Phosphatase activity tunes two-component system sensor detection threshold

Brian P. Landry1, Rohan Palanki1, Nikola Dyulgyarov1, Lucas A. Hartsough1 & Jeffrey J. Tabor1,2

Two-component systems (TCSs) are the largest family of multi-step signal transduction pathways in biology, and a major source of sensors for biotechnology. However, the input concentrations to which biosensors respond are often mismatched with application requirements. Here, we utilize a mathematical model to show that TCS detection thresholds increase with the phosphatase activity of the sensor histidine kinase. We experimentally validate this result in engineered Bacillus subtilis nitrate and E. coli aspartate TCS sensors by tuning their detection threshold up to two orders of magnitude. We go on to apply our TCS tuning method to recently described tetrathionate and thiosulfate sensors by mutating a widely conserved residue previously shown to impact phosphatase activity. Finally, we apply TCS tuning to engineer B. subtilis to sense and report a wide range of fertilizer concentrations in soil. This work will enable the engineering of tailor-made biosensors for diverse synthetic biology applications.
A central goal of synthetic biology is to program cells to sense and respond to chemical or physical inputs in desired ways. To this end, researchers develop genetically encoded sensors, often based on multi-step signal transduction pathways or one-component transcription factors that convert inputs of interest into biological signals such as gene expression. However, all biosensors respond to their cognate inputs over finite concentration ranges that are often mismatched with application demands.

Despite this challenge, there has been little focus on developing technologies for tuning biosensor detection windows. In two recent studies, the input concentrations required to activate *Escherichia coli* nitrate and hydrogen peroxide sensors by 50% (i.e., the detection thresholds, quantified by the parameter $K_{i/2}$) were decreased 412- and 15-fold by linking the respective sensors to the expression of a phage recombinase that inverts a segment of DNA into an orientation appropriate for transcription of an output gene. Though this approach is simple and modular, the recombination step is irreversible and delays sensor response by up to 15 h, making it incompatible with applications requiring dynamic or rapid responses. In a separate pair of yeast studies, RNA secondary structure design was used to lower the detection threshold of an engineered theophylline-responsive antiswitch from 10 mM to 1 mM, and protein expression level optimization was used to reduce the estradiol detection threshold of the mitogen-activated protein kinase (MAPK)/extracellular signal-regulated kinase pathway from 32 µM to 6.6 µM. However, antiswitches currently sense a limited number of inputs, and both of these approaches yield modest changes in detection threshold, limiting the utility of these strategies. Finally, computational design and directed evolution of ligand-binding transcription factors show promise for tuning sensor detection thresholds. However, these methods are time and labor intensive and require extensive domain-specific expertise, limiting their widespread use.

Two-component systems (TCSs) are an important source of sensors for synthetic biology. Tens of thousands of TCSs have been identified in bacterial genome sequences. Individual members of this family sense inputs as diverse as metal ions of particular oxidation states, respiratory electron acceptors, inorganic phosphate, heme, quorum sensing auto-inducers, antimicrobial peptides, simple sugars, gut polysaccharides derived from the diet or host, human and plant hormones, oxidative stress, physical contact, and specific wavelengths of light. Synthetic biologists have begun to repurpose light-sensing TCSs to function as sensors for optogenetics and chemical-sensing TCSs to engineer diagnostic gut bacteria, among other applications.

The prototypical TCS comprises two proteins: a sensor histidine kinase (SK) and a response regulator (RR) (Fig. 1a). The SK contains a (typically extracellular) N-terminal sensor domain that switches from an inactive to active conformation in the presence of the input. This conformational change is transmitted to a C-terminal cytoplasmic signaling region comprised of catalytic and adenosine triphosphate (ATP) binding (CA) and dimerization and histidine phosphotransfer (DHp) domains. The CA domain catalyzes the transfer of the gamma phosphoryl group from ATP to a conserved histidine residue within the DHp domain. The phosphorylated SK (SK–P) binds the RR via a DHp interaction interface, and transfers the phosphoryl group to a conserved RR aspartate. Phosphorylation activates the RR, driving it to modulate transcription from one or more output promoters. Many SKs are also bi-functional and dephosphorylate the phosphorylated RR (RR–P) (Fig. 1a). The presence of input increases the rate at which the RR is phosphorylated, decreases the rate at which the RR–P is dephosphorylated, or both. Many SK mutations, in both the DHp and CA domains, have been identified that decrease this phosphatase activity, resulting in increased RR–P level. When this increase is substantial, it results in leaky transcriptional output, i.e., output in the absence of input. However, the impact of these phosphatase-altering mutations on TCS detection thresholds has not been considered.

Here, we combine mathematical modeling with an experimental synthetic biology approach to show that mutations that alter SK phosphatase or kinase activity can be used to rationally tune TCS detection thresholds. We demonstrate that our method functions in Gram-negative and Gram-positive bacteria and in diverse chemical-sensing TCSs. We go on to demonstrate that a widely conserved residue can be mutated to tune the detection thresholds of two recently described TCSs for which signaling mutations have not yet been identified. Finally, we utilize *Bacillus subtilis* expressing wild-type and sensitivity-enhanced nitrate sensors to quantify a wide range of fertilizer levels in soil. These sensors could be used to control the expression of engineered nitrogen fixation pathways in order to achieve synthetic nitrogen homeostasis in soil.

