Exploring and understanding signal-response relationships and response dynamics of microbial two-component signaling systems.

Submitted by Munia Amin

to the University of Exeter as a thesis for the degree of
Doctor of Philosophy in Biological Sciences
December 2013

I certify that all material in this thesis which is not my own work has been identified and that no material has been previously submitted and approved for the award of a degree by this or any other University.

Munia Amin

THESIS ABSTRACT

Two-component signaling systems are found in bacteria, fungi and plants. They mediate many of the physiological responses of these organisms to their environment and display several conserved biochemical and structural features. This thesis identifies a potential functional role for two commonly found architectures in two-component signaling system, the split kinases and phosphate sink, which suggests that by enabling switch-like behaviors they could underlie physiological decision making.

I report that split histidine kinases, where autophosphorylation and phosphotransfer activities are segregated onto distinct proteins capable of complex formation, enable ultrasensitivity and bistability. By employing computer simulations and analytical approaches, I show that the specific biochemical features of split kinases "by design" enable higher nonlinearity in the system response compared to conventional two-component systems and those using bifunctional (but not split) kinases. I experimentally show that one of these requirements, namely segregation of the phosphatase activity only to the free form of one of the proteins making up the split kinase, is met in proteins isolated from *Rhodobacter sphaeroides*. While the split kinase I study from *R. sphaeroides* is specifically involved in chemotaxis, other split kinases are involved in diverse responses. Genomics studies suggest 2.3% of all chemotaxis kinases, and 2.8% of all kinases could be functioning as split kinases.

Combining theoretical and experimental approaches, I show that the phosphate sink motif found in microbial and plant TCSs allows threshold behaviors. This motif involves a single histidine kinase that can phosphotransfer reversibly to two separate response regulators and examples are found in bacteria, yeast and plants. My results show that one of the response

regulators can act as a “sink” or “buffer” that needs to be saturated before the system can generate significant responses. This sink, thereby allows the generation of a signal threshold that needs to be exceeded for there to be significant phosphoryl group flow to the other response regulator. Thus, this system can enable cells to display switch-like behavior to external signals. Using an analytical approach, I identify mathematical conditions on the system parameters that are necessary for threshold dynamics. I find these conditions to be satisfied in both of the natural systems where the system parameters have been measured. Further, by *in vitro* reconstitution of a sample system, I experimentally demonstrate threshold dynamics for a phosphate-sink containing two-component system.

This study provides a link between these architectures of TCSs and signal-response relationship, thereby enabling experimentally testable hypotheses in these diverse two-component systems. These findings indicate split kinases and phosphate as a microbial alternative for enabling ultrasensitivity and bistability - known to be crucial for cellular decision making. By demonstrating ultrasensitivity, threshold dynamics and their mechanistic basis in a common class of two-component system, this study allows a better understanding of cellular signaling in a diverse range of organisms and will open the way to the design of novel threshold systems in synthetic biology. Thus, I believe that this study will have broad implications not only for microbiologists but also systems biologists who aim to decipher conserved dynamical features of cellular networks.

Table of Contents

2 Split histidine kinases enable ultrasensitivity and bistability in two-component signaling networks

3 Phosphate sink containing two-component signaling systems as tunable threshold devices

4 General Discussion

List of Tables

List of Figures

ACKNOWLEDGEMENT

First of all, I am greatly indebted to my supervisors, Professor Orkun S. Soyer and Dr. Steven L. Porter. Throughout the years, their invaluable guidance, motivation, expertise and enthusiastic support have made this a thoughtful, rewarding and beneficial journey. They have truly been inspiration for me to become an independent researcher. I could never thought of doing mathematical modeling before starting this PhD! All credit goes to Professor Soyer to help me in understanding theoretical biology and also having confidence on me that I could be a modeller too. I am really grateful to Dr. Porter for introducing me to the techniques used in the laboratory and also his immense helps in day to day lab life. I could not have imagined having better advisors.

I consider it an honor to work in the 4th floor lab with my companions, a talented team of 'crazy' researchers. My special thanks go to my team mate Vanessa Francis and bench mate Matthew Robinson. I really appreciate Yvonne's company in the office. I am thankful to all the 4th floor people for providing the delicious cakes and cookies, which helped me to survive the long hours and weekends in the lab!

I have enjoyed the company of the people from Prof Orkun's group. Specially, I owe Dr. Francesco Montefusco my heartfelt appreciation for his kind help on parameter fitting. I also want to acknowledge Varun Kothamachu for being a great company while working on two-component system. It is been a pleasure to work with you all.

I would like to give a heartfelt thanks to our collaborator Dr. Elisenda Feliu for her kind help in deriving mathematical conditions used in chapter 3 and Dr. Birgit E Scharf for her kind support in sending necessary constructs and plasmids.

I am very grateful to the College of Life and environmental sciences for providing the stimulating training environment. I would also like to acknowledge the financial supports of Exeter Science Strategy, Systems Biology Theme, University of Exeter.

Last, but not the least, I would like to remember all my family members and friends back in Bangladesh for their continuous support, love and inspiration. My parents are the reason for whom I started this journey. My husband Muhammad is my all time support and love throughout the whole period. It might be really difficult to end this journey without his contribution in my life! I am very grateful to my brother and sisters for having faith on me always. How can I forget my precious gem of my life? He came this world during my PhD time and it was not surely that much easy, but Amr Eesaa, my dearest son you made my life complete. Despite of some troubles and noises, he has been the coolness of my tired eyes.

I would like to **dedicate** this on my parents, Mr. Ruhul Amin and Mrs. Salma Amin and my beloved grandmother Mrs. Monwara Awal, they are the people for whom I am today what I am…

ABBREVIATIONS

ATP	Adenosine-5'- triphosphate
CME	Chemical master equation
CRNT	Chemical reaction network theory
DNA	Deoxuribonucleic acid
HK	Histidine kinases
HOG	High-osmolarity glycerol
MAPK	Mitogen-activated protein kinase
ODE	Ordinary differential equation
REC	Receiver
RR	Response regulator
STYKs	Serine/threonine/tyrosine kinases
TCS	Two-component signaling
TCSNs	Two-component signaling networks

CHAPTER 1: Introduction

1.1 Microbial cell communication and decision making

Cells detect and respond to external cues, messages they receive and transmit using signaling networks that allow bacteria to adapt this behavior. Cellular decision making is the process through which cell takes information from its surroundings (including neighbouring cells), processes these data through complex signal transduction and genetic circuits, and modulate cellular phenotypes in response. Cellular decision making together with environmental sensing and cell-cell communication are three main processes underlying development and pattern formation from microbes to mammals. Cell communication and decision-making underpins microbial behaviour and is vital for their ability to survive in different environments, utilize environmental sources properly, infect plants, animals and humans and fight against the immune system. Two-component signalling (TCS) networks enable microbial communication and decision-making. Understanding these networks can thus lead to understanding of the molecular basis of behaviour, developing strategies to treat infections by pathogenic microbes and facilitating the engineering/reengineering of microbes for biological and biotechnological applications.

1.2 Overview of two-component signalling networks

TCS networks are generally comprised of two protein modules, a sensor protein or histidine kinase (HK) and its cognate response regulator (RR). The activity of the HK is controlled by an environmental stimulus, which controls the rate of autophosphorylation. Once phosphorylated, the HK transfers its phosphoryl group to a cognate RR, which in its phosphorylated form mediates the output of the signaling pathway (1) (Figure 1.1).

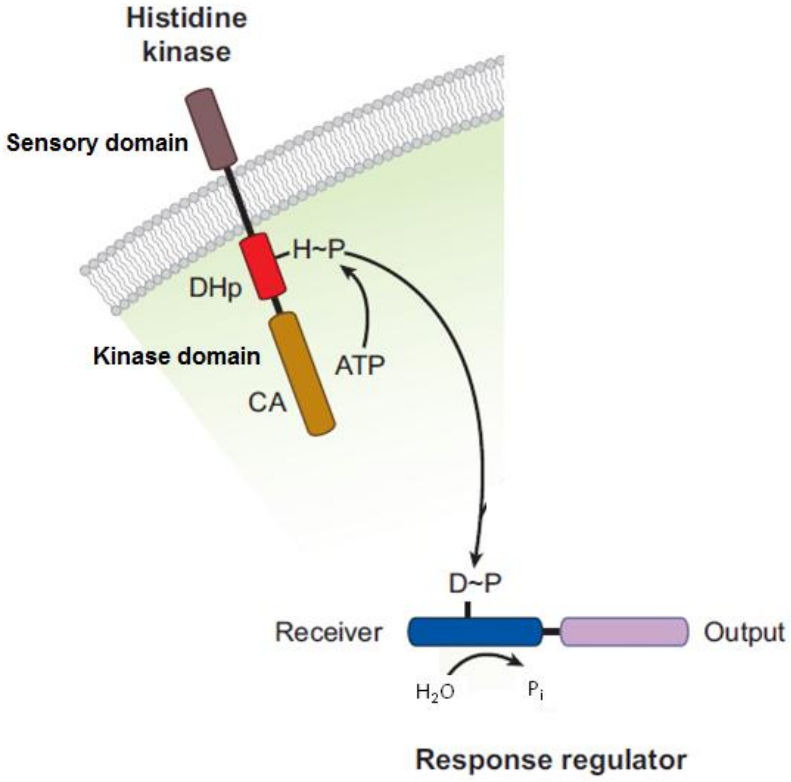

Figure 1.1: Phosphotransfer process occurring between the histidine kinase (HK) and cognate response regulator (RR). Adapted from (2). Following detection of its environmental stimulus, the HK autophosphorylates on a histidine residue. CA domain binds ATP and phosphorylates the histidine residue in the DHp domain. The phosphoryl group is then transferred from the HK to its cognate RR protein. The phosphorylated RR protein can control gene regulation by binding to promoter regions upstream of its target genes.

In Bacteria and Archaea, two component signaling systems are the principle devices for signaling. The first studies into the biochemical reactions in TCSs started in the model organism, *Escherichia coli* (3). Some bacteria extensively employ two component systems; for example, over 30 distinct HK-RR circuits operate in *E. coli* alone and *Myxococcus xanthus* has over 200 two-component systems (4); whereas, no likely HK homologs are found in *Mycoplasma genetilium* genome (5) suggesting that not all prokaryotes use two component

systems to the same extent as *E. coli*. In the case of eukaryotes, two component pathways constitute a small proportion of signalling systems: some (*Arabidopsis thaliana)* have around 50 two component systems whereas; *Saccharomyces cerevisiae* seems to employ only a single two component system (6).

1.3 Eukaryotic TCS

In fungi, TCS is involved in environmental stress responses (7, 8) and hypal development (9-11). In amoeba *Dictyostelium* and in plants, they mediate important processes, like cell growth, differentiation and osmoregulation (12, 13). Hybrid kinases are very common in eukaryotes; the only known exception is *Arabidopsis* ERS (14). In hybrid kinase system, the phosphoryl group is passed to the intramolecularly to a C-terminal receiver domain (conservation of the aspartate residue), similar to that found in response regulators and then to its RR usually via an Hpt domain (Figure 1.2).

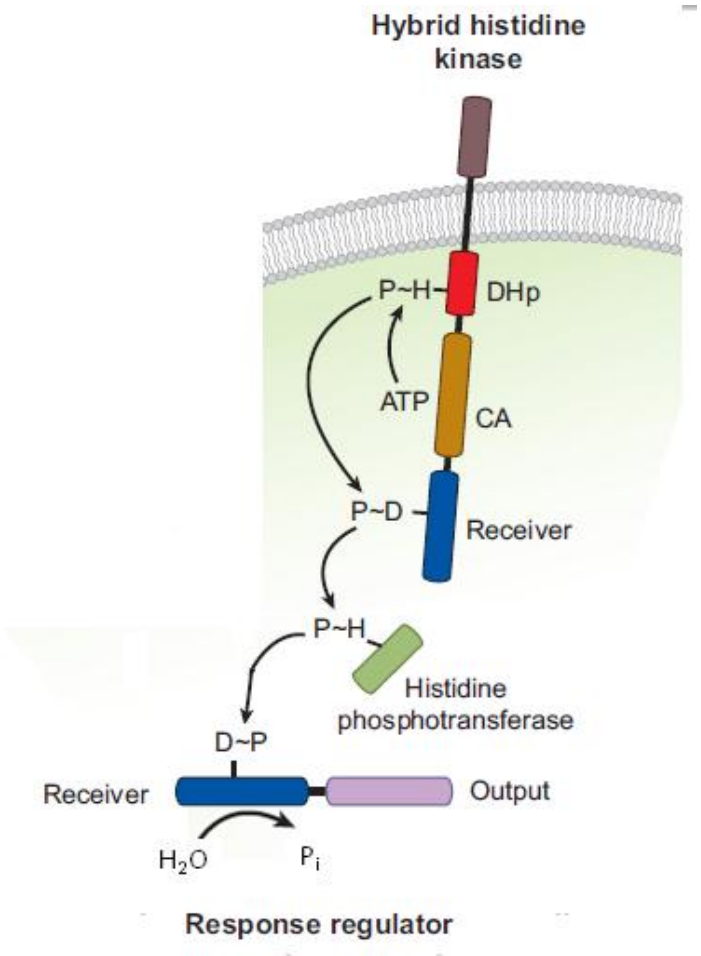

Figure 1.2: Phosphotransfer process occurring between the hybrid histidine kinase and its cognate response regulator (RR) (2). An environmental stimulus activates autophosphorylation of a hybrid HK. The phosphoryl group is then passed intramolecularly to a C-terminal receiver domain, similar to that found in response regulators. A histidine phosphotransferase (HPt) then shuttles the phosphoryl group from the hybrid kinase to a soluble response regulator containing an output domain.

Phylogenetic analyses suggest that the eukaryotic HKs evolved from a single bacterial source represented by a cluster of bacterial hybrid HKs (BarA, RcsC, ArcB) (15). There is only one known eukaryotic RR with a DNA-binding domain (*S. cerevisiae* SKN7) (16); although, bacterial RRs are mainly transcription factors (at least 25 of 32 in *E. coli*). Other signaling

components that are themselves regulated by two-component proteins can potentially effect the final response in eukaryotes. This more-complex strategy allows a greater number of potential steps for regulation, facilitating signal transmission from the cytoplasm to the nucleus, where transcription takes place. The modular domains of the proteins are conserved in terms of their structures and functions in both eukaryotic and bacterial TCS networks. Therefore, characterization of individual components generates a strong basis for understanding other family members.

1.4 Applications of TCS system

TCS systems enable bacteria to sense, respond, and adapt to a wide range of environments, stressors, and growth conditions. These pathways have been adapted to respond to a wide variety of stimuli, including nutrients, cellular redox state, changes in osmolarity, quorum signals, antibiotics, temperature, chemoattractants, pH and more (17). Most bacteria will contain multiple TCS and these can be involved in the regulation of a large number of genes or sets of genes. Indeed any one TCS may interact with one or more other TCS, and activation of a single TCS can induce both negative and positive regulation of different sets of genes. Understanding and manipulating TCS pathways are required for beneficial environmental applications such as nitrogen fixation and bioremediation for agricultural purposes, although most research has prioritized TCSs of pathogens (18). Because of their complete absence from animals (although present in some eukaryotes such as yeasts (19) and plants (20)), two-component proteins have been targeted for the development of antibacterial drugs (21).

1.5 Biochemical basis of TCS

1.5.1 Signal transduction

The chemistry of basic two component pathway involves the following events:

1. Autophosphorylation: HK-His +ATP ⟺ HK-His-P+ADP

2. Phosphotransfer: HK-His-P+RR-Asp ⟺ HK-His+ RR-Asp-P

3. Dephosphorylation: RR-Asp-P +H_2O ⟺ RR-Asp

Generally HKs function as homodimers and autophosphorylate. Many are associated with the cytoplasmic membrane and contain a periplasmic sensory input domain which is coupled to cytoplasmic catalytic kinase domain (22). Biochemical and mutagenesis studies have shown that an ATP-dependent autophosphorylation reaction is catalyzed by the core kinase domain in which one subunit of the dimer phosphorylates a specific His residue. The activity of a RR is regulated by phosphotransfer from its cognate HK-P. Autodephosphorylation of RR-P then takes place allowing signal termination. Some HKs are also found to have phosphatase activities towards their cognate RRs (23, 24). RRs are phosphorylated on an aspartate residue and most of them contain two domains: a receiver domain which is fused to an output domain that has output activity e.g. DNA binding transcription factor (25). However, the activity of the output domain is controlled by phosphorylation of the receiver domain.

1.5.2 Histidine kinases

In a canonical two component signaling pathway, sensor HKs can sense the extracellular stimuli and transmit the information to their cognate RRs. The autophosphorylation activity of HKs can be controlled by the input signals to the sensing domain. The HK catalyzes an ATP-dependent autophosphorylation of its conserved histidine residue within the HK dimerization

domain and then the phosphorylated HK transfers the phosphoryl group to the specific aspartate residue within the receiver domain of the cognate RR. Two-component pathways are controlled by the ability of the HK to the phosphorylation state of the cognate RR. Many HKs have phosphatase activity towards their cognate RRs in addition to their ability to phosphotransfer (26-28). Additional complexity has been also appeared in these pathways where several HKs can phosphorylated a single RR or a single HK can phosphorylate several RRs (29-31).

1.5.2.1 Sensory domain: The N-terminal periplasmic sensory domain of HKs can detect external stimuli either directly or indirectly (32). There are some HKs that contain cytoplasmic sensory domains which can detect intracellular changes. In general, diverse sensory domains of HKs show little primary sequence similarities which is consistent with the ability to respond to a wide range of stimuli. In the simplest mechanism, the sensory domain interacts with the stimulus molecule directly for signal perception, for example, in the control of respiratory nitrate reductase synthesis, the HK NarX directly senses periplasmic nitrate (33). There are also indirect ways of sensing stimuli, for example, in nitrogen sensing the cytoplasmic HK NtrB responds to the uridylylation state of PII, which is controlled by the relative levels of 2-oxoglutarate and glutamine (which reflects cellular nitrogen levels) (34). However, in many cases the specific stimulus remains unknown for the HKs. A number of periplasmic and cytoplasmic sensory domains have been identified in recent times (35-40).

1.5.2.2 ATP binding catalytic domain (CA) and histidine containing phosphotransfer domain (DHp): The length of the core kinase is ~350 amino acid and is essential for ATP binding and autophosphorylation. There are five conserved sequence motifs are present in HKs. The core catalytic ATP binding domain (CA domain) of HKs consists of a conserved set of sequence motifs, named as N, G1, F and G2 boxes. These specific sequences are needed for

Mg^{+2} and ATP binding (41, 42). The conserved phosphorylatable His substrate is a part of H-box sequence motif bound in the DHp domain. In CheA, unlike classical HKs, the phosphorylated histidine residue is present on a separate N-terminal Hpt domain (named P1). The sequence around the CheA phosphorylation site does not resemble the H box of most HKs.

1.5.3 Response regulators

Response regulators are found at the terminal end of the pathway, acting as phosphorylation-activated switches that generate output responses. RRs can transfer phosphoryl group from phosphorylated kinase to its conserved Asp residue. Most of the RRs possess dephosphorylation activities, limiting the lifetime of their active states. Most RRs have two domains: a conserved N-terminal receiver (REC) domain and a variable C-terminal output domain.

1.5.3.1 Receiver domain: The REC domain has approximately 125 amino acids which contain the aspartic acid residue that accepts the phosphoryl group from the phosphorylated histidine kinase in an Mg^{2+-}dependent reaction (43). The receiver domain of any response regulator generally exhibits 20-30 % amino acid sequence homology to other receiver domain, and also contains several invariant residues. These include a pair of acidic residues (an aspartate and a glutamate residue) near the N-terminus of the domain, one near the centre and a lysine residue near the C-terminus border. These conserved residues cluster within an acidic pocket that serves as the site for phosphorylation. The gain of a phosphoryl group within this domain generally activates the output domain located at the C-terminus, which consequently results in activation and/or (44) repression of a given set of genes (45). The conserved REC domains can also be found within hybrid HKs or as isolated proteins within phosphorelay pathways.

1.5.3.2 Output domain: Output domains are diverse in terms of their structure and function. Hence, their regulation is controlled by the receiver domains through different mechanisms. Based on the homology of their DNA-binding domains, the majority of RRs (around two third) are transcription factors with output domains that can be divided into three major subfamilies:

The OmpR/PhoB domains: a novel subclass of winged-helix transcription factors was discovered by crystal structures of the DNA-binding domain of OmpR (46, 47). The fold, conserved in all members of the subfamily (48,49), contains a recognition helix that interacts with the major groove of DNA and flanking loops or "wings" that are proposed to contact the minor groove.

The NarL domains: a four-helix fold for the 62-residue DNA-binding domain was found by the crystal structure of NarL (50). The fold contains a typical helix-turn-helix motif that has allowed postulation of specific interactions between residues of the recognition helix and bases in the NarL heptamer (50).

The NtrC ATPase-coupled transcription factors: The output region of this subfamily contains two domains: an ATPase domain and a helix-turn- helix DNA-binding domain (51-53). NtrC dimers are able to bind to DNA (54) and upon phosphorylation, oligomerize into octamers (55). Oligomerization stimulates ATP hydrolysis (56, 57).

In most response regulators, phosphorylation of the receiver domain causes a conformational change that is propagated to the output domain of the response regulator, which then brings about an appropriate response. However, in the case of CheY, there is no output domain; the phosphorylated receiver domain is capable of binding directly to the FliM component of the switch complex of the flagellar motor, bringing about a change in the direction of flagellar

rotation (58). Not all output domains include DNA binding activities; for example, the output domain of the chemotaxis response regulator, CheB, is a protein methylesterase (59).

1.5.4 Signal termination

Signal termination is a necessity of all signalling pathways. Once the original stimulus is removed then the signalling pathway should return to its previous state, allowing it to respond properly to further stimuli. Within two component signal transduction systems, the input domains deactivate the HKs once the stimulus is removed. Response regulators have an autodephosphorylation activity, which hydrolyses the phosphoaspartate bond. This dephosphorylation reaction has a half time varying from seconds (the chemotaxis proteins CheY and CheB) to hours (the osmoregulating response regulator OmpR) (60, 61). In some systems, the rate of response regulator dephosphorylation is accelerated by other proteins. A specific protein phosphatase is involved in the most common mechanism. Frequently, these phosphatases are encoded on the same polypeptide chain as the HK, and are inversely regulated by the stimuli. For example, NtrB has two activities; a HK activity and an NtrC phosphatase activity. The autophosphorylation of the HK is activated in response to low levels of unmodified PII (indicative of low levels cellular nitrogen) and the phosphatase is activated by high levels of unmodified PII (indicative of high levels of cellular nitrogen) (62). In *E. coli* chemotaxis a dedicated phosphatase protein, CheZ, binds to the response regulator CheY and catalyses the autodephosphorylation reaction of CheY-P (60).

1.6 Diverse architectures of TCS

Evolutionary processes seem to have exploited the modular structure of these TCS proteins to produce a distinct set of biochemical features and network structures that reoccur in diverse TCS. Therefore, different two-component systems control RR phosphorylation through

somewhat different mechanisms. Phosphorelays, bifunc[ptional HKs, sink RRs, and split HKs are some examples of this diversity.

1.6.1 Phosphorelays These are composed of several proteins (or domains), which act as a relay between the HK and RR (Figure 1.3). All phosphorelays characterised to date have a length of four, where the intermediary layers are composed of a response regulator receiver domain (REC) and a histidine domain (Hpt) resulting in a HK–REC–Hpt–RR relay. Signal transduction in this pathway starts when histidine kinase autophosphorylates a histidine residue upon signal stimulation. The phosphoryl group is in turn transferred to an aspartic acid residue in the receiver domain (REC). Subsequently, the phosphoryl group is transferred to the phosphotransmitter (Hpt) and then to the receiver domain of the response regulator (RR), which generates the ultimate output response. The core characteristics of phosphorelay length, and presence and location of hydrolysis and reverse phosphorylation reactions are combined with extra features in different systems. For example, additional RRs can be found at the end of the relay (63, 64), on the other hand, a bifunctional HK can act as both a kinase and a phosphatase (64), and some nested relays are also found within transcriptional feedback loops (65,66). Beside bacteria, phosphorelays are also found in eukaryotic microbes, such as yeast, and in plants, and are shown to be involved in the regulation of virulence (67), sporulation (68, 69), stress responses (70) and cytokinin signalling (71). A detailed study of these systems is lacking with the exception of the sporulation phosphorelay from *B. subtilis*, which has been shown to receive several signals on its different layers (68, 69). While it has been suggested that such signal integration is the main functional role of phosphorelays (69, 72), there are relays that do not involve signal integration (67).

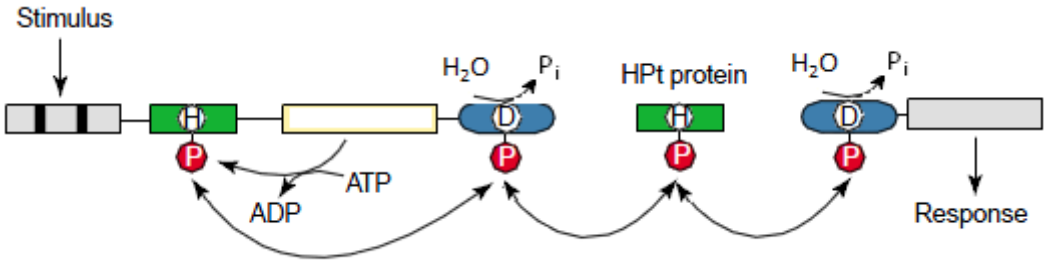

Figure 1.3: Mechanism in phosphorelays**.** A phosphorelay system usually begins with a hybrid HK that has an additional RR regulatory domain at the C-terminus. The scheme generally involves a His-containing phosphotransfer (HPt) protein that serves as a His-phosphorylated intermediate and more than one His–Asp phosphoryl transfer reaction takes place in the system. Adapted from (73).

1.6.2 Bifunctional kinases These HKs display both phosphatase and kinase activity towards their cognate RR. Therefore, they can transfer phosphoryl group to their RRs and also facilitate dephosphorylation of those RRs in the signal cascade (Figure 1.4). The input stimuli can regulate either the kinase or phosphatase activity of the bifunctional HK. There are several examples of bi-functional HKs including the DegS – DegU network in *Bacillus subtilis* (27), the VanS - VanR network of *Enterococcus faecium* (74), and the FixL – FixJ network of *Sinorhozium meliloti* (75), CheA3/CheA4-CheY6 network in *Rhodobacter sphaeroides* (30). Among them, the most-studied bifunctional HK is the osmosensor EnvZ from *E. coli*, which regulates the cognate RR, OmpR (28). It has been proposed that an osmotic signal regulates the ratio of the kinase to the phosphatase activity of EnvZ to modulate the level of cellular OmpR-P primarily by altering the phosphatase activity (76). Hence, when not active for autophosphorylation, bifunctional histidine kinases can effectively suppress any inadvertent cross phosphorylation of their cognate regulators by other kinases (77).

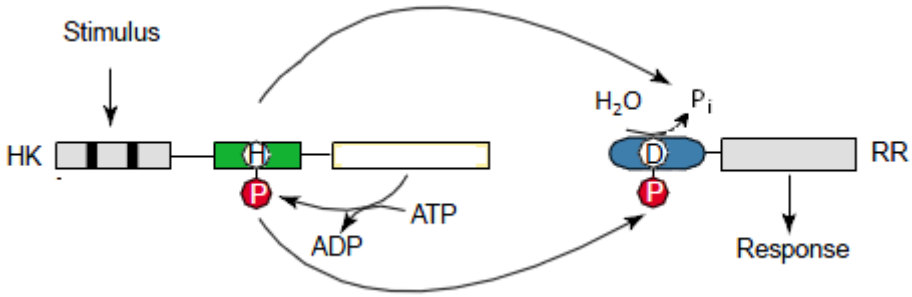

Figure 1.4: Mechanism in Bifunctional kinase**.** A bifunctional HK system in which the kinase gets autophosphorylated and it can also dephosphorylate its cognate RR. Adapted from (73).

1.6.3 Phosphate Sinks This architectural motif has been identified in several microbial and plant systems (31, 78-81). In this pathway, two RRs can receive phosphoryl group from a single HK ((Figure 1.5). There has been previous research demonstrating that such systems feature competition between the two RRs for the phosphoryl group from the HK (31). It has been shown that one of the two RRs displays high affinity for phosphoryl group from the HK, but has a low rate of reverse phosphorylation back to the HK (31, 78-81). This RR can thus function as a sink with respect to the other one. This sink mechanism is well described in *Sinorhizobium meliloti* chemotaxis pathway, *Helicobacter pyroli* chemotaxis pathway and in yeast osmoregulation (31, 78, 80). In *S. meliloti* for example, two response regulators CheY1 and CheY2 are phosphorylated by their kinase CheA. The main RR, CheY2 in its phosphorylated form can bind to flagellar motor and alters rotation. CheY2-P can also back transfer its phosphoryl group to CheA, which in turn phosphorylates the other RR, CheY1. Given its higher phosphorylation rate (from HK) and low reverse phosphorylation rate (to HK), CheY1 is proposed to act as a sink for the phosphoryl group from CheY2-P (31). This is expected to allow faster signal termination, and it is proposed that this sink mechanism replaces the function of a dedicated phosphatase for CheY2 in *S. meliloti* (31).

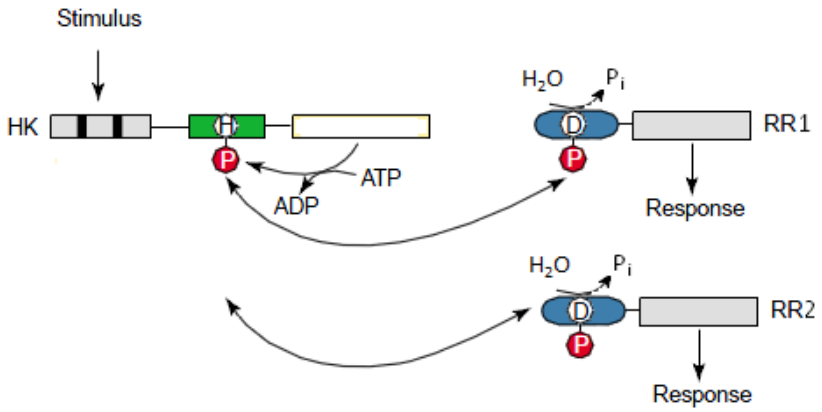

Figure 1.5: Mechanism in phosphate sink. A phosphate sink system where two RRs compete each other for the phosphate group from the same HK. Adapted from (73).

1.6.4 Split Histidine Kinases Split kinases are a complex where the ATP binding and phosphotransfer activities of a conventional HK are split onto two distinct proteins (Figure 6). They are predicted in several bacterial genomes (82, 83) and are biochemically characterized in *Rhodobacter sphaeroides* (84, 85). In this organism, the split kinase system is composed of CheA3 and CheA4, which form a bipartite histidine kinase that phosphorylates the response regulator CheY6 (30). CheA4 lacks the phosphorylatable P1 domain, whereas CheA3 lacks the dimerization (P3) and catalytic kinase (P4) domains. Neither CheA3 nor CheA4 can autophosphorylate when incubated separately with ATP; however, when a mixture of CheA3 and CheA4 is incubated with ATP, then CheA3 becomes phosphorylated, indicating that these proteins can act as a histidine kinase only by forming a complex (30). Activated by incoming signals, the P4 domain of CheA4 binds ATP and phosphorylates the P1 domain of CheA3. Subsequently, CheA3-P acts as a phosphodonor for its cognate response regulator, CheY6 (30), which control flagellar rotation (86). In essence split kinases are unusual bifunctional

HKs, where the autophosphorylation and subsequent phosphotransfer and phosphatase activities are encoded on two separate proteins.

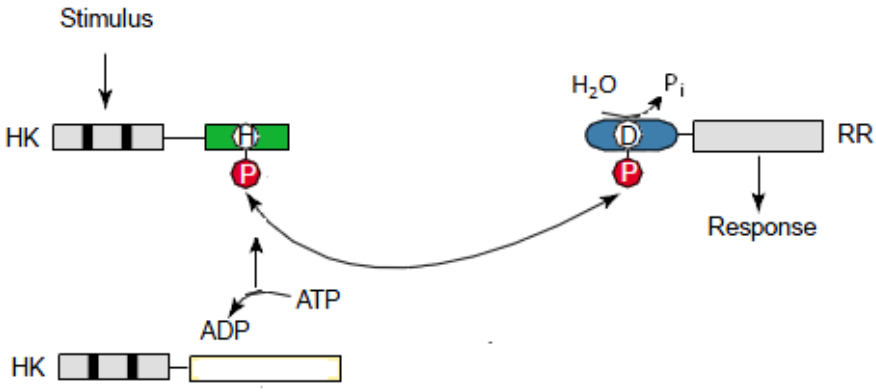

Figure 1.6: Mechanism in split kinase. A split HK system where two HKs form a complex together so that they can act as a conventional kinase and then transfer phosphoryl group to the cognate RR. Adapted from (73).

1.7.1 Sensory pathways involving split kinases

1.7.1.1 Chemotaxis in *Rhodobacter sphaeroides*

Rhodobacter sphaeroides has much more complex chemotaxis signalling pathway with multiple copies of the signalling proteins encoded by three major chemosensory operons and two sets of flagellar genes (fla1 and fla2) (87) compared to the single pathway of *E. coli*. Cells using the Fla1 flagellum have on average one flagellum per cell which rotates unidirectionally. The rotation of this flagellum is controlled by the proteins encoded by *cheOp2* and *cheOp3*; whereas polar flagellum Fla2 is controlled by *cheOp1* (87). In *R. sphaeroides* the signal cascade proteins controlling rotation of the Fla1 flagellum are localized and organized into two distinct sensory clusters: one chemotaxis system is polarly localized while the other forms

cytoplasmic clusters and the signals from the both clusters is needed for chemotaxis (88) (Figure 1.7).

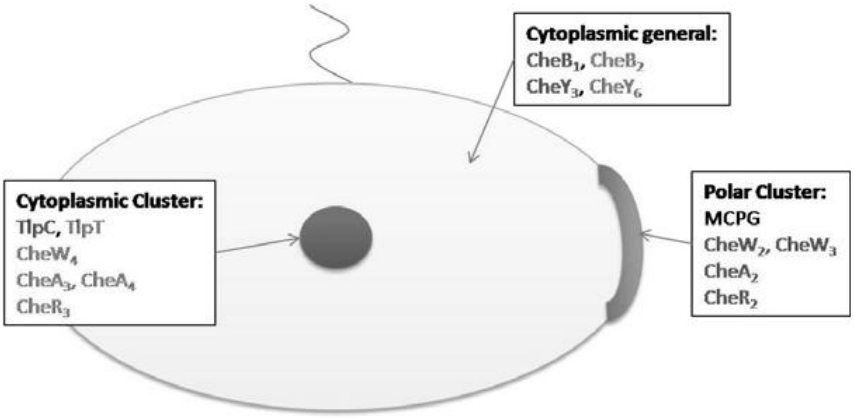

Figure 1.7: The distribution of chemosensory proteins in *R. sphaeroides.* There are two distinct clusters. Most of the proteins from *che*Op3 localise to the cytoplasmic cluster whereas those from *che*Op2 locate to the polar cluster. Adapted from (88).

R. sphaeroides has three CheA homologues that are essential for Fla1 driven chemotaxis – CheA2, CheA3 and CheA4 (92). CheA2 localises to the polar chemoreceptor cluster and has a similar domain structure as *E. coli* CheA. CheA3 and CheA4 are both in the cytoplasmic cluster. CheA3 and CheA4 together form an unconventional split kinase, in the sense that one (CheA4) lacks the conserved P1 containing the autophosphorylatable histidine residue whereas the other (CheA3) lacks the catalytic kinase domains (P3 and P4). These proteins can act as a conventional histidine kinase only by forming a complex (30) in which CheA4 acts as the histidine kinase and phosphorylates the P1 domain of CheA3, CheA3-P then catalyses phosphotransfer to the response regulators (Figure 1.8).

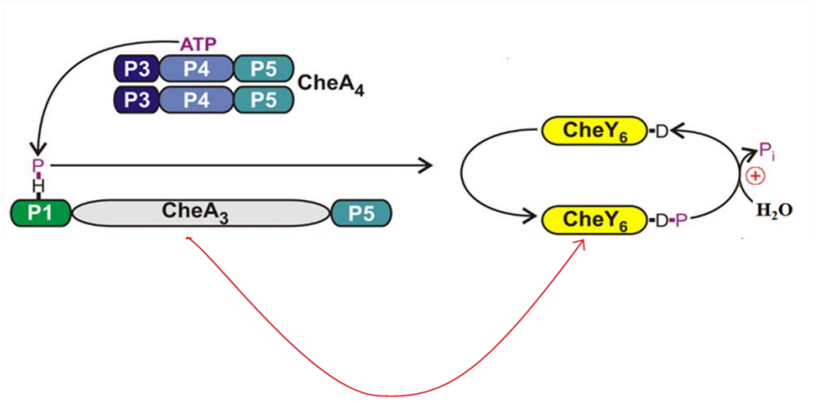

Figure 1.8: A split kinase and bifunctional kinase/phosphatase in bacterial chemotaxis. A CheA4 dimer phosphorylates the P1 domain of CheA3. Then CheA3-P acts as a phosphodonor for CheY6. CheY6-P can autodephosphorylate; however, CheA3 acts as a phosphatase for CheY6-P (red arrow) and can speed up the rate of dephosphorylation. Taken from (86).

However, the CheAs show specific phosphotransfer to the different response regulators *in vitro* (89). CheA2-P can activate all eight chemotaxis response regulators; whereas CheA3-P is specific for CheY1, CheY6 and CheB2. CheY6, CheY4 and CheY3 all bind to the *fla1* motor switch protein, FliM; in all cases phosphorylation enhances binding (90). However, CheY6 with either CheY3 or CheY4 are required for chemotaxis in *vivo* (88). CheY6 is predominantly phosphorylated by CheA3-P which is located in the cytoplasmic cluster. Phosphorylated CheY6 can switch the component of the motor upon binding and change in the direction. In addition, CheA3 has an aspartyl-phosphate phosphatase activity that is specific for CheY6-P (and this is thought to be required for rapid signal termination and hence, necessary for the chemotactic response (86)).

1.7.2 Sensory pathways involving phosphate sinks

1.7.2.1 Chemotaxis in *Sinorhizobium meliloti*

Another diverse TCS architecture is found in soil bacteria, *Sinorhizobium meliloti*. In this organism, the chemotaxis signal transduction system consists of a single HK, CheA and two RRs, CheY1 and CheY1 and there is no such dedicated phosphatase like *E.coli* CheZ. Deletion mutations in CheA, CheY1, CheY2 and both CheY1, CheY2 confirmed that these three proteins are the main regulators in controlling flagellar rotation (91). From that study, it was shown that CheY2 can act as a main regulator which binds the motor and regulates the rotation. However, there is a moderate effect on chemotaxis due to CheY1 mutation, suggesting the role of CheY1 is to compete with CheY2 for phosphorylation by CheA as the phenotype resembles an *E.coli* CheZ mutant which takes a much longer time for signal termination. These phenotypic observations lead to biochemical analysis of the phosphorylation/dephosphorylation events that take place in the system and it was found that CheY1 can act as phosphate sink for the main RR, CheY2 (31). In this organism, the two response regulators CheY1 and CheY2 are phosphorylated by their cognate kinase CheA. Both CheYs can also undergo reverse phosphotransfer, where they return their phosphoryl group to CheA. The phosphoryl group from excess CheY2-P (and to a much lesser extent from CheY1-P) is shuttled back to CheA, which in turn phosphorylates free CheY1 (31) (Figure 1.9). Reverse-phosphorylation via CheA thus accelerates the deactivation of CheY2-P and this is how a phosphate sink can mimic the activity of a 'traditional phosphatase'. More recently, another small protein CheS has been found to work in this pathway by enhancing the interaction between CheY1 and CheA (92). It has been proposed that CheS directly or indirectly promotes CheY1 dephosphorylation and thus make the sink more efficient in the system (92).

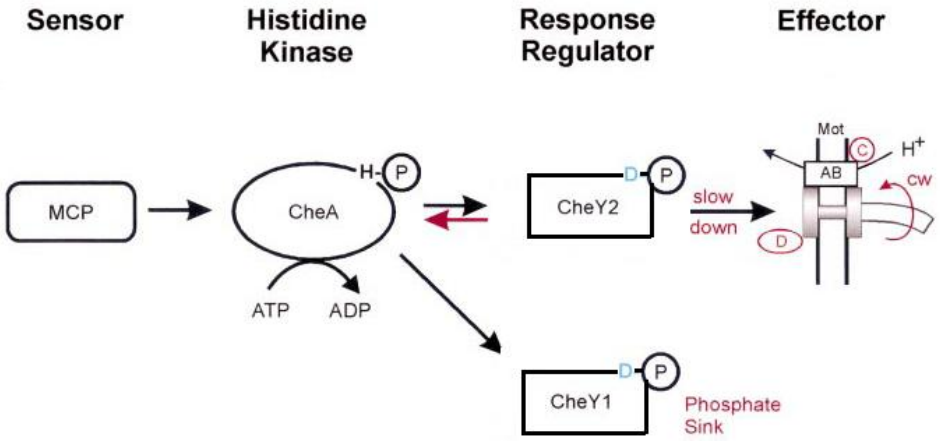

Figure 1.9: A phosphate sink system in *S.meliloti*. HK**,** CheA get autophosphorylated and two RRs, CheY1 and CheY2 compete for the phosphate from the single kinase CheA. CheY2 as a main RR can bind to the motor and alter rotation speed while CheY1 can act as a phosphate sink in this pathway. Adapted from (93).

1.7.2.2 Osmoregulation in yeast

Although TCS networks are not as common in eukaryotes as in bacteria, but they are still found in some key pathways, for example in yeast osmoregulation pathway. In *Saccharomyces cerevisiae*, high osmolarity activates the high-osmolarity glycerol (HOG) response mitogen-activated protein (MAP) kinase pathway (94-96). The downstream HOG MAPK pathway is regulated by the upstream phosphorelay signaling transduction system which is homologous to bacterial two component network (95, 96). A membrane bound histidine kinase Sln1 phosphorylates itself at its histidine residue and the phosphoryl group is then transferred to the aspartic residue of its receiver domain. Subsequently the phosphotransmitter Ypd1 is phosphorylated by the phosphorylated kinase and finally phosphoryl group is transferred to the response regulator Ssk1. High osmotic pressure inhibits the Sln1 histidine kinase activity and

thus maintains lower level of phosphorylated RR, Ssk1. Then unphosphorylated Ssk1 can activate HOG MAPK pathway which in turn can activate the transcription of those genes needed for the response to high osmolarity (Figure 1.10). In yeast, another RR, Skn7 can be phosphorylated by the same signal cascade, using Sln1 as the HK and Ypd1 as the intermediate phosphodonor (Figure 9). Skn7p has a DNA-binding domain homologous to heat shock transcription factors, as well as a receiver domain, acting as a transcription factor for genes involved in various stress related responses (97-99).

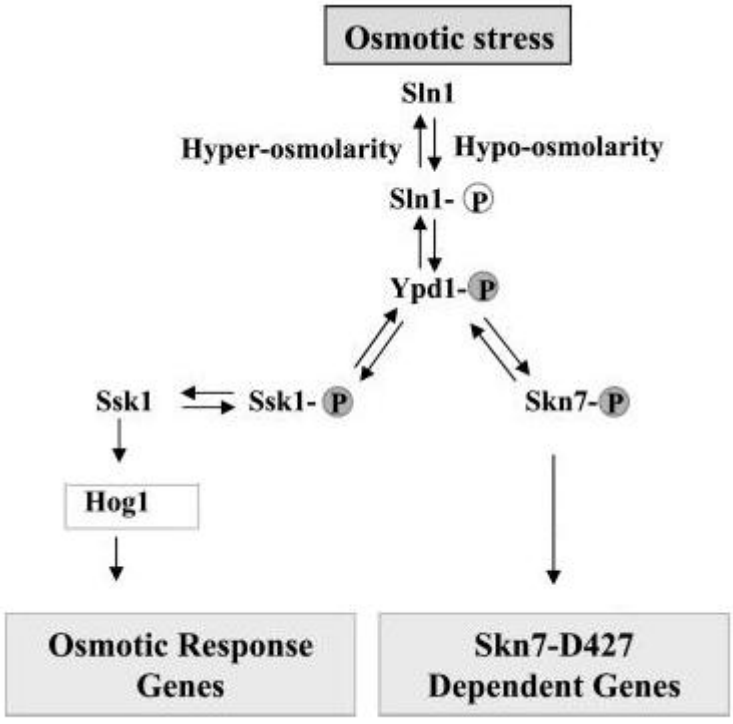

Figure 1.10: Yeast osmoregulation pathway. *HOG1*–dependent gene expression is activated upon accumulation of dephosphorylated Sln1 (Sln1) and *SKN7*-dependent gene expression is activated upon accumulation of phosphorylated Sln1 (Sln1-P) and aspartyl phosphorylation of Skn7p. Adapted from (100).

1.8 Dynamical and signal-response relationship studies of TCS

Significant analyses of TCS networks have been carried out but these are mainly focused on genomic identification of their components and experimental identification of the biochemical reactions among their components in selected model organisms. Collecting a broad understanding of a TCS network in a specific model organism through the combined application of these approaches requires lots of research efforts. Hence, it might be difficult to have a broad and predictive understanding of signaling in microbes, which often utilise structurally diverse TCS architectures for the same function. Therefore, understanding the signal-response relationships in diverse TCS architectural motifs by combining systems and synthetic biology approaches would be possibly a more efficient and safer route to a broad and predictive understanding of microbial signaling.

The physiological responses of cells to external and internal cues are driven by genes and proteins interacting in complex networks. Dynamical properties of those complex networks are difficult to understand only by intuitive reasoning. Recent advances in theoretical biology have shown that biological networks can be perfectly modeled in mathematical terms. These models then throw light on the design principles of biological systems and create predictive ideas that can be verified experimentally (101).

In multidisciplinary research, a key issue is to understand the connections between network structure and network dynamics (102). The multidisciplinary research community studying biological networks has recently seen important progress towards understanding the implications of structural features for network dynamics and functions (103, 104). As, these networks are characterized by their dynamic behaviour, the extensive biochemical knowledge

about these systems is predominantly represented in a static and qualitative manner by drawing arrows connecting interacting components of the network. It was stated correctly in a recent editorial in *Nature:* "But, to really understand the biochemical networks thus represented, one needs to have numbers attached to the arrows" (105). A first step in this direction of analyzing the dynamics is the simulation of these networks. Here, the qualitative reaction scheme is translated into a set of parameterized differential equations. Then, mathematical and computational methods are applied for finding the solution of the governing those equations in the model. The investigation of the steady-state solution as well as the time-dependent solution is carried out which explains how the system will be at the final stage. Therefore, the functional behaviour of a signaling cascade is defined by the way it acts to transmit an input signal (arising from a stimulus) into an output response (usually the expression level of an output protein of interest). They may be characterized by their input–output behavior either by considering steady-state response in concentration space, or transient dynamics in time space. Signal-response curves (106) capture the concentration-space, steady-state behaviour and allow researchers to determine the behavior of the pathway. Signal-response curves are generally formulated in terms of the response to a given signal. Around the core HK-RR interaction, different two-component systems have diverse architectures, which could underpin specific signal processing capabilities. The signal-response relationship can be obtained by measuring both the transient dynamics and the steady state level of system output (i.e. phosphorylated RR) for different inputs (i.e. signal sensed by the HK). The kinetics of the biochemical reactions that make up the specific TCS network structure can determine the relation between the signal and the response and potentially give rise to remarkable complexity even in the simplest implementations of two-component networks (68,77). A detailed understanding of this relation can be achieved for the chemotaxis network of *Escherichia coli* (107), however, this understanding does not permit quantitative prediction

of chemotaxis responses in other bacteria (108), which harbour structurally diverse chemotaxis networks (82). The diversity in TCS networks is the result of evolution extending the core HK-RR reaction both through duplication and diversification of specific proteins and domains, and through acquisition of new genes by horizontal transfer (109). Evolution can also fine tune signal-response relationships by adjusting reaction kinetics and dynamical features such as stability of complexes formed during certain biochemical reactions. In certain cases, these "small" changes can have drastic effects on the overall network and the signal response relationship it mediates (77, 110).

1.8.1 Dynamical and signal-response relationship studies in diverse architectures of TCS

As stated in section 1.7, most of the structural diversity in TCS networks can be found in four common structural arrangements (TCS-motifs); phosphorelays, phosphate sinks, bi-functional and split HKs. In a theoretical analysis, a generic model of phosphorelays was developed (111). This analysis showed that the level of phosphorylated RR responds in a linear fashion to incoming signals, while intermediary layers of the relay display ultrasensitivity. In the case of ultrasensitivity, the response of the system is low until signal levels increase above a certain threshold, after which the response increases disproportionately to reach a high level. The ultrasensitivity of intermediate layers allows these to act as a noise-filter, so that the final layer achieves a significant signal-to-noise ratio that is higher than all the other layers. Further, it was found that this relay structure to favour signal integration through additional phosphatase action on intermediary layers. Ultrasensitivity and noise-filtering can be improved with increasing relay length but this saturated at a relay length of four, which is the maximum length found so far in nature.

Bifunctional HKs have been shown to enable robustness in system output with respect to fluctuations in the amount of the signaling proteins (112,113). Also, bifunctionality is predicted to suppress cross-talk among different TCSNs (77,114). Further, theoretical work indicates that bifunctional HKs can generate flexible signal-response relationships (115,116) and allow higher signal amplification compared to monofunctional HKs that lack phosphatase activity under certain parameter regimes (77).

To date, there are no theoretical or experimental analyses of the signal-response relation in systems with phosphate sinks and split kinase. Although a split kinase system is the part of the overall chemotaxis network in *R. sphaeroides* where several CheA and CheY proteins are found, the exact role of these proteins is still unclear. A recent mathematical modelling study suggested that CheY6 acts as a phosphate sink with respect to other CheYs and enable the bifunctional kinase/phosphatase activity of CheA3 to integrate and tune the sensory output of each signaling cluster to produce a balanced response (117). These architectures are commonly found. There are numerous examples of phosphate sinks found in bacteria and yeast, for example *S. meliloti* and *H. pylori* chemotaxis pathway and yeast osmoregulation pathway (31, 78, 80). The split kinase motif is also common, with over 700 possible examples found in the Genbank database (118).

As the highly modular TCSNs are used by bacteria to control many of their physiological responses, it will be valuable to explore mechanisms which can enable specific response dynamics in these systems and to determine the evolutionary drivers that were responsible for their emergence. This would increase our ability to better understand microbial signaling and exploit it in synthetic biology applications.

1.9 Mathematical approaches to modeling biochemical reaction network

Many biochemical systems such as signal transduction pathways, enzymatic reaction networks and gene regulatory networks are modelled as biochemical reaction networks. To describe the dynamics of these systems, mathematical methods and computational tools have been used. Specially, mathematical modeling of biochemistry has been revisited by 'systems biologists'.

Changes in the concentrations or molecular numbers of biochemical species occur through various reactions in a cell of living organisms. Researchers have tried to find appropriate mathematical and computational ways of modelling to describe such changes. A common way for describing chemical reaction is the mass action approximation. This simply states that the rate of a reaction is equal to a constant multiplied by the product of the concentration of the reactants (119). The time-dependent dynamics of such reaction networks has been traditionally modelled deterministically using differential equations. For large scale biological models which have sufficiently many species and reactions, deterministic description is generally accurate, and mathematical and computational methods have been developed for finding the solution of the governing equation for deterministic models. The investigation of the steady-state solution as well as the time-dependent solution is also important in that it shows how the system will be at the final stage and the stability of the system at the equilibrium. The steady-state solution plays a key role in many subsequent treatments of coupled multienzyme systems (120), allosteric regulation in the concerted MWC (Monod, Wyman, and Changeux) (121), or induced-fit KNF models (Koshland, Nemethy, and Filmer) (122). Throughout the 1960's and 1970's, advances in research were made by combining graph theory, differential equation and chemical reaction network theory, on the existence and importance of the steady state solution (123-126). Especially, Feinberg determined a very important property about the steady-state solution (126); According his theorem 'one can determine the existence and uniqueness of the

stable steady-state for a general class of reaction networks that satisfy an easy-checkable topological property'. It can be usefully applied to large complex networks whose dynamical properties at the equilibrium are difficult to analyse.

