Reading List
Two-component Systems and cyclic-diGMP-mediated Regulation
December 18, 2007
Tod J. Merkel

1. Bijlsma, J. J., and E. A. Groisman. 2003. Making informed decisions: regulatory interactions between two-component systems. *Trends Microbiol* 11:359-366.
2. Romling, U., and D. Amikam. 2006. Cyclic di-GMP as a second messenger. *Curr Opin Microbiol* 9:218-228.
3. Romling, U., M. Gomelsky, and M. Y. Galperin. 2005. C-di-GMP: the dawning of a novel bacterial signalling system. *Mol Microbiol* 57:629-639.
4. Calva, E., and R. Oropeza. 2006. Two-component signal transduction systems, environmental signals, and virulence. *Microb Ecol* 51:166-176.
5. Stock, A. M., V. L. Robinson, and P. N. Goudreau. 2000. Two-component signal transduction. *Annu Rev Biochem* 69:183-215.
6. Hoch, J. A. 2000. Two-component and phosphorelay signal transduction. *Curr Opin Microbiol* 3:165-170.

Making informed decisions: regulatory interactions between two-component systems

Jetta J.E. Bijlsma and Eduardo A. Groisman

Department of Molecular Microbiology, Howard Hughes Medical Institute, Washington University School of Medicine, Campus Box 8230, 660 S. Euclid Avenue, St Louis, MO 63110, USA

Bacteria experience a multitude of stresses in the complex niches they inhabit. These stresses are denoted by specific signals that typically alter the activity of individual two-component regulatory systems. Processing of multiple signals by different two-component systems occurs when multiple sensors interact with a single response regulator, by phosphatases interrupting phosphoryl transfer in phosphorelays and through transcriptional and post-transcriptional mechanisms. Gaining insight into these mechanisms provides an understanding at the molecular level of bacterial adaptation to ever changing multifaceted environments.

A crucial question in signal transduction is: how does an organism integrate multiple inputs into the most appropriate action? Most of the present knowledge on bacterial gene regulation is based on studies conducted in laboratory settings, where conditions are carefully controlled and, typically, only one stress is applied at a time. Although this reductionist approach has been successful in elucidating several regulatory mechanisms, it does not take into consideration the multitude of simultaneous stresses that bacteria experience in the niches they occupy. We are now beginning to uncover how bacteria adjust their behavior in response to multiple cues from complex environments.

Bacteria sense environmental conditions using a variety of systems, which include secondary metabolites and ions, such as cyclic adenosine monophosphate and Fe^{2+}, and regulatory proteins, such as catabolite regulatory protein (CRP) and ferric uptake regulator (Fur), whose affinity for DNA changes on ligand binding. However, two-component systems are the predominant form of signal transduction used by bacteria to respond to environmental stresses [1,2]. Two-component systems usually consist of a sensor protein (or histidine kinase) and a cognate response regulator. Sensor proteins use adenosine triphosphate (ATP) to autophosphorylate at a conserved histidine residue, often in response to a specific signal. The high-energy phosphoryl group is then transferred to a conserved aspartate residue in the response regulator, often altering its ability to bind target DNA sequences (Fig. 1a) [1,2]. Certain response regulators have the ability to phosphorylate from small molecular weight phosphoryl donors, such as acetyl phosphate and carbamoyl phosphate [3,4]. In addition, the vast majority of sensor kinases exhibit phosphatase activity towards their cognate response regulators. Although most two-component systems entail a single His-to-Asp phosphotransfer event between a histidine kinase and a response regulator, there is a subset that consists of a three-step His-Asp-His-Asp phosphorelay (Fig. 1b) [1,2]. The extra phosphorylatable domains in a phosphorelay can be present in separate proteins or be part of multi-domain (i.e. hybrid) sensor proteins, which, in either case, will ultimately donate the

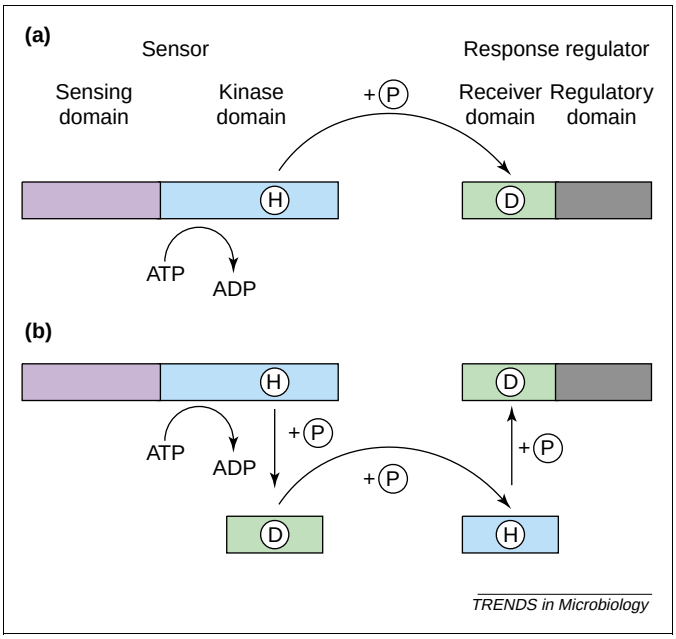


Fig. 1. Two-component systems consist of a sensor protein (histidine kinase) that transfers a high-energy phosphoryl group to the response regulator, which is often a transcription factor. (a) The sensor protein autophosphorylates on a conserved histidine residue (H) in the kinases domain (blue), often in response to signal sensing in the sensing domain (purple). The high-energy phosphoryl group (P) is then transferred to a conserved aspartate (D) residue in the receiver domain (green) of the response regulator, often changing the ability of its regulatory domain (grey) to bind DNA. (b) The His-Asp-His-Asp three-step phosphorelay consists of the transfer of the phosphoryl group from the conserved histidine in the sensor to additional phosphorylatable aspartate and histidine residues in auxiliary domains before transfer to the aspartate in the receiver domain of the response regulator. Many different arrangements of these extra phosphorylatable domains exist within proteins and pathways. They are either present in separate proteins, a representative example of which is depicted here, or part of multi-domain (hybrid) sensors.

Corresponding author: Eduardo A. Groisman (groisman@borcim.wustl.edu).

http://timi.trends.com 0966-842X/03/$ - see front matter doi:10.1016/S0966-842X(03)00176-8

phosphoryl group to the response regulator controlling gene transcription (Fig. 1b).

How is specificity between cognate sensor/regulator pairs maintained in organisms harboring multiple two-component systems? Based on conserved residues around the phosphorylatable histidine, sensor proteins have been divided into two major families and several minor ones [5]. The conserved residues of sensor proteins are proposed to function as 'anchors' that mediate the interaction with conserved residues around the phosphorylatable aspartate in the response regulator, bringing the catalytic residues of both molecules in close proximity [6,7]. Specificity between a cognate sensor/regulator pair is then mediated by hypervariable residues around the phosphorylatable aspartate in the response regulators [6,8]. This notion is strengthened by the observation that mutations in the hypervariable region of the response regulator PhoB increase its affinity for the non-cognate sensor VanS [9].

Two-component systems are ideally suited to function in the integration of multiple signals because the membrane location of the sensor protein allows transduction of environmental cues, and the high degree of specificity of each sensor/regulator pair ensures a physiological response. Different two-component systems can process multiple signals into a particular response by interacting with one another and/or by phosphorylating from small molecular weight phosphodonors [10,11]. This article will discuss: (1) how multiple sensors interact with a single response regulator (Fig. 2a); (2) the phosphorylation of response regulators by non-cognate sensors (Fig. 2b); (3) the ability of phosphatases to interrupt the phosphoryl flow in His-Asp-His-Asp phosphorelays (Fig. 2a) and (4) transcriptional and post-transcriptional control mechanisms (Fig. 2c, Fig. 2d).

Multiple sensors acting on a single response regulator

Different signals can be fed into the same regulatory pathway if the phosphorylated state of a single response regulator is modified by several cognate sensors, each responding to a different signal. The best-studied example of signal integration is bacterial chemotaxis, where swimming behavior is modified in response to changes in the concentration of different substrates. These changes are sensed by distinct chemoreceptors that, through the coupling protein CheW, alter the phosphorylated state of the histidine kinase CheA, resulting in phosphorylation of the response regulators CheY and CheB [12]. Phosphorylation of CheY promotes a conformational change that affects its ability to interact with the flagellar motor, altering flagellar rotation from smooth to tumble. A dedicated aspartyl phosphatase, termed CheZ, accelerates the dephosphorylation of phospho-CheY. Phosphorylation of CheB stimulates the methylesterase activity of CheB towards the chemoreceptors, which leads to downregulation of CheA autokinase activity, thereby providing a negative feedback [12].

Escherichia coli has five chemoreceptors, which respond to specific ligands [12] and are clustered at the bacterial poles; these chemoreceptors form signaling teams consisting of trimers of the same or different dimeric chemoreceptors [13,14]. The chemoreceptor clusters form a complex with CheA and CheW [14,15], thus enabling simultaneous processing of multiple signals. Furthermore, ligand binding by one chemoreceptor influences the surrounding chemoreceptors, leading to increased stimulation of CheA kinase activity [14,16,17].

A similar convergence of sensor proteins controls sporulation in the Gram-positive bacterium *Bacillus subtilis*. Here, five related histidine kinases – KinA, KinB, KinC, KinD and KinE – phosphorylate the response regulator Spo0F (Fig. 2a) [18]. The Kin proteins exhibit sequence similarity in the residues surrounding the phosphorylatable histidine residue that will donate the phosphoryl group to the conserved aspartatyl residue in Spo0F but differ in their putative sensing domains, indicating that they respond to different ligands or conditions. Consistent with this notion, the severity of the sporulation defect displayed by mutants in KinA, KinB or KinC depends on the growth conditions [19].

Quorum sensing governs bioluminescence in the marine bacterium *Vibrio harveyi* in response to two different autoinducers: AI-1, specifically produced and detected by *V. harveyi*, and AI-2, a product of the metabolism of a wide variety of microorganisms [20]. These autoinducers are sensed by two different hybrid histidine kinases: LuxN, which senses AI-1, and LuxQ, which responds to AI-2 via the periplasmic protein LuxP. These sensors converge onto a histidine-containing phosphotransfer protein termed LuxU, which will ultimately donate the phosphate to the response regulator LuxO [21]. Phosphorylated LuxO activates a repressor that turns off transcription of the bioluminescence genes. The potential for AI-1 and AI-2 to control independently the phosphorylated state of LuxO could allow *V. harveyi* to respond not only to its own cell density but also to that of other microbes. However, when both autoinducers are present, expression of the LuxO-controlled genes changes 10- to 100-fold more than when only one of them is present, suggesting that the quorum-sensing circuit of *V. harveyi* functions as a 'coincidence detector' that discriminates between the presence of both autoinducers and all other conditions (i.e. the presence of either autoinducer or no autoinducer at all) [21].

E. coli harbors a pair of two-component systems – NarQ/NarP and NarX/NarL – that exhibit sequence similarity, most strikingly around the phosphorylatable histidine for the sensors NarX and NarQ, and around the phosphorylatable aspartate residue for the regulators NarL and NarP [22,23]. Both NarX and NarQ respond to nitrate and nitrite and can phosphorylate both NarP and NarL. Nitrate is detected at lower concentrations than nitrite by NarX [24], which mainly dephosphorylates NarL when nitrite is the signal. Phosphorylation of the NarL and NarP proteins results in non-equivalent responses because these regulators recognize slightly different target sequences [25]. Thus, small differences in signal preference by the NarX and NarQ sensors, coupled to distinct promoter recognition by the NarL and NarP regulators, allows the integration of a range of signals (i.e. the concentrations of both nitrate and nitrite) into distinct responses.

Does 'cross-talk' occur between non-cognate two-component systems?

Despite the high degree of specificity revealed by structural and mutational analysis of two-component system proteins, certain histidine kinases have been shown to interact with and modify the phosphorylated state of non-cognate regulators in a process referred to as 'cross-talk' [26]. For example, the PhoB/PhoR system of *E. coli* is normally switched on in low phosphate conditions (Fig. 2b). However, in strains lacking the sensor PhoR, PhoB-activated genes can be induced by glucose and pyruvate in a process mediated by the non-cognate sensor CreC (Fig. 2b), or by the low molecular weight phosphodonor acetyl-phosphate [27]. The PhoB protein can be phosphorylated, even by the sensor protein VanS from the Gram-positive species *Enterococcus faecalis* [28], although the efficiency of this reaction is four orders of magnitude lower than that exhibited by cognate proteins [29], reflecting a non-physiological interaction. Other cases of 'cross-talk' include phosphorylation of the response regulator OmpR (which normally responds to osmolarity changes sensed by its cognate sensor EnvZ) by the sensor ArcB, which is activated in anaerobiosis [30].

It has been proposed that two-component systems use 'cross-talk' to integrate multiple signals that normally activate separate two-component systems [26]. However, the physiological relevance of 'cross-talk' is presently unclear because phosphotransfer between non-cognate sensors and regulators has been found only *in vitro*, or *in vivo* when the non-cognate proteins are present in high gene dosage and/or when the cognate sensor is missing. These are crucial issues because the efficiency of the interaction between a response regulator and a non-cognate sensor and/or low molecular weight phosphodonors appears to be too low for 'cross-talk' to occur when these proteins are present at normal levels [29,31]. Moreover, sensor proteins that exhibit phosphatase activity towards their cognate response regulators will abrogate spurious activation when the inducing signalsare absent. For example, a mutant allele of the *Salmonella* response regulator PhoP that can be activated (i.e. phosphorylated) from acetyl phosphate promotes gene transcription in the absence of PhoP's cognate sensor PhoQ [32]. This activation is independent of the concentration of Mg^{2+}, which is the signal controlling the PhoQ protein. However, high Mg^{2+} abolishes this activation in the presence of the PhoQ protein, reflecting the Mg^{2+}-promoted PhoQ-mediated dephosphorylation of phospho-PhoP. This indicates that 'cross-talk' might be better evaluated using mutant sensors that retain phosphatase activity but cannot serve as phosphodonors. Furthermore, a modeling study predicts that phosphatase activity provides sensors with a mechanism to avoid spurious 'cross-talk', and suggests that the lack of phosphatase activity in the histidine kinase CheA allows the incorporation of cues provided by other two-component systems into bacterial chemotaxis [33].

Two important studies raise further questions about the occurrence of physiological 'cross-talk'. One study [34] explored 'cross-talk' among the *E. coli* UhpA/UhpB, PhoB/PhoR, NtrB/NtrC and ArcA/ArcB two-component systems. In this study, the expression of target genes for each system was investigated when two or more signals specific for these systems were provided [34]. 'Cross-talk' could be detected only towards the response regulator NtrC when its cognate sensor NtrB was mutated. The results of the second study [35] suggest that the specificity of the interaction between sensor and regulator pairs appears to prevent non-cognate partners from interacting: the response regulators NtrC and DctD from *Rhizobium meliloti* have only four amino acids in common around the conserved aspartate residue, which is in contrast to the 14 residues shared between the equivalent region of the *Rhizobium* DctD and *E. coli* NtrC proteins. Thus, although the *Rhizobium* DctD/DctB system can 'cross-talk' with the *E. coli* NtrB/NtrC system in *E. coli*, 'cross-talk' has not been detected between the *R. meliloti* NtrB/NtrC and DctD/DctB systems [35]. This implies that the *R. meliloti* NtrC protein has diverged from the *E. coli* NtrC, possibly to avoid non-physiological interactions with the DctCD system. Cumulatively, the studies described above suggest that 'cross-talk' is rare, if it occurs at all in a way that is significant biologically.

Phosphatases can interrupt phosphoryl flow in a phosphorelay

The phosphorylated state of a response regulator is determined not only by its ability to phosphorylate from its cognate sensor kinase but, potentially, also by specific phosphatases that respond to additional signals. The His-Asp-His-Asp phosphorelay is ideally suited to integrate extra cues because the phosphorylation state of the several domains can be disrupted at different points in the cascade. The signal transduction cascade controlling *B. subtilis* sporulation is the paradigmatic example of integration of multiple signals into a phosphorelay. Triggered by environmental and nutritional conditions that are incompatible with bacterial growth, the sporulation phosphorelay begins with phosphorylation of the response regulator Spo0F by any of five different kinases (KinA, KinB, KinC, KinD or KinE), it is followed by phosphotransfer from Spo0F to Spo0B and ends with phosphotransfer to Spo0A (Fig. 2a), a transcription factor that activates the genes necessary for sporulation and represses genes involved in competence development [36,37].

Two families of aspartyl phosphatases control the sporulation phosphorelay: the Rap family members RapA, RapB and RapE [36,38] that specifically dephosphorylate Spo0F, and the phosphatases Spo0E, YnzD and YisI that target Spo0A [39]. Transcription of the *rapE* gene is promoted during vegetative growth, ensuring that sporulation does not take place in optimal growing conditions. Transcription of the RapA and RapB phosphatases is induced by a two-component system – ComA/ComP – that governs the development of competence, an alternative state that is incompatible with sporulation, ensuring that sporulation and competence do not develop at the same time [36,38]. In addition to their transcriptional control, the activity of Rap phosphatases is inhibited by specific pentapeptides derived from larger proteins that are encoded in the same operon and co-transcribed with the *rap* genes [37,38]. These larger proteins are secreted

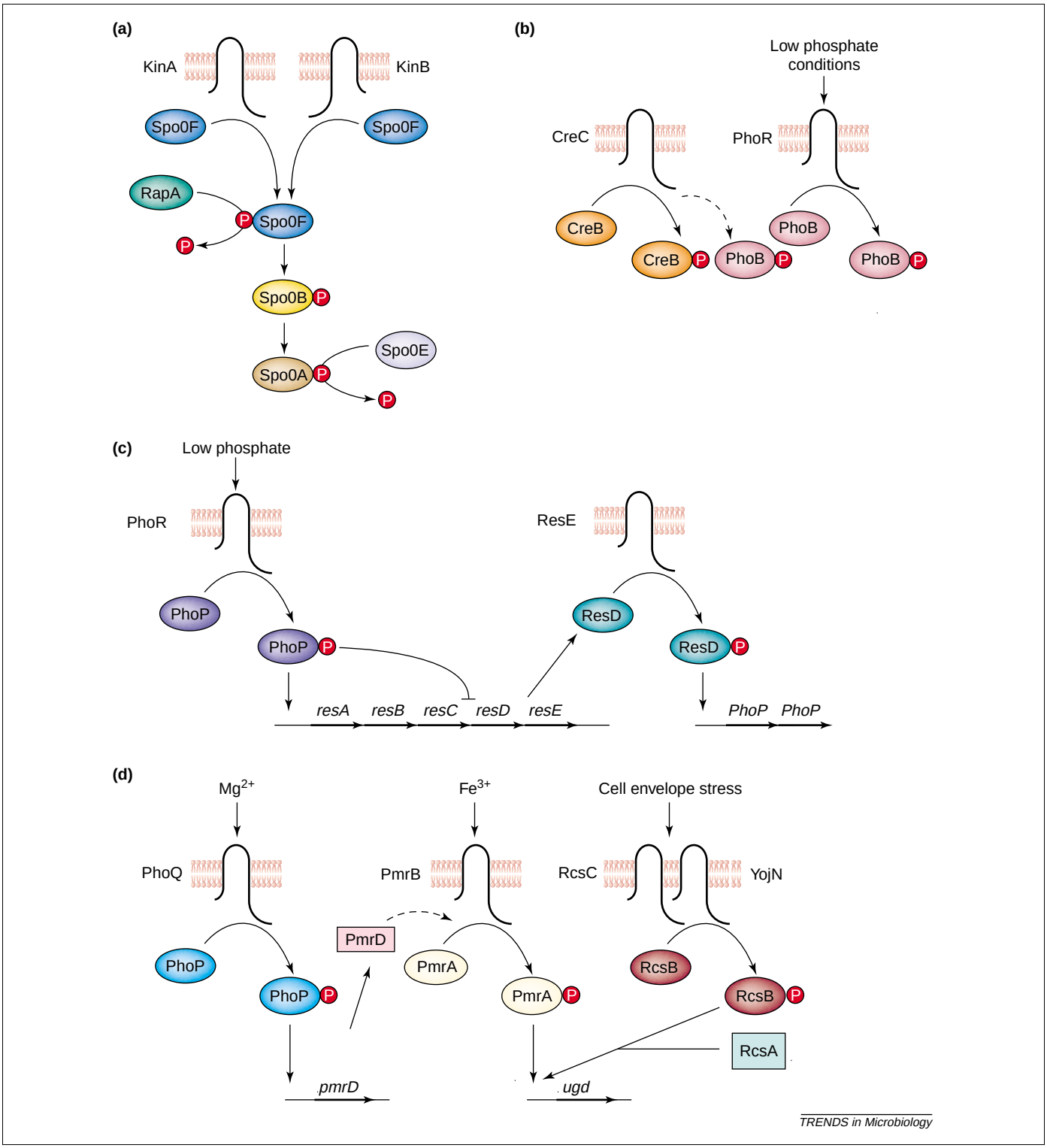


Fig. 2. Different two-component systems can process multiple signals into a particular response by interacting with each other. (a) Multiple cognate sensors can change the phosphorylated state of a single protein. In *B. subtilis*, any of five kinases (KinA, KinB, KinC, KinD and KinE) can phosphorylate the response regulator Spo0F. Sporulation is controlled by the phosphorylated state of Spo0A, which receives its phosphoryl group from Spo0B via a phosphorelay that starts with the phosphorylation of Spo0F by a kinase (here, only KinA and KinB are depicted for clarity) and is followed by phosphotransfer from Spo0F to Spo0B. This His-Asp-His-Asp phosphorelay provides several entry points for signal integration. Two families of aspartyl phosphatases dephosphorylate Spo0F and Spo0A and thus also determine their phosphorylation state. The Rap family consists of RapA, RapB and RapE (RapA is depicted here), and acts on Spo0F. The other family consists of Spo0E, YnszD and YisI (Spo0E is depicted here), and acts on Spo0A. (b) Sensors can interact with and modify the phosphorylated state of non-cognate response regulators in a process referred to as 'cross-talk'. In *Escherichia coli*, the histidine kinase PhoR normally phosphorylates the response regulator PhoB in response to low phosphate conditions. However, in the absence of the sensor PhoR, the non-cognate sensor CreC can phosphorylate PhoB. The efficiency of this interaction is probably lower than that between the cognate protein pairs, as indicated by the dashed line. (c) Transcriptional control of one two-component system by another two-component system leads to the transcription of the genes controlled by the first system being dependent on the signals for both systems. In *Bacillus subtilis*, transcription of the ResD/ResE system, which regulates genes required for anaerobic respiration, is controlled by the PhoP/PhoR system, which responds to phosphate starvation. When phosphate levels are low, transcription of *resDE* genes depends on the response regulator PhoP, which activates the promoter in front of the *resABCDE* operon and represses the constitutive promoter in front of *resDE*. In addition, the ResD protein is necessary for full activation of the PhoP/PhoR system, indicating that the ResD/ResE and PhoP/PhoR systems are connected through a positive feedback loop. This figure

and processed before the peptides are imported back into the cytoplasm by a specific oligopeptide permease. When these peptides accumulate beyond a certain level, they inhibit the RapA and RapB phosphatase activities, thereby allowing phosphotransfer to proceed from Spo0F to Spo0B. As peptide accumulation can be affected at several steps, this might provide a timing mechanism for the development of sporulation [38]. Thus, the regulatory pathways leading to competence or sporulation are closely intertwined, indicating that the bacteria constantly monitor the signals received and make a decision based on the relative weight each signal has at any given moment in time, which is determined by the relative activity of several kinases and phosphatases [36,37,40].

In Gram-negative bacteria, His-Asp-His-Asp phosphorelays typically consist of hybrid sensors. Sequential phosphotransfer within the same hybrid sensor must occur before transfer to the response regulator that mediates transcription [41,42]. The role of the extra signaling domains in hybrid sensors does not appear to be in integrating additional signals but in controlling other processes. For example, they determine the specificity for the cognate response regulator in the case of the sensors BvgS and EvgS [43] and regulate the autokinase activity of the sensor VirA [44]. The only phosphatase acting on two-component system proteins identified in *E. coli* – termed SixA – does not exhibit sequence identity to the *B. subtilis* phosphatases and acts on the histidine phosphotransfer domain of the hybrid sensor ArcB [45]. The *sixA* gene appears to be constitutively expressed and its deletion does not significantly interfere with the correct regulation of ArcA-regulated genes [46]. Thus, integration of signals through the His-Asp-His-Asp phosphorelay is used in *B. subtilis* sporulation but its prevalence in other systems and organisms awaits further investigation.

Transcriptional and post-transcriptional mechanisms

An intrinsic property of two-component systems is that gene transcription requires both the presence of the response regulator and the signal that promotes its activation, which is sensed by the cognate histidine kinase. Thus, when a two-component system is transcriptionally controlled by another two-component system, the genes regulated by the first system will be expressed only when the signals activating both systems are present. For example, transcription of the ResD/ResE system, which regulates the genes required for anaerobic respiration in *B. subtilis*, is controlled by the PhoP/PhoR system, which responds to phosphate starvation (Fig. 2c) [47]. When phosphate levels are low, transcription of *resDE* genes depends on both response regulators PhoP and ResD (which is expressed at low levels from a PhoP-independent promoter). In addition, the ResD protein is necessary for full activation of the PhoP/PhoR system, indicating that the ResD/ResE and PhoP/PhoR systems are connected through a positive feedback loop (Fig. 2c) which allows the products under control of ResD/ResE to incorporate the phosphate imported by the products controlled by PhoP/PhoR into ATP [47]. Microarray studies carried out in *E. coli* and *B. subtilis* have identified several cases in which a two-component system is under the transcriptional control of another two-component system [48,49], suggesting that the requirement for two signals for gene expression could be a common feature in two-component signal transduction.

Expression of certain genes is dependent on the simultaneous activation of several regulatory systems. In *Ralstonia solanacearum*, expression of the Eps virulence factor is governed by a LysR-type regulator and a pair of two-component systems [50]. The VsrA/VsrD system and the LysR-type regulator are required for expression of a protein that, together with another two-component system (termed VsrC/VsrB), promotes transcription of the *eps* gene [50,51]. Thus, *eps* is expressed only when the signals for both two-component systems and the LysR-type regulator are present, which effectively integrates three signals into one output.

It is often the case that particular genes are needed under different stress conditions. Expression of such genes can be accomplished by the presence of binding sites for several regulatory proteins in their promoters, or by post-transcriptional mechanisms that activate two-component systems even in the absence of their specific signals. An example of the first scenario is the regulation of the *Salmonella ugd* gene: it is transcriptionally regulated by the PmrA/PmrB system, which responds to extracytoplasmic Fe^{3+}, and by the RcsC/YojN/RcsB phosphorelay, which is activated by cell envelope stress (Fig. 2d) [52]. The *ugd* gene is necessary for the PmrA-controlled modification of the lipopolysaccharide and the RcsB-dependent production of the colanic acid capsule. The presence of distinct binding sites for the PmrA and RcsB response regulators in the *ugd* promoter allows for two different signals to activate the *ugd* gene independently.

In addition to the Fe^{3+} and cell envelope stress signals, low Mg^{2+} promotes transcription of the *ugd* gene in a novel process that does not require binding of the Mg^{2+}-responsive regulator PhoP to the *ugd* promoter (Fig. 2d) [53]. Low Mg^{2+}, which is sensed by the sensor kinase PhoQ, promotes expression of the PhoP-activated PmrD protein, which activates the PmrA/PmrB system at a post-transcriptional level. Thus, even in the absence of Fe^{3+}, which is the specific signal sensed by the PmrB protein, *ugd*, like other PmrA-activated genes, is expressed in response to the low Mg^{2+} signal. By contrast, the PmrD protein is not required for the Fe^{3+}-promoted activation of the PmrA/PmrB system. This enables *Salmonella* to express independently all PmrA-regulated gene products in response to two signals – high Fe^{3+} and low Mg^{2+}.

has been adapted from Birkey *et al.* [47]. (d) The presence of binding sites for several regulatory proteins in the promoter of an effector gene, and post-transcriptional mechanisms lead to the independent transcription of this gene in response to different conditions. In *Salmonella enterica*, the *ugd* gene is under the transcriptional control of the PmrA/PmrB system, which responds to extracytoplasmic Fe^{3+}, and the RcsC/YojN/RcsB system, which responds to envelope stress. Transcription of the *ugd* gene is also promoted in response to low Mg^{2+}, the signal for the sensor PhoQ. Upon phosphorylation, PhoP switches on expression of the PmrD protein, which activates the PmrA/PmrB system at a post-transcriptional level.

Box 1. Autoregulation of two-component systems

Many two-component systems positively regulate their own expression. Such systems typically harbor (at least) two promoters: a constitutive promoter that provides background levels of expression, producing enough sensor and regulator proteins to detect and respond to the presence of signal, and an autoregulated promoter that contains a binding site for the response regulator and is responsible for the production of additional sensor and regulator molecules [56–58]. If constitutive promoters supply sufficient levels of two-component proteins to react to a signal, what is the function of autoregulation?

Autoregulation could provide bacteria with a 'memory' of previous encounters with a signal because the increased amounts of sensor and response regulator proteins would still be present after the signal disappears [59]. Indeed, the PhoB/PhoR system reacted faster to phosphate limitation in organisms previously exposed to this condition. This 'learning' behavior was dependent on the presence of elevated levels of the PhoB and PhoR proteins, being abolished by constitutive expression of the PhoB/PhoR system [59].

Autoregulation might be necessary for those two-component systems in which the response regulator controls a number of target binding sites that is too large relative to the amount of response regulator made from the constitutive promoter. A survey of well characterized systems of *E. coli* and *Salmonella* supports a correlation between autoregulation and the presence of multiple regulated promoters: PhoQ/PhoP, PmrB/PmrA, CpxA/CpxR, PhoR/PhoB and NtrB/NtrC are all autoregulated and control at least four promoters, whereas the UhpB/UhpA system is not known to be autoregulated and appears to control only the *uhpT* gene.

Autoregulation could facilitate the hierarchical organization of a regulon so that certain genes are expressed by the phosphorylated response regulator produced from a constitutive promoter, whereas others are strictly dependent on higher levels of response regulator produced from an autoregulated promoter [60]. A temporal pattern of gene expression is observed when the BvgA/BvgS system of *Bordetella pertusis* is activated. A group of genes that includes *bvgAS* is expressed immediately upon activation when BvgA levels are still low; whereas transcription of another group requires a dramatic increase in BvgA levels generated from autoregulation [56]. The *bipA* gene represents a third group, which is expressed at intermediate concentrations of BvgA because its promoter harbors both high- and low-affinity binding sites for phospho-BvgA [61]. Thus, both the affinity and the position of a BvgA binding site determine the sequential expression of BvgA-regulated genes. This has led to the suggestion that BvgA/BvgS, and potentially other two-component systems, could function as rheostats rather than switches [61].

In addition, autoregulation could provide a threshold for gene activation: only when a signal persists will it promote sufficient levels of phosphorylated response regulator for gene regulation to take place [62]. Although some of the discussed roles for autoregulation are mutually exclusive, they could all be true because autoregulation could be used for different reasons in different systems.

Connections shape the architecture of regulatory networks

Two-component systems can be regarded either as switches whereby changes in phosphorylation of response regulators determine the 'on' or 'off' state [40], or as rheostats [54], whereby the phosphorylation leads to gradual differences in the expression of target genes (see Box 1). Regardless of how a single two-component system functions, the outcome of the interaction between two-component systems can be described with Boolean operators in analogy to neural networks [55].

The independent activation of a particular response regulator by different signals can be described by the Boolean operator OR. For example, high Fe^{3+} OR low Mg^{2+} promote transcription of PmrA-regulated genes. This Boolean operator can also be used to denote the activation of a particular promoter by different response regulators, and for multiple sensors converging onto a single response regulator. The transcriptional control of a two-component system by another might be defined with the Boolean operator AND: both the low phosphate signal controlling PhoB AND the anaerobic conditions activating ResE must be present to ensure ResD/ResE-promoted expression. Likewise, the Boolean operator AND describes the regulatory cascade leading to the transcription of the virulence factor *eps* because all three signals need to be present for transcription to occur. The circuitry of bioluminescence production in *V. harveyi* is indicative of the Boolean operator OR: either AI-1 OR AI-2 can induce expression of LuxO- regulated genes; however, if the synergistic effects of these autoinducers on LuxO-controlled expression is the only physiologically relevant output, the Boolean operator AND would best describe this circuit. Finally, the function of phosphatases in a phosphorelay represents the Boolean operator NOT: the response regulator Spo0A will be phosphorylated if the signals for the Rap phosphatases are NOT present.

In summary, although bacteria use a variety of ways to integrate multiple signals via two-component systems, transcriptional control seems to be the most prevalent way by which the different outputs are connected. Microarray, ChIP-chip, genetic studies and regulatory circuit simulations are uncovering novel regulatory interactions which, ultimately, might make it possible to predict bacterial behavior under different circumstances, provided that we know all signals that are present in a particular environment.

Acknowledgements

We thank four anonymous reviewers for valuable comments on our manuscript. Our research on two-component systems is supported, in part, by grants AI42236 and AI49561 from the National Institutes of Health to E.A.G. who is an Investigator of the Howard Hughes Medical Institute. J.J.E.B has been supported by the Netherlands Organization for Scientific Research (NWO) and is the recipient of an EMBO long-term fellowship.

References

1 Stock, A.M. *et al.* (2000) Two-component signal transduction. *Annu. Rev. Biochem.* 69, 183–215
2 Hoch, J.A. and Silhavy, T.J. (1995) *Two-Component Signal Transduction*, ASM Press
3 Lukat, G.S. *et al.* (1992) Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. *Proc. Natl. Acad. Sci. U. S. A.* 89, 718–722
4 McCleary, W.R. *et al.* (1993) Is acetyl phosphate a global signal in *Escherichia coli*? *J. Bacteriol.* 175, 2793–2798
5 Fabret, C. *et al.* (1999) Two-component signal transduction in *Bacillus subtilis*: how one organism sees its world. *J. Bacteriol.* 181, 1975–1983
6 Zapf, J. *et al.* (2000) A transient interaction between two phosphorelay proteins trapped in a crystal lattice reveals the mechanism of molecular recognition and phosphotransfer in signal transduction. *Structure Fold. Des.* 8, 851–862
7 Tzeng, Y.L. and Hoch, J.A. (1997) Molecular recognition in signal transduction: the interaction surfaces of the Spo0F response regulator

with its cognate phosphorelay proteins revealed by alanine scanning mutagenesis. *J. Mol. Biol.* 272, 200–212
8 Hoch, J.A. and Varughese, K.I. (2001) Keeping signals straight in phosphorelay signal transduction. *J. Bacteriol.* 183, 4941–4949
9 Haldimann, A. *et al.* (1996) Altered recognition mutants of the response regulator PhoB: a new genetic strategy for studying protein-protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* 93, 14361–14366
10 Lawhon, S.D. *et al.* (2002) Intestinal short-chain fatty acids alter *Salmonella typhimurium* invasion gene expression and virulence through BarA/SirA. *Mol. Microbiol.* 46, 1451–1464
11 McCleary, W.R. and Stock, J.B. (1994) Acetyl phosphate and the activation of two-component response regulators. *J. Biol. Chem.* 269, 31567–31572
12 Bourret, R.B. and Stock, A.M. (2002) Molecular information processing: lessons from bacterial chemotaxis. *J. Biol. Chem.* 277, 9625–9628
13 Ames, P. *et al.* (2002) Collaborative signaling by mixed chemoreceptor teams in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 99, 7060–7065
14 Shimizu, T.S. *et al.* (2000) Molecular model of a lattice of signalling proteins involved in bacterial chemotaxis. *Nat. Cell Biol.* 2, 792–796
15 Levit, M.N. *et al.* (2002) Organization of the receptor-kinase signaling array that regulates *Escherichia coli* chemotaxis. *J. Biol. Chem.* 277, 36748–36754
16 Kim, S.H. *et al.* (2002) Dynamic and clustering model of bacterial chemotaxis receptors: structural basis for signaling and high sensitivity. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11611–11615
17 Falke, J.J. and Hazelbauer, G.L. (2001) Transmembrane signaling in bacterial chemoreceptors. *Trends Biochem. Sci.* 26, 257–265
18 Jiang, M. *et al.* (2000) Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*. *Mol. Microbiol.* 38, 535–542
19 LeDeaux, J.R. *et al.* (1995) Different roles for KinA, KinB, and KinC in the initiation of sporulation in *Bacillus subtilis*. *J. Bacteriol.* 177, 861–863
20 Bassler, B.L. (2002) Small talk. Cell-to-cell communication in bacteria. *Cell* 109, 421–424
21 Mok, K.C. *et al.* (2003) *Vibrio harveyi* quorum sensing: a coincidence detector for two autoinducers controls gene expression. *EMBO J.* 22, 870–881
22 Rabin, R.S. and Stewart, V. (1992) Either of two functionally redundant sensor proteins. NarX and NarQ, is sufficient for nitrate regulation in *Escherichia coli* K-12. *Proc. Natl. Acad. Sci. U. S. A.* 89, 8419–8423
23 Rabin, R.S. and Stewart, V. (1993) Dual response regulators (NarL and NarP) interact with dual sensors (NarX and NarQ) to control nitrate- and nitrite-regulated gene expression in *Escherichia coli* K-12. *J. Bacteriol.* 175, 3259–3268
24 Lee, A.I. *et al.* (1999) Signal-dependent phosphorylation of the membrane-bound NarX two-component sensor-transmitter protein of *Escherichia coli*: nitrate elicits a superior anion ligand response compared to nitrite. *J. Bacteriol.* 181, 5309–5316
25 Darwin, A.J. *et al.* (1997) Differential regulation by the homologous response regulators NarL and NarP of *Escherichia coli* K-12 depends on DNA binding site arrangement. *Mol. Microbiol.* 25, 583–595
26 Wanner, B.L. (1992) Is cross regulation by phosphorylation of two-component response regulator proteins important in bacteria? *J. Bacteriol.* 174, 2053–2058
27 Wanner, B.L. (1993) Gene regulation by phosphate in enteric bacteria. *J. Cell Biochem.* 51, 47–54
28 Fisher, S.L. *et al.* (1995) Cross-talk between the histidine protein kinase VanS and the response regulator PhoB, Characterization and identification of a VanS domain that inhibits activation of PhoB. *J. Biol. Chem.* 270, 23143–23149
29 Fisher, S.L. *et al.* (1996) Kinetic comparison of the specificity of the vancomycin resistance VanS for two response regulators, VanR and PhoB. *Biochemistry* 35, 4732–4740
30 Matsubara, M. *et al.* (2000) Tuning of the porin expression under anaerobic growth conditions by His-to-Asp cross-phosphorelay through both the EnvZ-osmosensor and ArcB-anaerosensor in *Escherichia coli*. *Genes Cells* 5, 555–569
31 McCleary, W.R. (1996) The activation of PhoB by acetylphosphate. *Mol. Microbiol.* 20, 1155–1163
32 Chamnongpol, S. and Groisman, E.A. (2000) Acetyl phosphate-dependent activation of a mutant PhoP response regulator that functions independently of its cognate sensor kinase. *J. Mol. Biol.* 300, 291–305
33 Alves, R. and Savageau, M.A. (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
34 Verhamme, D.T. *et al.* (2002) Investigation of *in vivo* cross-talk between key two-component systems of *Escherichia coli*. *Microbiol.* 148, 69–78
35 Allaway, D. *et al.* (1995) NtrBC-dependent expression from the Rhizobium meliloti dctA promoter in *Escherichia coli*. *FEMS Microbiol. Lett.* 128, 241–245
36 Perego, M. (1998) Kinase-phosphatase competition regulates Bacillus subtilis development. *Trends Microbiol.* 6, 366–370
37 Lazazzera, B.A. and Grossman, A.D. (1998) The ins and outs of peptide signaling. *Trends Microbiol.* 6, 288–294
38 Perego, M. and Brannigan, J.A. (2001) Pentapeptide regulation of aspartyl-phosphate phosphatases. *Peptides* 22, 1541–1547
39 Perego, M. (2001) A new family of aspartyl phosphate phosphatases targeting the sporulation transcription factor Spo0A of *Bacillus subtilis*. *Mol. Microbiol.* 42, 133–143
40 Msadek, T. (1999) When the going gets tough: survival strategies and environmental signaling networks in *Bacillus subtilis*. *Trends Microbiol.* 7, 201–207
41 Kwon, O. *et al.* (2000) Phosphorelay as the sole physiological route of signal transmission by the arc two-component system of *Escherichia coli*. *J. Bacteriol.* 182, 3858–3862
42 Jourlin, C. *et al.* (1997) Transphosphorylation of the TorR response regulator requires the three phosphorylation sites of the TorS unorthodox sensor in *Escherichia coli*. *J. Mol. Biol.* 267, 770–777
43 Perraud, A.L. *et al.* (1998) Specificity of the BvgAS and EvgAS phosphorelay is mediated by the C-terminal HPt domains of the sensor proteins. *Mol. Microbiol.* 27, 875–887
44 Chang, C.H. *et al.* (1996) Pleiotropic phenotypes caused by genetic ablation of the receiver module of the *Agrobacterium tumefaciens* VirA protein. *J. Bacteriol.* 178, 4710–4716
45 Ogino, T. *et al.* (1998) An *Escherichia coli* protein that exhibits phosphohistidine phosphatase activity towards the HPt domain of the ArcB sensor involved in the multistep His-Asp phosphorelay. *Mol. Microbiol.* 27, 573–585
46 Matsubara, M. and Mizuno, T. (2000) The SixA phospho-histidine phosphatase modulates the ArcB phosphorelay signal transduction in *Escherichia coli*. *FEBS Lett.* 470, 118–124
47 Birkey, S.M. *et al.* (1998) Pho signal transduction network reveals direct transcriptional regulation of one two-component system by another two-component regulator: *Bacillus subtilis* PhoP directly regulates production of ResD. *Mol. Microbiol.* 30, 943–953
48 Oshima, T. *et al.* (2002) Transcriptome analysis of all two-component regulatory system mutants of *Escherichia coli* K-12. *Mol. Microbiol.* 46, 281–291
49 Kobayashi, K. *et al.* (2001) Comprehensive DNA microarray analysis of *Bacillus subtilis* two-component regulatory systems. *J. Bacteriol.* 183, 7365–7370
50 Huang, J. *et al.* (1995) A complex network regulates expression of eps and other virulence genes of *Pseudomonas solanacearum*. *J. Bacteriol.* 177, 1259–1267
51 Garg, R.P. *et al.* (2000) Multicomponent transcriptional regulation at the complex promoter of the exopolysaccharide I biosynthetic operon of *Ralstonia solanacearum*. *J. Bacteriol.* 182, 6659–6666
52 Mouslim, C. and Groisman, E.A. (2003) Control of the *Salmonella ugd* gene by three two-component regulatory systems. *Mol. Microbiol.* 47, 335–344
53 Kox, L.F. *et al.* (2000) A small protein that mediates the activation of a two-component system by another two-component system. *EMBO J.* 19, 1861–1872
54 Russo, F.D. and Silhavy, T.J. (1993) The essential tension: opposed reactions in bacterial two-component regulatory systems. *Trends Microbiol.* 1, 306–310
55 Hellingwerf, K.J. *et al.* (1995) Signal transduction in bacteria: phospho-neural network(s) in *Escherichia coli*? *FEMS Microbiol. Rev.* 16, 309–321

56 Scarlato, V. *et al.* (1990) Positive transcriptional feedback at the bvg locus controls expression of virulence factors in Bordetella pertussis. *Proc. Natl. Acad. Sci. U. S. A.* 87, 6753–6757
57 Soncini, F.C. *et al.* (1995) Transcriptional autoregulation of the *Salmonella typhimurium* phoPQ operon. *J. Bacteriol.* 177, 4364–4371
58 Kato, A. *et al.* (1999) Molecular characterization of the PhoP-PhoQ two-component system in *Escherichia coli* K-12: identification of extracellular Mg2 + -responsive promoters. *J. Bacteriol.* 181, 5516–5520
59 Hoffer, S.M. *et al.* (2001) Autoamplification of a two-component regulatory system results in "learning" behavior. *J. Bacteriol.* 183, 4914–4917
60 Cotter, P.A. and DiRita, V.J. (2000) Bacterial virulence gene regulation: an evolutionary perspective. *Annu. Rev. Microbiol.* 54, 519–565
61 Deora, R. *et al.* (2001) Diversity in the Bordetella virulence regulon: transcriptional control of a Bvg-intermediate phase gene. *Mol. Microbiol.* 40, 669–683
62 Raivio, T.L. *et al.* (1999) The Cpx envelope stress response is controlled by amplification and feedback inhibition. *J. Bacteriol.* 181, 5263–5272

ELSEVIER

Cyclic di-GMP as a second messenger

Ute Römling[1] and Dorit Amikam[2]

In many bacteria bis-(3′,5′)-cyclic dimeric guanosine monophosphate (c-di-GMP) signaling determines the timing and amplitude of complex biological processes from biofilm formation and virulence to photosynthesis. Thereby, the tightly regulated temporal and spatial activity patterns of GGDEF and EAL domain proteins, which synthesize and degrade c-di-GMP, respectively, are currently being resolved. Although details of the mechanisms of c-di-GMP signaling are not yet determined, the recent presentation of PilZ as a candidate c-di-GMP binding-domain opens the field for experimental investigations. Besides its role as an intracellular signaling molecule in bacteria, c-di-GMP also acts as an intercellular signaling molecule between prokaryotes and also has effects in eukaryotes that could provide a perspective in cancer treatment.

