAMP domain immediately N-terminal to the DHp and CA domains. We then measured covariation between all possible pairs of positions within the alignment using MIp, a scoring metric based on mutual information (MI) that includes a correction for covariation that arises from phylogenetic relationships and random noise.[31] Within the 25 most highly covarying pairs, 20 pairs were within the DHp domain and 5 pairs were within the HAMP domain. None of the HAMP residue pairs were in physical contact based on the Af1503 HAMP NMR structure (see Methods and Supp. Fig. 1). In contrast, 15 DHp pairs had atoms within 5.5 Å of one another and 7 of these pairs involved residues on different chains based on the EnvZ NMR structure. Six of the seven interchain DHp contacts identified by covariation mapped to the region of the DHp domain distal from the HAMP and near the loops of the four-helix bundle (Fig. 3A). We refer to this region, the lower half of the helix bundle, as the "base" of the bundle, and we refer to the upper half as the membrane-proximal region (Fig. 3A). Interestingly, recent crystal structures of the histidine kinase DesK indicate that the membrane-proximal region of the DHp domain exhibits significant

structural plasticity, consistent with the notion that residues in the more static base region may be important for maintaining dimerization and perhaps encoding specificity.[24]

To test the role of residues at the base of the DHp domain in dimerization specificity, we used a series of chimeric proteins in which portions of EnvZ were replaced with the corresponding residues from RstB. These same chimeric proteins were previously studied in the context of histidine kinase-response regulator interaction specificity (J. Skerker, M.T. Laub, unpublished data).[28] In each of the four chimeras, progressively more of the base of the EnvZ helical bundle was replaced with the corresponding residues from RstB. The proteins were designed such that equivalent portions of alpha helices 1 and 2 were replaced in each case (Fig. 4A, B). Each of the four chimeric DHp domains was fused to the EnvZ CA domain.

We tested the chimeric proteins for homodimerization using FRET competition, as above, and observed that each was able to homodimerize (Table 1, Supp. Fig. 2, 5). We then assessed whether the chimeras lost the ability to interact with $EnvZ_{HDC}$. Interestingly, chimera 1 bound $EnvZ_{HDC}$ with a K_d of 0.3 μM, whereas chimeras 2, 3, and 4 did not interact with $EnvZ_{HDC}$ under the conditions tested (Fig. 4C, Table 1). Chimeras 2–4, but not chimera 1, share nine residues from RstB that replace the corresponding residues of EnvZ (Fig. 4A). Our results thus implicated some or all of these residues in establishing homodimerization specificity.

Specificity determinants can be further localized within the chimera 1 and 2 regions

Chimera 2 showed dimerization behavior similar to RstB, in that it did not hetero-associate with $EnvZ_{HDC}$ but maintained an ability to homo-associate. To further localize specificity determinants we again used the covariation analysis described above. Strongly covarying residue pairs in the chimeria-2 region included 252–270 and 256–270, making up a 3-residue cluster (cluster 1, green in Fig. 3A, B), as well as 255–266 and 255–269, making another 3-residue cluster (cluster 2, purple in Fig. 3A, B). These six residues mediate interchain contacts in the NMR structure of EnvZ. Cluster 1 is located in the hydrophobic core of the DHp helical bundle whereas cluster 2 is located adjacent to the chimera 1 loop. To examine the contribution covarying residue pairs make to homodimerization specificity, we made mutations in $EnvZ_{DC}$ that substituted EnvZ residues with the corresponding ones from RstB. For each three-residue cluster we made each single mutant as well as two double mutants in which the covarying residues were both substituted with the corresponding residues from RstB.

Using the FRET competition assay, we tested each cluster-1 mutant for homodimerization and for interaction with $EnvZ_{HDC}$. Unlike chimera 2, none of the double mutants showed high homospecificity (Table 1, Supp. Table 1, 2, Supp. Fig. 2, 3). In fact, cluster 1 mutations, particularly the I252L/I270L double mutant, destabilized homodimer formation. We hypothesized that the effect of these mutations may be context dependent and require concomitant changes in the loop residues that comprise chimera 1. Indeed, introducing the double mutation I252L/I270L into the context of chimera 1 gave a highly homospecific protein with high homodimer stability (K_d = 0.1 μM) that did not heterodimerize significantly with $EnvZ_{HDC}$ (K_d = 20 μM) (Table 1, Supp. Fig. 5, 6). This analysis thus localized determinants of homospecificity to twelve residues at the base of the EnvZ DHp bundle: the ten residues in chimera 1 and the two residues at positions 252 and 270. Interestingly, a significant part of the chimera-1 context effect can be explained by the role of residue T256. Changing this residue to Leu (as in RstB), along with the mutations I252L and I270L, produced a homospecific mutant, I252L/T256L/I270L. These three residues, which strongly covaried with each other, thus exert a strong influence on homodimerization specificity in EnvZ (Table 1, Supp. Fig. 5, 6).

Similarly, we found examples of mutations in cluster 2 that were destabilizing when made individually but were accommodated in the context of chimera 1. The double mutant A255R/S269A had no effect on homo or hetero-associations, but A255R/L266E was significantly destabilized as a homodimer, an effect that was attributable to L266E alone (Table 1, Supp. Table 3, 4). Because position 266 covaries strongly with position 259 which is within the chimera 1 loop, we tested whether mutating M259 in EnvZ to serine, as found in RstB, would rescue homodimerization stability, but it did not (Supp. Table 4). However, the L266E substitution made in the context of chimera 1 did form a stable homodimer (K_d = 0.9 μM) and a slightly less stable heterodimer with $EnvZ_{HDC}$ (K_d = 4.4 μM), yielding modest homodimerization specificity (Table 1, Supp. Fig. 5, 6). Our results thus suggest that the loops present in chimera 1 are either very stabilizing, or provide an important context for the residues in the chimera-2 region.

Dimerization specificity switch mutants do not affect phosphotransfer specificity

We had created three homospecific kinases, I252L/T256L/I270L, chimera 1 + L266E, and chimera 1 + I252L/I270L. Although each kinase was capable of homodimerizing, we wanted to test whether each mutant protein could still catalyze autophosphorylation and phosphotransfer to response regulators. Previous studies have shown that wild-type EnvZ and RstB exhibit a kinetic preference for phosphotransfer *in vitro* to their cognate regulator substrates OmpR and RstA, respectively.[4] Further, chimera 1 was previously shown to have EnvZ-like substrate specificity while chimera 2 was shown to have altered substrate specificity such that it phosphorylated RstA but not OmpR (Fig. 5A). We tested our three most homospecific kinases for autophosphorylation and for phosphotransfer to OmpR and RstA (Fig. 5B). Each mutant was competent for autophosphorylation, providing further evidence that the mutant kinases homodimerize, because EnvZ, and likely RstB, autophosphorylate exclusively as dimers.[9,32] Each mutant also maintained the phosphotransfer specificity of EnvZ, and phosphorylated OmpR but not RstA. Thus, the mutations characterized here affect dimerization specificity without disrupting phosphotransfer specificity.

Discussion

In this study we investigated the dimerization specificity of a pair of histidine-kinase paralogs from *E. coli*, EnvZ and RstB. Using FRET to measure interactions, we showed that the EnvZ and RstB cytoplasmic domains preferentially self-associate, and we identified a set of residues in EnvZ that can be replaced with the corresponding residues from RstB to enable orthogonal homodimerization. Below, we discuss the possible roles of different parts of histidine-kinase structures in determining dimerization specificity and how our results illustrate a plausible evolutionary pathway for establishing histidine-kinase dimer homospecificity following gene duplication events.

Structural and functional modularity of histidine kinases

The strong preference of $EnvZ_{HDC}$ and $RstB_{HDC}$ for homodimerization *in vitro* is consistent with a mechanism in which molecular recognition establishes pathway isolation in two-component signaling. The HAMP and DHp domains likely comprise the dimerization interface and thereby determine the stability and specificity of histidine kinases. To localize the structural determinants of homodimerization specificity, we first made constructs in which most of the HAMP domain was removed. The K_D estimated for $EnvZ_{DC}$ was similar to that of $EnvZ_{HDC}$, consistent with previous thermal denaturation studies.[33] Either the EnvZ HAMP domain does not contribute significantly to dimer stability or it does not fold properly *in vitro*.[33] Nevertheless $EnvZ_{DC}$ preferentially homodimerized, with no evidence of a heterodimer forming with $RstB_{DC}$, even using high concentrations of the latter. These

findings suggest that, at least for EnvZ, the DHp domain plays a critical role in determining homodimerization specificity.

