PROTEIN DESIGN

De novo design of protein logic gates

Zibo Chen[1,2]*, Ryan D. Kibler[1,2]†, Andrew Hunt[3]†, Florian Busch[4,5]†, Jocelynn Pearl[6]†‡, Mengxuan Jia[4,5], Zachary L. VanAernum[4,5], Basile I. M. Wicky[1,2], Galen Dods[7], Hanna Liao[6]‡, Matthew S. Wilken[6], Christie Ciarlo[6], Shon Green[6], Hana El-Samad[7,8], John Stamatoyannopoulos[6,9,10], Vicki H. Wysocki[4,5], Michael C. Jewett[3,11,12], Scott E. Boyken[1,2]‡, David Baker[1,2,13]§

The design of modular protein logic for regulating protein function at the posttranscriptional level is a challenge for synthetic biology. Here, we describe the design of two-input AND, OR, NAND, NOR, XNOR, and NOT gates built from de novo–designed proteins. These gates regulate the association of arbitrary protein units ranging from split enzymes to transcriptional machinery in vitro, in yeast and in primary human T cells, where they control the expression of the *TIM3* gene related to T cell exhaustion. Designed binding interaction cooperativity, confirmed by native mass spectrometry, makes the gates largely insensitive to stoichiometric imbalances in the inputs, and the modularity of the approach enables ready extension to three-input OR, AND, and disjunctive normal form gates. The modularity and cooperativity of the control elements, coupled with the ability to de novo design an essentially unlimited number of protein components, should enable the design of sophisticated posttranslational control logic over a wide range of biological functions.

Protein-protein interactions are ubiquitous in cellular decision-making and controlling them will be increasingly important in synthetic biology (*1*–*4*). Although protein interactions are central to natural biological circuits, efforts to create new logic circuits have focused on control at the level of DNA (*5*, *6*), transcription (*7*–*18*), or RNA (*13*, *19*–*22*). Recently, protein-based circuits have been generated by rewiring native signaling pathways (*23*–*28*), bringing proteins together with coiled coils (*29*), or creating protease cascades (*30*, *31*); however, these circuits were constructed from a limited pool of building blocks, which hinders their scalability. The ability to de novo design protein-based logic gates modulating arbitrary protein-protein interactions could open the door to new protein-based control systems in and out of cells.

In principle, it should be possible to design a wide range of logic gates de novo using a set of heterodimeric molecules. For example, given hypothetical heterodimer pairs ***A*:*A′***, ***B*:*B′***, and ***C*:*C′***, an AND gate modulating the association of ***A*** with ***C′*** can be constructed by genetically fusing ***A′*** and ***B***, and ***B′*** and ***C***: association occurs only in the presence of both ***A′*-*B***, and ***B′*-*C*** (here and below "**:**" denotes noncovalent interaction and "-" a genetic fusion through flexible linkers). Several building block properties are desirable for constructing such associative logic gates. First, there should be many mutually orthogonal heterodimeric pairs so that gate complexity is not limited by the number of individual elements. Second, the building blocks should be modular and similar in structure so that differences in building block shape and other properties do not have to be considered when constructing the gates. Third, single building blocks should be able to bind to multiple partners with different and tunable affinities, allowing inputs to perform negation operations by disrupting preexisting lower-affinity interactions. Fourth, the interactions should be cooperative so that gate activation is not sensitive to stoichiometric imbalances in the inputs. In the above AND gate, for example, if the interactions are not cooperative, then a large excess of ***A′*-*B*** will pull the equilibrium toward partially assembled complexes (***A′*-*B*** with either ***A*** or ***B′*-*C*** but not both), which will limit gate activation.

Here, we explored the possibility of designing logic gates satisfying all four of the above criteria using de novo–designed protein heterodimers with hydrogen bond network–mediated specificity (*32*). Sets of mutually orthogonal designed heterodimers (DHDs, hereafter referred to by numbers, e.g., ***1*** and ***1′*** form one cognate pair; table S1) with hydrogen bond network–mediated specificity (e.g., see Fig. 1A, inset) are available for logic gate construction, satisfying condition 1 (orthogonality). The heterodimeric interfaces all share the same four helix bundle topology (Fig. 1A), satisfying condition 2 (modularity). The shared interaction interface allows a limited amount of cross-talk between pairs, leading to a hierarchy of binding affinities, satisfying condition 3 (multiple binding specificities). Inspired by cooperative systems in nature (*33*, *34*), we sought to achieve condition 4 (cooperativity) by constructing the monomer fusions (***A′*-*B*** and ***B′*-*C*** in the above example) in such a way that the interaction surfaces (with ***A*** and ***C′***) are buried within the fusions. The free energy required to expose these buried interfaces would oppose gate activation, and we reasoned that the system could be tuned so that the sum of the binding energies of the two partners, but not either one alone, would be sufficient to overcome this barrier, ensuring cooperative gate activation. If condition 2 (modularity) holds, then a single scheme for ensuring cooperativity could in principle work for a wide range of gate configurations.

To explore the design of cooperative building blocks, we focused on the simple system ***A*** + ***A′*-*B*** + ***B′*** (we refer to this as induced dimerization below, ***A*** and ***B′*** as the monomers, and ***A′*-*B*** as the dimerizer). If binding is not cooperative, then the amount of the trimeric complex decreases when ***A′*-*B*** is in stoichiometric excess relative to ***A*** and ***B′***: the formation of intermediate dimeric species of the dimerizer binding to either of the monomers competes with formation of trimeric complexes. On the contrary, if binding is cooperative such that no binding to either monomer occurs in the absence of the other, then the amount of trimeric complex formed becomes insensitive to an excess of the dimerizer. A simple thermodynamic model of the effect of binding cooperativity on the stoichiometry dependence of such induced dimerization systems (Fig. 1B and see modeling section in the supplementary materials) shows that, as the binding cooperativity decreases, there is a corresponding decrease in the population of full trimeric complexes at high dimerizer concentrations (Fig. 1C).

