Spatial tethering of kinases to their substrates relaxes evolutionary constraints on specificity

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Capra, Emily J., Barrett S. Perchuk, Orr Ashenberg, Charlotte A. Seid, Hana R. Snow, Jeffrey M. Skerker, and Michael T. Laub. “Spatial Tethering of Kinases to Their Substrates Relaxes Evolutionary Constraints on Specificity.” Molecular Microbiology 86, no. 6 (December 2012): 1393–1403.</th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1111/mmi.12064">http://dx.doi.org/10.1111/mmi.12064</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Wiley Blackwell</td>
</tr>
<tr>
<td>Version</td>
<td>Author's final manuscript</td>
</tr>
<tr>
<td>Accessed</td>
<td>Thu Jun 15 21:52:12 EDT 2017</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/86127">http://hdl.handle.net/1721.1/86127</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Creative Commons Attribution-Noncommercial-Share Alike</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td><a href="http://creativecommons.org/licenses/by-nc-sa/4.0/">http://creativecommons.org/licenses/by-nc-sa/4.0/</a></td>
</tr>
</tbody>
</table>
Spatial tethering of kinases to their substrates relaxes evolutionary constraints on specificity

Emily J. Capra¹, Barrett S. Perchuk¹, Orr Ashenberg¹, Charlotte A. Seid¹, Hana R. Snow¹, Jeffrey M. Skerker³, Michael T. Laub¹,²,⁴

¹Department of Biology  
²Howard Hughes Medical Institute  
³Massachusetts Institute of Technology  
⁴corresponding author: laub@mit.edu, 617-324-0418

Keywords: histidine kinase, coevolution, cross-talk, protein evolution

Running Title: Specificity of hybrid histidine kinases
Summary

Signal transduction proteins are often multidomain proteins that arose through the fusion of previously independent proteins. How such a change in the spatial arrangement of proteins impacts their evolution and the selective pressures acting on individual residues is largely unknown. We explored this problem in the context of bacterial two-component signaling pathways, which typically involve a sensor histidine kinase that specifically phosphorylates a single cognate response regulator. Although usually found as separate proteins, these proteins are sometimes fused into a so-called hybrid histidine kinase. Here, we demonstrate that the isolated kinase domains of hybrid kinases exhibit a dramatic reduction in phosphotransfer specificity in vitro relative to canonical histidine kinases. However, hybrid kinases phosphotransfer almost exclusively to their covalently attached response regulator domain, whose effective concentration exceeds that of all soluble response regulators. These findings indicate that the fused response regulator in a hybrid kinase normally prevents detrimental cross-talk between pathways. More generally, our results shed light on how the spatial properties of signaling pathways can significantly affect their evolution, with additional implications for the design of synthetic signaling systems.
Introduction

Cells can sense and respond to a remarkable diversity of signals and stimuli. This sensory capability typically involves a limited number of signal transduction protein families that have expanded through gene duplication. Although the relative ease of duplication and divergence has enabled cells to dramatically expand their signaling repertoires, the use of highly related signaling proteins has a significant cost, or risk. Cells must avoid detrimental cross-talk and ensure the fidelity of information flow through different signaling pathways. How the specificity of each signaling pathway is determined and how it evolves following gene duplication events are important problems that remain incompletely understood.

In bacteria, the dominant form of signal transduction is known as two-component signaling and typically involves a sensor histidine kinase that can autophosphorylate and then transfer its phosphoryl group to a cognate response regulator, which effects changes in cellular physiology or behavior (Stock et al., 2000) (Fig. 1A). Two-component signaling genes have undergone extensive duplication and horizontal transfer, such that most species possess tens or hundreds of these pathways (Galperin, 2005). Previous work has shown that the interaction between a histidine kinase and its cognate response regulator is highly specific with limited cross-talk between pathways in vivo (Capra et al., 2012, Fisher et al., 1996, Grimshaw et al., 1998, Laub & Goulian, 2007, Skerker et al., 2005). This specificity is determined predominantly at the level of molecular recognition rather than relying on cellular factors such as scaffolds. Consequently, a histidine kinase preferentially phosphorylates its cognate response regulator in vitro, relative to all other response regulators (Skerker et al., 2005).