**Fig. 1** Mathematical model reveals that SK phosphatase activity tunes TCS detection threshold. **a** Diagram of a canonical TCS. **b** Model simulations of the relationship between TCS input concentration and transcriptional output rate (i.e., transfer function) wherein SK phosphatase activity is varied between 1% and 10,000% of wild-type (Supplementary Note 1). Detection threshold ($K_{i/2}$), or the input concentration where transcriptional output is half-maximal, increases with phosphatase activity. There is no trade-off between detection threshold and dynamic range for intermediate changes in phosphatase activity; however, a trade-off emerges for strong changes in phosphatase activity (Supplementary Fig. 1).
Results

Mathematical model of TCS detection threshold. We hypothesized that TCS detection thresholds could be tuned by introducing mutations that alter SK kinase or phosphatase activity without compromising the overall response (i.e., dynamic range, or ratio of output in saturating versus zero input) of the system. Specifically, we considered that the detection threshold of a TCS occurs at the particular RR~P concentration that elicits a half-maximal output promoter response (i.e., RR~P_1/2). For any input concentration, the corresponding RR~P concentration is set by the ratio of SK kinase to phosphatase activity. Thus, we reasoned that mutations that enhance kinase or reduce phosphatase activity should result in RR~P_1/2 being reached at a lower input concentration, thereby reducing TCS detection threshold. The opposite should also be true: TCS detection thresholds should increase with kinase-reducing or phosphatase-enhancing mutations. Furthermore, if a mutation is sufficiently weak that the window of altered RR~P concentrations still traverses the range to which the output promoter is sensitive, there should be little effect on TCS dynamic range.

To examine this hypothesis, we first utilized a previous mathematical model of TCS signaling. We parameterized the model with the best available in vivo experimental values of TCS reaction rates as determined for the well-studied inorganic phosphate-sensing TCS PhoRB. Then, we set the phosphatase reaction rates as determined for the well-studied inorganic model with the best available in vivo experimental values of TCS activity should result in RR~P_1/2 being reached at a lower input concentration, thereby reducing TCS detection threshold. The opposite should also be true: TCS detection thresholds should increase with kinase-reducing or phosphatase-enhancing mutations. Furthermore, if a mutation is sufficiently weak that the window of altered RR~P concentrations still traverses the range to which the output promoter is sensitive, there should be little effect on TCS dynamic range. We also found that mutating kinase activity had the reciprocal effect to that of modulating phosphatase activity, with increasing kinase activity decreasing the detection threshold and decreasing kinase activity increasing the detection threshold (Supplementary Fig. 1). However, our primary goal is to decrease TCS detection thresholds, and it is easier to identify mutations that decrease rather than increase enzymatic activity. Thus, we chose to focus on decreasing phosphatase activity as opposed to increasing kinase activity. This decision is supported by mutational screens of SK activity that have found that decreases in phosphatase activity are much more common than increases in kinase activity.

Tuning the detection threshold of a nitrate sensor. To examine our modeling results experimentally, we selected two point mutations, C415R and D558V, that decrease the phosphatase activity of the E. coli nitrate-activated SK NarX via different mechanisms and to different extents. C415R targets the DHP interaction interface, weakens the interaction between NarX and its cognate RR NarL, and causes a moderate reduction in phosphatase activity. On the other hand, D558V targets the CA domain and is thought to decrease phosphatase activity more strongly than C415R. However, because its impact has been measured only with gene expression assays, it is also possible that D558V may increase kinase activity. We measured the nitrate detection thresholds of a wild-type NarXL that we engineered to function in Bacillus subtilis, and its corresponding C415R and D558V variants (Fig. 2a; Supplementary Fig. 2). The wild-type system exhibits a relatively high K_1/2 of 762 μM (95% confidence interval (CI) 629–963 μM) (Fig. 2b). On the other hand, the medium strength C415R mutation decreases the value substantially (K_1/2 = 22 μM, 95% CI 16–33 μM) (Fig. 2b). The strong D558V mutation reduces it even further (K_1/2 = 6 μM, 95% CI 0–23 μM) (Fig. 2b).

Dynamic range is commonly reported as the primary performance metric for biosensors. The C415R and D558V versions of our nitrate sensor exhibit decreased dynamic range due to increased minimum output levels (Fig. 2b). Thus, we individually optimized SK and RR expression levels in these mutated sensors in an effort to maximize the dynamic range for each (Supplementary Fig. 3). Consistent with our modeling results, maximal dynamic range decreases from 1909-fold (wild type), to 78-fold (C415R) and 2-fold (D558V) (Supplementary Figs. 1, 3). On the other hand, the amplitude range, or difference between maximum and minimum output, may be a more useful
performance metric for many applications. While the amplitude range of our wild-type nitrate sensor is 24,652 molecules of equivalent fluorescein (MEFL) (21 MEFL to 24,664 MEFL), it increases to 65,402 MEFL (2,508 MEFL to 67,910 MEFL) for C415R and 31,294 MEFL (34,758 MEFL to 66,052 MEFL) for D558V (Fig. 2b). These results provide compelling initial support for our approach.

