MINIREVIEW

Two-Component Signal Transduction in Bacillus subtilis: How One Organism Sees Its World†

CÉLINE FABRET,* VICTORIA A. FEHER, AND JAMES A. HOCH*

Division of Cellular Biology, Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, California 92037

Variability and adaptability are crucial characteristics of organisms possessing the ability to survive and prosper in a wide variety of environmental conditions. The most adaptable bacteria contain a large reservoir of genetic information encoding biochemical pathways designed to cope with a variety of environmental situations. Organisms that have the genetic capability to respond to altered conditions do so when stimulated by specific signals. Recognition of specific signals and conversion of this information into specific transcriptional or behavioral responses is the essence of signal transduction.

A mechanism commonly found in bacteria for signal transduction is the two-component system (23, 26). Its basis is the conversion of signal recognition to a chemical entity, i.e., a phosphoryl group, that modifies the functional activity of proteins. Signal recognition and transduction are the province of the sensor histidine kinase component of the system. This protein has separable sensor and histidine phosphotransferase domains that function to recognize (bind) the signal, causing the kinase to autophosphorylate a histidine residue of the phosphotransferase domain (Fig. 1). The phosphoryl group is subsequently transferred to the second component protein, the response regulator, where it resides as an acyl phosphate of an aspartic acid residue. The response regulator consists of the phosphorylatable aspartate domain and an output domain that is activated to carry out its function by conformational or, perhaps, electrostatic alterations induced by the phosphoryl group. In most cases, the response regulator is a transcription activator for genes whose products are specifically utilized to respond to the unique nature of a given input signal. In the chemotaxis system of bacteria, the response regulator determines the direction of rotation of the flagellar motor. The basics of the signal transduction mechanism remain the same regardless of the input signal or the function of the response regulator.

This phosphoryl group-based signal transduction mechanism exists in two major conformations in microorganisms: the two-component system and a four-component system termed the phosphorelay (Fig. 1). Signal interpretation and transduction by histidine kinases are the same in both, but the target of the kinase in a phosphorelay is a single-domain response regulator consisting of only the phosphorylatable aspartate domain. This phosphorylated protein serves as a substrate for a phosphotransferase that transfers the phosphoryl group to a response regulator-transcription factor. The phosphotransferase is transiently phosphorylated on a histidine during this process. In a phosphorelay, the phosphoryl group is transferred in the order His-Asp-His-Asp, which differs from the His-Asp series of a two-component system. In the first-discovered phosphorelay used to initiate sporulation in Bacillus subtilis, all of the components (domains) reside on different proteins (4). Subsequently discovered phosphorelays in bacteria, fungi, and plants use composite proteins where the kinase and first response regulator domain and sometimes the phosphotransferase domain are contiguous within a single polypeptide chain (1, 6). The sporulation initiation phosphorelay is a signal integration circuit that processes both positive and negative signals, which suggests that phosphorelays are used where a number of opposing signals must be interpreted by the signal transduction system (22).

GENOME ANALYSIS OF TWO-COMPONENT SYSTEMS

Knowing the complete sequence of the B. subtilis chromosome allowed analysis of the number and kinds of two-component systems in this organism (14). The structural and functional principles for these analyses were the conserved ATP-binding site characteristic of sensor histidine kinases in conjunction with a conserved histidine motif and the overall similarity of the phosphorylated aspartate domains of response regulators (23). Using these criteria, 36 histidine kinases and 34 response regulators were found among the open reading frames identified in the genome (see Table 1). Comparing the kinases found to those of the distantly related gram-negative microorganism Escherichia coli revealed that none of the B. subtilis enzymes were composite kinases in which a phosphorylatable response regulator domain was contiguous with the kinase polypeptide. E. coli has five of these composite kinases that are believed to function in phosphorelays similar to the sporulation phosphorelay (19, 20).

