

## 22. ENVIRONMENTAL SENSING

2 January 2023

In order to survive, reproduce, and physiologically adjust in appropriate ways at the correct times, nearly all species constantly monitor their internal and external environments. Assessment of extracellular conditions usually involves trans-membrane proteins, with an external domain serving as an environmental sensor, and an internal domain transmitting signals to messenger proteins that further elicit appropriate cellular responses. These signal-transduction (ST) pathways may involve multiple steps, but information exchange almost always involves a series of chemical and/or physical changes in the pathway participants. ST systems are central to the nervous systems of metazoans, but for unicellular species, they are the nervous system.

Unraveling the features of both single molecules and ensembles of them is key to understanding the function and evolutionary properties of communication systems at the cellular level. At the single-molecule level, ST processes are digital in the sense that each molecular participant exhibits a finite number of effectively discrete phenotypes, e.g., active vs. inactive conformational states. At the whole-cell level, such information will be distributed over all of the relevant ST molecules in the cell, providing a more graded and accurate assessment of environmental states.

The three main features of all ST pathways are sensitivity, accuracy, and speed. First, external chemical concentrations are often in the  $\mu\text{M}$  range (Chapter 18), so cells have to make decisions based on encounters with small numbers of molecules. Second, each ST pathway is devoted to a specific environmental stimulus (or small set of them), the menu of which is very broad, including organic metabolites, inorganic nutrients, markers of pathogens, atmospheric gases, osmolarity, and antibiotics. With up to several dozen pathways operating simultaneously within the confines of single cells, avoidance of crosstalk among pathways is critical to maintaining coherent cellular responses. Third, the chemical systems involved must generate responses on appropriate time scales. Too slow a response can leave a cell in a compromised physiological state, but too rapid a response can be a major energy drain on the bewildered cell.

Many of the central players in signal transduction are enzymes, so understanding the operation of such systems requires an appreciation of the basic features of enzyme kinetics. But this is not enough. As reviewed in Foundations 19.1, the kinetics of simple one-enzyme systems are such that there is generally a smooth, hyperbolic relationship between substrate concentration and enzyme output. In contrast, ST pathways are generally constructed in ways that can yield much sharper responses of the whole system to external ligand concentrations. In extreme cases, this amplification of external signals can lead to near switch-like behavior in phenotypic outputs.

ST systems invoke many evolutionary questions that remain to be answered in a

convincing fashion. First, as the accurate transmission of environmental information along chains of pathway molecules is key to signal transduction, the recurring theme of the evolution of cohesive molecular languages again becomes central (Chapters 15, 19, and 21). Second, because the recording and erasing of information is energy demanding, questions emerge about the critical threshold above which the gain in information is offset by the energetic cost of building and maintaining a communication system. Third, the innate capacity of many ST systems to generate populations of individuals with discrete alternative states (in the absence of genetic variation) raises questions as to whether such systems are exploited by natural selection to operate as bet-hedging strategies, as opposed to being inadvertent by-products of the structure of ST networks.

In the following pages, these issues will be explored mainly from the standpoint of bacterial ST systems, which owing to their simplicity have been studied in much more detail than the more complicated ST networks typically operating in eukaryotes. A broad overview of the biology of ST systems, more focused on eukaryotes, is given by Lim et al. (2015). Drawing from many examples, Wan and Jékely (2021) have argued that eukaryotes owe their success to the evolution of more diverse and complex sensory systems, implying that these are also more refined in terms of speed and accuracy. However, as noted multiple times in previous chapters, success in the eye of the eukaryotic beholder is often a false caricature of the actual situation. Increased complexity need not mean increased functionality, and it remains unclear if prokaryotic sensory systems are less efficient in terms of time and/or energy, let alone less accurate than those of eukaryotes.

## Bacterial Signal-transduction Systems

Relative to the complex ST systems of eukaryotes (below), those of bacteria typically have simple enough structures that their operational features can be dissected in detail. The simplest type of bacterial ST mechanism is the so-called one-component system, which consists of just a single protein (usually a cytosolic transcription factor) with an input domain serving as a signal receptor (sensor) and an output domain transmitting information to a receiver (Figure 22.1). Almost always, the incoming signal is a small ligand molecule that allosterically modifies the protein in such a way as to activate the response domain, which then induces transcription in one or more downstream target genes (Ulrich et al. 2005).

The second most common mechanism of signal transduction in bacteria is the two-component system (Stock et al. 2000; Capra and Laub 2012), the operation of which always involves post-translational modification of protein participants. The first component in such systems, the signal receptor, is generally a histidine kinase (HK) embedded in the cell membrane. Kinases are enzymes that catalyze the transfer of a terminal phosphate from ATP to an amino acid on a target protein. The extracellular domain of a HK receives environmental information, usually in the form of a small ligand that induces autophosphorylation (addition of a phosphoryl group,  $\text{PO}_3^-$ ) of a specific histidine residue on the internal domain (Figure 22.1). The phosphoryl group is then transferred to a specific aspartate residue on the sec-

ond (intracellular) component, known as the response regulator (RR). This transfer elicits a conformational change in the RR that in turn induces a specific cellular response, usually with the RR operating as a transcription factor. Almost all HK and RR proteins in such systems have homodimeric structures.

There are thus six key determinants of specificity in a two-component system: the ligand-binding, phosphotransfer, and dimerization domains of the HK, and the receiver, DNA-binding, and dimerization domains of the RR. Moreover, most HK proteins are bifunctional in that when not phosphorylated, they operate as phosphatases on their cognate RR proteins. The ratio of kinase to phosphatase activity dictates the output of the pathway. In *E. coli*, an average HK protein is present as 10 to 100 molecules per cell, whereas the cognate response regulators are generally 10 to 100× more abundant.

More complex phosphorelay systems exist in some bacteria, the simplest of which consist of three components (Figure 22.1). Here, as with two-component systems, a membrane-bound HK molecule first autophosphorylates at an internal histidine residue after receiving an appropriate extracellular signal. But in this case, the phosphoryl group is then transferred to a secondary receiver domain, often on the same molecule (usually an arginine residue). Given the presence of both a transfer and acceptor domain in the same molecule, such a protein is often referred to as a hybrid kinase. A separate protein transfers the phosphoryl group to the final acceptor, again a cytosolic RR protein. A number of variants on this type of pathway are known, including chains of transfers from a histidine to an arginine to another histidine, and so on. The *Bacillus subtilis* sporulation control system is an example of a chain involving a series of four proteins (Sonenshein 2000).

**Origin and diversification.** The numbers of both one- and two-component systems in different bacterial species scale with roughly the square of genome size (Ulrich et al. 2005; Alm et al. 2006; Capra and Laub 2012). One-component systems are about 7× more abundant than two-component systems, with most bacterial species containing dozens (in a few cases, hundreds) of such systems in total (Figure 22.2). The HK and RR genes for most two-component systems reside in the same operons, and hence are coexpressed. Nevertheless, many cases are known in which single members of an operon are duplicated, and as a consequence on average, orphan HK and RR genes are nearly as abundant as those in operons (Burger and van Nimwegen 2008). In addition, although they are a minority, many-to-one and one-to-many HK-RR systems exist (Goulian 2010), with the numbers of HK proteins in a genome usually being up to twice the number of RR proteins (Figure 22.2). This implies that multiple signals are often transmitted through the same response regulator.

Given the very precise mode of operation of all such systems – initial phosphorylation of a histidine on one protein, followed by phosphotransfer to an aspartate on another, it is likely that this was the ancestral state of a primordial two-component system from which most others were subsequently derived by duplication and divergence. Why histidine and arginine became the chosen amino acids remains unclear. Why phosphorylation was settled upon as the key form of post-translational modification is also unclear, although the two negatively charged oxygens associated with each phosphate group provide opportunities for the modification of protein structure

by binding with positively charged residues. Such structural changes can then be linked to alterations in protein function in a binary fashion.

The fact that the vast majority of one- and two-component systems acquire environmental information via small-molecule binding, and elicit a final output via transcriptional regulation motivates the suggestion that one-component systems provide the evolutionary seeds of two-component systems (Ulrich et al. 2005). However, any such transition requires several modifications: minimally, the insertion of a histidine kinase domain and a receptor domain, the physical separation of these two domains into two separate proteins, and the acquisition of a trans-membrane domain by the HK. Moreover, there is no compelling reason to rule out the alternative possibility that one-component systems are often derived and simplified versions of two-component systems.

**Coevolutionary integration of components.** The multiplicity of ST systems within cells and their operation by pairwise communication raise many questions about their evolutionary properties. One of the central issues concerns the mechanisms by which individual systems avoid the risks of miscommunication with noncognate systems. Each HK generally communicates with a specific RR, although some low level of error must occur. Indeed, some systems exhibit a low degree of crosstalk when examined *in vitro* (Yamamoto et al. 2005), and crosstalk can be greatly enhanced if cognate partners are eliminated *in vivo*, owing to the release from competitive binding (Siryaporn and Goulian 2008). Pathway insulation is in part an ingrained consequence of the biochemical nature of bacterial two-component pathways, in particular the editing-like properties of the HK molecules. When the proper signal for a particular pathway is lacking, its HK remains unphosphorylated and acts primarily as a phosphatase for the cognate RR, thereby tending to erase inadvertent phosphorylation of an RR by noncognate HKs.

Mutual HK-RR recognition is generally a function of coevolutionary changes accumulated on just a small number of amino-acid residues at phosphotransfer domains, typically  $< 6$  sites on each protein (Li et al. 2003; Laub and Goulian 2007; Weigt et al. 2009; Capra et al. 2012b). As a consequence, the requirements for specificity rewiring are not high. For example, alterations of just three amino-acid residues of the *E. coli* HK EnvZ (involved in osmoregulation) are sufficient to both eliminate its ability to recognize its cognate RR and to confer full specificity toward one particular noncognate RR (Skerker et al. 2008; Capra et al. 2010; Nocedal and Laub 2022). Just single amino-acid changes can lead to mutual recognition of cognate and noncognate RRs, showing that under the right situations, a HK can gain an ability to recognize a novel RR without relinquishing its initial partner. Indeed, Siryaporn et al. (2010) found that single amino-acid substitutions in an *E. coli* sensor kinase called CpxA (involved in envelope stress response) can cause the efficiency of signaling to a noncognate RR to even exceed that of the latter's cognate HK.

In a broader attempt to understand the degree of recognition-motif degeneracy, Podgornaia and Laub (2015) made constructs of all possible  $20^4 = 160,000$  amino-acid motifs for the four key recognition residues in *E. coli* protein kinase PhoQ, a signal receptor for external magnesium concentration. From this pool of variants, 1659 were found to be functional. Extending this analysis even further to include

the recognition motif on the cognate response regulator PhoP, McClune et al. (2019) found 58 unique PhoQ-PhoP motif combinations that yielded fully functional systems that were also completely insulated from the native system.

Given that many two-component systems operate in an essentially one-to-one manner, these kinds of observations also suggest the capacity for substantial neutral coevolutionary drift between interacting motifs, even in the face of selection for conserved function. Recalling the theory formally evaluated in Chapter 21, such systems drift in coupled regulatory vocabularies is expected in simple pairwise interactions, so long as the maintenance of a strong degree of mutual interaction is retained (Lynch and Hagner 2015). In principle, this could then lead to the evolution of incompatibilities among mixtures of orthologous HKs and RRs from different taxa, in the absence of any within-species functional changes.

The few experimental attempts to shed light on this matter have yielded mixed results. On the one hand, divergence of the recognition motifs of the components of the bacterial PhoR-PhoB system (involved in phosphate regulation) between members of the Alphaproteobacteria and Gammaproteobacteria is sufficient to nearly completely prevent crosstalk (Capra et al. 2012b). On the other hand, two studies of other systems have shown that the HK gene from a different bacterial phylum can complement the loss of the orthologous *E. coli* gene (Tabatabai and Forst 1995; Ballal et al. 2002). Likewise, the conserved ability to phosphorylate orthologous substrate proteins from distantly related species has been noted for a different class of bacterial signaling proteins, the tyrosine kinases, in this case despite the lack of obvious sequence homology (Shi et al. 2014). These kinds of observations are not necessarily incompatible with a hypothesis of neutral systems drift, although they do highlight uncertainties in the degree to which the evolution of sequence motifs in the individual participants are mutually constrained.

**Emergence of new pathways.** Although these types of observations make clear that two-component systems can be rewired with only a small number of changes, the challenges in establishing an entirely new ST system are numerous. A common idea is that the evolution of a novel system initiates with duplication of both members of the pair, a scenario made plausible in bacteria by the frequent joint occurrence of cognate HK and RR genes in the same operon and by the high degree of congruence between the phylogenetic trees of HK and RR genes conjoined within operons (Koretke et al. 2000). Linkage within operons ensures that both members of an interacting pair will be coexpressed from their time of origin, an essential ingredient for coevolutionary reinforcement.

However, ultimate preservation by neofunctionalization requires the integration of a new signal input and/or output into a system in a way that avoids crosstalk with the ancestral system (Figure 22.3). This, in turn, requires divergence in the HK-RR communication language used within the duplicated pairs (Capra et al. 2012b). Evidence suggests that the amino-acid sequences within HK-RR interface regions evolve at high rates at least in the early stages of post-duplication divergence (Rowland and Deeds 2014), and dramatic changes in recognition motifs are known to have accumulated among duplicated systems within the Gammaproteobacteria (Capra et al. 2012b).

While these observations are potentially consistent with diversifying selection,

modifications at the HK-RR phosphotransfer interface are not sufficient for the insulation of two pathways. There is also a need for changes in the dimerization interfaces of both the HK and RR molecules to prevent heterodimerization between the diverging copies. Consistent with more general observations on the relative simplicity of binding interfaces in multimeric enzymes (Chapter 13), experimental work suggests that changes involving fewer than four amino-acid residues in dimerization interfaces can suffice to establish a new homodimerization group (Ashenberg et al. 2011).

Capra and Laub (2012) and Rowland and Deeds (2014) have suggested that all of these crosstalk interactions must be removed before new input/output functions are acquired, arguing that in large bacterial populations even mutations with very mildly deleterious crosstalk effects would be immediately removed by selection. If this is indeed the case, then a transition to a novel signaling pathway would require an early order of events that is essentially neutral with little to no impact on the overall performance of the ancestral system.

A key difficulty with this hypothesis is the series of steps that must be accomplished – two losses of heterodimerization potential, one for the HK and one for the RR; loss of cross-phosphotransfer potential; and the emergence of at least one new HK-RR phosphotransfer interaction. During the period of time in which this series of events is achieved, both systems must also avoid nonfunctionalizing mutations (which would remove the entire system from selection). It may help that most bacterial populations are so large that most first-step variants (as well as double mutants) are always maintained by recurrent mutation. These might then provide the staging grounds for the emergence of downstream mutations, which can then proceed to fixation without any bottleneck in population fitness (Chapter 5). However, the population-genetic conditions necessary for the origin of an insulated, coevolving pair of proteins need to be worked out formally to resolve these numerous open questions (Foundations 15.1 provides a starting point).

Finally, a plausible case can be made that more complex phosphorelay systems (Figure 22.1) arise by fusion of the HK and RR components of two-component systems to produce a hybrid kinase. In principle, this may involve nothing more than deletion of an intergenic region within an operon, provided the open-reading frames of both components remain intact (Zhang and Shi 2005; Cock and Whitworth 2007). Such a starting point would facilitate the evolution of a new signaling system, as the HK and RR are open to communication from the onset. In addition, their enforced proximity within the same molecule would reduce the necessity of high affinity between the pair, thereby enhancing the likelihood of evolutionary motif divergence. Consistent with this idea, empirical work has shown that when the kinase domain of a phosphorelay system is disconnected from its receiver domain, the level of crosstalk increases substantially (Wegener-Feldbrügge and Søgaard-Andersen 2009; Capra et al. 2012a). Once established with novel communication motifs with little potential for crosstalk, such a phosphorelay system might then revert back to the structure of a two-component system by a fission event.