To more rigorously validate TCS tuning, we next developed a strategy to continuously vary phosphatase activity in live cells (Fig. 3a). Specifically, we expressed wild-type NarX and NarX (C415R) under two different chemically inducible promoters and utilized green fluorescent protein (GFP) fusions and quantitative flow cytometry to map the relationship between inducer and SK levels (Fig. 3b; Supplementary Figs. 4-6). Then, we used different inducer combinations to achieve NarX+NarX (C415R) expression ratios between 0% and 100% at a constant total SK expression level (NarX + NarX (C415R)) (Fig. 3b). Assuming NarX and NarX (C415R) function identically outside of their different phosphatase activities, tuning their expression ratio in this way enables us to continuously vary phosphatase activity between mutant and

---

**Fig. 3** Detection threshold decreases continuously with phosphatase activity. **a** Cartoon of the genetic system engineered for the iso-SK experiment. Both NarX and NarX(C415R) are expressed under inducible promoters in the same cell. **b** IPTG and xylose are used to express different levels of NarX and NarX(C415R) resulting in different total SK concentrations (Supplementary Figure 6). We selected induction levels that result in 775 MEFL (Methods) of total SK expression (black line), while evenly varying the percentage of NarX compared to total SK levels between 0 and 100% (blue points). Requisite inducer concentrations were calculated by inverting the Hill fit (inducer = \( \sqrt{\frac{r}{C_{138}}/C_{138}} \)) in (Supplementary Fig. 6). **c** Nitrate transfer functions for the iso-SK strain when induced with IPTG and xylose to the 11 different percentages of wild-type NarX from (b). Points represent the mean and error bars the standard error of the mean (SEM) of experiments on three separate days. **d** The relationship between the percent NarX and the \( K_{1/2} \) of the Hill function fits in (c). Points represent the \( K_{1/2} \) values and error bars the 95% confidence intervals of the best fits.
Due to the language model's limitations, the natural text of the entire document is not accurately generated. Please refer to the raw text for the complete content.
exhibiting the largest fold activation (Supplementary Figs. 10, 11). All of the mutations that we tested lower the detection threshold (Fig. 5d–g). In the case of TtrSR, $K_{1/2}$ varies between 35.6 μM (95% CI 27–48 μM) for wild-type and 1.5 μM (95% CI 1.3–1.9 μM) for the strongest mutant (Fig. 5d). For ThsSR, $K_{1/2}$ varies between 192 μM (95% CI 138–305 μM) and 35 μM (95% CI 24–60 μM) (Fig. 5e). Because the CA domain is involved in the kinase and phosphatase reactions, further characterization is needed to determine which enzymatic activity, or activities, have been changed by these GXGXG mutations. Interestingly, we found that the TtrS(L627A) mutant not only decreased the detection threshold from 35.6 μM to 2.4 μM, but it also increased the dynamic range from 15- to 21-fold and the amplitude range from 1377 MEFL (100 MEFL to 1477 MEFL) to 2095 MEFL (105 MEFL to 2200 MEFL) (Fig. 5f). Conversely, decreasing the detection threshold of the thiosulfate sensor twofold with L547T resulted in a decrease in dynamic range from 34- to 13-fold and an increase in amplitude range from 19,390 MEFL (596 MEFL to 19,986 MEFL) to 23,482 MEFL (1905 MEFL to 25,387 MEFL) (Fig. 5g). We conclude that mutating the second GXGXG residue is a simple strategy for tuning the detection thresholds of diverse TCSs.

**Application of TCS tuning to fertilizer biosensing.** Finally, we set out to demonstrate a proof-of-principle application for TCS tuning. Nitrate is the primary source of nitrogen used by crops, and a major component of fertilizer. However, over-application of fertilizer causes billions of dollars in damage per year to human health and the environment45. Recently, synthetic biologists have expressed bacterial nitrogen fixation pathways, which ultimately convert atmospheric N$_2$ into nitrate, in non-native host bacteria46. However, heterologous production of nitrogen fixation pathways in soil bacteria could also lead to nitrate overproduction. To prevent this outcome, genetic feedback control systems wherein bacteria sense a wide range of soil nitrate levels...
and induce nitrogen fixation pathways only to the extent that they are needed are highly desirable.

To demonstrate such a sensing capability, we incubated *B. subtilis* engineered to express our wild-type and C415R NarXL systems in soil spiked with various amounts of a nitrate standard, and measured the resulting superfolder GFP (sfGFP) fluorescence values via flow cytometry (Methods; Fig. 6a; Supplementary Fig. 12). Then, we used the resulting data to generate a standard curve relating sfGFP fluorescence to soil nitrate concentration (Supplementary Fig. 12; Supplementary Note 2). Then, we added different amounts of commercial fertilizer, rather than nitrate, to the soil (Methods; Supplementary Fig. 12). Using the standard curves, we compared the amount of nitrate reported by each of our sensor systems to the amount specified by the manufacturer (Supplementary Note 2). Indeed, the wild-type NarXL system enables estimation of fertilizer levels within twofold of the manufacturer value between the tested values of 31.6 μM and 562 μM nitrate, while the C415R system allows accurate detection between 5.62 μM and 562 μM (Fig. 6b). This experiment demonstrates that we can use TCS tuning to engineer bacteria to sense a large range of nitrate concentrations in a complex soil environment. Such broad-range sensing could be coupled with nitrogen fixation pathways to maintain soil nitrate at ideal levels in different agricultural contexts.