The CheY protein is the single example in E. coli of a response regulator consisting of only the phosphorylatable aspartate domain. Three of these were found in B. subtilis. These include a close homologue of CheY as well as Spo0F, of the sporephorylation phosphorelay, and YneI, a protein of unknown function. While most response regulators are transcription factors dependent on phosphorylation for activity, phosphorylation of these single-domain response regulators serves other functions, such as interaction with the flagellar motor in the case of CheY, or as phosphointermediates in the phosphorelay in the case of Spo0F.
FIG. 1. Schematic view of two-component and phosphorelay systems. Activation signals recognized by sensor domains of histidine kinases result in autophosphorylation of a histidine in the histidine phosphotransferase domain (His PTase). The phosphoryl group (P) is transferred directly to the phosphorylated aspartate domain (PA) of a response regulator in a two-component system, causing a conformational change that activates the output domain. In a phosphorelay, the phosphoryl group is transferred to a PA domain that serves as a substrate for a phosphotransferase whose role is to transfer the phosphoryl group to the PA domain of a response regulator. Note that all the steps are reversible in many systems, which may result in dephosphorylation in the absence of a signal.

FIG. 2. Classification of kinases by the sequence around the phosphorylated histidine. Kinases were sorted into classes on the basis of the sequence relationships of the residues on either side of the phosphorylated histidine. The classes were related to their response regulator based on the homology of the response regulator output domain to those of E. coli (19). Note the orphan kinases of group IIIB were related to NtrB of E. coli through the homology of the residues surrounding the histidine to NtrB, not through homologies to NtrC, a response regulator which does not exist in B. subtilis. Gaps introduced to maximize alignment are indicated by the dots.
CLASSIFICATION OF HISTIDINE KINASES AND RESPONSE REGULATORS

The classification of histidine kinases and response regulators into related groups was not accomplished on the basis of overall protein homology. The sensor domains of histidine kinases differ greatly, possibly reflecting the diversity of molecules sensed by the microorganism. Histidine kinases are characterized by the presence of a conserved ATP-binding site that, while it distinguishes them from other proteins, does little to differentiate them. Examination of the region around the histidine that becomes phosphorylated was more informative. The histidine motifs fell into five homology classes of which two, IIA and IIB, were closely related (Fig. 2).

Response regulators present a special problem in classification since the phosphorylated aspartate domain is highly conserved in them. This suggests the observed conservation of amino acids occurs because of structural similarities of this domain. High-resolution structural studies of the CheY response regulator of *E. coli* and Spo0F of *B. subtilis* showed a remarkable similarity in structure between these two molecules. Despite the conservation of amino acids, alanine-scan-ning mutagenesis studies of Spo0F revealed that only a small number of residues around the active site determine specificity of interaction with other components of the signaling pathway (29). Thus, amino acid similarity per se is a valid criterion for functional relatedness but does not allow distinction among response regulators.

Most of the response regulators could be classified by the relatedness of their output domains. Structural determinations of this domain of the *E. coli* OmpR and NarL response regulators provided a basis for relating similarity to structure. Alignment of the C-terminal domains of *B. subtilis* response regulators with the amino acid sequence of the OmpR DNA-binding domain revealed a group of response regulators with high homology to *E. coli* OmpR (Fig. 3). The most informative conserved amino acids were the residues making up the hydrophobic core of this domain (17). All of the response regulators falling in this group were paired with a kinase classified as group IIA by the homology around the histidine residue. One exception to this rule is YccH, which has weak similarity to OmpR (Fig. 3). A similar study using the *E. coli* NarL output domain identified nine response regulators with high homology for those residues required for proper folding of the domain (2) (Fig. 4). Interestingly, all of these response regulators are paired with a kinase of group II. None of the kinases from group II or IIA were paired with a response regulator of a different type with the possible exception of YccG. Using this method of analysis, 23 of the response regulators were found to be related to either OmpR or NarL. Comparison of the entire catalytic domain of the kinases to *E. coli* kinases revealed that class II kinases were most related to NarX homologues and class IIA kinases were most related to EnvZ homologues as expected (data not shown). Since classification of the kinases was based on homology around the phosphorylated histidine and this region most certainly interacts with the active-site region of the phosphorylated aspartate domain of response regulators, the simplest conclusion for the observed relationships is that the catalytic domain of the kinase and both domains of the response regulator evolved as a unit from a common ancestor. Consistent with this conclusion is the observation that gene order in the transcription unit in which they reside is preserved within classes (see Table 1). The origins of the diverse sensor domains of...
the kinases remain to be uncovered, but clear subgroups exist within each group with sensor domains of similar size and membrane configuration. Some of the kinases within subgroups clearly evolved from a common progenitor (e.g., PhoR and ResE).

