1. Introduction

Bacteria monitor and adapt to varying environmental stimuli by numerous signal transduction systems like chemotaxis sensory systems (Park *et al.*, 2006), anti-sigma or sigma factor pairs (Ravio and Silhavy, 1999), Ser/Thr- Kinases (Kenney, 1997) and so called two-component systems (TCSs) (Stock *et al.*, 2000). Signal transduction systems regulate the majority of cellular activities including the metabolism, development, host-recognition, biofilm production, virulence, and antibiotic resistance of human pathogens. Currently the “Microbial Signal Transduction database” (MiST2) contains 966 complete and 157 draft bacterial and archaeal genomes, which collectively contain more than 245,000 signal transduction proteins (Ulrich and Zhulin, 2010).

1.1. Two-component systems

Due to their highly conserved structures, TCSs are easily to identify in genomic databases and are present in more than 95% of the bacterial genomes sequenced to date, as well as, archaea and eukaryotic organisms such as plants and fungi. In the model for Gram-negative bacteria *Escherichia coli* (*E. coli*), 26 functional TCSs have been reported (KEGG database).

A prototypical TCS is composed of a membrane-integrated histidine kinase (HK, also known as sensor kinase) and a cytoplasmic response regulator (RR) (Figure 1). The HK senses stimuli and is capable of autophosphorylation in response to environmental input signals to mediate the output responses. The HK, which is phosphorylated on a conserved histidine residue, transfers the phosphoryl group to the aspartic acid on the receiver domain (REC) of the RR (Gao and Stock, 2009). Phosphorylation induces a conformational change in the receiver domain that results in activation of the effector domain (also called output domain) that initials the responses. The domains of two-component proteins are modular and can be integrated into proteins and pathways in a variety of ways, but the core structures and activities are maintained (Stock *et al.*, 2000).

The functional state of these two components is controlled by three phosphotransfer reactions (Figure 1): (1) the autophosphorylation of a conserved histidine in the kinase core of the HK, (2) the phosphotransfer to a conserved aspartate in the REC of the RR, and (3) dephosphorylation of the RR to set the system back to the prestimulus state (Parkinson, 1993; Stock *et al.*, 1995). The phosphatase can be an intrinsic property (autophosphatase) of the RR or a phosphoprotein phosphatase activity of the kinase towards the regulator (Mascher *et al.*, 2006).

