This manuscript has been accepted for publication in *Annual Review of Microbiology* (2019, vol. 73).

**STRUCTURAL BASIS OF RESPONSE REGULATOR FUNCTION**

Rong Gao, Sophie Bouillet, Ann M. Stock

Center for Advanced Biotechnology and Medicine, Department of Biochemistry and Molecular Biology, Rutgers – Robert Wood Johnson Medical School, Piscataway, New Jersey 08854, USA; email: rgao@cabm.rutgers.edu, sb1722@cabm.rutgers.edu, stock@cabm.rutgers.edu

**Author ORCID Numbers:**

Gao: 0000-0002-9245-4952

Bouillet: 0000-0002-0113-3169

Stock: 0000-0002-0446-8079

**Corresponding Author:** Ann Stock

CABM

679 Hoes Lane West

Piscataway, NJ 08854

+1 848-445-9812

Shortened Title: Response Regulator Structure Function
Keywords

histidine kinase, phosphorylation, response regulator, signal transduction, transcription factor, two-component system
Abstract

Response regulators function as the output components of two-component systems, which couple the sensing of environmental stimuli to adaptive responses. Response regulators typically contain conserved receiver (REC) domains that function as phosphorylation-regulated switches to control the activities of effector domains that elicit output responses. This modular design is extremely versatile, enabling different regulatory strategies tuned to the needs of individual signaling systems. This review summarizes functional features that underlie response regulator function. An abundance of atomic resolution structures and complementary biochemical data have defined the mechanisms for response regulator enzymatic activities, revealed trends in regulatory strategies utilized by response regulators of different subfamilies and provided insights into interactions of response regulators with their cognate histidine kinases. Among the hundreds of thousands of response regulators identified, variations abound. This article provides a framework for understanding structural features that enable function of canonical response regulators and a basis for distinguishing non-canonical configurations.
1. INTRODUCTION

Two-component systems (TCSs) are the predominant multi-step signaling pathways in bacteria. Two conserved proteins, a histidine protein kinase (HK) and response regulator (RR), constitute the core system (Figure 1a). Autophosphorylation of the HK at a conserved His provides a high-energy phosphoryl group that is transferred to a conserved Asp in the RR, resulting in RR activation (99). Stimuli, sensed either directly or indirectly by the HK, regulate opposing autophosphorylation/phosphotransfer and RR phosphatase activities of the HK, thus determining the level of phosphorylation of the RR and the output response. A large variety of input domains in HKs and output domains in RRs allows the coupling of an almost limitless array of chemical or physical stimuli to diverse output responses. The conserved proteins and the systems themselves are extremely versatile. An enormous range of variations in protein activities, domain architectures and system configurations allows adaptation of TCSs to the needs of specific signaling systems.

TCSs are found in bacteria, archaea (41) and eukaryotes such as slime molds, yeast, and plants (2). While great diversity exists among all TCSs it should be noted that some specific trends in system configurations are found in different organisms. The Pfam database (35) lists >342,000 entries of proteins containing the conserved receiver (REC, named Response_reg in Pfam) domain that characterizes RRs and >650 structures of such proteins in the Protein Data Bank (PDB) (18). In addition to their presence in RRs, REC domains are also found in hybrid HKs that function within phospho-relay systems involving multiple phosphotransfer steps, accounting for ~10% of REC domain-containing proteins. The scope of this review will focus on bacterial RRs that mediate output of TCSs. Among these RRs, variations abound. The plasticity of the REC domain and the versatility of TCS architecture has driven the evolution of RRs that are uniquely adapted and fine-tuned for function in specific pathways. Out of necessity, this review will focus on canonical RRs with full
understanding, but lack of specific acknowledgement of individual cases, that for every feature noted, exceptions exist. This review provides an overview of the structure of RRs and features that provide conserved enzymatic activities in REC domains. Beyond the core domain structures and common enzymatic mechanisms, RRs display great variation in domain arrangements that provide a variety of mechanistically distinct regulatory mechanisms. However, as structures have accumulated, trends in regulatory strategies used by different subfamilies of RRs have begun to emerge and these will be explored in this review.

2. RESPONSE REGULATOR ARCHITECTURE

REC Domains

RRs are defined by the presence of a conserved REC domain. The REC domain fold is composed of five α helices surrounding a central five-stranded parallel β-sheet with a 21345 topology (Figure 1b). Sequence identity between REC domains is usually 20-30%, showing great variations (Figure 1c). The central hydrophobic β-strands, β1, β3, β4 and β5, are more conserved than the peripheral helices and loops. The most conserved residues include the site of phosphorylation, an Asp at the C-terminus of β3, and several other residues in the β-α loops, which constitute the active site. The highly conserved phosphorylation site and the variable peripheral sequences allow RRs to couple phosphorylation to diverse effector functions.

Effector Domains

RRs typically function as the output components of signaling pathways. Regulatory REC domains can be linked either covalently or non-covalently to a great diversity of effectors and thus control numerous diverse responses. Figure 2 summarizes the distribution of the major RR effector domains.
identified in the Pfam database. Approximately one fourth of RRs (23%) consist solely of a REC domain. These single-domain RRs (SDRRs) regulate intermolecular effectors such as the chemotaxis system regulator CheY, which binds to the flagellar motor component FliM controlling flagellar rotation. CheY-like proteins found in chemotaxis-like or chemosensory systems mediate responses other than chemotaxis, for example biofilm formation regulated by Cle proteins (77). Other functions for SDRRs include acting as intermediates in phosphorelays (e.g. the general stress response proteins SdrG and MrrA, and the sporulation protein Spo0F) (45, 63), as phosphate sinks to control phosphate flux within phosphorelays (e.g. Rhizobium meliloti CheY1) (97), as allosteric regulators of HKs (e.g. Caulobacter crescentus DivK) (24, 117) or as protease adapters (e.g. C. crescentus CpdR) (56).

The largest class of RRs (64.5%) is composed of RRs containing a DNA-binding domain (DBD) with subfamilies defined by different DBD folds. The OmpR subfamily (29% of RRs) contains a winged-helix effector domain (68), the NarL subfamily (19%) a four-helix DNA-binding HTH domain (73), the NtrC subfamily (7%) a Fis-type HTH domain fused to an AAA+ ATPase domain and the LytTR (5.5%) a predominantly β fold (94). The abundance of RRs in this class likely reflects the importance of transcriptional regulation as a response to environmental change. A small class of RRs (1%) harbors an RNA-binding domain belonging to the ANTAR subfamily of anti-termination factors such as AmiR (79).

RRs with enzymatic domains account for ~8% of RRs. A major group within this class (3% of RRs) is the regulators of cyclic di-GMP including cyclases (GGDEF) and phosphodiesterases (EAL, HD-GYP) (92). A second major group (2%) is the chemotaxis methylesterase CheB proteins (31). A variety of different enzymatic domains have been identified in other RRs (3%) such as the PP2C phosphatase domain of RsbY and the hybrid kinase domain of FrzE (30, 54). RRs with protein
binding domains account for ~1% of RRs. Examples in this class include CheV, with a CheW-like domain that connects chemoreceptors to the chemotaxis histidine kinase CheA; RssB, which regulates turnover of the stress response sigma factor RpoS; and PhyR, a regulator of the general stress response that contains an extracytoplasmic function (ECF) sigma factor-like domain (1, 11, 66). While RRs often contain a simple REC-effector domain architecture they can have complex domain organizations with additional signaling domains including PAS, GAF, HisKa, HATPase, etc. The large variety of RRs identified to date emphasizes the versatility of the REC domain with no apparent limits on the types of effector domains that can be controlled by this phosphorylation-activated switch.

3. ENZYMATIC ACTIVITIES OF RRs

Conserved Phosphorylation Site

One of the defining characteristics of the REC domain is its highly conserved phosphorylation site. Due to the lability of the high-energy acyl phosphate, structural characterization of the active site often relies on berylliofluoride (BeF$_3^-$) that noncovalently binds to the phosphorylation site Asp and serves as a mimic of the phosphoryl group (119). Different BeF$_3^-$-bound RR structures reveal a conserved active site with a network of hydrogen bonds (Figure 3a). The carboxylate side chains of the acidic residue duo (DD) at the β1-α1 loop participate in coordinating a Mg$^{2+}$ required for catalysis. Sidechains of two additional residues, a Thr/Ser (T) at the β4-α4 loop and a Lys (K) at the β5-α5 loop, together with backbone atoms of non-conserved active site residues, coordinate phosphate oxygens in the phosphorylated REC domain.

RR phosphorylation level, the ultimate determining factor of TCS output for canonical RRs, is regulated by multiple enzyme activities, including phosphotransfer from HKs, dephosphorylation...
by auxiliary phosphatases or bifunctional HKs that also possess RR phosphatase activity,
autophosphorylation by small-molecule phosphodonsors such as phosphoramidate and acetyl phosphate (AcP) (67, 115), and autodephosphorylation activities. Similar phosphorylation sites with almost identical positioning of active site residues have been observed in numerous RR or RR complex structures (21, 81, 86, 107). A pentavalent phosphorus intermediate is believed to be a common transition state for different activities of the REC domain (Figure 3b). Both phosphorylation and dephosphorylation reactions can proceed through either an associative or a dissociative mechanism depending on how tight or extended the transition state is. A phosphodonor or water molecule needs to be in line with the acyl phosphate bond that is to be formed or broken. The phosphotransferase or phosphatase helps position these molecules to further enhance the reaction rates. In all cases, the majority of residues involved in coordinating the phosphorus intermediate are from the REC domain. Thus, most RRs are catalytically competent of autophosphorylation and autodephosphorylation in the absence of any enzymatic protein partner and in vitro analyses of these reactions often provide insights into RR regulatory mechanisms.

**Autodephosphorylation and Autophosphorylation**

Despite the highly conserved active site geometry, RRs show large variations in rates of autophosphorylation and autodephosphorylation (87, 101, 102). For example, autodephosphorylation rate constants of the REC domain range over six orders of magnitude, giving phosphorylation half-lives of seconds to hours and even days (101). Such great diversity is partly attributed to several variable residues surrounding the active site, such as positions D+2 (two residues C-terminal to the conserved D), T+1 and T+2 (one and two residues C-terminal to T) (Figure 3b and 3c). Backbone atoms of D+2 and T+1 directly form hydrogen bonds with the phosphoryl group while the charge, size and hydrophobicity of sidechains at these three positions may affect the energy barrier of the
transition state, facilitate the positioning or block the in-line path of the phosphodonor or attacking water (53, 80, 102, 121). In more than half of RR sequences, a limited number of amino acid combinations are preferred at these positions and the distribution of preferred amino acids correlates with RR effector subfamilies (53, 80). REC enzyme activities are therefore suggested to co-evolve with effector domain regulation and these residues represent functional sites for modulating the stability of RR phosphorylation to match the timescale of individual TCS output responses.

Phosphorylation by small molecules, particularly AcP, has long been suggested to be physiologically relevant in some RRs to couple TCS output to global conditions (115). RR phosphorylation, often in the absence of the cognate HK, can be influenced by cellular AcP levels that are sensitive to the metabolic state of cell. Because the phosphorylation rate by AcP is usually much slower than the rate of phosphotransfer by the cognate HK and often offset by the phosphatase activity of bifunctional HKs or other auxiliary phosphatases (43, 57), the contribution of phosphorylation by AcP to TCS output is typically minimal in wild-type cells. However, given the great diversity of TCSs, AcP can play a significant role in some systems, specifically those with a fast RR autophosphorylation rate and/or a slow phosphatase rate (59, 85).