Three kinase-response regulator pairs were classified in group I, and they were related to the Others B group of *E. coli* (19). Similarly, four pairs were classified in group IV and were related to the Others A group of *E. coli*. CheY and CheA are alone in their classification. YneI is an orphan single-domain response regulator for which there is no basis for classification. Finally, at the level of the kinases, the C-terminal domains of YccG, YbdK, and YcbA do not align perfectly with the other members of their class.

---

**FIG. 4.** Relationships of response regulators to the output domain of NarL. Amino acid sequences of response regulators were compared to the sequence of the output domain of NarL of *E. coli*. The shaded residues are crucial for correct folding of the domain (2). Gaps introduced to maximize alignment are indicated by the dashes.

---

**FIG. 5.** Schematic structures of the kinases. Groups were determined from the homology of the residues surrounding the phosphorylated histidine of the histidine phosphotransferase domain. Related domains are the same color, and green rectangles are likely transmembrane segments.
ORPHAN KINASES OF CLASS IIIB

Five kinases originally grouped into class IIIB by their relatedness around the phosphorylated histidine were present on the chromosome without a linked response regulator gene. Two of these orphan kinases, KinA and KinB, are the major kinases responsible for phosphate input into the phosphorelay initiating sporulation (28). KinC was identified as a kinase that phosphorylates mutant forms of the Spo0A response regulator allowing bypass of the phosphorelay in sporulation (13, 15). YkvD is known to phosphorylate certain Spo0F mutants and bypass both KinA and KinB (12). YkrQ has not been implicated in the phosphorelay. These five kinases have residues around the phosphorylated histidine with sequence homology to those found in the NtrB kinase of E. coli. There is no formal equivalent to the NtrB-NtrC two-component system in B. subtilis, and the regulation of nitrogen metabolism is very different in the two organisms. In addition, none of the B. subtilis response regulators contain an NtrC-like ATPase domain required for $\sigma^{54}$ activity despite the presence of $\sigma^{54}$ and genes transcribed by it. The orphan kinases of class IIIB have no known relationship to nitrogen metabolism, and their sequence similarity around the phosphorylated histidine suggests they may all act as transducers of different signals in sporulation (11).

REGULATORY FUNCTIONS OF TWO-COMPONENT SYSTEMS

Several two-component systems have been extensively studied in B. subtilis and the genes they regulate are known. They include such systems as CheA-CheY in chemotaxis (24), PhoR-PhoP in phosphate regulation (27), ResE-ResD in anaerobic gene activation (21), ComP-ComA in competence (9), and DegS-DegU in degradative enzyme regulation (5). The CitS-CitT system may be involved in Mg$^{2+}$/citrate transport based on its close similarity to a system in E. coli, and LytS-LytT may
be involved in autolysis regulation (Table 1). The remainder of the systems identified from genome analysis have so little similarity to characterized systems from other organisms that a tentative functional assignment is unwarranted.

In a directed gene knockout study of the response regulators of the unknown two-component systems shown in Table 1, only the YycG-YycF system was found to be essential for growth (7). The other response regulator null mutations did not noticeably affect colony morphology, growth, or sporulation on laboratory media. It is probably safe to conclude that most of the unknown two-component systems shown in Table 1, only if a cow happened to stop for a bite.

