Cell
Leading Edge

CellPress

Review

Programmable protein circuit design

Zibo Chen[1] and Michael B. Elowitz[1,2,*]
[1]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA
[2]Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA 91125, USA
*Correspondence: melowitz@caltech.edu
https://doi.org/10.1016/j.cell.2021.03.007

SUMMARY

A fundamental challenge in synthetic biology is to create molecular circuits that can program complex cellular functions. Because proteins can bind, cleave, and chemically modify one another and interface directly and rapidly with endogenous pathways, they could extend the capabilities of synthetic circuits beyond what is possible with gene regulation alone. However, the very diversity that makes proteins so powerful also complicates efforts to harness them as well-controlled synthetic circuit components. Recent work has begun to address this challenge, focusing on principles such as orthogonality and composability that permit construction of diverse circuit-level functions from a limited set of engineered protein components. These approaches are now enabling the engineering of circuits that can sense, transmit, and process information; dynamically control cellular behaviors; and enable new therapeutic strategies, establishing a powerful paradigm for programming biology.

INTRODUCTION

In cells, molecular circuits of interacting genes, proteins, and other biomolecules sense signals transmit and process information, and control cellular processes (Benenson, 2012). These circuits function at a level between the lifeless chemistry of individual molecules and the vibrant dynamic behavior of the living cell. A powerful aspect of molecular circuits is their programmability. By controlling the regulatory interactions among a set of molecular building blocks, synthetic biologists can engineer diverse circuits that perform a broad range of dynamic functions in living cells, similar to the way one can create many different electronic circuits from a handful of electrical components.

Synthetic circuit design was initially dominated by gene regulation, because of the relative ease with which transcriptional interactions could be controlled (Cameron et al., 2014). However, it has now become feasible to design post-translational circuits composed of interacting proteins. Protein circuits can interact directly with endogenous protein-level pathways in the cell to sense or manipulate cell function. They can often respond more rapidly than gene regulation circuits, which require slow and stochastic steps of transcription and translation and whose response times are typically limited by mRNA and protein decay rates (Eldar and Elowitz, 2010). They can also operate across distinct cellular compartments, including the cytoplasm, nucleus, and plasma membrane, among others. Synthetic protein circuits could sense and respond to complex disease states, enabling new therapeutic strategies (Hong et al., 2020). Encoded on viral vectors or even as mRNA, such therapeutic circuits could, in principle, be transiently delivered to a host cell without permanent genetic modification to enhance gene and cell therapies. Thus, there is growing interest in developing the capability for rapid, predictable engineering of programmable protein circuits.

The goal of this review is to survey how far we have come, and how far we have yet to go, toward this goal. We first highlight fundamental molecular circuit design principles developed in the context of nucleic acid circuits with implications for protein circuit design. We then review diverse synthetic protein circuits, illustrating the ways in which these principles come together to enable core functions including sensing, transmitting, and processing signals as well as the ability to dynamically control cellular behaviors. We discuss how these circuits, made from individual protein components, can be combined to create programmable protein-based therapeutic devices. Lastly, we provide an outlook on opportunities for the field going forward. Our emphasis here is on generalizable, enabling strategies for protein circuit design, particularly in eukaryotic cells. Other recent reviews have covered transcriptional (Xie and Fussenegger, 2018), DNA (Scalise and Schulman, 2019), protein-DNA (Jusiak et al., 2016), and protease-based (Chung and Lin, 2020) circuits.

GENERAL PRINCIPLES OF PROTEIN CIRCUIT DESIGN

Today's protein circuits build on concepts established in earlier work in the field of molecular computing. Early theoretical studies recognized an analogy between sets of interacting molecular components and electronic circuits (Bray, 1990, 1995; Monod and Jacob, 1961; Rössler, 1974a, 1974b; Sugita, 1963), pointing out that molecular interaction specificity, playing a role analogous to wiring in electronics, could allow molecular circuits to provide a similarly broad range of computational capabilities.

Early programmable molecular circuits focused on nucleic acids, whose base-pairing interactions allow predictable design based on validated thermodynamic and kinetic models.

Please cite this article in press as: Chen and Elowitz, Programmable protein circuit design, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.007

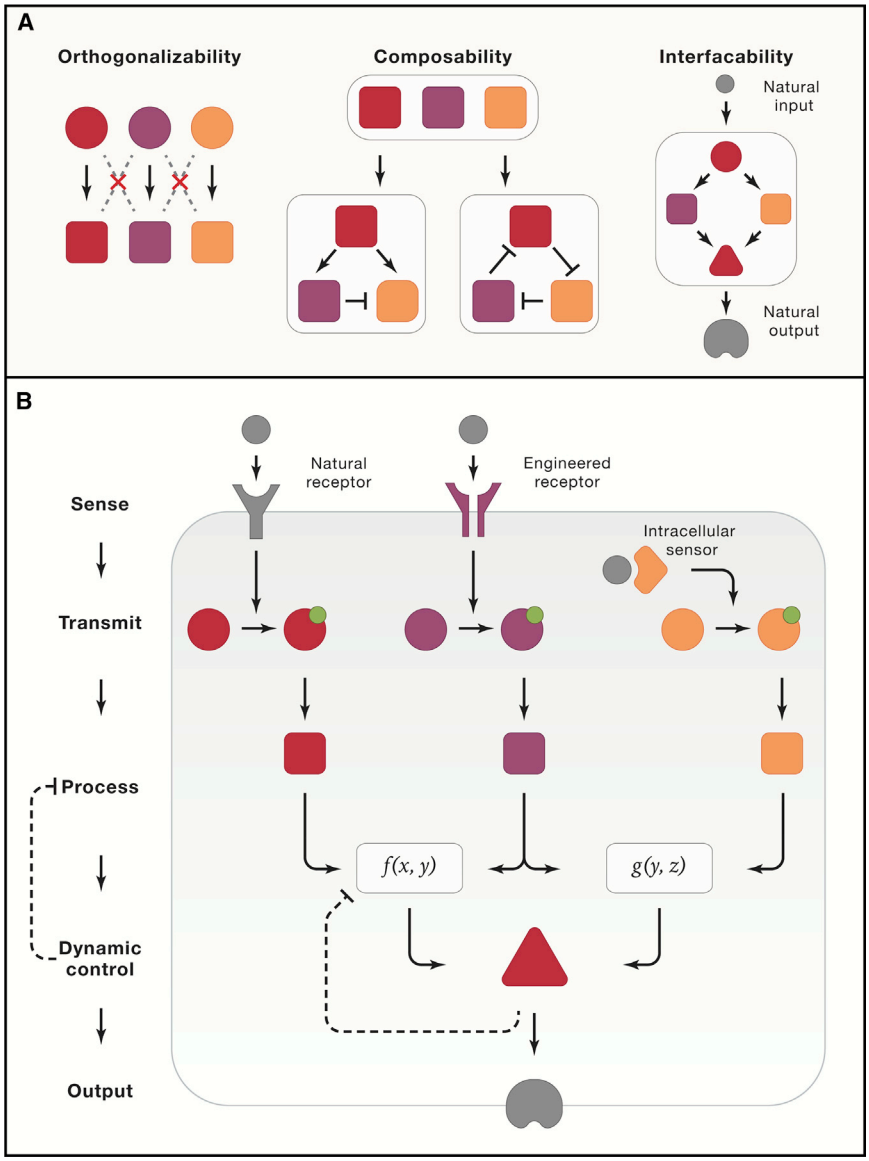

Figure 1. Protein circuits built from modular components can perform diverse functions

(A) Ideal protein circuit components should be orthogonalizable, to allow multiple versions of the same operation without crosstalk (left); composable, so that the output of one component can serve as an input for the next (middle); and interfaceable with endogenous cellular proteins (right).

(B) A generic protein circuit operating in a living cell can sense extracellular and intracellular signals, transmit sensed signals to desired subcellular locations or proteins, process information, provide dynamic control, and produce programmed output. Here, gray shapes represent endogenous proteins, while colored shapes represent synthetic or engineered proteins. Proteins involved in each level of the circuit are distinguished by different shapes. Green circles represent modifications such as phosphorylation.

Paradigm-establishing work in 1982 suggested that designed DNA strands could programmably self-assemble into designed shapes (Seeman, 1982). This was followed by a pioneering demonstration that parallel computation by DNA molecules could solve the directed Hamiltonian path problem in test tubes (Adleman, 1994). The last decades have seen a series of seminal accomplishments, including DNA building blocks that use self-assembly to execute various algorithms (Barish et al., 2005; Wagenbauer et al., 2017; Wei et al., 2012a; Woods et al., 2019), DNA circuits that carry out neural network computations (Cherry and Qian, 2018; Qian et al., 2011), large-scale and high-density memory storage using DNA (Ceze et al., 2019), implementation of chemical reaction networks (Chen et al., 2013; Srinivas et al., 2017), DNA robots that programmably sort cargo molecules (Thubagere et al., 2017) or transport payloads to tumor cells (Li et al., 2018), RNA switches that implement post-transcriptional control (Green et al., 2014), and RNA classifiers (Xie et al., 2011) for cancer cell identification, each expanding the envelope of possible computations achievable by nucleic acids. These studies demonstrated that nucleic acids could provide a versatile toolbox for biomolecular computation.

Nucleic acids are relatively easy to program in a predictable manner, based on their ability to bind, and even cleave, each other. However, they are challenging to interface with endogenous protein pathways in the living cell. By contrast, proteins are more difficult to predictably design, but possess a much larger potential repertoire of activities and interactions including binding, cleavage, ligation, allosteric modulation, and chemical modification. Overcoming this trade-off requires engineered protein circuit components with three key features (Figure 1A). First, they should be orthogonalizable, able to exist in multiple variants that operate equivalently but independently, with minimal crosstalk (Figure 1A, left panel). This property allows the creation of multiple similarly functioning modules that can operate "in parallel." Second, they should be composable, or able to regulate other components of their own type (Figure 1A, middle panel). This property allows one to string them together "in series" to create feedforward and feedback loops that can process signals. In addition, ideal protein components should also be able to directly interface with endogenous cellular proteins (e.g., through allosteric regulation or post-translational modification), to allow sensing and control of cellular pathways (Figure 1A, right panel). Engineered protein components that embody these principles should facilitate design and construction of diverse protein-level devices with a broad set of capabilities.

In addition to these principles, protein circuits, like other computational systems, require four key circuit-level capabilities:

Please cite this article in press as: Chen and Elowitz, Programmable protein circuit design, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.007

CellPress

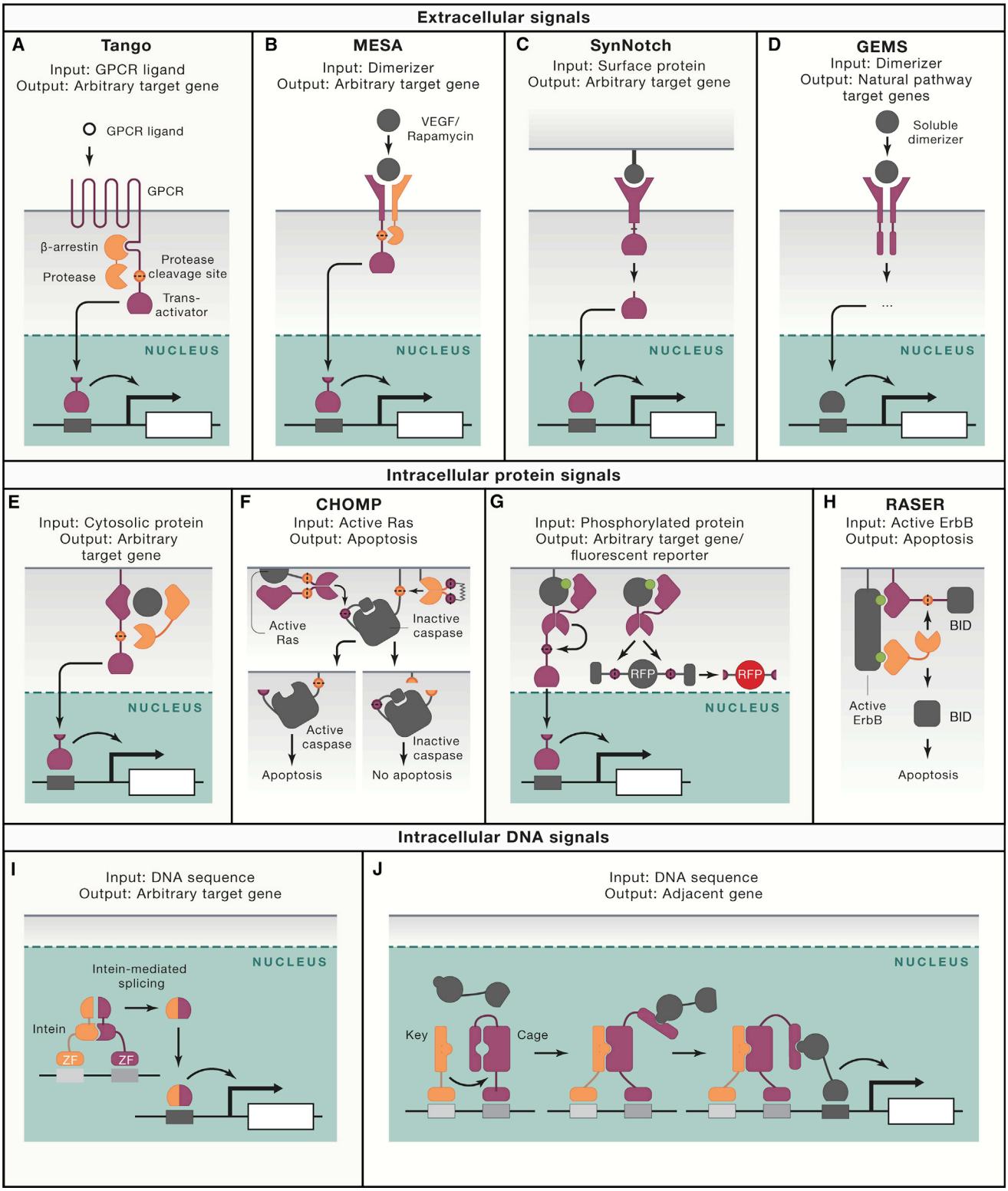

Figure 2. Engineered protein systems sense extracellular and intracellular targets

Existing protein-based sensors use colocalization-induced reconstitution or cleavage as the common theme to detect signals.