We hypothesized that a folded four-helix bundle–like state of the ***A′*-*B*** dimerizer could oppose binding to either ***A*** or ***B′*** because the relatively hydrophobic interacting surfaces would likely be sequestered within the folded structure (fig. S1A). We tested different flexible linker lengths connecting ***A′*** with ***B*** using heterodimers ***1*:*1′*** and ***2*:*2′*** as a model system.

[1]Department of Biochemistry, University of Washington, Seattle, WA 98195, USA. [2]Institute for Protein Design, University of Washington, Seattle, WA 98195, USA. [3]Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL 60208, USA. [4]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH 43210, USA. [5]Resource for Native Mass Spectrometry Guided Structural Biology, The Ohio State University, Columbus, OH 43210, USA. [6]Altius Institute for Biomedical Sciences, Seattle, WA 98195, USA. [7]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA. [8]Chan-Zuckerberg Biohub, San Francisco, CA 94158, USA. [9]Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA. [10]Department of Medicine, Division of Oncology, University of Washington, Seattle, WA 98109, USA. [11]Chemistry of Life Processes Institute, Northwestern University, Evanston, IL 60208, USA. [12]Center for Synthetic Biology, Northwestern University, Evanston, IL 60208, USA. [13]Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA.
*Present address: Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA. †These authors contributed equally to this work. ‡Present address: Lyell Immunopharma, Inc., Seattle, WA 98109, USA.
§Corresponding author. Email: dabaker@uw.edu

Downloaded from http://science.sciencemag.org/ on April 4, 2020

At all linker lengths tested (between 0 and 24 residues), constructs were folded and stable in circular dichroism (CD) guanidine hydrochloride denaturation experiments, with unfolding free energies >13 kcal/mol (Fig. 1D, fig. S2, and table S3). Although ***1′-2′*** dimerizer constructs with short linkers of 0 and 2 residues, or with a very long 24-residue linker, could be purified as monomers (fig. S1B), they were prone to aggregation, perhaps due to domain swapping. By contrast, designs with 6 and 12 residue linkers remained largely monomeric (table S4). Small-angle x-ray scattering (SAXS) experiments (*35*) indicated that their hydrodynamic radii are close to those of folded four-helix bundle DHDs (Fig. 1E and table S2). Linkers in this length range likely allow the two monomers (***1′*** and ***2′***) to fold back on each