The kinases, with the exception of five, are believed to be connected to the last membrane-spanning domain and ends at the phosphotransferase domain itself could be sites of kinase activation or deactivation. Their presence in a PAS domain is known to be present in the cytoplasmic linker, although that function remains obscure. In the case of KinC, a PAS domain may indeed be ligand-binding signal input (data not shown). The intermediate-length linkers from YclK, YvqE, YvqB, YrkQ, YesM, and YbdK are related and have a conserved sequence DEIGXhyA (hy is any hydrophobic residue) beginning about 40 residues distal to the last transmembrane region (Fig. 6). This sequence is also found in ResE and YycG. A second conserved sequence, GhyhyAhyhyXDXTE (E denotes a glutamic acid residue) and begins at loop positions spanning between secondary structure elements β-strand 1 to α-helix 1 (β1-α1), β2-α2, β3-α3, β4-α4, and β5-α5 are illustrated individually as members of groups I, II, IIIA, and IIIB, IV and collectively in the top bar graph for comparison. The residue type is coded by color: acidic in red (D and E), basic in dark blue (K and R), hydroxyl in rose (S and T), polar in light blue (H, N, and Q), hydrophobic in yellow (C, I, L, V, and M), aromatic in green (Y, F, and W), and structural in gray (A, G, and P). The numbering scheme is based on the Spo0F sequence, and conserved residues essential for phosphorylation are D10, D11, D54 (the phosphohydryl site), T82, and K104 by this numbering. Residues previously described as conserved (C) by alignment of response regulators from all organisms (30) are denoted. Response regulator sequences (14) were aligned with Clustal W (10) and illustrated with the Excel 98 (Microsoft) and Adobe Illustrator programs (Adobe).

### CYTOPLASMIC LINKERS BETWEEN SENSOR AND HISTIDINE PHOSPHOTRANSFERASE DOMAINS

The sensor domains of membrane kinases are connected to the histidine phosphotransferase domains through a cytoplasmic linker that starts at the end of the last membrane-spanning domain and ends at the phosphotransferase histidine motif. These linkers are of variable size but roughly fall into three length classes: ≤40, 60 to 80, and 130 to 170 amino acids. The shortest linkers are clearly related to one another and fall into two subgroups: (i) YvqD and KinB and (ii) YvqT, YocF, and YdfH (data not shown). The intermediate-length linkers from YclK, YvqE, YvqB, YrkQ, YesM, and YbdK are related and have a conserved sequence DEIGXhyA (hy is any hydrophobic residue) beginning about 40 residues distal to the last transmembrane region (Fig. 6). This sequence is also found in ResE and YycG. A second conserved sequence, GhyhyAhyhyXDXTE (E denotes a glutamic acid residue) and begins at loop positions spanning between secondary structure elements β-strand 1 to α-helix 1 (β1-α1), β2-α2, β3-α3, β4-α4, and β5-α5 are illustrated individually as members of groups I, II, IIIA, and IIIB, IV and collectively in the top bar graph for comparison. The residue type is coded by color: acidic in red (D and E), basic in dark blue (K and R), hydroxyl in rose (S and T), polar in light blue (H, N, and Q), hydrophobic in yellow (C, I, L, V, and M), aromatic in green (Y, F, and W), and structural in gray (A, G, and P). The numbering scheme is based on the Spo0F sequence, and conserved residues essential for phosphorylation are D10, D11, D54 (the phosphohydryl site), T82, and K104 by this numbering. Residues previously described as conserved (C) by alignment of response regulators from all organisms (30) are denoted. Response regulator sequences (14) were aligned with Clustal W (10) and illustrated with the Excel 98 (Microsoft) and Adobe Illustrator programs (Adobe).
changes in response regulator residues involved in individual specificity are most likely to result in altered interaction with kinases of the same group. As a corollary, if cross talk between kinase-response regulator pairs is of regulatory significance, it is likely to occur only within a group.

**PERSPECTIVES**

It is now becoming clear how two-component systems work. The basic chemical mechanisms of phosphotransfer, the structure of active domains, and the requirements for histidine kinase-response regulator interaction are relatively easy to study, and therefore, much is known. With a few exceptions, the nature of the signals activating these kinases remains obscure. The structure of histidine kinases, with the exception of some kinase fragments, and the mechanism of signal-activated autophosphorylation remain mysteries. The cellular roles of most two-component systems and the genes they activate are unknown. It is safe to conclude that there is much to learn about bacterial responses to their environment and how these systems help to mediate that response.

**ACKNOWLEDGMENT**

This work was supported, in part, by a grant GM19416 from the National Institute of General Medical Sciences, National Institutes of Health, USPHS.

**REFERENCES**