Addresses
[1] Karolinska Institutet, Microbiology and Tumor Biology Center (MTC), Box 280, SE-171 77 Stockholm, Sweden
[2] Department of Biotechnology and Environmental Sciences, Tel-Hai Academic College, Tel-Hai, and Sharett Institute of Oncology, Hadassah University Medical Center, Ein-Kerem, Jerusalem, Israel

Corresponding author: Römling, Ute (ute.romling@ki.se)

Current Opinion in Microbiology 2006, **9**:218–228

This review comes from a themed issue on
Cell regulation
Edited by Werner Goebel and Stephen Lory

Available online 10th March 2006

1369-5274/$ – see front matter

DOI 10.1016/j.mib.2006.02.010

Introduction

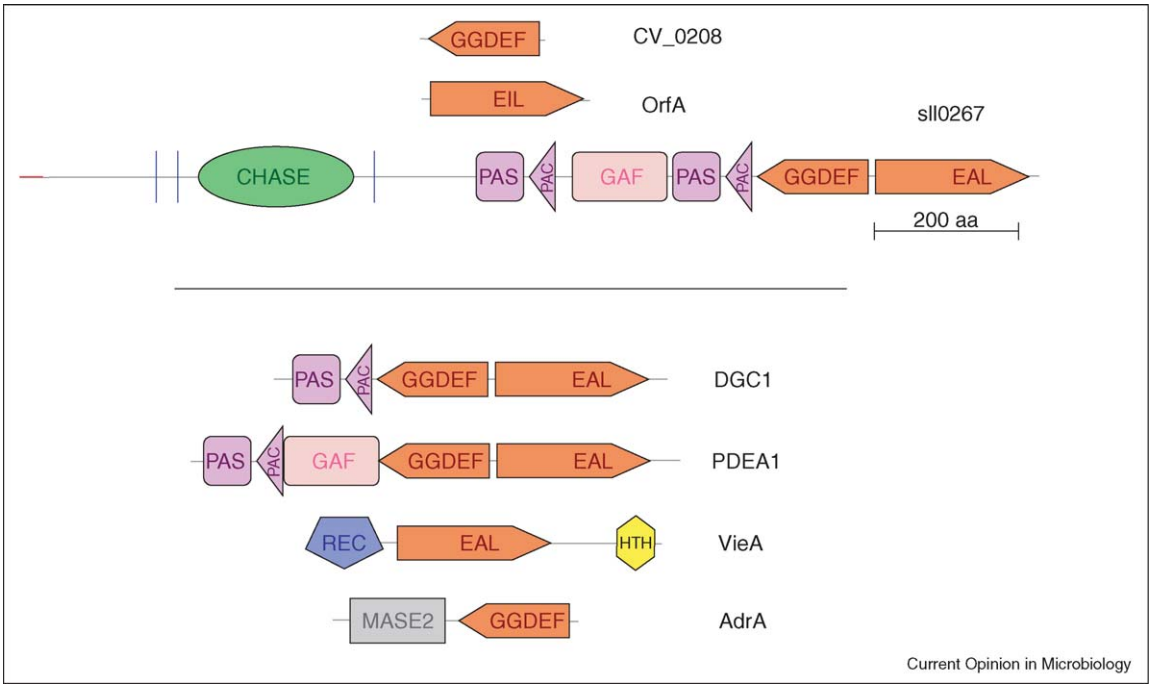
The molecular structure of bis-(3',5')-cyclic dimeric guanosine monophosphate (cyclic di-GMP or c-di-GMP) and its presence in bacteria was discovered 19 years ago by Benziman *et al.* [1] after long investigations to detect the allosteric activator of the cellulose synthase in the fruit-degrading bacterium *Gluconacetobacter xylinus*. In a search for regulators of bacterial development 11 years ago, Hecht and Newton [2] described the response regulator PleD required for subsequent loss of motility and morphological changes in the motile swarmer cell of the aquatic bacterium *Caulobacter crescentus*. The unorthodox response regulator PleD did not harbor a C-terminal DNA-binding domain, but rather a novel output domain that was termed GGDEF after a motif that is highly conserved in six additional proteins of unknown function from various bacteria. Three years later the function of the GGDEF domain in bacterial development was partly uncovered when Benziman's group discovered that each enzyme that performed either biosynthesis or breakdown of c-di-GMP in *G. xylinus* contained highly homologous GGDEF and EAL domains [3]. When large-scale sequencing showed the abundance of GGDEF and EAL domain proteins, which constitute two of the largest super families of proteins present in the bacterial kingdom [4], the allosteric activator of the cellulose synthase c-di-GMP was finally recognized as a general secondary signaling molecule in bacteria. Concurrently, more phenotypes for genes encoding GGDEF domain proteins were characterized as a result of the interest in the genetic basis of the multicellular behavior of bacteria, commonly termed biofilm formation [5,6,7•].

In this review, we describe the significant recent progress of knowledge concerning c-di-GMP metabolism and signaling. The enzymatic activities of the GGDEF and EAL domain have been characterized, the impact of the c-di-GMP signaling pathway in multicellular behavior and virulence has been well established, and details of signaling through GGDEF and EAL domain proteins are becoming apparent. The molecular mechanisms of c-di-GMP action still remain to be resolved, but the recent proposal of a c-di-GMP binding-domain paves the way for experimental investigations.

The role of GGDEF and EAL domains in c-di-GMP signaling

Although known to be involved in c-di-GMP turnover for some time, the precise role of the GGDEF and EAL domains in c-di-GMP turnover remained obscure. Previously characterized di-guanylate cyclases and phosphodiesterases from *G. xylinus* have a similar domain structure with highly homologous GGDEF and EAL domains [3]. Recently, the predicted enzymatic activities of these two domains were experimentally verified. Purified GGDEF domain proteins that were derived from bacterial species originating from diverse branches of the phylogenetic tree were found to possess di-guanylate cyclase activity (Figure 1) [8•,9,10]. When tested, substrate specificity was restricted to GTP and the product outcome to c-di-GMP [8•,10], demonstrating that c-di-GMP is produced by a hyperthermophilic marine bacterium of the Thermotoga branch, as well as γ-proteobacteria such as *Escherichia coli*, a commensal of humans. As with the well-characterized adenylate cyclases, diguanlylate cyclases function as dimers or trimers [10,11], whereby oligomerization is an intrinsic property of the GGDEF domain [10]. Although able to oligomerize, an individual GGDEF domain

Figure 1



Domain structure of GGDEF and EAL domain proteins. The simplest proteins consist of only a GGDEF (as illustrated by CV_0208 from *Chromobacterium violaceum*) or an EAL (as illustrated by OrfA from plasmid pLEB513 from *Lactococcus lactis*) domain. The GGDEF domain is approximately 180 amino acids long, whereas the EAL domain consists of 260 residues. One of the longest GGDEF and EAL domain proteins is the 1578 amino acid long sll0267 from *Synechocystis* species PCC6803. Proteins DGC1 (*G. xylinus*), CC3396 (*C. crescentus*), AdrA (*S.* Typhimurium) and VieA (*V. cholerae*) were previously investigated and are described in the text. Abbreviations: CHASE, cyclase/histidine kinases-associated sensory domain; GAF, domain present in phytochromes and cGMP-specific phosphodiesterases; PAS/PAC, Per (periodic clock protein), Arnt (aryl hydrocarbon receptor nuclear translocator protein) and Sim (single-minded protein) domain; MASE2, membrane-associated sensor type 2; Rec, cheY-homologous receiver domain.

showed only residual enzymatic activity, but a higher unspecific activity in hydrolyzing GTP to GDP [10]. Because GGDEF and EAL domains do not usually stand alone, but are part of proteins with multiple domains (Figure 1) [5,7•], this finding underscored the modulation of the di-guanylate cyclase activity of the GGDEF domain through signal sensing. In the comparatively simple GGDEF domain proteins investigated so far, phosphorylation of an N-terminal receiver domain or the presence of a GAF domain stimulated the activity by two orders of magnitude [8•,10], which ensures dynamic adjustment of c-di-GMP signaling.

Recent studies have also unambiguously assigned phosphodiesterase activity to the EAL domain with consensus residues [12•,13•,14,15]. Thereby, c-di-GMP is hydrolyzed to pGpG, whereas hydrolysis of pGpG to GMP is significantly slower and not considered a physiologically relevant activity of the domain. As reported before [1], phosphodiesterase activity required the divalent cation Mg^{2+} or Mn^{2+}, but was strongly inhibited by Ca^{2+} or Zn^{2+} ions. As with GGDEF domain proteins, N-terminal signaling and sensing domains modulated the substrate affinity of the EAL domain [12,13], but phosphodiesterase activity was also observed without phosphorylation of an N-terminal receiver domain [14].

The role of unconventional GGDEF and EAL domains

Alignment of GGDEF and EAL domains identified a set of consensus residues clustering in several regions throughout the domains [4,7•]. However, extrapolation of the exact function of additional GGDEF and EAL domains from previous investigations might not be entirely justified, because substrate specificity of nucleotide cyclases and phosphodiesterases can be readily manipulated by exchange of one or two amino acids [16,17]. GGDEF and EAL domains from proteins with experimentally demonstrated *in vitro* or *in vivo* di-guanylate and c-di-GMP phosphodiesterase activity possessed the most invariable residues in agreement with the consensus profile for the domains [8•], although the basic requirements for both activities are not entirely clear yet [13•]. For example, the GGDEF domains of three phosphodiesterases from *G. xylinus* possess a conserved GGDEF motif and deviate from the consensus sequence only in the functionally assigned lysine at position 20, which is replaced by arginine. STM3388

apparently has functional GGDEF and EAL domains, however di-guanylate cyclase activity was assigned to this protein [18•,19]. However, first, functions of GGDEF and EAL domains with unconventional amino acid replacements need to be resolved. Biochemical characterization has shown that in *C. crescentus* the GGDEF domain of CC3396 (Figure 1) with an unorthodox GEDEF motif and several additional amino acid changes in residues conserved in functional di-guanylate cyclases lacked detectable di-guanylate cyclase activity, but retained GTP-binding activity [12•]. Binding of GTP to the GEDEF domain enhanced the phosphodiesterase activity of the C-terminal EAL domain by lowering the substrate affinity for c-di-GMP. Hence, GEDEF is a sensor domain responding to the intracellular GTP concentration that serves as a physiological signal of the nutrient status, at least in some bacteria. However, conservation of the GGDEF motif is not a prerequisite for di-guanylate cyclase activity, because GGDEF domains in functional phosphodiesterases have retained this characteristic GGDEF motif [3].

Consistent with the EAL-like domain structure of STM1344, an apparent phosphodiesterase activity was observed when the cytoplasmic c-di-GMP concentration of the *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium) wild type strain was compared with the STM1344 knockout and the STM1344 overexpressing strain [20]. This finding stimulated the authors to rename STM1344 to *cdgR* (c-di-GMP regulator). In contradiction, other investigations of the effects of STM1344 on the cellular concentration of c-di-GMP observed apparent di-guanylate cyclase activity for STM1344 (R Simm and U Römling, unpublished). Central invariant residues of the EAL consensus sequence are not conserved in STM1344. Thus, the role of STM1344 in c-di-GMP metabolism and signaling is not clear yet.

C-di-GMP signaling regulates sessility versus motility

Low intracellular c-di-GMP levels, in the range of 0.6 pmol mg^{-1} cells, have been found in several bacteria [12•,18•,21,22]. Consequently, when produced in physiological amounts, the c-di-GMP concentration is not expected to disturb GTP metabolism, because the GTP concentration is 100-fold higher than the c-di-GMP concentration [23••]. However, the intracellular concentration of c-di-GMP can be readily modulated by over expression of GGDEF and EAL domain proteins, thus achieving saturation and depletion of the compartmentalized c-di-GMP pool(s) that are created by multiple GGDEF or EAL domain proteins in the cell [23••,24]. Using this strategy to control the c-di-GMP concentration, a correlation between c-di-GMP levels and bacterial behavior could be established (Figure 2a) [23••]. In *S.* Typhimurium high c-di-GMP concentrations stimulated biofilm formation, the production of adhesive surface organelles, such as cellulose and curli fimbriae, and suppressed motility [18•,23••]. These consequences of high c-di-GMP concentrations on bacterial behavior are most precisely described as locking the cells in a sessile state. In contrast, low c-di-GMP concentrations inhibited biofilm formation and the production of adhesive surface organelles; and stimulated swimming and swarming motility. The correlation between high c-di-GMP concentrations and biofilm formation, and low c-di-GMP concentrations and motility, has been well established by over expression of GGDEF and EAL domain proteins in many bacteria [23••,25].

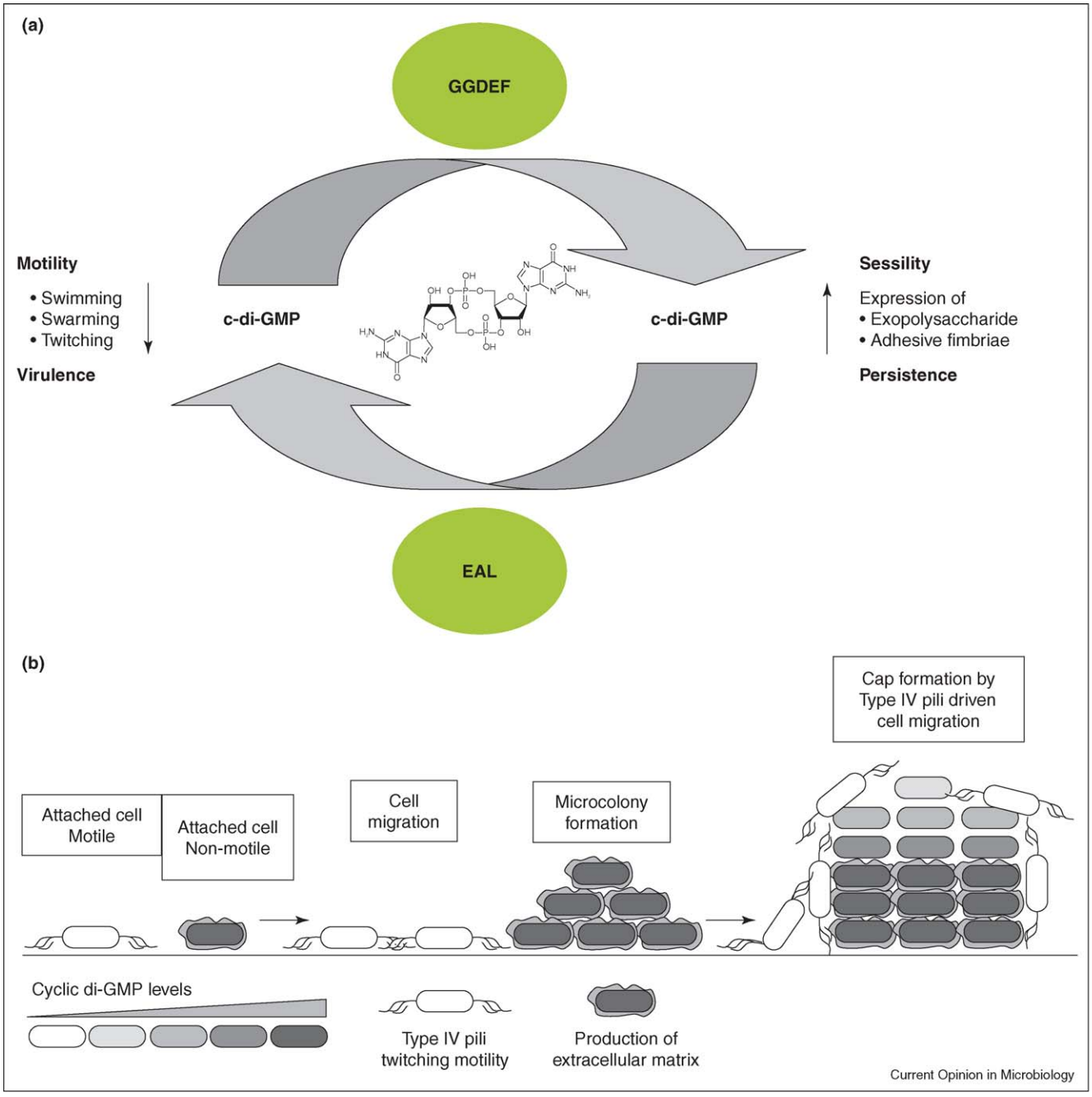
A traditional view is that in a cell community the transition is between biofilm (multicellular) behavior and planktonic (single) cells. However, our more precise interpretation of the data above takes into consideration that c-di-GMP mediates the transition between sessility and motility, even in a single cell. This regulatory concept of c-di-GMP signaling resolves these apparent. Not all multicellular behavior and production of an extracellular matrix has been correlated with high c-di-GMP concentrations. For example, swarming motility, which occurs at low c-di-GMP concentrations, is a multicellular behavior dependent of cell–cell interactions, morphological differentiation into elongated swarmer cells and production of 'slime' [26].

Interestingly, experimental observations have shown that development of a sophisticated mushroom-shaped biofilm architecture requires highly regulated motility by flagella or type IV pili that is regulated at the level of individual cells (Figure 2b) [27]. Thereby, non-motile bacteria form the mushroom stalks, while migrating bacteria form the mushroom caps by climbing the stalks. The regulatory concept of c-di-GMP signaling predicts the c-di-GMP signaling to be substantially different in individual cells in a biofilm depending on whether cells move or are embedded in an extracellular matrix. Thus, c-di-GMP signaling takes place on a nano-scale determined by the microenvironment around the cell [7•,23••].

Flexibility of c-di-GMP signaling

High variability in the regulatory pattern of c-di-GMP signaling is observed in unrelated species, but also in closely related species and in individual strains of the same species. When it comes to the details, motility is affected in various ways by c-di-GMP signalling. In different species, motility is regulated by distinct GGDEF and EAL domain proteins. What's more, differential regulation of motility by homologous GGDEF and EAL domain proteins is observed in closely related organisms. A constitutive swimming phenotype is mediated by the lack of certain GGDEF domain proteins in *Caulobacter crescentus* and *Pseudomonas putida*. The response regulator PleD, which has a diguanylate cyclase output domain, is required for flagella shedding in the later growth phase of *C. crescentus* [28]. MorA, a complex

Figure 2



Regulator concepts of c-di-GMP metabolism and signalling **(a)** Regulatory concept of c-di-GMP metabolism and signalling on the population level. GGDEF domains with consensus residues synthesize c-di-GMP from two molecules of GTP. EAL domains with consensus residues cleave the c-di-GMP molecule into pGpG. High c-di-GMP levels promote sessility aided by the production of adhesive extracellular matrix components such as polysaccharides (i.e. cellulose [56] and *Vibrio* polysaccharide (VPS) [24]) and fimbriae [curli fimbriae of *S. typhimurium* [18•] and Cup fimbriae of *P. aeruginosa* [43] (A Meissner *et al*., abstract MPP025, 2nd Common Congress of the DGHM and VAAM, Göttingen, September 2005)]. Environmental and host persistence is also predicted to be promoted by high c-di-GMP levels [57]. Conversely, low c-di-GMP levels promote motility behavior (swimming, swarming and twitching motility [23••]) and virulence [31]. **(b)** Regulatory concept of c-di-GMP signaling at the individual cell level. Cells adherent to a surface can display various c-di-GMP concentrations depending on whether they display twitching motility or produced an adhesive extracellular matrix. The c-di-GMP signaling concept was developed based on observations made by Klausen *et al*. [27], who described the requirement of movement of individual cells for the development of a sophisticated architecture in a biofilm.

GGDEF and EAL domain protein with various sensory domains, suppressed flagella biosynthesis in the early growth phase of *P. putida* and affected flagella number leading to hyper-flagellated cells in the transition phase [29]. However, MorA has no effect on flagellation in *Pseudomonas aeruginosa* when observed under the same conditions.

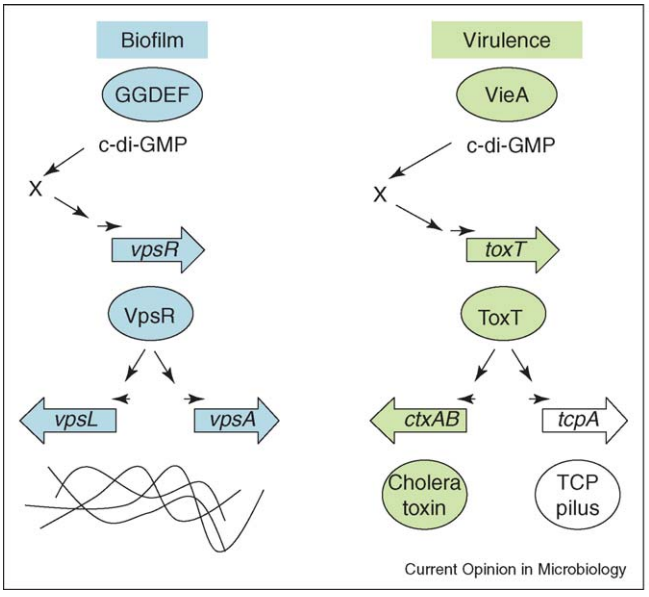
When it comes to virulence, VieA, a response regulator with c-di-GMP phosphodiesterase output activity, is required for full colonization of the gut of the infant mouse by a classic biotype of *Vibrio cholerae*, but is not required for colonization by the El Tor biotype [30•]. These data provide compelling evidence that c-di-GMP signaling fine-tunes the timing and amplitude of complex biosynthesis pathways, including motility, virulence and biofilm formation as well as processes such as photosynthesis [31].

The genomic context of GGDEF and EAL domain proteins reflects the flexibility of c-di-GMP signalling. The EAL domain protein PvrR contributes to virulence of *P. aeruginosa* strain PA14 in the mouse burn model and in plants [32]. The gene encoding PvrR, which is possibly required for the exceptionally broad virulence potential of PA14, is located on a genetic island. GGDEF and EAL domain proteins are also found encoded for on plasmids and transposons [4,33]. Genes encoding GGDEF and EAL domain proteins are predicted to be frequently individual transcriptional entities (Simm R and Römling U, unpublished), which indicates the highly adaptable regulatory flexibility and an exchange potential at the level of the individual GGDEF or EAL domain protein.

Role of c-di-GMP signaling in virulence

As well as its role in biofilm formation, c-di-GMP metabolism plays a major role in virulence. Screens for *in vivo* expressed genes discovered GGDEF and EAL domain proteins in *V. cholerae* and *Vibrio vulnificus* [30•,34]. The VieSAB three-component system is required for maximal expression of *ctxAB*, which encodes cholera toxin, *in vivo* and during growth under *in vitro* conditions in *V. cholerae* [30•]. Expression of VieA (Figure 1), a c-di-GMP specific phosphodiesterase, ensures maintenance of low c-di-GMP levels under certain growth conditions. A single amino acid exchange in the EAL motif (E170A) of VieA abolishes its enzymatic activity and raises the intracellular levels of c-di-GMP sixfold [14,24]. Investigation of the regulatory network leading to transcription of *ctxA* showed that the decrease in c-di-GMP levels mediated by VieA positively affected transcription of ToxT, a transcriptional regulator that directly activates expression of *ctxAB* (Figure 3). Other than this master regulator of virulence gene expression, none of the regulatory factors upstream in the regulatory cascade was affected by modulation of the c-di-GMP concentration. Conversely, elevated c-di-GMP levels in the *vieA* knockout mutant stimulated biofilm formation through enhanced transcription of *Vibrio* exopolysaccharide synthesis (*vps*) genes by the transcriptional activator VpsR [24]. This inverse regulation of biofilm formation and virulence (Figure 3) suggests that persistence in the environment and survival within the host are inversely coupled through c-di-GMP signaling.

Figure 3



Biofilm formation and virulence are inversely regulated by c-di-GMP concentrations in *Vibrio cholerae* [24,31]. The intracellular c-di-GMP concentration is regulated by the phosphodiesterase VieA and an unknown di-guanylate cyclase (GGDEF). High c-di-GMP concentrations stimulate biofilm formation by activation of *vpsR* transcription. VpsR is a NtrC-like positive regulator of *vps* biosynthesis genes *vpsA–vpsK* and *vpsL–vpsQ*. Expression of cholera toxin is repressed. Low c-di-GMP concentrations enhance virulence by activation of *toxT* transcription, which subsequently leads to enhanced expression of cholera toxin. Expression of TCP pili is not affected in strains with non-functional *vieA,* although ToxR, a virulence factor regulator of *V. cholerae* also activates TCP expression.

In *Bordetella pertussis* suppression of genes of unknown functions by the EAL domain protein BvgR significantly aids virulence in the mouse aerosol challenge model [35]. Genes repressed by BvgR might be required for *B. pertussis* to persist in or to survive outside of the host. Therefore inverse regulation of environmental survival and host virulence by c-di-GMP signaling is not restricted to *V. cholerae*, but might be a general regulatory concept of c-di-GMP signaling.

An *in vivo* screen for genes required for *S.* Typhimurium to resist oxidative killing by phagocytes recovered STM1344, an EAL-domain like protein, as the sole output [20]. The participation of STM1344 in resistance to

phagocyte oxidase was confirmed by *in vitro* susceptibility of the STM1344 mutant to hydrogen peroxide. However, the STM1344 mutant killed macrophages earlier and was more cytotoxic than the wild type. Cytotoxicity was dependent on bacterial uptake and in part on SipB, an effector protein of type III secretion, which is required for rapid cell death of macrophages [36]. Although over expression of STM1344 partly suppressed secretion of SipB, the molecular mechanism of rapid killing of macrophages through the lack of STM1344 in *S.* Typhimurium is not clear yet.

The complexity of c-di-GMP signaling

Because most bacteria harbor more than one GGDEF or EAL domain protein (the top score is held by *Vibrio vulnificus* with 66 GGDEF and 33 EAL domain proteins), a tight regulation of c-di-GMP signaling by complex spatial and temporal activity patterns of GGDEF and EAL domain proteins has been predicted [7•].

The rdar morphotype is a multicellular behavior of *S.* Typhimurium characterized by the expression of the polymeric extracellular matrix components cellulose and curli fimbriae [33]. Biosynthesis of cellulose and curli fimbriae requires activation by CsgD, a transcriptional regulator of the LuxR superfamily (Figure 4). At least three GGDEF domain proteins are active during development of the rdar morphotype [18•,37]. During the early growth phase on agar plates, the complex GGDEF–EAL domain protein STM2123 was required for expression of CsgD [18•]. The major contribution of STM3388, another complex GGDEF–EAL domain protein, to CsgD expression occurred at a later growth stage of rdar morphotype development. The effects of STM2123 and STM3388 are additive; in a double knockout, 42% to 69% reduction of CsgD expression was observed depending on the growth phase. As a consequence of CsgD downregulation by STM2123 and STM3388 in the single and double knockout mutants, expression of cellulose and curli fimbriae is diminished [18•]. However, the c-di-GMP signaling network in rdar morphotype development is more complex. CsgD subsequently regulates the expression of AdrA (Figure 1), another GGDEF domain protein [37]. AdrA is required for activation of cellulose biosynthesis by *S.* Typhimurium grown on agar plates [37]. AdrA also partly regulated the biosynthesis of curli fimbriae downstream of CsgD expression [18•]. This finding indicated that AdrA might be responsible for biosynthesis events on the membrane and connects, for the first time, the biosynthesis pathways for cellulose and curli fimbriae downstream of CsgD.

Despite the substantial effects of STM2123 and STM3388 on CsgD expression, c-di-GMP concentration was not significantly altered throughout the growth phase in a double mutant as compared with the wild type. AdrA is responsible for over 60% of the cellular c-di-GMP produced at 16 hours of growth [18•]. Despite the major contribution of AdrA to the intracellular c-di-GMP concentration, AdrA hardly affects the expression of CsgD. This experimental finding demonstrated for the first time the restricted temporal and spatial compartmentalization of the c-di-GMP signaling.

Specificity of c-di-GMP signaling

Although not rigorously tested, anecdotal evidence suggested that di-guanylate cyclase activities of GGDEF domain proteins are variable and that variable amounts of c-di-GMP are produced by different GGDEF domain proteins [18•,22]. Thereby the activity of GGDEF domain proteins is also regulated at the level of gene expression. A promoter up-mutation, which leads to a threefold higher CsgD expression, raises the c-di-GMP concentration threefold as a result of the over expression of AdrA (Simm R and Römling U, unpublished), thus making AdrA the dominant contributor to the c-di-GMP concentration in the cell. In this situation a feedback loop of AdrA on CsgD expression is observed (Figure 4).

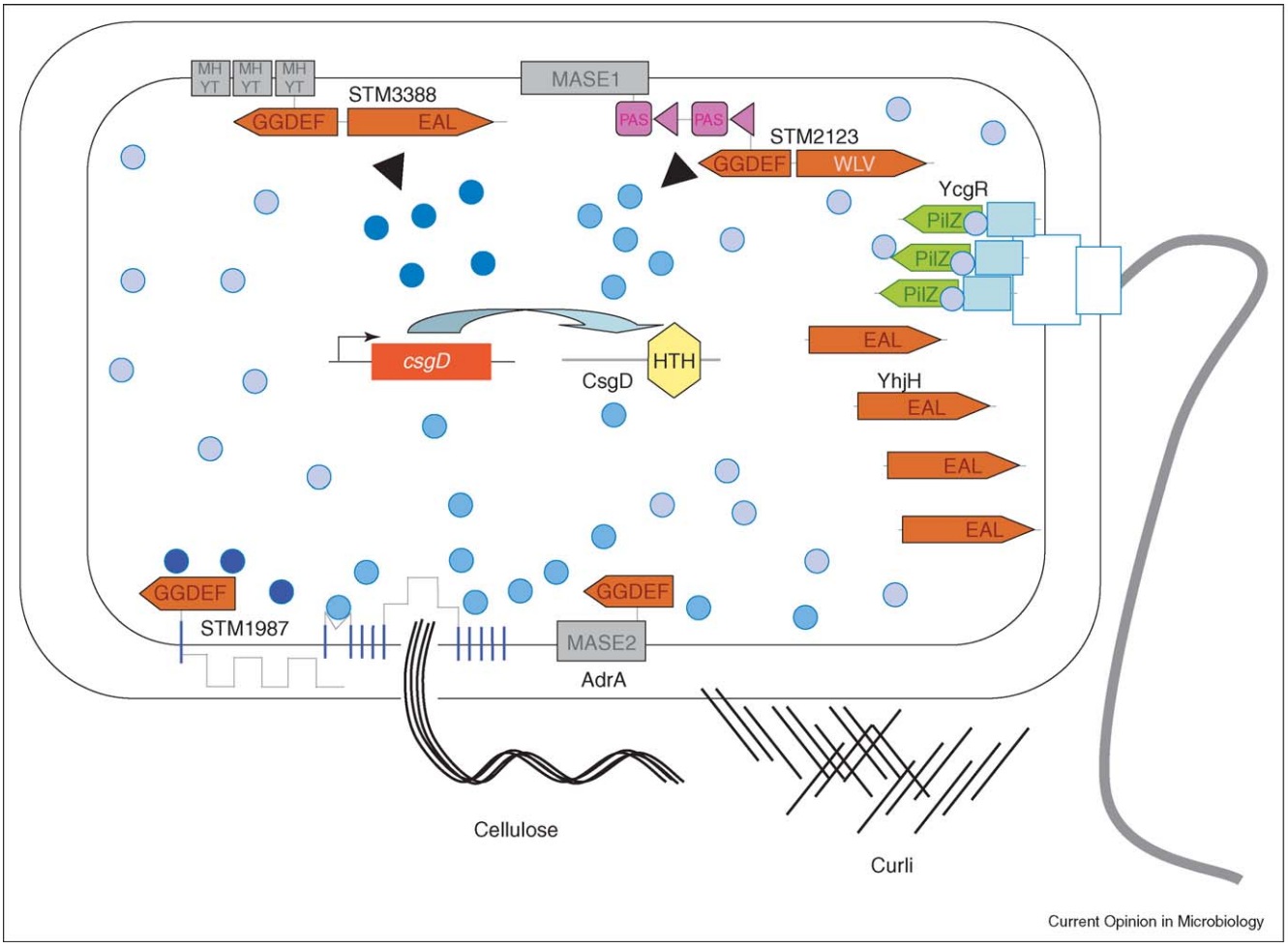
Consequently, over expression of GGDEF domain proteins and enhanced c-di-GMP production fills up the various spatially separated c-di-GMP pools resulting in the observation of the most obvious phenotype, pronounced cellulose production in *S.* Typhimurium [19]. This biological consequence of c-di-GMP action *in vivo* is a rapid and probably the most sensitive test for di-guanylate cyclase activity because the total cellular c-di-GMP levels might rise only twofold to fivefold when respective proteins are over expressed [18•,22], whereas knockout of GGDEF or EAL domain proteins might not necessarily lead to a phenotype [19].

1. PilZ — first proposal of a c-di-GMP binding domain

Allosteric activation of the cellulose synthase by c-di-GMP in membrane fractions provided the first evidence for a c-di-GMP binding site, whereby the activation activity was purified in close association with the catalytic subunit of the cellulose synthase BcsA [1]. Early experimental work based on cross-linking studies indicated BcsB, the second gene in the cellulose biosynthesis operon *bcs*, to be the c-di-GMP-binding protein [38]. However, recent re-evaluation based on the topology of BcsB and phylogenetic distribution of BcsB in comparison with GGDEF and EAL domain proteins made BcsB an unlikely candidate for the c-di-GMP-binding protein [7•].

Recently, Amikam and Galperin presented the 118 amino acid long PilZ domain as a candidate c-di-GMP binding domain [39••]. More than 600 PilZ domains were found, whereby the phylogenetic distribution of PilZ domains grossly follows those of the GGDEF and EAL domains. Most importantly, PilZ domain proteins are part of

Figure 4



Model illustrating the role of c-di-GMP signaling in rdar morphotype development and swimming motility in *S*. Typhimurium. Rdar morphotype expression is stimulated with c-di-GMP production. The GGDEF and EAL domain proteins STM2123 and STM3388 are di-guanylate cyclases producing c-di-GMP [18•]. As a consequence, c-di-GMP enhances the level of expression of CsgD by an unknown mechanism. Production of c-di-GMP by AdrA primarily activates biosynthesis of cellulose [23••,37] and partially contributes to the production of curli fimbriae downstream of CsgD expression [18•], when cells are grown on agar plates. STM1987 activates cellulose biosynthesis in a nutrient-deficient medium [19]. C-di-GMP presumably binds to the PilZ domain of the cellulose synthase BcsA, thus activating cellulose biosynthesis [39••]. Feedback regulation of c-di-GMP produced by AdrA on CsgD expression was seen at a higher AdrA expression level, but not in the wild type strain *S*. Typhimurium UMR1. c-di-GMP signaling also plays a role in swimming motility: expression of the EAL domain protein YhjH is required to stimulate swimming motility in a *hns* mutant of *E*. *coli* and in wild type *S*. Typhimurium [23••,42]. Swimming motility was also stimulated by deletion of *ycgR* which encodes a PilZ domain protein. As a hypothesis, a YcgR–c-di-GMP complex inhibits the motor function of the flagella, whereas degradation of cylic-di-GMP by YhjH relieves inhibition of motility by YcgR. The different shades of blue circles represent c-di-GMP produced by different GGDEF or EAL domain proteins. The EAL domain of STM2123 with non-conserved residues is assigned the same symbol as the EAL domain with consensus sequences, but amino acid deviations are indicated (WLV instead of EAL). Other domains and abbreviations: HTH, helix-turn-helix DNA binding motif of GerE family; MASE1, membrane-associated sensor; MASE2, membrane-associated sensor; MHYT, integral membrane sensory domain with conserved MHYT amino acid pattern; PAS/PAC, Per (periodic clock protein), Arnt (aryl hydrocarbon receptor nuclear translocator protein) and Sim (single-minded protein) domain.

biosynthesis pathways that are regulated by c-di-GMP. Location of a PilZ domain at the C-terminal end of bacterial cellulose synthases BcsA is consistent with observations of the late Benziman that the majority of c-di-GMP was membrane-bound to the cellulose synthase of 600 kDa and to a 200 kDa protein complex [21], which could correspond to a dimer of BcsA or a BcsA–BcsB fusion protein [40].

Type IV pili mediated twitching motility is another phenotype that is regulated by c-di-GMP (Figure 2) [23••]. The name-giving protein PilZ is a stand-alone protein in the type IV pili biosynthesis pathway without assigned function. Bacterial cells lacking PilZ or FimX, a multidomain GGDEF–EAL domain protein [41], showed similar phenotypes; significantly reduced twitching motility, no intact surface pili, but relatively normal

levels of intracellular pilin fimbrial subunits. Inactivation of *ycgR*, encoding a protein with C-terminal PilZ domain, reversed the motility defect of an *hns* (encoding an abundant nucleoid-associated protein) mutant in *E. coli* [42]. Because over expression of YhjH, a stand-alone EAL domain protein with phosphodiesterase activity, had the same effect as knocking out *ycgR*, the data are consistent with the hypothesis that YcgR serves as a cellular receptor for c-di-GMP, which either protects c-di-GMP from degradation and/or functions as a YcgR–c-di-GMP complex in protein–protein interactions (Figure 4).

C-di-GMP signaling does not stand alone

The c-di-GMP signaling pathway is interconnected with other signaling pathways and responds to environmental cues [5,7•]. C-di-GMP signaling is integrated into phosphotransfer signaling. Several GGDEF and EAL domain proteins are response regulators, the activity of which is determined by phosphorylation of a conserved aspartate in the N-terminal receiver domain [8•,9,10,43].

Deletion of Dos, a GGDEF and EAL domain protein with phosphodiesterase activity [13•], leads to a 26-fold increase in cAMP levels when *E. coli* was grown under aerobic, but not under anaerobic, conditions [44]. Because Dos has been demonstrated to harbor c-di-GMP specific phosphodiesterase activity [13•], the cAMP and c-di-GMP signaling pathways are directly or indirectly connected through Dos.

GGDEF domain proteins also connect polyamine signaling with c-di-GMP metabolism. It has been reported that NspS, a homologue of the periplamic spermidine-binding protein PotD from *E. coli*, enhances biofilm formation of *V. cholerae* in response to norspermidine, a triamine [45]. The genetic and phenotypic data presented in this work are in agreement with the hypothesis that binding of norspermidine to NspS leads to an interaction with the integral membrane protein MbdA through its periplasmic domain. This interaction subsequently inhibits the output activity of MbdA proposed to be a phosphodiesterase as a result of the presence of a GGDEF and EAL domain and phenotypic analysis.

The aminoglycoside antibiotic tobramycin at subinhibitory concentrations is another signal that directly or indirectly affects c-di-GMP metabolism [46••]. Tobramycin treatment leads to induction of biofilm formation and reduction of swimming motility in *P. aeruginosa*. The tobramycin-induced bacterial behavior was mediated by the membrane-bound EAL domain protein Arr (aminoglycoside response regulator, PA2818), which also contributes to biofilm-specific aminoglycoside resistance. The effect of Arr, which has apparent phosphodiesterase activity, on bacterial behavior is somewhat unexpected, because the correlation between high c-di-GMP levels and increased biofilm formation has been well established in many bacteria including *P. aeruginosa* [23••,27]. However, c-di-GMP metabolism is complex and the cellular c-di-GMP concentration was not established in this study.

Although c-di-GMP is considered an intracellular second messenger, application of c-di-GMP from the outside of a bacterial cell affected virulence and biofilm formation in *Staphylococcus aureus* [47,48] and fimbrial expression in *P. aeruginosa* (A Meissner *et al.*, abstract MPP025, 2nd Common Congress of the DGHM and VAAM, Göttingen, September 2005, unpublished). Because nucleotides do not penetrate membranes, this suggests the presence of a c-di-GMP receptor and/or uptake system that waits to be identified.

C-di-GMP signaling in eukaryotes — a link to cancer treatment?

Small molecules uniquely produced by bacteria have been shown to interact with eukaryotic signaling systems. Independent studies demonstrated that treatment with c-di-GMP impaired the proliferation of various cancer cell lines *in vitro* [49,50,51•]. Innovative studies performed by Amikam and collegues [49,50] more than a decade ago showed that c-di-GMP treatment exerted a marked increase of (H^3)thymidine incorporation in the human acute lymphoblastic leukemia cell line Molt 4 and the human T-cell leukemia cell line Jurkat. Further analysis revealed that c-di-GMP-treatment locked the majority of cells in the S (DNA synthesis) phase of the cell cycle, which lead to impaired cell replication. C-di-GMP treated cells had smaller cell size and augmented DNA content, which is a characteristic of S phase cells.

Recent studies by Karaolis *et al.* [51•] confirmed the previously observed inhibition of cancer cell proliferation by c-di-GMP. C-di-GMP exerted a marked inhibitory effect on basal proliferation and acetylcholine- and epidermal growth factor-induced proliferation of the human colon cancer cell line H508, although growth factor induced proliferation was also significantly inhibited by cGMP and 5'-GMP. Evidence for apoptosis of cells by various analytical methods was not found in either study [49,51•].

Molecular basis of growth inhibition of cancer cells by c-di-GMP

The Ras proteins are GTPases that cycle between GTP-bound active and GDP-bound inactive state, thereby functioning as molecular switches of signaling pathways that regulate cell proliferation, cell differentiation and survival [52]. Mutated forms of p21Ras, the classic Ras protein, are found in 30% of all human cancers underscoring the central role of Ras in the cellular physiology. Because the effects of c-di-GMP treatment resembled the effects of activated p21Ras, binding of c-di-GMP to Ras was examined in lymphoblastoid cells. c-di-GMP was found to specifically bind to the p21Ras oncoprotein

[49,50]. Furthermore, another consequence of Ras action, elevated expression of the T-cell receptor CD4, was also observed when the lymphoblastoid cell line Jurkat was treated with c-di-GMP. c-di-GMP-induced up regulation of CD4 was specific and not obtained with other guanine nucleotides.