**HK-mediated Activities**

RR phosphatase activities mediated by HKs or auxiliary phosphatases are believed to function through positioning a water molecule and stimulating the intrinsic RR autophosphatase activity (17). Structure of CheY3 complexed with the phosphatase CheX (86) reveals an amide side chain inserting into the RR active site and forming a hydrogen bond with the attacking water molecule (Figure 3c). Similar positioning of the amide from a Gln or Asn residue has been observed in different phosphatases with distinct structures, such as CheX (86), CheZ (123) and RapH (81), as well as the phosphatase state of DesK, a bifunctional HK from the second largest HK subfamily,
HisKA_3 (107). The Gln-containing sequence motif, DXXXQ, is located immediately after the phospho-accepting His residue in HKs from the HisKA_3 subfamily. The His residue does not appear to be required for RR dephosphorylation despite its close proximity to the phosphatase-essential Gln (52). For the largest HK subfamily, HisKA, the His residue may play a role in RR dephosphorylation (62, 125) but the exact phosphatase mechanism is less clear. An EXXN/T motif similar to CheX-like phosphatases has been identified and the conserved Asn/Thr residue is suggested to be the catalytic residue (52). However, among the available structures of RR complexes with HKs of the HisKA subfamily, the side chain of the Asn/Thr residue is not at a similar position as the Asn/Gln in other phosphatases and is distant from the active site Asp. A dual engagement model has been suggested involving both His and Asn/Thr residues positioning the catalytic water molecule (62). Roles of the two residues may differ for individual HisKA proteins depending on structural details and structures unequivocally capturing the phosphatase state are needed to elucidate the mechanism.

Loss of RR phosphorylation can result from autodephosphorylation, dephosphorylation by the HK and back-transfer to the cognate HK (34, 98), i.e. the reverse reaction of phosphotransfer. The back-transferred histidyl phosphoryl group can be further transferred to ADP (46, 100) or to other RRs that function as phosphate sinks to modulate phosphorylation levels (3, 98, 103). Back-transfer is suggested to accompany an associative transfer mechanism that has a tight RR phosphorylation transition state and appears to be evolved for bi-directional transfer common in CheA or other HPt-containing phosphorelay proteins (107, 122). Unidirectional phosphotransfer with little back-transfer, observed in several canonical HKs, is linked to a dissociative mechanism and the asymmetry of the Mg$^{2+}$ position in an extended transition state (107). The dissociative mechanism can be distinguished by a long distance between the conserved Asp and His residues in
an HK-RR complex structure. It remains to be investigated whether there is any domain preference
or sequence signature that distinguishes the two mechanisms and determines the relative rates of
forward- and back-transfer. Residues affecting the transition state stability, such as non-conserved
residues at D+2, T+1 and T+2 positions, also modulate the HK-catalyzed phosphotransfer rate (53, 
107). Due to multiple phosphorylation and dephosphorylation reactions present simultaneously,
quantitation of individual activities is often complicated by interference from other activities.
Furthermore, RR activities measured in vitro with truncated cytoplasmic fragments of 
transmembrane HKs require careful examination and may differ greatly from their full-length
counterparts in the cellular environment (43).

4. REGULATORY MECHANISMS IN RR SUBFAMILIES

As a phosphorylation-activated switch between inactive and active conformations, the REC domain
mediates effector functions through intramolecular and/or intermolecular interactions. Structures of
individual REC and effector domains usually undergo subtle changes upon phosphorylation, but the
overall structures may vary greatly because of different domain arrangements. A typical RR
regulatory mechanism is exemplified by Staphylococcus aureus VraR (Figure 4a), one of a few RRs
with full-length protein structures available in both phosphorylated (or BeF$_3^-$-bound) and
unphosphorylated states (60). The monomeric unphosphorylated VraR adopts a closed conformation
with extensive contacts between the REC and DBD effector domains, holding the DNA-recognition
helix at a position unfavorable for dimerization on DNA. Phosphorylation results in an extended
conformation with a flexible linker between the two domains and an altered REC surface that
promotes RR dimerization for DNA binding.
Negative regulation in the inactive state and positive regulation in the active state, as shown for VraR, are two common mechanisms mediated by the REC domain (Figure 4b). For positive regulation, the REC domain facilitates the effector domain function, as shown for many DBD-containing RRs in which dimerization of the phosphorylated REC domains is thought to promote DNA binding and transcription regulation. Effector domains of some RRs, such as NtrC1 (58) and CheB (31), are catalytically competent when alone, but are inhibited by the REC domain in the inactive state. The two mechanisms are not exclusive and many RRs use both. Regulatory details for individual RRs show great variations that may have been evolved to adapt each protein to its unique structure and function. For example, in inactive states, interactions between REC and effector domains can differ dramatically (Figure 4b), even within the same subfamily of RRs (42).

Functional sites, such as the DNA-recognition helix for DBDs and the active site for enzymatic effector domains, can be buried within the REC-effector interface or exposed but held at unfavorable positions by tight interactions. Extended conformations with few interdomain contacts have also been observed for a few RRs that are believed to employ a positive regulatory mechanism.

**Inactive and Active Conformational States**

The REC interaction surface that mediates effector domain function is usually distant from the phosphorylation site. A dynamic allosteric mechanism allows conformational changes at the phosphorylation site to propagate to the distal interaction surface. The REC domain samples different allosteric conformations and exists in equilibrium between inactive and active conformations, with phosphorylation shifting the equilibrium. A “Y-T coupling” allosteric mechanism was initially described for CheY (25, 124) and several other RRs (14, 47).

Rearrangement of the conserved T at the phosphorylation site (Figure 3a) is believed to correlate with the rotameric conformation of a conserved Tyr/Phe (Y) residue in the β5 strand. In the active
the aromatic side chain is oriented toward the interior of the REC domain, distinct from the outward position in the inactive state (Figure 5a), resulting in alteration of the α4-β5-α5 face, a surface widely used by many RRs for interdomain interactions. Because of the readily recognizable position of the two residues, they are often used to classify structures as inactive or active, with the caveat that the conformational change involves a broad surface of the REC domain and Y-T coupling is not the only allosteric mechanism (20, 109).

X-ray crystallography has been central to understanding RR regulatory mechanisms although conformations trapped in crystals represent only static snapshots of RR conformational dynamics and can be influenced by experimental conditions, crystal lattice contacts and high protein concentrations used in crystallization. A particular interaction interface or dimerization/oligomerization mode observed in crystal structures, such as domain-swapped dimers observed in several RRs (17, 26, 55), may not be physiologically relevant and requires complementary experiments for validation. Nevertheless, an increasing number of RR structure snapshots start to reveal different states of conformational trajectories and trends of prevalent regulatory strategies in RR subfamilies.

DBD-containing RRs, the largest class of RRs, have the greatest number of X-ray structures available, making it possible to analyze the allosteric conformational features within individual RR subfamilies (Table 1). All structures crystalized with phosphorylated or BeF₃-bound RRs display inward orientations of the side chain at the conserved Y position while conformations of unphosphorylated RRs are diverse. For unphosphorylated RRs in the OmpR and NtrC subfamilies, both outward and inward orientations of Y residues have been observed, with the inward orientation more readily observed for the REC domain alone than for multi-domain proteins (Table 1). This is consistent with NMR studies suggesting that unphosphorylated RRs exist in equilibrium between
inactive and active conformations (110) while interactions with the effector domain shift the
equilibrium to the inactive state (27). The inward orientation of Y is predominant for most NarL
subfamily members irrespective of phosphorylation status, thus is unlikely indicative of the active
state, but rather a result of packing the α4-β5-α5 face with an accessory β6 strand, a structural
feature found in many RRs of the NarL subfamily (20, 60, 83, 106). Therefore, the Y-T coupling
mechanism is not universal in all RRs. Even for NtrC in which orientation switching has been
observed, interconversion of the aromatic side chain has been suggested not to be involved in
allosteric regulation because its rate of conversion is faster than the rate of active/inactive state
conversion (109). RRs appear to be highly plastic for allosteric regulation with diverse mechanisms
matching their sequence and structural features.

The Y and T residues are two of many residues that may participate in allosteric regulation.
As discussed earlier, the phosphorylation site involves an intricate hydrogen-bond network with
several residues from the β1-α1, β3-α3, β4-α4 and β5-α5 loops. It has been recognized that
interactions, such as salt bridges and van der Waals contacts, of residues at D+1, T+1, T+2, K and
other positions can impact loop conformations and propagate the conformational changes to different
areas of the REC domain (5, 14, 20, 44, 70, 106). Conformations with only subsets of active site
residues at proper hydrogen-bonding positions have been discovered for many RRs (37, 95, 106) and
are often referred to as meta-active states. A single RR can have multiple meta-active states with
different combinations of loop conformations, as shown for NtrX (37). Unphosphorylated RRs can
exist in equilibrium between the active state and multiple inactive or meta-active conformations and
there are multiple pathways for transition to the active state (37, 44, 89). NMR relaxation data also
support a model of segmental motions of multiple allosteric residues for activation of CheY, instead
of a strict two-state switching model (70).
Just as phosphorylation shifts the equilibrium for effector domain regulation, the reverse scenario, with output domains affecting phosphorylation, has also been observed (8, 37). It has been suggested that adoption of an active conformation precedes phosphorylation (44, 53).

Phosphorylation can be accelerated by any effect that increases the population of active conformation, such as relieving inhibition from the output domain by truncation, DNA binding in DBD-containing RRs, dimerization at high protein concentrations and HK-RR interaction (8, 9, 27, 28, 37, 66). Surface changes caused by these different interactions can propagate through the connecting βα loops to the phosphorylation site and other regions of the REC domain, increasing the active state population. One extraordinary example is DesR in which the HK-RR interaction can stabilize the active state of DesR and promote dimerization (66). Besides the phosphorylation-activated switch, interaction with the HK also functions as an allosteric switch to increase the population of active conformation for transcription activation. The same principle may be the basis for many phosphorylation-independent regulatory mechanisms.

**Dimerization Modes and Corresponding Conformational Changes**

Phosphorylation-promoted conformational changes are routinely identified by comparing X-ray or NMR structures obtained for unphosphorylated and BeF$_3$-bound RRs. As shown in Figure 5a, for the stand-alone RR CheY, significant differences in positions of backbone atoms occur at the β4-α4, β5-α5 loops and part of the α4-β5-α5 surface where CheY binds the FliM effector protein to regulate flagellar rotation. It is unsurprising that structural elements of the REC domain undergoing the largest rearrangements upon phosphorylation often correlate with protein-protein interaction surfaces, particularly, the dimer interface for DBD-containing RRs. Certain dimerization modes are popular for specific RR subfamilies with corresponding conformational changes.
A dimerization mode using the complete α4-β5-α5 surface is predominant in the OmpR subfamily (Figure 5b). Dimer interaction involves salt bridges between pairs of highly conserved charged residues within the α4-β5-α5 region (104, 105). For the DrrB protein shown in Figure 5b, conformational changes span the entire dimer surface. A Tyr residue is prevalent at the Y position (111) and switching of Tyr orientations has been observed. In several RR (40, 78, 91), the hydroxyl group of Tyr is in close proximity to a polar residue from the DBD (Figure 5b right), suggesting a potential role in effector inhibition. This is unlikely to be a common mechanism given the wide diversity of domain arrangements within the subfamily (8). RR from the OmpR subfamily often recognize tandem DNA sites and two DBDs bind DNA in a head-to-tail manner (15, 48, 64, 76). The translational symmetry of DBDs coupled with the rotational symmetry of the REC domain predicts a flexible linker or different REC-DBD interfaces for individual RR monomers within the RR-DNA complex. Such asymmetry of REC-DBD interfaces has been observed for KdpE and PmrA (65, 76). However, NMR studies suggest that the REC-DBD interaction observed for PmrA in the crystal structure is transient in solution (65). Additionally, alternative dimers involving α1-α5 (6, 71) or other surfaces (13) have also been discovered for unphosphorylated RR but the physiological relevance of these dimers awaits further studies.