(A) Tango couples GPCR activation to the receptor recruitment of protease for cleavage, resulting in release of the transactivator from the membrane for target gene expression.

(legend continued on next page)

Please cite this article in press as: Chen and Elowitz, Programmable protein circuit design, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.007

sensing external and internal inputs, transmitting information from one component to another, processing signals, and exerting dynamic control (Figure 1B). In the next sections, we discuss each of these capabilities. When individual circuits span more than one of these four functional categories, as they often do, we include them in the category that they best illustrate.

SENSING SIGNALS

Sensing of input signals, including extracellular or intracellular proteins, nucleic acids, and small molecules, is the first layer in many circuits. The ideal protein-based sensing system would be able to specifically, quantitatively, and non-destructively detect an arbitrary input in the presence of background noise and convert the sensing event to an internal signal, such as a protein concentration. It should also be able to detect fold change in input signals, in addition to their absolute values (Goentoro et al., 2009). Recent work has achieved different aspects of this vision.

Sensing extracellular signals

Sensing extracellular signals requires discriminating extracellular targets from background molecules and transducing the detection event across the cell membrane. Ideally, a single sensing framework could be equipped with modular extracellular sensing domains and intracellular effector domains. Several synthetic sensors link binding of an extracellular input to activation of intracellular viral proteases. Proteases offer high target site specificity, orthogonality to bacterial and mammalian proteins, and the ability to function similarly in diverse cellular contexts. Equally important, proteolysis is a simple, direct, and powerful way to alter protein activities (Fernandez-Rodriguez and Voigt, 2016).

An early design, termed Tango, used G-protein-coupled receptors (GPCRs) for their innate input specificity, but coupled them to membrane-recruited proteases that divert output to other downstream effectors (Barnea et al., 2008). To achieve this, transactivators are fused to GPCRs, separated by a protease cleavage site. Upon ligand binding, activated GPCRs recruit a chimeric fusion protein made of β-arrestin, which binds activated GPCRs, and a protease, which cleaves the transactivators off the GPCRs, allowing them to translocate to the nucleus and activate target genes (Figure 2A). Subsequent designs modified the system by swapping the locations of the protease and cleavage site (Kipniss et al., 2017) and incorporating split proteases, which become active only upon signal-dependent reconstitution (Djannatian et al., 2011). An alternative approach, termed modular extracellular sensor architecture (MESA), took advantage of the ability of an extracellular input to heterodimerize two receptor subunits, one fused to a protease and the other to a protease-releasable transcription factor domain. Dimerization leads to proteolytic release of the transcription factor and activation of downstream target genes (Daringer et al., 2014; Hartfield et al., 2017; Schwarz et al., 2017) (Figure 2B). These systems have been effective as reporters for ligands of GPCRs or other native receptors. However, their reliance on natural receptor domains limits the range of inputs they can detect.

The synthetic Notch (SynNotch) system offered more flexibility for coupling molecular recognition of natural and synthetic inputs to arbitrary target activation (Morsut et al., 2016). SynNotch receptors comprise customizable extracellular sensing and intracellular effector domains connected by the transmembrane linker from the natural Notch receptor. Engagement with a protein ligand on the target cell surface leads to intramembrane proteolysis and release of the intracellular domain (Figure 2C). The modularity of SynNotch receptors has facilitated applications ranging from specific T cell targeting (Roybal et al., 2016a, 2016b) to multicellular patterning (Toda et al., 2018, 2020). Another system based on the same principle, called transneuronal control of transcription (TRACT), was developed to detect intercellular contacts in embryos (Huang et al., 2016) and trace neuronal connections in fruit flies (Huang et al., 2017).

SynNotch demonstrated remarkable programmability in sensing extracellular signals. However, it is only able to sense target proteins expressed on cell surfaces or that can mediate cell contact, due to the pulling force required to activate signaling. By contrast, other platforms such as "synthekines" and generalized extracellular molecule sensors (GEMSs) can sense soluble extracellular proteins and generate customized outputs (Moraga et al., 2017; Scheller et al., 2018). Synthekines, for example, are artificial cytokines that can dimerize cytokine receptor subunits to induce signaling. In GEMSs, the extracellular

(B) In MESA, extracellular proteins or small-molecule ligands inducibly dimerize the membrane receptors, thereby colocalizing the intracellularly attached protease and its target cleavage site, followed by release and nuclear relocation of the transactivator for gene expression.
(C) The SynNotch receptor works by binding to a surface antigen on the target cell and inducing a cleavage event to release the transcription factor for gene expression.
(D) In GEMSs, a soluble dimerizer ligand dimerizes the engineered receptors to activate certain natural pathways for target gene expression.
(E) An intracellular protein is sensed by colocalizing its two nanobodies, each attached to a protease or its cleavage sequence, resulting in the release of the transactivator from the membrane for gene activation.
(F) A Ras activation sensor works by reconstituting a split protease (purple) via the interaction between active Ras proteins and Ras binding domains (purple). The reconstituted protease proteolytically activates a caged caspase, leading to apoptosis (lower left). To establish an input threshold, a second membrane-tethered protease (yellow) can inhibit caspase activation by proteolytically releasing the caspase from the membrane (lower right). The two proteases can also mutually inhibit each other through cleavage sites (yellow and purple circles).
(G) Protein phosphorylation is sensed by the binding between a phosphorylated protein and its cognate binder, thereby reconstituting a split protease, which either cleaves off a membrane bound transactivator for gene expression or activates a caged red fluorescent protein.
(H) A sensor for the active ErbB protein works by colocalization of its two cognate binders that recognizes phosphorylated sites, resulting in release of a membrane bound Bid protein via protease cleavage, which subsequently mediates apoptosis.
(I) Adjacent DNA sequences are sensed by zinc finger (ZF) domains with genetically fused split inteins and transactivators. Colocalization of the ZF domains results in intein-mediated splicing of the transactivators that activate the expression of a reporter gene.
(J) Colocalization of the "key" and "cage" domains of the LOCKR system at adjacent DNA-binding sites leads to uncaging of a custom recruitment domain for binding of a transcriptional activator, resulting in expression of a neighboring gene.

Please cite this article in press as: Chen and Elowitz, Programmable protein circuit design, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.007

region of a mutated erythropoietin receptor is fused to affinity domains that dimerize in the presence of ligands, while its intracellular domain is fused to signal transduction domains from common pathways such as the Janus kinase-signal transducer and activator of transcription (JAK-STAT) and mitogen-activated protein kinase (MAPK) pathways. Ligand-induced homodimerization of the engineered receptor by extracellular ligands leads to activation of the intracellular domains and their target pathways (Figure 2D). Thus, SynNotch in principle allows arbitrary gene expression following the detection event, but is limited to membrane-bound targets, while synthekines and GEMSs can detect soluble extracellular proteins, but are limited in what downstream pathways they can activate. A hybrid of these two approaches would generalize extracellular sensing to allow robust coupling between arbitrary sensing targets and endogenous gene expression.

Sensing intracellular proteins

Cells encode important signals, such as DNA damage, in the concentrations and states of diverse intracellular proteins, such as p53 (Lakin and Jackson, 1999). Engineered viral proteases provide a flexible platform to sense such signals. For example, fusing distinct target-binding domains to a protease and a cleavage-releasable effector allows effector release to occur only in the presence of the target (Figure 2E). Siciliano et al. (2018) demonstrated this design using two nanobodies that bind distinct epitopes of the same target protein to create the necessary target protein "sandwich." This approach successfully detected protein disease markers involved in hepatitis C virus infection, human immunodeficiency virus infection, and Huntington's disease. One limitation in this scheme is its requirement for two nanobodies that bind different target protein epitopes in a non-exclusive and non-perturbative manner. Recent advances in the rapid development of nanobodies could help to overcome this issue for specific endogenous targets (Wellner et al., 2020). Furthermore, this scheme could be particularly well suited to fully synthetic protein circuits, where one can include artificial epitopes with known binders.

Reversible chemical modifications of proteins enable a huge range of natural signal processing capabilities inside cells (Arkin and Ross, 1994). How can such modifications be sensed by protein circuits? By fusing one half of a split tobacco etch virus protease (TEVP) to a target protein and the other to a protein that conditionally binds the phosphorylated target protein, Wehr et al. (2008) showed that they could conditionally reconstitute active TEVP in response to target protein phosphorylation. Once reconstituted, TEVP could activate or release other proteins engineered to contain cognate cleavage sites (Figure 2G). Together, these examples establish colocalization-dependent reconstitution of a split effector protein as a versatile and reliable approach to detect intracellular proteins or protein states. This approach can be extended to sense active forms of cancer protein markers (Figures 2F and 2H), as discussed below in Therapeutic protein devices.

Sensing nucleic acids

RNA species are informative about the gene expression state of a cell, while the presence or copy number of a DNA locus can be informative about the disease state of a cell, e.g., tumor versus normal. Therefore, it is of interest to engineer protein-based sensors that can detect nucleic acids with high specificity. Converting the sensing event to a customized signal could be achieved by binding of two domains of a split protein to adjacent sites on a target DNA, where they reconstitute a functional protein. One scheme used zinc fingers to bind adjacent DNA target sites, with their colocalization triggering intein-based *trans*-splicing to generate a functional transcriptional activator that in turn activated a reporter gene (Slomovic and Collins, 2015) (Figure 2I). In a separate study, Chaikind et al. (2012) fused split methyltransferase domains to zinc fingers that recognize adjacent DNA-binding sites to achieve sequence-specific DNA methylation. More recently, Kirkpatrick et al. (2020) used the same principle of adjacent recruitment for detection. However, instead of reconstituting a split effector protein, they took advantage of an engineered protein activation system called latching orthogonal cage/key proteins (LOCKR) in which one protein, termed the "key," can release a second protein domain from autoinhibition by a "cage" domain (Langan et al., 2019). Here, recruitment of the key successfully released the adjacent target peptide sequence from the "cage" domain, allowing it to recruit a transcriptional activating domain, effectively converting sequence sensing to transcriptional activity (Kirkpatrick et al., 2020) (Figure 2J). The modular nature of LOCKR in principle allows arbitrary peptide domains to be autoinhibited and released and should enable orthogonal recruitment of a diverse set of transcriptional effectors, allowing sequence-specific transcriptional modulation in cells. These designs demonstrate the many ways in which colocalization can be converted to measurable output signals to detect DNA sequences. Nevertheless, further work is still necessary to increase sensitivity and specificity and to extend these approaches to RNA sensing as a complementary approach to CRISPR-based sensing schemes (van Dongen et al., 2020).

Sensing schemes are critically necessary for emerging cell- and protein-based therapeutics, but several challenges still limit their capabilities. First, existing sensors rely on highly specific target-binding domains. However, for many targets, such domains do not yet exist, are insufficiently specific, or have other undesirable properties such as large size. Nanobodies (Beghein and Gettemans, 2017), designed ankyrin repeat proteins (DARPins) (Stumpp et al., 2008), and designed protein binders (Chevalier et al., 2017) may partly address these issues. Second, even when a specific target-binding domain exists, it is not always clear how binding can be converted into an arbitrary downstream protein-level signal (protein degradation, post-translational modification, etc.). While the systems mentioned above are flexible in some respects, they each produce a single type of output signal (e.g., reconstitution or release of an effector protein). Advances in engineering split proteins that reconstitute upon colocalization could provide a generalizable modality for customizing output signals (Dolberg et al., 2021). Strategies for conditional de-inhibition (uncaging) of proteins (Stein and Alexandrov, 2014) may open up additional flexibility in signal conversion. Third, current technology relies on genetically fusing effector domains onto the proteins of interest or having nanobodies that directly bind to surface hotspots of target proteins.

Please cite this article in press as: Chen and Elowitz, Programmable protein circuit design, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.007

Both of these methods raise the concern of potentially interfering with the folding or function of the proteins they detect. This issue is especially relevant for the intracellular sensing of proteins or protein modifications. Methods that couple transient binding with signal amplification hold the promise of reducing perturbations to the target protein.

INFORMATION TRANSMISSION

Once information is sensed, protein circuits must propagate it, convert it among distinct molecular carriers (e.g., from protein concentration to phosphorylation), and physically transport it from one cellular compartment to another. These information transmission steps can be implemented either directly, through specific protein-protein interactions, or indirectly, through scaffolds that recruit source and destination molecules to the same site, facilitating their interactions.

Direct signal transmission

The most direct way to control transmission is through a protein-protein interaction specificity "code" in which different amino acid sequence variants on one protein specifically interact with corresponding sequences on another protein. A classic example of such a code occurs in bacterial two-component signal transduction systems, in which histidine kinases transfer phosphates specifically to cognate response regulators. A handful of specificity-determining residues control which histidine kinase interacts with which response regulator, usually in an orthogonal, or one-to-one, manner. Altering the specificity-determining residues is sufficient to rewire kinase connections (Skerker et al., 2008). Furthermore, analysis of the coding space suggests that it can accommodate a large number of orthogonal communication channels (McClune et al., 2019). The well-studied specificity code, in combination with the modularity of two-component systems (Schmidl et al., 2019) and their absence in higher eukaryotes (Capra and Laub, 2012), would seem to make it ideal as a synthetic mammalian signal transmission system. In fact, Hansen et al. (2014) transplanted three two-component pathways to HEK293 cells and found that although the pathways lost ligand sensitivity, they retained mutual orthogonality. Subsequent work made these two-component systems ligand activatable in mammalian cells, enabling them to respond to a series of small molecules in a dose-dependent manner, while remaining unresponsive to common endogenous molecules found in mammalian cells (Mazé and Benenson, 2019; Scheller et al., 2020) (Figure 3A). These studies suggest that the two-component system has the potential to serve as a generalizable and scalable framework for orthogonal signal transmission in mammalian cells.

Could one engineer a similar specificity-determining code from scratch? There are numerous examples of analysis and reprogramming of protein-protein interaction specificity (Aakre et al., 2015; Kapp et al., 2012; Zarrinpar et al., 2003). Here, we focus on a few examples of direct information transmission using *de-novo*-designed programmable protein specificity systems, with demonstrated orthogonalizability. Early efforts to design specific protein-protein interactions focused on coiled coils due to their structural simplicity. Researchers used a combination of hydrophobic and hydrophilic (polar) interactions, and peripheral charges to program desired binding specificities (Acharya et al., 2006; Fletcher et al., 2012; Gonzalez et al., 1996; Gradišar and Jerala, 2011; Havranek and Harbury, 2003; Lumb and Kim, 1995; Reinke et al., 2010); however, the limited size of the binding interface between coiled coils restricted the diversity of specificity-determining residues that could fit and therefore limited the number of possible orthogonal pairs that could be designed.

