Chapter 15

Bacterial Two-Component Systems: Structures and Signaling Mechanisms

Shuishu Wang

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/48277

1. Introduction

Two-component systems (TCS) are ubiquitous among bacteria. They play essential roles in signaling events in bacteria, such as cell-cell communication, adaptation to environments, and pathogenesis in the case of pathogens. Due to their absence in humans and other mammals, TCS proteins are considered potential targets for developing new antibiotics. Most bacterial TCSs consist of two proteins, a sensor histidine kinase (HK) and a response regulator (RR). The HK senses specific signals, and that leads to activation of the kinase activity and autophosphorylation of a conserved histidine residue. The phosphoryl group is subsequently transferred to a cognate response regulator to activate its activities. Most RRs are transcription regulators and turn on or off gene transcriptions in response to the signals received by sensor HKs. Recent years have seen a rapid expansion of structural data of bacterial TCS proteins. In this chapter, I will review structures of HKs and RRs and discuss their structure-function relationship and their signaling mechanisms. The chapter contains three main sections: structures of histidine kinases, structures of response regulators, and structures of complexes between a histidine kinase and a response regulator. I will conclude the chapter with the implications of these structural and functional data on developments of new therapeutics against bacterial pathogens. Analysis of structures and reaction mechanisms will focus on the extracytosolic sensor HKs and OmpR/PhoB subfamily transcription regulator RRs.

2. Structures of histidine kinases

The prototypical histidine kinase is a homodimeric integral membrane protein (Figure 1). Each protomer has two transmembrane (TM) helices with the N-terminus in the cytosol. An extracytosolic sensor domain lies between the two TM helices. After the second TM helix is a HAMP domain, which is commonly found in Histidine kinases, Adenylyl cyclases, Methyl-

accepting chemotaxis proteins, and Phosphatases [1], hence the name. The HAMP domain connects TM2 to the dimerization and histidine phosphorylation domain (often abbreviated as DHp). A catalytic and ATP-binding (CA) domain lies at the carboxyl terminus. The combination of DHp and CA is sometimes referred to as the kinase domain. The TM helices, HAMP domains, and DHp domains are all involved in homodimerization. The sensor domain is likely to dimerize in the context of the entire protein. Sensor domains share little sequence identity, as is expected from their diverse functions of sensing various signals. However, available structures of sensor domains fall into a few common structural folds, suggesting conserved signal sensing mechanisms. The HAMP, DHp, and CA domains are common modules of HKs and have well conserved structures and sequences, especially DHp and CA, whose sequences contain several conserved motifs. There is an absolutely conserved histidine residue that is phosphorylated, and the phosphoryl group is then donated to an RR, in response to signals sensed by the sensor domain. How the sensor domain regulates the kinase activities is still unknown, due to the lack of full-length structures of the transmembrane sensor HKs. However, a large accumulation of structures of isolated domains in recent years has started to shed light on possible molecular mechanisms of the signal transduction. In this section, I will summarize these structures and discuss their conformational changes that transmit signals from the sensor domain to the kinase domain.

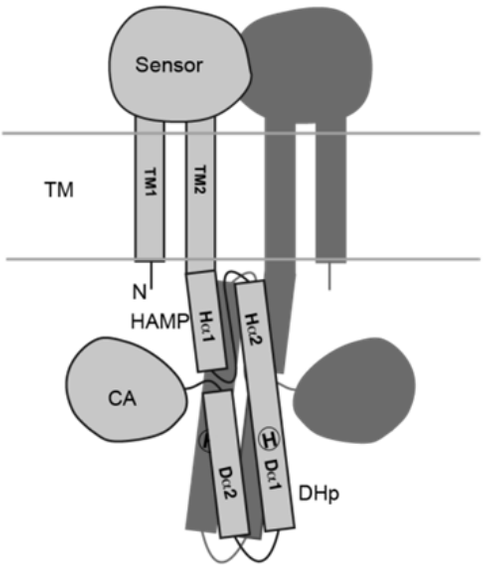

Figure 1. Schematic diagram of the modular structure of the prototypical sensor histidine kinase. At the N-terminus, a sensor domain lies between the two transmembrane (TM) helices. A HAMP domain follows TM2 with its second helix (Hα2) connected to helix Dα1 of the DHp domain. The catalytic and ATP-binding (CA) domain is at the C-terminus. The position of the phosphorylation site histidine is marked with an "H". The HAMP and DHp domains form homodimers.

2.1. Structures of sensor domains

Sensor domains of histidine kinases are located in cytosol, in membrane, or outside of cell membrane (extracytosolic). Currently, there is little structural information of the membrane-

embedded sensor domains. The prototypical HK has an extracytosolic sensor domain that senses extracellular signals or conditions in the cell envelope. These sensor domains have highly diverse sequences. However, most of the known structures of extracytosolic sensor domains fall into three distinct structural folds, mixed αβ, all-helical, and β-sandwich. Unlike extracytosolic sensor domains, many cytosolic sensor domains can be annotated on the sequence level as PAS or GAF domains, which have related structural folds and are named from their occurrence in Period circadian, Aryl hydrocarbon receptor nuclear translocator, and Single-minded proteins (PAS) [2], or in cGMP-regulated cyclic nucleotide phosphodiesterases, Adenylate cyclases, and the bacterial transcriptional regulator FhlA (GAF) [3].

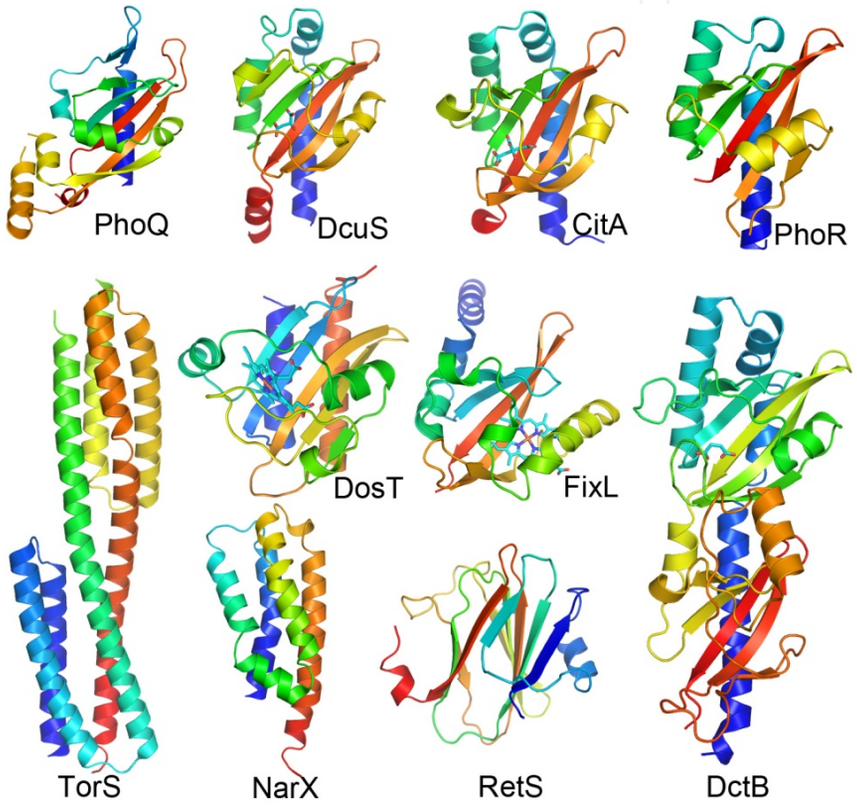

Figure 2. Structures of sensor domains. The structures are each colored in a rainbow spectrum from blue to red from the N- to C-terminus, respectively. The DosT (PDB code 2VZW, *M. tuberdulosis*) and FixL (1EW0, *R. meliloti*) sensor domains are located in the cell cytosol and belong to GAF and PAS domains, respectively; the remainders are extracytosolic domains. The PhoQ (3BQ8, *Escherichia coli*) and DcuS (3BY8, *E. coli*), and PhoR (3CWF, *Bacillus subtilis*), and CitA (2J80, *Klebsiella pneumoniae*) sensor domains have a PDC fold. DctB (3BY9, *Vibrio cholerae*) has two tandem PDC domains. The structures are shown in similar orientations with the N- and C-termini facing down for easy comparison, except DosT and FixL, which are shown with their central β-sheet having the same orientation as the PDC domains but their N- and C-termini facing other directions. The ligands and heme groups of DcuS, CitA, DctB, DosT, and FixL are shown as sticks. All of them are bound at the central β-sheet. NarX (3EZH, *E. coli*) has an all-helical sensor domain, while TorS (3I9Y, *Vibrio parahaemolyticus*) has double helical domains. The RetS (3JYB, *Pseudomonas aeruginosa*) sensor domain is currently the sole structural representative of β-sandwich fold extracytosolic sensors. The ribbon diagrams were prepared with PyMOL (Schrodinger, LLC).

DosS and DosT from *Mycobacterium tuberculosis* are soluble sensor HKs that have dual GAF domains located in the cytosol. The first GAF domain binds a heme group for sensing redox and hypoxia conditions of the cell cytoplasm to regulate the dormancy regulon *Dos*, which is believed by some researchers to allow *M. tuberculosis* to survive in a latent state for decades [4]. The structures of the N-terminal GAF domains of DosS [5] and DosT [6] reveal highly similar structures with a central five-stranded antiparallel β-sheet flanked on one side by the first and last helices and on the other side by some loops and short helices (Figure 2, DosT). The heme group is bound in the center of the β-sheet with the plane of the heme perpendicular to the β-sheet. The PAS domain of FixL from *Rhizobium meliloti* has the same topology of the central β-sheet, but has different α-helices (Figure 2) [7]. The heme binds at a different location of the sheet in a parallel manner.

The most common structural fold of the extracytosolic sensor domains is the mixed αβ fold, which has an identical topology to that of PAS domains [8]. The structures consist of a central 5-stranded antiparallel β-sheet, flanked by α-helices on both sides (Figure 2, top row). These extracytosolic sensor domains are slightly different in structure from the PAS and GAF domains. Their structures fall into a group by themselves and are named as PDC domains, referring to first three structures of this group, i.e. sensor domains of PhoQ, DcuS, and CitA [9]. Among the PDC domains, the structures are similar to each other, especially the N-terminal helix and the position of the C-terminus. The PDC domains have a long N-terminal helix, possibly continuous from the TM1 helix. The N- and C-termini of the domain are next to each other, and both are facing the same direction toward the membrane for connecting to the transmembrane helices. In some structures, there is a short helix at the C-terminus, possibly continuous to the TM2 helix.

Some HK sensor domains are all-helical, represented by those of NarX [10] and TorS [11] (Figure 2). The NarX sensor domain is an antiparallel four-helix bundle. Most of the helices have kinks or bends, suggesting mobility. TorS has two antiparallel four-helix bundles stacked along their bundle axes, with the second one inserted between the last two helices of the first bundle. Their mechanisms for signal sensing are different. NarX binds directly to nitrite and nitrate; while TorS interacts with a periplasmic binding protein TorT to detect trimethylamine-N-oxide.

The crystal structure of the sensor domain of RetS from *P. aeruginosa* reveals a β-sandwich fold (Figure 2) [12]. RetS works with two other HKs, LadS and GacS, to regulate genes for type III secretion system controlling acute infection and those for drug-resistant biofilm formation controlling chronic infection. The RetS sensor domain has two antiparallel β-sheets, with topologies β1-β3-β8-β5-β6 and β2-β9-β4-β7, stacking back-to-back. The β-sandwich fold structure resembles carbohydrate-binding modules.

2.2. Structure and function of HAMP domains

HAMP domains are widely occurring domains of prokaryotic transmembrane receptors. They follow the last TM helix and thus connect the extracytosolic sensory domain to the

cytosolic signaling domains. The sequences of HAMP domains have heptad repeats, in which hydrophobic residues occupy positions a and d (Figure 3). The structures of HAMP domains are dimers with each protomer of ~50 residues forming two helices (referred to as Hα1 and Hα2) connected by a linker of ~14 residue. The four helices from two protomers associate to form a parallel four-helix bundle. The linker residues have an extended structure that spans the length of the four-helix bundle [13, 14]. Other than the hydrophobic side chains of the heptad repeats and a conserved glycine residue at the beginning of the linker, HAMP domains do not have strong sequence conservation. Nevertheless, the domains can be exchanged, yielding hybrid proteins that are still functional [15, 16], which suggests that HAMP domains share a conserved mechanism for propagating signals.

