Robust and tunable signal processing in mammalian cells via engineered covalent modification cycles

Ross D. Jones[1, 2], Yili Qian[2, 3], Katherine Ilia[1, 2], Benjamin Wang[2, 4], Michael T. Laub[2, 4, 5], Domitilla Del Vecchio[* 2, 3], and Ron Weiss[* 1, 2, 6]

[1]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

[2]Synthetic Biology Center, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

[3]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

[4]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

[5]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

[6]Electrical Engineering and Computer Science Department, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

[*]Correspondence should be addressed to D.D.V. (ddv@mit.edu). or R.W. (rweiss@mit.edu)

Abstract

Rewiring signaling networks imparts cells with new functionalities that are useful for engineering cell therapies and directing cell development. While much effort has gone into connecting extracellular inputs to desired outputs, less has been done to control the signal processing steps in-between. Here, we develop synthetic signal processing circuits in mammalian cells using proteins derived from bacterial two-component signaling pathways. First, we isolate kinase and phosphatase activities from the bifunctional histidine kinase EnvZ and demonstrate tunable phosphorylation control of the response regulator OmpR via simultaneous phosphoregulation by an EnvZ kinase and phosphatase. We show that modulation of phosphatase expression at the mRNA and protein levels via miRNAs and small molecule-regulated degradation domains, respectively, can effectively tune kinase-to-output responses. Further, we implement a novel phosphorylation-based miRNA sensor that effectively classifies cell types and enables cell type-specific kinase-output signaling responses. Finally, we implement a tunable negative feedback controller by co-expressing the kinase-driven output gene with the small molecule-tunable phosphatase, substantially reducing both gene expression noise and sensitivity to perturbations at the transcriptional and translational level. Our work lays the foundation for establishing tunable, precise, and robust control over cell behavior with synthetic signaling networks.

1 Introduction

Across all organisms, sensing and processing of environmental factors is critical for growth, proliferation, and survival[1]. Engineering of mammalian cells to transmute specific intracellular and extracellular inputs into desirable output behaviors has broad applications in cell therapy[2–7] and in the engineering of stem cells, tissues, and organoids[8–11]. To date, much work has been done to rewire natural signaling pathways and engineer synthetic receptors that sense extracellular inputs[6,12,13]. However, little work has been done to control the downstream signal processing that ultimately converts such signal inputs into changes in gene expression outputs.

An ideal engineered signaling system would have tunable input/output responses, low output noise, and drive gene expression that is robust to perturbations. Nearly all engineered receptors to date have utilized either native intracellular signaling domains or proteolytic mechanisms to transduce extraceullular signals[13]. Interfacing with the cell's natural signaling networks has been a powerful method to rewire signaling pathways[13], but it is difficult to modulate signaling between natural receptors and their gene expression targets due to the complexity of natural signaling networks in mammalian cells. Alternatively, using proteolysis to liberate gene regulators from the plasma membrane enables regulation independent from the cell signaling context through non-native proteins such as dCas9 or tTA[13]. However, since the effector proteins are irreversibly released from the receptor, the ability to tune the input-output response is limited to creating sequence variants or tuning the concentrations of receptor components[13–15]. Recently, synthetic receptors comprising extracellular receptors or dimerization domains fused to a bacterial two-component signaling (TCS) protein were shown to successfully transmute ligand inputs to TCS-regulated transcriptional outputs in mammalian cells[16,17].

The use of TCS proteins in synthetic mammalian signaling networks has the potential for creating tunable, robust signaling circuits that do not cross-react with existing networks in mammalian cells. TCS pathways typically comprise a transmembrane sensor protein called a histidine kinase (HK) and a cognate intracellular effector protein called a response regulator (RR). In response to specific signal inputs, the HK autophosphorylates on a conserved histidine residue and then transfers the phosphoryl group to a conserved asparate residue in the receiver (Rec) domain of the RR (referred to as the HK's kinase activity). Once phosphorylated, most RRs carry out transcriptional regulation, though other modes of regulation are possible[18,19]. Unlike typical eukaryotic receptors, in the absence of signal inputs, most HKs catalyze removal of the phosphoryl group from their cognate RRs (referred to the HK's phosphatase activity)[19,20]. The presence of signal input alters the conformational state of the HK, thereby tuning its relative kinase and phosphatase activities[21–23]. The bifunctional nature of HKs is important for insulating TCS pathways from off-target interactions[24–29] as well as for increasing the responsiveness to signal inputs[30]. The recently developed TCS/GPCR receptors work by coupling ligand-induced G-protein localization to the receptor to HK kinase activity and thus RR-driven gene expression[16]. TCS pathways are ubiquitous in bacteria, but are rare in eukaryotes

generally and absent in animals specifically[31]. The lack of any known examples of histidine-aspartate phosphorelays in mammalian cells thus strongly suggests that the introduced TCS signaling pathways are insulated from mammalian signaling pathways[16,32].

In phosphorylation cascades, reversing the phosphorylation of a substrate by a kinase via a phosphatase can impart tunability and robustness to perturbations into the system[33–38]. Here we introduce a framework for engineering signal processing circuits in mammalian cells based on synthetic covalent modification cycles built with bacterial TCS proteins (Figure 1). To develop these circuits, we isolate monofunctional kinases and phosphatases from the bifunctional *E. coli* HK EnvZ[39], then use specific phosphorylation and dephosphorylation of EnvZ's cognate RR OmpR to regulate downstream gene expression. First, we illustrate the tunability of this system by using the level of an EnvZ phosphatase to shift the sensitivity of OmpR-driven gene expression output to the levels of an EnvZ kinase. Further, we show that kinase-to-output dose responses can be tuned by regulating phosphatase expression with miRNAs and small molecule-inducible degradation domains. We then build upon this tunability to create novel phosphorylation-based genetic sensors that are capable of cell type classification and enable cell-type specific signaling responses.

A major challenge for developing synthetic genetic circuits is context-dependence due to factors such as off-target binding of gene regulators and loading of gene expression resources, which can perturb gene expression levels[40,41]. Currently, there is a lack of synthetic signaling circuits that are robust to such context effects in mammalian cells. To introduce robustness to such perturbations, we develop a feedback controller in which an EnvZ phosphatase is co-expressed with the output gene, thereby catalyzing reversal of phosphorylated (active) OmpR back to its inactive form. To introduce tunability into the strength of feedback and set level of output expression, we utilize our phosphatase fused to a small molecule-inducible degradation domain. The addition of feedback substantially reduces cell-to-cell noise in output expression and mitigates the effects of off-target translational repression and transcriptional resource loading on the signaling input-output response. Overall, we present the design and characterization of phosphorylation-regulated genetic modules that will enable tunable, precise, and robust control of signaling outputs in mammalian cells, with applications in engineering therapeutic cells, stem cells, and organoids.

2 Results

2.1 Isolation of kinase and phosphatase activity from EnvZ

As a model system for engineering synthetic signal processing circuits, we utilized the well-characterized EnvZ-OmpR TCS pathway from *E. coli*[42]. Both *in vitro* and *in vivo* in bacteria, it has been shown that the kinase and phosphatase activities of the bifunctional HK EnvZ can be isolated through various mutations[43–49], truncations[50,51],

and domain rearrangements[52]. In mammalian cells, it was shown that wild-type (WT) EnvZ is constitutively active[32], indicating that it has net-kinase activity, but may still retain some phosphatase activity. To create monofunctional kinases and phosphatases from EnvZ in mammalian cells, we generated several variants of EnvZ using established mutations, truncations, domain rearrangments, and novel combinations thereof (Figure 2 & Supplementary Figure 4).

To test for kinase activity, we evaluated the ability of EnvZ variants to activate an OmpR-driven reporter when transfected into HEK-293FT cells (Supplementary Figure 5). OmpR-activated promoters were made by placing three to nine OmpR binding sites upstream of a minimal CMV promoter or a synthetic minimal promoter (YB_TATA,[32], referred to as minKB), of which the 6x$OmpR_{BS}$-minCMV variant was chosen for use in most downstream experiments (Supplementary Figure 6). OmpR-driven gene expression induced by full-length mutants of EnvZ are shown in Figure 2a. From this initial screen, we identified two variants, EnvZm2 [T247A] and EnvZm2[AAB], the latter having an extra DHp domain fused to EnvZ[223+][52], that induced higher levels of output expression than WT EnvZ, suggesting that their phosphatase activity is reduced. Variants expected to be deficient in ATP binding or autophosphorylation based on previous studies in bacteria were indeed found to lack activation of OmpR-VP64, indicating that in mammalian cells they also lack kinase activity (Supplementary Figure 5). Moving forward, we used EnvZm2 as our kinase of choice due to the highly conserved ability of the T247A mutation to reduce or eliminate phosphatase activity in other HKs[49,53].

To test for phosphatase activity, we co-expressed EnvZm2 with OmpR-VP64 to generate phosphorylated OmpR-VP64 (P-OmpR-VP64), then evaluated the ability of EnvZ variants to deactivate expression of an OmpR-driven reporter (Supplementary Figure 7). While several EnvZ variants that were predicted to be phosphatases based on previous studies indeed showed deactivation of OmpR-driven expression at high concentrations, this deactivation was comparable to that of a variant that is predicted to lack any catalytic activity (EnvZm0m1m2m3 [H243A/D244A/T247A/N343K]) (Supplementary Figure 7). Thus, it is possible that these variants were primarily inhibiting output expression through sequestration of P-OmpR-VP64 from its target promoter, rather than through dephosphorylation. Indeed, high dosages of such variant (EnvZm1, [T247A]) can reduce "leaky" activation of the output reporter by non-phosphorylated OmpR-VP64, indicating that the observed reduction in output can occur absent dephosphorylation (Supplementary Figure 8). Notably, at both low (Supplementary Figure 7) and high (Supplementary Figure 8) dosages of the variant EnvZ[A] (DHp domain only), we found no apparent phosphatase activity, contrasting with the original report[51]. Only the variant EnvZm1[AAB], having an extra DHp domain fused to EnvZ[223+] with the mutation [D244A] in both DHp domains, was found to deactivate OmpR-driven expression more strongly than EnvZm0m1m2m3 (which lacks catalytic activity) (Supplementary Figure 7), suggesting EnvZm1[AAB] has real phosphatase activity in mammalian cells. However, at higher dosages of EnvZm1[AAB] and in the absence of EnvZm2, OmpR-VP64 appears to become activated, indicating that this variant may still retain some kinase activity (Supplementary Figure 8).

Because of the constitutive kinase activity of WT EnvZ and the lack of clear monofunctional phosphatase activity by purported phosphatase variants of EnvZ, we hypothesized that in mammalian cells, EnvZ takes a structural conformation that is unfavorable for phosphatase activity. To generate a functional phosphatase, we therefore followed the approach of Wang *et al.*[54] to force the alpha helices in the DHp domain into fixed rotational states using GCN4 leucine zippers (Figure 2c & Supplementary Figure 4). We generated a library of 10 rotationally-locked variants of EnvZ (EnvZt#1-10) with and without a mutation to eliminate ATP binding and hence kinase activity (m3 – [N343A])[44]. As expected, we observed a range of OmpR-driven gene expression levels that depend on the putative rotational angle of the DHp domain (Figure 2c). Interestingly, compared to WT EnvZ, all of the EnvZt# variants yielded equivalent or weaker output activation by OmpR-VP64, while also reducing EnvZm2-induced expression by at least 3-fold (Figure 2c). Comparing the exact levels of output with and without EnvZm2 in Figure 2c, we found that EnvZt# variants are capable of overwriting the initial phosphorylation state of OmpR-VP64 to ultimately set a defined level of output (Supplementary Figure 9). Most strikingly, all of the EnvZm3t# variants showed potent and nearly identical deactivation of OmpR-driven expression regardless of their rotational conformation, suggesting that all the GCN4-fused truncations posess similar phosphatase activities.

To more quantitatively compare the activation and deactivation of OmpR-driven expression by each of the EnvZ variants described above, we fit simple first-order models to estimate the dosages of each variant needed for half-maximal activation or deactivation ($K_{1/2}$) of the output (Supplementary Figure 10). Notably, the EnvZm3t# variants deactivated output expression at 2- to 3-fold lower dosages than our previous best putative phosphatase, EnvZm1[AAB], and at 10- to 20-fold lower dosages than the enzymatically null variant EnvZm0m1m2m3 (Supplementary Figure 10), suggesting more potent phosphatase activity. Moving forward, we chose to use EnvZm3t10 as our phosphatase because it has one of the lowest values of $K_{1/2}$ among all EnvZ variants and completely deactivates the output down to basal levels (Figure 2 & Supplementary Figure 10).

To ensure that the observed putative phosphatase activity is not explained by the formation of partially or completely inactive heterodimers between any putative phosphatases and EnvZm2, we repeated the experiments described above with CpxA in place of EnvZm2 (Supplementary Figure 11a). CpxA has weak off-target kinase activity for OmpR[28], and broadly, heterodimerization between different HKs is rare[55]. In the presence of CpxA, the putative phosphatases similarly, and in some cases more potently, deactivate OmpR-driven expression (Supplementary Figure 11b-c). Thus, the observed output deactivation is independent of how OmpR-VP64 is phosphorylated.

Direct verification of EnvZm3t10 phosphatase activity *in vivo* is challenging due to the acid-lability of phosphohistidine and phosphoaspartate bonds[56,57] and lack of commercial antibodies against P-OmpR. To verify that EnvZm3t10 acts as a phosphatase, we thus carried out additional control experiments. The deactivation of OmpR-driven output by EnvZm3t10 is abolished when adding mutations predicted to eliminate its phosphatase

activity, or using constitutively active variants of OmpR-VP64 (Supplementary Figure 12). Thus, the observed putative phosphatase activity is not caused by blocking interactions between the kinase and OmpR-VP64, nor by sequestration of OmpR-VP64. It is thus unlikely that EnvZm3t10 is acting through a mechanism other than direct dephosphorylation of P-OmpR-VP64.

2.2 Tuning kinase-output responses via phosphatase activity

Tunability in synthetic genetic devices has often been achieved through small molecule regulation[58–60], competitive binding[61], or sequestration of protein or RNA species[62–64]. With both a kinase and phosphatase, we constructed a tunable genetic device based on covalent modification cycles (Figure 3a). Factors that affect either the levels or activities of the kinase (u_K) or phosphatase (u_P) are considered inputs to the device. The output of the devices are the transcriptional and translational products driven by OmpR-VP64. To evaluate the tunability of our engineered covalent modification cycle, we compared the level of OmpR-VP64-driven output across combinations of kinase and phosphatase levels, with the phosphatase level regulated at the DNA, RNA, and protein levels (Figure 3b-d).