## Interconvertible Proteins and Ultrasensitivity

Despite their relatively simple modular structures, ST pathways exhibit an array of unusual features at the biochemical and cellular level. Central to understanding such behavior is the concept of an interconvertible protein whose active vs. inactive states are defined by the presence/absence of post-translational modifications. As already discussed, the most common case by far is the phosphorylation/dephosphorylation cycle, in which a specific ATP-dependent kinase attaches a phosphoryl group to a particular amino-acid residue on the interconvertible protein, and a specialized phosphatase is responsible for the reverse reaction. In the simplest bacterial systems, the same enzyme is often used for both the addition and removal of the modification, but in eukaryotes different enzymes are generally deployed in each transformation. The joint activities of kinases and phosphatases, along with the concentration of the intermediate protein (here viewed as a response regulator), determine the fractional activity of the latter, which ultimately dictates the cellular response.

In the following discussion, the three proteins will be denoted as F (forward converter, e.g., a kinase), R (reverse converter, e.g., a phosphatase), and I (interconvertible protein) (Figure 22.4). Although such systems have discrete on/off states at the single-molecule level, this is not the case for the entire ensemble of molecules at the cellular level. Instead, the proportional levels of alternative forms of I (active  $I_a$ , and inactive  $I_i$ ) can fall over an essentially continuous range of 0.0 to 1.0, depending on the activities of the converter enzymes (F and R). In one range of parameter space, the forward (kinase) reaction will dominate, and the majority of I will exist in its active form, whereas for other parameter values, the reverse (phosphatase) reaction will dominate, rendering the average molecule of I inactive. As will be revealed below, however, the real novelty of such systems is their capacity to generate switch-like behavior, from nearly completely off to nearly completely on with just a small change in signal concentration.

The degree to which the fractional activation of I depends on the concentration of the external signal, in the form of a ligand  $S_F$ , dictates the magnitude of the overall cellular response. Recall that with standard two-parameter Michaelis-Menten enzyme kinetics, there is a fairly gradual response of the system output to the substrate concentration, describable with just two parameters (Chapter 19). In contrast, with a triad of interacting proteins, more than ten parameters, including the kinetic coefficients of enzymes F and R, determine the partitioning of the total concentration of I, denoted  $[I_T]$ , into its active and inactive forms (Figure 22.4; Foundations 22.1). As a consequence, the level of I activation can exhibit a far richer array of behaviors than possible with basic enzyme kinetics, even though both enzymes F and R behave as Michaelis-Menten enzymes with I as their substrate.

Consider the situation in which the active vs. inactive form of converter enzyme F depends on whether it is bound to its ligand  $S_F$  (Figure 22.4), and recall that with Michaelis-Menten enzymes, the rate of a reaction is hyperbolically related to the substrate concentration,  $[S_F]$ , as described with Equation 19.1.4. With increasing  $[S_F]$ , enzyme F is expected to be increasingly converted to its active form  $F_a$ . If, however,  $F_a$  feeds into a loop involving the interconvertible enzyme I, the fraction of active enzyme,  $I^* = [I_a]/[I_T]$ , can reach much higher levels than the fraction of active F at low levels of the input substrate  $S_F$  (Figure 22.5, upper panel). In other words, the signal from the external ligand can be substantially amplified. This is because the total concentration of I constitutes a closed system, enabling  $F_a$  to cumulatively

convert  $I_i$  to  $I_a$ .

The degree of amplification also depends on the kinetic parameters of the reverse converter, which provides the only route back to  $I_i$ . As the forward conversion reaction increasingly overwhelms the reverse reaction, the phosphorylation reaction dominates, and  $[I_a] \rightarrow [I_T]$ . On the other hand, as the kinetic efficiency or amount of enzyme R increases, the phosphatase reaction increasingly dominates, and the system can converge to situations in which I can never attain a fully active state, even at the highest concentrations of the external ligand (Figure 22.5, upper panel).

A key assumption underlying the preceding results is that the total concentration of I is substantially below the half-saturation constants for the forward and reverse reactions (Stadtman and Chock 1977), so that the active forms of enzymes F and R are not saturated by their substrate. In this case, a simple expression can be obtained for the fraction of activated intermediate enzyme,

$$I^* = \frac{\kappa_F [F_a]}{\kappa_R [R_a] + \kappa_F [F_a]}, \quad (22.1)$$

where  $\kappa_x$  is the kinetic efficiency of enzyme x operating on substrate I (from Equations 22.1.7a,b). In this nonsaturating case, the forward and reverse rates of conversion of I are both linearly related to their substrate concentrations, and  $I^*$  is independent of the total concentration of intermediate enzyme,  $[I_T]$ , in the system. Furthermore, because the amounts of active converter enzymes,  $[F_a]$  and  $[R_a]$ , are Michaelis-Menten functions of their ligand concentrations (Equations 22.2.3a,b),  $I^*$  is also a conventional hyperbolic function, in this case of  $\kappa_F [F_a]$ .

Goldbeter and Koshland (1981) found that with increasing concentration of I in the system (so that the responses of  $F_a$  and  $R_a$  to their substrate concentrations are no longer linear), the response of  $I^*$  to ligand concentrations is no longer hyperbolic or independent of  $[I_T]$ . Rather, the steepness of the activation response to ligand concentrations elevates dramatically with increasing  $[I_T]$ , in the extreme becoming an effectively stepwise process (Figure 22.5, lower panel). This sharp response, often referred to as zero-order ultrasensitivity, arises because high levels of I allow the converter enzymes to operate at maximum capacity, thereby sharpening their responses near the threshold between the kinase and phosphatase domains on the scale of ligand concentrations. At low levels of the external ligand, the reverse (phosphatase) reaction dominates, and with a high level of I pushes the rate of a conversion to  $I_i$  to the maximum. At high levels of the external ligand, the forward reaction overwhelms the reverse reaction, pushing  $I_a$  to the maximum level.

To summarize, the use of an interconvertible-enzyme system can alter both the sensitivity and the amplitude of response of a signaling system to the input ligand concentration. Most notably, for sufficiently high concentrations of the central enzyme I, near switch-like behavior of the population of active I molecules arises. Thus, should a selective scenario exist in which switch-like behavior is advantageous, mutational fine tuning of the kinetic parameters of the enzymes underlying the ST system, combined with the maintenance of a sufficiently high level of I, can provide an evolutionary path towards such behavior.

Here, it should be emphasized that it is an open question as to whether the switch-like behavior implied by the mathematics of these kinds of systems commonly occurs in cells, let alone is promoted by selection. In fact, zero-order ultrasensitivity has not yet been directly demonstrated in *in vivo* signal-transduction

systems (Blüthgen 2006). Moreover, as pointed out by Ortega et al. (2002) and Xu and Gunawardena (2012), such an extreme response requires that the phosphorylation/dephosphorylation reactions are intrinsically irreversible, such that a predominating kinase can literally drive I to the point at which all molecules are in the active state. With most enzymes, a high product concentration generally drives reactions in their reverse direction, leading to a steady-state situation in which both active and inactive molecules coexist within a cell. The situation is even more complicated in bacterial two-component systems, where the same enzyme often serves both the F and R functions, rendering the maintenance of high substrate concentrations for both enzymatic functions (necessary for ultrasensitivity) difficult to achieve.

**The cost of signal transduction.** Acquisition, processing, and propagation of information requires energy, whether via a computer (Landauer 1988) or by the nervous system of a metazoan (Mehta and Schwab 2012; Niven 2016; Kempes et al. 2017; Levy and Calvert 2021). Quiescent nerve tissue consumes energy just to maintain a steady response capacity, and the same is true for the interconvertible enzymes at the heart of ST systems. In addition to the structural costs, there is the additional matter that the relay signal (the relative concentrations of active and inactive I molecules) must be constantly adjusted by the simultaneous running of two pathways (phosphorylation and dephosphorylation) in opposing directions in order to convey information on the cellular environment.

For this reason, the dynamics of the proteins underlying ST systems are often referred to as push-and-pull or futile cycles – even when the system settles into a steady-state in a constant environment, active and inactive I molecules are continuously being interconverted, leading to a cyclical flux. However, the word futile is a bit of a misnomer here, in that phosphorylation/dephosphorylation cycles are the price a cell must pay to keep a ST system in a constant state of readiness.

Because each phosphorylation event requires hydrolysis of an ATP molecule, a rough idea of the cost of maintaining a particular level of system activity can be obtained by noting that at steady state, the reciprocal rates of activation/inactivation of I molecules must be completely balanced and equal to the rate of ATP consumption. The cost of running a signal-transduction system can then be obtained from measures of *in vivo* rates of ATP consumption by the pathway kinases. For example, Shacter et al. (1984) noted that the flux through hepatic (liver) pyruvate kinase is  $V_{\text{ATP}} = 20$  to  $200 \mu\text{M}/\text{min}$ , depending on the level of activation of the intermediate protein. Assuming a cell volume of  $\sim 5000 \mu\text{m}^3 = 5 \times 10^{-12}$  liters, a cell division time of 24 hours, and converting moles to number of molecules using Avogadro’s number, the total rate of ATP consumption per cell by this kinase is then of order  $10^{10}$  to  $10^{12}$  molecules/cell division.

As usual, a quantitative understanding of what this energetic cost means to the cell requires information on the total cellular energy budget. From Chapter 8, for a cell of this size and cell-cycle length, total cell maintenance costs are  $\sim 2 \times 10^{13}$  ATP hydrolyses/cell cycle, implying that a single mammalian ST system can demand up to 5% of a cell’s basal-maintenance energy budget. From a knowledge of the concentration of the kinase and phosphatase in this system and their molecular sizes ( $\sim 2600$  and  $\sim 1430$  amino acids, respectively), the construction cost of the entire system can be shown to be  $\sim 2 \times 10^{11}$  ATPs (not including the cost of the

interconvertible enzyme). Thus, in this particular example, the costs of running and building the system are of the same order of magnitude.

Estimates for other mammalian kinases derived in Goldbeter and Koshland (1987) are in this rough range as well. Studies with yeast, ciliates, and zebrafish embryos have shown that such expenditures in cell signaling can be sufficient enough to be discerned as heat-dissipation oscillations during the eukaryotic cell cycle (Poole et al. 1973; Lloyd et al. 1978; Rodenfels et al. 2019).

**Similarities and differences in eukaryotic systems.** Eukaryotic signal transduction systems are generally much more complex than those in bacteria. Although kinases and phosphatases are still broadly utilized, different amino acids typically serve as the sites of phosphorylation – usually serine, threonine, and tyrosine, with the latter largely confined to metazoans. Moreover, whereas bacterial histidine kinases produce phosphoramidates by phosphorylating side-chain nitrogen atoms, eukaryotes phosphorylate oxygen atoms on serine, threonine, and tyrosine residues, creating phosphoesters, which are much more energetically stable. The reasons for this shift in target residues is unclear, but one possibility is that enhanced stability is essential for activated molecules that have to travel larger distances in eukaryotic cells. Some two-component systems involving autophosphorylating histidine kinases are known in plants, fungi, and slime molds, but they are absent from a number of eukaryotic lineages, notably metazoans (Loomis et al. 1997; Schaller et al. 2011).

Given that eukaryotes are derived from archaea, it is of interest to know the nature of ST systems in the latter, but there is an unfortunate void of knowledge here. It is known that many archaea are entirely lacking histidine kinases, and that perhaps all have serine/threonine kinases (Makarova et al. 2017). Most unusual is the apparent exploitation of KaiC-like ATPases in a wide variety of signaling pathways across the archaeal phylogeny. Recalling from Chapter 18 that KaiC is at the heart of the phosphorylation/dephosphorylation cycle that forms the circadian clock in cyanobacteria, the latter may have been acquired by horizontal transfer from archaea.

The numbers of kinases in eukaryotes scale with roughly the square of the total number of proteins, similar to what is seen in bacteria, and substantial lineage-specific expansions of particular families have arisen (Anantharaman et al. 2007). However, in eukaryotes, each of the interacting enzymes typically engages with multiple substrate proteins (Figure 22.6), often more than a dozen, leading to complex networks quite unlike the well-insulated systems of bacteria. Hence, the elegant molecular dissections that have been accomplished for bacterial ST systems are a rarity for eukaryotes. Nonetheless, for the few simple eukaryotic systems that have been investigated, many of the principles outlined above for bacteria with respect to motif evolution still apply. For example, a transcription factor involved in the response to amino-acid starvation in the yeast *S. cerevisiae* (Gcn4) is targeted for degradation by a specialized kinase (Pho85), but whereas the same system operates in the yeast *Candida albicans*, the cross-species components are incapable of molecular recognition (Gildor et al. 2005). As discussed above for bacteria, this kind of coevolutionary wandering of the motif language used in information transfer in the face of conserved function is consistent with the operation of mutually constrained systems drift.

There can, however, be limits to the degree to which such wandering extends. For example, Zarrinpar et al. (2003) found that a kinase (Pbs2) involved in the osmoregulation pathway in *S. cerevisiae* interacts specifically with one particular membrane-bound sensor protein, despite the presence in this species of 26 other related sensor proteins with rather similar recognition sequences. However, when orthologs of these off-target proteins from other distantly related species were presented to Pbs2, strong cross-species recognition often occurred. This suggests that in this case there has been significant negative selection within yeast to avoid off-target interactions, with the suppressing motifs diverging among lineages.

Unlike the situation in bacterial ST systems where just a single amino-acid residue is typically modified on the intermediate protein to elicit a response, eukaryotic ST proteins commonly have multiple phosphosites (Chapter 14). Because the activity of the modified enzyme can require a complete set of phosphorylated sites, and the sequential ordering of marks may follow a rigid recipe, this introduces novel twists to the types of models discussed in the preceding section. Contrary to the switch-like behavior found with simple systems with single-site modifications, the use of multiple phosphorylation sites for activation leads to a more graded response. Although there is still a threshold level of substrate below which the system is inactive, there can be a simple Michaelis-Menten-like response above the threshold (Gunawardena 2005). Thus, although the reason for the use of multiple phosphosites in eukaryotic ST-pathway enzymes remains unclear, refined enhancement of switch-like behavior does not seem to be a viable hypothesis.

Limited attention has been given to the consequences of the kinds of shared utilization of kinases/phosphatases illustrated in Figure 22.6 (left). However, focusing on the simple case of enzymes with single modifiable sites, Rowland et al. (2012) found that coupled systems can often behave in a transitive fashion, such that if one intermediate substrate saturates the controlling enzymes in a way that leads to switch-like ultrasensitivity, all other connected intermediate enzymes will behave in the same way. In effect, saturation by one intermediate substrate alters the controlling enzymes (F and R) to fully active states, governing the entire system. This can even happen when all multiple intermediate substrates are below saturation levels, provided the aggregate is sufficient for saturation. Whether these kinds of collective effects are evolved mechanisms for coordinated ultrasensitivity in eukaryotic cells or simple inadvertent consequences of complex networks remains unclear.

One of the most pronounced differences between eukaryotic and bacterial ST systems is the extended length of the former, with a relay of three kinases being particularly common (Figure 22.6, right). For example, MAP (mitogen-activated protein) kinase kinase kinases phosphorylate MAP kinase kinases, which in turn phosphorylate MAP kinases that finally transmit information to a response regulator. Many variants of the MAP kinase family are deployed in wide variety of eukaryotic cellular processes, and their efficiency of operation and degree of insulation is enhanced by the use of scaffold proteins that physically link all three layers of each pathway into single complexes.

