Comparative analysis of prototype two-component systems with either bifunctional or monofunctional sensors: differences in molecular structure and physiological function

Rui Alves¹,²,³ and Michael A. Savageau¹*†
¹Department of Microbiology and Immunology, University of Michigan Medical School, 5641 Medical Science Building II Ann Arbor, MI 48109–0620, USA.
²Grupo de Bioquímica e Biologia Teóricas, Instituto Rocha Cabral, Caçada Bento da Rocha Cabral 14, 1250 Lisboa, Portugal.
³Programa Gulbenkian de Doutoramentos em Biologia e Medicina, Departamento de Ensino, Instituto Gulbenkian de Ciencia, Rua da Quinta Grande 6, 1800 Oeiras, Portugal.

Summary

Signal transduction by a traditional two-component system involves a sensor protein that recognizes a physiological signal, autophosphorylates and transfers its phosphate, and a response regulator protein that receives the phosphate, alters its affinity toward specific target proteins or DNA sequences and causes change in metabolic activity or gene expression. In some cases the sensor protein, when unphosphorylated, has a positive effect upon the rate of dephosphorylation of the regulator protein (bifunctional sensor), whereas in other cases it has no such effect (monofunctional sensor). In this work we identify structural and functional differences between these two designs. In the first part of the paper we use sequence data for two-component systems from several organisms and homology modelling techniques to determine structural features for response regulators and for sensors. Our results indicate that each type of reference sensor (bifunctional and monofunctional) has a distinctive structural feature, which we use to make predictions regarding the functionality of other sensors. In the second part of the paper we use mathematical models to analyse and compare the physiological function of systems that differ in the type of sensor and are otherwise equivalent. Our results show that a bifunctional sensor is better than a monofunctional sensor both at amplifying changes in the phosphorylation level of the regulator caused by signals from the sensor and at attenuating changes caused by signals from small phosphodonors. Cross-talk to or from other two-component systems is better suppressed if the transmitting sensor is monofunctional, which is the more appropriate design when such cross-talk represents pathological noise. Cross-talk to or from other two-component systems is better amplified if the transmitting sensor is bifunctional, which is the more appropriate design when such cross-talk represents a physiological signal. These results provide a functional rationale for the selection of each design that is consistent with available experimental evidence for several two-component systems.

Introduction

Two-component systems (TCS) are signal transduction modules that exist mainly in bacteria but have also been found in fungi and plants (recently reviewed in Hellingwerf et al., 1998; see also Parkinson, 1993; Hoch and Silhavy, 1995; Perego and Hoch, 1996; Schaller, 1997; Posas et al., 1998). Two-component systems differ from eukaryotic phosphorylation cascades like the MAP kinase cascade. In the former, ATP is only consumed in the first step of the cascade to provide for autophosphorylation of a histidine residue in the sensor, whereas in the latter, ATP is consumed at each step in the cascade to provide for phosphorylation of the protein at that step. More than 180 different TCS have been identified in bacteria (Kadner, 1996; see also http://www.genome.ad.jp/kegg/regulation.html) and over one thousand putative sensors and response regulators in over one hundred organisms have had their gene sequence determined (e.g. http://www.expasy.ch/srs5/ http://www-fp.mcs.anl.gov/~gaasterland/genomes.html). Two-component systems are often involved in complex circuits that exhibit enormous variation in design. The physiological importance of systems that differ in the type of sensor and are otherwise equivalent. Our results show that a bifunctional sensor is better than a monofunctional sensor both at amplifying changes in the phosphorylation level of the regulator caused by signals from the sensor and at attenuating changes caused by signals from small phosphodonors. Cross-talk to or from other two-component systems is better suppressed if the transmitting sensor is monofunctional, which is the more appropriate design when such cross-talk represents pathological noise. Cross-talk to or from other two-component systems is better amplified if the transmitting sensor is bifunctional, which is the more appropriate design when such cross-talk represents a physiological signal. These results provide a functional rationale for the selection of each design that is consistent with available experimental evidence for several two-component systems.
TCS provides strong motivation to search for underlying design principles that would allow one to rationalize the variations in design. The search is in its infancy and there are undoubtedly many design principles that remain to be discovered.

Here we focus on a specific example of a simple design principle involving TCS composed of a sensor protein and a response regulator protein. These traditional TCS can be considered the prototype for a motif that has been elaborated on and modified in more complex cases that will not be considered here. [For example, we will not address cases in which there are specific aspartate phosphatases (e.g. Perego and Hoch, 1996; Blat et al., 1998) that are independent of the sensor protein. In some prototype cases the sensor protein, when unphosphorylated, has a positive effect upon the rate of dephosphorylation of the regulator protein (bifunctional sensor) whereas in other cases it has no such effect (monofunctional sensor). For example, the TCS involved in the regulation of chemotaxis via the CheA/CheY cascade in Escherichia coli has a monofunctional sensor (reviewed in Eisenbach, 1996), whereas the TCS involved in the regulation of osmotic pressure via the EnvZ/OmpR cascade has a bifunctional sensor (reviewed in Pratt et al., 1996). These bifunctional and monofunctional sensor proteins exhibit different specificities in molecular structure that, by homology modelling, are found to reoccur in many other TCS.

In the first part of this paper we use sequence data for other TCS from several organisms, together with homology modelling techniques, to predict structural features and in turn functionality of their sensor proteins. These predictions are confirmed in a few cases for which there is independent biochemical or genetic evidence for functionality. Our results suggest an association between a sensor’s mode of action (monofunctional or bifunctional) and the structure of its ATP-binding domain. In the second part of this paper we analyse and compare the functional effectiveness of TCS with either bifunctional or monofunctional sensors by using a technique known as Mathematically Controlled Comparison. This technique, which is analogous to a well-controlled experiment, determines the differences in the physiology between alternative designs that are otherwise equivalent. This approach reveals qualitative differences (independent of specific values for the parameters of the systems) as well as quantitative differences (statistical tendencies when a large sample of numerical values for the parameters is examined). We also obtain results that discriminate between different types of cross-talk to and from TCS. Our results suggest physiological situations that favour each of the two sensor designs. Hence, these results provide a functional rationale for the selection of each design. Experimental evidence for several TCS is discussed in the light of this rationale.

Results

Abstractions of the alternative designs for a prototype TCS are represented in Fig. 1. The unphosphorylated sensor protein \( S(X) \) becomes phosphorylated \( (S^*,X) \) in response to changes in a physiological signal \( (Q,X) \). Autophosphorylation of the sensor is achieved with consumption of ATP. Once phosphorylated, the sensor loses its phosphate group, either by auto-dephosphorylation (although this occurs on a time scale that is not of interest here) or by transfer of the phosphate to an aspartate residue in the regulator protein \( (R^*,X) \) (e.g. Weiss and Magasanik, 1988; Sanders et al., 1989; Hsing et al., 1998; Jung and Altendorf, 1998; Jiang and Ninfa, 1999). This covalent modification causes changes in the response regulator, thereby altering its affinity toward specific target proteins or DNA sequences \( (T) \), and this in turn leads to changes in metabolic activity or gene expression. The dephosphorylation of response regulators (e.g. Perego and Hoch, 1996) is, in some cases, enhanced by the unphosphorylated form of the sensor protein (bifunctional sensor, Fig. 1A). In other cases, the sensor protein lacks this ability (monofunctional sensor, Fig. 1B). In either case, the regulator protein also can be phosphorylated by other mechanisms \( (Q,X) \), e.g. by small phosphodonor like acetyl-phosphate, but not by ATP (e.g. Lukat et al., 1992; McClery, 1996; Mayover et al., 1999). Cross-talk from the sensor proteins of other two-component systems has been demonstrated with purified proteins in vitro (Ninfa et al., 1988) and with overexpressed proteins in vivo (Grob et al., 1994), but the physiological significance of these results has been questioned (Wanner, 1992).

Structure of regulators

Structures of some response regulators have been determined and can be used as templates to predict the 3D structure of other regulators by homology modelling. The known structures of response regulators have receiver domains composed of alternating \( \alpha \)-helices and \( \beta \)-sheets. The 3D shape resembles a barrel composed of five \( \alpha / \beta \) sequential segments, with parallel \( \beta \)-sheets (blue) in the middle, surrounded by the \( \alpha \)-helices (red) shown in Fig. 2. In the protein classification database SCOP (Murzin et al., 1995) response regulators are classified within the flavodoxin-like folds. The aspartate residue that is phosphorylated is at the tip of one of the internal \( \beta \)-sheets, almost in the loop that connects it to the next \( \alpha \)-helix (Fig. 2A). These structural features appear to be at least partially conserved in the receiver domains of all the response regulators we have been able to model (Fig. 2B, Table 1). Comparison of the experimentally determined structures of
response regulators did not reveal structural features that might indicate any distinction between response regulators of TCS with either bifunctional sensors or monofunctional sensors. Similar results also were obtained when we examined the modelled structures for response regulators. Although the phosphorylation domain of all response regulators whose structure has been determined or modelled so far have similar structures this does not allow us to conclude that the hundreds to thousands more that are as of yet unresolved all share the same structure. There are known cases where protein domains with very similar sequence have very different folds and cases where protein domains with very different sequence have very similar folds. Table 1 may be used as a guide to indicate response regulators that are likely to have similar structures. In the effort to explore the fold space of response regulators, it will probably be more informative to concentrate on resolving structures of response regulators that are not present in Table 1.
Structure of sensors

Partial structures for both types of sensors also have been determined and can be used as templates to predict the 3D structure of other sensors by homology modelling. Structures of the soluble domains of the CheA sensor (monofunctional, Fig. 3A) as well as the EnvZ sensor (bifunctional, Fig. 4A) have been determined. The CheA transmitter domain is of the type HPt, reminiscent of the PTS phosphotransferase systems. The EnvZ sensor has an HK transmitter domain. In the protein classification database SCOP (Murzin et al., 1995) sensor histidine kinases are classified within the ROP-like folds ("four helices; dimers of identical alpha-hairpin subunits; bundle, closed, left-handed twist"). Even though the domains of the EnvZ sensor responsible for its bifunctionality are not defined, an HK transmitter domain appears to be essential for the increase in the rate of dephosphorylation of the response regulator (Zhu et al., 2000). In the same work, Zhu et al. (2000) also showed that this increase is much less for the bifunctional sensors if their catalytic ATP-binding domain is truncated from their transmitter domain.
Alternative regulatory designs for two-component systems

Fig. 4. Three-dimensional structure for the catalytic domain of bifunctional sensors.
A. Structure determined by X-ray crystallography for the EnvZ sensor from Escherichia coli (PDB reference identifier 1BXD).
B. Structure predicted by homology modelling for the putative O66656 sensor from Aquifex aeolicus.

Table 1. Putative response regulator proteins that yielded a predicted structure by homology modelling techniques.