Other evidence supports an important role for the DHp domain in establishing dimerization specificity for histidine kinases more broadly. For example, this notion is consistent with the observation that many histidine kinases do not have HAMP domains. In fact, only ~24% of histidine kinases have a HAMP domain N-terminal to the DHp domain.[35] Some kinases have a DHp domain immediately adjacent to the last transmembrane domain, with no additional domains, whereas others have a PAS or GAF domain in this region. Some kinases have multiple domains preceding the DHp domain, for example, in *B. subtilis* KinA there are three PAS domains.[36] Although HAMP, PAS, and GAF domains can homodimerize and hence could contribute to specificity, these domains are often involved in transmitting signals from an extracellular or periplasmic domain to the kinase domain. A recent NMR study of full-length *E. coli* histidine kinase DcuS showed that increased disorder and flexibility in the N-terminal helix of the cytoplasmic PAS domain increased kinase activity.[37] Similarly, there is evidence that HAMP domains may fluctuate between two different packing states, which may correlate with different signaling states.[20,21,38,39] The DHp domain thus appears to be the most static part of the kinase structure and, consequently, well suited to enforcing dimerization specificity.

Covariation analysis also supports a primary role for DHp domains in establishing homodimer specificity. Of the 100 most strongly covarying residue pairs in our histidine-kinase sequence alignment, the number between residues within the HAMP domain was close to that expected at random (28 vs. 22) and the number between the HAMP and DHp domains was significantly less than expected at random (4 vs. 50, p-value $< 10^{-24}$). By contrast, the number of highly covarying pairs within the DHp domain was significantly more than expected (68 vs. 28, $p < 10^{-16}$). We also assessed whether the strongly covarying residues within the DHp and HAMP domains were positioned to make contacts across the dimer interface. Among the 100 most strongly covarying pairs, the number of interchain pairs within contact distance in the DHp domain was greatly enriched relative to that expected by chance (7 vs. 0.7, $p < 10^{-5}$) whereas the number of interchain contact pairs within the HAMP domain was not as significantly enriched (4 vs. 1.1, $p \approx 0.02$). The enrichment of covarying residues mediating interchain contacts within the DHp domain suggests that this domain might be particularly important for dimerization specificity.

Results from co-variation analysis are based on thousands of kinase sequences and are in no way specific to EnvZ. For a residue pair to show significant co-variation, a very large number of proteins must contribute to the signal. Thus, although we have experimentally verified the importance of the DHp domain for dimerization specificity in only one kinase, EnvZ, the above analysis suggests this finding could generalize to many other kinases. Amino-acid covariation within the HAMP domains was also previously examined, and three clusters of residues identified were proposed to affect signaling and transitions between conformational states.[40]

Modularity within the DHp domain

The DHp four-helix bundle can be divided into the membrane proximal region and the base (Fig. 3A). Covariation analysis identified more interchain residue pairs in the base, arguing that this region might be more important for specificity. The importance of this region is also supported by recent structural analyses showing that there is more variability in the membrane proximal part of the DHp structure than in the base. For example, in *T. maritima* TM0853, the first three turns of helix 1 in the DHp are unfolded when the kinase is bound to its cognate response regulator, but folded into a 2-helix coiled coil when the kinase is in the unbound state.[25] For the *B. subtilis* histidine kinase DesK, conformational states likely

corresponding to the kinase, phosphatase, and phosphotransfer states of the kinase were crystallized, and only in the putative phosphatase state was the N-terminus of the DHp folded into a coiled coil.[24] The smaller amount of conformational diversity at the base of the DHp domain makes it more likely to mediate and specify selective dimerization.

Within the base of the DHp domain, the structure of EnvZ also appears to maintain a degree of modularity. It is notable that all four EnvZ/RstB chimeras examined were able to form stable homodimers, indicating that significant portions of the RstB bundle could be transplanted into an EnvZ context without disrupting the domain. However, different layers of the helix bundle are not completely insulated from one another, as changes in the chimera 1 loop significantly influence mutations in the chimera-2 region. The base of EnvZ also contains another level of modularity with respect to protein-protein interactions. This region is important for both homodimerization and for phosphotransfer to cognate response regulators.[28] However, residues we identified as important for establishing the dimerization specificity of EnvZ (252, 256, 270, 266) do not overlap with residues identified as important for phosphotransfer specificity (250, 254, 255, 269). These two sets of residues are positioned close together in the base of the DHp four-helix bundle, with the more solvent exposed sites governing regulator specificity and the more buried sites influencing dimerization specificity (Fig. 5C).

Localizing specificity determinants in the DHp domain using covariation analysis

To identify individual residues in the base of the EnvZ DHp domain that are sufficient to alter dimerization specificity, we used amino-acid covariation analysis. The sites identified with the greatest effects on specificity were positions 252, 256, 270, and 266 in EnvZ, although we do not rule out that additional specificity residues may be localized to other parts of the structure (Fig. 3B). Making substitutions in EnvZ with the corresponding residues from RstB at these covarying sites destabilized both homodimerization and heterodimerization with $EnvZ_{HDC}$. However, introducing the same mutations in the context of chimera 1 (chimera 1 + I252L/I270L and chimera 1 + L266E), which includes the loops at the base of the DHp domain, produced kinases that maintained stable self-association and only weakly hetero-associated with $EnvZ_{HDC}$ (Table 1). Replacing only the loop of EnvZ, as in chimera 1, produced a kinase that stably homodimerized but continued to heterodimerize with EnvZ. We thus speculate that the chimera 1 mutations may alter the structure near the base of the bundle, better accommodating substitutions I252L, I270L and L266E, and thereby stabilizing the homodimer and destabilizing the heterodimer with EnvZ. An alternative possibility is that instead of being structurally coupled, the loop region in chimera 1 and substitutions I252L, I270L, and L266E do not influence one another. Replacing the loop in EnvZ with that in RstB could instead simply stabilize the dimer, allowing it to accommodate destabilization from chimera 2 substitutions.

Although chimera 1 introduced 10 amino-acid changes into EnvZ, it may only be a subset of these residues that influence substitutions introduced in chimera 2. For example, two positions within chimera 1, 256 and 259, covaried strongly with positions 252, 270 and 266 within chimera 2. The substitution T256L along with I252L and I270L led to more homospecificity than I252L and I270L alone. With only three mutations, this kinase achieved a notable fraction of the full homospecificity found in chimera 1 + I252L/I270L.

Our results, and those reported previously, demonstrate the power of using amino-acid coevolution analyses to guide the identification of critical specificity-determining residues.[28,29] Here, covariation analyses highlighted a small number of interchain contacts, among the dozens of possible contacts in the DHp interface, many of which significantly influenced dimerization specificity. Identifying interchain contacts, however, relied on the EnvZ NMR structure. Also, not all of the covarying positions affected specificity when

mutations were made at those sites. The covariation analysis also may have missed important residues within the loop at the base of the DHp domain. As noted previously, these loop regions do not align well, making their analysis by covariation methods difficult.[28] Finally, it is clear that the sequence context in which mutations are made can matter. For instance, adding the double mutant I252L/I270L into EnvZ versus chimera 1 resulted in homodimers with very different stabilities. These observations underscore the importance of combining analyses of coevolution with structural analysis and experimental studies.

Evolutionary implications

Paralogous kinases generated through gene duplication can initially physically associate. Establishing a new and distinct homodimerizing kinase requires a series of mutations in one or both kinases that prevent heterodimerization but retain function. The mutants of EnvZ created here, which homodimerize but no longer interact with EnvZ, represent possible evolutionary intermediates. In our most homospecific mutants, I252L/T256L/I270L, chimera 1 + I252L/I270L, and chimera 1 + L266E, we had to change both residues in the loop and residues unique to chimera 2. One interesting and plausible evolutionary pathway for kinase dimerization specificity is that neutral sequence drift in the loop preceded the more destabilizing changes in the helix bundle base (positions 252, 270, and 266). This scenario would be consistent with our observations that mutations I252L/I270L and L266E alone destabilized both the homodimers and the heterodimers with EnvZ, but were accommodated in homodimers after introducing either T256L or the entire chimera 1 loop region.