First, we titrated both kinase and phosphatase levels by dosing in different amounts of plasmid DNA per sample using poly-transfection[65] (Figure 3b). From the 2D input-output map (left), we can see that the level of output increases gradually with the ratio of kinase to phosphatase dosages. As the dosage of phosphatase increases, the amount of kinase needed to activate the output increases (center), indicating a decreased sensitivity to kinase input levels. Likewise, as the level of kinase increases, the amount of phosphatase needed to deactivate the output also increases (right). Both results were in accordance with standard models of covalent modification cycles[33] (see Supplementary Note 1 for our derivation).

Following the above results, we predicted that we could tune output expression through modulation of phosphatase expression (Figure 3c). To do so, we fused the phosphatase to small molecule-inducible degradation domains (DDs) DDd[66] and DDe[67], which are stabilized by addition of trimethoprim (TMP) and 4-hydroxytamoxifen (4-OHT), respectively. N-terminal fusions of both DDd and DDe showed the highest fold-changes in output expression upon addition of the cognate small molecule (Supplementary Figure 13); we chose to move forward with DDd/TMP for further testing due to lower background signal than DDe/4-OHT. Titration of both the kinase dosage and TMP concentration shows that the output is high only when the kinase is high and TMP is low (left). Addition of TMP decreases the sensitivity of the output to kinase (center) and addition of kinase decreases the sensitivity of the output to TMP (right). The response of the TMP-tuned design to kinase and TMP levels depends on the initial level of phosphatase in the cell. If the level of phosphatase initiates too high, the miRNA cannot suppress it enough to enable output induction by the kinase; conversely, if the initial level of phosphatase is too low, the kinase dominates

the covalent modification cycle even without any miRNA added (Supplementary Figure 14). Thus, there is an optimal level of phosphatase at which the ability of TMP to induce deactivation of gene expression is maximized.

Overall, these data demonstrate the ability of our engineered covalent modification cycle to function as a flexible building block for creating highly tunable signaling functions within mammalian cells.

2.3 Engineered, cell type-specific signaling responses

In addition to ectopically-expressed factors, endogenous cellular factors can also be inputs to the kinase (u_K) and phosphatase (u_K) in our engineered covalent modification cycle, enabling cell state sensing. One application for this sensing is in cell type classification, in which cellular factors such as TFs[68] and/or miRNAs[69] that are differentially expressed across cell types are used to limit gene expression to certain cell types. In miRNA-based classifiers, "low miRNAs" are expected to be low in the target cell type, and directly repress the output gene and/or regulators that activate its expression in non-target cell types[69]. In practice, "low miRNA" sensors are relatively simple to construct and their functions can be predicted through well-developed models of combinatorial miRNA activities[70]. Conversely, "high miRNAs" are expected to be high in the target cell type, and indirectly activate an output gene through double inversion[69] (repression of repressors). This double-inversion makes optimization of high miRNA sensors more challenging[65,71], with previous implementations being highly sensitive to the absolute level of the intermediate repressor molecule[65]. Since the output of the covalent modification cycle depends approximately on the ratio of kinase to phosphatase activity, rather than the absolute level of either (Figure 3b & Supplementary Figure 19), we built and tested a novel miRNA-based cell type classifier using covalent modification cycles (Figure 4a).

As a proof of concept, we placed four target sites for a cancer-associated miRNA, miRNA-21-5p (miR-21, T21), in the UTRs of the gene expressing the phosphatase (Figure 4b). As we have shown previously, miR-21 is differentially expressed in HeLa cells compared to HEK-293 cells, and can be used to distinguish the cell lines with high accuracy[65,69]. As a control, we placed four target sites for the synthetic miR-FF4 (TFF4)[72] in the UTRs of the phosphatase instead of T21. In cells expressing miR-21, we expected the phosphatase to be knocked down, thereby dramatically shifting the covalent modification cycle to favor phosphorylation of OmpR-VP64 and thus activation of the output. An endoribonuclease (endoRNase)-based incoherent feedforward loop (iFFL)[73] was included to constrain cell-to-cell variance in the expression level of OmpR-VP64 (Supplementary Figure 16). This was needed because P-OmpR has only a ~10-30-fold lower affinity for OmpR's DNA binding sites compared to OmpR[74], which we validated in HEK-293FT cells (Supplementary Figure 17). Thus, there is an ideal level of OmpR at which the on-off response is maximized. Due to the high DNA dosage variance of transfections, only a small subset of cells typically receive the ideal dosage of OmpR-VP64. With the feedforward controller, we can significantly reduce DNA dosage as a source of gene expression variance[73], and thereby increase the fraction of cells that receive the 'ideal' dosage of

OmpR-VP64 to improve selectivity (Supplementary Figure 16).

First, we considered the effect of miR-21 on the kinase-output dose-response curve. We predicted that endogenous expression of miR-21 could selectively sensitize output expression to kinase levels in HeLa cells. Indeed, when the phosphatase is present and highly expressed, it suppresses output induction by the kinase in HEK cells and in HeLa cells with the TFF4 target sites (rather than T21 sites) in the UTRs of the phosphatase (Figure 4d, left). Conversely, in HeLa cells with the T21 circuit variant, the kinase can induce output expression nearly as sensitively as conditions without the phosphatase (Figure 4d, right). Note that without the iFFL, the output expression has higher 'leaky' background expression at low ratios of kinase to phosphatase dosages (Supplementary Figures 18 & 19). Depending on the phosphatase dosage, the T21 variant in HeLa cells has between 10- to 1000-fold higher sensitivity to kinase input than the TFF4 variant (Supplementary Figure 22). Thus, these results illustrate a novel application of miR sensors for cell-type specific tuning of signaling responses.

To optimize our sensor for cell type classification, we followed the approach of Gam *et al.*[65] to systematically compare the percent of cell positive for output expression at different ratios of each circuit component using poly-transfection. We found that a 1:1:0.5 ratio of K:P:Output plasmids (the latter of which was co-delivered with the CasE/OmpR-VP64 iFFL) maximized classification accuracy for the T21 vs TFF4 variants in HeLa cells (Supplementary Figure 20). At this ratio, we obtained a significant ~50% increase in cells positive for the output reporter between the circuit variants in HeLa cells and a ~55% increase between HeLa and HEK-293 for the T21 variant (Figure 4d). The area under the curve (AUC) of the receiver operating characteristic (ROC) curve of the circuit was 0.83 ± 0.01 when comparing T21 vs TFF4 variants in HeLa cells and 0.93 ± 0.01 when comparing the T21 variant in HEK-293 vs HeLa cells (Supplementary Figure 20). Examining various combinations of dosages of the kinase, phosphatase, and output reporter, we found that the AUC of the resulting ROC-like curve of our phosphorylation-based classifier (0.93 ± 0.04 – Supplementary Figure 21) is higher than that of our recently-optimized transcriptional repressor-based classifier (0.84 – see SI Fig. 16 in Gam *et al.*[65]) for discriminating HEK vs HeLa cells, indicating improved performance for cell-type classification.

2.4 Design of a phosphorylation-based feedback controller

The response of expressed genes to their extracellular (or intracellular) inputs are often stochastic and thus imprecise across individual cells[75,76]. In addition, the intracellular context affects the level of gene expression induced by signaling[40,41], due to factors such as off-target interactions among engineered genes[58] or resource competition among genes[73,77,78]. To remedy these issues and enable the construction of genetic devices that enforce precise and robust signaling responses across cells, we developed a phosphorylation-based negative feedback controller of gene expression (Figure 5a). In both natural[79] and synthetic[60,80,81] systems, negative feedback reduces

the cell-to-cell variance of gene expression in response to signal inputs. Negative feedback has also been used to make gene expression robust to perturbations that affect processes within the feedback loop[63,64,82]. In our controller, the phosphatase is co-expressed with the output gene via a 2A linker[83] and suppresses its own production via dephosphorylation of P-OmpR-VP64 (Figure 5b). We further utilize the TMP/DDd-regulated phosphatase to tune the feedback strength through TMP.

The level of output set by the controller is dictated by the competitive phosphorylation of OmpR-VP64 by the kinase and phosphatase. In an ideal system operating with both enzymes saturated, the concentrations of the phosphatase and the output species become insensitive to disturbances affecting their gene expression processes (see Model Box).

As TMP selectively regulates phosphatase but not output stability, it can be used as an input to the controller to tune the strength of the feedback. Under the ideal conditions presented above and as long as OmpR-VP64 has not saturated the output promoter, the relationship between the levels of kinase and output is independent of both the exact mechanism by which OmpR-VP64 activates output expression as well as of any perturbations in the transcription and translation processes of the output/phosphatase (see Model Box).

To evaluate the performance of the feedback controller, we first measured the kinase-output responses for open loop (OL) and closed loop (CL) variants. The OL system was made by replacing the phosphatase with Fluc2, which has no effect on OmpR phosphorylation (Figure 2b-c). Since the presence of negative feedback reduces the level of output expression for a given input level of kinase, we tested several OL variants in which the amount of output reporter in transfections was reduced by 3x, 9x, 27x, or 81x (respectively referred to as Fluc2/3, Fluc2/9, Fluc2/27, and Fluc2/81). We define kinase responsiveness as the maximal fold-change in output between the absence and presence of kinase. The kinase responsiveness of the OL systems varies from ~10- to ~55-fold, whereas that of the CL system is ~3.5-fold for the variant without DDd fused to the phosphatase (Figure 5c, left – see Supplementary Figures 23-26 for full poly-transfection scheme and data). Adding DDd to the phosphatase increases the responsiveness to ~7.6-fold without TMP, or 6.4-fold for the lowest non-zero amount of TMP that we tested: 0.001 μM (Figure 5c, right). The kinase responsiveness of the CL system decreases as more TMP is added (and thus the phosphatase is stabilized) to the point of approximately matching that of the non-DDd CL system (Figure 5c, right). The maximum output level of the DDd CL system is up to 10-fold higher than that of the non-DDd CL system and within ~5-fold of that of the OL system. Thus, tuning the feedback strength via TMP allows the CL system to recover much of the dynamic range of the OL system.

In the absence of kinase, we see similar levels of noise in output expression for all OL and CL variants; however, as the dosage of kinase input is increased, we observe a decrease in noise for CL variants and an increase in noise for OL variants (Figure 5d). At high dosages of kinase, the output noise for OL devices decreases again, but does not reach the low noise achieved in CL devices. The higher noise in OL systems can be attributed to a more digital-like

transition in output expression per cell as the kinase dosage increases, whereas in CL systems we observe a smooth, unimodal shift in output expression per cell (Figure 5e, see Supplementary Figure 27 for all variants). The decrease in noise in CL expression as a function of increasing kinase likely can be attributed to the increasing concentration of P-OmpR-VP64 on which the phosphatase can actuate negative feedback. Interestingly, tuning the feedback strength with TMP appears to have little effect on the level of output noise observed (Supplementary Figure 28), suggesting that the degradation of the phosphatase does not push the system into a regime whereby the negative feedback is significantly negated. Compiling together all CL and OL variants and measuring noise as a function of output level, we can see that the noise in the OL systems peaks at intermediate absolute levels of output (regardless of the kinase dosage needed to achieve such an output level for a given OL variant), whereas the noise in the CL systems decreases as the output increases due to the factors described above (Figure 5f). The pattern of noise in the OL variants can potentially be explained by stochastic transcriptional variation among cells when the output promoter is not saturated – a phenomenon that the CL system would be able to sense and respond to in order to suppress noise.

2.5 Robustness to perturbations via feedback control

According to our mathematical modeling comparing the OL and CL circuits, the presence of negative feedback is expected to impart robustness to perturbations that affect the gene expression process of the output protein (see Model Box). We thus tested the capability of the CL system to impart robustness of output expression levels to two perturbations that model off-target regulation and resource loading (Figure 6a). For off-target regulation, we expressed miR-FF4, which binds and cleaves a target site placed in the 3' UTR of the output/phosphatase mRNA, thereby causing mRNA degradation. For resource loading, we expressed Gal4-VPR, which strongly sequesters transcriptional resources and thereby reduces transcription of other genes[73]. In addition to the modeled effects, these perturbations are useful because they affect output production both before (Gal4-VPR) and after (miR-FF4) transcription, so we can compare the ability of the CL system to respond to perturbations at different stages of expression. Robustness is quantified through fold-changes to perturbations and the robustness score (100% minus the percent deviation from the unperturbed level).

As expected, we found that the CL system is indeed more robust to miR-FF4 and Gal4-VPR perturbations than comparable OL variants (Figure 6b-e). Detailed comparisons of the response of all OL and CL variants to both perturbations are provided in Supplementary Figures 29 & 30. For illustration, we highlight and compare two OL and two CL variants with similar basal output levels in the absence of kinase (Fluc2, Fluc2/3, EnvZm3t10, DDd-EnvZm3t10 + 0.001 μM TMP – Figure 6b). Without kinase, there is little difference between the effects of miR-FF4 and Gal4-VPR on the OL and CL systems (Figure 6c, left panels), consistent with the expected lack of feedback actuation in the absence of P-OmpR-VP64 and our earlier findings of similar levels of noise in the same

regime (Figure 5d). At higher kinase input levels, the fold-changes in output expression for the CL variants in response to both perturbations are substantially less than those of the OL variants (Figure 6c, right panels). The relative decrease in fold-changes as a function of kinase input dosage is clearly illustrated in Figure 6d for two levels of miR-FF4 and Gal4-VPR perturbations that knock down the OL systems to similar degrees. At medium-to-high kinase input levels, the feedback controller can sustain the output level to within 2-4-fold of the nominal (unperturbed) levels, improving significantly over the 6-10-fold changes observed in the OL systems. The smaller relative decreases in fold-changes of the CL variants to Gal4-VPR with increasing kinase may result from generally weaker effects of transcriptional resource sequestration on basal transcription vs activated transcription[84], causing the perturbation to be weaker in the low-kinase regime where the CL system is more susceptible to perturbations.

Since both miR-FF4 and Gal4-VPR knock down gene expression, a full comparison of their effects on the OL and CL systems must account for the reduction in nominal output level caused by negative feedback. This is because lower nominal output levels will naturally have a reduced measurable dynamic range of knockdown due to detection limits imposed by the autofluorescence background. Comparing the robustness scores of all CL and OL variants at the same miR-FF4 and Gal4-VPR dosages as in Figure 6d, we can see that the CL systems are nearly always more robust than the OL systems for a given nominal output level (Figure 6e). The only substantial overlap in the plots between the OL and CL systems occurs at low kinase inputs to the CL system. Quantitatively, for a given nominal output level, we see a 20-30 percentage point increase in robustness score for the CL systems compared to the OL variants. Comparisons across additional dosages of each perturbation show similar results (Supplementary Figures 31 & 32). Thus, our phosphorylation-based feedback controller is capable of reducing the impact of perturbations both before and after transcription on the expression of the output gene. Coupled with the reduction in noise (Figure 5), these data indicate that the feedback controller can successfully impart precise, tunable, and robust control over gene expression in mammalian cells.