Frequent movement of molecules causes important stochastic effects in reaction networks with small number of reactions and molecular species (127). Therefore, stochastic modelling and probabilistic methods can be used for describing the system. Stochastic models mainly are based on the chemical master equation (CME). Stochastic modeling has come to the attention of molecular biologists through studies like single-molecule enzymology (128-131) and live-cell analysis of stochastic processes in living cells, such as gene transcription (132-134) and protein translation (135-138), but it is also true that many physiological processes can be described quite well using deterministic models. It might be easier to analyze for relationships among rate constants or initial protein concentrations and product dynamics (e.g., sensitivity analysis) using deterministic models. Deterministic models may explore fundamental concepts in modeling cellular biochemistry more simply that rely on a simplified representation of space. In many cases, such ordinary differential equation (ODE) models are entirely sufficient as a modeling formalism, and their relative simplicity promotes detailed model analysis, representation of elaborate mechanisms and multi protein networks, and proper comparison of model-based prediction of experimental data (139).

1.9.1 Overview of modeling approaches used in the thesis

Protein network can be described by a mathematical model consisting of ordinary differential equations (ODEs). Bifurcation analysis is a mathematical technique that can determine of the stability of a system with respect to a parameter (140, 141). This can describe the dependence of a state variable on a continuous change in a chosen system parameter. A bifurcation is taken place when there is a change in the number or the stability of solutions of a system. For

example, steady-state solutions for the values of the dependent variables may appear, disappear, change stability, or multiple steady-state solutions may coexist. The coexistence of multiple steady-state solutions is known as multistability. Here, I created the ODEs model and solutions derived from the model give traces of the deterministic behaviour of the concentrations of biochemical species over time, from which the input–output (signal–response) behaviour was computed. Sensitivity analyses were also carried out as they are rapid and easily performed. A complex parameter space generally has too many dimensions to be explored thoroughly by bifurcation analysis. Therefore, sensitivity analysis is needed to determine a class of control parameters, defined as those that strongly affect the stimulus response of a system. Bifurcation analysis can then focus, at least initially, on characterizing how the dynamics are altered by larger changes in the subset of control parameters. Various kinds of tools are available which allow the construction of qualitative biochemical pathway models using kinetic data and their simulation and analysis (142). However, I used XPPAUT and Oscill8 to perform bifurcation analysis and time-course analysis. I also used Chemical reaction theory toolbox for checking the system whether it can be bistable or not (Figure 11).

Chemical reaction network theory toolbox is a framework for modeling the evolution of chemical concentrations resulting from simultaneously occurring chemical reactions. A key feature of the theory is the relationship between the graphical structure of the reaction network and the resulting dynamics. A strong emphasis, consequently, is placed on results which hold regardless of the parameter values of the network, i.e. results which depend on the network structure alone. Biochemical models that exhibit bistability are of interest to biologists and mathematicians simililarly. Chemical reaction network theory can provide conditions for the existence of bistability, and on the other hand can rule out the possibility of emergence of multiple steady states. Systematic study of mass-action kinetics models– which may or may

not admit multiple steady states–constitutes chemical reaction network theory (CRNT), pioneered by Horn, Jackson, and Feinberg (126, 143). Certain classes of networks, such as those of deficiency zero, do not exhibit multistationarity. A generalization of deficiency-zero systems is the class of toric dynamical systems which have a unique steady state (144). Also, there are conditions that are sufficient for establishing whether a network allows multiple steady states. The CRNT Toolbox developed by Feinberg and improved by Ellison implements the Deficiency One and Advanced Deficiency Algorithms (145, 146); this software is available online (147). For a large class of systems, the CRNT Toolbox either provides a possibility for multiple steady states or concludes that it is impossible.

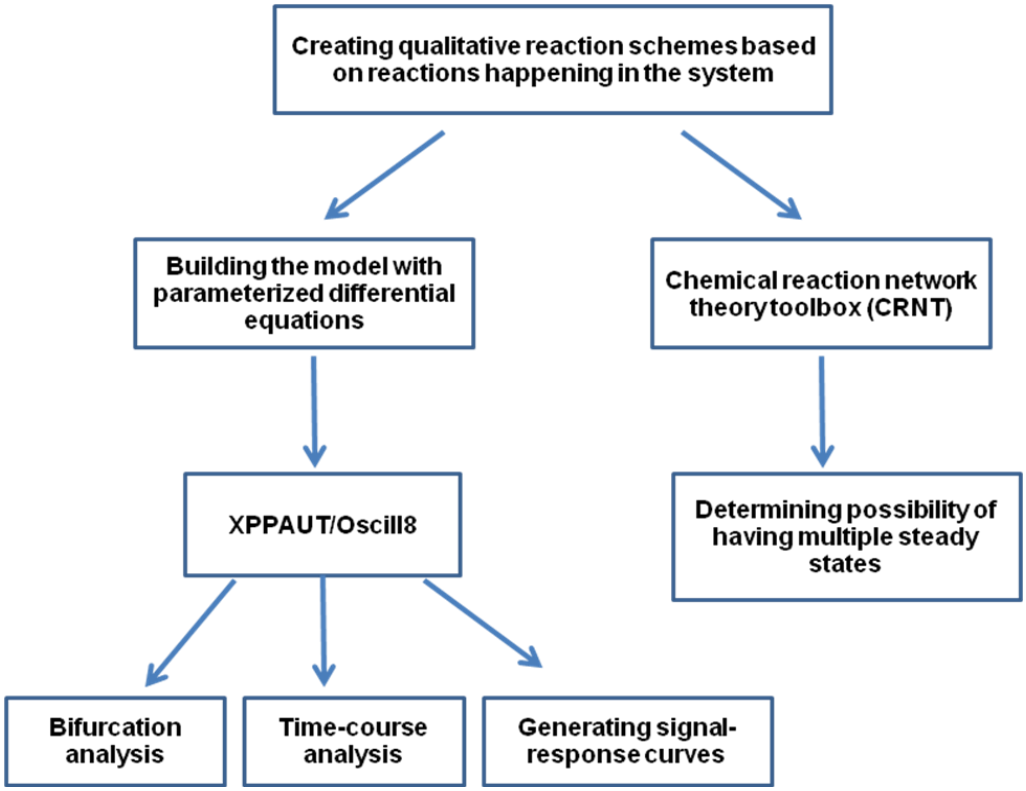

Figure 1.11: A workflow diagram of using mathematical model and tools for analyses, carried out in the thesis.

1.10 Introduction to the Current Study

The main focus of this thesis is the exploring and understanding signal-response relationships and response dynamics of microbial two-component signaling systems. This thesis is submitted for examination for the award of a PhD at the University of Exeter. These are briefly described below.

In chapter 2, I explore the functional role of "split kinases" where the ATP binding and phosphotransfer activities of a conventional histidine kinase are split onto two distinct proteins that form a complex. I show that this unusual configuration enable ultrasensitivity and bistability in the signal-response relationship of the resulting system. These dynamics are displayed under a wide parameter range but only when specific biochemical requirements are met (the kinase activity cannot be increased without reducing the phosphatase activity and vice versa). I experimentally show that one of these requirements, namely segregation of the phosphatase activity predominantly onto the free form of one of the proteins making up the split kinase, is met in proteins isolated from *Rhodobacter sphaeroides*. This chapter was published in *PLoS Computational Biology* (Amin M, Porter S and Soyer OS, 2013). This study provides a linkage between response dynamics, behavior and system architectures.

In chapter 3, I present the signal-response relationship arising from a diverse motif found in two-component signaling. In this motif two response regulators (RRs) can compete with each other for the phosphoryl group from the single kinase (HK), whereby one of the RRs acts as a phosphate sink towards the other (i.e. output RR). I first show that this motif allows rapid signal termination, under the experimentally observed parameters from examples of these systems from bacteria and yeast. Secondly, I demonstrated that phosphate sink containing two-

component systems display a sigmoidal signal-response relationship. I identify two mathematical conditions on system parameters that are necessary for sigmoidal signal-response relationships and define key parameters that control threshold levels and sensitivity of the signal-response curve. I confirm these findings experimentally, by *in vitro* reconstitution of the one HK-two RR motif found in the *S. meliloti* chemotaxis pathway and derive an experimental signal-response curve. I find that the presence of the sink RR can control the level of sigmoidality experimentally and also that an auxiliary protein shown to bind to the HK can further tune the signal-response relationship. These findings show that the one HK-two RR motif allows bacteria and yeast to implement tunable switch-like signal processing and provides an ideal basis for developing threshold devices for synthetic biology applications. This study is submitted in *PLoS Computational Biology* (Amin *et al*., 2013). Except for deriving the mathematical conditions on system parameters and Ordinary differential equation (ODE) model of yeast, I contributed the rest of the paper.

1.11 References

1. Stock AM, Robinson VL, Goudreau PN (2000) Two-component signal transduction. Annu Rev Biochem 69: 183-215.
2. Laub MT, Goulian M (2007). Specificity in two-component signal transduction pathways. Annu. Rev. Genet., 41, 121-145.
3. Ninfa AJ, Magasanik B (1986) Covalent modification of the glnG product, NRI, by the glnL product, NRII, regulates the transcription of the glnALG operon in *Escherichia coli*. Proc Natl Acad Sci USA 83: 5909–5913.
4. Shi X, Wegener-Feldbrügge S, Huntley S, Hamann N, Hedderich R, Søgaard-Andersen L (2008) A bioinformatics and experimental analysis of proteins of two-component systems *in Myxococcus xanthus*. J. Bacteriol., 190, 613 - 624.

5. Fraser CM, et al. (1995) The minimal gene complement of *Mycoplasma genitalium*. Science 270, 397–403.

6. Caren Chang, Richard C Stewart (1998) Two-component system: Regulation of Diverse Signaling Pathways in Prokaryotes and Eukaryotes. Plant Physiol. 117: 723–731.

7. Wurgler-Murphy SM, Saito H (1997) Two-component signal transducers and MAPK cascades. Trends Biochem. Sci. 22, 172–176.

8. Li S, et al. (1998) The yeast histidine protein kinase, Sln1p, mediates phosphotransfer to two response regulators, Ssk1p and Skn7p. EMBO J. 17, 6952–6962.

9. Calera JA, et al. (2000) Defective hyphal development and avirulence caused by a deletion of the SSK1 response regulator gene in *Candida albicans*. Infect. Immun. 68, 518–525.

10. Calera JA, Calderone R (1999) Flocculation of hyphae is associated with a deletion in the putative C-HK1 two-component histidine kinase gene from *Candida albicans*. Microbiology 145, 1431–1442.

11. rikantha T, et al. (1998) The two-component hybrid kinase regulator CaNIK1 of *Candida albicans*. Microbiology 144, 2715–2729.

12. Thomason P, et al. (1999) Taking the plunge: terminal differentiation in *Dictyostelium*. Trends Genet. 15, 15–19.

13. Urao T, et al. (2000) Two-component systems in plant signal transduction. Trends Plant Sci. 5, 67–74.

14. Hua J, Chang C, Sun Q, Meyerowitz EM (1995) Ethylene insensitivity conferred by *Arabidopsis* ERS gene. Science. 22; 269(5231):1712–1714.

15. Pao GM, Saier MH Jr (1997) Nonplastid eukaryotic response regulators have a monophyletic origin and evolved from their bacterial precursors in parallel with their cognate sensor kinases. J Mol Evol 44: 605.

16. Brown JL, North S, Bussey H (1993) SKN7, a yeast multicopy suppressor of a mutation affecting cell wall β-glucan assembly, encodes a product with domains homologous to prokaryotic two-component regulators and to heat shock transcription factors. J Bacteriol. 175:6908–6915.
17. Beier D, Gross R (2006) Regulation of bacterial virulence by two-component systems. Curr. Opin. Microbiol. 9, 143–152.
18. Kruppa M, Calderone R (2006) Two-component signal transduction in human fungal pathogens. FEMS Yeast Res. 6, 149–159.
19. Mizuno T (2005) Two-component phosphorelay signal transduction systems in plants: from hormone responses to circadian rhythms. Biosci. Biotechnol. Biochem. 69, 2263–2276.
20. Stephenson K, Hoch JA (2004) Developing inhibitors to selectively target two-component and phosphorelay signal transduction systems of pathogenic microorganisms. Curr. Med. Chem. 11, 765–773.
21. Bourret RB, Davagnino J, Simon MI (1993) The carboxy-terminal portion of the CheA kinase mediates regulation of autophosphorylation by transducer and CheW. J. Bacteriol. 175, 2097–2101.
22. Igo M, Ninfa AJ, Stock JB, Silhavy TJ (1989) Phosphorylation and dephosphorylation of a bacterial transcriptional activator by atransmembrane receptor. Genes Dev 3: 1725–1734.
23. Makino K, Shinagawa H, Amamura M, Kawamoto T, Yamada M, Nakata A (1989) Signal transduction in the phosphate regulon of *Escherichia coli* involves phosphotransfer between PhoR and PhoB proteins. J Mol Biol 210: 551–559.
24. Parkinson JS, Kofoid EC (1992) Communication modules in bacterial signaling proteins. Annu Rev Genet 26: 71–112.

25. Keener J, Kustu S (1988) Protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: roles of the conserved amino-terminal domain of NTRC. Proc Natl Acad Sci U S A. 85(14):4976–4980.
26. Lois AF, Weinstein M, Ditta GS, Helinski DR (1993) Autophosphorylation and phosphatase activities of the oxygen-sensing protein FixL of *Rhizobium meliloti* are coordinately regulated by oxygen. J Biol Chem. 25; 268(6):4370–4375.
27. Dahl MK, et al. (1992) The phosphorylation state of the DegU response regulator acts as a molecular switch allowing either degradative enzyme synthesis or expression of genetic competence in *Bacillus subtilis*. J Biol Chem 267, 14509-14514.
28. Zhu Y, Qin L, Yoshida T, Inouye M (2000) Phosphatase activity of histidine kinase EnvZ without kinase catalytic domain. Proc Natl Acad Sci USA, 97:7808-7813.
29. Mascher T, Helmann JD, Unden G (2006) Stimulus perception in bacterial signal-transducing histidine kinases. Microbiol Mol Biol Rev; 70:910–38.
30. Porter SL, Armitage JP (2004) Chemotaxis in *Rhodobacter sphaeroides* requires an atypical histidine protein kinase. J Biol Chem 279: 54573-54580.
31. Sourjik V, Schmitt R (1998) Phosphotransfer between CheA, CheY1, and CheY2 in the chemotaxis signal transduction chain of *Rhizobium meliloti*. Biochemistry 37: 2327-2335.
32. Cho US, Bader MW, Amaya MF, Daley ME, Klevit RE, Miller SI, Xu W (2006) Metal bridges between the PhoQ sensor domain and the membrane regulate transmembrane signaling. J Mol Biol; 356:1193–206.
33. Cavicchioli R, Chiang RC, Kalman LV, Gunsalus RP (1996) Role of the periplasmic domain of the *Escherichia coli* NarX sensor-transmitter protein in nitrate-dependent signal transduction and gene regulation. Mol Microbiol 21: 901-911.

34. Magasanik, B. (1993) The regulation of nitrogen-utilization in enteric bacteria. J Cell Biochem 51: 34-40.

35. Neiditch MB, Federle MJ, Pompeani AJ, Kelly RC, Swem DL, Jeffrey PD, Bassler BL, Hughson FM (2006) Ligand-induced asymmetry in histidne sensor kinase complex regulates quorum sensing. Cell; 126:1095–108.

36. Vakonakis I, Klewer DA, Williams SB, Golden SS, LiWang AC (2004) Structure of the N-terminal domain of the circadian clock-associated histidine kinase SasA. J Mol Biol; 342:9–17.

37. Key J, Hefti M, Purcell EB, Moffat K (2007) Structure of the redox sensor domain of Azotobacter vinelandii NifL at atomic resolution: signaling, dimerization, and mechanism. Biochemistry; 46:3614–23.

38. Marina A, Mott C, Auyzenberg A, Hendrickson WA, Waldburger CD (2001) Structural and mutational analysis of the PhoQ histidine kinase catalytic domain. J Biol Chem; 276:41,182–41,190.

39. Bilwes AM, Quezada CM, Croal LR, Crane BR, Simon MI (2001) Nucleotide binding by the histidine kinase CheA. Nat Struct Biol; 8:353–60.

40. Marina A, Waldburger CD, Hendrickson WA (2005) Structure of the entire cytoplasmic portion of a sensor histidine-kinase protein. EMBO J; 24:4247–59.

41. Stock JB, et al. (1989) Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiol. Rev. 53, 450–490.

42. Bilwes AM et al. (1999) Structure of CheA, a signal-transducing histidine kinase. Cell 96, 131–141.

43. Lukat GS, Stock AM, Stock JB (1990) Divalent metal ion binding to the CheY protein and its significance to phosphotransfer in bacterial chemotaxis. Biochemistry 29, 5436-5442.

44. Lange R, Wagner C, Saizieu A, Flint N, Molnos J, Stieger M, Caspers P, Kamber M, Keck Wand Amrein K (1999) Domain organization and molecular characterization of 13 two-component systems identified by genome sequencing of *Streptococcus pneumoniae*. Gene 237: 223-34.
45. Hakenbeck R, Stock JB (1996) Analysis of two-component signal transduction systems involved in transcriptional regulation. Methods in Enzymology 273: 281-300.
46. Kondo H, Nakagawa A, Nishihira J, Nishimura Y, Mizuno T, Tanaka I (1997) *Escherichia coli* positive regulator OmpR has a large loop structure at the putative RNA polymerase interaction site. Nature Struct. Biol., 4, 28–31.
47. Martinez-Hackert E, Stock AM (1997). The DNA-binding domain of OmpR: crystal structures of a winged helix transcription factor. Structure 5, 109-24
48. Mizuno T, Tanaka I (1997) Structure of the DNA-binding domain of the OmpR family of response regulators. Mol Microbiol. 24:665–667.
49. Martinez-Hackert E, Stock AM (1997). Structural relationships in the OmpR family of winged-helix transcription factors. J Mol Biol 269, 301-12.
50. Baikalov I, Schroder I, Kaczor-Grzeskowiak M, Grzeskowiak K, Gunsalus RP, Dickerson RE (1996). Structure of the *Escherichia coli* response regulator NarL. Biochemistry 35(34); 11053-61.
51. Kustu S, North AK, Weiss DS (1991) Prokaryotic transcriptional enhancers and enhancer-binding proteins. Trends Biochem. Sci. 16: 397–402
52. Morett E, Segovia L (1993) The sigma 54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. J Bacteriol 175: 6067–6074.

53. Osuna J, X Soberon, E Morett (1997) A proposed architecture for the central domain of the bacterial enhancer-binding proteins based on secondary structure prediction and fold recognition. Protein Sci. 6:543-555.

54. Weiss V, Claverie-Martin F, Magasanik B (1992) Phosphorylation of nitrogen regulator I of *Escherichia coli* induces strong cooperative binding to DNA essential for activation of transcription. Proc Natl Acad Sci U S A. 1;89 (11):5088–5092.

55. Wyman C, Rombel I, North AK, Bustamante C, Kustu S (1997) Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. Science 275:1658–1661.

56. Weiss DS, Batut J, Klose KE, Keener J, Kustu S (1991) The phosphorylated form of the enhancer-binding protein NtrC has an ATPase activity that is essential for activation of transcription. Cell. 67:155–167.

57. Austin S, Dixon R (1992) The prokaryotic enhancer binding protein NTRC has an ATPase activity which is phosphorylation and DNA dependent. EMBO J. 11:2219–2228.

58. Welch M, Oosawa K, Aizawa SI, Eisenbach, M. (1993) Phosphorylationdependent binding of a signal molecule to the flagellar switch of bacteria. Proc Natl Acad Sci U S A 90: 8787-8791.

59. Simms SA, Keane MG, Stock J (1985) Multiple forms of the CheB methylesterase in bacterial chemosensing. J Biol Chem 260: 10161-10168.

60. Hess JF, Oosawa K, Kaplan N, Simon MI (1988) Phosphorylation of three proteins in the signaling pathway of bacterial chemotaxis. Cell 53: 79-87.

61. Igo MM, Ninfa AJ, Stock JB, Silhavy TJ (1989) Phosphorylation and dephosphorylation of a bacterial transcription activator by a transmembrane receptor. Genes & Devel 3: 1725-1734.

62. Magasanik, B. (1996) Regulation of nitrogen utilization. In *Escherichia coli* and *Salmonella*. Neidhardt, F.C. (ed). Washington, DC. ASM Press, pp. 1344-1356.

63. Janiak-Spens F, Cook PF, West AH (2005) Kinetic analysis of YPD1-dependent phosphotransfer reactions in the yeast osmoregulatory phosphorelay system. Biochemistry 44: 377-386.

64. Chen YE, Tsokos CG, Biondi EG, Perchuk BS, Laub MT (2009) Dynamics of two phosphorelays controlling cell cycle progression in *Caulobacter crescentus*. Journal of Bacteriology 191: 7417–29.

65. Chastanet A, Vitkup D, Yuan GC, Norman TM, Liu JS, et al. (2010) Broadly heterogeneous activation of the master regulator for sporulation in *Bacillus subtilis.* Proc Natl Acad Sci U S A 107: 8486–8491.

66. Narula J, Devi SN, Fujita M, Igoshin OA (2012) Ultrasensitivity of the *Bacillus subtilis* sporulation decision. Proc Natl Acad Sci U S A 109: E3513–E3522.

67. Cotter PA, Jones AM (2003) Phosphorelay control of virulence gene expression in Bordetella. Trends Microbiol 11, 367-373.

68. Bischofs IB, Hug JA, Liu AW, Wolf DM, Arkin AP (2009) Complexity in bacterial cell-cell communication: quorum signal integration and subpopulation signaling in the *Bacillus subtilis* phosphorelay. Proc Natl Acad Sci U S A 106: 6459-6464.

69. Burbulys D, et al. (1991) Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. Cell 64, 545-552.

70. Posas F, et al. (1996) Yeast HOG1 MAP kinase cascade is regulated by a multistep phosphorelay mechanism in the SLN1-YPD1-SSK1 "two-component" osmosensor. Cell 86, 865-875.

71. To JP, Kieber JJ (2008) Cytokinin signaling: two-components and more. Trends Plant Sci 13, 85-92.

72. Appleby JL, et al. (1996) Signal transduction via the multi-step phosphorelay: not necessarily a road less traveled. Cell 86, 845-848.
73. West AH, Stock AM (2001) Histidine kinases and response regulator proteins in two-component signaling systems. Trends Biochem. Sci. 26: 369-376.
74. Arthur M, Molinas C, Courvalin P (1992) The VanS-VanR two-component regulatory system controls synthesis of depsipeptide peptidoglycan precursors in *Enterococcus faecium* BM4147. J Bacteriol 174, 2582-2591.
75. Weinstein M, et al. (1992) Isolation of phosphorylation-deficient mutants of the *Rhizobium meliloti* two-component regulatory protein, FixJ. Mol Microbiol 6, 2041-2049.
76. Jin T, Inouye M (1993) Ligand binding to the receptor domain regulates the ratio of kinase to phosphatase activities of the signaling domain of the hybrid *Escherichia coli* transmemebrane receptor, Taz1. J Mol Biol232: 484–492
77. Alves R, Savageau MA (2003) Comparative analysis of prototype two-component systems with either bifunctional or monofunctional sensors: differences in molecular structure and physiological function. Mol Microbiol 48, 25-51.
78. Jiménez-Pearson MA, et al. (2005) Phosphate flow in the chemotactic response system of *Helicobacter pylori*. Microbiology 151, 3299-3311.
79. Wegener-Feldbrügge, S, Søgaard-Andersen, L (2009) The atypical hybrid histidine protein kinase RodK in *Myxococcus xanthus*: spatial proximity supersedes kinetic preference in phosphotransfer reactions. J Bacteriol 191, 1765-1776.
80. Posas F, Saito H (1998) Activation of the yeast SSK2 MAP kinase kinase kinase by the SSK1 two-component response regulator. EMBO J 17:1385–1394.
81. Lohrmann J, Harter K (2002) Plant two-component signaling systems and the role of response regulators. Plant Physiol 128:363–369.

82. Wuichet K, Zhulin IB (2010) Origins and diversification of a complex signal transduction system in prokaryotes. Sci Signal 3: ra50.
83. Hamer R, Chen PY, Armitage JP, Reinert G, Deane CM (2010) Deciphering chemotaxis pathways using cross species comparisons. BMC Syst Biol 4: 3.
84. Porter SL, Roberts MAJ, Manning CS, Armitage JP (2008) A bifunctional kinase-phosphatase in bacterial chemotaxis. Proc Natl Acad Sci U S A 105: 18531-18536.
85. Porter SL, Wadhams GH, Armitage JP (2011) Signal processing in complex chemotaxis pathways. Nat Rev Microbiol 9: 153-165.
86. Porter SL, Wadhams GH, Martin AC, Byles ED, Lancaster DE, et al. (2006) The CheYs of *Rhodobacter sphaeroides*. J Biol Chem 281: 32694-32704.
87. Porter SL, Wadhams GH, Armitage JP (2008) *Rhodobacter sphaeroides*: complexity in chemotactic signaling. Trends in Microbiology. 16, 251-260.
88. Wadhams GH, Warren AV, Martin AC, Armitage JP (2003) Targeting of two signal transduction pathways to different regions of the bacterial cell. Mol. Microbiol. 50, 763-770.
89. Porter SL, Armitage JP (2002) Phosphotransfer in *Rhodobacter sphaeroides* chemotaxis. J. Mol Biol. 324, 35-45.
90. Ferre A, de la Mora J, Ballado T, Camarena L, Dreyfus G (2004) Biochemical study of multiple CheY response regulators of the chemotactic pathway of *Rhodobacter sphaeroides*. J Bacteriol 186:5172–5177.
91. Sourjik V, R Schmitt (1996) Different roles of CheY1 and CheY2 in the chemotaxis of Rhizobium meliloti. Mol. Microbiol. 22:427-436.
92. Dogra G, et al. (2012) *Sinorhizobium meliloti* CheA complexed with CheS exhibits enhanced binding to CheY1 resulting in accelerated CheY1-P dephosphorylation. J Bacteriol 194(5):1075-1087.

93. Schmitt, R (2002) Sinorhizobial chemotaxis: a departure from the enterobacterial paradigm. Microbiology 148:627-631.

94. Brewster JL, de Valoir T, Dwyer ND, Winter E, Gustin MC (1993) An osmosensing signal transduction pathway in yeast. Science *259*, 1760–1763.

95. Maeda T, Wurgler-Murphy SM, Saito H (1994) A two-component system that regulates an osmosensingMAP kinase cascade in yeast. Nature *369*, 242–245.

96. Maeda T, Takekawa M, Saito H (1995) Activation of yeast PBS2 MAPKK by MAPKKKs or by binding of an SH3-containing osmosensor. Science *269*, 554–558.

97. Brown JL, H Bussey, RC Stewart (1994) Yeast Skn7p functions in a eukaryotic two-component regulatory pathway. EMBO J. 13:5186–5194.

98. Krems B, C Charizanis, KD Entian (1996) The response regulatorlike protein Pos9/Skn7 of *Saccharomyces cerevisiae* is involved in oxidative stress resistance. Curr. Genet. 29:327–334.

99. Morgan BA, GR Banks, WM Toone, D Raitt, S Kuge, LH Johnston (1997) The Skn7 response regulator controls gene expression in the oxidative stress response of the budding yeast *Saccharomyces cerevisiae.* EMBO J. 16:1035–1044.

100. Li S, S Dean, Z Li, J Horecka, RJ Deschenes, JJ Fassler (2002) The eukaryotic two-component histidine kinase Sln1p regulates OCH1 via the transcription factor, Skn7p. Mol. Biol. Cell 13:412-424.

101. Tyson JJ, Katherine CC, Novak B (2003) Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Curr Opin Cell Biol 15:221-231.

102. SH Strogatz (2001) Exploring complex networks. (Invited Insight article) Nature 410, 268-276.

103. R Milo, S Shen-Orr, S Itzkovitz, N Kashtan, D Chklovskii, U Alon (2002) Network motifs: simple building blocks of complex networks. Science, 298, pp. 824–827

104. Gardner TS, Di Bernardo D, Lorenz D, Collins JJ (2003). Inferring genetic networks and identifying compound mode of action via expression profiling. Science, 301(5629), 102-105.

105. Campbell P (1999) Can physics deliver another biological revolution? Nature 397, p. 89.

106. Weiss R, Basu S, Hooshangi S et al. (2003) Genetic circuit building blocks for cellular computation, communications, and signal processing. Nat Comput 2:47–84.

107. Sourjik, V, Berg, HC (2002) Binding of the Escherichia coli response regulator CheY to its target measured in vivo by fluorescence resonance energy transfer. Proc Natl Acad Sci U S A 99, 12669-12674.

108. Soyer OS (2010) The promise of evolutionary systems biology: lessons from bacterial chemotaxis. Sci Signal 3, pe23.

109. Alm E, Huang K, Arkin A (2006) The evolution of two-component systems in bacteria reveals different strategies for niche adaptation. PLoS Comput Biol 2, e143.

110. Soyer OS, et al. (2006) Signal transduction networks: topology, response and biochemical processes. J Theor Biol 238, 416-425.

111. Csikász-Nagy A, Cardelli L, Soyer OS (2010) Response dynamics of phosphorelays suggest their potential utility in cell signaling. J R Soc Interface 8(57):480-488.

112. Batchelor E, Goulian M (2003) Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system. Proc Natl Acad Sci U S A 100: 691-696.

113. Shinar G, Milo R, Martínez MR, Alon U (2007) Input output robustness in simple bacterial signaling systems. Proc Natl Acad Sci U S A 104: 19931-19935.

114. Groban, ES, et al. (2009) Kinetic buffering of cross talk between bacterial two-component sensors. J Mol Biol 390, 380-393.

115. Igoshin OA, Alves R, Savageau MA (2008) Hysteretic and graded responses in bacterial two-component signal transduction. Mol Microbiol 68: 1196-1215.

116. Tiwari A, Ray JC, Narula J, Igoshin OA (2011) Bistable responses in bacterial genetic networks: designs and dynamical consequences. Math Biosci 231: 76-89.

117. Tindall MJ, Porter SL, Maini PK, Armitage JP (2010) Modeling chemotaxis reveals the role of reversed phosphotransfer and a bi-functional kinase-phosphatase. PLoS Comput Biol 6.

118. Amin M, Porter SL, Soyer OS (2013) Split histidine kinases enable ultrasensitivity and bistability in two-component signaling networks. PLoS Comp Biol 9(3): e1002949.

119. F Horn, R Jackson (1972) General mass action kinetics. Archive for Rational Mechanics and Analysis, 47(2), 81–116.

120. Goldbeter A, Koshland DE Jr (1981) An amplified sensitivity arising from covalent modification in biological systems. Proc Natl Acad Sci 78: 6840–6844.

121. Monod J, Wyman J, Changeux JP (1965) On the nature of allosteric transitions: A plausible model. J Mol Biol 12: 88–118.

122. Koshland DE Jr, Nemethy G, Filmer D (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry 5: 365–385.

123. F Heineken, H Tsuchiya, R Aris (1967) On the mathematical status of the pseudo-steady-state hypothesis of biochemical kinetics, Mathematical Biosciences 1, 95–113.

124. F Horn (1972) Necessary and sufficient conditions for complex balancing in chemical kinetics, Archive for Rational Mechanics and Analysis, 49(3) 172– 186.

125. H G Othmer (1976) Nonuniqueness of equilibria in closed reacting systems. Chemical Engineering Science 31 993–1003.

126. Feinberg, M.: Lectures on chemical reaction networks. Notes of lectures given at the Mathematics Research Center of the University of Wisconsin in 1979, http://www.che.eng.ohio-state.edu/_feinberg/LecturesOnReactionNetworks.

127. CV Rao, MW Wolf, AP Arkin (2002) Control, exploitation and tolerance of intracellular noise. Nature, 420 231–237.

128. Ishijima A, Doi T, Sakurada K, Yanagida T (1991) Sub-piconewton force fluctuations of actomyosin in vitro. Nature 352: 301–306.

129. Cai L, Friedman N, Xie XS (2006) Stochastic protein expression in individual cells at the single molecule level. Nature 440: 358–362.

130. Finer JT, Simmons RM, Spudich JA (1994) Single myosin molecule mechanics: Piconewton forces and nanometre steps. Nature 368: 113–119.

131. Kim S, Blainey PC, Schroeder CM, Xie XS (2007) Multiplexed single-molecule assay for enzymatic activity on flowstretched DNA. Nat Methods 4: 397–399.

132. Golding I, Cox EC (2004) RNA dynamics in live Escherichia coli cells. Proc Natl Acad Sci 101: 11310–11315.

133. Elf J, Li GW, Xie XS (2007) Probing transcription factor dynamics at the single-molecule level in a living cell. Science 316: 1191–1194.

134. Zenklusen D, Larson DR, Singer RH (2008) Single-RNA counting reveals alternative modes of gene expression in yeast. Nat Struct Mol Biol 15: 1263–1271.

135. Choi PJ, Cai L, Frieda K, Xie XS (2008) A stochastic singlemolecule event triggers phenotype switching of a bacterial cell. Science 322: 442–446.

136. Munro JB, Altman RB, O'Connor N, Blanchard SC (2007) Identification of two distinct hybrid state intermediates on the ribosome. Mol Cell 25: 505–517.

137. Agirrezabala X, Lei J, Brunelle JL, Ortiz-Meoz RF, Green R, Frank J (2008) Visualization of the hybrid state of tRNA binding promoted by spontaneous ratcheting of the ribosome. Mol Cell 32: 190–197.

138. Julian P, Konevega AL, Scheres SH, LazaroM, Gil D,Wintermeyer W, Rodnina MV, Valle M. 2008. Structure of ratcheted ribosomes with tRNAs in hybrid states. Proc Natl Acad Sci 105: 16924–16927.

139. WW Chen, M Niepel, PK Sorger (2010) Classic and contemporary approaches to modeling biochemical reactions Genes Dev., 24, pp. 1861–1875.

140. Guckenheimer J, P Holmes (1983) Nonlinear Oscillations, Dynamical Systems and Bifurcation of Vector Fields. Springer, New York.

141. Strogatz SH (1994) Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering. Addison-Wesley, Reading, MA.

142. Alves R, Antunes F, Salvador S (2006) Tools for kinetic modeling of biochemical networks. Nat Biotechnol, 24:667-72.

143. Horn F, Jackson R (1972) General mass action kinetics. Arch. Rat. Mech. Anal. 47 81–116.

144. Craciun G, Dickenstein A, Shiu A, Sturmfels B Toric dynamical systems. Available from arXiv: 0708.3431.

145. Ellison P (1998) The advanced deficiency algorithm and its applications to mechanism discrimination. PhD Thesis, University of Rochester.

146. Feinberg M (1995) The existence and uniqueness of steady-states for a class of chemical reaction networks. Arch. Ration. Mech. Anal. 132, 311–370.

147. Ellison, P., Feinberg, M. CRNT Toolbox. Available from http://www.che.eng.ohio-state.edu/_feinberg/crnt/.

CHAPTER 2

Split histidine kinases enable ultrasensitivity and bistability in two-component signaling networks

2.1 Introduction

Bacterial responses to many external stimuli are underpinned by two-component signaling networks (TCSNs). These are found in most bacterial species and are also present in Archaea, eukaryotic microbes, and plants [1,2]. TCSNs are built upon the core reactions involving a histidine kinase (HK) that autophosphorylates on a conserved histidine residue in response to a signal, and a cognate response regulator (RR) that is activated when the HK phosphorylates one of its conserved aspartate residues [3]. Evolutionary processes seem to have exploited the modular structure of these proteins to produce a distinct set of biochemical features and network structures that reoccur in diverse TCSNs; bifunctional HKs [4], sink RRs [5], phosphorelays [6] and split HKs [7]. In order to achieve a broad and predictive understanding of bacterial signaling, it is important to assess whether these features enable specific signaling dynamics and properties [8].

There has already been progress towards this goal. Firstly, bifunctional HKs, which display both phosphatase and kinase activity towards their cognate RR, enable robustness in system output with respect to fluctuations in the amount of these signaling proteins [4,9] and reduce cross-talk among different TCSNs [10,11]. Further, theoretical work indicates that bi-functional HKs can generate flexible signal-response relationships [12, 13] and allow higher signal amplification compared to monofunctional HKs that lack phosphatase activity [10]. Secondly, sink RRs, which compete with another RR for phosphoryl groups from a single

cognate HK, are suggested to allow faster response termination [5, 14]. Finally, phosphorelays, which contain several proteins (or domains) acting as a relay between the HK and RR, are suggested to integrate several signals received on their different layers [15-17] and implement both ultrasensitive and linear responses [18,19]. Taken together, these studies suggest that specific biochemical and structural features in TCSNs could enable specific functional roles.

Of the different features of TCSNs, split kinases are predicted in several bacterial genomes [1,2] and are biochemically characterized in *Rhodobacter sphaeroides* [7,20]. In this organism, the split kinase system is composed of CheA3 and CheA4, which form a bipartite histidine kinase that phosphorylates the response regulator CheY6 [21] (Figure 2.1). CheA4 lacks the phosphorylatable P1 domain, whereas CheA3 lacks the dimerization (P3) and catalytic kinase (P4) domains. Neither CheA3 nor CheA4 can autophosphorylate when incubated separately with ATP; however, when a mixture of CheA3 and CheA4 is incubated with ATP, then CheA3 becomes phosphorylated, indicating that these proteins can act as a histidine kinase only by forming a complex [21]. Activated by incoming signals, the P4 domain of CheA4 binds ATP and phosphorylates the P1 domain of CheA3. Subsequently, CheA3-P acts as a phosphodonor for its cognate response regulator, CheY6 [21], which controls flagellar rotation [22].

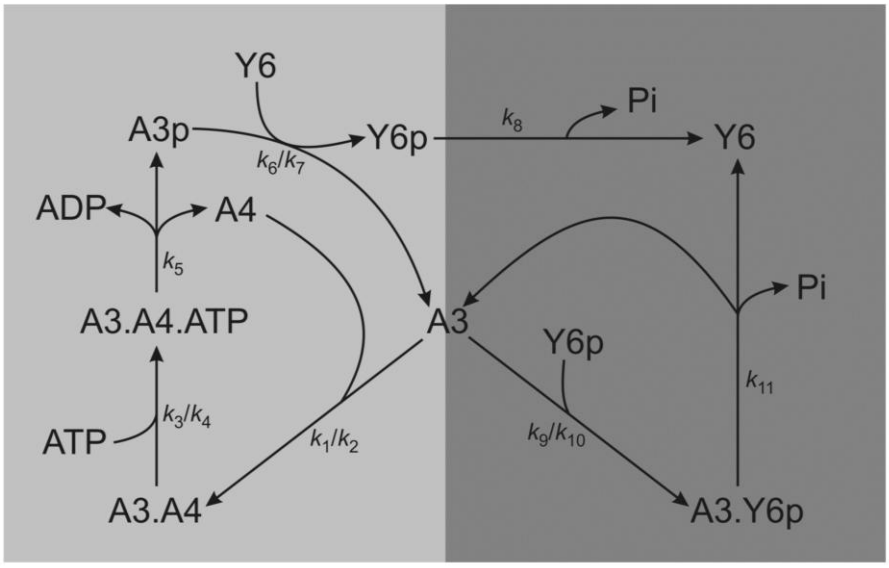

Figure 2.1 A cartoon diagram of the CheA3-CheA4-CheY6 split kinase system. The diagram is arranged so to highlight the role of free CheA3 acting as a branching point for the two arms that form competing cycles leading to phosphorylation and dephosphorylation of CheY6. Rate constants are shown on the relevant reactions. In the case of reversible reactions, two rate constants are given ($k_{forward}/k_{reverse}$).

In vivo, CheA3 and CheA4 co-localize to the cytoplasmic chemotaxis cluster [23] and are both essential for chemotaxis [7,24]. CheA3 and CheA4 bind to the cytoplasmic cluster via their P5 domains [25]. Whilst part of this cluster, CheA3 and CheA4 dynamically interact with one another. To allow phosphorylation of CheA3, the P4 domain of CheA4 must transiently bind to the P1 domain of CheA3 (in the subsequent analysis we refer to this complex as CheA3:CheA4). Once phosphorylated, the P1 domain of CheA3 is released by CheA4, and CheA3-P can then donate its phosphoryl group to the corresponding response regulator CheY6 [21, 26]. In addition to its phosphotransfer function, CheA3 is also a phosphatase towards CheY6-P [7]. cheA3 mutants retaining phosphotransfer functions but lacking phosphatase activity do not support chemotaxis, similarly, cheA3 mutants retaining phosphatase activity

but lacking phosphotransfer activity also fail to support chemotaxis, indicating that chemotaxis requires both activities of CheA3 [7,21]. In addition, to being phosphorylated and dephosphorylated by the split kinase comprising CheA3 and CheA4 [21], CheY6 is also phosphorylated by CheA2 at the polar chemotaxis cluster [27].

Despite this wealth of information, the general role of split kinases in bacterial signaling is not clear. In essence split kinases are unusual bifunctional HKs, where the autophosphorylation and subsequent phosphotransfer and phosphatase activities are encoded on two separate proteins. Since the complex formed by these proteins is functionally equivalent to a bifunctional HK, it is not clear what the role of splitting biochemical activities in this way might be. Using the biochemical reactions of CheA3, CheA4, and CheY6 as a model system, we developed a mathematical model and analyzed the response dynamics mediated by this split kinase. Repeating this analysis with a bifunctional HK and a conventional HK-RR pair featuring a separate phosphatase, we found that in contrast to these configurations, split kinases enable ultrasensitivity and bistability in the signal-response relationship. We show that these dynamical features are maintained under a wide parameter range, provided certain biochemical assumptions are met. These requirements indicate that the source of ultrasensitivity and bistability in split kinases is the inverse coupling between their kinase and phosphatase activities; i.e. the kinase activity cannot be increased without reducing the phosphatase activity and vice versa. Through measurements of phosphatase activity, we show that this condition is met in the *R. sphaeroides* system *in vitro*. These findings suggest that bacteria might be utilizing split kinases as a means of implementing ultrasensitivity and bistability in cellular decision making.

2.2 Results

2.2.1 Construction of a mathematical model of a split kinase

Since our aim is to study the general response dynamics that split kinases can mediate, we use the CheA3, CheA4, and CheY6 triplet as a model system and study its dynamics in isolation through *in vitro* experiments, numerical simulation and analytical approaches. We developed a mathematical model of the system and parameterized it with *in vitro* and *in vivo* measured kinetic rates and protein concentrations respectively (see Methods and Table 2.1). We then analyzed the response dynamics of the resulting model and its variants both through numerical simulations and deriving analytical solutions of steady state behavior using approximations and the chemical network theory[28,29] (see Methods and Text S 2.1 in Appendix A). In the subsequent sections, we use the terms free CheA3 and free CheA3-P to indicate CheA3 species where the P1 domain is not interacting with the P4 domain of CheA4; *in vivo*, however, these species are expected to be always joined to the chemotaxis cluster by their P5 domains.

Parameter	**Description**	**Value**	**Unit**	**Ref**
k1	On rate for binding of CheA3 and CheA4	100	(μMs^{-1})	[21] see also Results
k_2	Off rate for binding of CheA3 and CheA4	10	s^{-1}	[21] see also Results
k_3	Forward rate for phosphorylation complex	1	$(\mu Ms)^{-1}$	[21]
k_4	Reverse rate for phosphorylation complex	39	s^{-1}	[21]

k_5	K_{cat} for phosphorylation of CheA3 by CheA4	varied	s^{-1}	
k_6	CheA3-P to CheY6 Phosphotransfer	0.775	$(\mu Ms)^{-1}$	[21]
k_7	CheA3-P to CheY6 Reverse phosphotransfer	0.00283	$(\mu Ms)^{-1}$	[21]
k_8	Autodephosphorylation	0.169	s^{-1}	[7]
k_9	Association of phosphatase assisted dephosphorylation complex	5.6	$(\mu Ms)^{-1}$	[48]
$k1_0$	Dissociation of phosphatase assisted dephosphorylation complex	0.04	s^{-1}	[48]
$k1_1$	K_{cat} for phosphatase assisted dephosphorylation	2.5	s^{-1}	See Methods
$[A3]_{tot}$	Total concentration of CheA3	90	µM	[7]
$[A4]_{tot}$	Total concentration of CheA4	40	µM	[34]
$[Y6]_{tot}$	Total concentration of CheY6	225	µM	[34]
[ATP]	Total concentration of ATP	1000	µM	

Table 2.1: Literature source and parameter values used in the analysis of the basic model.

2.2.2 The input-output relationship for the split kinase shows ultrasensitivity and bistability

A primary property of interest for any signal transduction system is the signal-response relationship it implements [30]. To analyze the signal-response relationship in systems featuring a split kinase, we defined the system response as the steady state level of phosphorylated CheY6 (CheY6-P) at a given signal level, and derived this relationship for different parameters and biochemical assumptions (see Methods). This analysis revealed that when assuming free CheA3 as the sole phosphatase for CheY6-P, the system has a high potential for displaying ultrasensitivity and bistability (Figure 2.2 and Figures S2.1 – S2.3 in Appendix A). Both of these dynamics result in a switch-like behavior; the response of the system is low until signal levels increase above a certain threshold, after which the response increases disproportionately to reach a high level (e.g. Figure 2.2A). In the case of bistability, the low and high response levels correspond to stable states of the system, separated by an unstable region, resulting in abrupt switching dynamics and hysteresis (i.e. the switching threshold is different depending on the past state of the system).

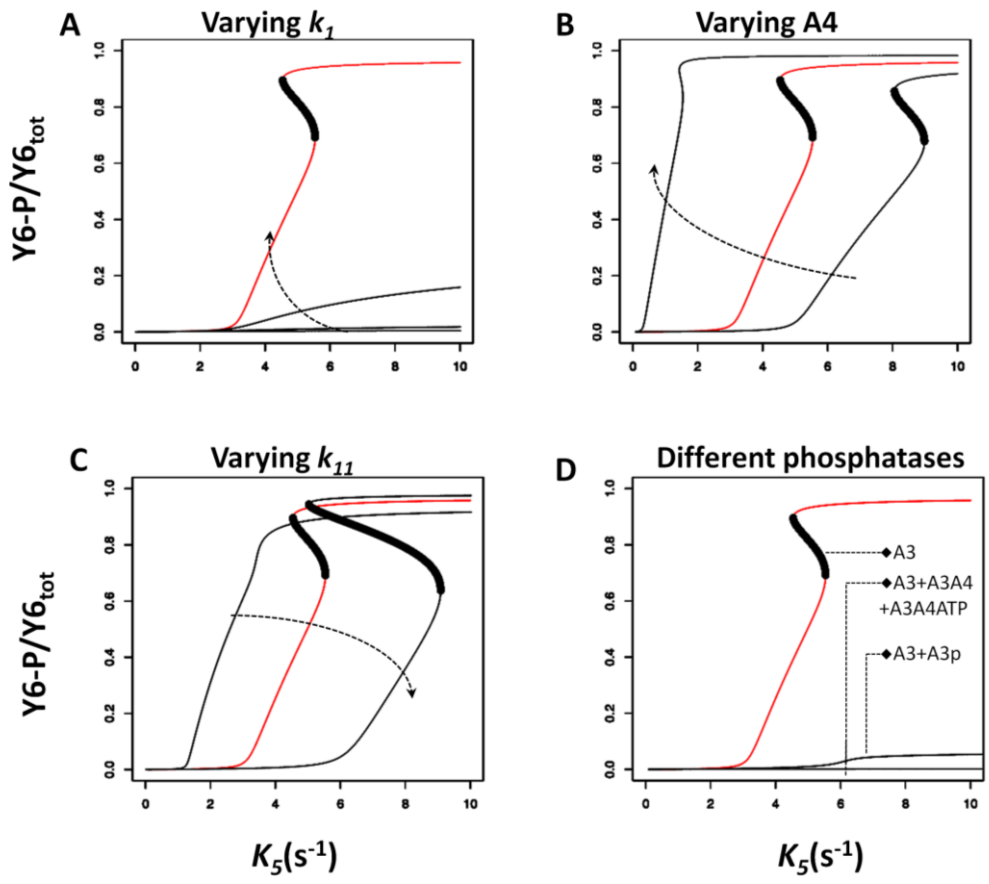

Figure 2.2 Effects of varying key parameters of the model and addition of different phosphatases. The x- and y-axis show the signal (*k5*) level and the corresponding steady state CheY6-P level respectively. Each panel shows a signal-response analysis for varying model parameters (A-C) or the inclusion of additional phosphatases (D). The results of the basic model are shown in red. Where present, the dark region indicates the region of unstable steady states and hence the presence of bistability. Arrows on panels A, B and C indicate increasing value of the changed parameter. (A) The on rate (*k1*) for CheA3:CheA4 complex formation was varied from basic model value [100(μMs)$^{-1}$] to 10, 1, and 0.208. (B) Concentration of CheA4 was varied from 30 μM, 40 μM (basic model) and 80 μM. (C) The rate of CheA3 mediated dephosphorylation of CheY6-P ($k1_1$) was varied from 1s^{-1}, 2.5 s^{-1} (basic model) and 5s^{-1}. (D) The basic model has free CheA3 as the sole phosphatase; the effect of having either

CheA3-P or CheA3:CheA4 and CheA3:CheA4: ATP as additional phosphatases is shown. See also Figures S2.1-2.4 in Appendix A for additional sensitivity analyses.

The *in vitro* and *in vivo* measured kinetic rates and protein concentrations from *R. sphaeroides* constitute "biologically meaningful" values that could be representative for two-component systems in general. To analyze the potential effects of these rates on the observed nonlinearity of the signal-response relationship, we have performed a sensitivity analysis by varying the base parameter values over a large range and quantifying the shape of the resulting signal-response curve (see Methods). This analysis shows that the level of ultrasensitivity in the signal-response relationship is most sensitive to the parameters controlling the complex formation between CheA3:CheA4 (*k1*) and the dephosphorylation of phosphorylated CheY6 (*k9* and $k1_1$) (Figure 2.2 and Figures S2.1 – S2.3 in Appendix A). The association rate constant (*k1*/*k2*) we used in the basic model is approximately 500-fold higher than that measured *in vitro*, using purified *R. sphaeroides* proteins [21]. We still consider this high value "biologically relevant" as *in vivo* conditions can result in confining of split kinase components to small regions of the cell, resulting in much higher effective concentrations than are attainable under the *in vitro* conditions as used in [21]. For example, in *R. sphaeroides*, CheA3 and CheA4 localize to the cytoplasmic chemoreceptor cluster [23], which - using immunogold electron microscopy - is estimated to occupy less than 5% of the cross-sectional area of the cell [31]. Assuming a spherical shape for both the cell and this cluster, the volume of the latter could be estimated to be approximately 1% of the total cell volume. Thus, the effective concentrations of CheA3 and CheA4 in this cluster could be increased by as much as 100-fold, resulting in a significantly higher effective association rate constant than measured *in vitro* (up to 10,000 fold).

Besides parameter values, several modeling choices could also alter the finding of bistability and ultrasensitivity arising in a split kinase system. In particular, the basic model presented above assumes that free CheA3 is the sole phosphatase in the system (besides the intrinsic autodephosphorylation activity of CheY6-P). Relaxing this assumption and considering increasing phosphatase activity by the CheA3:CheA4 and CheA3:CheA4:ATP complexes (see Text S1, section 1), significantly reduced ultrasensitivity in the system (Figure 2.2D and S2.4 in Appendix A). In contrast, the presence of ultrasensitivity was much more robust to increasing phosphatase activity by CheA3p (Figure 2.2D, S2.4 and S2.5 in Appendix A). Another mechanistic choice in the modeling of the split kinase system is the fate of the CheA3:CheA4 complex after phosphorylation of CheA3. In the basic model analyzed in Figure 2.2, this is modeled as phosphorylation resulting in the dissociation of the complex and release of CheA4 and CheA3-P. An alternative would be that the CheA3:CheA4 complex remains intact post phosphorylation, resulting in a CheA3-P:CheA4 complex (see Text S2.1, section 2 in Appendix A). When we assume the presence of CheA3-P:CheA4 complex that can phosphotransfer to CheY6, bistability was lost, but not ultrasensitivity (Figure S2.6 in Appendix A). Finally, we found that including an additional (monofunctional, non-split) kinase in the model, as seen for example in *R. sphaeroides* CheA2 (see Text S2.1, section 3 in Appendix A), does not affect the ultrasensitivity but can result in the loss of bistability (Figure S2.7 in Appendix A).