The loop at the base of the four-helix bundle is the least conserved region of the DHp domain, and changes in the loop could place residues in the helix bundle base in a new structural context while being neutral with respect to interaction specificity. Chimera 1, which represents a loop swap between EnvZ and RstB, is neutral with respect to kinase dimerization specificity (Table 1), and with respect to phosphotransfer specificity (Fig. 5A). The notion that neutral mutations must accumulate first to provide a context for additional, selectively advantageous mutations has been suggested for a number of systems.[41–43]

The structural and functional modularity of EnvZ also informs models of how histidine kinases evolve. The ability to modulate dimerization specificity and kinase-regulator interaction specificity independently, each with a relatively small number of mutations, increases the evolvability of this large and important gene family. Coupled with mechanisms that add or replace sensory and signaling domains, facile routes to achieving the broad diversity of histidine kinases are not difficult to imagine.

Methods

Constructing pENTR clones and expression clones for histidine kinases

The Gateway recombinational cloning system (Invitrogen) was used to generate pENTR clones for all constructs. EnvZ and RstB constructs containing the HAMP, DHp, and CA domains ($EnvZ_{HDC}$ = amino acids 179–450, $RstB_{HDC}$ = amino acids 157–433) or the DHp and CA domains ($EnvZ_{DC}$ = amino acids 223–450, $RstB_{DC}$ = amino acids 201–433), as well as full-length response regulators OmpR and RstA were cloned previously.[4] Single-mutant EnvZ constructs were generated using QuikChange site-directed mutagenesis (Stratagene). Double and triple mutant EnvZ clones were generated using PCR-based site-directed mutagenesis, with DpnI digestion preceding blunt end ligation.[44] pENTR clones of chimeras 1–4 were constructed previously.[28]

The $EnvZ_{HDC}$ mutant in which the HAMP domain was replaced with the RstB HAMP was generated using splicing-by-overlap-extension PCR. In the first round the RstB HAMP and $EnvZ_{DC}$ were individually amplified and gel purified. The products were then mixed in a 1:1 molar ratio and amplified in a second round of PCR with outer primers and gel purified. The final product was cloned into pENTR/D-TOPO vector following the manufacturer's protocol (Invitrogen).

For constructs used as competitors in the FRET assays and for pull-down assays, a tandem FLAG tag (DYKDDDDKDYKDDDDKGSG) was introduced at the N-terminus of the appropriate kinase using PCR with a forward primer containing the tandem tag.[45] PCR products were then cloned into the pENTR/D-TOPO vector.

The Gateway recombinational cloning system was used to move genes from pENTR clones into IPTG inducible expression vectors with suitable purification tags. pENTR clones were recombined into either pHIS-MBP-DEST, pTRX-HIS-DEST, or pHIS-DEST destination vectors using the Gateway LR clonase reaction.[4]

Plasmids encoding N-terminal ECFP (pRG31) and N-terminal monomeric mYFP (pRG88) along with a His_6 tag were a generous gift from the laboratory of A.M. Stock.[8] EnvZ and RstB pENTR clones were amplified with forward and reverse primers containing NheI and NotI restriction sites and ligated into pRG31 and pRG88 vectors digested with NheI and NotI. All constructs were sequence verified.

Protein expression and purification

Proteins were expressed and purified as previously described.[4] Briefly, expression vectors were transformed into *E. coli* BL21-Tuner cells. Colonies were picked and grown up in 1 L LB at 37 °C. After reaching OD_{600} ~0.6, cells were either induced with 0.3 mM IPTG for 4 hrs at 30 °C (for strains expressing proteins with His_6-MBP, TRX-His_6, or His_6 tags) or with 0.5 mM IPTG for 12–16 hrs at 18 °C (for strains expressing proteins tagged with CFP or YFP). Induced cells were then pelleted and stored at −80 °C. His_6-tagged proteins were purified using Ni-NTA agarose beads and final protein aliquots were stored in storage buffer (10 mM HEPES-KOH [pH 8.0], 50 mM KCl, 10% glycerol, 0.1 mM EDTA, 1 mM DTT) at −80 °C. Protein concentrations were determined using the Edelhoc method (1x PBS, 7 M GuHCl, pH 7.4) with absorbance measured at 280 nm[46], except for fluorescent proteins for which concentration was measured by absorbance at 433 nm for CFP (ε 32500 M^{-1} cm^{-1}) or at 514 nm for YFP (ε 83400 M^{-1} cm^{-1}).[47]

Pull-down assay

For pull-down assays, 2.5 μM $FLAG_2$-histidine kinase was mixed with 12.5 μM MBP-histidine kinase in HEPES buffer (10 mM HEPES-KOH, 50 mM KCl, 0.1 mM EDTA, pH 8.0). Non-specific bead-binding controls included only the MBP-tagged histidine kinase. Proteins were equilibrated for two hours at room temperature on an end-over-end rotator to allow time for subunit exchange before adding 40 μL of anti-FLAG M2 affinity gel (50% slurry, Sigma) that had been washed once in HEPES buffer. The protein mixture and beads were incubated for 30 minutes at 4 °C and then washed 3 times with HEPES buffer. Protein was competitively eluted using five column volumes of 3X-FLAG peptide (100 μg/mL, Sigma). Eluant was concentrated using StrataClean binding resin (Stratagene), run on 10% Tris-HCl SDS-PAGE gel (Bio-Rad), and visualized with Coomassie staining.

FRET competition binding assay

Equimolar mixtures of CFP-histidine kinase and YFP-histidine kinase at 0.5, 5, or 20 μM, and $FLAG_2$-histidine kinase were placed in 96-well plates (Corning), covered with a foil

seal, and incubated for eight hours at 30 °C. Plates were incubated for eight hours because in a kinetic study of 5 μM CFP-$RstB_{HDC}$ mixed with 5 μM YFP-$RstB_{HDC}$, the FRET ratio required five hours to reach equilibrium (data not shown), although most other mixtures examined reached equilibrium within an hour. Fluorescence was measured using a Varioskan plate reader at 30 °C with three channels monitored: donor channel (excite 433 nm, emit 475 nm), acceptor channel (excite 488 nm, emit 527 nm), and FRET channel (excite 433 nm, emit 527 nm). For each well, 30 measurements were made and then averaged. The FRET ratio was calculated as the ratio of FRET channel signal to donor channel signal. To correct for crosstalk and bleed-through, a corrected FRET ratio was calculated as $(F_m-F_a(A_m/A_a)-F_d(D_m/D_d)) / D_m$ where D=donor channel, A=acceptor channel, F=FRET channel, d=donor sample alone, a=acceptor sample alone, and m=mix of donor and acceptor samples and unlabeled kinase.[48] The FRET ratio was measured when making a qualitative assessment of an interaction. The numerator of the corrected FRET ratio, $F_m-F_a(A_m/A_a)-F_d(D_m/D_d)$, referred to as the corrected FRET emission signal, was the signal used in fitting K_d values. K_d values were measured in duplicate.

Fitting equilibrium dissociation constants for homodimers and heterodimers

The experimental setup of mixing unlabeled $FLAG_2$-kinase (U) with an equimolar mixture of CFP-kinase (C) and YFP-kinase (Y) can be described by the following six equilibrium reactions: $2C\leftrightarrow C_2$, $2Y\leftrightarrow Y_2$, $2U\leftrightarrow U_2$, $C+Y\leftrightarrow CY$, $C+U\leftrightarrow CU$, and $Y+U\leftrightarrow YU$. To measure homodimer K_d values, the FRET competition was performed with kinases C, Y, and U having the same kinase sequences. For example, we mixed unlabeled $EnvZ_{HDC}$ with CFP-$EnvZ_{HDC}$ and YFP-$EnvZ_{HDC}$. We made the simplifying assumption that C_2, Y_2, and U_2 have the same equilibrium dissociation constant, K_d, and that CY, CU, and YU have the same dissociation constant, $K_d/2$ (heterodimer twice as likely as homodimer). The decrease in corrected FRET emission signal as the FRET complex CY is competed off by U is directly proportional to the concentration of FRET complex, and the concentration of the FRET complex is determined by the single fitting parameter K_d. A simulation of the FRET experiment was performed in MATLAB, where the system of ordinary differential equations describing the six reactions was integrated until they reached equilibrium. The initial conditions were given by the concentrations of C, Y, and U and the single fitting parameter was the K_d. The concentrations of C and Y were 0.5 μM, except in the case of $RstB_{DC}$, where it was 20 μM. Starting with an initial guess for the K_d, the R^2 between the simulated data and the experimental data was calculated. The K_d was varied in 0.1 μM increments to maximize the R^2, referred to as R^2_{max} To estimate the range of K_d values that fit the data equally well, we calculated the lower and upper limits of K_d values that fit the data with $0.95R^2_{max}$.