3 Model Box

Here we develop a mathematical model to show that covalent modification cycle (CMC)-mediated feedback enables the expression level of a regulated gene to be robust to disturbances. In particular, for a fixed kinase level (K_t), we treat the genetic circuit shown in Figure 6 as feedback interconnection of two dynamical processes with input/output (I/O): an engineered CMC that takes phosphatase concentration (P_t) as input and outputs P-OmpR-VP64 concentration (X^*), and a gene expression process that takes X^* as input to produce the phosphatase P_t as output. We use a standard Goldbeter-Koshland model[33] for the dynamics of the CMC:

$$\frac{\mathrm{d}}{\mathrm{d}t}X^* = \theta_k \frac{(X_t - X^*)K_t}{(X_t - X^*) + K_{M,k}} - \theta_p \frac{X^* P_t}{X^* + K_{M,p}}, \tag{1}$$

where θ_k and θ_p are catalytic rate constants of the kinase and the phosphatase, respectively, $K_{M,k}$ and $K_{M,p}$ are their respective Michaelis-Menten constants, and X_t is the total amount of OmpR-VP64 (*i.e.*, OmpR-VP and P-OmpR-VP). The expression of P_t is regulated by an OmpR-activated promoter, which gives rise to the following dynamics:

$$\frac{\mathrm{d}}{\mathrm{d}t}P_t = \alpha(1-w)\phi(X^*) - \gamma P_t, \tag{2}$$

where α is the production rate of P_t that lumps the rate constants for transcription, translation, and mRNA decay, $\phi(\cdot)$ is a Hill function satisfying $\phi' > 0$ for all X^*, γ is the protein decay rate constant, and $0 \leq w < 1$ is a disturbance that models the fold change in production rate of P_t, which could either arise from indirect transcriptional repression via resource loading or from direct post-transcriptional repression via miRNA (see Figure 6). The output from this feedback-regulated gene is $Y = \rho P_t$, since the output protein and phosphatase are co-transcribed but produced as separate proteins using a 2A-linker. We find that the relative sensitivity of output to disturbance w for this closed-loop system (1)-(2) at a given output level Y is

$$\mathcal{S}_{\mathrm{CL}}(Y) = \frac{1}{Y}\cdot\left|\frac{\mathrm{d}Y}{\mathrm{d}w}\right| = \frac{1}{1-w}\left[1+\frac{\alpha}{\gamma}(1-w)\left|\frac{\mathrm{d}}{\mathrm{d}Y}(\phi\circ h)\right|\right]^{-1}. \tag{3}$$

where h is the transfer curve of the CMC. In comparison, when the CMC in (1) is not connected with (2), the relative sensitivity of y to disturbance w for the open-loop system (2) is $\mathcal{S}_{\mathrm{OL}} = \frac{1}{1-w}$. Hence, we have $\mathcal{S}_{\mathrm{CL}} < \mathcal{S}_{\mathrm{OL}}$ for all y regardless of where the sensitivity is evaluated. This implies that the closed-loop system is always more robust than the open-loop system to disturbance w. To enable near-perfect adaptation to w, it is sufficient to increase $T := \left|\frac{\mathrm{d}}{\mathrm{d}y}(\phi\circ h)\right| = \left|h'\cdot\phi'\right|$. In particular, if $T \to \infty$, then $\mathcal{S}_{\mathrm{CL}} \to 0$, implying that the closed-loop system can perfectly adapt to w. Specifically, for any fixed X^* and y, there exists sufficiently small $K_{M,p}$ and sufficiently large X_t to make $|h'|$ arbitrarily large. On the other hand, to ensure T is large, $|\phi'|$ must not be too small. This requires us to design the system so that the OmpR-activated promoter is not saturated. Hence, the K_D of binding between phospohrylated OmpR and its target promoter must not be too small[38]. Promoter saturation limits the ability of the output to respond to changes in OmpR phosphorylation, and thus can limit the benefit of the negative feedback to achieve robustness to perturbations. Under the ideal operating conditions described above, both enzymes are saturated because $K_{M,p}$ is small and X_t is large. Specifically, if $K_{M,p} \ll X^*$ and $X_t \gg K_{M,k}$, equation (1) can be approximated by $\mathrm{d}X^*/\mathrm{d}t = \theta_k K_t - \theta_p Y/\rho$, leading to quasi-integral feedback control[38].

4 Discussion

Here, we developed tunable, precise, and robust signaling circuits in mammalian cells using engineered covalent modification cycles derived from bacterial two-component signaling (TCS) proteins (Figure 1). We first screened engineered variants of the *E. coli* histidine kinase (HK) EnvZ to isolate kinase and phosphatase activity from the

bifunctional protein (Figure 2). We demonstrated the tunability in kinase-induced gene expression responses conferred by small molecule-inducible expression of a phosphatase (Figure 3). Building upon this tunability, we showed that endogenous miRNAs can be used to create cell type-specific signaling responses through knockdown of phosphatase expression. Co-expressing the phosphatase with the output, we created a tunable negative feedback loop that reduces both the cell-to-cell variation and sensitivity to perturbations of kinase-induced gene expression (Figure 5 & 6).

Combined with recent advances in utilizing TCS proteins to engineer synthetic receptors in mammalian cells[16,17] and rewire the specificity of response regulators (RRs)[85], our results will enable the construction of sophisticated synthetic signaling systems in mammalian cells. While much work has so far focused on synthetic receptor engineering[13], incorporation of downstream signal processing moieties to improve signaling pathway function has only recently begun to be explored[86]. In particular, the ability to easily tune signaling pathway activity through phosphatase expression and the ability to robustly control the downstream gene expression processes will facilitate the creation of synthetic signaling systems that can operate across diverse cellular contexts. In addition, connecting signaling pathway activity to endogenous gene regulation, such as through miRNA regulation of pathway components, will facilitate applications in guiding differentiation or programming signaling to adapt to different cellular states.

The high degree of orthogonality among existing TCS pathways[87–89] and the relative ease of finding new orthogonal HK/RR pairs[90] indicates that TCS pathways will be a bountiful source of orthogonal signaling pathways for use in mammalian cells. To support this effort, we identified several HK-RR pairs that show good orthogonality in mammalian cells (Supplementary Figure 1-3). Though TCS is absent in animals[31], histidine and aspartate phosphorylation is more prevalent than previously thought[57]. The lack of observed histidine to aspartate phosphotransfer in animals likely guarantees orthogonality between TCS pathways and existing signaling networks in animal cells, though future work will be needed to examine possible cross-talk.

From our testing of EnvZ variants for isolated kinase or phosphatase activities, we found that the phosphatase state of EnvZ was nearly equally competent across all fixed rotational conformation of its DHp domains (Figure 2 & Supplementary Figure 10). The autokinase ability of an HK is dependent on the rotational state of the DHp domain[22,54,91,92]. While the phosphatase state of HKs is relatively unstable and subject to conformational changes caused by signaling[21,93], our data suggests that EnvZ can nonetheless dephosphorylate P-OmpR regardless of the rotational state of its DHp domain. As signal inputs can affect the autokinase activity, phosphatase activity, or both of an HK[19,42], future work will be needed to determine the degree to which the observation of a robust phosphatase state in EnvZ generalizes to other HKs.

Through the implementation of feedback control via covalent modification cycles, we have opened the door to creating increasingly precise and robust responses in engineered signaling pathways. Reducing cell-to-cell variation

in signaling output can be critical for ensuring that cells in a population make uniform, rather than multi-modal or stochastic, decisions. Reducing sensitivity of output expression to perturbations further constrains cellular decision-making and will help to ensure that engineered signaling systems can operate across diverse cell types[73]. In the future, it may be possible to further improve the robustness to perturbations conferred by our feedback controller. To achieve near-perfect adaptation to perturbations, the system parameters need to be tuned such that it can operate as a quasi-integral feedback controller[38,94]. We identified that the K_M of the phosphatase is likely similar to or higher than the K_D of P-OmpR, reducing the efficacy of the feedback (see Model Box). Increasing $\frac{K_D}{K_M}$ helps to ensure that the phosphatase is saturated with P-OmpR while the output promoter is not, both of which are critical conditions for the feedback to work in a quasi-integral manner[38] (see Model Box). Further discussion of possible future approaches to achieve quasi-integral feedback control with our system are discussed in Supplementary Note 2.

In natural systems, feedback control plays a critical role in regulating signaling pathway activities. Both negative and positive feedback are common in TCS pathways[95–97]. As with the robustness to perturbations conferred by our feedback controller, negative feedback in natural and engineering TCS pathways in bacteria allows for adaptation to signal inputs[36,87,96,98,99]. A conceptually similar controller to our design is found in bacterial chemotaxis, in which feedback control via reversible methylation of the receptor protein Tar enables near-perfect adaptation of flagellar motion to chemoattractants[94,100]. Another close analog of our controller can also be found in the human ERK1/2 MAPK (mitogen-activated protein kinase) pathway[101]. In this pathway, Mek is analogous to our HK kinase, Erk is analogous to OmpR-VP64 (though Erk itself only indirectly activates transcription through its targets[102,103]), and the Erk-induced phosphatases DUSP5/6 are analogous to our HK phosphatase[101,104]. Negative feedback in both natural and engineered systems has been shown to convert digital, multimodal input-output responses to more graded, linear, and uniform responses[79,80,105]. Likewise, our feedback controller is capable of imparting graded, uniform activation of gene expression in the cell population. These examples highlight how the topology of our feedback controller finds use in natural systems and will thus serves as a key building block for future synthetic signaling pathways.

In addition to feedback control, natural signaling pathways also incorporate constitutive phosphatase and regulators thereof to tune signaling functions across diverse cell types. For example, signaling through the T cell receptor (TCR) is regulated by several inhibitory receptors such as CD45 and phosphatases such as PTPN22, which help to suppress TCR pathway activation unless sufficiently high stimulus is encountered[106]. In developing thymocytes, miR-181a-5p suppresses the expression of PTPN22, thereby allowing for TCR pathway stimulation at lower antigen affinities, providing critical signals for survival and development towards mature T cells[107,108]. In mature T cells, a variety of miRNAs regulate TCR signaling, other signaling pathways, the cell cycle, and secretion, thereby tuning the immunological responses of T cells to their environments[109] Thus, tunable phosphatases and miRNA-regulated signaling responses like what we developed can be powerful tools for achieving stage-specific control of differentiation and tuning cell behavior in different contexts. In addition, future designs may incorporate

miRNAs that regulating kinase expression to provide a further layer of tunability.

In conclusion, we have engineered synthetic covalent modification cycles in mammalian cells using proteins derived from bacterial TCS. The signaling responses of these cycles are tunable and can be made precise and robust through negative feedback control. In the future, the customizable signaling responses enabled through our platform may be combined with engineered receptors[16] and modular effectors[85] to engineer complex signaling pathways that transmute extracellular inputs to various intracellular functions in mammalian cells. Such engineered signaling pathways will enable precise sensing and responding to cell-cell signaling, with applications in engineering cell therapies and programming development of stem cells into specific cells, tissues, and organoids.

5 Methods

Modular plasmid cloning scheme

Plasmids were constructed using a modular Golden Gate strategy similar to previous work in our lab[65,110]. Briefly, basic parts ("Level 0s" [pL0s] – insulators, promoters, 5'UTRs, coding sequences, 3'UTRs, and terminators) were created via standard cloning techniques. Typically, pL0s were generated via PCR (Q5 and OneTaq hot-start polymerases, New England BioLabs (NEB)) followed by In-Fusion (Takara Bio) or direct synthesis of shorter inserts followed by ligation into pL0 backbones. Oligonucleotides were synthesized by Integrated DNA Technologies (IDT) or SGI-DNA. pL0s were assembled into transcription units (TUs – "Level 1s" [pL1s]) using BsaI Golden Gate reactions (10-50 cycles between 16degC and 37degC, T4 DNA ligase). TUs were assembled into multi-TU plasmids ("Level 2s" [pL2s]) using SapI Golden Gate reactions. All restriction enzymes and T4 ligase were obtained from NEB. Plasmids were transformed into Stellar *E. coli* competent cells (Takara Bio). Transformed Stellar cells were plated on LB agar (VWR) and propagated in TB media (Sigma-Aldrich). Carbenicillin (100 μg/mL), kanamycin (50 μg/mL), and/or spectinomycin (100 μg/mL) were added to the plates or media in accordance with the resistance gene(s) on each plasmid. All plasmids were extracted from cells with QIAprep Spin Miniprep and QIAGEN Plasmid Plus Midiprep Kits. Plasmid sequences were verified by Sanger sequencing at Quintara Biosciences. Genbank files for each plasmid and vector backbone used in this study are provided in Supplementary Data. Plasmid sequences were created and annotated using Geneious (Biomatters).

In addition to the above, we devised a new scheme for engineering synthetic promoters using what we call "Level Sub-0" (pSub0) plasmids. The approach for creating promoters from pSub0 vectors is illustrated in Figure 33. In short, promoters are divided into up to 10 pSub0 fragments. Because the core elements of a promoter are typically at the 3' end, we made the pSub0 position vectors start with the 3'-most element and move towards the 5' of the promoter. Promoter position 1 (pP1) contains the transcription start site (TSS) – the +1 position for transcription initiation. pP1 can also optionally contain transcriptional repressor binding sites (not done in this study). pP2

contains the TATA box and other upstream core promoter elements[111–113] as desired. Many of the pP1 and pP2 sequences were derived from the minimal promoters studied by Ede *et al.*[114]. Because the spacing between the TATA box and +1 site are critical[115], we broke apart each minimal promoter at equivalent positions such that they can be interchanged. pP1 and pP2 are generally created via PCR reactions using the base pSub0 backbone as a template and adding the insert via primer overhangs and In-Fusion cloning. Position 3-10 (pP3-10) are 'enhancer' positions, wherein we generally encode binding sites for transcriptional activators (such as the RRs in this study), or sequences from constitutive promoters (not done in this study). pP3-10 plasmids are made by directly ligating annealed primers into pSub0 pP3-10 backbones. The annealed primers are synthesized with 4 bp offsets at each end to naturally create overhangs when annealed. All pSub0 plasmids include BsaI binding sites in an analogous position to pL0s, such that pSub0s can be used directly in place of pL0s when generating pL1s (the overhangs are compatible for up to four pSub0 inserts, see Supplementary Table 1). pSub0s use BsaI for cloning, they cannot be efficiently inserted into pL0 backbones using Golden Gate. To most efficiently clone pSub0s into pL0 P.2 (level 0 promoter) plasmids, we first perform a Golden Gate reaction with the pSub0s separately from the pL0 backbone, then ligate the Golden Gate product with a pre-fragmented and gel-extracted pL0 backbone.