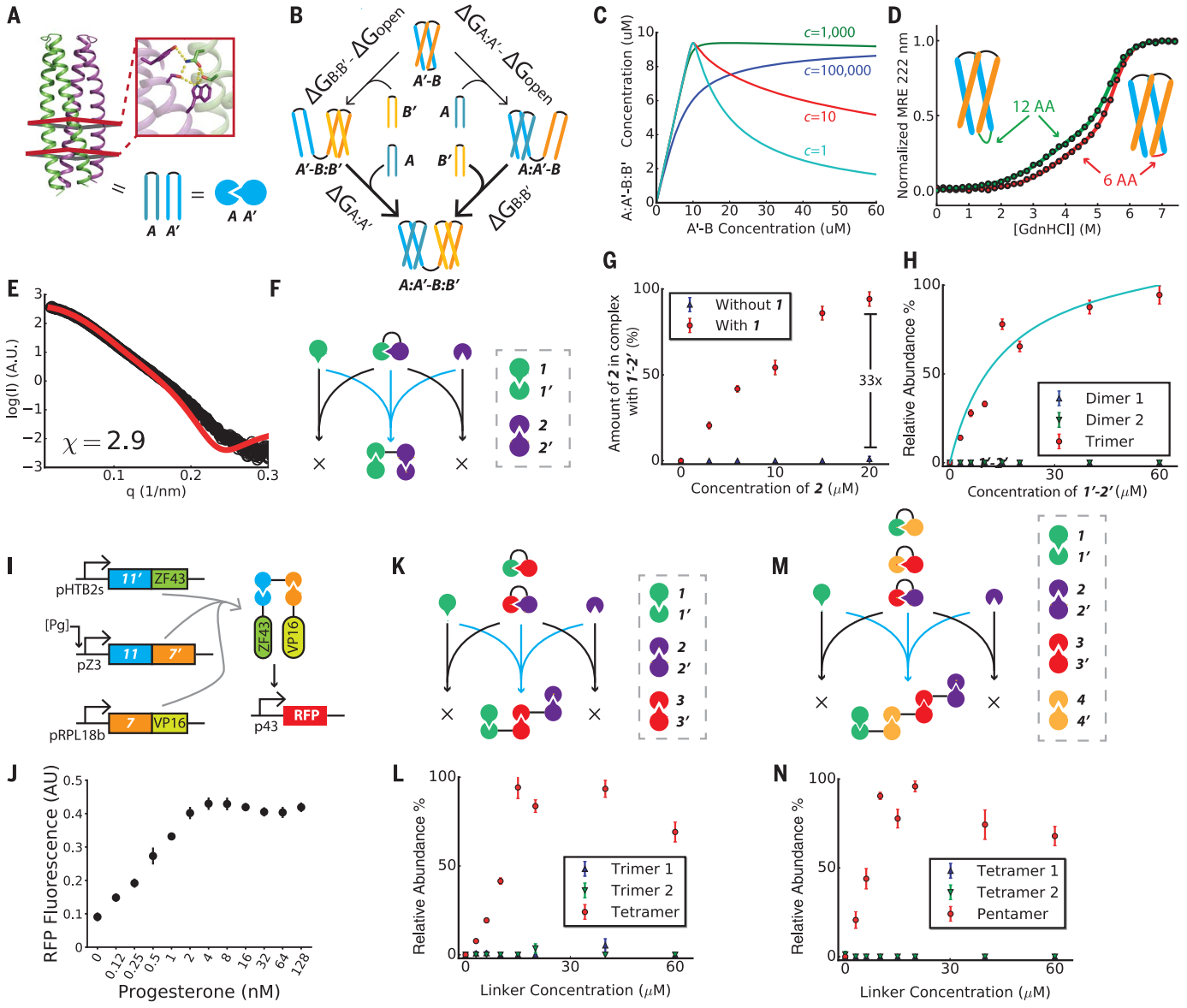

Fig. 1. Cooperativity of CIPHR logic gates. (**A**) Left: Backbone structure of ***A:A′*** heterodimer building block, with its hydrogen bond network shown in the inset. Bottom: Shorthand representations used throughout figures. (**B**) Thermodynamic cycle describing the induced dimerization system. (**C**) Simulation of the induced dimerization system under thermodynamic equilibrium. ***A*** and ***B′*** monomers were held constant at 10 μM each while titrating in various initial amounts of the ***A′-B*** dimerizer proteins. If binding is not cooperative (small *c*), then the final amount of trimeric complexes decreases when the dimerizer protein is in excess. (**D**) Equilibrium denaturation experiments monitored by CD for designs with 6- and 12-amino acid (AA) linkers. Circles represent experimental data, and lines are fits to the three-state unimolecular unfolding model. (**E**) Experimental SAXS profile of ***1′-2′*** with a six-residue linker (in black) fitted to the calculated profile of ***1:1′*** heterodimer (in red). (**F**) Schematic of induced dimerization system (with a six-residue linker); experimental results in (G) and (H). (**G**) nMS titration of ***2*** against 10 μM of ***1′-2′*** in the presence (red) or absence (blue) of 10 μM of ***1***. (**H**) nMS titration of ***1′-2′*** against 10 μM each of ***1*** and ***2***. Dimer 1 and 2 refer to partial dimeric complexes consisting of the dimerizer binding to either of the monomers. For comparison, the thermodynamic model result with *c* = 991,000 is shown in cyan. (**I**) Schematic of testing of the induced dimerization system in yeast, with in vivo results shown in (**J**). Pg, progesterone. (**K**) Two-input AND gate schematic, with nMS titration results shown in (**L**). Trimer 1 and 2 refer to partial trimeric complexes of the two dimerizer proteins binding to either one of the monomers. (**M**) A three-input AND gate, with nMS titration results shown in (**N**). Tetramer 1 and 2 refer to partial tetrameric complexes of the three dimerizer proteins binding to either one of the monomers. All error bars are standard deviations of *n* = 3 independent replicates.

Downloaded from http://science.sciencemag.org/ on April 4, 2020

other such that the largely hydrophobic interaction surfaces are buried against each other; such a structure would have to partially unfold for ***1′-2′*** to interact with either ***1*** or ***2***. The magnitude of the unfolding energy (ΔG_{open} in Fig. 1B) determines the extent of cooperativity for the gate. We selected linker lengths of 6, 10, or 12 residues for all of the following experiments.

We studied the cooperativity of the induced dimerizer system in vitro using native mass spectrometry (nMS, *36, 37*), which can directly measure the populations of different oligomeric species in a sample (tables S5 to S8; for calibration curve, see fig. S3). We first measured the extent to which ***1*** activates the binding of ***2*** to ***1′-2′*** (Fig. 1F). ***1***, ***2***, and ***1′-2′*** were separately expressed in *Escherichia coli* and purified. At 10 μM of ***1′-2′*** and 20 μM of 2, we observed a 33 fold increase in binding between ***2*** and ***1′-2′*** upon addition of 10 μM of ***1*** (Fig. 1G), a fold increase comparable to naturally occurring allosteric systems (*33*). To assess the sensitivity of binding to stoichiometric imbalance, 10 μM ***1*** and ***2*** were titrated with increasing concentrations of ***1′-2′*** (Fig. 1H) and the species formed were determined by nMS. The heterotrimeric ***1:1′-2′:2*** complex was observed over a wide range of ***1′-2′*** concentrations (Fig. 1H). Even in the presence of a 6-fold excess of ***1′-2′***, there was no decrease in the amount of ***1:1′-2′:2*** formed, and neither ***1:1′-2′*** nor ***1′-2′:2*** was detected (Fig. 1H). We define a cooperativity parameter *c* as the ratio of the affinities in the presence and absence of the other monomer, which in our model directly