How and why ST pathways with extra steps evolve remains unclear. Armbruster et al. (2014) argue that additional steps enable pathways to integrate out the effects of environmental noise on the elicited response, but again, this leaves unanswered the question as to why bacteria would not take advantage of such possibilities, especially

given the likelihood that small bacterial cells may be more subject to stochastic effects than their larger eukaryotic counterparts. An alternative possibility is that, as with gene regulatory systems (Chapter 21), more complex systems passively emerge in lineages experiencing elevated levels of random genetic drift.

Finally, eukaryotes harbor another broad class of intracellular messaging systems, the so-called G proteins, used in a wide range of cellular activities, including vesicle transport and import/export through nuclear pores (Chapter 15). These proteins operate in a quite different way than those noted above, through binding of GTP rather than via phosphorylation of amino acids. Nonetheless, the kinetics of the overall systems follow the same general plan as the interconvertible proteins noted above. G proteins have alternative on/off states driven by opposing enzymes responsible for GTP addition and removal, which in turn lead to conformational changes in the substrate G protein: writers called guanine nucleotide exchange factors (GEFs) add GTP to the G protein, putting it in the active state, whereas erasers called GTPase-activator proteins (GAPs) hydrolyze the GTP to GDP. Most GEFs are membrane-bound G-protein-coupled receptors (GPCRs) that become activated upon binding an appropriate ligand, presenting a still further analogy with the types of systems noted above. A whole realm of theoretical investigation, involving all of the issues noted in Chapter 21 for gene expression, awaits exploration here (Kapp et al. 2012).

## Chemotaxis

The most ornate ST systems in bacteria drive chemotaxis, which directs motility towards particular chemo-attractants (or away from repellants). Although such systems have a kinase and a response regulator at their core, the output is modulated by up to nine other participants (Figure 22.7). The system operates like a “bacterial eye.” Rather than being membrane-bound, the histidine kinase, CheA, is located at the base of a sensory complex, connected to a special set of chemoreceptors (called methyl-accepting chemotaxis proteins or MCPs) residing at the cell surface. Rather than functioning as a transcription factor, the response regulator, CheY, interacts directly with the base of the flagellum, influencing the direction of flagellar rotation and thereby the swimming behavior. This allows for a much more rapid response than with conventional two-component systems that regulate gene expression.

The MCPs typically sit in one or two large hexagonal arrays at the cell surface, organized in a honeycomb-like form, with thousands of trimers of receptor dimers forming the vertices and acting as relays to CheA (Briegel et al. 2009). In *E. coli*, there are five forms of MCPs, with various sensitivities to different ligands such as sugars or amino acids. As the different forms are mixed together within the arrays, this allows the cell to simultaneously process complex information about the environment. Cooperative interactions between adjacent elements sharpen the response (Mello et al. 2004). Absent from eukaryotes, this type of sense organ appears to have been horizontally transferred to some archaeal lineages (Briegel et al. 2015), even though the archaeal flagellum evolved independently (Chapter 16).

Although the complex kinetic and dynamical features of the *E. coli* chemo-

taxis system have been worked out in considerable detail (Barkai and Leibler 1997; Keymer et al. 2006; Mello and Tu 2007; Bitbol and Wingreen 2015; Colin and Sourjik 2017), these technicalities will not be covered here. The key point to appreciate is that bacterial chemotactic responses are achieved by comparing a current ligand concentration with that in the recent past, reflected in part by the level of receptor methylation. Specialized methyltransferases and methylesterases regulate the methylation level of four key residues on the MCPs, providing them with a capacity for adjusting the level of sensitivity over a broad range of ligand concentrations. High levels of methylation result from high levels of ligand concentration, but by altering ligand affinity, this resensitizes the system to higher levels of chemoattractant that would otherwise be saturating. Thus, the system operates similarly to the way in which a vertebrate eye adjusts to different levels of light, enabling the cell to maintain a constant sensitivity to changes in ligand concentration regardless of the absolute ligand concentration.

Transmission of information on the external ligand concentration to the modified response regulator is accomplished through the response mechanisms noted above. In this case, CheA and then CheY become phosphorylated / dephosphorylated in the absence / presence of ligand binding, and the resultant switch dictates whether the flagellum rotates in a clockwise vs. counterclockwise fashion. When phosphorylated, CheY becomes bound to the flagellar base, which causes tumbling and a change in swimming direction, whereas CheY disconnects when dephosphorylated by phosphatase CheZ, inducing the swimming that propels the cell forward. This sort of guided behavior then leads to a biased random walk in the direction of a perceived chemical gradient.

As noted above, the maintenance of the information utilized in ST systems requires an energy expenditure, and in this rapidly responding CheA/CheY system, the costs are quite high. For an *E. coli* cell, the opposing processes of phosphorylation and dephosphorylation result in the consumption of  $\sim 5 \times 10^6$  ATPs per hour (Lan et al. 2012; Govern and ten Wolde 2014), which is about 10% of the cost of swimming in this species (Chapter 16). Additional ATPs must also be consumed in the opposing methylation and demethylation reactions operating on the MCPs, although the numbers are unclear.

Approximately 50% of bacterial species have a chemotaxis system, although the architectural features vary widely, with some having additional proteins outside of the core CheA and CheY (Wuichet and Zhulin 2010; Abedrabbo et al. 2017). Whereas *E. coli* has five types of receptor proteins, the number ranges from 1 to 30 in other bacteria (Wadhams et al. 2005). For example, the purple photosynthetic bacterium *Rhodospirillum rubrum* has nine different receptor proteins, as well as four versions of CheA and six of CheY, which engage in nonrandom crosstalk, presumably broadening the capacity for environmental differentiation (Porter and Armitage 2002, 2004).

A clear example of the evolutionary rewiring of such systems is revealed by the contrast between the *E. coli* network, described above, and that in the soil bacterium *Bacillus subtilis*. Whereas interaction of phosphorylated CheY with the flagellar motor induces clockwise rotation and tumbling in *E. coli*, it induces counterclockwise rotation and directional swimming in *B. subtilis*; and whereas the phosphatase CheZ dephosphorylates CheY-P in *E. coli*, this function is carried out by a flagellar motor

protein in *B. subtilis* (Szurmant et al. 2004; Yang et al. 2015).

Chemotaxis enables organisms to move up resource gradients, thereby leading to an elevation of cell growth and clonal expansion. However, the mechanism of accrual of any such advantage may be more nuanced. In an otherwise homogeneous environment, populations of organisms at the edge of their range will indirectly generate a gradient of a chemo-attractant through their own activities, thereby causing continued migration (Adler 1966). Indeed, Cremer et al. (2019) found that even under conditions in which resources are nonlimiting to growth, *E. coli* still migrate towards a chemo-attractant of no nutritional value. This leads to range expansion, which increases the overall population growth rate, as the migrating cells at the leading edge extend the clonal range over a greater area, while the laggards utilize the still plentiful nutrients in the space left behind. Clones modified to be insensitive to chemo-gradients experience lower overall growth rates because more cells experience local resource limitation.

Notably, however, if the attractant itself is the primary nutrient, this effect is not seen. In this case, if the nutrient is nonlimiting, the strong self-produced gradient of attractant necessary for expansion does not occur, and if it is too limiting, the migrating wave depletes the nutrients remaining for any laggards (Cremer et al. 2019). Thus, although the usual view is that chemotaxis has evolved as a mechanism for moving towards more immediately beneficial conditions, these observations suggest a role for simply expanding into unoccupied locations even under nutrient replete conditions.

It is unclear whether the kinds of spatial structure that can arise on a completely unoccupied solid surface, as employed in these experiments, generalize to other settings. Notably, when *E. coli* is grown in a well-mixed liquid environment (which prevents the development of chemical gradients), cells increase their investment in motility when grown in nutrient-poor conditions, consistent with the idea that such a shift is a searching mechanism for more nutrient-rich situations (Ni et al. 2020). In addition, laboratory populations of *E. coli* exhibit higher levels of chemotaxis towards amino acids that serve as more nutritional resources, although such a correlation does not exist for *B. subtilis* (Yang et al. 2015). Such mixed results may exist because utility in a laboratory setting need not reflect the conditions under which differential chemosensitivity evolved in nature. For example, whereas amino acids are commonly used as nutrients in the intestinal bacterium *E. coli*, they may simply serve as indicators of other resources in the soil bacterium *B. subtilis*.

Much less attention has been given to the mechanisms of chemotaxis in eukaryotes, outside of issues related to cell migration and signaling in metazoan development. Unlike small bacteria, whose arrayed chemoreceptors monitor nutrients in a temporal manner, larger eukaryotic cells populate their entire surfaces with receptors and are able to sense spatial concentration gradients of  $< 5\%$  from the front to the rear of the cell (van Haastert and Postma 2007).

The best eukaryotic example comes from the slime mold *Dictyostelium discoideum*. During times of nutrient scarcity, the amoeboid cells of *Dictyostelium* aggregate into multicellular slugs, with recruitment being induced by waves of cyclic AMP emanating from the aggregation center about every 6 minutes. Surrounding cells respond by producing pseudopods in the direction of the front of the plume, and continue to do so even after the wavefront has passed such that the cells are

confronted with a downward gradient. Skoge et al. (2014) found that the solution to this “back of the wave” problem involves memory-like processes associated with positive feedback. As a wave approaches, the front of the cell is sensitized, whereas the back of the cell is desensitized, and the message to move forward persists for several minutes owing to the slow decay of the positive-feedback mechanism. The kinetics of the response system appear to have coevolved with the signaling system, as exposure to waves with periodicities exceeding 6 minutes leads to reversals in migratory behavior.

**Accuracy of environmental assessment.** To improve fitness, the environmental sensing mechanisms of cells must provide accurate information on the concentrations of relevant ligands in the surrounding medium. However, the capacity to make environmental assessments resides in the degree to which signal receptors on cell surfaces bind to their ligands, which is an inherently stochastic process, owing to fluctuations in molecular arrival times and binding success at individual receptors. This raises significant questions about the conditions under which chemoreception can actually convey accurate information about environmental conditions. These problems were first analyzed by Berg and Purcell (1977), who viewed the expected degree of occupancy of a receptor as a counting mechanism for assessing ligand concentrations (Foundations 22.2).

Considering the features of a single receptor molecule, the basis for their primary result starts with the assumption that the long-term average probability of occupancy of a receptor is described by a function of Michaelis-Menten form  $p = c_0/(K_D + c_0)$ , where  $c_0$  is the environmental concentration of the ligand, and  $K_D$  is the dissociation constant, equivalent to the concentration at which the receptor has a 50% probability of being bound (Chapter 19). Rearrangement of this expression shows how  $c_0$  is predicted by  $p$ . For low ligand concentrations,  $c_0 \ll K_D$ , the relationship between  $p$  and  $c_0$  is essentially linear ( $p \simeq c_0/K_D$ ), but with increasing  $c_0$ , the expected degree of occupancy approaches saturation (i.e.,  $p \simeq 1.0$ ). To be most effective in transmitting information, i.e., to maximize the response of  $p$ , a receptor should have a dissociation constant larger than the typical environmental concentration.

Owing to the transient nature of binding, individual receptors have binary states at any particular time (bound or unbound). Thus, accurate assessment of information via the degree of occupancy requires a long enough time for the averaging of repeated instances of binding and unbinding. Assuming the cell continuously monitors the environment for a time period  $T$ , a measure of the error in inference of the true environmental concentration ( $c_0$ ) is provided by the coefficient of variation (CV, ratio of the standard deviation of the inferred concentration to  $c_0$ )

$$\frac{\sigma_c}{c_0} = \sqrt{\frac{1}{2Dr_sc_0(1-p)T}}, \quad (22.2)$$

where  $D$  is the diffusion coefficient for the ligand,  $r_s$  is the radius of the receptor at the cell surface (the target size), and  $p$  is the function of  $c_0$  defined above (Berg and Purcell 1977).

Full derivation of this expression is given in Foundations 22.2, but its final structure meets intuitive expectations. The quantity  $4Dr_sc_0$  is the rate at which a

ligand particle diffuses to a receptor, and a longer  $T$  means that the receptor can integrate information over a longer series of bound and unbound states. Thus, the level of noise in assessment scales negatively with the encounter rate and time, but positively with the degree of occupancy  $p$ . As  $p \rightarrow \infty$ , the surface receptor becomes saturated, providing little quantitative information on the environmental state, i.e.,  $\sigma_c/c_0 \rightarrow \infty$ . As noted in Foundations 22.2, given typical estimates of  $D$ , even for quite low ligand concentrations, a measurement duration of several seconds can be sufficient to reduce the CV to near 0.01, i.e., to a level of  $\sim 99\%$  accuracy in the estimation of the true environmental concentration.

The statistical relationship conveyed by Equation 22.2 is just one of many possible criteria for evaluating the accuracy of environmental monitoring, in this case a time-averaged fractional occupancy of receptors. Arguing that only the length of unbound periods provides information on the environmental concentration of ligand, Endres and Wingreen (2009) showed that if the cell were instead somehow able to sense the duration of unbound intervals and use this alone as an estimate of  $c_0$ , the uncertainty in Equation 22.2 would be reduced by a factor of  $1/\sqrt{2}$ . Still another alternative arises if the “counter” resides within the cellular interior, in which case the result in Equation 22.2 would need to be multiplied by 1.6 (Berg and Purcell 1977); the noise is elevated in this case because molecules transiently trapped within the cell (as opposed to being released into the environment) can be recounted, effectively reducing the number of independent evaluations per unit time. Finally, whereas the preceding calculations assume that the cell is evaluating a constant environment, gradient sensing (i.e., monitoring the rate of change of ligand concentration, as in swimming up a gradient, or contrasting the inferred concentration at two ends of a stationary cell) might be employed. Endres and Wingreen (2008) found that gradient sensing at the cell surface yields a measure identical to Equation 22.2, whereas monitoring inside the cell leads to a  $2.9\times$  increase in the noise level, again emphasizing the advantages of monitoring at the cell surface.

Notably, all of the above refinements only change Equation 22.2 by a constant multiplier. However, all of these measures ignore the biochemical aspects of binding to receptors, making the assumption that the whole sensing process is essentially diffusion limited; the necessary modification for including the former is described in Foundations 22.2. In addition, it should be emphasized that the measure of noise outlined here does not necessarily translate linearly to that expected after transmission to the downstream response regulator, an issue taken up by Mehta and Schwab (2012). Although uncertainty remains as to how cells actually count, these varied formulations illustrate an array of potential mechanisms that may be just as accurate as human decision-making processes.

Questions remain as to the optimal spatial configuration of collections of sensor molecules (Iyengar and Rao 2014). Spatial clustering reduces the sampling error in the vicinity of the array, whereas the spreading of receptors across the cell surface improves average sensing in spatially variable environments. Even in an environment that is spatially uniform on the scale of the length of the cell, sensor aggregation can enhance information transfer into the cellular interior if cooperative interactions exist among adjacent receptor molecules, as appears to be the case for bacterial sensor arrays (Briegel et al. 2009). Although information on the key microanatomical features remains to be determined, with  $n$  effectively independent receptors, the

denominator of Equation 22.3 would just be multiplied by  $\sqrt{n}$ . With cooperation among receptors, the denominator needs to be multiplied by  $\sqrt{n^x}$  with  $x > 1$ .

There remain many unanswered evolutionary questions in this area. Relative to the situation with the nervous systems of metazoans, how much of the energy budgets of single-celled organisms is devoted to environmental monitoring and decision making? What allocation of resources to environmental sensors optimally balances the costs of production of such molecules and the advantages accrued? Confronted with increasing levels of environmental variation, at what point does sensory overload and the energetic cost of running a chemoreception system offset the advantages of tracking environmental changes? What role does the timescale of environmental fluctuations (e.g., within vs. between generations) play in these processes? Evolutionary theory relevant to these questions can be found in Lynch and Gabriel (1987), Lan et al. (2012), and Govern and ten Wolde (2014), but to be of full use, this work will need to be integrated with the known cell biological features of chemosensory systems.