Cell culture

HEK-293 cells (ATCC), HEK-293FT cells (Thermo Fisher), and HeLa cells (ATCC) were maintained in Dulbecco's modified Eagle media (DMEM) containing 4.5 g/L glucose, L-glutamine, and sodium pyruvate (Corning) supplemented with 10% fetal bovine serum (FBS, from VWR). All cell lines used in the study were grown in a humidified incubator at 37deg and 5% CO_2. All cell lines tested negative for mycoplasma.

Transfections

Cells were cultured to 90% confluency on the day of transfection, trypsinized, and added to new plates simultaneously with the addition of plasmid-transfection reagent mixtures (reverse transfection). Transfections were performed in 384-, 96-, 24-, or 6-well pre-treated tissue culture plates (Costar). Following are the volumes, number of cells, and concentrations of reagents used for 96-well transfections; for 384-, 24- and 6-well transfections, all values were scaled by a factor of 0.2, 5, or 30, respectively. 120 ng total DNA was diluted into 10 μL Opti-MEM (Gibco) and lightly vortexed. For poly-transfection experiments, the DNA dosage was subdivided equally among each complex (*e.g.* for two complexes, we delivered 60 ng DNA in each, 40 ng for three complexes, *etc.*) The transfection regent was then added and samples were lightly vortexed again. The DNA-reagent mixtures were incubated for 10-30 minutes while cells were trypsinized and counted. After depositing the transfection mixtures into appropriate wells, 40,000 HEK-293, 40,000 HEK-293FT, or 10,000 HeLa cells suspended in 100 μL media were added. Lipofectamine 3000 was used at a ratio of 2 μL P3000 and 2 μL Lipo 300 per 1 μg DNA. PEI MAX (Polysciences VWR) was used at a ratio of 3 μL PEI per 1 μg DNA. FuGENE6 (Promega) was used at a ratio of 3 μL FuGENE6 per 1 μg DNA. Viafect (Promega) was used at a ratio of 3 μL Viafect per 1 μg DNA. The media of the

transfected cells was not replaced between trasnfection and data collection. For all transfections with TMP (Sigma-Aldrich) or 4-OHT (Sigma-Aldrich), the small molecules were added concurrently with transfection complexes. In each transfection reagent-DNA complex, we included a hEF1a-driven transfection marker to indicate the dosage of DNA delivered to each cell.

Luciferase assays and analysis

To measure RR-driven luminescence output in Supplementary Figure 1, we used the Promega Nano-Glo Dual-Luciferase Reporter Assay System, following the manufacturer's instructions. Briefly, 6,000 HEK-293FT cells were transfected using the FuGENE6 reagent with 25 ng total DNA comprising the plasmids hPGK:Fluc2 (pGL4.53), an hEF1a-driven HK, an hEF1a-driven RR, an RR-driven promoter expressing NanoLuc, and filler DNA at 5 ng each. The cells were cultured in 20 uL DMEM supplemented with 10% FBS in 384-well plates with solid white walls and bottoms (Thermo Fisher) to facilitate luminescence measurements. 48 hours post-transfection, cells were removed from the incubator and allowed to cool to room temperature. 20 μL of ONE-Glo EX Reagent was added directly to the cultures, and cells were incubated for 3 minutes on an orbital shaker at 900 revolutions per minute (RPM). Fluc2 signal was measured on a BioTek Synergy H1 hybrid reader, with an integration time of 1 s. 20 μL of NanoDLR Stop & Glo Reagent was then added, and cells were again incubated for 3 minutes on an orbital shaker at 900 RPM. After waiting an additional 10 minutes following shaking, NanoLuc signal was measured on the same BioTek plate reader, with an integration time of 1 s. NanoLuc signals were normalized by dividing by the Fluc2 signals, thereby accounting for differences in transfection efficiency among wells.

Identification of optimal orthogonal TCS pairs

To identify the optimal set of orthogonal TCS interactions, we ran a MATLAB script to score all possible combinations of 4-7 HK-RR protein pairs. The script uses a scoring function to evaluate each particular subset of HKs and RRs. The data input into the scoring function is a matrix of output expression levels driven by the RRs in the presence of the selected HKs. The scoring function first identifies a reference value for each row and column by iteratively finding the maximum value in the matrix, blocking off the rest of the values in its row and column, then repeating until each row and column has one reference value. The reference value is then divided by the rest of the values in its row and column, and the quotients are multiplied together to give a score. The scores for each reference value are then again multiplied together to get a final score for a particular combination of HKs and RRs. After iterating through all possible such combinations, the highest final score for a given submatrix size is selected. The method gave qualitatively orthogonal combinations for up to 7 TCS pairs; we thus present the optimized 7-matrix in Supplementary Figure 1.

Flow cytometry

To prepare samples in 96-well plates for flow cytometry, the following procedure was followed: media was aspirated, 50 μL PBS (Corning) was added to wash the cells and remove FBS, the PBS was aspirated, and 40 μL

Trypsin-EDTA (Corning) was added. The cells were incubated for 5-10 minutes at 37deg C to allow for detachment and separation. Following incubation, 80 μL of DMEM without phenol red (Gibco) with 10% FBS was added to inactivate the trypsin. Cells were thoroughly mixed to separate and suspend individual cells. The plate(s) were then spun down at $400 \times g$ for 4 minutes, and the leftover media was aspirated. Cells were resuspended in 170 μL flow buffer (PBS supplemented with 1% BSA (Thermo Fisher), 5 mM EDTA (VWR), and 0.1% sodium azide (Sigma-Aldrich) to prevent clumping). For prepping plates of cells with larger surface areas, all volumes were scaled up in proportion to surface area and samples were transferred to 5 mL polystyrene FACS tubes (Falcon) after trypsinization. For standard co-transfections, 10,000-50,000 cells were collected per sample. For the poly-transfection experiment and transfections into cells harboring an existing lentiviral integration, 100,000-200,000 cells were collected per sample.

For all experiments, samples were collected on a BD LSR Fortessa equipped with a 405nm laser with 450/50nm filter ('Pacific Blue') for measuring TagBFP or EBFP2, 488 laser with 530/30 filter ('FITC') for measuring EYFP or mNeonGreen, 561nm laser with 582/15nm filter ('PE') or 610/20nm filter ('PE-Texas Red') for measuring mKate2 or mKO2, and 640 laser with 780/60nm filter ('APC-Cy7') for measuring iRFP720. 500-2000 events/s were collected either in tubes via the collection port or in 96-well plates via the high-throughput sampler (HTS). All events were recorded and compensation was not applied until processing the data (see below).

Flow cytometry data analysis

Analysis of flow cytometry data was performed using our MATLAB-based flow cytometry analysis pipeline (`https://github.com/Weiss-Lab/MATLAB_Flow_Analysis`). Basic processing steps follow the procedures described previously[73]. In addition, we frequently utilized our new poly-transfection technique and associated methods[65] to characterize and optimize circuits. Poly-transfection enables rapid and accurate assessment of dose-response curves for genetic components[65], such as the kinases and phosphatases in our circuits. Full schematics describing each poly-transfection experiment are shown in the SI (*e.g.* Supplementary Figure 5a).

Multi-dimensional binning of poly-transfection data was performed by first defining bin edges in each dimension (*i.e.* for the transfection markers for each poly-transfection complex), then assigning each cell to a bin where the cell's expression of these markers was less-than-or-equal-to the high bin edges and greater-than the low bin edges. Bins with three or fewer cells were ignored (values set to NaN in the MATLAB code) to avoid skewing by outliers in sparsely-populated samples (*e.g.* HeLa cells). Such binning is demonstrated via colorization of cells by their bin assignment in the SI (*e.g.* Supplementary Figure 5b). In order to avoid the artefact of negative fold-changes, non-positive fluorescence values were discarded prior to making measurements on binned or gated populations. In the second and third experimental repeats of the miRNA-dependent signaling/classifier data in Figure 4 and Supplementary Figures 18-22, a newly-prepared Output Marker plasmid was later discovered to have ~8-fold lower concentration than expected due to a measurement error on the nanodrop. To account for this, the bins for the Output

Marker in those samples are shifted down by 10x (so as to match the same bin boundaries as in the first repeat).

To find the optimal ratio of components in the miR-21 sensor for high cell classification accuracy, we scanned ratios between 1000:1 to 1:1000 of K:P and output plasmid:K/P, roughly halving the ratio between steps. At each combination of ratios, a trajectory was computed and all cells within 0.25 biexponential units of the trajectory based on euclidean distance were recorded. Accuracy was computed as described below, and accuracy values were compared across all ratios for each experimental repeat. From this scanning of trajectories at different ratios of components, we found that a 1:1:0.5 ratio of K:P:Output plasmid gave the highest accuracy. This optimal trajectory was used to sub-sample cells for display in Figure 4f & Supplementary Figure 20, finding percent positive for output in Figure 4g and calculating ROCs/AUCs in Supplementary Figure 20.

In the case of simple co-transfections and sub-sampled trajectories, cells were considered to be transfected if they were positive for the output/transfection marker *or* the output reporter. When computing summary statistics from binned data, such thresholding is unnecessary since binning already isolates the cell sub-population for measurement.

Calculation of cell classification metrics

Sensitivity was defined as the percent of cells positive for the output reporter in HeLa cells transfected with the T21 circuit variant. Specificity was defined as 100 minus the percent of cells positive for the output in HeLa cells with the TFF4 variant or in HEK-293 cells with the T21 variant. The former was considered the more ideal comparison for evaluating classification performance due to higher overall expression of the circuit in HeLa cells compared to HEKs (Supplementary Figure 18). Accuracy was computed by averaging sensitivity and specificity.

ROC curves in Supplementary Figure 20 were generated by scanning thresholds starting at -10^8, then 0, then 15 log-spaced steps between 10^3 and 10^8. The AUCs were computed individually for each experimental repeat by trapezoidal area approximation using the MATLAB function 'trapz()' (`https://www.mathworks.com/help/matlab/ref/trapz.html`). The AUC-like curves in Supplementary Figure 21 were computed by fitting data from each experimental repeat with a bi-normal classification model in MATLAB (see below for details of the fitting algorithm used).

Calculation of p-values

P-values shown in Figure 4 were computed using the MATLAB function 'ttest()' (`https://www.mathworks.com/help/stats/ttest.html`). Samples were paired per experimental repeat and the test was two-tailed.

Calculation of fold-changes and robustness scores

For quantifying the effects of EnvZ variants and perturbations, we measured fold-changes by dividing the median output level of each sample by that of the equivalent sample in the absence of the EnvZ variant or perturbation. For

perturbation experiments, the level of output absent perturbation is referred to as the nominal output level.

$$\text{Fold-}\Delta(\text{Input/perturbation bin}x) = \frac{\text{Output(Input/perturbation bin}x)}{\text{Output(Input/perturbation bin1)}} \tag{4a}$$

Where log_2-transformed fold-changes are shown for experiments with multiple repeats, the values shown are the mean of the log_2-transformed fold-changes, rather than the log_2-transformation of the mean of the fold-changes. This order of operations ensures that standard deviations of the fold-changes can be computed directly on the log_2-transformed scale.

We computed robustness scores from the fold-changes using the formulae below:

$$\text{Robustness(Perturbation bin}x) = 100 \cdot \left(1 - \left|1 - \text{Fold-}\Delta(\text{Perturbation bin}x)\right|\right) \tag{5a}$$

Quantification of cell-to-cell output variance

To measure noise, we computed the interquartile range (IQR) of the output distributions. As we chose the median to represent the middle of the distribution, the IQR is a corresponding non-parametric measurement of noise. Since gene expression noise in approximately log-distributed, we $\log_1 0$-transformed the data prior to computing the IQR. As with calculations of the medians, negative fluorescent values were discarded when computing the IQR to avoid artefacts.

Model fitting

Where possible, fluorescent reporters were used to estimate the concentration of a molecular species for the purpose of model fitting.

For fitting all models, we used the MATLAB function 'lsqcurvefit()' (`https://www.mathworks.com/help/optim/ug/lsqcurvefit.html`), which minimizes the sum of the squares of the residuals between the model and the data. In general, fits were made with cells subsampled from bins, as indicated for each figure. In Supplementary Figure 21, the fits were made using the true/false positive rates for each bin. Fits were always performed individually per experimental repeat, then means and standard deviations were computed for individual fit parameters.

Goodness of fit was measured by computing the normalized root-mean-square error CV(RMSE) using the following formula:

$$\text{CV(RMSE)} = \frac{\sqrt{\frac{1}{\bar{y}} \sum_i (y(x_i) - f(x_i))^2}}{\bar{y}}$$

Where $y(x_i)$ is the value of the data at the input value x_i, $\bar{y}$ is the mean of y for all values of x, and $f(x_i)$ is the function output at input value x_i.

Fitting functions:

Activation of transcription by OmpR-VP64:

$$y = \alpha_0 + (\alpha - \alpha_0)\frac{x^2}{K_{1/2}^2 + x^2} \tag{6}$$

The cooperativity of OmpR was assumed to be two because it forms a dimer once phosphorylated to bind DNA[18,116].

Activation of OmpR-VP64-driven expression by kinase: (see Supplementary Note 1 for more details):

$$y = \alpha_0 + (\alpha - \alpha_0)\frac{x^2}{K_{1/2}^2 + x^2} \tag{7}$$

Deactivation of OmpR-VP64 by phosphatase:

$$y = \alpha_0 + (\alpha - \alpha_0)\frac{K_{1/2}^2}{K_{1/2}^2 + x^2} \tag{8}$$

While OmpR-VP64 has not been completely tuned over to P-OmpR-VP64, the amount of P-OmpR-VP64 is assumed to be proportional to the level of kinase because the production rate is only dependent on the kinase. In the presence of the phosphatase, the decay rate becomes overwritten by the dephosphorylation reaction. Thus, these proteins can be plugged directly into the OmpR-VP64 activation function, such that the kinase is proportional to OmpR and the phosphatase is inversely so. Because the of the inversion, the phosphatase function becomes a repression-form Hill function.