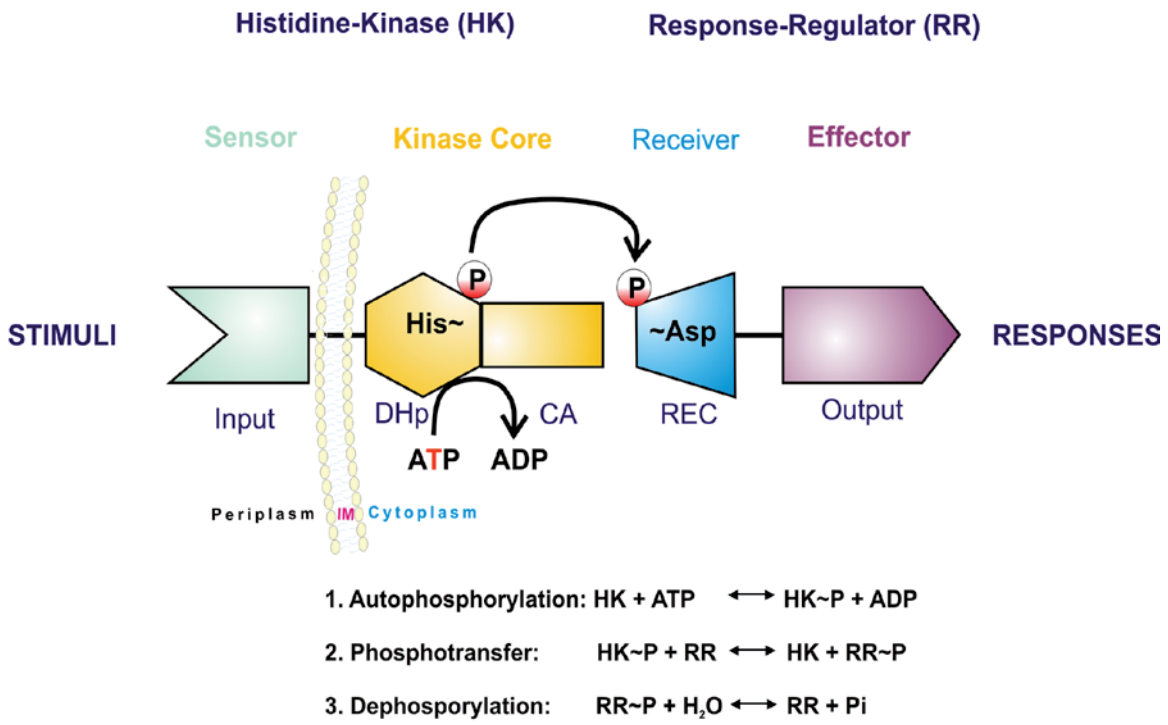

Figure 1 Modular design of a prototypic two-component system (TCS). A typical HK consists of a variable N-terminal sensor domain (also called input domain) connected to a conserved C-terminal kinase core that catalyzes phosphorylation. The kinase core is composed of the DHp (dimerization and histidine phosphotransfer) domain with the phosphoaccepting His and the CA (catalytic and ATP binding) domain. A typical RR consists of a conserved N-terminal receiver domain (REC) and a variable C-terminal output domain.

1.1.1. Structural insight into the functions of histidine kinases

HKs sense the stimulus and mediate the phosphotransfer signalling pathways to allow cells to adapt to the environment (Gao and Stock, 2009). HKs contain a variable sensor domain and a conserved kinase core (Figure 1).

The sensor domain of most membrane integral HKs is formed by an extracellular loop integrated between two membrane domains (Mascher *et al.*, 2006) and detects a specific stimulus (Cheung and Hendrickson, 2010). However, the mechanistic details of signal recognition and transmission for most stimuli are unclear. Recent crystallographic studies revealed that sensor domains comprise only a small number of discrete structural classes despite their extreme sequence diversity (Cheung and Hendrickson, 2010). For extracellular sensor domains three structural classes have been classified (Figure 2): (1) mixed alpha-beta folds also named PDC domain followed by the initial three extracellular sensor domains structures first determined from PhoQ (Cheung *et al.*, 2008; Cho *et al.*, 2006), DcuS (Cheung and Hendrickson, 2008; Pappalardo *et al.*, 2003), and CitA (Reinelt *et al.*, 2003; Sevvana *et al.*, 2008) which sense divalent ions, certain C_4-dicarboxylates, and citrate, respectively; (2) all alpha folds like NarX (Cheung and Hendrickson, 2009) and TorS (Moore and Hendrickson, 2009), and (3) periplasmic binding proteins (PBP) like folds exemplified by HK29s (Cheung *et al.*, 2009).

For the PDC domain class sub-domain folds have been classified. These include PAS (sensory input domain named after eukaryotic Per, ARNT and SIM proteins), GAF (small ligand binding sensory input domain found in some cGMP-regulated phosphodiesterases, adenlyl cyclases, and transcription factor FhlA), CACHE (Found in Ca^{2+} channels and chemotaxis receptors), CHASE (Cyclase/His kinase-associated sensing extracellular) (Mascher *et al.*, 2006) and PHY (phytochrome) domains (Essen *et al.*, 2008; Yang *et al.*, 2008).

However, it is not clear whether the structure of a signal domain correlates with ligand selectivity. Thus, further structural and functional studies are needed to decrypt the mechanism of signal recognition in TCSs in general.