<table>
<thead>
<tr>
<th>Organism</th>
<th>SWISSPROT entry no.</th>
<th>Name</th>
<th>Organism</th>
<th>SWISSPROT entry no.</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aeromas jandaei</td>
<td>P97015</td>
<td>P97015</td>
<td>Bacillus megaterium</td>
<td>P39486</td>
<td>YGD1</td>
</tr>
<tr>
<td>Agrobacterium</td>
<td>Q07783</td>
<td>CHVI</td>
<td>Bacillus subtilis</td>
<td>P45709</td>
<td>CCDB</td>
</tr>
<tr>
<td>Agrobacterium tumefaciens</td>
<td>Q7545</td>
<td>VFRG</td>
<td></td>
<td>P37583</td>
<td>CHEY</td>
</tr>
<tr>
<td>Alcaligenes eutrophus</td>
<td>Q44430</td>
<td>NTRR</td>
<td></td>
<td>P14204</td>
<td>CMA1</td>
</tr>
<tr>
<td>Aquifex aeolicus</td>
<td>Q44006</td>
<td>CZCR</td>
<td></td>
<td>O34534</td>
<td>CITT</td>
</tr>
<tr>
<td>Anabaena sp.</td>
<td>P39048</td>
<td>PatA</td>
<td></td>
<td>Q02115</td>
<td>LXR</td>
</tr>
<tr>
<td>Aquifex aeolicus</td>
<td>O87395</td>
<td>OrrA</td>
<td></td>
<td>Q03163</td>
<td>RESD</td>
</tr>
<tr>
<td>Archeoglobus fulgidus</td>
<td>O66551</td>
<td>NtrC1</td>
<td></td>
<td>P06654</td>
<td>SP0A</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>Q9ZWS8</td>
<td>Q9ZWS8</td>
<td></td>
<td>O31432</td>
<td>YBDJ</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>Q9ZWS7</td>
<td>Q9ZWS7</td>
<td></td>
<td>P40759</td>
<td>YCBB</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>Q9ZWSK0</td>
<td>Q9ZWSK0</td>
<td></td>
<td>P42244</td>
<td>YCBL</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>Q9ZWJ9</td>
<td>Q9ZWJ9</td>
<td></td>
<td>P70955</td>
<td>YCCH</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>Q9ZWS9</td>
<td>Q9ZWS9</td>
<td></td>
<td>P94413</td>
<td>YCLJ</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>Q080365</td>
<td>O80365</td>
<td></td>
<td>O31517</td>
<td>YESN</td>
</tr>
<tr>
<td>Azorhizobium</td>
<td>O82798</td>
<td>O82798</td>
<td></td>
<td>O34903</td>
<td>YKOG</td>
</tr>
<tr>
<td>Azorhizobium</td>
<td>O29221</td>
<td>CheB</td>
<td></td>
<td>O34723</td>
<td>YOCG</td>
</tr>
<tr>
<td>Azorhizobium</td>
<td>O28035</td>
<td>O28035</td>
<td></td>
<td>P54443</td>
<td>YRKP</td>
</tr>
<tr>
<td>Azorhizobium</td>
<td>O28381</td>
<td>O28381</td>
<td></td>
<td>O34951</td>
<td>YTSQA</td>
</tr>
<tr>
<td>Azorhizobium</td>
<td>O28799</td>
<td>O28799</td>
<td></td>
<td>O05251</td>
<td>YUMF</td>
</tr>
<tr>
<td>Azorhizobium</td>
<td>O28887</td>
<td>O28887</td>
<td></td>
<td>O06978</td>
<td>YVCP</td>
</tr>
<tr>
<td>Azorhizobium</td>
<td>O29012</td>
<td>O29012</td>
<td></td>
<td>Q9LL43</td>
<td>YVFU</td>
</tr>
<tr>
<td>Azorhizobium</td>
<td>O29199</td>
<td>O29199</td>
<td></td>
<td>O32192</td>
<td>YYQA</td>
</tr>
<tr>
<td>Azorhizobium</td>
<td>O29220</td>
<td>O29220</td>
<td></td>
<td>O32197</td>
<td>YYQC</td>
</tr>
<tr>
<td>Azorhizobium</td>
<td>O29800</td>
<td>O29800</td>
<td></td>
<td>P94504</td>
<td>YVRH</td>
</tr>
<tr>
<td>Azorhizobium</td>
<td>O30252</td>
<td>O30252</td>
<td></td>
<td>P42421</td>
<td>YXDJ</td>
</tr>
<tr>
<td>Azorhizobium</td>
<td>O04849</td>
<td>NtrX</td>
<td></td>
<td>P51844</td>
<td>YXJL</td>
</tr>
<tr>
<td>Azorhizobium</td>
<td>P26487</td>
<td>FixJ</td>
<td></td>
<td>P37478</td>
<td>YYCF</td>
</tr>
<tr>
<td>Bacillus thuringiensis</td>
<td>P45671</td>
<td>NtrC</td>
<td>Bacillus thuringiensis</td>
<td>P52935</td>
<td>Spo0A</td>
</tr>
<tr>
<td>Bacteroides</td>
<td>Q44929</td>
<td>GTCR</td>
<td>Bacteroides</td>
<td>P52942</td>
<td>Spo0F</td>
</tr>
<tr>
<td>Bordetella pertussis</td>
<td>P54662</td>
<td>DEGU</td>
<td>thetaïotalicron</td>
<td>P00936</td>
<td>RTEB</td>
</tr>
<tr>
<td>Bordetella pertussis</td>
<td>P52929</td>
<td>SP0A</td>
<td>Bordetella pertussis</td>
<td>P16574</td>
<td>BvgA</td>
</tr>
</tbody>
</table>

© 2003 Blackwell Publishing Ltd, Molecular Microbiology, 48, 25–51
<table>
<thead>
<tr>
<th>Organism</th>
<th>SWISSPROT entry no.</th>
<th>Name</th>
<th>Organism</th>
<th>SWISSPROT entry no.</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borrelia burgdorferi</td>
<td>O88117</td>
<td>RisA</td>
<td>P07027</td>
<td>UVRY</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1380</td>
<td>O51380</td>
<td>P97172</td>
<td>YEDW</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1501</td>
<td>O51501</td>
<td>P33356</td>
<td>YEHT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1517</td>
<td>O51517</td>
<td>P52076</td>
<td>YGIX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1615</td>
<td>O51615</td>
<td>P77350</td>
<td>YLCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1704</td>
<td>O51704</td>
<td>O86183</td>
<td>NtrC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P72489</td>
<td>RegR</td>
<td>acidaminophilium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P10576</td>
<td>NTRc</td>
<td>O32610</td>
<td>O32610</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q45272</td>
<td>NwsB</td>
<td>Q01473</td>
<td>RCAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS2845</td>
<td>PhoB</td>
<td>Q78425</td>
<td>YC27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P31908</td>
<td>HoxA</td>
<td>O4918</td>
<td>ARCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P23221</td>
<td>FixJ</td>
<td>P76777</td>
<td>CPXR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q92ZH51</td>
<td>CtrA</td>
<td>P44845</td>
<td>NARP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>O67996</td>
<td>BvrR</td>
<td>P45189</td>
<td>PhoB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51380</td>
<td>O51380</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51501</td>
<td>O51501</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51517</td>
<td>O51517</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51615</td>
<td>O51615</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51704</td>
<td>Eubacterium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O88117</td>
<td>RisA</td>
<td>P07027</td>
<td>UVRY</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1380</td>
<td>O51380</td>
<td>P97172</td>
<td>YEDW</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1501</td>
<td>O51501</td>
<td>P33356</td>
<td>YEHT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1517</td>
<td>O51517</td>
<td>P52076</td>
<td>YGIX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1615</td>
<td>O51615</td>
<td>P77350</td>
<td>YLCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1704</td>
<td>O51704</td>
<td>O86183</td>
<td>NtrC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P72489</td>
<td>RegR</td>
<td>acidaminophilium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P10576</td>
<td>NTRc</td>
<td>O32610</td>
<td>O32610</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q45272</td>
<td>NwsB</td>
<td>Q01473</td>
<td>RCAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS2845</td>
<td>PhoB</td>
<td>Q78425</td>
<td>YC27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P31908</td>
<td>HoxA</td>
<td>O4918</td>
<td>ARCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P23221</td>
<td>FixJ</td>
<td>P76777</td>
<td>CPXR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q92ZH51</td>
<td>CtrA</td>
<td>P44845</td>
<td>NARP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>O67996</td>
<td>BvrR</td>
<td>P45189</td>
<td>PhoB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51380</td>
<td>O51380</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51501</td>
<td>O51501</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51517</td>
<td>O51517</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51615</td>
<td>O51615</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51704</td>
<td>Eubacterium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O88117</td>
<td>RisA</td>
<td>P07027</td>
<td>UVRY</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1380</td>
<td>O51380</td>
<td>P97172</td>
<td>YEDW</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1501</td>
<td>O51501</td>
<td>P33356</td>
<td>YEHT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1517</td>
<td>O51517</td>
<td>P52076</td>
<td>YGIX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1615</td>
<td>O51615</td>
<td>P77350</td>
<td>YLCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1704</td>
<td>O51704</td>
<td>O86183</td>
<td>NtrC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P72489</td>
<td>RegR</td>
<td>acidaminophilium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P10576</td>
<td>NTRc</td>
<td>O32610</td>
<td>O32610</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q45272</td>
<td>NwsB</td>
<td>Q01473</td>
<td>RCAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS2845</td>
<td>PhoB</td>
<td>Q78425</td>
<td>YC27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P31908</td>
<td>HoxA</td>
<td>O4918</td>
<td>ARCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P23221</td>
<td>FixJ</td>
<td>P76777</td>
<td>CPXR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q92ZH51</td>
<td>CtrA</td>
<td>P44845</td>
<td>NARP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>O67996</td>
<td>BvrR</td>
<td>P45189</td>
<td>PhoB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51380</td>
<td>O51380</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51501</td>
<td>O51501</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51517</td>
<td>O51517</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51615</td>
<td>O51615</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51704</td>
<td>Eubacterium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O88117</td>
<td>RisA</td>
<td>P07027</td>
<td>UVRY</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1380</td>
<td>O51380</td>
<td>P97172</td>
<td>YEDW</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1501</td>
<td>O51501</td>
<td>P33356</td>
<td>YEHT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1517</td>
<td>O51517</td>
<td>P52076</td>
<td>YGIX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1615</td>
<td>O51615</td>
<td>P77350</td>
<td>YLCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS1704</td>
<td>O51704</td>
<td>O86183</td>
<td>NtrC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P72489</td>
<td>RegR</td>
<td>acidaminophilium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P10576</td>
<td>NTRc</td>
<td>O32610</td>
<td>O32610</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q45272</td>
<td>NwsB</td>
<td>Q01473</td>
<td>RCAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS2845</td>
<td>PhoB</td>
<td>Q78425</td>
<td>YC27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P31908</td>
<td>HoxA</td>
<td>O4918</td>
<td>ARCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P23221</td>
<td>FixJ</td>
<td>P76777</td>
<td>CPXR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q92ZH51</td>
<td>CtrA</td>
<td>P44845</td>
<td>NARP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>O67996</td>
<td>BvrR</td>
<td>P45189</td>
<td>PhoB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51380</td>
<td>O51380</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51501</td>
<td>O51501</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51517</td>
<td>O51517</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51615</td>
<td>O51615</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9XXDD4</td>
<td>O51704</td>
<td>Eubacterium</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Cont.
Table 1. Cont.

<table>
<thead>
<tr>
<th>Organism</th>
<th>SWISSPROT entry no.</th>
<th>Name</th>
<th>Organism</th>
<th>SWISSPROT entry no.</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pseudomonas aureofaciens</strong></td>
<td>Q9X307</td>
<td>GACA</td>
<td><strong>Pseudomonas pseudomallei</strong></td>
<td>O31395</td>
<td>IRLR</td>
</tr>
<tr>
<td><strong>Pseudomonas putida</strong></td>
<td>Q52201</td>
<td>PPRA</td>
<td><strong>Pseudomonas syringae</strong></td>
<td>Q02540</td>
<td>COPR</td>
</tr>
<tr>
<td><strong>Pseudomonas solanacearum</strong></td>
<td>O07846</td>
<td>VI8R</td>
<td><strong>Pseudomonas tolaasi</strong></td>
<td>Q34175</td>
<td>PHEN*</td>
</tr>
<tr>
<td><strong>Rhizobium leguminosarum</strong></td>
<td>Q9X574</td>
<td>CELR</td>
<td><strong>Rhizobium mellii</strong></td>
<td>P55701</td>
<td>Y4XI</td>
</tr>
<tr>
<td><strong>Rhodobacter sphaeroides</strong></td>
<td>O33553</td>
<td>CHER*</td>
<td><strong>Rhodobacter sulfidophilus</strong></td>
<td>O82868</td>
<td>REGA</td>
</tr>
<tr>
<td><strong>Rhodococcus erythropolis</strong></td>
<td>O9274</td>
<td>ERSR</td>
<td><strong>Rickettsia prowazekii</strong></td>
<td>O05971</td>
<td>OmpR</td>
</tr>
<tr>
<td><strong>Saccaromyces cerevisiae</strong></td>
<td>O83302</td>
<td>CopR</td>
<td><strong>Salmoneilla dublin</strong></td>
<td>B34096</td>
<td>BASR</td>
</tr>
<tr>
<td><strong>Salmoneilla typhimurium</strong></td>
<td>P04042</td>
<td>CHEB</td>
<td><strong>Salmonella enterica</strong></td>
<td>P06657</td>
<td>CHEY</td>
</tr>
<tr>
<td><strong>Shigella dysenteriae</strong></td>
<td>P22104</td>
<td>TCTD</td>
<td><strong>Shigella flexneri</strong></td>
<td>P45607</td>
<td>PhoB</td>
</tr>
<tr>
<td><strong>Shigella flexneri</strong></td>
<td>P26319</td>
<td>FIMZ</td>
<td><strong>Staphylococcus aureus</strong></td>
<td>P96456</td>
<td>LYSTS</td>
</tr>
<tr>
<td><strong>Streptomyces coelicolor</strong></td>
<td>O60496</td>
<td>Streptomyces</td>
<td><strong>Streptomyces griseus</strong></td>
<td>O92417</td>
<td>YC45</td>
</tr>
<tr>
<td><strong>Streptomyces hygroscopicus</strong></td>
<td>Q54292</td>
<td>Streptomyces</td>
<td><strong>Streptomyces lividans</strong></td>
<td>P72471</td>
<td>P72471</td>
</tr>
<tr>
<td><strong>Streptomyces lividans</strong></td>
<td>Q54292</td>
<td>Streptomyces</td>
<td><strong>Streptomyces pactum</strong></td>
<td>Q08421</td>
<td>Q08421</td>
</tr>
<tr>
<td><strong>Streptomyces peucetius</strong></td>
<td>Q52201</td>
<td>STP3</td>
<td><strong>Streptomyces violaceonear</strong></td>
<td>Q05243</td>
<td>SPHR</td>
</tr>
<tr>
<td><strong>Streptomyces violaceus</strong></td>
<td>Q54039</td>
<td>STP4</td>
<td><strong>Synechococcus sp.</strong></td>
<td>O68523</td>
<td>NBLR</td>
</tr>
<tr>
<td><strong>Synechocystis sp.</strong></td>
<td>P26384</td>
<td>PILG</td>
<td><strong>Thermotoga maritima</strong></td>
<td>P23501</td>
<td>PILH</td>
</tr>
<tr>
<td><strong>Thermotoga maritima</strong></td>
<td>Q00934</td>
<td>PHOB</td>
<td><strong>Thiocystis violacea</strong></td>
<td>Q56127</td>
<td>RCSB</td>
</tr>
<tr>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio harveyi</strong></td>
<td>Q56128</td>
<td>RCSB°</td>
</tr>
<tr>
<td><strong>Vibrio harveyi</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thiocystis violacea</strong></td>
<td>P04042</td>
<td>CHEB</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thermotoga maritima</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio harveyi</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thiocystis violacea</strong></td>
<td>P04042</td>
<td>CHEB</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thermotoga maritima</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio harveyi</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thiocystis violacea</strong></td>
<td>P04042</td>
<td>CHEB</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thermotoga maritima</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio harveyi</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thiocystis violacea</strong></td>
<td>P04042</td>
<td>CHEB</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thermotoga maritima</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio harveyi</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thiocystis violacea</strong></td>
<td>P04042</td>
<td>CHEB</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thermotoga maritima</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio harveyi</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thiocystis violacea</strong></td>
<td>P04042</td>
<td>CHEB</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thermotoga maritima</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio harveyi</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thiocystis violacea</strong></td>
<td>P04042</td>
<td>CHEB</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thermotoga maritima</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio harveyi</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thiocystis violacea</strong></td>
<td>P04042</td>
<td>CHEB</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thermotoga maritima</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio harveyi</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thiocystis violacea</strong></td>
<td>P04042</td>
<td>CHEB</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thermotoga maritima</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio harveyi</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thiocystis violacea</strong></td>
<td>P04042</td>
<td>CHEB</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Thermotoga maritima</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio harveyi</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
<tr>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
<td><strong>Vibrio cholerae</strong></td>
<td>Q56128</td>
<td>RCSC°</td>
</tr>
</tbody>
</table>