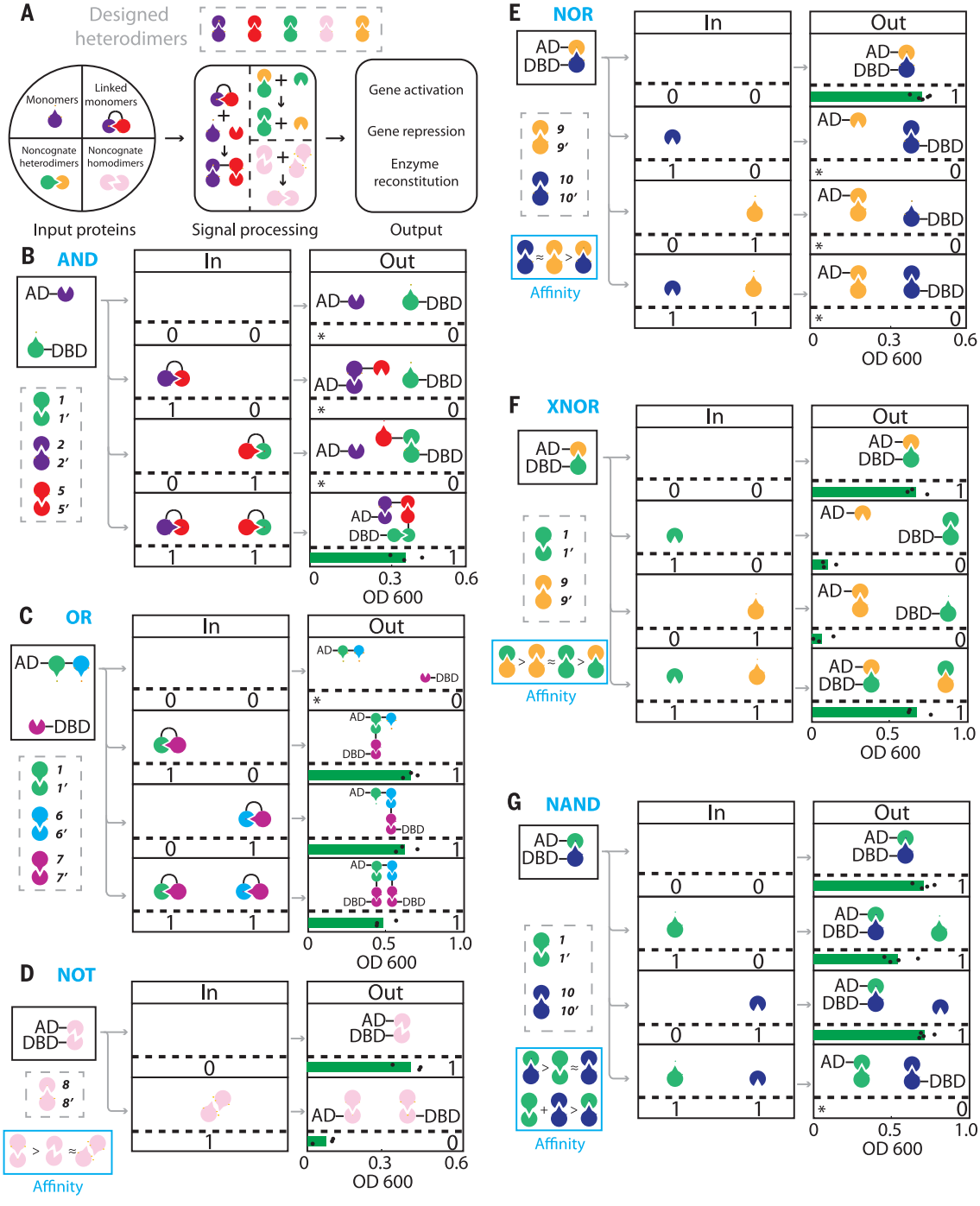

Fig. 2. CIPHR two-input logic gates. (**A**) CIPHR gates are built from DHDs (top) with monomers or covalently connected monomers as inputs (left); some gates use only the designed cognate interactions (left side of middle panel), whereas others take advantage of observed binding affinity hierarchies (right side of middle panel). (**B** and **C**) Two-input AND (B) and OR (C) CIPHR logic gates based on orthogonal DHD interactions. (**D** to **G**) NOT (D), NOR (E), XNOR (F), and NAND (G) CIPHR logic gates made from multispecific and competitive protein binding. For each gate, black dots represent individual Y2H growth measurement corrected over background growth, with their average values shown in green bars. Components in gray boxes indicate the DHD pairs used. Blue boxes indicate affinity gradients. *No yeast growth over background. "0" and "1" in the middle and right blocks represent different input states and expected outputs, respectively.

Downloaded from http://science.sciencemag.org/ on April 4, 2020

relates to the free energy of opening of the dimerizer ($c = e^{\Delta G_{open}/RT}$; see supplementary materials). The estimated c value from fits of the thermodynamic model to nMS data (Fig. 1H, cyan line) is 991,000 ± 21 (for reference, the c value of the naturally occurring N-Wasp system is 350 but system differences complicate quantitative comparisons). This value of c corresponds to ΔG_{open} of 8.2 kcal/mol, which is about half the measured unfolding free energy of ***1′-2′*** (table S3), suggesting that binding may not require complete unfolding of the four-helix bundle state of the dimerizer.

To investigate the cooperativity of the induced dimerizer system in living cells, we used a two-hybrid–like assay in yeast. ***11′*** was fused to the DNA-binding domain ZF43 (*14*), ***7*** to the transactivation domain VP16, and the dimerizer ***11-7′*** was placed under the control of a progesterone-responsive element. Association of the DNA-binding and activation domains results in transcription of red fluorescent protein (RFP) (Fig. 1I). Treating cells with increasing amount of progesterone resulted in up to a 4.5-fold increase in RFP signal, with only a small drop at saturating progesterone concentrations (Fig. 1J). On the basis of calibration curves, under these conditions, ***11-7′*** is expected to be in >5-fold molar excess over ***11′*** and ***7*** (fig. S4), suggesting that ***11-7′*** binds cooperatively to ***11′*** and ***7*** in cells. Thus, the cooperativity of the dimerizer system makes it robust to fluctuating component stoichiometries in cells.

With dimerizers displaying cooperative binding, we reasoned that the lack of dependence on stoichiometric excesses of one of the components should extend to more complex gates. Using nMS, we investigated the cooperativity of a two-input AND gate constructed with the two dimerizers ***1′-3′*** and ***3-2′*** as inputs and monomers ***1*** and ***2*** brought together by the two inputs (Fig. 1K). As the concentration of the 2 inputs increased, the amount of heterotetrameric complex plateaued at a stoichiometry of 2:1, and then largely remained constant with a small drop at molar ratio of 6:1. Only very small amounts of partial complexes (heterotrimers and heterodimers) were observed, further indicating high cooperativity (Fig. 1L). We constructed a three-input AND gate from ***1′-4′***, ***4-3′***, and ***3-2′***, which together control the association of ***1*** and ***2*** (Fig. 1M). Similar to the two-input AND gate, the abundance of full, pentameric complexes only decreased slightly at greater than stoichiometric concentrations of inputs with no detectable competing tetrameric complexes (Fig. 1N).