### Phenotypic Bimodality and Bet Hedging

ST systems provide one means for physiological acclimation within the lifespan of an individual cell. An alternative mechanism for dealing with environmental stochasticity is to generate phenotypic diversity independent of current environmental information. Here, the focus is not on genetic polymorphism, but on the production of variable offspring by individual genotypes. This second option provides a potential advantage in that a segment of a clonal population may be immediately poised to deal with an environmental shift, but this comes at the cost of being maladapted at other times. Moreover, the possibility exists that all phenotypes will be suboptimal on some occasions. If phenotypic diversification among clonal progeny is to be promoted by natural selection, it must increase the long-term genotypic growth rate relative to that in other clones.

Although phenotypic distributions of many biological traits are continuous in form (Lynch and Walsh 1998), striking cases of discrete bimodal states are known in microbes, e.g., dispersing vs. sedentary states, vegetative reproduction vs. spore formation, and activation vs. silencing of metabolic pathways. In some situations, different lineages of genetically identical cells can become trapped in alternative states for indefinite periods, even after the initiating environmental signal has dissipated. Such a condition is known as hysteresis.

There are a number of ways by which such phenotypic switching can be modulated by the types of ST pathway architecture noted above. Consider the case of ultrasensitivity illustrated in Figure 22.5. Should cells straddle the threshold point, either because of stochastic internal cellular variation (e.g., molecular inheritance and/or transcriptional noise) or external environmental variation in ligand concentrations, individuals would receive entirely different messages for downstream phenotypic modification. In other words, adding noise to the otherwise deterministic model noted above will lead to some level of phenotypic switching. In this case, the duration of dwell times in alternative states would depend on the magnitude of

fluctuations and the degree to which they are sustained.

Sustained bimodality can arise when there is positive feedback between the activated intermediate substrate and its activation enzyme, as shown graphically in Figure 22.8. In the absence of feedback, the rate of activation of the intermediate substrate declines smoothly to 0.0 as the fraction of active intermediate enzyme ( $I^*$ ) approaches 1.0. In this case, with the rate of deactivation increasing smoothly with increasing  $I^*$ , a single intersection between these opposing functions leads to a single stable steady-state point for  $I^*$  (Figure 22.8a). When  $I^*$  exceeds this point, the rate of deactivation exceeds that of activation, and  $I^*$  declines, and vice versa if the starting point is below the equilibrium steady-state.

If, however, there is positive feedback between  $I^*$  and the activating enzyme, the form of the activation function can be altered in such a way that there are up to three alternative equilibria for  $I^*$  (Figure 22.8b). In this case, when  $I^*$  is sufficiently low, increasing  $I^*$  further accelerates its own production by positive feedback, but eventually the rate of activation must decline (as in the case of no feedback, as a natural consequence of the reduction in inactivated substrate). If the inflection in the activation function leads to three intersections with the deactivation function, the intermediate equilibrium will be unstable, with deviations in either direction resulting in movement towards the flanking equilibria, both of which are stable. Thus, depending on their starting states, cells will gravitate towards one or the other alternative stable states, and remain there until fluctuations in the internal and/or external environments shift  $I^*$  into an alternative basin of attraction. Such a system is said to exhibit bistability.

Assuming that the different equilibrium levels of  $I^*$  are sufficient to lead to altered downstream patterns of gene expression, bistability of an underlying ST system provides a basis for eliciting discrete differences in phenotypic states of otherwise genetically identical cells. The relative frequencies of alternative states will depend on the underlying enzyme kinetics of the system and the magnitude and frequency of fluctuations in the governing parameters.

Bistability can arise under a number of other scenarios, including those involving inhibition, provided the number of inhibitory steps is even. If, for example,  $I_a$  inhibits the reverse enzyme, while  $I_i$  inhibits the forward enzyme (Figure 22.4), a situation can arise in which one of the modifying enzymes, but not both simultaneously, can be common. In addition, many bacterial ST systems exhibit autoregulation, with the response regulator (RR) activating the transcription of the operon containing it (Gao and Stock 2013). This too can generate bistability – when the RR level is high, RR gene expression remains high because of the positive feedback loop; but when the RR level is low, the concentration remains at the basal level of expression (Ferrell 2002; Igoshin et al. 2008; Hermsen et al. 2011; Ram and Goulian 2013).

**Adaptive fine-tuning vs. inadvertent by-products of pathway structure.** Taken together, these theoretical results indicate that without any direct selection at all, the basic structure of ST systems endows genotypically uniform populations of cells with a capacity to develop bistable phenotypic polymorphism. As described below, this is reflected in a number of dramatic dimorphisms in cell morphology and/or behavior. However, less visible, molecular-level shifts may also commonly arise

from the structural underpinnings of ST systems.

This being said, while being conducive to bistability, feedback-containing networks do not guarantee it (Cherry and Adler 2000; Angeli et al. 2004; Nichol et al. 2016). The conditions for existence of dual equilibria and the relative sizes of the basins of attraction for the alternative stable points (when they do exist) are sensitive to the underlying kinetic parameters of the pathway constituents. For example, changes in the elevation of the activation/deactivation response curve in Figure 22.8b can lead to there being just a single intersection, implying monostability. This then suggests the feasibility of the evolutionary fine-tuning of underlying features of ST systems via the selection of appropriate mutations to favor phenotypic switching vs. uniformity. Indeed, because bistability may often be deleterious, with one or both cellular states far from optimal, selection may often operate to move the key kinetic parameters of ST systems to levels that minimize the chances of phase shifting (Hermsen et al. 2011).

A striking example of the types of processes that govern the dynamics of bistability is known for the bacterium *Bacillus subtilis*, which stochastically switches between a motile single-celled state and a nonmotile, chained state (Norman et al. 2013). In this system, switches to alternative states are governed by a double-negative feedback loop. One protein confers the motile, colonizing state, and the other promotes biofilm formation, but each protein suppresses the expression of the other. The molecular details can be found in the original paper. The focus here is simply on the statistical properties of the phase-shifting processes, which are quite different in the two directions.

The *B. subtilis* motile state is memory-less, in the sense that once initiated, there is a constant probability of switching to the chained state at each subsequent cell division. Letting this probability be  $p$ , the probability of switching after the first cell division is  $p$ , after the second division is  $(1-p)p$ , and after the  $n$ th division is  $(1-p)^n p$ . This is an exponential (roughly L-shaped) distribution, with the mean (and standard deviation) of the number of generations to switching both equal to  $1/p$ , in this case  $\sim 81$  cell divisions. Mechanistically, the stochastic switching appears to be due to rare random fluctuations of the concentrations of proteins underlying the double-negative feedback loop, allowing the previously silenced proteins to escape suppression. Notably, although the distribution of switching times suggests randomness at the population level, the behavior of closely related cells (e.g., sister cells) is correlated, owing to shared effects inherited from the maternal cell (Kaufmann et al. 2007; see Chapter 9).

The distribution of switching times for *B. subtilis* chains is quite different, being approximately normal (bell-shaped), with a mean of  $\sim 8$  cell divisions, implying a tight degree of regulation. However, the underlying molecular mechanism for chain termination is again fairly simple – upon chain initiation, there is a substantial pulse of production of material involved in cell adhesion, which is then diluted over subsequent divisions until a minimum threshold is reached. Taken together, these results illustrate remarkably simple molecular paths to dramatic developmental changes – stochastic molecular switches influencing alternative master regulators for flagellar production and cell adhesion.

One of the key evolutionary questions concerning the dynamics of switching behavior concerns the relative longevities of the alternative phase states and the de-

gree to which natural selection molds them in relationship to the scale of temporal variation in environmental conditions favoring the alternative states. Not surprisingly, inspired by the frequent assumption that all aspects of biodiversity must be a product of natural selection, a considerable amount of attention has been given to the idea that bistability is an adaptively evolved “bet-hedging” strategy enabling individual genotypes to maximize their long-term fitness without resorting to potentially costly mechanisms for short-term physiological acclimation (Kussell and Leibler 2005; Smits et al. 2006; Veening et al. 2008; Wei et al. 2014; Norman et al. 2015). For reasons of tractability, most of the theoretical work has focused on simple systems with just two discrete environments, each lasting for a time period in excess of cell generation lengths, and with two discrete phenotypes, each better adapted to an alternative environment.

In a temporally variable environment, the genotype with the highest long-term exponential growth rate will be favored, and for the simple two-state / two-environment model, the optimal average random phenotypic switching time is equal to the average of the periods between shifts between the two environmental states, provided there are equally large (but opposite in sign) selective pressures on the two alternative phenotypes in the alternative environments (Thattai and van Oudenaarden 2004; Kussell et al. 2005; Salathé et al. 2009; Gaál et al. 2010). On the other hand, if the selection differential between the two environments is sufficiently large, stochastic switching can be disfavored, as the monomorphic genotype favored in the environment with large effects can overwhelm the smaller, short-lived disadvantage in the opposite environment.

Some aspects of this model have been explored with an experimental evaluation of a system involving two alternative growth phenotypes in the yeast *S. cerevisiae* (Acar et al. 2008). Here, strains were engineered to switch between two physiological states that yielded different growth rates in two alternative environments. The strain that rapidly shifted from one state to the other experienced an early advantage whenever the environment shifted, as it quickly produced progeny adapted in the changed environment. However, as the duration between environmental shifts was lengthened, the slow-shifting variant gained an advantage, owing to its reduced production of the maladapted type.

Although somewhat contrived, these results show that the long-term advantage of phenotypic switching depends on the temporal dynamics of both environmental change and phenotypic response. However, it is prudent to consider that most environmental variables are continuous in nature, and can vary on both within- and between-generation time scales, as well as across spatial scales, and these kinds of alternative scenarios can lead to rather different expectations (Lynch and Gabriel 1987). Thus, without direct empirical evidence, there is no justification for assuming that all instances of phenotypic polymorphisms reflect adaptive fine-tuning.

Finally, in all of the preceding discussion, it has been assumed that phenotypic switching is an intrinsic feature of a cellular network, occurring without regard to the current environmental state. However, using environmental cues, organisms can, in principle, reinforce various phenotypic outcomes. This might happen, for example, through appropriate epigenetic modifications such as DNA methylation and/or histone modifications if these somehow encouraged individuals expressing a particular phenotype to produce offspring with an elevated frequency of the same

phenotype (Xue and Leibler 2016). Provided that the individuals with inappropriate phenotypes are removed by selection, such a system can then lead to a form of trans-generational acclimation that superficially appears like learning or the inheritance of acquired characteristics. Such transient shifts in mean phenotype without any underlying genetic change will resort back to an alternative phenotype distribution upon environmental change (see Foundations 9.5).

## Summary

- Unicellular organisms respond to external environmental stimuli through the use of signal-transduction (ST) pathways that relay information from the cell surface to intracellular effectors, such as transcription factors. In most species, dozens to hundreds of such systems are specialized to different environmental indicators. In bacteria, such information relay systems often involve just two proteins – an external sensor and a response regulator.
- ST systems are modular in nature, being based on several several small motifs that specify proper communication between sensor and regulator proteins to the exclusion of members of other parallel pathways. The vast majority of such systems operate via additions and removals of phosphoryl groups on the participating proteins.
- Owing to the simplicity of these communication systems, rewiring of ST pathways is readily accomplished by changes in just a few key amino acids. Although this opens up opportunities for the establishment of novel signaling pathways following gene duplication, it also promotes the neutral drift of the recognition vocabulary in the absence of selection for altered functions.
- Despite the fact that the individual proteins driving ST systems operate as conventional Michaelis-Menten enzymes, the pathways through which they operate are often structured in such a way as to potentially generate very steep responses to external ligand concentrations. In some cases, the response approaches switch-like behavior wherein the downstream target is essentially 100% on or off when the ligand concentration is above vs. below the threshold value.
- The continuous operation of opposing phosphorylation/dephosphorylation reactions at the heart of ST systems imposes a substantial energetic cost of processing and transmitting information, even in a constant environment.
- Owing to fluctuations in the arrival and binding of external signals to cell-surface receptors, environmental sensing is also an inherently noisy process. Noise buffer-

ing is facilitated by increasing the numbers of receptors and setting the binding / unbinding kinetics to levels that allow the cell to repeatedly make independent measures of the degree of receptor binding at rates that exceed the internal cellular response.

- Eukaryotic ST systems tend to be much more complicated than those in bacteria, commonly with expansions to cascades of multiple intermediate steps, use of multiple phosphosites per communicating molecule, and kinases and phosphatases cross-talking with multiple interacting partners. It remains unclear whether this complexity enhances the speed, efficiency, or accuracy of environmental assessment, and whether differences between prokaryotic vs. eukaryotic ST systems have been driven by adaptive forces.
- Chemotaxis provides a rapid mechanism for adjusting the direction of bacterial motility in response to environmental gradients. These systems, which vary in structure among species, also often have a simple built-in feedback mechanisms for adjusting sensitivity to the prevailing environmental state, much like the visual systems of metazoans adjust to different light levels.
- The structure of ST systems is such that the addition of positive feedback loops (or pairs of negative feedbacks) can give rise to bistable responses to external ligand concentrations. By this means, genetically uniform populations can generate dimorphic populations of cells, potentially enhancing long-term genotypic fitness in environments presenting certain levels of variation, but also yielding maladaptive responses in inappropriate environments.

**Foundations 22.1. Behavior of a monocycling system.** A key issue with respect to an interconvertible enzyme (I) is the degree of activity expected under various conditions. The total concentration of the enzyme  $[I_T]$  partitions into subsets of active and inactive molecules,  $[I_a]$  and  $[I_i]$ , to a degree that depends on the relative concentrations of the active forms of converter enzymes (denoted F and R, respectively, for forward and reverse reactions). The fractions of active vs. inactive converter enzymes depend in turn on the concentrations of their ligands and their affinities for them. Because the two converter enzymes push the interconvertible enzyme in opposite directions, the relative concentrations of their active forms dictate the level of activity of enzyme I.

Here, we consider the steady-state situation in which the concentrations of both converter enzymes and their ligands are kept constant by ambient cellular conditions. Initially, we further assume that the fractions of both converter enzymes tied up with the interconvertible enzyme are negligible, which requires that the latter not be at a saturating level. Under these conditions (first-order rate kinetics), the active fractions of both converting enzymes will reach steady-state levels independent of the amount of enzyme I, and determined only by the rates of association and dissociation with their ligands. Using the terms defined in Figure 22.4, for the forward enzyme, a steady state requires that the rate of production of the active (a) enzyme from inactive (i) enzyme equals the flux in the opposite direction (resulting from the deactivation of  $F_a$ ),

$$k_{a,F}[F_i][S_F] = k_{d,F}[F_a]. \quad (22.1.1)$$

Noting that the total concentration of forward enzyme in the system is

$$[F_T] = [F_i] + [F_a], \quad (22.1.2)$$

solving these two equations leads to the steady-state concentration of the active forward enzyme

$$[F_a] = \frac{[F_T][S_F]}{k_{D,F} + [S_F]}, \quad (22.1.3a)$$

where  $k_{D,F} = k_{d,F}/k_{a,F}$  is the dissociation constant of enzyme  $F$ . Likewise, the equilibrium concentration of the active form of the reverse enzyme is

$$[R_a] = \frac{[R_T][S_R]}{k_{D,R} + [S_R]}. \quad (22.1.3b)$$

Provided the concentrations of the converter enzymes are at steady state, the alternative forms of the central enzyme I will also attain steady state. This occurs when the rate of production of active from inactive I equals the rate in the opposite direction. Using the familiar Michaelis-Menten formulations (Chapter 19), these forward and reverse reaction rates can be written as

$$V_F = \frac{k_{cat,F}[F_a][I_i]}{k_{S,F} + [I_i]}, \quad (22.1.4a)$$

and

$$V_R = \frac{k_{cat,R}[R_a][I_a]}{k_{S,R} + [I_a]}. \quad (22.1.4b)$$

Letting the total concentration of interconvertible enzyme in the system be

$$[I_T] = [I_i] + [I_a], \quad (22.1.5a)$$

the quantity of interest is the fraction of molecules that are in the active state,

$$I^* = [I_a]/[I_T]. \quad (22.1.5b)$$

The general solution can be obtained by setting Equations 22.1.4a,b equal to each other, letting  $[I_i] = [I_T] - [I_a]$ , and solving for the level of  $[I_a]$  that satisfies the equality.