It is important to note that the basic model and all of these variants arising from specific modeling choices are “nested” in the sense that the basic model can be recovered through appropriate choice of parameters (e.g. setting dephosphorylation activity of CheA3p very low). In line with this observation, we find that the basic model and all of the alternative structures discussed so far can be analytically shown to possess the “ability” to attain

bistability (see Methods). More particularly, each of the chemical reaction systems arising from these models have the capacity for multiple steady states according to the higher deficiency theorem [32,29]; i.e. these chemical systems permit bistability for some set of non-zero parameter values and under the assumption of mass action kinetics (see Text S2.2 in Appendix A).

2.2.3 Segregation of kinase and phosphatase activities allows ultrasensitivity and bistability

Taken together, these analyses suggest that the ability of a split kinase to mediate ultrasensitivity and bistability relates to the segregation of kinase and phosphatase activities. To better understand how this relates to ultrasensitivity and bistability, we simulated the time evolution of the system in the presence of step signals. As expected from the ultrasensitive signal-response relationship, system response (i.e. increase in free CheY6-P) was low for step-signals below the threshold and displayed a sudden large jump for step-signals crossing the threshold (Figure 2.3). Before the threshold, increasing signal levels resulted in an increase in the CheA3:CheY6-P complex, while the crossing of the threshold and subsequent increases in signal caused it to decrease. The reason for this behavior is that before the threshold there is enough free CheA3 in the system to bind and dephosphorylate the CheY6-P that is formed, while after crossing of the threshold there is no free CheA3 left in the system (Figure 2.3).

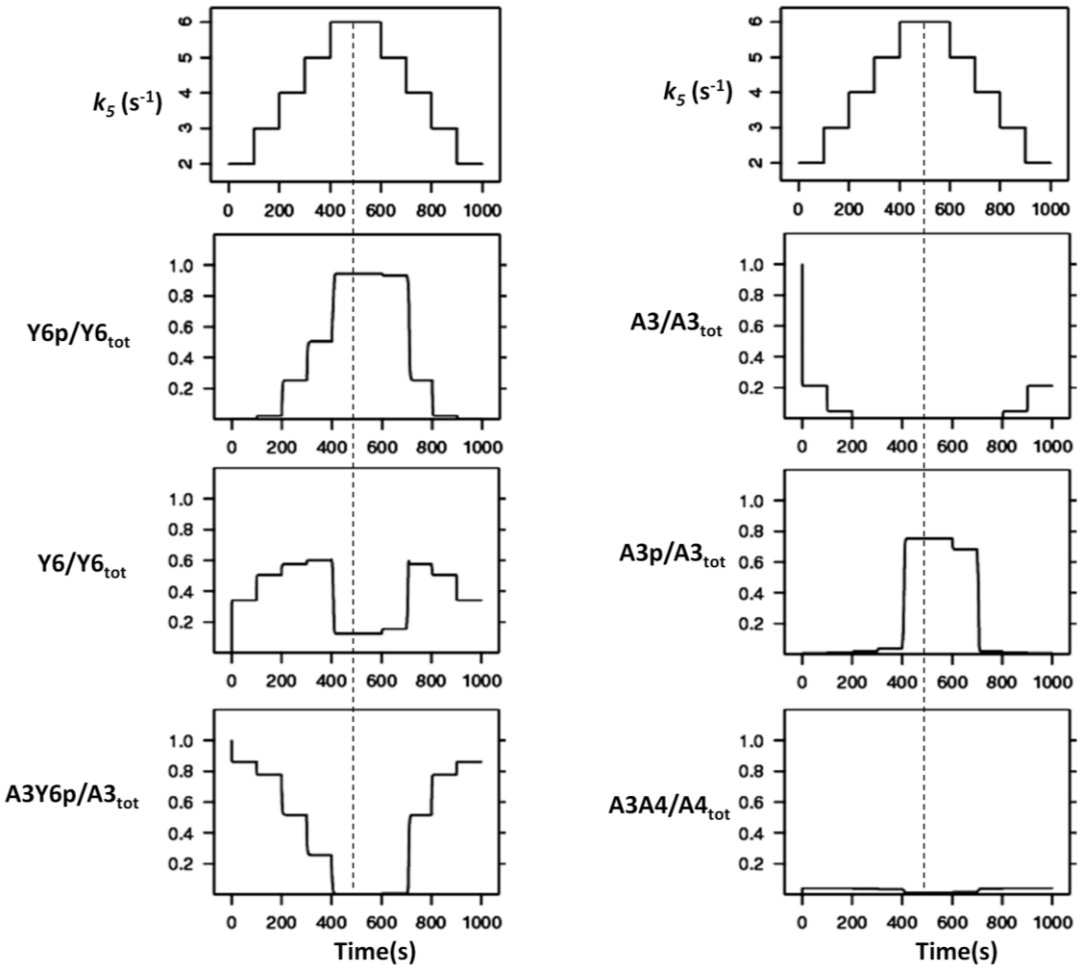

Figure 2.3 Time-course analyses. The model is simulated with increasing and decreasing signal levels (*k5*) in course of time. *k5* is increased from 2 to 6 and decreased in similar fashion at indicated time points (top most, left panel), and changes in each species were measured (as indicated on each panel). The dotted line represents the highest signal level, with equal signal steps on each side of it. The noted asymmetry around this line shows the presence of hysteresis in the system. The x- and y-axis represent time and species concentration respectively, where the latter is normalized by the appropriate total protein levels.

These observations can be understood if we consider the cyclic nature of the reactions in this system as shown in Figure 2.1. The free CheA3 can be seen as a branching point in the system, with one branch leading to binding to CheA4 and ultimately to more CheY6

phosphorylation (phosphorylation branch), while the other leading to binding to CheY6-P and subsequent dephosphorylation (dephosphorylation branch). While the phosphorylation branch is regulated externally of the system by signals sensed by the cytoplasmic cluster (i.e. through altering *k3* and/or *k5*), the dephosphorylation branch is controlled internally by the covalent modification of CheY6. This results in a dynamical motif that is similar to that seen in metabolic branching points and that can embed ultrasensitivity [33]. The split kinase system can embed a high level of nonlinearity as it contains both an inverse coupling of the two branches themselves (via CheY6) and their regulation (via CheA3). At low signals, these two branches allow enough free CheA3 in the system so to result in equally fast phosphorylation and dephosphorylation of CheY6. As the signal increases, however, the rate of the phosphorylation branch increases, while at the same time shutting down the dephosphorylation branch. In other words, the phosphorylation and dephosphorylation branches are coupled inversely, such that the kinase activity cannot be increased without reducing the phosphatase activity and vice versa. These dynamics can be observed in Figure 2.3; the loss of free CheA3 in the system coincides with an abrupt increase in CheA3-P and CheY6-P, while the CheA3:CheA4 complex maintains a fast turnover. This dynamical picture also explains the parameter effects observed in Figure 2 (and Figures S2.1-S2.4 in Appendix A). For example, the decrease in ultrasensitivity from the reduction of CheA3-CheA4 association rate constant (*k1*) can be explained by a slowing down of the phosphorylation branch. Similarly, the decrease in ultrasensitivity from the inclusion of additional phosphatase activity via species other than free CheA3 can be explained by its perturbing effects on the inverse coupling between the phosphorylation and dephosphorylation branches (Figure S2.4 and S2.5 in Appendix A). It must also be noted that the total level of CheA4 in the cell allows additional (internal) control on the dynamics of the system (Figure 2.2B and Figure S2.3 in Appendix A), through its effects on the phosphorylation branch.

To further test whether the inverse coupling of kinase and phosphatase activities through free CheA3 is the underpinning mechanism of ultrasensitivity, we considered dynamics in two alternative models where such coupling is missing; (i) a bifunctional HK that is not split, and (ii) a traditional HK that is neither bifunctional nor split, with a dedicated auxiliary phosphatase for the phosphorylated RR. An analytical treatment of the dynamics arising in the former scenario suggests that non-split bifunctional HKs (where the phosphorylated/non-phosphorylated HK acts as kinase/phosphatase on its cognate response regulator) gives rise to hyperbolic signal-response relationships and provides the system with robustness towards variations in component concentrations [9]. For the latter scenario (e.g. CheA-CheY-CheZ system found in the E. coli chemotaxis system) we developed a simplified model and solved it for the steady state levels of phosphorylated response regulator. We compared this analytical solution to that derived from a simplified model of a split kinase system (see Text S2.1, section 4 in Appendix A). This analytical treatment shows that the latter displays a higher level of nonlinearity for the steady state expression of phosphorylated RR. More importantly, we find that of the three possible alternative structures - bifunctional and split, monofunctional and split, bifunctional and non-split - only the chemical reaction system arising from the bifunctional and split kinase have the capacity for multiple steady states according to the higher deficiency theorem [32,29] (see Text S2.3-2.6 in Appendix A for detailed results). Taken together, these analytical findings show that for bistable and ultrasensitive dynamics to be realized in a split kinase system, both bifunctionality of the HK and the splitting of these two functionalities (i.e. kinase and phosphatase activity) are needed.

2.2.4 Experimental verification that free CheA3 is a better phosphatase than CheA3:CheA4

As shown above, the ability of the split kinase to achieve both segregation and inverse coupling of kinase and phosphatase activities requires that free CheA3 is the predominant phosphatase with other CheA3 containing species (in particular CheA3:CheA4 and CheA3:CheA4:ATP) showing much lower phosphatase activity. Testing this requirement, or directly the level of ultrasensitivity *in vivo,* is complicated both by the presence of additional components in the system and our lack of knowledge of the signal identity in split kinase systems studied to date. As an alternative, and to achieve an approximate test of our theoretical understanding of split kinase response dynamics, we performed *in vitro* measurements of CheY6-P dephosphorylation in the presence of CheA3 and CheA4. In these experiments we used a purified phosphorylated P1 domain of CheA3 (CheA3P1-P) as the sole phosphodonor in the environment. As CheA3P1-P is known to lack phosphatase activity [7], this setup allows us to test directly the phosphatase activity of free CheA3 and the CheA3:CheA4 complex. If kinase and phosphatase activities are segregated into the complexed and free CheA3 respectively, these measurements should reveal a decrease of phosphatase activity with increasing CheA4 concentration, as this would sequester free CheA3 into the CheA3:CheA4 complex. In contrast, such an effect would be absent if the CheA3:CheA4 complex possessed the same level of phosphatase activity as free CheA3. We found evidence for such a decrease, with increasing CheA4 concentrations reducing the rate of CheA3 mediated dephosphorylation of CheY6-P (Figure 2.4 and Figure S2.8 in Appendix A). To rule out the possibility of any interference from free CheA4, we have also confirmed the lack of dephosphorylation activity by CheA4 (Figure 2.4B). This observation qualitatively matches predictions from a specific model of this *in vitro* experimental setup where we assumed phosphatase activity to be restricted to only free CheA3 (see Text S2.1 in Appendix

A and Figure 4). These experimental findings strongly suggest that the CheA3:CheA4 complex has much lower phosphatase activity than free CheA3.

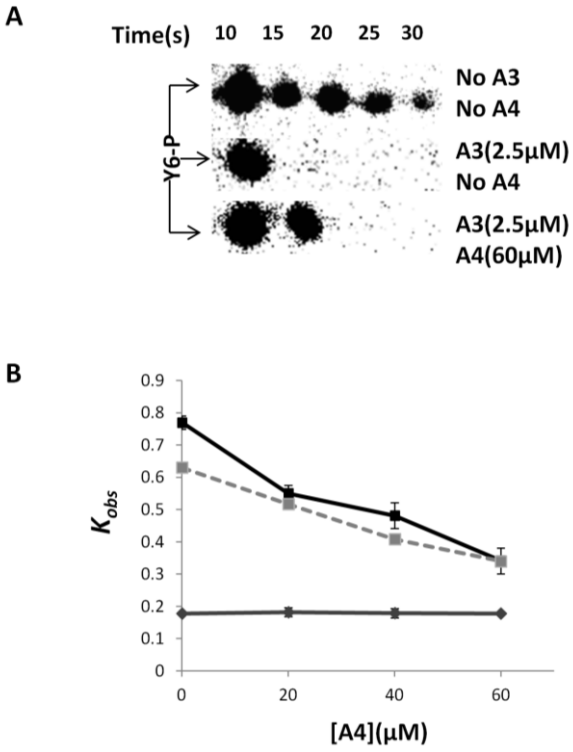

Figure 2.4 Measurement of CheY6-P dephosphorylation rates under different conditions (as indicated). An excess of CheY6 was phosphorylated using CheA3P1-P as phosphodonor. The phosphotransfer reaction was complete within 10s of adding CheY6 to the reaction mixture. Subsequently the decay in CheY6-P levels was followed over time. (A) Phosphorimages showing the decay in CheY6-P levels over time. (B) Graph comparing the observed pseudo-first order rate constant (k_{obs}) for CheY6-P dephosphorylation with and without CheA3 and CheA4. The values predicted by the modeling are shown with a dashed line, while the experimentally measured values are shown in black. Results from a control experiment (without CheA3 and solely CheA4) is shown in grey. Error bars show the standard error of the mean obtained from eight independent experiments.

2.3 Discussion

Two component signaling systems mediate many of the physiological responses of bacteria and display several conserved biochemical and structural features. Here, we analyzed how one such feature, the split kinase, affects response dynamics. Our theoretical treatment proved that the chemical reaction system arising from a bifunctional split kinase gives rise to the possibility of bistability, whereas systems arising from bifunctional, non-split and monofunctional, split kinases lack such capability (unless featuring dead-end complex formation [12]). Sampling the parameter space around kinetic rates and protein concentrations measured in (or estimated from) *R. sphaeroides*, we found that a split kinase system set in a "biologically relevant" parameter regime has potential for an ultrasensitive and bistable signal-response relationship. These nonlinear dynamics arise from the bifunctional and split nature of the kinase, which introduce a branching point into the system between phosphorylation and dephosphorylation reactions. Thus, the level of ultrasensitivity (and emergence of bistability) in the system is determined by the parameters and the biochemical mechanisms found in the reaction cycles linked to this branching point.

We found that the one crucial biochemical aspect enabling ultrasensitivity and bistability in the split kinase system is the predominant allocation of phosphatase activity to the free protein (rather than any of the complexes in the system). Using *in vitro* phosphotransfer assays in the CheA3-CheA4-CheY6 split kinase system isolated from *R. sphaeroides*, we found support for free CheA3 being the principal phosphatase in that system (Figure 2.4). It remains to be shown whether this system enables ultrasensitivity or bistability *in vivo*. The theoretical findings of this study suggest that the switch-like dynamics resulting from ultrasensitivity and bistability could be relevant in the physiological context of the CheA3-CheA4-CheY6 system, which is involved in the integration of cytoplasmic and extracellular signals for proper chemotaxis [7,

34]. It would be plausible for example, if the switching dynamics described here allowed cells to override external chemotaxis signals in favor of internal signals such as those related to metabolism, which could contribute to motility decisions [35-37]. As shown in Figure 2.2, several internal parameters of the system, including the total expression level of CheA4, allow control of the dynamics mediated through CheA3:CheA4 and might enable further tuning of such decision making mechanisms.

While our results highlight split kinases as a potential strategy for implementing ultrasensitivity in bacterial two-component systems, it is not the only one. Previous theoretical studies have found that ultrasensitivity can be achieved in phosphorelays [18, 19], in classical HK-RR systems embedding specific spatial dynamics [38] and in systems with bifunctional HKs, where unphosphorylated HKs and RR form a dead-end complex that is incapable of HK autophosphorylation [12, 39]. These findings suggest that there are several diverse structural, spatial and dynamics that are possible in bacterial two-component systems and that have the potential to enable nonlinear response dynamics. Our theoretical findings extend this list with split kinase systems. Further, we provide experimental support for a condition that increases their potential for generating ultrasensitivity and bistability. Such responses are known to be common in eukaryotes and can enable decision making at the cellular level [40-42]. Thus, it is perhaps not surprising that bacterial signaling systems harbor mechanisms to enable similar levels of ultrasensitivity.

Although rare, split kinases are found in several other bacteria. A recent study looking at CheAs identified 11 split CheAs (2.3 %) versus 470 complete CheAs (97.7%) in fully sequenced non-redundant genomes [1]. In addition to these split CheAs, there is the potential for other HKs to be split where the HisKA (dimerization and histidine phosphotransfer) and

the catalytic HATPase (histidine kinase ATPase) domains are found on separate proteins. *In vitro* studies of the osmosensing histidine kinase, EnvZ, have shown that it possible to split the HATPase and HisKA domains onto separate polypeptides whilst retaining their activity [43]. Interrogation of the SMART database reveals that of the 42417 proteins containing HisKA domains (dimerization and histidine phosphotransferase), 1556 (3.66%) lack a HATPase (histidine kinase ATPase) domain (expect value <0.01) and of these, 711 (1.7%) have the phosphatase sequence motif (HE/DxxN/T;[44] and could therefore be split bi-functional kinases. The results presented here suggest that cells may use such split kinases to allow high sensitivity and bistability enabling switch-like physiological responses to environmental stimuli.

As the highly modular TCSNs are used by bacteria to control many of their physiological responses, it will be valuable to explore other mechanisms which can enable specific response dynamics in these systems and to determine the evolutionary drivers that were responsible for their emergence. This would increase our ability to better understand microbial signaling and exploit it in synthetic biology applications.

2.4 Models and methods

2.4.1 A mathematical model for a split kinase

To model the CheA3-CheA4-CheY6 split kinase system, we considered its dynamics in isolation of other cellular components. The reactions in this system that we have included in the "basic model" are (see also alternative reaction schemes shown in Text S1);

$$A3 + A4 \underset{k_2}{\overset{k_1}{\rightleftarrows}} A3A4$$

$$A3A4 + ATP \underset{k_4}{\overset{k_3}{\rightleftarrows}} A3A4ATP \xrightarrow{k_5} A3p + A4 + ADP$$

$$A3p + Y6 \underset{k_7}{\overset{k_6}{\rightleftarrows}} A3 + Y6p$$

$$Y6p \xrightarrow{k_8} Y6 + P_i$$

$$A3 + Y6p \underset{k_{10}}{\overset{k_9}{\rightleftarrows}} A3Y6p \xrightarrow{k_{11}} A3 + Y6 + P_i$$

where A3, A4, Y6 stand for CheA3, CheA4 and CheY6 respectively and the -p suffix represents phosphorylated forms of these proteins. Variant models which include additional CheY6-P de-phosphorylation reactions involving alternative phosphatases such as CheA3-P, and CheA3:CheA4 complex are shown in supplementary text S1, and their effects are analyzed in Figure 2D and S4. The above "basic model" reaction scheme can be used to derive a system of ordinary differential equations (ODEs), which describe the changes in concentrations of proteins over time;

$$\frac{d[A3p]}{dt} = k_5 \cdot [A3A4ATP] + k_7 \cdot [A3] \cdot [Y6p] - k_6 \cdot [A3p] \cdot [Y6p]$$

$$\frac{d[A3A4]}{dt} = k_1 \cdot [A3] \cdot [A4] + k_4 \cdot [A3A4ATP] - [A3A4] \cdot \left(k_2 + k_5 \cdot [ATP]\right)$$

$$\frac{d[A3A4ATP]}{dt} = k_3 \cdot [A3A4] \cdot [ATP] - [A3A4ATP] \cdot \left(k_4 + k_5\right)$$

$$\frac{d[A3Y6p]}{dt} = k_9 \cdot [A3] \cdot [Y6p] - [A3Y6p] \cdot \left(k_{10} + k_{11}\right)$$

$$\frac{d[Y6p]}{dt} = k_{10} \cdot [A3Y6p] + k_6 \cdot [A3p] \cdot [Y6] - [Y6p] \cdot \left(k_7 \cdot [A3] + k_8 + k_9 \cdot [A3]\right)$$

In addition, we have a set of three conservation equations;

$$[Y6]_{tot} = [Y6] + [Y6p] + [A3Y6p]$$

$$[A3]_{tot} = [A3] + [A3p] + [A3A4] + [A3A4ATP] + [A3Y6p]$$

$$[A4]_{tot} = [A4] + [A3A4] + [A3A4ATP]$$

To analyze the behavior of the split kinase motif with increasing signal, we simulated the incoming signals from receptors as an increase in the autophosphorylation rate of the kinase (*k5*). The model was parameterized with data from literature (see Table 2.1). In the case of the

dephosphorylation of CheY6-P by CheA3, we derived the relevant parameters ($k9$, $k1_0$, and $k1_1$) through fitting simulation data to previously published *in vitro* dephosphorylation measurements [7]. Fitting was done using a hybrid genetic algorithm (functions ga and fmincon from the MATLAB Global Optimization Toolbox).

We numerically integrated the model to derive time course and steady state signal-response relationships. The latter analysis gives the steady state CheY6-P level at a given signal ($k5$) where signal was taken as the rate of autophosphorylation of split kinase and allows deriving a so-called signal-response curve. This curve is found by numerically integrating the system to steady state at a fixed signal level and then numerically "following" this steady state (i.e. steady state CheY6-P level), while changing the signal. This analysis is equal to allowing the system to reach steady state under different signal values. Both time course and signal-response analyses were performed using the software packages XPPAUT (http://www.math.pitt.edu/~bard/xpp/xpp.html) and Oscill8 (http://oscill8.sourceforge.net/).

2.4.2 Sensitivity analysis. We have quantified the sensitivity of the shape of the signal-response curves to variations in each of the parameters from their described base values (Table 2.1) and in a biologically relevant range. For these analyses, we measured the "sigmoidality" of the signal-response curve, RS, as its maximum slope (s_{max}) multiplied by the signal level at which this slope occurs ($k5_s$) (i.e. RS = $k5_s \times s_{max}$). This measure is similar to the "response coefficent", which measures the slope between 90% and 10% saturation [33], but is better able to distinguish between hyperbolic and sigmoidal dose-response curves. For each parameter, we varied it in a wide range around its basic value and measured "sigmoidality" of the resulting dose-response curves, as well as the maximum response of the system (Figures S2.1-

S2.3 in Appendix A). The same analysis is also applied for alternative models featuring additional phosphatase species (Figure S2.4 in Appendix A).

2.4.3 Analytical comparison of different models. To perform a formal check for the potential of bistability in the different models (discussed in the main text and Supplementary Information), we have utilized the chemical network theory [28, 29]. This theory provides several analytical tests that can provide a definite answer on the possibility of existence of multiple stationary states in a given reaction network. We have applied these tests to the basic and alternative models we had devised using the Chemical Network Tool v2.2 (http://www.chbmeng.ohio-state.edu/~feinberg/crntwin/). The model files used with this tool and describing the chemical reaction systems, as well as the analytical results from the tool are provided as Text S2.2-2.4 in Appendix A.

2.4.4 Plasmid and strains. See Table 2.2 for the plasmids and strains used. *E. coli* strains were grown in LB medium at 37^{o}C. Antibiotics were used at concentrations of 100 μg ml-1 for ampicillin and 25 μg ml-1 for kanamycin, where needed. *E. coli* M15pRep4 cells were made competent using the calcium chloride technique [45]. Transformations were performed according to [46].

Strains/plasmid	Description	Source/Reference
E.coli strains M15pREP4	Expression host containing pREP4; kanamycin resistant	Qiagen
pQE30	IPTG inducible expression vector. Introduces RGS(H)6 at the N terminus of the expressed protein. Confers ampicillin resistance	Qiagen
pQE60	IPTG inducible expression vector. Introduces RGS(H)6 at the C terminus of the expressed protein. Confers ampicillin resistance	Qiagen
pQE60A3P1	Plasmid for overexpressing C-terminally His-tagged CheA3P1 from *R.sphaeroides*. pQE60 derivative	[7]
pQEY6	Plasmid for overexpressing C-terminally His-tagged CheY6 from *R.sphaeroides*. pQE30 derivative	[24]
pQEA3	Plasmid for overexpressing C-terminally His-tagged CheA3 from *R.sphaeroides*. pQE30 derivative	[21]
pQEA4	Plasmid for overexpressing C-terminally His-tagged CheA4 from *R.sphaeroides*. pQE30 derivative	[21]

Table 2.2: Plasmids and strains used and the associated literature source.

2.4.5 Protein purification. His tagged *R. sphaeroides* CheA3, CheA4, CheA3P1 and CheY6 proteins were purified as described previously [47]. Protein purity and concentration was measured as described in [24]. Purified proteins were stored at -20^{o}c.

2.4.6 Preparation of CheA3P1-^{32}P. CheA3P1 was phosphorylated using [γ-^{32}P] ATP and CheA4 and purified as described before with the following modifications [7]. Proteins were phosphorylated in reactions performed at 20°C in phosphotransfer buffer (50 mM Tris HCl, 10% (v/v) glycerol, 5 mM $MgCl_2$, 150 mM NaCl, 50 mM KCl, 1 mM DTT, pH 8.0). The final reaction volumes were 2 ml. For production of CheA3P1-^{32}P, reaction mixtures contained 300 μM CheA3P1 and 20 μM CheA4. Reactions were initiated by addition of 2 mM [γ-^{32}P] ATP (specific activity 14.8 GBq mmol^{-1}; PerkinElmer). After 1 hour incubation, samples were purified by using Ni-NTA columns (Qiagen) as described previously for unphosphorylated His-tagged CheA3 [47]. This purification step removed the unincorporated ATP and also removed the CheA4 protein from the CheA3P1-^{32}P preparation. Purified proteins were stored at -20°C.

2.4.7 Measurement of CheY6-P dephosphorylation rate. Assays were performed at 20 ^{o}C in phosphotransfer buffer. Purified CheA3P1-^{32}P was used as the phosphodonor. An excess of CheY6 (100 μM) was added to 30 μM of purified CheA3P1-^{32}P in the presence of 2.5 μM CheA3 and 0-60 μM CheA4. Following the addition of CheY6, reaction aliquots of 10 μl were taken at the indicated time points and quenched immediately in 10 μl of 2 X SDS-PAGE loading dye(7.5% (w/v) SDS, 90 mM EDTA, 37.5 mM Tris HCl, 37.5% glycerol, 3% (v/v) β-

mercaptoethanol, pH 6.8). Quenched samples were analyzed using SDS-PAGE and phosphorimaging as described previously [24].

2.5 References

1 Wuichet K, Zhulin IB (2010) Origins and Diversification of a Complex Signal Transduction System in Prokaryotes. Sci Signal 3: ra50.

2 Hamer R, Chen PY, Armitage JP, Reinert G, Deane CM (2010) Deciphering chemotaxis pathways using cross species comparisons. BMC Syst Biol 4: 3.

3 Stock AM, Robinson VL, Goudreau PN (2000) Two-component signal transduction. Annu Rev Biochem 69: 183-215.

4 Batchelor E, Goulian M (2003) Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system. Proc Natl Acad Sci U S A 100: 691-696.

5 Sourjik V, Schmitt R (1998) Phosphotransfer between CheA, CheY1, and CheY2 in the chemotaxis signal transduction chain of Rhizobium meliloti. Biochemistry 37: 2327-2335.

6 Hoch JA (2000) Two-component and phosphorelay signal transduction. Curr Opin Microbiol 3: 165-170.

7 Porter SL, Roberts MAJ, Manning CS, Armitage JP (2008) A bifunctional kinase-phosphatase in bacterial chemotaxis. Proc Natl Acad Sci U S A 105: 18531-18536.

8 Soyer OS, Creevey CJ (2010) Duplicate retention in signalling proteins and constraints from network dynamics. J Evol Biol. 11: 2410-21.

9 Shinar G, Milo R, Martínez MR, Alon U (2007) Input output robustness in simple bacterial signaling systems. Proc Natl Acad Sci U S A 104: 19931-19935.

10 Alves R, Savageau MA (2003) Comparative analysis of prototype two-component systems with either bifunctional or monofunctional sensors: differences in molecular structure and physiological function. Mol Microbiol 48: 25-51.

11 Groban ES, Clarke EJ, Salis HM, Miller SM, Voigt CA (2009) Kinetic buffering of cross talk between bacterial two-component sensors. J Mol Biol 390: 380-393.

12 Igoshin OA, Alves R, Savageau MA (2008) Hysteretic and graded responses in bacterial two-component signal transduction. Mol Microbiol 68: 1196-1215.

13 Tiwari A, Ray JC, Narula J, Igoshin OA (2011) Bistable responses in bacterial genetic networks: designs and dynamical consequences. Math Biosci 231: 76-89.

14 Jiménez-Pearson MA, Delany I, Scarlato V, Beier D (2005) Phosphate flow in the chemotactic response system of Helicobacter pylori. Microbiology 151: 3299-3311.

15 Bischofs IB, Hug JA, Liu AW, Wolf DM, Arkin AP (2009) Complexity in bacterial cell-cell communication: quorum signal integration and subpopulation signaling in the Bacillus subtilis phosphorelay. Proc Natl Acad Sci U S A 106: 6459-6464.

16 Burbulys D, Trach KA, Hoch JA (1991) Initiation of sporulation in Bacillus subtilis is controlled by a multicomponent phosphorelay. Cell 64: 545-552.

17 Appleby JL, Parkinson JS, Bourret RB (1996) Signal transduction via the multi-step phosphorelay: not necessarily a road less traveled. Cell 86: 845-848.

18 Kim J-RR, Cho K-HH (2006) The multi-step phosphorelay mechanism of unorthodox two-component systems in E. coli realizes ultrasensitivity to stimuli while maintaining robustness to noises. Comput Biol Chem 30: 438-444.

19 Csikász-Nagy A, Cardelli L, Soyer OS (2010) Response dynamics of phosphorelays suggest their potential utility in cell signalling. J R Soc Interface 8(57): 480-8.

20 Porter SL, Wadhams GH, Armitage JP (2011) Signal processing in complex chemotaxis pathways. Nat Rev Microbiol 9: 153-165.

21 Porter SL, Armitage JP (2004) Chemotaxis in *Rhodobacte sphaeroides* requires an atypical histidine protein kinase. J Biol Chem 279: 54573-54580.

22 Porter SL, Wadhams GH, Martin AC, Byles ED, Lancaster DE, et al. (2006) The CheYs of *Rhodobacte sphaeroides*. J Biol Chem 281: 32694-32704

23 Wadhams GH, Warren AV, Martin AC, Armitage JP (2003) Targeting of two signal transduction pathways to different regions of the bacterial cell. Mol Microbiol 50: 763-770.

24 Porter SL, Warren AV, Martin AC, Armitage JP (2002) The third chemotaxis locus of *Rhodobacte sphaeroides* is essential for chemotaxis. Mol Microbiol 46: 1081-1094.

25 Scott KA, Porter SL, Bagg EAL, Hamer R, Hill JL, et al. (2010) Specificity of localization and phosphotransfer in the CheA proteins of *Rhodobacte sphaeroides*. Mol Microbiol 76: 318-330.

26 Bell CH, Porter SL, Strawson A, Stuart DI, Armitage JP (2010) Using structural information to change the phosphotransfer specificity of a two-component chemotaxis signalling complex. PLoS Biol 8: e1000306.

27 Porter SL, Armitage JP (2002) Phosphotransfer in *Rhodobacte sphaeroides* chemotaxis. J Mol Biol 324: 35-45.

28 Craciun, G., Y. Z. Tang, and M. Feinberg.(2006) Understanding bistability in complex enzyme-driven reaction networks. Proc. Natl Acad Sci USA 103:8697-8702.

29 Shinar G. & Feinberg M. (2012) Concordant Chemical Reaction Networks and the Species-Reaction Graph. Mathematical biosciences. Aug 21, epub ahead of print.

30 Tyson JJ, Chen KC, Novak B (2003) Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Curr Opin Cell Biol 15: 221-231.

31 Harrison DM, Skidmore J, Armitage JP, Maddock JR (1999) Localization and environmental regulation of MCP-like proteins in *Rhodobacte sphaeroides*. Mol Microbiol 31: 885-892.

32 Feinberg, M. (1987) Chemical reaction network structure and the stability of complex isothermal reactors. I. The deficiency zero and deficiency one theorems. Chem. Eng. Science, 42, 2229-2268.

33 LaPorte DC, Walsh K, Koshland DE (1984) The branch point effect. Ultrasensitivity and subsensitivity to metabolic control. J Biol Chem 259: 14068-14075.

34 Tindall MJ, Porter SL, Maini PK, Armitage JP (2010) Modeling chemotaxis reveals the role of reversed phosphotransfer and a bi-functional kinase-phosphatase. PLoS Comput Biol 6.

35 Goldstein RA, Soyer OS (2008) Evolution of taxis responses in virtual bacteria: non-adaptive dynamics. PLoS Comput Biol 4: e1000084.

36 Egbert MD, Barandiaran XE, Di Paolo EA (2010) A minimal model of metabolism-based chemotaxis. PLoS Comput Biol 6: e1001004.

37 Soyer OS, Goldstein RA (2011) Evolution of response dynamics underlying bacterial chemotaxis. BMC Evol Biol 11: 240.

38 van Albada SB, Ten Wolde PR (2009) Differential affinity and catalytic activity of CheZ in E. coli chemotaxis. PLoS Comput Biol 5: e1000378.

39 Salvado B, Vilaprinyo E, Karathia H, Sorribas A, Alves R (2012) Two component systems: physiological effect of a third component. PLoS One. 7(2):e31095.

40 Goldbeter A, Koshland DE (1981) An amplified sensitivity arising from covalent modification in biological systems. Proc Natl Acad Sci U S A 78: 6840-6844.

41 Huang CYF, Ferrell JE (1996) Ultrasensitivity in the mitogen-activated protein kinase cascade. Proc Natl Acad Sci U S A 93: 10078-10083.

42 Kholodenko BN (2006) Cell-signalling dynamics in time and space. Nature Reviews Molecular Cell Biology 7: 165-176.

43 Park H, Saha SK, Inouye M (1998) Two-domain reconstitution of a functional protein histidine kinase. Proc Natl Acad Sci U S A 95: 6728-6732.

44 Huynh TN, Stewart V (2011) Negative control in two-component signal transduction by transmitter phosphatase activity. Mol Microbiol 82: 275-286.

45 Sambrook J, Russell JB (2001) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press.

46 Hanahan D (1983) Studies on transformation of Escherichia coli with plasmids. J Mol Biol 166: 557.

47 Porter SL, Wadhams GH, Armitage JP (2007) *In vivo* and *in vitro* analysis of the *Rhodobacte sphaeroides* chemotaxis signaling complexes. Methods Enzymol 423: 392-413.

48 Silversmith RE, Levin MD, Schilling E, Bourret RB (2008) Kinetic characterization of catalysis by the chemotaxis phosphatase CheZ. Modulation of activity by the phosphorylated CheY substrate. J Biol Chem 283: 756-765.

CHAPTER 3

Phosphate sink containing two-component signaling systems as tunable threshold devices

3.1 Introduction

Cells process external cues in order to produce appropriate responses that ensure survival and efficient proliferation. They achieve this goal through a myriad of signaling and gene regulatory networks, which implement specific signal processing capabilities such as switch-like threshold dynamics, logic gates, oscillations, and noise filtering (1-8). Understanding the architecture and response dynamics of these systems is of fundamental value, providing us with a better insight into cell biology and allowing us to engineer *de novo* biological systems. The field of synthetic biology exploits the understanding and components from natural systems to rationally design synthetic systems that implement specific signaling dynamics. So far, this led to the development of oscillatory systems (9, 10), systems with threshold dynamics (1, 11-12) and logic gates (13-15). In most cases, these studies use transcriptional regulation to implement the desired dynamics, while a few studies have explored the possibility of extending synthetic design approaches to signaling networks (16-18).

Bacterial systems are particularly attractive for attempting synthetic engineering of signaling networks. All bacteria and certain eukaryotic microbes and plants utilize the so-called two-component signaling systems for signal transduction (19-21). In their most simple implementation, these systems consist of a histidine protein kinase (HK) and a response regulator (RR). The activity of the HK is controlled by an environmental stimulus, which controls the rate of autophosphorylation. Once phosphorylated, the HK transfers its phosphoryl group to a cognate RR, which in its phosphorylated form mediates the output of

the signaling pathway (21). The phosphotransfer reaction is at the core of all two-component systems, and regulating its specificity could allow direct control over microbial (and to some extent plant) physiology, as well as creating synthetic signaling systems. Thus, several studies have attempted to decipher the coupling specificity of HK and RR proteins (22-25) and have generated chimeric HKs with specified and controllable inputs (26-30). More recently, scaffolding of HK and RR proteins has been shown to allow significant control over the phosphotransfer specificity (17).

Generating synthetic systems with specified signal processing capabilities, however, requires a deeper understanding of system properties such as the signal-response relationship they embed. Around the core HK-RR interaction, different two-component systems have diverse architectures, which could underpin specific signal processing capabilities. For example, the commonly observed phosphorelays, where the flow of phosphoryl groups from the HK to the RR is relayed through several proteins, are believed to allow signal integration and specific response properties such as control of noise and ultrasensitivity (31-35). Other architectural features such as presence of a bifunctional or a split HK, formation of specific "dead-end" complexes and also transcriptional feedbacks have been shown to allow ultrasensitivity and even bistability (36-38). Of such different architectural features, one that has not attracted much attention is the "sink" system, where two RRs can compete for the phosphoryl group from a single HK. This architectural motif has been identified in several microbial and plant systems (39-43). In the *Sinorhizobium meliloti* chemotaxis pathway, the two response regulators CheY1 and CheY2 are phosphorylated by their cognate kinase CheA. Of these, only CheY2 in its phosphorylated form can bind to the flagellar motor and control its rotation (39). Both CheYs can also perform reverse phosphotransfer, where they return the phosphoryl group to CheA. Given its high phosphorylation rate (from HK), low reverse phosphorylation

rate (to HK), and the observation that the *S. meliloti* chemotaxis system lacks a dedicated phosphatase, it is proposed that CheY1 acts as a sink that accelerates dephosphorylation of CheY2 (39). A similar situation is described in the *Rhodobacter sphaeroides* and *Helicobacter pylori* chemotaxis pathways (40, 41) and the yeast osmoregulation pathway (43, 44). In the latter case, the HK, SLN1 autophosphorylates in response to changes in the membrane structure and phosphorylates two downstream RRs, SSK1 and SKN7. *In vitro* phosphotransfer studies found similar dynamics as in the *S. meliloti* chemotaxis pathway with SKN7 displaying significant reverse phosphotransfer to SLN1, while SSK1 showing no such activity (44). Interestingly, both SSK1 and SKN7 are functionally active in this system, with SSK1 activating the downstream HOG1 MAP kinase cascade (45,46) and SKN7 acting as a transcription factor for genes involved in various stress related responses (47,48).

Here, we use mathematical and experimental approaches to identify the full signal processing capabilities of this two-component system. We first develop a generic model of the one HK – two RR motif and perform both analytical and simulation-based analyses. These reveal that this system is capable of both enhancing signal termination time and implementing a threshold signal-response relationship, i.e. the system displays a sigmoidal signal-response relationship in which the steady state levels of the phosphorylated output RR remains low until a threshold level of signal is crossed. We then verify these dynamics experimentally by *in vitro* re-constitution of the two-component proteins from the chemotaxis pathway of *S. meliloti*. Using this *in vitro* setup, we further show that specific properties of the threshold dynamics can be controlled through the concentrations of the core components, as well as through presence of an auxiliary protein that is known to bind the HK in *S. meliloti* (49). These findings allow better understanding of the physiological responses mediated by phosphate sink-containing

two-component systems in microbes and plants, and will facilitate design of synthetic threshold devices using two-component signalling proteins.

3.2 Results

3.2.1 Analysis of response dynamics in the one HK – two RR motif

While the implementation of the phosphate sink motif in diverse two-component systems could differ in the molecular details of the proteins involved and their exact kinetic rates, the sink mechanisms can be formulated as a general architectural motif (Figure 3.1A and S1A); a two-component system comprising a single HK and two RRs, namely the output-RR and the sink-RR (as referred to, in the rest of the text). We have developed a generic model of this motif and parameterized it using experimental measurements from the reaction kinetics of the *S. meliloti* chemotaxis and yeast osmoregulation systems (see *Methods*). To monitor temporal dynamics in the presence of a signal, we simulated two conditions, one with the sink-RR and one without the sink-RR. Using the "controlled comparison" approach (50), we simulated each scenario at a signal level that resulted in 90% phosphorylation of the output-RR at steady state. The signal was then removed and the half-time for the decay of phosphorylated output-RR measured. We found that under the experimentally measured parameters, the presence of the sink-RR decreases the half-time for the output-RR dephosphorylation by more than 2 fold in both *S. meliloti* and yeast (Figure 3.1B and S3.1B in Appendix B). These simulation results are consistent with previous experimental results (39), which led to the sink hypothesis, and show that in the experimentally observed parameter regime, a sink-RR can accelerate the dephosphorylation of the output-RR.

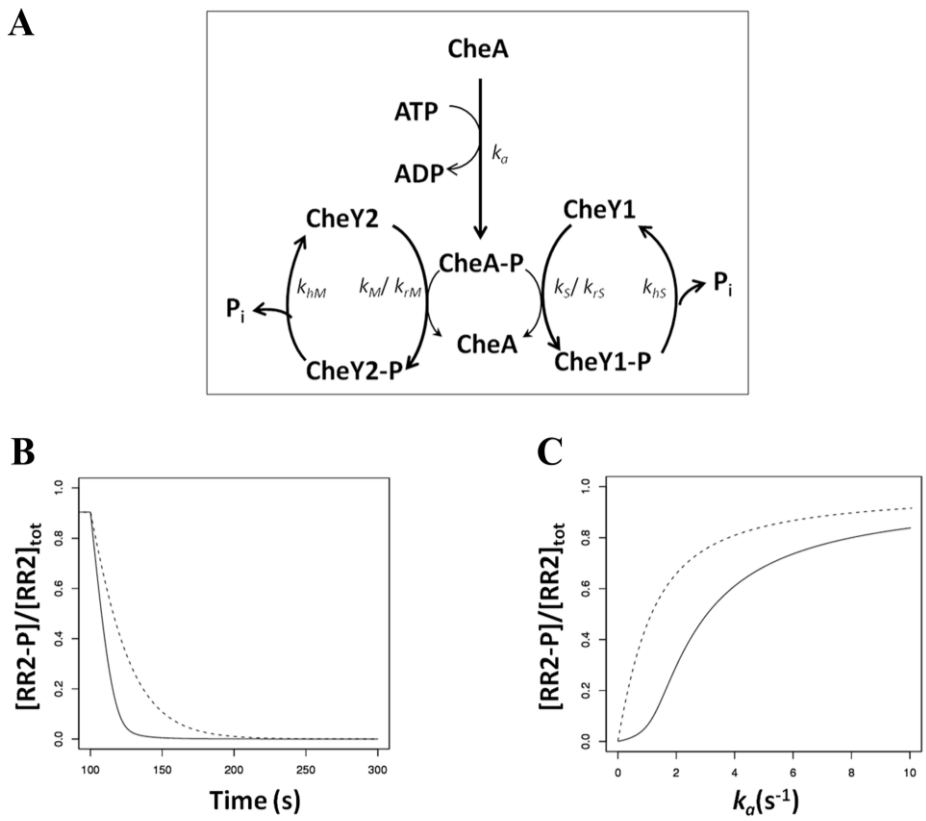

Figure 3.1 The one HK – two RR motif as seen in the *S. meliloti* chemotaxis signaling pathway (**A**) A cartoon diagram of the CheA-CheY1-CheY2 system. The diagram is arranged to highlight the role of CheY1 as a phosphate sink for CheY2. Rate constants are shown on the relevant reactions. In the case of reversible reactions, two rate constants are given as $k_{forward}$ and $k_{reverse}$. (**B)** Role of the sink, RR1 (CheY1) in signal termination (i.e. dephosphorylation of RR2 (CheY2)). The x- and y-axis show the time and the corresponding steady state levels of phosphorylated RR2, respectively. A value of k_a was selected that resulted in ~90% of the total RR2 being phosphorylated at steady state. At t=0, k_a was reduced to zero and the progress of the reaction to the new steady state simulated. The solid line represents the presence of the sink, while the dashed line shows the absence of the sink. **(C)** Signal-response relationship in the presence (solid line) and absence (dashed line) of sink, RR1 (CheY1). The x- and y-axis

show the signal (k_a) and the corresponding steady state level of phosphorylated RR2 (CheY2), respectively.

3.2.2 The one HK – two RR motif can exhibit a sigmoidal signal-response relationship

Besides temporal dynamics, another key characteristic of any signaling system is the signal-response relationship it implements, i.e. the steady state output of the system for any given signal level (51). Focusing again on experimentally measured parameters, we found that the presence of the sink-RR changes the signal-response relationship in the system from hyperbolic to sigmoidal (Figures 3.1C and S3.1C in Appendix B). In other words, the presence of the sink-RR allows threshold dynamics in these natural systems, whereby the steady state level of the phosphorylated output-RR remains low until a threshold signal level is reached. At the threshold point, the steady state level of phosphorylated output-RR is highly sensitive to small changes in signal.

To understand better whether the sensitivity and threshold levels in the sigmoidal signal-response curve can be controlled, and by which parameters, we performed a sensitivity analysis around experimentally measured kinetic rates from *S. meliloti* and yeast (Figures 3.2, S3.4, S3.5 and S3.2 in Appendix B). This revealed several kinetic features for ensuring a sigmoidal signal-response relationship (see below for exact necessary conditions). For example, we found that a key kinetic feature is for phosphotransfer to the sink-RR (parameter k_S) to be faster than reverse phosphotransfer from the sink-RR back to the HK (parameter k_{rS}). Under this condition, the steady state phosphorylation level of output-RR remains low until the sink-RR is almost fully phosphorylated (Figure S3.3 in Appendix B), resulting in a high level of sigmoidality in the signal-response curve (Figures 3.2A and S3.2A in Appendix B).

We also found that both the sharpness of the sigmoidal signal-response relationship and the threshold signal level can be controlled through changes in parameters. In particular, the phosphotransfer rate constant between the HK and sink-RR (Figures 3.2A and S3.2A in Appendix B), and the autodephosphorylation rate constant of the sink-RR (Figures S3.4 and S3.2 in Appendix B) can affect the sharpness of the signal-response curve, while the threshold signal level is determined by the amount of sink present (Figures 3.2B and S3.2B in Appendix B).

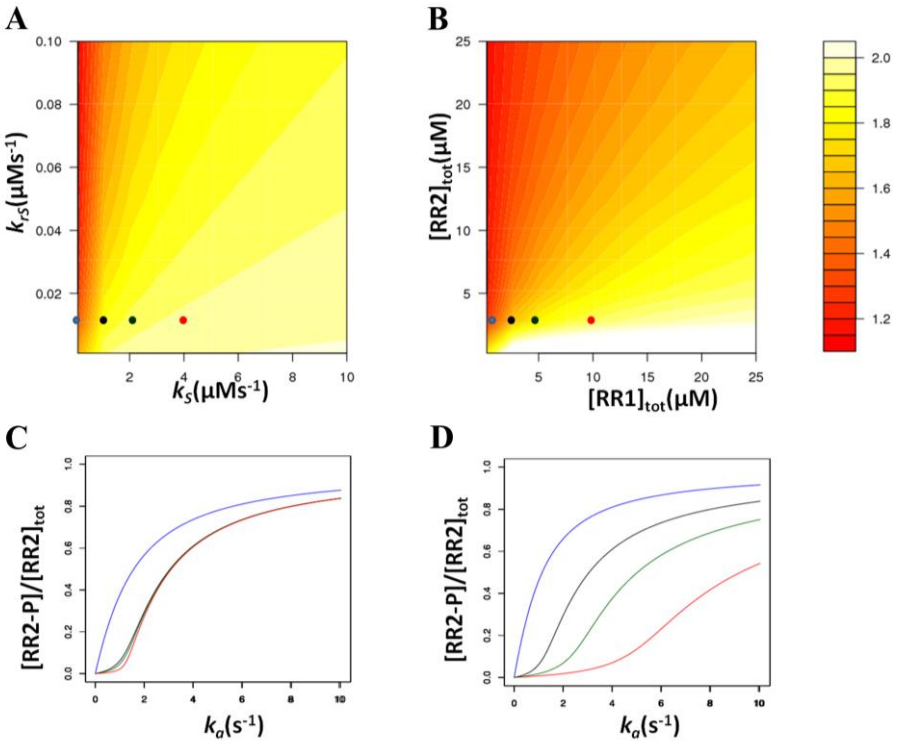

Figure 3.2 The effect of parameter changes on the "sigmoidality" of the signal-response curve. The level of sigmoidality, Hill coefficient, is shown as a heat map on each panel. (**A**) Effect of varying the forward and reverse phosphotransfer rates for the sink RR (CheY1; x-axis; k_S and y-axis; k_{rS}). (**B**) Effect of varying the total concentration of the output RR (CheY2; y-axis) and sink RR (CheY1; x-axis). (**C** and **D**) Signal-response curves for models

corresponding to parameter values indicated as colored circles on the heat maps, in which the black circle represents the basic model and the other circles show the models with varying parameter values.

3.2.3 Necessary conditions for the one HK – two RR motif to exhibit sigmoidal signal-response relationships

To understand more completely the effects of parameters on the signal-response curve, we derived an analytical description for this curve and computed its second derivative at zero (see Appendix B section 1). The second derivative at zero can be used as an indicator of sigmoidal or hyperbolic nature of the signal-response relationship; a hyperbolic shape of the signal-response curve implies that the second derivative of this function is constantly negative on its domain (i.e. positive signals), while a sigmoidal shape implies that the second derivative is initially positive and then it changes sign. Thus, the sign of the second derivative of the signal-response curve at zero can be taken as a test for sigmoidality (52). Using this approach we found two *necessary* conditions on the parameters of the system for achieving a sigmoidal signal-response relationship (i.e. conditions that are required for a positive second derivative at zero): (*i*) $k_S \cdot k_{hS} \cdot [RR1]_{tot} \neq 0$ *and (ii)* $k_S > k_{rS}$, where k_{hS} is the autodephosphorylation rate constant of the sink-RR, $[RR1]_{tot}$ is the total amount of sink-RR, and k_S (k_M) and k_{rS} (k_{rM}) are the forward and reverse phosphotransfer rate constants of the sink-RR (output-RR) respectively (see *Supplementary Information*). The first condition shows that the sink-RR is *necessary* for the system to exhibit sigmoidality. Provided these two conditions are satisfied and, additionally $k_{rS}/k_S < k_{rM}/(k_M+k_{rM})$, having high concentrations of the HK and the sink-RR (i.e. the RR with no/weak reverse phosphotransfer to the HK), and low concentration of the output-RR further ensures sigmoidality. It is important to note that experimentally measured values from both the *S. meliloti* chemotaxis and yeast osmoregulation systems fit with these

analytical conditions for sigmoidality (see Tables 1 and S1). We found that these analytical results on the necessary conditions for the sigmoidality of the signal-response relationship are further simplified when assuming complex formation in the phosphotransfer reactions (*Supplementary Information*). In particular, the second condition (i.e. of having $k_S/k_{rS} > 1$) is not a strict requirement for the second derivative of the signal-response curve at zero to attain a positive value. In this extended model, the second necessary condition becomes either $k_S/k_{rS} > 1$ or $k_S/k_{rS} > (k_{yM}-k_{yS}) / k_{yrS}$, where k_{yM}, k_{yS}, k_{yrS} are the inverse of the Michaelis-Menten constants of the added complexes in the forward phosphotransfer reactions of the sink-RR and output-RR, and the reverse phosphotransfer reaction of the sink-RR, respectively (see *Supplementary Information*). We conclude that for sigmoidality to arise, the quotient k_S/k_{rS} must be larger than some quantity that depends on the parameters of the system and, further, sigmoidality cannot arise simply by the introduction of complex forming reactions in a system without a sink-RR.

The finding that achieving a sigmoidal signal-response relationship for the single HK- two RR system is facilitated by the presence of complexes, prompted us to use the chemical reaction network toolbox (53) to analytically assess the potential of bistability. We found that when the phosphotransfer reactions are modelled as bi-molecular reactions, the system is not capable of bistability (see in Appendix B*)*. However, when considering complex formation and alternative reaction schemes involving the different possible binding events among the HK, the two RRs and their complexes, we found that a certain scenario allows for the presence of bistability in the system (see in Appendix B). In this scenario, the HK can bind to both of the RRs, irrespective of its own phosphorylation state and the phosphorylation states of the two RRs. The resulting system contains four complexes between the

phosphorylated/unphosphorylated HK and the phosphorylated/unphosphorylated RRs, and can permit bistability under certain parameter regimes (see in Appendix B).