The HAMP domain of Af1503, a hypothetical receptor from the archaeon *Archaeoglobus fulgidus* [14] is the first HAMP structure determined (Figure 3). The packing of the 4 helices is different from the expected conventional knobs-into-holes geometry, in which a residue from one helix (knob) packs into a space surrounded by four side chains of the facing helix (hole). Instead, the helical packing of the Af1503 HAMP domain is described as complementary x-da, in which the x- and da-layers alternate along the helices to form mixed x-da layers. In the x-layer, side chains point straight at the central supercoil axis; and in the da-layer, side chains point side ways to form an interacting ring of residues enclosing a central cavity. This non-canonical packing geometry is stabilized by the small side chain of alanine at residue 291 [17]. Mutations of A291 with amino acids of larger side chain convert the helical packing to the knobs-into-holes conformation through a concerted axial rotation of all four helices of ~26° (Figure 3). It was hence proposed that signals propagate through the TM helices and HAMP by helix rotation around an axis perpendicular to the membrane, a so-called gearbox model [18].

The crystal structure of a tri-HAMP unit from the *P. aeruginosa* soluble receptor Aer2 shows that the three HAMP domains each form a parallel four-helix bundle similar to HAMP of Af1503 (Figure 3) [13]. This confirms that the parallel four-helix bundle is likely universal among HAMP domains. The top HAMP domain (HAMP1) is separated from HAMP2 by a helical insert that connects Hα2 of HAMP1 to Hα1 of HAMP2 by a slightly kinked helical insert. HAMP2 and HAMP3 are contiguous and form a concatenated di-HAMP structure. There are two distinct packing configurations among the three HAMP domains. HAMP1 and HAMP3 are similar to each other and resemble that of Af1503. The middle HAMP2 domain, however, has a unique helical packing, in which there is an offset of the helical register between α1 and α2 of half a helical turn (~2-3 Å). This results in side chain staggering of the buried hydrophobic core between the two helices of the same polypeptide chain. The two α1 helices coil around each other, so do the two α2 helices; but the two coiled-coils have limited interactions with each other. The unique packing of HAMP2 leads to a slightly different possible mechanism of signaling propagation by the HAMP domain: a combination of vertical movement and helical rotation, also known as screw-like motion.

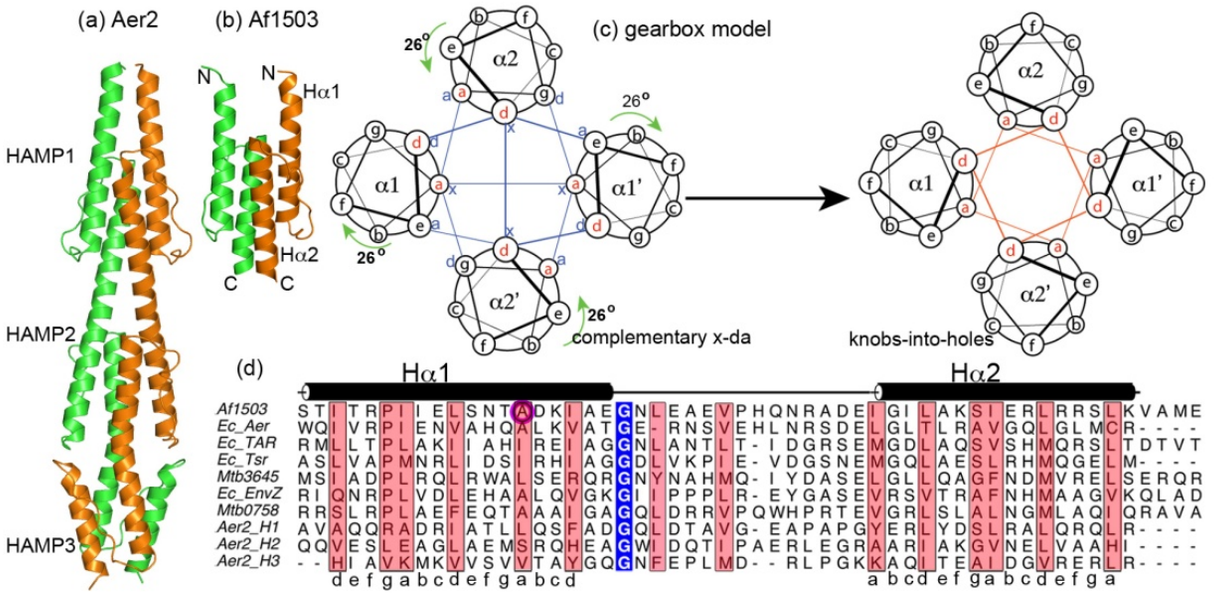

Figure 3. HAMP domain structures, a gearbox model for signal propagation, and sequences. The crystal structure of a tri-HAMP domain from Aer2 (Pdb code 3LNR) shows that all three HAMP domains have a four-helix bundle structure (a), which is similar to that of the HAMP domain of Af1503 (2L7H) (b). Each four-helix bundle is a homodimer of a parallel two-helix coiled-coil of helix Hα1 and Hα2. Af1503 has a helical packing of complementary x-da, which can be converted to the canonical knobs-into-holes supercoil by a 26 º rotation of each helix (c). (d) Sequence alignment of HK HAMP domains shows heptad repeats with hydrophobic side chains at positions a and d. Conserved hydrophobic residues are highlighted in red. There is a highly conserved Gly at the beginning of the linker, highlighted in blue. Residue A291 in Af1503 is circled. Aer2_H1, Aer2_H2, and Aer2_H3 in (d) are sequences for Aer2 HAMP1, HAMP2, and HAMP3, respectively.

2.3. The dimerization and phosphorylation domain

The dimerization and phosphorylation domain is a conserved domain of histidine kinases. This domain is often abbreviated as DHp for dimerization and histidine phosphotransfer. It is also referred to as dimerization domain or histidine kinase domain A. The sequence of DHp contains the H box that is part of the signature motifs defining histidine kinases (Figure 4). The histidine residue in the H box motif is absolutely conserved and is the site for autophosphorylation and subsequent transfer of the phosphoryl group to downstream proteins. This domain participates in dimerization of the histidine kinase. DHp is located in the cytosol and directly follows the HAMP domain, and it is immediately followed by the CA domain in the prototypical histidine kinases (Figure 1).

The structure of the DHp domain consists of two α-helices (referred to as Dα1 and Dα2), that form an antiparallel coiled-coil (also referred to as helical hairpin), and two protomers associate to form a four-helix bundle (Figure 5). The conserved histidine residue is on helix Dα1. Similar to HAMP domains, sequences of the DHp domains have heptad repeats. Several structures of DHp have been determined. The NMR structure of the *E. coli* EnvZ DHp shows a four-helix bundle that is relatively straight [19]. The phosphorylation site H243 is solvent exposed and located in the middle of helix Dα1, which is more mobile than

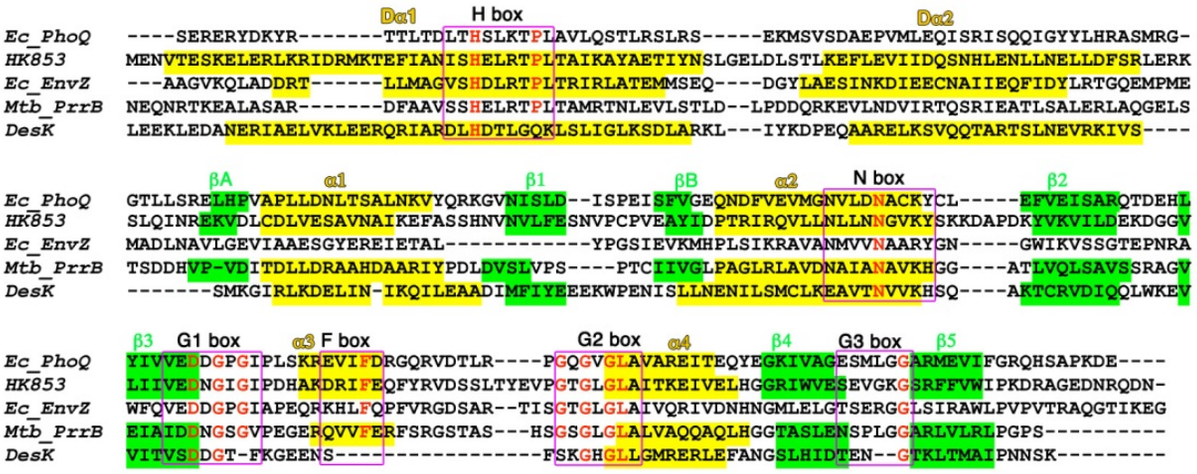

Figure 4. Sequences of the DHp and CA domains of HKs. Conserved sequence motifs H, N, G1, F, G2, and G3 boxes are in magenta boxes. Helices are highlighted in yellow, and β-strands in green.

helix Dα2, especially around the conserved H-box residues. The intra-subunit interface is mostly hydrophobic; the dimer interface, however, contains two acidic clusters. A related phosphotransferase Spo0B from *B. subtilis* has a similar dimerization domain of four-helix bundle [20]. The active site residue H30 is in the middle of helix Dα1, with its side chain protruding to the surface, similar to that of the EnvZ DHp domain.

The structure of the entire cytoplasmic portion of HK853, a putative sensor HK from *Thermotoga maritima*, shows that the DHp domain forms a four-helix bundle slightly different from that of EnvZ [21]. The HK853 four-helix bundle has a larger twist angle of ~25°, compared to that of EnvZ, which is unusually straight and parallel. The connections between hairpin helices have different topologies between the two structures. EnvZ is unique in this respect as all other known DHp structures have the same topology as HK853 (Figure 5). HK853 does not have a HAMP domain but has a linker of 22 residues between TM2 and DHp. This linker region forms a helix, possibly continuous from TM2 to Dα1. The two helices of the dimer associate to form a left-handed coiled-coil on top of the four-helix bundle. There is a kink in Dα1 induced by P265, which is near the phosphorylation site H260 and is well conserved among HKs. This kink is common among all DHp structures, even those without a proline, such as Spo0B and DesK (Figure 5).

2.4. ATP-binding domain

In prototypical sensor histidine kinases, the ATP-binding domain is located at the carboxyl terminus of the polypeptide (Figure 1). The ATP-binding domain is often abbreviated as CA for catalytic and ATP-binding. This domain binds ATP, which then donates its γ-phosphate group to the conserved histidine residue on the DHp domain. Sequences of the CA domains have well conserved N, G1, F, G2, and G3 box sequence motifs (Figure 4). Together with the H box at the DHp domain, they are the markers to annotate histidine kinases from DNA sequences.