Structurally repeating protein helical bundles and the Rosetta protein design software together extend the interaction surface and allow exploration of a more diverse set of specificity-determining residues (Huang et al., 2014). These advances allowed *de novo* design of a protein-protein interaction code based on accurate placement of buried hydrogen bond networks at the binding interfaces between structurally repeating protein helical bundles (Figure 3B), where the hydrogen bond networks provide binding specificity and hydrophobic interactions contribute to affinity, reminiscent of base pairing in DNA (Boyken et al., 2016). Chen et al. (2019) used this strategy to create sets of mutually orthogonal designed heterodimers (DHDs), including 14 pairs analyzed *in vitro* and 6 analyzed *in vivo*. These components represent a flexible and powerful toolkit to control information transmission within cells via protein-protein interactions. A separate study engineered a specificity code based on a natural interaction between a colicin endonuclease and a cognate immunity protein (Netzer et al., 2018). In this case, the authors used Rosetta to explore perturbations at the protein backbone level, followed by synthesis and analysis of different potential binding partners. This approach produced a broader "many-to-many" network of interactions with a range of affinities. The full coding space for both of these systems may be substantially larger than explored in their initial demonstrations (McClune and Laub, 2020).

These synthetic examples demonstrate remarkable control of protein interaction specificity via rational design of backbone and side-chain interactions, but have remained limited to protein-protein binding. Advances in co-evolution-based protein design, which mines evolutionary correlations in protein sequence to design functional enzymes (Anishchenko et al., 2020; Riesselman et al., 2019; Russ et al., 2020), could help to design even more powerful signal transmission systems involving post-translational modifications and signal amplification.

Indirect signal transmission

In natural circuits, scaffold proteins route signals by recruiting otherwise weakly interacting proteins into close proximity (Good et al., 2011). Engineered scaffold systems can similarly allow controlled routing through swappable recruitment domains. In a pioneering study, Park et al. (2003) engineered such a chimeric scaffold protein by fusing scaffolds involved in yeast mating and osmolarity response pathways, diverting mating input signals to activate osmostress response (Figure 3C). In another study, the Fus3 and Hog1 proteins, which are involved in yeast MAPK pathways for pheromone- and osmotic-pressure-sensing, respectively, were spliced into hybrids that cross-wire and diversify the signal transmission capacities of both pathways (Mody et al., 2009).

Please cite this article in press as: Chen and Elowitz, Programmable protein circuit design, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.007

CellPress

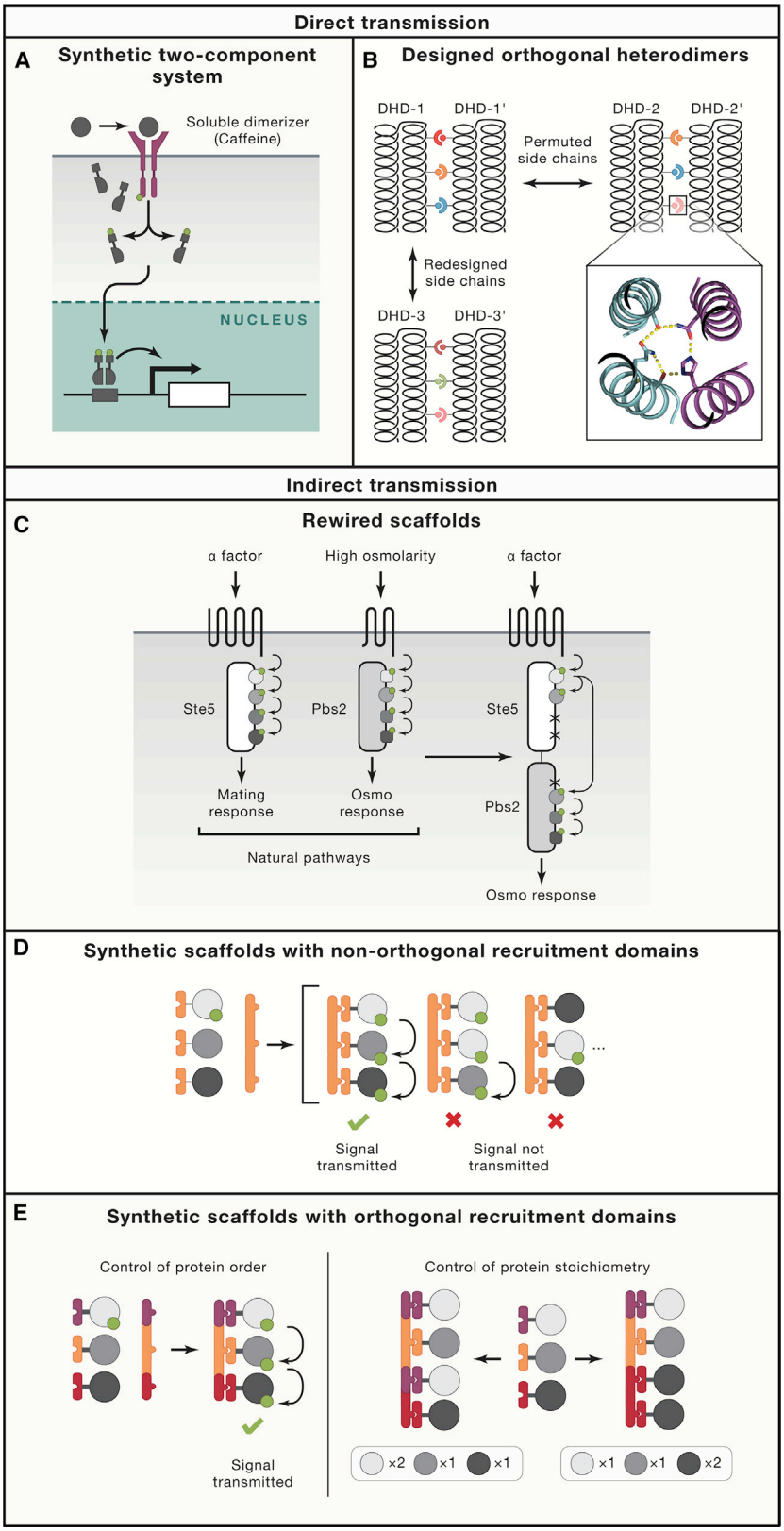

Figure 3. Direct and indirect modes of protein-based information transmission

(A) A bacterial two-component system is transplanted into mammalian cells. The system uses the direct information transmission mode, where activated receptor dimers phosphorylate downstream response regulators, which homodimerize and translocate into the nucleus with fused transactivators for gene activation.

(B) *De-novo*-designed protein heterodimers (DHDs) use synthetic specificity code to allow orthogonal binding. Designed hydrogen bond networks (example in dashed box) at the binding interfaces provide a rich repertoire of specificity code, and a change in the code results in mutually orthogonal DHD pairs (from the DHD 1:1′ pair to DHD 3:3′ pair). The same set of code can be modularly permuted on the same DHD backbone, providing another way to generate orthogonality (from the DHD 1:1′ pair to DHD 2:2′ pair).

(C) Genetic fusion of the scaffold proteins involved in yeast mating and osmostress response pathways rewires information transmission from α factor to osmostress response. In (C–E), green dots represent transmitted signals, e.g., phosphorylation. Signals are transmitted between adjacent units unidirectionally from lighter to darker shading.

(D) Synthetic scaffold proteins with non-orthogonal recruitment domains result in a large number of unproductive arrangements hindering the desired information transmission pathway.

(E) Orthogonal recruitment domains allow control of protein order and composition on the scaffolds.

Please cite this article in press as: Chen and Elowitz, Programmable protein circuit design, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.007

Similar scaffold-engineering approaches have been useful for rewiring and understanding dynamic signaling systems. In one case, a transcriptional activation domain was fused to endogenous signaling proteins that translocate into the nucleus once the target pathway is activated (Krawczyk et al., 2020). For pathways with available nuclear translocators, this approach can reroute signaling pathway activation to the expression of user-defined genes. In other work, Gordley et al. (2016) constructed synthetic phospho-regulon motifs by fusing docking and substrate peptides involved in the MAPK pathway and created fast-acting synthetic phosphorylation circuits. By coupling these to slower transcriptional regulations, the authors studied the effect of dual-timescale network dynamics on cell fate decisions. A semi-synthetic scaffold can be constructed by genetically fusing two heterodimerizing proteins to a natural scaffold protein and a target protein, respectively, such that the target protein can be recruited to the scaffold. Multiple studies used this scheme to alter natural pathway behaviors (Bashor et al., 2008; Groves et al., 2016; Wei et al., 2012b). Together, these examples demonstrate that the physical hybridization of engineered and endogenous proteins can serve as a simple and modular method to rewire information transmission in cells.

Compared with these approaches, designing synthetic scaffold proteins from scratch could provide more flexibility and better insulation from endogenous pathways. Fully synthetic scaffolds have been created from tandem repeats of heterodimerization domains to recruit two (Aper et al., 2018; Whitaker et al., 2012), three (Ryu and Park, 2015), or more (Dueber et al., 2009) proteins to the same scaffold with no predefined order (Figure 3D). More precise control of the order and composition of recruited proteins on a scaffold (Figure 3E), a feature crucial for metabolic engineering, was recently achieved using designed orthogonal protein heterodimers (Chen et al., 2019). These *de-novo*-designed scaffolds offer the additional advantages of being compact (each heterodimer pair is around 150 amino acids) and cooperative (Chen et al., 2020), avoiding signal degradation when the scaffold protein concentration is too high (Levchenko et al., 2000). Together, these examples demonstrate that synthetic scaffold proteins can mediate customized signal transmission.

SIGNAL PROCESSING

Signal processing operations such as logic, amplification, and analog-to-digital conversion enable powerful computational capabilities. Researchers have recently combined the principles of orthogonalizability and composability to implement these capabilities in different ways.

Logic

Logic is ubiquitous in cell signaling, allowing cells to selectively respond only under certain input combinations. In one early study, the *in vitro* assembly of specific heteromeric complexes of designed coiled coils could occur only when certain combinations of peptides were present, effectively implementing different Boolean logic functions (Ashkenasy and Ghadiri, 2004; Ashkenasy et al., 2004). Although the harsh *in vitro* reaction conditions make this scheme unlikely to work *in vivo*, it served as an early proof of principle that a group of proteins with predefined interaction specificity could carry out logic operations.

In living cells, viral proteases represent ideal components for protein-based signal processing. These proteases recognize and cleave specific, compact target sequences in diverse protein contexts, can be engineered to stabilize or destabilize other proteins, and are found in many variants with orthogonal cleavage-site specificities. Viral proteases were used to selectively control target proteins *in vitro* (Holt and Kwong, 2020; Stein and Alexandrov, 2014; Stein et al., 2017) and in bacteria (Fernandez-Rodriguez and Voigt, 2016; Moser et al., 2018). However, it had remained unclear whether they could be engineered to regulate one another in a composable manner within cells and thereby allow the design of diverse protein circuits from a small set of components. Two systems, termed circuits of hacked orthogonal modular proteases (CHOMP) and split-protease-cleavable orthogonal-coiled coil-based (SPOC) logic circuits, demonstrated protease circuit design in mammalian cells (Fink et al., 2019; Gao et al., 2018) (Figure 4A). Protease-protease regulation was achieved through a ''zipper clipper'' mechanism in which complementary binding domains such as leucine zippers reconstitute split protease subunits. By ''clipping'' the zippers off of target protease subunits, an upstream regulatory protease can irreversibly deactivate them. Alternatively, positive protease-protease regulation can be achieved by reconstituting a split protease through cleavage of an autoinhibitory domain. Additionally, proteases can regulate each other by cleaving off or revealing fused degrons (Figure 4A). In this way, proteases can be chained into regulatory cascades, combined to generate diverse binary logic operations, and further assembled into incoherent feedforward loops to produce tunable analog filters (Gao et al., 2018).

A complementary and reversible system, cooperatively inducible protein heterodimer (CIPHR), was created using a set of *de-novo*-designed orthogonal protein heterodimers (Chen et al., 2019). CIPHR regulates colocalization of arbitrary protein units via two complementary protein interaction modalities: competitive binding among DHD pairs dissociates preformed protein complexes to achieve the negation operation, whereas rationally fused monomers from different heterodimer pairs bridge the specific association of target protein units to achieve activation (Figure 4B). Because protein binding appears to be independent of cellular context, CIPHR performs logic operations in cell-free extracts, yeast, and T cells (Chen et al., 2020). The ability to *de novo* design a virtually unlimited set of protein components to mediate non-covalent interactions allows orthogonal systems to be produced, whereas the demonstrated composability of viral proteases to activate/inhibit each other enables construction of multi-layered calculations. A combination of these two approaches should allow the creation of ideal protein-based processors with further increased scalability to allow more powerful computations (Figure 4C).