Because of its effect on cancer cells, c-di-GMP might be a novel therapeutic agent in the treatment of cancer. Further studies to test the efficacy of c-di-GMP in an animal cancer model are suggested [51•].

Conclusions

The regulation of cellular physiology, morphogenesis and complex behavioral processes by cyclic nucleotides has long been thought to be restricted to eukaryotes. However, the recent recognition of the c-di-GMP signaling system in bacteria has demonstrated that developmental processes and virulence traits are subject to equally fine-tuned regulation based on second messenger signaling by a cyclic nucleotide. Many more similarities arise between the eukaryotic and prokaryotic signaling systems. The complexity of c-di-GMP metabolism tailored by GGDEF and EAL domain proteins (and other yet to be characterized functional entities such as the HD-GYP domain [4] with predicted c-di-GMP hydrolase activity) reflects the fine-tuned spatial and temporal regulation of cyclic nucleotides by nucleotide cyclase and phosphodiesterase superfamilies in eukaryotes [53]. The molecular mechanisms of c-di-GMP action in bacteria might be as diverse as targets of cyclic nucleotide binding in eukaryotes with binding sites in ion channels, protein kinases, phosphodiesterases and guanosine exchange factors [53]. PilZ, the first (predicted) c-di-GMP binding domain is already recognized as an allosteric binding site in enzymes such as cellulose synthases and protein phosphatases and in transcription factors. PilZ domains are also found in proteins that are part of secretion systems and in response regulators that connect phosphotransfer with c-di-GMP binding. Because extracellular c-di-GMP is sensed by certain bacteria, specific receptors and uptake systems might exist. Last but not least, c-di-GMP signaling extends to the eukaryotic system where c-di-GMP treatment inhibited the growth of cancer cells. We can only expect exciting novel details of the c-di-GMP signaling pathway to be discovered in the near future.

Update

The complexity of c-di-GMP signalling became apparent in a comprehensive survey of the role of all chromosomally encoded GGDEF or EAL domain proteins in two commonly used *P. aeruginosa* strains [54•]. Di-guanylate cyclase and phosphodiesterase activity was assigned to a subset of GGDEF and EAL domain proteins, whereas, apart from one exception; the 15 GGDEF-EAL domain proteins did not display either activity. Cytotoxicity and biofilm formation phenotypes were also mediated by a subset of GGDEF and EAL domain proteins with no inverse correlation indicating that c-di-GMP signalling might be highly localized.

A more detailed picture of the effects of c-di-GMP signalling is made possible through whole genome expression analysis. Overexpression of the di-guanylate cyclase YddV in *E. coli* revealed that expression of approximately 4.5% of the genes was affected by high intracellular c-di-GMP concentration [55]. Genes encoding for products involved in energy metabolism, transport and binding, and cell envelope related proteins were most often regulated by c-di-GMP.

Acknowledgements

Work in the laboratory of UR is supported by the Karolinska Institutet (elitforskartjänst to UR), Vetenskapsrådet (621-2004-3979) and the European Commission (MEST-CT-2004-008475). DA was supported by Tel-Hai Academic College, Tel-Hai, and Sharett Institute of Oncology, Hadassah University Medical Center, Israel.

This review is dedicated to Prof Moshe Benziman of the Hebrew University of Jerusalem, beloved husband of Dorit Amikam. He left a scientific legacy as the pioneer who discovered the novel nucleotide c-di-GMP. We cherish his memory with love and admiration.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Ross P, Weinhouse H, Aloni Y, Michaeli D, Weinberger-Ohana P, Mayer R, Braun S, de Vroom E, van der Marel GA, van Boom JH, Benziman M: **Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid**. *Nature* 1987, **325**:279-281.

2. Hecht GB, Newton A: **Identification of a novel response regulator required for the swarmer-to-stalked-cell transition in *Caulobacter crescentus***. *J Bacteriol* 1995, **177**:6223-6229.

3. Tal R, Wong HC, Calhoon R, Gelfand D, Fear AL, Volman G, Mayer R, Ross P, Amikam D, Weinhouse H *et al.*: **Three *cdg* operons control cellular turnover of cyclic di-GMP in *Acetobacter xylinum*: genetic organization and occurrence of conserved domains in isoenzymes**. *J Bacteriol* 1998, **180**:4416-4425.

4. Galperin MY, Nikolskaya AN, Koonin EV: **Novel domains of the prokaryotic two-component signal transduction systems**. *FEMS Microbiol Lett* 2001, **203**:11-21.

5. Jenal U: **Cyclic di-guanosine-monophosphate comes of age: a novel secondary messenger involved in modulating cell surface structures in bacteria?** *Curr Opin Microbiol* 2004, **7**:185-191.

6. D'Argenio DA, Miller SI: **Cyclic di-GMP as a bacterial second messenger**. *Microbiology* 2004, **150**:2497-2502.

7. • Römling U, Gomelsky M, Galperin MY: **C-di-GMP: The dawning of a novel bacterial signalling system**. *Mol Microbiol* 2005, **57**:629-639.
This review summarizes the current research on c-di-GMP metabolism and signalling, and addresses the most significant open questions for future research.

8. • Paul R, Weiser S, Amiot NC, Chan C, Schirmer T, Giese B, Jenal U: **Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain**. *Genes Dev* 2004, **18**:715-727.
The first biochemical demonstration of di-guanylate cyclase activity of a GGDEF domain. The authors also showed the coupling between activa-

tion and spatial localization of PleD in the cell for localized c-di-GMP production.

9. Hickman JW, Tifrea DF, Harwood CS: **A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels**. *Proc Natl Acad Sci USA* 2005, **102**:14422-14427.

10. Ryjenkov DA, Tarutina M, Moskvin OV, Gomelsky M: **Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain**. *J Bacteriol* 2005, **187**:1792-1798.

11. Chan C, Paul R, Samoray D, Amiot NC, Giese B, Jenal U, Schirmer T: **Structural basis of activity and allosteric control of diguanylate cyclase**. *Proc Natl Acad Sci USA* 2004, **101**:17084-17089.

12. • Christen M, Christen B, Folcher M, Schauerte A, Jenal U: **Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP**. *J Biol Chem* 2005, **280**:30829-30837.
As well as the characterization of the phosphodiesterase activity of the EAL domain protein CC3396, the authors demonstrate activation of phosphodiesterase by binding of GTP to a non-functional GGDEF domain located N-terminal of the EAL domain in CC3396.

13. • Schmidt AJ, Ryjenkov DA, Gomelsky M: **The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains**. *J Bacteriol* 2005, **187**:4774-4781.
Together with [12•], this paper shows the first biochemical characterization of phosphodiesterase activity of EAL domains.

14. Tamayo R, Tischler AD, Camilli A: **The EAL domain protein VieA is a cyclic diguanylate phosphodiesterase**. *J Biol Chem* 2005, **280**:33324-33330.

15. Bobrov AG, Kirillina O, Perry RD: **The phosphodiesterase activity of the HmsP EAL domain is required for negative regulation of biofilm formation in *Yersinia pestis***. *FEMS Microbiol Lett* 2005, **247**:123-130.

16. Tucker CL, Hurley JH, Miller TR, Hurley JB: **Two amino acid substitutions convert a guanylyl cyclase, RetGC-1, into an adenylyl cyclase**. *Proc Natl Acad Sci USA* 1998, **95**:5993-5997.

17. Richter W, Unciuleac L, Hermsdorf T, Kronbach T, Dettmer D: **Identification of substrate specificity determinants in human cAMP-specific phosphodiesterase 4A by single-point mutagenesis**. *Cell Signal* 2001, **13**:159-167.

18. • Kader A, Simm R, Gerstel U, Morr M, Römling U: **Hierarchical involvement of various GGDEF domain proteins in rdar morphotype development of *Salmonella enterica* serovar Typhimurium**. *Mol Microbiol* 200610.1111/j.1365-2958.2006.05123.x.
The first experimental evidence of the distinct functions of concomitantly expressed GGDEF and EAL domain proteins in the rdar morphotype development of *Salmonella* Typhimurium. The temporal and spatial compartmentalization of c-di-GMP signaling was demonstrated. In a mutant with a knockout of the diguanylate cylcase AdrA, the c-di-GMP pool was reduced by 60% and this did not affect the c-di-GMP responsive expression of CsgD.

19. Garcia B, Latasa C, Solano C, Portillo FG, Gamazo C, Lasa I: **Role of the GGDEF protein family in *Salmonella* cellulose biosynthesis and biofilm formation**. *Mol Microbiol* 2004, **54**:264-2772.

20. Hisert KB, MacCoss M, Shiloh MU, Darwin KH, Singh S, Jones RA, Ehrt S, Zhang Z, Gaffney BL, Gandotra S *et al.*: **A glutamate-alanine-leucine (EAL) domain protein of *Salmonella* controls bacterial survival in mice, antioxidant defence and killing of macrophages: role of cyclic diGMP**. *Mol Microbiol* 2005, **56**:1234-1245.

21. Weinhouse H, Sapir S, Amikam D, Shilo Y, Volman G, Ohana P, Benziman M: **c-di-GMP-binding protein, a new factor regulating cellulose synthesis in *Acetobacter xylinum***. *FEBS Lett* 1997, **416**:207-211.

22. Simm R, Fetherston JD, Kader A, Romling U, Perry RD: **Phenotypic convergence mediated by GGDEF-domain-containing proteins**. *J Bacteriol* 2005, **187**:6816-6823.

23. •• Simm R, Morr M, Kader A, Nimtz M, Römling U: **GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility**. *Mol Microbiol* 2004, **53**:1123-1134.
This study describes synthesis and degradation of c-di-GMP by GGDEF and EAL domain proteins in *S.* Typhimurium, *E. coli* and *P. aeruginosa*. A general concept for the regulation of bacterial behavior by c-di-GMP signaling is offered.

24. Tischler AD, Camilli A: **Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation**. *Mol Microbiol* 2004, **53**:857-869.

25. Kovacikova G, Lin W, Skorupski K: **Dual regulation of genes involved in acetoin biosynthesis and motility/biofilm formation by the virulence activator AphA and the acetate-responsive LysR-type regulator AlsR in *Vibrio cholerae***. *Mol Microbiol* 2005, **57**:420-433.

26. Fraser GM, Hughes C: **Swarming motility**. *Curr Opin Microbiol* 1999, **2**:630-635.

27. Klausen M, Aaes-Jorgensen A, Molin S, Tolker-Nielsen T: **Involvement of bacterial migration in the development of complex multicellular structures in Pseudomonas aeruginosa biofilms**. *Mol Microbiol* 2003, **50**:61-68.

28. Aldridge P, Paul R, Goymer P, Rainey P, Jenal U: **Role of the GGDEF regulator PleD in polar development of *Caulobacter crescentus***. *Mol Microbiol* 2003, **47**:1695-1708.

29. Choy WK, Zhou L, Syn CK, Zhang LH, Swarup S: **MorA defines a new class of regulators affecting flagellar development and biofilm formation in diverse *Pseudomonas* species**. *J Bacteriol* 2004, **186**:7221-7228.

30. • Tischler AD, Camilli A: **Cyclic diguanylate regulates *Vibrio cholerae* virulence gene expression**. *Infect Immun* 2005, **73**:5873-5882.
The authors showed that virulence properties of *V. cholerae*, such as the expression of cholera toxin, are stimulated by low c-di-GMP levels.

31. Thomas C, Andersson CR, Canales SR, Golden SS: **PsfR, a factor that stimulates *psbAI* expression in the cyanobacterium *Synechococcus elongatus* PCC 7942**. *Microbiology* 2004, **150**:1031-1040.

32. He J, Baldini RL, Deziel E, Saucier M, Zhang Q, Liberati NT, Lee D, Urbach J, Goodman HM, Rahme LG: **The broad host range pathogen Pseudomonas aeruginosa strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes**. *Proc Natl Acad Sci USA* 2004, **101**:2530-2535.

33. Römling U: **Characterization of the rdar morphotype, a multicellular behaviour in Enterobacteriaceae**. *Cell Mol Life Sci* 2005, **62**:1234-1246.

34. Osorio CG, Crawford JA, Michalski J, Martinez-Wilson H, Kaper JB, Camilli A: **Second-generation recombination-based *in vivo* expression technology for large-scale screening for *Vibrio cholerae* genes induced during infection of the mouse small intestine**. *Infect Immun* 2005, **73**:972-980.

35. Merkel TJ, Stibitz S, Keith JM, Leef M, Shahin R: **Contribution of regulation by the *bvg* locus to respiratory infection of mice by *Bordetella pertussis***. *Infect Immun* 1998, **66**:4367-4373.

36. Hersh D, Monack DM, Smith MR, Ghori N, Falkow S, Zychlinsky A: **The Salmonella invasin SipB induces macrophage apoptosis by binding to caspase-1**. *Proc Natl Acad Sci USA* 1999, **96**:2396-2401.

37. Römling U, Rohde M, Olsen A, Normark S, Reinköster J: **AgfD, the checkpoint of multicellular and aggregative behaviour in *Salmonella typhimurium* regulates at least two independent pathways**. *Mol Microbiol* 2000, **36**:10-23.

38. Mayer R, Ross P, Weinhouse H, Amikam D, Volman G, Ohana P, Calhoon RD, Wong HC, Emerick AW, Benziman M: **Polypeptide composition of bacterial cyclic diguanylic acid-dependent cellulose synthase and the occurrence of immunologically crossreacting proteins in higher plants**. *Proc Natl Acad Sci USA* 1991, **88**:5472-5476.

39. •• Amikam D, Galperin MY: **PilZ domain is part of the bacterial c-di-GMP binding protein**. *Bioinformatics* 2006, **22**:3-6.

Identification of PilZ, a candidate c-di-GMP binding domain, by bioinformatic analysis, supported by extensive literature search, opened the door to experimentally elucidate the molecular mechanisms of c-di-GMP signaling.

40. Saxena IM, Brown RM Jr: **Identification of a second cellulose synthase gene (*acsAII*) in *Acetobacter xylinum***. *J Bacteriol* 1995, **177**:5276-5283.

41. Huang B, Whitchurch CB, Mattick JS: **FimX, a multidomain protein connecting environmental signals to twitching motility in *Pseudomonas aeruginosa***. *J Bacteriol* 2003, **185**:7068-7076.

42. Ko M, Park C: **Two novel flagellar components and H-NS are involved in the motor function of *Escherichia coli***. *J Mol Biol* 2000, **303**:371-382.

43. Kulasekara HD, Ventre I, Kulasekara BR, Lazdunski A, Filloux A, Lory S: **A novel two-component system controls the expression of *Pseudomonas aeruginosa* fimbrial cup genes**. *Mol Microbiol* 2005, **55**:368-380.

44. Yoshimura-Suzuki T, Sagami I, Yokota N, Kurokawa H, Shimizu T: **DOS(Ec), a heme-regulated phosphodiesterase, plays an important role in the regulation of the cyclic AMP level in *Escherichia coli***. *J Bacteriol* 2005, **187**:6678-6682.

45. Karatan E, Duncan TR, Watnick PI: **NspS, a predicted polyamine sensor, mediates activation of *Vibrio cholerae* biofilm formation by norspermidine**. *J Bacteriol* 2005, **187**:7434-7443.

46. •• Hoffman LR, D'Argenio DA, MacCoss MJ, Zhang Z, Jones RA, Miller SI: **Amino-glycoside antibiotics induce bacterial biofilm formation**. *Nature* 2005, **436**:1171-1175.
This study demonstrated that the aminoglycoside antibiotic tobramycin enhances biofilm formation by signaling through the EAL domain protein Arr.

47. Karaolis DK, Rashid MH, Chythanya R, Luo W, Hyodo M, Hayakawa Y: **c-di-GMP (3′-5′-cyclic diguanylic acid) inhibits *Staphylococcus aureus* cell-cell interactions and biofilm formation**. *Antimicrob Agents Chemother* 2005, **49**:1029-1038.

48. Brouillette E, Hyodo M, Hayakawa Y, Karaolis DK, Malouin F: **3′,5′-cyclic diguanylic acid reduces the virulence of biofilm-forming Staphylococcus aureus strains in a mouse model of mastitis infection**. *Antimicrob Agents Chemother* 2005, **49**:3109-3113.

49. Amikam D, Steinberger O, Shkolnik T, Ben-Ishai Z: **The novel cyclic dinucleotide 3′-5′ cyclic diguanylic acid binds to p21ras and enhances DNA synthesis but not cell replication in the Molt 4 cell line**. *Biochem J* 1995, **311**:921-927.

50. Steinberger O, Lapidot Z, Ben-Ishai Z, Amikam D: **Elevated expression of the CD4 receptor and cell cycle arrest are induced in Jurkat cells by treatment with the novel cyclic dinucleotide 3′,5′-cyclic diguanylic acid**. *FEBS Lett* 1999, **444**:125-129.

51. • Karaolis DK, Cheng K, Lipsky M, Elnabawi A, Catalano J, Hyodo M, Hayakawa Y, Raufman JP: **3′,5′-Cyclic diguanylic acid (c-di-GMP) inhibits basal and growth factor-stimulated human colon cancer cell proliferation**. *Biochem Biophys Res Commun* 2005, **329**:40-45.
Ten-year-old findings were confirmed in this study, which investigated the effect of c-di-GMP treatment on cancer cell proliferation. Growth of the colon cancer cell line H508 was inhibited without apparent cytotoxicity suggesting that c-di-GMP might be suitable for cancer treatment.

52. Giehl K: **Oncogenic Ras in tumour progression and metastasis**. *Biol Chem* 2005, **386**:193-205.

53. Beavo JA, Brunton LL: **Cyclic nucleotide research — still expanding after half a century**. *Nat Rev Mol Cell Biol* 2002, **3**:710-718.

54. • Kulesekara H, Lee V, Brencic A, Liberati N, Urbach J, Miyata S, Lee DG, Neely AN, Hyodo M, Hayakawa Y *et al.*: **Analysis of Pseudomonas aeruginosa diguanylate cyclases and phosphodiesterases reveals a role for bis-(3′-5′)-cyclic-GMP in virulence**. *Proc Natl Acad Sci USA* 2006, **103**:2839-2844.
This is the first comprehensive analysis of the entire pool of GGDEF and EAL domain proteins in *P. aeruginosa* revealed task distribution in biofilm formation, cytotoxicity and virulence.

55. Mendez-Ortiz MM, Hyodo M, Hayakawa Y, Membrillo-Hernandez J: **Genome wide transcriptional profile of *Escherichia coli* in response to high levels of the second messenger c-di-GMP**. *J Biol Chem* 200610.1074/jbc.M510701200.

56. Römling U: **Molecular biology of cellulose production in bacteria**. *Res Microbiol* 2002, **153**:205-212.

57. Goodman AL, Kulasekara B, Rietsch A, Boyd D, Smith RS, Lory S: **A signaling network reciprocally regulates genes associated with acute infection and chronic persistence in *Pseudomonas aeruginosa***. *Dev Cell* 2004, **7**:745-754.

Molecular Microbiology (2005) **57**(3), 629–639 doi:10.1111/j.1365-2958.2005.04697.x

MicroReview

C-di-GMP: the dawning of a novel bacterial signalling system

Ute Römling,[1]* Mark Gomelsky[2] and Michael Y. Galperin[3]

[1]*Microbiology and Tumor Biology Center, Karolinska Institutet, Box 280, SE-17177 Stockholm, Sweden.*
[2]*Department of Molecular Biology, University of Wyoming, Laramie, WY 82071, USA.*
[3]*NCBI, NLM, National Institutes of Health, Bethesda, MD 20894, USA.*

Summary

Bis-(3′-5′)-cyclic dimeric guanosine monophosphate (c-di-GMP) has come to the limelight as a result of the recent advances in microbial genomics and increased interest in multicellular microbial behaviour. Known for more than 15 years as an activator of cellulose synthase in *Gluconacetobacter xylinus*, c-di-GMP is emerging as a novel global second messenger in bacteria. The GGDEF and EAL domain proteins involved in c-di-GMP synthesis and degradation, respectively, are (almost) ubiquitous in bacterial genomes. These proteins affect cell differentiation and multicellular behaviour as well as interactions between the microorganisms and their eukaryotic hosts and other phenotypes. While the role of GGDEF and EAL domain proteins in bacterial physiology and behaviour has gained appreciation, and significant progress has been achieved in understanding the enzymology of c-di-GMP turnover, many questions regarding c-di-GMP-dependent signalling remain unanswered. Among these, the key questions are the identity of targets of c-di-GMP action and mechanisms of c-di-GMP-dependent regulation. This review discusses phylogenetic distribution of the c-di-GMP signalling pathway in bacteria, recent developments in biochemical and structural characterization of proteins involved in its metabolism, and biological processes affected by c-di-GMP. The accumulated data clearly indicate that a novel ubiquitous signalling system in bacteria has been discovered.

Accepted 18 April, 2005. *For correspondence. E-mail ute.romling@mtc.ki.se; Tel. (+46) 8524 87319; Fax (+46) 833 0744.

Embarrassment of the riches: GGDEF and EAL domains in bacterial genomes

The principal reason for studying GGDEF and EAL domains – and, ultimately, for this review – is quite simple: there are just too many of them. The protein domains designated GGDEF and EAL are encoded in the genomes from diverse branches of the phylogenetic tree of bacteria. The domain names originate from the conserved sequence motifs, Gly–Gly–Asp–Glu–Phe and Glu–Ala–Leu; however, domain sizes are much larger, i.e. approximately 180 and 240 residues – for GGDEF and EAL respectively. Public protein databases currently contain over 2200 proteins with one domain or another, which makes GGDEF and EAL the most numerous domains, whose functions until recently have not been understood. Some databases still list GGDEF and EAL as 'domains of unknown function', DUF1 and DUF2 (e.g. SMART, http://smart.embl.de), so that annotation of these domains as uncharacterized often shows up in databases. In this review we describe the significant progress achieved in biochemical and physiological characterization of these protein domains. It has now been shown that GGDEF and EAL are involved in synthesis and hydrolysis, respectively, of bis-(3′-5′)-cyclic dimeric guanosine monophosphate (c-di-GMP) (Fig. 1). This unusual cyclic nucleotide is emerging as an important bacterial second messenger that has been overlooked despite decades of intensive microbiology research (see also D'Argenio and Miller, 2004; Galperin, 2004; Jenal, 2004).

Phylogenetic distribution of the GGDEF and EAL domains

GGDEF and EAL domains are found in the deeply branching phyla of bacteria, such as Thermotogales and Aquificales, but are not present in proteins encoded by the genomes of any Archaea or Eukarya (not counting highly diverged and presumably misassigned sequences; Galperin, 2004). This suggests that c-di-GMP-mediated signalling is an exclusively bacterial trait. However, c-di-GMP-mediated signalling pathways are not equally abundant in diverse bacterial species (Fig. 2). The genome of *Escher-*

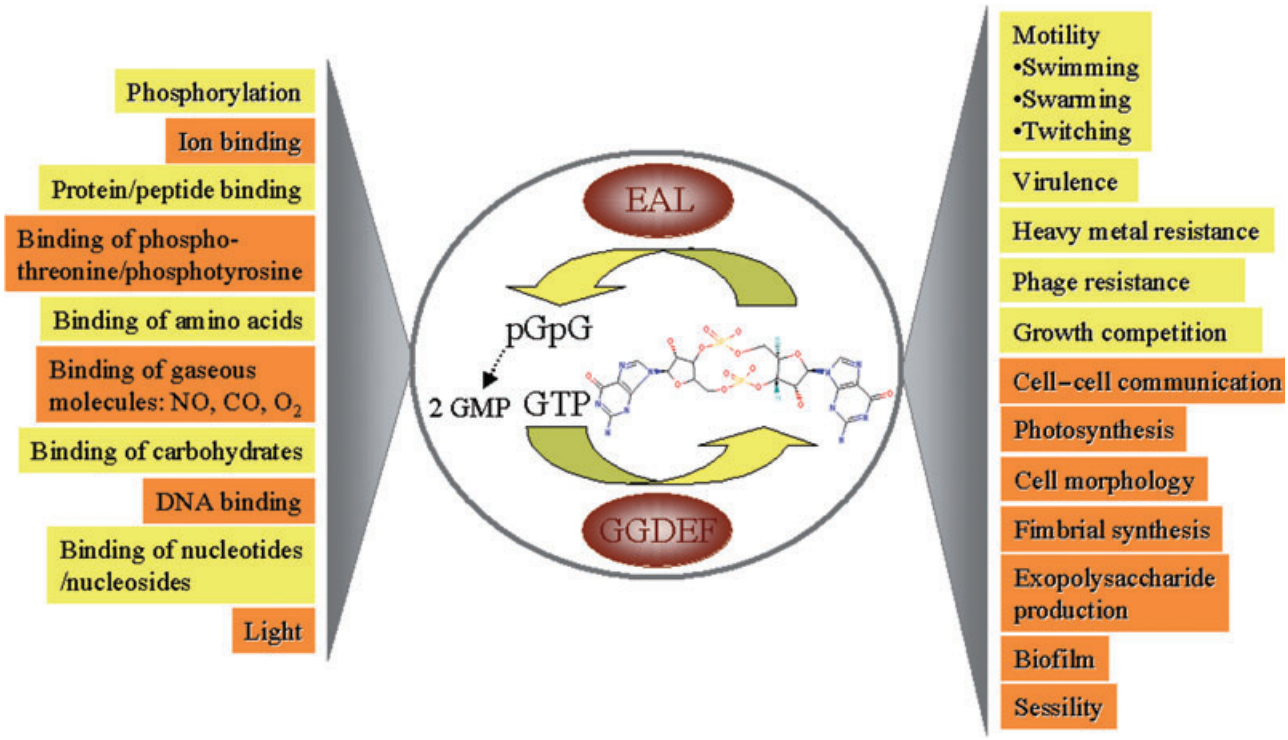


Fig. 1. Known input signals and output of c-di-GMP metabolism. GGDEF and EAL domains conduct the turnover of c-di-GMP. pGpG is degraded to two GMP by an unknown phosphodiesterase. Various domains N-terminal of GGDEF or EAL receive and transmit the input signals (left): phosphorylation [Rec, CheY-homologous receiver domain; HisKA, histidine kinase A (phosphoacceptor) domain; HATPase, histidine kinase-like ATPases; Hpt, histidine phosphotransfer domain], ion binding (haemerythrin, HHE cation-binding domain), protein/peptide binding (TRP, tetratrico peptide repeats; CBS, domain in cystathionine beta-synthase and other proteins; CHASE, cyclase/histidine kinases-associated sensory domain), binding of phosphothreonine/phosphotyrosine (FHA, fork-head-associated domain), binding of amino acids (PBPb, bacterial periplasmic substrate-binding proteins), binding of gaseous molecules [PAS/PAC, Per (periodic clock protein), Arnt (Ah receptor nuclear translocator protein), Sim (single-minded protein); haemerythrin], binding of carbohydrates (7TMR-DISMED2, 7TM receptors with diverse intracellular signalling modules), DNA binding (HTH–LUXR, helix–turn–helix, Lux regulon), binding of nucleotides/nucleosides (GAF, domain present in phytochromes and cGMP-specific phosphodiesterases; cNMP, cyclic nucleotide-monophosphate binding domain; CBS), light (BLUF, sensor of blue-light using FAD). The output behaviour by variation of c-di-GMP concentration is shown on the right.

ichia coli K12, for example, encodes 19 proteins with the GGDEF domain and 17 with the EAL domain, whereas *Bacillus subtilis* has four and three, respectively, and even the tiny genome of *Rickettsia prowazekii* encodes one of each (Galperin *et al.*, 1999; 2001; Galperin, 2004). The current champion, *Vibrio vulnificus*, encodes 66 proteins with the GGDEF domain and 33 with the EAL domain, with the human pathogen *Vibrio cholerae* coming not far behind (Galperin, 2004). Among bacteria, GGDEF and EAL domains are absent in proteins encoded by the sequenced genomes of *Bacteroidetes*, *Chlamydiales* and *Fusobacteria* (Fig. 2).

Flexibility through modularity

Existing experimental evidence suggests that GGDEF and EAL are soluble cytoplasmic domains. Although few single-domain proteins exist, the GGDEF and EAL domains are usually found in multidomain proteins. Often a GGDEF and an EAL domain occur in the same protein, whereby the GGDEF domain is located N-terminally from the EAL domain (few exceptions exist). Extending the complexity, the GGDEF and EAL domains are located C-terminally from, often multiple, sensory and signal transduction domains (Fig. 1). The different N-terminal domains are capable of acquiring a wide range of signals, i.e. phosphorylation, protein binding, binding of gaseous molecules, light. A significant fraction of the GGDEF and/or EAL domains is linked to cytoplasmic sensory domains such as PAS and GAF involved in binding of small molecular ligands or protein–protein interactions (Taylor and Zhulin, 1999; Hurley, 2003). Another sizable fraction is linked to N-terminal periplasmic or integral membrane sensory domains whose ligand binding specificity is unknown (Nikolskaya *et al.*, 2003; Zhulin *et al.*, 2003). In some cases, periplasmic domains belong to well-characterized families of solute-binding proteins, for example, those involved in amino acid binding, suggesting that the

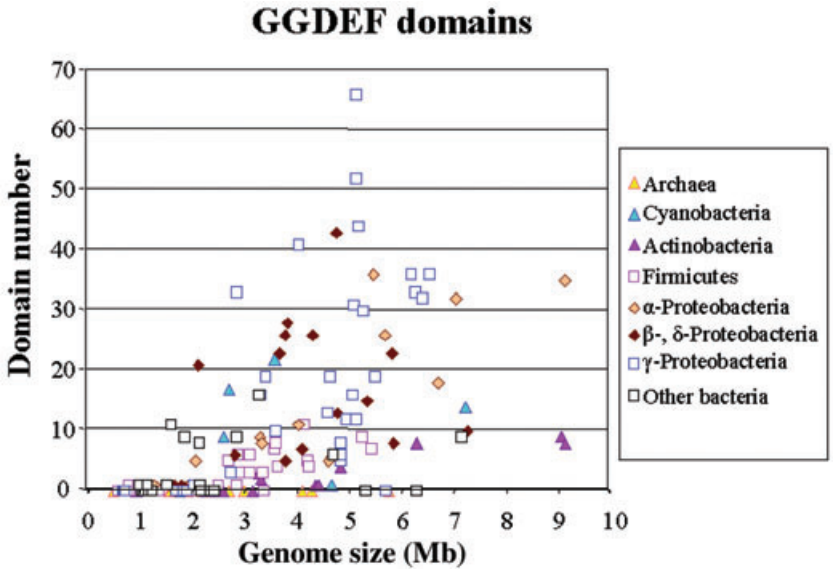


Fig. 2. Phylogenetic distribution of the GGDEF domains in sequenced prokaryotic genomes.

presence of an amino acid can modulate the turnover of c-di-GMP in the same fashion as they modulate activities of histidine kinases. Thus, the overall architecture of GGDEF and/or EAL domain proteins (sensor + output domain) is similar to that of sensor histidine kinases and methyl-carrier chemotaxis proteins (Galperin, 2004).

Biochemical activity of the GGDEF domains

A key paper in 1998 from Moshe Benziman's group (Tal *et al*., 1998) provided the first connection between enzymatic activities of proteins containing GGDEF and EAL domains and turnover of c-di-GMP (Fig. 1), which has been identified as an allosteric activator of cellulose synthase in the bacterium *Gluconacetobacter xylinus* (formerly *Acetobacter xylinum*) more than 15 years ago (Ross *et al*., 1987). Tal *et al*. (1998) cloned and characterized six proteins of almost identical domain composition, namely PAS-GGDEF-EAL and PAS-GAF-GGDEF-EAL, where PAS and GAF are structurally related ligand-binding domains, the first of which typically binds haem and flavins (Taylor and Zhulin, 1999), and the second one cAMP and cGMP (Hurley, 2003). Of these six paralogous proteins, three had c-di-GMP synthetase (diguanylate cyclase, DGC) activity and three had c-di-GMP phosphodiesterase (PDE-A) activity, hydrolysing c-di-GMP to linear diguanylate pGpG (Tal *et al*., 1998). Further hydrolysis of pGpG to two 5′-GMP molecules, referred to as PDE-B activity by Tal *et al*. (1998), was apparently catalysed by unrelated enzymes.

This combination of activities led to the proposal that GGDEF domain, which is distantly related to mammalian adenylate cyclases (Pei and Grishin, 2001), might be responsible for c-di-GMP production (Galperin *et al*., 1999; 2001; Ausmees *et al*., 2001; Galperin, 2004). These suggestions have been supported by a variety of genetic data (Ausmees *et al*., 2001; Aldridge *et al*., 2003; Simm *et al*., 2004; Tischler and Camilli, 2004) and, recently, by direct biochemical experiments (Paul *et al*., 2004; Ryjenkov *et al*., 2005). Specifically, Jenal and colleagues demonstrated the DGC activity of the purified *Caulobacter crescentus* PleD response regulator protein (CheY-CheY-GGDEF; Paul *et al*., 2004), while Gomelsky and co-workers showed the same GTP-dependent DGC activity for another six randomly chosen GGDEF domain proteins from representatives of diverse branches of the bacterial phylogenetic tree, namely, *Thermotogae*, *Deinococcus-Thermus*, *Cyanobacteria*, *Spirochaetes*, α- and γ-divisions of the *Proteobacteria* (Fig. 3; Ryjenkov *et al*., 2005). None of the GGDEF domain proteins showed activity with any other nucleotides, indicating that they are dedicated to c-di-GMP synthesis (Ryjenkov *et al*., 2005).

Mechanistic aspects of c-di-GMP formation

The crystal structure of the PleD protein (Hecht and Newton, 1995; Aldridge *et al*., 2003; Paul *et al*., 2004) in the presence of its reaction product, c-di-GMP, was recently solved (Protein Data Bank access code 1w25; Chan *et al*., 2004). As predicted by Pei and Grishin (2001), the structure of its GGDEF domain is very similar to that of eukaryotic adenylyl cyclases and includes a putative GTP binding catalytic site formed by the residues of the GGDEF motif as well as other conserved residues of the GGDEF domain (Fig. S1A in *Supplementary material*). The three-dimensional (3D) structure of the GGDEF domain readily explained why most amino acid substitutions, such as changing either of the Gly residues in the

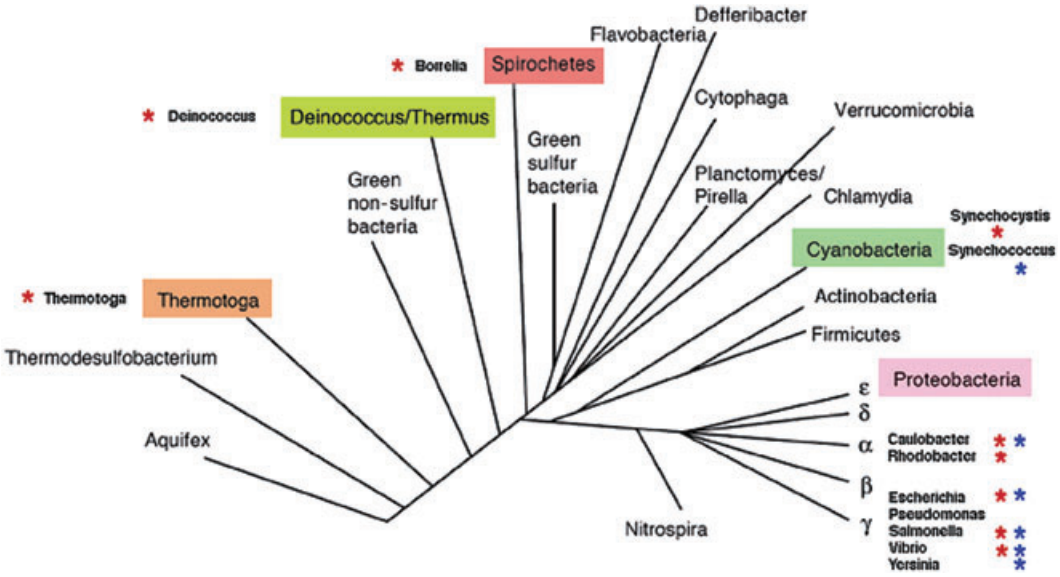


Fig. 3. Functionally verified GGDEF and EAL domains in different bacterial phyla. Blue star indicates species for which DGC and/or PDE activities have been demonstrated *in vivo*. Red star indicates species for which DGC and/or PDE activities have been demonstrated *in vitro*.

GGDEF motif to Ala (Kirillina *et al*., 2004), abrogate the activity of this domain. The GGDEF domain appeared to contain half of the catalytic site where a single GTP molecule may be bound, suggesting that the enzymatically active form of DGC is a dimer of two GGDEF domains, as seen in the crystal structure (Chan *et al*., 2004). The phosphorylation of one of the CheY-like receiver domains of PleD is envisioned to trigger formation of such an active dimer with a catalytic site formed between the two subunits. The proposed mechanism is in agreement with biochemical observations that individually expressed and purified GGDEF domains from various proteins have propensity to form dimers. However, such dimers possess only low-level DGC activity, while the presence of activated sensor domains (GAF or phosphorylated CheY-like receiver domain) results in significantly higher DGC activity (Paul *et al*., 2004; Ryjenkov *et al*., 2005). Noteworthy, sensor histidine kinases and chemotaxis methyl-carrier proteins, like GGDEF domains, function as dimers, suggesting possible common themes in the mechanisms of regulation of these otherwise different signalling systems.

Biochemical activity of the EAL domains

While these findings left the EAL domain (Fig. S1B) as the most plausible candidate for the PDE-A activity seen by Tal *et al*. (1998) and Chang *et al*. (2001) in *G. xylinus* proteins, there still remained a possibility that both GGDEF and EAL domains are required for the PDE-A activity. Although *in vivo* experiments showed that proteins containing EAL but not GGDEF domain were capable of c-di-GMP degradation (Simm, 2004; Tischler and Camilli, 2004), they could not fully resolve the issue of whether EAL is sufficient for PDE-A activity or just reversing the activity of GGDEF domains present in the same cell. Schmidt *et al*. (2005) showed that purified EAL domains from *E. coli*, YahA and Dos (Delgado-Nixon *et al*., 2000) proteins possessed PDE-A activity, i.e. hydrolysed c-di-GMP into linear dimeric GMP, pGpG. The earlier report that *E. coli* Dos protein had cAMP phosphodiesterase activity (Sasakura *et al*., 2002) could not be confirmed. PDE-A activity was dependent on Mg^{2+} or Mn^{2+}, strongly inhibited by Ca^{2+} and did not require protein oligomerization. EAL domains and full-length EAL domain proteins also catalysed the subsequent hydrolysis of the reaction product, pGpG, into two 5′-GMP molecules. However, the latter PDE-B activity was several orders of magnitude slower and probably not physiologically relevant. This is consistent with the observation by the Benziman group (Ross, 1987; Tal *et al*., 1998) and suggests that hydrolysis of pGpG is catalysed by a different enzyme(s).

One can envisage that pGpG is rapidly degraded by cellular nucleases. However, as production and hydrolysis of pGpG are apparently uncoupled, a regulatory role of pGpG as a player in the c-di-GMP-mediated response or as a signalling molecule on its own is also possible, as has been found for other nucleotides and dinucleotides (Ismail *et al*., 2003; Gralla, 2005).

Another domain with PDE-A activity?

Several bacterial genomes encode proteins with the GGDEF but not proteins with EAL domains. The most conspicuous case is *Thermotoga maritima*, which

encodes nine proteins with the GGDEF domain but not a single protein with the EAL domain. This led to the suggestion that this bacterium could have proteins with PDE-A activity but with another domain, i.e. HD-GYP (Galperin *et al*., 1999; 2001; Galperin, 2004). Whether HD-GYP has c-di-GMP hydrolase activity remains to be tested but other bacteria, including *Treponema pallidum*, *Campylobacter jejuni* and several other important human pathogens, have a protein with a GGDEF domain but do not have proteins with either EAL or HD-GYP domains. If these GGDEF domain proteins are active, how do these pathogens reduce their c-di-GMP levels? As a possibility, c-di-GMP is degraded by phosphodiesterases with relaxed substrate specificity in the respective bacteria.

So many domains, so few phenotypes

The first phenotype resulting from a mutation inactivating a gene encoding a GGDEF domain protein was detected when cell differentiation was analysed in *C. cresentus* (Hecht and Newton, 1995). However, it was not until the systematic genetic analysis of multicellular bacterial behaviour was initiated that mutants in genes encoding GGDEF and EAL domain proteins started to pop up more frequently (Ausmees *et al*., 1999; Jones *et al*., 1999; Römling *et al*., 2000; Boles and McCarter, 2002; D'Argenio *et al*., 2002). The realization that multicellular behaviour can be assayed by simply testing for adherence to and growth on polystyrene or glass surfaces, by pellicle formation at the liquid–air interface or by expression of distinct colony morphologies, greatly facilitated high throughput screening for mutants and hence lead to the functional identification of GGDEF and/or EAL domain proteins.

The peculiar colony morphology, which shows a surprising convergence among bacteria, has been termed wrinkled, rugose or rdar (Fig. 4; Römling *et al*., 1998; Boles and McCarter, 2002; D'Argenio *et al*., 2002; Spiers *et al*., 2002; Rashid *et al*., 2003; Kearns *et al*., 2005) and results from the presence of extracellular matrix components, which allow bacterial cells to form highly structured entities. The extracellular matrix consists of exopolysaccharides and other extracellular matrix components, often called 'slime', if the components identity is uncertain. Cellulose and an acetylated cellulose derivative were the first identified exopolysaccharides, the production of which was shown to be activated by GGDEF domain-containing proteins in such diverse bacteria as *G. xylinus*, *E. coli*, *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium), *Rhizobium leguminosarum* and *Pseudomonas fluorescence* (Tal *et al*., 1998; Ausmees *et al*., 1999; Zogaj *et al*., 2001; Spiers *et al*., 2002). Adhesive fimbriae represent yet another class of extracellular matrix components whose synthesis depends on the GGDEF domain proteins (D'Argenio *et al*., 2002; Römling, 2005). Apparently, production of GGDEF domain proteins and the resulting elevated c-di-GMP concentrations favour the production of adhesive matrix components, which leads to elevated multicellular behaviour.

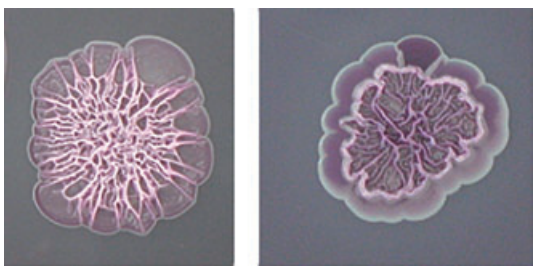


Fig. 4. Rdar colony morphologies of *S.* Typhimurium and *E. coli.* Colonies were grown on LB without salt agar at 37°C for 24 h (*S.* Typhimurium ATCC14028) and at 28°C for 48 h (*E. coli* DSM6601). The agar medium contained Congo Red (40 μg ml^{-1}) and Coomassie brilliant blue (20 μg ml^{-1}).

Consequently, decrease of c-di-GMP concentration leads to the opposite behaviour. Consistent with the function of EAL domain proteins as PDE-A, biofilm formation and autoaggregation are all suppressed upon overproduction of EAL domain proteins, but various types of motility are activated (Hecht and Newton, 1995; Drenkard and Ausubel, 2002; Kirillina *et al*., 2004; Simm *et al*., 2004; Tischler and Camilli, 2004). Swimming motility defect resulting from loss of flagella motion in an *hns* mutant of *E. coli* was overcome by producing an EAL domain protein (Ko and Park, 2000). This finding showed that downregulation of c-di-GMP concentrations leads to functional activation of structural components that can be uncoupled from the synthesis of the respective structures.