3.2.4 Experimental verification of the sigmoidal signal-response relationship in a one HK – two RR motif

To test the model findings experimentally, we re-constituted *in vitro* the CheA, CheY1 (sink-RR) and CheY2 (output-RR) proteins from *S. meliloti*. *In vivo*, CheA kinase activity is controlled by interaction with the signaling domain of chemoreceptor proteins (54). Since it is experimentally difficult to re-constitute chemoreceptors in the *in vitro* system, we varied the kinase activity of CheA by varying the concentration of its substrate, ATP, as a proxy for the *in vivo* signal. This allowed us to monitor the steady state levels of phosphorylated CheY1 and CheY2 at different levels of kinase activity, i.e. to derive an experimental signal-response curve. We found excellent quantitative agreement between the signal-response curves resulting from the model and experiments. In the presence (absence) of CheY1, the steady state levels of phosphorylated CheY2 displayed a sigmoidal (hyperbolic) relation with increasing ATP levels (Figure 3.3). Thus, these experiments strongly suggest that the *S. meliloti* one HK – two RR motif displays a sigmoidal signal-response relationship *in vivo* and could potentially function as a threshold device.

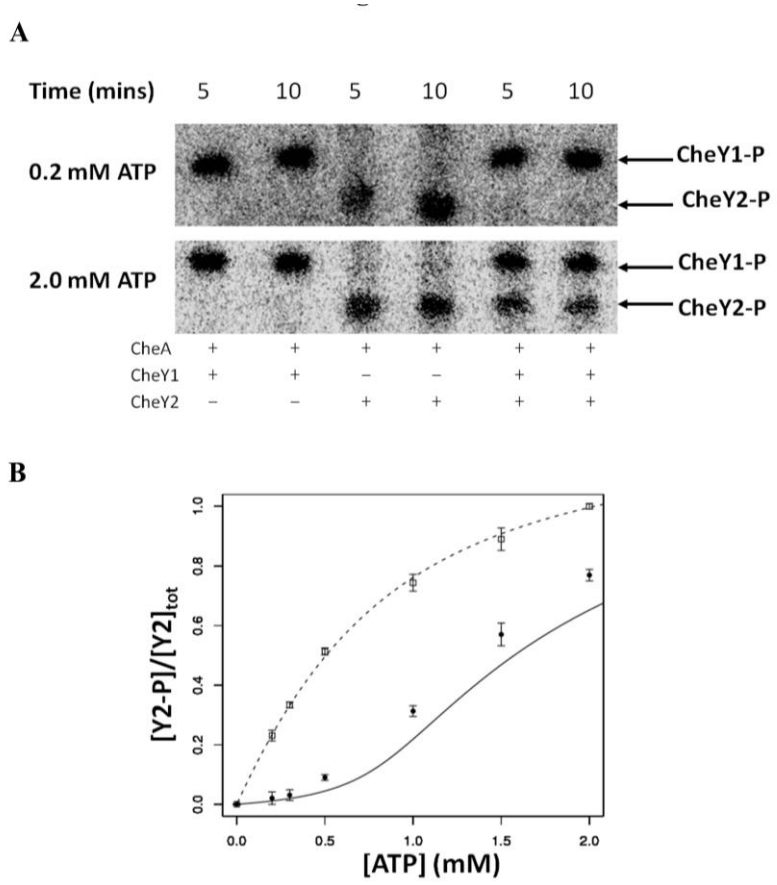

Figure 3.3 Experimental validation for the role of the sink RR in shaping the signal-response curve. The steady-state level of phosphorylated CheY2 was measured in the presence or absence of the sink (i.e. CheY1) at different ^{32}P-ATP concentrations. (**A**) Phosphorimages showing phosphorylated CheY2 levels in the presence or absence of CheY1 at low (0.2 mM) and high (2 mM) ATP levels. The indicated quantity of [γ–^{32}P] ATP was added to a reaction mixture containing 10 μM CheA, 2.5 μM CheY2, and where indicated 2.5 μM CheY1. (**B**) Graph comparing the observed steady state levels of phosphorylated CheY2 with and without the sink, CheY1. The phosphorylated CheY2 levels predicted by the model are shown with a solid line (in absence of sink) and with a dashed line (in presence of sink), while the experimentally measured values are shown by squares (in absence of sink) and circles (in presence of sink). Error bars show the standard error of the mean obtained from three independent experiments.

3.2.5 CheS sharpens the signal-response curve

In the *S. meliloti* system, the behavior of the sink-RR (CheY1) was found to be altered by a small auxiliary protein, CheS (49). In particular, it was shown that CheY1 binds 100-fold more strongly to the CheA-CheS complex than to CheA alone and that the decay of phosphorylated CheA (CheA-P) in the presence of CheY1 is faster with CheS than without. This lead to the suggestion that CheS might directly or indirectly promote CheY1 dephosphorylation and thus make the sink-RR more efficient in allowing signal termination (49). In light of our results, an alternative explanation for how CheS could reduce CheA-P levels is that CheS may accelerate phosphotransfer from CheA-P to CheY1 (i.e. the forward phosphotransfer to the sink, controlled by the parameter k_S) and thus enhancing the possibility of the analytical conditions for sigmoidality to be fulfilled (see above).

Towards obtaining a better understanding of the role of CheS in the system and quantifying its potential effects on the signal-response curve, we first re-constituted CheS in the *in vitro* assay along with CheA, CheY1 and CheY2. We found that the presence of CheS in the system resulted in the sharpening of the signal-response curve (Figure 3.4). We found that the observed effects of CheS can be explained by the model either by increasing the rate of the phosphotransfer reaction between CheA and CheY1 (k_s) (Figure 3.4) or the dephosphorylation rate of CheY1-P (k_{hs}) (Figure S6). The former model alteration better fits the experimentally observed sharpening of the signal-response curve (Figure 3.4), suggesting that the CheS effect on the sigmoidality of the signal-response relationship might be due to increasing the rate of the phosphotransfer reaction between CheA and CheY1. Regardless of which of these mechanisms is employed by CheS, it functions to sharpen the threshold of the sigmoidal signal-response curve given by the system comprising CheA, CheY1 and CheY2.

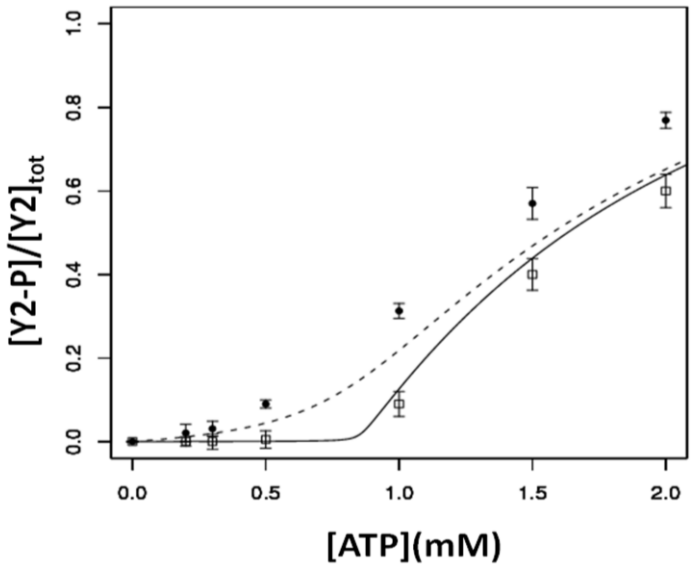

Figure 3.4 Effect of CheS on the signal-response curve. The x- and y-axis show the ATP level and the corresponding steady state level of phosphorylated CheY2, respectively. The phosphorylated CheY2 levels predicted by the model are shown with a dashed line (absence of CheS) and with a solid line (presence of CheS; where phosphotransfer reaction between CheA and CheY1, k_s was increased 100 fold), while the experimentally measured values are shown in circles and squares on respective graph. See also Figure S6 for an alternative approach to modeling the presence of CheS. Error bars show the standard error of the mean obtained from three independent experiments.

3.3 Discussion

We have analyzed the system dynamics of an architectural motif found in bacterial two-component signalling pathways where a single HK can reversibly phosphorylate two RRs. We have shown that this one HK-two RR motif can accelerate signal termination, i.e. act as a sink, as hypothesized before (39), but more interestingly, allows the system to exhibit a sigmoidal signal-response relationship. We have shown that such threshold behavior is observed under

experimentally measured parameters from the *S. meliloti* chemotaxis and yeast osmoregulation pathways. Further, theoretical analyses showed that the presence of a sigmoidal signal-response relationship *necessitates* two conditions on the system; *(i) the sink-RR to be present and* (*ii*) $k_S > k_{rS}$, where k_S and k_{rS} are the forward and reverse phosphotransfer rate constants of the sink-RR respectively. By reconstituting the one HK – two RR motif from the *S. meliloti* chemotaxis pathway *in vitro*, we verified these findings experimentally, showing that the system displays a sigmoidal signal-response relationship, and that the auxiliary protein, CheS, can modulate sensitivity levels by sharpening the response threshold.

These findings have important implications for understanding bacterial physiology and designing synthetic signaling circuits. In broad terms, the findings of this study will have implications for any two-component signaling circuit where multiple response regulators compete for phosphorylation by a single HK, i.e. where the one HK – two RR motif is implemented. This includes the majority of bacterial chemotaxis systems (which employ CheY and CheB as response regulators), fungal osmoregulatory circuits (39-42) and certain plant signaling systems (43). This study indicates that these systems might be acting as a threshold device, whereby cells commit to a specific outcome only above certain signal thresholds. Alternatively, the threshold behavior could be used for regulating the noise characteristics of the system (34, 55). It is important to note however, that the one HK – two RR architectural motif is *able* to display sigmoidal signal-response relationships, but does not preclude hyperbolic relationships. In other words, this motif cannot be taken as proof for threshold behavior but should be taken as indicative and be considered in experimental design when analyzing the response dynamics in associated signaling systems.

Synthetic biology has so far concentrated on designing small circuits based on transcriptional regulation. While two-component proteins have been recognized as potential candidates for

synthetic design, the main efforts have concentrated on engineering chimeric proteins and interaction specificity (15, 17, 26-30). Our findings show that a system dynamics perspective can allow understanding of the signal processing capabilities of natural bacterial signaling pathways and new avenues for reengineering these. Exploiting the single HK - two RR system in the construction of synthetic signaling circuits will require coupling of an appropriate output (e.g. an RR that can as a transcription factor) to a useful signal that can control HK activity. This could be accomplished through mutational alterations on the signal and output of an existing natural system (such as the one used here), using chimeric proteins, or by artificially engineering phosphate sinks into existing two-component systems.

Two-component proteins are highly modular, and evolution seems to have exploited this feature in creating diverse architectures in signaling. Studies like this one should allow us to understand these functionalities and ultimately lead to their exploitation in synthetic biology.

3.4 Methods

3.4.1 A mathematical model for a phosphate sink

To model the one HK - two RR motif, the dynamics was considered in isolation of other cellular components. The reactions in this system that we have included in the model are;

$$HK + ATP \underset{k_2}{\overset{k_1}{\rightleftharpoons}} HK \cdot ATP \xrightarrow{k_a} HKp + ADP$$
$$HKp + RR1 \underset{k_{rS}}{\overset{k_S}{\rightleftharpoons}} RR1p + HK$$
$$HKp + RR2 \underset{k_{rM}}{\overset{k_M}{\rightleftharpoons}} RR2p + HK$$
$$RR1p \xrightarrow{k_{hS}} RR1 + P_i$$
$$RR2p \xrightarrow{k_{hM}} RR2 + P_i$$

where HK, RR1, and RR2 stand for CheA, CheY1 and CheY2 respectively in the *S. meliloti* chemotaxis system (Figure 3.1) and for SLN1, SSK1 and SKN7 in the yeast osmoregulation

system (Figure S3.1). The -p suffix represents phosphorylated forms of these proteins. The above reaction scheme can be used to derive a system of ordinary differential equations (ODEs), which describe the changes in concentrations of proteins over time;

$$\frac{d[HKp]}{dt} = [HK.ATP]\cdot k_a + [RR1p]\cdot[HK]\cdot k_{rS} + [RR2p]\cdot[HK]\cdot k_{rM}$$
$$-[RR1]\cdot[HKp]\cdot k_S - [RR2]\cdot[HKp]\cdot k_M$$
$$\frac{d[HK.ATP]}{dt} = [HK]\cdot[ATP]\cdot k_1 - [HK.ATP]\cdot(k_2 + k_a)$$
$$\frac{d[RR1p]}{dt} = [RR1]\cdot[HKp]\cdot k_S - [RR1p]\cdot[HK]\cdot k_{rS} - [RR1p]\cdot k_{hS}$$
$$\frac{d[RR2p]}{dt} = [RR2]\cdot[HKp]\cdot k_M - [RR2p]\cdot[HK]\cdot k_{rM} - [RR2p]\cdot k_{hM}$$

In addition, we have three conservation equations;

$$[HK]_{tot} = [HK] + [HK_p] + [HK\cdot ATP]$$
$$[RR1]_{tot} = [RR1] + [RR1_p]$$
$$[RR2]_{tot} = [RR2] + [RR2_p]$$

To analyze the behavior of the system with increasing signal, the incoming signals were simulated (e.g. chemoreceptors in case of the chemotaxis system or membrane alterations in the yeast system) as an increase in the autophosphorylation rate constant of the HK (k_a). The model was parameterized with data from literature (see Table 3.1). In the case of the *S. meliloti* chemotaxis system the parameters for phosphotransfer to CheY1 and CheY2 (k_S, k_{rS}, k_M *and* k_{rM}) were derived through fitting the simulation data to previously published *in vitro* experiments (39). Fitting was done using a hybrid genetic algorithm (functions ga and fmincon from the MATLAB Global Optimization Toolbox).

3.4.2 Temporal simulations and signal-response curve. The model was numerically integrated to derive time course and steady state signal-response relationships. The latter

analysis gives the steady state phosphorylated RR levels at a given signal (k_a), where signal was taken as the rate constant of HK autophosphorylation and allows deriving a so-called signal-response curve. This curve is found by numerically integrating the system to steady state at a fixed signal level and then numerically "following" this steady state, while changing the signal. This analysis is equivalent to allowing the system to reach steady state under different signal values. Both time course and signal-response analyses were performed using the software packages XPPAUT (http://www.math.pitt.edu/~bard/xpp/xpp.html) and Oscill8 (http://oscill8 .sourceforge.net). An explicit description of the inverse of the signal-response curve was also obtained, using a recently developed recursive technique (34, 56) (see in Appendix B). The resulting analytical function for the signal-response curve was then used to verify the results of the numerical approach and to derive the necessary conditions that the parameters must fulfill for the signal-response curve to be sigmoidal. This analytical approach is also used to extend the analysis to the case with complex formation (see in Appendix B).

3.4.3 Measuring "sigmoidality" of signal-response curves and sensitivity of this feature to parameters. To measure sigmoidality of the signal-response curve the Hill coefficient was used as previously described (57, 58). The Hill coefficient is measured as ln81/ln(S90/S10) where S90 and S10 are the signal levels for achieving 90 and 10 percent of output saturation respectively. Using alternative measures, such as the maximum value of the response coefficient across the signal domain produces qualitatively similar results as those shown in Figures 3.2, S3.4 and S3.5. To quantify the sensitivity of sigmoidality of the signal-response curve to variations in each of the parameters, these were varied from their described experimentally measured values (Table 3.1) and in a biologically relevant range. Each parameter was varied around its basic value up/down 10-fold and the "sigmoidality" of the resulting signal-response curves measured.

3.4.4 Experimental design. The CheA, CheY1, and CheY2 system was reconstituted *in vitro* to measure the signal-response curve in the presence/absence of CheY1. For this, phosphorylated CheY2 levels were measured under increasing ATP levels as a proxy for signal. The protein concentrations used for these experiments were 10 µM, 2.5 µM, 2.5 µM for CheA, CheY1 and CheY2 respectively. This gives a ratio of 4:1:1, which is different from the *in vivo* measured ratio of 1:10:10 (39), but the higher HK concentration gave increased capacity to measure CheY1 and CheY2 phosphorylation levels at low levels of signal. It was found that the exact ratio among these proteins does not alter the conclusions on the shape of signal-response curve and sigmoidality is not affected by altering the level of sink- and output-RRs with respect to the level of the HK (Figure S3.5.)

3.4.5 Plasmids and strains. See Table 3.2 for the plasmids and strains used. *E. coli* strains were grown in LB medium at 37 ^{o}C. Antibiotics were used at concentrations of 100 µg ml^{-1} for ampicillin, 34 µg ml^{-1} for chloramphenicol and 25 µg ml^{-1} for kanamycin, where needed. *E. coli* M15pREP4 cells were made competent using the calcium chloride technique (59). Transformations were performed according to (60).

3.4.6 Protein purification. His tagged *S. meliloti* CheA, CheA-CheS, CheY1 and CheY2 proteins were purified as described previously (49). Protein purity and concentration was measured as described in (61). Purified proteins were stored at -20^{o}c.

3.4.7 Preparation of CheA-^{32}P and CheA-^{32}P-CheS. CheA-^{32}P and CheA-^{32}P-CheS were phosphorylated using [γ-^{32}P] ATP and purified as described before (62), but with the following modifications: Proteins were phosphorylated in reactions performed at 20°C in phosphotransfer buffer (50 mM Tris HCl, 10% (v/v) glycerol, 5 mM $MgCl_2$, 150 mM NaCl, 50mM KCl, 1 mM DTT, pH 8.0). The final reaction volumes were 2 ml. Reactions were

initiated by addition of 2 mM [γ-^{32}P] ATP (specific activity 14.8 GBq mmol^{-1}; PerkinElmer). After 1 hour incubation, samples were purified by using Ni-NTA columns (Qiagen) as described previously (63). This purification step removed the unincorporated ATP from the CheA-^{32}P and CheA-^{32}P-CheS preparation. Purified proteins were stored at -20°C.

3.4.8 Measurement of CheY2-P at different ^{32}P- ATP concentrations with and without CheS and CheY1: Assays were performed at 20^{o}C in phosphotransfer buffer. Either CheA (10 μM) or CheA-CheS (10 μM) was added to 2.5 μM of purified CheY2 in the presence and absence of 2.5 μM CheY1 under different ATP concentrations. Following the addition of ^{32}P - ATP, reaction aliquots of 10 μl were taken at the indicated time points and quenched immediately in 10 μl of 2 X SDS-PAGE loading dye (7.5% (w/v) SDS, 90 mM EDTA, 37.5 mM Tris HCl, 37.5 % glycerol, 3 % (v/v) β- mercaptoethanol, pH 6.8). Quenched samples were analyzed using SDS-PAGE and phosphorimaging as described previously (64).

Table 3.1: The parameters used for the model of the *S. meliloti* phosphate sink

Parameter	Description	Value	Unit	Reference
k_1	Forward rate constant for phosphorylation complex	1	$(\mu Ms)^{-1}$	[39]
k_2	Reverse rate constant for phosphorylation complex	100	s^{-1}	[39]
k_a	k_{cat} for phosphorylation of CheA	Varied	s^{-1}	
k_S	CheA-P to CheY1 (sink RR) phosphotransfer	1	$(\mu Ms)^{-1}$	Fitted to data from [39] (see *Methods*)
k_{rS}	CheA-P to CheY1 Reverse phosphotransfer	0.01	$(\mu Ms)^{-1}$	Fitted to data from [39] (see *Methods*)
k_M	CheA-P to CheY2 (main RR) phosphotransfer	2	$(\mu Ms)^{-1}$	Fitted to data from [39] (see *Methods*)

k_{rM}	CheA-P to CheY2 Reverse phosphotransfer	1	$(\mu Ms)^{-1}$	Fitted to data from [39] (see *Methods*)
k_{hS}	Autodephosphorylation of CheY1 (sink RR)	0.05	s^{-1}	[39]
k_{hM}	Autodephosphorylation of CheY2 (main RR)	0.06	s^{-1}	[39]
[A]tot	Total concentration of CheA	10	µM	Based on ratios from [39] (see *Methods*)
[Y1]tot	Total concentration of CheY1	2.5	µM	Based on ratios from [39] (see *Methods*)
[Y2]tot	Total concentration of CheY2	2.5	µM	Based on ratios from [39] (see *Methods*)

Table 3.2: The strains and plasmids used in this study

Strains/plasmid	Description	Source/Reference
E. coli strain M15pREP4	Expression host containing pREP4; kanamycin resistant	Qiagen
pQE30	IPTG inducible expression vector. Introduces $RGS(H)_6$ at the N terminus of the expressed protein. Confers ampicillin resistance	Qiagen
pQE60	IPTG inducible expression vector. Introduces $RGS(H)_6$ at the C terminus of the expressed protein. Confers ampicillin resistance	Qiagen
pRU1735 (pQE60Y1)	Plasmid for overexpressing C-terminally His-tagged CheY1 from *S. meliloti*. pQE60 derivative	[39]
pRU1736 (pQE60Y2)	Plasmid for overexpressing C-terminally His-tagged CheY2 from *S. meliloti*. pQE60 derivative	[39]
pRU1742 (pQEA)	Plasmid for overexpressing N-terminally His-tagged CheA from *S. meliloti*. pQE30 derivative	[39]
pBS174 (pET27bmodA/S)	Plasmid for coexpressing *S. meliloti* N-terminally His-tagged CheA and CheS. pET27bmod derivative	[49]

3.5 References

1. Buchler NE, Cross FR (2009) Protein sequestration generates a flexible ultrasensitive response in a genetic network. *Mol Syst Bio* 5:272
2. Tyson JJ, Novak B eds. (2012) Irreversible transitions, bistability and checkpoint controls in the eukaryotic cell cycle: a systems-level understanding. *Elsevier*, San Diego, CA.
3. Tyson JJ, Novak B (2010) Functional motifs in biochemical reaction networks. *Annu Rev Phys Chem* 61:219-240
4. Battogtokh D and Tyson JJ (2004) Bifurcation analysis of a model of the budding yeast cell cycle. *Chaos* 14:653-661.
5. Novak B, Tyson JJ (2008) Design principles of biochemical oscillators. *Nat Rev Mol Cell Biol* 9:981-991
6. Tyson JJ, Albert R, Goldbeter A, Ruoff P, Sible JC (2008) Biological switches and clocks. *J R Soc Interface* 5:S1-S8
7. Kaimachnikov NP, Kholodenko BN (2009) Toggle switches, pulses and oscillations are intrinsic properties of the Src activation/deactivation cycle. *FEBS J* 276:4102-18.
8. Chickarmane V, Kholodenko BN, Sauro HM (2007) Oscillatory dynamics arising from competitive inhibition and multisite phosphorylation. *J Theor Biol* 244(1):68-76.
9. Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335-338.
10. Koseska A, Volkov E, Kurths J (2011) Synthetic multicellular oscillatory systems: controlling protein dynamics with genetic circuits. *Physica Scripta* 84(4): 045007.
11. Hsu C, et al. (2012) Stochastic signalling rewires the interaction map of a multiple feedback network during yeast evolution. *Nat Commun* 3:682.

12. Khalil AS, Lu TK, Bashor CJ, Ramirez CL, Pyenson NC, Joung JK,Collins JJ (2012) A synthetic biology framework for programming eukaryotic transcription functions. *Cell* 150:647-658.
13. Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403:339–342.
14. Siuti P, Yazbek J, Lu TK (2013) Synthetic circuits integrating logic and memory in living cells. *Nat Biotech* 31:448–452.
15. Friedland AE, et al. (2009) Synthetic Gene Networks that Count. *Science* 324:1199–1202.
16. Wang B, Kitney RI, Joly N, Buck M (2011) Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nat Commun* 2:508.
17. Whitaker WR, Davis SA, Arkin AP, Dueber JE (2012) Engineering robust control of two-component system phosphotransfer using modular scaffolds. *Proc Natl Acad Sci USA* 109:18090-18095.
18. Peisajovich SG, Garbarino JE, Wei P, Lim W A (2010) Rapid diversification of cell signaling phenotypes by modular domain recombination. *Sci Signal* 328(5976):368.
19. Wuichet K, Zhulin IB (2010) Origins and Diversification of a Complex Signal Transduction System in Prokaryotes. *Sci Signal* 3: ra50.
20. Hamer R, Chen PY, Armitage JP, Reinert G Deane CM (2010) Deciphering chemotaxis pathways using cross species comparisons. *BMC Sys Biol* 4:3.
21. Stock AM, Robinson VL, Goudreau PN (2000) Two-component signal transduction. *Annu Rev Biochem* 69:183-215.
22. Casino P, Vicente R, Alberto M (2009) Structural Insight Into Partner Specificity and Phosphoryl Transfer in Two-component Signal Transduction. *Cell* 139:325-336.
23. Scott KA, et al. (2010) Specificity of localization and phosphotransfer in the CheA proteins of *Rhodobacter sphaeroides*. *Mol Microbiol* 76:318-330.

24. Podgornaia AI, Laub MT (2013) Determinants of specificity in two-component signal transduction**.** *Curr Opin Microbiol* 16:156-62.
25. Skerker JM, et al. (2008) Rewiring the specificity of two-component signal transduction systems. *Cell* 133:1043-54
26. Ninfa AJ (2010) Use of two-component signal transduction systems in the construction of synthetic genetic networks. *Curr opin microbiol* 13(2):240-245.
27. Levskaya A, et al. (2005) Synthetic biology: Engineering *Escherichia coli* to see light. *Nature* 438:441-442.
28. Moon TS, et al. (2011) Construction of a genetic multiplexer to toggle between chemosensory pathways in *Escherichia coli*. *J Mol Biol* 406(2):245-227.
29. Yoshida T, Phadtare S, Inouye M (2007) The design and development of Tar-EnvZ chimeric receptors. *Methods Enzymol* 423:166-83.
30. Möglich A, Ayers RA, Moffat K (2009) Design and signaling mechanism of light-regulated histidine kinases. *J Mol Biol* 385(5):1433-44.
31. Bischofs IB, Hug JA, Liu AW, Wolf DM, Arkin AP (2009) Complexity in bacterial cell-cell communication: quorum signal integration and subpopulation signaling in the *Bacillus subtilis* phosphorelay. *Proc Natl Acad Sci USA* 106:6459-6464.
32. Csikász-Nagy A, Cardelli L, Soyer OS (2010) Response dynamics of phosphorelays suggest their potential utility in cell signalling. *J R Soc Interface* 8(57):480-488.
33. Jatin N, Devi SN Fujita, M, Igoshin OA (2012). Ultrasensitivity of the *Bacillus subtilis* Sporulation Decision. *Proc Natl Acad Sci USA* 109:20196-20197.
34. Knudsen M, Feliu E, Wiuf C (2012) Exact analysis of intrinsic qualitative features of phosphorelays using mathematical models. *J theor biol* 300:01.007.
35. Arnaud C, et al.(2010) Broadly heterogeneous activation of the master regulator for sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci USA* 107:8486–8491.

36. Amin M, Porter SL, Soyer OS (2013) Split histidine kinases enable ultrasensitivity and bistability in two-component signaling networks. *PLoS Comp Biol* 9(3): e1002949.

37. Igoshin OA, Alves R, Savageau MA (2008) Hysteretic and graded responses in bacterial two-component signal transduction. *Mol Microbiol* 68:1196-1215.

38. Tiwari A, Ray JC, Narula J, Igoshin OA (2011) Bistable responses in bacterial genetic networks: designs and dynamical consequences. *Math Biosci* 231:76-89.

39. Sourjik V, Schmitt R (1998) Phosphotransfer between CheA, CheY1, and CheY2 in the chemotaxis signal transduction chain of *Rhizobium meliloti*. *Biochem* 37:2327-2335.

40. Tindall MJ, Porter SL, Maini PK, Armitage JP (2010) Modeling chemotaxis reveals the role of reversed phosphotransfer and a bi-functional kinase-phosphatase. *PLoS Comp Biol* 6:e1000896.

41. Jiménez-Pearson MA, Delany I, Scarlato V, Beier D (2005). Phosphate flow in the chemotactic response system of *Helicobacter pylori*. *Microbiol* 151(10):3299-3311.

42. Posas F, et al. (1996) Yeast HOG1 MAP kinase cascade is regulated by a multistep phosphorelay mechanism in the SLN1-YPD1-SSK1 "two-component" osmosensor. *Cell* 86:865-875

43. Lohrmann J, Harter K (2002) Plant two-component signaling systems and the role of response regulators. *Plant Physiol* 128:363–369

44. Fabiola J-S, Cook PF, West AH (2005). Kinetic analysis of YPD1-dependent phosphotransfer reactions in the yeast osmoregulatory phosphorelay system. *Biochem* 44(1):377-86.

45. Posas F, Saito H (1998) Activation of the yeast SSK2 MAP kinase kinase kinase by the SSK1 two-component response regulator. *EMBO J* 17:1385–1394.

46. Horie T, Tatebayashi K, Yamada R, Saito H (2008) Phosphorylated Ssk1 prevents unphosphorylated Ssk1 from activating the Ssk2 MAP kinase kinase kinase in the yeast HOG osmoregulatory pathway. *Mol Cell Biol* 28:5172–5183.
47. Brown JL, Bussey H, Stewart RC (1994) Yeast Skn7p functions in a eukaryotic two-component regulatory pathway. *EMBO J* 13:5186–5194.
48. Krems B, Charizanis C, Entian KD (1996) The response regulator-like protein Pos9/Skn7 of *Saccharomyces cerevisiae* is involved in oxidative stress resistance. *Curr Genet* 29:327–334.
49. Dogra G, et al. (2012) *Sinorhizobium meliloti* CheA complexed with CheS exhibits enhanced binding to CheY1 resulting in accelerated CheY1-P dephosphorylation. *J Bacteriol* 194(5):1075-1087.
50. Alves R, Savageau MA (2000) Extending the method of mathematically controlled comparison to include numerical comparisons. *Bioinformatics* 16:786–798
51. Tyson JJ, Katherine CC, Novak B (2003) Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr Opin Cell Biol* 15:221-231.
52. Kothamachu VB, Feliu E, Wiuf C, Cardelli L, Soyer OS (2013) Phosphorelays provide tunable signal processing capabilities for the cell. *PLoS Comp Biol in press.*
53. Hazelbauer GL, Falke JJ, Parkinson JS (2008) Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem Sci* 33(1):9-19
54. Shinar G, Feinberg M (2012) Concordant chemical reaction networks and the species-reaction graph. *Math Biosci* 241(1):1-23.
55. Shibata T, Koichi F (2005) Noisy signal amplification in ultrasensitive signal transduction. *Proc Natl Acad Sci USA* 102(2): 331–336.
56. Feliu E, Knudsen M, Andersen LN, Wiuf C (2012) An algebraic approach to signaling cascades with N layers. *Bull Math Biol* 74:45-72.

57. Zhang Q, Bhattacharya S, Andersen ME (2013) Ultrasensitive response motifs: basic amplifiers in molecular signalling networks. *Open Biol* rsob.130031.
58. Goldbeter A, Koshland DE (1981) An amplified sensitivity arising from covalent modification in biological systems. *Proc Natl Acad Sci USA* 78:6840–6844.
59. Sambrook J Russell JB (2001) Molecular cloning: A laboratory manual. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press.
60. Hanahan D (1983) Studies on transformation of Escherichia coli with plasmids. *J Mol Biol* 166:557.
61. Porter SL, Armitage JP (2002) Phosphotransfer in *Rhodobacter sphaeroides* chemotaxis. *J Mol Biol* 324(1):35-45.
62. Porter SL, Roberts MAJ, Manning CS, Armitage JP (2008) A bifunctional kinase-phosphatase in bacterial chemotaxis. *Proc Natl Acad Sci USA* 105:18531-18536.
63. Porter SL, Wadhams GH, Armitage JP (2007) *In vivo* and *in vitro* analysis of the *Rhodobacter sphaeroides* chemotaxis signaling complexes. *Methods Enzymol* 423:392-413.
64. Porter SL, Warren AV, Martin AC, Armitage JP (2002) The third chemotaxis locus of Rhodobacter sphaeroides is essential for chemotaxis. *Mol Microbiol* 46:1081-1094.

CHAPTER 4: General Discussion

In this thesis, I characterised two structurally diverse motifs found in TCS networks (the split HK motif and the phosphate sink motif). Genomic studies suggest that there are over 700 split kinases found in bacteria (SMART database) (1); on the other hand, phosphate sink motifs are found commonly in bacteria, plants and yeast (2-4). I used both theoretical and experimental approaches to elucidate the underlying mechanisms of TCS signal processing by the networks employing these motifs. I demonstrated the ultrasensitive and bistable behavior in TCS network featuring either split HKs or phosphate sink. Also, this study identifies key parameters and proteins, required for ultrasensitivity in these pathways. In addition, I reported *in vitro* experimental conditions and validations of our theoretical findings that generate ultrasensitive/sigmoidal signal-response relationships in TCS.

4.1 Summary of findings

In the first study (**Chapter 2**), we developed a mathematical model and analyzed the response dynamics mediated by the split kinase motif, using the biochemical reactions of CheA3, CheA4, and CheY6 from the *R. sphaeroides* chemotaxis pathway as a model system. Repeating this analysis with a bifunctional HK and a conventional HK-RR pair featuring a separate phosphatase, we found that in contrast to these configurations, split kinases enable ultrasensitivity and bistability in the signal-response relationship. We show that ultrasensitivity and bistability are maintained under a wide parameter range. A key requirement for ultrasensitivity and bistability in split kinases is the inverse coupling between their kinase and phosphatase activities such that kinase and phosphatase activity are conducted

by separate complexes (Figure 2D and 3 in chapter 2). We also show that this condition is observed in the *R. sphaeroides* system *in vitro* through measurements of phosphatase activity, which showed that binding of CheA4 to CheA3 reduced the phosphatase activity of CheA3 i.e. the CheA3.CheA4 complex has less phosphatase activity than free CheA3 (Figure 4 in chapter 2).

In the second study presented in this thesis (**Chapter 3**), we first developed a generic model of the one HK – two RR (phosphate sink) motif and performed both analytical and simulation-based analyses. These revealed that this system is capable of both enhancing signal termination time and implementing a threshold signal-response relationship, i.e. the system displays a sigmoidal signal-response relationship in which the steady-state levels of the phosphorylated output RR remains low until a threshold level of signal is crossed (Figure 1and S1 in chapter 3). We then verified these dynamics experimentally by *in vitro* re-constitution of the two-component proteins from the chemotaxis pathway of *S. meliloti* (Figure 3 in chapter 3). Using this *in vitro* setup, we further demonstrated that the sharpness of the threshold behavior can be controlled through the concentrations of the core components, as well as through presence of an auxiliary protein that is known to bind the HK in *S. meliloti* (Figure 4 in chapter3)(5).

4.2 Insights and perspectives of the study on signaling networks

In order to achieve a broad and predictive understanding of bacterial signal processing and communication, it is important to assess whether structural diversity within signaling networks enables specific signaling dynamics and properties (6). Research on the diverse architectures of TCSs and associated protein dynamics, their functions and the underlying regulatory processes, are areas of active investigation which can expand our understanding of the complex regulatory networks in terms of their signal- response relationship in this pathway.

Therefore, for understanding dynamics in signaling networks, we need to identify model systems and study the role of specific biochemical reactions. Here, we explored two diverse architectural TCS motifs, using three different model systems: the chemotaxis pathways of *R. sphaeroides* and *S.meliloti* and yeast osmoregulation system. We considered biochemical reactions happening in those systems in our mathematical models to identify their functions in shaping the signal-response relationship.

This study provides the first analysis of the split HK mediated signal-response relationship and response dynamics. Our model shows that a split kinase system set in a "biologically relevant" parameter regime enables an ultrasensitive and bistable signal-response relationship. We found that the predominant allocation of phosphatase activity to the free CheA3 protein (rather than any of the complexes in the system) in the split kinase system is one significant biochemical feature that generates ultrasensitivity and bistability in this signaling pathway. If free CheA3 was the major phosphatase, then the system has the potential to show bistable and ultrasensitive input-output relationships (Figure 2D in chapter 2).

We further showed experimental support for free CheA3 being the principal phosphatase in that system (Figure 4 in chapter 2). *In vitro* phosphotransfer assays in the CheA3-CheA4-CheY6 split kinase system isolated from *R. sphaeroides*, we experimentally measured the effect of CheA4 on the activity of the CheA3 phosphatase, and found that CheA4 reduces the activity of the CheA3 phosphatase indicating that indeed the CheA3 is the main species with phosphatase activity. This experimental result is consistent with the theoretical findings showing bistability and ultrasensitivity in the split kinase system. Whether this system displays ultrasensitivity and bistable responses *in vivo* still needs to be investigated, although such experiments are presently difficult due to our lack of our knowledge of the exact ligands that bind to the chemoreceptors and control split kinase signaling.

We identified some key parameters and alternative models that can alter the response dynamics and hence affect the ultrasensitivity and bistability (see section 2.4.2). These nonlinear dynamics arise from the bifunctional and split nature of the kinase, which introduces a branching point into the system between the phosphorylation and dephosphorylation reactions. Thus, the level of ultrasensitivity (and emergence of bistability) in the system is determined by the parameters and the biochemical mechanisms found in the reaction cycles linked to this branching point. For example, the total concentration of CheA4 allows the control of the dynamics mediated via formation of the CheA3:CheA4 complex. CheA3 mediated dephosphorylation of CheY6-P are also found to be very crucial in altering the dynamics (Figure 2 in chapter 2). We showed that some alternate biochemical/structural assumptions made in the model might not affect the ultrasensitivity that much but could seem to affect the bistability in the system (Figure 2D, S4, S6 and S7 in chapter 2). For example, bistability is maintained in the system for a significant range of phosphotransfer rates from an additional kinase to CheY6, and that ultrasensitivity is maintained in even a larger range for this parameter (Figure S7 in chapter 2). This raises the interesting possibility that cross-talk between a non-split kinase and a split-kinase can confer nonlinearity on the system output of the non-split kinase (in this case CheA2), which it would not have, when operating on its own. These all can potentially participate in further tuning cellular decision making mechanisms.

We also demonstrate through theoretical treatment that the chemical reaction system arising from a bifunctional split kinase gives rise to the possibility of bistability. In contrast, we found that this bistability was lost for systems comprising a bifunctional, non-split kinase or a monofunctional, split kinase with separate phosphatase [unless featuring dead-end complex formation (7)]. Therefore, in conjunction with parameter sensitivity analyses, the analytical results from the Chemical reaction network theory toolbox has allowed us to conclude that the

split kinase architecture allows high level of ultrasensitivity and bistability as long as phosphatase activity is mainly confined to the free form of one of the proteins making up the split kinase (section 2.10).

As, the CheA3-CheA4-CheY6 system is involved in the integration of cytoplasmic and extracellular signals for proper chemotaxis (8, 9), the switch-like dynamics resulting from ultrasensitivity and bistability (observed from the theoretical analyses) might play an important role in the physiological context. Two distinct sensory clusters: one chemotaxis system is polarly localized while the other forms cytoplasmic clusters; are found in *R. sphaeroides* chemotaxis system and it has been thought that polar clusters sense external stimuli while cytoplasmic clusters might sense metabolic signals. As, CheY6 is predominantly phosphorylated by CheA3-P which is located in the cytoplasmic cluster, it would be plausible if the switching dynamics described here allowed cells to override external chemotaxis signals in favor of internal signals such as those related to metabolism, which could contribute to motility decisions (10-12).

Genomics studies suggest 2.3% of all chemotaxis kinases, and 2.8% of all kinases could be functioning as split kinases (1). The SMART database shows that 1.7% of the proteins with HisKA domains could form part of a split bifunctional kinase. These genomic analyses indicate that cells may use such split kinases to allow high sensitivity and bistability enabling switch-like physiological responses to environmental stimuli.

In the second study, we used a systems biology approach to explore the factors that control ultrasensitivity and threshold dynamics in two-component signaling networks containing a phosphate sink motif. We used mathematical modeling to predict the behavior of the system having a phosphate sink, such as rapid signal termination and sigmoidality/ultrasensitivity. We

also used the proteins from *S. meliloti* in an *in vitro* system to validate our theoretical findings experimentally.

Our model was used to investigate the role of a sink in the response dynamics. Firstly, we found that, in agreement with a previous study, the phosphate sink RR is able to speed up the signal termination by enhancing dephosphorylation of the main RR (2) (Figure 1 and S1 in chapter 3). However, our second finding, which is novel, is that systems containing a phosphate sink have the potential to show a ultrasensitive signal response relationship (Figure 1 and S1 in chapter 3). Therefore, the role of a sink could be to generate an ultrasensitive response coupled with rapid signal termination in the system. We also provide experimental support for an ultrasensitive signal response relationship in the presence of a sink in this pathway (Figure 3 in chapter 3). As, ultrasensitive responses are central to many complex biological behaviors (13), perhaps phosphate sink containing bacterial signaling network could potentially display similar switch functions.

We demonstrated that varying the relative concentrations of the RRs and the measured kinetic rates from *S. meliloti* chemotaxis and yeast osmoregulation pathway under parameter spaces enable tuning the threshold behavior in the system featuring a phosphate sink (Figure 2, S2 and S4 in chapter 3). Further, theoretical analyses showed that the presence of sigmoidal signal-response relationship *necessitates* two conditions on the system; *(i)the sink-RR to be present and* (*ii*) $k_S > k_{rS}$, where k_S and k_{rS} are the forward and reverse phosphotransfer rate constants of the sink-RR respectively. An ultrasensitive activation profile can be tuned in a signaling cascade through controlling protein expression levels (14). Thus our system is very promising for building up a tunable synthetic cascade.

We also identified the role of an auxiliary protein CheS (5) in the dynamics of this pathway. This small protein was found to play a key role in sharpening the response and we have

investigated its role again both theoretically and experimentally. The regulatory mechanism of this protein in this pathway is still not confirmed, as there are several possibilities that it can enhance the binding of CheY1 to CheA or/and it can accelerate the dephosphorylation of phosphorylated CheY1. However, we modeled each of these possible roles of CheS and compared them with our experimental data. We found that in each case, CheS sharpens the threshold of the ultrasensitive signal-response curve (Figure 4 and S6 in chapter 3) but that the experimental data fits better to the model where CheS enhances the binding of CheY1 to CheA (Figure 4 in chapter 3). By allowing complex formation and alternative reaction schemes in this system featuring a single HK and two RRs, we found that a certain scenario where four complexes between the phosphorylated/unphosphorylated HK and the phosphorylated/unphosphorylated RRs, generates bistability in the system (section 3.3.3).

4.3 Conclusions and significance

TCS networks are the principal device for cell signaling in bacteria. Communication and decision-making underpins bacterial behaviour including their ability to infect plants, animals and humans, fight off the immune system, survive harsh environmental conditions through spore and biofilm formation, and utilise environmental resources efficiently, both as an individual, and as a community, through cooperation. The signal-response relationships in four common structural arrangements that capture most of the structural diversity found in the TCS networks, decipher their functional role and evolutionary significance, and exploit them for engineering an adaptable synthetic signalling module. The insight gained from these systems will allow translating genomics-derived knowledge of TCS networks into an ability to predict their signal-response relationships, extending experimental results obtained in model organisms to other bacteria, designing novel intervention strategies with bacterial infection and enhancing the use of bacterial signalling components.

In this study, we found that two-component systems containing either split kinases or phosphate sink have the potential to show ultrasensitivity and bistability. Overall we found that split HK motif regulates ultrasensitivity and bistability due to the inverse coupling of kinase and phosphatase activities; for phosphate sinks, if the sink RR is more competitive for phosphoryl groups than the main RR, then a response is not produced until a threshold signal is reached where the sink is filled. These findings were never suggested before and are not intuitively evident from the knowledge of the functions of the individual proteins and only by combining modeling and experiments; it has been possible to gain this level of understanding. However, competition-based ultrasensitivity and threshold-generating mechanisms have been described before in regulatory networks as "molecular titration" (15), in metabolic networks as the "branch-point effect" (16), in signal transduction cascade due to sequestration (17) and in a specific signaling protein (the eukaryotic mitotic regulator Wee1) that has multiple phosphorylation sites (18). Therefore, this study suggests that these systems could be acting as a threshold device, whereby cells commit to a particular outcome only above certain signal thresholds. Alternatively, the noise characteristics of the system can be regulated by such threshold behavior (19, 20).

Ultrasensitive switches are potential engineering targets as they can produce various important cellular behaviors, such as amplification, threshold, oscillation etc (21-23). Our system provides all the potentiality to act as an ultrasensitive and threshold device in a natural process. But it would be really a useful device tool if we can apply synthetic biology to re-engineer those biological functions through designing the synthetic cascade. Our findings show that a system dynamics perspective can allow understanding and engineering specific system dynamics from relatively few two-component proteins. While our *in vitro* implementation of a single HK-two RR system can already be seen as a synthetic threshold

device, exploiting this motif fully in synthetic biology applications will require coupling of an appropriate output (e.g. an RR that can as a transcription factor) to specific signals. This could be done through mutational alterations on the signal and output of an existing natural system (such as the one used here), or by implementing this architectural motif with existing chimeric proteins (fusion of a chemoreceptor with a sensor kinase) or transcription factors.

The principle signal transduction system of prokaryotes called the two-component system (TCS) is a one-step phosphorelay system composed of a histidine kinase (HK) and a response regulator (RR) while the main signal transduction system of eukaryotes is a multi-step system composed of serine/threonine/tyrosine kinases (STYKs). These systems are also different in their phosphorylation mechanisms. HK in the TCS transfers its own phosphate group to the response regulator protein while STYKs phosphorylate other proteins using ATP. It has been proposed that STYKs are more appropriate for signaling cascades than HK because of their efficient regulation of duration of response signals and secondly, in the case of multi-step signaling cascade, the transphosphorylation mechanism of STYKs is faster in signaling than that of HKs (24). In eukaryotes, the mitogen-activated protein kinase (MAPK) cascade system (MCS) is the most common signal transduction system, composed of three kinases (MAPK kinase kinase (MAPKKK), MAPK kinase (MAPKK), and MAPK). A phosphorylated MAPKKK activates MAPKKs by phosphorylating two conserved serine residues, and then the phosphorylated MAPKK activates MAPKs by phosphorylating the conserved threonine and tyrosine residues. Finally, the phosphorylated MAPK phosphorylates and regulates several cellular proteins and nuclear transcription factors (25-30). Ultrasensitivity in MAPK activation is a critical, systems level property and it can produce the fundamental nonlinearity required to achieve stable, potentially irreversible cellular decisions. Also, a number of experimental studies of bistable MAPK activation report ultrasensitive activation profiles (31-33).

Therefore, such switch like responses is known to be common in eukaryotes which can enable decision making at the cellular level (25, 34, 35). In this study, split kinases and phosphate sink motifs set examples as microbial alternative for enabling ultrasensitivity and bistability - response behaviors known to be essential for cellular decision making.

This study bridges between these proteins and response dynamics which also lead to test the theoretical concepts through experiments in real systems. I believe that this study will have broad implications not only for microbiologists but also systems and synthetic biologists who aim to decipher conserved dynamical features of cellular networks involved in decision making. Overall, it is our hope that the work presented here, will aid in the formulation of concrete hypotheses about the dynamic nature of TCS involving split HKs and phosphate sinks. Therefore, they will serve as a launching pad for future investigations to explore more of such diverse architectures in cellular processes.

The studies provided in this thesis represent major steps towards unravelling the dynamic nature and functions of diverse architectural motifs in TCS. While we have found clear relationships between cell signaling, communication, response dynamics and behaviors, we have just started to explore what we believe could be an exciting new area at the interface of different motifs of TCS, response dynamics and signal-response relationships. Clearly much remains to be learned, but to conclude, we can state that these studies have broadened our understanding of how ultrasensitive and bistable responses can be generated and regulated in two-component signaling systems.

4.4 References

1. Wuichet K, Zhulin IB (2010) Origins and diversification of a complex signal transduction system in prokaryotes. Sci Signal 3: ra50.

2. Sourjik V, Schmitt R (1998) Phosphotransfer between CheA, CheY1, and CheY2 in the chemotaxis signal transduction chain of *Rhizobium meliloti*. Biochem 37:2327-2335.

3. Posas F, et al. (1996) Yeast HOG1 MAP kinase cascade is regulated by a multistep phosphorelay mechanism in the SLN1-YPD1-SSK1 "two-component" osmosensor. Cell 86:865-875

4. Lohrmann J, Harter K (2002) Plant two-component signaling systems and the role of response regulators. Plant Physiol 128:363–369.

5. Dogra G, et al. (2012) *Sinorhizobium meliloti* CheA complexed with CheS exhibits enhanced binding to CheY1 resulting in accelerated CheY1-P dephosphorylation. J Bacteriol 194(5):1075-1087.

6. Soyer OS, Creevey CJ (2010) Duplicate retention in signaling proteins and constraints from network dynamics. J Evol Biol. 11: 2410-21.

7. Igoshin OA, Alves R, Savageau MA (2008) Hysteretic and graded responses in bacterial two-component signal transduction. Mol Microbiol 68: 1196-1215.

8. Porter SL, Roberts MAJ, Manning CS, Armitage JP (2008) A bifunctional kinase-phosphatase in bacterial chemotaxis. Proc Natl Acad Sci U S A 105: 18531-18536.

9. Tindall MJ, Porter SL, Maini PK, Armitage JP (2010) Modeling chemotaxis reveals the role of reversed phosphotransfer and a bi-functional kinase-phosphatase. PLoS Comput Biol 6.

10. Goldstein RA, Soyer OS (2008) Evolution of taxis responses in virtual bacteria: non-adaptive dynamics. PLoS Comput Biol 4: e1000084.

11. Egbert MD, Barandiaran XE, Di Paolo EA (2010) A minimal model of metabolism-based chemotaxis. PLoS Comput Biol 6: e1001004.

12. Soyer OS, Goldstein RA (2011) Evolution of response dynamics underlying bacterial chemotaxis. BMC Evol Biol 11: 240.

13. Tyson JJ, Chen KC, Novak B (2003) Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Curr Opin Cell Biol, 15:221-231.

14. O'Shaughnessy EC, Palani S, Collins JJ, Sarkar CA (2011) Tunable signal processing in synthetic MAP kinase cascades. Cell 144: 119–131.

15. Buchler NE, Louis M (2008) Molecular titration and ultrasensitivity in regulatory networks. J Mol Biol 384, 1106-1119.

16. LaPorte DC, Walsh K, Koshland DE (1984)The branch point effect. Ultrasensitivity and subsensitivity to metabolic control. J Biol Chem 259, 14068-14075.

17. Bluthgen N, Bruggermann FJ, Legewie S, Herzel H, Westerhoff HV, Kholodenko BN(2006) Effects of sequestration on signal transduction cascades. FEBS J. **273**, 895–906.

18. Kim SY, Ferrell JE (2007) Substrate competition as a source of ultrasensitivity in the inactivation of Wee1. Cell 128, 1133-1145.

19. Knudsen M, Feliu E, Wiuf C (2012) Exact analysis of intrinsic qualitative features of phosphorelays using mathematical models. J theor biol 300:01.007.

20. Shibata T, Koichi F (2005) Noisy signal amplification in ultrasensitive signal transduction. Proc Natl Acad Sci USA 102(2): 331–336.

21. Chickarmane V, Kholodenko BN, Sauro HM (2007) Oscillatory dynamics arising from competitive inhibition and multisite phosphorylation. J Theor Biol 244(1):68-76.

22. Hsu C, et al. (2012) Stochastic signaling rewires the interaction map of a multiple feedback network during yeast evolution. Nat Commun 3:682.

23. Buchler NE, Cross FR (2009) Protein sequestration generates a flexible ultrasensitive response in a genetic network. Mol Syst Bio 5:272.

24. Choi HS, Kim JR, Lee SW, Cho KH (2008) Why have serine/threonine/tyrosine kinases been evolutionarily selected in eukaryotic signaling cascades? Comput Biol Chem. 32:218–221.

25. Huang CY, Ferrell Jr JE (1996) Ultrasensitivity in the mitogen-activated protein kinase cascade. Proc Natl Acad Sci USA 93, 10078–10083.

26. Kolch W (2000) Meaningful relationships: the regulation of the Ras/Raf/MEK/ERK pathway by protein interactions. Biochem. J. 351 (Pt 2), 289–305.

27. Schoeberl B, Eichler-Jonsson C, Gilles ED, Muller G (2002) Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. Nat. Biotechnol. 20, 370–375.

28. Wurgler-Murphy SM, Saito H (1997) Two-component signal transducers and MAPK cascades. Trends Biochem. Sci. 22, 172–176.

29. Yoon S, Seger R (2006) The extracellular signal-regulated kinase: multiple substrates regulate diverse cellular functions. Growth Factors 24, 21–44.

30. Zhang W, Liu HT (2002) MAPK signal pathways in the regulation of cell proliferation in mammalian cells. Cell Res. 12, 9–18.

31. Bhalla US, Ram PT, and Iyengar R (2002). MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network. Science 297, 1018–1023.

32. Ferrell JE, Machleder EM (1998). The biochemical basis of an all-ornone cell fate switch in Xenopus oocytes. Science 280, 895–898.