Analog-to-digital conversion

Ultrasensitive responses convert analog input signals to digital all-or-none outputs, an operation that is critical for background noise suppression, accurate detection of molecular targets

Please cite this article in press as: Chen and Elowitz, Programmable protein circuit design, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.007

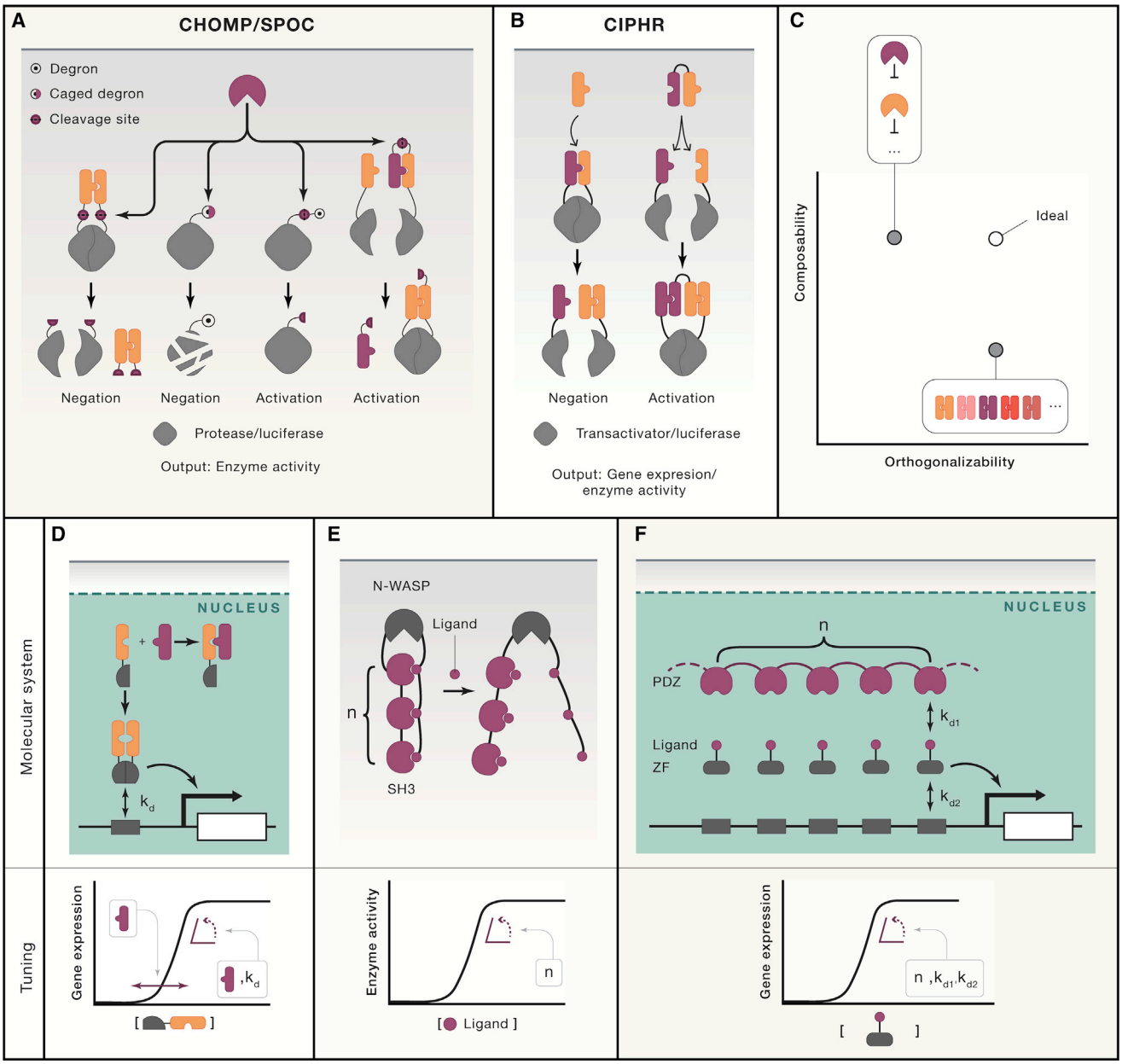

Figure 4. Protein-based processors carry out logic operations and analog-to-digital conversions

(A) In protease-based logic gates CHOMP and SPOC, proteases can carry out the negation operation by cleaving off the fused dimerization domains or revealing a caged degron on other proteases and activate each other by cleaving off a fused degron or an inhibition domain that prevents binding of dimerization domains for split protease or luciferase reconstitution.

(B) The designed heterodimer-based logic gate CIPHR carries out the negation operation via competitive binding to separate a preformed dimer and the activation operation by bridging two otherwise non-interacting monomers and reconstituting fused split protein domains.

(C) CHOMP/SPOC circuits allow proteases to directly inhibit and activate each other, offering composability. CIPHR makes use of potentially unlimited *de-novo*-designed protein heterodimers, enabling orthogonalizability. A combination of these two schemes should result in an ideal system with full scalability.

(D) An analog-to-digital converter that makes use of intermolecular sequestration, where a homodimer with fused DNA-binding domains can be separated by a higher affinity inhibitory-binding domain. The concentration of the inhibitor controls the threshold and the slope of response, with the latter being additionally tunable by the DNA-binding affinity of fused DNA-binding domains.

(E) Intramolecular sequestration of the N-WASP domain by fused tandem repeats of SH3-ligand pairs results in ultrasensitivity where the slope is controlled by the number of SH3-ligand interactions.

(F) Ultrasensitive transcriptional activation is achieved by multivalent cooperative binding between a tandem repeat of PDZ domains that recruits multiple PDZ ligand-ZF fusions. Ultrasensitivity is tunable by the PDZ ligand and ZF-DNA-binding affinities and the number of PDZ domains in tandem.

Please cite this article in press as: Chen and Elowitz, Programmable protein circuit design, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.007

such as antigens, and dynamic behaviors such as oscillation and multistability. In natural circuits, ultrasensitivity is realized through cooperativity, stoichiometric inhibitors, and futile phosphorylation cycles, among other mechanisms (Ferrell and Ha, 2014a, 2014b, 2014c).

An ideal synthetic ultrasensitive protein module would allow independent tuning of response threshold and the level of ultrasensitivity, defined as the fold change in response for a given fold change in input. A simple but powerful mechanism to achieve this is molecular titration, in which an inhibitory molecule stoichiometrically binds to and inhibits a target. This scheme was realized synthetically using nucleic acids (Kim et al., 2006). Applying it to proteins, in one study, an inhibitor blocked activity of a transcription factor by tightly binding to it and preventing it from forming functional homodimers, such that the transcription factor became active only when its concentration exceeded that of the inhibitor (Buchler and Cross, 2009). By tuning the inhibitor concentration and the affinity for DNA binding, one could vary both the concentration threshold for activation and the ultrasensitivity of the response (Figure 4D). Similar effects can be achieved by engineering reciprocal proteolytic inactivation. In this system, two proteases inactivate one another, causing a thresholding effect, in which each protease has to exceed a threshold set by the other to effectively cleave its target (Gao et al., 2018).

Intramolecular sequestration can also lead to ultrasensitivity. Inspired by the autoinhibition-based cooperative binding in the neural Wiskott-Aldrich syndrome protein (N-WASP) switch from nature, Dueber et al., (2007) constructed a synthetic ultrasensitive switch where the N-WASP output domain was caged by flanking repeats of Src Homology 3 (SH3) domains and their peptide ligands fused on either side. The degree of ultrasensitivity, and to a lesser extent the threshold, were shown to be tunable by the number of SH3 interaction modules (Figure 4E). Additionally, Bashor et al. (2019) took advantage of multivalent cooperative binding, using a polymer of PDZ domains to recruit multiple synthetic zinc fingers fused to PDZ ligands. Three independently tunable parameters, namely the binding affinities between PDZ domains and PDZ ligands, between zinc fingers and DNA, and the number of PDZ repeats, allowed control of ultrasensitivity (Figure 4F). These examples demonstrated intermolecular sequestration, intramolecular inhibition, and multivalent cooperativity as viable means to generate ultrasensitivity at the protein level. Further developments in protein engineering should allow other ultrasensitivity schemes (e.g., multisite phosphorylation, positive feedback, phosphorylation cascades [O'Shaughnessy et al., 2011]) to be realized in cells, each compatible with various types of circuit modules requiring specific input and output formats.

Toward scalable processing

In nucleic acid circuits, one elegant reaction—strand displacement—enabled an astonishing variety of powerful molecular programs. In strand displacement, a single invading DNA strand binds to an exposed complementary sequence on a double-helical DNA duplex and kicks off a competing strand. This simple interaction is composable: the strand that is kicked off by one input strand can act as the input for a subsequent strand displacement. It is also orthogonalizable: many independent strand displacement reactions can occur in parallel. It has therefore been used as the basis for a wide range of complex molecular programs (Green et al., 2014; Zhang and Seelig, 2011). Is it possible to create an analogous operation at the protein level?

Recently, single-step strand displacement reactions were implemented in proteins using coiled coils (Dai et al., 2019; Gröger et al., 2017; Groth et al., 2018). Although not yet fully scalable, the ability to add more orthogonal components to this system could further expand its capabilities. A similar principle was also used to design a powerful DNA-templated displacement system. Lebar et al. (2018) used transcription activator-like effectors (TALEs) that recognize target genomic DNA sequences and control gene expression. When bound, these factors inhibit the binding of other TALEs to adjacent sequences in a polar (directional) manner. This allows them to perform logic operations (Lebar et al., 2018). In this system, protein-protein interactions are organized along a designed DNA sequence, and steric hindrance among TALEs controls the computation being performed. This scheme is powerful but requires gene expression steps to achieve multi-stage programs. Additionally, the requirement for a DNA strand with specific sequences limits the portability of this system. While we currently lack a protein-based system with the scalability of DNA-based strand displacement systems, improvements in the design of orthogonal protein pairs, together with method development to better understand and design protein binding kinetics, should advance the field toward this goal.

The signal processing schemes discussed above all assume that the relevant signal is encoded in the concentration of a particular protein species. However, natural (especially mammalian) sensing systems appear to use promiscuous (many-to-many) interactions between sets of ligand and receptor variants to selectively respond to complex combinations of their inputs (Antebi et al., 2017). Computational approaches indicate that competition to form a variety of protein complexes with different activities can perform complex signal processing operations (Su et al., 2020). It will be interesting to see whether these principles can be adapted to enable synthetic circuits with similar functions.

DYNAMIC CONTROL

Dynamic control systems permit robust adaptation to the environment, oscillations and time-based regulation, and multistability, the basis for cellular memory, among other behaviors (Sowa et al., 2015). Recent work suggests that synthetic protein-based circuits could provide similar functionality.

Adaptation

Many biological circuits maintain constant output level across a broad dynamic range of steady-state input concentrations, while also responding transiently to input perturbations (Barkai and Leibler, 1997). By incorporating feedback and feedforward loops (Ma et al., 2009), researchers have begun to engineer protein circuits with similar adaptive capabilities.

One recent study showed how the LOCKR technology, introduced above, could be used to create a protein-level feedback loop that adapts to changes in its input (Figure 5A). In this system, a "key" protein exposes an otherwise caged degron on a transcription factor, causing its degradation. The key is under

Please cite this article in press as: Chen and Elowitz, Programmable protein circuit design, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.007

CellPress

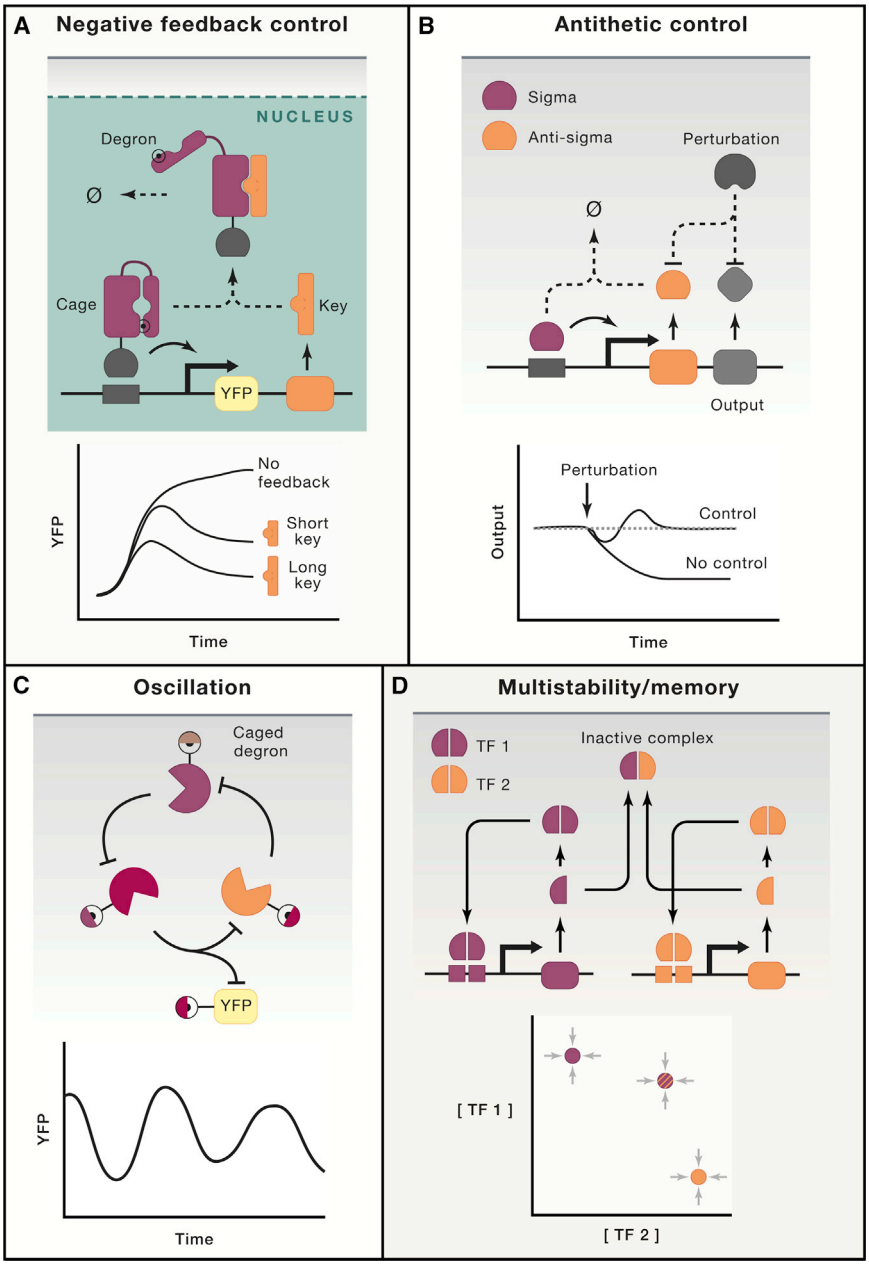

Figure 5. Protein-based dynamic control systems enable adaptation, oscillation, and memory

(A) A negative-feedback system based on LOCKR. The cage is fused to a transactivator, which activates expression of a yellow fluorescent protein (YFP) reporter and the key. Binding between the cage and the key results in exposure of a sequestered degron and subsequent degradation of the cage-transactivator fusion, resulting in reduced expression level of the YFP reporter (top). The strength of the negative feedback is tunable by the length of the key protein, which controls its binding affinity to the cage (bottom).