Interestingly, overproduction of or mutations affecting the GGDEF and EAL domain proteins do not lead to an 'all or nothing' phenotype. Individual GGDEF/EAL domain proteins modulate biofilm architecture, indicating differential activity in the individual cells within the biofilm (Huber *et al*., 2002; Bomchil *et al*., 2003). Such spatial heterogeneity is not surprising, as biofilms, with their water channels and complex architecture, contain various microenvironmental niches. Moreover, biofilms are not static, but consist of both moving and sessile cells (Klausen *et al*., 2003). Thus, it is reasonable to assume that the decision to stay in the biofilm or to move away is made on a nanoscale by individual cells based on the assessment of their local microenvironment.

The role of GGDEF and EAL domain proteins goes beyond production of extracellular matrix components and stimulation of motility. In *C. crescentus*, the GGDEF-domain protein PleD controls not only flagellum ejection but also cell morphology (stalk formation; Aldridge and Jenal 1999). Further, cell-to-cell communication in *Myxo-*

coccus xanthus and *Burkholderia cepacia* required to coordinate multicellular behaviour is affected by GGDEF domain proteins (Gronewold and Kaiser, 2001; Huber *et al*., 2002), while the EAL domain protein YhjH has been shown to affect growth competition between different strains of *S. enterica* (Rychlik *et al*., 2002).

Functions beyond multicellular behaviour

GGDEF/EAL domain proteins appear to alter the expression of other GGDEF and EAL domain proteins (Kirillina *et al*., 2004) as well as protein composition in membranes. Early studies of uncharacterized *E. coli* proteins now known to contain the EAL domain suggested that they might play a role in resistance to bacteriophages and heavy metal ions (Brown *et al*., 1986; Chae and Yoo, 1986). In the cyanobacterium *Synechococcus elongatus*, the production of photopigments and the photosynthetic reaction centre protein is modulated by the response regulator PsfR that contains a GGDEF domain (Thomas *et al*., 2004). It is not entirely clear whether the above described phenotypes are related to multicellular behaviour.

In addition, GGDEF and EAL domain proteins are apparently involved in controlling interactions of pathogens with their various hosts ranging from plants to animals (Merkel and Stibitz, 1995; Milton *et al*., 1995; Ausmees *et al*., 1999). An EAL domain-containing protein is required to suppress virulence-inhibiting genes in *Bordetella pertussis* (Merkel *et al*., 1998a). Furthermore, some GGDEF/EAL domain-containing proteins were shown to be produced by the bacterial pathogens *V. cholerae* and *Vibrio vulnificus* exclusively during infection of mice or humans, but not in laboratory media (Camilli and Mekalanos, 1995; Lee *et al*., 2001; Kim *et al*., 2003). The list of processes regulated by GGDEF/EAL domain proteins will undoubtedly continue to grow.

Functions of GGDEF-EAL domain fusions

As the GGDEF domain is sufficient for DGC activity and the EAL domain for PDE-A activity, what is the impact of proteins containing an N-terminal GGDEF and C-terminal EAL domain? Is one of the domains non-functional or can these proteins switch between the DGC and PDE-A states depending on conditions?

GGDEF-EAL domain proteins whose activities were tested *in vitro* (Tal *et al*., 1998; M. Tarutina, D.A. Ryjenkov and M. Gomelsky, unpublished) show only one activity, suggesting that one of the two domains is inactive. Sequence analysis of the EAL domains from the three DGCs from *G. xylinus* revealed differences in several highly conserved amino acid motifs preserved in the EAL domains from active PDE-As (Schmidt *et al*., 2005). Similarly, some GGDEF domains in GGDEF-EAL domain proteins with PDE-A activity deviate from the consensus sequence and different subclasses of GGDEF domains could be distinguished (Pei and Grishin, 2001). However, the molecular basis for inactivity is not entirely clear and the assignment of amino acids definitely required for enzymatic functionality requires saturating analyses. The hypothesis of inactive domains, if proven, would explain the biochemical paradox of the GGDEF-EAL domain proteins. However, it would pose new questions, i.e. what roles do enzymatically inactive domains play? What is the function of proteins that contain one or even two seemingly enzymatically inactive domains?

However, some *in vivo* observations suggest that a GGDEF-EAL domain protein may possess both DGC and PDE-A activities. For example, the knockout of diguanylate cyclase 1 (*dgc-1*), which is responsible for 80% of the c-di-GMP production in *G. xylinus* (Tal *et al*., 1998; Bae *et al*., 2004), showed decreased cellulose production under some growth conditions but, surprisingly, elevated cellulose production under others (Bae *et al*., 2004). Furthermore, ScrC, a GGDEF-EAL domain protein from *Vibrio parahaemolyticus*, has the opposite function on two transcriptional fusions monitoring swarming and biofilm formation when expressed alone or in its operon context *scrABC* (Boles and McCarter, 2002). Inactivation of homologous GGDEF-EAL domain-containing proteins in different bacteria sometimes causes opposite phenotypes. For example, the knockout of *yciR* (STM1703) of *S.* Typhimurium caused cellulose overproduction, indicating that YciR has PDE-A activity (Garcia *et al*., 2004), whereas, a mutation in *yciR* diminished biofilm formation in *B. cepacia*, indicating that the encoded protein has DGC activity (Huber *et al*., 2002). Clearly, additional work is needed to determine the full enzymatic capacity of proteins containing both GGDEF and EAL domains.

The challenge of fine-tuned networking

The inactivation of genes encoding GGDEF and EAL domain proteins frequently modulates the amplitude of a phenotype or regain of function is achieved under different environmental conditions, but seldom causes a complete phenotype change (Aldridge and Jenal, 1999; Gronewold and Kaiser, 2001; Garcia *et al*., 2004; Thomas *et al*., 2004). For example, cellulose biosynthesis in *S. enterica* on LB agar plates without salt is activated by the GGDEF domain protein AdrA (Römling *et al*., 2000), while in a carbon source-rich, trace element-poor medium, STM1987 fulfils this function (Garcia *et al*., 2004). Thus, the functionality of the two GGDEF domain proteins to mediate the phenotype 'cellulose biosynthesis' was restricted to certain environmental conditions. This finding brings up the question of how to manage a

branched network of DGCs and PDE-As and avoid unwanted cross-talk between different c-di-GMP targets in a single bacterial cell. We speculate that several mechanisms might work in parallel to accomplish specificity of c-di-GMP action. One involves temporal and environmental control of GGDEF and EAL domain protein expression on a transcriptional and post-transcriptional level. Cell density-dependent expression of GGDEF and EAL domain proteins was observed by DNA array analysis (Johnson *et al*., 2005). Temperature-regulated expression of the GGDEF domain protein HmsT by proteolysis controls the haemin storage phenotype of *Yersinia pestis* (Perry *et al*., 2004). The second possible mechanism relies on the dependence of enzymatic activities of DGCs and PDE-As on environmental or intracellular stimuli. In line with this suggestion, the activities of PleD and of the Rrp1 protein from *Borrelia burgdorferi* are strongly dependent on the phosphorylation status of the CheY domains (Chan *et al*., 2004; Ryjenkov *et al*., 2005). Hence, a given enzyme produces or degrades c-di-GMP only when a specific signal has been received. A third mechanism probably involves colocalization of enzymes of c-di-GMP turnover with their targets. Two pieces of experimental evidence are in line with this scenario. (i) *G. xylinus* DgcA and PdeA proteins were found to co-purify with cellulose synthase (Ross *et al*., 1987). (ii) *C. crescentus* PleD is localized near the base of the flagellum, where its DGC activity is required for flagellum ejection. A fourth potential mechanism is a feedback inhibition of the enzymatic activity by the product, c-di-GMP. Such inhibition by c-di-GMP has been observed for PleD (Paul *et al*., 2004), in which an allosteric binding site for c-di-GMP is formed jointly by the signalling and GGDEF domain (Chan *et al*., 2004). Whether this mechanism is unique to PleD or common among DGCs remains to be tested. However, significantly different *in vivo* activities of c-di-GMP formation have been observed among proteins (R. Simm and U. Römling, unpublished results), suggesting that a broad range of enzymatic activities is present.

The intracellular levels of c-di-GMP are in the micromolar range or lower, as judged by measurement in several proteobacterial species (Weinhouse *et al*., 1997; Simm *et al*., 2004; Tischler and Camilli, 2004). Even if a reliable procedure for quantification is set up, the significance of a measurement of the average cytoplasmic c-di-GMP concentration is not entirely obvious. C-di-GMP concentrations probably fluctuate locally, as discussed above. The measured and real concentrations of c-di-GMP in the cytoplasm might differ. For example, the majority of the c-di-GMP in *G. xylinus* is apparently bound by a yet undefined membrane protein (Weinhouse *et al*., 1997) and is probably released in response to certain signals, which might lead to very short cytoplasmic c-di-GMP spikes. Benziman and colleagues have shown that only a fraction of the cellular c-di-GMP concentration is required to achieve almost optimal activation of cellulose biosynthesis (Tal *et al*., 1998), suggesting that various biological processes might require different thresholds of c-di-GMP concentrations.

Networking – yes! – but where?

Where do GGDEF/EAL domain proteins stand in the established regulatory network of signal transduction in bacteria. A subfraction of GGDEF/EAL domain proteins is intimately linked to two-component signalling systems. In some bacteria, genes encoding EAL domain proteins are coexpressed with sensor kinase and response regulator genes (Merkel *et al*., 1998b; Tischler and Camilli, 2004; Kulasekara *et al*., 2005). Many functionally characterized GGDEF/EAL domain proteins have an N-terminal receiver domain (Gronewold and Kaiser, 2001; Drenkard and Ausubel, 2002) that is phosphorylated by cognate sensor kinases (Aldridge *et al*., 2003; Kulasekara *et al*., 2005). In *S. enterica*, production of the GGDEF domain protein AdrA is transcriptionally activated by the response regulator CsgD (Römling *et al*., 2000). These examples show that the phosphotransfer network established by two-component systems integrates components of the c-di-GMP signalling pathway on several levels.

In other instances, EAL domain proteins might be directly involved in gene regulation, although there are no experimental data that support this idea. However, several proteobacterial proteins contain EAL domains and predicted DNA-binding domains, and two of them have experimentally verified PDE-A activity (Tischler and Camilli, 2004; Schmidt *et al*., 2005). What, if any, is the relation of their PDE-A activity to DNA binding? One working hypothesis that arises is that binding of the small molecule c-di-GMP to the EAL domain affects DNA binding or transcriptional activation.

Mechanism(s) of c-di-GMP action – an entirely unresolved question

Studies of the c-di-GMP-mediated signalling in bacteria are still in their infancy and there are more questions than answers. The most intriguing question is the nature of the target(s) of c-di-GMP action. In their original studies of the regulation of *G. xylinum* cellulose synthase, Benziman and colleagues detected c-di-GMP binding to a 67 kDa soluble protein whose N-terminal region is very similar to an internal sequence in the cellulose synthase β-subunit, BcsB (Mayer *et al*., 1991). This led to the suggestion that c-di-GMP binding is a property of the cellulose synthase itself. However, subsequent studies revealed that c-di-GMP could have been interacting with a 200 kDa

membrane-bound protein complex (Weinhouse *et al*., 1997) that has not been further characterized.

Could there be a single c-di-GMP-binding adaptor protein, akin to CRP for cAMP, or are there multiple targets for c-di-GMP action? A single adaptor would be expected to have the same phylogenetic distribution as the GGDEF and EAL domains but no such protein could be identified (M.Y. Galperin, unpubl. obs.). Therefore, c-di-GMP is likely to operate through several different targets.

Practical importance of c-di-GMP

There is an urgent need to overcome the intrinsically high resistance of biofilm-forming pathogens to traditional antimicrobial agents. Here, one strategy is to develop therapies that interfere with the formation of the multicellular communities. Biofilm formation of *Pseudomonas aeruginosa*, *S. enterica*, *E. coli*, *V. cholerae* and *Y. pestis* has already been demonstrated to be manipulable by varying c-di-GMP concentrations (Kirillina *et al*., 2004; Simm *et al*., 2004; Tischler and Camilli, 2004). As biofilm formation of those important human pathogens is associated with chronic disease, the carrier state or transmission, c-di-GMP metabolism is emerging as an attractive target for clinical intervention. On the other hand, c-di-GMP added extracellularly to *Staphylococcus aureus* inhibited cell biofilm formation and attachment to human epithelial cells suggesting that c-di-GMP itself could be used as a therapeutic agent to interfere with multicellular behaviour and adherence mechanisms (Karaolis *et al*., 2005a). These findings once again demonstrated the link between basic microbiological research and practical aspects of antibacterial defence as prominently stated in a recent open letter by US microbiologists (Altman *et al*., 2005). However, effects of c-di-GMP treatment with a potential for therapeutical interference in cancer were also observed in eukaryotic cells (Amikam *et al*., 1995; Steinberger *et al*., 1999; Karaolis *et al*., 2005b). The observation of effectively applied extracellular c-di-GMP brings forth the intriguing possibility that specific receptors or uptake systems for c-di-GMP exist in pro- and eukaryotes.

Concluding remarks

It is evident that there are more questions than answers regarding c-di-GMP-dependent signalling. However, we are in the midst of an explosion of interest in c-di-GMP. The next few years will likely prove crucial for establishing the scope of targets of this novel second messenger and determining its mechanism(s) of action. C-di-GMP has risen to prominence due to the massive microbial genome sequencing and an increased interest in the natural bacterial lifestyle, the biofilm. We find ourselves at the dawning of a novel signalling system in bacteria. C-di-GMP is rapidly earning its place alongside its better-understood cousins, cAMP and cGMP, as a ubiquitous second messenger.

Acknowledgements

Work in the laboratory of U.R. is supported by the Karolinska Institutet (Elitforskartjänst), the Mukoviszidose e.V. and Vetenskapsrådet Grant 621-2004-3979. Work on c-di-GMP in the laboratory of M.G. is supported by the NSF Grant MCB-0316270.

Note added in proof

After this review was completed, phosphodiesterase activity of the EAL domain proteins has been reported in two more papers by Hisert *et al*. (2005) A glutamate-alonine-leucine (EAL) domain protein of *Salmonella* controls bacterial survival in mice, antioxidant defence and killing of macrophages: role of cyclic oliGMP. *Mol Microbiol* **56:** 1234–1245. And Bobrov *et al*. (2005) The phosphodiesterase activity of the HmsP EAL domain is required for negative regulation of biofilm formation in *Yersinia pestis*. *FEMS Microbiol Lett*, in press.

References

Aldridge, P., and Jenal, U. (1999) Cell cycle-dependent degradation of a flagellar motor component requires a novel-type response regulator. *Mol Microbiol* **32:** 379–391.

Aldridge, P., Paul, R., Goymer, P., Rainey, P., and Jenal, U. (2003) Role of the GGDEF regulator PleD in polar development of *Caulobacter crescentus*. *Mol Microbiol* **47:** 1695–1708.

Altman, S., Bassler, B.L., Beckwith, J., Belfort, M., Berg, H.C., Bloom, B., *et al.* (2005) An open letter to Elias Zerhouni. *Science* **307:** 1409–1410.

Amikam, D., Steinberger, O., Shkolnik, T., and Ben-Ishai, Z. (1995) The novel cyclic dinucleotide 3′-5′-cyclic diguanylic acid binds to p21ras and enhances DNA synthesis but not cell replication in the Molt 4 cell line. *Biochem J* **311:** 921–927.

Ausmees, N., Jonsson, H., Hoglund, S., Ljunggren, H., and Lindberg, M. (1999) Structural and putative regulatory genes involved in cellulose synthesis in *Rhizobium leguminosarum* bv. *trifolii*. *Microbiology* **145:** 1253–1262.

Ausmees, N., Mayer, R., Weinhouse, H., Volman, G., Amikam, D., Benziman, M., *et al.* (2001) Genetic data indicate that proteins containing the GGDEF domain possess diguanylate cyclase activity. *FEMS Microbiol Lett* **204:** 163–167.

Bae, S.O., Sugano, Y., Ohi, K., and Shoda, M. (2004) Features of bacterial cellulose synthesis in a mutant generated by disruption of the diguanylate cyclase 1 gene of *Acetobacter xylinum* BPR 2001. *Appl Microbiol Biotechnol* **65:** 315–322.

Boles, B.R., and McCarter, L.L. (2002) *Vibrio parahaemolyticus scrABC*, a novel operon affecting swarming and capsular polysaccharide regulation. *J Bacteriol* **184:** 5946–5954.

Bomchil, N., Watnick, P., and Kolter, R. (2003) Identification and characterization of a Vibrio cholerae gene, mbaA, involved in maintenance of biofilm architecture. *J Bacteriol* **185:** 1384–1390.

Brown, N.L., Misra, T.K., Winnie, J.N., Schmidt, A., Seiff, M., and Silver, S. (1986) The nucleotide sequence of the mercuric resistance operons of plasmid R100 and transposon Tn*501*: further evidence for *mer* genes which enhance the activity of the mercuric ion detoxification system. *Mol Gen Genet* **202:** 143–151.

Camilli, A., and Mekalanos, J.J. (1995) Use of recombinase gene fusions to identify *Vibrio cholerae* genes induced during infection. *Mol Microbiol* **18:** 671–683.

Chae, K.S., and Yoo, O.J. (1986) Cloning of the lambda resistant genes from *Brevibacterium albidum* and *Proteus vulgaris* into *Escherichia coli*. *Biochem Biophys Res Commun* **140:** 1101–1105.

Chan, C., Paul, R., Samoray, D., Amiot, N.C., Giese, B., Jenal, U., *et al.* (2004) Structural basis of activity and allosteric control of diguanylate cyclase. *Proc Natl Acad Sci USA* **101:** 17084–17089.

Chang, A.L., Tuckerman, J.R., Gonzalez, G., Mayer, R., Weinhouse, H., Volman, G., *et al.* (2001) Phosphodiesterase A1, a regulator of cellulose synthesis in *Acetobacter xylinum*, is a heme-based sensor. *Biochemistry* **40:** 3420–3426.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res* **14:** 1188–1190.

D'Argenio, D.A., and Miller, S.I. (2004) Cyclic di-GMP as a bacterial second messenger. *Microbiology* **150:** 2497–2502.

D'Argenio, D.A., Calfee, M.W., Rainey, P.B., and Pesci, E.C. (2002) Autolysis and autoaggregation in *Pseudomonas aeruginosa* colony morphology mutants. *J Bacteriol* **184:** 6481–6489.

Delgado-Nixon, V.M., Gonzalez, G., and Gilles-Gonzalez, M.A. (2000) Dos, a heme-binding PAS protein from *Escherichia coli*, is a direct oxygen sensor. *Biochemistry* **39:** 2685–2691.

Drenkard, E., and Ausubel, F.M. (2002) *Pseudomonas* biofilm formation and antibiotic resistance are linked to phenotypic variation. *Nature* **416:** 740–743.

Galperin, M.Y. (2004) Bacterial signal transduction network in a genomic perspective. *Environ Microbiol* **6:** 552–567.

Galperin, M.Y., Natale, D.A., Aravind, L., and Koonin, E.V. (1999) A specialized version of the HD hydrolase domain implicated in signal transduction. *J Mol Microbiol Biotechnol* **1:** 303–305.

Galperin, M.Y., Nikolskaya, A.N., and Koonin, E.V. (2001) Novel domains of the prokaryotic two-component signal transduction systems. *FEMS Microbiol Lett* **203:** 11–21.

Garcia, B., Latasa, C., Solano, C., Portillo, F.G., Gamazo, C., and Lasa, I. (2004) Role of the GGDEF protein family in *Salmonella* cellulose biosynthesis and biofilm formation. *Mol Microbiol* **54:** 264–277.

Gralla, J.D. (2005) *Escherichia coli* ribosomal RNA transcription: regulatory roles for ppGpp, NTPs, architectural proteins and a polymerase-binding protein. *Mol Microbiol* **55:** 973–977.

Gronewold, T.M., and Kaiser, D. (2001) The *act* operon controls the level and time of C-signal production for *Myxococcus xanthus* development. *Mol Microbiol* **40:** 744–756.

Hecht, G.B., and Newton, A. (1995) Identification of a novel response regulator required for the swarmer-to-stalked-cell transition in *Caulobacter crescentus*. *J Bacteriol* **177:** 6223–6229.

Huber, B., Riedel, K., Kothe, M., Givskov, M., Molin, S., and Eberl, L. (2002) Genetic analysis of functions involved in the late stages of biofilm development in *Burkholderia cepacia* H111. *Mol Microbiol* **46:** 411–426.

Hurley, J.H. (2003) GAF domains: cyclic nucleotides come full circle. *Sci STKE* **2003:** PE1.

Ismail, T.M., Hart, C.A., and McLennan, A.G. (2003) Regulation of dinucleoside polyphosphate pools by the YgdP and ApaH hydrolases is essential for the ability of *Salmonella enterica* serovar *typhimurium* to invade cultured mammalian cells. *J Biol Chem* **278:** 32602–32607.

Jenal, U. (2004) Cyclic di-guanosine-monophosphate comes of age: a novel secondary messenger involved in modulating cell surface structures in bacteria? *Curr Opin Microbiol* **7:** 185–191.

Johnson, M.R., Montero, C.I., Conners, S.B., Shockley, K.R., Bridger, S.L., and Kelly, R.M. (2005) Population density-dependent regulation of exopolysaccharide formation in the hyperthermophilic bacterium *Thermotoga maritima*. *Mol Microbiol* **55:** 664–674.

Jones, H.A., Lillard, J.W., Jr, and Perry, R.D. (1999) HmsT, a protein essential for expression of the haemin storage (Hms+) phenotype of *Yersinia pestis*. *Microbiology* **145:** 2117–2128.

Karaolis, D.K., Rashid, M.H., Chythanya, R., Luo, W., Hyodo, M., and Hayakawa, Y. (2005a) c-di-GMP (3′-5′-cyclic diguanylic acid) inhibits *Staphylococcus aureus* cell–cell interactions and biofilm formation. *Antimicrob Agents Chemother* **49:** 1029–1038.

Karaolis, D.K., Cheng, K., Lipsky, M., Elnabawi, A., Catalano, J., Hyodo, M., *et al.* (2005b) 3′,5′-Cyclic diguanylic acid (c-di-GMP) inhibits basal and growth factor-stimulated human colon cancer cell proliferation. *Biochem Biophys Res Commun* **329:** 40–45.

Kearns, D.B., Chu, F., Branda, S.S., Kolter, R., and Losick, R. (2005) A master regulator for biofilm formation by *Bacillus subtilis*. *Mol Microbiol* **55:** 739–749.

Kim, Y.R., Lee, S.E., Kim, C.M., Kim, S.Y., Shin, E.K., Shin, D.H., *et al.* (2003) Characterization and pathogenic significance of *Vibrio vulnificus* antigens preferentially expressed in septicemic patients. *Infect Immun* **71:** 5461–5471.

Kirillina, O., Fetherston, J.D., Bobrov, A.G., Abney, J., and Perry, R.D. (2004) HmsP, a putative phosphodiesterase, and HmsT, a putative diguanylate cyclase, control Hms-dependent biofilm formation in *Yersinia pestis*. *Mol Microbiol* **54:** 75–88.

Klausen, M., Aaes-Jorgensen, A., Molin, S., and Tolker-Nielsen, T. (2003) Involvement of bacterial migration in the development of complex multicellular structures in *Pseudomonas aeruginosa* biofilms. *Mol Microbiol* **50:** 61–68.

Ko, M., and Park, C. (2000) Two novel flagellar components and H-NS are involved in the motor function of *Escherichia coli*. *J Mol Biol* **303:** 371–382.

Kulasekara, H.D., Ventre, I., Kulasekara, B.R., Lazdunski, A., Filloux, A., and Lory, S. (2005) A novel two-component system controls the expression of *Pseudomonas aeruginosa* fimbrial *cup* genes. *Mol Microbiol* **55:** 368–380.

Lee, S.H., Butler, S.M., and Camilli, A. (2001) Selection for *in vivo* regulators of bacterial virulence. *Proc Natl Acad Sci USA* **98:** 6889–6894.

Mayer, R., Ross, P., Weinhouse, H., Amikam, D., Volman, G., Ohana, P., *et al.* (1991) Polypeptide composition of bacterial cyclic diguanylic acid-dependent cellulose synthase and the occurrence of immunologically crossreacting proteins in higher plants. *Proc Natl Acad Sci USA* **88:** 5472–5476.

Merkel, T.J., and Stibitz, S. (1995) Identification of a locus required for the regulation of bvg-repressed genes in *Bordetella pertussis*. *J Bacteriol* **177:** 2727–2736.

Merkel, T.J., Stibitz, S., Keith, J.M., Leef, M., and Shahin, R. (1998a) Contribution of regulation by the bvg locus to respiratory infection of mice by *Bordetella pertussis*. *Infect Immun* **66:** 4367–4373.

Merkel, T.J., Barros, C., and Stibitz, S. (1998b) Characterization of the *bvgR* locus of *Bordetella pertussis*. *J Bacteriol* **180:** 1682–1690.

Milton, D.L., Norqvist, A., and Wolf-Watz, H. (1995) Sequence of a novel virulence-mediating gene, *virC*, from *Vibrio anguillarum*. *Gene* **164:** 95–100.

Nikolskaya, A.N., Mulkidjanian, A.Y., Beech, I.B., and Galperin, M.Y. (2003) MASE1 and MASE2: two novel integral membrane sensory domains. *J Mol Microbiol Biotechnol* **5:** 11–16.

Paul, R., Weiser, S., Amiot, N.C., Chan, C., Schirmer, T., Giese, B., *et al.* (2004) Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. *Genes Dev* **18:** 715–727.

Pei, J., and Grishin, N.V. (2001) GGDEF domain is homologous to adenylyl cyclase. *Proteins* **42:** 210–216.

Perry, R.D., Bobrov, A.G., Kirillina, O., Jones, H.A., Pedersen, L., Abney, J., *et al.* (2004) Temperature regulation of the hemin storage (Hms+) phenotype of *Yersinia pestis* is posttranscriptional. *J Bacteriol* **186:** 1638–1647.

Rashid, M.H., Rajanna, C., Ali, A., and Karaolis, D.K. (2003) Identification of genes involved in the switch between the smooth and rugose phenotypes of *Vibrio cholerae*. *FEMS Microbiol Lett* **227:** 113–119.

Römling, U. (2005) Characterization of the rdar morphotype, a multicellular behaviour in Enterobacteriaceae. *CLMS* **62:** 1–13.

Römling, U., Bian, Z., Hammar, M., Sierralta, W.D., and Normark, S. (1998) Curli fibers are highly conserved between *Salmonella typhimurium* and *Escherichia coli* with respect to operon structure and regulation. *J Bacteriol* **180:** 722–731.

Römling, U., Rohde, M., Olsen, A., Normark, S., and Reinköster, J. (2000) AgfD, the checkpoint of multicellular and aggregative behaviour in Salmonella typhimurium regulates at least two independent pathways. *Mol Microbiol* **36:** 10–23.

Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., *et al.* (1987) Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. *Nature* **325:** 279–281.

Rychlik, I., Martin, G., Methner, U., Lovell, M., Cardova, L., Sebkova, A., *et al.* (2002) Identification of *Salmonella enterica* serovar Typhimurium genes associated with growth suppression in stationary-phase nutrient broth cultures and in the chicken intestine. *Arch Microbiol* **178:** 411–420.

Ryjenkov, D.A., Tarutina, M., Moskvin, O.M., and Gomelsky, M. (2005) Cyclic diguanylate is a ubiquitous signaling molecule in Bacteria: insights into biochemistry of the GGDEF protein domain. *J Bacteriol* **187:** 1792–1798.

Sasakura, Y., Hirata, S., Sugiyama, S., Suzuki, S., Taguchi, S., Watanabe, M., *et al.* (2002) Characterization of a direct oxygen sensor heme protein from *Escherichia coli*. Effects of the heme redox states and mutations at the heme-binding site on catalysis and structure. *J Biol Chem* **277:** 23821–23827.

Schmidt, A.J., Ryjenkov, D.A., and Gomelsky, M. (2005) Ubiquitous protein domain EAL encodes cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. *J Bacteriol* (in press).

Simm, R., Morr, M., Kader, A., Nimtz, M., and Römling, U. (2004) GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol Microbiol* **53:** 1123–1134.

Spiers, A.J., Kahn, S.G., Bohannon, J., Travisano, M., and Rainey, P.B. (2002) Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. I. Genetic and phenotypic bases of wrinkly spreader fitness. *Genetics* **161:** 33–46.

Steinberger, O., Lapidot, Z., Ben-Ishai, Z., and Amikam, D. (1999) Elevated expression of the CD4 receptor and cell cycle arrest are induced in Jurkat cells by treatment with the novel cyclic dinucleotide 3′,5′-cyclic diguanylic acid. *FEBS Lett* **444:** 125–129.

Tal, R., Wong, H.C., Calhoon, R., Gelfand, D., Fear, A.L., Volman, G., *et al.* (1998) Three *cdg* operons control cellular turnover of cyclic di-GMP in *Acetobacter xylinum*: genetic organization and occurrence of conserved domains in isoenzymes. *J Bacteriol* **180:** 4416–4425.

Taylor, B.L., and Zhulin, I.B. (1999) PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol Mol Biol Rev* **63:** 479–506.

Thomas, C., Andersson, C.R., Canales, S.R., and Golden, S.S. (2004) PsfR, a factor that stimulates psbAI expression in the cyanobacterium *Synechococcus elongatus* PCC 7942. *Microbiology* **150:** 1031–1040.

Tischler, A.D., and Camilli, A. (2004) Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation. *Mol Microbiol* **53:** 857–869.

Weinhouse, H., Sapir, S., Amikam, D., Shilo, Y., Volman, G., Ohana, P., *et al.* (1997) c-di-GMP-binding protein, a new factor regulating cellulose synthesis in *Acetobacter xylinum*. *FEBS Lett* **416:** 207–211.

Zhulin, I.B., Nikolskaya, A.N., and Galperin, M.Y. (2003) Common extracellular sensory domains in transmembrane receptors for diverse signal transduction pathways in bacteria and archaea. *J Bacteriol* **185:** 285–294.

Zogaj, X., Nimtz, M., Rohde, M., Bokranz, W., and Römling, U. (2001) The multicellular morphotypes of *Salmonella*

typhimurium and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. *Mol Microbiol* **39:** 1452–1463.

Supplementary material

The following supplementary material is available for this article online:

Fig. S1. Consensus sequence of the GGDEF (A) and EAL (B) domains.

Appendix S1. Suggestion for a nomenclature of GGDEF and EAL domains.

Microbial Ecology

Two-Component Signal Transduction Systems, Environmental Signals, and Virulence

E. Calva and R. Oropeza

Instituto de Biotecnología, UNAM, Cuernavaca, Morelos 62210, Mexico

Received: 29 April 2005 / Accepted: 19 September 2005 / Online publication: 31 January 2006

Abstract

The relevance toward virulence of a variety of two-component signal transduction systems is reviewed for 16 pathogenic bacteria, together with the wide array of environmental signals or conditions that have been implicated in their regulation. A series of issues is raised, concerning the need to understand the environmental cues that determine their regulation in the infected host and in the environment outside the laboratory, which shall contribute toward the bridging of bacterial pathogenesis and microbial ecology.

Introduction

The two-component signal transduction systems (TCS) in bacteria are constituted by a membrane-bound sensor histidine kinase that perceives environmental stimuli and a response regulator that affects gene expression (Fig. 1). Upon sensing an environmental signal, the kinase becomes autophosphorylated and then transfers the phosphate to the response regulator which, in turn, binds to DNA regulatory sequences affecting gene expression. Paradigms of such TCS include the NtrB/NtrC system involved in nitrogen assimilation; the chemotactic system CheA/CheY, although the latter interacts with the flagellar switch and not with the DNA; the porin regulon EnvZ/OmpR; and the system for sporulation control KinA,KinB/Spo0A. A multiple-step phosphorelay pathway can also be found among TCS, as in the virulence BvgS/BvgA system [38].

TCS have been implicated in virulence in a number of bacteria. Moreover, various environmental signals or conditions have been invoked to influence such TCS (Table 1). Major questions in this research area involve not only understanding all the environmental signals affecting each of these systems, but also elucidating which are relevant in each environmental niche, including various compartments in the host. Bacteria usually will have to survive and proliferate on mucous surfaces, competing with commensals that usually inhabit on them. Then, they must invade the tissues of the host, grow, and proliferate, only to encounter nonspecific and specific host immune responses. One can hence envision a remarkable number of potential signals that might be key in each step of pathogenesis, ranging from ion concentration to complex macromolecular structures, and encompassing pH, osmolarity, oxygen availability, bile salts, among many possibilities. The various steps in the adaptation to a wide array of environmental niches must require not only the interaction of virulence factors with host tissues and cells, but also the tight regulation of metabolic processes.

Some important questions that come to mind are as follows. Are the signals determined so far that influence the expression of these systems under laboratory conditions relevant in the host? Which are the preferred environments for each pathogen outside the host, and which are the relevant signals? How can we design experiments to determine the relevant environmental signals?

Table 1 and the following summary of TCS in pathogenic bacteria are intended to encourage this line of thinking on these various aspects of the field.

Bordetella pertussis. In *B. pertussis*, the causal agent of whooping cough, and in *B. bronchiseptica*, the BvgS protein (from B*ordetella v*irulence *g*ene) is the transmembrane sensor and BvgA the response regulator that mediate the transition between two distinct phenotypic phases: the Bvg^+ phase, characterized by the expression of adhesins and proteinaceous toxins and of the *vir* activated genes (*vag*s), and the Bvg– phase,

Correspondence to: E. Calva; E-mail: ecalva@ibt.unam.mx

 DOI: 10.1007/s00248-005-0087-1 • Volume 51, 166–176 (2006)

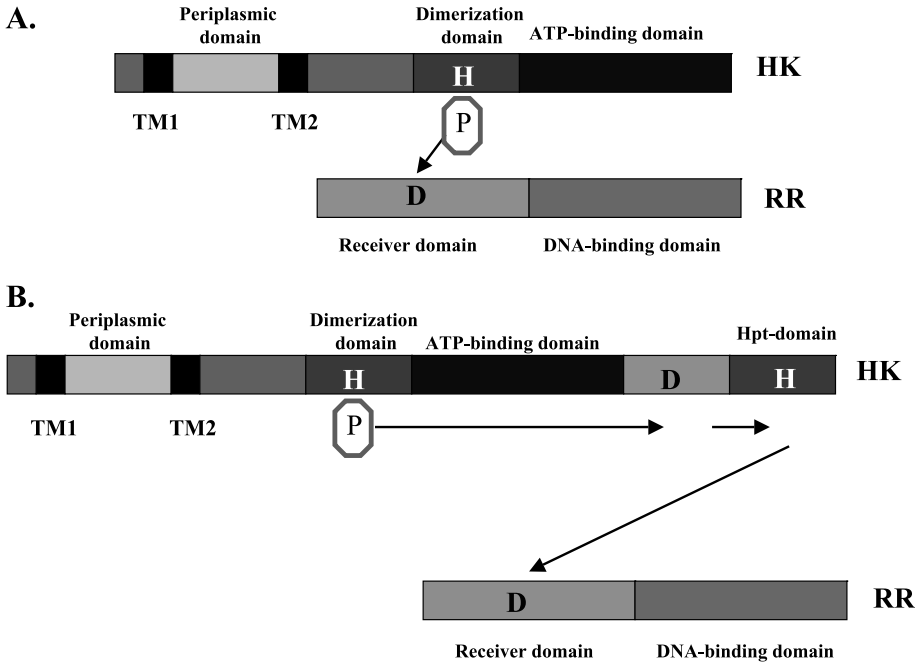


Figure 1. Signal transduction two-component systems in bacteria. (A) The simple prototype phosphotransfer pathway from one histidine kinase sensor (HK) to the response regulator (RR), as for EnvZ/OmpR and PhoQ/PhoP. (B) Multiple-step phosphorelay pathway involving a phosphotransfer scheme of His-Asp-His-Asp, as for BvgS/BvgA. TM: transmembrane domain; Hpt: histidine-phosphotransfer domain.

characterized by motility in *B. bronchiseptica* and by expression of the *vrg* loci in *B. pertussis*. A third intermediate stage has been identified, Bvg^i, characterized by a BvgS mutant locked into this state. It is considered that the Bvg^+ phase is required for respiratory tract colonization and that the Bvg– phase is required for survival and multiplication under nutrient-limiting conditions; moreover, it has been assumed that the phenotypic alteration occurs at some stage in the human host during infection, as no known environmental or animal reservoir has been found for *B. pertussis* [52].

The intermediate state has been hypothesized to be involved in facilitating transmission between hosts. Interestingly, a mutant locked in this state persisted at wild-type levels only in the upper respiratory tract, and a locked mutant also carrying a deletion mutation in *bipA* (the gene for the Bvg^i-phase-specific polypeptide), a cell-surface protein with homology to *Yersinia* invasin and enteropathogenic *Escherichia coli* intimin at the N terminus, displayed a reduced ability to colonize the nasal cavity of mice compared to the single locked mutant. Moreover, it has been observed that the concentration of magnesium and nicotinic acid and temperature regulate these phases: low magnesium and nicotinic acid concentrations, and 37°C, favor the Bvg^+ state [52].

In a recent study [81], the time course of expression of various *bvg*-mediated genes was explored *in vivo* in the mouse. The results were consistent with the *in vitro* characterization of the promoter for the *fha* gene as early, for the *prn* gene as middle, and for *cya* as a late promoter. The filamentous agglutinin is coded by *fha*, *prn* codes for pertactin, an adhesin, and *cya* for adenylate cyclase toxin. Interestingly, the results are suggestive of the presence in mice of an environment that is more highly inducing than the *in vitro* conditions, although the *in vivo* environmental signals sensed by BvgS are still unknown. That is, even the rich laboratory media, used for culturing *B. pertussis*, do not appear to provide the fullest set of signals or conditions for the maximum expression of these genes. In this study, the recombinase-based *in vivo* technology (RIVET) system was used, based on the monitoring of the *in vivo* (or *in vitro*) induction of the TnpR resolvase, under the gene promoter sequences of interest.

Enterococcus faecalis. *E. faecalis* is a Gram-positive bacterium that has emerged as a leading cause of nosocomial infections. It has a great capacity to resist and adapt to many environmental stresses, as it is a nonsporulating microorganism. Seventeen TCS and a single orphan response regulator have been identified in the genome and four have been characterized as being induced by environmental stress, including high (50°C) or low temperature (10°C), and bile salts [33, 47]. Interestingly, one of these TCS, the EtaS/EtaR system (for *e*nterococcal *t*wo-component system *a*), whose expression is induced at 50°C or by bile salts [47] and is homologous to the PhoS/PhoP TCS from *B. subtilis*, has been shown to be involved in virulence, as mutation of the putative response regulator gene, *etaR*, resulted in delayed killing and a 50% higher lethal dose in the mouse peritonitis model [78].

Table 1. Two-component signal transduction systems (TCS) in pathogenic bacteria

TCS sensor/regulator	*Bacterium*	*Criteria for virulence*	*Environmental signals/conditions*[a]	*References*
BvgS/BvgA	*Bordetella pertussis*	Attenuated mutants	Nicotinic acid, $[Mg^{2+}]$, temperature	[52, 81]
EtaS/EtaR	*Enterococcus faecalis*	Attenuated mutant	50°C, bile salts	[47, 78]
ExpS/ExpA	*Erwinia carotovora*	Attenuated mutant	?	[10, 17]
PehS/PehR	*Erwinia carotovora*	Attenuated mutants	$[Ca^{2+}]$	[21]
PmrB/PmrA	*Erwinia carotovora*	Attenuated mutants	$[Fe^{3+}]$, pH	[40]
PhoQ/PhoP	*Erwinia chrysanthemi*	Induced *in planta*; acid-sensitive mutant	$[Mg^{2+}]$, *in planta*	[50]
ArcB/ArcA	*Haemophilus influenzae*	Attenuated mutant	Anaerobiosis	[12]
Sensor ORFs HP244, HP165, HP1364; Regulator ORF HP1365	*Helicobacter pylori*	Attenuated mutants	?	[64, 71]
SenX3/RegX3	*Mycobacterium tuberculosis*	Attenuated mutants	Anaerobiosis	[65, 70]
PhoQ/PhoP	*Mycobacterium tuberculosis*	Attenuated mutant	$[Mg^{2+}]$	[27, 66]
MprB/MprA	*Mycobacterium tuberculosis*	Attenuated mutants	?	[86, 87]
DevS/DevR,	*Mycobacterium tuberculosis*	Altered virulence of mutants	Hypoxia	[6, 11, 51, 65]
TcrY/TcrX, TrcS/TrcR, KdpD/KdpE	*Mycobacterium tuberculosis*	Enhanced virulence of mutants	Trc: early exponential phase Kdp: K^+ limitation	[65]
EnvZ/OmpR	*Salmonella typhimurium*	Attenuated mutants	Osmolarity	[9, 14, 49]
SsrA/SsrB	*Salmonella typhimurium*	Attenuated mutant; located on SPI-II	Cation chelation, acidic pH	[25, 35, 36, 45, 48, 73, 74]
PhoQ/PhoP	*Salmonella typhimurium*	Attenuated mutants	$[Mg^{2+}]$, $[Ca^{2+}]$, cationic antimicrobial peptides	[2, 20, 24, 44, 54]
PmrB/PmrA/PmrD	*Salmonella typhimurium*	Dependent on PhoQ/PhoP; attenuated *pmrF* mutant	Acidic pH, high $[Fe^{3+}]$	[31, 43, 77, 84]
BarA/SirA (UvrY)	*Salmonella typhimurium*	Attenuated mutant	Carbon metabolite	[1, 67, 42, 68, 79]
CpxA/CpxR	*Shigella flexneri*, *Salmonella typhimurium*	Regulation of capsule, streptolysin and streptokinase expression, positive regulation of *hilA*	pH	[56–59]
EnvZ/OmpR	*Shigella flexneri*	Attenuated mutants for epithelial cell invasion and spreading, and for virulence in the Sereny test	Osmolarity	[4, 5]
AgrC/AgrA	*Staphylococcus aureus*	Exotoxin production	Seven to nine aminoacyl peptide	[60, 61]
SaeS/SaeR	*Staphylococcus aureus*	Nuclease and coagulase production	High salt, low pH, glucose, subinhibitory antibiotics	[26, 60]
SrrA/SrrB	*Staphylococcus aureus*	Toxic shock syndrome toxin production; diminished virulence upon overexpression	Oxygen levels	[60, 69, 85]
CovS/CovR	*Streptococcus pyogenes*, *S. agalactiae*	Altered virulence of mutants	$[Mg^{2+}]$	[16, 29, 30, 34, 41, 46]
Ihk/Irr	*Streptococcus pyogenes*	Evasion of PMN-mediated killing	PMN contact, innate immune system, radical oxygen series	[18, 82]
(ToxS)ToxR	*Vibrio cholerae*	Expression of toxin and colonization genes; diminished virulence of mutant strain	Amino acids, carbon dioxide, bile salts, osmolarity	[8, 13, 32, 37, 39, 55, 72, 75, 76]
LuxO	*Vibrio cholerae*	Attenuated mutant strain	?	[53, 80, 88]
OmpR	*Yersinia enterocolitica*	Attenuated mutant strain	Stress signals	[7, 15]
PhoP	*Yersinia pestis*, *Y. pseudotuberculosis*	Attenuated mutant strain	$[Mg^{2+}]$ Stress signals	[28, 63]
YsrR/YsrS	*Yersinia enterocolitica*	Control of a type III secretion system	[NaCl]	[83]

[a] Environmental signals that were shown to function under laboratory conditions or were associated to function.

