Published in final edited form as:

Mol Microbiol. 2008 September ; 69(6): 1336–1340. doi:10.1111/j.1365-2958.2008.06379.x.

Signal transduction meets systems biology: deciphering specificity determinants for protein-protein interactions

Robert B. Bourret^{1,*}

¹ Department of Microbiology & Immunology, University of North Carolina, Chapel Hill, NC, 27599-7290, USA

Abstract

Two recent papers [Gao *et al. Mol. Microbiol.* **69**, 1358 (2008); Skerker *et al. Cell* **133**, 1043 (2008)] describe investigations into the specificity of protein-protein interactions that occur during signal transduction by two-component regulatory systems. This MicroCommentary summarizes and provides context for the reported findings. The results offer insights into molecular determinants that provide specificity to maintain signal separation and thus prevent deleterious crosstalk between pathways, as well as the potential extent and nature of interactions that may combine signals to achieve beneficial cross regulation among pathways. The methods employed are suitable for application to other systems.

Keywords

cross regulation; crosstalk; protein-protein interactions; response regulator; sensor kinase; two-component regulatory systems

Fundamental biological principles from the microbial world

Microorganisms are the dominant form of life on planet Earth by many criteria. Microbial populations are astronomical in number ($>10^{30}$ prokaryotes and $\sim 10\times$ more phage) (Rohwer, 2003; Whitman *et al.*, 1998), occupy essentially all possible environmental niches (including those not tolerated by macroscopic organisms) (Rothschild and Mancinelli, 2001), comprise a significant fraction of the Earth's biomass (more than plants) (Whitman *et al.*, 1998), and strongly influence geochemical cycles (Falkowski *et al.*, 2008). Perhaps most remarkably, the fantastic and almost incomprehensible variety of organisms visible to the unaided human eye constitutes an insignificant fraction of the diversity of life as assessed by 16S ribosomal RNA gene sequences (Pace, 1997). Thus if one wishes to discover fundamental biological principles generally applicable to all forms of life, it is logical to investigate microorganisms. Two recent papers, one in this issue of *Molecular Microbiology* from Ann Stock's group (Gao *et al.*, 2008) and another in *Cell* from the laboratories of Michael Laub and Mark Goulian (Skerker *et al.*, 2008), provide noteworthy illustrations of the potential knowledge to be gained from scrutinizing microbes. Together, the papers show how the use of broad systems biology approaches to investigate signal transduction in *Escherichia coli* can provide insight into even more general questions, in this case how a protein can specifically recognize its correct binding partner in the presence of many closely related competitors. Both the methods employed and the results obtained are of great interest.

*For correspondence. Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, NC 27599-7290, USA. E-mail bourret@med.unc.edu; Tel. (+1) 919 966 2679; FAX (+1) 919 962 8103.

Signal transduction by two-component regulatory systems

Living cells monitor parameters of interest in their environment, create internal representations of such stimuli, and use the encoded information to implement appropriate adaptive responses to changing conditions. This process of signal transduction is universal and occurs in both prokaryotic and eukaryotic cells, whether from unicellular or multicellular organisms. Two-component regulatory systems [reviewed in (Gao *et al.*, 2007)] are employed by a particularly wide range of organisms (Bacteria, Archaea, eukaryotic microorganisms, and plants) to carry out signal transduction. The most basic form of two-component system is composed of a sensor kinase and a response regulator (Fig. 1A). Each component typically contains conserved domains that facilitate communication between the two partners, as well as unique input or output domains specific for a particular pathway. Detection of environmental stimuli by the sensor kinase regulates the extent of phosphorylation of a conserved cytoplasmic domain termed DHp. Information thus encoded in the form of phosphoryl groups is then transferred to the conserved receiver domain of the response regulator. Phosphorylation results in a conformational change in the receiver domain, which in turn controls the activity of an output domain. Genomic sequencing has revealed more than 40 different types of output domains in response regulators, but two-thirds of response regulators probably bind DNA to regulate transcription and half of them contain winged helix-turn-helix output domains such as in the well-studied paradigms OmpR and PhoB (Galperin, 2006). Several members of the predominant OmpR/PhoB subclass of response regulators are known to dimerize via their receiver domains upon phosphorylation. Dimerization facilitates binding of the output domains to repeated DNA sequences in the promoter regions of regulated genes.