(B) Antithetic control provides perfect adaptation. The bacterial sigma factor activates expression of the target protein (gray) and the anti-sigma factor, which binds to the sigma factor and renders it inactive (top). Perturbations applied to the target protein and anti-sigma factor result in reduced level of the target protein without antithetic control, but perfect adaptation back to the original level with antithetic control (bottom).

(C) Three proteases that inhibit each other by exposing caged degrons (top) lead to oscillatory behavior in bacteria (bottom).

(D) In the synthetic MultiFate system, transcription factors homodimerize to activate their own expression and heterodimerize to inhibit each other (top). This circuit design can produce multiple stable states with different levels of expression of the transcription factors (bottom).

the control of the transcription factor, generating a negative feedback loop that adapts to changes in inputs (Ng et al., 2019). This work showcased modular tuning of feedback strength via rational design of protein-binding affinities, making the control circuit programmable at the protein level. Another study addressed the challenge of concentration tracking, in which one cellular protein concentration tracks that of another, despite noise. To perform this task in bacterial cells, Hsiao et al. (2015) designed a synthetic negative feedback circuit whose input species was a scaffold that reconstituted two-component signaling. Increased signaling in turn activated production of an anti-scaffold protein, sequestering the input, and completing a negative feedback loop.

These two circuits demonstrated feedback with tunable setpoints, but did not achieve robust perfect adaptation. That is, their responses did not return exactly to their setpoint across a wide range of steady-state input levels and biochemical parameter values. Using principles of control theory recently allowed the design of an "antithetic" integral feedback system that tracks and integrates the deviation from a target value over time (Aoki et al., 2019) (Figure 5B). While the circuit is mostly transcriptional, it achieves integral feedback through a protein-level mechanism: dimerization and mutual annihilation of a sigma factor (specialized bacterial transcriptional regulators) and its cognate anti-sigma factor (an inhibitor that specifically binds to and inhibits the sigma factor). In fact, an analogous architecture could potentially be implemented entirely at the protein level, enabling more rapid adaptation by omitting transcription and translation steps. Together, the examples mentioned above highlight the role of mutual annihilation and inducible degradation as two simple and powerful protein-level mechanisms to achieve adaptation.

Oscillation

Many biological processes, such as the cell cycle and circadian clock, are periodic, sequentially advancing the cell from one

Please cite this article in press as: Chen and Elowitz, Programmable protein circuit design, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.007

stage or phase to the next before restarting again from the beginning. Reconstitution of the natural three-protein (KaiABC) phosphorylation-based oscillator at the core of the cyanobacterial circadian clock demonstrated that an all-protein oscillator is not only possible but also achievable with just a few protein species (Nakajima et al., 2005; Rust et al., 2007), provoking the challenge of engineering analogous synthetic protein oscillators.

One way to explore oscillations is by transplanting systems into new hosts. Recently, Zhang et al. (2017) transplanted components of the nuclear factor κB (NF-κB) pathway, which is known to oscillate in mammalian cells under some conditions (Hoffmann et al., 2002), to yeast cells, where its component levels could be systematically varied without perturbing cellular functions. Oscillations, resulting from a time-delayed negative feedback that combines transcriptional and post-translational regulation, were observed in the new host, implying that key protein-level interactions preserved their functions and could be used to implement this dynamic function. A separate study achieved a fully protein-level oscillator in bacteria. Using three orthogonal viral proteases, researchers created a negative feedback loop, resembling the repressilator (Elowitz and Leibler, 2000), in which each protease specifically cleaves the next protease in the cycle, revealing a degron and thereby targeting that protease for degradation (Gao et al., 2019). This circuit generated oscillations for up to 6 hours, representing a key step toward a fully protein-based oscillator (Figure 5C). Creating tunable protein-level oscillator modules that can drive arbitrary target proteins should enable analysis of natural dynamic responses and enable temporal control in engineered circuits.

Memory

Memory allows cells to alter their behavior depending on their own individual history. Recent work has demonstrated different ways that protein-based circuits could be used to store and read out information encoded in the states of proteins or DNA, providing a foundation for protein-based memory.

Prions provide a particularly direct way to implement protein-level memory in cells. Prion-containing proteins can persist in either soluble or aggregated states. Exploiting this property, Newby et al. (2017) engineered prion-containing proteins that could be reversibly switched among soluble or aggregated states and could also be read out through a synthetic transcription factor fused to the protein, which is sequestered by aggregation and therefore active only in the soluble state. Using these systems, the authors showed that yeast cells could "remember" a heat-shocked event for more than 10 generations.

In contrast to the prion system, where the same protein represents both the "storage medium" and the stored information, chromatin regulatory systems can store information in DNA or histone modifications. In natural epigenetic systems, chromatin regulators actively propagate these modifications, ensuring their stable mitotic inheritance (Allis and Jenuwein, 2016). To understand the minimal features required for such a *cis*-acting epigenetic memory system, Park et al. (2019) sought to engineer an orthogonal and synthetic chromatin-based epigenetic memory system in mammalian cells. They borrowed an epigenetic modification, N^6-methyladenine, widely used in bacteria, but not commonly found in mammalian cells, along with protein domains that recognize this mark or catalyze its formation. Individually fusing these domains to DNA-binding proteins provided "read" and "write" functionality, while combining them together coupled reading to writing, propagating the modification for multiple cell generations. This work thus demonstrated a fully synthetic and generalizable epigenetic memory system.

In complementary work, Zhu et al. (2021) engineered transcription factors that positively autoregulate their own expression as homodimers and inhibit one another's activity through heterodimerization (Figure 5D). The authors demonstrated that 3 transcription factors could produce 7 distinct states, each stable for weeks of continuous culture. Because it implements inhibition through engineered protein-level interaction domains, this architecture scales, generating exponentially increasing numbers of cell states as one introduces additional transcription factors. Together, these approaches provide a foundation for engineering memory in future protein circuits.

THERAPEUTIC PROTEIN DEVICES

The abilities to sense, transmit, process, and dynamically respond to protein-level inputs, discussed above, together open up the possibility of engineering therapeutic devices that specifically target cells in disease states. Such therapeutic circuits could be designed and validated rapidly compared with conventional drugs. Here, we focus on therapeutically relevant protein circuits and assemblies that exemplify the promise of rational design.

Logic-gated cancer immunotherapy

As a proof of principle for a therapeutic circuit, Gao et al. (2018) exploited the composability of multiple engineered proteases (Figure 4A) to create circuits that sense activation of the Ras oncogene and conditionally trigger cell death when it exceeds a threshold set by a second protease (Figure 2F). The entire circuit, which could be encoded as a single multi-protein "gene", selectively killed cells with elevated Ras activation, while largely sparing co-cultured off-target cells. In a separate study, Chung et al. (2019) targeted the ErbB receptor found to be active in cancer cells. Instead of inhibiting its activity, they used the activation of ErbB receptors to recruit and co-localize designed proteases to the cell membrane, which then selectively cleave off membrane-tethered cargo proteins to activate downstream responses, including apoptosis and expression of endogenous anti-cancer genes (Figure 2H). Both of these schemes use protease-based circuits to conditionally link otherwise unrelated endogenous input and output pathways—a programmable "classification" strategy that is distinct from pharmacological inhibition of disease-driving proteins. Developing sensors that can detect a broad spectrum of cancer markers or pathway activities, expanding the toolkit of signal transmission and processing operations described above, and improving the efficiency of *in vivo* circuit delivery, should allow rapidly engineerable therapeutic circuits.

One major application of protein-based circuits is combinatorial recognition of cancer cells based on multiple surface antigens. One scheme combines sensing, cross-membrane information transmission, and signal processing to selectively activate chimeric antigen receptor (CAR) T cells in the presence of target tumor cells (Hong et al., 2020). In addition to the

Please cite this article in press as: Chen and Elowitz, Programmable protein circuit design, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.007

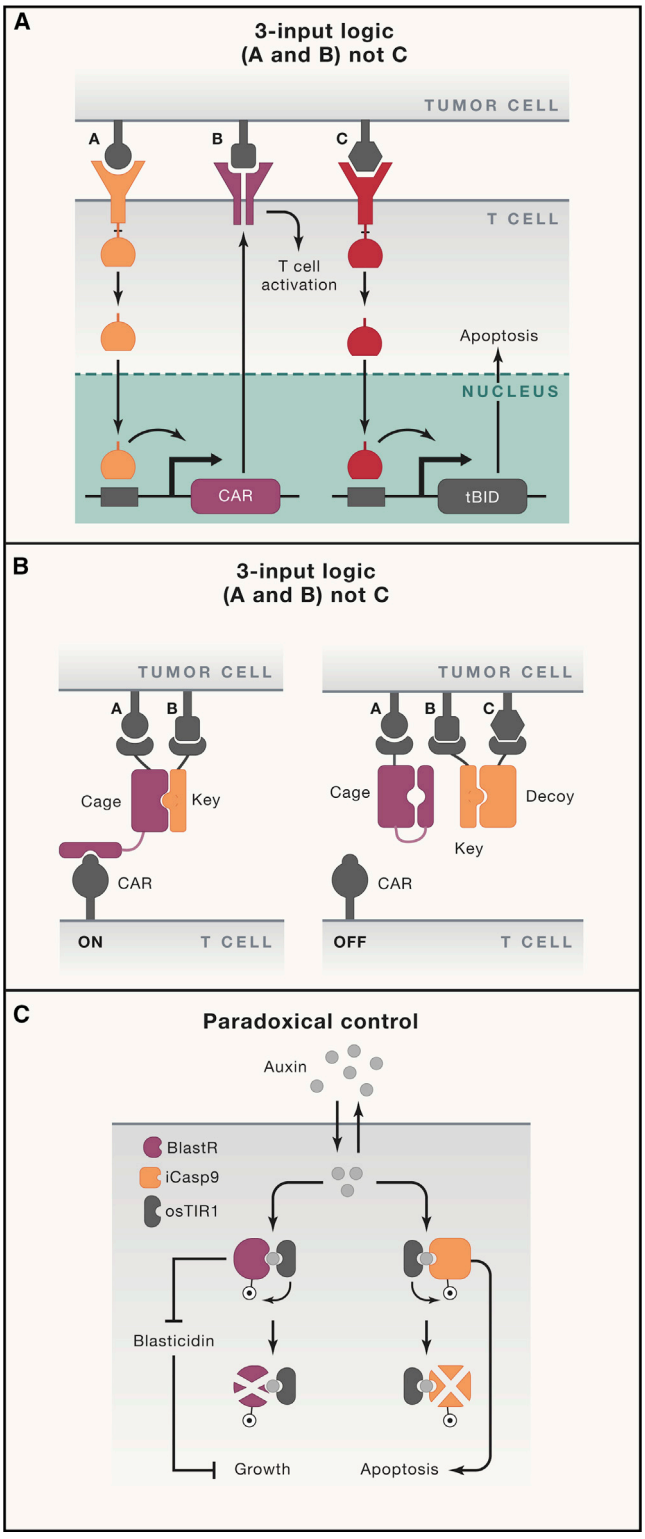

Figure 6. Protein circuits can provide useful capabilities for cell-based therapeutics

(A) In the SynNotch "(A AND B) NOT C" logic gate, antigen A activates a SynNotch receptor, causing expression of a chimeric antigen receptor (CAR), which recognizes the second antigen B and triggers the T cell response. Recognition of a third antigen, C, by an additional SynNotch, leads to expression of truncated BID (tBID) to trigger apoptosis.

(B) A co-LOCKR circuit carries out the "(A AND B) NOT C" logic for tumor surface antigen recognition. The presence of antigens A and B recruits their cognate targeting domains with fused key and cage domains, resulting in the uncaging of an epitope recognized by the CAR on a T cell and its subsequent activation. An additional antigen C recruits a decoy domain that sequesters the key, thereby re-caging the epitope on the cage domain, rendering the T cell inactive.

(C) The paradoxical circuit design enables robust population control. Auxin induces Oryza sativa Transport Inhibitor Response1 (osTIR1)-mediated degradation of a blasticidin S deaminase ($Blast^R$), which enables cell survival in media containing blasticidin, and iCasp9, which promotes apoptosis. This setup prevents cheater mutations from escaping regulation.

previously mentioned SynNotch platform that performs AND logic and only activates T cells in the presence of two tumor surface markers (Roybal et al., 2016b), an inhibitory CAR was incorporated into CAR T cells to identify surface markers unique to healthy cells and block T cell activation. This "AND-NOT" logic reduced off-target cytotoxicity (Fedorov et al., 2013). Similarly, OR logic was used to activate the T cell when either of two target antigens is present, providing resistance to antigen escape. Protein level strategies to install OR logic in CAR T cells include co-expression of two different CARs in each T cell (Ruella et al., 2016) and tandem fusion of two recognition domains into a hybrid CAR (Hegde et al., 2016; Zah et al., 2016, 2020). Recently, CAR T cells with two orthogonal SynNotch constructs were engineered to perform (A AND B) NOT C (Figure 6A), 3-input AND, and 3-input OR logic operations (Williams et al., 2020).

Two other recent studies have extended CAR approaches to achieve multi-input and externally controlled protein logic. Cho et al. (2018) engineered "zipCAR" proteins that use modular extracellular binding domains made of designed coiled-coil heterodimers (Synzips). T cells with zipCAR domains could be directed to tumor cells with desired surface antigens using adaptor proteins consisting of complementary Synzips fused to antibodies. This scheme naturally enables OR logic using multiple synzip-antibody fusions as well as more complex AND logic by using two distinct zipCARs (Cho et al., 2018).

To detect co-expression of multiple surface proteins, Lajoie et al. (2020) used *de-novo*-designed protein switches that were modified to activate and expose an epitope only when co-localized on the cell surface (co-LOCKRs). Introduced as purified proteins, they allow CAR T cells to conditionally target cells co-expressing two antigens and not expressing a third ([A and B] not C logic) within a mixed population of cells (Lajoie et al., 2020) (Figure 6B). The post-translational nature of co-LOCKR allows it to simultaneously interrogate multiple proteins on the surface of the same cell without intervening transcriptional regulatory steps. Taken together, these systems add new capabilities to the CAR T paradigm, allow control with extracellular proteins, and limit antigen escape. It will be exciting to see how these rapidly expanding approaches can be translated to the clinic.