The NtrC subfamily of RR displays two major modes of dimerization, with each involving different subsets of the α4-β5-α5 face (Figure 5c). The α4-β5 dimer interface is often seen in phosphorylated RR (Table 1) or unphosphorylated RR with meta-active conformations (32, 37, 84). Another β5-α5 dimer is associated with inactive conformations (36, 84), as indicated by the outward orientation of the Phe residue at the Y position. Dimer interaction centers around the β5-α5 face, sometimes with a slightly tilted α5 helix in contact with both β5 and α4. Formation of the inactive β5-α5 dimer is believed to inhibit the AAA+ ATPase output domain from oligomerization.
and transcription initiation (10, 84). For the LytTR subfamily, structural information is limited and a dimer also involving the α4-β5 subset of the α4-β5-α5 face has been observed for ComE (16).

Diverse dimerization modes have also been observed in the NarL subfamily. An α4-β5 dimer interface and the phosphorylation-dependent orientation of the Phe residue have been reported for the REC domain of FixJ (14). However, most RRs from the NarL subfamily show an α1-α5 dimer interface (Table 1) distinct from the widely used α4-β5-α5 surface (7, 20, 29, 60, 82, 106). The dimer interface mainly involves the α1 helix, the β1-α1 and β5-α5 loops (Figure 5d).

Correspondingly, significant structural rearrangements spanning to the α1 helix have been observed when apo- and BeF₃⁻-bound structures are compared. As discussed earlier, many RRs of the NarL subfamily display a constitutive inward orientation of the residue at the Y position, thus this residue is unlikely to be involved in allosteric regulation. Instead, different switching mechanisms involving the T residue and other residues at α1, the β1-α1, β4-α4 and β5-α5 loops have been proposed for different RRs, such as VraR, DesR and RcsB (20, 60, 106). In addition to the α1-α5 dimer interface, the α4-β5-α5 surface is also remodeled to alter the positioning of the accessory β6 and α6 that are directly linked to the effector domain. Another dimer mode involving β6 and α6 is common (Figure 5d, right) and several RRs, such as VraR (60), DesR (106), Spr1814 (82) and LiaR (29), show both α1-α5 and β6-α6 dimer modes within a single crystal. It has been suggested that the β6-α6 dimer may help RRs to form higher order oligomers on DNA to bind arrays of DNA-binding sites (106).

DNA-bound structures of RcsB again indicate an asymmetric dimer with different rotation symmetry axes for the REC and DBDs (38). Different relative positions of the REC and DBDs are frequently observed for different full length RRs. Asymmetry is believed to result from either different REC-DBD interactions (20, 38) or a flexible linker that enables different disposition or even domain-swap of individual domains trapped in crystals (29, 60).
5. RESPONSE REGULATOR – HISTIDINE KINASE INTERACTIONS

**HK Conformational States**

HKs have a modular architecture with a large variety of sensory domains (extracellular, transmembrane or cytoplasmic) linked to a conserved catalytic domain by one or more signal transducing domains (12, 69) (Figure 1a). The cytoplasmic enzymatic core of HKs consists of a dimerization histidine phosphotransfer (DHp) domain and a catalytic ATP-binding (CA) domain. The DHp domain contains the conserved phosphorylatable His residue and mediates dimerization of the HK to form a 4-helix bundle that is essential for its activity (Figure 6). The DHp dimer is flanked by the two α/β catalytic CA domains containing the kinase active sites that catalyze phosphoryl transfer from ATP to the His residues. Autophosphorylation of the HK provides a high-energy phosphoryl group for subsequent phosphotransfer to the RR. Most HKs also mediate dephosphorylation of their cognate RRs in order to modulate the output response.

Large conformational changes in HKs occur during transitions between kinase, phosphotransfer and phosphatase states. In the HisKA_3 subfamily of HKs (e.g. DesK of *Bacillus subtilis*), the switch between states involves a large rotation of the DHp domains, a mechanism that is less pronounced in the HisKA subfamily. Specific regulatory mechanisms have been described for individual HKs, including the stabilization of the phosphatase state by ligand binding (e.g. c-di-GMP binding on CckA) and the inhibition of the phosphatase state by a pH-dependent conformational switch (33, 62). In contrast to the different conformations of HKs, structures of HK-REC domain complexes determined to date indicate that REC domains bound to either phosphotransfer or phosphatase states of HKs have similar conformations, with the REC domain displaying a meta-active conformation, as defined in the previous section.
Partner Recognition and Specificity

Specificity of HK-RR pairs is crucial for the integrity of signaling pathways, given the large number of different TCSs that typically exist in a single cell. Specificity within a pair is mediated by recognition residues that participate in HK-RR binding. These residues are located in helix $\alpha_1$ of the REC domain and in the two $\alpha$ helices of the HK DHp domain that contribute to the binding interface.

Signaling pathways have been successfully rewired by substituting as few as three co-evolved specificity residues (88, 96). Additional strategies are employed to further ensure specificity of the pairs. For example, the phosphatase activity of HKs eliminates non-specific phosphorylation by non-cognate kinases or small molecule phosphodonors such as AcP. The low-abundance of HKs relative to RRss also minimizes cross-phosphorylation between non-cognate pairs (57).

In an HK-RR complex, the REC domain inserts its $\alpha_1$ helix and $\beta_5-\alpha_5$ loop between the two helices of the DHp domain, primarily contacting the DHp $\alpha_1$ helix that contains the phosphorylatable His residue. This conserved mode of binding buries the active site of the REC domain at the HK-RR interface (Figure 6). Thus, conserved residues of both the HK and RR form a substantial subset of the interface residues, with specificity being determined by a relatively small set of variable residues. In addition to these specificity residues in $\alpha_1$ of the REC domain and the two helices of the DHp domain, additional variable contacts can involve the $\beta_2-\alpha_2$, $\beta_3-\alpha_3$ and $\beta_4-\alpha_4$ loops of the REC domain interacting with the CA domain, the C-terminal region of the DHp and the DHp-CA linker of the other protomer of the dimer (Figure 6) (21, 72, 74, 88, 107, 108, 114, 118, 122). Other contacts can occur between the REC domain and additional domains within the HK, increasing the specificity and/or stability of the pair (e.g. contacts between the PAS domain of ThkA and the RR TrrA) (118). These elements suggest that beyond the conserved REC-DHp interface that buries the active site, the binding interface can vary greatly in different HK-RR pairs.
A model based on structures of HK FixL and full-length RR FixJ suggests that the RR effector domain is free to move and does not participate in binding to the HK (116). However, a lack of structures of full-length RRs bound to HKs leaves open the question as to how effector domains might be positioned relative to REC domains and domains of the HK. Given the great diversity in REC-effector domain arrangements in different inactive RRs, it is possible that diversity exists in the ways effector domains interact in different HK-RR complexes. It is easy to envision how an effector domain could participate in the stability of the complex and/or the efficiency of catalysis. Indeed, the meta-active conformation of the REC domain observed in HK-REC domain complexes might be a consequence of isolated REC domains being unhindered by REC-effector interactions that could bias conformational states.

**Phosphotransfer State**

Structural and biochemical studies of DesK-DesR complexes suggest that one HK dimer binds to one RR molecule in an asymmetric conformation. The DesK-DesR structure and a docking model of CpxA-CpxR in a putative phosphotransfer state show asymmetry within the HK dimer, with the CA domain of one monomer bound to the DHp domain of the other monomer, leaving the second CA domain untethered with enough space for RR binding to the DHp. Structural data from the DesK-DesR complex and from CpxA-CpxR model suggest that this highly dynamic state is coupled with autokinase activity, involves back-and-forth movements of the DHp and CA domains, and is modulated by REC domain binding (72, 107).

**Phosphatase State**

In contrast to the asymmetrical domain arrangements associated with phosphotransfer, HK-RR complexes in putative dephosphorylation states have symmetrical arrangements of the HK DHp and CA domains. The REC domains are also arranged symmetrically, each interacting with a DHp
domain with a 2:2 HK:RR stoichiometry (62, 107). The HK in its phosphatase state is more rigid
than in the phosphotransfer state, with the N-terminal coiled-coil of the DHp domains being highly
stable.

Despite the large number of structures recently determined, the mechanistic details of the
enzymatic reactions involving HKs and RRs are not fully understood. Discrimination between the
phosphotransfer and phosphatase states that have been trapped in crystal structures still remains
challenging. The situation is further complicated by distinct phosphatase mechanisms used by HKs
of different subtypes. A major limitation of current studies is the use of truncated proteins, most
notably the cytoplasmic domains of transmembrane HKs lacking sensor and transmembrane regions
that in intact proteins control the signaling states of HKs, and REC domains of RRs in the absence of
effector domains that are known to influence their conformational equilibria. While conserved
features of HK-RR interactions are beginning to be elucidated, variations on the theme are
anticipated. Similar to many other aspects of TCS structure and function, it is likely that the great
variety of domain architectures in TCS proteins enable different modes of HK-RR interactions and
regulatory mechanisms that are adapted to individual HK-RR pairs.

6. NON-CANONICAL MODES OF RR REGULATION

The canonical mechanism for regulation of RR activity involves phosphorylation at a conserved Asp
in the REC domain, stabilizing an active conformation that enables effector domain function.
Regulation of RR activity, in addition to or in place of Asp phosphorylation, potentially can be
achieved in many different ways including post-translational modifications at other sites that bias the
conformational equilibrium of the REC domain or directly alter effector domain function,
interactions that lower the energetic barrier for transition to an active conformation, ligand binding
to the RR, and regulation of expression to control RR levels and thus dimerization or oligomerization. Indeed, each of these strategies has been observed and a few examples are described below.

Reversible protein acetylation is known to regulate metabolic enzymes and transcription in bacteria (49) and was documented in CheY more than 20 years ago (90). Enzyme catalyzed acetylation of CheY at K91 and K106 increases in response to acetate and promotes clockwise flagellar rotation, with acetylation at K91 proposed to alter conformational dynamics of the β4α4 loop (39). Autoacetylation with AcCoA as acetyl donor at other sites in CheY is thought to link chemotaxis to the metabolic state of the cell (120). Another well-characterized example of acetylation occurs in E. coli transcription factor RcsB. Acetylation of K154, a residue in the helix-turn-helix motif that interacts with a phosphate in the DNA backbone, disrupts DNA binding, down-regulates expression of the acid stress response genes, and compromises survival in acidic conditions (22, 51). It should be noted that the stoichiometry of acetylation has not been determined, thus physiological effects observed with mutations that mimic fully unmodified or acetylated states might overestimate regulation that occurs in wild-type cells. A proteomics study in E. coli documented acetylation of seven RRs: ArcA, CheY, CpxR, EvgA, NarL, OmpR and RcsB (22). The E. coli genome encodes 26 putative Gcn5 N-acetyltransferases (GNATs) and a single known deacetylase (CobB) (49). While the best-characterized acetyltransferase, PatZ, is proposed to mediate a global strategy linking regulation to metabolic state, the large number of acetyltransferases raises the possibility of substrate specificity that might enable system-specific signaling mechanisms.

Signal-induced, HK-dependent, phosphorylation-independent monomer to dimer activation has been observed for several RR transcription factors. In the extensively studied B. subtilis system, binding of DesR to DesK promotes an active conformation of DesR. At sufficiently high RR
concentrations and with a slow active to inactive state transition, dimerization can occur upon release of the RR from the HK with subsequent stabilization of the dimer upon DNA binding (106). Similar phosphorylation-independent activation of OmpR by HK EnvZ has been observed in acidic conditions (23). Cyanobacterial transcription factor NblR contains a conserved Asp, but lacks other residues necessary for phosphorylation. HK NblS is required for phosphorylation-independent activation of NblR. However, no NblS-NblR interactions have been detected and the activating monomer to dimer transition is postulated to be promoted by another protein partner (93).