To measure a heterodimer K_d, the FRET experiment was performed with C and Y as the same kinase and U a different kinase. As a result, three dissociation constants determine the equilibrium: the two homodimer dissociation constants for C_2 (same as for Y_2) and U_2, and the heterodimer dissociation constant for CU (same for YU). Because the homodimer K_d values are measured in separate experiments, the heterodimer K_d is the only fitting parameter. The fitting was done in a manner analogous to the fitting of homodimer K_d values.

To estimate the homodimer K_d for $RstB_{DC}$, we increased the concentration of CFP-$RstB_{DC}$ and YFP-$RstB_{DC}$ to 20 μM. We found that even in the presence of 50-fold more unlabeled $RstB_{DC}$, there was a significant amount of corrected FRET emission signal remaining. This was attributed to direct binding between the ECFP and mYFP at these higher concentrations. The corrected FRET emission signal for the mix of purified ECFP and mYFP, each at 20 μM, was comparable to the signal in the $RstB_{DC}$ experiment at 50-fold more unlabeled

$RstB_{DC}$. We corrected for this background level of binding between CFP and YFP by subtracting the CFP and YFP binding acceptor emission from the $RstB_{DC}$ experiment acceptor emissions. After this background correction, fitting the homodimer K_d was carried out as described above. Such a background correction was unnecessary when fluorescent protein concentrations were 0.5 μM as the magnitude of CFP and YFP binding signal was only ~1–5% of the signal in the binding experiments involving kinases.

Phosphotransfer assay

Phosphotransfer assays were performed as previously described.[4] Briefly, purified $FLAG_2$-tagged histidine kinases and thioredoxin tagged response regulators were diluted to 5 μM in storage buffer plus 5 mM $MgCl_2$. Autophosphorylation was initiated with the addition of 500 μM ATP and 5 μCi [γ^{32}P]ATP (Amersham Biosciences, 6,000 Ci/mmol) to the kinase, and reactions were incubated for one hour at 30 °C. To initiate phosphotransfer, response regulator was added to the kinase in a 2.5 μM:2.5 μM ratio at room temperature and the reaction was stopped after 10 seconds by adding sample buffer. Samples were then run on a 10% Tris-HCl SDS-PAGE gel (Bio-Rad). The gel was exposed to a phosphor screen for two hours at room temperature and then scanned with a Storm 860 imaging system (Amersham Biosciences) at 50 μm resolution.

Histidine-kinase sequence alignments

Sequence alignments were built using hidden Markov models obtained from the Pfam database.[49,50] A sequence alignment of kinases with the domain architecture HAMP-DHp-CA was constructed by taking histidine-kinase sequences containing a DHp domain (PF00512) and identifying, using HMMER (http://hmmer.org), the subset of those sequences containing a HAMP domain (PF00672) followed by a DHp domain followed by a CA domain (PF02518). The sequence alignment was subjected to a 90% sequence identity cutoff (no pair of sequences could share more than 90% sequence identity) and a 10% gap cutoff (columns with >10% gaps were removed). The final alignment of sequences contained 4,272 sequences.

To calculate whether highly covarying positions in the multiple sequence alignment preferentially occurred in HAMP or DHp domains, position pairs were first divided into three classes: both residues within the HAMP domain (total number of pairs, $m = 1225$), both residues within the DHp domain ($m = 1596$), or one residue from each domain ($m = 2850$). The process of choosing the observed number of pairs from one class, k, is described by the hypergeometric distribution. I.e., the probability that k pairs out of n pairs are from a single class, when selected randomly without replacement from a set of N pairs containing m pairs of that class, is given by $P(X=k)=\binom{m}{k}\binom{N-m}{n-k}/\binom{N}{n}$. In our case $N = 5671$ pairs, and $n = 100$ pairs. Of these 100 pairs, 28 pairs are between HAMP residues, 68 pairs are between DHp residues, and 4 pairs are between HAMP and DHp residues (these are the values of k for the 3 classes). The expected number of covarying pairs from a class is given by the mean of this distribution, nm/N. This null model, which assumes that covarying pairs are distributed randomly over domains, acts to correct for domain size because longer domains are expected to contain more covarying pairs. The probability that k or more covarying pairs from the same class were chosen from a distribution characterized by the above null model is $P(X \geq k)$.

The expected number of interchain contacts within the HAMP (PDB id: 2asw) or DHp (PDB id: 1joy) domains and the significance of seeing the observed number of contacts were modeled in the same manner as above. The null hypothesis was that interchain contacts were randomly distributed across a rank-ordered list of covarying pairs. When counting the

observed number of interchain contacts, as before the top 100 covarying pairs out of 5671 total pairs were considered. Covarying pairs were divided into three classes: HAMP domain interchain contacts ($k = 4$, $m = 63$), DHp domain interchain contacts ($k = 7$, $m = 39$), or none of the above.

Covariation and structure

For the HAMP-DHp-CA multiple sequence alignment, mutual information with the average product correction, MIp, was measured for every pair of positions within the HAMP and DHp domains.[31] When analyzing covariation, a MIp score threshold of 0.1 was chosen (Supp. Fig. 1). The 25 highest scoring pairs were mapped onto the structure of either the Af1503 HAMP domain or the EnvZ DHp domain.[20,23] Distances were measured between heavy atoms, and the distance between a pair of residues was measured in either an intrachain context or an interchain context. In an intrachain context, residues are on the same monomer in the dimer whereas in an interchain context, residues are on opposite monomers in the dimer. A pair was classified as a contact if the minimum distance between the residues was less than 5.5 Å.

Research Highlights

E. coli histidine-kinase paralogs EnvZ and RstB specifically homodimerize *in vitro*.

Covariation analysis predicts sites that influence homodimer vs. heterodimer formation.

Local mutations in EnvZ are sufficient to impart RstB-like homodimerization.

Mutant kinases illustrate possible paths in the evolution of pathway isolation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the laboratory of A.M. Stock for providing plasmid reagents. We thank members of the Keating laboratory (especially T.S. Chen, C. Negron, L. Reich, A. W. Reinke, and V. Potapov) and the Laub laboratory (especially E.J. Capra, B.S. Perchuk, and C.G. Tsokos) for helpful discussions. We thank the BioMicro Center of the Massachusetts Institute of Technology for use of the Varioskan plate reader. We thank the HHMI-MIT Summer Research Program in Chemical Biology for support of K. Rozen-Gagnon. This work was funded by National Institutes of Health award GM067681 to A.E. Keating, a National Science Foundation CAREER Grant to M.T. Laub, and the National Science Foundation GRFP fellowship to O. Ashenberg. M.T.L. is an Early Career Scientist of the Howard Hughes Medical Institute. We used computer resources provided by National Science Foundation award 0821391.