---

a. These response regulators are fused to sensors as part of a composite TCS consisting of a sensor-regulator protein, a sensor-regulator-sensor protein, or a sensor-regulator-sensor-regulator protein.

© 2003 Blackwell Publishing Ltd, Molecular Microbiology, 48, 25–51
We have made predictions regarding the monofunctional or bifunctional character of the sensor for many TCS (Table 2). In most cases, TCS have been identified by sequence similarity, and definitive biochemical evidence regarding the monofunctional or bifunctional character of their sensor is lacking. Our predicted molecular models for these sensors reveal a consistent difference in the folding of their catalytic domain: the fold resembles either that of the monofunctional template or that of the bifunctional template.

We also have tested our predictions with the TCS for which biochemical and genetic evidence regarding the monofunctional or bifunctional character of their sensor is available and homology modelling was possible (Table 3). The sensor protein has an ATP-binding domain that contains a small sequence of amino acids giving rise to a 3D structure known as the ATP lid. This structure is highly organized and its position shifts, enclosing the nucleotide or releasing it from its binding site. The secondary structure of the ATP lid is mostly $\alpha$-helical. If the sensor is known to be bifunctional, we find that its ATP lid resembles that of EnvZ (Fig. 3B). Approximately in the middle of the amino acid sequence, the $\alpha$-helix fold is interrupted by a small non-$\alpha$-helical T-loop. If the sensor protein is known to be monofunctional, we find a different folding of the ATP lid in which the three-dimensional structure resembles the ATP-binding domain of CheA (Fig. 4B). In this case, the $\alpha$-helical structure of the ATP lid is interrupted by two

<table>
<thead>
<tr>
<th>Organism</th>
<th>SWISSPROT entry no.</th>
<th>Name</th>
<th>Prediction</th>
<th>Organism</th>
<th>SWISSPROT entry no.</th>
<th>Name</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agrobacterium tumefaciens</td>
<td>O07737</td>
<td>ChvG</td>
<td>Bi</td>
<td>Calothrix vugleri</td>
<td>O52937</td>
<td>O52937</td>
<td>Bi</td>
</tr>
<tr>
<td><strong>Aquifex aeolicus</strong></td>
<td>O6656</td>
<td>O6656</td>
<td>Bi</td>
<td>Candida albicans</td>
<td>O42696</td>
<td>Can1</td>
<td>Bi</td>
</tr>
<tr>
<td><strong>Arabidopsis thaliana</strong></td>
<td>P49333</td>
<td>ETR1</td>
<td>Lid $^d$</td>
<td>Caulobacter</td>
<td>O98922</td>
<td>HK1</td>
<td>Lid $^d$</td>
</tr>
<tr>
<td><strong>Archeoglobus fulgidus</strong></td>
<td>O22267</td>
<td>O22267</td>
<td>Lid $^d$</td>
<td>crescentus</td>
<td>Q9X688</td>
<td>CdxA</td>
<td>Lid $^d$</td>
</tr>
<tr>
<td><strong>Date: Table 2. Putative sensor proteins that yielded a predicted structure by homology modelling techniques.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 2003 Blackwell Publishing Ltd, Molecular Microbiology, 48, 25–51
### Table 2. Cont.

<table>
<thead>
<tr>
<th>Organism</th>
<th>SWISSPROT entry no.</th>
<th>Name</th>
<th>Prediction</th>
<th>Organism</th>
<th>SWISSPROT entry no.</th>
<th>Name</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Halobacterium salinarium</td>
<td>Q48297</td>
<td>CheA like</td>
<td>Mono</td>
<td>Pseudomonas aeruginosa</td>
<td>O34206</td>
<td>KINB</td>
<td>Bi*</td>
</tr>
<tr>
<td>Helicobacter pylori</td>
<td>Q92KE9 Q92KE9</td>
<td>Mono</td>
<td>Pseudomonas pylorimimica</td>
<td>O68596</td>
<td>O68596</td>
<td>Lid*</td>
<td></td>
</tr>
<tr>
<td>Lactobacillus sake</td>
<td>Q82929 Q82929</td>
<td>Bi</td>
<td>Pseudomonas putida</td>
<td>O07831</td>
<td>BzP</td>
<td>Lid*</td>
<td></td>
</tr>
<tr>
<td>Lactococcus lactis</td>
<td>Q48675 NisK</td>
<td>Bi</td>
<td>Pseudomonas solanacearum</td>
<td>O07845</td>
<td>O07845</td>
<td>Bi*</td>
<td></td>
</tr>
<tr>
<td>Listeria monocytogenes</td>
<td>Q48768 CheA</td>
<td>Mono</td>
<td>Pseudomonas tolaasi</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Methanobacterium thermoautotrophicum</td>
<td>MT901 MT901</td>
<td>Lидi</td>
<td>Rhodobacter capsulatus</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Mycobacterium leprae</td>
<td>P54883 YV16</td>
<td>Bi</td>
<td>Rhodobacter capsulatus</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Mycobacterium tuberculosis</td>
<td>PT32572 KD02</td>
<td>Bi</td>
<td>Rhodobacter sphaeroides</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Neisseria meningitides</td>
<td>Q9K12 Nix</td>
<td>Lидi</td>
<td>Rhodobacter sphaeroides</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Neurospora crassa</td>
<td>O93851 NIK1</td>
<td>Lидi</td>
<td>Serratia plymuthica</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Nostoc punctiforme</td>
<td>Q92693 Q92693</td>
<td>Lидi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Proteus mirabilis</td>
<td>O85662 RsC</td>
<td>Lидi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Pseudomonas aeruginosa</td>
<td>Q51453 FLES</td>
<td>Lидi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus aureus</td>
<td>Q9XCM6 YYCG</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Streptomyces hygroscopicus</td>
<td>Q54955 Q54955</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P72792 SQ792</td>
<td>Lидi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P43182 OS1P</td>
<td>Lидi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37321 P37321</td>
<td>Lиди</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37365 P37365</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37385 P37385</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37386 P37386</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37387 P37387</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37388 P37388</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37389 P37389</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37390 P37390</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37391 P37391</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37392 P37392</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37393 P37393</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37394 P37394</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37395 P37395</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37396 P37396</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37397 P37397</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37398 P37398</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37399 P37399</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37400 P37400</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37401 P37401</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37402 P37402</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37403 P37403</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37404 P37404</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37405 P37405</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37406 P37406</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37407 P37407</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37408 P37408</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37409 P37409</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37410 P37410</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>Synechocystis sp.</td>
<td>P37411 P37411</td>
<td>Bi</td>
<td>Synechocystis sp.</td>
<td>O25153</td>
<td>25153</td>
<td>Mono</td>
<td></td>
</tr>
</tbody>
</table>

© 2003 Blackwell Publishing Ltd, Molecular Microbiology, 48, 25–51

Alternative regulatory designs for two-component systems
small non- \( \alpha \)-helical loops, with the \( \alpha \)-helix stretch in between the two loops laying almost perpendicular to the other two \( \alpha \)-helical stretches.

**Qualitative functional differences**

The concentrations of the input signal molecules (\( X_i \), primary input, such as a chemical gradient of nutrients in the CheA case or osmotic pressure changes in the EnvZ case, and \( X_{\text{opt}} \), secondary input, such as acetyl phosphate) and the total concentrations of sensor protein (\( X = X_1 + X_2 \)) and regulator protein (\( X = X_3 + X_4 \)) are determined by external influences that are independent of changes within the system. Thus, these are defined as **independent variables**. By contrast, the concentrations of the sensor proteins (\( X_i \), phosphorylated and \( X_i \), unphosphorylated) and of the regulator proteins (\( X_i \), phosphorylated and \( X_i \), unphosphorylated) are determined by the values of the independent variables and by the system’s internal dynamical behaviour. These variables are defined as **dependent state variables** in our models.

Protein synthesis and degradation occur on a time-scale that is much slower than that of phosphorylation, phosphotransfer and dephosphorylation in the TCS modules. Thus, on the time scale of interest here, one can ignore gene regulation and consider the total amount of sensor protein (\( S_{\text{Total}} \), \( X \)) and the total amount of regulator protein (\( R_{\text{Total}} \), \( X \)) to be conserved quantities (i.e. \( X_i + X_i = X_i = \text{constant} \) and \( X_i + X_i = X_i = \text{constant} \)).

Thus, when considering changes in the dependent variables, in this and the following section, we need only emphasize the phosphorylated forms of the sensor protein \( X_i \) and the regulator protein \( X_i \). The corresponding changes in the unphosphorylated forms, \( X_i \) and \( X_i \), have the same magnitude but opposite sign. If for any reason, one of the \( X \) species in one of the conserved pairs increases by some amount, then the other \( X \) species of the pair must decrease by exactly the same amount. For example, amplification of signals at the level of the unphosphorylated proteins is equal to the negative of that at the level of the phosphorylated proteins. This amplification can be measured by the logarithmic gain \( L(X_i, X) \) (see **Experimental procedures** section), which is defined as the percentage change in a dependent variable \( X_i \) in

<table>
<thead>
<tr>
<th>Sensor name</th>
<th>Functionality</th>
<th>Predicted</th>
<th>Observed</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CheA</td>
<td>Mono</td>
<td>Mono</td>
<td>Ninfa et al. (1991)</td>
<td></td>
</tr>
<tr>
<td>(E. coli)</td>
<td>Bi</td>
<td>Bi</td>
<td>Kanamaru et al. (1989)</td>
<td></td>
</tr>
<tr>
<td>EmZ</td>
<td>Bi</td>
<td>Bi</td>
<td>Jung et al. (1997)</td>
<td></td>
</tr>
<tr>
<td>(E. coli)</td>
<td>Bi</td>
<td>Bi</td>
<td>Wright et al. (1993)</td>
<td></td>
</tr>
<tr>
<td>KdpD</td>
<td>Bi</td>
<td>Bi</td>
<td>Shi et al. (1999)</td>
<td></td>
</tr>
<tr>
<td>(E. faecalis)</td>
<td>Bi</td>
<td>Bi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VannSA</td>
<td>Bi</td>
<td>Bi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(B. subtilis)</td>
<td>Bi</td>
<td>Bi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PhoR</td>
<td>Bi</td>
<td>Bi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(B. subtilis)</td>
<td>Bi</td>
<td>Bi</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Inferred from gene expression results.
b. Shows an increase in the rate of dephosphorylation only for the soluble domain of PhoR in the presence of ADP or ATP.

c. These are partially modelled structures that closely resemble the ATP-binding domain of sensor EnvZ. The resulting model has the ATP lid and a portion, but not all, of the remaining ATP binding domain of the sensor EnvZ.
d. These are partially modelled structures that only slightly resemble the ATP-binding domain of sensor EnvZ. The resulting model has an ATP lid similar to that of EnvZ, but lacks the remainder of the EnvZ-like ATP binding domain.

e. This is a case in which the sensor protein, NRI, is monofunctional unless it is bound with a second protein, PII, in which case it becomes bifunctional.
response to a one per cent change in an independent variable \( X_i \). Thus, \( \Delta L(X_i X_j) = -L(X_i X_j) \) and \( L(X_i X_j) = -L(X_i X_j) \), where \( X_i \) is any independent variable of the model. Similarly, the parameter sensitivities in the levels of the unphosphorylated proteins in response to parameter fluctuations are equal to the negative of those in the levels of the phosphorylated proteins. These parameter sensitivities can be measured by the expression \( S(X_p, X_j) \) (see Experimental procedures section), which is defined as the percentage change in a dependent variable \( X_j \) in response to a one per cent change in a parameter \( p_i \). Thus, \( \frac{\Delta S(X_p, X_j)}{\Delta p_i} \) and \( S(X_i, X_j) = -S(X_i, p_j) \) where \( p_i \) is any parameter of the model.