Cell population control

In the immune system, numerous cell types collectively coordinate responses to challenges through a combination of intracellular processing circuits and intercellular signaling. Protein circuits could implement analogous synthetic immune-inspired systems. This paradigm will minimally require quorum sensing and

Please cite this article in press as: Chen and Elowitz, Programmable protein circuit design, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.007

population control circuits that enable cells to sense and control their own population density using "private" communication channels that do not crosstalk with endogenous pathways. Pioneering synthetic population control circuits focused on bacterial cells, where core components of the circuits were placed under transcriptional control (You et al., 2004). Recent work demonstrated protein-based circuits that combine a synthetic auxin-signaling system with sensing and feedback capabilities to achieve population control in mammalian cells (Ma et al., 2020). In this system, the authors created a two enzyme "sending" pathway for constitutive biosynthesis of the plant hormone auxin from tryptophan. Using an auxin-inducible degron, they further connected auxin sensing to regulation of cell growth in "receiving" cells. A critical concern in population control is selection for "cheater" mutants that lose the ability to sense signals and thereby escape control. Seminal work by Hart et al. (2014) and Karin and Alon (2017) in natural cytokine systems showed that a cytokine counterintuitively stimulates both cell proliferation and death and leads to suppression of likely mutational escape paths. Implementing an analogous circuit synthetically successfully limited cell density for at least 43 days in cell culture (Figure 6C). Future iterations of this circuit, transposed to *in vivo* settings, could enable density-dependent responses and avoid overgrowth of therapeutic cells.

Taken together, protein-based circuits exhibit remarkable programmability both inside and outside of cells. Translating these systems into therapeutics will benefit from expanding their sensing capabilities, improving delivery systems, and addressing questions about potential immunogenicity for extracellular proteins. Looking ahead, we anticipate protein-based circuits emerging as a powerful class of smart and programmable therapeutics by themselves and as essential components of future cell therapies.

OUTLOOK

Each level of protein circuitry discussed above—sensing, transmission, processing, and dynamic control—has now been implemented, with protein components that are at least partially orthogonalizable, composable, and interfaceable with natural components. Nevertheless, realizing the full potential of protein circuits will require additional advances. First, while we now have protein-based sensors of specific extracellular and intracellular targets, many operating through reconstitution of protein domains (Figure 2), we still lack a fully general system that can sense arbitrary protein inputs and conditionally activate different types of protein outputs for downstream transmission. Second, current signal transmission systems repurpose natural scaffolds, often transplanted from other species, limiting the number of independent signals that can be routed in a given system (Figure 3). Synthetic proteins with designed interaction specificities promise to enable multiple signal transmission channels orthogonal to one another and to host cell pathways, expanding computational power. Third, existing protein signal processing systems have demonstrated logic operations and analog-to-digital conversion (Figure 4). Some systems also exhibit multiple orthogonal variants (DHDs) and composable interactions (engineered proteases). Nevertheless, components that are not only orthogonalizable and composable but also utilize post-translational modifications to enable reversible regulation, signal restoration, and integration are needed if synthetic protein circuits are to more closely approach the power and specificity of their endogenous counterparts. Fourth, most of the circuits discussed here build upon specific one-to-one protein-protein interactions. However, natural eukaryotic signaling systems often comprise sets of promiscuously interacting proteins that, together, can perform complex computations through mechanisms loosely analogous to neural networks (Su et al., 2020). The ability to rationally design similar promiscuous network architectures could lead to distinct paradigms for biochemical computation that exceed the capabilities of simpler Boolean logic. And fifth, advances in engineering other protein-protein interactions such as degradation and phosphorylation should expand our ability to interface synthetic and endogenous protein circuits.

Looking ahead, a flexible protein circuit design framework would provide a basis for future programmable sensing and therapeutic devices. In a cell-free context, protein circuits incorporating sensing and amplification could allow multiplexed detection of diverse target proteins. Intracellularly, as therapeutics, protein circuits, perhaps delivered with viruses or as mRNAs (Sahin et al., 2014), could target disease states, defined by combinations of protein levels or activities, more specifically than molecular drugs. Designed post-translational modifications, when coupled with energy expenditure (e.g., phosphorylation), could enable proofreading mechanisms (Hopfield, 1974) that will be necessary to ensure accurate detection and transmission in protein-based therapeutic circuits. Scalable protein circuits could enable combinatorial tumor surface antigen recognition that is robust against antigen escape and off-target activation. Finally, at the multicellular level, combining sensing, processing, and feedback capabilities should allow customized protein circuits for cell-cell communication, pattern formation, and tissue generation, going beyond what have been demonstrated in recent studies (Stapornwongkul et al., 2020; Toda et al., 2020).

Advances in protein circuit design will depend on parallel advances in protein engineering. Rational design of protein conformational changes (Langan et al., 2019; Wei et al., 2020) should allow individual proteins to play more active roles as circuit components. Proteins designed to respond to small molecules could provide additional layers of control via induced dimerization and degradation (Foight et al., 2019; Rakhit et al., 2014; Shui et al., 2021). At the supramolecular level, proteins could act as information-bearing particles to program their own self-assembly (Winfree, 1996). Orthogonal protein interaction domains will allow programmable self-assembly of proteins. This includes designed protein phase separation (Banani et al., 2016; Fisher and Elbaum-Garfinkle, 2020; Schuster et al., 2018), which can be rationally engineered from weak multi-valent protein interactions to provide intracellular compartments for protein circuit operations, allowing noise reduction (Klosin et al., 2020), orthogonalization (Reinkemeier et al., 2019), and more complex protein-level computation.

In DNA circuits, a lively interplay of basic research and applications has elucidated fundamental questions about self-assembly and molecular computation while also generating useful nucleic acid tools and devices. A similar dynamic in the younger field of protein circuit design should provoke fundamental questions about circuit architecture, dynamics, and function, while enabling new generations of programmable biological devices.

ACKNOWLEDGMENTS

We thank E. Winfree, J. Gregrowicz, F. Horns, L. Chong, S. Boyken, A. Askary, and M. Budde for helpful discussions. This work was supported by National Institutes of Health grant R01 MH116508, the Allen Discovery Center program under award number UWSC10142, a Paul G. Allen Frontiers Group-advised program of the Paul G. Allen Family Foundation, and by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under award number R01EB030015. Z.C. was supported by the Damon Runyon Cancer Research Foundation DRG-2388-20 and is a fellow of the Burroughs Wellcome Fund Career Awards at the Scientific Interface. M.B.E. is a Howard Hughes Medical Institute investigator. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

DECLARATION OF INTERESTS

Z.C. and M.B.E. are inventors on patents and patent applications related to protein circuit design.

REFERENCES

Aakre, C.D., Herrou, J., Phung, T.N., Perchuk, B.S., Crosson, S., and Laub, M.T. (2015). Evolving new protein-protein interaction specificity through promiscuous intermediates. Cell *163*, 594–606.

Acharya, A., Rishi, V., and Vinson, C. (2006). Stability of 100 homo and heterotypic coiled-coil a-a' pairs for ten amino acids (A, L, I, V, N, K, S, T, E, and R). Biochemistry *45*, 11324–11332.

Adleman, L.M. (1994). Molecular computation of solutions to combinatorial problems. Science *266*, 1021–1024.

Allis, C.D., and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. Nat. Rev. Genet. *17*, 487–500.

Anishchenko, I., Chidyausiku, T.M., Ovchinnikov, S., Pellock, S.J., and Baker, D. (2020). De novo protein design by deep network hallucination. bioRxiv. https://doi.org/10.1101/2020.07.22.211482.

Antebi, Y.E., Linton, J.M., Klumpe, H., Bintu, B., Gong, M., Su, C., McCardell, R., and Elowitz, M.B. (2017). Combinatorial Signal Perception in the BMP Pathway. Cell *170*, 1184–1196.e24.

Aoki, S.K., Lillacci, G., Gupta, A., Baumschlager, A., Schweingruber, D., and Khammash, M. (2019). A universal biomolecular integral feedback controller for robust perfect adaptation. Nature *570*, 533–537.

Aper, S.J.A., den Hamer, A., Wouters, S.F.A., Lemmens, L.J.M., Ottmann, C., Brunsveld, L., and Merkx, M. (2018). Protease-activatable scaffold proteins as versatile molecular hubs in synthetic signaling networks. ACS Synth. Biol. *7*, 2216–2225.

Arkin, A., and Ross, J. (1994). Computational functions in biochemical reaction networks. Biophys. J. *67*, 560–578.

Ashkenasy, G., and Ghadiri, M.R. (2004). Boolean logic functions of a synthetic peptide network. J. Am. Chem. Soc. *126*, 11140–11141.

Ashkenasy, G., Jagasia, R., Yadav, M., and Ghadiri, M.R. (2004). Design of a directed molecular network. Proc. Natl. Acad. Sci. USA *101*, 10872–10877.

Banani, S.F., Rice, A.M., Peeples, W.B., Lin, Y., Jain, S., Parker, R., and Rosen, M.K. (2016). Compositional Control of Phase-Separated Cellular Bodies. Cell *166*, 651–663.

Barish, R.D., Rothemund, P.W.K., and Winfree, E. (2005). Two computational primitives for algorithmic self-assembly: copying and counting. Nano Lett. *5*, 2586–2592.

Barkai, N., and Leibler, S. (1997). Robustness in simple biochemical networks. Nature *387*, 913–917.

Barnea, G., Strapps, W., Herrada, G., Berman, Y., Ong, J., Kloss, B., Axel, R., and Lee, K.J. (2008). The genetic design of signaling cascades to record receptor activation. Proc. Natl. Acad. Sci. USA *105*, 64–69.

Bashor, C.J., Helman, N.C., Yan, S., and Lim, W.A. (2008). Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. Science *319*, 1539–1543.

Bashor, C.J., Patel, N., Choubey, S., Beyzavi, A., Kondev, J., Collins, J.J., and Khalil, A.S. (2019). Complex signal processing in synthetic gene circuits using cooperative regulatory assemblies. Science *364*, 593–597.

Beghein, E., and Gettemans, J. (2017). Nanobody Technology: A Versatile Toolkit for Microscopic Imaging, Protein-Protein Interaction Analysis, and Protein Function Exploration. Front. Immunol. *8*, 771.

Benenson, Y. (2012). Biomolecular computing systems: principles, progress and potential. Nat. Rev. Genet. *13*, 455–468.

Boyken, S.E., Chen, Z., Groves, B., Langan, R.A., Oberdorfer, G., Ford, A., Gilmore, J.M., Xu, C., DiMaio, F., Pereira, J.H., et al. (2016). De novo design of protein homo-oligomers with modular hydrogen-bond network-mediated specificity. Science *352*, 680–687.

Bray, D. (1990). Intracellular signalling as a parallel distributed process. J. Theor. Biol. *143*, 215–231.

Bray, D. (1995). Protein molecules as computational elements in living cells. Nature *376*, 307–312.

Buchler, N.E., and Cross, F.R. (2009). Protein sequestration generates a flexible ultrasensitive response in a genetic network. Mol. Syst. Biol. *5*, 272.

Cameron, D.E., Bashor, C.J., and Collins, J.J. (2014). A brief history of synthetic biology. Nat. Rev. Microbiol. *12*, 381–390.

Capra, E.J., and Laub, M.T. (2012). Evolution of two-component signal transduction systems. Annu. Rev. Microbiol. *66*, 325–347.

Ceze, L., Nivala, J., and Strauss, K. (2019). Molecular digital data storage using DNA. Nat. Rev. Genet. *20*, 456–466.

Chaikind, B., Kilambi, K.P., Gray, J.J., and Ostermeier, M. (2012). Targeted DNA methylation using an artificially bisected M.HhaI fused to zinc fingers. PLoS ONE *7*, e44852.

Chen, Y.-J., Dalchau, N., Srinivas, N., Phillips, A., Cardelli, L., Soloveichik, D., and Seelig, G. (2013). Programmable chemical controllers made from DNA. Nat. Nanotechnol. *8*, 755–762.

Chen, Z., Boyken, S.E., Jia, M., Busch, F., Flores-Solis, D., Bick, M.J., Lu, P., VanAernum, Z.L., Sahasrabuddhe, A., Langan, R.A., et al. (2019). Programmable design of orthogonal protein heterodimers. Nature *565*, 106–111.

Chen, Z., Kibler, R.D., Hunt, A., Busch, F., Pearl, J., Jia, M., VanAernum, Z.L., Wicky, B.I.M., Dods, G., Liao, H., et al. (2020). De novo design of protein logic gates. Science *368*, 78–84.

Cherry, K.M., and Qian, L. (2018). Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. Nature *559*, 370–376.

Chevalier, A., Silva, D.-A., Rocklin, G.J., Hicks, D.R., Vergara, R., Murapa, P., Bernard, S.M., Zhang, L., Lam, K.-H., Yao, G., et al. (2017). Massively parallel de novo protein design for targeted therapeutics. Nature *550*, 74–79.

Cho, J.H., Collins, J.J., and Wong, W.W. (2018). Universal Chimeric Antigen Receptors for Multiplexed and Logical Control of T Cell Responses. Cell *173*, 1426–1438.e11.

Chung, H.K., and Lin, M.Z. (2020). On the cutting edge: protease-based methods for sensing and controlling cell biology. Nat. Methods *17*, 885–896.

Chung, H.K., Zou, X., Bajar, B.T., Brand, V.R., Huo, Y., Alcudia, J.F., Ferrell, J.E., Jr., and Lin, M.Z. (2019). A compact synthetic pathway rewires cancer signaling to therapeutic effector release. Science *364*, eaat6982.

Dai, Y., Abbasi, K., Bandyopadhyay, S., and Liu, C.C. (2019). Dynamic Control of Peptide Strand Displacement Reaction Using Functional Biomolecular Domain for Biosensing. ACS Sens. *4*, 1980–1985.

Daringer, N.M., Dudek, R.M., Schwarz, K.A., and Leonard, J.N. (2014). Modular extracellular sensor architecture for engineering mammalian cell-based devices. ACS Synth. Biol. *3*, 892–902.

Djannatian, M.S., Galinski, S., Fischer, T.M., and Rossner, M.J. (2011). Studying G protein-coupled receptor activation using split-tobacco etch virus assays. Anal. Biochem. *412*, 141–152.

Dolberg, T.B., Meger, A.T., Boucher, J.D., Corcoran, W.K., Schauer, E.E., Prybutok, A.N., Raman, S., and Leonard, J.N. (2021). Computation-guided optimization of split protein systems. Nat. Chem. Biol. https://doi.org/10.1038/s41589-020-00729-8.