33. Santos SD, Verveer PJ, Bastiaens PI (2007). Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate.Nat. Cell Biol. 9, 324–330.

34. Goldbeter A, Koshland DE (1981) An amplified sensitivity arising from covalent modification in biological systems. Proc Natl Acad Sci U S A 78: 6840-6844.

35. Kholodenko BN (2006) Cell-signaling dynamics in time and space. Nature Reviews Molecular Cell Biology 7: 165-176.

Appendix A

In this appendix additional information, analyses and figures needed for chapter 2 are described for better understanding.

S2.1 Alternative models and their analyses

S2.1.1 Models considering additional species with phosphatase activity

In the basic model describing a split kinase system we have assumed that only free CheA3 has phosphatase activity towards the phosphorylated response regulator CheY6. Here, we relax this assumption by considering additional molecular species with phosphatase activity. We create two alternative models where we separately consider the ability of phosphorylated and complexed CheA3 to act as a phosphatase. These models contain one and two additional reactions respectively, in addition to those reactions considered in the basic model. Below, we list these additional reactions and the resulting ordinary differential equations (ODEs) for each model. Model parameters are given in Table S2.1 and are mostly derived from the basic model parameters. The effect of having these additional phosphatases on signal-response relationship is shown in Figure S2.4.

S2.1.2 Model with CheA3-CheA4 and CheA3-CheA4-ATP complexes as phosphatases Additional reaction;

$$A3A4 + Y6p \underset{k_{13}}{\overset{k_{12}}{\rightleftarrows}} A3A4Y6p \xrightarrow{k_{14}} A3A4 + Y6 + P_i$$

$$A3A4ATP + Y6p \underset{k_{16}}{\overset{k_{15}}{\rightleftarrows}} A3A4ATPY6p \xrightarrow{k_{17}} A3A4ATP + Y6 + P_i$$

which, combined with the original reactions listed in chapter 2, results in the following new set of ODEs;

$$\frac{d[A3p]}{dt} = k_5 \cdot [A3A4ATP] + k_7 \cdot [A3] \cdot [Y6p] - k_6 \cdot [A3p] \cdot [Y6]$$
$$\frac{d[A3A4]}{dt} = k_1 \cdot [A3] \cdot [A4] + k_4 \cdot [A3A4ATP] + [A3A4Y6p] \cdot (k_{13} + k_{14}) - [A3A4] \cdot \left(k_2 + k_3 \cdot [ATP] + k_{12} \cdot [Y6p]\right)$$
$$\frac{d[A3A4ATP]}{dt} = k_3 \cdot [A3A4] \cdot [ATP] + [A3A4ATP] \cdot (k_{16} + k_{17}) - [A3A4ATP] \cdot \left(k_4 + k_5 + k_{15} \cdot [Y6p]\right)$$
$$\frac{d[A3Y6p]}{dt} = k_9 \cdot [A3] \cdot [Y6p] - [A3Y6p] \cdot \left(k_{10} + k_{11}\right)$$
$$\frac{d[A3A4Y6p]}{dt} = k_{12} \cdot [Y6p] \cdot [A3A4] - [A3A4Y6p] \cdot (k_{13} + k_{14})$$
$$\frac{d[A3A4ATPY6p]}{dt} = k_{15} \cdot [Y6p] \cdot [A3A4ATP] - [A3A4ATPY6p] \cdot (k_{16} + k_{17})$$
$$\frac{d[Y6p]}{dt} = k_{10} \cdot [A3Y6p] + k_6 \cdot [A3p] \cdot [Y6] + k_{13} \cdot [A3A4Y6p] + k_{16} \cdot [A3A4ATPY6p]$$
$$-[Y6p] \cdot \left(k_7 \cdot [A3] + k_8 + k_9 \cdot [A3] + k_{12} \cdot [A3A4] + k_{15} \cdot [A3A4ATP]\right)$$

S2.1.3 Model with CheA3p as phosphatase

Additional reactions;

$$A3p + Y6p \underset{k_{19}}{\overset{k_{18}}{\rightleftarrows}} A3pY6p \xrightarrow{k_{20}} A3p + Y6 + P_i$$

which, combined with the original reactions listed in the main text, result in the following new set of ODEs;

$$\frac{d[A3p]}{dt} = k_5 \cdot [A3A4ATP] + k_7 \cdot [A3] \cdot [Y6p] + [A3pY6p] \cdot (k_{19} + k_{20}) - k_6 \cdot [A3p] \cdot [Y6] - k_{18} \cdot [A3p] \cdot [Y6p]$$
$$\frac{d[A3A4]}{dt} = k_1 \cdot [A3] \cdot [A4] + k_4 \cdot [A3A4ATP] - [A3A4] \cdot \left(k_2 + k_3 \cdot [ATP]\right)$$
$$\frac{d[A3A4ATP]}{dt} = k_3 \cdot [A3A4] \cdot [ATP] - [A3A4ATP] \cdot \left(k_4 + k_5\right)$$
$$\frac{d[A3Y6p]}{dt} = k_9 \cdot [A3] \cdot [Y6p] - [A3Y6p] \cdot \left(k_{10} + k_{11}\right)$$
$$\frac{d[A3pY6p]}{dt} = k_{18} \cdot [A3p] \cdot [Y6p] - [A3pY6p] \cdot \left(k_{19} + k_{20}\right)$$
$$\frac{d[Y6p]}{dt} = k_{10} \cdot [A3Y6p] + k_6 \cdot [A3p] \cdot [Y6] + k_{19} \cdot [A3pY6p] - [Y6p] \cdot \left(k_7 \cdot [A3] + k_8 + k_9 \cdot [A3] + k_{18} \cdot [A3p]\right)$$

S2.1.4 Model with an alternative reaction scheme

In the basic model describing a split kinase system and discussed in chapter 2, we have

assumed that the phosphorylation of the CheA3 by CheA4 results in the dissociation of the CheA3:CheA4 complex. Here, we relaxed this assumption to create an alternative model. In this model, we allowed for the possibility that phosphorylated CheA3 remains in complex with CheA4 and that this CheA3p:CheA4 complex is also capable of acting as phosphatase towards CheY6p (corresponding reaction rates k'_5, k'_6 and k''_6). We find that having these reactions in the model does not affect the level of ultrasensitivity but can lead to loss of bistability (Figure S2.6). Note, that besides these reactions, this alternative model is the same as the basic model and only considers phosphatase activity by free CheA3. Model parameters are given in Table S2.2 and are mostly derived from the basic model parameters. As in the basic model (Figure 2.2D and S2.4), considering alternative phosphatases in this alternative model significantly reduces ultrasensitivity and leads to loss of bistability (data not shown). This alternative model contains the following reactions;

$$A3 + A4 \underset{k_2}{\overset{k_1}{\rightleftarrows}} A3A4$$

$$A3A4 + ATP \underset{k_4}{\overset{k_3}{\rightleftarrows}} A3A4ATP \xrightarrow{k_5} A3p + A4 + ADP$$

$$A3p + A4 \underset{k'_2}{\overset{k'_1}{\rightleftarrows}} A3pA4$$

$$A3A4ATP \xrightarrow{k'_5} A3pA4 + ADP$$

$$A3p + Y6 \underset{k_7}{\overset{k_6}{\rightleftarrows}} A3 + Y6p$$

$$A3pA4 + Y6 \underset{k'_7}{\overset{k'_6}{\rightleftarrows}} A3A4 + Y6p$$

$$A3pA4 + Y6 \underset{k''_7}{\overset{k''_6}{\rightleftarrows}} A3 + A4 + Y6p$$

$$Y6p \xrightarrow{k_8} Y6 + P_i$$

$$A3 + Y6p \underset{k_{10}}{\overset{k_9}{\rightleftarrows}} A3Y6p \xrightarrow{k_{11}} A3 + Y6 + P_i$$

resulting in the following set of ODEs;

$$\frac{d[A3A4]}{dt} = [A3]\cdot[A4]\cdot k_1 - [A3A4]\cdot k_2 + [A3A4ATP]\cdot k_4 - [A3A4]\cdot[ATP]\cdot k_3$$
$$+[A3pA4]\cdot[Y6]\cdot k'_6 - [A3A4]\cdot[Y6p]\cdot k'_7$$
$$\frac{d[A3A4ATP]}{dt} = [A3A4]\cdot[ATP]\cdot k_3 - [A3A4ATP]\cdot(k_4 + k_5 + k'_5)$$
$$\frac{d[A3p]}{dt} = [A3A4ATP]\cdot k_5 + [Y6p]\cdot[A3]\cdot k_7 - [A3p]\cdot([Y6]\cdot k_6 + [A4]\cdot k'_1)$$
$$+[A3pA4]\cdot k'_2$$
$$\frac{d[A3pA4]}{dt} = [A3p]\cdot[A4]\cdot k'_1 - [A3pA4]\cdot(k'_2 + [Y6]\cdot k'_6 + [Y6]\cdot k''_6) + [A3A4ATP]\cdot k'_5$$
$$+[A3A4]\cdot[Y6p]\cdot k'_7 + [A3]\cdot[A4]\cdot[Y6p]\cdot k''_7$$
$$\frac{d[A3Y6p]}{dt} = [A3]\cdot[Y6p]\cdot k_9 - [A3Y6p]\cdot(k_{10} + k_{11})$$
$$\frac{d[Y6p]}{dt} = [A3p]\cdot[Y6]\cdot k_6 - [Y6p]\cdot[A3]\cdot k_7 - [A3A4]\cdot[Y6p]\cdot k'_7 + [A3pA4]\cdot([Y6]\cdot k'_6 + [Y6]\cdot k''_6)$$
$$-[A3]\cdot[A4]\cdot[Y6p]\cdot k''_7 + [A3Y6p]\cdot k_{10} - [A3]\cdot[Y6p]\cdot k_9 - [Y6p]\cdot k_8$$

S2.1.5 Model with additional kinase, CheA2

In the basic model describing a split kinase system we have only considered phosphorylation of the response regulator (i.e. CheY6) by the split kinase. *In vivo*, cross-talk from other kinases could also result in the phosphorylation of the response regulator. For example, in *Rhodobacter sphaeroides*, another kinase, CheA2 is known to phosphorylate CheY6 [27]. Here, we determine the effect of having such an additional kinase on the response dynamics generated by the split kinase. We created a model having this additional kinase activity and analysed the signal-response relationship in the system under a range of phosphotransfer rates from such an additional kinase (Figure S2.7). Model parameters are given in Table S2.3 and are mostly derived from the basic model parameters. This model contains two additional reactions:

$$A2 + ATP \underset{k^*_4}{\overset{k^*_3}{\rightleftarrows}} A2\cdot ATP \xrightarrow{k^*_5} A2p$$
$$A2p + Y6 \underset{k^*_7}{\overset{k^*_6}{\rightleftarrows}} Y6p + A2$$

which, combined with the original reactions listed in the main text, results in the following new set of ODEs;

$$\frac{d[A3A4]}{dt} = A3 \cdot A4 \cdot k_1 - A3A4 \cdot k_2 + A3A4ATP \cdot k_4 - A3A4 \cdot ATP \cdot k_3$$
$$\frac{d[A3A4ATP]}{dt} = A3A4 \cdot ATP \cdot k_3 - A3A4ATP \cdot (k_4 + k_5)$$
$$\frac{d[A3p]}{dt} = A3A4ATP \cdot k_5 + Y6p \cdot A3 \cdot k_7 - A3p \cdot Y6 \cdot k_6$$
$$\frac{d[A2p]}{dt} = A2ATP \cdot k*_5 + Y6p \cdot A2 \cdot k*_7 - A2p \cdot Y6 \cdot k*_6$$
$$\frac{d[A2ATP]}{dt} = A2 \cdot ATP \cdot k*_3 - A2ATP \cdot (k*_4 + k*_5)$$
$$\frac{d[A3Y6p]}{dt} = Y6p \cdot A3 \cdot k_9 - A3Y6p \cdot (k_{10} + k_{11})$$
$$\frac{d[Y6p]}{dt} = A3p \cdot Y6 \cdot k_6 - Y6p \cdot A3 \cdot k_7 + A3Y6p \cdot k_{10} - Y6p \cdot A3 \cdot k_9$$
$$+A2p \cdot Y6 \cdot k*_6 + Y6p \cdot A2 \cdot k*_7 - Y6p \cdot k_8$$

S2.2 Analytical solutions for simplified systems with a bifunctional, split kinase vs. split kinase with a stand-alone phosphatase.

Besides using the chemical reaction network theory to analse different models (see discussion in chapter 2), we have also derived analytical solutions for a simplified reaction scheme for a bifunctional split kinase and also for a monofunctional split kinase with a stand-alone phosphatase (i.e. where dephosphorylation of the response regulator is mediated by a separate phosphatase).

S2.2.1 Simplified reaction scheme and analytical solution for a system with bifunctional, split kinase.

In this simplified scheme, we assume that all phosphotransfer and dephosphorylation reactions occur very fast and that complex formation can be ignored. The reaction scheme we consider is;

$$A3 + A4 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} A3A4$$
$$A3A4 + ATP \xrightarrow{k_5} A3p + A4 + ADP$$
$$A3p + Y6 \xrightarrow{k_6} A3 + Y6p$$
$$Y6p + A3 \xrightarrow{k_9} Y6 + A3 + P_i$$
$$Y6p \xrightarrow{k_8} Y6 + P_i$$

which results in the following ODEs;

$$\frac{d[A3A4]}{dt} = [A3]\cdot[A4]\cdot k_1 - [A3A4]\cdot(k_2 + k_5)$$
$$\frac{d[A3p]}{dt} = [A3A4]\cdot k_5 - [A3p]\cdot[Y6]\cdot k_6$$
$$\frac{d[Y6p]}{dt} = [A3p]\cdot[Y6]\cdot k_6 - [Y6p]\cdot[A3]\cdot k_9 - [Y6p]\cdot k_8$$

We first define the conservation relations in the system:

$$[A3]_{tot} = [A3] + [A3p] + [A3A4]$$
$$[A4]_{tot} = [A4] + [A3A4]$$
$$[Y6]_{tot} = [Y6] + [Y6p]$$

At steady state, all of the above ODEs would be equal to zero, allowing us to derive the steady state expression for phosphorylated CheY6. Following simple algebra, we arrive at a quartic equation;

$$A3 \cdot A4 \cdot k_1 - A3A4 \cdot (k_2 + k_5) = 0 \qquad (1)$$

$$A3A4 \cdot k_5 - A3p \cdot Y6 \cdot k_6 = 0 \qquad (2)$$

$$A3p \cdot Y6 \cdot k_6 - Y6p \cdot A3 \cdot k_9 - Y6p \cdot k_8 = 0 \qquad (3)$$

Solving equation 2 and 3, we get

$$A3p \cdot Y6 \cdot k_6 = Y6p \cdot A3 \cdot k_9 + Y6p \cdot k_8$$
$$Y6p \cdot (A3_{tot} - A3p - A3A4) \cdot k_9 + Y6p \cdot k_8 = A3p \cdot Y6 \cdot k_6$$

$$A3p \cdot Y6 \cdot k_6 = A3A4 \cdot k_5$$
$$A3p = \frac{A3A4 \cdot k_5}{Y6 \cdot k_6}$$

$$Y6p \cdot (A3_{tot} - \frac{A3A4 \cdot k_5}{Y6 \cdot k_6} - A3A4) \cdot k_9 + Y6p \cdot k_8 = A3A4 \cdot k_5$$
$$\frac{Y6p \cdot A3_{tot} \cdot Y6 \cdot k_6 \cdot k_9 - Y6p \cdot A3A4 \cdot k_5 \cdot k_9 - Y6p \cdot Y6 \cdot A3A4 \cdot k_6 \cdot k_9 + Y6p \cdot Y6 \cdot k_8 \cdot k_6}{Y6 \cdot k_6} = A3A4 \cdot k_5$$
$$Y6p \cdot A3_{tot} \cdot Y6 \cdot k_6 \cdot k_9 - Y6p \cdot A3A4 \cdot k_5 \cdot k_9 - Y6p \cdot Y6 \cdot A3A4 \cdot k_6 \cdot k_9 + Y6p \cdot Y6 \cdot k_8 \cdot k_6 = A3A4 \cdot Y6 \cdot k_5 \cdot k_6$$
$$A3A4 \cdot (Y6p \cdot k_5 \cdot k_9 + Y6p \cdot Y6 \cdot k_6 \cdot k_9 + Y6 \cdot k_5 \cdot k_6) = Y6p \cdot A3_{tot} \cdot Y6 \cdot k_6 \cdot k_9 + Y6p \cdot Y6 \cdot k_8 \cdot k_6$$
$$A3A4 = \frac{Y6p \cdot A3_{tot} \cdot Y6 \cdot k_6 \cdot k_9 + Y6p \cdot Y6 \cdot k_8 \cdot k_6}{Y6p \cdot k_5 \cdot k_9 + Y6p \cdot Y6 \cdot k_6 \cdot k_9 + Y6 \cdot k_5 \cdot k_6}$$
$$A3A4 = \frac{Y6p \cdot A3_{tot} \cdot Y6 \cdot k_6 \cdot k_9 + Y6p \cdot Y6 \cdot k_8 \cdot k_6}{N} [N = Y6p \cdot k_5 \cdot k_9 + Y6p \cdot Y6 \cdot k_6 \cdot k_9 + Y6 \cdot k_5 \cdot k_6]$$

(4)

$$A3p = (\frac{Y6p \cdot A3_{tot} \cdot Y6 \cdot k_6 \cdot k_9 + Y6p \cdot Y6 \cdot k_8 \cdot k_6}{Y6p \cdot k_5 \cdot k_9 + Y6p \cdot Y6 \cdot k_6 \cdot k_9 + Y6 \cdot k_5 \cdot k_6}) \cdot \frac{k_5}{Y6 \cdot k_6}$$
$$A3p = \frac{Y6p \cdot A3_{tot} \cdot k_5 \cdot k_6 \cdot k_9 + Y6p \cdot k_5 \cdot k_8 \cdot k_6}{N \cdot k_6} \qquad (5)$$

From equation 1,

$$A3 \cdot A4 \cdot k_1 = A3A4 \cdot (k_2 + k_5)$$

$$A3 \cdot A4 = A3A4 \cdot \frac{(k_2 + k_5)}{k_1}$$

$$(A3_{tot} - A3p - A3A4) \cdot (A4_{tot} - A3A4) = A3A4 \cdot a[a = \frac{(k_2 + k_5)}{k_1}]$$

$$A3_{tot} \cdot A4_{tot} - A3p \cdot A4_{tot} - A3A4 \cdot (A4_{tot} + A3_{tot} + a) + A3p \cdot A3A4 + A3A4^2 = 0$$

(6)

Putting value of A3p and A3A4 from eq 4&5 into 6, we get

$$A3_{tot} \cdot A4_{tot} - (\frac{Y6p \cdot A3_{tot} \cdot k_5 \cdot k_6 \cdot k_9 + Y6p \cdot k_5 \cdot k_8 \cdot k_6}{N \cdot k_6}) \cdot A4_{tot} - (\frac{Y6p \cdot A3_{tot} \cdot Y6 \cdot k_6 \cdot k_9 + Y6p \cdot Y6 \cdot k_8 \cdot k_6}{N}) \cdot (A4_{tot} + A3_{tot} + a)$$
$$+(\frac{Y6p \cdot A3_{tot} \cdot k_5 \cdot k_6 \cdot k_9 + Y6p \cdot k_5 \cdot k_8 \cdot k_6}{N \cdot k_6}) \cdot (\frac{Y6p \cdot A3_{tot} \cdot Y6 \cdot k_6 \cdot k_9 + Y6p \cdot Y6 \cdot k_8 \cdot k_6}{N}) +$$
$$(\frac{Y6p \cdot A3_{tot} \cdot Y6 \cdot k_6 \cdot k_9 + Y6p \cdot Y6 \cdot k_8 \cdot k_6}{N})^2 = 0$$

$$A3_{tot} \cdot A4_{tot} \cdot N^2 \cdot k_6 - N \cdot Y6p \cdot c - N \cdot k_6 \cdot (A4_{tot} + A3_{tot} + a) \cdot (A3_{tot} \cdot k_6 \cdot k_9 + k_8 \cdot k_6) \cdot Y6p \cdot Y6$$
$$\frac{+(A3_{tot} \cdot k_5 \cdot k_6 \cdot k_9 + k_5 \cdot k_8 \cdot k_6) \cdot Y6p \cdot \cdot (Y6p \cdot A3_{tot} \cdot Y6 \cdot k_6 \cdot k_9 + Y6p \cdot Y6 \cdot k_8 \cdot k_6) + k_6 \cdot (A3_{tot} \cdot k_6 \cdot k_9 + k_8 \cdot k_6)^2 \cdot (Y6p \cdot Y6)^2}{N^2 \cdot k_6} = 0$$

$$b \cdot N^2 - N \cdot Y6p \cdot c - N \cdot d \cdot Y6p \cdot Y6$$
$$+e \cdot Y6p \cdot (Y6p \cdot A3_{tot} \cdot Y6 \cdot k_6 \cdot k_9 + Y6p \cdot Y6 \cdot k_8 \cdot k_6) + f \cdot (Y6p \cdot Y6)^2 = 0$$

Where following were given by,

$$[b = A3_{tot} \cdot A4_{tot} \cdot k_6,$$
$$c = A4_{tot} \cdot (A3_{tot} \cdot k_5 \cdot k_6 \cdot k_9 + k_5 \cdot k_8 \cdot k_6),$$
$$d = (A4_{tot} + A3_{tot} + a) \cdot k_6 \cdot (A3_{tot} \cdot k_6 \cdot k_9 + k_8 \cdot k_6),$$
$$e = (A3_{tot} \cdot k_5 \cdot k_6 \cdot k_9 + k_5 \cdot k_8 \cdot k_6),$$
$$f = k_6 \cdot (A3_{tot} \cdot k_6 \cdot k_9 + k_8 \cdot k_6)^2]$$

Putting back the value of *N* and solving it we get,

$$Y6p^4 \cdot (k_6^2 \cdot k_9^2 \cdot b - k_6 \cdot k_9 \cdot d + f)$$
$$+Y6p^3 \cdot (2 \cdot k_5 \cdot k_6^2 \cdot k_9 \cdot b - 2 \cdot k_5 \cdot k_6 \cdot k_9^2 \cdot b - 2 \cdot Y6_{tot} \cdot k_6^2 \cdot k_9^2 \cdot b + k_6 \cdot k_9 \cdot c - k_5 \cdot k_6 \cdot d + 2 \cdot Y6_{tot} \cdot k_6 \cdot k_9 \cdot d$$
$$+k_5 \cdot k_9 \cdot d - A3_{tot} \cdot k_6 \cdot k_9 \cdot e - k_6 \cdot k_8 \cdot e - 2 \cdot Y6_{tot} \cdot f)$$
$$+Y6p^2 \cdot (k_5^2 \cdot k_9^2 \cdot b + Y6_{tot}^2 \cdot k_6^2 \cdot k_9^2 \cdot b + k_5^2 \cdot k_6^2 \cdot b + 2 \cdot Y6_{tot} \cdot k_5 \cdot k_6 \cdot k_9^2 \cdot b - 4 \cdot Y6_{tot} \cdot k_5 \cdot k_6^2 \cdot k_9 \cdot b$$
$$-2 \cdot k_5^2 \cdot k_6 \cdot k_9 \cdot b - k_5 \cdot k_9 \cdot c - Y6_{tot} \cdot k_6 \cdot k_9 \cdot c + k_5 \cdot k_6 \cdot c - Y6_{tot} \cdot k_5 \cdot k_9 \cdot d - Y6_{tot}^2 \cdot k_6 \cdot k_9 \cdot d$$
$$+2 \cdot Y6_{tot} \cdot k_5 \cdot k_6 \cdot d + A3_{tot} \cdot Y6_{tot} \cdot k_6 \cdot k_9 \cdot e + Y6_{tot} \cdot k_6 \cdot k_8 \cdot e + Y6_{tot}^2 \cdot f)$$
$$+Y6p \cdot (2 \cdot Y6_{tot}^2 \cdot k_5 \cdot k_6^2 \cdot k_9 \cdot b - 2 \cdot Y6_{tot} \cdot k_5^2 \cdot k_6^2 \cdot b + 2 \cdot Y6_{tot} \cdot k_5^2 \cdot k_6 \cdot k_9 \cdot b - Y6_{tot} \cdot k_5 \cdot k_6 \cdot c - Y6_{tot}^2 \cdot k_5 \cdot k_6 \cdot d)$$
$$+Y6_{tot}^2 \cdot k_5^2 \cdot k_6^2 \cdot b = 0$$

The emergence of the quartic expression for the steady state level of phosphorylated CheY6 indicates the potential of this system to reach bistability and high level of nonlinearity even without considering complex formation. To confirm bistability, we have analysed the model shown above and a similar one (see Text S2.2) using chemical reaction network theory. This confirmed the potential of bistability in both of these models (see also discussion in chapter 2). Furthermore, we have analysed the above simplified model by evaluating the analytical solution over the same signal range as for the basic model. For reactions that were modeled as bi- or uni-molecular both in the basic model and this simplified model, we have used the parameters as in the basic model. For reactions that were modeled via complex formation in the basic model (e.g. the A3p mediated dephosphorylation of Y6p), we have explored different parameter values. In line with the results of the chemical reaction network theory, this analysis confirmed that the modeled system displays bistability (i.e. multiple permissible steady states) in a biologically permissible parameter regime.

S2.2.2 Simplified reaction scheme and analytical solution for a system with a monofunctional, split kinase and stand-alone phosphatase.

As before, we assume that all phosphotransfer and dephosphorylation reactions occur very fast and ignore the formation of complexes. The reaction scheme we consider is;

$$A3 + A4 \underset{k_2}{\overset{k_1}{\rightleftarrows}} A3A4$$
$$A3A4 + ATP \xrightarrow{k_5} A3p + A4 + ADP$$
$$A3p + Y6 \xrightarrow{k_6} A3 + Y6p$$
$$Y6p + X \xrightarrow{k_9} Y6 + X + P_i$$
$$Y6p \xrightarrow{k_8} Y6 + P_i$$

We first define the conservation relations in the system:

$$[A3]_{tot} = [A3] + [A3p] + [A3A4]$$
$$[A4]_{tot} = [A4] + [A3A4]$$
$$[Y6]_{tot} = [Y6p] + [Y6]$$
$$[X]_{tot} = [X]$$

which results in the following ODEs;

$$\frac{d[A3A4]}{dt} = [A3]\cdot[A4]\cdot k_1 - [A3A4]\cdot(k_2 + k_5)$$
$$\frac{d[A3p]}{dt} = [A3A4]\cdot k_5 - [A3p]\cdot[Y6]\cdot k_6$$
$$\frac{d[Y6p]}{dt} = [A3p]\cdot[Y6]\cdot k_6 - [Y6p]\cdot[X]\cdot k_9 - [Y6p]\cdot k_8$$

At steady state, all of the above ODEs would be equal to zero, allowing us to derive the steady state expression for phosphorylated CheY6. Following simple algebra, we arrive at a cubic equation;

$$A3\cdot A4\cdot k_1 - A3A4\cdot(k_2 + k_5) = 0 \qquad (1)$$

$$A3A4\cdot k_5 - A3p\cdot Y6\cdot k_6 = 0 \qquad (2)$$

$$A3p\cdot Y6\cdot k_6 - Y6p\cdot X\cdot k_9 - Y6p\cdot k_8 = 0 \qquad (3)$$

Solving equation 2 and 3, we got

$$Y6p = \frac{A3A4\cdot k_5}{X_{tot}\cdot k_9 + k_8}$$
$$Y6p = A3A4\cdot a$$
$$A3A4 = \frac{Y6p}{a}$$

Where a is given by; $a = \dfrac{k_5}{X_{tot}\cdot k_9 + k_8}$

From equation 2,

$$A3p = \frac{Y6p\cdot c}{Y6}$$

Where c is given by; $c = \dfrac{k_5}{k_6\cdot a}$

From equation 1, we get

$$A3A4 \cdot b = (A3_{tot} - A3p - A3A4) \cdot (A4_{tot} - A3A4) \qquad (4)$$

Putting A3A4 and A3p values in equation 4, we get

$$Y6p^3 \cdot Y6_{tot} + Y6p^2 \cdot (-A3_{tot} \cdot a - b \cdot a - c \cdot a - A4_{tot} \cdot a - Y6_{tot}) +$$
$$Y6p \cdot (b \cdot a \cdot Y6_{tot} + A4_{tot} \cdot c \cdot a^2 + A3_{tot} \cdot A4_{tot} \cdot a^2 +$$
$$Y6_{tot} \cdot A4_{tot} \cdot a + A3_{tot} \cdot a \cdot Y6_{tot}) - A3_{tot} \cdot Y6_{tot} \cdot A4_{tot} \cdot a^2 = 0$$

Where b is given by; $b = \frac{(k_2 + k_5)}{k_1}$

The emergence of the cubic expression for the steady state level of phosphorylated CheY6 indicates less nonlinearity in the system compared with the system with a bifunctional split kinase (previous section). A numerical analysis using this analytical expression (as done in the previous section), shows that in the similar parameter ranges where the previous model shows bistability, this one does not. Again, this is inline and as expected from the results of the chemical reaction network theory, which shows no possibility of bistability in this model (see chapter 2 and Text S2.3).

S2.3 Mathematical model of the phosphotransfer experiments

We developed a mathematical model of the specific *in vitro* experimental setup used to test whether CheA4 can inhibit the phosphatase activity of CheA3. In particular, these experiments employed a truncated form of CheA3, CheA3P1, that lacks phosphatase activity and that can be isolated in a fully phosphorylated form (7). We mixed CheA3P1-P with CheY6 in the absence of ATP and monitored phosphotransfer to CheY6 and its subsequent dephosphorylation by CheA3. In the model, CheA3P1-P was assumed to have the same phosphotransfer kinetics as CheA3. We also assumed that CheA3P1 and CheA3P1-P can bind to CheA4 at the same rate as CheA3. The resulting set of reactions in the system are;

$$A3 + A4 \underset{k_2}{\overset{k_1}{\rightleftarrows}} A3A4$$

$$P1 + A4 \underset{k'_2}{\overset{k'_1}{\rightleftarrows}} A4P1$$

$$P1p + A4 \underset{k''_2}{\overset{k''_1}{\rightleftarrows}} A4P1p$$

$$A3p + Y6 \underset{k_7}{\overset{k_6}{\rightleftarrows}} A3 + Y6p$$

$$A4P1p + Y6 \underset{k`_7}{\overset{k`_6}{\rightleftarrows}} A4P1 + Y6p$$

$$Y6p \xrightarrow{k_8} Y6 + P_i$$

$$A3 + Y6p \underset{k_{10}}{\overset{k_9}{\rightleftarrows}} A3Y6p \xrightarrow{k_{11}} A3 + Y6 + P_i$$

$$A3A4 + Y6p \underset{k_{13}}{\overset{k_{12}}{\rightleftarrows}} A3A4Y6p \xrightarrow{k_{14}} A3A4 + Y6 + P_i$$

giving rise to the following ODEs;

$$\frac{d[A3A4]}{dt} = A3 \cdot A4 \cdot k_1 - A3A4 \cdot k_2 - A3A4 \cdot Y6p \cdot k_{12} + A3A4Y6p \cdot (k_{13} + k_{14})$$

$$\frac{d[A4P1]}{dt} = P1 \cdot A4 \cdot k'_1 - A4P1 \cdot k'_2 + A4P1p \cdot Y6 \cdot k`_6 - A4P1 \cdot Y6p \cdot k`_7$$

$$\frac{d[P1]}{dt} = A4P1 \cdot k'_2 - P1 \cdot A4 \cdot k'_1 + P1p \cdot Y6 \cdot k_6 - P1 \cdot Y6p \cdot k_7$$

$$\frac{d[A4P1p]}{dt} = A4 \cdot P1p \cdot k''_2 - A4P1p \cdot k''_1 - A4P1p \cdot Y6 \cdot k`_6 + A4P1 \cdot Y6p \cdot k`_7$$

$$\frac{d[A3Y6p]}{dt} = A3 \cdot Y6p \cdot k_9 - A3Y6p \cdot (k_{10} + k_{11})$$

$$\frac{d[A3A4Y6p]}{dt} = A3A4 \cdot Y6p \cdot k_{12} - A3A4Y6p \cdot (k_{13} + k_{14})$$

$$\frac{d[Y6p]}{dt} = P1p \cdot Y6 \cdot k_6 - P1 \cdot Y6p \cdot k_7 + A4P1p \cdot Y6 \cdot k`_6 - A4P1 \cdot Y6p \cdot k`_7 + A3Y6p \cdot k_{10} - A3 \cdot Y6p \cdot k_9$$
$$+ A3A4Y6p \cdot k_{13} - A3A4 \cdot Y6p \cdot k_{12} - Y6p \cdot k_8$$

We numerically solved this system using parameter values given in Table S2.4 and in the presence of different levels of CheA4. By fitting a first-order exponential decay curve to this simulation data, we estimated the half-time of phosphorylated CheY6 (k_{obs}) shown in Figure 4. Under the assumption that CheA4 and CheA3:CheA4 complex are not capable of CheY6-P dephosphorylation, this model predicts that increasing CheA4 levels would slow the CheY6-P dephosphorylation kinetics by sequestering free CheA3. We

found that this model provides a good qualitative match to the experimental observations (Figure2. 4 in chapter 2).

S2.4 Results of the analytical analysis of models

These contain the reaction system considered and the report produced with the Chemical Network Tool v2.2 (http://www.chbmeng.ohio-state.edu/~feinberg/crntwin/). In these reaction systems A, B, Y and X stand for CheA3, CheA4, CheY6 and an hypothetical separate phosphatase respectively.

S24.1 Results of the analytical analysis of the basic model

BASIC REPORT

==========

Reaction network:

A + B <-> AB

AB <-> Ap + B

Ap + Y <-> Yp + A

Yp + A <-> YpA

YpA -> Y + A

Yp -> Y

Graphical Properties

==========

Number of complexes = 9

Number of linkage classes = 3:

Linkage class no. 1: {A + B, AB, Ap + B}

Linkage class no. 2: {Ap + Y, Yp + A, YpA, Y + A}

Linkage class no. 3: {Yp, Y}

Number of TERMINAL strong linkage classes = 3:

Strong linkage class no. 1: {A + B, AB, Ap + B}

Strong linkage class no. 2: {Y + A}

Strong linkage class no. 3: {Y}

Number of NON-TERMINAL strong linkage classes = 2:

Strong linkage class no. 4: {Ap + Y, Yp + A, YpA}

Strong linkage class no. 5: {Yp}

The network is neither reversible nor weakly reversible.

Rank Information

==========

Rank of entire network = 4

Deficiency Information

==========

Deficiency of entire network = 2

Deficiency of linkage class no. 1 = 0

Deficiency of linkage class no. 2 = 0

Deficiency of linkage class no. 3 = 0

Analysis

========

This is a deficiency two network. It is an excellent candidate for application of HIGHER DEFICIENCY THEORY (tailored mostly to networks with deficiencies greater than one).

Whether results will be obtained, will depend on whether or not the reaction network has certain additional structural attributes that help reduce the problem to a study of systems of linear inequalities.

If a network is "good", higher deficiency theory will determine, either affirmatively or negatively, whether there are positive rate constant values such that the corresponding mass action differential equations admit multiple (positive) steady states. If the answer is affirmative, higher deficiency theory will generate a sample set of rate constants and a pair of distinct steady states that are consistent with those rate constants.

If a network is "bad", some additional nonlinear analysis might be required, and the program might not be able to ascertain the network's capacity for multiple positive steady states. If definite conclusions can be reached they they will be reported. Otherwise the program will tell you that it cannot reach a conclusion.

Higher deficiency theory will also determine, either affirmatively or negatively, whether there can exist a set of rate constants such that the corresponding mass action differential equations admit a positive steady state having a zero eigenvalue (corresponding to an eigenvector in the stoichiometric subspace). When the answer is affirmative, the theory will produce such a set of rate constants, a positive steady state, and an eigenvector (in the stoichiometric subspace) corresponding to an eigenvalue of zero. Results of this kind are contained after running the Zero Eigenvalue Report.

HIGHER DEFICIENCY REPORT: NoName1

==========

Analysis

========

Taken with mass action kinetics, the network DOES have the capacity for multiple steady states. That is, there are rate constants that give rise to two or more positive (stoichiometrically compatible) steady states.

S2.4.2 Results of the analytical analysis of a model with a monofunctional kinase

BASIC REPORT: NoName1

==========

Reaction network:

Ap + Y <-> Yp + A

Yp + A <-> YpA

Yp -> Y

A <-> Ap

Graphical Properties

==========

Number of complexes = 7

Number of linkage classes = 3:

Linkage class no. 1: {Ap + Y, Yp + A, YpA}

Linkage class no. 2: {Yp, Y}

Linkage class no. 3: {A, Ap}

Number of TERMINAL strong linkage classes = 3:

Strong linkage class no. 1: {Ap + Y, Yp + A, YpA}

Strong linkage class no. 2: {A, Ap}

Strong linkage class no. 3: {Y}

Number of NON-TERMINAL strong linkage classes = 1:

Strong linkage class no. 4: {Yp}

The network is neither reversible nor weakly reversible.

Rank Information

==========

Rank of entire network = 3

Deficiency Information

==========

Deficiency of entire network = 1

Deficiency of linkage class no. 1 = 0

Deficiency of linkage class no. 2 = 0

Deficiency of linkage class no. 3 = 0

Analysis

========

This is a regular deficiency one network. It is an excellent candidate for application of DEFICIENCY ONE THEORY.

Deficiency one theory will determine, either affirmatively or negatively, whether there are positive rate constant values such that the corresponding mass action differential equations admit multiple (positive) steady states. If the answer is affirmative, deficiency one theory will

generate a sample set of rate constants and a pair of distinct steady states that are consistent with those rate constants. To get this informatoin, you should run the Deficiency One Report.

Deficiency one theory will also determine, either affirmatively or negatively, whether there can exist a set of rate constants such that the corresponding mass action differential equations admit a positive steady state having a zero eigenvalue (corresponding to an eigenvector in the stoichiometric subspace). When the answer is affirmative, the theory will produce such a set of rate constants, a positive steady state, and an eigenvector (in the stoichiometric subspace) corresponding to an eigenvalue of zero. To get this information, run the Zero Eigenvalue Report (after running the Deficiency Zero Report).

DEFICIENCY ONE REPORT: NoName1

==========

Analysis

========

Taken with mass action kinetics, the network CANNOT admit multiple positive steady states or a degenerate positive steady state NO MATTER WHAT (POSITIVE) VALUES THE RATE CONSTANTS MIGHT HAVE.

S2.4.3 Results of the analytical analysis of a model with a split kinase and a separate phosphatase

BASIC REPORT

==========

Reaction network:

A + B <-> AB

AB <-> Ap + B

Ap + Y <-> Yp + A

Yp + A <-> YpA

X + Yp <-> XYp

XYp -> X + Y

Yp -> Y

Graphical Properties

==========

Number of complexes = 11

Number of linkage classes = 4:

Linkage class no. 1: {A + B, AB, Ap + B}

Linkage class no. 2: {Ap + Y, Yp + A, YpA}

Linkage class no. 3: {X + Yp, XYp, X + Y}

Linkage class no. 4: {Yp, Y}

Number of TERMINAL strong linkage classes = 4:

Strong linkage class no. 1: {A + B, AB, Ap + B}

Strong linkage class no. 2: {Ap + Y, Yp + A, YpA}

Strong linkage class no. 3: {Y}

Strong linkage class no. 4: {X + Y}

Number of NON-TERMINAL strong linkage classes = 2:

Strong linkage class no. 5: {X + Yp, XYp}

Strong linkage class no. 6: {Yp}

The network is neither reversible nor weakly reversible.

Rank Information

==========

Rank of entire network = 5

Deficiency Information

==========

Deficiency of entire network = 2

Deficiency of linkage class no. 1 = 0

Deficiency of linkage class no. 2 = 0

Deficiency of linkage class no. 3 = 0

Deficiency of linkage class no. 4 = 0

Analysis

========

This is a deficiency two network. It is an excellent candidate for application of HIGHER DEFICIENCY THEORY (tailored mostly to networks with deficiencies greater than one).

Whether results will be obtained, will depend on whether or not the reaction network has certain additional structural attributes that help reduce the problem to a study of systems of linear inequalities.

HIGHER DEFICIENCY REPORT

==========

Analysis

========

Taken with mass action kinetics, the network CANNOT admit multiple positive steady states or a degenerate positive steady state NO MATTER WHAT (POSITIVE) VALUES THE RATE CONSTANTS MIGHT HAVE.

S2.4.4 Results of the analytical analysis of a model with a bifunctional, non-split kinase

BASIC REPORT

==========

Reaction network:

Ap + Y <-> A + Yp

A + Yp <-> AYp

AYp -> A + Y

Yp -> Y

A <-> Ap

Graphical Properties

==========

Number of complexes = 8

Number of linkage classes = 3:

Linkage class no. 1: {Ap + Y, A + Yp, AYp, A + Y}

Linkage class no. 2: {Yp, Y}

Linkage class no. 3: {A, Ap}

Number of TERMINAL strong linkage classes = 3:

Strong linkage class no. 1: {A, Ap}

Strong linkage class no. 2: {Y}

Strong linkage class no. 3: {A + Y}

Number of NON-TERMINAL strong linkage classes = 2:

Strong linkage class no. 4: {Ap + Y, A + Yp, AYp}

Strong linkage class no. 5: {Yp}

The network is neither reversible nor weakly reversible.

Rank Information

==========

Rank of entire network = 3

Deficiency Information

==========

Deficiency of entire network = 2

Deficiency of linkage class no. 1 = 0

Deficiency of linkage class no. 2 = 0

Deficiency of linkage class no. 3 = 0

Analysis

========

This is a deficiency two network. It is an excellent candidate for application of HIGHER DEFICIENCY THEORY (tailored mostly to networks with deficiencies greater than one).

Whether results will be obtained, will depend on whether or not the reaction network has certain additional structural attributes that help reduce the problem to a study of systems of linear inequalities.

HIGHER DEFICIENCY REPORT

==========

Analysis

========

Taken with mass action kinetics, the network CANNOT admit multiple positive steady states or a degenerate positive steady state NO MATTER WHAT (POSITIVE) VALUES THE RATE CONSTANTS MIGHT HAVE.

S2.4.5 Results of the analytical analysis of a model with a monofunctional, split kinase

BASIC REPORT

==========

Reaction network:

A + B <-> AB

AB <-> Ap + B

Ap + Y <-> Yp + A

Yp + A <-> YpA

Yp -> Y

Graphical Properties

==========

Number of complexes = 8

Number of linkage classes = 3:

Linkage class no. 1: {A + B, AB, Ap + B}

Linkage class no. 2: {Ap + Y, Yp + A, YpA}

Linkage class no. 3: {Yp, Y}

Number of TERMINAL strong linkage classes = 3:

Strong linkage class no. 1: {A + B, AB, Ap + B}

Strong linkage class no. 2: {Ap + Y, Yp + A, YpA}

Strong linkage class no. 3: {Y}

Number of NON-TERMINAL strong linkage classes = 1:

Strong linkage class no. 4: {Yp}

The network is neither reversible nor weakly reversible.

Rank Information

==========

Rank of entire network = 4

Deficiency Information

==========

Deficiency of entire network = 1

Deficiency of linkage class no. 1 = 0

Deficiency of linkage class no. 2 = 0

Deficiency of linkage class no. 3 = 0

Analysis

========

This is a regular deficiency one network. It is an excellent candidate for application of DEFICIENCY ONE THEORY.

Deficiency one theory will determine, either affirmatively or negatively, whether there are positive rate constant values such that the corresponding mass action differential equations admit multiple (positive) steady states. If the answer is affirmative, deficiency one theory will generate a sample set of rate constants and a pair of distinct steady states that are consistent with those rate constants. To get this informatoin, you should run the Deficiency One Report.

DEFICIENCY ONE REPORT

==========

Analysis

========

Taken with mass action kinetics, the network CANNOT admit multiple positive steady states or a degenerate positive steady state NO MATTER WHAT (POSITIVE) VALUES THE RATE CONSTANTS MIGHT HAVE.

S2.5 Figures

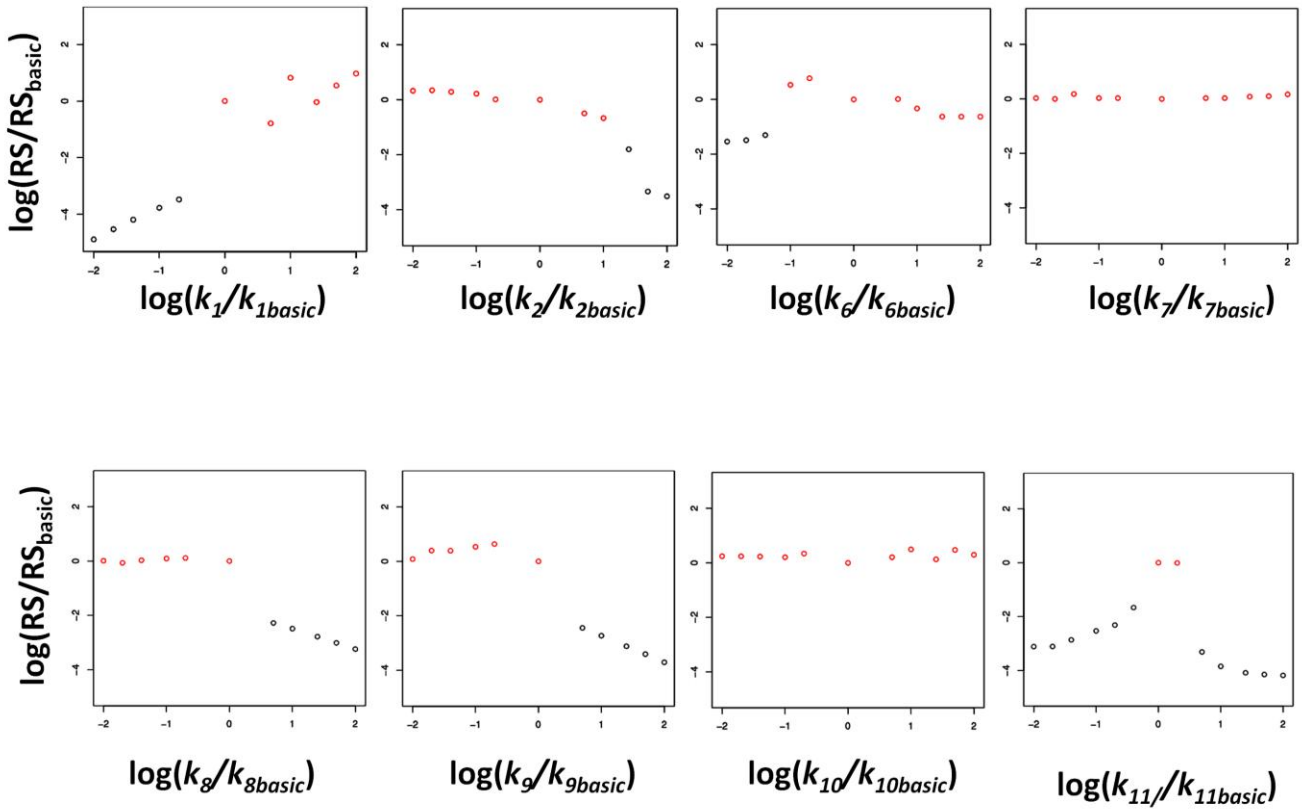

Figure S2.1: The sensitivity of the signal response curve "sigmoidality" to parameter changes. The "sigmoidality" of the signal-response curve, RS, is measured as its maximum slope (s_{max}) multiplied by the signal level at which this slope occurs ($k5_s$) (i.e. RS = $k5_s \times s_{max}$). On each panel, the y-axis shows the ratio of RS, resulting from models with different values of a specific parameter, to that resulting from the basic model. x-axis shows the ratio of this parameter value to its corresponding value in the basic model. Data points in red

indicates presence of bistability in the signal-response relationship. Note the log scale on both axes.

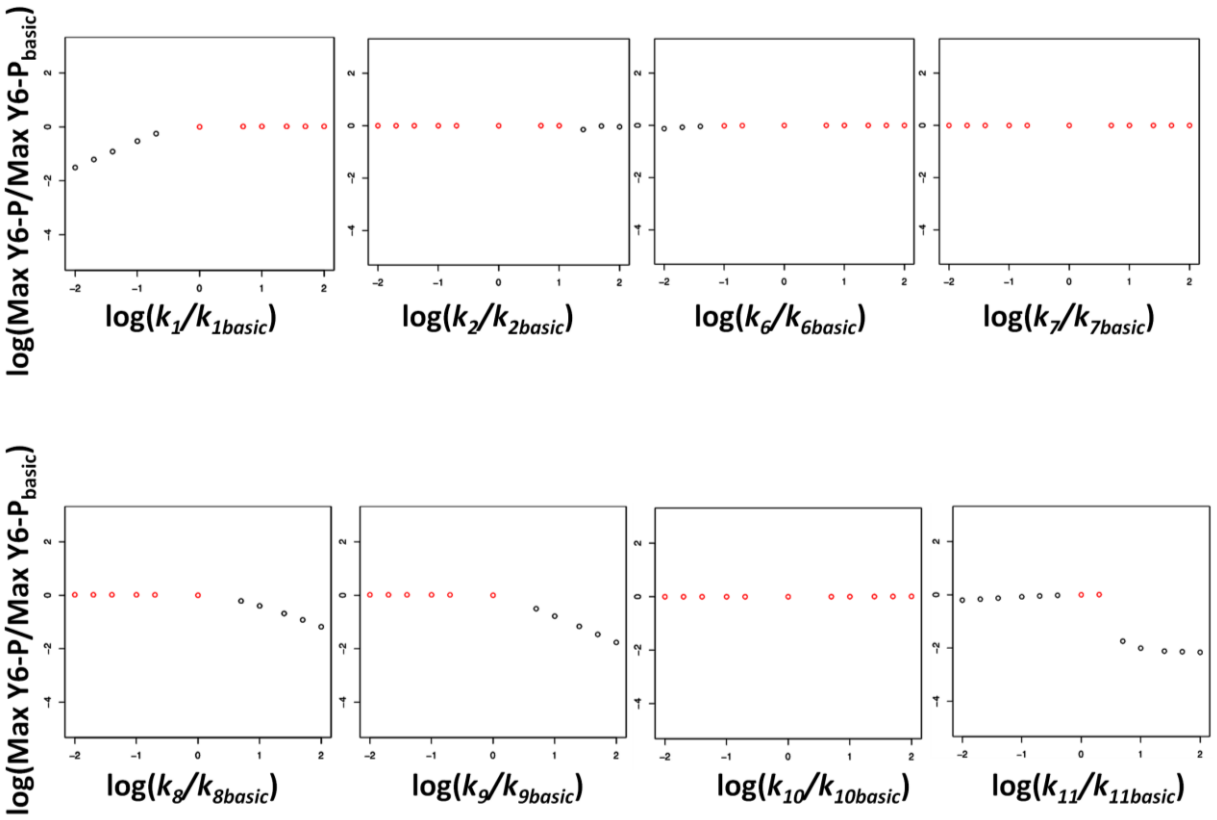

Figure S2.2: The sensitivity of the maximum phosphorylation level of CheY6 to parameter changes. On each panel, the y-axis shows the ratio of the maximal CheY6 phosphorylation, resulting from models with different values of a specific parameter, to that resulting from the basic model. x-axis shows the ratio of this parameter value to its corresponding value in the basic model. Data points in red indicates presence of bistability in the signal-response relationship. Note the log scale on both axes.