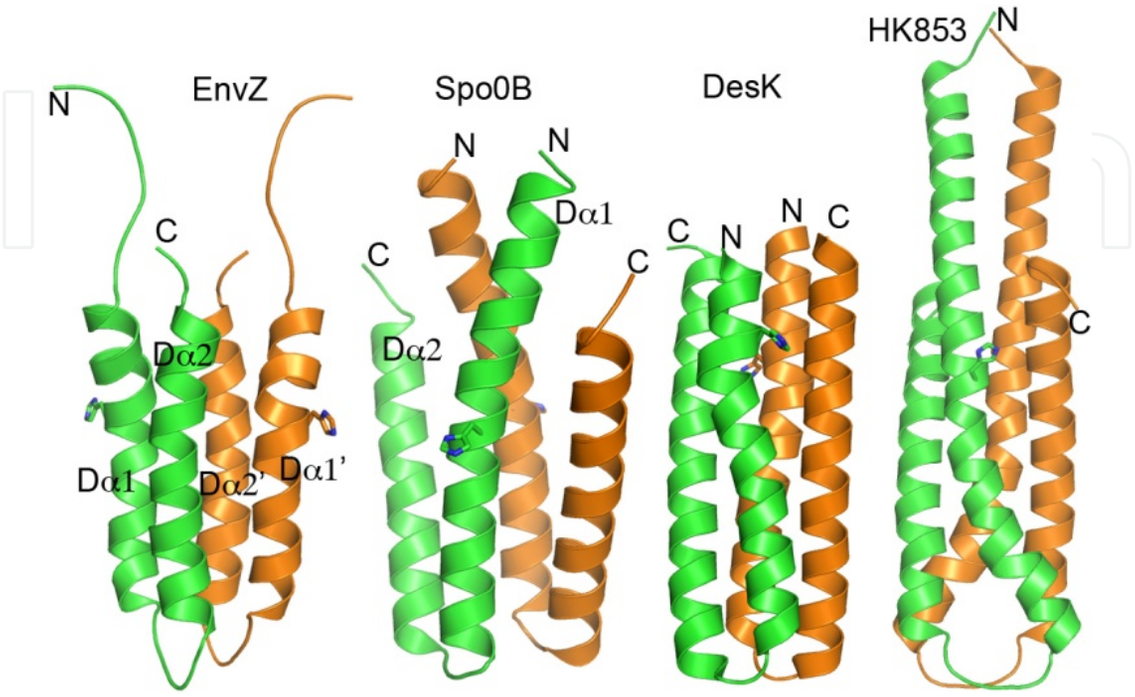

Figure 5. Structures of the DHp domains. All structures show an antiparallel helical bundle. The phosphorylation site histidine is shown as sticks. This histidine residue is located in the middle of Dα1. There is a kink of the helix right below the histidine residue, produced by a strongly conserved proline residue in EnvZ (PDB ID 1JOY) and HK853 (2C2A). DesK (3EHF) and Spo0B (1IXM) do not have this proline residue but still have a kink right below the histidine residue. Structures of DHp from Spo0B, DesK, and HK853 are a part of larger structures containing a C-terminal domain.

The structures of the CA domains are well conserved, as expected from their highly conserved sequences. The structures have an αβ sandwich fold containing two layers, a layer of mixed 5-stranded β-sheet and a layer of three α-helices (Figure 6). The CA domain of DesK from *B. subtilis* has the shortest sequence and represents the minimal core structure for this domain [22]. Most of other structures have additional α-helices, β-strands, and a longer ATP-lid (the loop covering the ATP-binding site). The ATP-binding site in DesK is much shallower due to the lack of additional structural elements and a shorter ATP-lid. Interestingly, an intact ATP molecule binds in the ATP-binding site with full occupancy, even though the phosphate groups are partially exposed to the protein surface. It is likely that in the crystal, the protein is in a conformation that shields any nucleophilic water from getting access to the phosphates, thereby protecting the ATP from hydrolysis during crystal growth. In this case, phosphotransfer will occur only in the presence of the DHp domain with the conserved histidine functioning as a nucleophile to transfer the γ-phosphate to the histidine.

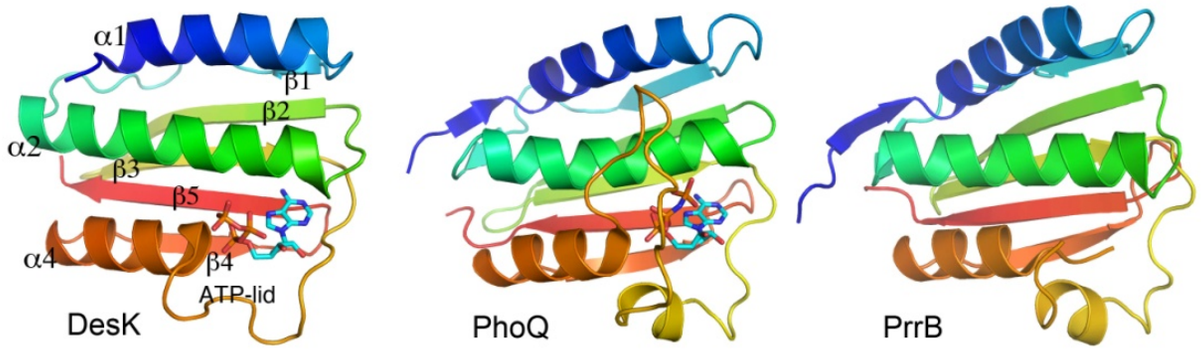

Figure 6. Ribbon diagrams of the ATP-binding domains. Each domain is colored in a rainbow spectrum with blue for the N-terminus and red for the C-terminus. Bound ATP in DesK (PDB code 3EHG) and AMPPNP in PhoQ (1ID0) are shown as sticks. The structure of PrrB (1YS3) from *M. tuberculosis* does not have a ligand bound, and part of the ATP-lid is disordered. Helices and β-strands are labeled on the structure of DesK CA. The short helix α3 in the ATP-lid is missing in DesK.

The ATP-binding site is at one end of the domain (Figure 6) and involves both absolutely conserved and partially conserved residues from the N, G1, F, G2, and G3 boxes (Figure 4). Residues from the G1 and G3 boxes are involved in binding the adenosine moiety; those from the N and G2 boxes contact both the adenosine moiety and the triphosphates-Mg^{2+} moiety. The structure of *E. coli* PhoQ in complex with AMPPNP reveals detailed binding interactions of the conserved boxes with ATP [23]. In the G1 box, the conserved Asp has a direct H-bond to the amino group and a water-mediated H-bond to N1 of adenine. In the N box, the side chains of conserved N389 and partially conserved N385 interact with the α-phosphate; the side chain oxygen of conserved N389 has a water-mediated H-bond to N1 of adenine; the partially conserved K392 and Y393 interact with the γ-phosphate through their side chains. The side chain of Y393 stacks with the adenine ring, while that of I420 from G1 box flanks the other face of the adenine ring. The side chains of partially conserved N385 (N box) and Q442 (G2 box) coordinate the divalent cation, which binds to the triphosphate group. DesK has similar binding interactions with the adenosine moiety, but slightly different binding interactions with the phosphate groups, although both proteins involve residues of the N and G2 boxes for binding the phosphate groups.

A long loop covering the ATP-binding site is called ATP-lid, which spans from the F box to the beginning of helix α4 and plays an important role in binding ATP. The ATP-lid is highly mobile; it has several glycine residues, including the conserved G2 box that contains three glycine residues (Figure 4). The bound ATP has extensive contacts with the ATP-lid residues, and thus ATP binding induces the closure of the ATP-lid. In the absence of ATP, this loop is partially disordered in crystal structures. Even in the presence of ATP, the ATP-lid shows high flexibility, indicated by high B-factors or in some cases is partially disordered in crystal structures [21-24]. The flexibility of the ATP-lid is important not only for binding ATP, but also for interacting with the DHp domain for the phosphotransfer reactions. The ATP-binding domain must adopt several positions relative to the DHp domain because the HK functions as an autokinase, phosphotransferase, or phosphatase in response to the environmental stimulates received. The flexibility of the ATP-lid allows the CA domain to

bind to different regions of DHp depending on the conformation of the DHp domain and the status of the ATP-binding site of the CA domain.

2.5. Signal transduction mechanism from sensor domain to the kinase domain

Signals flow from the sensor domain to the kinase domain by conformational changes throughout the entire protein. Environmental cues received by the sensor domain cause a conformational change in the sensor domain. That conformational change is transmitted through the transmembrane helices and the HAMP domain to the kinase domain. Most of isolated sensor domains are monomers in solution. However, they are expected to form dimers in the context of the whole protein because the DHp and HAMP domains are homodimers. Therefore, ligand-induced dimerization cannot be a mechanism of signal transduction through the cell membrane. Instead, there are structural changes of the sensor domains upon ligand binding that could trigger rotation or piston-like translation movements of the TM helices [9, 25].

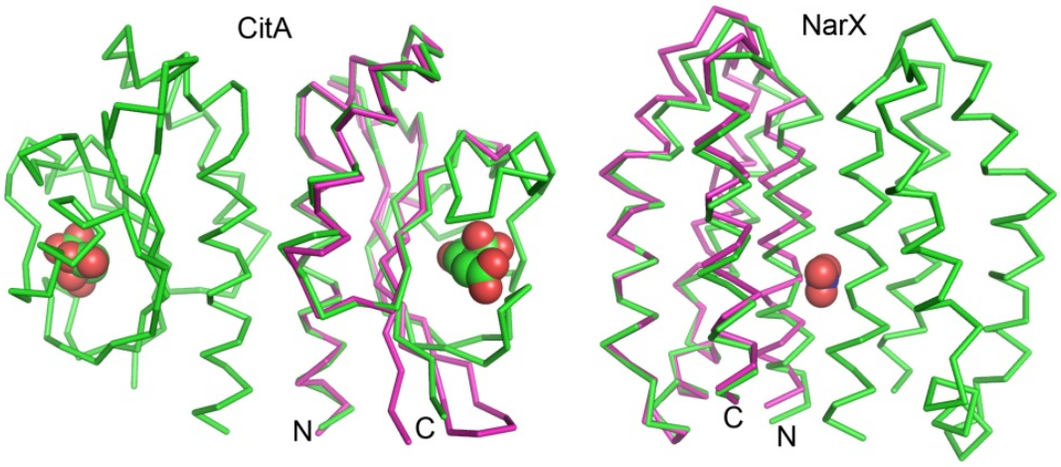

Figure 7. Piston model of ligand-induced conformational changes in the sensor domain. In both CitA (left panel, Pdb codes 2J80 and 2V9A) and NarX (right panel, 3EZH and 3EZI), ligand-bound structures are dimers in the crystal (green molecules), while ligand free forms are monomers (magenta color). Ligands are represented as space-filling models.

Comparison of structures of the sensor domain of CitA with and without citrate bound suggests a piston-type transmembrane signaling [8]. The sensor domain of CitA has a central 5-stranded anti-parallel β-sheet flanked by α-helices on both sides (Figure 2). The β-sheet curls up to one side where citrate binds. There are two loops, referred to as major loop and minor loop, which cover the citrate-binding site. The major loop is disordered in the citrate-free structure and becomes ordered in the presence of citrate in the binding site. Binding of citrate also induces more bent of the central β-sheet. This results in lifting of the C-terminus of the domain relative to the N-terminal helix (Figure 7). The C-terminus of the domain connects to the second transmembrane helix (TM2). Therefore, binding of citrate induces a piston-type movement of TM2 relative to TM1. Similar piston-type movement has been proposed for NarX from *E. coli*, whose sensor domain is a four-helix bundle that forms

a dimer when binding nitrate (Figure 7) [10]. Binding of nitrate induces the movement of the N-terminal helix toward the cell membrane, relative to the C-terminal helix. Since the N- and C-terminal helices are expected to connect to TM1 and TM2, respectively, the relative movement of the two helices would push down TM1 or lift up TM2. This movement of helices could also translate into rotations of the TM helices, if winding or unwinding of helical turns occurs.

A rotation or screw-like motion of TM2 has been proposed as the signal transduction mechanism from the sensory rhodopsin into the cytosol for HtrII [26]. Based on the structure of HtrII, the transmembrane helices of prototypical HKs are expected to form a four-helix bundle [27]. A rotation or screw-like motion of the TM helices for propagating signal through membrane is compatible with that proposed for the HAMP domain, which is often found between TM2 and the DHp domain. Structural and biochemical studies suggest that HAMP domains are not rigid structures. They can alternate between different structures through helical rotation [14] or a shift in register of coiled-coils and rotation [13]. Mutation of an alanine residue of the HAMP domain of Af1503 into larger side chain residues changes the helical packing gradually from complementary x-da to the canonical knobs-into-holes mode [17]. Through these conformational changes, the HAMP domain passes the signals from the sensor domain to the downstream kinase domain.