As mentioned below, the full solution is quite complicated, but as pointed out by Stadtman and Chock (1977) and Shacter-Noiman et al. (1983), provided the total amount of enzyme  $I$  in the system is small relative to the half-saturation constants in Equations 22.1.4a,b (nonsaturating conditions), the concentrations of  $I$  in the denominators of these equations can be ignored, and this leads to an expression of the form

$$I^* = \frac{\beta C}{1 + \beta C}, \quad (22.1.6)$$

where

$$\beta = \frac{\kappa_F [F_T]}{\kappa_R [R_T]} \quad (22.1.7a)$$

is the ratio of kinetic potentials of the forward and reverse converter enzymes, with  $\kappa_x = k_{cat,x}/k_{S,x}$  being the specificity constant of enzyme  $x$  (see Foundations 19.1), and

$$C = \frac{[S_F](k_{D,R} + [S_R])}{[S_R](k_{D,F} + [S_F])} \quad (22.1.7b)$$

is the ratio of degrees of saturation of the input reactions (see Equations 22.1.3a,b). Although Equation 22.1.6 has a simple hyperbolic form, the underlying function is quite complex, as it actually depends on ten different parameters (two each of the  $k_{cat}$ ,  $k_S$ , and  $k_D$  terms, and the concentrations of the two converter enzymes and their input ligands). Equation 22.3 in the main text gives an expression equivalent to Equation 22.1.6 in terms of the active concentrations of the forward and reverse enzymes.

Several significant points are revealed by Equation 22.1.6. First, for this case of low overall concentration of  $I$ ,  $I^*$  is independent of the total concentration  $[I_T]$ . Second, as in the case of simple Michaelis-Menten kinetics,  $I^*$  is a hyperbolic function, in this case of  $C$ . Although the latter is itself a complex function, inspection shows a hyperbolic relationship with either ligand concentration. Third, although  $I^* \rightarrow 1$  as  $(\beta C) \rightarrow \infty$ , because the relative concentrations of active enzymes are limited by the properties of the system (the total enzyme concentrations, total ligand concentrations, and the dissociation constants), there is an upper bound to  $\beta C$ . Thus, the maximum fractional activity of the central enzyme is generally  $< 1.0$  (Figure 22.5, upper panel).

Finally, a more general expression allowing for any concentration of  $I$  was derived by Goldbeter and Koshland (1981). In this case, the solution does depend on  $[I_T]$ , bringing the total number of relevant parameters to eleven, but can be written as a function of three composite parameters,

$$I^* = \frac{(\alpha - 1) - (k_F^* + \alpha k_R^*) + \sqrt{[(\alpha - 1) - (k_F^* + \alpha k_R^*)]^2 + 4\alpha(\alpha - 1)k_R^*}}{2\alpha}, \quad (22.1.8)$$

where  $\alpha = (k_{cat,F}[F_a])/(k_{cat,R}[R_a])$ ,  $k_F^* = k_{S,F}/[I_T]$ ,  $k_R^* = k_{S,R}/[I_T]$ , and  $[F_a]$  and  $[R_a]$  are defined by Equations 22.1.3a,b. Contrary to the limiting situation in which  $[I_T]$  is low, the relationship of  $I^*$  is no longer a simple hyperbola, as discussed in the main text (Figure 22.5, lower panel).

Additional types of systems, including those with inhibitor interactions and with linked (multicyclic) cycles, are explored in Chock and Stadtman (1977), Stadtman and Chock (1977), and Goldbeter and Koshland (1984). Not surprisingly, these exhibit

even richer behavior than those noted above. For the case of bacterial two-component systems, where the kinase often has a dual function as the phosphatase (the reverse converter enzyme in the above scheme), expressions similar in form to Equation 22.1.8 have been developed by Batchelor and Goulian (2003) and Rowland and Deeds (2014). An excellent overview of all of these models, and the logic underlying them, is provided by Qian (2007). Malaguti and ten Wolde (2021) extend things to time-varying signal concentrations.

---

**Foundations 22.2. Accuracy of environmental sensing.** A successful sensing system requires that the receptors be capable of assaying the current environmental state accurately enough to transmit a reliable signal to the downstream responders essential to elicit appropriate changes in cell behavior. Here, we consider the degree to which a single molecular receptor at the cell surface can assess the concentration of a ligand in the surrounding environment. The initial assumption is that the fractional time during which the receptor is bound to the external ligand ( $p$ ) provides the best information that the cell can utilize for environmental prediction. As will be seen below, however, this is by no means the only possible approach to the problem.

At any single point in time, the receptor is either occupied or not, so a single snapshot assessment provides little information. Over time, however, as ligand molecules become unbound, the receptor can make repeated assays of the environment, so that the average occupancy during a particular period becomes an estimate of  $p$ . As a measure of accuracy, we utilize the coefficient of variation (CV), which equals the ratio of the standard deviation ( $\sigma_x$ ) to the mean ( $\mu_x$ ) of repeated measures of a variable. Because it is easier to work with measures of variance, which is the square of the standard deviation, the derivations to follow will be based on the squared coefficient of variation,  $\sigma_x^2/\mu_x^2$ .

It is assumed here that on the time scale of environmental assessment, the cell resides in a homogeneous environment with constant ligand concentration ( $c_0$ ), so we are obtaining a pure measurement of environmental sensitivity based on the properties of the receptor molecule and its ligand molecule. If the environment is variable within the time-frame of environmental assessment, the variance of ligand concentration would need to be incorporated into the measure of noise derived below. The following derivations are based on the first presentation of the problem by Berg and Purcell (1977) and subsequent refinements by Kaizu et al. (2014). Some uncertainties about the precise nature of the final formulation are addressed in an excellent overview by Aquino et al. (2016).

To describe the temporal behavior of receptor occupancy, consider the stochastic differential equation

$$\frac{d\Gamma_t}{dt} = k_{\text{on}}c_0(1 - \Gamma_t) - k_{\text{off}}\Gamma_t + \epsilon_t, \quad (22.2.1)$$

where  $\Gamma_t$  denotes the occupancy (0 or 1) of a single receptor at time point  $t$ ,  $k_{\text{on}}$  is the rate of ligand binding to an unbound receptor (per unit of external concentration,  $c_0$ ),  $k_{\text{off}}$  is the rate of dissociation of a ligand molecule from a bound receptor, and  $\epsilon_t$  is a stochastic variable with mean zero. By setting the derivative to zero and solving, the equilibrium probability of occupancy is found to be

$$\bar{\Gamma} = p = \frac{k_{\text{on}}c_0}{k_{\text{on}}c_0 + k_{\text{off}}} = \frac{c_0}{c_0 + k_{\text{D}}}, \quad (22.2.2)$$

where  $k_{\text{D}} = k_{\text{off}}/k_{\text{on}}$ .

Although the cell perceives the environment through the act of ligand binding, the ultimate goal of environmental sensing is to obtain an estimate of the ligand

concentration ( $c$ ) that closely approximates the true concentration ( $c_0$ ). This requires an estimate of the variance among sample estimates of  $c_0$  inferred from the cell's readout  $\Gamma$ . To obtain this, we start with a general rule from statistics that the variance of a dependent variable is equal to the variance of a causal variable times the squared derivative of the first with respect to the second (Lynch and Walsh 1998, Appendix A), which in this case implies,

$$\sigma_{\Gamma}^2 = (\partial p / \partial c)^2 \cdot \sigma_c^2. \quad (22.2.3)$$

Rearranging and dividing by  $c_0^2$  yields our desired measure of accuracy, the squared coefficient of variation of inferred concentration,

$$\frac{\sigma_c^2}{c_0^2} = \frac{1}{c_0^2} \cdot \frac{\sigma_{\Gamma}^2}{(\partial p / \partial c)^2}. \quad (22.2.4)$$

From Equation 22.2.2, the partial derivative evaluated at  $c_0$  is

$$\frac{\partial p}{\partial c} = \frac{k_D}{(c_0 + k_D)^2}, \quad (22.2.5)$$

and substitution into Equation 22.2.4, after some rearrangement, leads to

$$\frac{\sigma_c^2}{c_0^2} = \frac{c_0^2}{p^4 k_D^2} \cdot \sigma_{\Gamma}^2. \quad (22.2.6)$$

The final step requires an expression for the variance in the mean occupancy  $\sigma_{\Gamma}^2$  over some period of time  $T$  of continuous assessment, as this ultimately the degree of accuracy of overall environmental assessment. This is complicated by the fact that the realized  $\Gamma$  at any one particular time is not independent of that in adjacent time periods, owing to the time spans between ligand binding and release. Taking these autocorrelations into consideration, Berg and Purcell (1977) showed that

$$\sigma_{\Gamma}^2 = \frac{2p(1-p)^2}{Tk_{\text{off}}}. \quad (22.2.7a)$$

Noting from Equation 22.2.2 that

$$k_{\text{on}}c_0(1-p) = k_{\text{off}}p,$$

Equation 22.2.7a can be equivalently written as

$$\sigma_{\Gamma}^2 = \frac{2p^2(1-p)}{Tk_{\text{on}}}c_0. \quad (22.2.7b)$$

Finally, substituting Equation 22.2.7b into 22.2.6, again with some downstream rearrangement, leads to a remarkably simple expression

$$\frac{\sigma_c^2}{c_0^2} = \frac{2}{pTk_{\text{off}}} = \frac{2}{(1-p)Tk_{\text{on}}c_0}. \quad (22.2.8)$$

The accuracy of assessment increases (i.e.,  $\sigma_c^2/c_0^2$  decreases) with increasing time over which the cell integrates environmental information, and also with increasing  $k_{\text{off}}$ . The latter feature arises because the inverse of  $k_{\text{off}}$  is equal to the average release time of ligands, which means that higher  $k_{\text{off}}$  allows the receptor to make more evaluations of the environment.

There are at least two other ways to express the accuracy. First, the average time between consecutive ligand-binding events is equal to the sum of the mean times for the length of binding to an occupied receptor and that of the time for an unoccupied receptor to accept another ligand, each of which is the reciprocal of the respective rate,

$$\tau_b = \frac{1}{k_{\text{off}}} + \frac{1}{k_{\text{on}}c_0}. \quad (22.2.9)$$

Noting that the mean number of expected bindings in interval  $T$  is  $\bar{N} = T/\tau_b$ , substitution of  $T = \bar{N}\tau_b$  and Equation 22.2.2 into 22.2.8 leads to

$$\frac{\sigma_c^2}{c_0^2} = \frac{2}{\bar{N}}, \quad (22.2.10)$$

showing that the squared CV of the cell's estimate of  $c_0$  is inversely proportional to the expected number of molecules bound during the assessment period (which itself is a function of time and ligand concentration).

Second,  $k_{\text{on}}$  is the inverse of the mean time to binding of an unoccupied receptor (per unit ligand concentration), which in turn is equal to the sum of expected times for particles to diffuse to the receptor ( $k_e$ ) and of binding upon contact ( $k_+$ ),

$$k_{\text{on}} = \left( \frac{1}{k_e} + \frac{1}{k_+} \right)^{-1} = \frac{k_e k_+}{k_e + k_+}. \quad (22.2.11)$$

Substituting Equation 22.2.11 into 22.2.8 yields

$$\frac{\sigma_c^2}{c_0^2} = \frac{2}{Tc_0(1-p)} \left( \frac{1}{k_e} + \frac{1}{k_+} \right). \quad (22.2.12)$$

Assuming that the receptor binding site can be approximated as disc of radius  $r_s$  on the cell surface, the encounter rate (per unit concentration) by diffusion is

$$k_e = 4Dr_s, \quad (22.2.13)$$

where  $D$  is the diffusion constant for the ligand. Berg and Purcell (1977) assumed the case of diffusion limitation, such that  $k_e \ll k_+$ , which reduces Equation 22.2.12 to

$$\frac{\sigma_c^2}{c_0^2} = \frac{1}{2Dr_s c_0 (1-p)T}. \quad (22.2.14)$$

To gain more quantitative insight into the accuracy of monitoring as inferred by Equation 22.2.14, let  $D = 10^{-5} \text{ cm}^2/\text{sec}$ , which closely approximates true values for single amino acids (with cations and anions having values only  $\sim 2\times$  higher; Chapter 7). Estimates of  $r_s$  for chemoreceptors are sparse, but can be inferred to be on the order of 2 nm ( $= 2 \times 10^{-7} \text{ cm}$ ) given that the total area of receptor arrays in a wide range of bacteria implies an area/receptor of  $\sim 100 \text{ nm}^2$  with most of the array space being empty (Briegleb et al. 2009). Supposing the dissociation constant  $k_D = k_{\text{off}}/k_{\text{on}}$  is such that  $p \simeq 0.5$  (Equation 22.2.2), and a ligand concentration of  $c_0 = 1 \mu\text{M} = 6 \times 10^{14} \text{ molecules/cm}^3$ , the coefficient of variation of measurement then becomes

$$\frac{\sigma_c}{c_0} \simeq \sqrt{\frac{1}{1000T}}$$

where the units of  $T$  are in seconds. Thus, monitoring a constant environment for just ten seconds is sufficient to reduce the level of estimation noise to 0.01 (i.e., a standard

deviation of the inferred concentration just 1% of the true value). Assuming the same  $p$ , with a 1000-fold lower concentration ( $c_0 = 1$  nM), the level of uncertainty will be increased by a factor of  $\sqrt{1000} \simeq 32$ , and to achieve a level of accuracy of 0.01,  $T$  has to be 1000-fold higher.

A broad overview of the biophysical constraints associated with other modes of environmental sensing (e.g., mechanoreception, vision, and hearing, all of which are exploited by metazoans) is given in Martens et al. (2015), who demonstrate that intrinsic limits associated with the scale of environmental noise and signal transmission restrict the utility of mechanosensing in an open-water environment to eukaryote-sized cells.