We explored the modular combination of DHDs (table S1) to generate a range of two-input cooperatively inducible protein heterodimer (CIPHR) logic gates. Monomers from individual DHDs were linked to effector proteins of interest by genetic fusion such that the inputs (linked heterodimer subunits) control colocalization or dissociation of the effector proteins. Taking advantage of previously measured all-by-all specificity matrices for the DHDs (*32*), we explored constructing gates from two interaction modalities: cognate binding between designed protein pairs and competitive binding involving multispecific interactions (Fig. 2A).

We began by constructing AND and OR gates, reading out gate function using a yeast-two-hybrid (Y2H) setup similar to previously described yeast-four-hybrid systems (*38*, *39*). To construct an AND gate, we fused ***2*** to the Gal4 activation domain (AD) and ***1*** to the Gal4 DNA-binding domain (DBD). In this scheme, the colocalization of AD and DBD, and the resulting transcriptional activation of the *His3* gene, should require the expression of both input proteins (***1′-5, 5′-2′***). Indeed, growth in medium lacking histidine required expression of both inputs (Fig. 2B). An OR gate was similarly constructed by linking the ***1-6*** fusion to the AD and ***7′*** to the DBD. Expression of either of the inputs, ***1′-7*** or ***6′-7***, resulted in growth by driving association of AD with DBD (Fig. 2C).

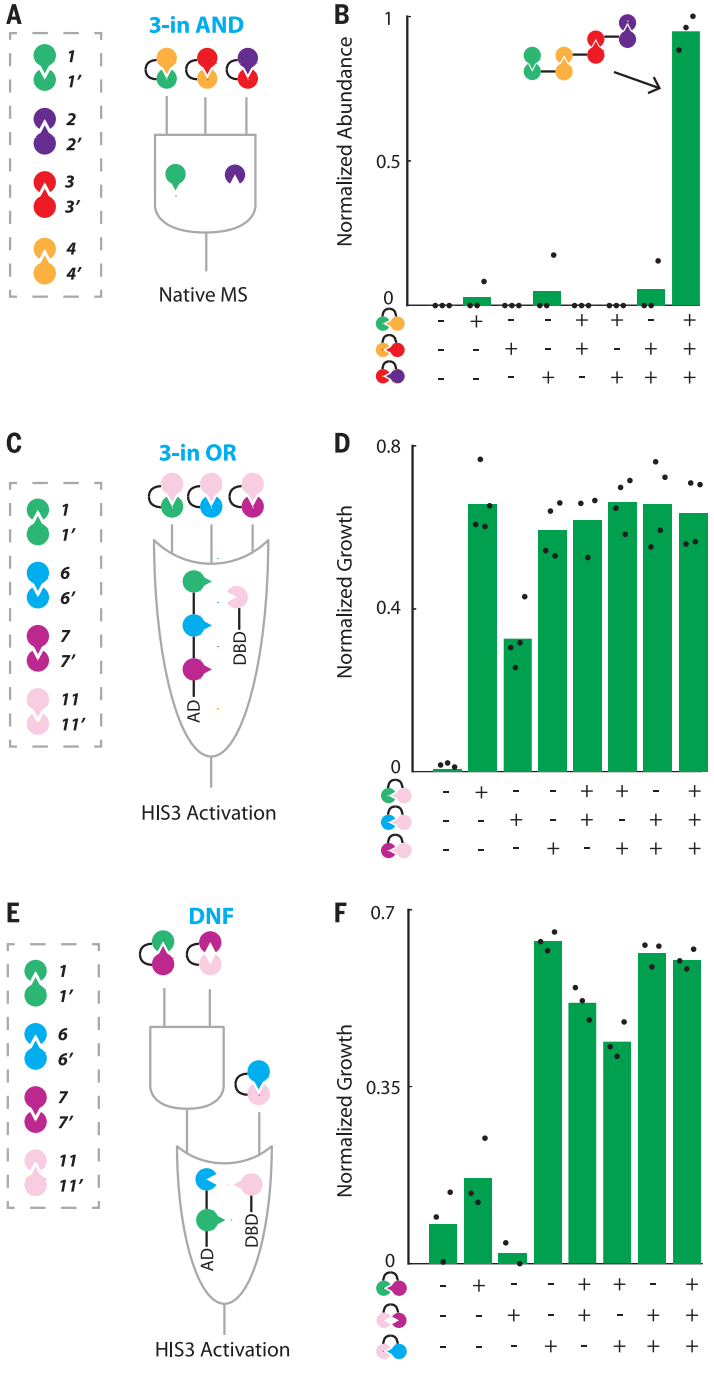

Fig. 3. Three-input CIPHR logic gates. (**A**) Schematic of a three-input AND gate. (**B**) nMS results indicating proper activation of the three-input AND gate only in the presence of all three inputs. (**C**) Schematic of a three-input OR gate. (**D**) Y2H results confirming activation of the three-input OR gate with any of the inputs. (**E**) Schematic of a DNF gate. (**F**) Y2H results confirming proper activation of the gate. For each gate, black dots represent individual measurements with their average values shown in green bars. For Y2H-based measurements [(D) and (F)], the growth measurements are corrected over background growth. Components in gray boxes indicate the DHD pairs used.