Dueber, J.E., Mirsky, E.A., and Lim, W.A. (2007). Engineering synthetic signaling proteins with ultrasensitive input/output control. Nat. Biotechnol. *25*, 660–662.

Dueber, J.E., Wu, G.C., Malmirchegini, G.R., Moon, T.S., Petzold, C.J., Ullal, A.V., Prather, K.L.J., and Keasling, J.D. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. Nat. Biotechnol. *27*, 753–759.

Eldar, A., and Elowitz, M.B. (2010). Functional roles for noise in genetic circuits. Nature *467*, 167–173.

Elowitz, M.B., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. Nature *403*, 335–338.

Fedorov, V.D., Themeli, M., and Sadelain, M. (2013). PD-1- and CTLA-4-based inhibitory chimeric antigen receptors (iCARs) divert off-target immunotherapy responses. Sci. Transl. Med. *5*, 215ra172.

Fernandez-Rodriguez, J., and Voigt, C.A. (2016). Post-translational control of genetic circuits using Potyvirus proteases. Nucleic Acids Res. *44*, 6493–6502.

Ferrell, J.E., Jr., and Ha, S.H. (2014a). Ultrasensitivity part I: Michaelian responses and zero-order ultrasensitivity. Trends Biochem. Sci. *39*, 496–503.

Ferrell, J.E., Jr., and Ha, S.H. (2014b). Ultrasensitivity part II: multisite phosphorylation, stoichiometric inhibitors, and positive feedback. Trends Biochem. Sci. *39*, 556–569.

Ferrell, J.E., Jr., and Ha, S.H. (2014c). Ultrasensitivity part III: cascades, bistable switches, and oscillators. Trends Biochem. Sci. *39*, 612–618.

Fink, T., Lonzarić, J., Praznik, A., Plaper, T., Merljak, E., Leben, K., Jerala, N., Lebar, T., Strmšek, Ž., Lapenta, F., et al. (2019). Design of fast proteolysis-based signaling and logic circuits in mammalian cells. Nat. Chem. Biol. *15*, 115–122.

Fisher, R.S., and Elbaum-Garfinkle, S. (2020). Tunable multiphase dynamics of arginine and lysine liquid condensates. Nat. Commun. *11*, 4628.

Fletcher, J.M., Boyle, A.L., Bruning, M., Bartlett, G.J., Vincent, T.L., Zaccai, N.R., Armstrong, C.T., Bromley, E.H.C., Booth, P.J., Brady, R.L., et al. (2012). A basis set of de novo coiled-coil peptide oligomers for rational protein design and synthetic biology. ACS Synth. Biol. *1*, 240–250.

Foight, G.W., Wang, Z., Wei, C.T., Jr Greisen, P., Warner, K.M., Cunningham-Bryant, D., Park, K., Brunette, T.J., Sheffler, W., Baker, D., and Maly, D.J. (2019). Multi-input chemical control of protein dimerization for programming graded cellular responses. Nat. Biotechnol. *37*, 1209–1216.

Gao, X.J., Chong, L.S., Kim, M.S., and Elowitz, M.B. (2018). Programmable protein circuits in living cells. Science *361*, 1252–1258.

Gao, C., Hou, J., Xu, P., Guo, L., Chen, X., Hu, G., Ye, C., Edwards, H., Chen, J., Chen, W., and Liu, L. (2019). Programmable biomolecular switches for rewiring flux in Escherichia coli. Nat. Commun. *10*, 3751.

Goentoro, L., Shoval, O., Kirschner, M.W., and Alon, U. (2009). The incoherent feedforward loop can provide fold-change detection in gene regulation. Mol. Cell *36*, 894–899.

Gonzalez, L., Jr., Woolfson, D.N., and Alber, T. (1996). Buried polar residues and structural specificity in the GCN4 leucine zipper. Nat. Struct. Biol. *3*, 1011–1018.

Good, M.C., Zalatan, J.G., and Lim, W.A. (2011). Scaffold proteins: hubs for controlling the flow of cellular information. Science *332*, 680–686.

Gordley, R.M., Williams, R.E., Bashor, C.J., Toettcher, J.E., Yan, S., and Lim, W.A. (2016). Engineering dynamical control of cell fate switching using synthetic phospho-regulons. Proc. Natl. Acad. Sci. USA *113*, 13528–13533.

Gradišar, H., and Jerala, R. (2011). De novo design of orthogonal peptide pairs forming parallel coiled-coil heterodimers. J. Pept. Sci. *17*, 100–106.

Green, A.A., Silver, P.A., Collins, J.J., and Yin, P. (2014). Toehold switches: de-novo-designed regulators of gene expression. Cell *159*, 925–939.

Gröger, K., Gavins, G., and Seitz, O. (2017). Strand Displacement in Coiled-Coil Structures: Controlled Induction and Reversal of Proximity. Angew. Chem. Int. Ed. Engl. *56*, 14217–14221.

Groth, M.C., Rink, W.M., Meyer, N.F., and Thomas, F. (2018). Kinetic studies on strand displacement in *de novo* designed parallel heterodimeric coiled coils. Chem. Sci. (Camb.) *9*, 4308–4316.

Groves, B., Khakhar, A., Nadel, C.M., Gardner, R.G., and Seelig, G. (2016). Rewiring MAP kinases in Saccharomyces cerevisiae to regulate novel targets through ubiquitination. eLife *5*, e15200.

Hansen, J., Mailand, E., Swaminathan, K.K., Schreiber, J., Angelici, B., and Benenson, Y. (2014). Transplantation of prokaryotic two-component signaling pathways into mammalian cells. Proc. Natl. Acad. Sci. USA *111*, 15705–15710.

Hart, Y., Reich-Zeliger, S., Antebi, Y.E., Zaretsky, I., Mayo, A.E., Alon, U., and Friedman, N. (2014). Paradoxical signaling by a secreted molecule leads to homeostasis of cell levels. Cell *158*, 1022–1032.

Hartfield, R.M., Schwarz, K.A., Muldoon, J.J., Bagheri, N., and Leonard, J.N. (2017). Multiplexing Engineered Receptors for Multiparametric Evaluation of Environmental Ligands. ACS Synth. Biol. *6*, 2042–2055.

Havranek, J.J., and Harbury, P.B. (2003). Automated design of specificity in molecular recognition. Nat. Struct. Biol. *10*, 45–52.

Hegde, M., Mukherjee, M., Grada, Z., Pignata, A., Landi, D., Navai, S.A., Wakefield, A., Fousek, K., Bielamowicz, K., Chow, K.K.H., et al. (2016). Tandem CAR T cells targeting HER2 and IL13Rα2 mitigate tumor antigen escape. J. Clin. Invest. *126*, 3036–3052.

Hoffmann, A., Levchenko, A., Scott, M.L., and Baltimore, D. (2002). The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. Science *298*, 1241–1245.

Holt, B.A., and Kwong, G.A. (2020). Protease circuits for processing biological information. Nat. Commun. *11*, 5021.

Hong, M., Clubb, J.D., and Chen, Y.Y. (2020). Engineering CAR-T Cells for Next-Generation Cancer Therapy. Cancer Cell *38*, 473–488.

Hopfield, J.J. (1974). Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. Proc. Natl. Acad. Sci. USA *71*, 4135–4139.

Hsiao, V., de los Santos, E.L.C., Whitaker, W.R., Dueber, J.E., and Murray, R.M. (2015). Design and implementation of a biomolecular concentration tracker. ACS Synth. Biol. *4*, 150–161.

Huang, P.-S., Oberdorfer, G., Xu, C., Pei, X.Y., Nannenga, B.L., Rogers, J.M., DiMaio, F., Gonen, T., Luisi, B., and Baker, D. (2014). High thermodynamic stability of parametrically designed helical bundles. Science *346*, 481–485.

Huang, T.-H., Velho, T., and Lois, C. (2016). Monitoring cell-cell contacts in vivo in transgenic animals. Development *143*, 4073–4084.

Huang, T.-H., Niesman, P., Arasu, D., Lee, D., De La Cruz, A.L., Callejas, A., Hong, E.J., and Lois, C. (2017). Tracing neuronal circuits in transgenic animals by transneuronal control of transcription (*TRACT*). eLife *6*, e32027.

Jusiak, B., Cleto, S., Perez-Piñera, P., and Lu, T.K. (2016). Engineering Synthetic Gene Circuits in Living Cells with CRISPR Technology. Trends Biotechnol. *34*, 535–547.

Kapp, G.T., Liu, S., Stein, A., Wong, D.T., Reményi, A., Yeh, B.J., Fraser, J.S., Taunton, J., Lim, W.A., and Kortemme, T. (2012). Control of protein signaling using a computationally designed GTPase/GEF orthogonal pair. Proc. Natl. Acad. Sci. USA *109*, 5277–5282.

Karin, O., and Alon, U. (2017). Biphasic response as a mechanism against mutant takeover in tissue homeostasis circuits. Mol. Syst. Biol. *13*, 933.

Kim, J., White, K.S., and Winfree, E. (2006). Construction of an in vitro bistable circuit from synthetic transcriptional switches. Mol. Syst. Biol. *2*, 68.

Kipniss, N.H., Dingal, P.C.D.P., Abbott, T.R., Gao, Y., Wang, H., Dominguez, A.A., Labanieh, L., and Qi, L.S. (2017). Engineering cell sensing and responses using a GPCR-coupled CRISPR-Cas system. Nat. Commun. *8*, 2212.

Kirkpatrick, R.L., Lewis, K., Langan, R.A., Lajoie, M.J., Boyken, S.E., Eakman, M., Baker, D., and Zalatan, J.G. (2020). Conditional Recruitment to a

DNA-Bound CRISPR-Cas Complex Using a Colocalization-Dependent Protein Switch. ACS Synth. Biol. *9*, 2316–2323.

Klosin, A., Oltsch, F., Harmon, T., Honigmann, A., Jülicher, F., Hyman, A.A., and Zechner, C. (2020). Phase separation provides a mechanism to reduce noise in cells. Science *367*, 464–468.

Krawczyk, K., Scheller, L., Kim, H., and Fussenegger, M. (2020). Rewiring of endogenous signaling pathways to genomic targets for therapeutic cell reprogramming. Nat. Commun. *11*, 608.

Lajoie, M.J., Boyken, S.E., Salter, A.I., Bruffey, J., Rajan, A., Langan, R.A., Olshefsky, A., Muhunthan, V., Bick, M.J., Gewe, M., et al. (2020). Designed protein logic to target cells with precise combinations of surface antigens. Science *369*, 1637–1643.

Lakin, N.D., and Jackson, S.P. (1999). Regulation of p53 in response to DNA damage. Oncogene *18*, 7644–7655.

Langan, R.A., Boyken, S.E., Ng, A.H., Samson, J.A., Dods, G., Westbrook, A.M., Nguyen, T.H., Lajoie, M.J., Chen, Z., Berger, S., et al. (2019). De novo design of bioactive protein switches. Nature *572*, 205–210.

Lebar, T., Verbič, A., Ljubetič, A., and Jerala, R. (2018). Polarized displacement by transcription activator-like effectors for regulatory circuits. Nat. Chem. Biol. *15*, 80–87.

Levchenko, A., Bruck, J., and Sternberg, P.W. (2000). Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties. Proc. Natl. Acad. Sci. USA *97*, 5818–5823.

Li, S., Jiang, Q., Liu, S., Zhang, Y., Tian, Y., Song, C., Wang, J., Zou, Y., Anderson, G.J., Han, J.-Y., et al. (2018). A DNA nanorobot functions as a cancer therapeutic in response to a molecular trigger in vivo. Nat. Biotechnol. *36*, 258–264.

Lumb, K.J., and Kim, P.S. (1995). A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. Biochemistry *34*, 8642–8648.

Ma, W., Trusina, A., El-Samad, H., Lim, W.A., and Tang, C. (2009). Defining network topologies that can achieve biochemical adaptation. Cell *138*, 760–773.

Ma, Y., Budde, M.W., Mayalu, M.N., Zhu, J., Murray, R.M., and Elowitz, M.B. (2020). Synthetic mammalian signaling circuits for robust cell population control. bioRxiv. https://doi.org/10.1101/2020.09.02.278564.

Mazé, A., and Benenson, Y. (2019). Artificial signaling in mammalian cells enabled by prokaryotic two-component system. Nat. Chem. Biol. *16*, 179–187.

McClune, C.J., and Laub, M.T. (2020). Constraints on the expansion of paralogous protein families. Curr. Biol. *30*, R460–R464.

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

Mody, A., Weiner, J., and Ramanathan, S. (2009). Modularity of MAP kinases allows deformation of their signalling pathways. Nat. Cell Biol. *11*, 484–491.

Monod, J., and Jacob, F. (1961). Teleonomic mechanisms in cellular metabolism, growth, and differentiation. Cold Spring Harb. Symp. Quant. Biol. *26*, 389–401.

Moraga, I., Spangler, J.B., Mendoza, J.L., Gakovic, M., Wehrman, T.S., Krutzik, P., and Garcia, K.C. (2017). Synthekines are surrogate cytokine and growth factor agonists that compel signaling through non-natural receptor dimers. eLife *6*, e22882.

Morsut, L., Roybal, K.T., Xiong, X., Gordley, R.M., Coyle, S.M., Thomson, M., and Lim, W.A. (2016). Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors. Cell *164*, 780–791.

Moser, F., Espah Borujeni, A., Ghodasara, A.N., Cameron, E., Park, Y., and Voigt, C.A. (2018). Dynamic control of endogenous metabolism with combinatorial logic circuits. Mol. Syst. Biol. *14*, e8605.

Nakajima, M., Imai, K., Ito, H., Nishiwaki, T., Murayama, Y., Iwasaki, H., Oyama, T., and Kondo, T. (2005). Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. Science *308*, 414–415.

Netzer, R., Listov, D., Lipsh, R., Dym, O., Albeck, S., Knop, O., Kleanthous, C., and Fleishman, S.J. (2018). Ultrahigh specificity in a network of computationally designed protein-interaction pairs. Nat. Commun. *9*, 5286.

Newby, G.A., Kiriakov, S., Hallacli, E., Kayatekin, C., Tsvetkov, P., Mancuso, C.P., Bonner, J.M., Hesse, W.R., Chakrabortee, S., Manogaran, A.L., et al. (2017). A Genetic Tool to Track Protein Aggregates and Control Prion Inheritance. Cell *171*, 966–979.e18.