Canonical histidine kinases harbor two highly-conserved domains, a dimerization and histidine phosphotransfer (DHp) domain and a catalytic and ATP binding (CA) domain. The DHp domain promotes homodimerization and harbors the histidine that is autophosphorylated by the CA domain. Response regulators also typically have two domains, a receiver domain and an output domain. The receiver domain contains a conserved aspartate that receives a phosphoryl group from the autophosphorylated kinase while the output domains are variable, but are often DNA-binding domains.
Phosphotransfer relies primarily on an interaction between the DHp domain of the kinase and the receiver domain of the regulator (Casino et al., 2009). The residues that determine the specificity of this interaction were identified through analyses of amino acid coevolution in large sets of cognate kinase-regulator pairs (Capra et al., 2010, Skerker et al., 2008). These studies pinpointed a small set of strongly coevolving residues that determine the specificity of two-component signaling proteins and that enable the rational rewiring of both the kinase and the regulator (Bell et al., 2010, Capra et al., 2010, Skerker et al., 2008).

The coevolution of specificity-determining residues in two-component signaling proteins is driven by negative selection against pathway cross-talk following gene duplication (Capra et al., 2012). The insulation of recently duplicated two-component proteins requires changes in the residues that govern molecular recognition, such that each cognate pair of signaling proteins continues interacting while avoiding cross-talk with the other pathway. In some cases, changes in the specificity residues of other two-component signaling proteins, that were not recently duplicated, are also necessary to achieve a system-wide insulation of all pathways in a given cell (Capra et al., 2012).

A common variant of two-component signaling involves hybrid histidine kinases, in which a conventional histidine kinase is fused to a receiver domain similar to those found in soluble response regulators (Fig. 1B). Hybrid kinases autoprophorylate and are thought to transfer the phosphoryl group intramolecularly to their receiver domains. The phosphoryl group can then be transferred to a histidine phosphotransferase and finally to a soluble response regulator, completing a phosphorelay. Hybrid histidine kinases are found in over 50% of all bacterial genomes and nearly 25% of all bacterial histidine kinases are hybrids (Wuichet et al., 2010). These hybrid kinases likely arise through the fusion of canonical, co-operonic histidine kinases and response regulators, and may further expand through gene duplication (Whitworth & Cock, 2009, Zhang & Shi, 2005).

Despite their prevalence, the phosphotransfer properties and specificity of hybrid kinases are poorly characterized relative to canonical histidine kinases. Here, we investigated the global phosphotransfer specificity of hybrid histidine kinases. We find that these hybrid kinases exhibit significantly reduced phosphotransfer specificity when liberated from their receiver domains. The covalently attached receiver domain thus normally serves as an intramolecular
phosphoacceptor and helps prevent unwanted cross-talk inside cells. Our data further indicate that, following the duplication of a hybrid kinase, there is reduced selective pressure to diversify the residues responsible for binding its attached response regulator domain, in stark contrast to canonical histidine kinases. In sum, we propose that the spatial arrangement of domains in hybrid histidine kinases strongly influences the evolution of these proteins with implications for understanding the evolution of multi-domain signaling proteins throughout biology and for designing synthetic circuits.
Results

Hybrid kinases show reduced amino acid coevolution between kinase and receiver domains

Analyses of amino acid coevolution using mutual information as a metric have helped pinpoint the residues that govern protein-protein interaction specificity in two-component signal transduction systems (Capra et al., 2010, Skerker et al., 2008). These analyses identified a small set of residues that map to the molecular interface formed during phosphotransfer (Casino et al., 2009), and were used to guide the rational rewiring of substrate specificity for the model histidine kinase EnvZ, validating their role in dictating specificity (Skerker et al., 2008). To assess whether the same residues coevolve in hybrid histidine kinases, we examined amino acid coevolution in a large set of hybrid kinases. This analysis was performed on a multiple sequence alignment containing 2681 hybrid histidine kinases, drawn from a wide phylogenetic range of organisms. This sequence alignment contained the DHp and CA domains of each hybrid kinase as well as its receiver domain, but omitted sensory domains. To measure coevolution we used a mutual information-based algorithm that helps adjust for phylogenetic and sampling biases in sequence alignments (Martin et al., 2005). Adjusted MI values were calculated for all possible pairs of positions within the sequence alignment (Fig. 1C, S1A-D). A similar analysis for canonical kinase-regulator pairs was used for comparison (Capra et al., 2010). The two alignments have similar entropy at each position, facilitating a comparison of mutual information scores (Fig S1E-F).