The kinetic behaviour of the models in Fig. 1 can be described by a system of differential equations [Equations (1) (2) and (2')], as outlined in the Experimental procedures section. Solving these equations allows us to quantify and to compare the systemic properties of the alternative designs shown in Fig. 1 based on the following functional considerations.

A signal transduction cascade should have a set of large logarithmic gains to amplify physiological signals and another set of small logarithmic gains that attenuate pathological noise. The cascade should be robust, i.e. it should function reproducibly despite perturbations in the values of the parameters that define the structure of the system. This is, by definition, equivalent to saying that the parameter sensitivities should, in general, be as low as possible. The steady state of the system should be stable and have a sufficient margin of stability, such that it will not become unstable when subjected to random fluctuations in the parameters of the system. If the margin of stability is small, a small change in a parameter of the system (e.g. ionic strength of the medium) may destroy the possibility of a stable steady state and make the system dysfunctional. Finally, the system should respond quickly to changes in its environment because otherwise the system is unlikely to be competitive in rapidly changing environments.

These properties, for each of the alternative models, are quantified and then compared by taking the ratio of the value for a property in the reference system (bifunctional sensor) to the corresponding value in the alternative system (monofunctional sensor). As we have analytical expressions for the steady-state properties, we can determine in some cases whether the ratio is always equal to one, less than one, or greater than one, independent of parameter values. With this approach, we have obtained the following qualitative results for steady-state concentrations and the amplification factors.

The concentrations (and rates of change) of the corresponding state variables in the two models can always be the same (see Calculating the constraints for external equivalence). Changes in the secondary signal \( X_i \) are amplified less in each of the corresponding state variables \([i.e. each L(X_i X_j) for i = 1, 2, 3 and 4]\) with the bifunctional design. Similarly, these changes are amplified less in the flux through each of the sensor pools \([i.e. L(V_s X_j) for i = 1 and 3]\) with the bifunctional design. This is shown in Table 4 by the ratio \( E \), which is always less than 1. On the other hand, amplification of the phosphorylated regulator \( X_j \) in response to a percentage change in the primary input signal \( X_i \) \([i.e. L(V_s X_j)\) is always greater with the bifunctional design. This is shown in Table 4 by the ratio \( B \), which is always greater than 1.0. In all other cases, the differences between the amplification factors of the alternative designs are dependent upon the specific numerical values of the parameters and can not be determined analytically.

Quantitative functional differences

Although some functional differences between TCS with the alternative sensor designs are analytically indeterminate, numerical comparisons can be used to determine quantitative differences in function, and thus to establish statistical tendencies in the differences. With this approach we have obtained the following quantitative results for signal amplification, robustness, margin of stability, response time, and phosphorylation/dephosphorylation ratio.

Numerical results for signal amplification of concentrations are shown in Fig. 5 and described in detail in the next paragraph. The data are represented as Density of Ratios plots for moving medians (as described in Experimental procedures). The moving median of the ratio for the bifunctional design (reference) to monofunctional design (alternative), shown on the vertical axis, is a function of the moving median of the amplification for the bifunctional design, shown on the horizontal axis. [In symbolic terms, \(<L(X_i X_j)_{br} / L(X_i X_j)_{br}>\) is plotted as a function of \(<L(X_i X_j)_{br} >\), where \( X_i \) indicates an independent variable, \( X_j \) a dependent variable and the angular brackets indicate averages.

The pattern of responses exhibited by the sensor protein is as follows. Overall, values for the amplification of the phosphorylated sensor signal \( X_i \) in response to a percentage change in the primary input signal \( X_j \) \([L(X_i X_j)]\) are similar in both designs. This is shown by the curve in Fig. 5A, which remains about 1.0. Values for the amplification of the phosphorylated sensor signal \( X_i \) in response to a percentage change in the secondary input signal \( X_j \) \([L(X_i X_j)]\) are also smaller in the bifunctional design (the curve remains below 1.0 in Fig. 5B). On the other hand, values for the amplification of the phosphorylated sensor signal \( X_i \) in response to a percentage change in either the total concentration of sensor protein \( X_j \) \([L(X_i X_j)]\) or the total concentration of regulator protein \( X_j \) \([L(X_i X_j)]\) are similar for the two designs (curves...
Table 4. Qualitative ratios of corresponding logarithmic gains for the reference system relative to the alternative system.

<table>
<thead>
<tr>
<th>Systemic property</th>
<th>X1</th>
<th>X2</th>
<th>V1</th>
<th>V2</th>
</tr>
</thead>
<tbody>
<tr>
<td>L(•, X5)</td>
<td>A</td>
<td>B &gt; 1</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>L(•, X6)</td>
<td>E &lt; 1</td>
<td>E &lt; 1</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>L(•, X7)</td>
<td>G</td>
<td>H</td>
<td>J</td>
<td>K</td>
</tr>
<tr>
<td>L(•, X8)</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>P</td>
</tr>
</tbody>
</table>

\[
\text{Ratio} = \frac{L(•, X)}{L(•, \text{Mono})}
\]

\[
A = 1 + \frac{-E_2 h_1 h_2}{E_3 (E_3 h_1 - h_2) + E_4 (E_3 h_1 - h_2) + h_1 h_2} < 1
\]

\[
B = 1 + \frac{E_2 h_1 (E_3 h_1 - h_2)}{E_3 (E_3 h_1 - h_2) + E_4 (E_3 h_1 - h_2)} > 1
\]

\[
C = 1 + \left( \frac{E_5 (E_5 h_1 - h_2) + E_6 (E_5 h_1 - h_2) + h_1 h_2}{E_5 (E_5 h_1 - h_2) + E_6 (E_5 h_1 - h_2)} \right) < 1
\]

\[
D = 1 + \frac{E_2 (E_3 h_1 + E_4 (E_3 h_1 - h_2))}{E_3 (E_3 h_1 + E_4 (E_3 h_1 - h_2)) - E_2 (E_3 h_1 - h_2)} < 1
\]

\[
E = 1 + \left( \frac{E_2 h_1}{E_3 (E_3 h_1 - h_2) + E_4 (E_3 h_1 - h_2)} \right) < 1
\]

\[
F = 1 + \left( \frac{E_2 h_1 + E_4 (E_3 h_1 - h_2)}{E_3 (E_3 h_1 - h_2) + E_4 (E_3 h_1 - h_2)} \right) > 1
\]

\[
G = 1 + \left( \frac{E_2 (E_3 h_1 + E_4 (E_3 h_1 - h_2))}{E_3 (E_3 h_1 + E_4 (E_3 h_1 - h_2))} \right) < 1
\]

\[
H = 1 + \left( \frac{E_2 (E_3 h_1 + E_4 (E_3 h_1 - h_2))}{E_3 (E_3 h_1 + E_4 (E_3 h_1 - h_2))} \right) > 1
\]

\[
J = 1 + \left( \frac{E_2 (E_3 h_1 + E_4 (E_3 h_1 - h_2))}{E_3 (E_3 h_1 + E_4 (E_3 h_1 - h_2))} \right) < 1
\]

\[
K = 1 + \left( \frac{E_2 (E_3 h_1 + E_4 (E_3 h_1 - h_2))}{E_3 (E_3 h_1 + E_4 (E_3 h_1 - h_2))} \right) > 1
\]

\[
L = 1 + \left( \frac{E_2 (E_3 h_1 + E_4 (E_3 h_1 - h_2))}{E_3 (E_3 h_1 + E_4 (E_3 h_1 - h_2))} \right) < 1
\]

\[
M = 1 + \left( \frac{E_2 (E_3 h_1 + E_4 (E_3 h_1 - h_2))}{E_3 (E_3 h_1 + E_4 (E_3 h_1 - h_2))} \right) < 1
\]

\[
N = 1 + \left( \frac{E_2 (E_3 h_1 + E_4 (E_3 h_1 - h_2))}{E_3 (E_3 h_1 + E_4 (E_3 h_1 - h_2))} \right) > 1
\]

\[
P = 1 + \left( \frac{E_2 (E_3 h_1 + E_4 (E_3 h_1 - h_2))}{E_3 (E_3 h_1 + E_4 (E_3 h_1 - h_2))} \right) > 1
\]
Alternative regulatory designs for two-component systems

Fig. 5. Comparison of signal amplification factors for concentrations in two-component systems with either a bifunctional or a monofunctional sensor. The logarithmic gains are compared in a density of ratios plot using moving medians (see Numerical analysis). On the x-axis are average values of a given logarithmic gain for the bifunctional design. On the y-axis are average values for the ratio of that logarithmic gain in the bifunctional design (Fig. 1A) over the corresponding logarithmic gain in the monofunctional design (Fig. 1B).

A. Logarithmic gain in phosphorylated sensor protein \( X_1 \) with respect to changes in the primary input signal \( X_0 \).
B. Logarithmic gain in \( X_0 \) with respect to changes in the secondary input signal \( X_2 \).
C. Logarithmic gain in \( X_1 \) with respect to changes in the total amount of sensor, \( X_0 \).
D. Logarithmic gain in \( X_1 \) with respect to changes in the total amount of regulator, \( X_0 \).
E. Logarithmic gain in phosphorylated regulator protein \( X_2 \) with respect to changes in the primary input signal \( X_0 \).
F. Logarithmic gain in \( X_2 \) with respect to changes in the secondary input signal \( X_0 \).
G. Logarithmic gain in \( X_2 \) with respect to changes in the total amount of sensor, \( X_0 \).
H. Logarithmic gain in \( X_2 \) with respect to changes in the total amount of regulator, \( X_0 \).

The regulator protein exhibits a different pattern of responses. Values for the amplification of the phosphorylated regulator signal \( X_2 \) in response to a percentage change in the primary input signal \( X_0 \) \([L(X_0,X_0)]\) are greater for the bifunctional design as can be seen by the curve in Fig. 5E, which is always above 1. For low median values of the gain \( L(X_0,X_0) \) in the bifunctional design, the differences in gain can be as large as 100%. For high median values, the differences are around 30%, with the gain in the bifunctional design being higher. This means that, for example, the bifunctional design will provide for larger changes in gene expression than would a monofunctional design for the same amount of osmotic pressure change in the EnvZ/OmpR system. Values for the amplification of the phosphorylated regulator signal \( X_2 \) in response to a percentage change in the secondary input signal \( X_0 \) \([L(X_2,X_0)]\) are smaller in the bifunctional design (Fig. 5F).

On the other hand, values for the amplification of the phosphorylated regulator signal \( X_2 \) in response to a percentage change in the total concentration of sensor protein \( X_1 \) \([L(X_0,X_1)]\) can be larger in either alternative depending on the parameter values (Fig. 5G). If the logarithmic gain \( L(X_0,X_1) \) in the bifunctional design is negative, then the amplification in the bifunctional design is smaller (in absolute value) as can be seen in Fig. 5G. If the logarithmic gain in the bifunctional design is positive, then the amplification in the monofunctional design is smaller. Values for the amplification of the phosphorylated regulator signal \( X_2 \) in response to a percentage change in the total concentration of regulator protein \( X_0 \) \([L(X_0,X_0)]\) also can be larger in either alternative depending on the parameter values (Fig. 5H). If the logarithmic gain \( L(X_0,X_0) \) in the bifunctional design is negative, then the amplification in the bifunctional design is larger (in absolute value); if the logarithmic gain in the bifunctional design is positive, then the amplification in the monofunctional design is larger. Even though it seems that this shift occurs at values of about −0.3 for \( L(X_0,X_0) \) in this case, the actual values are about zero. Because of the moving averaging technique there are residual ratios that keep the median above 1.0. When the values for \( L(X_0,X_0) \) change sign, the curves tend to exhibit a dip, which is a consequence of the moving average technique when positive and negative values are being averaged.