The challenges of crosstalk and opportunities of cross regulation

Evolution often selects for repeated use of a successful strategy to accomplish similar tasks. Thus, tens or hundreds of two-component systems might operate simultaneously in parallel in a single cell. Each system detects different stimuli, controls different biological processes, and operates on different timescales, but utilizes similar components and a common phosphochemistry. How, then, does a cell ensure that a given stimulus reproducibly yields the appropriate response, rather than provoking an irrelevant reaction? Imagine the chaos that would result in our homes if flipping a particular switch sometimes turned on a light, but on unpredictable occasions activated a toaster or a television instead. The household devices are all controlled by electricity, but rely on physically separate, insulated, and permanent connections between input and output to reliably deliver appropriate function. What “technologies” are employed by biological information processing systems to maximize high fidelity interactions and minimize crosstalk? In some situations, specificity is ensured by separation of closely related systems from one another in either space (compartmentalization within a cell) or time (different developmental stages). Another strategy is the use of scaffolding proteins to hold successive components of a pathway together, the functional equivalent of hard-wired connections. However, specificity of phosphotransfer in two-component systems is preserved in *in vitro* reactions consisting of just a sensor kinase and a response regulator (Skerker *et al.*, 2005; Yamamoto *et al.*, 2005). Thus, the proteins themselves contain the recognition determinants necessary to preclude crosstalk.

Although crosstalk is described above in detrimental terms, beneficial cross regulation probably also exists. Systems level analysis of transcription profiles (Oshima *et al.*, 2002) or growth phenotypes (Zhou *et al.*, 2003) for sets of mutants lacking each of the more than 30 two-component regulatory systems in *E. coli* provide clear evidence of interactions between

different pathways. Thus, the specificity of protein-protein recognition might be tuned to prevent some interactions across pathways yet allow others.

Signalling via changes in multimeric state

A change in the multimeric state of a protein can be used to transmit signals. For example, ligand binding by receptor tyrosine kinases results in dimerization and activation of kinase function. In two-component regulatory systems, sensor kinases remain homodimers throughout the signalling process, but some response regulators change oligomeric state upon phosphorylation, which in turn affects their output function. Oligomerization requires that proteins find their particular counterparts amid the distractions of multitudinous similar possibilities. To investigate the extent and specificity of response regulator dimerization, Gao *et al.* (2008) measured fluorescence resonance energy transfer (FRET) between cyan and yellow fluorescent proteins (FP) fused to the N-termini of the receiver domains of multiple response regulators (Fig. 1B). The N-terminal fusion location places the FP on the opposite side of the receiver domain from the active site, where the FP is least likely to interfere with response regulator function and also is likely to bring two FPs into close proximity with one another in cases where dimerization occurs through receiver domains. Because FRET signal strength (emission from YFP upon excitation of CFP) depends on the inverse of the distance between the FPs raised to the 6th power, FRET is potentially a very sensitive indicator of whether a reporter FP is in a monomer or a dimer. Gao *et al.* (2008) first used a wide variety of *in vivo* and *in vitro* tests to verify the functionality of FP-PhoB fusions and validate FRET as an assay for phosphorylation mediated dimerization. Interestingly, the kinetics of phosphorylation and dimerization were indistinguishable, indicating that phosphorylation was rate limiting. The authors next constructed CFP and YFP fusions to all 14 *E. coli* response regulators of the OmpR/PhoB subclass, mixed together both FP derivatives of each response regulator *in vitro*, and initiated phosphorylation. Because all response regulators tested to date can autophosphorylate using the small molecule phosphoramidate, the same phosphodonor could be used for every response regulator and potential complications of including sensor kinases (competing binding reactions, phosphatase activity) were avoided. Thirteen out of 14 pairs gave a FRET signal, which demonstrates that dimerization is a general property of OmpR/PhoB subclass response regulators and establishes a lower boundary of >90% on the sensitivity of the FRET method as an assay for response regulator dimerization. Gao *et al.* (2008) also applied their method to three response regulators (FixJ, NarL, NtrC) from other subclasses and in each case observed the result expected based on previous knowledge of individual multimerization properties.

Do response regulator heterodimers function as coincidence detectors?