We focused primarily on residue pairs in which one position corresponds to a site within the DHp or CA domains and the other to a site within the receiver domain. The overall shape of the distribution of adjusted MI values was similar for the canonical kinase-regulator pairs and the hybrid kinase-receiver domain pairs (Fig. S1C-D). However, the hybrid kinase distribution did not contain the same long tail seen in the canonical distribution. There are 12 pairs of amino acids in the canonical kinase-regulator alignment that have adjusted MI values greater than 3.5, which indicates significant coevolution. In contrast, in the hybrid kinase-receiver domain alignment, no residue pair had an MI value greater than 3.5, and only one pair had a value greater than 3.0 (Fig. 1C). The scores for residue pairs in the hybrid kinase alignment were not simply reduced relative to those from the canonical alignment. Of the 12 top-scoring residue pairs from the canonical kinase-regulator alignment, only 5 were included
in the top 12 scoring pairs from the hybrid kinase alignment. The other 7 had substantially reduced scores, falling throughout the distribution, although each had a positive score (Fig. 1D). This analysis suggests that hybrid kinases do not exhibit the same extensive amino acid coevolution between DHp and receiver domains as canonical kinase-regulator pairs.

**Hybrid kinases exhibit limited phosphotransfer specificity**

To determine whether the reduced coevolution in hybrid kinases translates into a difference in kinase specificity, we performed phosphotransfer profiling (Skerker et al., 2005). In this approach, a histidine kinase is autophosphorylated using $[\gamma-^{32}P]ATP$ and then systematically tested for phosphotransfer to a large panel of full-length response regulators or receiver domains, using SDS-PAGE and phosphorimaging. Robust phosphotransfer typically manifests both with a band corresponding to a phosphorylated response regulator and, sometimes, with depletion of the radiolabeled kinase band.

We profiled 10 different hybrid kinases from the $\alpha$-proteobacterium *C. crescentus*. In each case we purified an epitope-tagged construct harboring the DHp and CA domains, but not the receiver domain. We first profiled each kinase against the entire set of receiver domains from the 27 annotated *C. crescentus* hybrid kinases, using incubation times of 15 minutes (Fig. 2A-B, S2). Strikingly, most of the kinases phosphorylated several of the hybrid kinase receiver domains. In fact, some kinases phosphorylated the majority of the receiver domains. These profiles stand in sharp contrast to our results with canonical histidine kinases in which the phosphotransfer profiles were typically extremely sparse, with kinases phosphorylating a single cognate response regulator (Skerker et al., 2008, Skerker et al., 2005).

Interestingly, not all of the hybrid histidine kinases phosphorylated their own receiver domains. For example, the kinase CC0723 phosphorylated the receiver domains of CC3075 and CC2670, but not its own, even though other hybrid kinases were able to phosphorylate the CC0723 receiver domain. There were also several cases in which a hybrid kinase phosphorylated its own receiver domain, but did so more weakly than other receiver domains. For example, CC3191 phosphorylated the CC0921 receiver domain to a greater extent than its own (Fig. 2A, S4B). Thus, unlike canonical kinases for which the cognate response regulator
is usually the kinetically preferred target, hybrid kinases display a variety of behaviors, and
often harbor substantially less specificity.

Next, we profiled each of the 10 hybrid kinases against the entire set of 44 canonical, soluble
response regulators encoded in the *C. crescentus* genome (Fig. 2C, S3). Although these
profiles were sparser than those performed against the hybrid kinase receiver domains, there
were significant interactions observed with several of response regulators. For instance, the
kinase domain of CC2501 showed significant phosphotransfer to the regulators CheYIV,
DivK, and CC3015. There were also several response regulators that were phosphorylated by
multiple hybrid kinases, including CC0630, CC2576, CC3015, and CC3286. Finally, we
noted that two hybrid kinases, CC0723 and CC2324, showed stronger phosphotransfer to
CC0630 than to any of the hybrid kinase receiver domains, including their own. These
profiles reinforce the conclusion that hybrid kinases exhibit relaxed phosphotransfer
specificity and are fundamentally different in this respect from canonical histidine kinases.

*Physical attachment of a receiver domain reduces signaling cross-talk*

Although our data demonstrated a reduced specificity of hybrid kinases, these profiles were
performed using kinases that had been physically separated from their receiver domains. The
kinetic preference and phosphotransfer behavior of these liberated kinase domains likely
differ substantially from those of full-length hybrid kinases. For example, although the kinase
domain for CC0138 (ShkA) phosphorylated 16 receiver domains and 3 full-length response
regulators, previous studies have indicated that ShkA exclusively phosphorylates its own
receiver domain *in vivo* (Biondi *et al.*, 2006b). Similarly, although the kinase domain of
CC1078 (CckA) showed apparent promiscuity *in vitro* and phosphorylated the response
regulator PetR, there is no evidence of cross-talk to this regulator *in vivo* and CckA does not
activate PetR-dependent genes *in vivo* (Biondi *et al.*, 2006a). Thus, we propose that the high
local concentration of a covalently attached receiver domain normally allows this domain to
outcompete other response regulators for access to an autophosphorylated kinase domain.