---

**Literature Cited**

- Abedrabbo, S., J. Castellon, K. D. Collins, K. S. Johnson, and K. M. Ottemann. 2017. Cooperation of two distinct coupling proteins creates chemosensory network connections. *Proc. Natl. Acad. Sci. USA* 114: 2970-2975.
- Acar, M., J. T. Mettetal, and A. van Oudenaarden. 2008. Stochastic switching as a survival strategy in fluctuating environments. *Nature Genetics* 40: 471-475.
- Adler, J. 1966. Chemotaxis in bacteria. *Science* 153: 708-716.
- Alm, E., K. Huang, and A. Arkin. 2006. The evolution of two-component systems in bacteria reveals different strategies for niche adaptation. *PLoS Comput. Biol.* 2: e143.
- Anantharaman, V., L. M. Iyer, and L. Aravind. 2007. Comparative genomics of protists: new insights into the evolution of eukaryotic signal transduction and gene regulation. *Ann. Rev. Microbiol.* 61: 453-475.
- Angeli, D., J. E. Ferrell, Jr., and E. D. Sontag. 2004. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proc. Natl. Acad. Sci. USA* 101: 1822-1827.
- Aquino, G., N. S. Wingreen, and R. G. Endres. 2016. Know the single-receptor sensing limit? Think again. *J. Stat. Phys.* 162: 1353-1364.
- Armbruster, D., J. Nagy, and J. Young. 2014. Three level signal transduction cascades lead to reliably timed switches. *J. Theor. Biol.* 361: 69-80.
- Ashenberg, O., K. Rozen-Gagnon, M. T. Laub, and A. E. Keating. 2011. Determinants of homodimerization specificity in histidine kinases. *J. Mol. Biol.* 413: 222-235.
- Ballal, A., R. Heermann, K. Jung, M. Gassel, K. Apte, and K. Altendorf. 2002. A chimeric *Anabaena* / *Escherichia coli* KdpD protein (Anacoli KdpD) functionally interacts with *E. coli* KdpE and activates kdp expression in *E. coli*. *Arch. Microbiol.* 178: 141-148.
- Barkai, N., and S. Leibler. 1997. Robustness in simple biochemical networks. *Nature* 387: 913-917.
- Batchelor, E., and M. Goulian. 2003. Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system. *Proc. Natl. Acad. Sci. USA* 100: 691-696.
- Berg, H. C., and E. M. Purcell. 1977. Physics of chemoreception. *Biophys. J.* 20: 193-219.
- Bitbol, A. F., and N. S. Wingreen. 2015. Fundamental constraints on the abundances of chemotaxis proteins. *Biophys. J.* 108: 1293-1305.
- Blüthgen, N. 2006. Sequestration shapes the response of signal transduction cascades. *IUBMB Life* 58: 659-663.
- Briegel, A., D. R. Ortega, E. I. Tocheva, K. Wuichet, Z. Li, S. Chen, A. Müller, C. V. Iancu, G. E. Murphy, M. J. Dobro, et al. 2009. Universal architecture of bacterial chemoreceptor arrays. *Proc. Natl. Acad. Sci. USA* 106: 17181-17186.
- Briegel, A., D. R. Ortega, A. N. Huang, C. M. Oikonomou, R. P. Gunsalus, and G. J. Jensen. 2015. Structural conservation of chemotaxis machinery across Archaea and Bacteria. *Environ. Microbiol. Rep.* 7: 414-419.
- Burger, L., and E. van Nimwegen. 2008. Accurate prediction of protein-protein interactions from sequence alignments using a Bayesian method. *Mol. Syst. Biol.* 4: 165.

- Capra, E. J., and M. T. Laub. 2012. Evolution of two-component signal transduction systems. *Annu. Rev. Microbiol.* 66: 325-347.
- Capra, E. J., B. S. Perchuk, O. Ashenberg, C. A. Seid, H. R. Snow, J. M. Skerker, and M. T. Laub. 2012a. Spatial tethering of kinases to their substrates relaxes evolutionary constraints on specificity. *Mol. Microbiol.* 86: 1393-1403.
- Capra, E. J., B. S. Perchuk, E. A. Lubin, O. Ashenberg, J. M. Skerker, and M. T. Laub. 2010. Systematic dissection and trajectory-scanning mutagenesis of the molecular interface that ensures specificity of two-component signaling pathways. *PLoS Genet.* 6: e1001220.
- Capra, E. J., B. S. Perchuk, J. M. Skerker, and M. T. Laub. 2012b. Adaptive mutations that prevent crosstalk enable the expansion of paralogous signaling protein families. *Cell* 150: 222-232.
- Cherry, J. L., and F. R. Adler. 2000. How to make a biological switch. *J. Theor. Biol.* 203: 117-133.
- Chock, P. B., and E. R. Stadtman. 1977. Superiority of interconvertible enzyme cascades in metabolite regulation: analysis of multicyclic systems. *Proc. Natl. Acad. Sci. USA* 74: 2766-2770.
- Cock, P. J., and D. E. Whitworth. 2007. Evolution of prokaryotic two-component system signaling pathways: gene fusions and fissions. *Mol. Biol. Evol.* 24: 2355-2357.
- Colin, R., and V. Sourjik. 2017. Emergent properties of bacterial chemotaxis pathway. *Curr. Opin. Microbiol.* 39: 24-33.
- Cremer, J., T. Honda, Y. Tang, J. Wong-Ng, M. Vergassola, and T. Hwa. 2019. Chemotaxis as a navigation strategy to boost range expansion. *Nature* 575: 658-663.
- Endres, R. G., and N. S. Wingreen. 2008. Accuracy of direct gradient sensing by single cells. *Proc. Natl. Acad. Sci. USA* 105: 15749-15754.
- Endres, R. G., and N. S. Wingreen. 2009. Maximum likelihood and the single receptor. *Phys. Rev. Lett.* 103: 158101.
- Ferrell, J. E., Jr. 2002. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr. Opin. Cell Biol.* 14: 140-148.
- Gaál, B., J. W. Pitchford, and A. J. Wood. 2010. Exact results for the evolution of stochastic switching in variable asymmetric environments. *Genetics* 184: 1113-1119.
- Gao, R., and A. M. Stock. 2013. Evolutionary tuning of protein expression levels of a positively autoregulated two-component system. *PLoS Genet.* 9: e1003927.
- Gildor, T., R. Shemer, A. Atir-Lande, and D. Kornitzer. 2005. Coevolution of cyclin Pcl5 and its substrate Gen4. *Eukaryot. Cell* 4: 310-318.
- Goldbeter, A., and D. E. Koshland, Jr. 1981. An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci. USA* 78: 6840-6844.
- Goldbeter, A., and D. E. Koshland, Jr. 1984. Ultrasensitivity in biochemical systems controlled by covalent modification: interplay between zero-order and multistep effects. *J. Biol. Chem.* 259: 14441-14447.
- Goldbeter, A., and D. E. Koshland, Jr. 1987. Energy expenditure in the control of biochemical

- systems by covalent modification. *J. Biol. Chem.* 262: 4460-4471.
- Goulian, M. 2010. Two-component signaling circuit structure and properties. *Curr. Opin. Microbiol.* 13: 184-189.
- Govern, C. C., and P. R. ten Wolde. 2014. Optimal resource allocation in cellular sensing systems. *Proc. Natl. Acad. Sci. USA* 111: 17486-17491.
- Gunawardena, J. 2005. Multisite protein phosphorylation makes a good threshold but can be a poor switch. *Proc. Natl. Acad. Sci. USA* 102: 14617-14622.
- Hermesen, R., D. W. Erickson, and T. Hwa. 2011. Speed, sensitivity, and bistability in auto-activating signaling circuits. *PLoS Comput. Biol.* 7: e1002265.
- Igoshin, O. A., R. Alves, and M. A. Savageau. 2008. Hysteretic and graded responses in bacterial two-component signal transduction. *Mol. Microbiol.* 68: 1196-1215.
- Iyengar, G., M. and Rao. 2014. A cellular solution to an information-processing problem. *Proc. Natl. Acad. Sci. USA* 111: 12402-1247.
- Kaizu, K., W. de Ronde, J. Paijmans, K. Takahashi, F. Tostevin, and P. R. ten Wolde. 2014. The Berg-Purcell limit revisited. *Biophys. J.* 106: 976-985.
- Kapp, G. T., S. Liu, A. Stein, D. T. Wong, A. Reményi, B. J. Yeh, J. S. Fraser, J. Taunton, W. A. Lim, and T. Kortemme. 2012. Control of protein signaling using a computationally designed GTPase/GEF orthogonal pair. *Proc. Natl. Acad. Sci. USA* 109: 5277-5282.
- Kaufmann, B. B., Q. Yang, J. T. Mettetal, and A. van Oudenaarden. 2007. Heritable stochastic switching revealed by single-cell genealogy. *PLoS Biol.* 5: e239.
- Kempes, C. P., D. Wolpert, Z. Cohen, and J. Pérez-Mercader. 2017. The thermodynamic efficiency of computations made in cells across the range of life. *Philos. Trans. A Math. Phys. Eng. Sci.* 375: 20160343.
- Keymer, J. E., R. G. Endres, M. Skoge, Y. Meir, and N. S. Wingreen. 2006. Chemosensing in *Escherichia coli*: two regimes of two-state receptors. *Proc. Natl. Acad. Sci. USA* 103: 1786-1791.
- Koretke, K. K., A. N. Lupas, P. V. Warren, M. Rosenberg, and J. R. Brown. 2000. Evolution of two-component signal transduction. *Mol. Biol. Evol.* 17: 1956-1970.
- Kussell, E., R. Kishony, N. Q. Balaban, and S. Leibler. 2005. Bacterial persistence: a model of survival in changing environments. *Genetics* 169: 1807-1814.
- Kussell, E., and S. Leibler. 2005. Phenotypic diversity, population growth, and information in fluctuating environments. *Science* 309: 2075-2078.
- Lan, G., P. Sartori, S. Neumann, V. Sourjik, and Y. Tu. 2012. The energy-speed-accuracy tradeoff in sensory adaptation. *Nat. Physics* 8: 422-428.
- Landauer, R. 1988. Dissipation and noise immunity in computation and communication. *Nature* 335: 779-784.
- Laub, M. T., and M. Goulian. 2007. Specificity in two-component signal transduction pathways. *Annu. Rev. Genet.* 41: 121-145.
- Levy, W. B., and V. G. Calvert. 2021. Communication consumes 35 times more energy than computation in the human cortex, but both costs are needed to predict synapse number. *Proc. Natl. Acad. Sci. USA* 118: e2008173118.

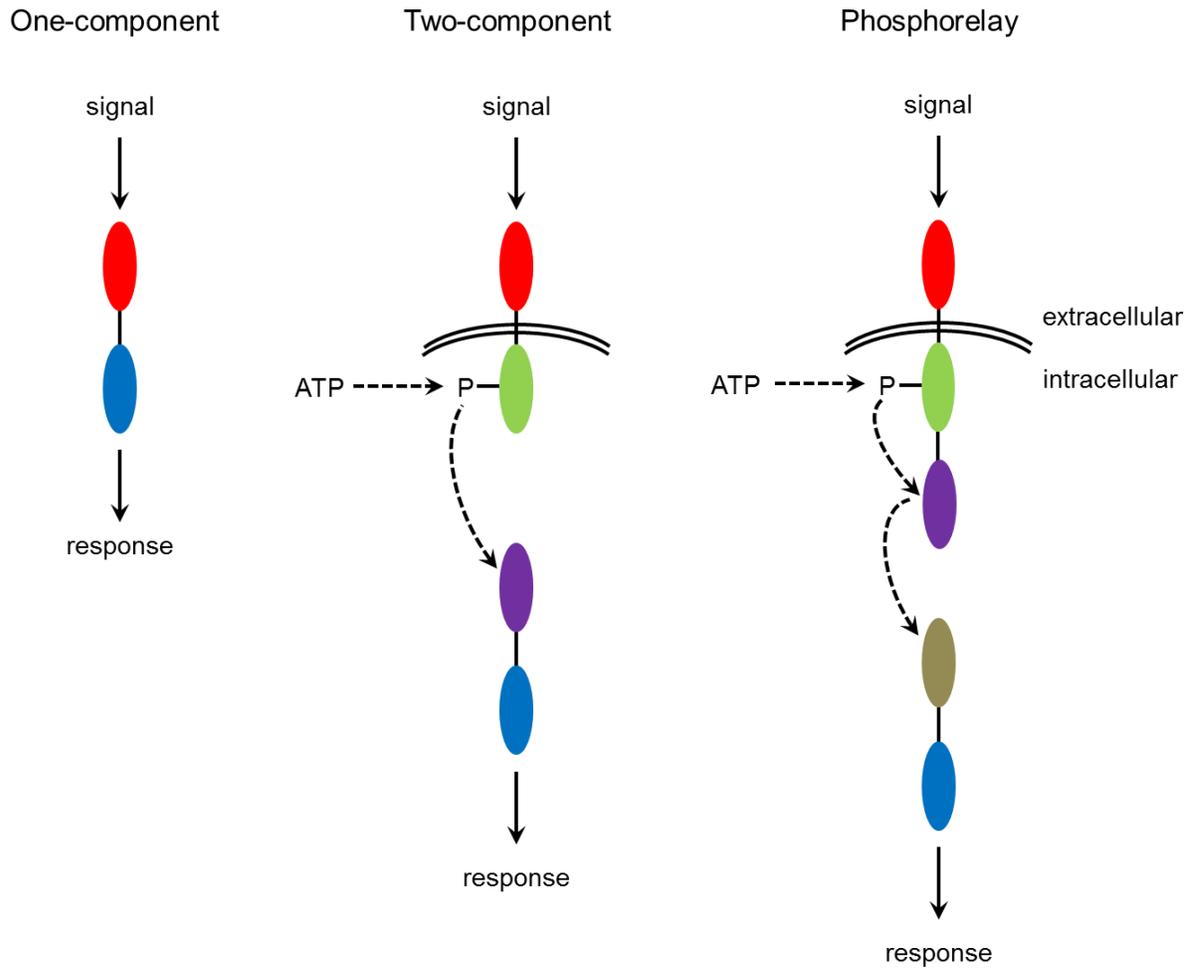
- Li, L., E. I. Shakhnovich, and L. A. Mirny. 2003. Amino acids determining enzyme-substrate specificity in prokaryotic and eukaryotic protein kinases. *Proc. Natl. Acad. Sci. USA* 100: 4463-4468.
- Lim, W., B. Mayer, and T. Pawson. 2014. *Cell Signaling*. Garland Science, New York, NY.
- Lloyd, D., C. A. Phillips, and M. Statham. 1978. Oscillations of respiration, adenine nucleotide levels and heat evolution in synchronous cultures of *Tetrahymena pyriformis* ST prepared by continuous-flow selection. *Microbiology* 106: 19-26.
- Loomis, W. F., G. Shaulsky, and N. Wang. 1997. Histidine kinases in signal transduction pathways of eukaryotes. *J. Cell Sci.* 110: 1141-1145.
- Lynch, M., and W. Gabriel. 1987. Environmental tolerance. *Amer. Natur.* 129: 283-303.
- Lynch, M., and K. Hagner. 2015. Evolutionary meandering of intermolecular interactions along the drift barrier. *Proc. Natl. Acad. Sci. USA* 112: E30-E38.
- Lynch, M., and J. B. Walsh. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Assocs., Inc., Sunderland, MA.
- Makarova, K. S., M. Y. Galperin, and E. V. Koonin. 2017. Proposed role for KaiC-like ATPases as major signal transduction hubs in Archaea. *mBio* 8: e01959-17.
- Malaguti, G., and P. R. ten Wolde. 2021. Theory for the optimal detection of time-varying signals in cellular sensing systems. *eLife* 10: e62574.
- Martens, E. A., N. Wadhwa, N. S. Jacobsen, C. Lindemann, K. H. Andersen, and A. Visser. 2015. Size structures sensory hierarchy in ocean life. *Proc. Biol. Sci.* 282: 20151346.
- McClune, C. J., A. Alvarez-Buylla, C. A. Voigt, and M. T. Laub. 2019. Engineering orthogonal signalling pathways reveals the sparse occupancy of sequence space. *Nature* 574: 702-706.
- Mehta, P., and D. J. Schwab. 2012. Energetic costs of cellular computation. *Proc. Natl. Acad. Sci. USA* 109: 17978-17982.
- Mello, B. A., L. Shaw, and Tu Y. 2004. Effects of receptor interaction in bacterial chemotaxis. *Biophys. J.* 87: 1578-1595.
- Mello, B. A., and Y. Tu. 2007. Effects of adaptation in maintaining high sensitivity over a wide range of backgrounds for *Escherichia coli* chemotaxis. *Biophys. J.* 92: 2329-2337.
- Ni, B., R. Colin, H. Link, R. G. Endres, and V. Sourjik. 2020. Growth-rate dependent resource investment in bacterial motile behavior quantitatively follows potential benefit of chemotaxis. *Proc. Natl. Acad. Sci. USA* 117: 595-601.
- Nichol, D., M. Robertson-Tessi, P. Jeavons, and A. R. Anderson. 2016. Stochasticity in the genotype-phenotype map: implications for the robustness and persistence of bet-hedging. *Genetics* 204: 1523-1539.
- Niven, J. E. 2016. Neuronal energy consumption: biophysics, efficiency and evolution. *Curr. Opin. Neurobiol.* 41: 129-135.
- Nocedal, I., and M. T. Laub. 2022. Ancestral reconstruction of duplicated signaling proteins reveals the evolution of signaling specificity. *eLife* 11: e77346.
- Norman, T. M., N. D. Lord, J. Paulsson, and R. Losick. 2013. Memory and modularity in cell-fate decision making. *Nature* 503: 481-486.