Downloaded from http://science.sciencemag.org/ on April 4, 2020

We explored the construction of additional Boolean logic gates by exploiting binding affinity hierarchies identified in all-by-all Y2H experiments (*32*). ***8*** alone formed a homodimer, but in the presence of ***8′*** it dissociated to form the ***8:8′*** heterodimer (fig. S5A). We constructed a NOT gate by fusing ***8*** to both AD and DBD; the ***8:8*** homodimer supported yeast growth but in the presence of coexpressed ***8′*** input protein, the interaction was broken and growth was slowed (Fig. 2D). On the basis of the affinity hierarchy ***9:9′*** ≈ ***10:10′*** > ***9:10′*** (fig. S5B), we constructed a NOR gate in which ***9*** was fused to the AD and ***10′*** to the DBD, with ***9′*** and ***10*** as the two inputs. Either or both inputs outcompeted the ***9:10′*** interaction and hindered yeast growth (Fig. 2E). On the basis of the affinity hierarchy ***9′:1′*** > ***9:9′*** ≈ ***1:1′*** > ***9:1*** (fig. S5B), an XNOR gate was constructed by fusing ***9*** to AD, ***1*** to DBD, and using ***9′*** and ***1′*** as the two inputs: the presence of either outcompeted the ***9:1*** binding and blocked growth, but when both were expressed they instead interacted with each other and growth was observed (Fig. 2F). Similarly, a NAND gate was designed based on the interaction hierarchy ***1′:10′*** > ***1:1′*** ≈ ***10:10′*** (fig. S5B). Neither ***1*** nor ***10*** alone could outcompete the ***1′:10′*** interaction and hence growth occured, but when both were expressed, the free energy of formation of both ***1:1′*** and ***10:10′*** outweighed that of ***1′:10′*** and growth was blocked (Fig. 2G).

We next investigated three-input CIPHR logic gates. We first used nMS to characterize a three-input AND gate (Fig. 1M) in which monomers ***1*** and ***2*** are brought into proximity by the three inputs ***1′-4′***, ***4-3′***, and ***3-2′***. We experimentally tested all eight possible input combinations (Fig. 3A) with both ***1*** and ***2*** present, quantifying all complexes using nMS. Consistent with three-input AND gate function, ***1*** and ***2*** only showed significant coassembly when all three inputs were present (Fig. 3B).

To test three-input CIPHR gate function in cells, we designed two additional gates using the same four pairs of DHDs and tested them by Y2H. To make a three-input OR gate, ***1′-6-7*** was fused to AD and ***11′*** to DBD. Any one of the three inputs (***11-1***, ***11-6′***, or ***11-7′***) connects the AD to the DBD through ***1′***, ***6***, or ***7***, respectively (Fig. 3C). Y2H results confirmed the expected behavior of this logic gate in cells: any of the input proteins induced cell growth (Fig. 3D). We additionally constructed a CIPHR-disjunctive normal form [DNF (***A*** AND ***B***) OR ***C***] gate by fusing ***1′-6*** to AD and ***11′*** to DBD with inputs ***11-7′***, ***7-1***, or ***11-6′*** (Fig. 3E). In Y2H