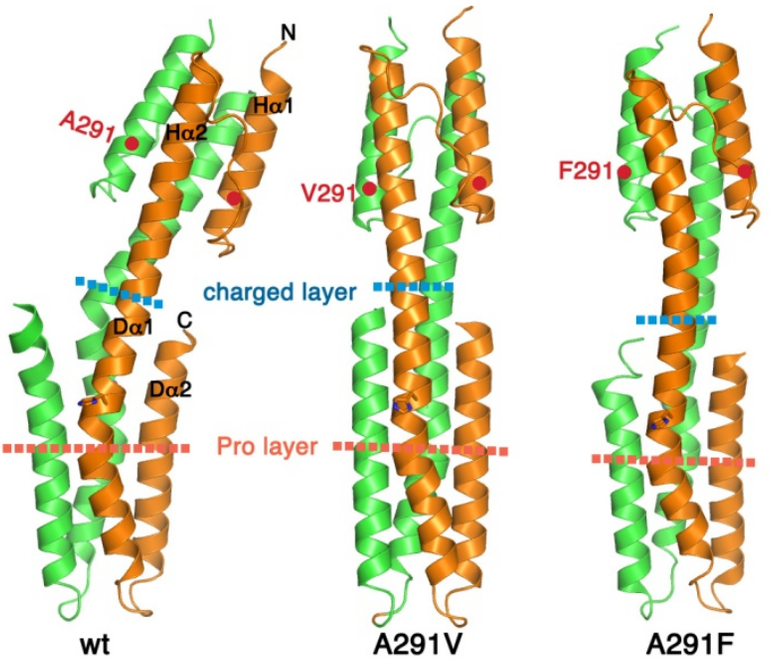

Figure 8. Crystal structures of chimeras of Af1503 HAMP fused to DHp of EnvZ. Mutations at A291 of HAMP cause conformational changes in the HAMP helical packing, which propagate to the DHp domain. The mutation sites are labeled with a red dot. Approximate positions of the charged layer and proline layer are marked. PDB codes for the structural models are 3ZRX (wt), 3ZRW (A291V), and 3ZRV (A291F).

Structures of chimeras of the Af1503 HAMP fused with the DHp domain of EnvZ show that conformational changes in the HAMP domain can trigger changes in the structure of the

downstream DHp domain [18]. Figure 8 shows three crystal structures of the chimeras. The HAMP domains preserve the same conformations as their corresponding isolated domains. Mutations of A291 into amino acids with larger hydrophobic side chains change the helical packing conformation. These conformational changes in the HAMP domain propagate to the DHp domain through the connecting helices. The helices linking the HAMP domain to the DHp domain do not have a strong hydrophobic core as in the four-helix bundles but are relatively flexible, showing variations among the structures with different mutations. Conformational variation is the most pronounced at the junction between the DHp four-helix bundle and the two-helix coiled-coil where there are charged residues (charged layer, Figure 8). The induced rotational movements in helix Dα1 gradually diminish at the conserved proline, which produces a kink in the helix. Helix Dα2 also changes conformation responding to changes of Dα1, and as a result the bundle radius differs above the proline layer among variant structures.

The structural changes in DHp induced by the HAMP domain are likely to alter the activities by regulating interactions between the DHp and CA domains. Despite the structural changes triggered by varying conformations in HAMP, the phosphorylation site histidine remains little perturbed, and thus altering accessibility of the conserved histidine is unlikely to be the mechanism of signal transmission. The lower half of the DHp helical bundle (below the proline layer) remains invariant among the structures. This region contains many residues important for binding the cognate RR and contributes the majority of the interface for RR binding in the HK-RR complex structures [19, 28]. Therefore, blocking direct binding of RR is unlikely to be a mechanism for signaling either. However, the region of DHp most variable among the structures (charged layer) plays an important role in interacting with the CA domain. In the structure of HK853-RR468 complex, residues F428 and L444 of the ATP-lid interact with an exposed hydrophobic core of DHp near the charged layer, and this interaction locks the ATP-binding domain in a kinase-inactive state [28]. It is likely that conformational changes triggered by signals transmitted from the HAMP domain either promote this DHp-CA interaction (phosphatase state) or release the CA domain for it to phosphorylate the conserved histidine (kinase state).

Several structures of the entire kinase domain (DHp + CA) are available and give insights into the interactions between the two sub-domains. The crystal structure of the entire cytoplasmic portion of HK853 from *T. maritima* shows a conformational state that is ready for the phosphotransfer reaction [21]. The HK853 cytoplasmic domain forms a dimer through the DHp domain (Figure 9a). The CA domain is connected to the DHp domain through a short linker. The crystal structure contains an ADPβN, from hydrolyzed AMPPNP, in the CA domain. The ATP-lid is flexible and partially disordered, despite the presence of the ADP analog in the ATP-binding site. The H260 side chain is fully exposed to the surface of the protein. There is a sulfate ion in the crystal structure bound next to the Nε atom of H260, mimicking the phosphorylated histidine. Contacts between DHp and CA are exclusively within one polypeptide chain and involve mostly conserved hydrophobic residues. The importance of the interface residues in the HK function is confirmed by site-directed mutagenesis studies [21]. In the DHp domain, the interface involves both α helices

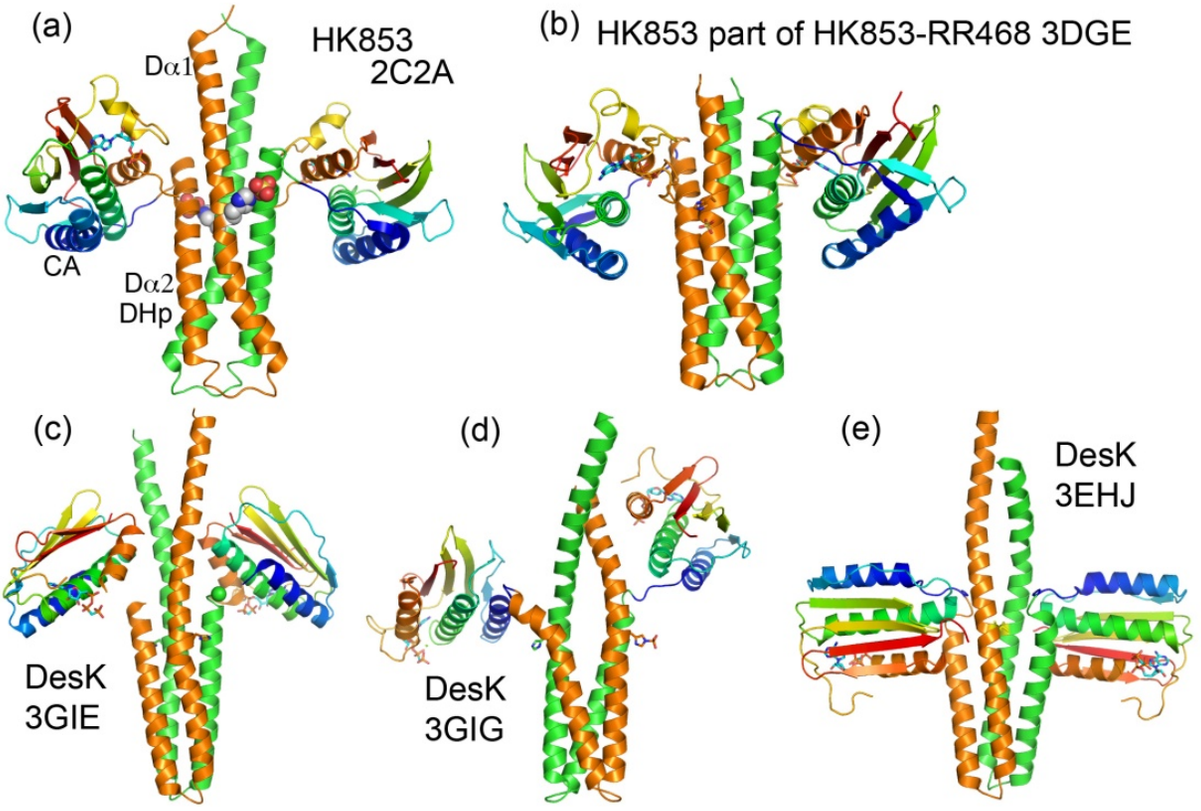

Figure 9. Structures of the entire kinase domain containing the DHp and CA subdomains. DHp domains are colored in orange for one subunit and green for the other. CA domains are colored in a rainbow spectrum. ATP analogs and the essential histidine side chains are shown as sticks except in (a) where the H260 side chains and a sulfate are shown as space-filling models. The hinge pivot residue G243 is shown as a green sphere in (c).

near the charged layer (Figures 8 and 9), and in the CA domain it involves the ATP-lid, helix α4, and residues at N-terminus of α2. The interactions are relatively weak, and thus the domain interface is dynamic, allowing the CA domain to move to other locations on DHp. The signals from the sensor domain control the interfacial stability through conformational changes of the helical bundle. The DHp-CA interface is destabilized for the kinase activity and stabilized for the phosphatase activity.

The crystal structure of HK853 complexed with RR468 reveals another binding interface between DHp and CA that is created by conformational changes in the helical packing [28]. In this structure, the CA domain moves closer to H260 of the same subunit (Figure 9b). This suggests that the autophosphorylation occurs in *cis*, although in the crystal structure, the ATP-lid interacts with the bound RR and blocks the access of the phosphorylation site histidine to ATP. The structure has a sulfate near each phosphoacceptor aspartate of RR, mimicking the phosphorylated RR. Therefore, the structure represents a conformational state poised for dephosphorylation of the RR. This phosphatase-competent state has a stronger domain interface between DHp and CA. A rotation of the helix Dα1 N-terminal coiled-coil region and a concerted inverse twist of Dα2, as in the proposed gearbox model for transmitting signals through HAMP domains (Figure 3), expose parts of the

hydrophobic core of DHp near the charged layer. This allows residues F428 and L444 of ATP-lid to interact with the exposed hydrophobic pocket of DHp of the same subunit, and thus allows the CA domain to move to a new binding site.

Structural plasticity of the HK proteins is essential to allow conformational changes in order to transmit signals received at the sensor domain to the kinase domain and switch its enzymatic activities. This structural plasticity is well demonstrated by the structures of the cytoplasmic domain of the *B. substilis* DesK, a thermo-sensing histidine kinase [24]. The DHp domain is highly mobile, adopting different conformations to allow the CA domain to bind at different positions in different functional states. The structures reveal three distinct conformational states with different helical packing of the DHp domain and the relative position of the CA domain. One conformational state represents the autokinase-competent state. This conformational state is characterized by the lack of specific interactions between the DHp and CA domains, allowing the CA domain to move to a position suitable for autophosphorylation. A single rigid-body rotation of the CA domain along the interdomain hinge residue G243 can bring the ATP γ-phosphate close to the side chain of the phosphorylation site histidine of the other protomer (Figure 9c). In addition, the contact surfaces are complementary to each other in both contour and charges. The structures are consistent with a *trans* phosphorylation mechanism. That the autokinase-competent state lacks specific interdomain interactions is consistent with observations that isolated DHp domains can be phosphorylated in the presence of isolated CA domains. A second conformational state is poised for the phosphotransfer reaction, represented by structures of a phosphorylated DesK and a phosphomimetic H188E mutant. Both structures show a similar asymmetric homodimer. One of the CA domains interacts with both Dα1 helices at the coiled-coil region. This interaction precludes the binding of the other CA domain, which makes no contacts with the DHp domain in the phosphorylated DesK structure (Figure 9d), but is completely disordered in the H188E-ADP structure. The helical bending at the phosphorylation site histidine is more pronounced partially due to the negatively charged phosphohistidine associating with conserved basic residues of the opposite protomer. The third conformational state is a phosphatase-competent state, inferred from the structures of the H188V mutant. This mutation abolishes the kinase activity but retains the wild-type level phosphatase activity. The V188 side chain is buried in the hydrophobic core of the four-helix bundle (Figure 9e). This removes the kink in Dα1 and allows a tight packing rigid helical bundle. The changes in the structure of DHp also expose a patch of hydrophobic surface on Dα1 near the charged layer for a tight association with the CA domain. The CA domain interacts with the same surface of DHp as seen in the structures of HK853 [21] and KinB [29].

3. Structures of response regulators

Response regulators are simpler in structure than histidine kinases. The prototypical RR has two domains: a receiver domain that accepts a phosphoryl group from a cognate HK and an effector domain that generates the outputs of the signaling events. The structure and sequence of receiver domains are well conserved. In contrast, effector domains are variable,

especially at the sequence level, reflecting their diverse output functions. The majority of RRs are transcription regulators with their effector domains as DNA-binding domains. A significant fraction of RRs have effector domains as enzymes. Many others have effector domains for binding RNA, ligands, or proteins to regulate bacterial cellular process at post-transcriptional and post-translational levels. There are also single domain RRs that have only the receiver domains such as CheY and Spo0F. RRs have been studied extensively for decades, and an extensive amount of structural data is available, as well as many well-written reviews [30]. I will focus my efforts on transcription regulator RRs and will give a summary on how phosphorylation of the receiver domain controls the activities of the effector domain with respect to their structure-function relationship.