**Discussion**

This work extends a growing suite of techniques for engineering TCSs to function as sensors for synthetic biology. First, literature searches or bioinformatics can be used to identify TCSs that sense inputs of interest. If a candidate TCS has a known output promoter, and functions in the desired host and environmental conditions, it can be used as an off-the-shelf sensor without further modifications. Otherwise, the sensor domain can potentially be swapped onto the SK of a second TCS that contains a reliable output promoter, resulting in the design of a chimeric sensor. Like all gene regulatory systems, TCSs can exhibit substantial ‘leakiness’ in the off state, or modest dynamic range. These performance features can be improved by redesigning the sequence of the output promoter and optimizing the expression levels of the SK and RR.

However, this workflow may produce sensors that do not respond appropriately to application-relevant input concentrations. For example, tetrathionate was previously shown to be elevated in the mouse colon during *Salmonella typhimurium*-induced inflammation. Following this report, Silver and colleagues used *S. typhimurium* TrtSR to activate a transcriptional memory circuit in order to engineer a gut bacterium that senses and remembers tetrathionate exposure in order to diagnose colon inflammation. However, despite 100% tetrathionate activation in vitro, most bacteria expressing this sensor device are not...
activated by inflammatory conditions in vivo. One possible reason for this discrepancy is that in vivo tetrahionate concentrations do not reach the S. typhimurium TrrSR detection threshold. Thus, by using TCS tuning to lower the detection threshold of TrrSR (Fig. 5f), it is possible that the performance of this diagnostic gut bacterium could be improved.

It is possible that nature uses phosphatase activity as a knob to tune TCS detection threshold as well. First, there are a wide range of SK residues that can be mutated to specifically alter phosphatase activity. This fact suggests that evolution can tune TCS detection thresholds, which could enable organisms to adapt to new niches with different input concentrations. Interestingly, few mutations have been discovered that increase phosphatase activity, or decrease kinase activity. As of currently, this fact restricts our TCS tuning method to applications where lower detection thresholds (i.e., increases in sensitivity) are needed. However, sensitivity decreases are also desirable in many synthetic biology applications, which motivates future work to identify appropriate mutations.

Additionally, some SKs interact with phosphatase-modulating auxiliary proteins. It is possible that these auxiliary proteins can tune the detection thresholds of the corresponding TCSs. Unlike SK mutations, they could also be dynamically induced or repressed in response to changing environmental or physiological conditions to temporarily adjust detection thresholds. This phenomenon is analogous to our use of chemically inducible promoters to adjust the NAR XL nitrate detection threshold in our iso-SK experiment (Fig. 3). These intriguing possibilities remain to be explored.

Finally, our approach may be extensible to other kinase pathways. For example, eukaryotes use MAPK cascades to sense and respond to important extracellular signals such as growth factors and immunomodulators. Threonine and tyrosine phosphatases modulate signaling through these pathways by dephosphorylating MAP kinases. Researchers have expressed variants of these phosphatases under synthetic feedback control to re-program pathway response dynamics. Alternatively, by constitutively expressing such phosphatases to different extents, or expressing phosphatases of different strengths, the detection thresholds of MAPK cascades could potentially be tuned.

In conclusion, we have demonstrated a simple, general strategy for tuning the detection threshold of TCSs—one of the largest and most diverse families of sensors in biology. Due to its effectiveness and ease of use, our method should have widespread applications in synthetic biology.

**Methods**

**DNA and bacterial strain construction.** Details of synthetic DNAs used in this work are given in Supplementary Data 1-4. All E. coli strains are expressed on extrachromosomal plasmids. All plasmids were assembled via Golden Gate cloning. Assembled plasmids were transformed into E. coli N80-10-β (New England Biolabs, cat no. C3019H). Ribosome binding site (RBS) strengths were calculated using the RBS calculator.

All B. subtilis systems are constructed as linear double-stranded DNA Integration Modules (IMs) and integrated into the chromosome. All IMs were assembled with Golden Gate cloning. Assembled DNA was amplified with PCR, transformed into B. subtilis 168 (BSCCID 1A1) and recombined into the chromosome using the two-step transformation protocol. B. subtilis genomic DNA was then purified (Promega, A1120) and used for subsequent transformations.

**Constructs.** E. coli B18 10-β and B. subtilis 168 were grown in LB Miller broth shaking at 250 rpm at 37 °C. Then, 50 µg ml⁻¹ ampicillin, 35 µg ml⁻¹ chloramphenicol, and 100 µg ml⁻¹ spectinomycin for E. coli and 100 µg ml⁻¹ spectinomycin, 0.5 µg ml⁻¹ erythromycin, 5 µg ml⁻¹ kanamycin for E. coli were added where appropriate. Transformed strains were stored in 15% glycerol stocks at −80 °C.

E. coli plasmids are available from Addgene using accession numbers listed in Supplementary Data 3. B. subtilis constructs are available from the Bacillus Genetic Stock Center using BGSC numbers listed in Supplementary Data 4.