References

1. Stiffler MA, Chen JR, Grantcharova VP, Lei Y, Fuchs D, Allen JE, Zaslavskaia LA, MacBeath G. PDZ domain binding selectivity is optimized across the mouse proteome. Science. 2007; 317:364–369. [PubMed: 17641200]
2. Newman JR, Keating AE. Comprehensive identification of human bZIP interactions with coiled-coil arrays. Science. 2003; 300:2097–2101. [PubMed: 12805554]
3. Chen L, Willis SN, Wei A, Smith BJ, Fletcher JI, Hinds MG, Colman PM, Day CL, Adams JM, Huang DC. Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function. Mol Cell. 2005; 17:393–403. [PubMed: 15694340]
4. Skerker JM, Prasol MS, Perchuk BS, Biondi EG, Laub MT. Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. PLoS Biol. 2005; 3:e334. [PubMed: 16176121]

5. Reinke AW, Grigoryan G, Keating AE. Identification of bZIP interaction partners of viral proteins HBZ, MEQ, BZLF1, and K-bZIP using coiled-coil arrays. Biochemistry. 2010; 49:1985–1997. [PubMed: 20102225]
6. Stock AM, Robinson VL, Goudreau PN. Two-component signal transduction. Annu Rev Biochem. 2000; 69:183–215. [PubMed: 10966457]
7. Laub MT, Goulian M. Specificity in two-component signal transduction pathways. Annu Rev Genet. 2007; 41:121–145. [PubMed: 18076326]
8. Gao R, Tao Y, Stock AM. System-level mapping of Escherichia coli response regulator dimerization with FRET hybrids. Mol Microbiol. 2008; 69:1358–1372. [PubMed: 18631241]
9. Yang Y, Inouye M. Intermolecular complementation between two defective mutant signal-transducing receptors of Escherichia coli. Proc Natl Acad Sci U S A. 1991; 88:11057–11061. [PubMed: 1662380]
10. Scheu PD, Liao YF, Bauer J, Kneuper H, Basche T, Unden G, Erker W. Oligomeric sensor kinase DcuS in the membrane of Escherichia coli and in proteoliposomes: chemical cross-linking and FRET spectroscopy. J Bacteriol. 2010; 192:3474–3483. [PubMed: 20453099]
11. Hidaka Y, Park H, Inouye M. Demonstration of dimer formation of the cytoplasmic domain of a transmembrane osmosensor protein, EnvZ, of Escherichia coli using Ni-histidine tag affinity chromatography. FEBS Lett. 1997; 400:238–242. [PubMed: 9001405]
12. Marina A, Waldburger CD, Hendrickson WA. Structure of the entire cytoplasmic portion of a sensor histidine-kinase protein. Embo J. 2005; 24:4247–4259. [PubMed: 16319927]
13. Bilwes AM, Alex LA, Crane BR, Simon MI. Structure of CheA, a signal-transducing histidine kinase. Cell. 1999; 96:131–141. [PubMed: 9989504]
14. Ninfa EG, Atkinson MR, Kamberov ES, Ninfa AJ. Mechanism of autophosphorylation of Escherichia coli nitrogen regulator II (NRII or NtrB): trans-phosphorylation between subunits. J Bacteriol. 1993; 175:7024–7032. [PubMed: 8226644]
15. Gao Z, Wen CK, Binder BM, Chen YF, Chang J, Chiang YH, Kerris RJ 3rd, Chang C, Schaller GE. Heteromeric interactions among ethylene receptors mediate signaling in Arabidopsis. J Biol Chem. 2008; 283:23801–23810. [PubMed: 18577522]
16. Grefen C, Stadele K, Ruzicka K, Obrdlik P, Harter K, Horak J. Subcellular localization and in vivo interactions of the Arabidopsis thaliana ethylene receptor family members. Mol Plant. 2008; 1:308–320. [PubMed: 19825542]
17. Goodman AL, Merighi M, Hyodo M, Ventre I, Filloux A, Lory S. Direct interaction between sensor kinase proteins mediates acute and chronic disease phenotypes in a bacterial pathogen. Genes Dev. 2009; 23:249–259. [PubMed: 19171785]
18. Forst S, Delgado J, Inouye M. Phosphorylation of OmpR by the osmosensor EnvZ modulates expression of the ompF and ompC genes in Escherichia coli. Proc Natl Acad Sci U S A. 1989; 86:6052–6056. [PubMed: 2668953]
19. Cheung J, Hendrickson WA. Sensor domains of two-component regulatory systems. Curr Opin Microbiol. 2010; 13:116–123. [PubMed: 20223701]
20. Hulko M, Berndt F, Gruber M, Linder JU, Truffault V, Schultz A, Martin J, Schultz JE, Lupas AN, Coles M. The HAMP domain structure implies helix rotation in transmembrane signaling. Cell. 2006; 126:929–940. [PubMed: 16959572]
21. Airola MV, Watts KJ, Bilwes AM, Crane BR. Structure of concatenated HAMP domains provides a mechanism for signal transduction. Structure. 2010; 18:436–448. [PubMed: 20399181]
22. Bick MJ, Lamour V, Rajashankar KR, Gordiyenko Y, Robinson CV, Darst SA. How to switch off a histidine kinase: crystal structure of Geobacillus stearothermophilus KinB with the inhibitor Sda. J Mol Biol. 2009; 386:163–177. [PubMed: 19101565]
23. Tomomori C, Tanaka T, Dutta R, Park H, Saha SK, Zhu Y, Ishima R, Liu D, Tong KI, Kurokawa H, Qian H, Inouye M, Ikura M. Solution structure of the homodimeric core domain of Escherichia coli histidine kinase EnvZ. Nat Struct Biol. 1999; 6:729–734. [PubMed: 10426948]
24. Albanesi D, Martin M, Trajtenberg F, Mansilla MC, Haouz A, Alzari PM, de Mendoza D, Buschiazzo A. Structural plasticity and catalysis regulation of a thermosensor histidine kinase. Proc Natl Acad Sci U S A. 2009; 106:16185–16190. [PubMed: 19805278]

25. Casino P, Rubio V, Marina A. Structural insight into partner specificity and phosphoryl transfer in two-component signal transduction. Cell. 2009; 139:325–336. [PubMed: 19800110]
26. Tanaka T, Saha SK, Tomomori C, Ishima R, Liu D, Tong KI, Park H, Dutta R, Qin L, Swindells MB, Yamazaki T, Ono AM, Kainosho M, Inouye M, Ikura M. NMR structure of the histidine kinase domain of the E. coli osmosensor EnvZ. Nature. 1998; 396:88–92. [PubMed: 9817206]
27. Capra EJ, Perchuk BS, Lubin EA, Ashenberg O, Skerker JM, Laub MT. Systematic dissection and trajectory-scanning mutagenesis of the molecular interface that ensures specificity of two-component signaling pathways. PLoS Genet. 2010; 6 e1001220.
28. Skerker JM, Perchuk BS, Siryaporn A, Lubin EA, Ashenberg O, Goulian M, Laub MT. Rewiring the specificity of two-component signal transduction systems. Cell. 2008; 133:1043–1054. [PubMed: 18555780]
29. Thattai M, Burak Y, Shraiman BI. The origins of specificity in polyketide synthase protein interactions. PLoS Comput Biol. 2007; 3:1827–1835. [PubMed: 17907798]
30. Park H, Saha SK, Inouye M. Two-domain reconstitution of a functional protein histidine kinase. Proc Natl Acad Sci U S A. 1998; 95:6728–6732. [PubMed: 9618480]
31. Dunn SD, Wahl LM, Gloor GB. Mutual information without the influence of phylogeny or entropy dramatically improves residue contact prediction. Bioinformatics. 2008; 24:333–340. [PubMed: 18057019]
32. Cai SJ, Inouye M. Spontaneous subunit exchange and biochemical evidence for trans-autophosphorylation in a dimer of Escherichia coli histidine kinase (EnvZ). J Mol Biol. 2003; 329:495–503. [PubMed: 12767831]
33. Kishii R, Falzon L, Yoshida T, Kobayashi H, Inouye M. Structural and functional studies of the HAMP domain of EnvZ, an osmosensing transmembrane histidine kinase in Escherichia coli. J Biol Chem. 2007; 282:26401–26408. [PubMed: 17635923]
34. Park H, Inouye M. Mutational analysis of the linker region of EnvZ, an osmosensor in Escherichia coli. J Bacteriol. 1997; 179:4382–4390. [PubMed: 9209057]
35. Szurmant H, White RA, Hoch JA. Sensor complexes regulating two-component signal transduction. Curr Opin Struct Biol. 2007; 17:706–715. [PubMed: 17913492]
36. Lee J, Tomchick DR, Brautigam CA, Machius M, Kort R, Hellingwerf KJ, Gardner KH. Changes at the KinA PAS-A dimerization interface influence histidine kinase function. Biochemistry. 2008; 47:4051–4064. [PubMed: 18324779]
37. Etzkorn M, Kneuper H, Dunnwald P, Vijayan V, Kramer J, Griesinger C, Becker S, Unden G, Baldus M. Plasticity of the PAS domain and a potential role for signal transduction in the histidine kinase DcuS. Nat Struct Mol Biol. 2008; 15:1031–1039. [PubMed: 18820688]
38. Zhou Q, Ames P, Parkinson JS. Mutational analyses of HAMP helices suggest a dynamic bundle model of input-output signalling in chemoreceptors. Mol Microbiol. 2009; 73:801–814. [PubMed: 19656294]
39. Ferris HU, Dunin-Horkawicz S, Mondejar LG, Hulko M, Hantke K, Martin J, Schultz JE, Zeth K, Lupas AN, Coles M. The Mechanisms of HAMP-Mediated Signaling in Transmembrane Receptors. Structure. 2011; 19:378–385. [PubMed: 21397188]
40. Dunin-Horkawicz S, Lupas AN. Comprehensive analysis of HAMP domains: implications for transmembrane signal transduction. J Mol Biol. 2010; 397:1156–1174. [PubMed: 20184894]
41. Ortlund EA, Bridgham JT, Redinbo MR, Thornton JW. Crystal structure of an ancient protein: evolution by conformational epistasis. Science. 2007; 317:1544–1548. [PubMed: 17702911]
42. Wang X, Minasov G, Shoichet BK. Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. J Mol Biol. 2002; 320:85–95. [PubMed: 12079336]
43. Soskine M, Tawfik DS. Mutational effects and the evolution of new protein functions. Nat Rev Genet. 2010; 11:572–582. [PubMed: 20634811]
44. Fisher CL, Pei GK. Modification of a PCR-based site-directed mutagenesis method. Biotechniques. 1997; 23:570–571. 574. [PubMed: 9343663]
45. Schneider F, Hammarstrom P, Kelly JW. Transthyretin slowly exchanges subunits under physiological conditions: A convenient chromatographic method to study subunit exchange in oligomeric proteins. Protein Sci. 2001; 10:1606–1613. [PubMed: 11468357]