Multiple strategies have been identified for regulation of orphan RRs that lack a conserved Asp and/or other conserved residues necessary for phosphorylation. In streptomycetes, two atypical RR transcription factors that lack residues necessary for phosphorylation, JadR1 and RedZ, are regulated by the end products of the antibiotic biosynthetic pathways they control. The antibiotic JdB binds directly to the JadR1 REC domain, disrupting DNA binding (112). A different strategy is used by Helicobacter pylori HP1043, which exists as a constitutively active dimer in vitro with a crystal structure similar to that of other activated OmpR subfamily members (50). Levels of HP1043 are regulated both transcriptionally and post-transcriptionally leading to speculation that control of expression of this constitutively active RR may be the sole mechanism for regulating HP1403 activity (75).

Combinations of these mechanisms create even more strategies. Streptomyces coelicolor GlnR, which regulates genes for nitrogen assimilation, is an orphan OmpR subfamily RR that lacks residues for Asp phosphorylation and forms a constitutive α4-β5-α5 dimer (61). GlnR is phosphorylated at 6 Ser/Thr sites in the DNA-binding domain under N-rich conditions, disrupting DNA binding. GlnR is also acetylated at multiple Lys residues in the DNA-binding domain, with acetylation enhancing DNA binding (4).
7. CONCLUDING REMARKS

When the structure of an RR REC domain was first reported thirty years ago, a central question was how this single conserved domain could regulate responses as diverse as flagellar rotation, transcription, and enzyme activity. The answer that emerged defined a mechanism that was both simple and versatile. The small α/β REC domain exists in equilibrium between two primary conformations with phosphorylation stabilizing an active conformation. This phosphorylation-regulated switch enables regulatory strategies via any type of activating or inhibitory macromolecular interactions that discriminate between the two states. Hundreds of structures of RRs have provided descriptions of the conformations of REC domains in inactive and active states, interactions with effector domains, DNA, HKs and auxiliary proteins. These structures provide a foundation for identifying conserved features as well as specific variations in individual RRs. Beyond the universally conserved enzymatic mechanisms facilitated by configurations at the active site, other features such as Y-T coupling, regions of conformational perturbations, domain arrangements and modes of DNA binding show distinct trends among RRs within specific subfamilies.

Numerous variations on most every feature of RRs have been observed and undoubtedly, many more remain to be discovered. The plasticity of the REC domain and versatility of RR design allow an almost unlimited array of adaptations to fit the needs of individual signaling systems. Defining details of how specific structural features impact function is important for interpreting the nuances of RR behavior within specific TCSs as well as for pursuit of applied projects such as development of antimicrobial therapeutics or the engineering of synthetic biosensing pathways.
Our current understanding of RR structure is largely informed by compiling information derived from many partial structural descriptions of different RRs. Very few structures are available for full-length multi-domain RRs in both inactive and active states. The crystal structures that do exist are constrained by the limitations of the methodology, specifically the capture of single discrete states that do not reflect conformational distributions and the potential promotion of inter- and/or intra-molecular interactions that are influenced by the high concentrations of proteins used in crystallization and/or stabilization of conformations necessary for crystallization and lattice interactions. While NMR studies have provided information about dynamics and conformational distributions in solution, size limitations have mostly precluded studies of full-length RRs, dimers, oligomers and complexes of RRs with HKs or other macromolecular partners. Thus, while structures unambiguously define allowable states, it is important to keep in mind the conformationally dynamic nature of RRs and the potential influence of associated domains and macromolecular partners when interpreting structures determined by methodologies with technical limitations. The emergence of new structural methods such as high-resolution cryo-electron microscopy promise to provide solutions to some of these challenges.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS
Research in the Stock laboratory is supported by grant R01 GM47958 from the National Institutes of Health. S.B. was supported in part by the Foundation pour la Recherche Médicale (FRM) Grant SPE201803005070.


70. McDonald LR, Boyer JA, Lee AL. 2012. Segmental motions, not a two-state concerted switch, underlie allostery in CheY. *Structure* 20:1363-73


84. Park S, Meyer M, Jones AD, Yennawar HP, Yennawar NH, Nixon BT. 2002. Two-component signaling in the AAA+ ATPase DctD: binding Mg^{2+} and BeF_{3} selects between alternate dimeric states of the receiver domain. FASEB J. 16:1964-66


TERMS AND DEFINITIONS

AcP  acetyl phosphate
BeF$_3$  beryllofluoride
CA  catalytic/ATP-binding (domain)
DBD  DNA-binding domain
DHp  dimerization/histidine phosphotransfer (domain)
ECF  extracytoplasmic function
HK  histidine kinase
HPt  histidine-containing phosphotransfer (domain)
PDB  Protein Data Bank
REC  receiver (domain of response regulator protein)
RR  response regulator protein
SDRR  single-domain response regulator
TCS  two-component system
FIGURE CAPTIONS

Figure 1. Common features of RRs. (a) Schematic diagram of the prototypical TCS pathway. (b) The conserved (βα)₅ fold of REC domains. Phosphorylation site residues and the phosphoryl group mimic, beryllofluoride (BeF₃⁻), are shown in sticks. (c) Sequence conservation of RECs. The profiled hidden Markov model (HMM) for the REC protein family (Pfam PF00072) is shown as sequence logos (113) with the secondary structure elements illustrated. Heights of individual stacked letters at each position correspond to information contents, reflecting the probability of observing the particular amino acids at each position. Phosphorylation site residues (stars), named after the most conserved amino acids (DD, D, T and K), are among the most conserved residues in REC domains.

Figure 2. Classification of RRs by their effector domains. The percentile distribution is indicated for RR effector functional classes and RR subfamilies. RR subfamilies are defined by effector domain folds identified in Pfam. Representative proteins are traditionally used to name the OmpR, NarL and NtrC subfamilies with effector domains named Trans_reg_C, GerE and Sigma54_activa in Pfam. Representative structures of each subfamily are shown with REC domains colored in grey and effector domains in colors (PDB ids: KdpE, 4KNY; RcsB, 5W43; LuxO, 5EP0; AgrA, 3BS1; AmiR, 1Q00; WspR, 3BRE; CheB, 1A2O).

Figure 3. Phosphorylation site of the REC domain. The phosphoryl group is positioned by a network of hydrogen bonds (dashed lines) with side chains of the highly conserved residues (orange) as well as backbone atoms of non-conserved residues (light pink). Non-conserved active site residues are labeled by their relative sequence positions to the nearest conserved residues, such as T+1, indicating
one residue C-terminal to the conserved Thr/Ser residue. (a) Phosphorylation site of the archetype RR CheY in the active conformation (PDB id: 1FQW) with residues that are differently positioned in the inactive conformation (2CHE) shown in cyan. (b) The putative trigonal bipyramidal transition state for both phosphorylation and dephosphorylation. X represents the leaving group of the phosphodonor for phosphorylation or the attacking water for dephosphorylation. (c) Surface view of the active site in the phosphatase-REC complex (3HZH).

Figure 4. RR regulatory strategies. (a) Distinct inactive and active RR conformations exemplified by full-length RR VraR (PDB ids: 4GVP, 4IF4). BeF$_3^-$ is shown in red spheres. (b) Schematic diagrams of RR regulatory mechanisms. Functional sites of effector domains, such as enzyme sites or DNA recognition regions, are shown as pink dots. These sites can be buried or exposed in a wide variety of inactive RR conformations and their activity depends on different interactions between the REC and effector domains. Phosphorylation of the RR can relieve REC domain inhibition, promote effector function or both. Representative RRs that utilize these strategies are indicated.

Figure 5. Phosphorylation-induced conformational changes in the REC domain. RR structures with or without BeF$_3^-$ were aligned using the conserved strands $\beta$1, $\beta$3, $\beta$4, and $\beta$5 to compute the average backbone RMSD per residue (a, right). RMSD values above the median + 2x MAD (median absolute deviation) are considered as significant conformational changes and the corresponding residues are colored blue. Non-REC structural elements are colored cyan. BeF$_3^-$ is shown as red spheres and residues involved in the potential Y-T coupling are shown as green or gray sticks. Representative protein structures shown for the (a) Stand-alone, (b) OmpR, (c) NtrC and (d) NarL
RR subfamilies are CheY (1F4V), DrrB (3NNS, 1P2F), NtrC1 (1ZY2, 1NY5) and VraR (4IF4). Y/T residues from the inactive CheY structure (2CHE) are differently positioned from the active structure. For the OmpR, NtrC and NarL subfamilies, two proteins from each subfamily were used for RMSD analyses and both showed similar regions of conformational changes. Substantial changes in α1 are also observed in one protein from the NtrC subfamily (pink).

Figure 6. Structure of the HK-RR complex. Ribbon (a) and surface (b) views of the HK856-RR468 complex (PDB id, 3DGE). Residues that determine the HK-RR interaction specificity (19, 96) are highlighted in light orange and cyan in the HK and RR, respectively. HK-RR contacts also involve other surface regions (grey), including both active sites of the HK (red) and the RR (pink).
Abstract

Response regulators function as the output components of two-component systems, which couple the sensing of environmental stimuli to adaptive responses. Response regulators typically contain conserved receiver (REC) domains that function as phosphorylation-regulated switches to control the activities of effector domains that elicit output responses. This modular design is extremely versatile, enabling different regulatory strategies tuned to the needs of individual signaling systems. This review summarizes functional features that underlie response regulator function. An abundance of atomic resolution structures and complementary biochemical data have defined the mechanisms for response regulator enzymatic activities, revealed trends in regulatory strategies utilized by response regulators of different subfamilies and provided insights into interactions of response regulators with their cognate histidine kinases. Among the hundreds of thousands of response regulators identified, variations abound. This article provides a framework for understanding structural features that enable function of canonical response regulators and a basis for distinguishing non-canonical configurations.
1. INTRODUCTION

Two-component systems (TCSs) are the predominant multi-step signaling pathways in bacteria. Two conserved proteins, a histidine protein kinase (HK) and response regulator (RR), constitute the core system (Figure 1a). Autophosphorylation of the HK at a conserved His provides a high-energy phosphoryl group that is transferred to a conserved Asp in the RR, resulting in RR activation (99). Stimuli, sensed either directly or indirectly by the HK, regulate opposing autophosphorylation/phosphotransfer and RR phosphatase activities of the HK, thus determining the level of phosphorylation of the RR and the output response. A large variety of input domains in HKs and output domains in RRs allows the coupling of an almost limitless array of chemical or physical stimuli to diverse output responses. The conserved proteins and the systems themselves are extremely versatile. An enormous range of variations in protein activities, domain architectures and system configurations allows adaptation of TCSs to the needs of specific signaling systems.

TCSs are found in bacteria, archaea (41) and eukaryotes such as slime molds, yeast, and plants (2). While great diversity exists among all TCSs it should be noted that some specific trends in system configurations are found in different organisms. The Pfam database (35) lists >342,000 entries of proteins containing the conserved receiver (REC, named Response_reg in Pfam) domain that characterizes RRs and >650 structures of such proteins in the Protein Data Bank (PDB) (18). In addition to their presence in RRs, REC domains are also found in hybrid HKs that function within phospho-relay systems involving multiple phosphotransfer steps, accounting for ~10% of REC domain-containing proteins. The scope of this review will focus on bacterial RRs that mediate output of TCSs. Among these RRs, variations abound. The plasticity of the REC domain and the versatility of TCS architecture has driven the evolution of RRs that
are uniquely adapted and fine-tuned for function in specific pathways. Out of necessity, this review will focus on canonical RRs with full understanding, but lack of specific acknowledgement of individual cases, that for every feature noted, exceptions exist.