Erwinia. *E. carotovora* and *E. chrysanthemi* are members of Enterobacteriaceae that cause soft-rot disease in a number of crops. Three TCS that were characterized in *E. carotovora* have been related to virulence. ExpS/ExpA (*ex*tracellular enzyme *p*roduction) bears similarity with the GacS/GacA (*g*lobal *ac*tivator) TCS of *Pseudomonas*; thus, both terminologies were used. Mutants in the *expA* gene showed reduced virulence on potato tubers and diminished production of extracellular enzymes [10, 17]. PehS/PehR [which controls the synthesis of polygalacturonase (Peh)] is the homologue of the PhoQ/PhoP TCS in *Salmonella* (see below). Mutants in both the *pehS* and *pehR* genes are attenuated for virulence in tobacco, and [Ca^{2+}] has been implicated as an environmental signal, specifically in the regulation of an endopolygalacturonase [21].

PmrA/PmrB TCS is the homologue of the one described in *Salmonella*, that is, it renders resistance to polymyxin B (see below). Mutants in both the *pmrA* and the *pmrB* genes result in attenuation in the potato tuber model; furthermore, this TCS responds to [Fe^{3+}] levels and acidic pH [40].

An acid-sensitive mutant has been isolated from *E. chrysanthemi* that maps in a gene very similar to *phoQ*, which codes for a TCS sensor protein involved in virulence in *Salmonella* (see below). The *E. chrysanthemi phoQ* mutant is of particular relevance as the plant apoplast is acidic, and thus bacterial pathogens must particularly resist and adapt to acidic pH. This *phoQ* gene was induced at low Mg^{2+} concentration and *in planta*, suggesting a role for the PhoQ/PhoP system in virulence [50].

Haemophilus influenzae. Life-threatening meningitis is caused by *H. influenzae* encapsulated type b strains. Carriage of unencapsulated *H. influenzae* in the nasopharyngeal tract is common, and a probable cause of infection in otitis media, sinusitis, and pneumonia. Thus, passage from the upper respiratory mucosa to the meninges requires successive adaptations to various environmental changes. Mutants in the *arcA* gene, coding for the response regulator of the ArcB/ArcA (*a*erobic *r*espiration *c*ontrol) system that is activated during the transition from aerobic to anaerobic growth, showed a reduced virulence in the mouse model (intraperitoneal injection), rendering a two-log increase in medium lethal dose. It also showed a marked sensitivity to bactericidal serum activity. Thus, difficulties in switching between the aerobic and anaerobic states might determine a physiological state that makes this bacterium sensitive to the serum complement membrane attack complex [12].

Helicobacter pylori. *H. pylori* is the causative agent of chronic type B gastritis and peptic ulcer disease in humans. Its genome contains sequences that code for only three histidine kinases and five response regulators. Of these, two open reading frames, HP166 and HP1043, coding for two of the response regulator genes, are essential for viability; deletion of a third response regulator gene, HP1021, resulted in a severe growth defect under *in vitro* culture [3]. Interestingly, deletion of open reading frame HP1365, coding for a response regulator, and of open reading frames HP244, HP165, and HP1364, coding for histidine kinases, resulted in mutants that are unable to colonize the stomach of Balb/c mice, an indication that these systems have an essential role in the virulence of *H. pylori*. Moreover, it has been postulated that genes under the control of the HP165/HP166 TCS that are not essential for colonization of the stomach do provide the bacteria with an advantage for survival under competitive conditions in the host [64]. The definition of the sets of genes regulated by these TCS and of the environmental signals that regulate their function is the subject of current research. Interestingly, however, phosphorylation of the receiver domains of HP1043 and of HP1021 does not appear be needed for their function, being thus dependent on a phosphorylation-independent action [71]. Hence, these receiver domains could be the result of the evolution of a typical TCS from a common ancestor that had to deal with a wide variety of environmental conditions, to adapt to the highly specialized niche encountered by *H. pylori*. It thus remains to be established whether HP1043 and HP1021 have a cognate sensor protein and, if so, to subsequently describe the mechanism by which the environmental signal is transduced.

Mycobacterium tuberculosis. *M. tuberculosis* is a pathogen that can persist for many years in the human host, where it can be found growing both intracellularly in macrophages, and extracellularly in the granuloma. Deletion of the SenX3–RegX3 TCS in *M. tuberculosis* resulted in a significant attenuation in the mouse model (intravenous); and an analysis of global gene expression was performed by competitive whole-genome microarray hybridization, between the wild-type and the deletion mutant [65]. Thus, 30 genes were found to be up-regulated and 68 down-regulated by SenX3–RegX3, although the direct effects were not distinguished from the indirect, and the identification of genes required for infection is (still) underway. In an independent study [70], mutants with knockouts of the *senX3* gene also showed reduced virulence in the mouse (intravenous).

Mutants in the *M. tuberculosis phoP* gene resulted in major changes in colony size and morphology, and dramatic changes in cording properties [66]. This is significant as noncording mutants have previously been observed to be attenuated for growth in mice [27]. In concordance, the *phoP* mutant showed impaired growth in bone marrow macrophages and in the spleen, lungs,

and liver of infected mice by the intravenous route [66]. This evidence points toward a major role of the PhoQ/PhoP system in virulence, although the issue of environmental signals affecting its function has not been fully addressed.

A mutant in the *Rv0981* locus was attenuated for growth in the spleen during the acute phase of infection and failed to establish a persistent infection. In the lungs, this mutant was not attenuated for growth but was unable to maintain viability and persist during the latent stage of infection. This phenomenon was tissue-specific, as no significant difference was observed in survival in the liver, between the wild-type and the mutant strain. Hence, *Rv0981* has now been designated as *mprA* (*m*ycobacterium *p*ersistence *r*egulator) [86]. Furthermore, mutagenesis of residue His249 in the MprB sensor and of Asp48 in the MprA regulator proteins, which abolish their ability to be phosphorylated *in vitro*, resulted in growth attenuation in murine macrophages, further supporting their role in virulence [87].

The DevS/DevR TCS was first described for *M. tuberculosis* by subtractive hybridization. The genes formed part of a subset that were expressed at higher levels in a virulent *M. tuberculosis* strain compared to its avirulent counterpart (*d*ifferentially *e*xpressed in a *v*irulent strain) [11]. The gene for the response regulator *devR* was also named *dosR* (for *d*ormancy *s*urvival regulator), as it was found to be up-regulated upon entry into dormancy and to be required to adapt to survival of hypoxia [6]. Furthermore, deletion of *devR* resulted in an alteration of virulence, as the mutants caused mice to die more rapidly (median survival time of 30.5 days) when compared to the wild-type (40.5 days), and the bacterial loads in the lung, liver, and spleen were significantly higher for the mutant. Moreover, whereas the wild-type was killed rapidly in macrophages, there was a rise in the mutant bacteria inside cultured mouse macrophages. This supports the notion for a role for the DevS/DevR TCS in modulating virulent activity [65]. Likewise, the deletion of other genes for TCS components, such as *tcrXY*, *trcS*, and *kdpDE*, also rendered hypervirulent mutants.

In contrast, in another study [51], an *M. tuberculosis devR* mutant was attenuated for virulence in the guinea pig, as it rendered an almost three-log decrease in the bacterial burden in the spleen and a decrease in gross lesions in the lungs, liver, and spleen as compared with the wild-type. Whether the differences in virulence behavior between both studies has to do with the animal model, the mode of inoculation, or some other factor is not known, although in both studies there was an effect on virulence.

Salmonella typhimurium. *S. typhimurium* is the causal agent of gastroenteritis in humans and of a typhoid-like infection in mice. *S. typhi* is the causal agent of typhoid fever, a systemic disease in humans. The *Salmonella* EnvZ/OmpR, a paradigm of TCS, was first described as the regulator of the expression of *o*uter *m*embrane *p*roteins, OmpC and OmpF [38]. It was implicated in virulence when *ompR* mutants were found to be highly attenuated in the mouse model both by the oral and intravenous routes [14]. As a double mutant in *ompC* and in *ompF* was less attenuated, the *ompR* mutant appears to be highly pleiotropic affecting a wider range of genes [9]. In this respect, a random screen for *Salmonella* mutants impaired in cytotoxicity toward the macrophage all located in *ompR*, thus pointing to a regulatory role toward *Salmonella* survival and escape from the macrophage [49]. The role of OmpR in virulence was further defined by its identification through signature-tagged mutagenesis (STM), a method devised to effectively identify microbial genes that are required for the survival and replication in an infected host organism [35].

More recently, OmpR has been found to be a regulator for the *ssrAB* (*s*ecretion *s*ystem *r*egulator) regulon, which codes for a TCS in salmonella patohogenicity island 2 (SPI-2) [48]. SPI-2 has been implicated in systemic disease and proposed to allow replication in macrophages as it was also selected by STM [35, 36]. Moreover, a mutant in *ssrA* was found to be significantly attenuated both via the oral and peritoneal routes in the mouse, and the bacterial load of a mutant in the *ssaJ* (*s*almonella *s*ecretion *a*pparatus) SPI-2 gene was severely reduced in the liver and spleen [73, 74]. Furthermore, mutants in either *ssrA* or *ompR* show diminished survival and replication in cultured mouse macrophages [48]. Low osmolarity, acidic pH, or absence of Ca^{2+} were found to be signals *in vitro* for the SsrA/SsrB-dependent expression of SPI-2 genes, although the EnvZ/OmpR system was found to be partially dependent on the response to these signals [25]. Similarly, by monitoring the activity of the *ssrA* and *ssaG* genes *in vitro*, it was found that they were induced upon ion chelation and by a shift from rich to acidic minimal medium, although both in an EnvZ- and an SsrA/SsrB-dependent manner [45]. Thus, although several signals have been implicated in laboratory media, the signals in other environments are still to be defined.

The *phoP* locus was initially defined as one of two loci (together with *phoN*) deemed necessary for the expression of a nonspecific acid phosphatase by *S. typhimurium* [44]. Mutants in the *phoP* gene, coding for the response regulator, and in the *phoQ* gene, coding for the sensor histidine kinase, were found to be severely attenuated for virulence in the mouse by the intraperitoneal route, as well as in the *pagC* gene (*p*hoP-*a*ctivated *g*ene) [54]. Moreover, *phoP* had been previously found to be required for resistance to microbicidal proteins from phagocytic cells [20]. Upon observation that PhoP controlled the *mgtA* and *mgtCB* genes, coding for two high-affinity Mg^{2+}

transporters whose expression is induced in low Mg^{2+} concentration, the expression of several *pag* genes was explored and found to be repressed at physiological concentrations of divalent ions. Thus, PhoQ was established as a Mg^{2+} sensor protein whose periplasmic sensing domain becomes modified upon binding of Mg^{2+}; moreover, a mutant *phoP* allele that is harbored by a strain attenuated for virulence in the mouse was found to be less responsive to Mg^{2+} [24].

Of the 25 or so loci regulated by PhoP, seven were found to be dependent on a functional PmrA protein, the response regulator of the PmrB/PmrA TCS which is, in turn, regulated by PhoQ/PhoP. The Pmr system regulates resistance to the antibiotic polymyxin B. Transcription of the PmrA-dependent loci were induced by either Mg^{2+} limitation or mild acidification, whereas transcription of a PmrA-independent gene was activated by Mg^{2+} limitation but not acid pH [77]. It has been recently established that the PmrD protein binds to the phosphorylated form of PmrA thus preventing its intrinsic dephosphorylation, and that induced by the cognate PmrB sensor kinase, thus promoting PmrA-mediated transcription. In addition, the expression of the PmrA-activated gene *pbgP* is promoted by high Fe^{3+} concentration in a PhoQ/PhoP-independent fashion [84]. Thus, the current model shows that the PmrB sensor responds to high Fe^{3+} concentrations independent of PhoQ/PhoP, and that responses to Mg^{2+} limitation are attributable to PhoQ [43, 84]. This is a good example of how a bacterium integrates multiple signals into a cellular response.

It is of interest to note that a mutant in the *pmrF* gene, of the *pmrHFIJKLM* operon that determines resistance to polymyxin by modifying the lipopolysacharide on the outer membrane, was attenuated for virulence in the mouse by the oral (but not by the intraperitoneal) route [31]. This suggests that the *pmr* operon has a role in the initial stages of the oral infection, perhaps by rendering resistance to cationic antimicrobial peptides (CAMPs), which are secreted by epithelia at mucosal and skin surfaces.

It has been shown that exposure to sublethal doses of CAMPs activates the PhoQ/PhoP and RpoS virulence regulons. CAMPs, which are also present as a non-oxidative killing activity in phagocyte vacuoles, have been proposed as an important environmental signal recognized by bacteria upon colonization of animal tissues [2].

The *s*almonella *i*nvasion *r*egulator gene, *sirA*, was isolated upon a random screen to identify genes that positively regulated the *prgHIJK* operon, which is required for crossing the intestinal mucosa in the initial stages of infection and is located in SPI-1, the *Salmonella* pathogenicity island-1 [42]. The SirA homologue in *E. coli* is UvrY; SirA is the response regulator for the BarA (*b*acterial *a*daptative *r*esponse) sensor regulator [68]. BarA has a predicted secondary structure similar to the BvgS hybrid sensor kinases, which contains both receiver and transmitter domains. SirA is a global regulator of pathogenicity that activates the expression of HilA, the transcriptional regulator of SPI-1 coded within this island that belongs to the OmpR/ToxR family of transcriptional regulators [1, 79]. The BarA/SirA TCS has been proposed to play a role in the switch between glycolytic and gluconeogenic carbon sources in the environment, although the exact nature of the physiological stimulus for BarA has not been identified so far [67].

Mutants in *sirA* have been determined to be defective in terms of their ability to invade HEp-2 cells, which was similar to the defect observed for *hilA* and *prgHIJK* mutants [42]. In addition, mutants in *sirA* as well as mutants in *hilA* showed diminished capability to cause intestinal secretory and inflammatory responses in a bovine ligated ileal loop model for gastroenteritis, although these were not attenuated for virulence in the mouse via the oral route [1].

Our group has been particularly interested in the study of OmpS1 and OmpS2 quiescent porins in *Salmonella*. The *ompS1* and *ompS2* genes are subject to a tight negative regulation: *ompS1* is part of the H-NS regulon and *ompS2* is positively regulated by the LeuO regulator of the LysR family. Both genes are regulated by the EnvZ/OmpR TCS, although *ompS1* has also an OmpR-independent promoter [19, 22, 62]. Thus, our current interest is to understand the environmental signals in the mouse that determine expression, as mutants in these porin genes and in their regulators are significantly reduced in virulence (Rodríguez-Morales *et al.*, unpublished data).

Shigella flexneri. *S. flexneri* is the causative agent of bacillary dysentery, an invasive disease of the lower gut where the bacteria enter and replicate within colonic epithelial cells and move between cells. *S. sonnei* is a causal agent of gastroenteritis. It has been observed that the CpxA sensor protein is a regulator for the *S. sonnei virF* master regulator gene in response to pH [58]. Moreover, CpxR is the cognate response regulator for *virF* regulation [59]. Interestingly, in *Salmonella*, CpxA (but not CpxR) activates the expression of the *hilA* locus at low pH, suggesting that CpxA can interact with other regulators [57]. In this respect, it has been recently proposed that the CpxA sensor positively regulates post-transcriptionally the InvE virulence regulator [56].

The EnvZ/OmpR TCS has been found to control the expression of a virulence invasion gene (*vir*), whose expression, in turn, was enhanced at high osmolarity; and a mutant in the *ompB* (*ompRenvZ*) operon and in *envZ* showed decreased invasion of epithelial cells and lacked virulence in the Sereny keratoconjuntivitis test

[4]. Moreover, the OmpC major porin, regulated by EnvZ/OmpR, was found to be required for invasion of cultured epithelial cells and intracellular spreading, and to affect virulence by the Sereny test [5].

Staphylococcus aureus. *S. aureus* is a common cause of many infections, including those of the skin, wounds, heart (endocarditis), bone (osteomyelitis), central nervous system, of pneumonia, and toxic shock syndrome. The AgrC/AgrA TCS (for *a*ccessory *g*ene *r*egulator) in *S. aureus* has been shown to regulate the virulon by up-regulating several exotoxin and capsular polysaccharide genes, and down-regulating several surface proteins [61]. This TCS responds to posttranslationally modified seven to nine aminoacyl residue peptides that are encoded by *agrD* within the *agrBDCA* operon. Thus, these peptides act as autoinducers and, in addition, as sensors of population density [60].

Three other TCS are involved in the regulation of the virulon: SaeS/SaeR (S. a*ureus* *e*xoprotein expression), ArlS/ArlR (*a*utolysis-*r*elated *l*ocus), and SrrA/SrrB (*s*taphylococcal *r*espiratory *r*esponse) [23, 60]. The *saePQRS* regulon is also autoinduced and regulates many extracellular protein genes, in particular the nuclease and the coagulase. The *sae* transcription pattern is affected by several environmental stimuli (such as high salt, low pH, glucose, and subinhibitory antibiotic concentrations), and also is growth-phase-dependent, in a rather complex regulatory network [26, 60].

Interestingly, overexpression of the *srrAB* operon cloned on a plasmid decreased virulence in the rabbit endocarditis model [69]. SrrA/SrrB appears to repress virulence factors under low-oxygen conditions; *srrAB* mutants are profoundly growth-defective in the absence of oxygen [60, 85].

Streptococcus. *S. pyogenes* (group A streptococci) can produce a variety of symptoms in humans, ranging from superficial wounds or pharyngeal mucosa infections to invasive infections of deep tissues or the bloodstream. *S. agalactiae* (group B streptococci) is a leading cause of invasive infections that lead to pneumonia, septicemia, and meningitis.

The CovS/CovR (*c*ontrol *o*f *v*irulence) TCS, also known as CsrS/CsrR (*c*apsule *s*ynthesis *r*egulator), negatively controls several proven or putative virulence factors, such as the hyaluronic acid capsule, cysteine protease (pyrogenic exotoxin), streptokinase, streptolysin S, and streptodornase, in group A streptococci. Accordingly, *csrR* mutants, but not the wild-type, produce necrotizing lesions in a mouse model of subcutaneous infections [29, 34]. Moreover, spontaneous mutants in the CovS/CovR TCS, which result in loss of function, have been observed to induce enhanced virulence in the mouse. Interestingly, coinoculation of the *csrR* mutant and the wild-type strains, in a 1:1 ratio, enhances growth of the wild-type by 3.5 orders of magnitude, possibly by creating a proper microenvironment for increased infection [16].

CovR has a pleiotropic effect, influencing transcription of as many as 15% of all chromosomal genes (n = 271). From these, 32 transcripts code for known or putative virulence-associated proteins. Some of the genes controlled by CovR *in vitro* were also controlled in the infected host; this was determined by performing RT–PCR of selected genes on total RNA from infected mouse tissues, 2 days after inoculation [29].

In contrast, mutants in the *covS covR* orthologues in group B streptococci were attenuated for virulence by the intraperitoneal route in the mouse; a *covR* mutant showed an altered expression of virulence factor genes such as those for a hemolysin/cytolysin, a C5 peptidase, and an unrelated cytolytic toxin; and a *covRS* mutant rendered an altered expression of 139 genes. Interestingly, the *covRS* mutant was hyperadherent to epithelial cells and showed an altered extracellular matrix [41, 46].

In terms of a possible environmental signal, it was observed that binding of Mg^{2+} to CovS (CsrS), in group A streptococci, resulted in the repression of CovR-regulated genes, the *hasABC* hyaluronic acid capsule genes, presumably by increasing phosphorylation of CovR. This is in contrast with the PhoQ/PhoP TCS in *Salmonella*, where binding of Mg^{2+} to PhoQ promotes dephosphorylation of PhoP, which renders it inactive as a transcriptional activator [30].

Interestingly, about 16% of the group A streptococcus genes (n = 276) were differentially transcribed, upon group A streptococcus interaction with polymorphonuclear (PMN) leukocytes, which are critical effectors of the human innate immune system. Loci associated with virulence were thus up-regulated; among them were genes that encode secreted proteins known to inhibit PMN phagocytosis or to modulate the innate immune response in the host. Importantly, genes encoding a TCS were up-regulated: *ihk* and *irr* [82]. These genes are located upstream of the *isp* gene, which codes for an *i*mmunogenic *s*ecreted *p*rotein, and were initially identified as homologues for the PhoP/PhoS TCS in *Bacillus subtilis*. Hence their denomination as i*sp*-adjacent *h*istidine *k*inase and i*sp*-adjacent *r*esponse *r*egulator [18].

In the same study, it was determined that the Ihk/Irr TCS contributes to evasion of human innate immunity as significantly more of an isogenic *irr* mutant was killed by human PMNs than the wild-type; and that this TCS enhances bacterial survival after phagocytosis by the PMNs. In addition, the *irr* gene was found to be highly expressed *in vivo* in human infections as determined by real-time PCR, on RNA from swab samples from the posterior pharynx of patients with acute group A streptococcus pharyngitis [82]. Hence, it appears that

PMN contact, factors of innate immunity, or radical oxygen species, or any combination of these, act as environmental stimuli for the Ihk/Irr TCS.

Vibrio cholerae. *V. cholerae*, the causal agent of Asiatic cholera, contains a repertoire of over 20 virulence genes that allow it to penetrate the mucous gel of the small intestine and adhere to the epithelium, where it multiplies and produces the potent cholera exotoxin that causes severe diarrhea associated with the disease. Such virulence regulon is coordinately expressed under the ToxR regulator. ToxR is a transmembrane DNA-binding protein whose carboxy-terminal domain interacts with ToxS, another transmembrane protein that appears to stabilize it in an optimal conformation. ToxR appears to function both as a sensor and as a response regulator, as its amino-terminal cytoplasmic portion shares homology with the response regulators of the TCS. ToxR, in turn, activates the gene for a second regulatory protein, ToxT, a member of the AraC family of regulators [76]. Mutants in *toxR* showed a diminished colonization capacity in human volunteers [37].

ToxT activates a number of virulence genes such as those coding for the toxin coregulated pilus (*tcp*), the accessory colonization factors (*acf*), and the cholera toxin genes, *ctxA* and *ctxB*. ToxR can also directly regulate some genes such as *ompU* and *ompT*, or even the *ctxA* and *ctxB* genes, by a ToxT-independent branch [8]. Expression of the ToxR regulon *in vitro* is affected by pH, temperature, and osmolarity. Nevertheless, the requirements for optimal expression vary between biotypes [13]. Other factors that have been found to influence the expression of the ToxR regulon *in vitro* are amino acids, carbon dioxide, and bile salts.

The observation that cholera toxin is highly expressed in a 10% CO_2 atmosphere might reflect an environmental microaerophilic condition found in the host [75]. Bile, an important constituent of the intestinal lumen, represses the expression of the cholera toxin and TCP genes while increasing motility, possibly through the interaction of other factors aside from ToxR [32]. More recently, it has been observed that cholera toxin expression can be enhanced in a ToxRS-dependent manner by various purified bile salts, at subbacteriocidal concentrations, even though crude bile inhibits cholera toxin production in a ToxT-dependent manner [39, 72]. These observations illustrate the possible complexity of signals and regulatory events regarding an environmental condition.

It is of interest that the ToxR regulon genes are coordinately regulated: the OmpU, TCP, and cholera toxin are optimally expressed at 30°C (and not at 37°C) and at pH 6.5 (and not at pH 8.0), when most of the growth of *V. cholerae* occurs in the upper intestine where the environment is thought to be alkaline, although it is transiently exposed to the acidic pH of the stomach [55]. This might simply reflect the difficulties encountered in extrapolating from *in vitro* to *in vivo* conditions.

The LuxO response regulator for quorum-sensing was found to control virulence, as a *luxO* mutant was profoundly affected in colonization in the infant mouse assay. Moreover, LuxO controls a variety of genes involved in pathogenesis, such as several *tcp* loci, the HapR regulatory gene that affects the secreted hemagglutinin protease, as well as biofilm formation. Thus, LuxO appears to function as a core regulator that coordinates virulence-related phenotypes [80, 88]. Interestingly, mutants in the genes coding for the sensor proteins for LuxO are not attenuated for virulence that is, in *luxS*, *luxP*, *luxQ* (System 2 for autoinducer 2), *cqsA*, *cqsS* (System 1 for autoinducer 1); nor in the gene coding for LuxU, a histidine phosphotransfer protein common to both systems. Moreover, double mutants were also not affected in virulence. This raises the remarkable possibility that LuxO regulates virulence through sensor proteins and signals other than those involved in quorum-sensing, through a TcpP/H-ToxT cascade [53, 80].

Yersinia. *Yersinia* constitutes an important genus of invasive enteropathogens: *Y. enterocolitica* causes enteritis and lymphadenitis in humans, *Y. pseudotuberculosis* is typically associated with acute infections of the mesenteric lymph nodes, and *Y. pestis*, considered a subtype of *Y. pseudotuberculosis*, is the causal agent of bubonic and pneumonic plague.

A *Y. enterocolitica* mutant in *ompR* was found to have an increased sensitivity to high osmolarity, high temperature, and low pH stresses *in vitro*. Furthermore, this mutant proved to be attenuated for virulence in the mouse by the oral route [15]. In another study, the *Y. enterocolitica ompR* mutant was also found to be more sensitive than the wild-type to the stresses mentioned and, in addition, to hydrogen peroxide. Moreover, the mutant had a lower survival in the macrophage [7].

Mutants in the gene coding for the PhoP response regulator in *Y. pestis* were found to have a decreased survival in murine macrophages, were slightly more sensitive to low pH and oxidative killing, and significantly more sensitive to high osmolarity. The *phoP* mutation also rendered a less virulent strain in the mouse [63]. Moreover, in a later study, it was found that a *phoP* mutation also resulted in a *Y. pseudotuberculosis* defective for survival and replication in macrophages: it is possible that the *phoP* mutants are unable to retard phagosome maturation both in *Y. pestis* and in *Y. pseudotuberculosis* [28].

The YsrS/YsrR TCS in *Y. enterocolitica* regulates the YsaE AraC-like regulator, which acts together with the SycB chaperone to regulate the *sycBypBCDA* operon, resembling the type-three secretion system (TTSS) coded

by SPI-1 in *Salmonella*. The expression of the operon was shown to be favored by low temperature (25°C); and high salt was observed to favor expression of the *ysaE* gene, which resulted in enhanced expression of the TTSS proteins [83].

Corollary

Over the past few years, it has been exciting to see the increase in the number of two-component signal transduction systems (TCS) that have been found to be instrumental in bacterial pathogenesis. Hence, the study of TCS is now enriched by trying to understand the actual environmental signals that trigger their response in hosts where they cause disease: this will certainly prove useful not only for further advancement in the field of bacterial pathogenesis, but of microbial ecology as well.

References

1. Ahmer, BMM, van Reeuwijk, J, Watson, PR, Wallis, TS, Heffron, F (1999) *Salmonella* SirA is a global regulator of genes mediating enteropathogenesis. Mol Microbiol 31: 971–982
2. Bader, MW, Navarre, WW, Shiau, W, Nikaido, H, Frye, JG, McClelland, M, Fang, FC, Miller, SI (2003) Regulation of *Salmonella typhimurium* virulence gene expression by cationic antimicrobial peptides. Mol Microbiol 50: 219–230
3. Beier, D, Frank, R (2000) Molecular characterization of two-component systems of *Helicobacter pylori*. J Bacteriol 182: 2068–2076
4. Bernardini, ML, Fontaine, A, Sansonetti, P (1990) The two-component regulatory system OmpR–EnvZ controls virulence of *Shigella flexneri*. J Bacteriol 172: 6274–6281
5. Bernardini, ML, Sanna, MG, Fontaine, A, Sansonetti, P (1993) OmpC is involved in invasion of epithelial cells by *Shigella flexneri*. Infect Immun 61: 2625–3635
6. Boon, C, Dick, T (2002) *Mycobacterium bovis* BCG response regulator essential for hypoxic dormancy. J Bacteriol 184: 6760–6767
7. Brzostek, K, Raczkowska, A, Zasada, A (2003) The osmotic sensor OmpR is involved in the response of *Yersinia enterocolitica* O:9 to environmental stresses and survival within macrophages. FEMS Microbiol Lett 228: 265–271
8. Champion, GA, Neely, MN, Brennan, MA, DiRita, VJ (1997) A branch in the ToxR regulatory cascade of *Vibrio cholerae* revealed by characterization of toxT mutant strains. Mol Microbiol 23: 323–331
9. Chatfield, SN, Dorman, CJ, Hayward, C, Dougan, G (1991) Role of ompR-dependent genes in *Salmonella typhimurium* virulence: mutants deficient in both OmpC and OmpF are attenuated *in vivo*. Infect Immun 59: 449–452
10. Cui, Y, Chattrjee, A, Chattrjee, AK (2001) Effects of the two-component system comprising GacA and GacS of *Erwinia carotovora* subsp. *carotovora* on the production of global regulatory rsmB RNA, extracellular enzymes, and $Harpin_{ECC}$. Mol Plant–Microb Interact 14: 516–526
11. Dasgupta, N, Kapur, V, Singh, KK, Das, TK, Sachdeva, S, Jyothisri, JS, Tyagi, JS (2000) Characterization of a two-component system, *devR–devS*, of *Mycobacterium tuberculosis*. Tuber Lung Dis 80: 141–159
12. DeSouza-Hart, JA, Blackstock, W, Di Modugno, V, Holland, IB, Kok, M (2003) Two-component systems in *Haemophilus influenzae*: a regulatory role for ArcA in serum resistance. Infect Immun 71: 163–172
13. DiRita, VJ, Neely, M, Taylor, RK, Bruss, PM (1996) Differential expression of the ToxR regulon in classical and El Tor biotypes of *Vibrio cholerae* is due to biotype-specific control over *toxT* expression. Proc Natl Acad Sci USA 93: 7991–7995
14. Dorman, CJ, Chatfield, S, Higgins, CF, Hayward, C, Dougan, G (1989) Characterization of porin and *ompR* mutants of a virulent strain of *Salmonella typhimurium*: *ompR* mutants are attenuated *in vivo*. Infect Immun 57: 2136–2140
15. Dorrell, N, Li, SR, Everest, PH, Dougan, G, Wren, BW (1998) Construction and characterisation of a *Yersinia enterocolitica* O8:*ompR* mutant. FEMS Microbiol Lett 165: 145–151
16. Engleberg, NC, Heath, A, Miller, A, Rivera, C, DiRita, VJ (2001) Spontaneous mutations in the CsrRS two-component regulatory system of *Streptococcus pyogenes* result in enhanced virulence in a murine model of skin and soft tissue infection. J Infect Dis 183: 1043–1054
17. Eriksson, ARB, Andersson, RA, Pirhonen, M, Palva, ET (1998) Two-component regulators involved in the global control of virulence in *Erwinia carotovora* subsp. *carotovora*. Mol Plant–Microb Interact 11: 743–752
18. Federle, MJ, McIver, KS, Scott, JR (1999) A response regulator that represses transcription of several virulence operons in the group A *Streptococcus*. J Bacteriol 181: 3649–3657
19. Fernández-Mora, M, Puente, JL, Calva, E (2004) OmpR and LeuO regulate the *Salmonella typhi ompS2* quiescent porin gene. J Bacteriol 186: 2909–2920
20. Fields, PI, Groisman, EA, Heffron, F (1989) A *Salmonella* locus that controls resistance to microbicidal proteins from phagocytic cells. Science 243: 1059–1062
21. Flego, D, Marits, R, Eriksson, ARB, Köiv, V, Karlsson, M, Heikinheimo, R, Palva, TE (2000) A two-component regulatory system, PehR–PehS, controls endopolygalacturonase production and virulence in the plant pathogen *Erwinia carotovora* subsp. *carotovora*. Mol Plant–Microb Interact 13: 447–455
22. Flores-Valdez, MA, Puente, JL, Calva, E (2003) Negative osmoregulation of the *Salmonella ompS1* porin gene independently of OmpR in an *hns* background. J Bacteriol 185: 6497–6506
23. Fournier, B, Hooper, DC (2000) A new two-component regulatory system involved in adhesion, autolysis, and extracellular proteolytic activity of *Staphylococcus aureus*. J Bacteriol 182: 3955–3964
24. García Véscovi, E, Soncini, F, Groisman, EA (1996) Mg^{2+} as an extracellular signal: environmental regulation of *Salmonella* virulence. Cell 84: 165–174
25. Garmendia, J, Beuzón, C, Ruiz-Albert, J, Holden, DW (2003) The roles of SsrA–SsrB and OmpR–EnvZ in the regulation of genes encoding the *Salmonella typhimurium* SPI-2 type III secretion system. Microbiology 149: 2385–2396
26. Giraudo, AT, Mansilla, C, Chan, A, Raspanti, C, Nagel, R (2003) Studies on the expression of regulatory loci *sae* in *Staphylococcus aureus*. Curr Microbiol 46: 246–250
27. Glickman, MS, Cox, JS, Jr, Jacobs, WR (2000) A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of *Mycobacterium tuberculosis*. Mol Cell 5: 717–727
28. Grabenstien, JP, Marceau, M, Pujol, C, Simonet, M, Bliska, JB (2004) The response regulator PhoP of *Yersinia pseudotuberculosis* is important for replication in macrophages and for virulence. Infect Immun 72: 4973–4984
29. Graham, MR, Smoot, LM, Migliaccio, CAL, Virtaneva, K, Sturdevant, DE, Porcella, SF, Federle, MJ, Adams, GJ, Scott, JR, Musser, JM (2002) Virulence control in group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and *in vivo* infection modeling. Proc Natl Acad Sci USA 99: 13855–13860

30. Gryllos, I, Levin, JC, Wessels, MR (2003) The CssR/CssS two-component system of group A *Streptococcus* responds to environmental Mg^{2+}. Proc Natl Acad Sci USA 100: 4227–4232
31. Gunn, JS, Ryan, SS, Velkinburgh, JCV, Ernst, RK, Miller, SI (2000) Genetic and functional analysis of a PmrA–PmrB-regulated locus necessary for lipopolysaccharide modification, antimicrobial peptide resistance, and oral virulence of *Salmonella enterica* serovar Typhimurium. Infect Immun 68: 6139–6146
32. Gupta, S, Chowdhury, R (1997) Bile affects production of virulence factors and motility of *Vibrio cholerae*. Infect Immun 65: 1131–1134
33. Hancock, L, Perego, M (2002) Two-component signal transduction in *Enterococcus faecalis*. J Bacteriol 184: 5819–5825
34. Heath, A, DiRita, VJ, Barg, NL, Engleberg, NC (1999) A two-component regulatory system, CsrR–CsrS, represses expression of three *Streptococcus pyogenes* virulence factors, hyaluronic acid capsule, streptolysin S, and pyrogenic exotoxin B. Infect Immun 67: 5298–5305
35. Hensel, M, Shea, JE, Gleeson, C, Jones, MD, Dalton, E, Holden, DW (1995) Simultaneous identification of bacterial virulence genes by negative selection. Science 269: 400–403
36. Hensel, M, Shea, JE, Waterman, SR, Mundy, R, Nikolaus, T, Banks, G, Vázquez-Torres, A, Gleeson, C, Fang, FC, Holden, DW (1998) Genes encoding putative effector proteins of the type III secretion system of *Salmonella* pathogenicity island 2 are required for bacterial virulence and proliferation in macrophages. Mol Microbiol 30: 163–174
37. Herrington, DA, Hall, RH, Losonsky, G, Mekalanos, JJ, Taylor, MM, Levine, MM (1988) Toxin, toxin-coregulated pili, and the *toxR* regulon are essential for *Vibrio cholerae* pathogenesis in humans. J Exp Med 168: 1487–1492
38. Hoch, JA, Silhavy, TJ (eds.) (1995) Two-Component Signal Transduction. American Society for Microbiology Press, Washington, DC
39. Hung, DT, Mekalanos, JJ (2005) Bile acids induce cholera toxin expression in *Vibrio cholerae* in a ToxT-independent manner. Proc Natl Acad Sci USA 102: 3028–3033
40. Hyytiäinen, H, Sjöblom, S, Palomäki, T, Tuikkala, A, Palva, TE (2003) The PmrA–PmrB two-component system responding to acidic pH and iron controls virulence in the plant pathogen *Erwinia carotovora* ssp. *carotovora*. Mol Microbiol 50: 795–807
41. Jiang, SM, Cieslewicz, MJ, Kasper, DL, Wessels, MR (2005) Regulation of virulence by a two-component system in group B *Streptococcus*. J Bacteriol 187: 1105–1113
42. Johnston, C, Pegues, DA, Hueck, CJ, Lee, CA, Miller, SI (1996) Transcriptional activation of *Salmonella typhimurium* invasion genes by a member of the phosphorylated response-regulator superfamily. Mol Microbiol 22: 715–727
43. Kato, A, Groisman, EA (2004) Connecting two-component regulatory systems by a protein that protects a response regulator from dephosphorylation by its cognate sensor. Genes Dev 18: 2302–2313
44. Kier, LD, Weppelman, RM, Ames, BN (1979) Regulation of nonspecific acid phosphatase in *Salmonella*: *phoN* and *phoP* genes. J Bacteriol 138: 156–161
45. Kim, CC, Falkow, S (2004) Delineation of upstream signaling events in the *Salmonella* pathogenicity island 2 transcriptional activation pathway. J Bacteriol 186: 4694–4704
46. Lamy, MC, Zouine, M, Fert, J, Vergassola, M, Couve, E, Pellegrini, P, Glaser, P, Kunst, F, Msadek, T, Tribu-Cuot, P, Poyart, C (2004) CovS/CovR of group B *Streptococcus*: a two-component regulatory system involved in virulence. Mol Microbiol 54: 1250–1268
47. Le Breton, YL, Boël, G, Benachour, A, Prévost, H, Auffray, Y, Rincé, A (2003) Molecular characterization of *Enterococcus faecalis* two-component signal transduction pathways related to environmental stresses. Environ Microbiol 5: 329–337
48. Lee, AK, Detweiler, CS, Falkow, S (2000) OmpR regulates the two-component system SsrA–SsrB in *Salmonella* pathogenicity island 2. J Bacteriol 182: 771–781
49. Lindgren, S, Stojiljkovic, I, Heffron, F (1996) Macrophage killing is an essential virulence mechanism of *Salmonella typhimurium*. Proc Natl Acad Sci USA 93: 4197–4201
50. Llama-Palacios, A, López-Solanilla, E, Poza-Carrión, C, García-Olmedo, F, Rodríguez-Palenzuela, P (2003) The *Erwinia chrysanthemi phoP–phoQ* operon plays an important role in growth at low pH, virulence and bacterial survival in plant tissues. Mol Microbiol 49: 347–357
51. Malhotra, V, Sharma, D, Ramanathan, VD, Shakila, H, Saini, DK, Chakravorty, S, Das, TK, Li, Q, Silver, RF, Narayanan, PR, Tyagi, JS (2004) Disruption of response regulator gene, devR, leads to attenuation in virulence of *Mycobacterium tuberculosis*. FEMS Microbiol Lett 231: 237–245
52. Martínez de Tejada, G, Cotter, PA, Heininger, U, Camilla, A, Akerley, BJ, Mekalanos, JJ, Miller, JF (1998) Neither the Bvg^- phase nor the *vrg6* locus of *Bordetella pertussis* is required for respiratory infection in mice. Infect Immun 66: 2762–2768
53. Miller, MB, Skorupski, K, Lenz, DH, Taylor, RK, Bassler, BL (2002) Parallel quorum sensing systems converge to regulate virulence in *Vibrio cholerae*. Cell 110: 303–314
54. Miller, SI, Kukral, AM, Mekalanos, JJ (1989) A two-component regulatory system (*phoP phoQ*) controls *Salmonella typhimurium* virulence. Proc Natl Acad Sci USA 86: 5054–5058
55. Miller, V, Mekalanos, JJ (1988) A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. J Bacteriol 170: 2575–2583
56. Mitobe, J, Arakawa, E, Watanabe, H (2005) A sensor of the two-component system CpxA affects expression of the type III secretion system through posttranscriptional processing of *invE*. J Bacteriol 187: 107–113
57. Nakayama, S-I, Kushiro, A, Asahara, T, Tanaka, R-I, Hu, L, Kopecko, DJ, Watanabe, H (2003) Activation of *hilA* expression at low pH requires the signal sensor CpxA, but not the cognate response regulator CpxR, in *Salmonella enterica* serovar Typhimurium. Microbiology 149: 2809–2817
58. Nakayama, S-I, Watanabe, H (1995) Involvement of *cpxA*, a sensor of a two-component regulatory system, in the pH-dependent regulation of expression of *Shigella sonnei virF* gene. J Bacteriol 177: 5062–5069
59. Nakayama, S-I, Watanabe, H (1998) Identification of *cpxR* as a positive regulator essential for expression of the *Shigella sonnei virF* gene. J Bacteriol 180: 3522–3528
60. Novick, RP (2003) Autoinduction and signal transduction in the regulation of staphylococcal virulence. Mol Microbiol 48: 1429–1449
61. Novick, RP, Ross, HF, Projan, SJ, Kornblum, J, Kreiswirth, B, Moghazeh, S (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. EMBO J 12: 3967–3975
62. Oropeza, R, Sampieri, CL, Puente, JL, Calva, E (1999) Negative and positive regulation of the non-osmoregulated *ompS1* porin gene in *S. typhi*: a novel regulatory mechanism that involves OmpR. Mol Microbiol 32: 243–252
63. Oyston, PCF, Dorrell, N, Williams, K, Li, S-R, Green, M, Titball, B, Wren, B (2000) The response regulator PhoP is important for survival under conditions of macrophage-induced stress and virulence in *Yersinia pestis*. Infect Immun 68: 3419–3425
64. Panthel, K, Dietz, P, Haas, R, Beier, D (2003) Two-component systems of *Helicobacter pylori* contribute to virulence in a mouse infection model. Infect Immun 71: 5381–5385
65. Parish, T, Smith, DA, Roberts, G, Betts, J, Stoker, NG (2003) The *senX3–regX3* two-component regulatory system of *Mycobacterium tuberculosis* is required for virulence. Microbiology 149: 1423–1435

66. Pérez, E, Samper, S, Bordas, Y, Guilhot, C, Gicquel, B, Martín, C (2001) An essential role for *phoP* in *Mycobacterium tuberculosis* virulence. Mol Microbiol 41: 179–187
67. Pernestig, A-K, Georgellis, D, Romeo, T, Suzuki, K, Tomenius, H, Normark, S, Melefors, Ö (2003) The *Escherichia coli* BarA–UvrY two-component system is needed for efficient switching between glycolytic and gluconeogenic carbon sources. J Bacteriol 185: 843–853
68. Pernestig, A-K, Melefors, Ö, Georgellis, D (2001) Identification of UvrY as the cognate response regulator for the BarA sensor kinase in *Escherichia coli*. J Biol Chem 276: 225–231
69. Pragman, AA, Yarwood, JM, Tripp, TJ, Schlievert, PM (2004) Characterization of virulence factor regulation by SrrAB, a two-component system in *Staphylococcus aureus*. J Bacteriol 186: 2430–2438
70. Rickman, L, Saldanha, JW, Hunt, DM, Hoar, DN, Colston, MJ, Millar, JBA, Buxton, RS (2004) A two-component signal transduction system with a PAS domain-containing sensor is required for virulence of *Mycobacterium tuberculosis* in mice. Biochem Biophys Res Commun 314: 259–267
71. Schär, J, Sickmann, A, Beier, D (2005) Phosphorylation-independent activity of atypical response regulators of *Helicobacter pylori*. J Bacteriol 187: 3100–3109
72. Schuhmacher, DA, Klose, KE (1999) Environmental signals modulate ToxT-dependent virulence factor expression in *Vibrio cholerae*. J Bacteriol 181: 1508–1514
73. Shea, JE, Beuzón, CR, Gleeson, C, Mundy, R, Holden, DW (1999) Influence of the *Salmonella typhimurium* pathogenicity island 2 type III secretion system on bacterial growth in the mouse. Infect Immun 67: 213–219
74. Shea, JE, Hensel, M, Gleeson, C, Holden, DW (1996) Identification of a virulence locus encoding a second type III secretion system in *Salmonella typhimurium*. Proc Natl Acad Sci USA 93: 2593–2597
75. Shimamura, T, Watanabe, S, Sasaki, S (1985) Enhancement of enterotoxin production by carbon dioxide in *Vibrio cholerae*. Infect Immun 49: 455–456
76. Skorupski, K, Taylor, RK (1997) Control of the ToxR virulence regulon in *Vibrio cholerae* by environmental stimuli. Mol Microbiol 25: 1003–1009
77. Soncini, F, Groisman, EA (1996) Two-component regulatory systems can interact to process multiple environmental signals. J Bacteriol 178: 6796–6801
78. Teng, F, Wang, L, Singh, KV, Murray, BE, Weinstock, GM (2002) Involvement of PhoP–PhoS homologs in *Enterococcus faecalis* virulence. Infect Immun 70: 1991–1996
79. Teplitski, M, Goodier, RI, Ahmer, BMM (2003) Pathways leading from BarA/SirA to motility and virulence gene expression in *Salmonella*. J Bacteriol 185: 7257–7265
80. Vance, RE, Zhu, J, Mekalanos, JJ (2003) A constitutively active variant of the quorum-sensing regulator LuxO affects protease production and biofilm formation in *Vibrio cholerae*. Infect Immun 71: 2571–2576
81. Veal-Carr, WL, Stiblitz, S (2005) Demonstration of differential virulence gene promoter activation *in vivo* in *Bordetella pertussis* using RIVET. Mol Microbiol 55: 788–798
82. Voyich, J, Sturdevant, DE, Braughton, KR, Kobayashi, SD, Lei, B, Virtaneva, K, Dorward, DW, Musser, JM, DeLeo, FR (2003) Genome-wide protective response used by group A *Streptococcus* to evade destruction by human polymorphonuclear leukocytes. Proc Natl Acad Sci USA 100: 1996–2001
83. Walker, KA, Miller, VL (2004) Regulation of the Ysa Type III secretion system of *Yersinia enterocolitica* by YsaE/SycB and YsrS/YsrR. J Bacteriol 186: 4056–4066
84. Wösten, MMSM, Kox, LFF, Chamnongpol, S, Soncini, F, Groisman, EA (2000) A signal transduction system that responds to extracellular iron. Cell 103: 113–125
85. Yarwood, JM, McCormick, JK, Schlievert, PM (2001) Identification of a novel two-component regulatory system that acts in global regulation of virulence factors of *Staphylococcus aureus*. J Bacteriol 183: 1113–1123
86. Zahrt, TC, Deretic, V (2001) *Mycobacterium tuberculosis* signal transduction system required for persistent infections. Proc Natl Acad Sci USA 98: 12706–12711
87. Zahrt, TC, Wozniak, C, Jones, D, Trevett, A (2003) Functional analysis of the *Mycobacterium tuberculosis* MprAB two-component signal transduction system. Infect Immun 71: 6962–6970
88. Zhu, J, Miller, MB, Vance, RE, Dziejman, M, Bassler, BL, Mekalanos, JJ (2002) Quorum-sensing regulators control virulence gene expression in *Vibrio cholerae*. Proc Natl Acad Sci USA 99: 3129–3134

Annu. Rev. Biochem. 2000. 69:183–215

TWO-COMPONENT SIGNAL TRANSDUCTION

Ann M. Stock, Victoria L. Robinson, and Paul N. Goudreau

Center for Advanced Biotechnology and Medicine and Howard Hughes Medical Institute, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, Piscataway, New Jersey 08854; e-mail: stock@cabm.rutgers.edu, robinson@cabm.rutgers.edu, paul.goudreau.b@bayer.com

Key Words gene expression, histidine protein kinase, phosphorylation, protein conformational change, response regulator

■ **Abstract** Most prokaryotic signal-transduction systems and a few eukaryotic pathways use phosphotransfer schemes involving two conserved components, a histidine protein kinase and a response regulator protein. The histidine protein kinase, which is regulated by environmental stimuli, autophosphorylates at a histidine residue, creating a high-energy phosphoryl group that is subsequently transferred to an aspartate residue in the response regulator protein. Phosphorylation induces a conformational change in the regulatory domain that results in activation of an associated domain that effects the response. The basic scheme is highly adaptable, and numerous variations have provided optimization within specific signaling systems. The domains of two-component proteins are modular and can be integrated into proteins and pathways in a variety of ways, but the core structures and activities are maintained. Thus detailed analyses of a relatively small number of representative proteins provide a foundation for understanding this large family of signaling proteins.