46. Edelhoch H. Spectroscopic determination of tryptophan and tyrosine in proteins. Biochemistry. 1967; 6:1948–1954. [PubMed: 6049437]

47. Shaner NC, Steinbach PA, Tsien RY. A guide to choosing fluorescent proteins. Nat Methods. 2005; 2:905–909. [PubMed: 16299475]

48. Gordon GW, Berry G, Liang XH, Levine B, Herman B. Quantitative fluorescence resonance energy transfer measurements using fluorescence microscopy. Biophys J. 1998; 74:2702–2713. [PubMed: 9591694]

49. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A. The Pfam protein families database. Nucleic Acids Res. 2009; 38:D211–D222. [PubMed: 19920124]

50. Eddy SR. Profile hidden Markov models. Bioinformatics. 1998; 14:755–763. [PubMed: 9918945]

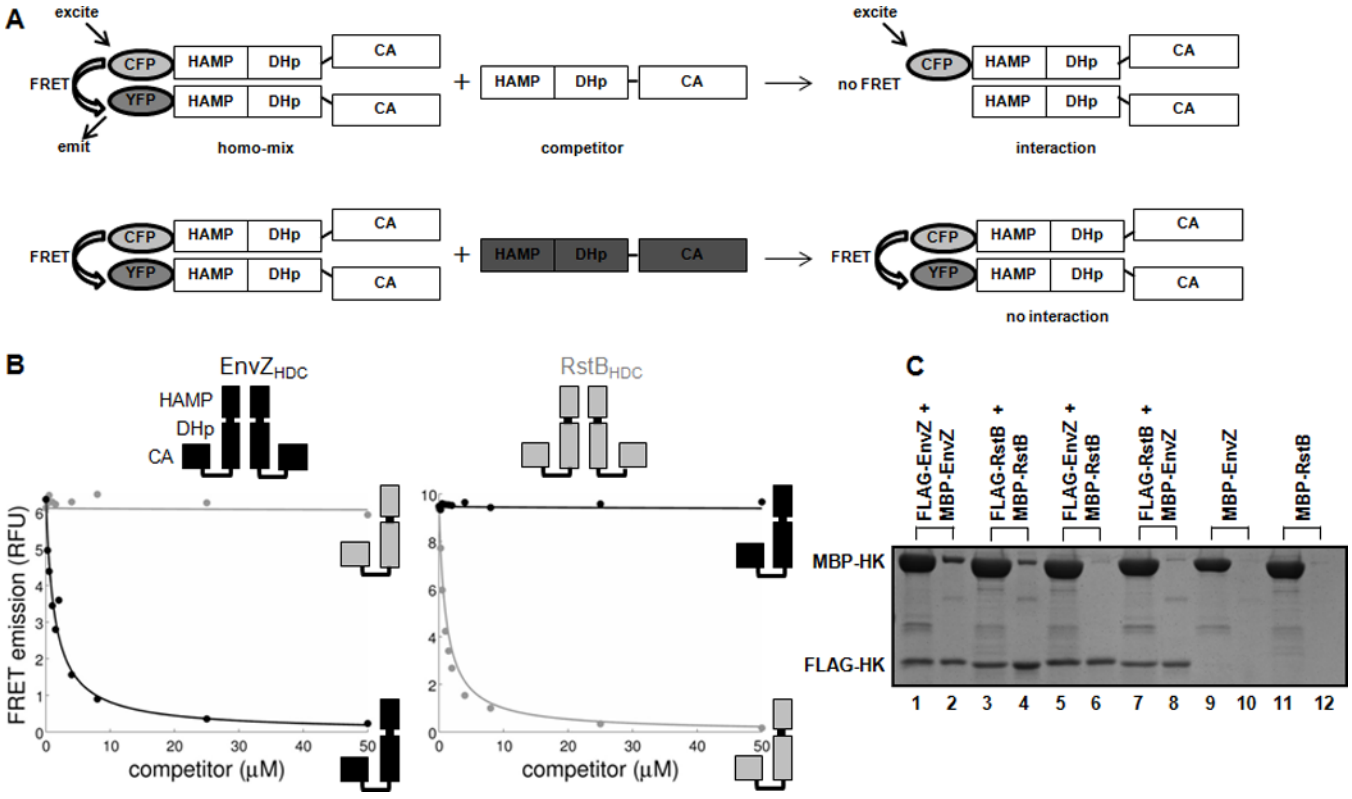

Figure 1. The Cytoplasmic Domains of *E. coli* Histidine Kinases EnvZ and RstB Specifically Homodimerize

(A) Schematic of *in vitro* FRET competition assay. FRET signal from a complex of CFP-kinase A and YFP-kinase A is reduced by interaction with an unlabeled competitor kinase. (B) FRET competition assay for homo-association and hetero-association of $EnvZ_{HDC}$ and $RstB_{HDC}$. Protein concentrations were 0.5 μM CFP-$EnvZ_{HDC}$ and 0.5 μM YFP-$EnvZ_{HDC}$ (left panel) or 0.5 μM CFP-$RstB_{HDC}$ and 0.5 μM YFP-$RstB_{HDC}$ (right panel). The curves are fit as described in the methods. Cartoons of EnvZ and RstB show the HAMP, DHp and CA domains. (C) Pull-down assay with purified $EnvZ_{HDC}$ and $RstB_{HDC}$. A mixture of FLAG-labeled kinase and MBP-labeled kinase was incubated and complexes were isolated using anti-FLAG beads. Odd lanes show inputs and even lanes show elutions. The following mixtures were assayed: FLAG-$EnvZ_{HDC}$ + MBP-$EnvZ_{HDC}$ (lane 2), FLAG-$RstB_{HDC}$ + MBP-$RstB_{HDC}$ (lane 4), FLAG-$EnvZ_{HDC}$ + MBP-$RstB_{HDC}$ (lane 6), FLAG-$RstB_{HDC}$ + MBP-$EnvZ_{HDC}$ (lane 8). Non-specific binding to beads was assessed for MBP-$EnvZ_{HDC}$ (lane 10) and MBP-$RstB_{HDC}$ (lane 12).