This review provides an overview of the structure of RRs and features that provide conserved enzymatic activities in REC domains. Beyond the core domain structures and common enzymatic mechanisms, RRs display great variation in domain arrangements that provide a variety of mechanistically distinct regulatory mechanisms. However, as structures have accumulated, trends in regulatory strategies used by different subfamilies of RRs have begun to emerge and these will be explored in this review.

2. RESPONSE REGULATOR ARCHITECTURE

REC Domains

RRs are defined by the presence of a conserved REC domain. The REC domain fold is composed of five α helices surrounding a central five-stranded parallel β-sheet with a 21345 topology (Figure 1b). Sequence identity between REC domains is usually 20-30%, showing great variations (Figure 1c). The central hydrophobic β-strands, β1, β3, β4 and β5, are more conserved than the peripheral helices and loops. The most conserved residues include the site of phosphorylation, an Asp at the C-terminus of β3, and several other residues in the β-α loops, which constitute the active site. The highly conserved phosphorylation site and the variable peripheral sequences allow RRs to couple phosphorylation to diverse effector functions.

Effector Domains

RRs typically function as the output components of signaling pathways. Regulatory REC domains can be linked either covalently or non-covalently to a great diversity of effectors and
thus control numerous diverse responses. Figure 2 summarizes the distribution of the major RR effector domains identified in the Pfam database. Approximately one fourth of RRs (23%) consist solely of a REC domain. These single-domain RRs (SDRRs) regulate intermolecular effectors such as the chemotaxis system regulator CheY, which binds to the flagellar component FliM controlling flagellar rotation. CheY-like proteins found in chemotaxis-like or chemosensory systems mediate responses other than chemotaxis, for example biofilm formation regulated by Cle proteins (77). Other functions for SDRRs include acting as intermediates in phosphorelays (e.g. the general stress response proteins SdrG and MrrA, and the sporulation protein Spo0F) (45, 63), as phosphate sinks to control phosphate flux within phosphorelays (e.g. *Rhizobium meliloti* CheY1) (97), as allosteric regulators of HKs (e.g. *Caulobacter crescentus* DivK) (24, 117) or as protease adapters (e.g. *C. crescentus* CpdR) (56).

The largest class of RRs (64.5%) is composed of RRs containing a DNA-binding domain (DBD) with subfamilies defined by different DBD folds. The OmpR subfamily (29% of RRs) contains a winged-helix effector domain (68), the NarL subfamily (19%) a four-helix DNA-binding HTH domain (73), the NtrC subfamily (7%) a Fis-type HTH domain fused to an AAA+ ATPase domain and the LytTR (5.5%) a predominantly β fold (94). The abundance of RRs in this class likely reflects the importance of transcriptional regulation as a response to environmental change. A small class of RRs (1%) harbors an RNA-binding domain belonging to the ANTAR subfamily of anti-termination factors such as AmiR (79).

RRs with enzymatic domains account for ~8% of RRs. A major group within this class (3% of RRs) is the regulators of cyclic di-GMP including cyclases (GGDEF) and phosphodiesterases (EAL, HD-GYP) (92). A second major group (2%) is the chemotaxis methylesterase CheB proteins (31). A variety of different enzymatic domains have been identified in other RRs (3%)
such as the PP2C phosphatase domain of RsbY and the hybrid kinase domain of FrzE (30, 54).

RRs with protein binding domains account for ~1% of RRs. Examples in this class include
CheV, with a CheW-like domain that connects chemoreceptors to the chemotaxis histidine
kinase CheA; RssB, which regulates turnover of the stress response sigma factor RpoS; and
PhyR, a regulator of the general stress response that contains an extracytoplasmic function (ECF)
sigma factor-like domain (1, 11, 66). While RRs often contain a simple REC-effector domain
architecture they can have complex domain organizations with additional signaling domains
including PAS, GAF, HisKa, HATPase, etc. The large variety of RRs identified to date
emphasizes the versatility of the REC domain with no apparent limits on the types of effector
domains that can be controlled by this phosphorylation-activated switch.

3. ENZYMATIC ACTIVITIES OF RRs

Conserved Phosphorylation Site

One of the defining characteristics of the REC domain is its highly conserved phosphorylation
site. Due to the lability of the high-energy acyl phosphate, structural characterization of the
active site often relies on beryllofluoride (BeF$_3^-$) that noncovalently binds to the phosphorylation
site Asp and serves as a mimic of the phosphoryl group (119). Different BeF$_3^-$-bound RR
structures reveal a conserved active site with a network of hydrogen bonds (Figure 3a). The
carboxylate side chains of the acidic residue duo (DD) at the $\beta$1-$\alpha$1 loop participate in
coordinating a Mg$^{2+}$ required for catalysis. Sidechains of two additional residues, a Thr/Ser (T) at
the $\beta$4-$\alpha$4 loop and a Lys (K) at the $\beta$5-$\alpha$5 loop, together with backbone atoms of non-conserved
active site residues, coordinate phosphate oxygens in the phosphorylated REC domain.
RR phosphorylation level, the ultimate determining factor of TCS output for canonical RRs, is regulated by multiple enzyme activities, including phosphotransfer from HKs, dephosphorylation by auxiliary phosphatases or bifunctional HKs that also possess RR phosphatase activity, autophosphorylation by small-molecule phosphodonors such as phosphoramidate and acetyl phosphate (AcP) (67, 115), and autodephosphorylation activities. Similar phosphorylation sites with almost identical positioning of active site residues have been observed in numerous RR or RR complex structures (21, 81, 86, 107). A pentavalent phosphorus intermediate is believed to be a common transition state for different activities of the REC domain (Figure 3b). Both phosphorylation and dephosphorylation reactions can proceed through either an associative or a dissociative mechanism depending on how tight or extended the transition state is. A phosphodonor or water molecule needs to be in line with the acyl phosphate bond that is to be formed or broken. The phosphotransferase or phosphatase helps position these molecules to further enhance the reaction rates. In all cases, the majority of residues involved in coordinating the phosphorus intermediate are from the REC domain. Thus, most RRs are catalytically competent of autophosphorylation and autodephosphorylation in the absence of any enzymatic protein partner and in vitro analyses of these reactions often provide insights into RR regulatory mechanisms.

**Autodephosphorylation and Autophosphorylation**

Despite the highly conserved active site geometry, RRs show large variations in rates of autophosphorylation and autodephosphorylation (87, 101, 102). For example, autodephosphorylation rate constants of the REC domain range over six orders of magnitude, giving phosphorylation half-lives of seconds to hours and even days (101). Such great diversity is partly attributed to several variable residues surrounding the active site, such as positions D+2
(two residues C-terminal to the conserved D), T+1 and T+2 (one and two residues C-terminal to T) (Figure 3b and 3c). Backbone atoms of D+2 and T+1 directly form hydrogen bonds with the phosphoryl group while the charge, size and hydrophobicity of sidechains at these three positions may affect the energy barrier of the transition state, facilitate the positioning or block the in-line path of the phosphodonor or attacking water (53, 80, 102, 121). In more than half of RR sequences, a limited number of amino acid combinations are preferred at these positions and the distribution of preferred amino acids correlates with RR effector subfamilies (53, 80). REC enzyme activities are therefore suggested to co-evolve with effector domain regulation and these residues represent functional sites for modulating the stability of RR phosphorylation to match the timescale of individual TCS output responses.

Phosphorylation by small molecules, particularly AcP, has long been suggested to be physiologically relevant in some RR s to couple TCS output to global conditions (115). RR phosphorylation, often in the absence of the cognate HK, can be influenced by cellular AcP levels that are sensitive to the metabolic state of cell. Because the phosphorylation rate by AcP is usually much slower than the rate of phosphotransfer by the cognate HK and often offset by the phosphatase activity of bifunctional HKs or other auxiliary phosphatases (43, 57), the contribution of phosphorylation by AcP to TCS output is typically minimal in wild-type cells. However, given the great diversity of TCSs, AcP can play a significant role in some systems, specifically those with a fast RR autophosphorylation rate and/or a slow phosphatase rate (59, 85).

HK-mediated Activities

RR phosphatase activities mediated by HKs or auxiliary phosphatases are believed to function through positioning a water molecule and stimulating the intrinsic RR autophosphatase activity
Structure of CheY3 complexed with the phosphatase CheX (86) reveals an amide side chain inserting into the RR active site and forming a hydrogen bond with the attacking water molecule (Figure 3c). Similar positioning of the amide from a Gln or Asn residue has been observed in different phosphatases with distinct structures, such as CheX (86), CheZ (123) and RapH (81), as well as the phosphatase state of DesK, a bifunctional HK from the second largest HK subfamily, HisKA_3 (107). The Gln-containing sequence motif, DXXXQ, is located immediately after the phospho-accepting His residue in HKs from the HisKA_3 subfamily. The His residue does not appear to be required for RR dephosphorylation despite its close proximity to the phosphatase-essential Gln (52). For the largest HK subfamily, HisKA, the His residue may play a role in RR dephosphorylation (62, 125) but the exact phosphatase mechanism is less clear. An EXXN/T motif similar to CheX-like phosphatases has been identified and the conserved Asn/Thr residue is suggested to be the catalytic residue (52). However, among the available structures of RR complexes with HKs of the HisKA subfamily, the side chain of the Asn/Thr residue is not at a similar position as the Asn/Gln in other phosphatases and is distant from the active site Asp. A dual engagement model has been suggested involving both His and Asn/Thr residues positioning the catalytic water molecule (62). Roles of the two residues may differ for individual HisKA proteins depending on structural details and structures unequivocally capturing the phosphatase state are needed to elucidate the mechanism.

Loss of RR phosphorylation can result from autodephosphorylation, dephosphorylation by the HK and back-transfer to the cognate HK (34, 98), i.e. the reverse reaction of phosphotransfer. The back-transferred histidyl phosphoryl group can be further transferred to ADP (46, 100) or to other RRs that function as phosphate sinks to modulate phosphorylation levels (3, 98, 103). Back-transfer is suggested to accompany an associative transfer mechanism.
that has a tight RR phosphorylation transition state and appears to be evolved for bi-directional
transfer common in CheA or other HPt-containing phosphorelay proteins (107, 122).
Unidirectional phosphotransfer with little back-transfer, observed in several canonical HKs, is
linked to a dissociative mechanism and the asymmetry of the Mg$^{2+}$ position in an extended
transition state (107). The dissociative mechanism can be distinguished by a long distance
between the conserved Asp and His residues in an HK-RR complex structure. It remains to be
investigated whether there is any domain preference or sequence signature that distinguishes the
two mechanisms and determines the relative rates of forward- and back-transfer. Residues
affecting the transition state stability, such as non-conserved residues at D+2, T+1 and T+2
positions, also modulate the HK-catalyzed phosphotransfer rate (53, 107). Due to multiple
phosphorylation and dephosphorylation reactions present simultaneously, quantitation of
individual activities is often complicated by interference from other activities. Furthermore, RR
activities measured in vitro with truncated cytoplasmic fragments of transmembrane HKs require
careful examination and may differ greatly from their full-length counterparts in the cellular
environment (43).

4. REGULATORY MECHANISMS IN RR SUBFAMILIES

As a phosphorylation-activated switch between inactive and active conformations, the REC
domain mediates effector functions through intramolecular and/or intermolecular interactions.
Structures of individual REC and effector domains usually undergo subtle changes upon
phosphorylation, but the overall structures may vary greatly because of different domain
arrangements. A typical RR regulatory mechanism is exemplified by Staphylococcus aureus
VraR (Figure 4a), one of a few RRs with full-length protein structures available in both
phosphorylated (or BeF$_3$-bound) and unphosphorylated states (60). The monomeric unphosphorylated VraR adopts a closed conformation with extensive contacts between the REC and DBD effector domains, holding the DNA-recognition helix at a position unfavorable for dimerization on DNA. Phosphorylation results in an extended conformation with a flexible linker between the two domains and an altered REC surface that promotes RR dimerization for DNA binding.