Formation of heterodimers between different members of the Fos/Jun eukaryotic transcription factor family results in combinations with various functional properties. To assess the specificity of response regulator dimerization, Gao *et al.* (2008) measured heterodimer formation using all possible combinations of their CFP and YFP fusions to 17 different response regulators. FRET interactions were not observed across subclass boundaries. Within the OmpR/PhoB subclass, FRET was not seen in either permutation of 87% of $(14 \times 13)/2 = 91$ possible pairings, so most dimerization appears to be specific. However, a FRET signal was observed for the remaining $12/91 = 13\%$ of possible OmpR/PhoB subclass heterodimers, although the interactions were weaker than for homodimers. The FRET assay might underestimate the extent of heterodimerization that occurs *in vivo*, because response regulators obtain phosphoryl groups from phosphoramidate much less efficiently than from sensor kinases, so the extent of *in vitro* phosphorylation may not have been sufficient to support allowed heterodimer formation in all cases.

BlpR and ComE response regulators of *Streptococcus pneumoniae* were previously proposed to form a heterodimer that binds to a hybrid DNA recognition site (Knutsen *et al.*, 2004). However, the results of Gao *et al.* (2008) are the first direct evidence of heterodimer formation by response regulators of which I am aware. Furthermore, the relatively high frequency of heterodimer formation raises the possibility of a physiologically relevant regulatory role in which heterodimers recognize different targets from either parent homodimer. If phosphorylation of both partners is necessary for dimerization under *in vivo* conditions, then heterodimer formation could act as a coincidence detector for simultaneous activation of two signalling pathways. In this context, it is intriguing that CpxR participated in seven of the 12 observed heterodimers. CpxR is known to assimilate a diversity of metabolic information (Wolfe *et al.*, 2008), so heterodimers might provide a previously unappreciated means of signal integration. To assess the physiological relevance of heterodimers, it would be helpful to identify their potential targets. The observation that target DNA binding caused homodimerization of FP-OmpR in the absence of phosphorylation (Gao *et al.*, 2008) suggests a plausible approach. It might be possible to find heterodimer binding sites by screening plasmid DNA from a genomic library for sequences that are able to provoke an *in vitro* FRET signal from a mixture of two FP-response regulator fusions.

Picking the proper phosphotransfer partner

The FRET assay provides a tool to assess the extent of interaction between two-component regulatory systems at the output end of the pathway. *In vitro* phosphotransfer assays previously revealed that the middle of the pathway has a high degree of fidelity (Skerker *et al.*, 2005; Yamamoto *et al.*, 2005). In new work, Skerker *et al.* (2008) have now probed deeper into the basis for specific recognition between sensor kinase and response regulator partners. They reasoned that if a residue in one protein changed over evolution, a compensatory change in the partner protein would be selected for to retain the specific interaction of the pair. Such covariation could be identified by calculating the mutual information content of all possible pairs consisting of one residue taken from an alignment of multiple sensor kinase sequences and the other from an alignment of multiple response regulator sequences. For example, if sensor kinase sequences are sorted by amino acid at position *i*, does a particular amino acid occur at response regulator position *j* more frequently in the partners of sensor kinases with say serine at position *i* than in those with lysine at position *i*? This method ideally requires a large set of matched sensor kinase and response regulator sequences, but it is not necessarily obvious from sequence information alone which sensor kinase interacts with which response regulator, and a relatively modest number of sensor kinase/response regulator pairings have been experimentally established. Skerker *et al.* (2008) cleverly exploited the knowledge that if the genes for a sensor kinase and a response regulator belong to the same operon, then the encoded proteins probably represent a matched pair, and were therefore able to analyze the conserved domains of ~1300 pairs from ~200 bacterial genomes.

It is noteworthy that this bioinformatics analysis does not require any structural information in order to identify residues important for specific interactions between two proteins. We do not yet have the structure of a sensor kinase/response regulator complex but, based on the related structure of the Spo0B histidine phosphotransferase complexed with the Spo0F response regulator (Zapf *et al.*, 2000), there is not a one to one correlation between interface contact residues in the structure and recognition residues identified by the covariation method. However, all residues predicted to be important for recognition specificity are close to the physical interface (Skerker *et al.*, 2008). Many more receiver domain residues covaried with residues in the DHp domain than in the CA domain of the sensor kinase. Skerker *et al.* (2008) constructed a variety of chimeric sensor kinase proteins and performed