**In vitro nitrate experiments.** In vitro nitrate induction experiments were conducted with B. subtilis 168 ΔΔβHIlEΔxam1 (ND464; Supplementary Fig. 2). C. (60 mM NaNO₃ with sodium succinate and potassium glutamate) containing 30 mM KH₂PO₄ (Fisher BioReagents, BP362-1), 70 mM K-HPO₄ (Fisher BioReagents, BP363-1), 25 mM (NH₄)₂SO₄ (Sigma, A4418-100mg), 10 mM MnSO₄ (Sigma-Aldrich, MT834-100G), 500 µM MgSO₄ (VWR, BDH2946-500G), 12.5 µM ZnCl₂ (Sigma, Z1753-500G), 254 µL L-tyrophan (Sigma-Aldrich, T0254- 25G), 22 mg L⁻¹ ammonium iron(III) citrate (Sigma-Aldrich, FS879-100G), 43.2 mM Potassium Glutamate (Alfa Aesar, A17232), 22.2 mM Sodium Succinate (Alfa Aesar, 33386), and 43.4 mM Glycerol (Fisher BioReagents, BP229-1) were used without antibiotics. Induction conditions were 25 mM NaNO₃ (Sigma-Aldrich, S5500), 10 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) (IBI Scientific IB02125), and 1% xylene (Alfa Aesar, A10643) unless otherwise noted. IPTG and xylene levels were chosen for optimal fold change of the NAR (D5585V) TCS (Supplementary Fig. 3). An overnight culture was inoculated from a 15% glycerol freezer stock and grown in 3 ml of media for 13–15 h. Cells were diluted to OD₆₀₀ = 3 × 10⁻⁵ with relevant inducers in a 500 µl volume in 24-well plates seeded with a tunin adhesive film (VWR, PB2910). Cells were grown to an OD₆₀₀ = 0.3 (approximately 6 h) and placed on ice prior to measuring via flow cytometry with a FL1 gain of 600. All growth was conducted shaking at 250 rpm at 37 °C.

**Aspartate experiments.** Aspartate induction experiments were conducted in E. coli BW29655 (BW28357 ΔΔmζ-ompR520::FRT); (FGC; CSTC #7934; Yale University). M9 media containing 1x M9 salts (42 mM Na₃HPO₄, 24 mM KH₂PO₄, 3.0 mM NaCl, 19 mM NH₄Cl, 0.3% FektoFerment M100), 2 mM MgCl₂, 0.05% SC media (BD, D8807-500G), and 0.1 mM CaCl₂ (Alfa Aesar, L13191) were used with 22.2 mM glucose (Avantor, 4908-06) as a carbon source, and 2 g L⁻¹ aspartic acid, 50 µg ml⁻¹ ampicillin, 35 µg ml⁻¹ chloramphenicol, 100 µg ml⁻¹ spectinomycin, 10 µM EDTA, 5 µg ml⁻¹ ampicillin, 25 µg ml⁻¹ kanamycin and 2 g L⁻¹ glycerol were used. Then, 3 ml of this medium in a 14 ml culture tube was inoculated to OD₆₀₀ = 5 × 10⁻² from a single use 15% glycerol stock stored at −80 °C containing cells frozen during exponential phase. Bacteria were grown for 2 h shaking at 250 rpm at 37 °C. Amino acids were then removed by centrifuging at 3220 × g for 5 min, resuspension in 5 ml of media without casamino acids, and centrifuging at 3220 × g for 5 min. The media without casamino acids was added to the culture and bacteria were grown for 2 h shaking at 250 rpm at 37 °C, placed on ice, and then measured via flow cytometry with a FL1 gain of 750.

**Computational analysis of the phosphatase hot spot residue.** To estimate the fraction of known SKs that contain the phosphatase hot spot residue, we first assembled a library of non-redundant SK sequences from 4861 NCBI (National Center for Biotechnology Information) RefSeq bacterial genomes using HMMERS. We used hmmsearch to identify all proteins that had a C-terminal kinase core composed of a single kinase domain (Pfam: HisKA, HisKA_2, HisKA_3, His_kinase, H_kinase_dim) followed by a HATPase_c domain (reporting threshold set to 12.0 for each). We eliminated SKs with non-canonical signaling architectures by requiring that each had at least a minimal sensing region (N terminal of the kinase core) and contained neither a Response_reg nor a histidine phosphotransfer domain (Hpt). This constraint resulted in 105,144 SK proteins. To eliminate redundant sequences from this pool, we used usearch to cluster the sequences according to a 60% sequence similarity threshold (using ‘cluster_fast’ and ‘sort_length’ parameters). The centroids of each cluster were then used as representatives of non-redundant SK sets, resulting in 56,855 proteins. We next created a hidden Markov model (HMM) representing the G2 box motif (Supplementary Fig. 9) by aligning 12 representative G2 box sequences and using hmmbuild to create a model. This model was then used with hmmsearch (default parameters) to identify SKs in the non-redundant set that match, yielding 38,966 SKs with putative G2 box motifs. Two additional criteria were used to eliminate false positives: (1) the putative G2 box must align to the correct region of the protein (C terminal to the HisKA domain), and (2) the G2 box must have G3 and G5 present when aligned to the HMM. Applying these constraints left 36,508 SKs remaining, constituting 64.21% of the full non-redundant SK data set. Finally, the distribution of residues in the second position of the G2 box motif were tabulated from these SKs.