3.1. Receiver domain structures

Receiver domains have well-conserved sequences and structures, suggesting a common signaling mechanism by two component systems to pass signals from HKs to their cognate RRs. The receiver domain is also known as the phosphorylation domain or regulatory domain. There is an invariable aspartate residue in the receiver domain that accepts a phosphoryl group from a cognate HK. Phosphorylation results in conformational changes in the domain, which is transmitted to the effector domain to regulate its activities. Receiver domains have a conserved $(\beta\alpha)_5$ fold (Figures 10 and 11), in which alternating β-strands and α-helices in the sequence fold into a central five-stranded parallel β-sheet surrounded by helices α1 and α5 on one side and helices α2, α3, and α4 on the other side. The β-sheet consists of mainly hydrophobic side chains that form a hydrophobic core. The α-helices are amphipathic and pack against the central β-sheet with their hydrophobic face. Most of the conserved residues are at the C-terminal ends of strands β1, β3, and β4. Helix α1 is involved in binding to the DHp domain of HK [28]. Although the structure and position of this helix is well conserved among RR structures, the amino acid sequence conservation is limited to the hydrophobic residues involved in packing against the hydrophobic core of the β-sheet (Figure 10). This helix is likely to play an important role in the specificity of HK-RR pairs.

The phosphorylation site is located at an acidic pocket near the C-termini of strands β1 and β3, where the conserved acidic side chains are clustered together. These acidic side chains are involved in binding a divalent cation, Mg^{2+} or Mn^{2+}, that is essential for the phosphorylation reaction [31]. An absolutely conserved aspartate at the end of the strand β3 is the phosphoacceptor. The loop β3-α3 forms a conserved structure that is held together by conserved residues: a Pro at 4 residues after the phosphorylation site aspartate and a Gly, which is at 4 residues after the Pro in sequence and the first residue of helix α3 (Figure 10).

The α4-β5-α5 face of the receiver domain has been proposed to be important for the function of RRs [32, 33]. Many RRs in the OmpR/PhoB subfamily are thought to be dimers in their active form. Structures of several RR receiver domains are found to form dimers through the α4-β5-α5 face when activated. Exposed side chains on this surface are primarily hydrophilic, suggesting that any interactions through this α4-β5-α5 face are likely to be dynamic. Comparison among receiver domain structures reveals that this face is considerably

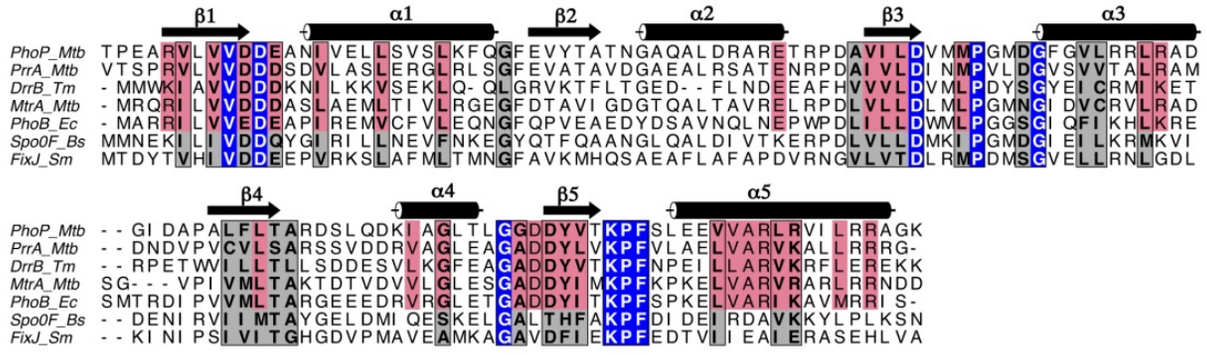

Figure 10. Sequence alignment of the response regulator receiver domains. The top five sequences, PhoP, PrrA and MtrA from *M. tuberculosis*, DrrB from *T. maritima*, and PhoB from *E. coli*, belong to the OmpR/PhoB subfamily. Spo0F from *B. subtilis* is a single domain RR. FixJ of *S. Meliloti* is in the NarL/FixJ subfamily. Identical residues for all sequences are shaded in blue. Highly conserved residues in the OmpR/PhoB subfamily are shaded in red. Moderately conserved residues across all sequences are shaded in gray.

variable. The sequence of helix α5 is well conserved among OmpR/PhoB subfamily RRs (Figure 10). Yet the length of this helix varies among known structures [34]. In the structure of PrrA from *M. tuberculosis* this helix has different lengths between two molecules in the crystal asymmetric unit [35]. Helix α4 and its flanking loops also have variable structures among RRs. Flexibility of the α4-β5-α5 elements is likely to be important for phosphorylation regulation of RR activities.

Activation of RRs by phosphorylation induces structural changes in the α4-β5-α5 face. This is accomplished by movements of two key residues, a conserved Thr/Ser at the C-terminus of β4 and a moderately conserved Tyr/Phe in the middle of β5 (Figure 11). These residues are referred to as switch residues. In the unphosphorylated RR, the side chain of Thr/Ser is oriented away from the phosphoacceptor, and that of Tyr/Phe extends outward toward the surface of the α4-β5-α5 face. Structures of activated receiver domains reveal that these side chains move in response to phosphorylation [36, 37]. All activated receiver domains of the OmpR/PhoB subfamily form a symmetric dimer through the same interface involving α4-β5-α5. Phosphorylation of the phosphoacceptor aspartate residue allows the side chain of the Thr/Ser residue to move closer to the phosphate group to make a favorable hydrogen bond. Repositioning the Thr/Ser side chain shifts the β4-α4 loop and helix α4, and makes the inward conformation of the Tyr/Phe residue more energetically favorable. Another conserved residue, a Lys in the β5-α5 loop, forms a salt bridge to the phosphoryl group and brings some modest shift in the β5-α5 loop [31]. This Lys side chain has a salt bridge to the side chain of phosphoacceptor aspartate in the unphosphorylated structure PhoP from *M. tuberculosis* [34]. It plays an important role in stabilizing the active site structure, but probably has only minor effects on structural switch between activated and non-activated RRs. Phosphorylation of NarL/FixJ subfamily RRs produces similar conformational changes in switch residues and results in dimerization through the α4-β5-α5 face, although the dimer interface are not identical to that of OmpR/PhoB RRs (Figure 11d). The difference in dimer interface might be due to lack of divalent cation in the crystal structure of FixJ [38].

Activation by BeF_3 and Mg^{2+} did not induce symmetric dimer formation of single domain response regulator RR468 [28]. However, the switch residues are in the active conformation as observed for all other RRs, suggesting that conformational changes of switch residues is the common mechanism for all RRs.

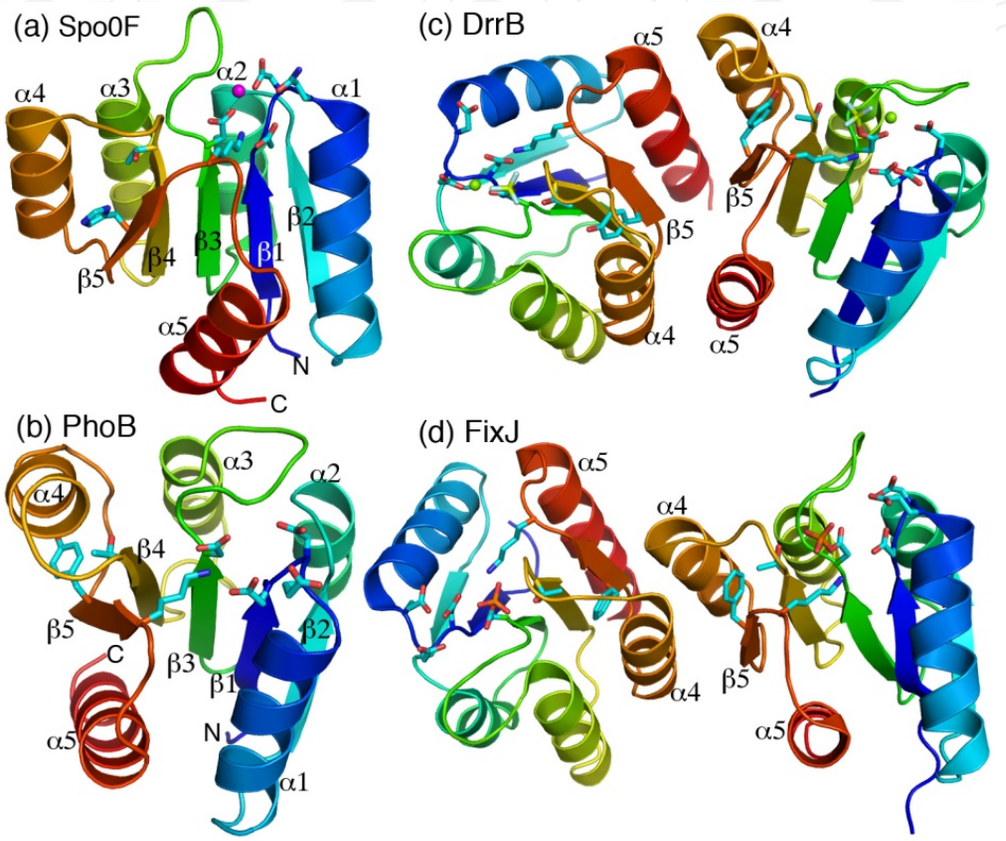

Figure 11. Structures of the receiver domains of response regulators. Each individual domain is colored from N- to C- terminus in a rainbow spectrum. Phosphorylation site residues and switch residues are shown as sticks. Spo0F (PDB code 1PEY) from *B. subtilis* binds a Mn^{2+} ion, shown in magenta. Both PhoB (1B00) from *E. coli* and Spo0B are non-activated. DrrB (3NNS) from *T. maritima* is activated with binding of BeF_3 and Mg^{2+}, shown as sticks and a green sphere, respectively. FixJ (1D5W) from *S. Meliloti* is activated by phosphorylation, but without divalent cations.

3.2. Structures of the effector domains

Unlike the receiver domains, the effector domains are very diverse, reflecting the diversity of cellular processes that are controlled by two-component systems. Effector domains are also referred to as output domains. They can be DNA-binding, RNA-binding, ligand-binding, or transporter output domains, or enzymes [30]. Except in archaea, which has almost 50% of RRs containing only the receiver domain, a great majority of bacterial RRs are transcription regulators that contain a C-terminal effector domain as a DNA-binding domain.

DNA-binding effector domains are grouped into several subfamilies based on predicted domain structures from amino acid sequences [30]. The largest subfamily is the OmpR/PhoB group, whose structure has a winged helix-turn-helix (wHTH) motif. The second largest subfamily is the NarL/FixJ group, which has a helix-turn-helix (HTH) DNA-binding motif. Another subfamily LytR/AgrA is also fairly common. A structure of a member of this subfamily, LytTR from *S. aureus*, shows that the DNA-binding domain consists mostly of β-strands [39]. The loops between the β-strands interact with DNA bases with side chains inserting into both major and minor grooves of DNA. Other subfamilies of DNA-binding response regulators are less common and are present only in certain types of bacterial species.