To further probe the effect of covalently attaching a receiver domain to a histidine kinase, we
focused on the hybrid kinase CC3191. We first compared the phosphotransfer behavior of the
CC3191 construct used in Fig. 2 that harbors the DHp and CA domains to a construct that
also contains the C-terminal receiver domain of CC3191. The kinase-only construct for CC3191 phosphorylated its own receiver domain in vitro, although it also phosphorylated the soluble response regulator CheYV at a similar rate (Fig. 2A, S4B). In contrast, the longer construct containing the C-terminal receiver domain no longer detectably phosphotransferred to CheYV (Fig. 3A, S4C). This result demonstrates that the receiver domain in a hybrid kinase normally prevents cross-talk between the kinase domain and other, soluble response regulators.

The suppression of cross-talk provided by a receiver domain could arise through steric hindrance or because the kinase domain is engaged in intramolecular phosphotransfer. To determine whether productive phosphotransfer contributes, we first generated a full-length CC3191 construct in which the phosphoaccepting aspartate (D563) in the receiver domain was mutated to alanine. This construct exhibited significantly more phosphotransfer to soluble CheYV than the wild-type CC3191 construct, indicating that engagement of the kinase domain in intramolecular phosphotransfer contributes to the suppression of cross-talk (Fig. 2B), although the receiver domain may also prevent cross-talk, in part, by occluding the binding of other regulators.

To further understand the contribution of a receiver domain to the prevention of cross-talk, we created chimeric hybrid kinases, fusing the kinase domain of CC3191 to a receiver domain from CheYIV or CC1182 (soluble response regulators) or from CC0026 or CC2670 (hybrid kinases). In our profiling studies, the liberated kinase domain of CC3191 had not detectably phosphorylated CheYIV, and had only weakly phosphorylated CC1182 and the receiver domain of CC2670, but it had strongly phosphorylated the receiver domain of CC0026 (Fig. 2C). To test whether these four chimeras could phosphotransfer intramolecularly from the CC3191 kinase domain to the heterologous receiver domain attached, we autophosphorylated each in buffer, acid, or base (Fig. 3A). Histidyl-phosphate bonds are sensitive to acid and aspartyl-phosphate bonds are sensitive to base (Fig. S4A). The phosphorylation of CC3191 was decreased in the presence of either acid or base, indicating that it was phosphorylated on both the histidine and aspartate. In contrast, the phosphorylation of CC3191(D563A) was primarily acid sensitive. Together, these patterns of acid/base sensitivity indicate that CC3191 normally autophosphorylates and transfers its phosphoryl group intramolecularly to its
receiver domain. We observed a similar pattern, consistent with intramolecular phosphotransfer, for the chimera CC3191-CC0026 and, to a lesser extent, CC3191-CC2670, but not CC3191-CheYIV or CC3191-1182. These findings are consistent with our results indicating that the CC3191 kinase domain alone can phosphorylate its own receiver domain and the receiver domains of CC0026 and CC2670, but not CC1182 or CheYIV (Fig. 2). These results also indicate that tethering non-cognate receiver domains to a histidine kinase is not always sufficient to promote phosphotransfer.

Next, we tested whether the four chimeras would phosphorylate, or cross-talk to, soluble CheYV. All four chimeras showed reduced phosphotransfer to CheYV compared to the CC3191 kinase-only construct (Fig. 3B, S4C), with the strongest suppression of cross-talk occurring with CC3191-CC2670 and CC3191-CC0026, the two chimeras that also demonstrated the most significant intramolecular phosphotransfer. Only the CC3191-CC0026 chimera, whose kinase and receiver domains displayed an interaction similar to that of CC3191-CC3191, both in isolation and when fused, completely prevented cross talk. Taken together, our results indicate that the receiver domain of a hybrid histidine kinase plays an important role in reducing, or eliminating, cross-talk with other response regulators by interacting with, and receiving phosphoryl groups from, the linked kinase domain.

**Hybrid kinases lacking their receiver domains likely cross-talk to other response regulators in vivo**

Previous work has shown that, with only a few exceptions, canonical histidine kinase-response regulator pairs are insulated from each other *in vivo* (Laub & Goulian, 2007, Skerker et al., 2005) and, importantly, that cross-talk between non-cognate pairs can be severely detrimental to an organism's fitness (Capra et al., 2012). We have shown here that many of the hybrid kinases, when separated from their receiver domains, interact readily with noncognate response regulators *in vitro*. Thus, we hypothesized that expressing only the kinase domain of a hybrid histidine kinase might induce cross-talk *in vivo* and affect the growth or fitness of cells.