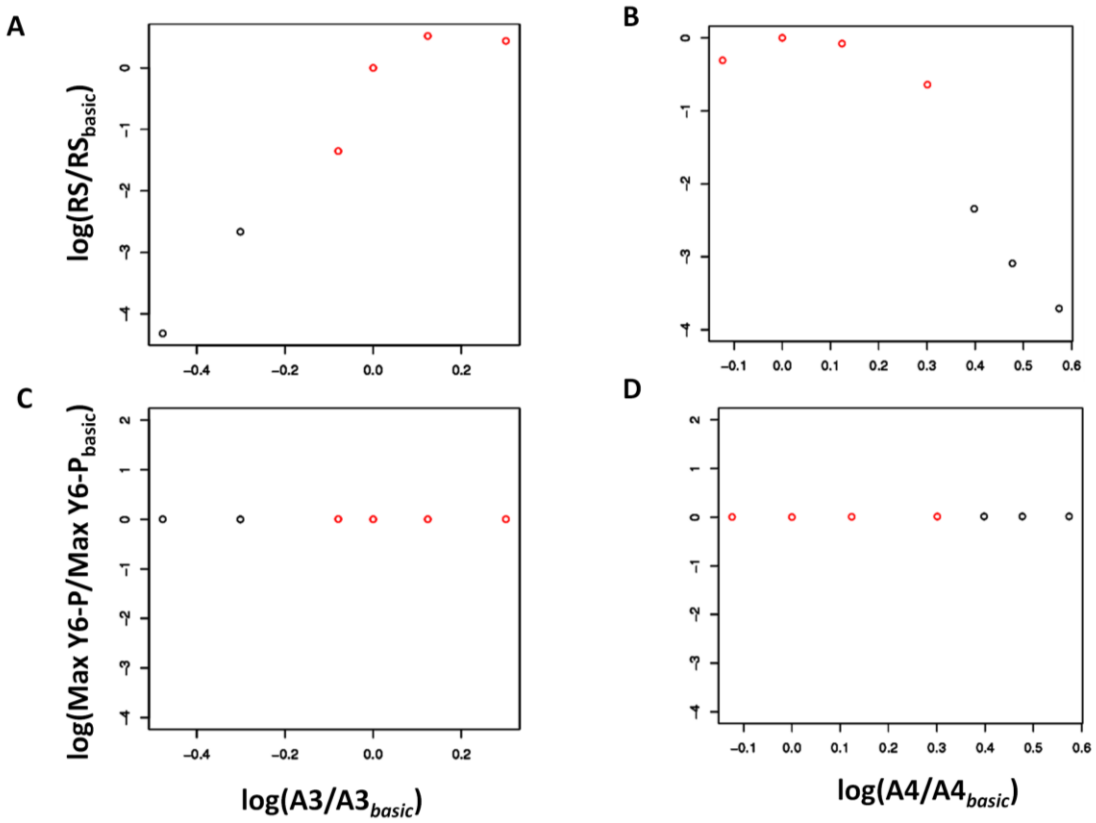

Figure S2.3: The sensitivity of the signal response curve "sigmoidality" to changes in the concentration of CheA3 (A) and CheA4 (B). The "sigmoidality" of the signal-response curve, RS, is measured as its maximum slope (s_{max}) multiplied by the signal level at which this slope occurs ($k5_s$) (i.e. RS = $k5_s \times s_{max}$). On panel A (B), the y-axis shows the ratio of RS, resulting from models with different values of CheA3 (CheA4) concentration, to that resulting from the basic model. x-axis shows the ratio of this concentration to its corresponding value in the basic model. Data points in red indicates presence of bistability in the signal-response relationship. The sensitivity of the maximum phosphorylation level of CheY6 to changes in the concentration of CheA3 (C) and CheA4 (D). On panel C (D), the y-axis shows the ratio of the maximal CheY6 phosphorylation, resulting from models with different values of CheA3 (CheA4) concentration, to that resulting from the basic model. x-axis shows the ratio of this concentration to its corresponding value in the basic model. Data points in red indicates presence of bistability in the signal-response relationship. Note the log scale on both axes on

all panels.

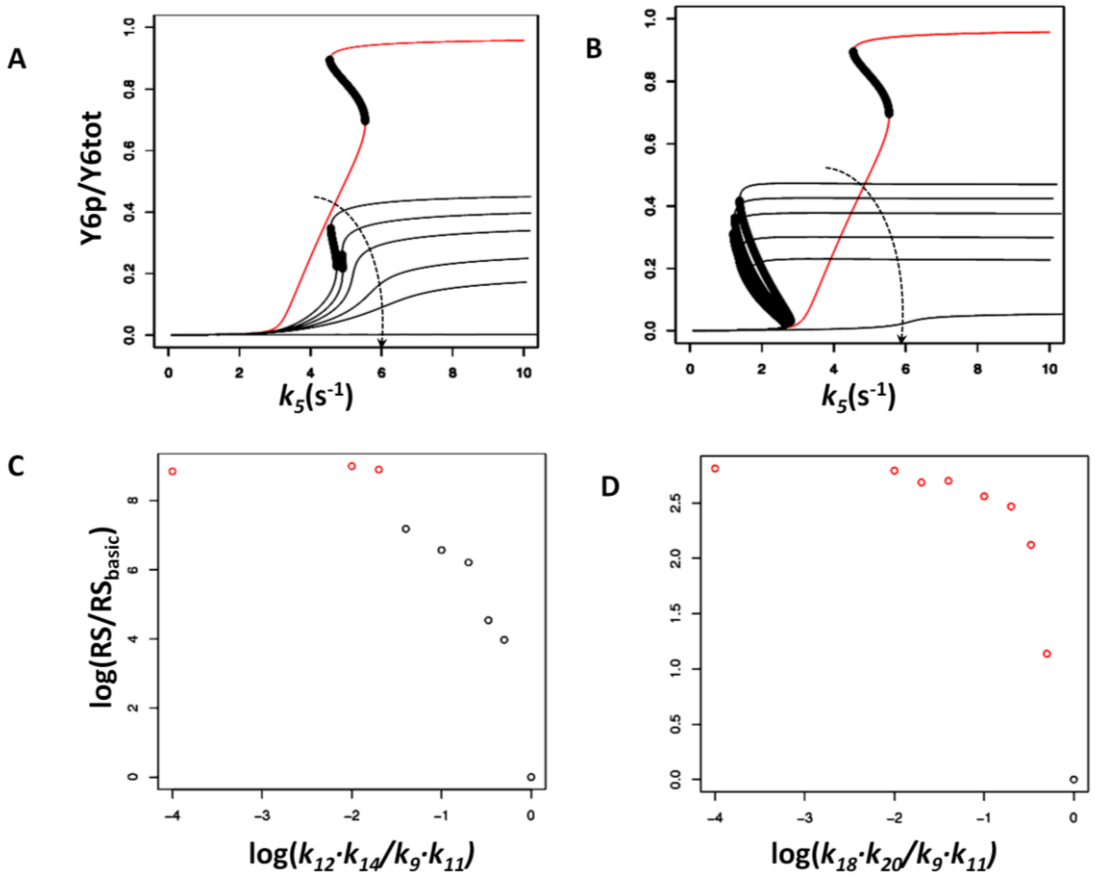

Figure S2.4: Analysis of signal-response relationship, in an alternative model considering phosphatase activity from additional species (see Supplementary Information, section 1). (A) Signal-response curves resulting from a model where both CheA3:CheA4 and CheA3:CheA4:ATP are considered to have phosphatase activity in addition to CheA3. For comparison, signal-response curve from the basic model is shown in red. Where present, the dark region indicates the region of unstable steady states and hence the presence of bistability. The different curves correspond to increasing levels of phosphatase activity (shown with the arrow) from the additional species. Phosphatase activity is varied in the same way for both CheA3:CheA4 and CheA3:CheA4:ATP by assuming that k_{on} and k_{cat} for these species are the same (i.e. $k1_2 = k1_5$ and $k1_4 = k1_7$) and by varying one set of rates simultaneously. The ratio between these rates ($k1_2$ and $k1_4$) to their corresponding values for

CheA3 ($k9$ and $k1_1$) is shown on the x-axis of panel C. (B) Signal-response curves resulting from a model where CheA-P is considered to have phosphatase activity in addition to CheA3. For comparison, signal-response curve from the basic model is shown in red. Where present, the dark region indicates the region of unstable steady states and hence the presence of bistability. The different curves correspond to increasing levels of phosphatase activity (shown with the arrow) from CheA3-P. Phosphatase activity is varied by changing both k_{on} and k_{cat} for CheA3-P (i.e. $k1_8$ and $k2_0$) simultaneously. The ratio between these rates ($k1_8$ and $k2_0$) to their corresponding values for CheA3 ($k9$ and $k1_1$) is shown on the x-axis of panel D. (C) The sensitivity of the signal response curve "sigmoidality" to increasing phosphatase activity from CheA3:CheA4 and CheA3:CheA4:ATP. The "sigmoidality" of the signal-response curve, RS, is measured as its maximum slope (s_{max}) multiplied by the signal level at which this slope occurs ($k5_s$) (i.e. RS = $k5_s \times s_{max}$). y-axis shows the ratio of RS, resulting from models with increasing phosphatase activity by additional species, to that of resulting from the basic model. X-axis shows the ratio of kinetic rates governing phosphatase activity ($k1_2$ and $k1_4$) to those in the basic model ($k9$ and $k1_1$). Data points in red indicates presence of bistability in the signal-response relationship. (D) The sensitivity of the signal response curve "sigmoidality" to increasing phosphatase activity from CheA3-P. The "sigmoidality" of the signal-response curve, RS, is measured as its maximum slope (s_{max}) multiplied by the signal level at which this slope occurs ($k5_s$) (i.e. RS = $k5_s \times s_{max}$). Y-axis shows the ratio of RS, resulting from models with increasing phosphatase activity by additional species, to that of resulting from the basic model. x-axis shows the ratio of kinetic rates governing phosphatase activity ($k1_8$ and $k2_0$) to those in the basic model ($k9$ and $k1_1$). Data points in red indicates presence of bistability in the signal-response relationship. Note the log scale on both axes in panels C and D.

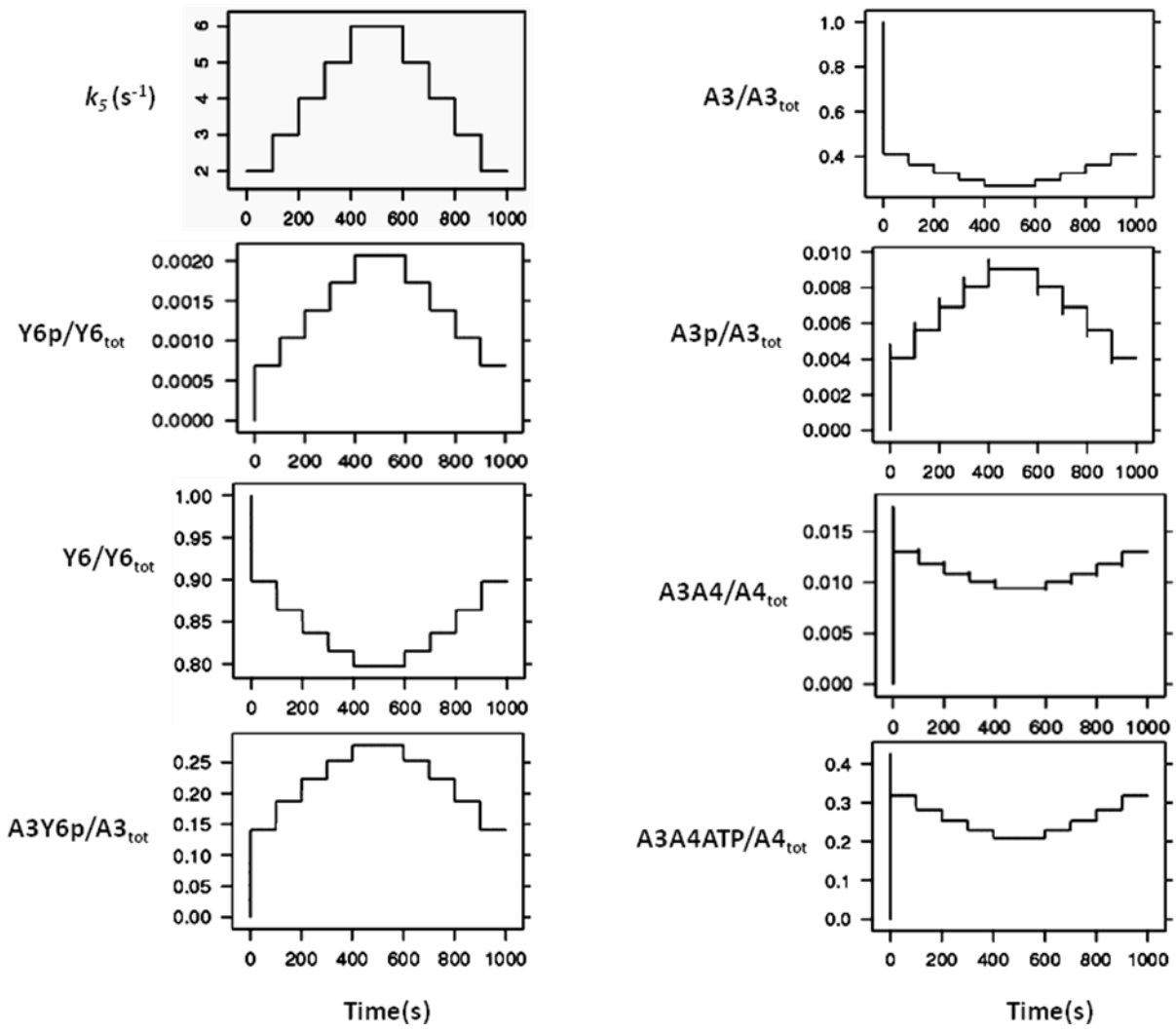

Figure S2.5: Time-course analysis using an alternative model where both CheA3:CheA4 and CheA3:CheA4:ATP are considered to have phosphatase activity in addition to CheA3 (see Supplementary Information, section 1). The model is simulated with increasing and decreasing signal levels (*k5*) in course of time. *k5* is increased from 2 to 6 and decreased in similar fashion at indicated time points (top most, left panel), and changes in each species were measured (as indicated on each panel). The x- and y-axis represent time and species concentration respectively, where the latter is normalized by the appropriate total protein levels.

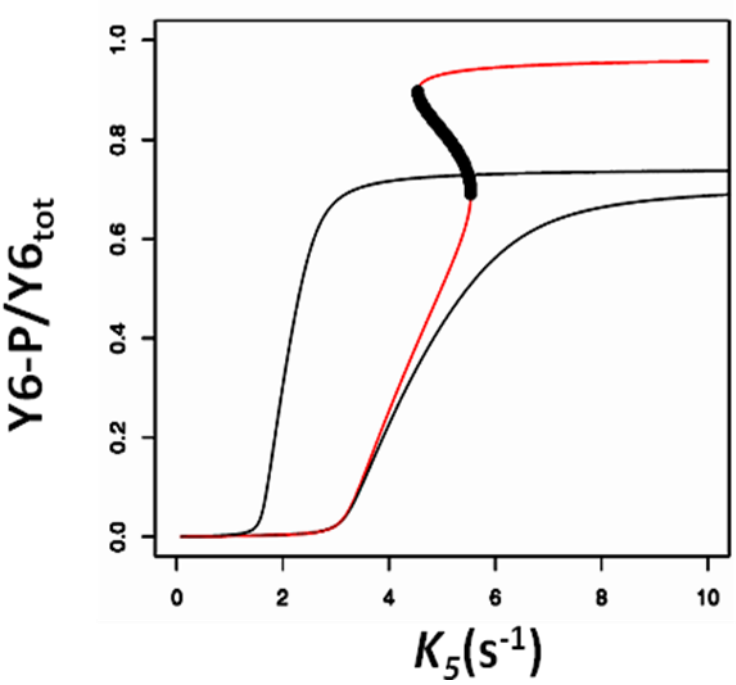

Figure S2.6: Signal-response curves resulting from an alternative model that allows for the possibility that phosphorylated CheA3 remains in complex with CheA4 and that this CheA3p:CheA4 complex is also capable of acting as phosphatase towards CheY6p (see Supplementary Information, section 2). The y-axis shows steady state Y6-P level normalised by total Y6, while x-axis shows signal (*k5*) level. Where present, a dark region indicates the region of unstable steady states and hence the presence of bistability. (a) The signal-response curve from the basic model (included for comparison). (b) Signal-response curve from the alternative model and simulating signal level through changing both k'_5 and *k5* simultaneously. (c) Signal-response curve from the alternative model and simulating signal level through changing *k5*, while k'_5=0.1 s-1.

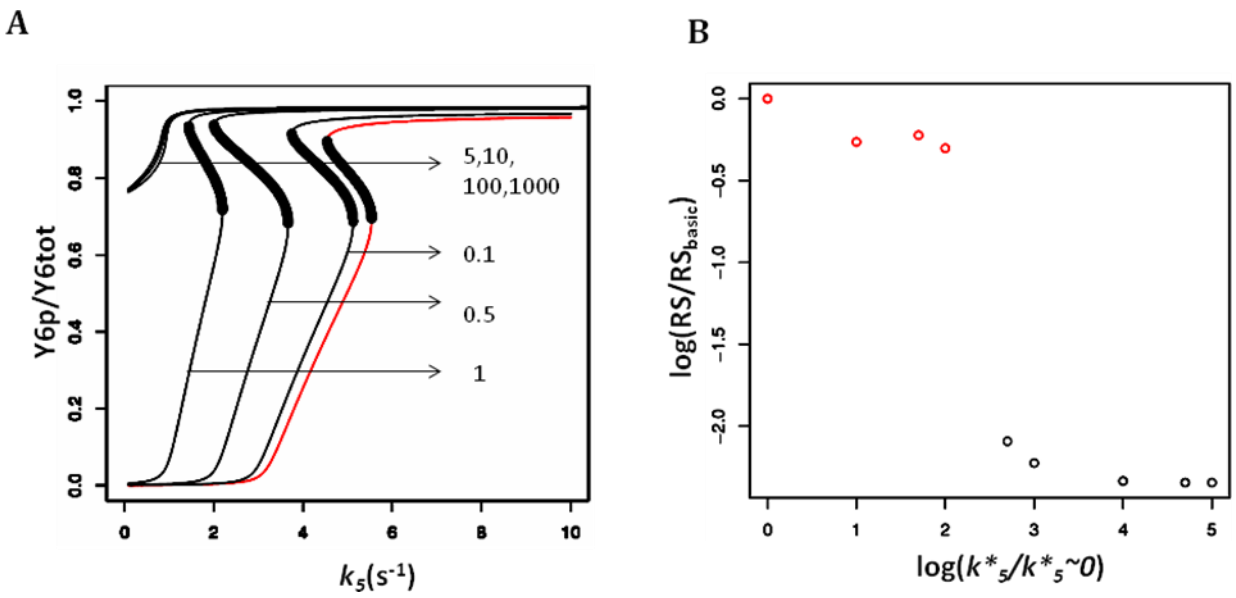

Figure S2.7: Analysis of signal-response relationship, in an alternative model considering additional kinase activity (see Supplementary Information, section 3). (A) Signal-response curves resulting from a model where additional kinase activity (from CheA2) is considered. For comparison, the signal-response curve from the basic model is shown in red. Where present, the dark region indicates the region of unstable steady states and hence the presence of bistability. The different curves correspond to increasing levels of autophosphorylation rates for CheA2 (i.e. increasing background signalling through CheA2). (B) The sensitivity of the signal-response "sigmoidality" with increasing background kinase activity (from CheA2). The "sigmoidality" of the signal-response curve, RS, is measured as its maximum slope (s_{max}) multiplied by the signal level at which this slope occurs ($k5_s$) (i.e. RS = $k5_s \times s_{max}$). y-axis shows the ratio of RS, resulting from models with increasing background kinase activity (k^*_5) to that of the case where such activity is minimal (i.e. $k^*_5 \sim 0$). Data points in red indicates presence of bistability in the signal-response relationship. Note the log scale on both axes.

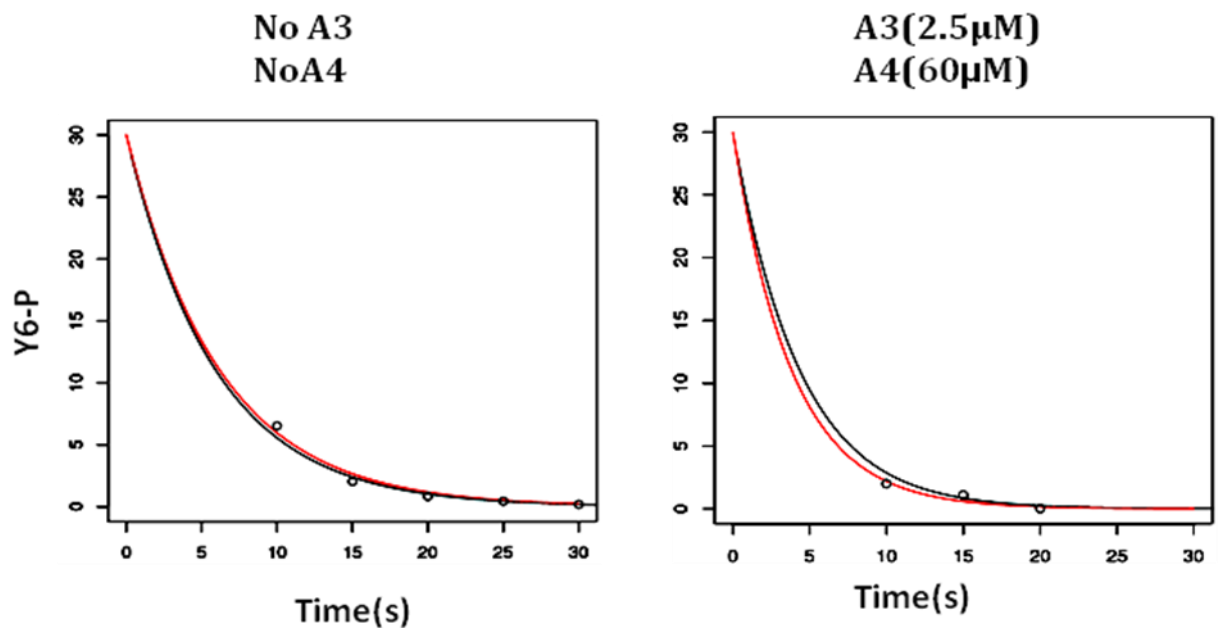

Figure S2.8: CheY6-P dephosphorylation time course data (circles) along with the fitted first-order exponential decay curves (red line) and simulated data (black line). The exponential fits are used to derive an estimate for overall CheY6p dephosphorylation rate (k_{obs}), which are shown in Figure 4.

S2.6 Tables

Parameter	Description	Value	Unit
k_9	Association of phosphatase (CheA3) assisted dephosphorylation complex	5.6	$(\mu Ms)^{-1}$
$k1_2$	Association of phosphatase (CheA3CheA4) assisted dephosphorylation complex	5.6	$(\mu Ms)^{-1}$
$k1_5$	Association of phosphatase (CheA3CheA4ATP) assisted dephosphorylation complex	5.6	$(\mu Ms)^{-1}$
$k1_8$	Association of phosphatase (CheA3-P) assisted dephosphorylation complex	5.6	$(\mu Ms)^{-1}$

$k1_0$	Dissociation of phosphatase (CheA3) assisted dephosphorylation complex	0.04	s^{-1}
$k1_3$	Dissociation of phosphatase assisted (CheA3CheA4) dephosphorylation complex	0.04	s^{-1}
$k1_6$	Dissociation of phosphatase (CheA3CheA4ATP) assisted dephosphorylation complex	0.04	s^{-1}
$k1_9$	Dissociation of phosphatase (CheA3-P) assisted dephosphorylation complex	0.04	s^{-1}
$k1_1$	K_{cat} for phosphatase (CheA3) assisted dephosphorylation	2.5	s^{-1}
$k1_4$	K_{cat} for phosphatase (CheA3CheA4) assisted dephosphorylation	2.5	s^{-1}
$k1_7$	K_{cat} for phosphatase (CheA3CheA4ATP) assisted dephosphorylation	2.5	s^{-1}
k_{20}	K_{cat} for phosphatase (CheA3-P) assisted dephosphorylation	2.5	s^{-1}

Table S2.1. Parameter values used for the models with additional phosphatases.

Parameter	Description	Value	Unit
k1	On rate for binding of CheA3 and CheA4	100	(μMs^{-1})
k'_1	On rate for binding of CheA3-P and CheA4	100	(μMs^{-1})
k_2	Off rate for binding of CheA3 and CheA4	10	s^{-1}
k'_2	Off rate for binding of CheA3-P and CheA4	10	s^{-1}
k5	K_{cat} for phosphorylation of CheA3 by CheA4	varied	s^{-1}
k'_6	CheA4/CheA3-P to CheY6 Phosphotransfer	0.775	(μMs) -1
k''_6	CheA4/CheA3-P to CheY6 Phosphotransfer	0.775	(μMs) -1

k_7	CheA3-P to CheY6 Reverse phosphotransfer	0.00283	(μMs) -1
k'_7	CheA4/CheA3-P to CheY6 Reverse phosphotransfer	0.00283	(μMs) -1
k''_7	CheA4/CheA3-P to CheY6 Reverse phosphotransfer	0.00283	(μMs) -1

Table S2.2. Parameter values used for the models with alternative reaction scheme.

Parameter	Description	Value	Unit
k_3	Forward rate for phosphorylation complex	1	$(\mu Ms)^{-1}$
k_4	Reverse rate for phosphorylation complex	39	s^{-1}
k^*_3	Forward rate for phosphorylation complex with CheA2	1	$(\mu Ms)^{-1}$
k^*_4	Reverse rate for phosphorylation complex with CheA2	39	s^{-1}
k_5	K_{cat} for phosphorylation of CheA3 by CheA4	varied	s^{-1}
k^*_5	K_{cat} for phosphorylation of CheA2	varied	s^{-1}
k^*_6	CheA2-P to CheY6 Phosphotransfer	0.775	$(\mu Ms)^{-1}$
k^*_7	CheA2-P to CheY6 Reverse phosphotransfer	0.00283	$(\mu Ms)^{-1}$

Table S2.3. Parameter values used for the models with additional kinase.

Parameter	Description	Value	Unit
k1	On rate for binding of CheA3 and CheA4	100	(μMs^{-1})
k_2	Off rate for binding of CheA3 and CheA4	10	s^{-1}
k'_1	On rate for binding of CheA3P1 and CheA4	100	(μMs^{-1})

k''_1	On rate for binding of CheA3P1-P and CheA4	100	(μMs^{-1})
k'_2	Off rate for binding of CheA3-P1 and CheA4	10	s^{-1}
k''_2	Off rate for binding of CheA3P1-P and CheA4	10	s^{-1}
k_6	CheA3-P to CheY6 Phosphotransfer	0.775	$(\mu Ms)^{-1}$
k_7	CheA3-P to CheY6 Reverse phosphotransfer	0.00283	$(\mu Ms)^{-1}$
k'_6	CheA4/CheA3P1-P to CheY6 Phosphotransfer	0.775	$(\mu Ms)^{-1}$
k'_7	CheA4/CheA3P1-P to CheY6 Reverse phosphotransfer	0.00283	$(\mu Ms)^{-1}$
k_8	Autodephosphorylation of CheY6-P	0.169	s^{-1}
k_9	Association of phosphatase assisted dephosphorylation complex	5.6	$(\mu Ms)^{-1}$
$k1_0$	Dissociation of phosphatase assisted dephosphorylation complex	0.04	s^{-1}
$k1_1$	k_{cat} for phosphatase assisted dephosphorylation	2.5	s^{-1}
$[A3]_{tot}$	Total concentration of CheA3	2.5	µM
$[A4]_{tot}$	Total concentration of CheA4	0,20,40,60	µM
$[Y6]_{tot}$	Total concentration of CheY6	100	µM
[A3P1-P]	Total concentration of CheA3P1-P	30	µM

Table S2.4. Parameter values used for the model of the *in vitro* experimental system.

Appendix B

In this appendix, additional text, analyses, figures and table are described related to chapter 3

.

List of contents

1 Mathematical analysis of the simple model

1.1 Reactions

The model has 6 species: one histidine kinase HK, two response regulators R1, R2 and their corresponding phosphorylated forms: HKp, R1p, R2p. The first model we analyze consists of the following reactions:

- Auto-phosphorylation of HK:

$$\mathrm{HK} \xrightarrow{k_a} \mathrm{HKp}$$

- Reversible phosphotransfer between HK and R1:

$$\mathrm{HKp} + \mathrm{R1} \xrightarrow{k_S} \mathrm{HK} + \mathrm{R1p} \qquad \mathrm{HK} + \mathrm{R1p} \xrightarrow{k_{rS}} \mathrm{HKp} + \mathrm{R1}$$

- Reversible phosphotransfer between HK and R2:

$$\mathrm{HKp} + \mathrm{R2} \xrightarrow{k_M} \mathrm{HK} + \mathrm{R2p} \qquad \mathrm{HK} + \mathrm{R2p} \xrightarrow{k_{rM}} \mathrm{HKp} + \mathrm{R2}$$

- Auto-dephosphorylation of the response regulators:

$$\mathrm{R1p} \xrightarrow{k_{hS}} \mathrm{R1} \qquad \mathrm{R2p} \xrightarrow{k_{hM}} \mathrm{R2}$$

1.2 System of ODEs

To simplify the notation, we write the concentration of each of the species as:

$$x_1 = [\mathrm{HK}] \qquad x_2 = [\mathrm{HKp}] \qquad x_3 = [\mathrm{R1}] \qquad x_4 = [\mathrm{R1p}] \qquad x_5 = [\mathrm{R2}] \qquad x_6 = [\mathrm{R2p}].$$

We let $x = (x_1, \ldots, x_6)$ be the vector of concentrations. We use mass-action kinetics to model the dynamics of the concentrations of the species in time with a system of ordinary differential equations (ODEs). We write the ODE system as the product of the stoichiometric matrix A and the vector of reaction rates. We order the reactions as shown above and the species analogously to the concentrations order $x_1, \ldots, x_6$. The stoichiometric matrix A and the vector of reaction rates $v(x)$ are:

$$A = \begin{pmatrix} -1 & 1 & -1 & 1 & -1 & 0 & 0 \\ 1 & -1 & 1 & -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 \end{pmatrix} \qquad v(x) = \begin{pmatrix} k_a x_1 \\ k_S x_2 x_3 \\ k_{rS} x_1 x_4 \\ k_M x_2 x_5 \\ k_{rM} x_1 x_6 \\ k_{hS} x_4 \\ k_{hM} x_6 \end{pmatrix}$$

We let $\dot{x}(t)$ denote $\frac{dx(t)}{dt}$ and drop the dependency on time t in the notation, that is, we write $\dot{x}_i$ and x_i for $\dot{x}_i(t)$ and $x_i(t)$ respectively. The system of ODEs modeling the dynamics of the concentrations in time is $\dot{x} = Av(x)$, that is:

$$\dot{x}_1 = k_M x_2 x_5 + k_S x_2 x_3 - k_{rM} x_1 x_6 - k_{rS} x_1 x_4 - k_a x_1 \tag{S1}$$

$$\dot{x}_2 = -k_M x_2 x_5 - k_S x_2 x_3 + k_{rM} x_1 x_6 + k_{rS} x_1 x_4 + k_a x_1 \tag{S2}$$

$$\dot{x}_3 = -k_S x_2 x_3 + k_{rS} x_1 x_4 + k_{hS} x_4 \tag{S3}$$

$$\dot{x}_4 = k_S x_2 x_3 - k_{rS} x_1 x_4 - k_{hS} x_4 \tag{S4}$$

$$\dot{x}_5 = -k_M x_2 x_5 + k_{rM} x_1 x_6 + k_{hM} x_6 \tag{S5}$$

$$\dot{x}_6 = k_M x_2 x_5 - k_{rM} x_1 x_6 - k_{hM} x_6 \tag{S6}$$

At steady state, $\dot{x}_i = 0$ for all $i = 1, \ldots, 6$.

The system has three conservation laws:

$$\dot{x}_1 + \dot{x}_2 = 0, \qquad \dot{x}_3 + \dot{x}_4 = 0, \qquad \dot{x}_5 + \dot{x}_6 = 0,$$

which reflect that the sum of species concentration in unphosphorylated and phosphorylated form is constant at each layer. These relations provide three equations at steady state:

$$\overline{H} = x_1 + x_2, \qquad \overline{R}_S = x_3 + x_4, \qquad \overline{R}_M = x_5 + x_6, \tag{S7}$$

for some positive total amounts $\overline{H}, \overline{R}_S, \overline{R}_M$. Due to these relations, three of the ODEs, one for each conservation law, are redundant for the computation of the steady states. We choose to disregard (S1), (S3) and (S5).

Therefore, the steady states of the system are the solutions to the steady state equations corresponding to (S2), (S4) and (S6) and the three conservation laws in (S7). That is, the steady states are the solutions to the following six equations:

$$\begin{aligned}
0 &= -k_M x_2 x_5 - k_S x_2 x_3 + k_{rM} x_1 x_6 + k_{rS} x_1 x_4 + k_a x_1 \\
0 &= k_S x_2 x_3 - k_{rS} x_1 x_4 - k_{hS} x_4 \\
0 &= k_M x_2 x_5 - k_{rM} x_1 x_6 - k_{hM} x_6 \\
0 &= x_1 + x_2 - \overline{H} \\
0 &= x_3 + x_4 - \overline{R}_S \\
0 &= x_5 + x_6 - \overline{R}_M.
\end{aligned}$$

The equations are easier to analyze if we replace the first equation by the sum of the first three

equations. This change does not alter the solutions to the system. The new system to be solved is:

$$0 = k_a x_1 - k_{hS} x_4 - k_{hM} x_6 \tag{S8}$$

$$0 = k_S x_2 x_3 - k_{rS} x_1 x_4 - k_{hS} x_4 \tag{S9}$$

$$0 = k_M x_2 x_5 - k_{rM} x_1 x_6 - k_{hM} x_6 \tag{S10}$$

$$0 = x_1 + x_2 - \overline{H} \tag{S11}$$

$$0 = x_3 + x_4 - \overline{R}_S \tag{S12}$$

$$0 = x_5 + x_6 - \overline{R}_M. \tag{S13}$$

1.3 Expressions for the steady states

We recursively write all concentrations at steady state as a function of x_6. We further impose that all concentrations are positive at steady state.

We **solve for** x_5 at steady state using (S13) and obtain

$$x_5 = \overline{R}_M - x_6. \tag{S14}$$

The right-hand side of this equality decreases in x_6. Further, x_5, x_6 are positive provided that $0 < x_6 < \overline{R}_M$. Let b_1 be this first upper bound for x_6 at a positive steady state, that is

$$b_1 := \overline{R}_M. \tag{S15}$$

We next **solve for** x_1, x_2 using (S11) and (S10) and obtain

$$x_1 = \frac{k_M \overline{H} x_5 - k_{hM} x_6}{k_M x_5 + k_{rM} x_6}, \qquad x_2 = \frac{x_6 \left(k_{rM} \overline{H} + k_{hM}\right)}{k_M x_5 + k_{rM} x_6}. \tag{S16}$$

At steady state x_2 is positive provided x_5, and x_6 are positive. For x_1 to be positive, we need $k_M \overline{H} x_5 - k_{hM} x_6 > 0$. Substituting x_5 by the expression in (S14), this is equivalent to require that

$k_M\overline{H}(\overline{R}_M - x_6) - k_{hM}x_6 > 0$. Isolating x_6 we obtain that the numerator of x_1 is positive if and only if $x_6 < b_2$, where

$$b_2 := \frac{k_M\overline{H}\,\overline{R}_M}{k_M\overline{H} + k_{hM}}. \tag{S17}$$

When $x_6 = b_2$ then $k_M\overline{H}(\overline{R}_M - x_6) - k_{hM}x_6 = 0$ and hence $x_1 = 0$.

Since $\frac{k_M\overline{H}}{k_M\overline{H}+k_{hM}} < 1$, we have that

$$b_2 = \frac{k_M\overline{H}\,\overline{R}_M}{k_M\overline{H} + k_{hM}} < \overline{R}_M = b_1$$

and hence the upper bound b_1 of x_6 in (S15) is larger than b_2 in (S17). Therefore $x_6 < b_1$ is satisfied if also $x_6 < b_2$ and the upper bound for x_6, b_1, can be ignored.

The expression for x_1 in (S16) decreases in x_6 and increases in x_5. Since x_5 decreases in x_6, we conclude that x_1 decreases in x_6. Since $x_2 = \overline{H} - x_1$ (see (S11)), we have that x_2 decreases in x_1 and hence x_2 increases in x_6.

We **solve for** x_3, x_4 using (S9) and (S12) and obtain

$$x_3 = \frac{(k_{rS}x_1 + k_{hS})\overline{R}_S}{k_S x_2 + k_{rS}x_1 + k_{hS}}, \qquad x_4 = \frac{x_2 k_S \overline{R}_S}{k_S x_2 + k_{rS}x_1 + k_{hS}} \tag{S18}$$

At steady state both x_3, x_4 are positive provided x_1, x_2 are positive. The expression for x_3 decreases in x_2 and increases in x_1. Since x_1 decreases in x_6 and x_2 increases in x_6 we conclude that x_3 decreases in x_6. Since from (S12) we have that $x_3 = \overline{R}_S - x_4$, we conclude that x_4 increases in x_6.

Summary: By iterative substitution, all concentrations at steady state are expressed as functions of x_6. Further, all steady state concentrations are positive if and only if

$$0 < x_6 < \frac{k_M\overline{H}\overline{R}_M}{k_M\overline{H} + k_{hM}}. \tag{S19}$$

1.4 Signal-response curve and maximal response

The value of x_6 at steady state is determined using the only equation not used so far, (S8), after writing all other concentrations as functions of x_6. This equation provides the analytical description of the inverse of the signal-response curve, written as $k_a = f(x_6)$.

Solving (S8) for k_a we obtain

$$k_a = f(x_6) = \frac{k_{hM}x_6 + k_{hS}x_4}{x_1}. \tag{S20}$$

Since x_1 decreases in x_6 and x_4 increases in x_6, the function f is increasing in x_6. For positive steady states, this function needs to be evaluated at the values of x_6 satisfying (S19). When $x_6 = 0$ then $k_a = 0$. As we observed above, when x_6 approaches b_2 we have that x_1 approaches 0, x_2 approaches $\overline{H}$, and hence using (S18) we have that x_4 approaches $\frac{k_S\overline{H}R_S}{k_S\overline{H}+k_{hS}}$. Therefore it follows from the expression (S20) that k_a tends to infinity when x_6 tends to the upper bound b_2.

We conclude that for x_6 fulfilling (S19), f is an increasing function with range $(0, +\infty)$. It follows that given a value of signal, $k_a > 0$, then there exists a unique value of x_6 fulfilling (S19) and $k_a = f(x_6)$. This is the steady state value of x_6 for the given input k_a. The concentrations of the other species at steady state are uniquely determined using their expressions as functions of x_6. As a consequence, we showed that multistationarity cannot occur.

The explicit expression of the function $f(x_6)$ in (S20) is:

$$f(x_6) = \frac{x_6 p_1(x_6) p_2(x_6)}{q_1(x_6) q_2(x_6)} \tag{S21}$$

where

$$q_1(x_6) = k_M(k_{rS}\overline{H} + k_{hS})(\overline{R}_M - x_6) + k_{rM}x_6(k_S\overline{H} + k_{hS}) + k_{hM}x_6(k_S - k_{rS})$$

$$q_2(x_6) = k_M\overline{H}(\overline{R}_M - x_6) - k_{hM}x_6$$

$$p_1(x_6) = \overline{R}_S k_S k_{hS}(k_{rM}\overline{H} + k_{hM}) + k_{hM}q_1(x_6)$$

$$p_2(x_6) = k_M(\overline{R}_M - x_6) + k_{rM}x_6.$$

The signal-response curve is the inverse of the function f. Since the function $f(x_6)$ increases, so does the signal-response curve. Further, we conclude that the upper bound b_2 is the level of phosphorylated R2 that the system attains when k_a increases infinitely. In other words,

$$r_{max} = \frac{k_M\overline{H}\,\overline{R}_M}{k_M\overline{H} + k_{hM}} \tag{S22}$$

is the maximal level of response of the system. The maximal response levels for HK and R1 are $\overline{H}$ and $\frac{k_S\overline{H}\,\overline{R}_S}{k_S\overline{H}+k_{hS}}$ respectively.

Observe that the maximal response (S22) does not depend on any of the rate constants or total amounts involving the sink-RR (R1). Therefore, the presence of a sink does not alter the maximal response, but might alter the shape of the signal-response curve. More generally, the inverse of the signal-response curve without the sink is:

$$f(x_6) = \frac{(k_M(\overline{R}_M - x_6) + k_{rM}x_6)k_{hM}x_6}{k_M\overline{H}(\overline{R}_M - x_6) - k_{hM}x_6} = \frac{k_{hM}x_6p_2(x_6)}{q_2(x_6)}. \tag{S23}$$

This is obtained by showing as above that in this case

$$k_a = \frac{k_{hM}x_6}{x_1}. \tag{S24}$$

Alternatively, we can set $\overline{R}_S = 0$ and all rate constants of the sink system to zero in the expression of f in (S21). Note that the expression of x_1 in (S16) does not depend on the presence of sink.

Therefore, for any fixed x_6, (S24) and (S20) satisfy the inequality

$$\frac{k_{hM}x_6}{x_1} < \frac{k_{hM}x_6 + k_{hS}x_4}{x_1}$$

after substituting x_1, x_4 for their expressions in x_6 and for any choice of rate constants of the sink system and $\overline{R}_S$. We deduce that given a response x_6, k_a in (S20) is larger to k_a in (S24). In other words, presence of the sink causes the system to require more signal to achieve the same level of response. If we plot the signal-response curve for a system with a sink, then the graph is below the signal-response curve of the corresponding system without sink, while keeping the common rates the same.

1.5 First derivative at zero

The first derivative of the signal-response curve at zero is computed as $1/f'(0)$ and takes the value:

$$\frac{k_M(k_{rS}\overline{H} + k_{hS})\overline{H}\overline{R}_M}{k_S k_{hS}(k_{rM}\overline{H} + k_{hM})\overline{R}_S + k_M k_{hM}(k_{rS}\overline{H} + k_{hS})\overline{R}_M}. \tag{S25}$$

The signal-response curve of the corresponding system without sink has first derivative at zero:

$$\frac{\overline{H}}{k_{hM}}.$$

Observe that (S25) can be rewritten as

$$\frac{k_M(k_{rS}\overline{H} + k_{hS})\overline{R}_M}{\frac{k_S k_{hS}(k_{rM}\overline{H}+k_{hM})\overline{R}_S}{k_{hM}} + k_M(k_{rS}\overline{H} + k_{hS})\overline{R}_M} \cdot \frac{\overline{H}}{k_{hM}},$$

and it becomes apparent that the first derivative at zero of the signal-response curve when the system has a sink is always smaller than the corresponding first derivative at zero of the signal-response curve of the system without a sign. Therefore, presence of the sink reduces the slope of the curve around zero.

1.6 Sigmoidal vs. hyperbolic

A hyperbolic curve has negative second derivative in all its domain, while a sigmoidal curve displays a change of sign of the second derivative. In particular, the second derivative at zero must be positive for a sigmoidal curve and negative for a hyperbolic curve.

The second derivative at zero of the signal-response curve can be computed from its inverse, f. In particular, the sign of the second derivative of the signal-response curve at zero is minus the sign of $f''(0)$ and is given by the sign of:

$$S = -k_S k_{hS} \overline{R}_S(\omega_1 \overline{H}(k_{rS} - k_S) + \omega_2(k_M \overline{H} + k_{hM})) - k_M k_{hM} \omega_2^2 \overline{R}_M \tag{S26}$$

where

$$\omega_1 = k_{rM}\overline{H} + k_{hM} \qquad \omega_2 = k_{rS}\overline{H} + k_{hS}.$$

For S to be positive, we necessarily need that

$$k_S k_{hS} \overline{R}_S \neq 0 \qquad \text{and} \qquad k_{rS} < k_S$$

as stated in the main text. These conditions are necessary for sigmoidality. In particular, presence of the sink is necessary. A simple system containing only one HK and one response regulator can only display hyperbolic signal-response curves.

The above conditions are necessary for sigmoidality. Sufficient conditions can also be given. For example, if

$$\frac{k_{rS}}{k_S} < \frac{k_{rM}}{k_M + k_{rM}}$$

then there exist total amounts $\overline{H}, \overline{R}_S$ large enough and $\overline{R}_M$ small enough such that the signal-response curve is sigmoidal. This follows from a general fact on polynomials. Consider a polynomial $p(x) = a_n x^n + \cdots + a_1 x + a_0$. If $a_n > 0$, then for x_0 large enough, $p(x_0) > 0$. Similarly, if $a_n < 0$, then for x_0 large enough, $p(x_0) < 0$. We consider the term S in (S26) as a polynomial in $\overline{H}$.

This polynomial has degree 2 and the coefficient of the term of degree 2 is

$$-k_S k_{hS}(k_{rS}k_M - k_{rM}k_S + k_{rM}k_{rS})\overline{R}_S - k_{rS}^2 k_{hM}k_M\overline{R}_M.$$

If $k_{rS}k_M - k_{rM}k_S + k_{rM}k_{rS} < 0$ or, equivalently, $\frac{k_{rS}}{k_S} < \frac{k_{rM}}{k_M+k_{rM}}$, then for $\overline{R}_S$ large and $\overline{R}_M$ small the coefficient of the polynomial is positive and we can use the general fact on polynomials to conclude that for $\overline{H}$ large enough, S is positive and hence sigmoidality occurs.

2 Mathematical analysis of the model with intermediates

We consider here the case where the model includes complex formation in the phosphotransfer reactions.

2.1 Reactions

The model has now 8 species: one histidine kinase HK, two response regulators R1, R2, their corresponding phosphorylated forms, HKp, R1p, R2p, and two intermediates Y_S, Y_M. The model consists of the following reactions:

- Auto-phosphorylation of HK:

$$\mathrm{HK} \xrightarrow{k_a} \mathrm{HKp}$$

- Reversible phosphotransfer between HK and R1 through the formation of a complex:

$$\mathrm{HKp}+\mathrm{R1} \xrightarrow{k_{aS}} Y_S \xrightarrow{k_{bS}} \mathrm{HK}+\mathrm{R1p} \qquad \mathrm{HK}+\mathrm{R1p} \xrightarrow{k_{brS}} Y_S \xrightarrow{k_{arS}} \mathrm{HKp}+\mathrm{R1}$$

- Reversible phosphotransfer between HK and R2:

$$\mathrm{HKp}+\mathrm{R2} \xrightarrow{k_{aM}} Y_M \xrightarrow{k_{bM}} \mathrm{HK}+\mathrm{R2p} \qquad \mathrm{HK}+\mathrm{R2p} \xrightarrow{k_{brM}} Y_M \xrightarrow{k_{arM}} \mathrm{HKp}+\mathrm{R2}$$

- Auto-dephosphorylation of the response regulators:

$$\text{R1p} \xrightarrow{k_{hS}} \text{R1} \qquad \text{R2p} \xrightarrow{k_{hM}} \text{R2}$$

2.2 System of ODEs

To simplify the notation, we write the concentrations of each of the species as:

$$x_1 = [\text{HK}] \quad x_2 = [\text{HKp}] \quad x_3 = [\text{R1}] \quad x_4 = [\text{R1p}] \quad x_5 = [\text{R2}] \quad x_6 = [\text{R2p}] \quad x_7 = [Y_S] \quad x_8 = [Y_M].$$

We let $x = (x_1, \ldots, x_8)$ be the vector of concentrations. We proceed as above to construct a system of ODEs for this model:

$$\dot{x}_1 = -k_{brM}x_1x_6 - k_{brS}x_1x_4 - k_a x_1 + k_{bM}x_8 + k_{bS}x_7 \tag{S27}$$

$$\dot{x}_2 = -k_{aM}x_2x_5 - k_{aS}x_2x_3 + k_a x_1 + k_{arM}x_8 + k_{arS}x_7 \tag{S28}$$

$$\dot{x}_3 = -k_{aS}x_2x_3 + k_{arS}x_7 + k_{hS}x_4 \tag{S29}$$

$$\dot{x}_4 = -k_{brS}x_1x_4 + k_{bS}x_7 - k_{hS}x_4 \tag{S30}$$

$$\dot{x}_5 = -k_{aM}x_2x_5 + k_{arM}x_8 + k_{hM}x_6 \tag{S31}$$

$$\dot{x}_6 = -k_{brM}x_1x_6 + k_{bM}x_8 - k_{hM}x_6 \tag{S32}$$

$$\dot{x}_7 = k_{aS}x_2x_3 + k_{brS}x_1x_4 - k_{arS}x_7 - k_{bS}x_7 \tag{S33}$$

$$\dot{x}_8 = k_{aM}x_2x_5 + k_{brM}x_1x_6 - k_{arM}x_8 - k_{bM}x_8. \tag{S34}$$

At steady state, $\dot{x}_i = 0$ for all $i = 1, \ldots, 8$. The system has also three conservation laws:

$$\dot{x}_1 + \dot{x}_2 + \dot{x}_7 + \dot{x}_8 = 0, \qquad \dot{x}_3 + \dot{x}_4 + \dot{x}_7 = 0, \qquad \dot{x}_5 + \dot{x}_6 + \dot{x}_8 = 0,$$

which provide three equations at steady state:

$$\overline{H} = x_1 + x_2 + x_7 + x_8, \qquad \overline{R}_S = x_3 + x_4 + x_7, \qquad \overline{R}_M = x_5 + x_6 + x_8 \tag{S35}$$

for positive total amounts $\overline{H}, \overline{R}_S$ and $\overline{R}_M$. Three of the steady state equations are redundant and are substituted by the conservation law equations. We keep the steady state equations corresponding to (S28), (S30), (S32), (S33) and (S34). Therefore, the steady states constrained to the conservation laws are given as the solutions to the equations:

$$\begin{aligned}
0 &= -k_{aM}x_2x_5 - k_{aS}x_2x_3 + k_a x_1 + k_{arM}x_8 + k_{arS}x_7 \\
0 &= -k_{brS}x_1x_4 + k_{bS}x_7 - k_{hS}x_4 \\
0 &= -k_{brM}x_1x_6 + k_{bM}x_8 - k_{hM}x_6 \\
0 &= k_{aS}x_2x_3 + k_{brS}x_1x_4 - k_{arS}x_7 - k_{bS}x_7 \\
0 &= k_{aM}x_2x_5 + k_{brM}x_1x_6 - k_{arM}x_8 - k_{bM}x_8 \\
0 &= x_1 + x_2 + x_7 + x_8 - \overline{H} \\
0 &= x_3 + x_4 + x_7 - \overline{R}_S \\
0 &= x_5 + x_6 + x_8 - \overline{R}_M.
\end{aligned} \tag{S36}$$

2.3 Sigmoidal vs. hyperbolic

We compute directly the second derivative of the signal-response curve at zero. First of all, we observe that when the signal is zero ($k_a = 0$), the steady state is:

$$x_1 = \overline{H}, \quad x_3 = \overline{R}_S, \quad x_5 = \overline{R}_M, \quad x_2 = x_4 = x_6 = x_7 = x_8 = 0.$$

We differentiate each equation in (S36) with respect to k_a. To simplify the notation, we write

$$p_i = \frac{dx_i(k_a)}{dk_a}$$

and evaluate the concentrations x_i at $k_a = 0$. We obtain a new system of equations:

$$\begin{aligned}
0 &= -k_{aS}\overline{R}_S\, p_2 - \overline{R}_M k_{aM}\, p_2 + k_{arM}\, p_8 + k_{arS}\, p_7 + \overline{H} \\
0 &= -k_{brS}\overline{H}\, p_4 + k_{bS}\, p_7 - k_{hS}\, p_4 \\
0 &= -k_{brM}\overline{H}\, p_6 + k_{bM}\, p_8 - k_{hM}\, p_6 \\
0 &= k_{brS}\overline{H}\, p_4 + k_{aS}\overline{R}_S\, p_2 - k_{arS}\, p_7 - k_{bS}\, p_7 \\
0 &= k_{brM}\overline{H}\, p_6 + k_{aM}\overline{R}_M\, p_2 - k_{arM}\, p_8 - k_{bM}\, p_8 \\
0 &= p_1 + p_2 + p_7 + p_8 \\
0 &= p_3 + p_4 + p_7 \\
0 &= p_5 + p_6 + p_8.
\end{aligned} \tag{S37}$$

The solutions in p_i are the derivatives at zero of each of the concentrations at steady state as a function of k_a. For instance, the **first derivative** of x_6, that is, of the signal-response curve, at zero is

$$\frac{k_{aM}k_{bM}\left(k_{arS}k_{brS}\overline{H} + (k_{arS}+k_{bS})k_{hS}\right)\overline{H}\,\overline{R}_M}{k_{aS}k_{bS}k_{hS}(k_{arM}k_{brM}\overline{H} + (k_{bS}+k_{bM})k_{hM})\overline{R}_S + k_{aM}k_{hM}k_{bM}(k_{arS}k_{brS}\overline{H} + (k_{arS}+k_{bS})k_{hS})\overline{R}_M}.$$

We introduce new rate constants that allow us to compare the systems with and without intermediates. That is, we define the inverse of the Michaelis-Menten constants of the intermediates:

$$k_{yS} = \frac{k_{aS}}{k_{arS}+k_{bS}}, \quad k_{yM} = \frac{k_{aM}}{k_{arM}+k_{bM}}, \quad k_{yrS} = \frac{k_{brS}}{k_{arS}+k_{bS}}, \quad k_{yrM} = \frac{k_{brM}}{k_{arM}+k_{bM}}$$

and let the new effective reaction rate constants for the phosphotransfer reactions be:

$$k_S = k_{bS}k_{yS}, \quad k_M = k_{bM}k_{yM}, \quad k_{rS} = k_{arS}k_{yrS}, \quad k_{rM} = k_{arM}k_{yrM}.$$

With the new constants, the first derivative at zero of the signal-response curve is

$$\frac{k_M(k_{rS}\overline{H} + k_{hS})\overline{R}_M\overline{H}}{k_S k_{hS}(k_{rM}\overline{H} + k_{hM})\overline{R}_S + k_M k_{hM}(k_{rS}\overline{H} + k_{hS})\overline{R}_M},$$

which is identical to the first derivative of the signal-response curve at zero without modeling of intermediates (given in (S25)). That is, with the identification of constants, the two derivatives are identical.