The bi-normal fitting function for ROC curves is included with our MATLAB flow cytometry analysis package on GitHub ('model_ROC.mat'). In short, the measurement of the fraction of cells positive for the output reporter is assumed to follow a normal distribution with $\mu_1 = 0$ and $\sigma_1 = 1$ for the negative observations (TFF4 or HEK cells in our case) and a normal distribution with unknown μ_2 and σ_2 for the positive observations (T21 in HeLa cells). μ_2 and σ_2 are fit such that the true positive rate for a given false positive rate approximates that of the data.

6 Data Availability

Sequences for all plasmids used in this study are provided as GenBank files in Supplementary Data. New plasmids used in this study will be available on Addgene upon publication. Raw .fcs files are available from the corresponding authors upon reasonable request.

7 Code Availability

General MATLAB code for use in .fcs file processing and analysis are available under an open-source license in our GitHub repository at `https://github.com/Weiss-Lab/MATLAB_Flow_Analysis`. Specific .m scripts for each experiment are available from the corresponding authors upon reasonable request.

8 Acknowledgements

We would like to acknowledge Douglas Lauffenburger, Ahmad Khalil, and Conor McClune for helpful discussion. Thanks to Bre DiAndreth for the MATLAB code to find optimal orthogonal TCS pairs; Jeremy Gam for the MATLAB code to fit ROC curves and plasmids; both Jin Huh and Kalon Overholt for plasmids and helpful discussion; and Nika Shakiba, Melody Wu, and Margaret Zhang for help with cloning pSub0 plasmids.

9 Author Contributions

R.D.J., Y.Q., D.D.V., and R.W. designed the study; R.W., D.D.V., and M.T.L. secured funding; R.D.J., K.I., and B.W. performed the experiments; R.D.J. and B.W. analyzed the data; Y.Q. and R.D.J. developed the mathematical models; R.D.J., Y.Q., D.D.V., and R.W. wrote the manuscript; R.D.J., Y.Q., K.I., B.W., M.T.L., D.D.V., and R.W. edited the manuscript.

10 Funding

This work was supported by the National Science Foundation (MCB-1840257) and the United States Air Force Office of Scientific Research (FA9550-14-1-0060).

11 Competing Interests Statement

The Massachusetts Institute of Technology has filed a patent application on behalf the inventors (R.D.J., J.H., and R.W.) of phosphorylation-based miRNA sensor design described (US Provisional Application No. 16/528,772) and a provisional application on behalf of the inventors (R.D.J., Y.Q., D.D.V., and R.W.) of the phosphorylation-based feedback controller design described. The remaining authors declare no conflict of interest.

References

1. Jordan, J. D., Landau, E. M. & Iyengar, R. Signaling networks: The origins of cellular multitasking. *Cell* **103,** 193–200 (2000).
2. Lipsitz, Y. Y., Timmins, N. E. & Zandstra, P. W. Quality cell therapy manufacturing by design. *Nature Biotechnology* **34,** 393–400 (2016).
3. Lim, W. A. & June, C. H. The Principles of Engineering Immune Cells to Treat Cancer. *Cell* **168,** 724–740 (2017).

4. Brenner, M. J., Cho, J. H., Wong, N. M. & Wong, W. W. Synthetic Biology: Immunotherapy by Design. *Annual Review of Biomedical Engineering* **20,** 95–118 (2018).

5. Kitada, T., DiAndreth, B., Teague, B. & Weiss, R. Programming gene and engineered-cell therapies with synthetic biology. *Science* **359,** eaad1067 (Feb. 2018).

6. Xie, M. & Fussenegger, M. Designing cell function: assembly of synthetic gene circuits for cell biology applications. *Nature Reviews Molecular Cell Biology* **19,** 507–525 (2018).

7. Scheller, L. & Fussenegger, M. From synthetic biology to human therapy: engineered mammalian cells. *Current Opinion in Biotechnology* **58,** 108–116 (2019).

8. Kamm, R. D. *et al.* Perspective: The promise of multi-cellular engineered living systems. *APL Bioengineering* **2,** 040901 (2018).

9. Madl, C. M., Heilshorn, S. C. & Blau, H. M. Bioengineering strategies to accelerate stem cell therapeutics. *Nature* **557,** 335–342 (2018).

10. Rossi, G., Manfrin, A. & Lutolf, M. P. Progress and potential in organoid research. *Nature Reviews Genetics* **19,** 671–687 (2018).

11. Tewary, M., Shakiba, N. & Zandstra, P. W. Stem cell bioengineering: building from stem cell biology. *Nature Reviews Genetics* **19,** 595–614 (2018).

12. Johnson, M. B., March, A. R. & Morsut, L. Engineering multicellular systems: Using synthetic biology to control tissue self-organization. *Current Opinion in Biomedical Engineering* **4,** 163–173 (2017).

13. Kojima, R., Aubel, D. & Fussenegger, M. Building sophisticated sensors of extracellular cues that enable mammalian cells to work as “doctors” in the body. *Cellular and Molecular Life Sciences* (2020).

14. Antebi, Y. E. *et al.* Combinatorial Signal Perception in the BMP Pathway. *Cell* **170,** 1184–1196.e24 (2017).

15. Edelstein, H. I. *et al.* Elucidation And Refinement Of Synthetic Receptor Mechanisms. *Synthetic Biology* **5,** 1–18 (2020).

16. Maze, A. & Benenson, Y. Artificial signaling in mammalian cells enabled by prokaryotic two-component system. *Nature Chemical Biology* **16,** 179–187 (2019).

17. Scheller, L. *et al.* Phosphoregulated orthogonal signal transduction in mammalian cells. *Nature Communications* **11,** 3085 (2020).

18. Galperin, M. Y. Structural classification of bacterial response regulators: Diversity of output domains and domain combinations. *Journal of Bacteriology* **188,** 4169–4182 (2006).

19. Gao, R. & Stock, A. M. Biological Insights from Structures of Two-Component Proteins. *Annual Review of Microbiology* **63,** 133–154 (2009).

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

21. 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).

22. Berntsson, O. *et al.* Sequential conformational transitions and α-helical supercoiling regulate a sensor histidine kinase. *Nature Communications* **8,** 1–7 (2017).

23. Jacob-Dubuisson, F., Mechaly, A., Betton, J. M. & Antoine, R. Structural insights into the signalling mechanisms of two-component systems. *Nature Reviews Microbiology* **16,** 585–593 (2018).

24. Stock, J. B., Ninfa, A. J. & Stock, A. M. *Protein phosphorylation and regulation of adaptive responses in bacteria* **4,** 450–490 (1989).

25. Wanner, B. L. Is cross regulation by phosphorylation of two-component response regulator proteins important in bacteria? *Journal of Bacteriology* **174,** 2053–2058 (1992).

26. Alves, R. & Savageau, M. A. Comparative analysis of prototype two-component systems with either bifunctional or monofunctional sensors: Differences in molecular structure and physiological function. *Molecular Microbiology* **48,** 25–51 (2003).

27. Laub, M. T. & Goulian, M. Specificity in Two-Component Signal Transduction Pathways. *Annual Review of Genetics* **41,** 121–145 (2007).

28. Siryaporn, A. & Goulian, M. Cross-talk suppression between the CpxA-CpxR and EnvZ-OmpR two-component systems in E. coli. *Molecular Microbiology* **70,** 494–506 (2008).

29. Groban, E. S., Clarke, E. J., Salis, H. M., Miller, S. M. & Voigt, C. A. Kinetic Buffering of Cross Talk between Bacterial Two-Component Sensors. *Journal of Molecular Biology* **390,** 380–393 (2009).

30. Yang, Y. & Inouye, M. *Requirement of both kinase and phosphatase activities of an Escherichia coli receptor (Taz1) for ligand-dependent signal transduction* 1993.

31. Capra, E. J., Perchuk, B. S., Skerker, J. M. & Laub, M. T. Adaptive mutations that prevent crosstalk enable the expansion of paralogous signaling protein families. *Cell* **150,** 222–232 (2012).

32. Hansen, J. *et al.* Transplantation of prokaryotic two-component signaling pathways into mammalian cells. *PNAS* **111,** 15705–15710 (2014).

33. Goldbeter, A. & Koshland, D. E. An amplified sensitivity arising from covalent modification in biological systems. *PNAS* **78,** 6840–6844 (1981).

34. Asthagiri, A. R. & Lauffenburger, D. A. A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model. *Biotechnology Progress* **17,** 227–239 (2001).

35. Jiang, P. *et al.* Load-induced modulation of signal transduction networks. *Science Signaling* **4,** ra67 (2011).

36. Chang, Y. C., Armitage, J. P., Papachristodoulou, A. & Wadhams, G. H. A single phosphatase can convert a robust step response into a graded, tunable or adaptive response. *Microbiology* **159,** 1276–1285 (2013).

37. Nilgiriwala, K. S., Jiménez, J., Rivera, P. M. & Del Vecchio, D. Synthetic Tunable Amplifying Buffer Circuit in E. coli. *ACS Synthetic Biology* **4,** 577–584 (2015).

38. Qian, Y. & Del Vecchio, D. Realizing "Integral Control" In Living Cells: How To Overcome Leaky Integration Due To Dilution? *Journal of The Royal Society Interface* **15,** 20170902 (2018).

39. Cai, S. J. & Inouye, M. EnvZ-OmpR interaction and osmoregulation in Escherichia coli. *Journal of Biological Chemistry* **277,** 24155–24161 (2002).

40. Del Vecchio, D. Modularity, context-dependence, and insulation in engineered biological circuits. *Trends in Biotechnology* **33,** 111–119 (2015).

41. Grunberg, T. W. & Del Vecchio, D. Modular Analysis and Design of Biological Circuits. *Current Opinion in Biotechnology* **63,** 41–47 (2020).

42. Stock, A. M., Robinson, V. L. & Goudreau, P. N. Two-Component Signal Transduction. *Annual Review of Biochemistry* **69,** 183–215 (2000).

43. Skarphol, K., Waukau, J. & Forst, S. A. Role of His243 in the phosphatase activity of EnvZ in Escherichia coli. *Journal of Bacteriology* **179,** 1413–1416 (1997).

44. Hsing, W., Russo, F. D., Bernd, K. K. & Silhavy, T. J. Mutations that alter the kinase and phosphatase activities of the two-component sensor EnvZ. *Journal of Bacteriology* **180,** 4538–4546 (1998).

45. Tanaka, T. *et al.* NMR Structure of the Histidine Kinase Domain of the E. coli Osmosensor EnvZ. *Nature* **396,** 88–92. arXiv: `9810036 [physics]` (1998).

46. Dutta, R., Yoshida, T. & Inouye, M. The critical role of the conserved Thr247 residue in the functioning of the osmosensor EnvZ, a histidine kinase/phosphatase, in Escherichia coli. *Journal of Biological Chemistry* **275,** 38645–38653 (2000).

47. Qin, L., Cai, S., Zhu, Y. & Inouye, M. Cysteine-scanning analysis of the dimerization domain of EnvZ, an osmosensing histidine kinase. *Journal of Bacteriology* **185,** 3429–3435 (2003).

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

714 49. Willett, J. W. & Kirby, J. R. Genetic and Biochemical Dissection of a HisKA Domain Identifies Residues
715 Required Exclusively for Kinase and Phosphatase Activities. *PLoS Genetics* **8,** e1003084 (2012).

716 50. Park, H. & Inouye, M. Mutational analysis of the linker region of EnvZ, an osmosensor in Escherichia coli.
717 *Journal of Bacteriology* **179,** 4382–4390 (1997).

718 51. Zhu, Y., Qin, L., Yoshida, T. & Inouye, M. Phosphatase activity of histidine kinase EnvZ without kinase
719 catalytic domain. *PNAS* **97,** 7808–7813 (2000).

720 52. Qin, L., Dutta, R., Kurokawa, H., Ikura, M. & Inouye, M. A monomeric histidine kinase derived from EnvZ,
721 an Echerichia coli osmosensor. *Molecular Microbiology* **36,** 24–32 (2000).

722 53. Huynh, T. N., Noriega, C. E. & Stewart, V. Conserved mechanism for sensor phosphatase control of
723 two-component signaling revealed in the nitrate sensor NarX. *PNAS* **107,** 21140–5 (2010).

724 54. Wang, B., Zhao, A., Novick, R. P. & Muir, T. W. Activation and inhibition of the receptor histidine kinase
725 AgrC occurs through opposite helical transduction motions. *Molecular Cell* **53,** 929–940 (2014).

726 55. Ashenberg, O., Rozen-Gagnon, K., Laub, M. T. & Keating, A. E. Determinants of homodimerization
727 specificity in histidine kinases. *Journal of Molecular Biology* **413,** 222–235 (2011).

728 56. Egger, L. a., Park, H. & Inouye, M. Signal transduction via the histidyl-aspartyl phosphorelay. *Genes to Cells*
729 **2,** 167–184 (1997).

730 57. Hardman, G. *et al.* Strong anion exchange-mediated phosphoproteomics reveals extensive human
731 non-canonical phosphorylation. *The EMBO Journal* **38,** 1–22 (2019).

732 58. Meyer, A. J., Segall-Shapiro, T. H., Glassey, E., Zhang, J. & Voigt, C. A. Escherichia coli "Marionette" strains
733 with 12 highly optimized small-molecule sensors. *Nature Chemical Biology* **15,** 196–204 (2019).

734 59. Foight, G. W. *et al.* Multi-input chemical control of protein dimerization for programming graded cellular
735 responses. *Nature Biotechnology* **37,** 1209–1216 (2019).

736 60. Szenk, M., Yim, T. & Balázsi, G. Multiplexed Gene Expression Tuning with Orthogonal Synthetic Gene
737 Circuits. *ACS Synthetic Biology* **9,** 930–939 (2020).

738 61. Donahue, P. S. *et al.* The COMET toolkit for composing customizable genetic programs in mammalian cells.
739 *Nature Communications* **11,** 779 (2020).

740 62. Hsiao, V., De Los Santos, E. L., Whitaker, W. R., Dueber, J. E. & Murray, R. M. Design and implementation
741 of a biomolecular concentration tracker. *ACS Synthetic Biology* **4,** 150–161 (2015).

742 63. Huang, H.-H., Qian, Y. & Del Vecchio, D. A quasi-integral controller for adaptation of genetic modules to
743 variable ribosome demand. *Nature Communications* **9,** 5415 (2018).

64. Aoki, S. K. *et al.* A universal rationally-designed biomolecular integral feedback controller for robust perfect adaptation. *Nature* **570,** 533–537 (2019).

65. Gam, J. J., DiAndreth, B., Jones, R. D., Huh, J. & Weiss, R. A 'poly-transfection' method for rapid, one-pot characterization and optimization of genetic systems. *Nucleic Acids Research* **47,** e106 (2019).