We tested this hypothesis by inducing expression of CC3191 lacking its C-terminal receiver domain in *C. crescentus* and assessing cellular growth in swarm plates. Wild-type *C.
crescentus cells can swim through low-percentage agar, creating a large circular colony, or swarm; defects in motility, chemotaxis, cell growth, or cell division can affect swarm size, making this a convenient assay for assessing gross cellular phenotype (Skerker et al., 2005). We found that cells producing the kinase-only portion of CC3191 produced a small swarm relative to the wild type without affecting growth or morphology. This observation is consistent with the notion that a kinase-only version of CC3191 inappropriately phosphotransfers to CheYV in vivo, as it does in vitro (Fig. 2C). In contrast, cells synthesizing either a full-length construct that contains the receiver domain or the receiver domain alone did not exhibit significant swarm phenotypes (Fig. 3C-D). The phenotype seen with cells expressing the kinase portion of CC3191 was dependent on autophosphorylation, as cells overexpressing a construct in which the conserved histidine was mutated to an alanine no longer exhibited a severe swarm phenotype.

We then tested the effects of overexpressing three other hybrid histidine kinases that we profiled above: CC0026, CC0138, and CC2670. Like CC3191, these kinases do not contain transmembrane domains. As with CC3191, overproducing the N-terminal and kinase domains of CC0138 and CC2670 led to a small swarm phenotype, whereas constructs containing both the kinase and receiver domains, or the receiver domain alone, did not (Fig. 3D, S4D). For the kinase-only constructs of CC0138 and CC2670, the phenotype was suppressed by substituting the phosphorylatable histidine with an alanine suggesting that autokinase activity is required for the small swarm phenotype. Unlike CC0138 and CC2670, cells synthesizing the kinase-only version of CC0026 did not exhibit a significant swarm phenotype. Notably, however, the kinase domain of CC0026 had not significantly phosphorylated any non-hybrid receiver domains in vitro (Fig. 2C). Taken together, these data are consistent with the idea that some hybrid kinases are promiscuous, but that their attached receiver domains normally help to prevent cross-talk with other response regulators in vivo.

**Hybrid histidine kinases are under reduced selective pressure to diversify**

Collectively, our results indicate that hybrid histidine kinases are subject to different selective pressures than canonical histidine kinases. We previously found that canonical histidine kinases and response regulators are under strong selective pressure to diversify their specificity residues following gene duplication, but are otherwise relatively static (Capra et
This diversification of specificity residues post-duplication is critical to preventing cross-talk and ultimately ensures the system-wide optimization of phosphotransfer specificity (Capra & Laub, 2012, Capra et al., 2012). Consistently, inspection of the six key specificity residues (those from α-helix 1 in the DHp domain) in genome-wide sets of canonical histidine kinases indicates fewer than three identities at these six positions in most pairwise comparisons (Fig. S5).

We extracted the corresponding six residues from each of 24 hybrid histidine kinases in C. crescentus (Fig. S5). Although there are 27 annotated hybrid kinases that contain CA and receiver domains, 3 did not have intact DHp domains. Strikingly, many of the 24 hybrid kinases share four, five, or even six identities at these positions with other hybrid kinases. This similarity does not arise simply because the hybrid kinases duplicated recently, as pairwise comparisons of the entire DHp and CA domains demonstrated extensive variability at other sites (Fig. S1E-F), resulting in significant separation in a neighbor-joining tree built from those domains (Fig. 4A).

The lack of variability at the sites corresponding to the six key specificity residues in canonical kinases was also evident in sequence logos for the 24 hybrid and 21 canonical kinases from C. crescentus (Fig. 4B). The logo for canonical kinases indicated relatively low conservation at each specificity position except the first, which may be constrained due to involvement in autophosphorylation (Capra et al., 2010, Casino et al., 2010). In contrast, the logo for hybrid kinases indicated higher conservation at each site.