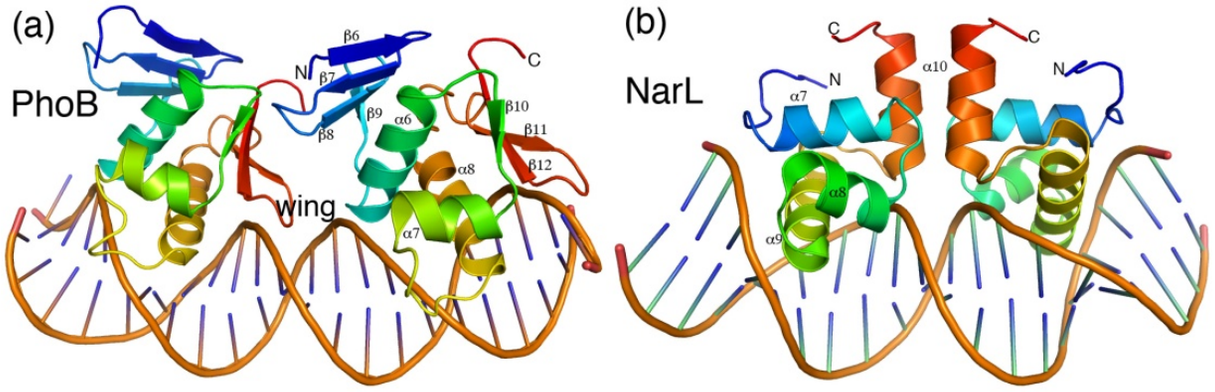

Figure 12. Structures of response regulator effector domains binding to DNA. (a) PhoB (1GXP) of *E. coli* binds to direct repeat Pho Box DNA as a tandem dimer. The PhoB effector domain has a winged HTH fold. (b) NarL (1JE8) of *E. coli* recognizes palindromic DNA sequences and forms a symmetric dimer through association of helix α10 on binding DNA.

The DNA-binding domain of the OmpR/PhoB subfamily has a conserved winged helix-turn-helix fold (Figure 12a). The structure starts with an N-terminal four-stranded anti-parallel β-sheet (β6 to β9), followed by three α-helices (α6 to α8) in the middle, and a β-hairpin (β11 and β12) at the C-terminus. A short strand (β10) between α6 and α7 assembles with the C-terminal β-hairpin to form a three-stranded anti-parallel β-sheet. The two β-sheets sandwich the three α-helices in the middle. The β-hairpin is called the wing, and α7, α8 and loop α7-α8 together constitute the helix-turn-helix motif. Helix α8 is the sequence recognition helix, which inserts into the major groove of DNA [40]. The β-hairpin (wing) binds in the minor groove of DNA. Loop α7-α8 is termed transaction loop because it is essential for transcription activation.

The C-terminal DNA-binding domain of the NarL/FixJ subfamily is a compact bundle of 4 α-helices (Figure 12). The second and third helices (α8 and α9) form a helix-turn-helix motif that interacts with the major groove of DNA [41, 42]. Helix α9 is the recognition helix that inserts into major groove of DNA with side chains interacting with DNA bases. The proteins bind DNA as a symmetric dimer to recognize palindromic DNA repeats. Helix α10 and the loop α7-α8 form the protein dimer interface in the protein-DNA complex. Isolated DNA-binding domains, however, are found to be monomeric in solution when not binding to DNA [42].

3.3. Structures of full-length response regulators and regulation mechanism

Full-length structures are available for the OmpR/PhoB and NarL/FixJ subfamilies of response regulators. Even though most RRs have only two domains, full-length structures prove to be difficult to obtain, especially for the OmpR/PhoB subfamily. This is most likely due to the dynamic nature of interactions between the receiver and effector domains. The domains are individually folded, and their associations are transient if they do associate with each other. The dynamic and transient nature of domain interactions is important for the regulation of DNA-binding activities. Signals transmitted from the HK through phosphorylation of the receiver domains must be able to relay to the DNA-binding domain by conformational changes of the receiver domain. Phosphorylation promotes dimerization of the receiver domain and in some cases releases the blockage of the DNA-binding element of the receiver domain. So far, all available structures of full-length RRs are non-activated and without DNA bound.

The full-length NarL structure in its unphosphorylated form reveals that the receiver domain is associated with the DNA-binding domain through an interface involving the DNA-binding elements [43]. The linker between the two domains forms a helix (α6) followed by a flexible loop that is disordered in the crystal structure (Figure 13). Helix α6 packs against α4 of the receiver domain and α10 of the effector domain. The domain interface also involves helices α8 and α9, locking the DNA-sequence-recognition helix α9 in a manner incapable of inserting into the DNA major groove. Phosphorylation is expected to dissociate the effector domain from the receiver domain, thus allowing it to bind DNA. Unphosphorylated NarL does not bind DNA in solution. NarL binds to inverted-repeat DNA sequences as a symmetric dimer, whose interface involves helix α10 and loop α7-α8 (Figure 12). Phosphorylation may also promote dimerization of the receiver domain and thus strengthens the NarL dimer.

Full-length structures of unphosphorylated StyR and DosR, which are members of the NarL/FixJ subfamily RRs, are also known (Figure 13). The receiver domain of StyR from *Pseudomonas fluorescens* has a conserved $(\beta\alpha)_5$ fold [44]. Its DNA-binding domain has the same fold as that of NarL. However, the interdomain linker forms a long helix continuous from α5 of the receiver domain that separates the N- and C-terminal domains by more than 16 Å, exposing the helix-turn-helix motif for binding DNA in the unphosphorylated state. Phosphorylation increases DNA-binding affinity by ~10-fold, presumably through dimerization of the receiver domain that brings two effector domains next to each other for binding to DNA inverted repeats. DosR from *M. tuberculosis* forms a dimer with a unique structure [45]. Its receiver domain has a $(\beta\alpha)_4$ fold instead of the $(\beta\alpha)_5$ fold observed for all other RRs (Figure 13b), even though the conserved residues in the β5 to α5 region are present in the DosR sequence. The sequence for the canonical β5 to α5 region forms a long helix, pairing with a helix of linker residues to form an antiparallel coiled-coil (Figure 13b). These two helices form an antiparallel four-helix bundle with the same helices from another subunit, accounting for a major part of the dimer interface. Helices α9 and α10 have extensive interactions with the receiver domain, locking the structure in an inactive conformation. Activation by phosphorylation presumably causes significant structural

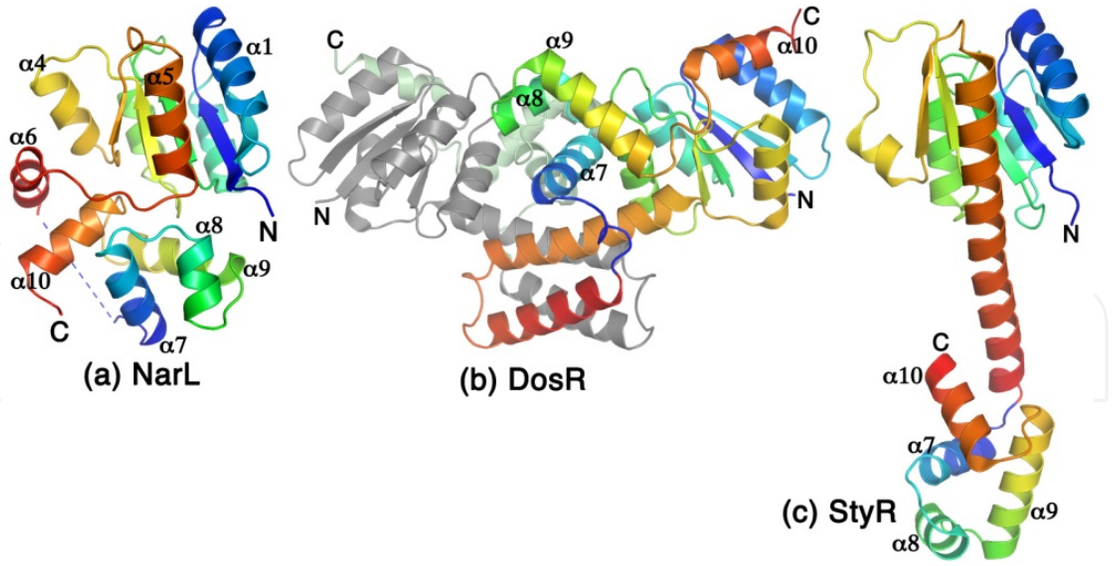

Figure 13. Full-length structures of the NarL/FixJ subfamily RRs. The receiver domains are colored in a rainbow spectrum from N-terminus in blue to the linker in red; the effector domains are separately colored in another rainbow spectrum from blue (α7) to red (α10). (a) NarL (PDB code 1RNL) has a tight packing between domains that blocks sequence recognition helix α9. Part of the linker that is disordered is depicted as a dashed line. (b) DosR (3C3W) forms a dimer, and one of the subunit is colored in gray for the receiver domain and the linker, and in pale green for the effector domain. (c) StyR (1ZN2) has a long helix α5 continuous into the linker that separates the two domains apart.

rearrangements that allows the receiver domain to form a dimer of $(\beta\alpha)_5$ fold and release the effector domain to bind DNA inverted repeats [46].

OmpR/PhoB subfamily RRs have been extensively studied in recent years. In addition to many isolated receiver and effector domain structures, there are six full-length RR structures available, DrrB [47] and DrrD [48] from *T. maritima*, and RegX3 [49], PrrA [35], MtrA [50], and PhoP [34] from *M. tuberculosis*. All structures are in the unphosphorylated forms. The interdomain interfaces are different among these structures. RegX3 forms a domain-swapped dimer in the crystal structure, exchanging the α4-β5-α5 unit between subunits. It is not clear if this domain-swapped dimer is physiologically relevant. However, consistent with the structures of the isolated receiver domains, the α4-β5-α5 unit has high mobility among the structures of full-length RRs, indicating the importance of the α4-β5-α5 face in regulating RR activities. All structures with interactions between the receiver domain and effector domain involve the α4-β5-α5 face in the interdomain interface (Figure 14).

The structures of DrrD and DrrB both reveal monomers in crystal and have interactions between domains that do not preclude DNA binding (Figure 14). The interdomain interactions are relatively weak, suggesting that in solution the domains could be free and thus phosphorylation activation mechanism is likely through dimerization of the N-terminal receiver domain. The structures of PrrA and MtrA show more extensive interdomain interactions that involve the α4-β5-α5 face and block the recognition helix. However, the interfaces are polar in nature, and the open conformation is likely to exist in solution for both proteins. All these structural data suggest that phosphorylation is likely to shift the equilibrium toward dimerization as a mechanism to activate the effector domain.

The structure of PhoP from *M. tuberculosis* supports this mechanism of activation through phosphorylation-induced dimerization of the receiver domain [34]. PhoP forms a dimer through the α4-β5-α5 face of the receiver domain (Figure 14). The effector domain is merely tethered to the receiver domain through a flexible linker, which is disordered in the crystal structure. This open conformation of unphosphorylated PhoP structure is consistent with observations that unphosphorylated PhoP binds DNA [51]. Because there is no direct interaction between the receiver domain and the effector domain, it is likely that phosphorylation regulates the DNA-binding activity by shifting the equilibrium toward dimer formation of the receiver domain [33, 44]. Although the PhoP dimer interface involves the α4-β5-α5 face, the dimer packing is less compact than the activated receiver dimers. This is because the switch residues are in the non-activated conformation, and hence the loops at the dimer interface are different in conformation from the activated structures [34]. By shifting the equilibrium toward dimer formation of the receiver domains, phosphorylation brings two DNA-binding domains in close proximity; the flexible domain linker allows the DNA-binding domains to bind two direct repeats of DNA sequence while the receiver domains adopt a symmetric dimer. The OmpR/PhoB subfamily RRs recognize DNA direct-repeat sequences as revealed in the PhoB effector domain structure in complex with DNA (Figure 12a). However, the isolated DNA-binding domain of PhoP is a monomer in solution and does not form any dimers in the crystal structure [52]. Recognition of DNA direct repeats can only be realized through dimerization of the receiver domain. Bringing two DNA-binding domains in close proximity by dimerization of the receiver domains can also increase DNA-binding affinity.