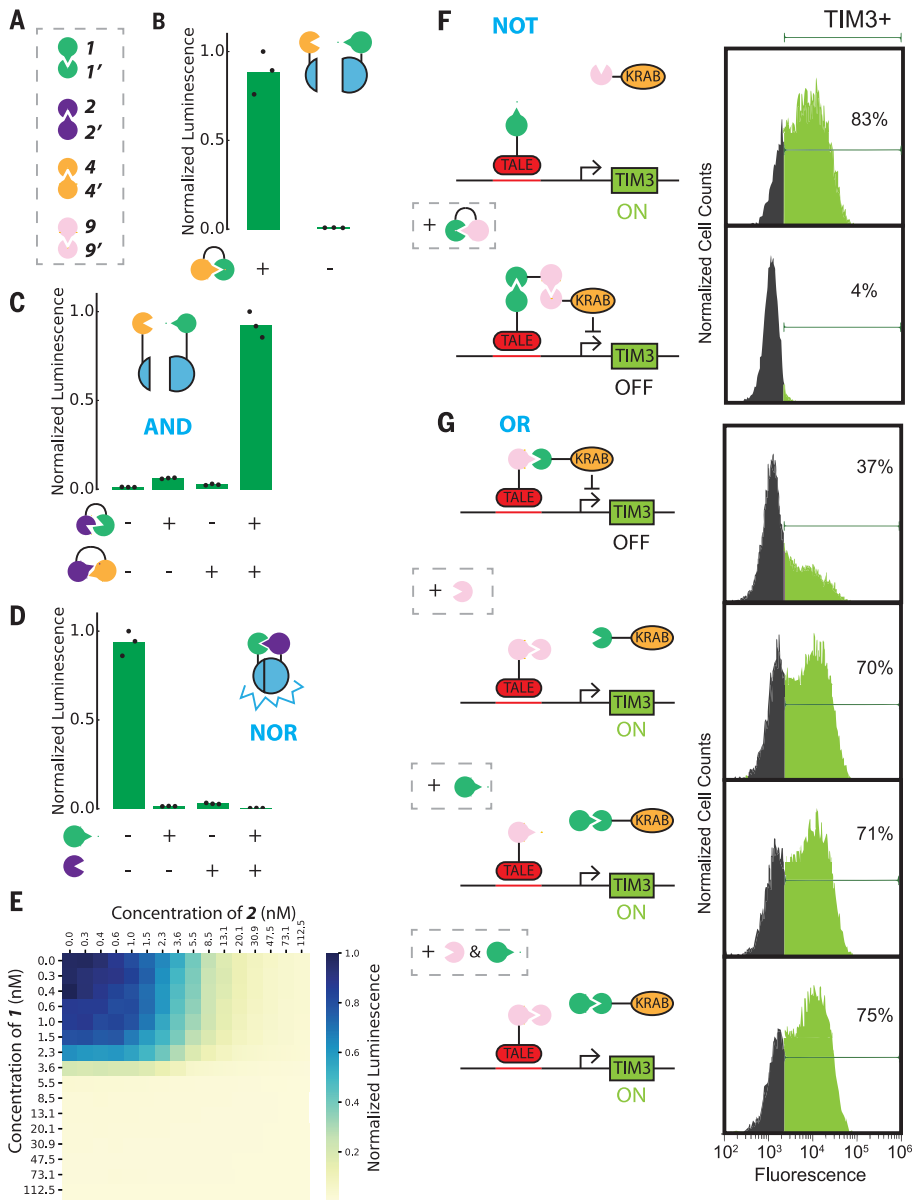

Fig. 4. Transferability of CIPHR logic gates. (**A**) Four pairs of DHDs were modularly combined to construct CIPHR logic gates that can be used to control different functions: catalytic activity of split luciferase (**B** to **E**) and gene expression in primary human T cells (**F** and **G**). (B) the induced dimerization system, (C) AND gate, and (D) NOR gate coupled to NanoBiT split luciferase system, tested by in vitro translation and monitoring luminescence. (E) In vitro titration of the two inputs of the NOR gate in D while keeping ***1′***-smBiT and ***2′***-lgBiT fixed at 5 nM. NOT gate (F) and OR gate (G) using a split TALE-KRAB repression system to control expression of TIM3 proteins in primary human T cells, tested by flow cytometry.

Downloaded from http://science.sciencemag.org/ on April 4, 2020

experiments, the DNF gate functioned as designed, with low yeast growth levels when no input or only one of the ***11-7′*** and ***7-1*** input proteins was present and high yeast growth levels otherwise (Fig. 3F).

To test the transferability of CIPHR logic gates, we explored the ability of CIPHR logic gates to reconstitute split enzyme activity by controlling the association of the two halves of the NanoBiT split luciferase system (*40*–*42*). Monomers from ***1:1′***, ***2:2′***, ***4:4′***, and ***9:9′*** (Fig. 4A) were fused in pairs to the two split domains (smBiT and lgBiT) and produced by in vitro transcription and translation, which facilitated a rapid testing cycle enabling the full 4×4 interaction affinity hierarchy to be determined by monitoring luciferase activity after mixing (fig. S6A). On the basis of this hierarchy, we constructed and experimentally verified an induced dimerization circuit with ***4***-smBiT, ***1***-lgBiT, and ***1′-4′*** as the input (Fig. 4B and fig. S6, C and D); characterization of the time dependence of the response revealed a 7-fold increase in signal 5 min after adding inputs (fig. S6D). We also constructed an AND gate with ***4***-smBiT, ***1***-lgBiT, and ***1′-2*** and ***2′-4′*** as the inputs (Fig. 4C) and a NOR gate with ***1′***-smBiT, ***2′***-lgBiT, and ***1*** and ***2*** as the inputs (Fig. 4D), both of which had the designed dependence of gate function (i.e., luciferase activity) on the inputs. We investigated the response of the NOR gate to varying concentrations of the inputs against the NanoBiT components held at 5 nM and found a sharp drop in signal around 5 nM for both inputs, consistent with NOR logic (Fig. 4E and fig. S6E).

Engineered T cell therapies are promising therapeutic modalities (*43*–*45*) but their efficacy for treating solid tumors is limited at least in part by T cell exhaustion (*46*, *47*). Immune checkpoint genes including *TIM3* are believed to play critical roles in modulating T cell exhaustion (*48*–*50*). To put the transcription of such proteins under the control of the CIPHR logic gates, we took advantage of potent and selective transcriptional repressors of immune checkpoint genes in primary T cells that combine sequence-specific transcription activator-like effector (TALE) DNA-binding domains with the Krüppel-associated box (KRAB) repressor domain (*51*). Repression activity is preserved in split systems pairing a DNA recognition domain fused to a DHD monomer with a repressor domain fused to the complementary DHD monomer (*51*). We reasoned that this system could be exploited to engineer programmable therapeutic devices by making the joining of the DNA recognition and transcriptional repression functionalities dependent on CIPHR gates. Use of a repressive domain effectively reverses the logic of CIPHR gates when expression level of the target gene is measured as the output.

To test the feasibility of this concept, we used a TALE-KRAB fusion engineered to repress the immune checkpoint gene *TIM3* (*51*). We designed a NOT gate with ***1*** fused to the TALE DNA recognition domain, ***9′*** fused to KRAB, and the ***1′-9*** dimerizer protein as the input (see fig. S7A for T cell DHD specificity matrix). In this scheme, ***1′-9*** brings KRAB to the promoter region bound by the TALE, therefore triggering repression of *TIM3* (Fig. 4F). Taking advantage of the interaction between ***9*** and ***1′***, we built an OR gate with ***9***-TALE and ***1′***-KRAB fusions; *TIM3* was repressed in the absence of inputs but upon addition of either ***9′*** or ***1***, the weaker ***9:1′*** interaction was outcompeted in favor of the stronger ***9:9′*** and ***1:1′*** interactions, restoring *TIM3* expression (Fig. 4G). These results suggest that the combination of CIPHR and TALE-KRAB systems could be directly applied to add signal-processing capabilities to adoptive T cell therapy.

The systematic design of logic gates described herein takes advantage of the strengths of de novo protein design. Because the building block heterodimers are designed de novo, many more components for gate construction with nearly identical overall topology can be generated than are available by repurposing biological motifs. The encoding of specificity using designed hydrogen bond networks enables a wide range of binding affinities between monomers with similar structures, which in turn allows the construction of more complex gates that are based on competitive binding. From the protein biophysics perspective, our results highlight the strong synergy between de novo design of protein complexes and nMS and, more generally, the ability of de novo protein design to generate complex cooperative assemblies. For example, detecting and quantifying the 33-fold activation of binding in Fig. 1G depended critically on the ability to resolve all species formed in solution by nMS. Analysis of the three-input logic gates in Fig. 3B required distinguishing the designed heteropentameric assemblies, which are composed of five distinct protein chains, from the very large number of alternative possible heterotetrameric, trimeric, and dimeric complexes. The ability to generate highly cooperative and well-defined assemblies composed of five distinct polypeptide chains demonstrates that de novo protein design is starting to approach the complexity of naturally occurring protein assemblies, which are responsible for much of biological function.