phosphotransfer assays to narrow down the region of the sensor kinase actually responsible for recognition of the partner response regulator. Changing the entire DHp domain was sufficient to completely alter specificity, as was changing the portion of the DHp domain on the C-terminal side of the histidine phosphorylation site. Remarkably, changing just three amino acids in the EnvZ sensor kinase to their RstB counterparts completely switched the phosphotransfer specificity from OmpR to RstA! Changing the same three residues was not sufficient to switch the specificity of EnvZ to that of four other sensor kinases. However, changing five residues identified by the covariation analysis in the DHp α 1 helix, two residues in the α 2 helix, and the intervening turn (a span of ~30 amino acids) in five different cases enabled EnvZ to efficiently transfer phosphoryl groups to a new response regulator and simultaneously lose the ability to transfer to OmpR. Furthermore, the engineered sensor kinases also exhibited altered specificity *in vivo*, as assayed by the expression of *gfp* reporter genes from appropriate promoters.

Skerker *et al.* (2008) have laid the foundation for rewiring two-component regulatory circuits at will, and have demonstrated an initial proof of concept. Some hurdles might remain, but the way forward appears clear. The reciprocal experiment of changing specificity by altering response regulator residues has not yet been reported, but the covariation analysis predicts that six residues in the α 1 helix and the α 1- β 2 loop of the receiver domain are likely to be important. Crosstalk or cross regulation might reflect partial matches at key specificity residues. Because the covariation analysis was based on a subset of bacterial sequences, it is not yet known whether the results are applicable to two-component systems in Archaea and/or Eukarya, or even bacterial systems in which the components are not encoded in the same operon. From a broader perspective, the covariation analysis method is generally applicable to protein families other than two-component regulatory systems and should become increasingly powerful as sequence databases continue to grow. Because residues can covary for reasons in addition to protein-protein interactions, such as other functions or common ancestry, experiments will be needed to identify the most critical residues and eliminate false positives from among the candidates predicted by bioinformatics analysis. Finally, another recent report in *Molecular Microbiology* indicates it also should be possible to rationally manipulate the timing characteristics (i.e. the lifetime of the phosphorylation activated state) of two-component regulatory system circuits at least 100x by altering the amino acids at two variable positions in the response regulator active site (Thomas *et al.*, 2008).

Prediction of branched pathways

A central challenge in an era of increasingly rapid genomic sequencing is the development of means to interpret large volumes of data. The microbial world is so immense that, for the foreseeable future, we are unlikely to know anything about most of it other than what can be gleaned from analysis of DNA sequences in the absence of experiments. Derivation of empirical rules for bioinformatics annotation is a useful step that could be followed later by investigation of the underlying molecular mechanisms. The work of Skerker *et al.* (2008) suggests that it might become possible to discern features of signalling circuit topology by inspection of predicted sensor kinase and response regulator amino acid sequences.

Many biological information-processing circuits have a branched rather than linear organization. For example, the *Bacillus subtilis* sporulation network integrates inputs from five different sensor kinases (KinA, KinB, KinC, KinD, KinE) into phosphorylation of a single response regulator, Spo0F. In bacterial chemotaxis, the CheA sensor kinase distributes information in the form of phosphoryl groups to two output branches via the CheY (excitation) and CheB (adaptation) response regulators. If the covariation analysis of Skerker *et al.* (2008) applies to two-component regulatory systems beyond sensor kinases

and response regulators that belong to the same operon, then KinABCDE should share residues allowing them all to specifically interact with Spo0F, and CheBY should share key residues that allow them to interact with CheA. Thus it may become possible to examine the amino acid sequences of all sensor kinases and response regulators from a particular organism for shared interaction residues to predict which components might participate in branched pathways.

Future experimental synergies

In addition to interesting results, Gao *et al.* (2008) and Skerker *et al.* (2008) describe useful methods with potential applications to each other's work. There is a good *in vitro* assay for phosphotransfer, but the *in vivo* assay of downstream gene expression is a few steps removed from the phosphotransfer reaction. The FRET assay for response regulator dimerization could be used as a measure of phosphotransfer *in vivo*. The FRET assay is useful to reveal dimerization, but covariation analysis in conjunction with available crystal structures could help identify candidate specificity residues in the $\alpha 4$ - $\beta 5$ - $\alpha 5$ interface. Such residues could be altered and the effects on dimer formation assessed with the FRET assay. A greater understanding of specificity determinants could lead to the ability to predict heterodimer partners from receiver domain amino acid sequences. Additional uses for the fascinating results and methods of Gao *et al.* (2008) and Skerker *et al.* (2008) seem likely to arise.