Numerical results for signal amplification in flux are shown in Fig. 6. The pattern of responses exhibited by flux through the pools of sensor protein is as follows. Values for the amplification of the flux \( V \) in response to a percentage change in the primary input signal \( X_0 \) \([L(V,X_0)]\) can be larger in either alternative depending on the parameter values (Fig. 6A). If the logarithmic gain \( L(V,X_0) \) remain very close to the y-axis value of 1.0 in Fig. 5C and D), except when signal amplification values are close to zero.

© 2003 Blackwell Publishing Ltd, Molecular Microbiology, 48, 25–51
Fig. 6. Comparison of signal amplification factors for fluxes in two-component systems with either a bifunctional or a monofunctional sensor. The logarithmic gains are compared in a density of ratios plot using moving medians (see Numerical analysis). On the y-axis are average values of a given logarithmic gain in the bifunctional design. On the x-axis are average values for the ratio of that logarithmic gain in the bifunctional design (Fig. 1A) over the corresponding logarithmic gain in the monofunctional design (Fig. 1B).

A. Logarithmic gain in flux through the pool of phosphorylated sensor protein $V_1$ with respect to changes in the primary input signal $X_5$.
B. Logarithmic gain in $V_1$ with respect to changes in the secondary input signal $X_4$.
C. Logarithmic gain in $V_1$ with respect to changes in the total amount of sensor, $X_5$.
D. Logarithmic gain in $V_1$ with respect to changes in the total amount of regulator, $X_6$.
E. Logarithmic gain in flux through the pool of phosphorylated regulator protein $V_2$ with respect to changes in the primary input signal $X_5$.
F. Logarithmic gain in $V_2$ with respect to changes in the secondary input signal $X_6$.
G. Logarithmic gain in $V_2$ with respect to changes in the total amount of sensor, $X_5$.
H. Logarithmic gain in $V_2$ with respect to changes in the total amount of regulator, $X_6$.

in the bifunctional design is negative, then the amplification in the bifunctional design is larger (in absolute value); if the logarithmic gain in the bifunctional design is positive, then the amplification in the monofunctional design is larger. Values for the amplification of the flux $V_1$ in response to a percentage change in the secondary input signal $X_6$ $[L(V_1, X_6)]$ are smaller in the bifunctional design (Fig. 6B). Values for the amplification of the flux $V_1$ in response to a percentage change in the total concentration of sensor protein $X_5$ $[L(V_1, X_5)]$ in the monofunctional design are always larger in absolute value (Fig. 6C). Values for the amplification of the flux $V_1$ in response to a percentage change in the total concentration of regulator protein $X_6$ $[L(V_1, X_6)]$ can be larger in either alternative depending on the parameter values (Fig. 6D). If the logarithmic gain $L(V_1, X_6)$ in the bifunctional design is negative, then the amplification in the monofunctional design is larger (in absolute value); if the logarithmic gain in the bifunctional design is positive, then the amplification in the monofunctional design is larger.

The pattern of responses exhibited by the flux through the pools of regulator protein is as follows. Values for the amplification of the flux $V_2$ in response to a percentage change in the primary input signal $X_5$ $[L(V_2, X_5)]$ can be larger in either alternative depending on the parameter values (Fig. 6E). If the logarithmic gain $L(V_2, X_5)$ in the bifunctional design is negative, then the amplification in the bifunctional design is larger (in absolute value); if the logarithmic gain in the bifunctional design is positive, then the amplification in the monofunctional design is larger. Even though it seems that this shift occurs at values of about $-0.8$ for $L(V_2, X_5)$ in this case, the actual values are about zero. Because of the moving averaging technique there are residual ratios that keep the median above 1.0. Values for the amplification of the flux $V_2$ in response to a percentage change in the secondary input signal $X_6$ $[L(V_2, X_6)]$ can be larger in either alternative depending on the parameter values (Fig. 6F). If the logarithmic gain $L(V_2, X_6)$ in the bifunctional design is negative, then the amplification in the monofunctional design is larger (in absolute value); if the logarithmic gain in the bifunctional design is positive, then the amplification in the monofunctional design is larger. Values for the amplification of the flux $V_2$ in response to a percentage change in the total concentration of sensor protein $X_5$ $[L(V_2, X_5)]$ is on average larger in the monofunctional design (Fig. 6G). Values for the amplification of the flux $V_2$ in response to a percentage change in the total concentration of regulator protein $X_6$ $[L(V_2, X_6)]$ can be larger in either alternative depending on the parameter values (Fig. 6H). If the logarithmic gain $L(V_2, X_6)$ in the bifunctional design is negative, then the amplification in the monofunctional design is larger (in absolute value); if the logarithmic gain in the bifunctional design is positive, then the amplification in the bifunctional design is larger.
design is larger. Even though it seems that this shift occurs at values of about 0.5 for $L(V_2, X_3)$ in this case, the actual values are about zero. Again, because of the moving averaging technique there are residual ratios that keep the median above 1.0. When the values for $L(V, X)$ change sign, the curves tend to exhibit a dip, which is a consequence of the moving average technique when positive and negative values are being averaged. The dip for the logarithmic gains in flux tends to be more pronounced than for the logarithmic gains in concentration.

Numerical results for robustness of the alternative designs are shown in Fig. 7. The data for aggregate parameter sensitivities are represented as Density of Ratios plots for moving medians. The moving median of the ratio for the bifunctional design (reference) to monofunctional design (alternative) is on the horizontal axis, and the moving median of aggregate parameter sensitivity for the bifunctional design is on the vertical axis. [In symbolic terms, $\langle S(X)_{\text{Bi}}/S(X)_{\text{Mono}} \rangle$ as a function of $\langle S(X)_{\text{Bi}} \rangle$, or $\langle S(V)_{\text{Bi}}/S(V)_{\text{Mono}} \rangle$ as a function of $\langle S(V)_{\text{Bi}} \rangle$, where $X$ indicates a dependent concentration, $V$ a dependent flux and the angular brackets indicate averages.] On average, the aggregate sensitivities of the sensor signals (Fig. 7A and C) are less in the bifunctional design than in the monofunctional design when the values for the parameter sensitivities are low, whereas the aggregate sensitivities of the regulator signals (Fig. 7B and D) are greater in the bifunctional design than in the monofunctional design under these conditions. The aggregate sensitivities of all the signals are lower in the monofunctional design when the values for the parameter sensitivities are high, which is the less physiologically relevant case. However, in all cases, the average differences in corresponding sensitivities between the monofunctional and the bifunctional system are very small. In the case of the sensor signals, this difference is negligible and the median ratio is for all practical purposes 1.0.

Numerical results for the stability margins of the alternative designs are shown in Fig. 8. These magnitudes, which correspond to the two critical Routh Criteria for stability, provide a measurement for the amount of perturbation that the system will tolerate before the steady state becomes unstable (see the Experimental procedure section). The bifunctional design has a larger margin of stability with respect to the first of the critical Routh criteria when this margin is small, which is when this margin is most important. When this margin becomes large, and its value becomes less important, the two designs have essentially the same value (Fig. 8A). The bifunctional design also has a larger margin of stability with respect to the second of the critical Routh criteria (Fig. 8B).

![Fig. 7. Comparison of robustness for two-component systems with either a bifunctional or a monofunctional sensor. The aggregate parameter sensitivities (see Steady-state solution and key systemic properties) are compared in a density of ratios plot using moving medians. On the x-axis are average values of a given aggregate sensitivity in the bifunctional design. On the y-axis are average values for the ratio of that aggregate sensitivity in the bifunctional design (Fig. 1A) over the corresponding aggregate sensitivity in the monofunctional design (Fig. 1B).](image)
Numerical results for the temporal responsiveness of the alternative designs are shown in Fig. 9. The response time, $t$, is defined as time required for the return to a steady state following a perturbation (see the Experimental procedures section). The data are represented as Density of Ratios plots for the raw data (Fig. 9A) and for moving medians (Fig. 9B). The ratio for the bifunctional design (reference) to monofunctional design (alternative) is on the vertical axis, and the response time for the bifunctional design is on the horizontal axis. [In symbolic terms, $<t_{Bi}/t_{mono}>$ as a function of $<t_{Bi}>$, where $t$ is the response time and the angular brackets indicate averages.] On average, the response time is slightly less for the bifunctional design when response times are small, but the differences between designs become insignificant as the response times become larger.

In this work we allowed the sensor and regulator each to have steady-state operating levels of phosphorylation between 0% and 100%. In practice, the random values for the ratios $f_{31} = -X_{10}/X_{30}$ and $f_{42} = -X_{20}/X_{40}$ in our ensemble range between $-0.0000001$ and $-999$. The amplification properties of the TCS with a bifunctional design are not influenced to any large extent by the steady-state operating value for phosphorylation of either the sensor or the regulator (data not shown), although more significant changes can be seen in some cases when the operating value for phosphorylation drops to near 0%.

Discussion
In this work we examined alternative designs for the sensor proteins of prototype TCS. A bifunctional sensor is characterized by two functions: (i) when phosphorylated, the sensor transfers its phosphate group to the response regulator; (ii) when unphosphorylated the sensor increases the dephosphorylation rate of the response regulator. A monofunctional sensor has only the first of these functions. Our results have identified both structural and functional attributes of these alternative designs.

Structural differences
The results of our homology modelling show a highly conserved structural feature (‘ATP lid’) that appears to be distinctive for each of the alternative designs (Figs 3 and 4). Whether or not the characteristic ATP lid is responsible...
for the bifunctionality of a sensor is unknown, but recent experiments suggest that this part of the sensor has a role in enhancing the rate of dephosphorylation of the response regulator for EnvZ/OmpR (Zhu et al., 2000) and for NRII-II/NRI (Pioszak and Ninfa, 2003). An experiment that exchanges the ATP lid of the EnvZ and CheA sensor proteins and assays the resulting proteins for bifunctional behaviour would help to resolve this issue. Whether this structural feature is responsible for the bifunctionality of a sensor or not, our rigorous comparative analysis of physiological function demonstrates that there is a clear basis for selection of monofunctional and bifunctional sensor proteins.

Functional differences

Protein levels for the majority of TCS are regulated on a slow time-scale by mechanisms affecting transcription. Regulation by these mechanisms, which have been studied elsewhere (Hlavacek and Savageau, 1995, 1996, 1997), is beyond the scope of this article, which focuses on the more rapid time scale of regulation within the TCS. Nevertheless, the results for the logarithmic gains in concentrations (Fig. 5C, D, G and H) and fluxes (Fig. 6C, D, G and H) with respect to changes in the total concentration of sensor ($X_1$) and regulator ($X_2$) provide some insight regarding the influence of changing protein levels.

The mathematically controlled comparisons in our study show that monofunctional and bifunctional designs for signalling within TCS differ on the basis of several criteria for functional effectiveness. The model with a bifunctional sensor has higher signal amplification in response to changes in the primary signal (Fig. 5E), $X_1$, as represented by $L(X_1, X_2)$. It has lower signal amplification in response to changes in the secondary signal (Fig. 5F), $X_2$, as represented by $L(X_2, X_1)$. If $X_2$ is considered to be the physiological signal for a TCS, then this implies that the model with the bifunctional sensor is more effective in responding to this signal. Robustness, except in the case of $V_2$, tends to be similar when the values for the parameter sensitivities are low (Fig. 7). Although, robustness for all the variables is greater in the bifunctional design when the values for the parameter sensitivities are high, systems with high parameter sensitivities are less likely to be biologically significant. The margins of stability for the steady state, as measured by both Routh criteria, are larger for the bifunctional design (Fig. 8). Response times are similar for the alternative designs (Fig. 9). These functional differences have implications for cross-talk among TCS.

Cross-talk to and from a TCS

The specificity of sensors and regulators is not absolute. Regulator proteins can respond to signals other than those transmitted by their cognate sensor, and sensor proteins can transmit signals to destinations other than their cognate regulator. In some cases, this cross-talk might be undesirable noise that should be minimized. Sensors that are homologous to the cognate sensor but are involved in distinct physiological responses may represent such a case. In other cases, this cross-talk might represent the physiological co-ordination of several processes that needs to be enhanced. Chemotaxis represents such a case, where the state of cellular metabolism needs to be taken into account before the cell migrates towards nutrient sources. It is less important for a cell that is already well fed to spend energy in migrating towards nutrients than if the cell is starving. The results in the previous sections allow us to identify designs that are appropriate for dealing with cross-talk in each of these contexts.