The kinase domains of hybrid histidine kinases are likely under less selective pressure than canonical kinases to diversify following gene duplication. The effective concentration of the attached receiver domain is high enough to ensure that a hybrid kinase will transfer its phosphoryl group intramolecularly and not to another regulator or receiver domain. Hence, after duplication of a hybrid kinase, the residues that bind to the receiver domain do not need to change to insulate the new proteins from one another, as occurs in canonical kinases (Fig. 5). Consistent with this hypothesis, many of the hybrid histidine kinases in C. crescentus, which were likely derived from a common ancestral gene through duplication and divergence, had similar specificity residues and exhibited similar phosphotransfer profiles when liberated from their receiver domains (Fig. 2B). One exception to this trend was CC1078 (CckA),
which had a distinct set of specificity residues relative to the other hybrid kinases and, consequently, had a significantly different phosphotransfer profile. Notably, CckA did not group with the other hybrid kinases in a tree of *Caulobacter* kinases (Fig. 4A) suggesting that CckA may be relatively ancient and not derived from a recent duplication.
The expansion of existing signaling protein families has enabled cells to rapidly evolve the ability to sense and response to a wide range of stimuli. In bacteria, two-component signaling proteins have expanded dramatically, such that most species encode dozens, and sometimes hundreds, of these proteins. For canonical pathways involving a single histidine kinase and response regulator, these pathways are exquisitely specific and a cognate response regulator can outcompete all other non-cognate regulators to receive phosphoryl groups from a given histidine kinase. Consequently, phosphotransfer profiles of canonical kinases have demonstrated that each possesses a strong kinetic preference for its cognate substrate (Skerker et al., 2005). This preference is determined by a small number of specificity-determining residues in both the kinase and regulator. These residues must coevolve to maintain a tight, specific interaction between cognate partners, particularly after a gene duplication event as a means of insulating the new pathways from one another (Fig. 5) (Capra et al., 2012).

In contrast to the canonical systems, we demonstrated here that kinase domains of hybrid kinases typically exhibit relaxed substrate specificity, often phosphorylating soluble response regulators or other receiver domains as well or better than they phosphorylate their own receiver domains. A similar observation was made previously in *Myxococcus xanthus* with a limited set of response regulators. In that case, the kinase domain of RodK was shown to preferentially phosphorylate the soluble regulator RokA relative to its own receiver domain, RodK-R3 even though the latter is the *in vivo* target of RodK (Wegener-Feldbrugge & Sogaard-Andersen, 2009).

Although hybrid kinases are more promiscuous on their own, our data indicate that the covalently attached receiver domain helps to prevent cross-talk with other cytoplasmic response regulators. The local concentration of an attached receiver domain likely exceeds the concentration of all soluble response regulators quite significantly. Consequently, intramolecular phosphotransfer from the kinase domain to the attached receiver domain will be strongly favored, thereby ensuring minimal cross-talk to other pathways.

The enforcement of intramolecular phosphotransfer specificity through spatial tethering of domains likely eliminates selective pressure to diversify the residues in a hybrid kinase that
mediate docking to the receiver domain. Hence, after a hybrid kinase duplicates, these residues either will not change or will change more rarely through processes such as genetic drift (Fig. 4B). The net result of the reduced rate of change is that for hybrid kinases in extant organisms, the interfacial residues show substantially reduced variability compared to the same set of residues in canonical histidine kinases.

The enforcement of phosphotransfer within hybrid kinases has also likely reduced the need for their kinase and receiver domains to coevolve (Fig. 1). Mutations that reduce or weaken the interaction of these domains are probably more easily tolerated because the domains are spatially tethered. By contrast, with canonical two-component pathways, the cognate proteins are under strong pressure to coevolve, as a means of maintaining their interaction and preventing interaction with non-cognate proteins. However, merely increasing the effective concentration of a receiver domain was not always sufficient to induce phosphotransfer from a kinase domain (Fig. 3A) indicating some requirement for molecular recognition and a proper pairing of interfacial residues. It may be that the fusion of domains in a hybrid kinase serves primarily to prevent cross talk, rather than driving phosphotransfer.

Why some two-component pathways involve hybrid histidine kinases instead of canonical kinases is not clear. Hybrid kinases are often involved in phosphorelays, and the additional number of components in a phosphorelay may create additional points for integrating signals (Burbuly et al., 1991). However, not all hybrid kinases necessarily participate in phosphorelays. Recent work with the hybrid kinase VirA from Agrobacterium tumefaciens suggests that the receiver domain binds the response regulator VirG, somehow stimulating its activity as a transcriptional activator (Wise et al., 2010). There are also hybrid kinases in some Gram-positive bacteria, such as Bacteroides thetaiotaomicron, that have DNA-binding domains C-terminal to their receiver domains, suggesting that these kinases may directly regulate transcription (Raghavan & Groisman, 2010). In short, although nearly a quarter of all kinases are of the hybrid variety, our understanding of their functions, properties, and advantages remains limited.