- Norman, T. M., N. D. Lord, J. Paulsson, and R. Losick. 2015. Stochastic switching of cell fate in microbes. *Annu. Rev. Microbiol.* 69: 381-403.
- Ortega, F., L. Acerenza, H. V. Westerhoff, F. Mas, and M. Cascante. 2002. Product dependence and bifunctionality compromise the ultrasensitivity of signal transduction cascades. *Proc. Natl. Acad. Sci. USA* 99: 1170-1175.
- Podgornaia, A. I., and M. T. Laub. 2015. Pervasive degeneracy and epistasis in a protein-protein interface. *Science* 347: 673-677.
- Poole, R. K., D. Lloyd, and R. B. Kemp. 1973. Respiratory oscillations and heat evolution in synchronously dividing cultures of the fission yeast *Schizosaccharomyces pombe* 972h. *Microbiol.* 77: 209-220.
- Porter, S. L., and J. P. Armitage. 2002. Phosphotransfer in *Rhodobacter sphaeroides* chemotaxis. *J. Mol. Biol.* 324: 35-45.
- Porter, S. L., and J. P. Armitage. 2004. Chemotaxis in *Rhodobacter sphaeroides* requires an atypical histidine protein kinase. *J. Biol. Chem.* 279: 54573-54580.
- Qian, H. 2007. Phosphorylation energy hypothesis: open chemical systems and their biological functions. *Annu. Rev. Phys. Chem.* 58: 113-142.
- Ram, S., and M. Goulian. 2013. The architecture of a prototypical bacterial signaling circuit enables a single point mutation to confer novel network properties. *PLoS Genet.* 9: e1003706.
- Rodenfels, J., K. M. Neugebauer, and J. Howard. 2019. Heat oscillations driven by the embryonic cell cycle reveal the energetic costs of signaling. *Dev. Cell* 48: 646-658.
- Rowland, M. A., and E. J. Deeds. 2014. Crosstalk and the evolution of specificity in two-component signaling. *Proc. Natl. Acad. Sci. USA* 111: 5550-5555.
- Rowland, M. A., W. Fontana, and E. J. Deeds. 2012. Crosstalk and competition in signaling networks. *Biophys. J.* 103: 2389-2398.
- Salathé, M., J. Van Cleve, and M. W. Feldman. 2009. Evolution of stochastic switching rates in asymmetric fitness landscapes. *Genetics* 182: 1159-1164.
- Schaller, G. E., S. H. Shiu, and J. P. Armitage. 2011. Two-component systems and their co-option for eukaryotic signal transduction. *Curr. Biol.* 21: R320-R330.
- Shacter, E., P. B. Chock, and E. R. Stadtman. 1984. Energy consumption in a cyclic phosphorylation/dephosphorylation cascade. *J. Biol. Chem.* 259: 12260-12264.
- Shacter-Noiman, E., P. B. Chock, and E. R. Stadtman. 1983. Protein phosphorylation as a regulatory device. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 302: 157-166.
- Shi, L., N. Pigeonneau, V. Ravikumar, P. Dobrinic, B. Macek, D. Franjevic, M.-F. Noirot-Gros, and I. Mijakovic. 2014. Cross-phosphorylation of bacterial serine/threonine and tyrosine protein kinases on key regulatory residues. *Front. Microbiol.* 5: 495.
- Siryaporn, A., and M. Goulian. 2008. Cross-talk suppression between the CpxA-CpxR and EnvZ-OmpR two-component systems in *E. coli*. *Mol. Microbiol.* 70: 494-506.
- Siryaporn, A., B. S. Perchuk, M. T. Laub, and M. Goulian. 2010. Evolving a robust signal transduction pathway from weak cross-talk. *Mol. Syst. Biol.* 6: 452.
- Skerker, J. M., B. S. Perchuk, A. Siryaporn, E. A. Lubin, O. Ashenberg, M. Goulian, and M. T.

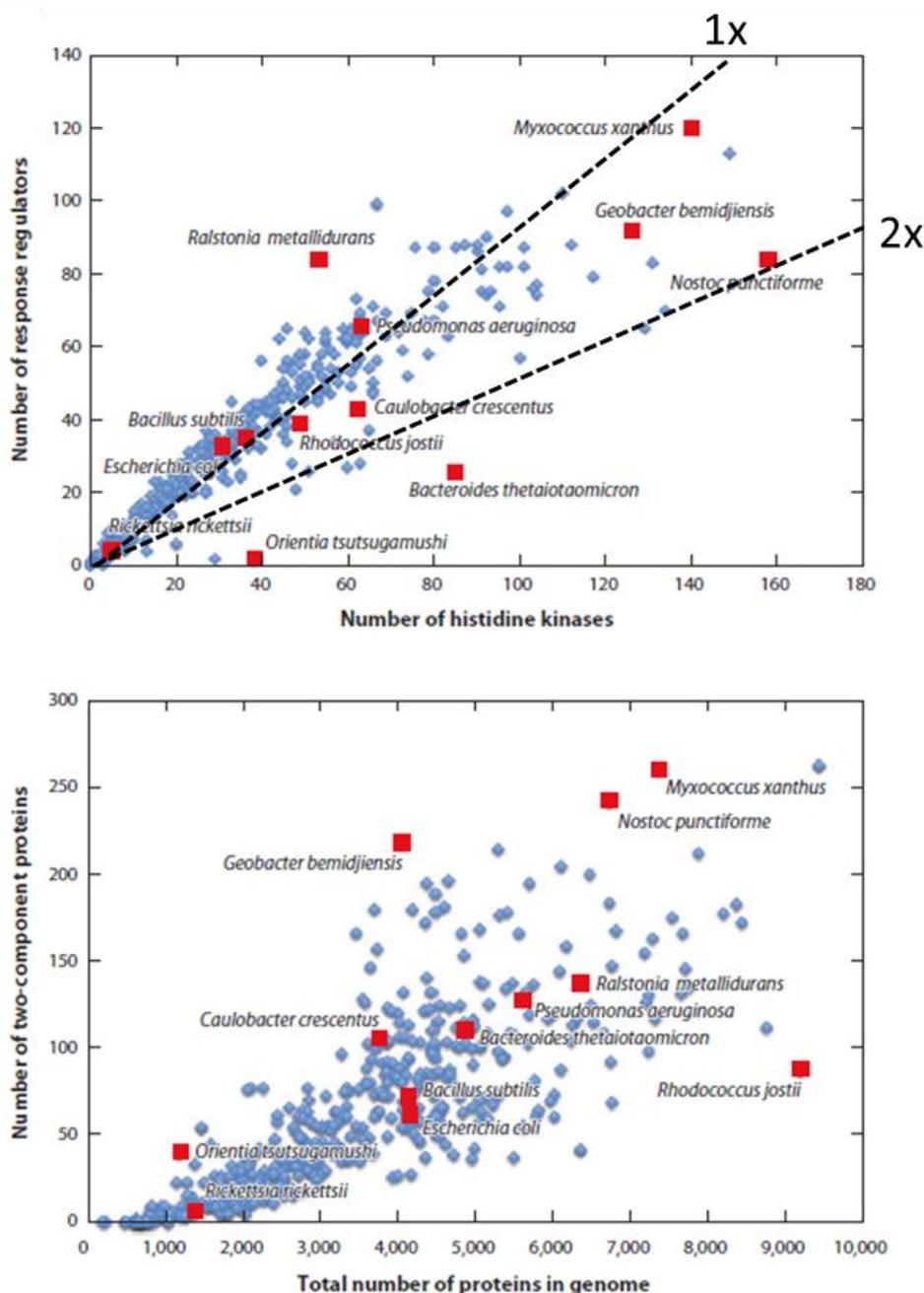
- Laub. 2008. Rewiring the specificity of two-component signal transduction systems. *Cell* 133: 1043-1054.
- Skoge, M., H. Yue, M. Erickstad, A. Bae, H. Levine, A. Groisman, W. F. Loomis, and W. J. Rappel. 2014. Cellular memory in eukaryotic chemotaxis. *Proc. Natl. Acad. Sci. USA* 111: 14448-144453.
- Smits, W. K., O. P. Kuipers, and J. W. Veening. 2006. Phenotypic variation in bacteria: the role of feedback regulation. *Nat. Rev. Microbiol.* 4: 259-271.
- Sonenshein, A. L. 2000. Control of sporulation initiation in *Bacillus subtilis*. *Curr. Opin. Microbiol.* 3: 561-566.
- Stadtman, E. R., and P. B. Chock. 1977. Superiority of interconvertible enzyme cascades in metabolic regulation: analysis of monocyclic systems. *Proc. Natl. Acad. Sci. USA* 74: 2761-2765.
- Stock, A. M., V. L. Robinson, and P. N. Goudreau. 2000. Two-component signal transduction. *Annu. Rev. Biochem.* 69: 183-215.
- Szurmant, H., T. J. Muff, and G. W. Ordal. 2004. *Bacillus subtilis* CheC and FliY are members of a novel class of CheY-P-hydrolyzing proteins in the chemotactic signal transduction cascade. *J. Biol. Chem.* 279: 21787-21792.
- Tabatabai, N., and S. Forst. 1995. Molecular analysis of the two-component genes, *ompR* and *envZ*, in the symbiotic bacterium *Xenorhabdus nematophilus*. *Mol. Microbiol.* 17: 643-652.
- Thattai, M., and A. van Oudenaarden. 2004. Stochastic gene expression in fluctuating environments. *Genetics* 167: 523-530.
- Ulrich, L. E., E. V. Koonin, and I. B. Zhulin. 2005. One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol.* 13: 52-56.
- van Haastert, P. J., and M. Postma. 2007. Biased random walk by stochastic fluctuations of chemoattractant-receptor interactions at the lower limit of detection. *Biophys. J.* 93: 1787-1796.
- Veening, J. W., W. K. Smits, and O. P. Kuipers. 2008. Bistability, epigenetics, and bet-hedging in bacteria. *Annu. Rev. Microbiol.* 62: 193-210.
- Wadhams, G. H., A. C. Martin, A. V. Warren, and J. P. Armitage. 2005. Requirements for chemotaxis protein localization in *Rhodobacter sphaeroides*. *Mol. Microbiol.* 58: 895-902.
- Wan, K. Y., G. Jékely. 2021. Origins of eukaryotic excitability. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 376: 20190758.
- Wegener-Feldbrügge, S., and L. Sogaard-Andersen. 2009. The atypical hybrid histidine protein kinase RodK in *Mycococcus xanthus*: spatial proximity supersedes kinetic preference in phosphotransfer reactions. *J. Bacteriol.* 191: 1765-1776.
- Wei, K., M. Moinat, T. R. Maarleveld, and F. J. Bruggeman. 2014. Stochastic simulation of prokaryotic two-component signalling indicates stochasticity-induced active-state locking and growth-rate dependent bistability. *Mol. Biosyst.* 10: 2338-2346.
- Weigt, M., R. A. White, H. Surmant, J. A. Hoch, and T. Hwa. 2009. Identification of direct residue contacts in protein-protein interaction by message passing. *Proc. Natl. Acad. Sci. USA* 106: 67-72.

- Wuichet, K., B. J. Cantwell, and I. B. Zhulin. 2010. Evolution and phyletic distribution of two-component signal transduction systems. *Curr. Opin. Microbiol.* 13: 219-225.
- Xu, Y., and J. Gunawardena. 2012. Realistic enzymology for post-translational modification: zero-order ultrasensitivity revisited. *J. Theor. Biol.* 311: 139-152.
- Xue, B., and S. Leibler. 2016. Evolutionary learning of adaptation to varying environments through a transgenerational feedback. *Proc. Natl. Acad. Sci. USA* 113: 11266-11271.
- Yamamoto, K., K. Hirao, T. Oshima, H. Aiba, R. Utsumi, and A. Ishihama. 2005. Functional characterization *in vitro* of all two-component signal transduction systems from *Escherichia coli*. *J. Biol. Chem.* 280: 1448-1456.
- Yang, Y., A. Pollard, C. Höfler, G. Poschet, M. Wirtz, R. Hell, and V. Sourjik. 2015. Relation between chemotaxis and consumption of amino acids in bacteria. *Mol. Microbiol.* 96: 1272-1282.
- Zarrinpar, A., S. H. Park, and W. A. Lim. 2003. Optimization of specificity in a cellular protein interaction network by negative selection. *Nature* 426: 676-680.
- Zhang, W., and L. Shi. 2005. Distribution and evolution of multiple-step phosphorelay in prokaryotes: lateral domain recruitment involved in the formation of hybrid-type histidine kinases. *Microbiology* 151: 2159-2173.

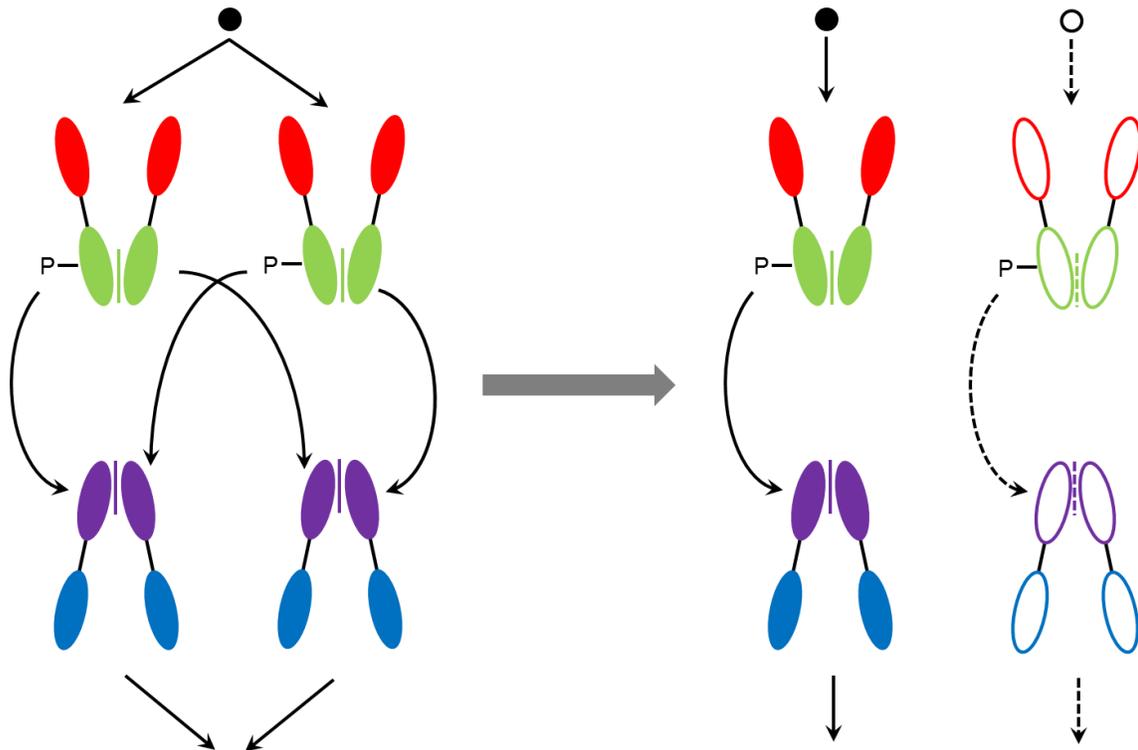
**Figure 22.1.** Idealized schematics for bacterial one-component, two-component, and phosphorelay systems for signal transduction. Different colored ovals denote protein domains (contained within the same protein if connected by a black line). Double lines denote cellular membranes. For the latter two systems, the covalently attached phosphate (P) group is derived from ATP after receipt of an external signal. Phosphorelays can have more complex structures than the one illustrated, with the first P transfer sometimes being to a separate protein, and with multiple players involved in longer chains of reaction. The ultimate response generally involves the operation of the final activated molecule in the pathway as a transcription factor, which binds the regulatory DNA of a target gene.