Cross-talk to a TCS module is represented by any secondary input signal ($Q_R$) coming from sources other than the regulator’s cognate sensor that causes a change in phosphorylation level of the module’s response regulator ($R^*$). The schematic diagrams in Fig. 1 explicitly represent the case in which the secondary signal is a small phosphodonor like acetyl phosphate, but not ATP. Cross-talk in this case is less amplified by the module with a bifunctional sensor (Fig. 1A) than by the module with the monofunctional sensor (Fig. 1B). Thus, the design with a bifunctional sensor is better at attenuating the cross-talk to the module, and this is appropriate when the cross-talk is physiologically undesirable. Conversely, the design with a monofunctional sensor is better at amplifying the cross-talk to the module, and this is appropriate when the cross-talk is a relevant physiological signal. A TCS with a design that could change from bifunctional, when only its primary signal conveyed physiologically relevant information, to monofunctional, when other signals should also be considered, would have an advantage in dealing with more complex situations in which there is a changing requirement for suppression or integration of secondary signals. (The NRI/NRII system of E. coli, which will be discussed below, may be one such example.)

Cross-talk to the module also can result from secondary input signals originating from other sensors. There are several formal possibilities shown schematically in Fig. 10. The analysis of these possibilities yields results similar to those already described (data not shown). A controlled comparison of the alternatives in Fig. 10A and B shows that a bifunctional design for the cognate sensor ($S_4$) is better at enhancing amplification of the regulator ($R$) response to the primary input signal ($Q_1$) while it is better at suppressing noise represented by the secondary input signal ($Q_2$). A controlled comparison of the alternatives in Fig. 10C and D shows a similar result. Thus, regardless
of the design for the non-cognate sensor, a bifunctional design for the cognate sensor results in better amplification of the primary input signal and better suppression of the secondary input signal. A controlled comparison of the alternatives in Fig. 10B and D shows that a bifunctional design for the non-cognate sensor (S₂) results in better amplification of the regulator (R) response to the secondary input signal (Q₂). Taken together, these results suggest that the design in Fig. 10C is preferred for enhancing amplification of the primary input signal while promoting attenuation of the secondary input signal. However, the design in Fig. 10A is preferred for enhanced cross-talk to achieve a more balanced integration of the two input signals.

Cross-talk from the module occurs when the cognate sensor transmits its signal to a non-cognate regulator protein. Again, there are several formal possibilities, shown schematically in Fig. 11, that we have analysed (data not shown). A controlled comparison of the alternatives in Fig. 11A and B shows that the design with a sensor that is bifunctional with respect to its cognate regulator (R₁) is better in two respects. It is better at amplifying the signal that is transmitted from the input Q to the primary output (T₁) and better at suppressing the signal that is transmitted to the secondary output (T₂). A controlled comparison of the alternatives in Fig. 11C and D shows a similar result. Thus, regardless of the design with respect to the non-cognate regulator, the design with a sensor that is bifunctional with respect to its cognate regulator is better at amplifying the signal transmitted to the primary output and at suppressing that to the secondary output. A controlled comparison of the alternatives in Fig. 11B and D shows that the design with a sensor that is bifunctional with respect to the non-cognate regulator (R₂) is better at amplifying the signal transmitted from the input Q to the secondary output (T₂). Taken together, these results suggest that the design in Fig. 11C is more effective in suppressing cross-talk from the module to the response regulators of other TCS. However, the design in Fig. 11A is more effective in enhanced cross-talk to achieve a more balanced set of response in both regulators.

Examples

The results in the previous sections suggest a rationale for selection of the two alternative sensor designs based on the physiology of the system in which the TCS is embedded. Assume that the output of the module is the phosphorylation level of the response regulator (X₂). Phosphorylation levels change as a response to changes in the input signals X₅ and X₆. Usually, X₅ is thought of as the input signal and X₆ is not considered. However, X₆ is also an input signal because it changes the phosphorylation level of the response regulator. Thus, we can consider the TCS in Fig. 1 as integrators of two signals, X₅ and X₆.

A sensor that is bifunctional with respect to its cognate regulator is better at amplifying the signal transmitted to the primary output and at suppressing that to the secondary output. A controlled comparison of the alternatives in Fig. 11B and D shows that the design with a sensor that is bifunctional with respect to the non-cognate regulator (R₂) is better at amplifying the signal transmitted from the input Q to the secondary output (T₂). Taken together, these results suggest that the design in Fig. 11C is more effective in suppressing cross-talk from the module to the response regulators of other TCS. However, the design in Fig. 11A is more effective in enhanced cross-talk to achieve a more balanced set of response in both regulators.

Fig. 10. Cross-talk to a common response regulator from two distinct sensor proteins.
A. Both sensors are bifunctional with respect to the common response regulator.
B and C. One sensor is bifunctional and the other is monofunctional with respect to the common response regulator.
D. Both sensors are monofunctional with respect to the common response regulator. See text for discussion.
When compared to the TCS with a monofunctional sensor, those with a bifunctional sensor maximize amplification of the signal $X_5$ and minimize amplification of the signal $X_6$. On the other hand, the TCS with a monofunctional sensor maximize $X_6$ amplification and minimize $X_5$ amplification when compared to otherwise equivalent TCS with a bifunctional sensor. Thus, in systems for which $X_5$ is the major signal and the influence of other signals (represented by $X_6$) needs to be minimized, the design with a bifunctional sensor should be selected; in systems for which the other signals need to be taken into account and integrated, the design with a monofunctional sensor should be selected. We will clarify these notions with a few examples.

The regulation of pore size in bacteria by changes in the osmolarity of the medium is mediated by a TCS with a bifunctional sensor. The pores in their cell membrane are composed of two different proteins. Subunit OmpF forms large pores, whereas OmpC forms smaller pores. EnvZ ($X_1$) is a membrane protein and the sensor for a TCS module. Changes in the osmolarity of the medium lead to changes in the EnvZ protein. In a high osmolarity medium, EnvZ increases its autophosphorylation rate. This in turn leads to a transfer of phosphate from EnvZ to the response regulator of the TCS, OmpR ($X_2$). OmpR is a transcription factor that binds DNA either in its phosphorylated (high affinity) or unphosphorylated (low affinity) form. In its phosphorylated form, OmpR dimerizes to increase its interaction with DNA. This leads to increased expression of OmpC and to decreased expression of OmpF (Pratt et al., 1996 for a review). The sensor in this system, EnvZ, is bifunctional (Igo et al., 1989), which can be rationalized in terms of our results as follows. Pore size should be determined exclusively by differences in osmolarity between the intracellular and the extracellular medium. If pore size were to be affected by other signals, such as changes in the levels of small phosphodonors, then osmotic balance could not be maintained and cell viability would be diminished. Thus, the bifunctional sensor design used in this TCS module is the one that maximizes amplification of changes in the osmotic pressure ($X_5$) and minimizes amplification of changes in other spurious signalling processes ($X_6$).

Chemotaxis in *E. coli* is mediated by a TCS module with a monofunctional sensor (see Eisenbach, 1996; for a review). In this case, the CheA sensor protein ($X_1$) transfers its phosphate to either of two response regulators, CheY or CheB ($X_2$). CheB is responsible for desensitizing the cell to chemical gradients, whereas CheY is responsible for changing the rate of cell tumbling so as to pro-
mote movement towards favourable concentrations. CheA is a monofunctional sensor, which means that CheY and CheB can be more effectively phosphorylated by other sources (Xₙ), either by phospho-donors or by other sensors, than they could if CheA were a bifunctional sensor. Thus, the internal metabolism of the bacterium regulating the levels of these phospho-donors is more likely to be involved in determining whether the cell will search for nutrients than it would be if CheA were a bifunctional sensor. [It must be emphasized that the chemotaxis system in _E. coli_ has a phosphatase protein, CheZ, that acts downstream of both response regulators. Another case where this situation also occurs is in the Spo0 phosphorelay in _B. subtilis_. These phosphatases are also the subject of regulation. However, in this work our goal has been to evaluate the effect of alternative sensor design on signal transmission within the prototype module and downstream aspects of design will not be considered here.]

FlbE/FlbD in _Caulobacter crescentus_ is a presumptive TCS module involved in flagellum assembly and cell cycle regulation (e.g., Wingrove and Gober, 1996). We are not aware of studies showing whether FlbE is a bifunctional or monofunctional sensor, but based on its involvement in the cell cycle we would predict that it would have a monofunctional sensor as several signals must be integrated to co-ordinate the timing of cell cycle events. Similar arguments apply to the co-ordination of sporulation events in _Bacillus subtilis_. The _Spo_ phosphorelay system appears to have a monofunctional sensor and to include different phosphatases that are specific for the different components of the relay (Perego and Hoch, 1996 for a review).

PhoR/PhoP is a TCS module with a monofunctional sensor involved in regulating expression of genes responsible for the transport of phosphate in _B. subtilis_. This module transduces signals generated by phosphate starvation (Xₙ). There is also cross-regulation between this module and PTS sugar systems (Xₙ) (Hulett, 1996; for a review). Thus, it is important for this TCS to sense other signals, beside the one coming from PhoR, and the design with a monofunctional sensor should be favoured. In fact, it has been shown that membrane-bound PhoR does not seem to influence the dephosphorylation rate of PhoB significantly (Shi et al., 1999) and thus it borders on the design of a monofunctional sensor. However, from Table 3, the PhoR sensor is predicted to be bifunctional which seems to contradict this result. A careful analysis of Shi _et al._ (1999) explains the apparent contradiction. A soluble version of PhoR (without the membrane spanning domains) enhances the dephosphorylation rate of PhoP only slightly in the absence of ATP or ADP but much more significantly in the presence of either of these molecules. Thus, the design of soluble PhoR is that of a bifunctional sensor as predicted from the homology modelling; however, this bifunctionality is effectively inhibited and transformed into monofunctionality by locating PhoR in the membrane.

If one accepts the rule we have suggested for the selection of monofunctional and bifunctional sensors, and the supportive evidence in the above cases where the physiological context can be interpreted in a fairly straightforward fashion, then one can go on to apply this rule in the interpretation of more complex systems. Two such cases in _E. coli_ are considered below, the NRI/NRII system involved in nitrogen fixation and the NarX/NarL and NarQ/NarP systems involved in nitrate and nitrite dependent gene expression.

The NRI/NRII system in _E. coli_ regulates nitrogen fixation and glutamine production. NRII is the sensor protein that phosphorylates the response regulator NRI. Under conditions of low nitrogen availability NRI upregulates the expression of glutamine synthase. This enzyme condenses nitrogen and glutamate to form glutamine. Glutamine increases the affinity of a third protein, PII, towards the sensor protein NRII, inhibiting NRII phosphorylation and creating a complex that binds phosphorylated NRI and increases NRI’s rate of dephosphorylation. Alone, NRII has little or no effect upon the rate of NRI dephosphorylation (Keener and Kustu, 1988). Thus, under normal nitrogen conditions (i.e. with normal levels of glutamine), this TCS is bifunctional and nitrogen fixation is more sensitive to regulation by glutamine levels than it would be if the module were monofunctional. Under nitrogen depletion (causing a decrease in the concentration of glutamine) PII does not bind NRII, which then becomes monofunctional. This causes the module to integrate the signals coming from glutamine/glutamate levels with those coming from other parts of metabolism, via the changes in the concentration of acetyl-phosphate, more efficiently than it would if NRII were bifunctional. Intuitively, one might think that, under nitrogen depletion, NRII should be bifunctional, in order to more efficiently respond to the nitrogen fixation needs of the cell and buffer against regulation of this fixation by other parts of metabolism. However, probably as a result of the central role of glutamate in amino acid biosynthesis, this is not so. Glutamate is a ubiquitous amino acid that is needed for the biosynthesis of all other amino acids and not just glutamine. Under nitrogen depleting conditions, it is important that the use of glutamate be co-ordinately regulated by the concentration of all amino acids, in order not to deplete the cell of some of them, by its overuse to produce glutamine. It has been reported that cell growth on glucose minimal medium containing arginine, a poor nitrogen source, is greatly decreased in mutants lacking either NRII or phosphate acetyl-transferase (E.C. 2.3.1.8) (Feng _et al._, 1992). This implies that phosphorylation by both NRII and acetyl-phosphate is important under these conditions, which agrees with a strong regulatory role for...
secondary signals in nitrogen fixation that get integrated through the NR1/NRII module. It is clear that the NR system is complex and that a more detailed model would provide additional insight.

Two TCS modules are involved in nitrate and nitrite dependent gene expression in *E. coli*, NarX/NarL and NarQ/NarP. NarX and NarQ are sensor proteins (X) that independently recognize nitrate in the medium. Each of these sensors can transfer phosphate to both regulators NarL and NarP (X) which have different specificities (Schroeder et al., 1994; Chiang et al., 1997). Whereas NarL regulates expression of genes encoding nitrate and fumarate reductases as well as nitrite exporter proteins, both NarL and NarP regulate expression of genes encoding nitrate reductase and an aerobically expressed proteins. NarQ and NarX are both sensors for NarL. NarQ has a higher affinity than NarX for phosphorylating NarL. However, NarX is more effective in enhancing the rate of NarL dephosphorylation (Schroeder et al., 1994). In light of our results this can be interpreted as a design that maximizes the signals transduced via NarX/NarL with respect to those transduced via NarQ/NarL, because NarX is bifunctional with respect to NarL whereas NarQ is not. It would be interesting to determine whether either of the sensors is bifunctional with respect to NarP. If not, this would indicate that NarP, in contrast to NarL, is designed to integrate signals coming from both NarX and NarQ.