Negative regulation in the inactive state and positive regulation in the active state, as shown for VraR, are two common mechanisms mediated by the REC domain (Figure 4b). For positive regulation, the REC domain facilitates the effector domain function, as shown for many DBD-containing RRs in which dimerization of the phosphorylated REC domains is thought to promote DNA binding and transcription regulation. Effector domains of some RRs, such as NtrC1 (58) and CheB (31), are catalytically competent when alone, but are inhibited by the REC domain in the inactive state. The two mechanisms are not exclusive and many RRs use both.

Regulatory details for individual RRs show great variations that may have been evolved to adapt each protein to its unique structure and function. For example, in inactive states, interactions between REC and effector domains can differ dramatically (Figure 4b), even within the same subfamily of RRs (42). Functional sites, such as the DNA-recognition helix for DBDs and the active site for enzymatic effector domains, can be buried within the REC-effector interface or exposed but held at unfavorable positions by tight interactions. Extended conformations with few interdomain contacts have also been observed for a few RRs that are believed to employ a positive regulatory mechanism.

**Inactive and Active Conformational States**
The REC interaction surface that mediates effector domain function is usually distant from the phosphorylation site. A dynamic allostERIC mechanism allows conformational changes at the phosphorylation site to propagate to the distal interaction surface. The REC domain samples different allostERIC conformations and exists in equilibrium between inactive and active conformations, with phosphorylation shifting the equilibrium. A “Y-T coupling” allostERIC mechanism was initially described for CheY (25, 124) and several other RRs (14, 47).

Rearrangement of the conserved T at the phosphorylation site (Figure 3a) is believed to correlate with the rotameric conformation of a conserved Tyr/Phe (Y) residue in the β5 strand. In the active state, the aromatic side chain is oriented toward the interior of the REC domain, distinct from the outward position in the inactive state (Figure 5a), resulting in alteration of the α4-β5-α5 face, a surface widely used by many RRs for interdomain interactions. Because of the readily recognizable position of the two residues, they are often used to classify structures as inactive or active, with the caveat that the conformational change involves a broad surface of the REC domain and Y-T coupling is not the only allostERIC mechanism (20, 109).

X-ray crystallography has been central to understanding RR regulatory mechanisms although conformations trapped in crystals represent only static snapshots of RR conformational dynamics and can be influenced by experimental conditions, crystal lattice contacts and high protein concentrations used in crystallization. A particular interaction interface or dimerization/oligomerization mode observed in crystal structures, such as domain-swapped dimers observed in several RRs (17, 26, 55), may not be physiologically relevant and requires complementary experiments for validation. Nevertheless, an increasing number of RR structure snapshots start to reveal different states of conformational trajectories and trends of prevalent regulatory strategies in RR subfamilies.
DBD-containing RRs, the largest class of RRs, have the greatest number of X-ray structures available, making it possible to analyze the allosteric conformational features within individual RR subfamilies (Table 1). All structures crystalized with phosphorylated or BeF$_3$-bound RRs display inward orientations of the side chain at the conserved Y position while conformations of unphosphorylated RRs are diverse. For unphosphorylated RRs in the OmpR and NtrC subfamilies, both outward and inward orientations of Y residues have been observed, with the inward orientation more readily observed for the REC domain alone than for multi-domain proteins (Table 1). This is consistent with NMR studies suggesting that unphosphorylated RRs exist in equilibrium between inactive and active conformations (110) while interactions with the effector domain shift the equilibrium to the inactive state (27). The inward orientation of Y is predominant for most NarL subfamily members irrespective of phosphorylation status, thus is unlikely indicative of the active state, but rather a result of packing the $\alpha 4$-$\beta 5$-$\alpha 5$ face with an accessory $\beta 6$ strand, a structural feature found in many RRs of the NarL subfamily (20, 60, 83, 106). Therefore, the Y-T coupling mechanism is not universal in all RRs. Even for NtrC in which orientation switching has been observed, interconversion of the aromatic side chain has been suggested not to be involved in allosteric regulation because its rate of conversion is faster than the rate of active/inactive state conversion (109). RRs appear to be highly plastic for allosteric regulation with diverse mechanisms matching their sequence and structural features.

The Y and T residues are two of many residues that may participate in allosteric regulation. As discussed earlier, the phosphorylation site involves an intricate hydrogen-bond network with several residues from the $\beta 1$-$\alpha 1$, $\beta 3$-$\alpha 3$, $\beta 4$-$\alpha 4$ and $\beta 5$-$\alpha 5$ loops. It has been recognized that interactions, such as salt bridges and van der Waals contacts, of residues at D+1,
T+1, T+2, K and other positions can impact loop conformations and propagate the conformational changes to different areas of the REC domain (5, 14, 20, 44, 70, 106).

Conformations with only subsets of active site residues at proper hydrogen-bonding positions have been discovered for many RRs (37, 95, 106) and are often referred to as meta-active states. A single RR can have multiple meta-active states with different combinations of loop conformations, as shown for NtrX (37). Unphosphorylated RRs can exist in equilibrium between the active state and multiple inactive or meta-active conformations and there are multiple pathways for transition to the active state (37, 44, 89). NMR relaxation data also support a model of segmental motions of multiple allosteric residues for activation of CheY, instead of a strict two-state switching model (70).

Just as phosphorylation shifts the equilibrium for effector domain regulation, the reverse scenario, with output domains affecting phosphorylation, has also been observed (8, 37). It has been suggested that adoption of an active conformation precedes phosphorylation (44, 53).

Phosphorylation can be accelerated by any effect that increases the population of active conformation, such as relieving inhibition from the output domain by truncation, DNA binding in DBD-containing RRs, dimerization at high protein concentrations and HK-RR interaction (8, 9, 27, 28, 37, 66). Surface changes caused by these different interactions can propagate through the connecting βα loops to the phosphorylation site and other regions of the REC domain, increasing the active state population. One extraordinary example is DesR in which the HK-RR interaction can stabilize the active state of DesR and promote dimerization (66). Besides the phosphorylation-activated switch, interaction with the HK also functions as an allosteric switch to increase the population of active conformation for transcription activation. The same principle may be the basis for many phosphorylation-independent regulatory mechanisms.
Dimerization Modes and Corresponding Conformational Changes

Phosphorylation-promoted conformational changes are routinely identified by comparing X-ray or NMR structures obtained for unphosphorylated and BeF$_3$-bound RRs. As shown in Figure 5a, for the stand-alone RR CheY, significant differences in positions of backbone atoms occur at the β4-α4, β5-α5 loops and part of the α4-β5-α5 surface where CheY binds the FliM effector protein to regulate flagellar rotation. It is unsurprising that structural elements of the REC domain undergoing the largest rearrangements upon phosphorylation often correlate with protein-protein interaction surfaces, particularly, the dimer interface for DBD-containing RRs. Certain dimerization modes are popular for specific RR subfamilies with corresponding conformational changes.

A dimerization mode using the complete α4-β5-α5 surface is predominant in the OmpR subfamily (Figure 5b). Dimer interaction involves salt bridges between pairs of highly conserved charged residues within the α4-β5-α5 region (104, 105). For the DrrB protein shown in Figure 5b, conformational changes span the entire dimer surface. A Tyr residue is prevalent at the Y position (111) and switching of Tyr orientations has been observed. In several RRs (40, 78, 91), the hydroxyl group of Tyr is in close proximity to a polar residue from the DBD (Figure 5b right), suggesting a potential role in effector inhibition. This is unlikely to be a common mechanism given the wide diversity of domain arrangements within the subfamily (8). RRs from the OmpR subfamily often recognize tandem DNA sites and two DBDs bind DNA in a head-to-tail manner (15, 48, 64, 76). The translational symmetry of DBDs coupled with the rotational symmetry of the REC domain predicts a flexible linker or different REC-DBD interfaces for individual RR monomers within the RR-DNA complex. Such asymmetry of REC-DBD interfaces has been observed for KdpE and PmrA (65, 76). However, NMR studies suggest that
the REC-DBD interaction observed for PmrA in the crystal structure is transient in solution (65).

Additionally, alternative dimers involving α1-α5 (6, 71) or other surfaces (13) have also been discovered for unphosphorylated RRs but the physiological relevance of these dimers awaits further studies.

The NtrC subfamily of RRs displays two major modes of dimerization, with each involving different subsets of the α4-β5-α5 face (Figure 5c). The α4-β5 dimer interface is often seen in phosphorylated RRs (Table 1) or unphosphorylated RRs with meta-active conformations (32, 37, 84). Another β5-α5 dimer is associated with inactive conformations (36, 84), as indicated by the outward orientation of the Phe residue at the Y position. Dimer interaction centers around the β5-α5 face, sometimes with a slightly tilted α5 helix in contact with both β5 and α4. Formation of the inactive β5-α5 dimer is believed to inhibit the AAA,+ ATPase output domain from oligomerization and transcription initiation (10, 84). For the LytTR subfamily, structural information is limited and a dimer also involving the α4-β5 subset of the α4-β5-α5 face has been observed for ComE (16).

Diverse dimerization modes have also been observed in the NarL subfamily. An α4-β5 dimer interface and the phosphorylation-dependent orientation of the Phe residue have been reported for the REC domain of FixJ (14). However, most RRs from the NarL subfamily show an α1-α5 dimer interface (Table 1) distinct from the widely used α4-β5-α5 surface (7, 20, 29, 60, 82, 106). The dimer interface mainly involves the α1 helix, the β1-α1 and β5-α5 loops (Figure 5d). Correspondingly, significant structural rearrangements spanning to the α1 helix have been observed when apo- and BeF₃⁻-bound structures are compared. As discussed earlier, many RRs of the NarL subfamily display a constitutive inward orientation of the residue at the Y position, thus this residue is unlikely to be involved in allosteric regulation. Instead, different switching
mechanisms involving the T residue and other residues at α1, the β1-α1, β4-α4 and β5-α5 loops have been proposed for different RRs, such as VraR, DesR and RcsB (20, 60, 106). In addition to the α1-α5 dimer interface, the α4-β5-α5 surface is also remodeled to alter the positioning of the accessory β6 and α6 that are directly linked to the effector domain. Another dimer mode involving β6 and α6 is common (Figure 5d, right) and several RR, such as VraR (60), DesR (106), Spr1814 (82) and LiaR (29), show both α1-α5 and β6-α6 dimer modes within a single crystal. It has been suggested that the β6-α6 dimer may help RR to form higher order oligomers on DNA to bind arrays of DNA-binding sites (106). DNA-bound structures of RcsB again indicate an asymmetric dimer with different rotation symmetry axes for the REC and DBDs (38).

Different relative positions of the REC and DBDs are frequently observed for different full length RR. Asymmetry is believed to result from either different REC-DBD interactions (20, 38) or a flexible linker that enables different disposition or even domain-swap of individual domains trapped in crystals (29, 60).

5. RESPONSE REGULATOR – HISTIDINE KINASE INTERACTIONS

HK Conformational States

HKs have a modular architecture with a large variety of sensory domains (extracellular, transmembrane or cytoplasmic) linked to a conserved catalytic domain by one or more signal transducing domains (12, 69) (Figure 1a). The cytoplasmic enzymatic core of HKs consists of a dimerization histidine phosphotransfer (DHp) domain and a catalytic ATP-binding (CA) domain. The DHp domain contains the conserved phosphorylatable His residue and mediates dimerization of the HK to form a 4-helix bundle that is essential for its activity (Figure 6). The DHp dimer is flanked by the two α/β catalytic CA domains containing the kinase active sites that
catalyze phosphoryl transfer from ATP to the His residues. Autophosphorylation of the HK provides a high-energy phosphoryl group for subsequent phosphotransfer to the RR. Most HKs also mediate dephosphorylation of their cognate RRs in order to modulate the output response.