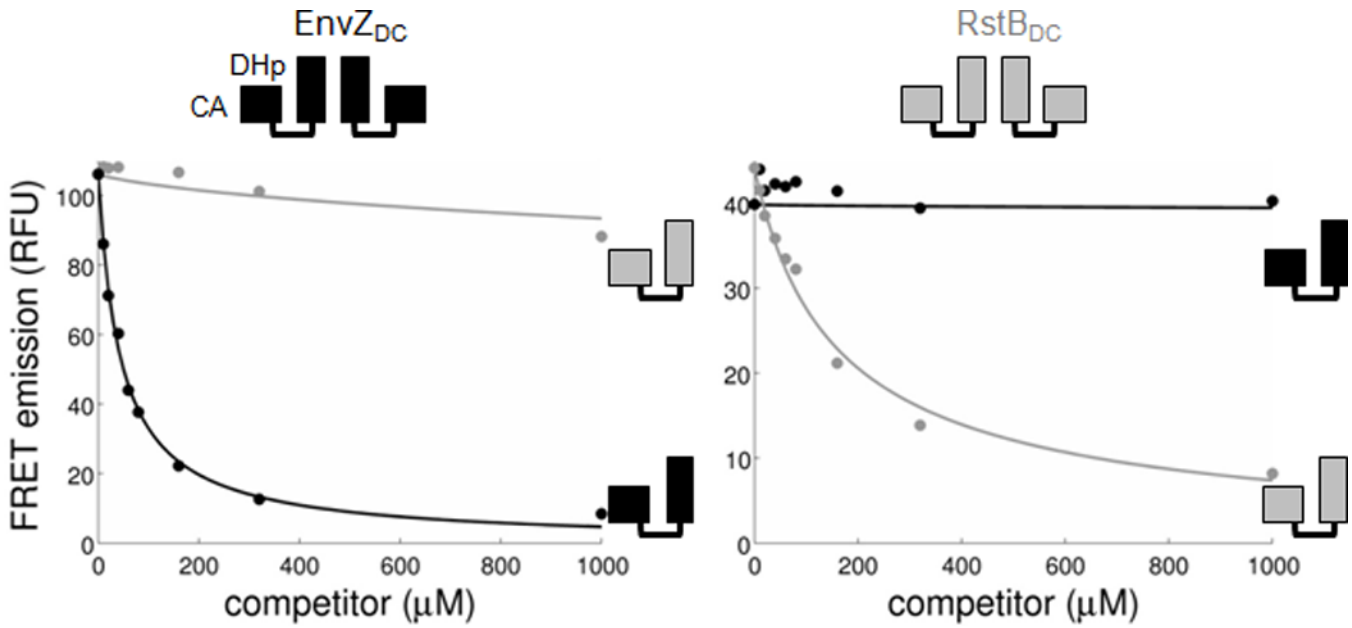

Figure 2. The DHp Domains of EnvZ and RstB Homodimerize

FRET competition assay for homo-association and hetero-association of $EnvZ_{DC}$ (left panel) and $RstB_{DC}$ (right panel). CFP and YFP fusion protein concentrations were 20 μM.

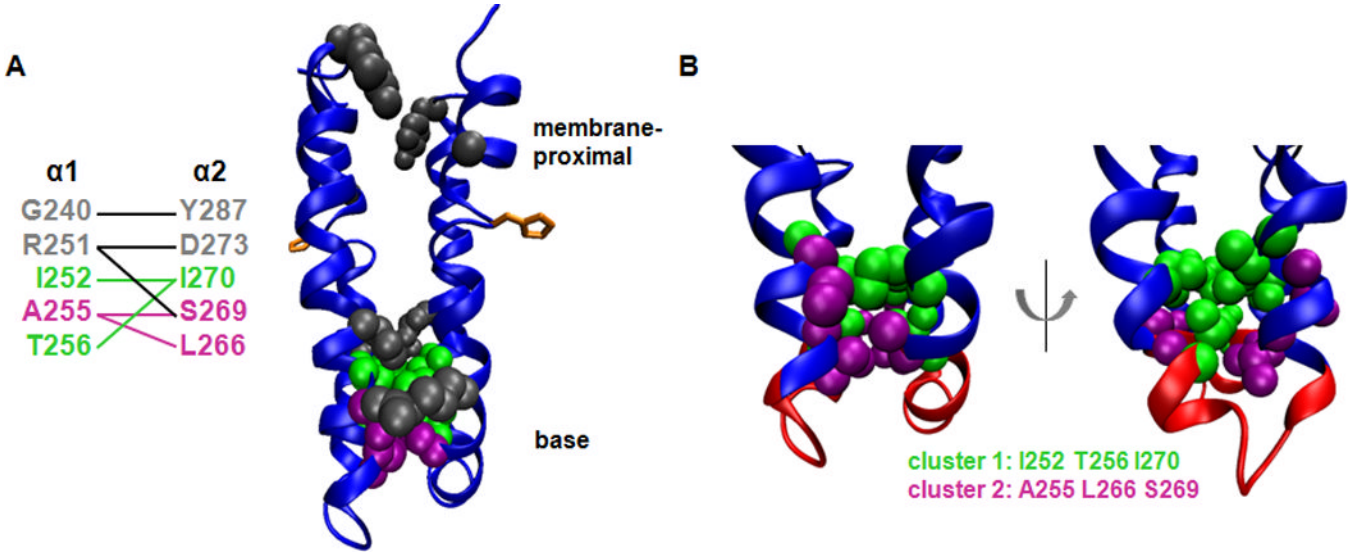

Figure 3. Covarying Pairs in the EnvZ Dimerization Interface

(A) Residues in EnvZ that highly covary and are within 5.5 Å across the dimerization interface are listed (at left) and shown on the NMR structure of EnvZ (at right).[23] Covarying residues in helices α1 and α2 are connected by lines. Cluster 1 residues are shown in green, cluster 2 residues are shown in purple and all other covarying, interchain residues are shown in grey. The conserved histidine is shown in orange. (B) Close-up views of clusters 1 and 2. The chimera-1-region backbone (residues 256–265) is in red; clusters 1 and 2 are outside that region.