It is clear from the examples above, and others, that two-component systems exhibit a diversity of design issues that are not well understood. Our analysis has contributed to the elucidation of just one of the variations in design, monofunctional versus bifunctional sensors. Based on our results, we have proposed the following rule: Relative to the transduction of cognate signals, bifunctional sensors enhance suppression of signals from non-cognate sources whereas monofunctional sensors enhance their integration. The experimental evidence we have discussed is consistent with this rule. This leads us to suggest that this rule may be useful in understanding whether other TCS are acting simply as transducers of the sensor signal or as integrators of signals coming from different sources.

**Experimental procedures**

**Structures for TCS**

Crystal or NMR structures have been determined for a limited number of TCS sensor and regulator proteins. In the PDB database one can find structures for a prototype bifunctional sensor (EnvZ: PDB reference identifiers 1B00, 1GXP and 1GXQ and 1QQI; Rcp1: PDB reference identifiers 1D5W, 1DBW, 1DCK and 1DCM; PhoB: PDB reference identifiers 1B00, 1GXP, 1GXQ and 1QQI; Rcp1: PDB reference identifiers 1B00, 1GXP, 1GXQ and 1QQI; Rcp1: PDB reference identifiers 13C and 1JLK). This is a very small portion of the total number of TCS with identified proteins. However, the known structures can be used in combination with sequence data and homology modelling techniques to predict the structures for these other proteins of TCS.

**Sequence data for proteins of TCS**

We searched GeneBank and Swissprot for Two Component Systems proteins. We also looked for proteins of TCS in the databases for sequenced microbial genomes (listed in MAGPIE at http://www-fp.mcs.anl.gov/~gaasterland/genomes.html). Whenever necessary, we translated the cDNA sequence into its corresponding amino acid sequence. The protein sequences that we extracted from the databases were then catalogued by organism. We did homology modelling of each protein by submitting the sequence to SWISS-MODEL (Guex and Peitsch, 1997). Those that were successfully modelled (at least for some domain of the protein) by this program are presented in Tables 1 and 2.

**Homology modelling**

The protein sequences were submitted via internet to the SwissModel server (http://swissmodel.expasy.org/; Guex and Peitsch, 1997). As templates for the homology modelling we have used files from the PDB database (http://www.rcsb.org/pdb/ Berman et al., 2000). The server attempted to generate a three-dimensional model for each sequence and returned the results via E-mail. In many cases there was not enough homology to any known structure to create a working model. In the remaining cases we have modelled (at least partially) the catalytic domain of the sensor as well as the receiver and DNA-binding domains of the regulator. In some cases we were also able to model the transmitter or the signal-sensing domain of the sensors. We then visualized and manipulated the models using SWISSPROTVIEWER (Guex and Peitsch, 1997).

**Mathematical representations**

We can describe the dynamical behaviour of a system exhibiting variations about any of its steady states by using a well-established power-law formalism for modelling biochemical systems (Savageau, 1969; Shiraishi and Savageau, 1992).
For the models in Fig. 1, the independent variables of our models are the concentrations of the input signal molecules \( (X_1, \text{primary} \text{ and } X_2, \text{secondary} ) \) and the concentration totals for sensor protein \( (X_7 = X_8 + X_9) \) and regulator protein \( (X_6 = X_4 + X_5) \). The dependent state variables, which are dependent upon the values of the independent variables and the internal dynamics of the system, are the concentrations of the sensor proteins \( (X_1, \text{phosphorylated} \text{ and } X_2, \text{unphosphorylated} ) \) and of the regulator proteins \( (X_4, \text{phosphorylated} \text{ and } X_5, \text{unphosphorylated} ) \).

There is one equation for each of the four dependent variables. For each dependent variable \( X_i \) its change in concentration with time \( (dX_i / dt) \) is given by the difference between the corresponding aggregate rate of production \( (V_i) \) and the aggregate rate of consumption \( (V_c) \). Each of these aggregate rates is represented by a mathematical rate law whose exact form is unknown, but whose arguments consist of all the variables that have a direct influence on the aggregate rate in question. For example, the aggregate rate law for consumption of the phosphorylated sensor protein in Fig. 1 is a function of its concentration \( X_i \) and of the concentration of the unphosphorylated regulator protein \( X_4 \) acting as a co-substrate.

Each aggregate rate law can be represented by a product of power-law functions, one for each argument in the rate law. This representation is guaranteed to be an accurate representation for some region of operation about a nominal steady state (e.g. see Shiraishi and Savageau, 1992). Thus, the equations for the first two dependent variables of the TCS with the bifunctional sensor (Fig. 1A) are the following (the dependent variables \( X_2 \) and \( X_3 \) will be treated below).

\[
\begin{align*}
\frac{dX_1}{dt} &= V_1 - V_2 = \alpha_1 X_1 \nu_1 X_1^{\nu_1} - \beta_1 X_1^{\beta_1} X_2^{\beta_2} \\
\frac{dX_2}{dt} &= V_1 - V_2 = \alpha_2 X_2 \nu_2 X_2^{\nu_2} - \beta_2 X_2^{\beta_2} X_3^{\beta_3} 
\end{align*}
\]

The corresponding equations for the TCS with the monofunctional sensor (Fig. 1B) are the same, except for the last term in Equation (2):

\[
\begin{align*}
\frac{dX_3}{dt} &= V_1 - V_2 = \alpha_3 X_3 \nu_3 X_3^{\nu_3} - \beta_3 X_3^{\beta_3} \\
\frac{dX_4}{dt} &= V_1 - V_2 = \alpha_4 X_4 \nu_4 X_4^{\nu_4} - \beta_4 X_4^{\beta_4}
\end{align*}
\]

The primed parameters [in Equation (2’)], whose significance will become evident below, have values that in general are different from the corresponding unprimed parameters in Equation (2).

The multiplicative parameters (rate constants), the \( \alpha \)'s for production and the \( \beta \)'s for consumption, influence the time scales of the reactions and are always non-negative. The subscript of a rate constant refers to the molecular species that is being produced or consumed. For example, \( \alpha_1 \) is the rate constant for the aggregate rate of production of \( X_1 \) (phosphorylated sensor kinase). The exponential parameters (kinetic orders), \( g \)'s for production and \( h \)'s for consumption, represent the direct influence of each variable on each aggregate rate law. These parameters need not have integer values, but can assume real values (positive, negative or zero). The first subscript of a kinetic order refers to the molecular species that is being produced or consumed; the second refers to the variable that has a direct influence on the aggregate rate of production or consumption. For example, the influences on \( V_1 \), the aggregate rate of consumption of \( X_1 \) (phosphorylated sensor protein), are represented by the kinetic orders \( h_{11} \), which is the kinetic order representing the influence of \( X_1 \) (phosphorylated sensor protein) acting as substrate in the aggregate rate of its own consumption, and \( h_{1w} \), which is the kinetic order representing the influence of \( X_1 \) (unphosphorylated response regulator) acting as a co-substrate in the aggregate rate of consumption of \( X_1 \).

It may seem more natural to represent the aggregate rate of consumption of \( X_1 \) by two separate terms such as \( \beta_1 X_1^{\beta_1} + \beta_2 X_1^{\beta_2} X_2^{\beta_2} \). However, it has been demonstrated that this is less accurate than the simple product of power-law functions described above for small variations about the steady state; this also tends to be the case for large variations, although this is not necessarily true in general (Voit and Savageau, 1987).

Protein synthesis and degradation occur on a time-scale that is in much slower than that of the catalytic activities within the TCS modules. Thus, on the time scale of interest here, the total amount of sensor protein \( (X_7, X_8, X_9) \) and the total amount of regulator protein \( (X_4, X_5, X_6) \) are considered to be conserved quantities. This permits one dependent variable to be expressed in terms of the conserved total minus the other dependent variable. Accordingly, one can write:

\[
\begin{align*}
X_3 &= X_8 - X_7 \\
X_4 &= X_6 - X_9
\end{align*}
\]

Because of these conservation relationships [Equations (3) and (4)], the rate of change in the concentration \( X_7 \) is equal in amount and opposite in sign to that in the concentration \( X_1 \) and the rate of change in the concentration \( X_8 \) is equal in amount and opposite in sign to that in the concentration \( X_2 \).

At any given steady state, one can represent the differences in Equations (3) and (4) by the following products of power-laws (see Sorribas and Savageau, 1989; for a more detailed explanation of the procedure)

\[
\begin{align*}
X_3 &= \gamma_3 X_7^{f_{31}} X_1^{f_{31}} \\
X_4 &= \gamma_4 X_8^{f_{42}} X_2^{f_{42}}
\end{align*}
\]

The new parameters in Equations (5) and (6) are defined as follows: \( f_{31} = (X_7 / X_8)(dX_7 / dX_8) \), \( f_{31} = (X_7 / X_9)(dX_7 / dX_9) \), \( f_{41} = (X_8 / X_6)(dX_8 / dX_6) \), \( f_{42} = (X_8 / X_7)(dX_8 / dX_7) \), \( f_{43} = (X_2 / X_3)(dX_2 / dX_3) \), \( f_{44} = (X_2 / X_4)(dX_2 / dX_4) \), \( f_{41} = (X_8 / X_6)(dX_8 / dX_6) \), \( f_{42} = (X_8 / X_7)(dX_8 / dX_7) \), \( f_{43} = (X_2 / X_3)(dX_2 / dX_3) \), \( f_{44} = (X_2 / X_4)(dX_2 / dX_4) \). The additional subscript \( 4 \) indicates the operating point at which the representation is made (in this case, the steady state). It must be emphasized that this representation is exact at the operating point. The multiplicative and exponential parameters in Equations (5) and (6) are analogous to the rate-constant and kinetic-order parameters in Equations (1), (2) and (2’). Thus, the \( \gamma \)'s are non-negative and the \( f \)'s are real (positive, negative or zero). The subscript of a \( \gamma \) parameter, and the first subscript of an \( f \) parameter, refers to the dependent variable that is being represented in terms of its paired variable and their sum. The second subscript of an \( f \) parameter refers to either the paired variable or the variable representing their sum.
The final form of the equations that will be used here is obtained by substituting the expressions in Equations (5) and (6) into Equations (1), (2) and (2'), and then redefining terms. Thus, one can write

\[ dX_i / dt = \alpha_{20} X_i^{1+s} X_i^{1+s} X_i^{1+s} - \beta_{20} X_i^{1+s} X_i^{1+s} X_i^{1+s} \]

\[ (7) \]

\[ dX_i / dt = \alpha_{20} X_i^{1+s} X_i^{1+s} X_i^{1+s} - \beta_{20} X_i^{1+s} X_i^{1+s} X_i^{1+s} \]

\[ (8) \]

and

where the new parameters in Equations (7), (8), and (8') are related to the original parameters in Equations (1), (2), (2'), (5) and (6) as follows: \( \alpha_{20} = \alpha_{11}^{2}, \; \beta_{20} = \beta_{11}^{2}, \; h_{21} = h_{11}^{2}, \; h_{21} = h_{21}^{2}, \; \alpha_{22} = \alpha_{12}^{2}, \; \beta_{22} = \beta_{12}^{2}, \; g_{21} = g_{11}^{2}, \; g_{21} = g_{21}^{2}, \; \beta_{22} = \beta_{21}^{2}, \; g_{21} = g_{21}^{2}, \; \beta_{22} = \beta_{22}^{2}, \; h_{21} = h_{21}^{2}, \; h_{21} = h_{21}^{2}, \) and \( h_{21} = h_{21}^{2}. \)

All of the parameters in Equations (7), (8) and (8') are positive except for \( h_{21}, \; h_{22}, \; g_{21} \) and \( h_{21}, \) which are negative. A negative exponent implies that an increase in the argument results in a decrease in the value of the power-law function. For example, an increase in \( X_i \) results in a decrease in the rate of consumption of \( X_i. \) This apparent inhibition of \( V_i \) by \( X_i \) is a result of the conservation relation among the different forms of the response regulator. Thus, an increase in \( X_i \) corresponds to an equivalent decrease in \( X_i, \) which actually causes a de-activation of \( V_i. \) Similar explanations account for the behaviour associated with the other negative kinetic orders.