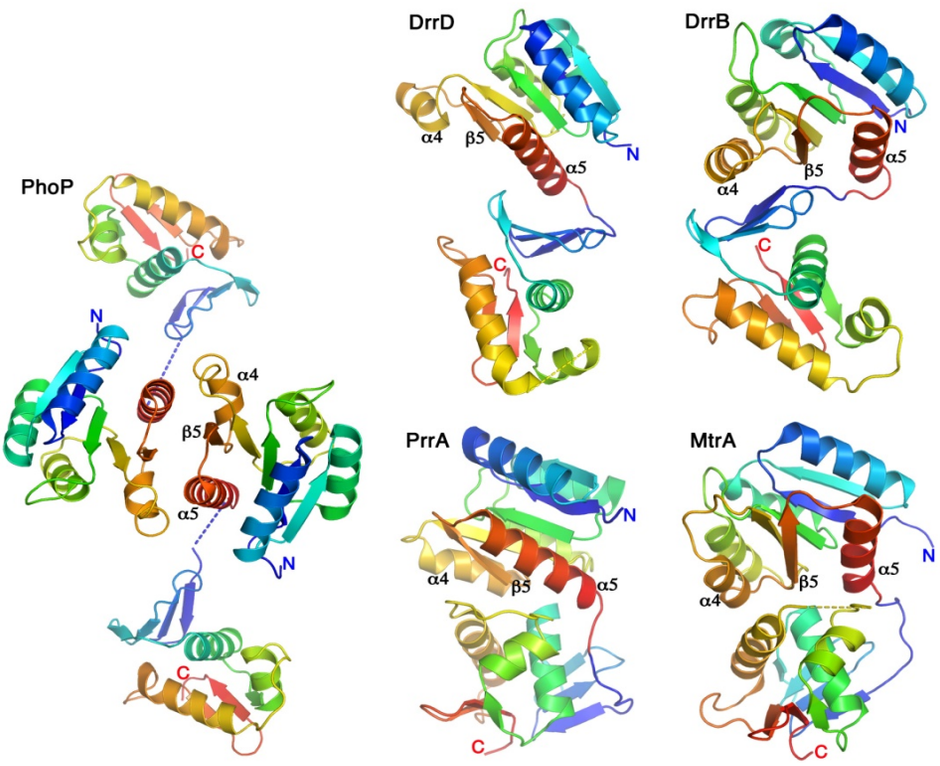

Figure 14. Full-length structures of the OmpR/PhoB subfamily RRs. Each domain is individually colored from N- to C-terminus in a rainbow spectrum. PhoP, PrrA, and MtrA are from *M. tuberculosis*; DrrD and DrrB are from *T. maritima*. The PDB codes are: PhoP, 3R0J; DrrD, 1KGS; DrrB, 1P2F; PrrA, 1YS6; MtrA, 2GWR.

4. Interactions between histidine kinases and response regulators

To complete the signaling events of two-component systems, histidine kinases, when activated, must associate with their cognate response regulators and transfer the phosphoryl group from the phosphorylated histidine to a conserved aspartate residue of RRs. Evidence suggests that HKs are also involved in dephosphorylation of RRs to turn off the signal cascade. The interactions between HK and RR must be specific because there is little cross-talk among TCS proteins. Yet, these interactions must be dynamic, as both proteins must adopt various conformations throughout the steps of the signaling event.

Several crystal structures of HK in complex with RR have been reported that offer insights into how HK and RR interact. A structure of the complex between HK853 and RR468 from *T. maritima* shows the two proteins engaged in a conformation for the phosphatase reaction [28]. The structure contains a dimer of the HK853 entire cytoplasmic portion and two RR468 molecules (Figure 15).

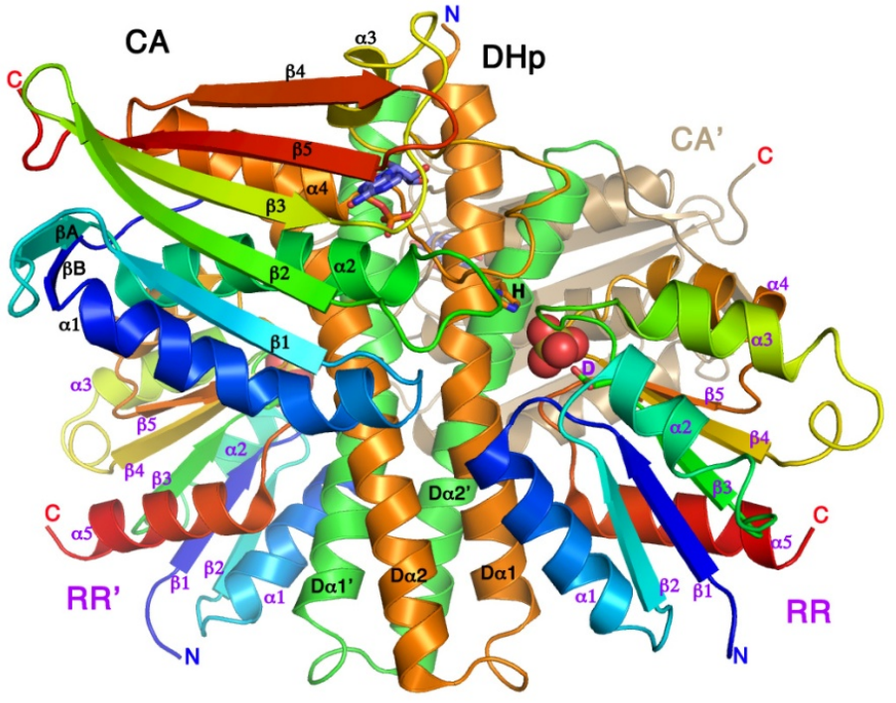

Figure 15. Structure of the HK853-RR468 complex (PDB code 3DGE). The DHp domain is colored in orange for one subunit and in green for the other. One CA domain (front) is colored in a rainbow spectrum; the other CA (back) is in sand color. Two RR molecules each individually colored in a rainbow spectrum. Secondary structural elements of HK are labeled in black, and those of RR are in purple. A sulfate between the essential histidine of HK and aspartate of RR is shown in a space-filling model.

The RR binds to the DHp domain on the helical hairpin below the phosphorylation site histidine. Helix α1 of RR interacts with both helices of the helical hairpin through hydrophobic side chains to form a six-helix bundle. Each α1 interacts only with helices of the same subunit of DHp. Loop β5-α5 also makes contacts with the DHp domain. In the

crystal structure, there is a sulfate ion bound between the phosphoacceptor aspartate of RR and the phosphorylation site histidine of DHp. Binding of the sulfate mimics phosphorylation and allows the switch residues of RR to adopt the active conformation. Consequently, both loops β3-α3 and β4-α4 are also in the active conformation. These two loops interact with the ATP-lid and loop α2-β2 of the CA domain. The interactions between HK853 and RR468 would preclude dimer formation of RR through the α4-β5-α5 face, despite that the RR is in an active conformation. RR468 is a single domain protein. The structure of activated RR468 with BeF_3^- shows changes in conformation of switch residues and loops β3-α3 and β4-α4, but it is a monomer [28]. This is different from OmpR/PhoB subfamily RRs, whose activated form promotes dimerization, or NarL/FixJ family RRs, whose activated form releases the effector domain for binding DNA and might also promote dimerization. Despite these differences, binding interactions of α1 of RR with the helical hairpin of DHp is likely conserved among TCS proteins. The structure of HK853-RR468 complex is in a phosphatase-competent state because the RR is phosphorylated and the ATP-lid blocks access of the γ-phosphate of ATP by the histidine side chain. The structure indicates that the phosphorylation site histidine could play a role in the phosphatase activity by orienting a water molecule as a nucleophile.

A complex structure of ThkA/TrrA, also from *T. maritima*, has a similar binding site on HK for RR [53], and the structure also represents a phosphatase-competent state because the CA domain is locked in a position away from the phosphorylation site histidine. Similar to RR468, TrrA is also a single domain RR. The resolution of the complex structure is relatively low at 3.8 Å, and high-resolution structures of individual domains were fit into the low-resolution electron density map of the ThkA-TrrA complex. TrrA has weak electron density, indicating its high mobility in the structure. The complex puts the phosphoacceptor aspartate of TrrA facing the phosphorylation site histidine of ThkA. However, TrrA has weaker interactions with DHp. Most of interactions involve the loops around the phosphoacceptor and the N-terminal end of α1 of TrrA with Dα1 of ThkA. Unlike that of RR468, helix α1 of TrrA does not have exposed hydrophobic side chains to have coiled-coil interactions with the four-helix bundle of ThkA. The binding position of RR on DHp is similar to that of the HK853-RR468 complex, suggesting that they share a common mechanism for the phosphotransfer reaction.

A structure of the Spo0F-Spo0B complex is related to HK-RR complexes in binding interactions and the mechanism of the catalyzed reactions. Spo0B is a phosphotransferase having a DHp domain similar to that of HKs, while Spo0F is a single domain RR [54, 55]. Spo0F is phosphorylated on its phosphoacceptor aspartate by sensor HK KinA. The phosphoryl group is then transferred to the phosphorylation site histidine on spo0B, which then transfers the phosphoryl group to the response regulator Spo0A. In the structure of the Spo0F-Spo0B complex, Spo0F binds to the four-helix bundle of Spo0B in a similar manner as the HK853-RR468 complex. Helix α1 of Spo0F packs against the helical hairpin region and makes up majority of the interface with DHp, confirming that helix α1 and the lower end of the DHp four-helix bundle as the binding interface between HK and RR.

5. Therapeutic potential against bacterial pathogens

Antibiotics resistance has been a major medical problem of modern medicine. Bacterial pathogens evolve mechanisms to become resistant to antibiotics, and strains resistant to multiple drugs or the last line of antibiotics, such as methicillin-resistant *S. aureus* and vancomycin-resistant *Enterococcus* have evolved and become a major challenge for modern medicine. Emergence of multi-drug resistant (MDR) and extremely drug resistant (XDR) tuberculosis (TB) has prompted the World Health Organization to declare TB as a global emergency.

Two component systems are major signaling proteins in bacteria. They are involved in every aspect of bacterial adaptation to their environments, including sensing the existence of antibiotics and regulating cellular responses to become drug resistant. Therefore, TCSs are potential drug targets for developing new antibiotics [56]. Unlike conventional antibiotics that directly target the proteins involved in essential cellular activities, drugs inhibiting TCSs target the upstream functions that regulate these essential proteins. Thus anti-TCS drugs work in a manner different from conventional drugs and are likely to be effective against drug-resistant bacterial pathogens. Because TCS proteins are absent in animals, drugs targeting TCSs can potentially have less toxicity.

Recent advancements in structural and functional characterization of TCSs have identified several potential targets for developing new antibiotics [56]. A PhoP-PhoQ system from *S. typhimurium* is a major virulence factor that is essential for this intracellular pathogen to establish infection. Inhibitors to the kinase activity of PhoQ were studied for their interactions with the protein and the inhibition mechanism by NMR and crystallography methods [57]. The structures can serve as a platform for rational design to improve the efficacy of the lead compounds. A PhoP-PhoR system in *M. tuberculosis* was found to be an important virulence factor [58]. PhoP has a point mutation in the avirulent H37Ra strain, and this mutation plays an important role in the loss of virulence of the avirulent strain [59]. Inhibitors targeting PhoP are likely to reduce the virulence of *M. tuberculosis*. With the availability of the PhoP structure [34], it is possible to do virtual screening or rational inhibitor design to search for compounds that will disrupt phosphorylation activation by interacting with the loops near the phosphoacceptor or with helix α1, or binding to the α4-β5-α5 face to prevent dimerization. Compounds inhibiting the effector domain binding DNA will also be efficient inhibitors. To develop inhibitors against PhoP-PhoR TCS, targeting the sensor domain of PhoR will be more specific and has the advantage that the compounds do not have to cross the cell membrane [58]. However, structural information of the PhoR protein is currently lacking.

Author details

Shuishu Wang
Department of Biochemistry and Molecular Biology, Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA

Acknowledgement

The author thanks Dr. Chou-Zen Giam, Dr. Linda Miallau, and Dr. Issar Smith for helpful comments and revisions of the manuscript. This work was supported by a grant GM079185 from the National Institute of Health of the United States.

6. References

[1] Aravind L, Ponting CP. The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins. FEMS Microbiol Lett. 1999;176(1):111-6.

[2] Hefti MH, Francoijs KJ, de Vries SC, Dixon R, Vervoort J. The PAS fold. A redefinition of the PAS domain based upon structural prediction. Eur J Biochem. 2004;271(6):1198-208.

[3] Martinez SE, Beavo JA, Hol WG. GAF domains: two-billion-year-old molecular switches that bind cyclic nucleotides. Mol Interv. 2002;2(5):317-23.

[4] Wayne LG, Sohaskey CD. Nonreplicating persistence of mycobacterium tuberculosis. Annu Rev Microbiol. 2001;55:139-63.