We compute now the **second derivative** of each variable at zero ($k_a = 0$) following the same procedure as above. We differentiate the equations (S37), evaluate the concentrations x_i at $k_a = 0$ and the first derivatives p_i at the solutions obtained in the first step. We obtain a new system of equations on the second derivatives of x_i with respect to k_a at zero and solve it. In particular, we obtain the second derivative of x_6 with respect to k_a at zero and conclude that the sign of the second derivative of the signal-response curve at zero agrees with the sign of the following term:

$$\begin{aligned} S_y = {} & \omega_1\omega_2 S - ((k_{yrM}k_M\overline{H} + k_{yM}\omega_1)\omega_2 - (k_{yrS}k_S\overline{H} + k_{yS}\omega_2)\omega_1)k_S k_{hS}\overline{H}\overline{R}_S\omega_1\omega_2 \\ & - ((k_{yrM}k_M\overline{H} + k_{yM}\omega_1)\omega_2\overline{R}_M + (k_{yrS}k_S\overline{H} + k_{yS}\omega_2)\omega_1\overline{R}_S)(k_M k_{hM}\omega_2^2\overline{R}_M \\ & + (k_{rS}\omega_1\overline{H} + k_{hM}\omega_2)k_S k_{hS}\overline{R}_S) \end{aligned} \tag{S38}$$

where

$$\omega_1 = k_{rM}\overline{H} + k_{hM} \qquad \omega_2 = k_{rS}\overline{H} + k_{hS}$$

and S corresponds to the system without intermediates, and was given in (S26) to be

$$S = -k_S k_{hS}(\omega_1\overline{H}(k_{rS} - k_S) + \omega_2(k_M\overline{H} + k_{hM}))\overline{R}_S - k_M k_{hM}\omega_2^2\overline{R}_M.$$

Observe that the terms not involving the Michaelis-Menten constants of the intermediates correspond precisely to the sigmoidality term S without intermediates.

For S_y to be positive one needs either S in (S26) to be positive or the blue term to be negative. The latter involves the Michaelis-Menten constants of the intermediates and, in particular, the condition $k_{rS} < k_S$ is not necessary anymore. What is still necessary is that

$$k_S k_{hS} \overline{R}_S \neq 0.$$

Let us look closely at the extra term that can be negative. Sigmoidality cannot occur if $k_{rS} > k_S$ and

$$(k_M k_{yrM} \overline{H} + k_{yM}\omega_1)\omega_2 - (k_S k_{yrS} \overline{H} + k_{yS}\omega_2)\omega_1 > 0.$$

The left-hand side of the inequality is written as a polynomial in $\overline{H}$ as

$$\begin{aligned} S_y = &\left(k_M k_{rS} k_{yrM} + k_{rM}(k_{rS}(k_{yM} - k_{yS}) - k_S k_{yrS})\right)\overline{H}^2 \\ &+ \left(k_M k_{hS} k_{yrM} + k_{hM}(k_{rS}(k_{yM} - k_{yS}) - k_S k_{yrS}) + k_{hS} k_{rM}(k_{yM} - k_{yS})\right)\overline{H} + k_{hM} k_{hS}(k_{yM} - k_{yS}). \end{aligned}$$

All the coefficients of this polynomial are positive if $k_{rS}(k_{yM} - k_{yS}) - k_S k_{yrS} > 0$, that is, if

$$\frac{k_S}{k_{rS}} < \frac{k_{yM} - k_{yS}}{k_{yrS}}. \tag{S39}$$

Hence sigmoidality cannot occur if $k_{rS} > k_S$ and (S39) holds. We conclude that necessary conditions for sigmoidality in the model with intermediates are that

(i) $k_S k_{hS} \overline{R}_S \neq 0$, and

(ii) $\frac{k_S}{k_{rS}} > \min\left(1, \frac{k_{yM} - k_{yS}}{k_{yrS}}\right)$.

In particular, if $k_{yM} < k_{yS}$, then we are left with the same necessary conditions as in the case

without intermediates. In terms of the original reaction rate constants, we note that

$$\frac{k_S k_{yrS} + k_{rS} k_{yS}}{k_{rS}} = \frac{k_{aS}}{k_{arS}} \qquad \text{and} \qquad \frac{k_S}{k_{rS}} = \frac{k_{aS} k_{bS}}{k_{arS} k_{brS}}. \tag{S40}$$

Then necessary conditions for sigmoidality are that

(i) $k_S k_{hS} \overline{R}_S \neq 0$, and

(ii) $\frac{k_{aS}}{k_{arS}} > \min\left(\frac{k_{brS}}{k_{bS}}, \frac{k_{aM}}{k_{arM}+k_{bM}}\right)$.

Finally, we look for conditions that guarantee the existence of sigmoidality for some total amounts. To this end we proceed as above and consider S_y as a polynomial in $\overline{H}$. The polynomial has degree five and the coefficient of highest degree is:

$$-\overline{R}_S k_S k_{hS} k_{rM} k_{rS} \left(k_M k_{rS} k_{yrM} - k_S k_{rM} k_{yrS} + k_{rM} k_{rS} (k_{yM} - k_{yS})\right). \tag{S41}$$

If the coefficient in is positive, then for $\overline{H}$ large enough, S_y is positive and as a consequence the signal-response curve is sigmoidal. The coefficient is positive if and only if

$$(k_M k_{yrM} + k_{rM} k_{yM}) k_{rS} < (k_S k_{yrS} + k_{rS} k_{yS}) k_{rM},$$

or equivalently (cf. (S40)), if and only if

$$\frac{k_{aM}}{k_{arM}} < \frac{k_{aS}}{k_{arS}}.$$

3 A mathematical model for the two-component system regulating yeast osmoregulation

To model the one HK-two RR system found in yeast osmoregulation, we considered its dynamics in isolation of other cellular components. In this system, the HK (SLN1) is a hybrid protein

composed of histidine kinase (HK) and receiver (RC) domains and there exists an additional histidine phosphotransfer (HPT) protein (YPD1). Together, these constitute a phosphorelay, where the YPD1 phosphotranfers to the two RRs, SSK1 and SKN7. We have modeled the hybrid HK as two separate proteins. The reactions in this system are:

$$\begin{aligned}
\mathrm{HK} &\xrightarrow{k_a} \mathrm{HKp} \\
\mathrm{RC} + \mathrm{HKp} &\underset{k_{rC}}{\overset{k_C}{\rightleftharpoons}} \mathrm{RCp} + \mathrm{HK} \\
\mathrm{HPT} + \mathrm{RCp} &\underset{k_{rT}}{\overset{k_T}{\rightleftharpoons}} \mathrm{HPTp} + \mathrm{RC} \\
\mathrm{RR1} + \mathrm{HPTp} &\underset{k_{rS}}{\overset{k_S}{\rightleftharpoons}} \mathrm{RR1p} + \mathrm{HPT} \\
\mathrm{RR2} + \mathrm{HPTp} &\underset{k_{rM}}{\overset{k_M}{\rightleftharpoons}} \mathrm{RR2p} + \mathrm{HPT} \\
\mathrm{RCp} &\xrightarrow{k_{hC}} \mathrm{RC} + \mathrm{Pi} \\
\mathrm{RR1p} &\xrightarrow{k_{hS}} \mathrm{RR1} + \mathrm{Pi} \\
\mathrm{RR2p} &\xrightarrow{k_{hM}} \mathrm{RR2} + \mathrm{Pi}
\end{aligned}$$

where HK, RC, HPT, RR1 and RR2 stand for SLN1, its receiver domain, YPD1, SSK1, and SKN7 respectively and the -p suffix represents phosphorylated forms of these proteins/domains. The above reaction scheme can be used to derive a system of ordinary differential equations, which describe the changes in concentrations over time. For the phosphorylated forms, the system is:

$$\frac{d[\text{HKp}]}{dt} = k_a[\text{HK}] + k_{rC}[\text{RCp}][\text{HK}] - k_C[\text{RC}][\text{HKp}]$$
$$\frac{d[\text{RCp}]}{dt} = k_C[\text{RC}][\text{HKp}] + k_{rT}[\text{HPTp}][\text{RC}] - k_{rC}[\text{RCp}][\text{HK}] - k_T[\text{HPT}][\text{RCp}] - k_{hC}[\text{RCp}]$$
$$\frac{d[\text{HPTp}]}{dt} = k_T[\text{HPT}][\text{RCp}] + k_{rS}[\text{RR1p}][\text{HPT}] + k_{rM}[\text{RR2p}][\text{HPT}] - k_{rT}[\text{HPTp}][\text{RC}]$$
$$- k_S[\text{RR1}][\text{HPTp}] - k_M[\text{RR2}][\text{HPTp}]$$
$$\frac{d[\text{RR1p}]}{dt} = k_S[\text{RR1}][\text{HPTp}] - k_{rS}[\text{RR1p}][\text{HPT}] - k_{hS}[\text{RR1p}]$$
$$\frac{d[\text{RR2p}]}{dt} = k_M[\text{RR2}][\text{HPTp}] - k_{rM}[\text{RR2p}][\text{HPT}] - k_{hM}[\text{RR2p}].$$

In addition, we have five conservation equations:

$$[\text{HK}]_{\text{tot}} = [\text{HK}] + [\text{HKp}]$$
$$[\text{RC}]_{\text{tot}} = [\text{RC}] + [\text{RCp}]$$
$$[\text{HPT}]_{\text{tot}} = [\text{HPT}] + [\text{HPTp}]$$
$$[\text{RR1}]_{\text{tot}} = [\text{RR1}] + [\text{RR1p}]$$
$$[\text{RR2}]_{\text{tot}} = [\text{RR2}] + [\text{RR2p}].$$

To analyze the behavior of the phosphate sink motif with increasing signal, we simulated the incoming signals from receptors as an increase in the auto-phosphorylation constant rate of the kinase (k_a). The model was parameterized with data from literature (see Supplementary Table 1). We numerically integrated the model to derive steady state signal-response relationships. The latter analysis gives the steady state level of phosphorylated RR2 at a given signal (k_a), where signal was taken as the constant rate of auto-phosphorylation of kinase and allows deriving a so-called signal-response curve. This curve is found by numerically integrating the system to steady state at a fixed signal level and then numerically "following" this steady state (i.e. steady state level of phosphorylated RR2), while changing the signal. This analysis is equivalent to allowing

the system to reach steady state under different signal values. Both signal-response analyses were performed using the software Oscill8 (http://oscill8.sourceforge.net/).

4. Analytical comparison of different models

To perform a formal check for the potential of bistability in the different models (discussed in the main text), we have utilized the Chemical Reaction Network Toolbox. The toolbox provides several analytical tests that can provide a definite answer on the possibility of existence of multiple stationary states in a given reaction network. We have applied these tests to the basic model and model with alternative reaction schemes; we had devised using the Chemical Reaction Network Toolbox v2.2 (http://www.crnt.osu.edu/CRNTWin). The model files used with this tool and describing the chemical reaction systems, as well as the analytical results from the tool are provided below.

Results of the analytical analysis of basic model

These contain the reaction system considered and the report produced with the Chemical Reaction Network Toolbox. In these reaction systems A, Y1 and Y2 stand for CheA, CheY1 and CheY2 respectively. P refers to phosphorylated form.

BASIC REPORT: NoName1

==========

Reaction network:

A -> AP

Y1P -> Y1

Y2P -> Y2

AP + Y1 <-> Y1P + A

AP + Y2 <-> Y2P + A

Remark: None.

———

Graphical Properties

==========

Number of complexes = 10

Number of linkage classes = 5:

Linkage class no. 1: {A, AP}

Linkage class no. 2: {Y1P, Y1}

Linkage class no. 3: {Y2P, Y2}

Linkage class no. 4: {AP + Y1, Y1P + A}

Linkage class no. 5: {AP + Y2, Y2P + A}

Number of TERMINAL strong linkage classes = 5:

Strong linkage class no. 1: {AP + Y1, Y1P + A}

Strong linkage class no. 2: {AP + Y2, Y2P + A}

Strong linkage class no. 3: {AP}

Strong linkage class no. 4: {Y1}

Strong linkage class no. 5: {Y2}

Number of NON-TERMINAL strong linkage classes = 3:

Strong linkage class no. 6: {A}

Strong linkage class no. 7: {Y1P}

Strong linkage class no. 8: {Y2P}

The network is neither reversible nor weakly reversible.

Rank Information

==========

Rank of entire network = 3

Deficiency Information

==========

Deficiency of entire network = 2

Deficiency of linkage class no. 1 = 0

Deficiency of linkage class no. 2 = 0

Deficiency of linkage class no. 3 = 0

Deficiency of linkage class no. 4 = 0

Deficiency of linkage class no. 5 = 0

Analysis

========

This is a deficiency two network. It is an excellent candidate for application of HIGHER DEFICIENCY THEORY (tailored mostly to networks with deficiencies greater than one).

Whether results will be obtained, will depend on whether or not the reaction network has certain additional structural attributes that help reduce the problem to a study of systems of linear inequalities.

If a network is "good", higher deficiency theory will determine, either affirmatively or negatively, whether there are positive rate constant values such that the corresponding mass action differential equations admit multiple (positive) steady states. If the answer is affirmative, higher deficiency theory will generate a sample set of rate constants and a pair of distinct steady states that are consistent with those rate constants.

If a network is "bad", some additional nonlinear analysis might be required, and the program might not be able to ascertain the network's capacity for multiple positive steady states. If definite conclusions can be reached they they will be reported. Otherwise the program will tell you that it cannot reach a conclusion.

Higher deficiency theory will also determine, either affirmatively or negatively, whether there can exist a set of rate constants such that the corresponding mass action differential equations admit a positive steady state having a zero eigenvalue (corresponding to an eigenvector in the stoichiometric subspace). When the answer is affirmative, the theory will produce such a set of rate constants, a positive steady state, and an

eigenvector (in the stoichiometric subspace) corresponding to an eigenvalue of zero. Results of this kind are contained after running the Zero Eigenvalue Report.

==========

HIGHER DEFICIENCY REPORT: NoName1

==========

Analysis

========

Taken with mass action kinetics, the network CANNOT admit multiple positive steady states or a degenerate positive steady state NO MATTER WHAT (POSITIVE) VALUES THE RATE CONSTANTS MIGHT HAVE.

Results of the analytical analysis of model with alternative reaction schemes

These contain the reaction system considered and the report produced with the Chemical Reaction Network Toolbox. In these reaction systems A, Y1 and Y2 stand for CheA, CheY1 and CheY2 respectively. P refers to phosphorylated form. The resulting system contains four complexes between the phosphorylated/unphosphorylated CheA and the phosphorylated/unphosphorylated CheY1 and CheY2.

BASIC REPORT: NoName1

==========

Reaction network:

A -> AP

A + Y1 <-> AY1

AY1 <-> APY1

APY1 <-> A + Y1P

APY1 <-> AP + Y1

AP + Y2 <-> APY2

APY2 <-> A + Y2P

APY2 <-> AY2

AY2 <-> A + Y2

Y1P -> Y1

Y2P -> Y2

Remark: None.

———

Graphical Properties

==========

Number of complexes = 16

Number of linkage classes = 5:

Linkage class no. 1: {A, AP}

Linkage class no. 2: {A + Y1, AY1, APY1, A + Y1P, AP + Y1}

Linkage class no. 3: {AP + Y2, APY2, A + Y2P, AY2, A + Y2}

Linkage class no. 4: {Y1P, Y1}

Linkage class no. 5: {Y2P, Y2}

Number of TERMINAL strong linkage classes = 5:

Strong linkage class no. 1: {A + Y1, AY1, APY1, A + Y1P, AP + Y1}

Strong linkage class no. 2: {AP + Y2, APY2, A + Y2P, AY2, A + Y2}

Strong linkage class no. 3: {AP}

Strong linkage class no. 4: {Y1}

Strong linkage class no. 5: {Y2}

Number of NON-TERMINAL strong linkage classes = 3:

Strong linkage class no. 6: {A}

Strong linkage class no. 7: {Y1P}

Strong linkage class no. 8: {Y2P}

The network is neither reversible nor weakly reversible.

Rank Information

==========

Rank of entire network = 7

Deficiency Information

==========

Deficiency of entire network = 4

Deficiency of linkage class no. 1 = 0

Deficiency of linkage class no. 2 = 0

Deficiency of linkage class no. 3 = 0

Deficiency of linkage class no. 4 = 0

Deficiency of linkage class no. 5 = 0

Analysis

This is a deficiency four network. It is a good candidate for application of HIGHER DEFICIENCY THEORY (tailored mostly to networks with deficiencies greater than one).

Whether results will be obtained, will depend on whether or not the reaction network has certain additional structural attributes that help reduce the problem to a study of systems of linear inequalities.

If a network is "good", higher deficiency theory will determine, either affirmatively or negatively, whether there are positive rate constant values such that the corresponding mass action differential equations admit multiple (positive) steady states. If the answer is affirmative, higher deficiency theory will generate a sample set of rate constants and a pair of distinct steady states that are consistent with those rate constants.

If a network is "bad", some additional nonlinear analysis might be required, and the program might not be able to ascertain the network's capacity for multiple positive steady states. If definite conclusions can be reached they they will be reported. Otherwise the program will tell you that it cannot reach a conclusion.

Higher deficiency theory will also determine, either affirmatively or negatively, whether there can exist a set of rate constants such that the corresponding mass action differential equations admit a positive steady state having a zero eigenvalue (corresponding to an eigenvector in the stoichiometric subspace). When the answer is affirmative, the theory will produce such a set of rate constants, a positive steady state, and an eigenvector (in the stoichiometric subspace) corresponding to an eigenvalue of zero. Results of this kind are contained after running the Zero Eigenvalue Report.

```
================================
    HIGHER DEFICIENCY REPORT: NoName1
    ================================

        Analysis
        ========
```

Taken with mass action kinetics, the network DOES have the capacity for multiple steady states. That is, there are rate constants that give rise to two or more positive (stoichiometrically compatible) steady states -- you'll see an example below -- and also rate constants for which there is a steady state having an eigenvector (in the stoichiometric subspace) corresponding to an eigenvalue of zero. (To construct rate constants that give a degenerate steady state, use the Zero Eigenvalue Report.)

A mass action system example is also given below:

Example No. 1: Multiple Steady States

The following mass action system gives rise to multiple steady states:

A ---53.264249-> AP

A + Y1 ---525.88864-> AY1

AY1 ---76.689969-> A + Y1

AY1 ---7356.1463-> APY1

APY1 ---6224.388--> AY1

APY1 ---7843.4039-> A + Y1P

APY1 ---1043.064--> AP + Y1

A + Y1P ---20534.231-> APY1

AP + Y1 ---132.87987-> APY1

AP + Y2 ---116491.28-> APY2

APY2 ---4951.8227-> AP + Y2

APY2 ---8571.821--> A + Y2P

APY2 ---140.24927-> AY2

A + Y2P ---5955.1979-> APY2

AY2 -------1-----> APY2

AY2 ---32.105055-> A + Y2

A + Y2 ---11290.218-> AY2

Y1P ---424.3858--> Y1

Y2P ---123.07748-> Y2

The steady states shown below are both consistent with the mass action system indicated.

Steady State No. 1	Species	Steady State No. 2
8.8663 E-2	A	0.17854626
1.2542 E-2	AP	0.10242547
2.8335553	Y1	2.5639068
5.2309 E-2	APY1	0.14219251
0.18275376	Y1P	0.27263657
0.15146312	Y2	6.1580 E-2
3.3672 E-2	APY2	0.12355565
0.44331725	Y2P	0.89273131
6.1580 E-2	AY1	0.15146312
4.7225921	AY2	4.2731781

Eigenvalues for Steady State No. 1

-5878.2169
-4452.4167
-2510.4006
-1299.6516
-392.65837
-145.57513
1.7915605

Steady State No. 1 is unstable.

Eigenvalues for Steady State No. 2

-6346.8262

-4927.26

-2866.2672

-1424.5087

-4.6933204

-355.14896

-189.03784

Steady State No. 2 is asymptotically stable.

5. Figures and table

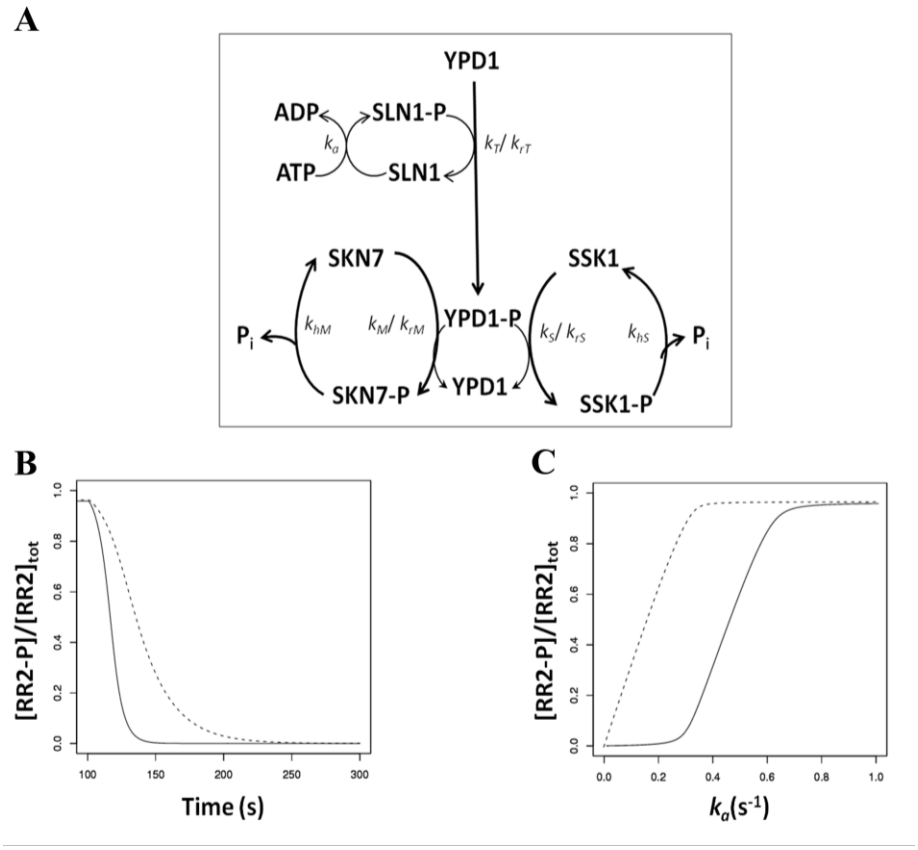

Figure S3.1. SSK1 is a phosphate sink for SLN7 in the yeast osmoregulation pathway (**A**) A cartoon diagram of the SLN1-YPD1-SSK1-SKN7 system. The diagram is arranged to highlight the role of the SSK1 as a phosphate sink for SKN7. Rate constants are shown on the relevant reactions. In the case of reversible reactions, two rate constants are given as $k_{forward}$ and $k_{reverse}$. (**B)** Role of the sink RR (SSK1) in dephosphorylation of SKN7-P (RR2-P**).** The x- and y-axis show the time and the corresponding phosphorylated RR2 (SKN7-P) level at steady-state respectively. A value of k_a was selected that resulted in ~90% of the total RR2 being phosphorylated at steady state. At t=0, k_a was reduced to zero and the progress of the reaction to the new steady state was simulated. Solid line represents the presence of the sink (i.e. SSK1), while dashed line shows the absence of the sink. **(C)** Signal-response curve in the

presence (solid line) and absence (dashed line) of the sink RR (SSK1). The x- and y-axis show the signal (k_a) level and the corresponding steady state level of phosphorylated SKN7 (RR2-P) respectively.

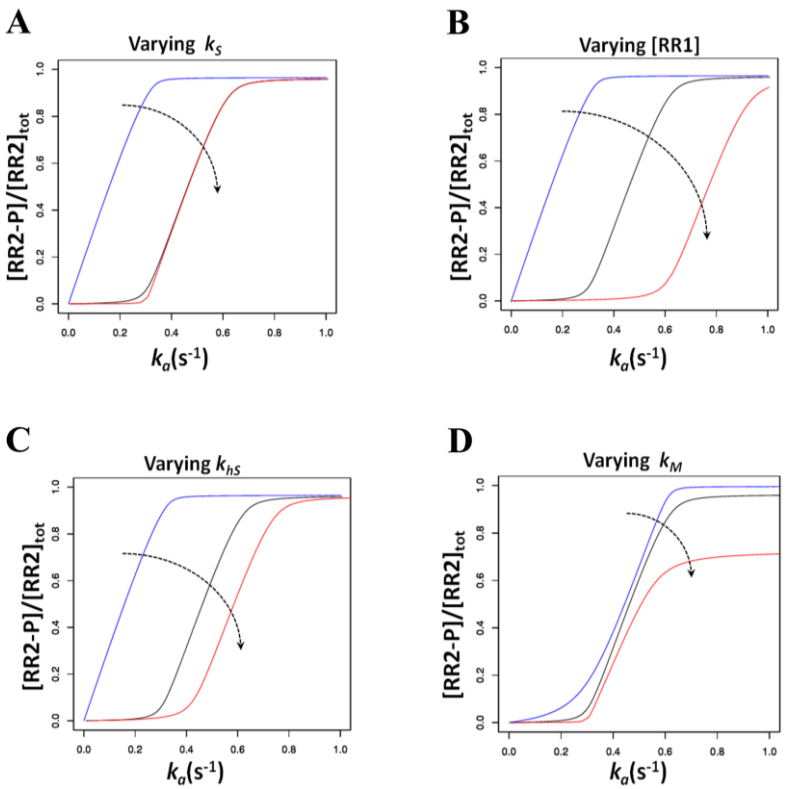

Figure S3.2. Effect of varying the key parameters in the yeast osmoregulation system on the shape of the signal-response curve. The x- and y-axis show the signal (k_a) level and the corresponding level of phosphorylated output RR (SKN7-P) at steady state respectively. Each panel shows a signal-response curve for different parameter values. The results of the basic model are shown in black. The arrow on each panel indicates increasing values of the changed parameter. (**A**) The forward phosphotransfer rate (k_S) for the sink RR was varied from basic model value [66.67(μMs)$^{-1}$] to 660, and 0. (**B**) Concentration of the sink RR was

set to 0μM, 1.5μM (basic model) and 3μM. (**C**) The rate of auto-dephosphorylation of sink RR-P (k_{hS}) was set to $0s^{-1}$, $0.5s^{-1}$ (basic model) and $1s^{-1}$. (**D**) The forward phosphotransfer rate (k_M) for the main RR, was set to $1(\mu Ms)^{-1}$(basic model), 0.5, and 10.

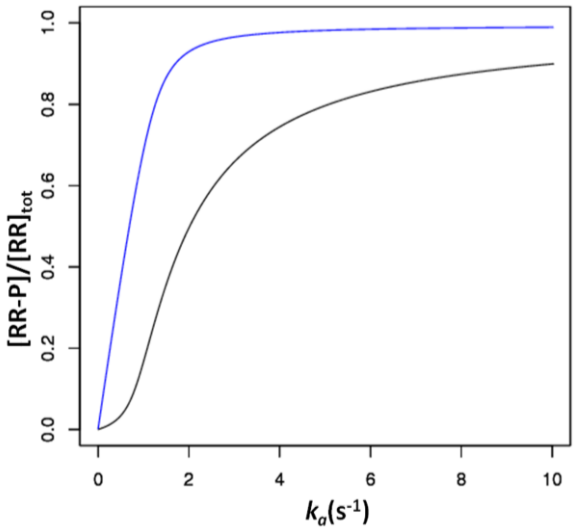

Figure S3.3. Signal-response relationship for the sink RR and the output RR in the *S. meliloti* system. The x- and y-axis show the signal (k_a) level and the corresponding steady state level of either phosphorylated sink (blue line) or main RR (black line).

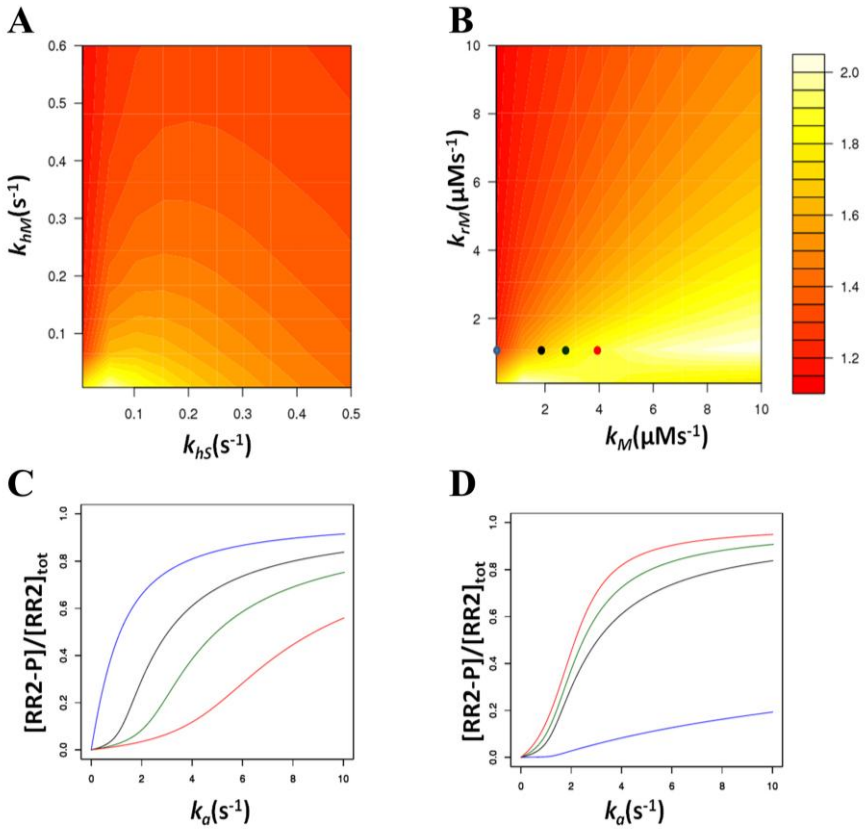

Figure S3.4. The effect of parameter changes on the signal-response curve of the *S. meliloti* system. The level of "sigmoidality" of the signal-response curve, Hill coefficient, is shown on each panel as a heat map. (**A**) Effect of varying the auto-dephosphorylation rate of the output RR (k_{hM}; y-axis) and sink RR (k_{hS}; x-axis). (**B**) Effect of varying the forward and reverse phosphotransfer rates to the output RR (CheY2; x-axis; k_M and y-axis; k_{rM}). (**C** and **D**) Signal-response curves for models corresponding to parameter values indicated as colored circles on the heat maps in which the black circle represents the basic model.

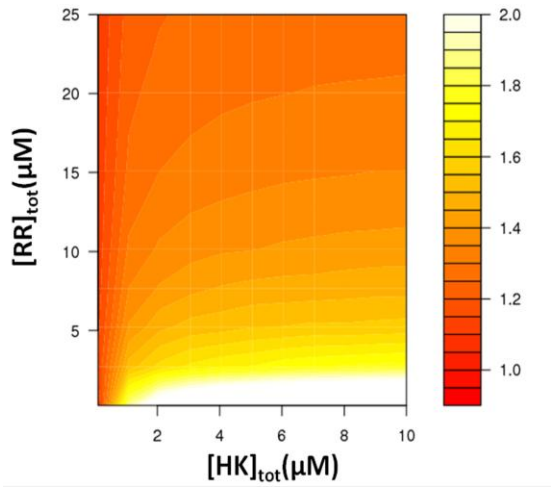

Figure S3.5. Effect of CheA, CheY1, and CheY2 total concentrations on the shape of the signal-response curve for the *S. meliloti* system. The level of "sigmoidality" of the signal-response curve, Hill coefficient, is shown on each panel as a heat map. The x-axis shows the total concentration of CheA, while the y-axis shows the total concentration of CheY1 and CheY2 (where $[CheY1]_{tot} = [CheY2]_{tot}$).

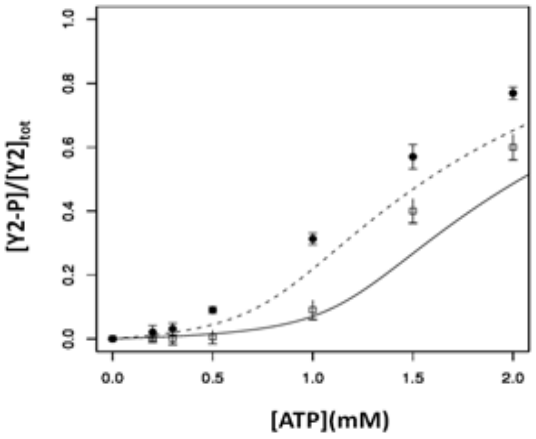

Figure S3.6. Effect of CheS on the signal-response curve. The x- and y-axis show the ATP level and the corresponding steady state phosphorylated CheY2 levels, respectively. The

phosphorylated CheY2 levels predicted by the model are shown with a dashed line (absence of CheS) and with a solid line (presence of CheS; where autodephosphorylation rate of CheY1, k_{hs} was increased to 0.08 s^{-1}), while the experimentally measured values are shown in circles and squares on respective graph. See also Figure 4 for an alternative approach to modeling the presence of CheS. Error bars show the standard error of the mean obtained from three independent experiments.

Table S3.1: The parameters used for the model of the yeast phosphate sink

Parameter	Description	Value	Unit	Reference
k_a	k_{cat} for phosphorylation of SLN1	Varied	s^{-1}	
k_C	SLN1-P to Receiver domain of SLN-1 Phosphotransfer	160	$(\mu Ms)^{-1}$	[38]
k_{rC}	SLN1-P to Receiver domain of SLN-1 Reverse Phosphotransfer	0	$(\mu Ms)^{-1}$	[38]
k_T	Phosphorylated Receiver domain of SLN-1 to YPD1 Phosphotransfer	20.7	$(\mu Ms)^{-1}$	[38]
k_{rT}	Phosphorylated Receiver domain of SLN-1 to YPD1 Reverse Phosphotransfer	29.5	$(\mu Ms)^{-1}$	[38]
k_S	YPD1-P to SSK1 (sink RR) Phosphotransfer	66.67	$(\mu Ms)^{-1}$	[38]
k_{rS}	YPD1-P to SSK1 Reverse phosphotransfer	0	$(\mu Ms)^{-1}$	[38]
k_M	YPD1-P to SKN7 (main RR) Phosphotransfer	1	$(\mu Ms)^{-1}$	[38]
k_{rM}	YPD1-P to SKN7 Reverse phosphotransfer	0.08	$(\mu Ms)^{-1}$	[38]
k_{hC}	Autodephosphorylation of Receiver domain of SLN1	0.05	s^{-1}	[38]
k_{hS}	Autodephosphorylation of SSK1 (sink RR)	0.05	s^{-1}	[38]

$[SLN]_{tot}$	Total conc. of SLN1	0.25	µM	[38]
$[Rec\ domain\ SLN]_{tot}$	Total conc. of Rec domain of SLN1	0.25	µM	[38]
$[YPD]_{tot}$	Total conc. of YPD	1.5	µM	[38]
$[SSK1]_{tot}$	Total conc. of SSK1	1.5	µM	[38]
$[SKN7]_{tot}$	Total conc. of SKN7	1.5	µM	[38]

OPEN ACCESS Freely available online

PLOS | COMPUTATIONAL BIOLOGY

Split Histidine Kinases Enable Ultrasensitivity and Bistability in Two-Component Signaling Networks

Munia Amin[1,2], Steven L. Porter[1]*, Orkun S. Soyer[2]*

1 Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom, **2** Systems Biology Program, College of Engineering, Computing and Mathematics, University of Exeter, Exeter, United Kingdom

Abstract

Bacteria sense and respond to their environment through signaling cascades generally referred to as two-component signaling networks. These networks comprise histidine kinases and their cognate response regulators. Histidine kinases have a number of biochemical activities: ATP binding, autophosphorylation, the ability to act as a phosphodonor for their response regulators, and in many cases the ability to catalyze the hydrolytic dephosphorylation of their response regulator. Here, we explore the functional role of "split kinases" where the ATP binding and phosphotransfer activities of a conventional histidine kinase are split onto two distinct proteins that form a complex. We find that this unusual configuration can enable ultrasensitivity and bistability in the signal-response relationship of the resulting system. These dynamics are displayed under a wide parameter range but only when specific biochemical requirements are met. We experimentally show that one of these requirements, namely segregation of the phosphatase activity predominantly onto the free form of one of the proteins making up the split kinase, is met in Rhodobacter sphaeroides. These findings indicate split kinases as a bacterial alternative for enabling ultrasensitivity and bistability in signaling networks. Genomic analyses reveal that up 1.7% of all identified histidine kinases have the potential to be split and bifunctional.

Citation: Amin M, Porter SL, Soyer OS (2013) Split Histidine Kinases Enable Ultrasensitivity and Bistability in Two-Component Signaling Networks. PLoS Comput Biol 9(3): e1002949. doi:10.1371/journal.pcbi.1002949

Editor: Jason M. Haugh, North Carolina State University, United States of America

Received June 26, 2012; **Accepted** January 11, 2013; **Published** March 7, 2013

Funding: We acknowledge the support of Exeter Science Strategy. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: S.Porter@exeter.ac.uk (SLP); O.S.Soyer@exeter.ac.uk (OSS)

Introduction

Bacterial responses to many external stimuli are underpinned by two-component signaling networks (TCSNs). These are found in most bacterial species and are also present in Archaea, eukaryotic microbes, and plants [1,2]. TCSNs are built upon the core reactions involving a histidine kinase (HK) that autophosphorylates on a conserved histidine residue in response to a signal, and a cognate response regulator (RR) that is activated when the HK phosphorylates one of its conserved aspartate residues [3]. Evolutionary processes seem to have exploited the modular structure of these proteins to produce a distinct set of biochemical features and network structures that reoccur in diverse TCSNs; bifunctional HKs [4], sink RRs [5], phosphorelays [6] and split HKs [7]. In order to achieve a broad and predictive understanding of bacterial signaling, it is important to assess whether these features enable specific signaling dynamics and properties [8].

There has already been progress towards this goal. Firstly, bifunctional HKs, which display both phosphatase and kinase activity towards their cognate RR, enable robustness in system output with respect to fluctuations in the amount of these signaling proteins [4,9] and reduce cross-talk among different TCSNs [10,11]. Further, theoretical work indicates that bi-functional HKs can generate flexible signal-response relationships [12,13] and allow higher signal amplification compared to monofunctional HKs that lack phosphatase activity [10]. Secondly, sink RRs, which compete with another RR for phosphoryl groups from a single cognate HK, are suggested to allow faster response termination [5,14]. Finally, phosphorelays, which contain several proteins (or domains) acting as a relay between the HK and RR, are suggested to integrate several signals received on their different layers [15–17] and implement both ultrasensitive and linear responses [18,19]. Taken together, these studies suggest that specific biochemical and structural features in TCSNs could enable specific functional roles.

Of the different features of TCSNs, split kinases are predicted in several bacterial genomes [1,2] and are biochemically characterized in *Rhodobacter sphaeroides* [7,20]. In this organism, the split kinase system is composed of CheA3 and CheA4, which form a bipartite histidine kinase that phosphorylates the response regulator CheY6 [21] (Figure 1). CheA4 lacks the phosphorylatable P1 domain, whereas CheA3 lacks the dimerization (P3) and catalytic kinase (P4) domains. Neither CheA3 nor CheA4 can autophosphorylate when incubated separately with ATP; however, when a mixture of CheA3 and CheA4 is incubated with ATP, then CheA3 becomes phosphorylated, indicating that these proteins can act as a histidine kinase only by forming a complex [21]. Activated by incoming signals, the P4 domain of CheA4 binds ATP and phosphorylates the P1 domain of CheA3. Subsequently, CheA3-P acts as a phosphodonor for its cognate response regulator, CheY6 [21], which controls flagellar rotation [22]. In vivo, CheA3 and CheA4 co-localize to the cytoplasmic chemotaxis cluster [23] and are both essential for chemotaxis [7,24]. CheA3 and CheA4 bind to the cytoplasmic cluster via their P5 domains [25]. Whilst part of this cluster, CheA3 and CheA4 dynamically interact with one another. To allow phosphorylation of CheA3,

Author Summary

Two-component signaling systems mediate many of the physiological responses of bacteria. In their core, these systems consist of a histidine kinase (HK) and a response regulator (RR) that it can phosphotransfer to. Around this core interaction, evolution has led to several conserved biochemical and structural features. In order to achieve a broad and predictive understanding of bacterial signaling, it is important to assess whether these features enable specific signaling dynamics and properties. Our study provides a potential functional role for one such feature, the split histidine kinases, where autophosphorylation and phosphotransfer activities of a conventional HK are segregated onto distinct proteins capable of complex formation. We show that that this unusual configuration can enable ultrasensitivity and bistability in signal transduction under specific biochemical conditions. We experimentally show that one of these requirements, namely segregation of the phosphatase activity predominantly onto the free form of one of the proteins making up the split kinase, is met in proteins isolated from *Rhodobacter sphaeroides*. Genomic studies suggest 1.7% of all histidine kinases could function as bifunctional split kinases. This study provides a linkage between these proteins and response dynamics, thereby enabling experimentally testable hypotheses in these systems.

the P4 domain of CheA4 must transiently bind to the P1 domain of CheA3 (in the subsequent analysis we refer to this complex as CheA3:CheA4). Once phosphorylated, the P1 domain of CheA3 is released by CheA4, and CheA3-P can then donate its phosphoryl group to the corresponding response regulator CheY6 [21,26]. In addition to its phosphotransfer function, CheA3 is also a phosphatase towards CheY6-P [7]. *cheA3* mutants retaining phosphotransfer functions but lacking phosphatase activity do not support chemotaxis, similarly, *cheA3* mutants retaining phosphatase activity but lacking phosphotransfer activity also fail to support chemotaxis, indicating that chemotaxis requires both activities of CheA3 [7,21]. In addition, to being phosphorylated and dephosphorylated by the split kinase comprising CheA3 and CheA4 [21], CheY6 is also phosphorylated by CheA2 at the polar chemotaxis cluster [27].

Despite this wealth of information, the general role of split kinases in bacterial signaling is not clear. In essence split kinases are unusual bifunctional HKs, where the autophosphorylation and subsequent phosphotransfer and phosphatase activities are encoded on two separate proteins. Since the complex formed by these proteins is functionally equivalent to a bifunctional HK, it is not clear what the role of splitting biochemical activities in this way might be. Using the biochemical reactions of CheA3, CheA4, and CheY6 as a model system, we developed a mathematical model and analyzed the response dynamics mediated by this split kinase. Repeating this analysis with a bifunctional HK and a conventional HK-RR pair featuring a separate phosphatase, we found that in contrast to these configurations, split kinases enable ultrasensitivity and bistability in the signal-response relationship. We show that these dynamical features are maintained under a wide parameter range, provided certain biochemical assumptions are met. These requirements indicate that the source of ultrasensitivity and bistability in split kinases is the inverse coupling between their kinase and phosphatase activities; i.e. the kinase activity cannot be increased without reducing the phosphatase activity and vice versa. Through measurements of phosphatase activity, we show that this condition is met in the *R. sphaeroides* system *in vitro*. These findings suggest that bacteria might be utilizing split kinases as a means of implementing ultrasensitivity and bistability in cellular decision making.

Results

Construction of a mathematical model of a split kinase

Since our aim is to study the general response dynamics that split kinases can mediate, we use the CheA3, CheA4, and CheY6 triplet as a model system and study its dynamics in isolation through *in vitro* experiments, numerical simulation and analytical approaches. We developed a mathematical model of the system and parameterized it with *in vitro* and *in vivo* measured kinetic rates and protein concentrations respectively (see *Methods* and Table 1). We then analyzed the response dynamics of the resulting model and its variants both through numerical simulations and deriving analytical solutions of steady state behavior using approximations and the chemical network theory [28,29] (see *Methods* and *Text S1*). In the subsequent sections, we use the terms free CheA3 and free CheA3-P to indicate CheA3 species where the P1 domain is not interacting with the P4 domain of CheA4; *in vivo*, however, these species are expected to be always joined to the chemotaxis cluster by their P5 domains.

The input-output relationship for the split kinase shows ultrasensitivity and bistability

A primary property of interest for any signal transduction system is the signal-response relationship it implements [30]. To analyze the signal-response relationship in systems featuring a split kinase, we defined the system response as the steady state level of phosphorylated CheY6 (CheY6-P) at a given signal level, and derived this relationship for different parameters and biochemical assumptions (see *Methods*). This analysis revealed that when assuming free CheA3 as the sole phosphatase for CheY6-P, the system has a high potential for displaying ultrasensitivity and bistability (Figure 2 and Figures S1, S2, S3). Both of these dynamics result in a switch-like behavior; the response of the system is low until signal levels increase above a certain threshold, after which the response increases disproportionately to reach a high level (e.g. Figure 2A). In the case of bistability, the low and high response levels correspond to stable states of the system, separated by an unstable region, resulting in abrupt switching dynamics and hysteresis (i.e. the switching threshold is different depending on the past state of the system).

The *in vitro* and *in vivo* measured kinetic rates and protein concentrations from *R. sphaeroides* constitute "biologically meaningful" values that could be representative for two-component systems in general. To analyze the potential effects of these rates on the observed nonlinearity of the signal-response relationship, we have performed a sensitivity analysis by varying the base parameter values over a large range and quantifying the shape of the resulting signal-response curve (see *Methods*). This analysis shows that the level of ultrasensitivity in the signal-response relationship is most sensitive to the parameters controlling the complex formation between CheA3:CheA4 (k_1) and the dephosphorylation of phosphorylated CheY6 (k_9 and k_{11}) (Figure 2 and Figures S1, S2, S3). The association rate constant (k_1/k_2) we used in the basic model is approximately 500-fold higher than that measured *in vitro*, using purified *R. sphaeroides* proteins [21]. We still consider this high value "biologically relevant" as *in vivo* conditions can result in confining of split kinase components to small regions of the cell, resulting in much higher effective concentrations than are attainable under the *in vitro* conditions as used in [21]. For example, in *R. sphaeroides*, CheA3 and CheA4 localize to the cytoplasmic chemoreceptor cluster [23], which - using immunogold

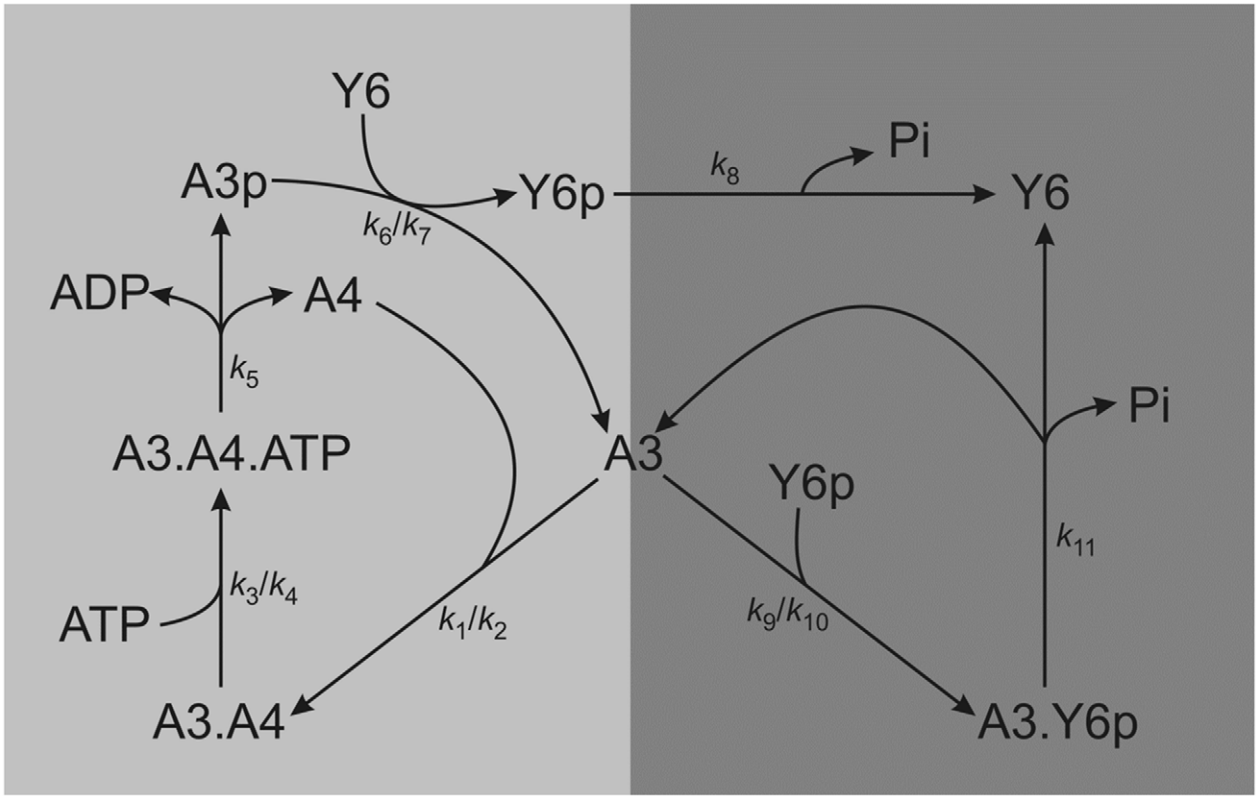

Figure 1. A cartoon diagram of the CheA3-CheA4-CheY6 split kinase system. The diagram is arranged so to highlight the role of free CheA3 acting as a branching point for the two arms that form competing cycles leading to phosphorylation and dephosphorylation of CheY6. Rate constants are shown on the relevant reactions. In the case of reversible reactions, two rate constants are given ($k_{forward}/k_{reverse}$).
doi:10.1371/journal.pcbi.1002949.g001

electron microscopy - is estimated to occupy less than 5% of the cross-sectional area of the cell [31]. Assuming a spherical shape for both the cell and this cluster, the volume of the latter could be estimated to be approximately 1% of the total cell volume. Thus, the effective concentrations of CheA3 and CheA4 in this cluster could be increased by as much as 100-fold, resulting in a significantly higher effective association rate constant than measured *in vitro* (up to 10,000 fold).

Table 1. Literature source and parameter values used in the analysis of the basic model.