CONTENTS

INTRODUCTION

Over a decade ago, the term "two-component" was coined to describe a new class of regulatory systems found in bacteria (1–3). To date, researchers have found hundreds of such systems in eubacteria, archaea, and a few eukaryotic organisms. Two-component systems serve as a basic stimulus-response coupling mechanism to allow organisms to sense and respond to changes in many different environmental conditions. These are sophisticated signaling systems marked by a highly modular design that has been adapted and integrated into a wide variety of cellular signaling circuits.

The prototypical system consists of a histidine protein kinase (HK), containing a conserved kinase core, and a response regulator protein (RR), containing a conserved regulatory domain. Extracellular stimuli are sensed by, and serve to modulate the activities of, the HK. The HK transfers a phosphoryl group to the RR, in a reaction catalyzed by the RR. Phosphotransfer to the RR results in activation of a downstream effector domain that elicits the specific response.

In this review, we summarize the current understanding of the basic biochemical and biophysical mechanisms of two-component signal transduction, with an emphasis on recent advances concerning structure/function aspects. Because of the enormous number of systems that have been identified to date and the pace at which new ones are currently being discovered, it is beyond the scope of this review to present a comprehensive cataloging of systems or components. A number of extensive reviews (4–6) and a monograph (7) have been published detailing the biology and chemistry of two-component systems. In addition to descriptions of individual systems in these sources, additional reviews have focused on well-characterized systems such as bacterial chemotaxis (8), aerobic/anaerobic

regulation in *Escherichia coli* (9, 10), the sporulation system of *Bacillus subtilis* (11, 12), and differentiation in *Caulobacter crescentus* (13–15) and *Myxococcus xanthus* (16, 17). The efforts of a large number of laboratories working with many different systems have gone into elucidation of the basic biology and chemistry of two-component signal transduction. The cumulative data have provided a basic understanding of how these systems transduce extracellular signals and elicit appropriate cellular responses.

PHOSPHOTRANSFER CHEMISTRY

Reactions

The chemistry of the basic two-component phosphoryl transfer signal transduction pathway involves three phosphotransfer reactions and two phosphoprotein intermediates (Figure 1*a*):

1. Autophosphorylation: HK-His + ATP $\Leftrightarrow$ HK-His$\sim$P + ADP
2. Phosphotransfer: HK-His$\sim$P + RR-Asp $\Leftrightarrow$ HK-His + RR-Asp$\sim$P
3. Dephosphorylation: RR-Asp$\sim$P + H_2O $\Leftrightarrow$ RR-Asp + P_i

The γ-phosphoryl group in ATP is first transferred to a conserved His side chain of the HK. The RR then catalyzes the transfer of this phosphoryl group from the phospho-His residue to a conserved Asp side chain within its own regulatory domain. Finally, the phosphoryl group is transferred from the phospho-Asp residue to water in a hydrolysis reaction. All three reactions require divalent metal ions, with Mg^{2+} presumably being the relevant cation in vivo.

Phosphohistidine Chemistry

The HKs are catalytically similar to Ser/Thr/Tyr protein kinases, but the chemistries differ. Ser/Thr/Tyr kinases create phosphoesters, whereas HKs create phosphoramidates. The hydrolysis of phosphoramidates has a significantly greater negative free energy than that of phosphoesters, and the use of these modifications in biological systems differs accordingly (18). The equilibrium for reaction 1 favors the unphosphorylated protein. Given the intracellular ATP/ADP ratio, only a small percentage of the HK population exists in a phosphorylated state. Thus, it is the flux of phosphoryl groups rather than stoichiometric phosphorylation that is relevant to the function of HKs.

The high-energy N$\sim$P bond is ideally suited for phosphoryl transfer. In this capacity, phospho-His occurs as a phosphoenzyme intermediate in proteins such as succinyl-CoA synthetase (19), pyruvate phosphate dikinase (20), and nucleoside diphosphate kinase (21, 22). Phospho-His is also used by the EI and EII enzymes of the phosphoenolpyruvate:sugar phosphotransferase system as phosphotransfer moieties, paralleling their use in two-component pathways (23, 24).

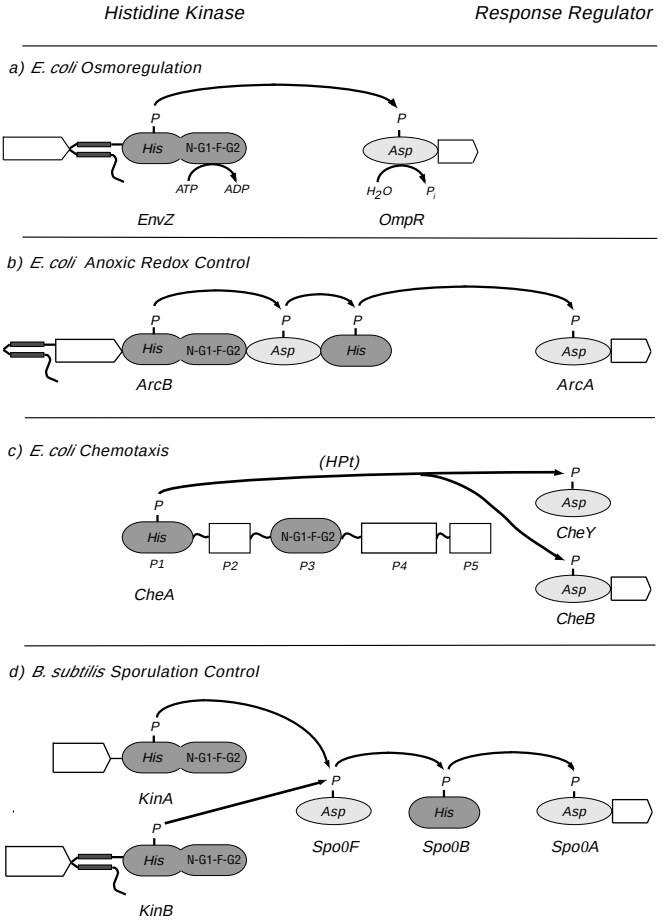


Figure 1 Schematic diagram depicting the modular organization of representative two-component systems. Asp-containing domains are colored *dark gray*, His-containing domains are colored *light gray*, and variable auxiliary domains are colored *white*. (*a*) The prototypical two-component pathway exemplified by the *E. coli* osmoregulatory system uses a single phosphoryl transfer event between the orthodox histidine protein kinase (HK) EnvZ and its cognate response regulator protein (RR) OmpR. (*b*) The *E. coli* Arc system illustrates a phosphorelay involving the hybrid HK ArcB. Depending on aerobic conditions, ArcA is capable of receiving a phosphoryl group from either the catalytic core or the His-containing phosphotransfer (HPt) domain of ArcB. (*c*) The *E. coli* chemotaxis pathway involves an atypical soluble HK CheA that phosphorylates either of two RRs, the single domain RR CheY and the methylesterase CheB. (*d*) The *B. subtilis* sporulation control system is a multicomponent His-Asp-His-Asp phosphorelay system in which all of the signaling domains are independent proteins. Spo0F receives a phosphoryl moiety from either KinA or KinB and subsequently transfers it to the HPt Spo0B, which then phosphorylates the terminal RR Spo0A.

The characteristic acid lability and alkaline stability of phospho-His has provided an experimental basis for its identification in proteins (25–27). Nuclear magnetic resonance (NMR) studies have also been used to detect phospho-His within an HK (28). Phosphorylation of His can occur on either the N1 or N3 position of the imidazole ring, and both forms have been found in proteins. However, in all HKs characterized to date, N3-phospho-His has been observed, or it has been inferred based on hydrolysis rates (26–29).

Phosphoaspartate Chemistry

Phosphorylation of Asp produces a high-energy acyl phosphate. Phospho-Asp enzyme intermediates have been observed in acetate kinase (30) and postulated for the large haloacid dehalogenase super family of hydrolases (comprising P-type ATPases, phosphatases, epoxide hydrolases, and L-2-haloacid dehalogenases) (31). The best characterized phospho-Asp intermediate is that found in the P-type ion-translocating ATPases (32–34). The free energy of hydrolysis of the phospho-Asp residue in these proteins is significantly different than that measured in small-molecule acyl phosphates (35). This has led to the hypothesis that the energy within the acyl phosphate bond may be used to drive conformational changes in proteins (32, 33). In this proposed role, phospho-Asp differs from phospho-Ser/Thr, which alters protein activity by local electrostatic effects (36, 37). Presumably, it is the propensity of phospho-Asp to effect long-range conformational changes that is exploited in RRs.

Acyl phosphates are rapidly hydrolyzed in both acidic and alkaline conditions and have half-lives of several hours in neutral conditions (38, 39). Identification of phospho-Asp is technically difficult and has been achieved in only a small number of RRs (40, 41). Characteristic hydrolysis rates under different conditions have been more frequently used to infer the presence of phospho-aspartyl residues in RRs (42, 43).

The lifetime of phospho-Asp within RRs varies significantly. Typical half-lives range from seconds to hours (3, 43–46). Many RRs have autophosphatase activity that decreases the lifetime of the phosphoprotein (3). In a few cases, the RR stabilizes the phospho-Asp, increasing the half-life significantly beyond that of a typical acyl phosphate. For example, the phospho-Asp in yeast SSK1 has a half-life of $\sim$2 days (47). Phospho-Asps in thermophilic proteins also appear to be stabilized, with half-lives at ambient temperature greater than their mesophilic counterparts (48, 49, 49a).

GENOMIC DISTRIBUTION

His-Asp vs Ser/Thr/Tyr Phosphorylation Pathways

Two-component systems are found in organisms of all domains: Eubacteria, Archaea, and Eukarya. However, their abundance in each domain differs substantially. His-Asp phosphotransfer systems account for the majority of signaling

pathways in eubacteria but are quite rare in eukaryotes, in which kinase cascades involving Ser/Thr and Tyr phosphorylation predominate. Although attempts have been made to rationalize the different distribution of signal transduction mechanisms, no persuasive explanation has been achieved. Clearly both phosphorylation schemes (His-Asp and Ser/Thr/Tyr) can function in both prokaryotes and eukaryotes. His-Asp phosphotransfer systems have been found in several eukaryotic organisms (for reviews, see 50–53), and Ser/Thr and Tyr kinases and phosphatases have been identified in bacteria (for a review, see 54). In fact, these modular signaling components appear to be intertwined within several signaling pathways. In eukaryotes, His-Asp phosphorelays are coupled to mitogen-activated protein (MAP) kinase cascades (55–57) and a cAMP-dependent protein kinase (58–60). In the cyanobacteria *Synechocystis* sp., the coupling of an RR domain to a Ser/Thr protein phosphatase domain and the clustering of genes encoding Ser/Thr kinases or phosphatases with those encoding two-component proteins have led to postulation that, within bacteria as well, multiple phosphorylation schemes may be combined in a single pathway (61, 62).

Occurrence in Prokaryotes

The availability of complete genome sequences has allowed definitive assessment of the prevalence of two-component proteins. In *E. coli* there are 30 HKs (5 of which are hybrid kinases) and 32 RRs (63). However, the number of two-component proteins differs greatly in different bacteria, ranging from 0 in *Mycoplasma genitalium* to 80 in *Synechocystis* sp., in which these proteins account for ~2.5% of the genome (61, 64). Preliminary analyses of other completed bacterial genomes have estimated the numbers of two-component proteins to be as follows: *B. subtilis*, 70 (65); *Haemophilus influenzae*, 9 (64); *Helicobacter pylori*, 11 (64); and *Thermotoga maritima*, 19 (65a). Similar analyses of two archaea have estimated the following: *Methanobacterium thermoautotrophicum*, 24 (66); and *Methanococcus jannaschii*, 0 (64).

Occurrence in Eukaryotes

In contrast to the hundreds of two-component proteins identified in prokaryotes, only a limited number have been found in eukaryotes. In the genome of the budding yeast *Saccharomyces cerevisiae*, there is only one phosphorelay system (SLN1-YPD1-SSK1, SKN7) involved in osmoregulation (56, 67–69). The fission yeast *Schizosaccharomyces pombe* contains an RR (MSC4) that regulates a stress-activated MAP kinase cascade (57). The pathogenic fungus *Candida albicans* contains at least two HKs involved in osmoregulation (CASLN1) (70) and hyphal development (COS1/CANIK1) (70, 71). Homologs of COS1/CANIK1 have been identified in the filamentous fungi *Neurospora crassa* (NIK-1) (72) and *Aspergillus nidulans* (ANNIK1) (71). The slime mold *Dictyostelium discoideum* contains at least 11 HKs (DOKA, DHKA-D, and ESTs) involved in a range of activities that include osmotic response and development (73–77). Two-component proteins are

not limited to eukaryotic microorganisms; they have also been found in plants, such as *Arabidopsis thaliana* (ETR1, ETR2, ERS, and EIN4) (55, 78) and tomato (NR) (79, 80), in which they regulate ethylene-mediated fruit ripening. The only completed animal genome, that of the worm *Caenorhabditis elegans* (81), has revealed no two-component proteins.

Differences Between Prokaryotic and Eukaryotic Components

Several features distinguish eukaryotic two-component systems from those of prokaryotes. Hybrid HKs that contain RR domains are rare in prokaryotes (5 of 30 in *E. coli*) (63), whereas eukaryotic HKs are almost exclusively hybrid kinases; the only known exception is *Arabidopsis* ERS (78). Specific sequences distinguish eukaryotic HKs and RR domains from those of prokaryotes (82), and phylogenetic analyses suggest that the eukaryotic HKs evolved from a single prokaryotic source represented by a cluster of bacterial hybrid HKs (BarA, RcsC, ArcB) (82). Prokaryotic RRs are predominantly transcription factors (at least 25 of 32 in *E. coli*), whereas there is only one known eurkaryotic RR with a DNA-binding domain (*S. cerevisiae* SKN7) (83). In eukaryotes (as in prokaryotes), the ultimate response of two-component pathways is generally regulation of gene expression. However, in eukaryotes, other signaling components that are themselves regulated by two-component proteins effect the final response. This more-complex scheme provides a greater number of potential steps for regulation and may also facilitate transmission of signals from the cytoplasm to the nucleus, where transcription occurs. Despite the differences in configurations of eukaryotic and prokaryotic two-component signaling systems, the structures and activities of the modular domains of the proteins are conserved, and characterization of individual components provides a solid foundation for understanding other family members.

STRUCTURE/FUNCTION OF HISTIDINE PROTEIN KINASES

Activities and Architecture

In typical two-component systems, sensor HKs monitor external stimuli and transmit this information to the RR by a phosphorylation event. Both prokaryotic and eukaryotic HKs contain the same basic signaling components, namely a diverse sensing domain and a highly conserved kinase core that has a unique fold, distinct from that of the Ser/Thr/Tyr kinase super family. The overall activity of the kinase is modulated by input signals to the sensing domain. HKs undergo an ATP-dependent autophosphorylation at a conserved His residue in the kinase core. Autophosphorylation is a bimolecular reaction between homodimers, in which one HK monomer catalyzes the phosphorylation of the conserved His residue in the second monomer

(28, 84–88). Unlike the typical protein kinase cascade in which one protein kinase phosphorylates multiple targets, in two-component systems, the RR stoichiometrically transfers the phosphoryl group from the phospho-HK to a conserved Asp residue in its regulatory domain. Therefore, control in two-component pathways is accomplished through the ability of the HK to regulate the phosphorylation state of the downstream RR. Besides directing the forward phosphorylation reaction, many HKs possess a phosphatase activity, enabling them to dephosphorylate their cognate RRs (89–91). These bifunctional HKs are commonly present in phosphotransfer pathways that need to be shut down quickly.

In the extremely diversified family of HKs, elegant signal transducers have been created from the simple combination of sensing, catalytic, and auxiliary domains. The modular nature of these proteins permits the structural architecture of individual HKs to be adapted to the specific needs of the signaling system. Members of the HK family range in size from $<$40 kDa to $>$200 kDa; the larger HKs consist of five or six structurally and functionally unique domains. Despite this diversity, HKs can be roughly divided into two classes: orthodox and hybrid kinases (6, 92; Figure 1).

Most orthodox HKs, exemplified by the *E. coli* EnvZ protein (Figure 1*a*), function as periplasmic membrane receptors. The osmosensor EnvZ has two transmembrane regions that separate the protein into a periplasmic N-terminal sensing domain and a cytoplasmic C-terminal catalytic region that is designated as the kinase core. Whereas EnvZ represents the most common membrane topology, other HKs contain multiple transmembrane segments. Examples include *Rhizobium meliloti* FixL, involved in controlling nitrogen fixation (93), and UhpB, part of the *E. coli* sugar transport system (94), which have four and eight transmembrane segments, respectively. Not all orthodox kinases are membrane bound. The chemotaxis kinase CheA (95) and the nitrogen regulatory kinase NtrB (96) are examples of soluble cytoplasmic HKs. Soluble HKs can be regulated by intracellular stimuli and/or interactions with cytoplasmic domains of other proteins.

The more elaborate hybrid kinases, found in some prokaryotic and most all eukaryotic systems, constitute the remainder of the HKs. These proteins contain multiple phosphodonor and phosphoacceptor sites. Instead of promoting a single phosphoryl transfer, hybrid kinases use multistep phosphorelay schemes. The overall complexity of the hybrid kinase structure allows different checks and inputs to be integrated into a signaling pathway. *E. coli* ArcB (Figure 1*b*), which functions in the anoxic redox control (Arc) system, has an architecture representative of most hybrid kinases (97). ArcB is composed of two N-terminal transmembrane regions followed by a kinase core, a domain similar to the regulatory domain of RRs, and finally a second His-containing region termed a His-containing phosphotransfer (HPt) domain. The great diversity of hybrid kinases is exemplified by TodS, a unique HK found in the *Pseudomonas putida* toluene degradation pathway (98). TodS is the only known HK containing two perfectly duplicated kinase cores, each containing all conserved HK motifs. TodS also has an N-terminal leucine zipper motif, which is uncommon in prokaryotes. Finally, TodS is a dual-sensing kinase, possessing not only an N-terminal toluene-sensing domain, but also a

putative oxygen-sensing (PAS) domain, identified based on sequence homology to the conventional *R. meliloti* HK FixL (98).

Kinase Catalytic Core

The unifying structural feature of the HK family is the characteristic kinase core composed of a dimerization domain and an ATP/ADP-binding phosphotransfer or catalytic domain (99). This concept began to evolve when deletion analysis of EnvZ indicated that the cytoplasmic domain of the protein contains at least two functionally and structurally separable domains (100). However, this novel kinase architecture was not defined until the overall topology of the HK core was determined from the crystal structure of the C-terminal half of the *T. maritima* CheA protein (101) and the NMR solution structures of both the catalytic (102) and dimerization (103) domains of the *E. coli* EnvZ protein.

The kinase core is ~350 amino acids in length and is responsible for binding ATP and directing kinase transphosphorylation. There are five conserved amino acid motifs present in both eukaryotic and prokaryotic HKs (4, 6, 92). The conserved His substrate is the central feature in the H box, whereas the N, G1, F, and G2 boxes define the nucleotide binding cleft (Figure 2). In most HKs, the H box is part of the dimerization domain. However, for some proteins, like CheA, the conserved His is located at the far N terminus of the protein in a separate HPt domain (Figure 1*c*). The N, G1, F, and G2 boxes are usually contiguous, but the spacing between these motifs is somewhat varied.

The fold of the HK catalytic domain is unlike any known Ser/Thr/Tyr kinase (reviewed in 99, 104), suggesting that the mechanism of action of this kinase may be different than that of other previously characterized prokaryotic and eukaryotic kinases. The main body of the domain is an α/β-sandwich fold consisting of five antiparallel β strands and three α helices (Figure 2*a*). This domain is structurally homologous to the ATPase domains of DNA gyrase B, MutL, and Hsp90 (101, 102). The ATP-binding site of the HKs consists of conserved residues from the N, G1, F, and G2 boxes. In both the CheA and EnvZ structures, this binding pocket is a highly flexible region of the protein. This flexibility may reflect conformational changes that accompany ATP binding. Crystallographic studies of MutL revealed structural shifts in the loops surrounding the nucleotide binding site upon ATP binding and hydrolysis. These changes modulate protein-protein interactions between MutL and other proteins involved in repair (105). Structural similarity has led to the proposal that HKs may undergo similar ATP-coupled conformational rearrangements (99).

Is the novel fold of the HK class of protein kinases indicative of specific catalytic requirements for phosphorylation of His rather than Ser/Thr/Tyr? The answer appears to be no. HK homologs are known in which the phosphorylated His is substituted by Ser (plant, cyanobacteria, and cryptophyte phytochromes) (106–109) or Tyr (*Caulobacter crescentus* DivL) (110). It seems more likely that the HK fold reflects a distinct evolutionary origin and, perhaps, constraints other than those associated with catalysis. The substrate residues of HKs lie within helical

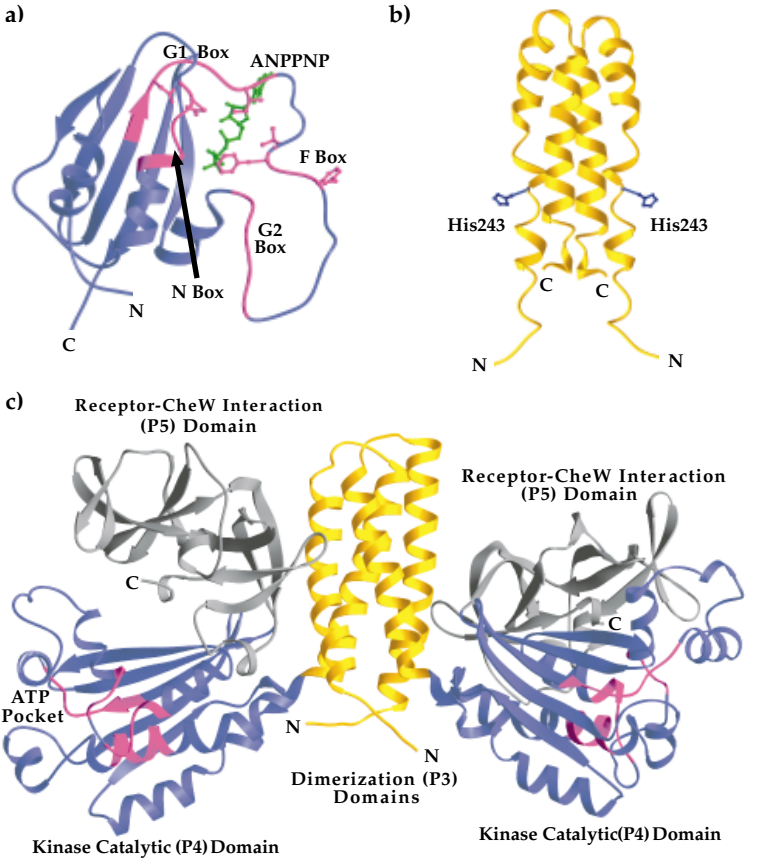


Figure 2 The histidine protein kinase core. The nuclear magnetic resonance structure of the (*a*) catalytic and (*b*) dimerization domains of *E. coli* EnvZ and the (*c*) crystal structure of the C-terminal half of *T. maritima* CheA illustrate the structural conservation of the kinase core. The catalytic domain (*blue*) contains four highly conserved sequence motifs (*magenta*) that form a binding site for ATP (nonhydrolyzable analog shown in *green*). The dimerization domain of EnvZ (*gold*), like that of most histidine protein kinases, contains a conserved His (*blue*) that is the site of phosphorylation. The fold of the dimerization domain of CheA is conserved, but the phosphorylated His is located in an auxiliary His-containing phosphotransfer domain (see Figures 1 and 3). This figure and all other structural images were prepared by using Ribbons software (110a).

structures, unlike the substrate residues of Ser/Thr/Tyr kinases, which are found in more extended secondary structure conformations (111–113). Different active site architecture may be advantageous in sterically facilitating interactions with structurally distinct substrates.

The dimerization domains of both EnvZ (103) and CheA (101) form antiparallel four-helix bundles, similar to the structure of the HPt domains discussed below.

The dimerization domain of EnvZ, like those of the majority of HKs, houses the conserved His, which is positioned midway along the exposed face of helix 1 (Figure 2*b*).

There are unanswered questions concerning the relative three-dimensional arrangement of the catalytic and dimerization domains. The only structural information for an intact kinase core is derived from the atypical kinase CheA, which lacks a contiguous H box (Figure 1*c*). In the crystal structure of the CheA dimer, catalytic domains (P4) protrude out from opposite sides of the central helical structure formed by the dimerization domains (P3) (Figure 2*c*). The P3 and P4 domains are connected through a short, flexible linker extending from the base of the four-helix bundle, with the active site of the catalytic domain facing away from the central core. This orientation may be well suited for transphosphorylation of the CheA HPt domain (P1) that is located at the N-terminus of the protein. However, such an arrangement, if maintained in conventional HKs, would place the H box of the dimerization domain far from the kinase active site. For typical kinases, it is necessary to postulate either a different domain arrangement than that found in CheA or a substantial conformational change that would bring the His and ATP into sufficiently close proximity for phosphotransfer.

Histidine-Containing Phosphotransfer Domain

A small number of two-component systems contain HPt domains. In prokaryotes, HPts are almost exclusively components of hybrid kinases, whereas in eukaryotes, they are found as separate proteins. HPt domains are ~120 amino acids in length and contain a His residue capable of participating in phosphoryl transfer reactions. The HPt domains do not exhibit kinase or phosphatase activity (114) and so are ideally suited to serve as specific cross-communication modules between different proteins. In recent years, the structures of four HPt domains/proteins have been solved. These are the P1 domain of *E. coli* CheA (115), the HPt domain of *E. coli* ArcB (116), the *B. subtilis* Spo0B protein (117), and the *S. cerevisae* YPD1 protein (118, 118a).

It is interesting that all of the HPts share a common four-helix bundle motif (Figure 3) despite their overall lack of sequence similarity. Moreover, the structural architecture of the HPt domain is homologous to the dimerization/His-containing domain of EnvZ, and as in the dimerization domain, the active site His in the HPt is located on a solvent-exposed helical face. Sequence similarity of specific residues surrounding the His residue has prompted postulation of structural and functional roles for these residues (118). Despite the overall fold conservation, specific differences in helix length and orientation in each HPt domain may provide structural features needed for individual functions. These structural variations likely result in modifications of surface properties designed to promote proper intermolecular contacts. The functional forms of YPD1, the ArcB HPt domain, and the CheA P1 domain are monomeric. They have somewhat irregular four-helix bundles, which may make them amenable to interactions with nonhelical partners. In contrast,

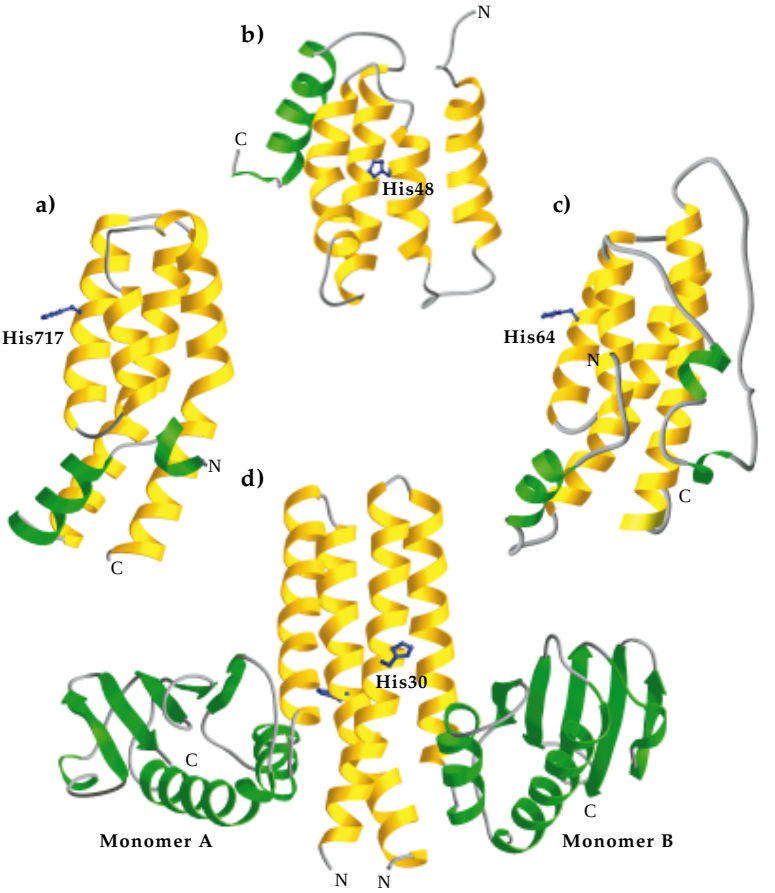


Figure 3 His-containing phosphotransfer (HPt) domains. The HPt domains represented by (*a*) the C-terminal domain of *E. coli* ArcB, (*b*) the P1 domain of *E. coli* CheA, (*c*) *S. cerevisiae* YPD1, and (*d*) *B. subtilis* Spo0B contain four-helix bundles within monomeric or dimeric structures. The conserved His (*blue*) in each structure protrudes from the four-helix bundle (*gold*). The asymmetry of these bundles coupled with additional structural elements (*green*) produces protein surfaces specific for cognate partners. The overall architecture of the Spo0B protein is similar to the CheA kinase core shown in Figure 1*c*.

Spo0B exists as a dimer and has a classic bundle that more closely resembles the EnvZ and CheA dimerization domains.

Sensing Domain

Environmental stimuli are detected either directly or indirectly by the N-terminal sensing domain of the HK. These diverse sensing domains share little primary sequence similarity, thus supporting the idea that they have been designed for

specific ligand/stimulus interactions. In numerous cases, the specific stimuli and mechanism of sensing remain undefined. For others, structural information, such as that obtained recently for the sensing domain of PhoQ (119), may help to define extracellular receptor/ligand interactions. However, the mode of transmission of the signal to the intracellular kinase core remains elusive.

Cytosolic sensing modules have also been integrated into HKs. Examples include the versatile PAS domains. These adaptable domains monitor changes in light, redox potential, oxygen, and small ligands, depending on their associated cofactor (120). PAS domains are small (~100 amino acids) and are typically located adjacent to the last transmembrane region of the sensing domain and N terminal to the kinase domain. The PAS domain of the HK FixL houses a heme group, the coordination state of which regulates kinase activity (121, 122). The crystal structure of this sensing domain (123) shows it to consist of a five-stranded antiparallel β barrel leading to a glutamine-rich helical region called the Q-linker, a motif that has been observed in other HKs (124). KinA, a soluble HK in the *B. subtilis* sporulation control pathway, is unusual in containing three PAS domains (120). Numerous other examples exist (120), highlighting the importance of PAS domains in two-component systems.

Linker Domain

In transmembrane HKs, the sensing domain is connected to the cytoplasmic kinase core through a transmembrane helix and a cytoplasmic linker. Although the least understood segment of any HK, multiple studies indicate that these linker regions are critical for proper signal transduction (125–129). The linker regions are of variable lengths, ranging from 40 to >180 amino acids (65). Computational analyses of sequence similarities of these regions revealed a 50-residue α-helical, coiled coil-like (CC) motif that, in most cases, directly precedes the H box of the kinase domain (130, 131). Similar repeats have also been identified in other multidomain proteins such as adenylyl cyclase, PP2C phosphatase, and diguanylate cyclase/phosphodiesterase (131). These repeats may promote intramolecular associations or, as suggested by a mutational analysis of the EnvZ linker region (127), correct structural alignment of monomers within the HK dimer. Alternatively, because of the dynamic nature of helices, these CCs might be used as a structural relay, sensing conformational changes in the periplasmic region of the kinase and communicating this signal to the kinase core.

STRUCTURE/FUNCTION OF RESPONSE REGULATORS

Activities and Architecture

In most prokaryotic systems, RRs are the terminal component of the pathway, functioning as phosphorylation-activated switches to effect the adaptive response. The RR catalyzes phosphoryl transfer from the phospho-His of the HK to a conserved

Asp in its own regulatory domain. Small molecules such as acetyl phosphate, carbamoyl phosphate, imidazole phosphate, and phosphoramidate can serve as phosphodonors to RRs (132), demonstrating that the RR can catalyze phosphoryl transfer independently of assistance from an HK. As discussed previously, most RRs also catalyze autodephosphorylation, limiting the lifetime of the activated state. Recent data support the view that phosphorylation promotes a conformational alteration affecting a large surface of the RR. This altered molecular surface facilitates a distinct set of inter- or intramolecular interactions through which the response is achieved. This basic scheme allows for a great variety of regulatory mechanisms optimized for the diverse effector functions of different systems.

Most RRs consist of two domains: a conserved N-terminal regulatory domain and a variable C-terminal effector domain (Figure 4). The majority of RRs are transcription factors with DNA-binding effector domains (25 of 32 in *E. coli*) (63). These DNA-binding domains can be subdivided into three major families, represented by OmpR, NarL, and NtrC (14, 7, and 4 members, respectively, in *E. coli*) (4, 63, 133). Not all RRs have DNA-binding domains. A few have C-terminal domains that function as enzymes, such as the chemotaxis methylesterase CheB (134) or the *Dictyostelium* cAMP phosphodiesterase RegA (58–60). Some RRs lack C-terminal effector domains altogether. The chemotaxis protein CheY associates in an intermolecular fashion with its effector protein FliM, a component of the flagellar motor (135). Another isolated regulatory domain is found in the *B. subtilis* sporulation factor Spo0F, which shuttles phosphoryl groups between the HK KinA and the HPt protein Spo0B (136). Thus, Spo0F functions similarly to the RR domains that are incorporated into hybrid kinases (137).

Regulatory Domain

The single-domain chemotaxis protein CheY has served as a representative model for RR regulatory domains (18, 138; Figure 4*a*). This 128-residue protein is a doubly wound α/β protein with a central five-stranded parallel β sheet surrounded by five α helices (139, 140). The site of phosphorylation in CheY is Asp57 (40), located in the solvent-exposed loop between β3 and αC. This residue lies adjacent to other acidic residues, Asp12 and Asp13 in CheY. Two other highly conserved residues, Thr 87 and Lys109, complete the cluster of conserved residues surrounding the active site of the regulatory domain.

The carboxylate side chains of the acidic cluster are involved in coordination of the Mg^{2+} that is required for phosphoryl transfer and dephosphorylation (141–143). Structures of Mg^{2+}-bound CheY indicate an octahedral coordination involving Asp57, Asp12, the backbone oxygen of Asn59, and three water molecules (144, 145). The coordination suggests a mechanism for phosphoryl transfer proceeding through a bipyramidal pentavalent phosphorus transition state (144, 146), an intermediate that is likely to be involved in the autodephosphorylation mechanism as well (141). Thr87 and Lys109, which are not absolutely required for

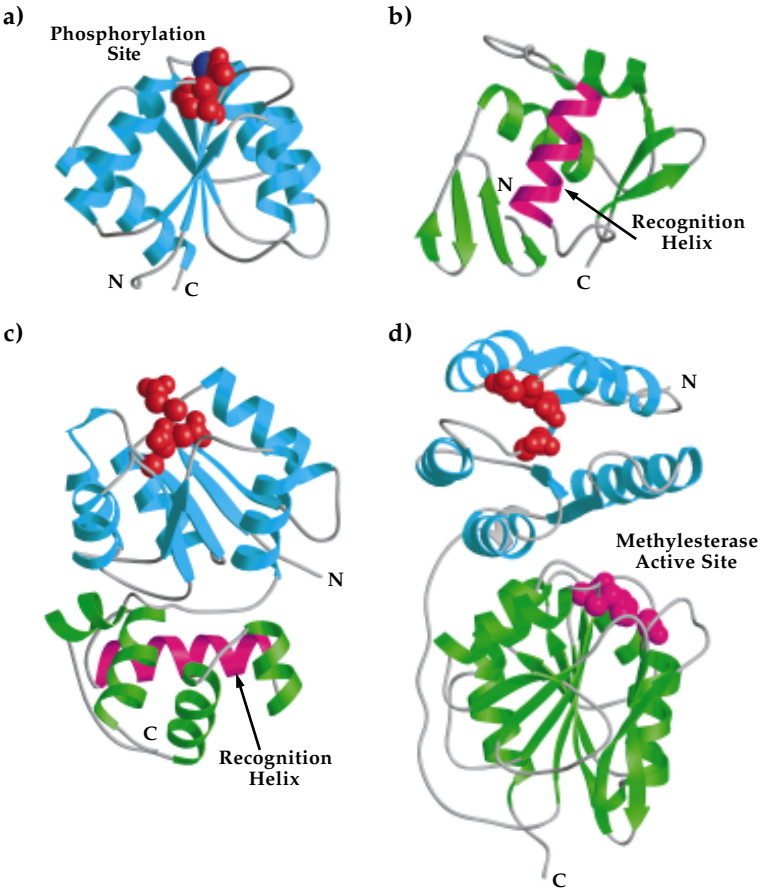


Figure 4 Response regulator protein (RR) family members. RR proteins commonly contain two domains, a conserved regulatory domain (*cyan*) and a variable effector domain (*green*). The conserved regulatory domain, represented by (*a*) *S. typhimurium* CheY, contains a cluster of conserved Asp residues (side chain and Cα atoms shown as *red spheres*) that bind Mg^{2+} (*dark blue*) and form the active site for phosphoryl transfer. The variable architectures of effector domains are illustrated by (*b*) the C-terminal domain of *E. coli* OmpR, (*c*) *E. coli* NarL, and (*d*) *S. typhimurium* methylesterase CheB. The transcription factors NarL and OmpR each contain a recognition helix (*magenta*), whereas CheB contains a catalytic triad (*magenta spheres*). In the intact response regulators NarL and CheB, juxtapositioning of the regulatory domains and the functional regions of the effector domains suggests a structural basis for inhibition.

phosphorylation/dephosphorylation, have been implicated as being involved in the phosphorylation-induced conformational change (142, 147, 148).

In addition to numerous structures of CheY proteins (139, 140, 144, 145, 147, 149–153), X-ray and/or NMR structures are available for the regulatory domains found in Spo0F (154, 155), NtrC (156), PhoB (157), NarL (158), CheB (159), Spo0A (49a), and FixJ (159a). The overall features of these proteins are similar to those of CheY. The most notable structural differences occur in differing lengths and conformations of surface loops (159) and in the orientations of helices relative to each other as they pack against the central β sheet (155).

Effector Domain

A comprehensive description of effector domains will not be attempted because of their great diversity. The majority of effector domains have DNA-binding activity and function to activate and/or repress transcription of specific genes. However, the specific DNA sequences that are recognized, the arrangement of binding sites, and the specific mechanism of transcriptional regulation differ for each RR, even within the same subfamily. Detailed analyses of individual RRs have revealed a great deal of complexity in the functioning of these transcription factors, as illustrated by representative members of the three major subfamilies.