Acknowledgments

I thank Ruth Silversmith for helpful comments on the text. National Institutes of Health grant GM050860 funds research on two-component regulatory systems in our laboratory. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

References

- Falkowski PG, Fenchel T, Delong EF. The microbial engines that drive Earth's biogeochemical cycles. *Science*. 2008; 320:1034–1039. [PubMed: 18497287]
- Galperin MY. Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J Bacteriol*. 2006; 188:4169–4182. [PubMed: 16740923]
- Gao R, Mack TR, Stock AM. Bacterial response regulators: versatile regulatory strategies from common domains. *Trends Biochem Sci*. 2007; 32:225–234. [PubMed: 17433693]
- Gao R, Tao Y, Stock AM. System-level mapping of *E. coli* response regulator dimerization with FRET hybrids. *Mol Microbiol*. 2008; 69:1358–1372. [PubMed: 18631241]
- Knutsen E, Ween O, Havarstein LS. Two separate quorum-sensing systems upregulate transcription of the same ABC transporter in *Streptococcus pneumoniae*. *J Bacteriol*. 2004; 186:3078–3085. [PubMed: 15126469]
- Oshima T, Aiba H, Masuda Y, Kanaya S, Sugiura M, Wanner BL, Mori H, Mizuno T. Transcriptome analysis of all two-component regulatory system mutants of *Escherichia coli* K-12. *Mol Microbiol*. 2002; 46:281–291. [PubMed: 12366850]
- Pace NR. A molecular view of microbial diversity and the biosphere. *Science*. 1997; 276:734–740. [PubMed: 9115194]
- Rohwer F. Global phage diversity. *Cell*. 2003; 113:141. [PubMed: 12705861]
- Rothschild LJ, Mancinelli RL. Life in extreme environments. *Nature*. 2001; 409:1092–1101. [PubMed: 11234023]
- Skerker JM, Prasol MS, Perchuk BS, Biondi EG, Laub MT. Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. *PLoS Biol*. 2005; 3:e334. [PubMed: 16176121]

- Skerker JM, Perchuk BS, Siryaporn A, Lubin EA, Ashenberg O, Goulian M, Laub MT. Rewiring the specificity of two-component signal transduction systems. *Cell*. 2008; 133:1043–1054. [PubMed: 18555780]
- Thomas SA, Brewster JA, Bourret RB. Two variable active site residues modulate response regulator phosphoryl group stability. *Mol Microbiol*. 2008; 69:453–465. [PubMed: 18557815]
- Whitman WB, Coleman DC, Wiebe WJ. Prokaryotes: the unseen majority. *Proc Natl Acad Sci U S A*. 1998; 95:6578–6583. [PubMed: 9618454]
- Wolfe AJ, Parikh N, Lima BP, Zemaitaitis B. Signal integration by the two-component signal transduction response regulator CpxR. *J Bacteriol*. 2008; 190:2314–2322. [PubMed: 18223085]
- Yamamoto K, Hirao K, Oshima T, Aiba H, Utsumi R, Ishihama A. Functional characterization in vitro of all two-component signal transduction systems from *Escherichia coli*. *J Biol Chem*. 2005; 280:1448–1456. [PubMed: 15522865]
- Zapf J, Sen U, Madhusudan Hoch JA, Varughese KI. A transient interaction between two phosphorelay proteins trapped in a crystal lattice reveals the mechanism of molecular recognition and phosphotransfer in signal transduction. *Structure*. 2000; 8:851–862. [PubMed: 10997904]
- Zhou L, Lei XH, Bochner BR, Wanner BL. Phenotype microarray analysis of *Escherichia coli* K-12 mutants with deletions of all two-component systems. *J Bacteriol*. 2003; 185:4956–4972. [PubMed: 12897016]

Fig. 1.

Two-component regulatory systems. Green colouring, conserved domains that define membership in two-component systems; rose colouring, unique domains; DHp, dimerization histidine phosphotransfer domain; CA, catalytic ATP-binding domain; receiver, receiver domain. (A) Signal transduction in a typical two-component regulatory system. (B) *In vitro* assay employed by Gao *et al.* (2008) to measure dimerization of OmpR class response regulators. CFP, cyan fluorescent protein; YFP, yellow fluorescent protein; FRET, fluorescence resonance energy transfer.