66. Iwamoto, M., Björklund, T., Lundberg, C., Kirik, D. & Wandless, T. J. A general chemical method to regulate protein stability in the mammalian central nervous system. *Chemistry and Biology* **17,** 981–988 (2010).

67. Miyazaki, Y., Imoto, H., Chen, L. & Wandless, T. Destabilizing domains derived from the human estrogen receptor. *Journal of the American Chemical Society* **134,** 3942–3945 (2012).

68. Nissim, L. & Bar-Ziv, R. H. A tunable dual-promoter integrator for targeting of cancer cells. *Molecular Systems Biology* **6,** 444 (2010).

69. Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R. & Benenson, Y. Multi-Input RNAi-Based Logic Circuit for Identification of Specific Cancer Cells. *Science* **333,** 1307–1311 (2011).

70. Gam, J. J., Babb, J. & Weiss, R. A mixed antagonistic/synergistic miRNA repression model enables accurate predictions of multi-input miRNA sensor activity. *Nature Communications* **9,** 1–12 (2018).

71. Mohammadi, P., Beerenwinkel, N. & Benenson, Y. Automated Design of Synthetic Cell Classifier Circuits Using a Two-Step Optimization Strategy. *Cell Systems* **4,** 207–218.e14 (2017).

72. Rinaudo, K. *et al.* A universal RNAi-based logic evaluator that operates in mammalian cells. *Nature Biotechnology* **25,** 795–801. arXiv: `arXiv:1011.1669v3` (2007).

73. Jones, R. D. *et al.* An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells. *Nature Communications,* 5690 (2020).

74. Head, C. G., Tardy, A. & Kenney, L. J. Relative binding affinities of OmpR and OmpR-phosphate at the ompF and ompC regulatory sites. *Journal of Molecular Biology* **281,** 857–870 (1998).

75. Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. & Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biology* **4,** 1707–1719 (2006).

76. Gutierrez, P. S., Monteoliva, D. & Diambra, L. Cooperative Binding of Transcription Factors Promotes Bimodal Gene Expression Response. *PLoS ONE* **7** (2012).

77. Qian, Y., Huang, H. H., Jiménez, J. I. & Del Vecchio, D. Resource Competition Shapes the Response of Genetic Circuits. *ACS Synthetic Biology* **6** (2017).

78. Frei, T. *et al.* Characterization and mitigation of gene expression burden in mammalian cells. *Nature Communications* **11,** 4641 (2020).

79. Sturm, O. E. *et al.* The mammalian MAPK/ERK pathway exhibits properties of a negative feedback amplifier. *Science Signaling* **3,** ra90 (2010).

80. Nevozhay, D., Adams, R. M., Murphy, K. F., Josic, K. & Balazsi, G. Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression. *PNAS* **106,** 5123–5128. arXiv: `arXiv:1408.1149` (2009).

81. Nevozhay, D., Zal, T. & Balázsi, G. Transferring a synthetic gene circuit from yeast to mammalian cells. *Nature Communications* **4,** 1451. arXiv: `9809069v1 [arXiv:gr-qc]` (2013).

82. Ng, A. H. *et al.* Modular and tunable biological feedback control using a de novo protein switch. *Nature* **572,** 265–269 (2019).

83. Szymczak, A. L. *et al.* Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nature Biotechnology* **22,** 589–594 (2004).

84. Berger, S. L., Cress, W., Cress, A., Triezenberg, S. J. & Guarente, L. Selective inhibition of activated but not basal transcription by the acidic activation domain of VP16: Evidence for transcriptional adaptors. English. *Cell* **61,** 1199–1208 (June 1990).

85. Schmidl, S. R. *et al.* Rewiring bacterial two-component systems by modular DNA-binding domain swapping. *Nature Chemical Biology* **15,** 690–698 (2019).

86. Muldoon, J. J. *et al.* Model-guided design of mammalian genetic programs. *Science Advances* **7,** eabe9375 (2021).

87. Yamamoto, K. & Ishihama, A. Transcriptional response of Escherichia coli to external zinc. *Journal of Bacteriology* **187,** 6333–6340 (2005).

88. Procaccini, A., Lunt, B., Szurmant, H., Hwa, T. & Weigt, M. Dissecting the specificity of protein-protein interaction in bacterial two-component signaling: Orphans and crosstalks. *PLoS ONE* **6** (2011).

89. Willett, J. W. *et al.* Specificity residues determine binding affinity for two-component signal transduction systems. *mBio* **4,** 1–11 (2013).

90. McClune, C. J., Alvarez-Buylla, A., Voigt, C. A. & Laub, M. T. Engineering orthogonal signalling pathways reveals the sparse occupancy of sequence space. *Nature* **574,** 702–706 (2019).

91. Albanesi, D. *et al.* Structural plasticity and catalysis regulation of a thermosensor histidine kinase. *Proceedings of the National Academy of Sciences of the United States of America* **106,** 16185–16190 (2009).

92. Casino, P., Rubio, V. & Marina, A. Structural Insight into Partner Specificity and Phosphoryl Transfer in Two-Component Signal Transduction. *Cell* **139,** 325–336 (2009).

93. Ferris, H. U. *et al.* Mechanism of regulation of receptor histidine kinases. *Structure* **20,** 56–66 (2012).

94. Yi, T.-M., Huang, Y., Simon, M. I. & Doyle, J. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *PNAS* **97,** 4649–4653 (Apr. 2000).

95. Groisman, E. A. Feedback Control of Two-Component Regulatory Systems. *Annual Review of Microbiology* **70,** 103–124 (2016).

96. Salazar, M. E., Podgornaia, A. I. & Laub, M. T. The small membrane protein MgrB regulates PhoQ bifunctionality to control PhoP target gene expression dynamics. *Molecular Microbiology* **102,** 430–445 (2016).

97. Stein, B. J., Fiebig, A. & Crosson, S. Feedback Control of a Two-Component Signaling System by an Fe-S-Binding Receiver Domain. *mBio* **11,** 1–16 (2020).

98. Hutchings, M. I., Hong, H. J. & Buttner, M. J. The vancomycin resistance VanRS two-component signal transduction system of Streptomyces coelicolor. *Molecular Microbiology* **59,** 923–935 (2006).

99. Yeo, W. S. *et al.* Intrinsic Negative Feedback Governs Activation Surge in Two-Component Regulatory Systems. *Molecular Cell* **45,** 409–421 (2012).

100. Alon, U., Surette, M. G., Barkai, N. & Leibler, S. Robustness in bacterial chemotaxis. *Nature* **397,** 168–171 (Jan. 1999).

101. Lake, D., Corrêa, S. A. L. & Müller, J. Negative feedback regulation of the ERK1/2 MAPK pathway. *Cellular and Molecular Life Sciences* **73,** 4397–4413 (2016).

102. Yoon, S. & Seger, R. The extracellular signal-regulated kinase: Multiple substrates regulate diverse cellular functions. *Growth Factors* **24,** 21–44 (2006).

103. McKay, M. M. & Morrison, D. K. Integrating signals from RTKs to ERK/MAPK. *Oncogene* **26,** 3113–3121 (2007).

104. Caunt, C. J. & Keyse, S. M. Dual-specificity MAP kinase phosphatases (MKPs): Shaping the outcome of MAP kinase signalling. *FEBS Journal* **280,** 489–504 (2013).

105. Nunns, H. & Goentoro, L. Signaling pathways as linear transmitters. *eLife* **7,** 1–37 (2018).

106. Salmond, R. J., Brownlie, R. J., Morrison, V. L. & Zamoyska, R. The tyrosine phosphatase PTPN22 discriminates weak self peptides from strong agonist TCR signals. *Nature Immunology* **15,** 875–883 (2014).

107. Li, Q. J. *et al.* miR-181a Is an Intrinsic Modulator of T Cell Sensitivity and Selection. *Cell* **129,** 147–161 (2007).

108. Gaud, G., Lesourne, R. & Love, P. E. Regulatory mechanisms in T cell receptor signalling. *Nature Reviews Immunology* **18,** 485–497 (2018).

109. Rodríguez-Galán, A., Fernández-Messina, L. & Sánchez-Madrid, F. Control of immunoregulatory molecules by miRNAs in T cell activation. *Frontiers in Immunology* **9,** 2148 (2018).

110. Duportet, X. *et al.* A platform for rapid prototyping of synthetic gene networks in mammalian cells. *Nucleic Acids Research* **42,** 13440–13451 (Nov. 2014).

111. Sandelin, A. *et al.* Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nature Reviews Genetics* **8,** 424–436 (2007).

112. Roy, A. L. & Singer, D. S. Core promoters in transcription: Old problem, new insights. *Trends in Biochemical Sciences* **40,** 165–171 (2015).

113. Haberle, V. & Stark, A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nature Reviews Molecular Cell Biology* **19,** 621–637 (2018).

114. Ede, C., Chen, X., Lin, M. Y. & Chen, Y. Y. Quantitative Analyses of Core Promoters Enable Precise Engineering of Regulated Gene Expression in Mammalian Cells. *ACS Synthetic Biology* **5,** 395–404 (2016).

115. Ponjavic, J. *et al.* Transcriptional and structural impact of TATA-initiation site spacing in mammalian core promoters. *Genome biology* **7,** R78 (2006).

116. Gao, R., Mack, T. R. & Stock, A. M. Bacterial response regulators: versatile regulatory strategies from common domains. *Trends in Biochemical Sciences* **32,** 225–234 (2007).

Figures

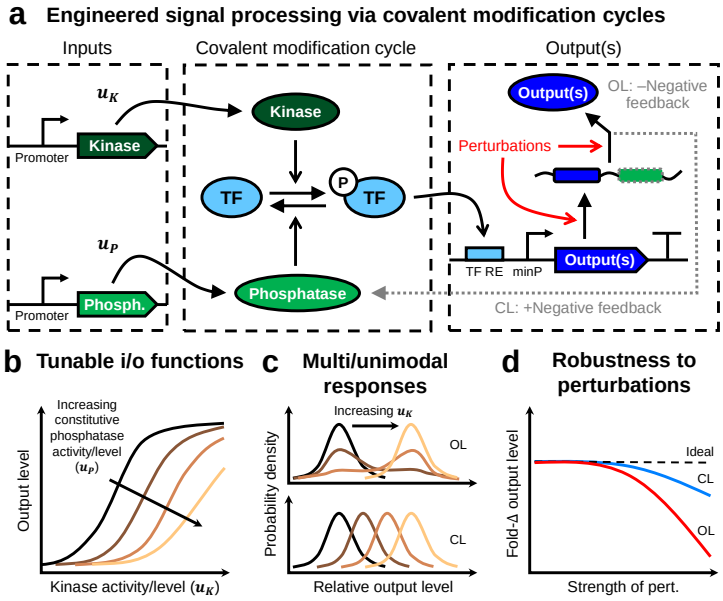

Figure 1

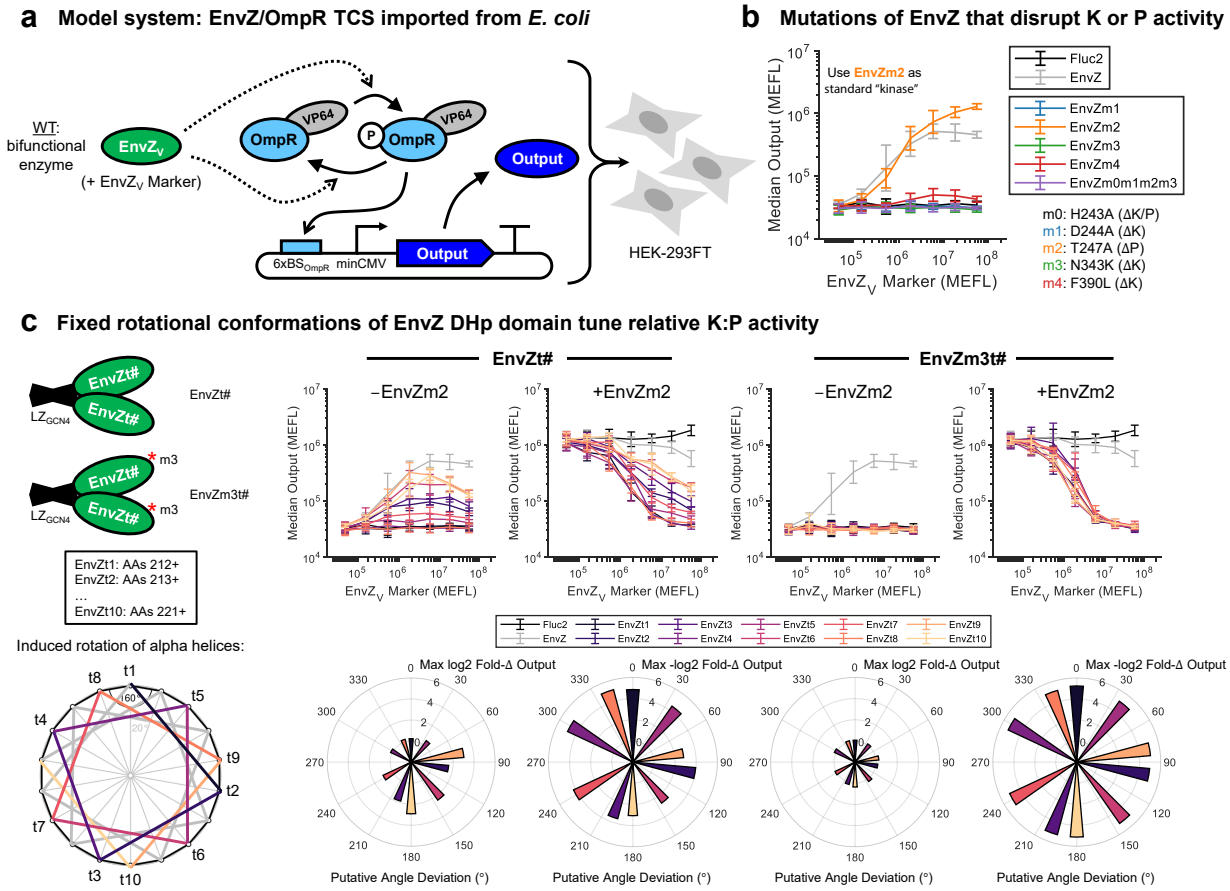
a Model system: EnvZ/OmpR TCS imported from E. coli
WT: bifunctional enzyme
(+ EnvZV Marker)
EnvZV
OmpR
VP64
P
OmpR
VP64
Output
6xBSOmpR minCMV
Output
HEK-293FT
b Mutations of EnvZ that disrupt K or P activity
Use EnvZm2 as standard "kinase"
Median Output (MEFL)
EnvZV Marker (MEFL)
Fluc2
EnvZ
EnvZm1
EnvZm2
EnvZm3
EnvZm4
EnvZm0m1m2m3
m0: H243A (ΔK/P)
m1: D244A (ΔK)
m2: T247A (ΔP)
m3: N343K (ΔK)
m4: F390L (ΔK)
c Fixed rotational conformations of EnvZ DHp domain tune relative K:P activity
EnvZt#
EnvZm3t#
LZGCN4
m3
EnvZt1: AAs 212+
EnvZt2: AAs 213+
...
EnvZt10: AAs 221+
Induced rotation of alpha helices:
EnvZt#
EnvZm3t#
−EnvZm2
+EnvZm2
−EnvZm2
+EnvZm2
Median Output (MEFL)
EnvZV Marker (MEFL)
Fluc2
EnvZ
EnvZt1
EnvZt2
EnvZt3
EnvZt4
EnvZt5
EnvZt6
EnvZt7
EnvZt8
EnvZt9
EnvZt10
Max log2 Fold-Δ Output
Max -log2 Fold-Δ Output
Putative Angle Deviation (°)