Parameter	Description	Value	Unit	Ref
k_1	On rate for binding of CheA3 and CheA4	100	($\mu M\ s^{-1}$)	[21] see also *Results*
k_2	Off rate for binding of CheA3 and CheA4	10	s^{-1}	[21] see also *Results*
k_3	Forward rate for phosphorylation complex	1	$(\mu M\ s)^{-1}$	[21]
k_4	Reverse rate for phosphorylation complex	39	s^{-1}	[21]
k_5	K_{cat} for phosphorylation of CheA3 by CheA4	varied	s^{-1}	
k_6	CheA3-P to CheY6 Phosphotransfer	0.775	$(\mu M\ s)^{-1}$	[21]
k_7	CheA3-P to CheY6 Reverse phosphotransfer	0.00283	$(\mu M\ s)^{-1}$	[21]
k_8	Autodephosphorylation	0.169	s^{-1}	[7]
k_9	Association of phosphatase assisted dephosphorylation complex	5.6	$(\mu M\ s)^{-1}$	[48]
k_{10}	Dissociation of phosphatase assisted dephosphorylation complex	0.04	s^{-1}	[48]
k_{11}	K_{cat} for phosphatase assisted dephosphorylation	2.5	s^{-1}	See *Methods*
$[A3]_{tot}$	Total concentration of CheA3	90	μM	[7]
$[A4]_{tot}$	Total concentration of CheA4	40	μM	[34]
$[Y6]_{tot}$	Total concentration of CheY6	225	μM	[34]
[ATP]	Total concentration of ATP	1000	μM	

doi:10.1371/journal.pcbi.1002949.t001

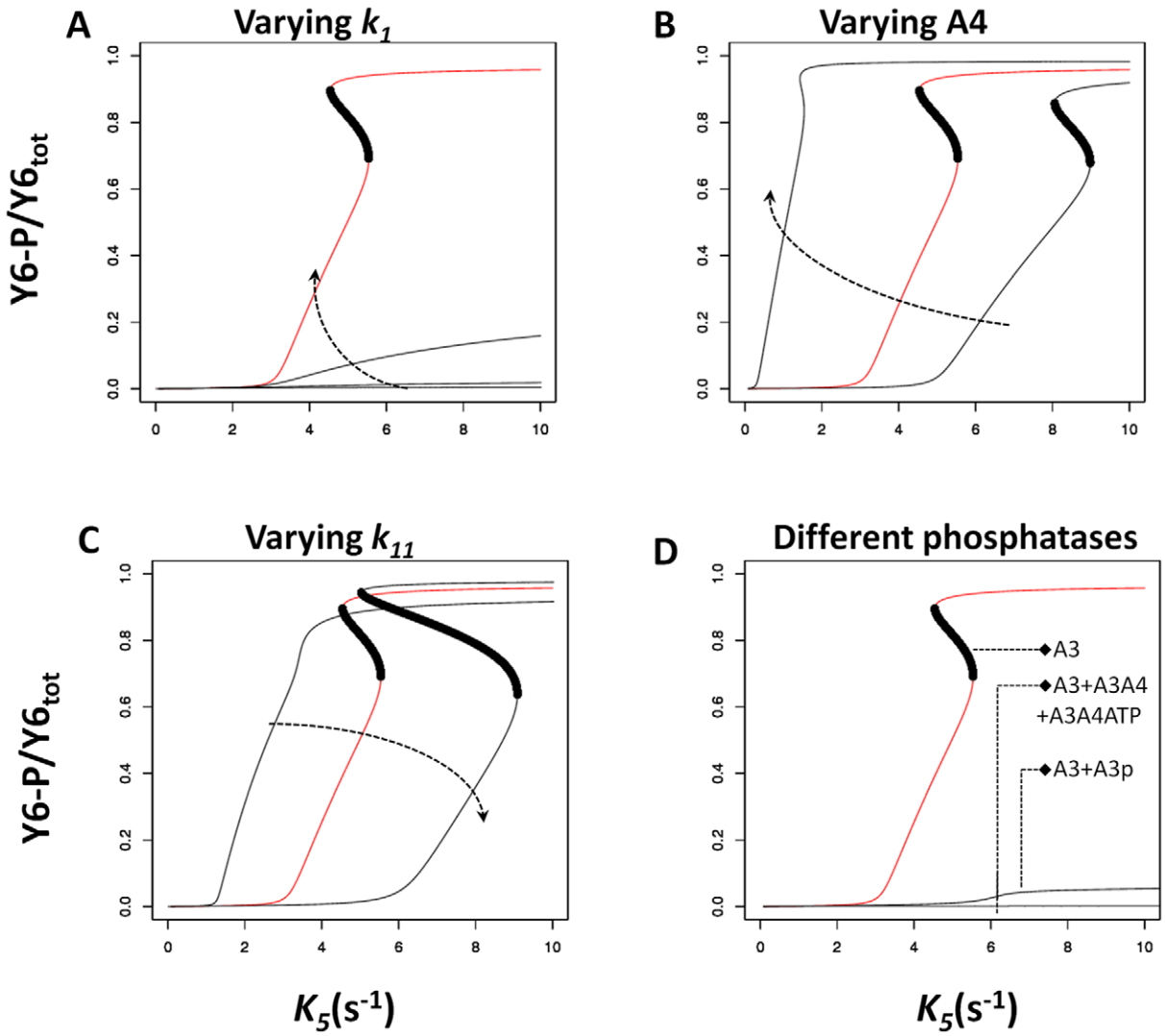

Figure 2. Effects of varying key parameters of the model and addition of different phosphatases. The x- and y-axis show the signal (k_5) level and the corresponding steady state CheY6-P level respectively. Each panel shows a signal-response analysis for varying model parameters (A–C) or the inclusion of additional phosphatases (D). The results of the basic model are shown in red. Where present, the dark region indicates the region of unstable steady states and hence the presence of bistability. Arrows on panels A, B and C indicate increasing value of the changed parameter. (**A**) The on rate (k_1) for CheA3:CheA4 complex formation was varied from basic model value [100(μMs)$^{-1}$] to 10, 1, and 0.208. (**B**) Concentration of CheA4 was varied from 30 μM, 40 μM (basic model) and 80 μM. (**C**) The rate of CheA3 mediated dephosphorylation of CheY6-P (k_{11}) was varied from 1 s^{-1}, 2.5 s^{-1} (basic model) and 5s^{-1}. (**D**) The basic model has free CheA3 as the sole phosphatase; the effect of having either CheA3-P or CheA3:CheA4 and CheA3:CheA4:ATP as additional phosphatases is shown. See also Figures S1, S2, S3, S4 for additional sensitivity analyses.
doi:10.1371/journal.pcbi.1002949.g002

Besides parameter values, several modeling choices could also alter the finding of bistability and ultrasensitivity arising in a split kinase system. In particular, the basic model presented above assumes that free CheA3 is the sole phosphatase in the system (besides the intrinsic autodephosphorylation activity of CheY6-P). Relaxing this assumption and considering increasing phosphatase activity by the CheA3:CheA4 and CheA3:CheA4:ATP complexes (see *Text S1, section 1*), significantly reduced ultrasensitivity in the system (Figure 2D and S4). In contrast, the presence of ultrasensitivity was much more robust to increasing phosphatase activity by CheA3p (Figure 2D, S4 and S5). Another mechanistic choice in the modeling of the split kinase system is the fate of the CheA3:CheA4 complex after phosphorylation of CheA3. In the basic model analyzed in Figure 2, this is modeled as phosphorylation resulting in the dissociation of the complex and release of CheA4 and CheA3-P. An alternative would be that the CheA3:CheA4 complex remains intact post phosphorylation, resulting in a CheA3-P:CheA4 complex (see *Text S1, section 2*). When we assume the presence of CheA3-P:CheA4 complex that can phosphotransfer to CheY6, bistability was lost, but not ultrasensitivity (Figure S6). Finally, we found that including an additional (monofunctional, non-split) kinase in the model, as seen for example in *R. sphaeroides* CheA2 (see *Text S1, section 3*), does not affect the ultrasensitivity but can result in the loss of bistability (Figure S7).

It is important to note that the basic model and all of these variants arising from specific modeling choices are "nested" in the sense that the basic model can be recovered through appropriate choice of parameters (e.g. setting dephosphorylation activity of CheA3p very low). In line with this observation, we find that the

basic model and all of the alternative structures discussed so far can be analytically shown to possess the "ability" to attain bistability (see *Methods*). More particularly, each of the chemical reaction systems arising from these models have the capacity for multiple steady states according to the higher deficiency theorem [32,29]; i.e. these chemical systems permit bistability for some set of non-zero parameter values and under the assumption of mass action kinetics (see *Text S2*).

Segregation of kinase and phosphatase activities allows ultrasensitivity and bistability

Taken together, these analyses suggest that the ability of a split kinase to mediate ultrasensitivity and bistability relates to the segregation of kinase and phosphatase activities. To better understand how this relates to ultrasensitivity and bistability, we simulated the time evolution of the system in the presence of step signals. As expected from the ultrasensitive signal-response relationship, system response (i.e. increase in free CheY6-P) was low for step-signals below the threshold and displayed a sudden large jump for step-signals crossing the threshold (Figure 3). Before the threshold, increasing signal levels resulted in an increase in the CheA3:CheY6-P complex, while the crossing of the threshold and subsequent increases in signal caused it to decrease. The reason for this behavior is that before the threshold there is enough free CheA3 in the system to bind and dephosphorylate the CheY6-P that is formed, while after crossing of the threshold there is no free CheA3 left in the system (Figure 3). These observations can be understood if we consider the cyclic nature of the reactions in this system as shown in Figure 1. The free CheA3 can be seen as a branching point in the system, with one branch leading to binding to CheA4 and ultimately to more CheY6 phosphorylation (phosphorylation branch), while the other leading to binding to CheY6-P and subsequent dephosphorylation (dephosphorylation branch). While the phosphorylation branch is regulated externally of the system by signals sensed by the cytoplasmic cluster (i.e. through altering *k3* and/or *k5*), the dephosphorylation branch is controlled internally by the covalent modification of CheY6. This results in a dynamical motif that is similar to that seen in metabolic branching points and that can embed ultrasensitivity [33]. The split kinase system can embed a high level of nonlinearity as it contains both an inverse coupling of the two branches themselves (via CheY6) and their regulation (via CheA3). At low signals, these two branches allow enough free CheA3 in the system so to result in equally fast phosphorylation and dephosphorylation of CheY6. As the signal increases, however, the rate of the phosphorylation branch increases, while at the same time shutting down the dephosphorylation branch. In other words, the phosphorylation and dephosphorylation branches are coupled inversely, such that the kinase activity cannot be increased without reducing the phosphatase activity and vice versa. These dynamics can be observed in Figure 3; the loss of free CheA3 in the system coincides with an abrupt increase in CheA3-P and CheY6-P, while the CheA3:CheA4 complex maintains a fast turnover. This dynamical picture also explains the parameter effects observed in Figure 2 (and Figures S1, S2, S3, S4). For example, the decrease in ultrasensitivity from the reduction of CheA3-CheA4 association rate constant (*k1*) can be explained by a slowing down of the phosphorylation branch. Similarly, the decrease in ultrasensitivity from the inclusion of additional phosphatase activity via species other than free CheA3 can be explained by its perturbing effects on the inverse coupling between the phosphorylation and dephosphorylation branches (Figure S4 and S5). It must also be noted that the total level of CheA4 in the cell allows additional (internal) control on the dynamics of the system (Figure 2B and Figure S3), through its effects on the phosphorylation branch.

To further test whether the inverse coupling of kinase and phosphatase activities through free CheA3 is the underpinning mechanism of ultrasensitivity, we considered dynamics in two alternative models where such coupling is missing; (i) a bifunctional HK that is not split, and (ii) a traditional HK that is neither bifunctional nor split, with a dedicated auxiliary phosphatase for the phosphorylated RR. An analytical treatment of the dynamics arising in the former scenario suggests that non-split bifunctional HKs (where the phosphorylated/non-phosphorylated HK acts as kinase/phosphatase on its cognate response regulator) gives rise to hyperbolic signal-response relationships and provides the system with robustness towards variations in component concentrations [9]. For the latter scenario (e.g. CheA-CheY-CheZ system found in the *E. coli* chemotaxis system) we developed a simplified model and solved it for the steady state levels of phosphorylated response regulator. We compared this analytical solution to that derived from a simplified model of a split kinase system (see *Text S1*, *section 4*). This analytical treatment shows that the latter displays a higher level of nonlinearity for the steady state expression of phosphorylated RR. More importantly, we find that of the three possible alternative structures - bifunctional and split, monofunctional and split, bifunctional and non-split - only the chemical reaction system arising from the bifunctional and split kinase have the capacity for multiple steady states according to the higher deficiency theorem [32,29] (see *Text S3–6* for detailed results). Taken together, these analytical findings show that for bistable and ultrasensitive dynamics to be realized in a split kinase system, *both* bifunctionality of the HK and the splitting of these two functionalities (i.e. kinase and phosphatase activity) are needed.

Experimental verification that free CheA3 is a better phosphatase than CheA3:CheA4

As shown above, the ability of the split kinase to achieve both segregation and inverse coupling of kinase and phosphatase activities requires that free CheA3 is the predominant phosphatase with other CheA3 containing species (in particular CheA3:CheA4 and CheA3:CheA4:ATP) showing much lower phosphatase activity. Testing this requirement, or directly the level of ultrasensitivity *in vivo*, is complicated both by the presence of additional components in the system and our lack of knowledge of the signal identity in split kinase systems studied to date. As an alternative, and to achieve an approximate test of our theoretical understanding of split kinase response dynamics, we performed *in vitro* measurements of CheY6-P dephosphorylation in the presence of CheA3 and CheA4. In these experiments we used a purified phosphorylated P1 domain of CheA3 (CheA3P1-P) as the sole phosphodonor in the environment. As CheA3P1-P is known to lack phosphatase activity [7], this setup allows us to test directly the phosphatase activity of free CheA3 and the CheA3:CheA4 complex. If kinase and phosphatase activities are segregated into the complexed and free CheA3 respectively, these measurements should reveal a decrease of phosphatase activity with increasing CheA4 concentration, as this would sequester free CheA3 into the CheA3:CheA4 complex. In contrast, such an effect would be absent if the CheA3:CheA4 complex possessed the same level of phosphatase activity as free CheA3. We found evidence for such a decrease, with increasing CheA4 concentrations reducing the rate of CheA3 mediated dephosphorylation of CheY6-P (Figure 4 and Figure S8). To rule out the possibility of any interference from free CheA4, we have also confirmed the lack of dephosphorylation activity by CheA4 (Figure 4B). This observation qualitatively matches predictions from a specific model of this *in vitro*

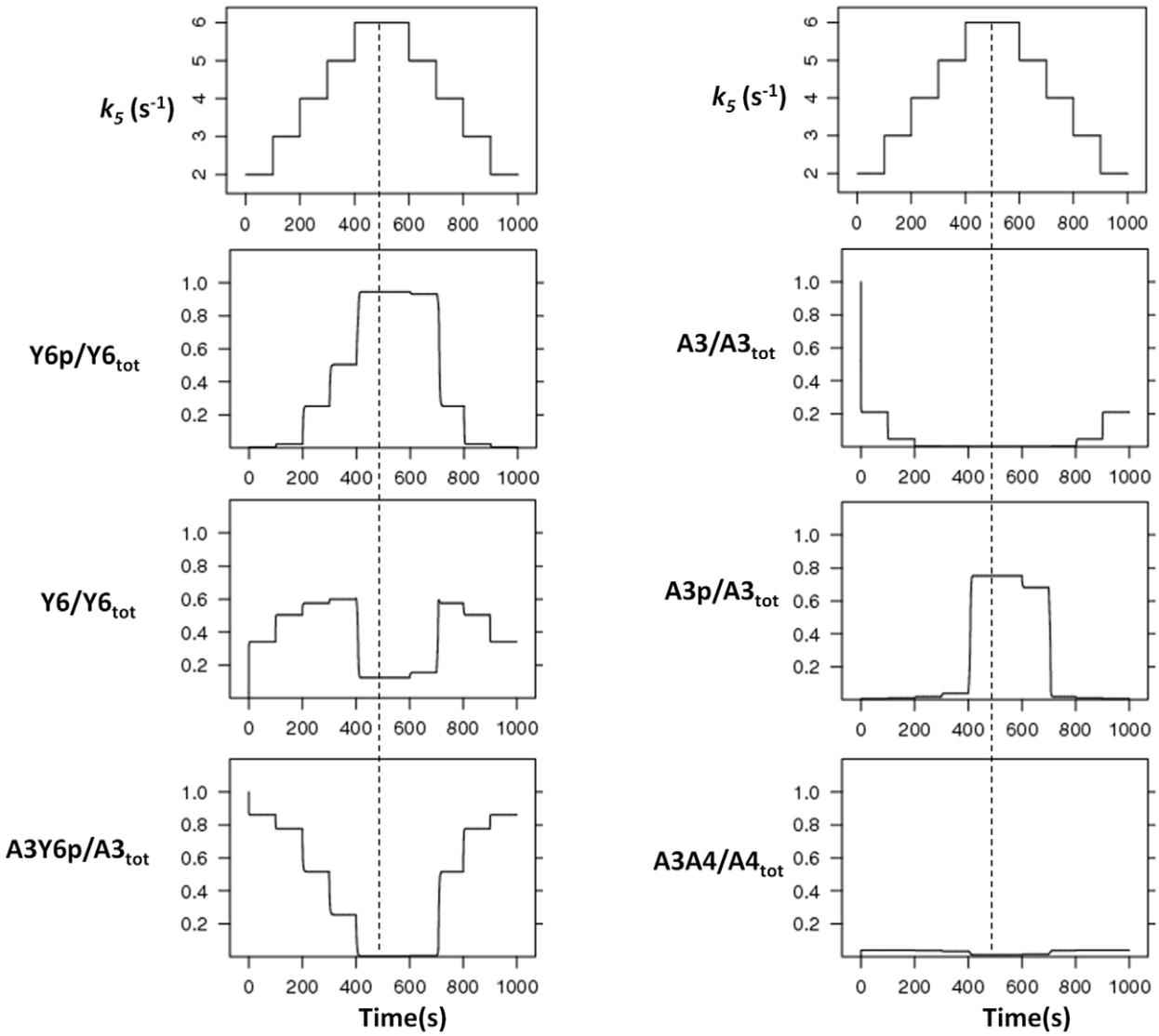

Figure 3. Time-course analyses. The model is simulated with increasing and decreasing signal levels (k_5) in course of time. k_5 is increased from 2 to 6 and decreased in similar fashion at indicated time points (top most, left panel), and changes in each species were measured (as indicated on each panel). The dotted line represents the highest signal level, with equal signal steps on each side of it. The noted asymmetry around this line shows the presence of hysteresis in the system. The x- and y-axis represent time and species concentration respectively, where the latter is normalized by the appropriate total protein levels.
doi:10.1371/journal.pcbi.1002949.g003

experimental setup where we assumed phosphatase activity to be restricted to only free CheA3 (see *Text S1* and Figure 4). These experimental findings strongly suggest that the CheA3:CheA4 complex has much lower phosphatase activity than free CheA3.

Discussion

Two component signaling systems mediate many of the physiological responses of bacteria and display several conserved biochemical and structural features. Here, we analyzed how one such feature, the split kinase, affects response dynamics. Our theoretical treatment proved that the chemical reaction system arising from a bifunctional split kinase gives rise to the possibility of bistability, whereas systems arising from bifunctional, non-split and monofunctional, split kinases lack such capability (unless featuring dead-end complex formation [12]). Sampling the parameter space around kinetic rates and protein concentrations measured in (or estimated from) *R. sphaeroides*, we found that a split kinase system set in a "biologically relevant" parameter regime has potential for an ultrasensitive and bistable signal-response relationship. These nonlinear dynamics arise from the bifunctional and split nature of the kinase, which introduce a branching point into the system between phosphorylation and dephosphorylation reactions. Thus, the level of ultrasensitivity (and emergence of bistability) in the system is determined by the parameters and the biochemical mechanisms found in the reaction cycles linked to this branching point.

We found that the one crucial biochemical aspect enabling ultrasensitivity and bistability in the split kinase system is the predominant allocation of phosphatase activity to the free protein (rather than any of the complexes in the system). Using *in vitro* phosphotransfer assays in the CheA3-CheA4-CheY6 split kinase system isolated from *R. sphaeroides*, we found support for free CheA3 being the principal phosphatase in that system (Figure 4). It

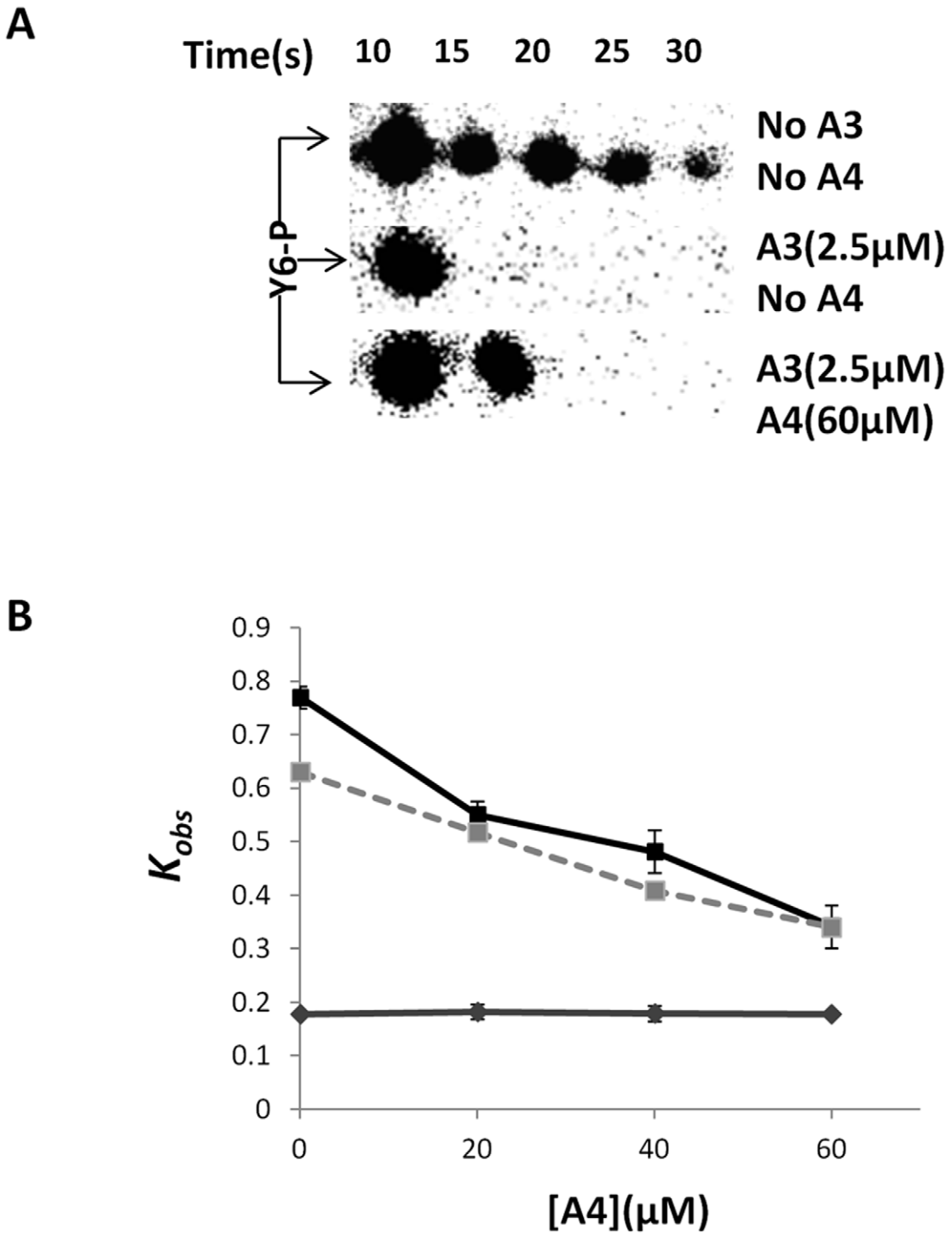

Figure 4. Measurement of CheY6-P dephosphorylation rates under different conditions (as indicated). An excess of CheY6 was phosphorylated using CheA3P1-P as phosphodonor. The phosphotransfer reaction was complete within 10 s of adding CheY6 to the reaction mixture. Subsequently the decay in CheY6-P levels was followed over time. (**A**) Phosphorimages showing the decay in CheY6-P levels over time. (**B**) Graph comparing the observed pseudo-first order rate constant (k_{obs}) for CheY6-P dephosphorylation with and without CheA3 and CheA4. The values predicted by the modeling are shown with a dashed line, while the experimentally measured values are shown in black. Results from a control experiment (without CheA3 and solely CheA4) is shown in grey. Error bars show the standard error of the mean obtained from eight independent experiments.
doi:10.1371/journal.pcbi.1002949.g004

remains to be shown whether this system enables ultrasensitivity or bistability *in vivo*. The theoretical findings of this study suggest that the switch-like dynamics resulting from ultrasensitivity and bistability could be relevant in the physiological context of the CheA3-CheA4-CheY6 system, which is involved in the integration of cytoplasmic and extracellular signals for proper chemotaxis

[7,34]. It would be plausible for example, if the switching dynamics described here allowed cells to override external chemotaxis signals in favor of internal signals such as those related to metabolism, which could contribute to motility decisions [35–37]. As shown in Figure 2, several internal parameters of the system, including the total expression level of CheA4, allow control of the dynamics mediated through CheA3:CheA4 and might enable further tuning of such decision making mechanisms.

While our results highlight split kinases as a potential strategy for implementing ultrasensitivity in bacterial two-component systems, it is not the only one. Previous theoretical studies have found that ultrasensitivity can be achieved in phosphorelays [18,19], in classical HK-RR systems embedding specific spatial dynamics [38] and in systems with bifunctional HKs, where unphosphorylated HKs and RR form a dead-end complex that is incapable of HK autophosphorylation [12,39]. These findings suggest that there are several diverse structural, spatial and dynamics that are possible in bacterial two-component systems and that have the potential to enable nonlinear response dynamics. Our theoretical findings extend this list with split kinase systems. Further, we provide experimental support for a condition that increases their potential for generating ultrasensitivity and bistability. Such responses are known to be common in eukaryotes and can enable decision making at the cellular level [40–42]. Thus, it is perhaps not surprising that bacterial signaling systems harbor mechanisms to enable similar levels of ultrasensitivity.

Although rare, split kinases are found in several other bacteria. A recent study looking at CheAs identified 11 split CheAs (2.3%) versus 470 complete CheAs (97.7%) in fully sequenced non-redundant genomes [1]. In addition to these split CheAs, there is the potential for other HKs to be split where the HisKA (dimerization and histidine phosphotransfer) and the catalytic HATPase (histidine kinase ATPase) domains are found on separate proteins. *In vitro* studies of the osmosensing histidine kinase, EnvZ, have shown that it possible to split the HATPase and HisKA domains onto separate polypeptides whilst retaining their activity [43]. Interrogation of the SMART database reveals that out of the 42417 proteins containing HisKA domains (dimerization and histidine phosphotransferase), 1556 (3.66%) lack a HATPase (histidine kinase ATPase) domain (expect value$<$0.01), and of these, 711 (1.7%) have the phosphatase sequence motif (HE/DxxN/T) [44] and could therefore be split bi-functional kinases. The results presented here suggest that cells may use such split kinases to allow high sensitivity and bistability enabling switch-like physiological responses to environmental stimuli.

As the highly modular TCSNs are used by bacteria to control many of their physiological responses, it will be valuable to explore other mechanisms which can enable specific response dynamics in these systems and to determine the evolutionary drivers that were responsible for their emergence. This would increase our ability to better understand microbial signaling and exploit it in synthetic biology applications.

Methods

A mathematical model for a split kinase

To model the CheA3-CheA4-CheY6 split kinase system, we considered its dynamics in isolation of other cellular components. The reactions in this system that we have included in the "basic model" are (see also alternative reaction schemes shown in *Text S1*);

$$A3 + A4 \underset{k_2}{\overset{k_1}{\rightleftarrows}} A3A4$$

$$A3A4 + ATP \underset{k_4}{\overset{k_3}{\rightleftarrows}} A3A4ATP \overset{k_5}{\rightarrow} A3p + A4 + ADP$$

$$A3p + Y6 \underset{k_7}{\overset{k_6}{\rightleftarrows}} A3 + Y6p$$

$$Y6p \overset{k_8}{\rightarrow} Y6 + P_i$$

$$A3 + Y6p \underset{k_{10}}{\overset{k_9}{\rightleftarrows}} A3Y6p \overset{k_{11}}{\rightarrow} A3 + Y6 + P_i$$

where *A3*, *A4*, *Y6* stand for CheA3, CheA4 and CheY6 respectively and the *-p* suffix represents phosphorylated forms of these proteins. Variant models which include additional CheY6-P de-phosphorylation reactions involving alternative phosphatases such as CheA3-P, and CheA3:CheA4 complex are shown in *supplementary text S1*, and their effects are analyzed in Figure 2D and S4. The above "basic model" reaction scheme can be used to derive a system of ordinary differential equations (ODEs), which describe the changes in concentrations of proteins over time;

$$\frac{d[A3p]}{dt} = k_5\cdot[A3A4ATP] + k_7\cdot[A3]\cdot[Y6p] - k_6\cdot[A3p]\cdot[Y6p]$$

$$\frac{d[A3A4]}{dt} = k_1\cdot[A3]\cdot[A4] + k_4\cdot[A3A4ATP] - [A3A4]\cdot(k_2 + k_5\cdot[ATP])$$

$$\frac{d[A3A4ATP]}{dt} = k_3\cdot[A3A4]\cdot[ATP] - [A3A4ATP]\cdot(k_4 + k_5)$$

$$\frac{d[A3Y6p]}{dt} = k_9\cdot[A3]\cdot[Y6p] - [A3Y6p]\cdot(k_{10} + k_{11})$$

$$\frac{d[Y6p]}{dt} = k_{10}\cdot[A3Y6p] + k_6\cdot[A3p]\cdot[Y6] - [Y6p]\cdot(k_7\cdot[A3] + k_8 + k_9\cdot[A3])$$

In addition, we have three conservation equations;

$$[Y6]_{tot} = [Y6] + [Y6p] + [A3Y6p]$$

$$[A3]_{tot} = [A3] + [A3p] + [A3A4] + [A3A4ATP] + [A3Y6p]$$

$$[A4]_{tot} = [A4] + [A3A4] + [A3A4ATP]$$

To analyze the behavior of the split kinase motif with increasing signal, we simulated the incoming signals from receptors as an increase in the autophosphorylation rate of the kinase (k_5). The model was parameterized with data from literature (see Table 1). In the case of the dephosphorylation of CheY6-P by CheA3, we derived the relevant parameters (k_9, k_{10}, and k_{11}) through fitting simulation data to previously published *in vitro* dephosphorylation measurements [7]. Fitting was done using a hybrid genetic algorithm (functions ga and fmincon from the MATLAB Global Optimization Toolbox).

We numerically integrated the model to derive time course and steady state signal-response relationships. The latter analysis gives the steady state CheY6-P level at a given signal (k_5) where signal

Table 2. Plasmids and strains used and the associated literature source.

Strains/plasmid	Description	Source/Reference
E.coli strain M15pREP4	Expression host containing pREP4; kanamycin resistant	Qiagen
pQE30	IPTG inducible expression vector. Introduces RGS(H)6 at the N terminus of the expressed protein. Confers ampicillin resistance	Qiagen
pQE60	IPTG inducible expression vector. Introduces RGS(H)6 at the C terminus of the expressed protein. Confers ampicillin resistance	Qiagen
pQE60A3P1	CheA3P1 expression plasmid. pQE60 derivative	[7]
pQEY6	CheY6 expression plasmid. pQE30 derivative	[24]
pQEA3	CheA3 expression plasmid. pQE30 derivative	[21]
pQEA4	CheA4 expression plasmid. pQE30 derivative	[21]

doi:10.1371/journal.pcbi.1002949.t002

was taken as the rate of autophosphorylation of split kinase and allows deriving a so-called signal-response curve. This curve is found by numerically integrating the system to steady state at a fixed signal level and then numerically "following" this steady state (i.e. steady state CheY6-P level), while changing the signal. This analysis is equal to allowing the system to reach steady state under different signal values. Both time course and signal-response analyses were performed using the software packages XPPAUT (http://www.math.pitt.edu/~bard/xpp/xpp.html) and Oscill8 (http://oscill8.sourceforge.net/).

Sensitivity analysis. We have quantified the sensitivity of the shape of the signal-response curves to variations in each of the parameters from their described base values (Table 1) and in a biologically relevant range. For these analyses, we measured the "sigmoidality" of the signal-response curve, *RS*, as its maximum slope (s_{max}) multiplied by the signal level at which this slope occurs (k_{5s}) (i.e. $RS = k_{5s} \bullet s_{max}$). This measure is similar to the "response coefficent", which measures the slope between 90% and 10% saturation [33], but is better able to distinguish between hyperbolic and sigmoidal dose-response curves. For each parameter, we varied it in a wide range around its basic value and measured "sigmoidality" of the resulting dose-response curves, as well as the maximum response of the system (Figures S1, S2, S3). The same analysis is also applied for alternative models featuring additional phosphatase species (Figure S4).

Analytical comparison of different models. To perform a formal check for the potential of bistability in the different models (discussed in the main text and *Supplementary Information*), we have utilized the chemical network theory [28,29]. This theory provides several analytical tests that can provide a definite answer on the possibility of existence of multiple stationary states in a given reaction network. We have applied these tests to the basic and alternative models we had devised using the Chemical Network Tool v2.2 (http://www.chbmeng.ohio-state.edu/~feinberg/crntwin/). The model files used with this tool and describing the chemical reaction systems, as well as the analytical results from the tool are provided as supplementary *Text S2–4*.

Plasmid and strains. See Table 2 for the plasmids and strains used. *E. coli* strains were grown in LB medium at 37°C. Antibiotics were used at concentrations of 100 μg ml−1 for ampicillin and 25 μg ml−1 for kanamycin, where needed. *E. coli* M15pRep4 cells were made competent using the calcium chloride technique [45]. Transformations were performed according to [46].

Protein purification. His tagged *R. sphaeroides* CheA3, CheA4, CheA3P1 and CheY6 proteins were purified as described previously [47]. Protein purity and concentration was measured as described in [24]. Purified proteins were stored at −20°c.

Preparation of CheA3P1-^{32}P. CheA3P1 was phosphorylated using [γ-^{32}P] ATP and CheA4 and purified as described before with the following modifications [7]. Proteins were phosphorylated in reactions performed at 20°C in phosphotransfer buffer (50 mM Tris HCl, 10% (v/v) glycerol, 5 mM $MgCl_2$, 150 mM NaCl, 50 mM KCl, 1 mM DTT, pH 8.0). The final reaction volumes were 2 ml. For production of CheA3P1-^{32}P, reaction mixtures contained 300 μM CheA3P1 and 20 μM CheA4. Reactions were initiated by addition of 2 mM [γ-^{32}P] ATP (specific activity 14.8 GBq mmol^{-1}; PerkinElmer). After 1 hour incubation, samples were purified by using Ni-NTA columns (Qiagen) as described previously for unphosphorylated His-tagged CheA3 [47]. This purification step removed the unincorporated ATP and also removed the CheA4 protein from the CheA3P1-^{32}P preparation. Purified proteins were stored at −20°C.

Measurement of CheY6-P dephosphorylation rate

Assays were performed at 20°C in phosphotransfer buffer. Purified CheA3P1-^{32}P was used as the phosphodonor. An excess of CheY6 (100 μM) was added to 30 μM of purified CheA3P1-^{32}P in the presence of 2.5 μM CheA3 and 0–60 μM CheA4. Following the addition of CheY6, reaction aliquots of 10 μl were taken at the indicated time points and quenched immediately in 10 μl of 2 X SDS-PAGE loading dye(7.5% (w/v) SDS, 90 mM EDTA, 37.5 mM Tris HCl, 37.5% glycerol, 3% (v/v) β- mercaptoethanol, pH 6.8). Quenched samples were analyzed using SDS-PAGE and phosphorimaging as described previously [24].

Supporting Information

Figure S1 The sensitivity of the signal response curve "sigmoidality" to parameter changes. The "sigmoidality" of the signal-response curve, *RS*, is measured as its maximum slope (s_{max}) multiplied by the signal level at which this slope occurs (k_{5s}) (i.e. $RS = k_{5s} \bullet s_{max}$). On each panel, the y-axis shows the ratio of *RS*, resulting from models with different values of a specific parameter, to that resulting from the basic model. x-axis shows the ratio of this parameter value to its corresponding value in the basic model. Data points in red indicates presence of bistability in the signal-response relationship. Note the log scale on both axes.
(TIF)

Figure S2 The sensitivity of the maximum phosphorylation level of CheY6 to parameter changes. On each panel, the y-axis shows the ratio of the maximal CheY6 phosphorylation, resulting from models with different values of a specific parameter, to that resulting from the basic model. x-axis shows the ratio of this

parameter value to its corresponding value in the basic model. Data points in red indicates presence of bistability in the signal-response relationship. Note the log scale on both axes.
(TIF)

Figure S3 The sensitivity of the signal response curve "sigmoidality" to changes in the concentration of CheA3 (**A**) and CheA4 (**B**). The "sigmoidality" of the signal-response curve, *RS*, is measured as its maximum slope (s_{max}) multiplied by the signal level at which this slope occurs (k_{5s}) (i.e. $RS = k_{5s} \bullet s_{max}$). On panel A (B), the y-axis shows the ratio of *RS*, resulting from models with different values of CheA3 (CheA4) concentration, to that resulting from the basic model. x-axis shows the ratio of this concentration to its corresponding value in the basic model. Data points in red indicates presence of bistability in the signal-response relationship. The sensitivity of the maximum phosphorylation level of CheY6 to changes in the concentration of CheA3 (**C**) and CheA4 (**D**). On panel C (D), the y-axis shows the ratio of the maximal CheY6 phosphorylation, resulting from models with different values of CheA3 (CheA4) concentration, to that resulting from the basic model. x-axis shows the ratio of this concentration to its corresponding value in the basic model. Data points in red indicates presence of bistability in the signal-response relationship. Note the log scale on both axes on all panels.
(TIF)

Figure S4 Analysis of signal-response relationship, in an alternative model considering phosphatase activity from additional species (see Supplementary Information, section 1). (**A**) Signal-response curves resulting from a model where both CheA3:CheA4 and CheA3:CheA4:ATP are considered to have phosphatase activity in addition to CheA3. For comparison, signal-response curve from the basic model is shown in red. Where present, the dark region indicates the region of unstable steady states and hence the presence of bistability. The different curves correspond to increasing levels of phosphatase activity (shown with the arrow) from the additional species. Phosphatase activity is varied in the same way for both CheA3:CheA4 and CheA3:CheA4:ATP by assuming that k_{on} and k_{cat} for these species are the same (i.e. $k_{12} = k_{15}$ and $k_{14} = k_{17}$) and by varying one set of rates simultaneously. The ratio between these rates (k_{12} and k_{14}) to their corresponding values for CheA3 (k_9 and k_{11}) is shown on the x-axis of panel C. (**B**) Signal-response curves resulting from a model where CheA-P is considered to have phosphatase activity in addition to CheA3. For comparison, signal-response curve from the basic model is shown in red. Where present, the dark region indicates the region of unstable steady states and hence the presence of bistability. The different curves correspond to increasing levels of phosphatase activity (shown with the arrow) from CheA3-P. Phosphatase activity is varied by changing both k_{on} and k_{cat} for CheA3-P (i.e. k_{18} and k_{20}) simultaneously. The ratio between these rates (k_{18} and k_{20}) to their corresponding values for CheA3 (k_9 and k_{11}) is shown on the x-axis of panel D. (**C**) The sensitivity of the signal response curve "sigmoidality" to increasing phosphatase activity from CheA3:CheA4 and CheA3:CheA4:ATP. The "sigmoidality" of the signal-response curve, *RS*, is measured as its maximum slope (s_{max}) multiplied by the signal level at which this slope occurs (k_{5s}) (i.e. $RS = k_{5s} \bullet s_{max}$). y-axis shows the ratio of *RS*, resulting from models with increasing phosphatase activity by additional species, to that of resulting from the basic model. X-axis shows the ratio of kinetic rates governing phosphatase activity (k_{12} and k_{14}) to those in the basic model (k_9 and k_{11}). Data points in red indicates presence of bistability in the signal-response relationship. (**D**) The sensitivity of the signal response curve "sigmoidality" to increasing phosphatase activity from CheA3-P. The "sigmoidality" of the signal-response curve, *RS*, is measured as its maximum slope (s_{max}) multiplied by the signal level at which this slope occurs (k_{5s}) (i.e. $RS = k_{5s} \bullet s_{max}$). Y-axis shows the ratio of RS, resulting from models with increasing phosphatase activity by additional species, to that of resulting from the basic model. x-axis shows the ratio of kinetic rates governing phosphatase activity (k_{18} and k_{20}) to those in the basic model (k_9 and k_{11}). Data points in red indicates presence of bistability in the signal-response relationship. Note the log scale on both axes in panels C and D.
(TIF)

Figure S5 Time-course analysis using an alternative model where both CheA3:CheA4 and CheA3:CheA4:ATP are considered to have phosphatase activity in addition to CheA3 (see Supplementary Information, section 1). The model is simulated with increasing and decreasing signal levels (k_5) in course of time. k_5 is increased from 2 to 6 and decreased in similar fashion at indicated time points (top most, left panel), and changes in each species were measured (as indicated on each panel). The x- and y-axis represent time and species concentration respectively, where the latter is normalized by the appropriate total protein levels.
(TIF)

Figure S6 Signal-response curves resulting from an alternative model that allows for the possibility that phosphorylated CheA3 remains in complex with CheA4 and that this CheA3p:CheA4 complex is also capable of acting as phosphatase towards CheY6p (see Supplementary Information, section 2). The y-axis shows steady state Y6-P level normalised by total Y6, while x-axis shows signal (k_5) level. Where present, a dark region indicates the region of unstable steady states and hence the presence of bistability. (a) The signal-response curve from the basic model (included for comparison). (b) Signal-response curve from the alternative model and simulating signal level through changing both k'_5 and k_5 simultaneously. (c) Signal-response curve from the alternative model and simulating signal level through changing k_5, while $k'_5 = 0.1\ s-1$.
(TIF)

Figure S7 Analysis of signal-response relationship, in an alternative model considering additional kinase activity (see Supplementary Information, section 3). (**A**) Signal-response curves resulting from a model where additional kinase activity (from CheA2) is considered. For comparison, the signal-response curve from the basic model is shown in red. Where present, the dark region indicates the region of unstable steady states and hence the presence of bistability. The different curves correspond to increasing levels of autophosphorylation rates for CheA2 (i.e. increasing background signalling through CheA2). (**B**) The sensitivity of the signal-response "sigmoidality" with increasing background kinase activity (from CheA2). The "sigmoidality" of the signal-response curve, *RS*, is measured as its maximum slope (s_{max}) multiplied by the signal level at which this slope occurs (k_{5s}) (i.e. $RS = k_{5s} \bullet s_{max}$). y-axis shows the ratio of *RS*, resulting from models with increasing background kinase activity (k^*_5) to that of the case where such activity is minimal (i.e. $k^*_5 \sim 0$). Data points in red indicates presence of bistability in the signal-response relationship. Note the log scale on both axes.
(TIF)

Figure S8 CheY6-P dephosphorylation time course data (circles) along with the fitted first-order exponential decay curves (red line) and simulated data (black line). The exponential fits are used to derive an estimate for overall CheY6p dephosphorylation rate (k_{obs}), which are shown in Figure 4.
(TIF)

Table S1 Parameter values used for the models with additional phosphatases.
(PDF)

Table S2 Parameter values used for the models with alternative reaction scheme.
(PDF)

Table S3 Parameter values used for the models with additional kinase.
(PDF)

Table S4 Parameter values used for the model of the *in vitro* experimental system.
(PDF)

Text S1 Supplementary information on alternative models and their analyses.
(PDF)

Text S2 Results of the analytical analysis of the basic model. The file contains the reaction system considered and the report produced with the Chemical Network Tool v2.2 (http://www.chbmeng.ohio-state.edu/~feinberg/crntwin/).
(DOC)

Text S3 Results of the analytical analysis of a model with a monofunctional kinase and a separate phosphatase. The file contains the reaction system considered and the report produced with the Chemical Network Tool v2.2 (http://www.chbmeng.ohio-state.edu/~feinberg/crntwin/).
(DOC)

Text S4 Results of the analytical analysis of a model with a monofunctional kinase. The file contains the reaction system considered and the report produced with the Chemical Network Tool v2.2 (http://www.chbmeng.ohio-state.edu/~feinberg/crntwin/).
(DOC)

Text S5 Results of the analytical analysis of a model with a bifunctional, non-split kinase. The file contains the reaction system considered and the report produced with the Chemical Network Tool v2.2 (http://www.chbmeng.ohio-state.edu/~feinberg/crntwin/).
(DOC)

Text S6 Results of the analytical analysis of a model with a monofunctional, split kinase. The file contains the reaction system considered and the report produced with the Chemical Network Tool v2.2 (http://www.chbmeng.ohio-state.edu/~feinberg/crntwin/).
(DOC)

Acknowledgments

We wish to thank Dr. Francesco Montefusco for his kind help on parameter fitting.

Author Contributions

Conceived and designed the experiments: MA SLP OSS. Performed the experiments: MA. Analyzed the data: MA SLP OSS. Contributed reagents/materials/analysis tools: MA SLP OSS. Wrote the paper: MA SLP OSS.

References

1. Wuichet K, Zhulin IB (2010) Origins and Diversification of a Complex Signal Transduction System in Prokaryotes. Sci Signal 3: ra50.
2. Hamer R, Chen PY, Armitage JP, Reinert G, Deane CM (2010) Deciphering chemotaxis pathways using cross species comparisons. BMC Syst Biol 4: 3.
3. Stock AM, Robinson VL, Goudreau PN (2000) Two-component signal transduction. Annu Rev Biochem 69: 183–215.
4. Batchelor E, Goulian M (2003) Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system. Proc Natl Acad Sci U S A 100: 691–696.
5. Sourjik V, Schmitt R (1998) Phosphotransfer between CheA, CheY1, and CheY2 in the chemotaxis signal transduction chain of *Rhizobium meliloti*. Biochemistry 37: 2327–2335.
6. Hoch JA (2000) Two-component and phosphorelay signal transduction. Curr Opin Microbiol 3: 165–170.
7. Porter SL, Roberts MAJ, Manning CS, Armitage JP (2008) A bifunctional kinase-phosphatase in bacterial chemotaxis. Proc Natl Acad Sci U S A 105: 18531–18536.
8. Soyer OS, Creevey CJ (2010) Duplicate retention in signalling proteins and constraints from network dynamics. J Evol Biol 11: 2410–21.
9. Shinar G, Milo R, Martínez MR, Alon U (2007) Input output robustness in simple bacterial signaling systems. Proc Natl Acad Sci U S A 104: 19931–19935.
10. Alves R, Savageau MA (2003) Comparative analysis of prototype two-component systems with either bifunctional or monofunctional sensors: differences in molecular structure and physiological function. Mol Microbiol 48: 25–51.
11. Groban ES, Clarke EJ, Salis HM, Miller SM, Voigt CA (2009) Kinetic buffering of cross talk between bacterial two-component sensors. J Mol Biol 390: 380–393.
12. Igoshin OA, Alves R, Savageau MA (2008) Hysteretic and graded responses in bacterial two-component signal transduction. Mol Microbiol 68: 1196–1215.
13. Tiwari A, Ray JC, Narula J, Igoshin OA (2011) Bistable responses in bacterial genetic networks: designs and dynamical consequences. Math Biosci 231: 76–89.
14. Jiménez-Pearson MA, Delany I, Scarlato V, Beier D (2005) Phosphate flow in the chemotactic response system of Helicobacter pylori. Microbiology 151: 3299–3311.
15. Bischofs IB, Hug JA, Liu AW, Wolf DM, Arkin AP (2009) Complexity in bacterial cell-cell communication: quorum signal integration and subpopulation signaling in the *Bacillus subtilis* phosphorelay. Proc Natl Acad Sci U S A 106: 6459–6464.
16. Burbulys D, Trach KA, Hoch JA (1991) Initiation of sporulation in *Bacillus subtilis* is controlled by a multicomponent phosphorelay. Cell 64: 545–552.
17. Appleby JL, Parkinson JS, Bourret RB (1996) Signal transduction via the multi-step phosphorelay: not necessarily a road less traveled. Cell 86: 845–848.
18. Kim J-RR, Cho K-HH (2006) The multi-step phosphorelay mechanism of unorthodox two-component systems in E. coli realizes ultrasensitivity to stimuli while maintaining robustness to noises. Comput Biol Chem 30: 438–444.
19. Csikász-Nagy A, Cardelli L, Soyer OS (2010) Response dynamics of phosphorelays suggest their potential utility in cell signalling. J R Soc Interface 8(57): 480–8.
20. Porter SL, Wadhams GH, Armitage JP (2011) Signal processing in complex chemotaxis pathways. Nat Rev Microbiol 9: 153–165.
21. Porter SL, Armitage JP (2004) Chemotaxis in *Rhodobacter sphaeroides* requires an atypical histidine protein kinase. J Biol Chem 279: 54573–54580.
22. Porter SL, Wadhams GH, Martin AC, Byles ED, Lancaster DE, et al. (2006) The CheYs of Rhodobacter sphaeroides. J Biol Chem 281: 32694–32704
23. Wadhams GH, Warren AV, Martin AC, Armitage JP (2003) Targeting of two signal transduction pathways to different regions of the bacterial cell. Mol Microbiol 50: 763–770.
24. Porter SL, Warren AV, Martin AC, Armitage JP (2002) The third chemotaxis locus of *Rhodobacter sphaeroides* is essential for chemotaxis. Mol Microbiol 46: 1081–1094.
25. Scott KA, Porter SL, Bagg EAL, Hamer R, Hill JL, et al. (2010) Specificity of localization and phosphotransfer in the CheA proteins of Rhodobacter sphaeroides. Mol Microbiol 76: 318–330.
26. Bell CH, Porter SL, Strawson A, Stuart DI, Armitage JP (2010) Using structural information to change the phosphotransfer specificity of a two-component chemotaxis signalling complex. PLoS Biol 8: e1000306.
27. Porter SL, Armitage JP (2002) Phosphotransfer in Rhodobacter sphaeroides chemotaxis. J Mol Biol 324: 35–45.
28. Craciun G, Tang YZ, Feinberg M (2006) Understanding bistability in complex enzyme-driven reaction networks. Proc. Natl Acad Sci USA 103:8697–8702.
29. Shinar G, Feinberg M (2012) Concordant Chemical Reaction Networks and the Species-Reaction Graph. Mathematical biosciences. E-pub ahead of print. doi: 10.1016/j.mbs.2012.08.002.
30. Tyson JJ, Chen KC, Novak B (2003) Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Curr Opin Cell Biol 15: 221–231.
31. Harrison DM, Skidmore J, Armitage JP, Maddock JR (1999) Localization and environmental regulation of MCP-like proteins in Rhodobacter sphaeroides. Mol Microbiol 31: 885–892.
32. Feinberg M (1987) Chemical reaction network structure and the stability of complex isothermal reactors. I. The deficiency zero and deficiency one theorems. Chem. Eng. Science 42: 2229–2268.
33. LaPorte DC, Walsh K, Koshland DE (1984) The branch point effect. Ultrasensitivity and subsensitivity to metabolic control. J Biol Chem 259: 14068–14075.

34. Tindall MJ, Porter SL, Maini PK, Armitage JP (2010) Modeling chemotaxis reveals the role of reversed phosphotransfer and a bi-functional kinase-phosphatase. PLoS Comput Biol 6: e1000896.
35. Goldstein RA, Soyer OS (2008) Evolution of taxis responses in virtual bacteria: non-adaptive dynamics. PLoS Comput Biol 4: e1000084.
36. Egbert MD, Barandiaran XE, Di Paolo EA (2010) A minimal model of metabolism-based chemotaxis. PLoS Comput Biol 6: e1001004.
37. Soyer OS, Goldstein RA (2011) Evolution of response dynamics underlying bacterial chemotaxis. BMC Evol Biol 11: 240.
38. van Albada SB, Ten Wolde PR (2009) Differential affinity and catalytic activity of CheZ in E. coli chemotaxis. PLoS Comput Biol 5: e1000378.
39. Salvado B, Vilaprinyo E, Karathia H, Sorribas A, Alves R (2012) Two component systems: physiological effect of a third component. PLoS One 7(2):e31095.
40. Goldbeter A, Koshland DE (1981) An amplified sensitivity arising from covalent modification in biological systems. Proc Natl Acad Sci U S A 78: 6840–6844.
41. Huang CYF, Ferrell JE (1996) Ultrasensitivity in the mitogen-activated protein kinase cascade. Proc Natl Acad Sci U S A 93: 10078–10083.
42. Kholodenko BN (2006) Cell-signalling dynamics in time and space. Nature Reviews Molecular Cell Biology 7: 165–176.
43. Park H, Saha SK, Inouye M (1998) Two-domain reconstitution of a functional protein histidine kinase. Proc Natl Acad Sci U S A 95: 6728–6732.
44. Huynh TN, Stewart V (2011) Negative control in two-component signal transduction by transmitter phosphatase activity. Mol Microbiol 82: 275–286.
45. Sambrook J, Russell JB (2001) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY, , USA: Cold Spring Harbor Laboratory Press.
46. Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166: 557.
47. Porter SL, Wadhams GH, Armitage JP (2007) In vivo and in vitro analysis of the *Rhodobacter sphaeroides* chemotaxis signaling complexes. Methods Enzymol 423: 392–413.
48. Silversmith RE, Levin MD, Schilling E, Bourret RB (2008) Kinetic characterization of catalysis by the chemotaxis phosphatase CheZ. Modulation of activity by the phosphorylated CheY substrate. J Biol Chem 283: 756–765.