OmpR, a well-characterized member of the largest subfamily of RRs, functions as both an activator and repressor to regulate differentially the expression of the *ompC* and *ompF* genes that encode outer membrane porin proteins. OmpR-DNA interactions have been extensively studied and have established a phosphorylation-regulated hierarchical binding of tandemly arranged OmpR dimers to the F and C boxes preceding the porin genes (160–166). Crystal structures of the DNA-binding domain of OmpR (167, 168) define a novel subclass of winged-helix transcription factors (Figure 4*b*). The fold, conserved in all members of the subfamily (169, 170), 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. Despite structural similarity, subfamily members have different modes of action. For instance, transcriptional activation by OmpR involves interaction with the α subunit of RNA polymerase (171, 172), whereas PhoB interacts with σ^{70} (173, 174).

Another subfamily of RRs is represented by NarL, a transcription factor that both activates and represses genes involved in nitrate and nitrite metabolism (175, 176). NarL-regulated operons are also regulated by the transcription factor Fnr and contain Fnr-binding sites as well as multiple, diversely arranged, "NarL heptamer"-binding sites (177, 178). The crystal structure of NarL has defined a four-helix fold for the 62-residue DNA-binding domain (158; Figure 4*c*). 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 (158).

The most structurally and perhaps functionally complex RR subfamily is that represented by the nitrogen regulatory protein NtrC, a transcriptional enhancer that

activates the σ^{54}-holoenzyme form of RNA polymerase (179, 180). The effector region of this subfamily consists of two domains: an ATPase domain and a helix-turn-helix DNA-binding domain (181–183). NtrC dimers, capable of binding to DNA (184), oligomerize into octamers upon phosphorylation (185). Oligomerization stimulates ATP hydrolysis (186, 187), which provides energy for open complex formation and activation of transcription (188).

Activation by Phosphorylation

The great diversity of effector domains raises the question of how a conserved regulatory domain can function to regulate so many different effector domain activities. From a large accumulation of data, an answer is beginning to emerge. The regulatory domains of RRs are thought to exist in equilibrium between two conformational states, inactive and active. Phosphorylation of the regulatory domain shifts the equilibrium toward the active form. The different molecular surfaces displayed in the two forms can facilitate specific protein-protein (or possibly protein-DNA) interactions. Any type of regulation that can be achieved through inter- or intramolecular interactions can potentially be exploited by this versatile family of proteins.

Accordingly, there are a number of different mechanisms for RR activation. Each mechanism is based on a distinct regulatory interaction(s) specific to the unphosphorylated and/or phosphorylated regulatory domain. In some cases, activation involves a relief of inhibition, as observed in RRs that can be activated by removing the N-terminal regulatory domain (134, 189–191). In other cases, the phosphorylated regulatory domain plays an active role. Phosphorylation can promote dimerization (192, 193), higher-order oligomerization (184, 185, 194), or interactions with other proteins (135, 195) or DNA (45, 160). Some proteins use a combination of these mechanisms (164, 196). Phosphorylation need not necessarily correspond to activation. In yeast osmoregulation, phosphorylated SSK1 is considered the "off" state (197).

Crystal structures of two intact RRs have provided a structural basis for inhibition of effector domain activity by the unphosphorylated regulatory domain. The regulatory domain of the transcription factor NarL blocks access of DNA to the recognition helix (158), and the regulatory domain of chemotaxis methylesterase CheB blocks access of the chemoreceptor substrate to the esterase active site (159; Figure 4*c*, *d*). In both proteins, it appears that phosphorylation-induced activation must involve a repositioning of the N- and C-terminal domains. Despite analogous mechanisms of regulation, different surfaces of the regulatory domains of NarL and CheB are found at the domain interfaces (159).

Structural analysis of a phosphorylated RR has been hindered by the short lifetime of the phosphorylated state. Data from a number of studies are consistent with a long-range conformational change (198–201), but direct evidence has, until recently, been elusive. NMR analyses of unphosphorylated and phosphorylated CheY have shown substantial chemical shift perturbations that map over a large

surface of the protein (199). However, crystal structures of active CheY mutants showed few structural perturbations, other than a different rotameric state of Tyr106 (147, 152).

The concept of an equilibrium between active and inactive states, supported by NMR dynamics studies of NtrC (200), provides an attractive explanation for these seemingly contradictory observations. If two conformations are accessible to both the unphosphorylated and phosphorylated proteins, then the solution conditions, such as those encountered in crystallization, could potentially influence the conformational equilibrium. The equilibrium hypothesis also aids in understanding the observation that phosphorylation-independent activating mutations found for one RR are not generalizable to the family (202–205). A propagated conformational change involves interactions of numerous residues. Even if a residue functions in an analogous manner in different proteins, alteration of any single residue in a given protein may or may not be sufficient to shift the equilibrium markedly between inactive and active states.

Recently determined structures of the phosphorylated regulatory domains of NtrC (206) and FixJ (159a) and of a stably modified phosphoprotein analog, phosphono-CheY (207; CJ Halkides, MM McEvoy, P Matsumura, K Volz, FW Dahlquist, manuscript in preparation) confirm that phosphorylation causes long-range structural perturbations, altering the molecular surface of the regulatory domain (Figure 5). The crystal structure of the CheY analog shows structural changes in αD, β5, and αE. The NMR structure of NtrC shows a displacement of β4, β5, αC, and αD away from the active site, thereby exposing a hydrophobic surface that is proposed to allow transmission of the activation signal to the effector domain (206). Phosphorylation does not alter the overall fold or result in any substantial changes in secondary structure. Rather, the secondary structure elements are slightly repositioned, causing backbone deviations of only a few angstroms. These changes, however, dramatically affect the molecular surface, altering both topological and electrostatic features.

Not surprisingly, the broad surface that is altered by phosphorylation overlaps regions that have previously been identified in many RRs as sites involved in phosphorylation-regulated protein-protein interactions (Figure 5). These surfaces include the interdomain interfaces of NarL (158) and CheB (159), as well as the interaction surfaces of CheY with HK CheA, phosphatase CheZ, and flagellar switch protein FliM (209–216); of Spo0F with HK KinA and HPt Spo0B (217); and of PhoB with HK PhoR (218). Additionally, in Spo0F, this surface is characterized by motions in the microsecond to millisecond time range, implying a conformationally dynamic region (219).

The phosphorylation-induced conformational change affects a large face of the regulatory domain, providing ample molecular surface to be exploited by multiple protein-protein interactions. Indeed, many RRs are involved in phosphorylation-modulated interactions with several different macromolecular targets. These include HKs, auxiliary phosphatases, effector domains, other regulatory domains within dimers, and possibly other components of the transcriptional machinery.

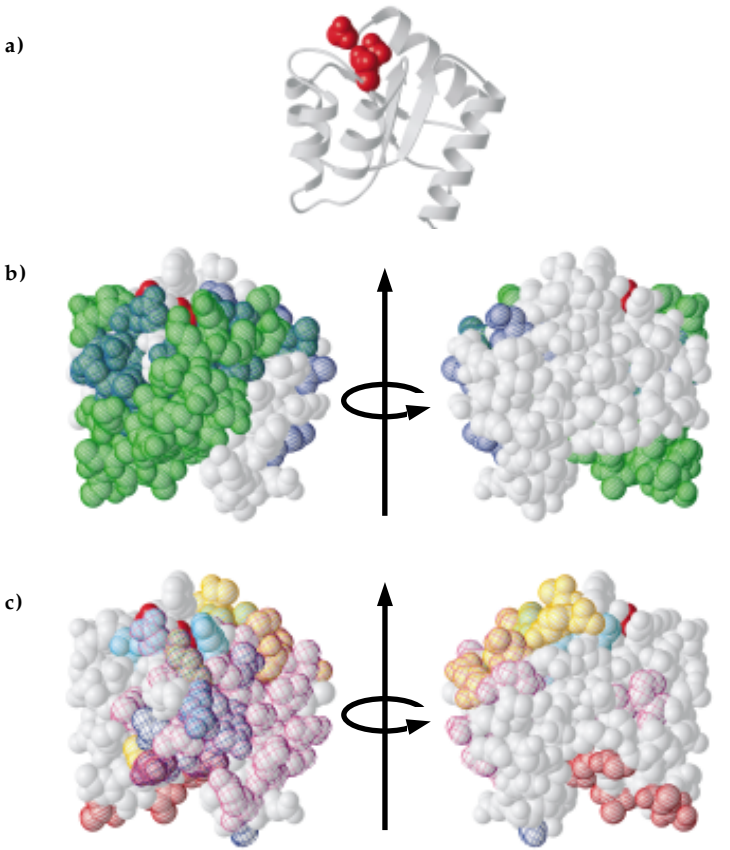


Figure 5 Correlation of phosphorylation-induced conformational changes with protein-protein interaction surfaces in the regulatory domains of response regulator proteins (RR). (*a*) A ribbon diagram of CheY with active-site Asp residues shown in *red*, establishes the orientation of the molecule depicted in subsequent space-filling representations. (*b*) Surfaces involved in phosphorylation-induced conformational changes are depicted in two views related by an 180° rotation about the indicated axis. Residues that differ in the nuclear magnetic resonance (NMR) structures of unphosphorylated and phosphorylated NtrC (206) are mapped in *green* onto the corresponding surface of CheY; residues that show significant NMR chemical shift perturbations in phosphorylated CheY are mapped in *blue* (199); and overlapping residues appear as *cyan*. (*c*) Protein-protein interaction surfaces in different RRs are mapped onto the surface of CheY. Residues involved in interactions of CheY with FliM (209, 213, 216), CheA (214, 215), and CheZ (212, 213, 216) are shown in *magenta*; interactions of Spo0F with KinA and Spo0B (217) in *cyan*; interactions of PhoB with PhoR (218) and a proposed dimerization surface of PhoB (157) in *gold*; the domain interface of NarL (158) in *red*; and the domain interface of CheB (159) in *blue*.

The emerging picture of RRs has significant parallels with the small G proteins of the Ras family. These switch proteins exist in two conformational states determined by the diphosphate or triphosphate state of the bound guanine nucleotide. Large surface regions differ in the "on" and "off" conformations of the G proteins. Subsets of these regions are used for specific interactions with many different signaling partners (220). Thus RRs, which have fundamental structural and catalytic similarity to small-G proteins (221–223), appear to have functional similarity as well.

SYSTEM ARCHITECTURE

The elegance of the two-component system is embodied in its modularity. This modularity extends into the individual systems themselves, in which the basic coupling of HKs and RRs has been adapted to accommodate the specific needs of the signaling pathway. The most commonly occurring phosphotransfer system is composed of a single HK and a single RR. However, there are other systems in which the combination of HK, RR, and HPt domains has been used to create a more complex signaling circuit known as a phosphorelay system. Whereas the classical phosphotransfer pathways are predominant in bacteria, the phosphorelay pathways are more prevalent in eukaryotes. The additional complexity of the phosphorelay system provides for multiple regulatory checkpoints as well as a means of communication between individual signaling pathways.

Phosphotransfer Systems

The majority of two-component systems have a very simplistic design. A transmembrane sensor HK uses a single phosphoryl transfer event to activate a cytoplasmic RR protein that elicits an appropriate adaptive response. The archetype of this basic His-Asp system is the *E. coli* EnvZ-OmpR osmosensing pathway responsible for modulating the expression of the outer membrane porin proteins OmpF and OmpC (224; Figure 1*a*). There are, of course, variations on this simple two-step scheme, in which multiple HKs phosphorylate the same RR or a single HK controls several RRs. For example, in the chemotaxis system, a single HK, CheA, competitively phosphorylates two RRs, CheB and CheY (225; Figure 1*c*). An even more complex scheme occurs in the system that controls nitrate/nitrite-responsive gene expression, in which two HKs, NarX and NarQ, regulate two RRs, NarL and NarP (225).

Phosphorelay Systems

More elaborate versions of the two-component system are the multiple phosphotransfer pathways known as phosphorelays (137, 227, 228). The architecture of the phosphorelay cascade extends the fundamental His-Asp sequence into a four-step reaction. The basic design of the phosphorelay incorporates five phosphoryl

transfer reactions and four phosphoprotein intermediates mechanistically linked by phosphoryl transfer events:

1. Autophosphorylation: HK-His_I + ATP ⇔ HK-His_I∼P + ADP
2. Phosphotransfer I: HK-His_I∼P + RR-Asp_I ⇔ HK-His_I + RR-Asp_I∼P
3. Phosphotransfer II: RR-Asp_I∼P + HPt-His_{II} ⇔ RR-Asp_I + HPt-His_{II}∼P
4. Phosphotransfer III: HPt-His_{II}∼P + RR-Asp_{II} ⇔ HPt-His_{II} + RR-Asp_{II}∼P
5. Dephosphorylation: RR-Asp_{II}∼P + H_2O ⇔ RR-Asp_{II} + P_i

His- and Asp-containing domains are used as phosphotransfer elements. They can exist as isolated domains or can be covalently coupled, as for hybrid kinases discussed previously.

The *B. subtilis* sporulation control system is an example of a His-Asp-His-Asp phosphorelay (136; Figure 1*d*). In this relay, multiple HKs function as phosphoryl donors to Spo0F. The phosphoryl group is subsequently transferred to the HPt protein Spo0B and finally to Spo0A, a transcription factor. Other well-characterized examples of phosphorelays include the *Bordetella pertussis* virulence control system BvgS/BvgA (229) and the *S. cerevisiae* osmoregulation system SLN1/YPD1/SSK1 (68). The multiple domains in phosphorelay systems provide potential for alternate pathways of phosphoryl transfer. In the hybrid kinase ArcB, either His-containing domain (the dimerization domain or the HPt) can be phosphorylated from ATP and can donate phosphoryl groups to the RR ArcA. Data suggest that different pathways are used in aerobic and anaerobic conditions (114).

Integration of Systems

As described above, basic signaling elements can be assembled into either simple phosphotransfer or more elaborate phosphorelay pathways. On yet another level, these distinct signaling pathways can be integrated into cellular networks. In *B. subtilis*, each characterized two-component system appears to interface with at least one other phosphotransfer pathway (230). One example of this integration is between the pathways controlling phosphate utilization (PhoR/PhoP), aerobic and anaerobic respiration (ResE/ResD), and sporulation (KinA-B/Spo0A). Respiration and phosphate utilization are coregulated; phospho-PhoP is required for expression of ResD and vice versa (231, 232). Furthermore, once the cell commits to sporulation, respiration and phosphate utilization are down-regulated. Phospho-Spo0A is a negative regulator of both phospho-ResD and phospho-PhoP and therefore mutually exclusive with both of these responses (231, 233).

In eukaryotes, two-component systems are commonly found as parts of larger signal transduction cascades. *S. cerevisiae* has a single phosphorelay system consisting of a hybrid HK (SLN1), an HPt domain (YPD1), and two RRs (SSK1 and SKN7) (68, 69, 234). Under normal physiological conditions, SSK1 is continually phosphorylated. However, in high osmolarity, SLN1 is less active, and SSK1 becomes dephosphorylated and subsequently activates the downstream

HOG1-dependent MAP kinase cascade (56, 197). Another example is the *Dictyostelium* osmoregulation system (58–60, 75). Phosphorylation of the RR RegA by the hybrid kinase DhkA causes inhibition of the cAMP-dependent phosphodiesterase activity of RegA, increasing cAMP levels, which, in turn, regulate the activity of protein kinase A.

Structural and functional conservation of two-component proteins suggests the potential for phosphotransfer between noncognate pairs. Although such transfer is commonly observed in vitro (27, 235), it appears to be rare in vivo. There are a few systems designed for phosphotransfer between noncognate pairs (176, 236). However, the issue of cross-talk outside such systems remains an open question (237). In at least one case, RR activation, initially attributed to cross-talk, was found to involve the phosphodonor acetyl phosphate (238, 239).

REGULATORY MECHANISMS

The sole purpose of two-component signal transduction systems is to allow for regulation; the signaling pathway merely provides steps at which the flow of information can be modulated. A great diversity of regulatory mechanisms has been overlaid on the central phosphotransfer/phosphorelay pathways, allowing optimization of signal transmission for the specific needs of each system. Some systems output a graded response, such as the EnvZ-OmpR system that mediates the differential expression of porin genes *ompF* and *ompC* (224). Others, such as the *B. subtilis* Spo system that controls commitment to sporulation, output an all-or-nothing response (240). Regardless of the output, both types of systems can involve a significant amount of regulation and often involve a number of auxiliary protein components. The primary targets for regulation are the activities of the HK and dephosphorylation of the RR.

Regulation of Histidine Kinase Activities

HKs can have two activities that determine the level of RR phosphorylation: autophosphorylation activity and RR phosphatase activity. Not all HKs possess phosphatase activity; for some HKs, regulation occurs exclusively at the level of autophosphorylation (241). In many other systems, it is RR phosphatase activity rather than autophosphorylation that is regulated (2, 90, 91, 242, 243). Either of these activities can be regulated directly or indirectly by stimuli.

In typical transmembrane HKs, sensing domains directly bind ligands or detect other physical stimuli. More complex schemes involve indirect detection of signals through interaction with other protein components. For instance, the autophosphorylation activity of the chemotaxis HK CheA, which forms a complex with chemoreceptors and an adapter protein CheW (244, 245), is either inhibited or stimulated by signals transmitted from the chemoreceptors (246, 247). The RR phosphatase activity of the soluble HK NtrB is regulated by an auxiliary protein

P_{II}, whose ability to interact with NtrB depends on its uridylylation state, a modification catalyzed by a uridylyltransferase enzyme, which is itself regulated by intracellular 2-ketoglutarate and glutamine (248). The ability of HK PhoR to mediate repression of the Pho regulon when phosphate is in excess, requires an intact phosphate uptake complex (PstA-C and PstS), together with an accessory protein, PhoU (249). These proteins are proposed to form a membrane-associated repression complex that regulates PhoR activity (250). Regulation of HK activities by intrinsic domains has also been found. The turgor sensor KdpD contains an auxiliary ATP-binding domain, and ATP binding, but not hydrolysis, is required for phosphatase activity (251).

Regulation of Response Regulator Dephosphorylation

As discussed previously, many RRs have autophosphatase activity, with dephosphorylation rates tuned to the time-scale requirements of the specific system. RR dephosphorylation is also influenced by the phosphatase activities of HKs that are regulated by a number of different mechanisms, as discussed above. In most instances, the phosphatase mechanism appears to be distinct from a reversal of phosphotransfer and does not require the H box His of the HK (176, 252–254). In some HKs (89, 251), but not others (46, 91), phosphatase activity is stimulated by ATP.

RR dephosphorylation can also be effected by auxiliary proteins. The *B. subtilis* sporulation system involves a set of highly regulated phosphatases (RapA, RapB, and RapE) that dephosphorylate Spo0F (255), and an unrelated phosphatase (Spo0E) that dephosphorylates Spo0A (256). In bacterial chemotaxis, an auxiliary protein, CheZ, oligomerizes with phospho-CheY and accelerates its dephosphorylation (195, 257, 258). The phosphoAsp phosphatases appear to have great specificity for their substrate proteins, raising the question of whether they directly participate in catalyzing hydrolysis or rather function by stimulating the autophosphatase activity of the RR.

Other Modes of Regulation

A small number of additional regulatory mechanisms are involved in specific systems. Phosphotransfer itself can be regulated. Within the *Agrobacterium tumefaciens* hybrid HK VirA, the C-terminal Asp-containing domain modulates the phosphotransfer ability of the kinase core by physically interacting with the autophosphorylation site (259). Additional regulatory potential exists in systems with an HK that can phosphorylate more than one RR. In these systems, competition for phosphoryl groups can influence activation of different branches of the signaling pathway (176, 225).

A novel regulatory mechanism involving multiple RRs has been postulated for the chemotaxis system of *R. meliloti* (260). This system contains two CheY proteins, CheY1 and CheY2, and lacks a phosphatase CheZ. Phosphorylated CheY2 is

responsible for the motor response, whereas CheY1 regulates the phosphorylation state of CheY2. In the absence of forward phosphotransfer, CheY1 serves as a "phosphatase" for phospho-CheY2, acting as a sink for phosphoryl groups that flow backwards in the pathway through CheA to CheY1.

All of the previously discussed regulatory mechanisms alter the level of phosphorylation of the RR. An alternative strategy is regulation of the level of the RR itself through control of gene expression. Many of the two-component systems that regulate transcription are subject to autoregulation. In these systems, the phosphorylated RR functions as an activator or repressor of the operon encoding the two-component proteins themselves (261–264).

POTENTIAL TARGETS FOR ANTIMICROBIAL THERAPY

The search for structurally unique antibiotics that inhibit new molecular targets has led researchers to prokaryotic two-component systems. Two-component systems are attractive for several reasons. First of all, they are widespread in bacteria and, so far, absent in mammals. Therefore, general HK or RR inhibitors could potentially be broad-spectrum antibiotics. Alternatively, by targeting specific HKs or RRs, selective inhibition may be achieved. The problem is that most two-component systems are nonessential. However, there is an interdependence not only among the proteins in these systems, but among the systems themselves. Cessation or slowing down of these intracellular networks may be a way to effect a cellular shutdown. Perhaps the most attractive reason for targeting two-component systems is that they are used by pathogenic bacteria to control the expression of virulence factors required for infectivity. Several well-characterized virulence systems are the *A. tumefaciens* VirA/VirG system, the *B. pertussis* BvgA/BvgS system, and the *Salmonella typhimurium* PhoP/PhoQ system. Interestingly, some bacteria have developed two-component systems that regulate resistance to certain chemotherapeutics. These include the vancomycin resistance systems in *Enterococcus faecalis* (VanR/VanS) (265) and *Streptococcus pneumoniae* (VncS/VncR) (266), as well as the system associated with tetracycline resistance in *Bacteroides fragilis* (RprX/RprY) (267).

A small number of investigations of two-component inhibitors have appeared in the literature. There are few reported natural inhibitors of two-component systems. Initially, derivatized unsaturated fatty acids were shown to be noncompetitive inhibitors of the autophosphorylation activity of KinA (268), as well as of CheA and NtrB (269). Limited success with these compounds spurred researchers to develop synthetic inhibitors (270). One class is the diphenol-methane compounds, originally identified because of their ability to inhibit NtrB in gram-positive bacteria (271). A number of compounds currently under investigation include hydrophobic tyramines (272), salicylanilides (273), triphenylalkyl derivatives, cyclohexenes, and benzoxazines (270). These compounds not only inhibit two-component

functions but also affect bacterial growth rates. The exact mechanism of action for most of these inhibitors has not been firmly established. Initial observations suggested that the target of these antibacterial compounds was the HK. However recent studies (274, 275) indicate that inhibition is occurring on multiple metabolic and biosynthetic levels. It may be also be possible to extend the use of these two-component inhibitors into the unicellular eukaryotic realm. Specifically, an HK inhibitor was shown to hinder proliferation of the human pathogens *Trypanasoma brucei rhodesiense* and *Leishmania donovani* by interfering with succinyl CoA phosphorylation (276).

SUMMARY AND PERSPECTIVES

In slightly over a decade since two-component systems were first described, numerous studies involving many different systems have provided a fundamental understanding of the essential features of phosphotransfer signal transduction pathways. The biochemical activities of the conserved protein components are known, and three-dimensional structures are available for all of the conserved modular domains. However, there are still several central questions regarding the functioning of HKs and RRs. In many systems, the stimuli sensed by the HKs are not well defined, and the molecular mechanism of signal transmission across the membrane from the sensing domain to the kinase core has not been determined. In RRs, the detailed mechanisms that couple phosphorylation-induced conformational changes in regulatory domains to activation of effector domains remain to be elucidated. In addition to these issues concerning the conserved components, there are many questions regarding regulation within individual pathways. As the number of characterized two-component systems grows, so does the inventory of variations on the basic scheme. Although, from one perspective, these system-specific regulatory mechanisms might be considered mere details, they are essential to the appropriate coupling of stimuli to adaptive responses within each system. With the basic framework in hand, researchers are well poised to advance the understanding of the extremely elegant and sophisticated regulation involved in two-component signaling systems.

ACKNOWLEDGMENTS

This review article reflects the collective views that have developed within the microbial signal transduction community through the years, and we are grateful to our colleagues for many insights and stimulating conversations. We thank Rick Dahlquist, Chris Halkides, Mitsu Ikura, Masayori Inouye, Dorothee Kern, Austin Newton, Mel Simon, and Ann West for providing manuscripts and/or coordinates before public release. Preparation of this review was supported by NIH grant GM47958 and the Howard Hughes Medical Institute.

LITERATURE CITED

1. Nixon BT, Ronson CW, Ausubel FM. 1986. *Proc. Natl. Acad. Sci. USA* 83: 7850–54
2. Ninfa AJ, Magasanik B. 1986. *Proc. Natl. Acad. Sci. USA* 83:5909–13
3. Hess JF, Oosawa K, Kaplan N, Simon MI. 1988. *Cell* 53:79–87
4. Stock JB, Ninfa AJ, Stock AM. 1989. *Microbiol. Rev.* 53:450–90
5. Bourret RB, Borkovich KA, Simon MI. 1991. *Annu. Rev. Biochem.* 60:401–41
6. Parkinson JS, Kofoid EC. 1992. *Annu. Rev. Genet.* 26:71–112
7. Hoch JA, Silhavy TJ. 1995. *Two-Component Signal Transduction*, p. 488. Washington, DC: Am. Soc. Microbiol.
8. Falke JJ, Bass RB, Butler SL, Chervitz SA, Danielson MA. 1997. *Annu. Rev. Cell Dev. Biol.* 13:457–512
9. Iuchi S, Weiner L. 1996. *J. Biochem.* 120:1055–63
10. Unden G, Bongaerts J. 1997. *Biochim. Biophys. Acta* 1320:217–34
11. Hoch JA. 1993. *Annu. Rev. Microbiol.* 47: 441–65
12. Perego M. 1998. *Trends Microbiol.* 6: 366–70
13. Domian IJ, Quon KC, Shapiro L. 1996. *Curr. Opin. Genet. Dev.* 6:538–44
14. Shapiro L, Losick R. 1997. *Science* 276: 712–18
15. Wu J, Newton A. 1997. *Mol. Microbiol.* 24:233–39
16. Kaplan HB, Plamann LA. 1996. *FEMS Microbiol. Lett.* 139:89–95
17. Ward MJ, Zusman DR. 1997. *Mol. Microbiol.* 24:885–93
18. Stock JB, Stock AM, Mottonen JM. 1990. *Nature* 344:395–400
19. Wolodko WT, Fraser ME, James MN, Bridger WA. 1994. *J. Biol. Chem.* 269: 10883–90
20. Spronk AM, Yoshida H, Wood HG. 1976. *Proc. Natl. Acad. Sci. USA* 73:4415–19
21. Parks RE Jr, Agarwal RP. 1973. In *The Enzymes*, ed. PD Boyer, pp. 307–34. New York: Academic
22. Morera S, Chiadmi M, LeBras G, Lascu I, Janin J. 1995. *Biochemistry* 34:11062–70
23. Postma PW, Lengeler JW, Jacobson GR. 1993. *Microbiol. Rev.* 57:543–94
24. McEvoy MM, Dahlquist FW. 1997. *Curr. Opin. Struct. Biol.* 7:793–97
25. Weigel N, Kukuruzinska MA, Nakazawa A, Waygood EB, Roseman S. 1982. *J. Biol. Chem.* 257:14477–91
26. Wylie D, Stock A, Wong C-Y, Stock J. 1988. *Biochem. Biophys. Res. Commun.* 151:891–96
27. Ninfa AJ, Ninfa EG, Lupas AN, Stock A, Magasanik B, Stock J. 1988. *Proc. Natl. Acad. Sci. USA* 85:5492–96
28. Surette MG, Levit M, Liu Y, Lukat G, Ninfa EG, et al. 1996. *J. Biol. Chem.* 271: 939–45
29. Amemura M, Makino K, Shinagawa H, Nakata A. 1990. *J. Bacteriol.* 172:6300–7
30. Anthony RS, Spector LB. 1972. *J. Biol. Chem.* 247:2120–25
31. Ridder IS, Dijkstra BWY. 1999. *Biochem. J.* 339:223–26
32. Jencks WP. 1980. *Adv. Enzymol.* 51: 75–106
33. Tanford C. 1984. *CRC Crit. Rev. Biochem.* 17:123–51
34. Aravind L, Galperin MY, Koonin EV. 1998. *Trends Biochem. Sci.* 23:127–29
35. Pickart CM, Jencks WP. 1984. *J. Biol. Chem.* 259:1629–43
36. Hurley JH, Dean AM, Sohl JL, Koshland DE Jr, Stroud RM. 1990. *Science* 249: 1012–16
37. Johnson LN, O'Reilly M. 1996. *Curr. Opin. Struct. Biol.* 6:762–69

38. Koshland DE Jr. 1952. *J. Am. Chem. Soc.* 74:2286–92
39. Di Sabato G, Jencks WP. 1961. *J. Am. Chem. Soc.* 83:4400–5
40. Sanders DA, Gillece-Castro BL, Stock AM, Burlingame AL, Koshland DE Jr. 1989. *J. Biol. Chem.* 264:21770–78
41. Sanders DA, Gillece-Castro BL, Burlingame AL, Koshland DE Jr. 1992. *J. Bacteriol.* 174:5117–22
42. Stock AM, Wylie DC, Mottonen JM, Lupas AN, Ninfa EG, et al. 1988. *Cold Spring Harbor Symp. Quant. Biol.* 53: 49–57
43. Weiss V, Magasanik B. 1988. *Proc. Natl. Acad. Sci. USA* 85:8919–23
44. Igo MM, Ninfa AJ, Silhavy TJ. 1989. *Genes Dev.* 3:598–605
45. Makino K, Shinagawa H, Amemura M, Kawamoto T, Yamada M, Nakata A. 1989. *J. Mol. Biol.* 210:551–59
46. Wright GD, Holman TR, Walsh CT. 1993. *Biochemistry* 32:5057–63
47. Janiak-Spens F, Sparling JM, Gurfinkel M, West AH. 1999. *J. Bacteriol.* 181: 411–17
48. Swanson RV, Sanna MG, Simon MI. 1996. *J. Bacteriol.* 178:484–89
49. Goudreau PN, Lee P-J, Stock AM. 1998. *Biochemistry* 37:14575–84
49a. Lewis RJ, Brannigan JA, Muchová K, Barák I, Wilkinson AJ. 1999. *J. Mol. Biol.* 294:9–15
50. Wurgler-Murphy SM, Saito H. 1997. *Trends Biochem. Sci.* 22:172–76
51. Chang C, Stewart RC. 1998. *Plant Physiol.* 117:723–31
52. Loomis WF, Shaulsky G, Wang N. 1997. *J. Cell Sci.* 110:1141–45
53. Loomis WF, Kuspa A, Shaulsky G. 1998. *Curr. Opin. Microbiol.* 1:643–48
54. Zhang C-C. 1996. *Mol. Microbiol.* 20: 9–15
55. Chang C, Kwok SF, Bleecker AB, Meyerowitz EM. 1993. *Science* 262:539–44
56. Maeda T, Wurgler-Murphy SM, Saito H. 1994. *Nature* 369:242–45
57. Shieh JC, Wilkinson MG, Buck V, Morgan BA, Makino K, Millar JB. 1997. *Genes Dev.* 11:1008–22
58. Shaulsky G, Escalante R, Loomis WF. 1996. *Proc. Natl. Acad. Sci. USA* 93: 15260–65
59. Shaulsky G, Fuller D, Loomis WF. 1998. *Development* 125:691–99
60. Thomason PA, Traynor D, Cavet G, Chang W-T, Harwood AJ, Kay RR. 1998. *EMBO J.* 17:2838–45
61. Mizuno T, Kaneko T, Tabata S. 1996. *DNA Res.* 3:407–14
62. Zhang C-C, Gonzalez L, Phalip V. 1998. *Nucleic Acids Res.* 26:3619–25
63. Mizuno T. 1997. *DNA Res.* 4:161–68
64. Mizuno T. 1998. *J. Biochem.* 123:555–63
65. Fabret C, Feher VA, Hoch JA. 1999. *J. Bacteriol.* 181:1975–83
65a. Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, et al. 1999. *Nature* 399:323–29
66. Smith DR, Doucette-Stamm LA, Deloughery C, Lee HM, Dubois J, et al. 1997. *J. Bacteriol.* 179:7135–55
67. Ota IM, Varshavsky A. 1993. *Science* 262: 566–69
68. Posas F, Wurgler-Murphy SM, Maeda T, Witten EA, Thai TC, Saito H. 1996. *Cell* 86:865–75
69. Ketela T, Brown JL, Stewart RC, Bussey H. 1998. *Mol. Gen. Genet.* 259:372–78
70. Nagahashi S, Mio T, Ono N, Yamada-Okabe T, Arisawa M, et al. 1998. *Microbiology* 144:425–32
71. Alex LA, Korch C, Selitrennikoff CP, Simon MI. 1998. *Proc. Natl. Acad. Sci. USA* 95:7069–73
72. Alex LA, Borkovich KA, Simon MI. 1996. *Proc. Natl. Acad. Sci. USA* 93:3416–21
73. Wang N, Shaulsky G, Escalante R, Loomis WF. 1996. *EMBO J.* 15:3890–98

74. Schuster SS, Noegel AA, Oehme F, Gerisch G, Simon MI. 1996. *EMBO J.* 15:3880–89
75. Zinda MJ, Singleton CK. 1998. *Dev. Biol.* 196:171–83
76. Singleton CK, Zinda MJ, Mykytka B, Yang P. 1998. *Dev. Biol.* 203:345–57
77. Thomason P, Traynor D, Kay R. 1999. *Trends Genet.* 15:15–19
78. Hua J, Chang C, Sun Q, Meyerowitz EM. 1995. *Science* 269:1712–14
79. Wilkinson JQ, Lanahan MB, Yen H-C, Giovannoni JJ, Klee HJ. 1995. *Science* 270:1807–9
80. Yen H-C, Lee S, Tanksley SD, Lanahan MB, Klee HJ, Giovannoni JJ. 1995. *Plant Physiol.* 107:1343–53
81. The *C. elegans* Sequencing Consortium. 1998. *Science* 282:2012–18
82. Pao GM, Saier MH Jr. 1997. *J. Mol. Evol.* 44:605–13
83. Brown JL, North S, Bussey H. 1993. *J. Bacteriol.* 175:6908–15
84. Pan SQ, Charles T, Jin S, Wu Z-L, Nester EW. 1993. *Proc. Natl. Acad. Sci. USA* 90:9939–43
85. Yang Y, Park H, Inouye M. 1993. *J. Mol. Biol.* 232:493–98
86. Wolfe AJ, Stewart RC. 1993. *Proc. Natl. Acad. Sci. USA* 90:1518–22
87. Ninfa EG, Atkinson MR, Kamberov ES, Ninfa AJ. 1993. *J. Bacteriol.* 175:7024–32
88. Swanson RV, Bourret RB, Simon MI. 1993. *Mol. Microbiol.* 8:435–41
89. Keener J, Kustu S. 1988. *Proc. Natl. Acad. Sci. USA* 85:4976–80
90. Aiba H, Mizuno T, Mizushima S. 1989. *J. Biol. Chem.* 264:8563–67
91. Lois AF, Weinstein M, Ditta GS, Helinski DR. 1993. *J. Biol. Chem.* 268:4370–75
92. Alex LA, Simon MI. 1994. *Trends Genet.* 10:133–38
93. Lois AF, Ditta GS, Helinski DR. 1993. *J. Bacteriol.* 175:1103–9
94. Island MD, Wei B-Y, Kadner RJ. 1992. *J. Bacteriol.* 174:2754–62
95. Stock A, Chen T, Welsh D, Stock J. 1988. *Proc. Natl. Acad. Sci. USA* 85:1403–7
96. MacFarlane SA, Merrick M. 1985. *Nucleic Acids Res.* 13:7591–606
97. Ishige K, Nagasawa S, Tokishita S, Mizuno T. 1994. *EMBO J.* 13:5195–202
98. Lau PCK, Wang Y, Patel A, Labbe D, Bergeron H, et al. 1997. *Proc. Natl. Acad. Sci. USA* 94:1453–58
99. Stock J. 1999. *Curr. Biol.* 9:R364–67
100. Park H, Saha SK, Inouye M. 1998. *Proc. Natl. Acad. Sci. USA* 95:6728–32
101. Bilwes AM, Alex LA, Crane BR, Simon MI. 1999. *Cell* 96:131–41
102. Tanaka T, Saha SK, Tomomori C, Ishima R, Liu D, et al. 1998. *Nature* 396:88–92
103. Tomomori C, Tanaka T, Dutta R, Park H, Saha SK, et al. 1999. *Nat. Struct. Biol.* 6:729–34
104. Robinson VL, Stock AM. 1999. *Structure Fold. Des.* 7:R47–53
105. Ban C, Junop M, Yang W. 1999. *Cell* 97:85–97
106. Yeh K-C, Wu S-H, Murphy JT, Lagarias JC. 1997. *Science* 277:1505–8
107. Elich TD, Chory J. 1997. *Cell* 91:713–16
108. Yeh KC, Lagarias JC. 1998. *Proc. Natl. Acad. Sci. USA* 95:13976–81
109. Fankhauser C, Chory J. 1999. *Curr. Biol.* 9:R123–26
110. Wu J, Ohta N, Zhao JL, Newton A. 1999. *Proc. Natl. Acad. Sci. USA.* 96:13068–73
110a. Carson M, 1991. *J. Appl. Crystallogr.* 24:958–61
111. Taylor SS, Knighton DR, Zheng J, Ten Eyck LF, Sowadski JM. 1992. *Annu. Rev. Cell Biol.* 8:429–62
112. Johnson LN, Lowe ED, Noble ME, Owen DJ. 1998. *FEBS Lett.* 430:1–11
113. Kreegipuu A, Blom N, Brunak S, Jarv J. 1998. *FEBS Lett.* 430:45–50
114. Tsuzuki M, Ishige K, Mizuno T. 1995. *Mol. Microbiol.* 18:953–62

115. Zhou H, Lowry DF, Swanson RV, Simon MI, Dahlquist FW. 1995. *Biochemistry* 34:13858–70
116. Kato M, Mizuno T, Shimizu T, Hakoshima T. 1997. *Cell* 88:717–23
117. Varughese KI, Madhusudan, Zhou XZ, Whiteley JM, Hoch JA. 1998. *Mol. Cell* 2:485–93
118. Xu Q, West AH. 1999. *J. Mol. Biol.* 292:1039–50
118a. Song HK, Lee JY, Lee MG, Moon J, Min K, et al. 1999. *J. Mol. Biol.* 293:753–61
119. Bingman C, Regelman A, Reyngold M, Stokes L, Hendrickson W, Waldburger C. 1999. *Structural analysis of the PhoQ extracellular sensor domain.* Presented at the BLAST V: Bact. Locomot. Signal Transduct., Cuernavaca, Mex.
120. Taylor BL, Zhulin IB. 1999. *Microbiol. Mol. Biol. Rev.* 63:479–506
121. Lukat-Rodgers GS, Rodgers KR. 1997. *Biochemistry* 36:4178–87
122. Miyatake H, Mukai M, Adachi S, Nakamura H, Tamura K, et al. 1999. *J. Biol. Chem.* 274:23176–84
123. Gong W, Hao B, Mansy SS, Gonzalez G, Gilles-Gonzalez MA, Chan MK. 1998. *Proc. Natl. Acad. Sci. USA* 95:15177–82
124. Wootton JC, Drummond MH. 1989. *Protein Eng.* 2:535–43
125. Atkinson MR, Ninfa AJ. 1992. *J. Bacteriol.* 174:4538–48
126. Collins LA, Egan SM, Stewart V. 1992. *J. Bacteriol.* 174:3667–75
127. Park H, Inouye M. 1997. *J. Bacteriol.* 179:4382–90
128. Fassler JS, Gray WM, Malone CL, Tao W, Lin H, Deschenes RJ. 1997. *J. Biol. Chem.* 272:13365–71
129. Jourlin C, Bengrine A, Chippaux M, Mejean V. 1996. *Mol. Microbiol.* 20: 1297–306
130. Singh M, Berger B, Kim PS, Berger JM, Cochran AG. 1998. *Proc. Natl. Acad. Sci. USA* 95:2738–43
131. Aravind L, Ponting CP. 1999. *FEMS Microbiol. Lett.* 176:111–16
132. Lukat GS, McCleary WR, Stock AM, Stock JB. 1992. *Proc. Natl. Acad. Sci. USA* 89:718–22
133. Albright LM, Huala E, Ausubel FM. 1989. *Annu. Rev. Genet.* 23:311–36
134. Simms SA, Keane MG, Stock J. 1985. *J. Biol. Chem.* 260:10161–68
135. Welch M, Oosawa K, Aizawa S-I, Eisenbach M. 1993. *Proc. Natl. Acad. Sci. USA* 90:8787–91
136. Burbulys D, Trach KA, Hoch JA. 1991. *Cell* 64:545–52
137. Appleby JL, Parkinson JS, Bourret RB. 1996. *Cell* 86:845–48
138. Volz K. 1993. *Biochemistry* 32:11741–53
139. Stock AM, Mottonen JM, Stock JB, Schutt CE. 1989. *Nature* 337:745–49
140. Volz K, Matsumura P. 1991. *J. Biol. Chem.* 266:15511–19
141. Lukat GS, Stock AM, Stock JB. 1990. *Biochemistry* 29:5436–42
142. Lukat GS, Lee BH, Mottonen JM, Stock AM, Stock JB. 1991. *J. Biol. Chem.* 266:8348–54
143. Needham JV, Chen TY, Falke JJ. 1993. *Biochemistry* 32:3363–67
144. Stock AM, Martínez-Hackert E, Rasmussen BF, West AH, Stock JB, et al. 1993. *Biochemistry* 32:13375–80
145. Bellsolell L, Prieto J, Serrano L, Coll M. 1994. *J. Mol. Biol.* 238:489–95
146. Herschlag D, Jencks WP. 1990. *J. Am. Chem. Soc.* 112:1942–50
147. Zhu X, Rebello J, Matsumura P, Volz K. 1997. *J. Biol. Chem.* 272:5000–6
148. Appleby JL, Bourret RB. 1998. *J. Bacteriol.* 180:3563–69
149. Moy FJ, Lowry DF, Matsumura P, Dahlquist FW, Krywko JE, Domaille PJ. 1994. *Biochemistry* 33:10731–42
150. Santoro J, Bruix M, Pascual J, Lopez E, Serrano L, Rico M. 1995. *J. Mol. Biol.* 247:717–25
151. Bellsolell L, Cronet P, Majolero M, Serrano L, Coll M. 1996. *J. Mol. Biol.* 257:116–28