Figure 2

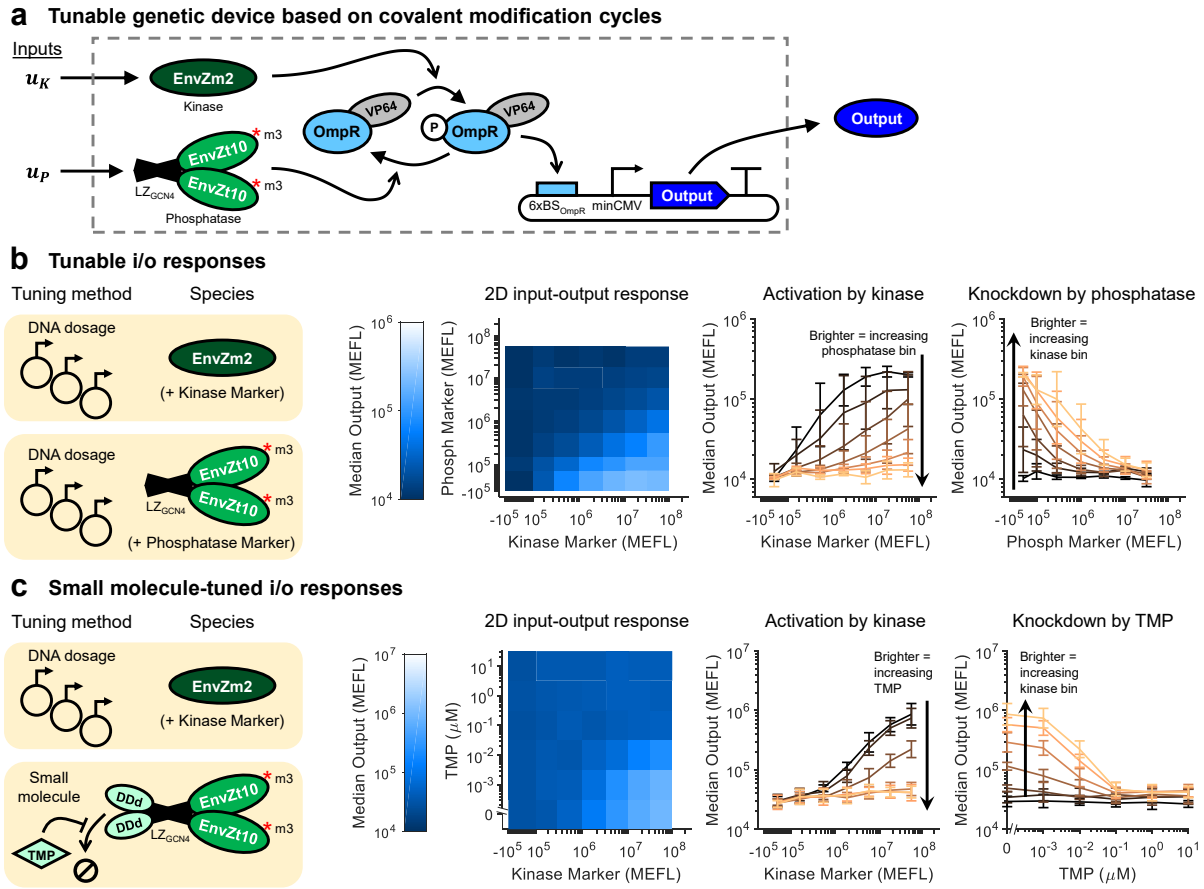
a Tunable genetic device based on covalent modification cycles
Inputs
EnvZm2
Kinase
EnvZt10
m3
LZGCN4
Phosphatase
OmpR
VP64
P
6xBSOmpR
minCMV
Output
b Tunable i/o responses
Tuning method
Species
DNA dosage
(+ Kinase Marker)
(+ Phosphatase Marker)
2D input-output response
Activation by kinase
Knockdown by phosphatase
Median Output (MEFL)
Phosph Marker (MEFL)
Kinase Marker (MEFL)
Brighter = increasing phosphatase bin
Brighter = increasing kinase bin
c Small molecule-tuned i/o responses
Small molecule
DDd
TMP
Knockdown by TMP
TMP (μM)
Brighter = increasing TMP

Figure 3

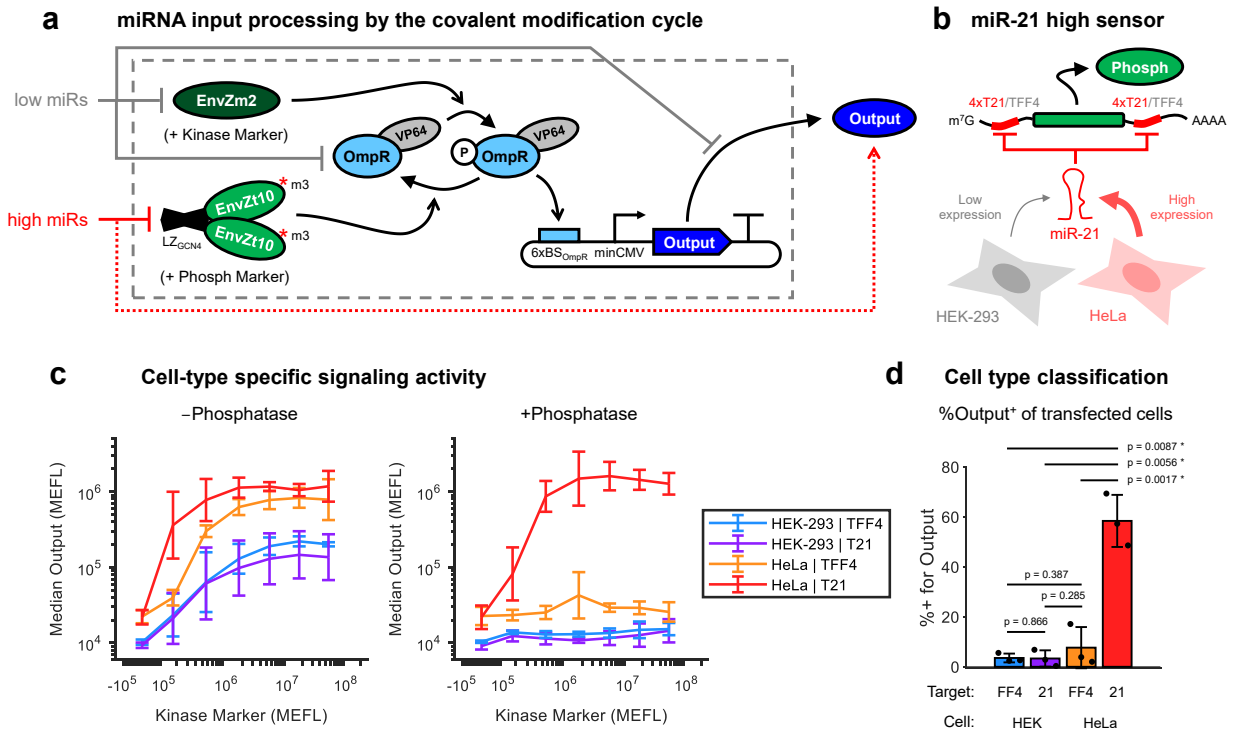

Figure 4

Figure Legends

Figure 1. Overview of engineered covalent modification cycle. **(a)** A covalent modification cycle (CMC) is composed of a substrate that is interconverted between an active and an inactive form by two different enzymes. Here we examine the CMC created by reversible phosphorylation/dephosphorylation of a transcription factor (TF) by a kinase and phosphatase. The inputs to this CMC, u_K and u_P, alter the production rate or catalytic rates of the kinase and phosphatase, respectively. The output(s) of the system are RNA and/or protein species produced in response to phosphorylation of the TF. Closed loop (CL) negative feedback control can be achieved by co-expressing a phosphatase with the output. Without the feedback, the expression of the outputs is open loop (OL). **(b)** The input/output (i/o) response of the system, *i.e.* the response of the TF-driven output(s) to kinase inputs (u_K), can be tuned via phosphatase inputs (u_P). **(c)** Negative feedback is expected to convert multimodal output responses into unimodal responses. **(d)** Negative feedback is expected to impart robustness to perturbations in the output production process.

Figure 2. Isolation of kinase and phosphatase activity from bifunctional histidine kinases. **(a)** Model system to construct a covalent modification cycle (CMC): EnvZ/OmpR proteins from *E. coli* two-component signaling

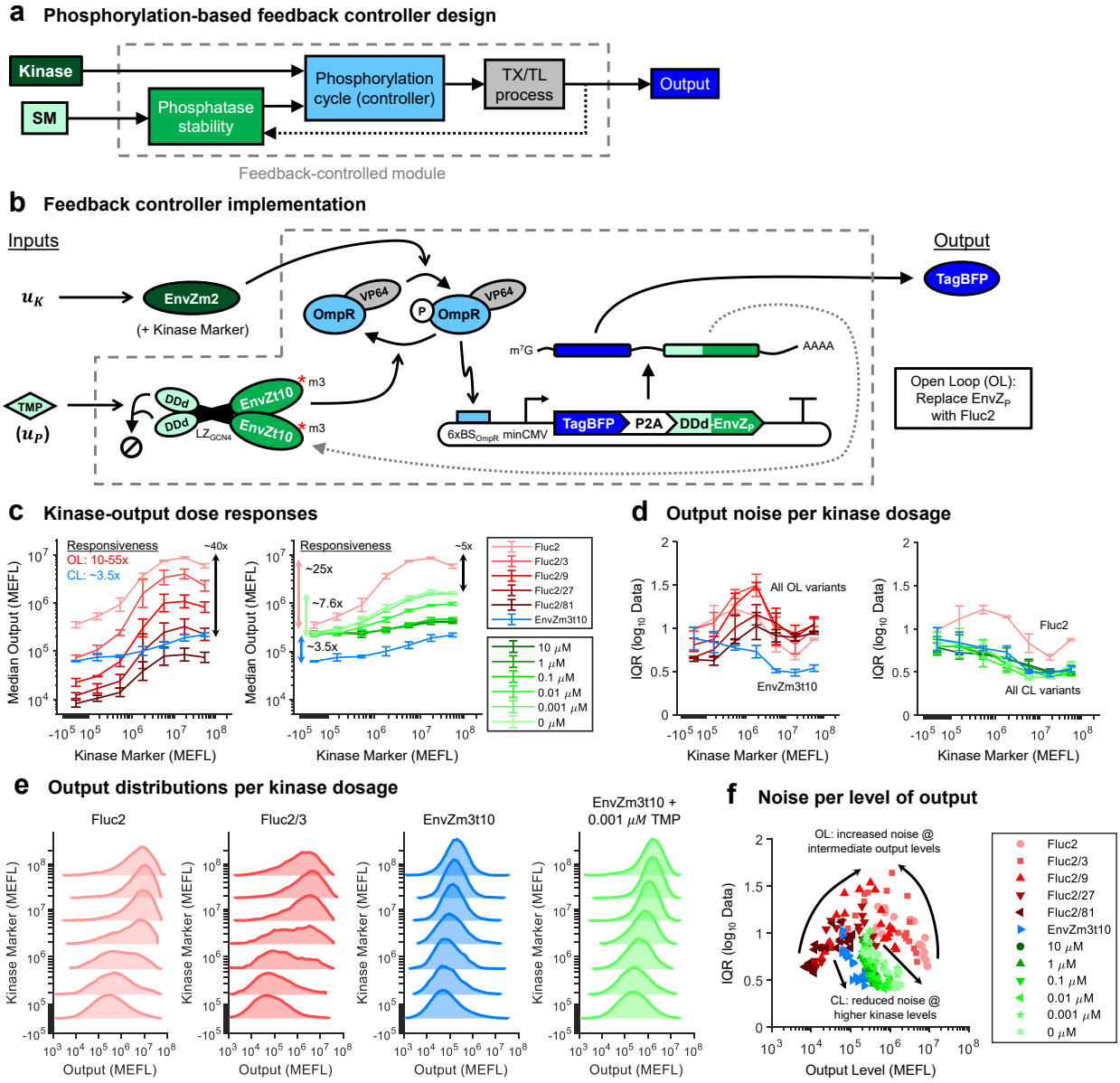
a Phosphorylation-based feedback controller design
Kinase
SM
Phosphatase stability
Phosphorylation cycle (controller)
TX/TL process
Output
Feedback-controlled module
b Feedback controller implementation
Inputs
Output
EnvZm2
(+ Kinase Marker)
OmpR
VP64
TagBFP
m7G
AAAA
TMP
DDd
EnvZt10
m3
LZGCN4
6xBSOmpR minCMV
TagBFP P2A DDd-EnvZP
Open Loop (OL): Replace EnvZP with Fluc2
c Kinase-output dose responses
Responsiveness
OL: 10-55x
CL: ~3.5x
~40x
~25x
~7.6x
~3.5x
~5x
Median Output (MEFL)
Kinase Marker (MEFL)
Fluc2
Fluc2/3
Fluc2/9
Fluc2/27
Fluc2/81
EnvZm3t10
10 μM
1 μM
0.1 μM
0.01 μM
0.001 μM
0 μM
d Output noise per kinase dosage
IQR (log10 Data)
All OL variants
EnvZm3t10
Fluc2
All CL variants
e Output distributions per kinase dosage
Fluc2
Fluc2/3
EnvZm3t10
EnvZm3t10 + 0.001 μM TMP
Kinase Marker (MEFL)
Output (MEFL)
f Noise per level of output
OL: increased noise @ intermediate output levels
CL: reduced noise @ higher kinase levels
IQR (log10 Data)
Output Level (MEFL)