Large conformational changes in HKs occur during transitions between kinase, phosphoryltransfer and phosphatase states. In the HisKA_3 subfamily of HKs (e.g. DesK of Bacillus subtilis), the switch between states involves a large rotation of the DHp domains, a mechanism that is less pronounced in the HisKA subfamily. Specific regulatory mechanisms have been described for individual HKs, including the stabilization of the phosphatase state by ligand binding (e.g. c-di-GMP binding on CckA) and the inhibition of the phosphatase state by a pH-dependent conformational switch (33, 62). In contrast to the different conformations of HKs, structures of HK-REC domain complexes determined to date indicate that REC domains bound to either phosphotransfer or phosphatase states of HKs have similar conformations, with the REC domain displaying a meta-active conformation, as defined in the previous section.

Partner Recognition and Specificity

Specificity of HK-RR pairs is crucial for the integrity of signaling pathways, given the large number of different TCSs that typically exist in a single cell. Specificity within a pair is mediated by recognition residues that participate in HK-RR binding. These residues are located in helix α1 of the REC domain and in the two α helices of the HK DHp domain that contribute to the binding interface. Signaling pathways have been successfully rewired by substituting as few as three co-evolved specificity residues (88, 96). Additional strategies are employed to further ensure specificity of the pairs. For example, the phosphatase activity of HKs eliminates non-specific phosphorylation by non-cognate kinases or small molecule phosphodonors such as AcP.
The low-abundance of HKs relative to RRs also minimizes cross-phosphorylation between non-cognate pairs (57).

In an HK-RR complex, the REC domain inserts its $\alpha_1$ helix and $\beta_5-\alpha_5$ loop between the two helices of the DHp domain, primarily contacting the DHp $\alpha_1$ helix that contains the phosphorylatable His residue. This conserved mode of binding buries the active site of the REC domain at the HK-RR interface (Figure 6). Thus, conserved residues of both the HK and RR form a substantial subset of the interface residues, with specificity being determined by a relatively small set of variable residues. In addition to these specificity residues in $\alpha_1$ of the REC domain and the two helices of the DHp domain, additional variable contacts can involve the $\beta_2-\alpha_2$, $\beta_3-\alpha_3$ and $\beta_4-\alpha_4$ loops of the REC domain interacting with the CA domain, the C-terminal region of the DHp and the DHp-CA linker of the other protomer of the dimer (Figure 6) (21, 72, 74, 88, 107, 108, 114, 118, 122). Other contacts can occur between the REC domain and additional domains within the HK, increasing the specificity and/or stability of the pair (e.g. contacts between the PAS domain of ThkA and the RR TrrA) (118). These elements suggest that beyond the conserved REC-DHp interface that buries the active site, the binding interface can vary greatly in different HK-RR pairs.

A model based on structures of HK FixL and full-length RR FixJ suggests that the RR effector domain is free to move and does not participate in binding to the HK (116). However, a lack of structures of full-length RRs bound to HKs leaves open the question as to how effector domains might be positioned relative to REC domains and domains of the HK. Given the great diversity in REC-effector domain arrangements in different inactive RRs, it is possible that diversity exists in the ways effector domains interact in different HK-RR complexes. It is easy to envision how an effector domain could participate in the stability of the complex and/or the
efficiency of catalysis. Indeed, the meta-active conformation of the REC domain observed in
HK-REC domain complexes might be a consequence of isolated REC domains being unhindered
by REC-effector interactions that could bias conformational states.

**Phosphotransfer State**

Structural and biochemical studies of DesK-DesR complexes suggest that one HK dimer binds to
one RR molecule in an asymmetric conformation. The DesK-DesR structure and a docking
model of CpxA-CpxR in a putative phosphotransfer state show asymmetry within the HK dimer,
with the CA domain of one monomer bound to the DHp domain of the other monomer, leaving
the second CA domain untethered with enough space for RR binding to the DHp. Structural data
from the DesK-DesR complex and from CpxA-CpxR model suggest that this highly dynamic
state is coupled with autokinase activity, involves back-and-forth movements of the DHp and CA
domains, and is modulated by REC domain binding (72, 107).

**Phosphatase State**

In contrast to the asymmetrical domain arrangements associated with phosphotransfer, HK-RR
complexes in putative dephosphorylation states have symmetrical arrangements of the HK DHp
and CA domains. The REC domains are also arranged symmetrically, each interacting with a
DHp domain with a 2:2 HK:RR stoichiometry (62, 107). The HK in its phosphatase state is more
rigid than in the phosphotransfer state, with the N-terminal coiled-coil of the DHp domains being
highly stable.

Despite the large number of structures recently determined, the mechanistic details of the
enzymatic reactions involving HKs and RRs are not fully understood. Discrimination between
the phosphotransfer and phosphatase states that have been trapped in crystal structures still
remains challenging. The situation is further complicated by distinct phosphatase mechanisms
used by HKs of different subtypes. A major limitation of current studies is the use of truncated proteins, most notably the cytoplasmic domains of transmembrane HKs lacking sensor and transmembrane regions that in intact proteins control the signaling states of HKs, and REC domains of RRs in the absence of effector domains that are known to influence their conformational equilibria. While conserved features of HK-RR interactions are beginning to be elucidated, variations on the theme are anticipated. Similar to many other aspects of TCS structure and function, it is likely that the great variety of domain architectures in TCS proteins enable different modes of HK-RR interactions and regulatory mechanisms that are adapted to individual HK-RR pairs.

6. NON-CANONICAL MODES OF RR REGULATION

The canonical mechanism for regulation of RR activity involves phosphorylation at a conserved Asp in the REC domain, stabilizing an active conformation that enables effector domain function. Regulation of RR activity, in addition to or in place of Asp phosphorylation, potentially can be achieved in many different ways including post-translational modifications at other sites that bias the conformational equilibrium of the REC domain or directly alter effector domain function, interactions that lower the energetic barrier for transition to an active conformation, ligand binding to the RR, and regulation of expression to control RR levels and thus dimerization or oligomerization. Indeed, each of these strategies has been observed and a few examples are described below.

Reversible protein acetylation is known to regulate metabolic enzymes and transcription in bacteria (49) and was documented in CheY more than 20 years ago (90). Enzyme catalyzed acetylation of CheY at K91 and K106 increases in response to acetate and promotes clockwise
flagellar rotation, with acetylation at K91 proposed to alter conformational dynamics of the β4α4 loop (39). Autoacetylation with AcCoA as acetyl donor at other sites in CheY is thought to link chemotaxis to the metabolic state of the cell (120). Another well-characterized example of acetylation occurs in *E. coli* transcription factor RcsB. Acetylation of K154, a residue in the helix-turn-helix motif that interacts with a phosphate in the DNA backbone, disrupts DNA binding, down-regulates expression of the acid stress response genes, and compromises survival in acidic conditions (22, 51). It should be noted that the stoichiometry of acetylation has not been determined, thus physiological effects observed with mutations that mimic fully unmodified or acetylated states might overestimate regulation that occurs in wild-type cells. A proteomics study in *E. coli* documented acetylation of seven RRs: ArcA, CheY, CpxR, EvgA, NarL, OmpR and RcsB (22). The *E. coli* genome encodes 26 putative Gcn5 N-acetyltransferases (GNATs) and a single known deacetylase (CobB) (49). While the best-characterized acetyltransferase, PatZ, is proposed to mediate a global strategy linking regulation to metabolic state, the large number of acetyltransferases raises the possibility of substrate specificity that might enable system-specific signaling mechanisms.

Signal-induced, HK-dependent, phosphorylation-independent monomer to dimer activation has been observed for several RR transcription factors. In the extensively studied *B. subtilis* system, binding of DesR to DesK promotes an active conformation of DesR. At sufficiently high RR concentrations and with a slow active to inactive state transition, dimerization can occur upon release of the RR from the HK with subsequent stabilization of the dimer upon DNA binding (106). Similar phosphorylation-independent activation of OmpR by HK EnvZ has been observed in acidic conditions (23). Cyanobacterial transcription factor NblR contains a conserved Asp, but lacks other residues necessary for phosphorylation. HK NblS is
required for phosphorylation-independent activation of NblR. However, no NblS-NblR
interactions have been detected and the activating monomer to dimer transition is postulated to
be promoted by another protein partner (93).

Multiple strategies have been identified for regulation of orphan RRs that lack a
conserved Asp and/or other conserved residues necessary for phosphorylation. In streptomyceses,
two atypical RR transcription factors that lack residues necessary for phosphorylation, JadR1 and
RedZ, are regulated by the end products of the antibiotic biosynthetic pathways they control. The
antibiotic JdB binds directly to the JadR1 REC domain, disrupting DNA binding (112). A
different strategy is used by Helicobacter pylori HP1043, which exists as a constitutively active
dimer in vitro with a crystal structure similar to that of other activated OmpR subfamily
members (50). Levels of HP1043 are regulated both transcriptionally and post-transcriptionally
leading to speculation that control of expression of this constitutively active RR may be the sole
mechanism for regulating HP1403 activity (75).

Combinations of these mechanisms create even more strategies. Streptomyces coelicolor
GlnR, which regulates genes for nitrogen assimilation, is an orphan OmpR subfamily RR that
lacks residues for Asp phosphorylation and forms a constitutive α4-β5-α5 dimer (61). GlnR is
phosphorylated at 6 Ser/Thr sites in the DNA-binding domain under N-rich conditions,
disrupting DNA binding. GlnR is also acetylated at multiple Lys residues in the DNA-binding
domain, with acetylation enhancing DNA binding (4).

7. CONCLUDING REMARKS

When the structure of an RR REC domain was first reported thirty years ago, a central question
was how this single conserved domain could regulate responses as diverse as flagellar rotation,
transcription, and enzyme activity. The answer that emerged defined a mechanism that was both simple and versatile. The small α/β REC domain exists in equilibrium between two primary conformations with phosphorylation stabilizing an active conformation. This phosphorylation-regulated switch enables regulatory strategies via any type of activating or inhibitory macromolecular interactions that discriminate between the two states. Hundreds of structures of RR s have provided descriptions of the conformations of REC domains in inactive and active states, interactions with effector domains, DNA, HKs and auxiliary proteins. These structures provide a foundation for identifying conserved features as well as specific variations in individual RR s. Beyond the universally conserved enzymatic mechanisms facilitated by configurations at the active site, other features such as Y-T coupling, regions of conformational perturbations, domain arrangements and modes of DNA binding show distinct trends among RR s within specific subfamilies.

Numerous variations on most every feature of RR s have been observed and undoubtedly, many more remain to be discovered. The plasticity of the REC domain and versatility of RR design allow an almost unlimited array of adaptations to fit the needs of individual signaling systems. Defining details of how specific structural features impact function is important for interpreting the nuances of RR behavior within specific TCSs as well as for pursuit of applied projects such as development of antimicrobial therapeutics or the engineering of synthetic biosensing pathways.