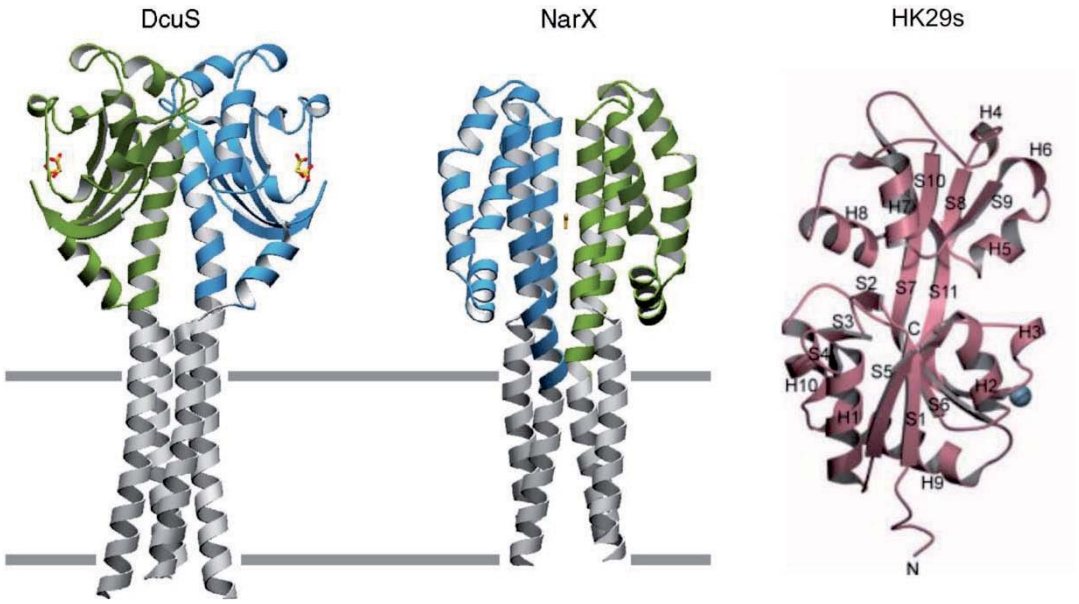

Figure 2 Classified structural classes of extracellular sensor domains. Dimeric sensor domain with modeled transmembrane helices (from Cheung and Hendrickson, 2010) Ribbon diagrams of PDC sensor DcuS and all-alpha sensor NarX. Dimeric sensor domains are modeled with membrane domains. Sensor domains are drawn as ribbon model in blue and green and the modeled membrane domains are coloured in grey non-protein moieties malate and nitrate are shown in ball-and-stick. **Structure of HK29s** (from Cheung *et al.*, 2010) Secondary structure elements are identified with strand (S) and helix (H) labels. The blue sphere shows the location of a single bound sodium ion.