**Tetrahionate and thiosulfate induction experiments.** Tetrahionate and thiosulfate induction experiments were conducted with E. coli BW28357 (CGSC# 7991, Yale University). M9 media were used with 1 x M9 salts, with 43.4 mM glycerol (Fisher BioReagents, BP229-1) as a carbon source, 2 g L⁻¹ casamino acids (EMD Millipore, 2240-500GM), 35 µg ml⁻¹ chloramphenicol, and 100 µg ml⁻¹ spectinomycin. For thiosulfate experiments, 200 µM IPTG and 20 ng ml⁻¹ aTc were used, and leaky expression of the TtrSR TCS without inducers was found to be sufficient. Ligand induction was achieved with K₂S₂O₃ (Sigma-Aldrich, P2926-25G) or Na₂S₂O₃ (Sigma-Aldrich, 217247-25G). The experiment was started by inoculating 3 ml of media in a 14 ml culture tube to OD₆₀₀ = 1 × 10⁻² from a single-use 15% glycerol stock. Amino acids and carbohydrates were added to the media. Cells were grown at 37 °C shaking at 250 rpm for 4 h, placed on ice, and then measured via flow cytometry with a FL1 gain of 600.
Soil nitrate experiments. Soil experiments were conducted with B. subtilis 168 SydHif.lamKmCheR (ND77; Supplementary Fig. 12). CSE media with 0.5% xylose and 3 μM IPTG were used without antibiotics in all experiments. IPTG and xylose levels were selected to achieve a large fold change of both wild-type and 150 MEFL for uninduced sample. The geometric mean of the population was used to calculate the standard curve created from the calibration beads measured on that day. The (1000 ordered events, a density gate was then applied to select the densest 10% of events.

To calculate sfGFP fluorescence, resulting in exaggerated fold change calculations.

To activate modeling and parameter estimation. All transfer function data were fit to an activating Hill equation

\[ y = \frac{\text{low} + \text{high} \times \frac{\text{input}}{\text{threshold}}}{} \]

using the LmFit python package. Here, \( y \) is the sfGFP fluorescence (MEFL), \( x \) is the concentration of inducer (μM), low is the sfGFP fluorescence at 0 μM input (MEFL), high is the maximum sfGFP fluorescence (MEFL), \( K_{\text{Hill}} \) is the Hill coefficient of inducer that gives rise to half-maximal sensor activation (μM), and \( n \) is the Hill coefficient. All transfer functions were experimentally measured on three separate days. Replicate data points were combined into a single data set. This set was fit to a Hill equation. To fit both low and high values well, the fit residuals at each data point were weighted by multiplying the residual by the inverse of the mean at that data point. The 95% confidence intervals of fit parameter values were calculated using the conf_interval function in LmFit, which executes the F-test. Fit parameters for all experiments in this study are shown in Supplementary Data.

Code availability. The code used to generate a model of a TCS is included as a supplementary file to this article.

Data availability. The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request. DNA sequences are available from GenBank and accession numbers can be found in Supplementary Data 3.

Received: 12 January 2018 Accepted: 22 March 2018
Published online: 12 April 2018

References

1. Brophy, J. A. N. & Voigt, C. A. Principles of genetic circuit design. Nat. Methods 11, 508–520 (2014).

2. Wang, Y.-H., Wei, K. Y. & Smolke, C. D. Synthetic biology: advancing the design of diverse genetic systems. Annu. Rev. Chem. Biomol. Eng. 4, 69–102 (2013).

3. Ang, J., Harris, E., Hussey, B. J., Kil, R. & McMillen, D. R. Tuning response curves for synthetic biology. ACS Synth. Biol. 2, 547–567 (2013).

4. Courbet, A., Endy, D., Renard, E., Molina, F. & Bonnet, J. Detection of pathological biomarkers in human clinical samples via amplifying genetic switches. Proc. Natl. Acad. Sci. USA 110, 7333–7338 (2013).

5. Rubens, R. J., Selvaggio, G. & Lu, T. K. Synthetic mixed-signal computation in living cells. Nat. Commun. 7, 11658 (2016).

6. Bayer, T. S. & Smolke, C. D. Programmable ligand-controlled riboregulators of eukaryotic gene expression. Nat. Biotechnol. 23, 337–343 (2005).

7. O’Shaughnessy, E. C., Palani, S., Collins, J. J. & Szarz, C. A. Tunable signal processing in synthetic microbial gene cascades. Cell 144, 119–131 (2011).

8. Taylor, N. D. et al. Engineering an allosteric transcription factor to respond to new ligands. Nat. Methods 13, 177–183 (2016).

9. Hawkins, A. C., Arnold, F. H., Stuermer, R., Hauer, B. & Leadbetter, J. R. Directed evolution of Vibrio fischeri LuxR for improved response to butanoyl-homoserine lactone. Appl. Environ. Microbiol. 73, 5775–5781 (2007).

10. Mascher, T., Helmann, J. D. & Unden, G. Stimulus perception in bacterial signal-transducing histidine kinases. Microbiol. Mol. Biol. Rev. 70, 910–938 (2006).