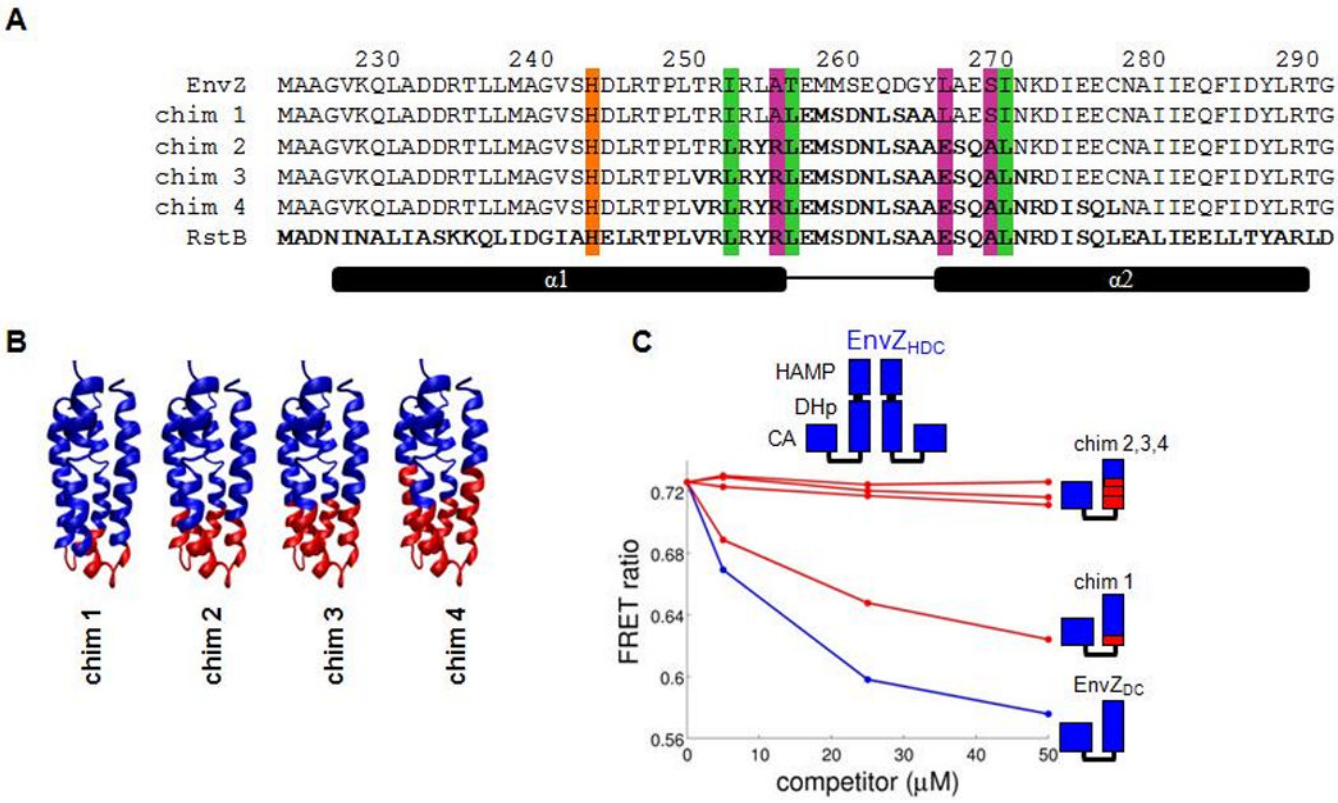

Figure 4. EnvZ-RstB Chimeric Proteins Isolate Dimerization Specificity to the DHp Domain Base

(A) Sequence alignment of the DHp domains of EnvZ, RstB and chimeras 1–4. Amino acids mutated, relative to EnvZ, to make the chimeras are shown in bold. The columns highlighted in color correspond to covarying residues in cluster 1 (I252, T256, I270) and cluster 2 (A255, L266, S269). The highlighted column in orange corresponds to the highly conserved histidine. Locations of the two helices in the DHp domain are indicated. (B) Models of chimeras 1–4 showing residues from EnvZ in blue and residues from RstB in red. The EnvZ NMR structure is used as a template. The chimeras are fused to the EnvZ CA domain, which is omitted in the models. (C) Interactions between $EnvZ_{DC}$ or chimeras 1–4 with $EnvZ_{HDC}$ (5 μM each of the CFP and YFP fusions).

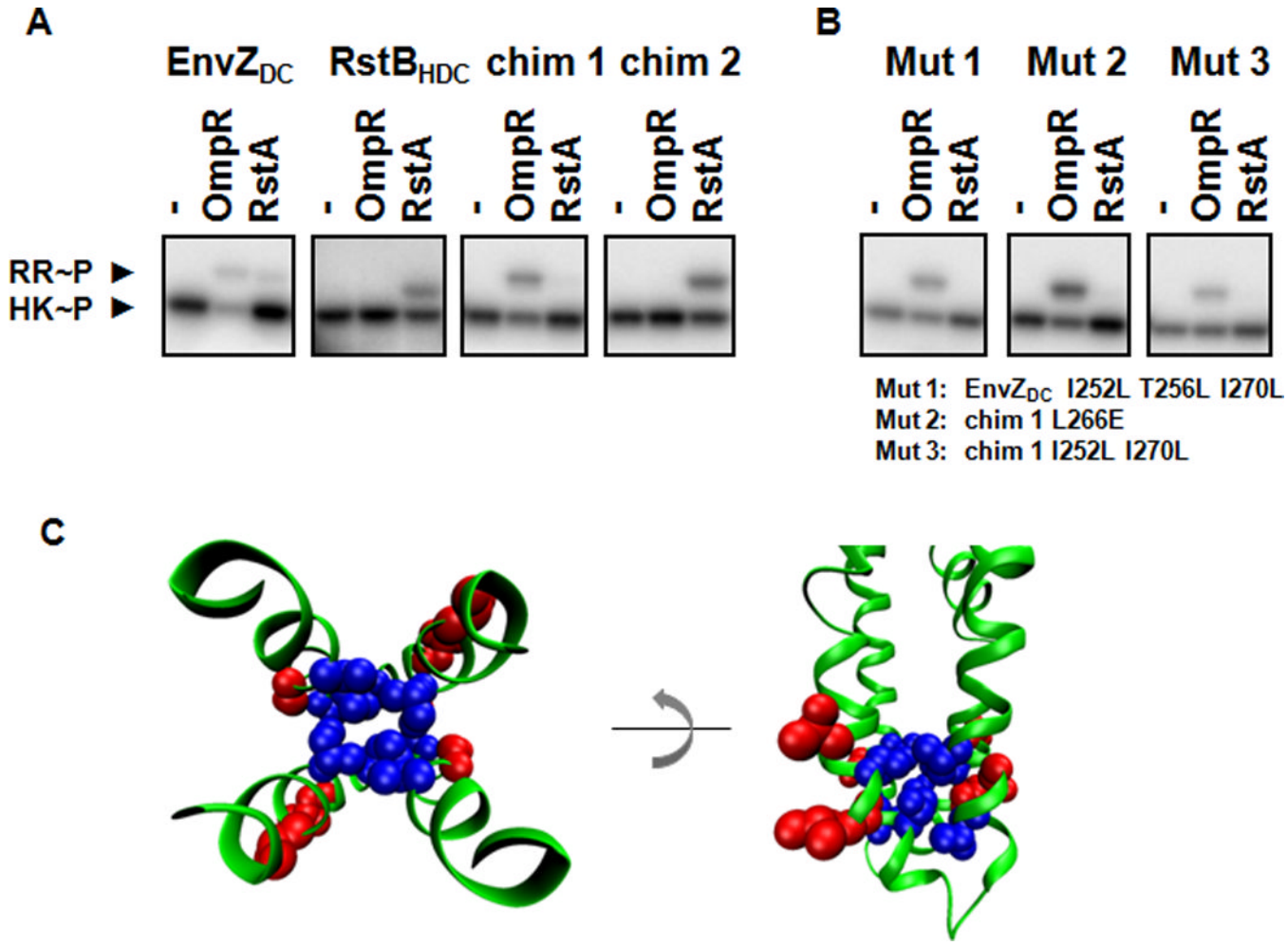

Figure 5. Phosphotransfer Specificity for Dimerization Specificity Switch Mutants

Each kinase (HK) was autophosphorylated with radiolabeled ATP and either incubated alone (−) or assayed for phosphotransfer to a response regulator (RR).

(A) $EnvZ_{DC}$ and $RstB_{HDC}$ phosphorylated their cognate response regulators, OmpR and RstA, respectively. Chimera 2 showed a switch in phosphotransfer specificity relative to chimera 1. (B) Phosphotransfer specificity of $EnvZ_{DC}$ dimerization specificity switch mutants Mut 1–3, defined as in the figure. (C) View from the base of the four-helix bundle in EnvZ, with positions affecting interaction specificity shown in space-filling form. Cluster 1 positions 252, 256, and 270, and cluster 2 position 266 (blue) are buried in the dimerization interface. Positions 250, 254, 255, and 269, previously identified to affect histidine kinase-response regulator phosphotransfer specificity (red) are solvent-exposed. The residues are mapped onto the EnvZ DHp domain.

Table 1

Dissociation constants for wildtype kinases, chimeras, and cluster 1 and 2 mutants (μM).[a]

	K_d (homodimer)	K_d (heterodimer with $EnvZ_{HDC}$)
wildtype kinases and chimeras		
$EnvZ_{HDC}$	0.4	
$RstB_{HDC}$	0.3	>50
$EnvZ_{DC}$	0.1	0.1
$RstB_{DC}$	270	
chim 1	0.1	0.3
chim 2	0.3	20.5
cluster 1 mutants		
$EnvZ_{DC}$ T256L I270L	1.7	1.7
$EnvZ_{DC}$ I252L I270L	23.0	22.1
$EnvZ_{DC}$ I252L T256L I270L	4.7	16.6
chim 1 I252L I270L	0.1	19.6
cluster 2 mutants		
$EnvZ_{DC}$ A255R L266E	>50	
$EnvZ_{DC}$ L266E	>50	
chim 1 L266E	0.9	4.4

[a] K_d values shown were averaged across duplicates. Limits in sensitivity of the FRET competition assay prevent measurement of homodimer K_d values tighter than 0.1 μM or heterodimer K_d values weaker than 50 μM.