Figure 5

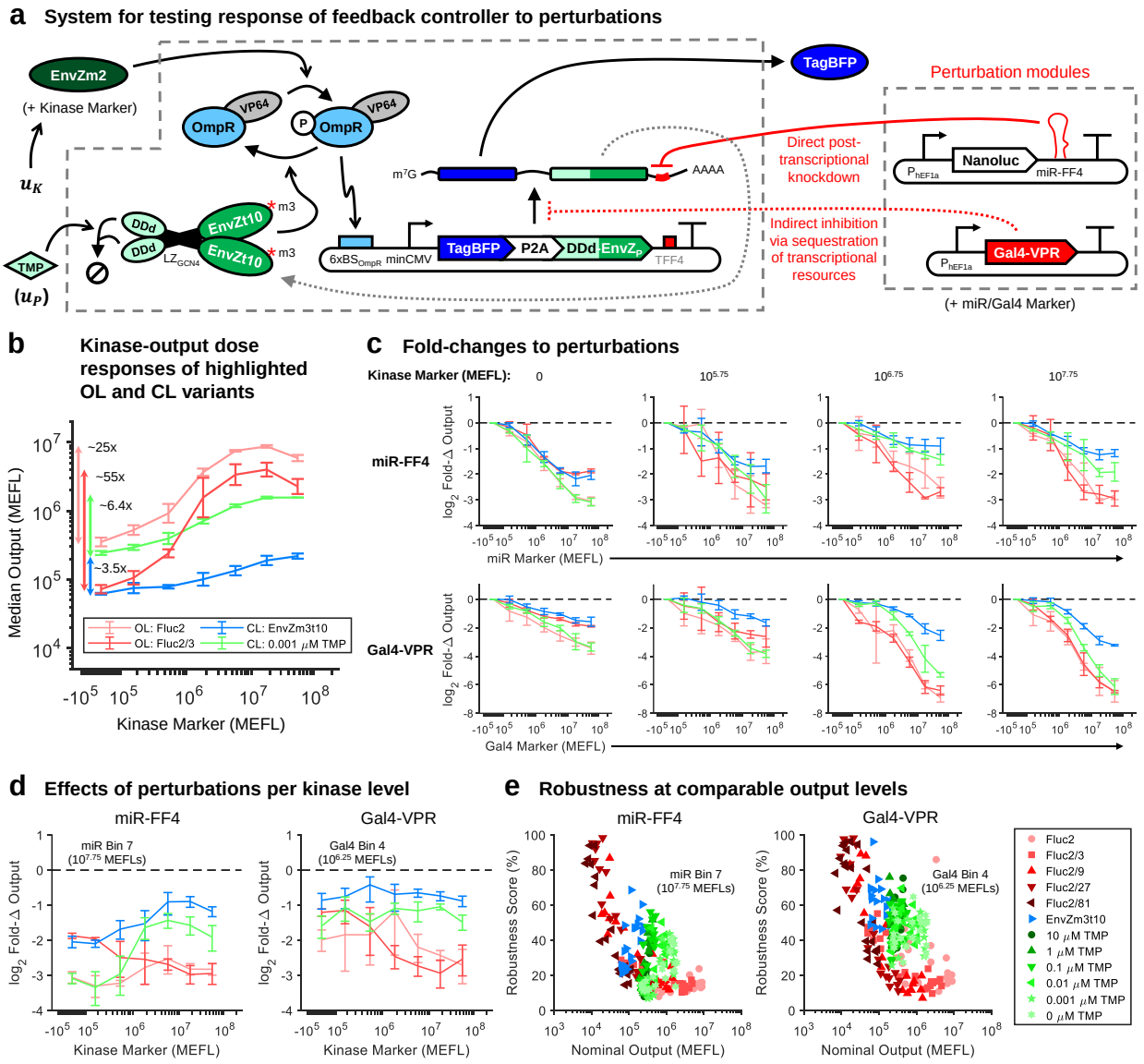
a System for testing response of feedback controller to perturbations
EnvZm2
(+ Kinase Marker)
u_K
TMP
(u_P)
OmpR
VP64
P
m3
DDd
LZ$_{GCN4}$
EnvZt10
m^{7}G
AAAA
6xBS$_{OmpR}$ minCMV
TagBFP
P2A
DDd-EnvZ$_P$
TFF4
Perturbation modules
Direct post-transcriptional knockdown
Indirect inhibition via sequestration of transcriptional resources
P$_{hEF1a}$
Nanoluc
miR-FF4
Gal4-VPR
(+ miR/Gal4 Marker)
b Kinase-output dose responses of highlighted OL and CL variants
Median Output (MEFL)
~25x
~55x
~6.4x
~3.5x
OL: Fluc2
OL: Fluc2/3
CL: EnvZm3t10
CL: 0.001 μM TMP
Kinase Marker (MEFL)
c Fold-changes to perturbations
Kinase Marker (MEFL): 0
$10^{5.75}$
$10^{6.75}$
$10^{7.75}$
miR-FF4
Gal4-VPR
log$_2$ Fold-Δ Output
miR Marker (MEFL)
Gal4 Marker (MEFL)
d Effects of perturbations per kinase level
miR-FF4
Gal4-VPR
miR Bin 7 ($10^{7.75}$ MEFLs)
Gal4 Bin 4 ($10^{6.25}$ MEFLs)
log$_2$ Fold-Δ Output
Kinase Marker (MEFL)
e Robustness at comparable output levels
miR-FF4
Gal4-VPR
miR Bin 7 ($10^{7.75}$ MEFLs)
Gal4 Bin 4 ($10^{6.25}$ MEFLs)
Robustness Score (%)
Nominal Output (MEFL)
Fluc2
Fluc2/3
Fluc2/9
Fluc2/27
Fluc2/81
EnvZm3t10
10 μM TMP
1 μM TMP
0.1 μM TMP
0.01 μM TMP
0.001 μM TMP
0 μM TMP

Figure 6

(TCS). EnvZ naturally exhibits both kinase and phosphatase activities, but favors one or the other depending on its conformation. EnvZ variants ($EnvZ_V$) are co-delivered to cells with $EnvZ_V$ Marker, a fluorescent reporter that indicates dosage per cell. **(b)** Evaluation of EnvZ mutants. Constitutively-expressed OmpR-VP64 (OmpR fused to the activation domain VP64) was co-transfected with a reporter plasmid comprising a promoter (P_{OmpR}) with 6xOmpR binding sites and a minimal CMV promoter driving TagBFP as the output. Each EvnZ variant ($EnvZ_V$) was poly-transfected against the other plasmids to evaluate the $EnvZ_V$-to-output dose-responses (see Supplementary Figure 5 for additional details). Further details about the mutants and other EnvZ variants tested are in Supplementary Figure 4. **(c)** Effect on EnvZ function by fixing its rotational conformation. GCN4 is fused directly to the N-terminus of EnvZ truncated between residues 212 and 221, thus connecting to the DHp domain and fixing the rotation of its alpha helices[54]. The top plots show the output response to each rotationally-locked variant. Samples were transfected with and without EnvZm2, which establishes a baseline of phosphorylated OmpR for testing dephosphorylation. EnvZm3t# have mutation 'm3' (N343K), which knocks out kinase activity. The radial plots in the bottom row indicate the maximum fold-change in output expression induced by each variant per the putative rotational conformation of the DHp domain, assuming 100° of rotation for each amino acid truncated between GCN4 and the DHp domain and setting EnvZt1 to 0°. All data were measured by flow cytometry at 48 hours post-transfection in HEK-293FT cells. All errorbars represent the mean ± s.d. of measurements from three experimental repeats.

Figure 3. Tuning input/output signaling response by modulating kinase and phosphatase levels. **(a)** Implementation of a covalent modification cycle with kinase (EnvZm2) and phosphatase (EnvZm3t10) variants of EnvZ. The expression level of the output can be tuned as a function of both enzymes, and inputs to each that affect their production rate (u_K and u_P). **(b)** Tuning output expression through different dosages of kinase and phosphatase DNA. The heatmap shows the median level of output for each combination of kinase and phosphatase DNA dosages, assayed with poly-transfection[65] (see Supplementary Figure 18 for full data). The line plots show the same data but broken out by rows or columns. Brighter lines correspond to bins with increasing phosphatase (left) or kinase (right). **(c)** Tuning output expression through small molecule-induced degradation of the phosphatase. DDd is fused to the N-terminus of the phosphatase (see Supplementary Figure 14 for different arrangements and comparison with DDe/4-OHT). Addition of TMP stabilizes the DDd-phosphatase fusion protein[66]. The data is extracted from the full poly-transfection results shown in Supplementary Figures 14 & 15, selecting the middle phosphatase bin (P Marker $\approx 10^6$). The line plots show the same data but broken out by rows or columns. Brighter lines correspond to samples with increasing TMP concentration (left) or bins with increasing kinase (right). All data were measured by flow cytometry at 48 hours post-transfection. HEK-293 cells were used for Panel (b) and HEK-293FT for Panel (c). All errorbars represent the mean ± s.d. of measurements from three experimental repeats. All heatmap values represent the mean of measurements from three experimental repeats.

Cell type-specific signaling responses using covalent modification cycles. (a) miRNA classifier design based on covalent modification cycles. miRNAs expected to be low in the target cell can be used to knock down the kinase, OmpR-VP64, and/or the output. miRNAs expected to be high in the target cell can be used to knock down the phosphatase, effectively increasing the output expression. Not shown for brevity, the level of OmpR-VP64 was optimized using a feedforward controller (Supplementary Figure 16). **(b)** Design of a miR-21 sensor for classification of HeLa cells. miR-21 knocks down phosphatase levels via 4x target sites in each of its 5' and 3' untranslated regions (UTRs). As a control, a variant was made with miR-FF4 target sites (TFF4) in place of the miR-21 target sites (T21), thus preventing knockdown by miR-21. miR-21 is differentially expressed in HeLa compared to HEK-293 cells[65,69]. **(c)** Cell type-specific signaling responses enabled by miRNA regulation of phosphatase expression. The data is extracted from the full poly-transfection data (Supplementary Figure 18), comparing the second-highest phosphatase bin (P Marker $\approx 10^7$) to the lowest (no phosphatase). **(d)** Comparison of the percent of transfected cells positive for the output for each circuit variant in HEK/HeLa cells. P-values are from two-tailed paired T-tests between each group of samples. Receiver-operator characteristic (ROC) curves are provided in Supplementary Figures 20 & 21. All data were measured by flow cytometry at 48 hours post-transfection. All errorbars represent the mean ± s.d. of measurements from three experimental repeats.

Figure 5. Design and implementation of a tunable phosphoryaltion-based feedback controller. **(a)** Block diagram of the feedback controller design. The phosphatase acts as an output sensor, and is fed back from the output to the phosphorylation cycle. The kinase sets the reference for output expression. The output responds to inputs both to the kinase and a small molecule (SM) regulator of phosphatase stability, the latter effectively serving to tune the feedback strength. **(b)** Implementation of the feedback controller. the kinase is EnvZm2, the phosphatase is DDd-EnvZm3t10, and the output is the fluorescent reporter TagBFP, and the output is 2A-linked[83] to the phosphatase to ensure coupled transcription. Addition of TMP stabilizes DDd-EnvZm3t10 and thereby increases the feedback strength. An open loop (OL) version of the system was made by replacing the phosphatase with the luminescent protein Fluc2. Since negative feedback reduces output expression, OL variants with reduced output level were created for comparison at equivalent output levels by reducing the copy number of output reporter by fractional amounts (1:3, 1:9, 1:27, and 1:81). The kinase was poly-transfected in a separate complex to the other plasmids to measure the dose-responses of the OL and closed loop (CL) systems (see Supplementary Figures 23 for details). **(c)** Dose-responses of OL and CL system outputs to kinase input levels. The range of responsiveness to kinase (max fold change ± kinase) are given for the CL and OL variants are indicated to the left of the lines. The fold-difference between max output levels for select OL and CL variants are indicated to the right of the lines. Dose-responses of the DDd-CL system are given in Supplementary Figure 28. **(d)** Quantification of output noise as a function of kinase input dosages. Because the output variance is log-distributed, the interquartile range (IQR) is computed on the

log-transformed data. **(e)** Comparison of output distributions for select OL and CL variants across kinase levels. The data is representative from the first experimental repeat. All OL and CL variants are compared in Supplementary Figure 27. **(f)** Noise as a function of median output levels for all CL and OL variants at all kinase inputs. The individual points are drawn from all experimental repeats. All data were measured by flow cytometry at 48 hours post-transfection in HEK-293FT cells. All errorbars represent the mean ± standard deviation of measurements from three experimental repeats.

Figure 6. Mitigation of perturbations via feedback control. (a) The CL and OL systems introduced in Figure 5 were tested against two perturbations: (i) indirect transcriptional inhibition via loading of transcriptional resource by Gal4-VPR and (ii) direct post-transcriptional knockdown by miR-FF4. The kinase, perturbations, and controllers were each poly-transfected in separate DNA-lipid complexes in order to measure the 2D dose-response of the OL and CL systems to the kinase and perturbations (see Supplementary Figures 23-26 for details). **(b)** Dose-responses of OL and CL systems highlighted in the following panels. The Fluc2 and Fluc2/3 OL variants were chosen since they have nearly identical output levels compared to the CL with and without DDd, respectively, in the absence of kinase. Dose-responses and detailed comparisons among all OL and CL variants are provided in Supplementary Figures 29-32. **(c)** Fold-changes (Fold-Δs) in output expression in response to miR-FF4 (top row) and Gal4-VPR (bottom row) perturbations. Each column represents an increasing amount of kinase input from left to right. The dashed lines indicate no fold-change (ideal). **(d)** Direct comparison of fold-changes to perturbations between OL and CL variants across kinase dosages. The data represents the maximum dosage of miR-FF4 (miR Marker ≈ $10^{7.75}$ MEFLs) and a dosage of Gal4-VPR with a comparable level of knockdown to the OL (Gal4 Marker ≈ $10^{6.25}$ MEFLs). **(f)** Robustness scores (100% − % deviation due to perturbations) for all OL and CL variants across each kinase input level at the same dosages of miR-FF4 and Gal4-VPR as highlighted in Panel (d). Nominal outputs indicate the level of output in the absence of any perturbations. The individual points are drawn from all experimental repeats. All data were measured by flow cytometry at 48 hours post-transfection in HEK-293FT cells. All errorbars represent the mean ± standard deviation of measurements from three experimental repeats.