[5] Cho HY, Cho HJ, Kim YM, Oh JI, Kang BS. Structural insight into the heme-based redox sensing by DosS from Mycobacterium tuberculosis. J Biol Chem. 2009;284(19):13057-67.

[6] Podust LM, Ioanoviciu A, Ortiz de Montellano PR. 2.3 A X-ray structure of the heme-bound GAF domain of sensory histidine kinase DosT of Mycobacterium tuberculosis. Biochemistry. 2008;47(47):12523-31.

[7] Miyatake H, Mukai M, Park SY, Adachi S, Tamura K, Nakamura H, et al. Sensory mechanism of oxygen sensor FixL from Rhizobium meliloti: crystallographic, mutagenesis and resonance Raman spectroscopic studies. J Mol Biol. 2000;301(2):415-31.

[8] Sevvana M, Vijayan V, Zweckstetter M, Reinelt S, Madden DR, Herbst-Irmer R, et al. A ligand-induced switch in the periplasmic domain of sensor histidine kinase CitA. J Mol Biol. 2008;377(2):512-23.

[9] Cheung J, Hendrickson WA. Sensor domains of two-component regulatory systems. Curr Opin Microbiol. 2010;13(2):116-23.

[10] Cheung J, Hendrickson WA. Structural analysis of ligand stimulation of the histidine kinase NarX. Structure. 2009;17(2):190-201.

[11] Moore JO, Hendrickson WA. Structural analysis of sensor domains from the TMAO-responsive histidine kinase receptor TorS. Structure. 2009;17(9):1195-204.

[12] Jing X, Jaw J, Robinson HH, Schubot FD. Crystal structure and oligomeric state of the RetS signaling kinase sensory domain. Proteins. 2010;78(7):1631-40.

[13] Airola MV, Watts KJ, Bilwes AM, Crane BR. Structure of concatenated HAMP domains provides a mechanism for signal transduction. Structure. 2010;18(4):436-48.

[14] Hulko M, Berndt F, Gruber M, Linder JU, Truffault V, Schultz A, et al. The HAMP domain structure implies helix rotation in transmembrane signaling. Cell. 2006;126(5):929-40.

[15] Appleman JA, Chen LL, Stewart V. Probing conservation of HAMP linker structure and signal transduction mechanism through analysis of hybrid sensor kinases. J Bacteriol. 2003;185(16):4872-82.
[16] Zhu Y, Inouye M. Analysis of the role of the EnvZ linker region in signal transduction using a chimeric Tar/EnvZ receptor protein, Tez1. J Biol Chem. 2003;278(25):22812-9.
[17] Ferris HU, Dunin-Horkawicz S, Mondejar LG, Hulko M, Hantke K, Martin J, et al. The mechanisms of HAMP-mediated signaling in transmembrane receptors. Structure. 2011;19(3):378-85.
[18] Ferris HU, Dunin-Horkawicz S, Hornig N, Hulko M, Martin J, Schultz JE, et al. Mechanism of regulation of receptor histidine kinases. Structure. 2012;20(1):56-66.
[19] Tomomori C, Tanaka T, Dutta R, Park H, Saha SK, Zhu Y, et al. Solution structure of the homodimeric core domain of Escherichia coli histidine kinase EnvZ. Nat Struct Biol. 1999;6(8):729-34.
[20] Varughese KI, Madhusudan, Zhou XZ, Whiteley JM, Hoch JA. Formation of a novel four-helix bundle and molecular recognition sites by dimerization of a response regulator phosphotransferase. Mol Cell. 1998;2(4):485-93.
[21] Marina A, Waldburger CD, Hendrickson WA. Structure of the entire cytoplasmic portion of a sensor histidine-kinase protein. Embo J. 2005;24(24):4247-59.
[22] Trajtenberg F, Grana M, Ruetalo N, Botti H, Buschiazzo A. Structural and enzymatic insights into the ATP binding and autophosphorylation mechanism of a sensor histidine kinase. J Biol Chem. 2010;285(32):24892-903.
[23] Marina A, Mott C, Auyzenberg A, Hendrickson WA, Waldburger CD. Structural and mutational analysis of the PhoQ histidine kinase catalytic domain. Insight into the reaction mechanism. J Biol Chem. 2001;276(44):41182-90.
[24] Albanesi D, Martin M, Trajtenberg F, Mansilla MC, Haouz A, Alzari PM, et al. Structural plasticity and catalysis regulation of a thermosensor histidine kinase. Proc Natl Acad Sci U S A. 2009;106(38):16185-90.
[25] Zhang Z, Hendrickson WA. Structural characterization of the predominant family of histidine kinase sensor domains. J Mol Biol. 2010;400(3):335-53.
[26] Gordeliy VI, Labahn J, Moukhametzianov R, Efremov R, Granzin J, Schlesinger R, et al. Molecular basis of transmembrane signalling by sensory rhodopsin II-transducer complex. Nature. 2002;419(6906):484-7.
[27] Goldberg SD, Clinthorne GD, Goulian M, DeGrado WF. Transmembrane polar interactions are required for signaling in the Escherichia coli sensor kinase PhoQ. Proc Natl Acad Sci U S A. 2010;107(18):8141-6.
[28] Casino P, Rubio V, Marina A. Structural insight into partner specificity and phosphoryl transfer in two-component signal transduction. Cell. 2009;139(2):325-36.
[29] Bick MJ, Lamour V, Rajashankar KR, Gordiyenko Y, Robinson CV, Darst SA. How to switch off a histidine kinase: crystal structure of Geobacillus stearothermophilus KinB with the inhibitor Sda. J Mol Biol. 2009;386(1):163-77.
[30] Galperin MY. Diversity of structure and function of response regulator output domains. Curr Opin Microbiol. 2010;13(2):150-9.

[31] Bourret RB. Receiver domain structure and function in response regulator proteins. Curr Opin Microbiol. 2010;13(2):142-9.
[32] Gao R, Stock AM. Molecular strategies for phosphorylation-mediated regulation of response regulator activity. Curr Opin Microbiol. 2010;13(2):160-7.
[33] Barbieri CM, Mack TR, Robinson VL, Miller MT, Stock AM. Regulation of response regulator autophosphorylation through interdomain contacts. J Biol Chem. 2010;285(42):32325-35.
[34] Menon S, Wang S. Structure of the response regulator PhoP from Mycobacterium tuberculosis reveals a dimer through the receiver domain. Biochemistry. 2011;50(26):5948-57.
[35] Nowak E, Panjikar S, Konarev P, Svergun DI, Tucker PA. The structural basis of signal transduction for the response regulator PrrA from Mycobacterium tuberculosis. J Biol Chem. 2006;281(14):9659-66.
[36] Bachhawat P, Stock AM. Crystal structures of the receiver domain of the response regulator PhoP from Escherichia coli in the absence and presence of the phosphoryl analog beryllofluoride. J Bacteriol. 2007;189(16):5987-95.
[37] Toro-Roman A, Wu T, Stock AM. A common dimerization interface in bacterial response regulators KdpE and TorR. Protein Sci. 2005;14(12):3077-88.
[38] Birck C, Mourey L, Gouet P, Fabry B, Schumacher J, Rousseau P, et al. Conformational changes induced by phosphorylation of the FixJ receiver domain. Structure. 1999;7(12):1505-15.
[39] Sidote DJ, Barbieri CM, Wu T, Stock AM. Structure of the Staphylococcus aureus AgrA LytTR domain bound to DNA reveals a beta fold with an unusual mode of binding. Structure. 2008;16(5):727-35.
[40] Blanco AG, Sola M, Gomis-Ruth FX, Coll M. Tandem DNA recognition by PhoB, a two-component signal transduction transcriptional activator. Structure. 2002;10(5):701-13.
[41] Maris AE, Kaczor-Grzeskowiak M, Ma Z, Kopka ML, Gunsalus RP, Dickerson RE. Primary and secondary modes of DNA recognition by the NarL two-component response regulator. Biochemistry. 2005;44(44):14538-52.
[42] Maris AE, Sawaya MR, Kaczor-Grzeskowiak M, Jarvis MR, Bearson SM, Kopka ML, et al. Dimerization allows DNA target site recognition by the NarL response regulator. Nat Struct Biol. 2002;9(10):771-8.
[43] Baikalov I, Schroder I, Kaczor-Grzeskowiak M, Grzeskowiak K, Gunsalus RP, Dickerson RE. Structure of the Escherichia coli response regulator NarL. Biochemistry. 1996;35(34):11053-61.
[44] Milani M, Leoni L, Rampioni G, Zennaro E, Ascenzi P, Bolognesi M. An active-like structure in the unphosphorylated StyR response regulator suggests a phosphorylation-dependent allosteric activation mechanism. Structure. 2005;13(9):1289-97.
[45] Wisedchaisri G, Wu M, Sherman DR, Hol WG. Crystal structures of the response regulator DosR from Mycobacterium tuberculosis suggest a helix rearrangement mechanism for phosphorylation activation. J Mol Biol. 2008;378(1):227-42.

[46] Wisedchaisri G, Wu M, Rice AE, Roberts DM, Sherman DR, Hol WG. Structures of Mycobacterium tuberculosis DosR and DosR-DNA complex involved in gene activation during adaptation to hypoxic latency. J Mol Biol. 2005;354(3):630-41.
[47] Robinson VL, Wu T, Stock AM. Structural analysis of the domain interface in DrrB, a response regulator of the OmpR/PhoB subfamily. J Bacteriol. 2003;185(14):4186-94.
[48] Buckler DR, Zhou Y, Stock AM. Evidence of intradomain and interdomain flexibility in an OmpR/PhoB homolog from Thermotoga maritima. Structure (Camb). 2002;10(2):153-64.
[49] King-Scott J, Nowak E, Mylonas E, Panjikar S, Roessle M, Svergun DI, et al. The structure of a full-length response regulator from Mycobacterium tuberculosis in a stabilized three-dimensional domain-swapped, activated state. J Biol Chem. 2007;282(52):37717-29.
[50] Friedland N, Mack TR, Yu M, Hung LW, Terwilliger TC, Waldo GS, et al. Domain orientation in the inactive response regulator Mycobacterium tuberculosis MtrA provides a barrier to activation. Biochemistry. 2007;46(23):6733-43.
[51] Gupta S, Sinha A, Sarkar D. Transcriptional autoregulation by Mycobacterium tuberculosis PhoP involves recognition of novel direct repeat sequences in the regulatory region of the promoter. FEBS Lett. 2006;580(22):5328-38.
[52] Wang S, Engohang-Ndong J, Smith I. Structure of the DNA-binding domain of the response regulator PhoP from Mycobacterium tuberculosis. Biochemistry. 2007;46(51):14751-61.
[53] Yamada S, Sugimoto H, Kobayashi M, Ohno A, Nakamura H, Shiro Y. Structure of PAS-linked histidine kinase and the response regulator complex. Structure. 2009;17(10):1333-44.
[54] Zapf J, Sen U, Madhusudan, Hoch JA, Varughese KI. A transient interaction between two phosphorelay proteins trapped in a crystal lattice reveals the mechanism of molecular recognition and phosphotransfer in signal transduction. Structure. 2000;8(8):851-62.
[55] Varughese KI, Tsigelny I, Zhao H. The crystal structure of beryllofluoride Spo0F in complex with the phosphotransferase Spo0B represents a phosphotransfer pretransition state. J Bacteriol. 2006;188(13):4970-7.
[56] Gotoh Y, Eguchi Y, Watanabe T, Okamoto S, Doi A, Utsumi R. Two-component signal transduction as potential drug targets in pathogenic bacteria. Curr Opin Microbiol. 2010;13(2):232-9.
[57] Guarnieri MT, Zhang L, Shen J, Zhao R. The Hsp90 inhibitor radicicol interacts with the ATP-binding pocket of bacterial sensor kinase PhoQ. J Mol Biol. 2008;379(1):82-93.
[58] Ryndak M, Wang S, Smith I. PhoP, a key player in Mycobacterium tuberculosis virulence. Trends Microbiol. 2008;16(11):528-34.
[59] Frigui W, Bottai D, Majlessi L, Monot M, Josselin E, Brodin P, et al. Control of M. tuberculosis ESAT-6 secretion and specific T cell recognition by PhoP. PLoS pathogens. 2008;4(2):e33.