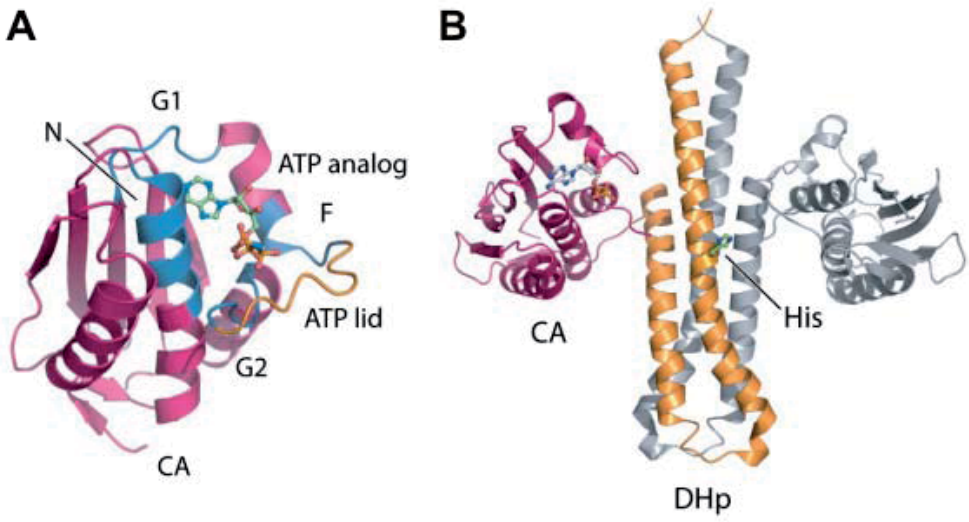

Figure 3 Structure of kinase core (from Gao and Stock, 2009)**. A.** Crystal structure of the CA domain of *Escherichia coli (E. coli)* PhoQ (PDB ID: 1ID0). Homology boxes crucial for ATP binding are shown in blue. A flexible ATP lid (orange) covers the bound ATP analog, AMPPNP. **B.** The entire kinase core of *Thermotoga maritime* HK853 (PDB ID: 2C2A). HK853 is dimeric with one monomer shown in orange and pink and the other in gray.

The sensor domain is linked to the kinase core by a transmitter domain immediately following the membrane domain. A classical transmitter domain is the HAMP linker which is a helical domain found in HKs, adenylyl cyclases, methyl-accepting chemotaxis proteins, and phosphatases (Aravind and Ponting, 1999; Hulko *et al.*, 2006). The catalytic active kinase core of a HK is composed of a conserved dimerization and histidine phosphotransfer (DHp) domain containing the histidine residue for phosphorylation (Figure 3B) and a highly well-conserved C-terminal catalytic (CA) domain that binds ATP (Tanaka *et al.*, 1998; Tomomori *et al.*, 1999).

It is generally accepted that the kinase core is only active as a dimer. Dimerization occurs over the DHp domain which comprises a four-helix bundle (Figure 3B and 4B). The CA domain protrudes on either side of the dimer helical stem (Marina *et al.*, 2005) and is capable of binding ATP and transfers the phosphoryl group *in trans* to the conserved histidine in the DHp domain of the other HK (Yang and Inouye, 1991). However, the dogma of *in trans* phophorylation has recently been overthrown by biochemical studies on the *Thermotoga maritima* HK853 that clearly demonstrated for this HK that autokinase reaction proceeds by a *cis* autophosphorylation mechanism within the HK subunit (Casino *et al.*, 2009).

The α/β sandwich fold of the CA domain (Figure 3A), a highly conserved ATP binding cavity, is defined by the conserved residues in the N, G1, F, and G2 boxes. Between the F and G2 boxes, a flexible region named the ATP lid can adopt different conformations upon nucleotide binding. In structures of PhoQ with bound ATP analogs (Bilwes *et al*., 2001; Marina *et al.*, 2001), the ATP lid is relatively ordered and covers the nucleotide (Bilwes *et al.*, 1999; Song *et al.*, 2004). Therefore, the conformational changes of the ATP-lid are proposed to couple the ATP binding to alterations of interdomain or protein-protein interactions (Pearl and Prodromou, 2006).

The structural details on the LuxPQ receptor strongly suggest that in response to a stimulus, the relative orientation of helices in the HK dimerization domain can reorient, via cogwheeling (rotation) and kinking (bending) (Figure 4) (Neiditch *et al.*, 2006). This movement results in repositioning of the DHp and CA domains with respect to each other and consequently effects HK activities (reviewed in Stewart, 2010).