The notion that spatial proximity can overcome relaxed specificity of signaling proteins is relevant in all cells. Multi-domain signaling proteins are quite common, particularly in eukaryotes. Additionally, some signal transduction proteins are spatially constrained through
the action of scaffolds. For example, in the *S. cerevisiae* pheromone pathway, the scaffold Ste5 enforces the proximity of three separate MAP kinases, helping to prevent them from inappropriately phosphorylating other substrates (Choi *et al.*, 1994). This spatial colocalization may, in turn, have relaxed evolutionary constraints on these MAP kinases.

Finally, our results suggest that information flow through two-component pathways could be rationally engineered by fusing together non-cognate kinases and regulators. Such an arrangement can also prevent unwanted cross-talk with other pathways. Indeed, we showed here that fusing heterologous receiver domains to a hybrid kinase was, in some cases, sufficient to allow phosphotransfer and prevent cross-talk with a soluble regulator. Synthetic scaffolds that bring non-cognate two-component signaling proteins in close proximity may also be used to promote phosphotransfer or prevent cross-talk. A similar approach of artificially colocalizing proteins has been applied in metabolic engineering studies, where enzymes have been tethered together to enhance the synthesis and yield of desired compounds (Dueber *et al.*, 2009).

In sum, our work has revealed new aspects of signaling protein evolution in bacteria that will likely inform similar evolutionary studies in other organisms and help guide efforts to construct synthetic signaling circuits.
Experimental Procedures

Sequence analyses
Histidine kinase and response regulator receiver domains were identified, aligned, and filtered as described previously (Capra et al., 2010). Hybrid kinases were defined as those proteins that had a single match to each of the three Pfam models: HisKA, HATPase_C, and Response_reg. The final alignment included 2681 hybrid kinases. Shannon entropy values were calculated for each position in the alignment. Mutual information for every pair of columns in the sequence alignment was calculated as previously reported (9). Raw and adjusted MI values are provided in Tables S1 and S2 and multiple sequence alignments are provided as Supporting Information. Sequence logos were built using WebLogo (weblogo.berkeley.edu). Neighbor-joining trees were built using the PHYLIP package and multiple sequence alignments built from the DHp domain of each canonical and hybrid histidine kinase in the C. crescentus genome.

Strain construction and growth conditions
E. coli and C. crescentus strains were grown as described previously (Skerker et al., 2005). Primers used are listed in Table S3. Full-length hybrid kinases and the kinase domains of hybrid kinases were amplified from genomic CB15N DNA and ligated into the Gateway pENTR vector (Invitrogen). Chimeric hybrid kinases were cloned by separately amplifying the kinase domain from CC3191 and the specified receiver domain, amplifying the chimeric sequence using splicing with overlap extension PCR and ligating the resulting product into pENTR. pENTR clones were moved into pDEST-His6-MBP or pDEST-TRX-His6 vectors for purification, or the pDEST-Pxyr-M2 vector derived from pJS71 for overexpression studies. Overexpression vectors were introduced into wild-type CB15N via electroporation.

Protein purification and phosphotransfer assays
Expression, protein purification, and phosphotransfer profiling experiments were carried out as described previously (Capra et al., 2012, Skerker et al., 2008, Skerker et al., 2005, Biondi et al., 2006a). All reactions used 500 µM ATP, and 0.5 µCi/µL [γ-32P]ATP. For phosphotransfer experiments in Fig. 3A, CC3191-HK was autophosphorylated under the same
conditions as the phosphotransfer profiles and then incubated with the given receiver domain in a 1:1 ratio for the time indicated. For phosphotransfer experiments in Fig. 3C, 2.5 μM of the specified kinase was mixed with 2.5 μM CheYV before ATP was added the reaction allowed to proceed for the indicated time before being stopped with the addition of 4X loading buffer. To test acid or base stability of phosphoryl groups, 5 μM of kinase was autophosphorylated at room temperature for 15 minutes. The reaction was then stopped by the addition of 4X loading buffer, and then buffer, 1 M HCl or 0.5 M NaOH was added. After 20 minutes, reactions were neutralized. All phosphotransfer experiments were analyzed by SDS-PAGE and phosphorimaging.

Acknowledgements

We thank Anna Podgornaia for helpful comments on the experiments and on the manuscript. This work was supported by an NSF CAREER award to MTL and an NSF GRFP award to EJC. MTL is an Early Career Investigator at the Howard Hughes Medical Institute.