As a practical matter, the results for the monofunctional design can be obtained directly from those for the bifunctional design simply by making the following exchanges: \( h_{21} \rightarrow 0, \; h_{22} \rightarrow 0, \) and \( \beta_{20} \rightarrow \beta_{20}, \; h_{22} \rightarrow h_{22}. \) Hence, in the following sections we shall focus on the bifunctional design, make these exchanges to obtain the corresponding results for the monofunctional design, and then make the appropriate comparisons.

**Steady-state solution and key systemic properties**

The equations describing the dynamic behaviour of the model in Fig. 1A. [Equations (7) and (8)] can be solved analytically for the steady state (Savageau, 1969; 1971a), where the rates of aggregate production and aggregate consumption for each metabolite are the same. By equating these aggregate rates, taking logarithms of both sides of the resulting equations and rearranging terms, one can write the steady-state equations as follows:

\[ a_{11} Y_1 - h_{12} Y_2 = b_1 - g_2 Y_2 - g_3 Y_1 + h_{12} Y_2 \]

\[ (9) \]

\[ a_{22} Y_2 + a_{22} Y_2 = b_2 - g_2 Y_2 - h_{22} Y_1 - g_3 Y_2 \]

\[ (10) \]

where \( Y_i = \log X_i, \; b_i = \log (k_i/\omega_0), \) and \( a_i = g_i - h_i. \) Equations (9) and (10) are single linear algebraic equations that can be solved for the dependent variables \( Y_1 \) and \( Y_2 \) in terms of the parameters of the system, the input variables \( Y_1 \) and \( Y_2, \) and the total concentrations of sensor and regulator proteins \( Y_i \) and \( Y_0. \) Thus,

\[ a_{11}, a_{22}, a_{12}, a_{21}, b_1, b_2, g_1, g_2, g_3, h_{11}, h_{12}, h_{21}, h_{22}, h_{23}, h_{24}, \] and \( g_{21}, g_{22}, g_{23}, g_{24}, \)

\[ (a_{11} b_1 + h_{12} b_2) - a_{22} g_1 Y_1 - h_{12} g_2 Y_2 - \]

\[ (a_{22} g_1 + h_{12} b_2) Y_1 + (a_{22} h_{12} b_1 + a_{11} g_2) Y_2 \]

\[ Y_i = \]

\[ a_{11} a_{22} + h_{12} a_{21} \]

\[ (11) \]

\[ (a_{11} b_1 - a_{22} b_2) + a_{22} g_1 Y_1 - a_{11} g_2 Y_2 + \]

\[ (a_{22} g_1 + a_{11} h_{12} b_2) Y_1 - (a_{11} h_{12} b_1 + a_{22} g_2) Y_2 \]

\[ Y_i = \]

\[ a_{11} a_{22} + h_{12} a_{21} \]

\[ (12) \]

Two types of coefficients, logarithmic gains and parameter sensitivities, can be used to characterize further the steady state of such models. Because steady-state solutions exist in explicit form [Equations (11) and (12)], we can calculate each of the two types of coefficients simply by taking the appropriate derivatives. Although the mathematical operations involved are the same in each case, it is important to keep in mind that the biological significance of the two types of coefficients is very different.

Logarithmic gains provide important information concerning the amplification or attenuation of signals as they are propagated through the system (Savageau, 1971a; Shiraishi and Savageau, 1992). For example,

\[ L(X_i, X_j) = \frac{d \log X_i}{d \log X_j} = \frac{d Y_i}{d Y_j} \] or \( L(Y_i, Y_j) = \frac{d \log Y_i}{d \log Y_j} \)

measures the percent change in the value of the dependent concentration variable \( X \) (or \( Y \) the flux through the pool of \( X \)) caused by a percentage change in the concentration of the input signal \( X \). A positive sign indicates that the changes are in the same direction (both increase or both decrease); a negative sign indicates that the changes are in the opposite direction (one increases while the other decreases).

Parameter sensitivities provide important information about system robustness, i.e. how sensitive the system is to perturbations in the structural determinants of the system (Savageau, 1971b; Shiraishi and Savageau, 1992). For example,

\[ S(Y_i, p_j) = \frac{d \log X_i}{d \log p_j} = p_j \frac{d Y_i}{d p_j} \] or \( S(V_i, p_j) = \frac{d \log V_i}{d \log p_j} \)

measures the per cent change in the value of the dependent concentration variable \( X \) (or \( V \) the flux through the pool of \( X \)) caused by a percentage change in the value of the parameter \( p_i \). Again, a positive sign indicates that the changes are in the same direction (both increase or both decrease); a negative sign indicates that the changes are in the opposite direction (one increases while the other decreases). The aggregate sensitivity of a given variable is defined as the Euclidean norm of the vector whose components are the individual parametric sensitivities for that variable. That is,

\[ S(X_j) = \sqrt{\sum_i S(X_i, p_j)^2} \] or \( S(V_i) = \sqrt{\sum_i S(V_i, p_j)^2} \)

These systems should be stable in the face of perturbations in their dependent state variables. That is, following a perturbation, the systems should return to their predisturbance
state. The local stability of the steady state can be determined by applying the Routh criteria (Dorf, 1992). The magnitude of the two critical Routh conditions can be used to quantify the margin of stability (Savageau, 1976). The two critical Routh conditions are given by

\[ F_1 a_{11} + F_2 a_{22} < 0 \]  
\[ F_1 F_2 (a_{11} a_{22} - a_{12} a_{21}) > 0 \]

where \( F_1 = V_{20}/X_{20} \) and \( F_2 = V_{20}/X_{20} \) are the reciprocal of the turnover times for the \( X \) and \( X \) pools respectively.

Systems should respond quickly to changes in their environment (Savageau, 1975). Thus, another key property of the system is its temporal response, which was determined by computer solution of the dynamic equations [Equations (7) (8) and (8')]. At time zero, each intermediate concentration was set to a value 20% less than its steady-state value. The concentrations were then followed as a function of time from this initial condition, and the time for all the concentrations to settle within 1% of their final steady-state value was calculated and denoted by the symbol \( \tau \).

**Mathematically controlled comparisons**

To determine the differences in systemic behaviour between the reference model (bifunctional sensor, Fig. 1A) and the alternative model (monofunctional sensor, Fig. 1B) we use a technique known as mathematically controlled comparison (Savageau, 1972, 1976; Irvine and Savageau, 1985; Alves and Savageau, 2000a). This technique introduces mathematical controls to ensure that the differences observed in the systemic behaviour of alternative models are a result of the specific differences in the design and not some accidental difference. The parameters of the alternative model are fixed relative to those of the reference model by introducing constraints to ensure that the two models are as nearly equivalent as possible from both an internal and an external perspective.

Only the step that accounts for the dephosphorylation of the regulator is allowed to differ between the reference model and the alternative model. Therefore, to establish internal equivalence (Savageau, 1972, 1976; Irvine, 1991) between the two designs, we require the values for the corresponding parameters of all other steps in the two models to be the same. The step that accounts for the dephosphorylation of the regulator is then the only step that differs between the reference model and the alternative model. If we reason that loss or gain of an activation site on the regulator protein comes about by mutation, and that this mutation can cause changes in all the parameters of the process, then a mutation that converts a bifunctional sensor to a monofunctional sensor would change the parameters.

\[ F_1 a_{11} + F_2 a_{22} < 0 \]  
\[ F_1 F_2 (a_{11} a_{22} - a_{12} a_{21}) > 0 \]

The kinetic-order parameters in the reference model, which is constrained, we require two different systemic properties to be invariant between the designs.

First, the value of \( h_{21} \) is fixed by requiring the total gain of the system, \( L(X_6, X_6) \), to be the same in both designs. This total gain is determined from the explicit solution for \( X_6 \) [\( Y_6 = \log X_6 \) in Equation (12)] by calculating the logarithmic gains as defined by the derivatives in Equation (13). Equating the results for each of the alternative designs allows the value of \( h_{21} \) to be calculated as a function of the kinetic-order parameters in the reference model, which is taken to be the bifunctional case.

\[ h_{21} = g_1 g_1 h_{21} + g_2 h_{21} - g_1 g_1 h_{21} + g_2 h_{21} - g_1 g_2 h_{21} - g_1 g_2 h_{21} \]

By using this constraint we ensure that the output signal of the response regulator \( X_6 \) is the same for both designs in responses to the aggregate of signals \( X_6 \) and \( X_6 \) that change its phosphorylation state. We also have constrained the alternative designs to have equal responses to the primary input signal \( X_6 \) [\( L(X_6, X_6) \)], or to the secondary input signal \( X_6 \) [\( L(X_6, X_6) \)]. The results of the comparative analysis in these other cases are qualitatively the same as those of the comparative analysis reported here (data not shown).

Second, the value of \( \beta_{20} \) is fixed by requiring the steady-state concentrations of the corresponding variables to be the same in both designs. By equating the explicit solution for \( X_5 \) [\( Y_5 = \log X_5 \) in Equation (12)] in each of the designs, and by utilizing the result in Equation (18), one is able to express the value of \( \beta_{20} \) in terms of the independent variables and parameters of the reference model.

\[
\log \beta_{20} = \left\{ \begin{array}{l}
g_1 h_{21} \log a_{20} - (g_1 g_2 - g_2 (h_{21} - h_{11}) \log h_{21}) + g_2 h_{21} \log (b_{21}/a_{11}) + g_1 g_2 h_{21} \log X_6 + X_4 \\
(\gamma_2 h_{21} (h_{21} - h_{11})) - g_1 g_2 h_{21} - g_2 h_{21} \log X_7 \\
(\gamma_3 h_{21} + g_2 h_{21}) \log X_6 \\
g_2 h_{21} (h_{21} - h_{11}) + g_1 (h_{21} - h_{21}) 
\end{array} \right.
\]

© 2003 Blackwell Publishing Ltd, Molecular Microbiology, 48, 25–51
In addition to making the steady-state value of $X_2$ equal in the alternative designs, this value of $b_{50}$ also makes the steady-state value of $X_1$ [and $X_2 \equiv X_1 - X_2$ and $X_3 \equiv X_1 - X_3$], and the steady-state values of the corresponding fluxes $V_2 \equiv V_2$ and $V_3 \equiv V_3$ equal in the alternative designs.

**Numerical analysis**

The analytical results give qualitative information that characterizes the effect of bifunctional versus monofunctional sensors in the models of Fig. 1. To obtain quantitative information about the comparisons, one must introduce specific values for the parameters and compare models (Alves and Savageau, 2000a). For this purpose we have randomly generated a large ensemble of parameter sets and selected $5 \times 10^5$ of these sets that define models consistent with various physical and biochemical constraints. These constraints include conservation of mass considerations, a requirement for positive signal amplification, and stability margins large enough to ensure local stability of the systems. A detailed description of these methods can be found in Alves and Savageau (2000b).

When applied to the current comparisons, the procedure is as follows. We select random values for all the unprimed parameters and for all the independent concentration variables ($X_1$ through $X_7$) in Equations (7) and (8). The values of the primed parameters in Equation (8’) are then fixed by the relationships in Equations (18) and (19). The analytical solution in Equations (11) and (12) determines the steady-state values for dependent state variables ($X_1, X_2, X_3$), complementary variables ($\dot{X}_1$ and $\dot{X}_3$), fluxes, logarithmic gains and parameter sensitivities. This information in turn determines the values for all the parameters following Equations (6) and (8). Taken together, this information determines the values for all the parameters in the original equations [Equations (1) (2) and (2’)]. Mathematica™ (Wolfram, 1997) was used for all the numerical procedures.

To interpret the ratios that result from our comparative analysis we use Density of Ratios plots as defined in Alves and Savageau (2000a). The primary density plots from the analysis we use Density of Ratios plots as defined in Alves and Savageau (2000a). The primary density plots from the analysis we use Density of Ratios plots as defined in Alves and Savageau (2000b). The secondary density plots are constructed from the primary density plots by the use of moving quantile techniques with a window size of 500. The slope in the secondary plot measures the degree of correlation between the quantities plotted on the x- and y-axes.

**Acknowledgements**

This work was supported in part by a joint PhD fellowship PRAXIS XXI/BD/9803/96 granted by PRAXIS XXI through Programa Gulbenkian de Doutoramentos em Biologia e Medicina (R.A.), U.S. Public Health Service Grant RO1-GM30054 from the National Institutes of Health (M.A.S.), and U.S. Department of Defense Grant N00014-97-1-0364 from the Office of Naval Research (M.A.S.). We thank Drs Susana Nery, Armindo Salvador, Vic DiRita and Alex Ninfa for critically reading early versions of this manuscript and making useful comments. We also thank an anonymous referee for a very careful and constructive review of the original manuscript. This work was completed while MAS was a guest at the Institut des Hautes Études Scientifiques, Bures-sur-Yvette, France. He thanks Drs A. Carbonme, M. Gromov and F. Képes for providing resources, intellectual stimulation and generous hospitality.

**References**


Hlavacek, W.S., and Savageau, M.A. (1996) Rules for cou-