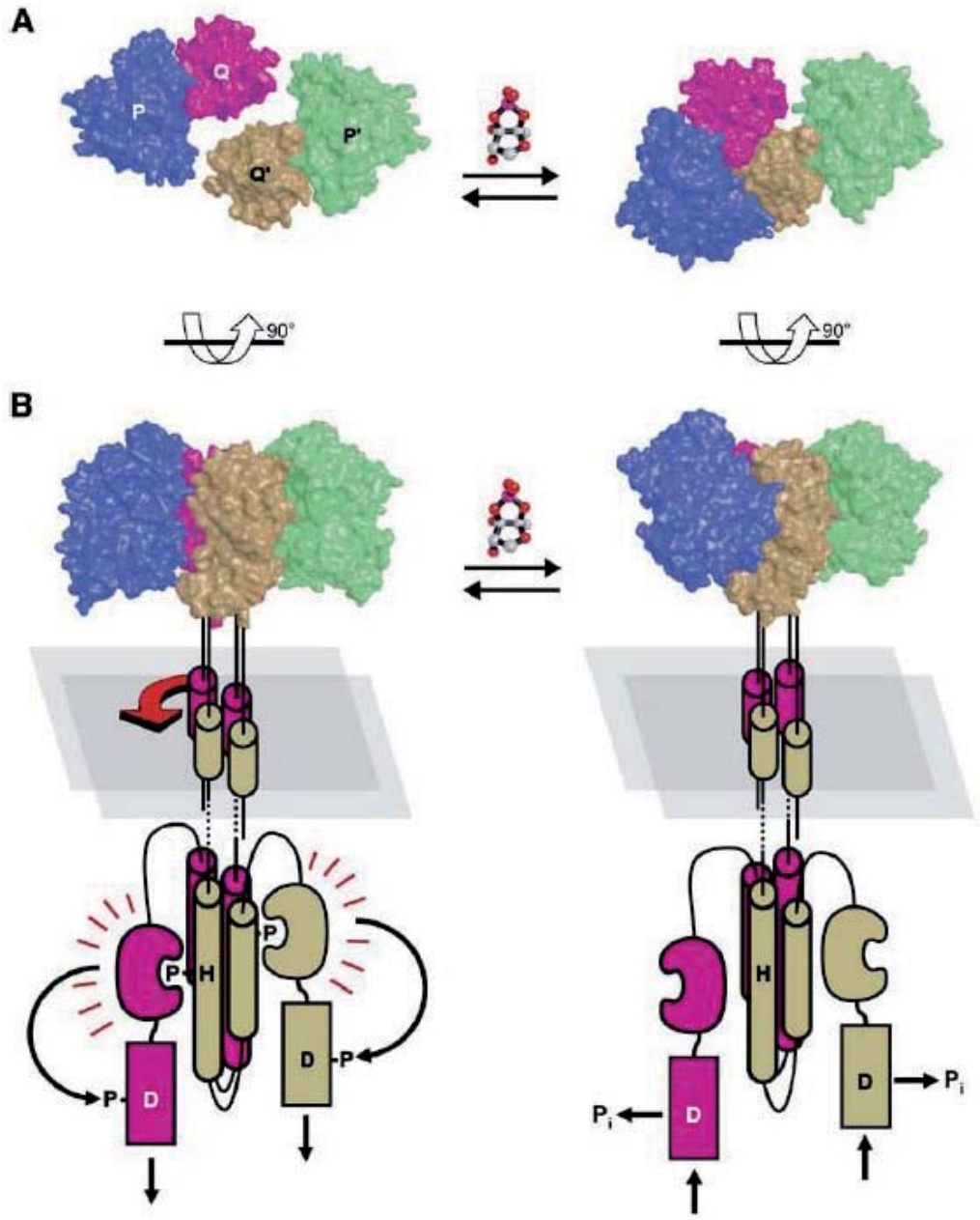

Figure 4 Proposed signal transduction mechanism for the LuxPQ receptor (from Neiditch *et al.*, 2006). The periplasmic binding protein LuxP and the sensor domain of the HK LuxQ form an open conformation in the absence of the quorum signal molecule, autoinducer 2 (AI-2) (left side), Binding of AI-2 induces a closed conformation that promots the formation of an asymmetric dimer (right side). This symmetry-breaking rotation presumably repositions the LuxQ membrane domains for signal transduction to the cytoplasmic domains. Structures are displayed from ‘‘above,’’ looking toward the bacterial inner membrane (A), or from the ‘‘side’’ (B).

1.1.2. Basic structures and functions of response regulators

Response regulators (RRs) have a modular architecture that consists of an N-terminal conserved receiver domain (REC) with conserved Asp residue and a variable effector domain which can function as regulator of protein activity, on RNA or as transcription regulator (Gao *et al.*, 2007). Michael Galperin's laboratory maintains an online database on RRs (Galperin, 2010). Here, I will focus on RRs that regulate gene transcription.