Figure Legends

Figure 1. Amino acid coevolution analysis of hybrid histidine kinases. (A) Diagram of canonical two-component signaling pathways and (B) phosphorelays, indicating the conserved domains in each protein. (C) Coevolving residues in cognate pairs of canonical histidine kinases and response regulators. Residue pairs with adjusted mutual information scores greater than 3.5 are listed, connected by lines (left), and shown in spacefilling on a structure of the *T. maritima* HK853-RR468 complex (right). The only pair in the hybrid kinase alignment with a score greater than 3.0 is highlighted. For clarity, only the DHp domain of HK853 is shown. Residue numbers correspond to positions within EnvZ and OmpR (see Fig. S1A-B). (D) Histogram of adjusted mutual information scores for all residue pairs in the hybrid histidine kinase alignment. Arrows indicate the residue pairs scoring higher than 3.5 in the analysis of canonical two-component proteins, with scores for these pairs in each alignment listed in the table.

Figure 2. Hybrid histidine kinases show reduced phosphotransfer specificity *in vitro*. (A) Phosphotransfer profiles for kinase domains from three *C. crescentus* hybrid histidine kinases against all 27 receiver domains from hybrid kinases. (B) Quantification of phosphotransfer profiles for 10 hybrid kinases against the 27 hybrid kinase receiver domains; for raw profile data, see Fig. S2. (C) Quantification of phosphotransfer profiles for 10 hybrid kinases against the 44 soluble *C. crescentus* response regulators; for raw profile data, see Fig. S3. For panels B-C, the ratio of receiver domain or response regulator band intensity to the autophosphorylated kinase band intensity was calculated and converted to color based on the legend shown. All phosphotransfer reactions were incubated 15 minutes.

Figure 3. Hybrid kinases lacking their receiver domains exhibit cross-talk. (A) Chimeric hybrid kinases were autophosphorylated in the presence of buffer, HCl, or NaOH to assess whether phosphoryl groups resided on the conserved histidine, aspartate, or both. (B) Chimeric hybrid kinases were autophosphorylated and then tested for phosphotransfer to soluble CheYV at the time points indicated. Error bars represent standard deviation from three independent replicates. Raw gel images are shown in Fig. S4C. The identity of domains in each chimeric kinase are listed. (C) Swarm plate assay for strains expressing each of the CC3191 constructs listed or vector alone. (D) Quantification of swarm sizes for strains
expressing various constructs for each of the four hybrid histidine kinases indicated. Swarm areas were measured and plotted relative to the empty vector control. Error bars represent standard deviations from three replicates. Swarm plate images are shown in Fig. S4D.

**Figure 4.** Specificity residues are conserved among hybrid histidine kinases. (A) An unrooted neighbor-joining tree of the *C. crescentus* kinases was built from an alignment of the DHp domains of all 24 hybrid and 21 canonical histidine kinases from *C. crescentus*. Hybrid kinases are labeled in red. (B) Sequence logos for the residues that dictate phosphotransfer specificity in canonical kinase-regulator pairs. Logos were built from an alignment of the 21 canonical histidine kinases and 44 soluble response regulators (top), and from an alignment of the 24 hybrid histidine kinases in *C. crescentus* (bottom).

**Figure 5.** Model for changes in specificity residues following duplication of canonical and hybrid histidine kinases. Ovals represent niches within sequence space, or the set of response regulators recognized by a given histidine kinase as determined by its specificity residues. Post-duplication, canonical kinases separate in sequence space to insulate the two pathways and prevent cross-talk. In contrast, hybrid kinases do not separate, as the tethered receiver domain effectively insulates the duplicated kinases against cross-talk.
A

signal → DHp → CA → PH → P → RD → output
histidine kinase

B

signal → DHp → CA → RD → H → P → Hpt → RD → output
hybrid histidine kinase
histidine phosphotransferase
response regulator

C

canonical: score > 3.5

HK

HK853

RR468

RD

D

residue pair | canonical score | hybrid score
---|---|---
250-16 | 6.2 | 2.5
251-13 | 3.9 | 0.4
251-15 | 4.2 | 2.7
251-19 | 3.9 | 2.2
254-19 | 5.8 | 3.3
255-19 | 4.3 | 2.5
257-108 | 4.2 | 1.2
258-22 | 4.1 | 0.9
258-23 | 4.3 | 1.0
269-15 | 4.4 | 0.5
272-15 | 3.5 | 0.2
273-15 | 4.5 | 1.3

adjusted MI score for hybrid kinases

# of residue pairs
canonical histidine kinases

HK1
HK2
HK1
HK2
HK1
HK3

ancestral state, pre-duplication

gene duplication

derived state, post-duplication

hybrid histidine kinases

HK1
HY1
HK1
HY1
HK1

HY2
HY1

HY2