11. Goh, E.-B. et al. Hierarchical control of anaerobic gene expression in Escherichia coli K-12: the nitrate-responsive NarX-NarL regulatory system represses synthesis of the fumarate-responsive DcuS-DcuR regulatory system. J. Bacteriol. 187, 4890–4899 (2005).

12. Gilles-Gonzalez, M.-A. & Gonzalez, G. Heme-based sensors: defining characteristics, recent developments, and regulatory hypotheses. J. Inorg. Biochem. 99, 1–22 (2005).

13. Hsieh, Y.-J. & Wanner, B. L. Global regulation by the seven-component Pi signaling system. FEMS Microbiol. Lett. 119, 291–297 (1994).

14. Martens, E. C., Chiang, H. C. & Gordon, J. I. Mucosal glycan foraging by multiple bacterial adrenergic sensors eavesdrop on host heme and modulates virulence. Cell. Host. Microbe 13, 198–203 (2011).

15. Torres, V. I. et al. A Staphylococcus aureus regulatory system that responds to host and modulates virulence. Cell. Host Microbe 1, 109–119 (2007).

16. Han, S. & Kang, I. J. Quorum sensing and signal-response systems in Gram-negative bacteria. Nat. Rev. Microbiol. 14, 576–588 (2016).

17. Hutchinson, M. L., Hons, H. J. & Buttner, M. J. T. The vancomycin resistance VanRS two-component signal transduction system of Streptomyces coelicolor. Mol. Microbiol. 59, 923–933 (2005).

18. Verhamme, D. T., Arents, J. C., Postma, P. W., Criel, A. & Wellingwerf, K. J. Glucose-6-phosphate-dependent phosphoryl flow through the Uhp two-component regulatory system. Microbiology 147, 3345–3352 (2001).

19. Sonnenburg, E. D. et al. A hybrid two-component system protein of a prominent human gut symbiont couples glycansensing in vivo to carbohydrate metabolism. Proc. Natl. Acad. Sci. USA 103, 8834–8839 (2006).

20. Martens, E. C., Chiang, H. C. & Gordon, J. I. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. Cell. Host. Microbe 4, 447–457 (2008).

21. Karavolos, M. H., Winzer, K., Williams, P. & Khan, C. M. A. Pathogen espionage: multiple bacterial adrenergic sensors eavesdrop on host communication systems. Mol. Microbiol. 87, 455–465 (2013).

22. Donoso, R. et al. Biochemical and genetic basis of indole-3-acetic acid (auxin phytohormone) degradation by the plant growth promoting rhizobacterium Paraburkholderia phytofirmans PsJN. Appl. Environ. Microbiol. 83, pii: 01991-16 (2016).

23. Singh, K. K. The Saccharomyces cerevisiae snl1p-ssk1p two-component system mediates response to oxidative stress and in an oxidant-specific fashion. Free Radic. Biol. Med. 29, 1043–1050 (2000).

24. Connor, J. R., Kuwada, N. J., Huangyutitham, V., Wiggins, P. A. & Harwood, C. S. Surface sensing and lateral subcellular localization of WspA, the receptor in a chemosensory-like system leading to c-di-GMP production. Mol. Microbiol. 86, 720–729 (2012).

25. Duanmu, D. et al. Marine algae and land plants share conserved phytochrome signaling systems. Proc. Natl. Acad. Sci. USA 111, 15827–15832 (2014).

26. Levskaya, A. et al. Synthetic biology: engineering Escherichia coli to see light. Nature 438, 441–442 (2005).

27. Ohlendorf, R., Vidavski, R. R., Eldar, A., Moffatt, K. & Most, C. A. Pathogen espionage: multiple bacterial adrenergic sensors eavesdrop on host communication systems. Mol. Microbiol. 87, 455–465 (2013).

28. Donoso, R. et al. Biochemical and genetic basis of indole-3-acetic acid (auxin phytomembrone) degradation by the plant growth promoting rhizobacterium Paraburkholderia phytofirmans PsJN. Appl. Environ. Microbiol. 83, pii: 01991-16 (2016).

29. O’Connor, J. R., Kuwada, N. J., Huangyutitham, V., Wiggins, P. A. & Harwood, C. S. Surface sensing and lateral subcellular localization of WspA, the receptor in a chemosensory-like system leading to c-di-GMP production. Mol. Microbiol. 86, 720–729 (2012).

30. Ramakrishnan, P. & Tabor, J. J. Repurposing synecocystis PCC8803 Urb–Urb as a UV-violet/green photoreversible transcriptional regulatory tool in E. coli. ACS Synth. Biol. 5, 733–740 (2016).

31. Ong, N. T. X. & Tabor, J. A. A miniaturized E. coli green light sensor with high dynamic range. Chembiochem Eur. J. Chem. Biol. (2018). https://doi.org/10.1002/cbic.201800007

32. Landry, B. P. & Tabor, J. J. Engineering diagnostic and therapeutic gut bacteria. Microbiol. Spectr. 2017.

33. Daefler, K. N.-M. et al. Engineering bacterial biosilicofluoride and tetrahionate sensors for detecting gut inflammation. Mol. Syst. Biol. 13, 923 (2017).
31. Riglar, D. T. et al. Engineered bacteria can function in the mammalian gut long term as live diagnostics of inflammation. *Nat. Biotechnol.* 35, 653–658 (2017).