1.1.2.1. Receiver domain of response regulators

The REC domains participate in three activities: catalysis of phosphoryl transfer from phosphorylated HKs to one of their own Asp residues, catalysis of autodephosphoralation and regulation of effector domain activity in a phosphorylation-dependent manner (Gao and Stock, 2009; West and Stock, 2001).

The structure and function of the RR REC domain is highly conserved. The REC domains typically consist of a α4-β5-α5 face (Figure 5).

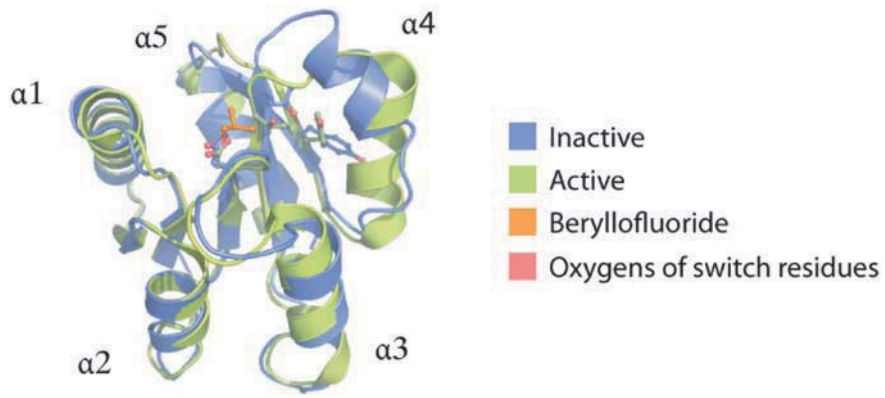

Figure 5 Conserved features of RR receiver domains (from Gao and Stock, 2009). The structures of inactive and active PhoB receiver (REC) domains are superimposed to illustrate the subtle conformational differences (PDB IDs: 1B00 and 1ZES). Switch residues are shown in ball-and-stick mode, and beryllofluoride is coloured orange.

A cluster of highly conserved acidic residues forming the active site is important in signal propagation and catalysis of phosphotransfer from HKs and auto-dephosphorylation (Stock *et al*., 2000; West and Stock, 2001). In addition to the site of phosphorylation, an Asp at the C terminal end of β3, the active site contains two additional acidic residues (Asp/Glu) in the β1-α1 loop position that coordinate an a divalent metal ion, commonly Mg^{2+}, that is required for both phosphotransfer and phosphatehydrolysis (Stock *et al.*, 1993; Gao *et al.*, 2007; Gao and Stock, 2009). Phosphorylation occurs at a conserved Asp residue, generating a high-energy acyl phosphate which provides energy coupling to drive a conformational change in the effector domain. The conserved mechanism couples phosphorylation at the active site to structural perturbations at the distal α4-β5-α5 surface (Gao and Stock, 2009). However, the precise mechanism of coupled communication between the REC domain and the effector domain is unknown, but may be mediated in part by a flexible interdomain region of variable length (Eldridge *et al*., 2002, Mattison *et al*., 2002, Walthers *et al*., 2003).

1.1.2.2. Effector domain of response regulators

The effector domain elicits the output response (reviewed in Gao and Stock, 2009). For *E. coli* K12 29 RRs regulating gene transcription have been identified and classified into four families, OmpR/PhoB (Martinez-Hackert and Stock, 1997), NarL (Baikalov *et al.*, 1996), NtrC (Lee *et al.*, 2003; Nixon *et al.*, 1986) and LytR (Nikolskaya and Galperin, 2002) (Galperin, 2006). Based on the mechanism of activation these RR-families have been further placed in two well defined classes (reviewed in Gao and Stock, 2009). In the first class, exemplified by OmpR, NarL and LytR, phosphorylation of the REC domain activates the response regulator by forming rotationally symmetric dimers that directly interact with DNA (Figure 6). The second class, represented by NtrC activates σ54 promoters and is characterized by an AAA^+ ATPase domain (Figure 6). Phosphorylation in the REC domain induces assembly of hexameric or heptameric rings competent for ATPase activity that induces open complex formation in RNA polymerases.