152. Jiang M, Bourret RB, Simon MI, Volz K. 1997. *J. Biol. Chem.* 272:11850–55
153. Usher KC, de la Cruz AF, Dahlquist FW, Swanson RV, Simon MI, Remington SJ. 1998. *Protein Sci.* 7:403–12
154. Madhusudan, Zapf J, Whiteley JM, Hoch JA, Xuong NH, Varughese KI. 1996. *Structure* 4:679–90
155. Feher VA, Zapf JW, Hoch JA, Whiteley JM, McIntosh LP, et al. 1997. *Biochemistry* 36:10015–25
156. Volkman BF, Nohaile MJ, Amy NK, Kustu S, Wemmer DE. 1995. *Biochemistry* 34:1413–24
157. Sola M, Gomis-Ruth FX, Serrano L, Gonzalez A, Coll M. 1999. *J. Mol. Biol.* 285:675–87
158. Baikalov I, Schröder I, Kaczor-Grzeskowiak M, Grzeskowiak K, Gunsalus RP, Dickerson RE. 1996. *Biochemistry* 35:11053–61
159. Djordjevic S, Goudreau PN, Xu QP, Stock AM, West AH. 1998. *Proc. Natl. Acad. Sci. USA* 95:1381–86
159a. Birck C, Mourey L, Gouet P, Fabry B, Schumacher J, et al. 1999. *Structure Fold. Des.* 7:1505–15
160. Aiba H, Nakasai F, Mizushima S, Mizuno T. 1989. *J. Biochem.* 106:5–7
161. Maeda S, Mizuno T. 1990. *J. Bacteriol.* 172:501–3
162. Rampersaud A, Harlocker SL, Inouye M. 1994. *J. Biol. Chem.* 269:12559–66
163. Pratt LA, Silhavy TJ. 1995. *Mol. Microbiol.* 17:565–73
164. Harlocker SL, Bergstrom L, Inouye M. 1995. *J. Biol. Chem.* 270:26849–56
165. Head CG, Tardy A, Kenney LJ. 1998. *J. Mol. Biol.* 281:857–70
166. Harrison-McMonagle P, Denissova N, Martínez-Hackert E, Ebright RH, Stock AM. 1999. *J. Mol. Biol.* 285:555–66
167. Kondo H, Nakagawa A, Nishihira J, Nishimura Y, Mizuno T, Tanaka I. 1997. *Nat. Struct. Biol.* 4:28–31
168. Martínez-Hackert E, Stock AM. 1997. *Structure* 5:109–24
169. Mizuno T, Tanaka I. 1997. *Mol. Microbiol.* 24:665–67
170. Martínez-Hackert E, Stock AM. 1997. *J. Mol. Biol.* 269:301–12
171. Garrett S, Silhavy TJ. 1987. *J. Bacteriol.* 169:1379–85
172. Matsuyama S, Mizushima S. 1987. *J. Mol. Biol.* 195:847–53
173. Makino K, Amemura M, Kim S-K, Nakata A, Shinagawa H. 1993. *Genes Dev.* 7:149–60
174. Kumar A, Grimes B, Fujita N, Makino K, Malloch RA, et al. 1994. *J. Mol. Biol.* 235:405–13
175. Iuchi S, Lin EC. 1987. *Proc. Natl. Acad. Sci. USA* 84:3901–5
176. Stewart V, Rabin RS. 1995. See Ref. 7, pp. 233–52
177. Tyson KL, Bell AI, Cole JA, Busby JW. 1993. *Mol. Microbiol.* 7:151–57
178. Stewart V. 1993. *Mol. Microbiol.* 9: 425–34
179. Kustu S, Santero E, Keener J, Popham D, Weiss D. 1989. *Microbiol. Rev.* 53: 367–76
180. Porter SC, North AK, Kustu S. 1995. See Ref. 7, pp. 147–58
181. Kustu S, North AK, Weiss DS. 1991. *Trends Biochem. Sci.* 16:397–402
182. Morett E, Segovia L. 1993. *J. Bacteriol.* 175:6067–74
183. Osuna J, Soberon X, Morett E. 1997. *Protein Sci.* 6:543–55
184. Weiss V, Claverie-Martin F, Magasanik B. 1992. *Proc. Natl. Acad. Sci. USA* 89: 5088–92
185. Wyman C, Rombel I, North AK, Bustamente C, Kustu S. 1997. *Science* 275: 1658–61
186. Weiss DS, Batut J, Klose KE, Keener J, Kustu S. 1991. *Cell* 67:155–67
187. Austin S, Dixon R. 1992. *EMBO J.* 11:2219–28
188. Wedel A, Kustu S. 1995. *Genes Dev.* 9:2042–52
189. Kahn D, Ditta G. 1991. *Mol. Microbiol.* 5:987–97

190. Huala E, Stigter J, Ausubel FM. 1992. *J. Bacteriol.* 174:1428–31
191. Grimsley JK, Tjalkens RB, Strauch MA, Bird TH, Spiegelman GB, et al. 1994. *J. Biol. Chem.* 269:16977–82
192. Fiedler U, Weiss V. 1995. *EMBO J.* 14: 3696–705
193. McCleary WR. 1996. *Mol. Microbiol.* 20:1155–63
194. Webber CA, Kadner RJ. 1997. *Mol. Microbiol.* 24:1039–48
195. Blat Y, Eisenbach M. 1994. *Biochemistry* 33:902–6
196. Anand GS, Goudreau PN, Stock AM. 1998. *Biochemistry* 37:14038–47
197. Posas F, Saito H. 1998. *EMBO J.* 17: 1385–94
198. Kenney LJ, Bauer MD, Silhavy TJ. 1995. *Proc. Natl. Acad. Sci. USA* 92:8866–70
199. Lowry DF, Roth AF, Rupert PB, Dahlquist FW, Moy FJ, et al. 1994. *J. Biol. Chem.* 269:26358–62
200. Nohaile M, Kern D, Wemmer D, Stedman K, Kustu S. 1997. *J. Mol. Biol.* 273: 299–316
201. Hwang I, Thorgeirsson T, Lee J, Kustu S, Shin YK. 1999. *Proc. Natl. Acad. Sci. USA* 96:4880–85
202. Bourret RB, Hess JF, Simon MI. 1990. *Proc. Natl. Acad. Sci. USA* 87:41–45
203. Kanamaru K, Mizuno T. 1992. *J. Biochem.* 111:425–30
204. Stewart RC. 1993. *J. Biol. Chem.* 268: 1921–30
205. Klose KE, Weiss DS, Kustu S. 1993. *J. Mol. Biol.* 232:67–78
206. Kern D, Volkman BF, Luginbühl P, Nohaile MJ, Kustu S, Wemmer DE. 1999. *Nature* 402:894–98
207. Halkides CJ, Zhu XY, Phillion DP, Matsumura P, Dahlquist FW. 1998. *Biochemistry* 37:13674–80
208. Deleted in proof
209. Sockett H, Yamaguchi S, Kihara M, Irikura VM, Macnab RM. 1992. *J. Bacteriol.* 174:793–806
210. Roman SJ, Meyers M, Volz K, Matsumura P. 1992. *J. Bacteriol.* 174:6247–55
211. Swanson RV, Lowry DF, Matsumura P, McEvoy MM, Simon MI, Dahlquist FW. 1995. *Nat. Struct. Biol.* 2:906–10
212. Sanna MG, Swanson RV, Bourret RB, Simon MI. 1995. *Mol. Microbiol.* 15: 1069–79
213. Zhu XY, Volz K, Matsumura P. 1997. *J. Biol. Chem.* 272:23758–64
214. McEvoy MM, Hausrath AC, Randolph GB, Remington SJ, Dahlquist FW. 1998. *Proc. Natl. Acad. Sci. USA* 95:7333–38
215. Welch M, Chinardet N, Mourey L, Birck C, Samama J-P. 1998. *Nat. Struct. Biol.* 5:25–29
216. McEvoy MM, Bren A, Eisenbach M, Dahlquist FW. 1999. *J. Mol. Biol.* 289: 1423–33
217. Tzeng Y-L, Hoch JA. 1997. *J. Mol. Biol.* 272:200–12
218. Haldimann A, Prahalad MK, Fisher SL, Kim S-K, Walsh CT, Wanner BL. 1996. *Proc. Natl. Acad. Sci. USA* 93:14361–66
219. Feher VA, Cavanagh J. 1999. *Nature* 400:289–93
220. Sprang SR. 1997. *Annu. Rev. Biochem.* 66:639–78
221. Artymiuk PJ, Rice DW, Mitchell EM, Willett P. 1990. *Protein Eng.* 4:39–43
222. Chen JM, Lee G, Murphy RB, Brandt-Rauf PW, Pincus MR. 1990. *Int. J. Pept. Protein Res.* 36:1–6
223. Stock JB, Lukat GS, Stock AM. 1991. *Annu. Rev. Biophys. Biophys. Chem.* 20: 109–36
224. Pratt LA, Silhavy TJ. 1995. See Ref. 7, pp. 105–27
225. Li J, Swanson RV, Simon MI, Weis RM. 1995. *Biochemistry* 34:14626–36
226. Rabin RS, Stewart V. 1993. *J. Bacteriol.* 175:3259–68
227. Parkinson JS. 1993. *Cell* 73:857–71
228. Perraud A-L, Weiss V, Gross R. 1999. *Trends Microbiol.* 7:115–20
229. Uhl MA, Miller JF. 1995. See Ref. 7, pp. 333–49

230. Msadek T. 1999. *Trends Microbiol.* 7: 201–7
231. Sun G, Birkey SM, Hulett FM. 1996. *Mol. Microbiol.* 19:941–48
232. Birkey SM, Liu W, Zhang X, Duggan MF, Hulett FM. 1998. *Mol. Microbiol.* 30:943–53
233. Hulett FM. 1996. *Mol. Microbiol.* 19: 933–39
234. Li S, Ault A, Malone CL, Raitt D, Dean S, et al. 1998. *EMBO J.* 17:6952–62
235. Fisher SL, Jiang W, Wanner BL, Walsh CT. 1995. *J. Biol. Chem.* 270:23143–49
236. Wanner BL. 1994. In *Phosphate in Microorganisms: Cellular and Molecular Biology*, ed. A Torriani-Gorini, E Yagil, S Silver, pp. 13–21. Washington, DC: Am. Soc. Microbiol.
237. Wanner BL. 1992. *J. Bacteriol.* 174: 2053–58
238. Feng J, Atkinson MR, McCleary W, Stock JB, Wanner BL, Ninfa AJ. 1992. *J. Bacteriol.* 174:6061–70
239. McCleary WR, Stock JB, Ninfa AJ. 1993. *J. Bacteriol.* 175:2793–98
240. Hoch JA. 1995. See Ref. 7, pp. 129–44
241. Borkovich KA, Kaplan N, Hess JF, Simon MI. 1989. *Proc. Natl. Acad. Sci. USA* 86:1208–12
242. Dahl MK, Msadek T, Kunst F, Rapoport G. 1992. *J. Biol. Chem.* 267:14509–14
243. Jung K, Tjaden B, Altendorf K. 1997. *J. Biol. Chem.* 272:10847–52
244. Gegner JA, Graham DR, Roth AF, Dahlquist FW. 1992. *Cell* 70:975–82
245. Liu Y, Levit M, Lurz R, Surette MG, Stock JB. 1997. *EMBO J.* 16:7231–40
246. Ninfa EG, Stock A, Mowbray S, Stock J. 1991. *J. Biol. Chem.* 266:9764–70
247. Borkovich KA, Alex LA, Simon MI. 1992. *Proc. Natl. Acad. Sci. USA* 89: 6756–60
248. Kamberov ES, Atkinson MR, Feng J, Chandran P, Ninfa AJ. 1994. *Cell. Mol. Biol. Res.* 40:175–91
249. Steed PM, Wanner BL. 1993. *J. Bacteriol.* 175:6797–809
250. Wanner BL. 1995. See Ref. 7, pp. 203–21
251. Jung K, Altendorf K. 1998. *J. Biol. Chem.* 273:17406–10
252. Kamberov ES, Atkinson MR, Chandran P, Ninfa AJ. 1994. *J. Biol. Chem.* 269:28294–99
253. Hsing W, Silhavy TJ. 1997. *J. Bacteriol.* 179:3729–35
254. Skarphol K, Waukau J, Forst SA. 1997. *J. Bacteriol.* 179:1413–16
255. Perego M, Hanstein C, Welsh KM, Djavakhishvili T, Glaser P, Hoch JA. 1994. *Cell* 79:1047–55
256. Ohlsen KL, Grimsley JK, Hoch JA. 1994. *Proc. Natl. Acad. Sci. USA* 91:1756–60
257. Hess JF, Bourret RB, Simon MI. 1988. *Nature* 336:139–43
258. Blat Y, Eisenbach M. 1996. *J. Biol. Chem.* 271:1226–31
259. Chang C-H, Zhu J, Winans SC. 1996. *J. Bacteriol.* 178:4710–16
260. Sourjik V, Schmitt R. 1996. *Mol. Microbiol.* 22:427–36
261. Ueno-Nishio S, Mango S, Reitzer LJ, Magasanik B. 1984. *J. Bacteriol.* 160: 379–84
262. Wanner BL, Chang BD. 1987. *J. Bacteriol.* 169:5569–74
263. Birkey SM, Sun G, Piggot PJ, Hulett FM. 1994. *Gene* 147:95–100
264. Soncini FC, Vescovi EG, Groisman EA. 1995. *J. Bacteriol.* 177:4364–71
265. Evers S, Courvalin P. 1996. *J. Bacteriol.* 178:1302–9
266. Novak R, Henriques B, Charpentier E, Normark S, Tuomanen E. 1999. *Nature* 399:590–93
267. Rasmussen BA, Kovacs E. 1993. *Mol. Microbiol.* 7:765–76
268. Strauch MA, deMendoza D, Hoch JA. 1992. *Mol. Microbiol.* 6:2909–17
269. Roychoudhury S, Zielinski NA, Ninfa AJ, Allen NE, Jungheim LN, et al. 1993. *Proc. Natl. Acad. Sci. USA* 90:965–69
270. Barrett JF, Hoch JA. 1998. *Antimicrob. Agents Chemother.* 42:1529–36

271. Domagala JM, Alessi D, Cummings M, Gracheck S, Huang L, et al. 1998. *Adv. Exp. Med. Biol.* 456:269–86
272. Barrett JF, Goldschmidt RM, Lawrence LE, Foleno B, Chen R, et al. 1998. *Proc. Natl. Acad. Sci. USA* 95:5317–22
273. Macielag MJ, Demers JP, Fraga-Spano SA, Hlasta DJ, Johnson SG, et al. 1998. *J. Med. Chem.* 41:2939–45
274. Deschenes RJ, Lin H, Ault AD, Fassler JS. 1999. *Antimicrob. Agents Chemother.* 43:1700–3
275. Hilliard JJ, Goldschmidt RM, Licata L, Baum EZ, Bush K. 1999. *Antimicrob. Agents Chemother.* 43:1693–99
276. Hunger-Glaser I, Brun R, Linder M, Seebeck T. 1999. *Mol. Biochem. Parasitol.* 100:53–59

Annual Review of Biochemistry
Volume 69, 2000

CONTENTS

Two-component and phosphorelay signal transduction

James A Hoch

Two-component and phosphorelay signal transduction systems are the major means by which bacteria recognize and respond to a variety of environmental stimuli. Recent results have implicated these systems in the regulation of a variety of essential processes including cell-cycle progression, pathogenicity, and developmental pathways. Elucidation of the structures of the interacting domains is leading to an understanding of the mechanisms of molecular recognition and phosphotransfer in these systems.

Addresses
Division of Cellular Biology, Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA 92037, USA; e-mail: hoch@scripps.edu

Current Opinion in Microbiology 2000, **3**:165–170

1369-5274/00/$ – see front matter

Abbreviation
PBP periplasmic-binding protein

Introduction

Bacteria are bombarded by all types of molecules that originate from their environment or are produced by their metabolism to which they must mount an appropriate response. In the majority of cases the response is transcriptional activation of genes whose products specifically cope with a given molecular signal. The essence of signal transduction is the conversion of signal recognition to gene activation or other cellular responses. The major means by which this is accomplished in bacteria is through two-component and phosphorelay signal transduction systems. Two-component systems are characterized by a sensor kinase consisting of a signal recognition domain (also known as 'input domain') with unique specificity coupled to an autokinase domain (Figure 1a). Signal binding causes activation of the autokinase resulting in ATP hydrolysis and phosphorylation of a histidine on a phosphotransferase sub-domain of the autokinase. The phosphotransferase sub-domain is monogamously mated with a regulator domain of a response regulator to which its phosphoryl group is transferred resulting in an aspartyl-phosphate. Regulator domains normally inhibit the output domain of the response regulator and phosphorylation relieves this inhibition freeing the output domain to carry out its function. That function is usually transcription activation, although enzymes or domains of other types may be controlled in this manner.

Phosphorelays differ from two-component systems by having a more convoluted signal transduction pathway using additional regulator and phosphotransferase domains. In the sporulation phosphorelay of *Bacillus subtilis*, depicted in Figure 1b, these domains are on individual proteins but in many systems the domains are associated with the kinase in a polydomain protein. Regardless, the signal transduction pathway remains as shown in Figure 1b and the phosphoryl group is transferred in the order His→Asp→His→Asp. The additional components give more targets for regulation of the pathway, especially by phosphatases that regulate phosphate flow by dephosphorylating the regulator domains. Phosphorelays are used for complex pathways such as sporulation or cell cycle control in bacteria and are used in eukaryotic signaling.

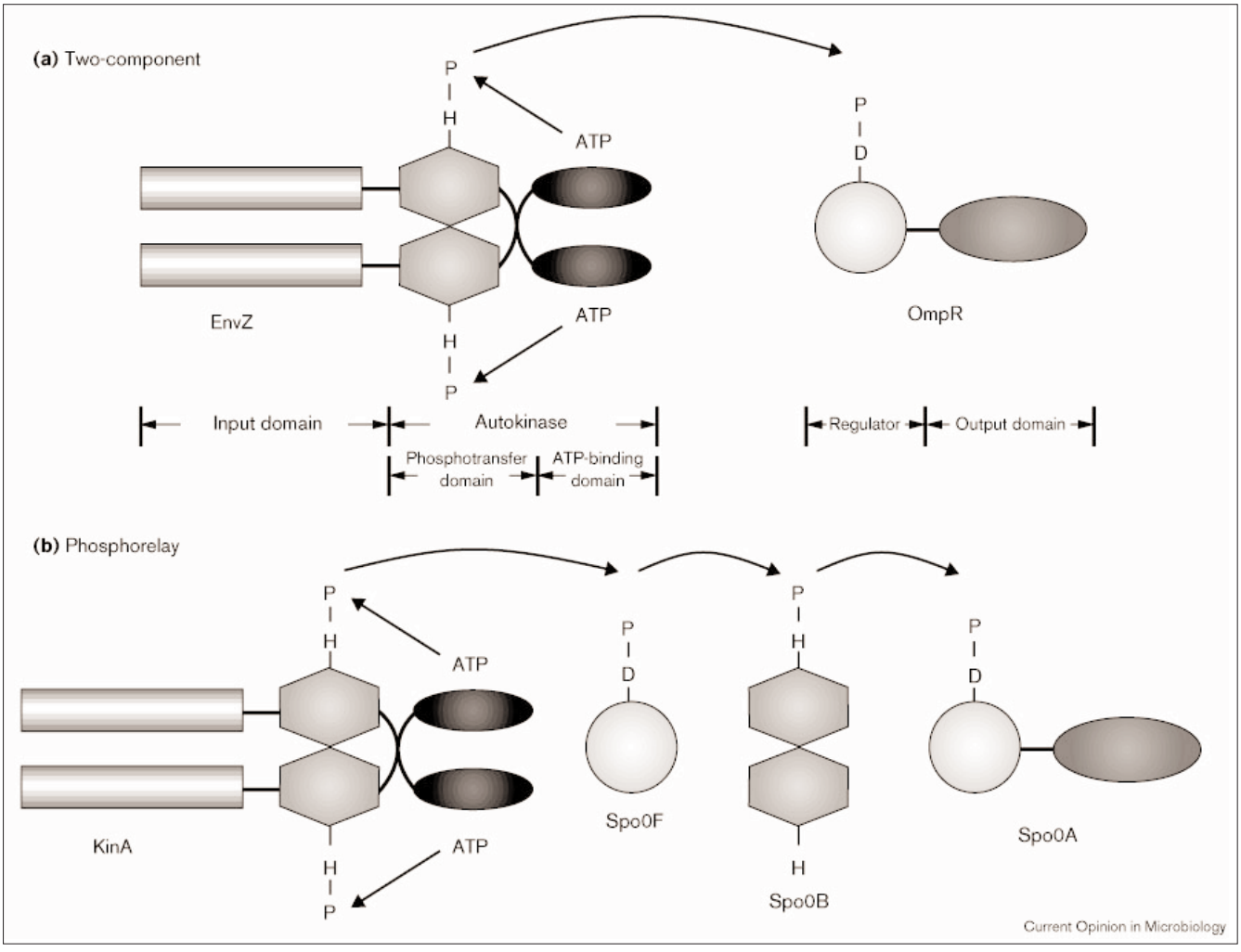
This short review of the year's activity in this area focuses on the structure and function of the systems while attempting to be comprehensive in its scope. Different viewpoints reflecting the interests of the authors may be obtained from the plethora of reviews on these signal transduction systems that appeared in the past year [1–5].

Domain structure

Kinases

Knowing the structure of the various domains comprising these systems is a prerequisite to understanding the mechanisms by which these systems work. Tremendous progress towards this goal has been made recently. Tanaka *et al.* [6•] solved the solution structure of the ATP-binding sub-domain of the autokinase of the *Escherichia coli* EnvZ histidine kinase (Figure 1). On the basis of amino acid homology, all kinases of two-component systems or phosphorelays appear to have identical structures for this sub-domain. A crystal structure of the Spo0B protein of the *B. subtilis* sporulation phosphorelay revealed that the histidine-containing phosphotransfer domain is a new type of four-helix bundle generated by the dimerization of two helices from each of the two subunits [7•]. A very similar structure was found for the phosphotransfer sub-domain of the EnvZ autokinase by NMR analyses [8]. Although the connectivity between the helices differs between the two structures, the four-helix bundles superpose (KI Varughese, JA Hoch, unpublished data). In the CheA kinase an identical four-helix structure was found to be used as a dimerization domain but not a phosphotransferase domain [9]. Crystal structure of the *Saccharomyces cerevisiae* Ypd1 phosphotransferase [10,11] showed identity to the previously described Hpt domain consisting of a four-helix bundle generated from a single polypeptide chain carrying the active site histidine and phosphotransferase functions but not functional as a dimerization domain. Thus, phosphotransferase domains appear to consist of two types: the Spo0B/EnvZ type serving as a dimerization domain, which is likely to be the type found in most kinases, and the Hpt type found in the CheA kinase and in phosphotransferase domains that serve the role of Spo0B in phosphorelays.

Figure 1



Comparison of the domain structure and signaling pathways of two-component systems and phosphorelays. **(a)** The EnvZ/OmpR system of *E. coli* is typical of the majority of two-component systems. Signal binding causes activation of the autokinase resulting in ATP hydrolysis and phosphorylation (P) of a histidine (H) on a phosphotransferase sub-domain of the autokinase. The phosphoryl group is transferred to the aspartyl group (D) on the response regulator. **(b)** The sporulation phosphorelay of *B. subtilis* has individual proteins for some domains that, in other systems, are fused into poly domain proteins.

Response regulators

Since the publication of the pioneering structure of CheY more than a dozen years ago [12], several regulator domains have been analyzed and all have virtually the same α/β structure as CheY. What has eluded detection is the structure of a phosphorylated regulator and an explanation of the means by which phosphorylation of the regulator disinhibits the output domain of response regulators. A crystal structure of the phosphorylated Spo0A regulator domain from *Bacillus stearothermophilus* has allowed dramatic advances toward the solution of these problems [13••]. The phosphoryl group was found to cause a large reorientation of the side chain of the threonine residue conserved in the active site of all regulators. This leads to alterations in the conformation of residues that pack onto the site-distal face of the threonine residue. NMR data predicted that this threonine controlled the conformational exchange of His101 of Spo0F, Tyr106 of CheY and Phe104 of CheB between buried and exposed positions [14•]. This change could lead to destabilization of the interdomain interactions that allow the unphosphorylated regulator to inhibit its output domain [14•,15]. However, at this point the explanation is informed speculation since a crystal structure for a complete phosphorylated response regulator has not been achieved.

Evolution of two-component systems

In bacteria with several dozen two-component systems, such as *E. coli* or *B. subtilis*, it is clear that the majority of the systems arose by gene duplication from one or more ancestral systems and evolved to acquire new input signal and output promoter specificities. In *B. subtilis*, for example,

two-component families derived from a single progenitor pair can be identified by the conservation of the sequences around the active site histidine of the kinase, the active site residues of the regulator, the structure of the output domain, and the relative chromosomal order of the genes for the kinase and response regulator [16]. Thus, within a family such as the OmpR-related family with 15 members, the phosphotransfer, ATP-domain, regulator, and output domains maintained the same structure and evolved by amino acid substitution to acquire new recognition specificities. A glaring exception to this peaceful evolution is the signal input domain. This domain is very heterogeneous in size and structure even within families. Although many of these domains are located in the membrane, the number of membrane spanning regions varies from one to eight or more, and even where two closely related paralogous kinases can be identified, the amino acid sequence of the periplasmic portion shows little homology to its paralog. What mutational mechanism leads to this hyper variability while the rest of the kinase protein remains unchanged?

The YycG kinase of *B. subtilis* is the only kinase readily identifiable between species since both the *Streptococcus pyogenes* and *Staphylococcus aureus* orthologs are highly identical in amino acid sequence [17,18]. However, the sequences of the periplasmic portions of the input domains are essentially randomized. This lack of sequence conservation is reminiscent of the portions of proteins that have no structural or enzymatic function other than to just be there. Perhaps many of the periplasmic domains of kinases are ancestral rather than functional. Some, such as that of CitS, certainly function as ligand receptors [19].

In addition to this seemingly patternless evolution in input domains, some histidine kinases acquire sub-domains that play a role in signaling. The most prevalent of these is the PAS domain that binds heme or FAD and serves as an oxygen or redox sensor [20]. The PAS domain and its associated PAC domain are very common in histidine kinases and may serve as binding sites for ligands with other signaling functions [21]. Another domain common to histidine kinases is the HAMP [22], type P [23], or Duf5 (Sanger Center nomenclature) domain that is located just cytoplasmic to transmembrane domains. This domain is also found in ligand-binding chemotaxis receptors suggesting that the domain may have a role in transmembrane signaling.

Just when all the action in histidine kinases seemed to be in the signal input domain, a kinase was discovered with a tyrosine substitution for histidine [24•]. This kinase from *Caulobacter crescentus* transfers phosphate to the CtrA response regulator involved in cell cycle control. In general, however, the phosphotransfer and ATP-binding domains of histidine kinases are boringly similar, although they can be sub-divided into related classes [1].

Recent studies have implicated ancestral relationships between the regulator domain of response regulators and other domains with regulatory or receptor function. Fukami-Kobayashi *et al.* [25] proposed that the periplasmic-binding proteins (PBPs) such as MalE and LacR of *E. coli* and related bacteria arose from a CheY-like protein by a process they called 'domain dislocation' involving dimerization by swapping one helix (helix 5) between two molecules followed by dimer fusion. A genealogical chart of all the PBPs could be constructed in this manner. A regulator domain lacking the phosphorylation site was found to control the activity of a ligand-regulated transcription antitermination complex from *Pseudomonas aeruginosa* [26]. A ligand-binding subunit interacts with the regulator domain to affect the activity of the complex. Thus regulator domains may transmit signals by sequestration as well as by phosphorylation.

Location, location, location

While this familiar refrain is a well-known axiom of the retail business, Shapiro and colleagues have shown how important location is to the function of histidine kinases in cell cycle control in *Caulobacter crescentus*. In this organism a response regulator — transcription factor CtrA — regulates key cell cycle events and its concentration as well as phosphorylation state are involved in its effects [27,28]. The CckA kinase phosphorylates CtrA and shows cell-cycle dependent dynamic movement to and from the cell poles. This dynamic positioning has been proposed to provide a new mechanism for control of the prokaryotic cell cycle [29••]. Recently the DivJ and PleC histidine kinases that regulate polar differentiation in *C. crescentus* were found to be differentially localized as a function of the cell cycle suggesting this may contribute to their regulatory functions [30•]. A detailed explanation of this system can be found in the review by Østerås and Jenal in this volume (pp 171–176). The concept of dynamic relocation of kinases opens new vistas to our imagination for the roles of input domains in sensor kinase activity.

Response regulator functions in bacteria

It is now clear there is hardly a pathway or a phenomenon that some bacterium has not adapted a two-component system to regulate its function. The list seems endless; the Sanger Center has 612 entries under response regulator at the time of this review. Even quorum sensing, the new darling of some microbiologists, is processed through or regulated by two-component systems in *Vibrio* [31], *Pseudomonas* [32], and perhaps many other species. Virulence and pathogenicity are adaptive responses of opportunistic pathogens requiring the sequential expression of new genetic information for successful invasion and growth in susceptible hosts. It is in adaptation to different environments that two-component systems and phosphorelays are especially useful and many pathogens have adopted them to regulate this process. Space only permits a superficial listing of pathogenic species and virulence systems that have been identified or developed in the past year. Some new notable involvements of two-component signaling systems in this area include

virulence in group A *Streptococcus* [33], *S. pneumoniae* [34], *Vibrio cholerae* [35], *Listeria monocytogenes* [36], and *Campylobacter jejuni* [37], among others. Some known virulence systems such as the Cpx system of *E. coli* [38,39] and the *Salmonella* pathogenicity system [40,41] have reached new heights of complexity and understanding. While two-component regulated antibiotic resistance such as vancomycin resistance is well known, a new phenotype of vancomycin tolerance has been found in *S. pneumoniae* [42]. Two-component systems are also implicated in virulence of *Candida albicans* [43,44]. The possibility of designing inhibitors of two-component systems as anti-infectives was enhanced by the discovery of essential systems in *B. subtilis* [17] and pathogenic Gram-positive bacteria [34].

Phosphorelays in developing systems

When the phosphorelay signal transduction pathway was discovered it was suggested that the increased complexity over a two-component system may reflect the need to integrate both positive and negative signals into the output of the pathway [45]. Indeed phosphorelays are used to regulate seemingly complex cellular phenomena such as development and cell cycle control in bacteria. More phosphorelay-based developmental pathways are being uncovered including heterocyst formation in cyanobacteria [46], sporulation timing in *Myxococcus xanthus* [47], development and fruiting body formation in *Dictyostelium* [48], and hyphal formation and properties in fungi including the human pathogen *C. albicans* [49].

In most lower eukaryotes the phosphorelays are minor signaling pathways, and they appear to be absent in mammals, whereas in plants they are used for several functions. *Arabidopsis thaliana* has at least 14 response regulators [50], at least 10 histidine kinases, and a set of HPt phosphotransferases characteristic of phosphorelays [51]. In addition to the ethylene response pathway regulating ripening and other aging properties of plants [52•,53], phosphorelays are implicated in cytokinin regulation and response to nitrogen starvation [54–56]. It will be of interest to see how many other properties these plant phosphorelays control and whether they show developmental and tissue specific expression.

Conclusions

Two-component and phosphorelay signal transduction pathways are woven in the fabric of signal responses from bacteria through to higher plants. From genomic inspection and mutational analyses these systems are being found to regulate a wide variety of pathways in response to a myriad of signals. With so many superficial conclusions being drawn from protein homologies about how pathways function based on the biochemistry, structure, and function of a few well-characterized bacterial systems, it becomes imperative that these systems are understood in detail. What do we know about these systems? We know the structure of regulator and output domains of response regulators, the structure of ATP domains and phosphotransferase domains of histidine kinases. We have a limited appreciation of the mechanism of phosphotransfer between the two proteins. What we do not know is how regulator domains control output domains and how phosphorylation interferes with this interaction, although we have some suspicions about how this works. We have no idea how signals activate kinases. Nothing is known of the structure of membrane-based signal input domains for the majority of kinases. Is membrane location important for their function and do they specifically associate with other membrane proteins or membrane-associated structures? The signals that activate all these kinases are a complete mystery with very few exceptions. The simple things appear to be done with the tough stuff yet to do.

Acknowledgements

This research was supported, in part, by grant GM19416 from the National Institutes of General Medical Sciences, National Institutes of Health, USPHS. This is publication 12952-MEM from The Scripps Research Institute.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Grebe TW, Stock JB: **The histidine protein kinase superfamily.** *Adv Microbiol Physiol* 1999, **41**:139-227.

2. Stock J: **Signal transduction: gyrating protein kinases.** *Curr Biol* 1999, **9**:R364-R367.

3. Pirrung MC: **Histidine kinases and two-component signal transduction systems.** *Chem Biol* 1999, **6**:R167-R175.

4. Dutta R, Qin L, Inouye M: **Histidine kinases: diversity of domain organization.** *Mol Microbiol* 1999, **34**:633-640.

5. Robinson VL, Stock AM: **High energy exchange: proteins that make or break phosphoramidate bonds.** *Structure* 1999, **7**:R47-R53.

6. • Tanaka T, Saha SK, Tomomori C, Ishima R, Liu D, Tong KI, Park H, Dutta R, Qin L, Swindells MB *et al.*: **NMR structure of the histidine kinase domain of the *E. coli* osmosensor EnvZ.** *Nature* 1998, **396**:88-92.

The first look at the structure of the ATP-binding domain of histidine kinases.

7. • Varughese KI, Madhusudan, Zhou XZ, Whiteley JM, Hoch JA: **Formation of a novel four-helix bundle and molecular recognition sites by dimerization of a response regulator phosphotransferase.** *Mol Cell* 1998, **2**:485-493.

A first look at the four-helix bundle/dimerization characteristics of histidine phosphotransfer domains. See also [8,9].

8. Tomomori C, Tanaka T, Dutta R, Park H, Saha SK, Zhu Y, Ishima R, Liu D, Tong KI, Kurokawa H *et al.*: **Solution structure of the homodimeric core domain of *Escherichia coli* histidine kinase EnvZ.** *Nat Struct Biol* 1999, **6**:729-734.

9. Bilwes AM, Alex LA, Crane BR, Simon MI: **Structure of CheA, a signal-transducing histidine kinase.** *Cell* 1999, **96**:131-141.

10. Song HK, Lee JY, Lee MG, Min JMK, Yang JK, Suh SW: **Insights into eukaryotic multistep phosphorelay signal transduction revealed by the crystal structure of Ypd1p from *Saccharomyce cerevisiae.*** *J Mol Biol* 1999, **293**:753-761.

11. Xu Q, West AH: **Conservation of structure and function among histidine-containing phosphotransfer (HPt) domains as revealed by the crystal structure of YPD1.** *J Mol Biol* 1999, **292**:1039-1050.

12. Stock AM, Mottonen JM, Stock JB, Schutt CE: **Three-dimensional structure of CheY, the response regulator of bacterial chemotaxis.** *Nature* 1989, **337**:745-749.

13. Lewis RJ, Brannigan JA, Muchová K, Barák I, Wilkinson AJ:
•• **Phosphorylated aspartate in the structure of a response regulator protein.** *J Mol Biol* 1999, **294**:9-15.
The first look at how phosphorylation affects the metal coordination and disposition of side chains of active site residues of response regulators with implications for signaling in this family of proteins.

14. Feher VA, Cavanagh J: **Millisecond-timescale motions contribute to**
• **the function of the bacterial response regulator protein Spo0F.** *Nature* 1999, **400**:289-293.
An NMR study showing residue and surface dynamics correlate with regions of response regulators critical for protein–protein interaction.

15. Djordjevic S, Stock AM: **Structural analysis of bacterial chemotaxis proteins: components of a dynamic signaling system.** *J Struct Biol* 1998, **124**:189-200.

16. Fabret C, Feher VA, Hoch JA: **Two-component signal transduction in *Bacillus subtilis*: how one organisms sees its world.** *J Bacteriol* 1999, **181**:1975-1983.

17. Fabret C, Hoch JA: **A two-component signal transduction system essential for growth of *Bacillus subtilis*: implications for anti-infective therapy.** *J Bacteriol* 1998, **180**:6375-6383.

18. Martin PK, Li T, Sun D, Biek DP, Schmid MB: **Role in cell permeability of an essential two-component system in *Staphylococcus aureus.*** *J Bacteriol* 1999, **181**:3666-3673.

19. Kaspar S, Perozzo R, Reinelt S, Meyer M, Pfister K, Scapozza L, Bott M: **The periplasmic domain of the histidine autokinase CitA functions as a highly specific citrate receptor.** *Mol Microbiol* 1999, **33**:858-872.

20. Gong W, Hao B, Mansy SS, Gonzalez G, Gilles-Gonzalez MA, Chan MK: **Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction.** *Proc Natl Acad Sci USA* 1998, **95**:15177-15182.

21. Taylor BL, Zhulin IB: **PAS domains: internal sensors of oxygen, redox potential, and light.** *Microbiol Mol Biol Rev* 1999, **63**:479-506.

22. Aravind L, Ponting CP: **The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signaling proteins.** *FEMS Microbiol Lett* 1999, **176**:111-116.

23. Williams SB, Stewart V: **Functional similarities among two-component sensors and methyl-accepting chemotaxis proteins suggest a role for linker region amphipathic helices in transmembrane signal transduction.** *Mol Microbiol* 1999, **33**:1093-1102.

24. Wu J, Ohta N, Zhao JL, Newton A: **A novel bacterial tyrosine kinase**
• **essential for cell division and differentiation.** *Proc Natl Acad Sci USA*, 1999, **96**:13068-13073.
A tyrosine substitution for histidine works as well. Are there more like this out there?

25. Fukami-Kobayashi K, Tateno Y, Nishikawa K: **Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history.** *J Mol Biol* 1999, **286**:279-290.

26. O'Hara BP, Norman RA, Wan PTC, Roe SM, Barrett TE, Drew RE, Pearl LH: **Crystal structure and induction mechanism of AmiC-AmiR: a ligand-regulated transcription antitermination complex.** *EMBO J* 1999, **18**:5175-5186.

27. Reisenauer A, Quon K, Shapiro L: **The CtrA response regulator mediates temporal control of gene expression during the *Caulobacter* cell cycle.** *J Bacteriol* 1999, **181**:2430-2439.

28. Domain IJ, Reisenauer A, Shapiro L: **Feedback control of a master bacterial cell-cycle regulator.** *Proc Natl Acad Sci USA* 1999, **96**:6648-6653.

29. Jacobs C, Domian IJ, Maddock JR, Shapiro L: **Cell cycle-dependent**
•• **polar localization of an essential bacterial histidine kinase that controls DNA replication and cell division.** *Cell* 1999, **97**:111-120.
The discovery that a histidine kinase required for cell cycle control shows dynamic membrane localization behavior dependent on the phase of the cell cycle.

30. Wheeler RT, Shapiro L: **Differential localization of two histidine**
• **kinases controlling bacterial cell differentiation.** *Mol Cell* 1999, **4**:683-694.
Dynamic partitioning in the cell membrane is found with two other histidine kinases involved in polar differentiation in the cell cycle. Is this the tip of the iceberg for membrane domains of histidine kinases?

31. Freeman JA, Bassler BL: **A genetic analysis of the function of LuxO, a two-component response regulator involved in a quorum sensing in *Vibrio harveyi.*** *Mol Microbiol* 1999, **31**:665-677.

32. Chancey ST, Wood DW, Pierson LS III: **Two-component transcriptional regulation of N-acyl-homoserine lactone production in *Pseudomonas aureofaciens.*** *Appl Environ Microbiol* 1999, **65**:2294-2299.

33. Federle MJ, McIver KS, Scott JR: **A response regulator that represses transcription of several virulence operons in the Group A *Streptococcus.*** *J Bacteriol* 1999, **181**:3649-3657.

34. Lange R, Wagner C, deSaizieu A, Flint N, Molnos J, Stieger M, Caspers P, Kamber M, Keck W, Amrein KE: **Domain organization and molecular characterization of 13 two-component systems identified by genome sequencing of *Streptococcus pneumoniae.*** *Gene* 1999, **237**:223-234.

35. Wong SM, Carroll PA, Rahme LG, Ausubel FM, Calderwood SB: **Modulation of expression of the ToxR regulon in *Vibrio cholerae* by a member of the two-component family of response regulators.** *Infect Immun* 1998, **66**:5854-5861.

36. Cotter PD, Emerson N, Gahan CG, Hill C: **Identification and disruption of lisRK, a genetic locus encoding a two-component signal transduction system involved in stress tolerance and virulence in *Listeria monocytogenes.*** *J Bacteriol* 1999, **181**:6840-6843.

37. Bras AM, Chatterjee S, Wren BW, Newell DG, Ketley JM: **A novel *Campylobacter jejuni* two-component regulatory system important for temperature-dependent growth and colonization.** *J Bacteriol* 1999, **181**:3298-3302.

38. Raivio TL, Popkin DL, Silhavy TJ: **The Cpx envelope stress response is controlled by amplification and feedback inhibition.** *J Bacteriol* 1999, **181**:5263-5272.

39. Dorel C, Vidal O, Progent-Combaret C, Vallet I, Lejeune P: **Involvement of the Cpx signal transduction pathway of *E. coli* in biofilm formation.** *FEMS Microbiol Lett* 1999, **181**:169-175.

40. Ernst RK, Guina T, Miller SI: **How intracellular bacteria survive: surface modifications that promote resistance to host innate immune responses.** *J Infect Dis* 1999, **179**:326-330.

41. Wosten MM, Groisman EA: **Molecular characterization of the PmrA regulon.** *J Biol Chem* 1999, **274**:27185-27190.

42. Novak R, Henriques B, Charpentier E, Normark S, Tuomanen E: **Emergence of vancomycin tolerance in *Streptococcus pneumoniae.*** *Nature* 1999, **399**:590-593.

43. Calera JA, Zhao XJ, De Bernardis F, Sheridan M, Calderone R: **A virulence of *Candida albicans* CaHK1 mutants in a murine model of hematogenously disseminated candidiasis.** *Infect Immun* 1999, **67**:4280-4284.

44. Yamada-Okabe T, Mio T, Ono N, Kashima Y, Matsui M, Arisawa M, Yamada-Okabe H: **Roles of three histidine kinase genes in hyphal development and virulence of the pathogenic fungus *Candida albicans.*** *J Bacteriol* 1999, **181**:7243-7247.

45. Burbulys D, Trach KA, Hoch JA: **The initiation of sporulation in *Bacillus subtilis* is controlled by a multicomponent phosphorelay.** *Cell* 1991, **64**:545-552.

46. Hagen, KD, Meeks JC: **Biochemical and genetic evidence for participation of DevR in a phosphorelay signal transduction pathway essential for heterocyst maturation in *Nostoc punctiforme* ATCC 29133.** *J Bacteriol* 1999, **181**:4430-4434.

47. Cho K, Zusman DR: **Sporulation timing in *Myxococcus xanthus* is controlled by the *espAB* locus.** *Mol Microbiol* 1999, **34**:714-725.

48. Thomason PA, Traynor D, Stock JB, Kay RR: **The RdeA-RegA system, a eukaryotic phospho-relay controlling camp breakdown.** *J Biol Chem* 1999, **274**:27379-27384.

49. Calera JA, Calderone R: **Flocculation of hyphae is associated with a deletion in the putative *CaHK1* two-component histidine kinase gene from *Candida albicans.*** *Microbiol* 1999, **145**:1431-1442.

50. Imamura A, Hanaki N, Nakamura A, Suzuki T, Taniguchi M, Kiba T, Ueguchi C, Sugiyama T, Mizuno T: **Compilation and characterization of *Arabiopsis thaliana* response regulators implicated in His-Asp phosphorelay signal transduction.** *Plant Cell Physiol* 1999, **40**:733-742.

51. Susuki T, Imamura A, Ueguchi C, Mizuno T: **Histidine-containing phosphotransfer (HPt) signal transducers implicated in His-to-Asp phosphorelay in *Arabidopsis*.** *Plant Cell Physiol* 1998, **39**:1258-1268.

52. • D'Agostino IB, Kieber JJ: **Phosphorelay signal transduction: the emerging family of plant response regulators.** *Trends Biochem Sci* 1999, **24**:452-456.
A cogent review of phosphorelay involvement in plant development.

53. Chang C, Shockey JA: **The ethylene-response pathway: signal perception to gene regulation.** *Curr Opin Plant Biol* 1999, **2**:352-358.

54. Kiba T, Taniguchi M, Imamura A, Ueguchi C, Mizuno T, Sugiyama T: **Differential expression of genes for response regulators in response to cytokinins and nitrate in *Arabidopsis thaliana.*** *Plant Cell Physiol* 1999, **40**:767-771.

55. Coello P, Polacco JC: **ARR6, a response regulator from *Arabidopsis*, is differentially regulated by plant nutritional status.** *Plant Sci* 1999, **143**:211-220.

56. D'Agostino IB, Kieber JJ: **Molecular mechanisms of cytokinin action.** *Curr Opin Plant Biol* 1999, **2**:359-364.