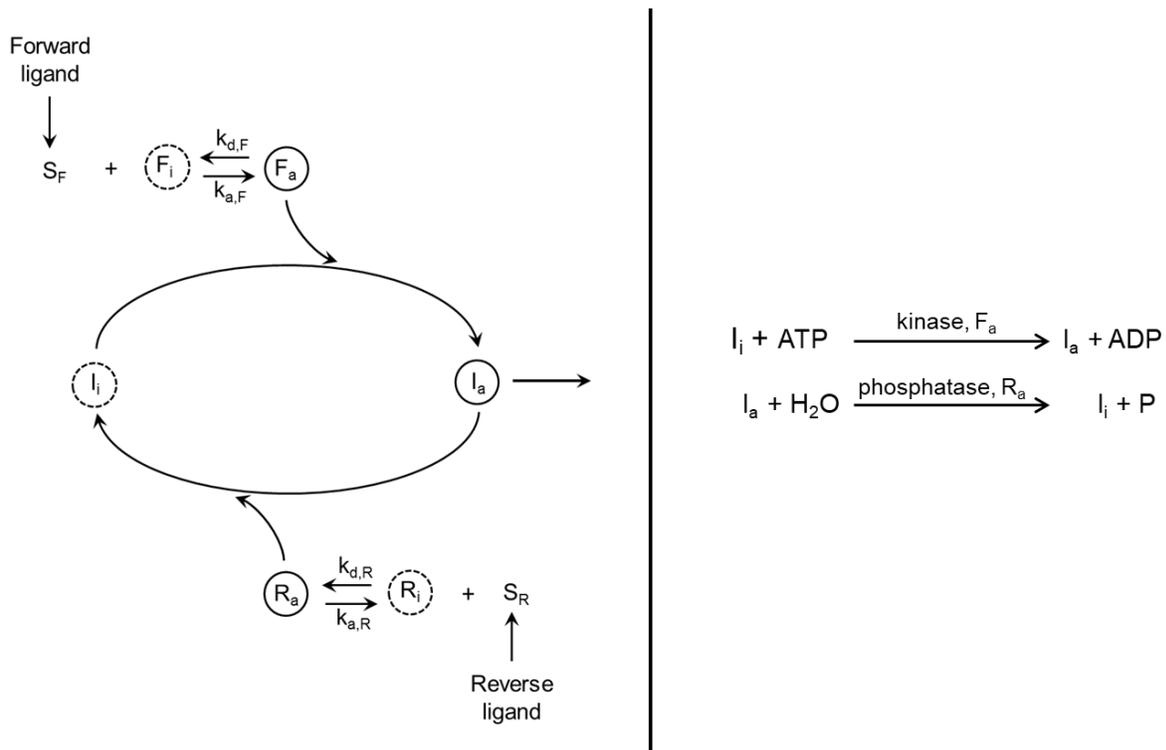
**Figure 22.2.** The number of response regulator (RR) and histidine kinases (HK) proteins involved in two-component systems in a wide range of bacterial species. **Above)** Scaling of the joint numbers of RRs and HKs within genomes. Diagonal lines denote positions of 1:1 and 1:2 ratios. **Below)** The number of proteins associated with two-components systems as a function of the total numbers of proteins encoded within genomes. From Capra and Laub (2012).



**Figure 22.3.** A general schematic for what might be necessary for the preservation and long-term functional divergence of a duplicated HK-RR pair. For the particular scenario noted, one system diverges and the other remains the same, although it is not necessary, and perhaps not even likely, that only one pair would undergo functionally significant evolutionary changes. If the two systems are to evolve to be completely insulated from each other: 1) new dimerization domains (vertical lines separating adjacent ovals) need to emerge for both the HK and RR proteins; and 2) a new feature of the phosphoryl transfer mechanism (large curved arrows) needs to be incorporated. Open and closed black dots denote two types of signaling ligands in the extracellular environment. The red and green ovals denote domains of the ancestral membrane-bound HK proteins, red being the sensory domain, and green being the autophosphorylation domain. Purple and blue ovals denote domains of the ancestral intracellular RR proteins, purple denoting the phosphotransfer domain, and blue denoting the DNA binding (or other output) domain. Right panel: Open vs. closed ovals and solid vs. dashed lines denote functional changes, with the system on the right no longer capable of crosstalk with the system on the left.



**Figure 22.4.** Generalized scheme for a monocyclic signal-transduction system. **Left)** An interconvertible protein  $I$  is transformed between active ( $a$ ) and inactive ( $i$ ) states by forward- and reverse-acting (F and R, respectively) enzymes, which themselves have active and inactive forms dependent on whether they are bound to their respective ligands,  $S_F$  and  $S_R$ . The  $k$  coefficients denote the association ( $a$ ) and dissociation ( $d$ ) constants between the ligands and the converting enzymes. **Right)** The reaction equations involving the interconvertible enzymes, with P denoting inorganic phosphate.



**Figure 22.5.** Response of the active fraction of an interconvertible protein ( $I^*$ ) to the concentration of ligand for the forward converting enzyme,  $[S_F]$ . Both plots are derived using Equation 22.1.8, with the upper plots denoting the limiting behavior when the concentration of  $I$  is well below the half-saturation constants of the converting enzymes, Equation 22.1.6. **Upper panel)** Results are given for the situation in which the concentration of the ligand for the reverse enzyme,  $[S_R]$ , is set equal to 1.0, with increasing levels of  $[S_F]$  (according to Equation 22.1.6, the results depend only on the ratio of these two concentrations). The dissociation constant for the forward enzyme and its ligand (equivalent to its half-saturation constant) is  $k_{D,F} = 1.0$ , with results given for different values of  $k_{D,R}$  (for the reverse enzyme). The parameter  $\beta$  is the ratio of kinetic potentials of the forward and reverse converter enzymes, as defined in Foundations 22.2. As  $\beta$  and  $k_{D,R}$  increase, autophosphorylation increasingly dominates and the response curves shift to the left. The black dotted line denotes the situation that would be expected if the response followed the Michaelis-Menten enzyme kinetics of the forward converter and its ligand. **Bottom panel)** Results are given for increasing total concentrations of the interconvertible enzyme. Both of the ligand concentrations and all of the catalytic coefficients and half-saturation constants are arbitrarily set to 1.0.

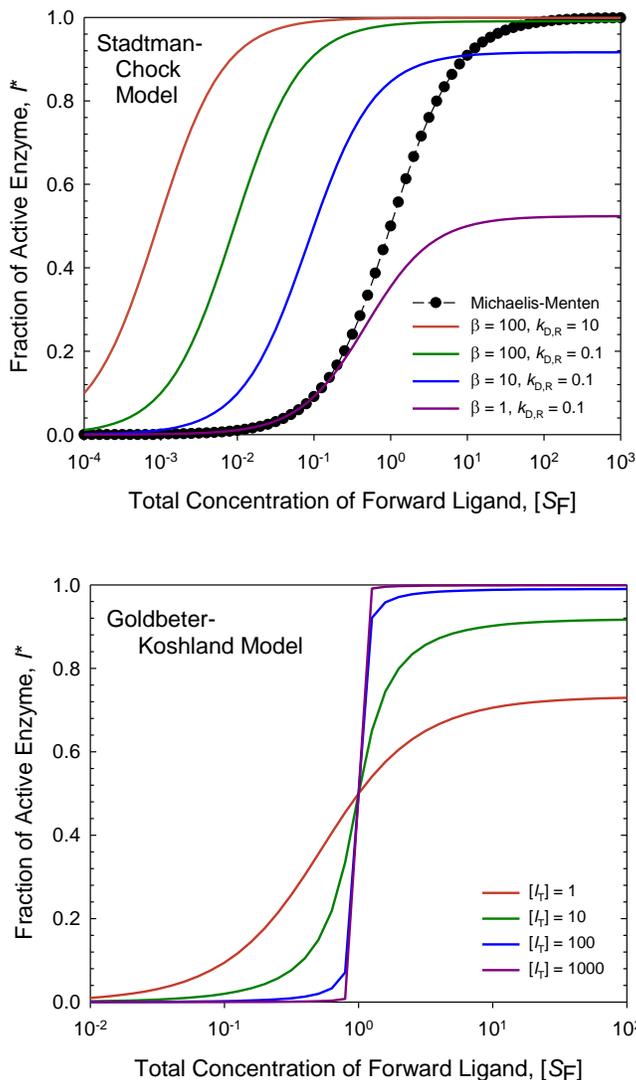
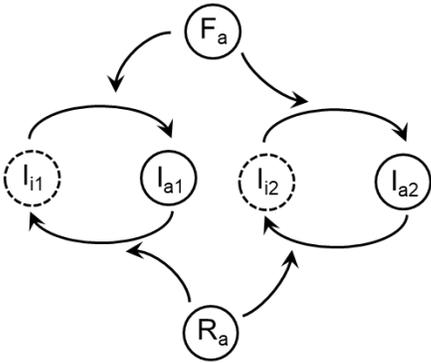
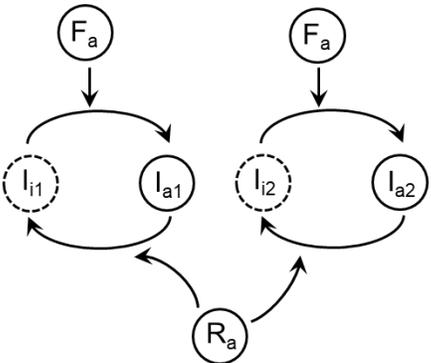


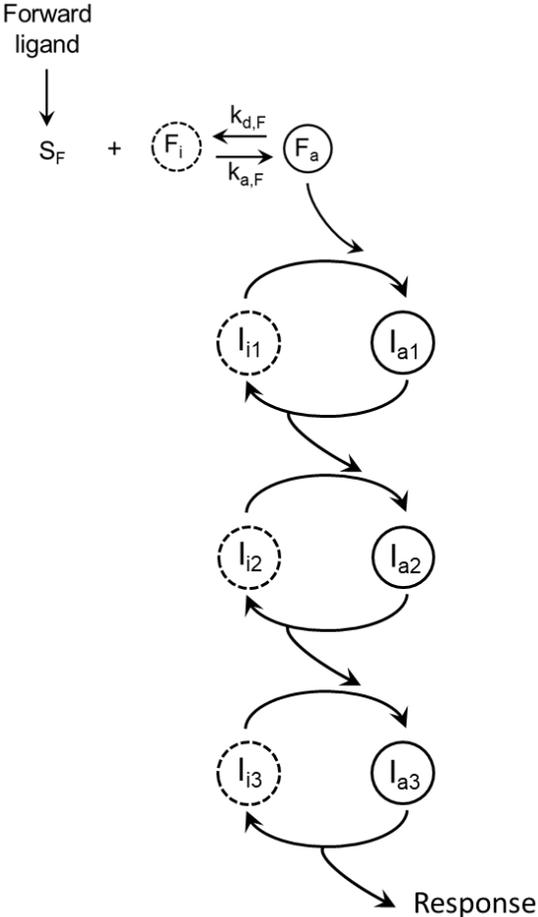
Figure 22.6. Three variants on the structure of signal-transduction pathways commonly found in eukaryotes, with notation as in Figure 22.4.



A kinase and a phosphatase sharing two intermediate substrates

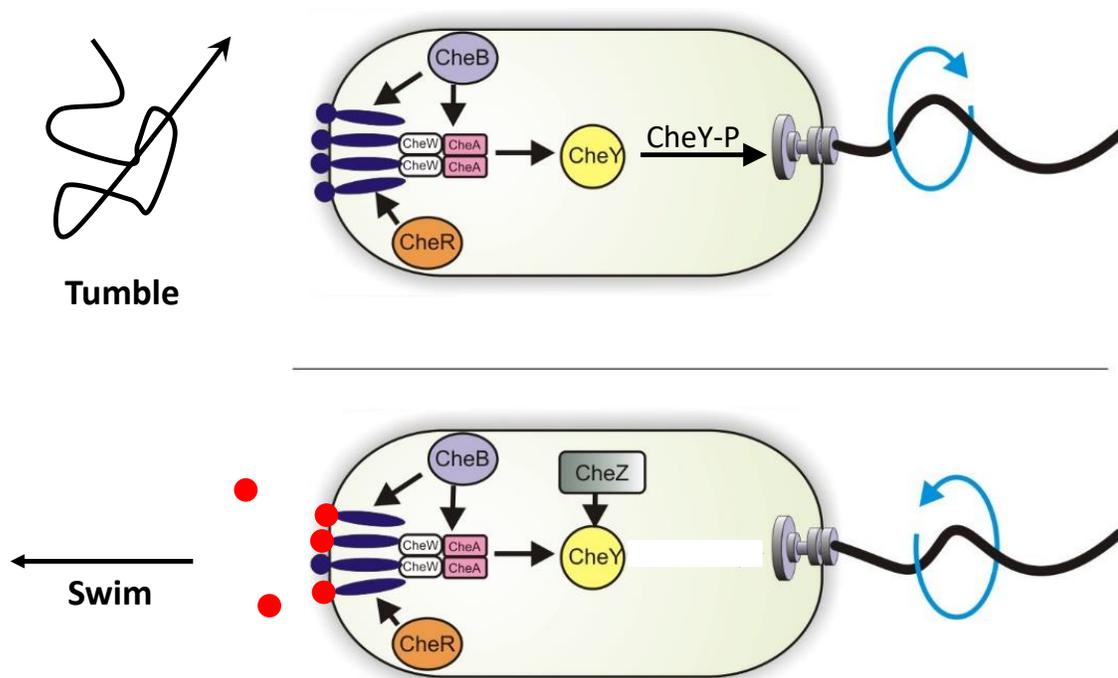


Two independently acting kinases, with one shared phosphatase



Triple cascade

**Figure 22.7.** Idealized schematic of the chemotaxis pathway in *E. coli*. The histidine kinase CheA is linked to the external sensors by another protein called CheW. In the absence of ligand binding, CheA becomes phosphorylated, thereby phosphorylating CheY, which binds to the base of the flagellum, inducing clockwise rotation and random tumbling. In the presence of ligand (red dots) binding and dephosphorylation of CheY, the flagellum rotates in a counter-clockwise fashion, causing the cell to propel forward in a directed manner. Proteins not specifically mentioned in the text: CheZ is a phosphatase that acts on CheY-P; CheR and CheB are methylases and demethylases that operate on the MCPs (methyl-accepting chemotaxis proteins; dark circles and ovals on the left) and modify their sensitivity to external ligand concentrations.



**Figure 22.8.** Phase diagrams for determining the equilibrium activity levels of interconvertible enzymes (I) subject to activation and inactivation cycles by enzymes F and R, respectively. The lines depict how the two rates change with increasing fraction of activated I. All points of intersection denote equilibria, but only the solid points are stable, as in these cases, deviations in both directions result in differences in activation and deactivation rates that return to the point. The fraction of activated I increases when the rate of activation (blue) exceeds the rate of deactivation, and vice versa under the opposite condition. **A)** With a system with no feedback, there is a single stable equilibrium. **B)** When activated enzyme feeds back positively to the forward enzyme, the dynamics are altered in such a way that there can be as many as three equilibria, with the central one being unstable (as deviations in either direction move the system to one of the alternative stable equilibria). Depending on the elevation and angularity of the activation curve, there might be only a single equilibrium in this case.