Ng, A.H., Nguyen, T.H., Gómez-Schiavon, M., Dods, G., Langan, R.A., Boyken, S.E., Samson, J.A., Waldburger, L.M., Dueber, J.E., Baker, D., and El-Samad, H. (2019). Modular and tunable biological feedback control using a de novo protein switch. Nature *572*, 265–269.

O'Shaughnessy, E.C., Palani, S., Collins, J.J., and Sarkar, C.A. (2011). Tunable signal processing in synthetic MAP kinase cascades. Cell *144*, 119–131.

Park, S.-H., Zarrinpar, A., and Lim, W.A. (2003). Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. Science *299*, 1061–1064.

Park, M., Patel, N., Keung, A.J., and Khalil, A.S. (2019). Engineering Epigenetic Regulation Using Synthetic Read-Write Modules. Cell *176*, 227–238.e20.

Qian, L., Winfree, E., and Bruck, J. (2011). Neural network computation with DNA strand displacement cascades. Nature *475*, 368–372.

Rakhit, R., Navarro, R., and Wandless, T.J. (2014). Chemical biology strategies for posttranslational control of protein function. Chem. Biol. *21*, 1238–1252.

Reinke, A.W., Grant, R.A., and Keating, A.E. (2010). A synthetic coiled-coil interactome provides heterospecific modules for molecular engineering. J. Am. Chem. Soc. *132*, 6025–6031.

Reinkemeier, C.D., Girona, G.E., and Lemke, E.A. (2019). Designer membraneless organelles enable codon reassignment of selected mRNAs in eukaryotes. Science *363*, eaaw2644.

Riesselman, A., Shin, J.-E., Kollasch, A., McMahon, C., Simon, E., Sander, C., Manglik, A., Kruse, A., and Marks, D. (2019). Accelerating Protein Design Using Autoregressive Generative Models. bioRxiv. https://doi.org/10.1101/757252.

Rössler, O.E. (1974a). Chemical Automata in Homogeneous and Reaction-Diffusion Kinetics. In Physics and Mathematics of the Nervous System, M. Conrad, W. Guttinger, and M. Dal Cin, eds. (Springer), pp. 399–418.

Rössler, O.E. (1974b). A Synthetic Approach to Exotic Kinetics (With Examples). In Physics and Mathematics of the Nervous System, M. Conrad, W. Guttinger, and M. Dal Cin, eds. (Springer), pp. 546–582.

Roybal, K.T., Williams, J.Z., Morsut, L., Rupp, L.J., Kolinko, I., Choe, J.H., Walker, W.J., McNally, K.A., and Lim, W.A. (2016a). Engineering T Cells with Customized Therapeutic Response Programs Using Synthetic Notch Receptors. Cell *167*, 419–432.e16.

Roybal, K.T., Rupp, L.J., Morsut, L., Walker, W.J., McNally, K.A., Park, J.S., and Lim, W.A. (2016b). Precision Tumor Recognition by T Cells With Combinatorial Antigen-Sensing Circuits. Cell *164*, 770–779.

Ruella, M., Barrett, D.M., Kenderian, S.S., Shestova, O., Hofmann, T.J., Perazzelli, J., Klichinsky, M., Aikawa, V., Nazimuddin, F., Kozlowski, M., et al. (2016). Dual CD19 and CD123 targeting prevents antigen-loss relapses after CD19-directed immunotherapies. J. Clin. Invest. *126*, 3814–3826.

Russ, W.P., Figliuzzi, M., Stocker, C., Barrat-Charlaix, P., Socolich, M., Kast, P., Hilvert, D., Monasson, R., Cocco, S., Weigt, M., and Ranganathan, R. (2020). An evolution-based model for designing chorismate mutase enzymes. Science *369*, 440–445.

Rust, M.J., Markson, J.S., Lane, W.S., Fisher, D.S., and O'Shea, E.K. (2007). Ordered phosphorylation governs oscillation of a three-protein circadian clock. Science *318*, 809–812.

Ryu, J., and Park, S.-H. (2015). Simple synthetic protein scaffolds can create adjustable artificial MAPK circuits in yeast and mammalian cells. Sci. Signal. *8*, ra66.

Sahin, U., Karikó, K., and Türeci, Ö. (2014). mRNA-based therapeutics–developing a new class of drugs. Nat. Rev. Drug Discov. *13*, 759–780.

Scalise, D., and Schulman, R. (2019). Controlling Matter at the Molecular Scale with DNA Circuits. Annu. Rev. Biomed. Eng. *21*, 469–493.

Scheller, L., Strittmatter, T., Fuchs, D., Bojar, D., and Fussenegger, M. (2018). Generalized extracellular molecule sensor platform for programming cellular behavior. Nat. Chem. Biol. *14*, 723–729.

Scheller, L., Schmollack, M., Bertschi, A., Mansouri, M., Saxena, P., and Fussenegger, M. (2020). Phosphoregulated orthogonal signal transduction in mammalian cells. Nat. Commun. *11*, 3085.

Schmidl, S.R., Ekness, F., Sofjan, K., Daeffler, K.N.-M., Brink, K.R., Landry, B.P., Gerhardt, K.P., Dyulgyarov, N., Sheth, R.U., and Tabor, J.J. (2019). Rewiring bacterial two-component systems by modular DNA-binding domain swapping. Nat. Chem. Biol. *15*, 690–698.

Schuster, B.S., Reed, E.H., Parthasarathy, R., Jahnke, C.N., Caldwell, R.M., Bermudez, J.G., Ramage, H., Good, M.C., and Hammer, D.A. (2018). Controllable protein phase separation and modular recruitment to form responsive membraneless organelles. Nat. Commun. *9*, 2985.

Schwarz, K.A., Daringer, N.M., Dolberg, T.B., and Leonard, J.N. (2017). Rewiring human cellular input-output using modular extracellular sensors. Nat. Chem. Biol. *13*, 202–209.

Seeman, N.C. (1982). Nucleic acid junctions and lattices. J. Theor. Biol. *99*, 237–247.

Shui, S., Gainza, P., Scheller, L., Yang, C., Kurumida, Y., Rosset, S., Georgeon, S., Correia, B.E., et al. (2021). A rational blueprint for the design of chemically-controlled protein switches. bioRxiv. https://doi.org/10.1101/2021.01.21.427547.

Siciliano, V., DiAndreth, B., Monel, B., Beal, J., Huh, J., Clayton, K.L., Wroblewska, L., McKeon, A., Walker, B.D., and Weiss, R. (2018). Engineering modular intracellular protein sensor-actuator devices. Nat. Commun. *9*, 1881.

Skerker, J.M., Perchuk, B.S., Siryaporn, A., Lubin, E.A., Ashenberg, O., Goulian, M., and Laub, M.T. (2008). Rewiring the specificity of two-component signal transduction systems. Cell *133*, 1043–1054.

Slomovic, S., and Collins, J.J. (2015). DNA sense-and-respond protein modules for mammalian cells. Nat. Methods *12*, 1085–1090.

Sowa, S.W., Gelderman, G., and Contreras, L.M. (2015). Advances in synthetic dynamic circuits design: using novel synthetic parts to engineer new generations of gene oscillations. Curr. Opin. Biotechnol. *36*, 161–167.

Srinivas, N., Parkin, J., Seelig, G., Winfree, E., and Soloveichik, D. (2017). Enzyme-free nucleic acid dynamical systems. Science *358*, eaal2052.

Stapornwongkul, K.S., de Gennes, M., Cocconi, L., Salbreux, G., and Vincent, J.-P. (2020). Patterning and growth control in vivo by an engineered GFP gradient. Science *370*, 321–327.

Stein, V., and Alexandrov, K. (2014). Protease-based synthetic sensing and signal amplification. Proc. Natl. Acad. Sci. USA *111*, 15934–15939.

Stein, V., Nabi, M., and Alexandrov, K. (2017). Ultrasensitive Scaffold-Dependent Protease Sensors with Large Dynamic Range. ACS Synth. Biol. *6*, 1337–1342.

Stumpp, M.T., Binz, H.K., and Amstutz, P. (2008). DARPins: a new generation of protein therapeutics. Drug Discov. Today *13*, 695–701.

Su, C.J., Murugan, A., Linton, J.M., Yeluri, A., Bois, J., Klumpe, H., Antebi, Y.E., and Elowitz, M.B. (2020). Ligand-receptor promiscuity enables cellular addressing. bioRxiv. https://doi.org/10.1101/2020.12.08.412643.

Sugita, M. (1963). Functional analysis of chemical systems in vivo using a logical circuit equivalent. II. The idea of a molecular automation. J. Theor. Biol. *4*, 179–192.

Thubagere, A.J., Li, W., Johnson, R.F., Chen, Z., Doroudi, S., Lee, Y.L., Izatt, G., Wittman, S., Srinivas, N., Woods, D., et al. (2017). A cargo-sorting DNA robot. Science *357*, eaan6558.

Toda, S., Blauch, L.R., Tang, S.K.Y., Morsut, L., and Lim, W.A. (2018). Programming self-organizing multicellular structures with synthetic cell-cell signaling. Science *361*, 156–162.

Toda, S., McKeithan, W.L., Hakkinen, T.J., Lopez, P., Klein, O.D., and Lim, W.A. (2020). Engineering synthetic morphogen systems that can program multicellular patterning. Science *370*, 327–331.

van Dongen, J.E., Berendsen, J.T.W., Steenbergen, R.D.M., Wolthuis, R.M.F., Eijkel, J.C.T., and Segerink, L.I. (2020). Point-of-care CRISPR/Cas nucleic acid detection: Recent advances, challenges and opportunities. Biosens. Bioelectron. *166*, 112445.

Wagenbauer, K.F., Sigl, C., and Dietz, H. (2017). Gigadalton-scale shape-programmable DNA assemblies. Nature *552*, 78–83.

Wehr, M.C., Reinecke, L., Botvinnik, A., and Rossner, M.J. (2008). Analysis of transient phosphorylation-dependent protein-protein interactions in living mammalian cells using split-TEV. BMC Biotechnol. *8*, 55.

Wei, B., Dai, M., and Yin, P. (2012a). Complex shapes self-assembled from single-stranded DNA tiles. Nature *485*, 623–626.

Wei, P., Wong, W.W., Park, J.S., Corcoran, E.E., Peisajovich, S.G., Onuffer, J.J., Weiss, A., and Lim, W.A. (2012b). Bacterial virulence proteins as tools to rewire kinase pathways in yeast and immune cells. Nature *488*, 384–388.

Wei, K.Y., Moschidi, D., Bick, M.J., Nerli, S., McShan, A.C., Carter, L.P., Huang, P.-S., Fletcher, D.A., Sgourakis, N.G., Boyken, S.E., et al. (2020). Computational design of closely related proteins that adopt two well-defined but structurally divergent folds. Proc. Natl. Acad. Sci. USA *117*, 7208–7215.

Wellner, A., McMahon, C., Gilman, M.S.A., Clements, J.R., Clark, S., Nguyen, K.M., Ho, M.H., Shin, J.-E., Feldman, J., Hauser, B.M., et al. (2020). Rapid generation of potent antibodies by autonomous hypermutation in yeast. bioRxiv. https://doi.org/10.1101/2020.11.11.378778.

Whitaker, W.R., Davis, S.A., Arkin, A.P., and Dueber, J.E. (2012). Engineering robust control of two-component system phosphotransfer using modular scaffolds. Proc. Natl. Acad. Sci. USA *109*, 18090–18095.

Williams, J.Z., Allen, G.M., Shah, D., Sterin, I.S., Kim, K.H., Garcia, V.P., Shavey, G.E., Yu, W., Puig-Saus, C., Tsoi, J., et al. (2020). Precise T cell recognition programs designed by transcriptionally linking multiple receptors. Science *370*, 1099–1104.

Winfree, E. (1996). On the computational power of DNA annealing and ligation. DIMACS Series in Discrete Mathematics and Theoretical Computer Science (American Mathematical Society), pp. 199–221.

Woods, D., Doty, D., Myhrvold, C., Hui, J., Zhou, F., Yin, P., and Winfree, E. (2019). Diverse and robust molecular algorithms using reprogrammable DNA self-assembly. Nature *567*, 366–372.

Xie, M., and Fussenegger, M. (2018). Designing cell function: assembly of synthetic gene circuits for cell biology applications. Nat. Rev. Mol. Cell Biol. *19*, 507–525.

Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R., and Benenson, Y. (2011). Multi-input RNAi-based logic circuit for identification of specific cancer cells. Science *333*, 1307–1311.

You, L., Cox, R.S., 3rd, Weiss, R., and Arnold, F.H. (2004). Programmed population control by cell-cell communication and regulated killing. Nature *428*, 868–871.

Zah, E., Lin, M.-Y., Silva-Benedict, A., Jensen, M.C., and Chen, Y.Y. (2016). T Cells Expressing CD19/CD20 Bispecific Chimeric Antigen Receptors Prevent Antigen Escape by Malignant B Cells. Cancer Immunol. Res. *4*, 498–508.

Zah, E., Nam, E., Bhuvan, V., Tran, U., Ji, B.Y., Gosliner, S.B., Wang, X., Brown, C.E., and Chen, Y.Y. (2020). Systematically optimized BCMA/CS1 bispecific CAR-T cells robustly control heterogeneous multiple myeloma. Nat. Commun. *11*, 2283.

Zarrinpar, A., Park, S.-H., and Lim, W.A. (2003). Optimization of specificity in a cellular protein interaction network by negative selection. Nature *426*, 676–680.

Zhang, D.Y., and Seelig, G. (2011). Dynamic DNA nanotechnology using strand-displacement reactions. Nat. Chem. *3*, 103–113.

Zhang, Z.-B., Wang, Q.-Y., Ke, Y.-X., Liu, S.-Y., Ju, J.-Q., Lim, W.A., Tang, C., and Wei, P. (2017). Design of Tunable Oscillatory Dynamics in a Synthetic NF-κB Signaling Circuit. Cell Syst. *5*, 460–470.e5.

Zhu, R., del Rio-Salgado, J.M., Garcia-Ojalvo, J., and Elowitz, M.B. (2021). Synthetic multistability in mammalian cells. bioRxiv. https://doi.org/10.1101/2021.02.10.430659.