Our current understanding of RR structure is largely informed by compiling information derived from many partial structural descriptions of different RR s. Very few structures are available for full-length multi-domain RR s in both inactive and active states. The crystal structures that do exist are constrained by the limitations of the methodology, specifically the
capture of single discrete states that do not reflect conformational distributions and the potential
promotion of inter- and/or intra-molecular interactions that are influenced by the high
concentrations of proteins used in crystallization and/or stabilization of conformations necessary
for crystallization and lattice interactions. While NMR studies have provided information about
dynamics and conformational distributions in solution, size limitations have mostly precluded
studies of full-length RRs, dimers, oligomers and complexes of RRs with HKs or other
macromolecular partners. Thus, while structures unambiguously define allowable states, it is
important to keep in mind the conformationally dynamic nature of RRs and the potential
influence of associated domains and macromolecular partners when interpreting structures
determined by methodologies with technical limitations. The emergence of new structural
methods such as high-resolution cryo-electron microscopy promise to provide solutions to some
of these challenges.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that
might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Research in the Stock laboratory is supported by grant R01 GM47958 from the National
Institutes of Health. S.B. was supported in part by the Foundation pour la Recherche Médicale
(FRM) Grant SPE201803005070.
LITERATURE CITED


   Curr. Opin. Microbiol. 13:142-49

   Protein Data Bank (PDB): The single global macromolecular structure archive. Methods 
   Mol. Biol. 1607:627-41

   crosstalk enable the expansion of paralogous signaling protein families. Cell 150:222-32

   Conformational dynamism for DNA interaction in the Salmonella RcsB response regulator. 
   Nucleic Acids Res. 46:456-72

   phosphoryl transfer in two-component signal transduction. Cell 139:325-36

   acetylation affects acetate metabolism, motility and acid stress response in Escherichia coli. 
   Mol. Syst. Biol. 10:762

   activation of OmpR drives acid and osmotic stress responses in single bacterial cells. Nat. Commun. 8:1587

   governed by a repurposed bacterial histidine kinase. PLoS Biol. 12:e1001979

   CheY. J. Mol. Biol. 297:543-51


70. McDonald LR, Boyer JA, Lee AL. 2012. Segmental motions, not a two-state concerted switch, underlie allostery in CheY. *Structure* 20:1363-73


TERMS AND DEFINITIONS

AcP acetyl phosphate
BeF₃ berylliofluoride
CA catalytic/ATP-binding (domain)
DBD DNA-binding domain
DHp dimerization/histidine phosphotransfer (domain)
ECF extracytoplasmic function
HK histidine kinase
HPt histidine-containing phosphotransfer (domain)
PDB Protein Data Bank
REC receiver (domain of response regulator protein)
RR response regulator protein
SDRR single-domain response regulator
TCS two-component system
**FIGURE CAPTIONS**

**Figure 1.** Common features of RRs. (a) Schematic diagram of the prototypical TCS pathway. (b) The conserved (βα)₅ fold of REC domains. Phosphorylation site residues and the phosphoryl group mimic, beryllofluoride (BeF₃⁻), are shown in sticks. (c) Sequence conservation of RECs. The profiled hidden Markov model (HMM) for the REC protein family (Pfam PF00072) is shown as sequence logos (113) with the secondary structure elements illustrated. Heights of individual stacked letters at each position correspond to information contents, reflecting the probability of observing the particular amino acids at each position. Phosphorylation site residues (stars), named after the most conserved amino acids (DD, D, T and K), are among the most conserved residues in REC domains.

**Figure 2.** Classification of RRs by their effector domains. The percentile distribution is indicated for RR effector functional classes and RR subfamilies. RR subfamilies are defined by effector domain folds identified in Pfam. Representative proteins are traditionally used to name the OmpR, NarL and NtrC subfamilies with effector domains named Trans_reg_C, GerE and Sigma54_activa in Pfam. Representative structures of each subfamily are shown with REC domains colored in grey and effector domains in colors (PDB ids: KdpE, 4KNY; RcsB, 5W43; LuxO, 5EP0; AgrA, 3BS1; AmiR, 1QO0; WspR, 3BRE; CheB, 1A2O).

**Figure 3.** Phosphorylation site of the REC domain. The phosphoryl group is positioned by a network of hydrogen bonds (dashed lines) with side chains of the highly conserved residues (orange) as well as backbone atoms of non-conserved residues (light pink). Non-conserved
active site residues are labeled by their relative sequence positions to the nearest conserved residues, such as T+1, indicating one residue C-terminal to the conserved Thr/Ser residue. (a) Phosphorylation site of the archetype RR CheY in the active conformation (PDB id: 1FQW) with residues that are differently positioned in the inactive conformation (2CHE) shown in cyan. (b) The putative trigonal bipyramidal transition state for both phosphorylation and dephosphorylation. X represents the leaving group of the phosphodonor for phosphorylation or the attacking water for dephosphorylation. (c) Surface view of the active site in the phosphatase-REC complex (3HZH).

Figure 4. RR regulatory strategies. (a) Distinct inactive and active RR conformations exemplified by full-length RR VraR (PDB ids: 4GVP, 4IF4). BeF$_3^-$ is shown in red spheres. (b) Schematic diagrams of RR regulatory mechanisms. Functional sites of effector domains, such as enzyme sites or DNA recognition regions, are shown as pink dots. These sites can be buried or exposed in a wide variety of inactive RR conformations and their activity depends on different interactions between the REC and effector domains. Phosphorylation of the RR can relieve REC domain inhibition, promote effector function or both. Representative RRs that utilize these strategies are indicated.

Figure 5. Phosphorylation-induced conformational changes in the REC domain. RR structures with or without BeF$_3^-$ were aligned using the conserved strands $\beta$1, $\beta$3, $\beta$4, and $\beta$5 to compute the average backbone RMSD per residue (a, right). RMSD values above the median $+$ 2x MAD (median absolute deviation) are considered as significant conformational changes and the
corresponding residues are colored blue. Non-REC structural elements are colored cyan. BeF$_3$ is shown as red spheres and residues involved in the potential Y-T coupling are shown as green or gray sticks. Representative protein structures shown for the (a) Stand-alone, (b) OmpR, (c) NtrC and (d) NarL RR subfamilies are CheY (1F4V), DrrB (3NNS, 1P2F), NtrC1 (1ZY2, 1NY5) and VraR (4IF4). Y/T residues from the inactive CheY structure (2CHE) are differently positioned from the active structure. For the OmpR, NtrC and NarL subfamilies, two proteins from each subfamily were used for RMSD analyses and both showed similar regions of conformational changes. Substantial changes in $\alpha$1 are also observed in one protein from the NtrC subfamily (pink).

Figure 6. Structure of the HK-RR complex. Ribbon (a) and surface (b) views of the HK856-RR468 complex (PDB id, 3DGE). Residues that determine the HK-RR interaction specificity (19, 96) are highlighted in light orange and cyan in the HK and RR, respectively. HK-RR contacts also involve other surface regions (grey), including both active sites of the HK (red) and the RR (pink).
Table 1. Dimerization modes and orientations of the conserved aromatic residue in RR subfamilies.

<table>
<thead>
<tr>
<th>Family</th>
<th>P</th>
<th>Position of Y/F (inward)</th>
<th>Active Dimer Interface</th>
<th>Alt. Dimer Interfaces</th>
</tr>
</thead>
<tbody>
<tr>
<td>OmpR</td>
<td>+</td>
<td>10/10&lt;sup&gt;a&lt;/sup&gt;</td>
<td>α&lt;sub&gt;4&lt;/sub&gt;-β&lt;sub&gt;5&lt;/sub&gt;-α&lt;sub&gt;5&lt;/sub&gt; (10/10) (1zes, 1xhf, 4uhk, 4s04, 6br7,...)</td>
<td>N.A.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>α&lt;sub&gt;4&lt;/sub&gt;-β&lt;sub&gt;5&lt;/sub&gt;-α&lt;sub&gt;5&lt;/sub&gt; (1zes, 1xhf, 4uhk, 4s04, 6br7,...)</td>
<td>N.A.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3/9 (MD)</td>
<td>α&lt;sub&gt;4&lt;/sub&gt;-β&lt;sub&gt;5&lt;/sub&gt;-α&lt;sub&gt;5&lt;/sub&gt; (MD 3/9, Rec 16/18) (1xhe, 2hqr, 3nhz, 3r0j, 4knr, 4uhs,...)</td>
<td>α1-α5 (2/27) (1b0o, 4uhs)</td>
</tr>
<tr>
<td>NtrC</td>
<td>+</td>
<td>6/6</td>
<td>α&lt;sub&gt;4&lt;/sub&gt;-β&lt;sub&gt;5&lt;/sub&gt; (4/5) (1l5y, 1zy2, 2jrl, 4d6y)</td>
<td>β5-α5 (1/5) (2vui)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>α&lt;sub&gt;4&lt;/sub&gt;-β&lt;sub&gt;5&lt;/sub&gt; (4/5) (1l5y, 1zy2, 2jrl, 4d6y)</td>
<td>β5-α5 (1/5) (2vui)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0/4 (MD)</td>
<td>α&lt;sub&gt;4&lt;/sub&gt;-β&lt;sub&gt;5&lt;/sub&gt; (MD 0/4, Rec 2/7) (3cfy, 4d6x)</td>
<td>N.A.</td>
</tr>
<tr>
<td>NarL</td>
<td>+</td>
<td>8/8</td>
<td>α&lt;sub&gt;1&lt;/sub&gt;-α&lt;sub&gt;5&lt;/sub&gt; (7/8) (4lf4, 4ldz, 4zmr, 5hev, 5o8z, 4e7p, 4le0)</td>
<td>β6-α6 (6/8) (4lf4, 4ldz, 4zmr, 5hev)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>α&lt;sub&gt;1&lt;/sub&gt;-α&lt;sub&gt;5&lt;/sub&gt; (7/8) (4lf4, 4ldz, 4zmr, 5hev, 5o8z, 4e7p, 4le0)</td>
<td>β6-α6 (6/8) (4lf4, 4ldz, 4zmr, 5hev)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6/8 (MD)</td>
<td>α&lt;sub&gt;1&lt;/sub&gt;-α&lt;sub&gt;5&lt;/sub&gt; (MD 4/8, Rec 5/8) (1ml, 5vxn, 4hve, 4le1, 3eul, 3b2n...)</td>
<td>β6-α6 (7/16) (4hve, 4le1, 2qsj, 5f64...)</td>
</tr>
<tr>
<td>LytTR</td>
<td>+/-&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1/1 (MD)</td>
<td>α&lt;sub&gt;4&lt;/sub&gt;-loop-β&lt;sub&gt;5&lt;/sub&gt; (4cbv, 4ml3)</td>
<td>N.A.</td>
</tr>
</tbody>
</table>

a: Numbers are for individual protein fragments. The REC domain or multi-domain (MD) fragments of the same protein are individually counted.
b: Mixed positioning. Protein chains contain both inward and outward orientations of Y within a single crystal symmetry unit.
c: Phosphorylation status is mimicked with D to E or D to A mutations.
Figure 1

**a**

Phosphoryl transfer

Stimuli → Histidine Kinase → Responses

**b**

Response Regulator

**c**

Graph showing interactions between different components.
Figure 2
Figure 3

(a) BeF$_3$-T$_{D+2}$Mg$^{2+}$DD1 DD2

(b) X = HK or small donor (phosphorylation)
X = OH$_2$ (dephosphorylation)

(c) Phosphatase

X = HK or small donor (phosphorylation)
X = OH$_2$ (dephosphorylation)
Figure 4

(a) **Inactive**

Inactive Rec DNA binding domain Recognition helix Linker

(b) **Inactive**

Inactive CheB NarL PrrA MtrA RitR DrrB VraR DrrD ChrA FixJ

Inactive Enzyme Inhibition Relief DctD, NtrC1, LuxO, ...

Active Enzyme Positive Regulation KdpE, PmrA, RcsB, ...

Inactive Enzyme
Figure 5

- **a** Figure showing a structural diagram of FilM, with residues labeled α4 and α5, and a graph depicting Phos. Dependent Backbone RMSD (Å) over residue number.
- **b** Diagram illustrating the OmpR subfamily, showing a 4-5-5 Dimer with α5, α4, and α5' labeled, and a DNA-binding domain with α4 labeled.
- **c** Diagram for the NtrC subfamily, featuring a 4-5 Dimer with α4 and α5 labeled, and a 5-5 Dimer with α5 labeled.
- **d** Diagram for the NarL subfamily, displaying a 1-5 Dimer with α1 labeled, and an Alternative Dimer with α1 labeled.
Figure 6