32. Bhate, M. P., Molnar, K. S., Goulian, M. & DeGrado, W. F. Signal transduction in histidine kinases: insights from new structures. *Structure* 23, 981–994 (2015).

33. Huynh, T. N. & Stewart, V. Negative control in two-component signal transduction by transmitter phosphatase activity. *Mol. Microbiol.* 82, 275–286 (2011).

34. Atkinson, M. R. & Ninfa, A. J. Mutational analysis of the bacterial signal-transducing protein kinase/phosphatase nitrogen regulator II (NRII or NtrB). *J. Bacteriol.* 175, 7016–7023 (1993).

35. Wanner, B. L. in *Escherichia coli and Salmonella: Cellular and Molecular Biology* (eds Neidhardt, F. C. et al.) 2, 1357–1381 (ASM Press, Washington, 1996).

36. Zhu, Y. & Inouye, M. The role of the G2 box, a conserved motif in the histidine kinase superfamily, in modulating the function of EnvZ. *Mol. Microbiol.* 45, 653–663 (2002).

37. Qin, L., Cai, S., Zhu, Y. & Inouye, M. Cysteine-scanning analysis of the dimerization domain of EnvZ, an osmosensing histidine kinase. *J. Bacteriol.* 185, 3439–3443 (2003).

38. Capra, E. J. et al. Systematic dissection and trajectory-scanning mutagenesis of the molecular interface that ensures specificity of two-component signaling pathways. *PLoS Genet.* 6, e1001220 (2010).

39. Huynh, T. N., Nortjea, C. E. & Stewart, V. Missense substitutions reflecting regulatory control of transmitter phosphatase activity in two-component signalling. *Mol. Microbiol.* 88, 459–472 (2013).

40. Russo, F. D. & Silhavy, T. J. EnvZ controls the concentration of phosphorylated OmpR to mediate osmoregulation of the porin genes. *J. Bacteriol.* 172, 567–580 (1991).

41. Batchelor, E. & Goulian, M. Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system. *Proc. Natl. Acad. Sci. USA* 100, 691–696 (2003).

42. Gao, R. & Stock, A. M. Probing kinase and phosphatase activities of two-component systems in vivo with concentration-dependent phosphorylation profiling. *Proc. Natl. Acad. Sci. USA* 110, 672–677 (2013).

43. Iqbal, O. A., Alves, R. & Savageau, M. A. Hysteretic and graded responses in bacterial two-component signal transduction. *Mol. Microbiol.* 68, 1196–1215 (2008).

44. Jin, T. & Inouye, M. Ligand binding to the receptor domain regulates the ratio of kinase to phosphatase activities of the signaling domain of the hybrid Escherichia coli transmembrane receptor, Tar1. *J. Mol. Biol.* 232, 484–492 (1993).

45. Sutton, M. A. et al. *The European Nitrogen Assessment: Sources, Effects and Policy Perspectives* (Cambridge University Press, Cambridge, 2011).

46. Smanski, M. J. et al. Functional optimization of gene clusters by combinatorial design and assembly. *Mol. Microbiol.* 126, 1241–1249 (2014).

47. Tabor, J. J., Levkay, A. & Voigt, C. A. Multichromatic control of gene expression in Escherichia coli. *J. Mol. Biol.* 405, 315–324 (2011).

48. Ong, N. T. X., Olson, E. J. & Tabor, J. J. Engineering an E. coli near-infrared light sensor. *ACS Synth. Biol.* 7, 240–248 (2018).

49. Olson, E. J., Hartsgough, L. A., Landry, B. P., Shroff, R. & Tabor, J. J. Characterizing bacterial gene circuit dynamics with optically programmed expression signals. *Nat. Methods* 11, 449–455 (2014).

50. Baumgartner, J. W. et al. Transmembrane signalling by a hybrid protein: communication from the domain of the chemoreceptor Trg that recognizes sugar-binding proteins to the kinase/phosphatase domain of osmosensor EnvZ. *J. Bacteriol.* 176, 1157–1163 (1994).

51. Usuki, R. et al. Activation of bacterial porin gene expression by a chimeric signal transducer in response to aspartate. *Science* 245, 1246 (1989).

52. Schmidt, S. R., Sheth, R. U., Wu, A. & Tabor, J. J. Refactoring and optimization of light-switchable Escherichia coli two-component systems. *ACS Synth. Biol.* 3, 820–831 (2014).

53. Winter, S. E. et al. Gut inflammation provides a respiratory electron acceptor for Salmonella. *Nature* 467, 426–429 (2010).

54. Jeong, D.-W. et al. The auxiliary protein complex SaePQ activates the phosphorylation activity of sensor kinase SaeS in the SaeRS two-component system of Staphylococcus aureus. *Mol. Microbiol.* 86, 331–348 (2012).

55. Kim, E. K. & Choi, E.-J. Pathological roles of MAPK signaling pathways in human diseases. *Biochim. Biophys. Acta* 1802, 396–405 (2010).

56. Caunt, C. J. & Keyse, S. M. Dual-specificity MAP kinase phosphatases (MKPs). *FEBS J.* 280, 489–504 (2013).