Unlike nucleic acid–based logic gates, CIPHR gates can be directly coupled to arbitrary protein actuation domains, offering greater diversity in the types of functional outputs. We illustrate here the coupling to transcriptional activation and repression and split enzyme reconstitution; in principle, any function that can be modulated by protein-protein association can be put under the control of the CIPHR gates. Because the designed components are hyperstable proteins and no additional cellular machinery is required, the gates should function in a wide range of conditions inside and outside of cells (here, we have demonstrated function with purified components in cell-free extracts, yeast cells, and T cells). The small size of DHDs and thus their genetic payload makes them attractive for mammalian cell engineering. The sophistication of the circuits could be further increased by proteolytic activation as in recent elegant studies using protease-based protein circuits (*30*, *31*); our purely protein interaction–based circuits have advantages in bioorthogonality, demonstrated scalability to three inputs, composability (the output, like the input and the computing machinery, consists of interactions between building blocks with common design features), and extensibility because an essentially unlimited repertoire of heterodimeric building blocks can be created using de novo design.

REFERENCES AND NOTES

1. R. Nussinov, *Mol. Biosyst.* **8**, 22–26 (2012).
2. A. W. Reinke, J. Baek, O. Ashenberg, A. E. Keating, *Science* **340**, 730–734 (2013).
3. Y. E. Antebi *et al.*, *Cell* **170**, 1184–1196.e24 (2017).
4. B. Z. Harris, W. A. Lim, *J. Cell Sci.* **114**, 3219–3231 (2001).
5. G. Seelig, D. Soloveichik, D. Y. Zhang, E. Winfree, *Science* **314**, 1585–1588 (2006).
6. L. Qian, E. Winfree, *Science* **332**, 1196–1201 (2011).
7. M. B. Elowitz, S. Leibler, *Nature* **403**, 335–338 (2000).
8. T. S. Gardner, C. R. Cantor, J. J. Collins, *Nature* **403**, 339–342 (2000).
9. A. Tamsir, J. J. Tabor, C. A. Voigt, *Nature* **469**, 212–215 (2011).
10. P. Siuti, J. Yazbek, T. K. Lu, *Nat. Biotechnol.* **31**, 448–452 (2013).
11. J. Bonnet, P. Yin, M. E. Ortiz, P. Subsoontorn, D. Endy, *Science* **340**, 599–603 (2013).
12. B. H. Weinberg *et al.*, *Nat. Biotechnol.* **35**, 453–462 (2017).
13. S. Ausländer, D. Ausländer, M. Müller, M. Wieland, M. Fussenegger, *Nature* **487**, 123–127 (2012).
14. A. S. Khalil *et al.*, *Cell* **150**, 647–658 (2012).
15. N. Roquet, A. P. Soleimany, A. C. Ferris, S. Aaronson, T. K. Lu, *Science* **353**, aad8559 (2016).
16. L. B. Andrews, A. A. K. Nielsen, C. A. Voigt, *Science* **361**, eaap8987 (2018).
17. B. Angelici, E. Mailand, B. Haefliger, Y. Benenson, *Cell Rep.* **16**, 2525–2537 (2016).
18. J. J. Lohmueller, T. Z. Armel, P. A. Silver, *Nucleic Acids Res.* **40**, 5180–5187 (2012).
19. A. A. Green, P. A. Silver, J. J. Collins, P. Yin, *Cell* **159**, 925–939 (2014).
20. A. A. Green *et al.*, *Nature* **548**, 117–121 (2017).
21. K. Rinaudo *et al.*, *Nat. Biotechnol.* **25**, 795–801 (2007).
22. L. Wroblewska *et al.*, *Nat. Biotechnol.* **33**, 839–841 (2015).
23. S.-H. Park, A. Zarrinpar, W. A. Lim, *Science* **299**, 1061–1064 (2003).
24. P. L. Howard, M. C. Chia, S. Del Rizzo, F.-F. Liu, T. Pawson, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11267–11272 (2003).
25. B. Groves, A. Khakhar, C. M. Nadel, R. G. Gardner, G. Seelig, *eLife* **5**, e15200 (2016).
26. L. Morsut *et al.*, *Cell* **164**, 780–791 (2016).
27. J. E. Dueber, B. J. Yeh, K. Chak, W. A. Lim, *Science* **301**, 1904–1908 (2003).
28. J. E. Dueber, E. A. Mirsky, W. A. Lim, *Nat. Biotechnol.* **25**, 660–662 (2007).
29. A. J. Smith, F. Thomas, D. Shoemark, D. N. Woolfson, N. J. Savery, *ACS Synth. Biol.* **8**, 1284–1293 (2019).

Downloaded from http://science.sciencemag.org/ on April 4, 2020

30. X. J. Gao, L. S. Chong, M. S. Kim, M. B. Elowitz, . *Science* **361**, 1252–1258 (2018).
31. T. Fink *et al*., *Nat. Chem. Biol.* **15**, 115–122 (2019).
32. Z. Chen *et al*., *Nature* **565**, 106–111 (2019).
33. K. E. Prehoda, J. A. Scott, R. D. Mullins, W. A. Lim, *Science* **290**, 801–806 (2000).
34. B. Yu *et al*., *Cell* **140**, 246–256 (2010).
35. K. N. Dyer *et al*., *Methods Mol. Biol.* **1091**, 245–258 (2014).
36. B. T. Ruotolo, C. V. Robinson, *Curr. Opin. Chem. Biol.* **10**, 402–408 (2006).
37. Z. L. VanAernum *et al*., *Nat. Protoc.* **15**, 1132–1157; Rapid online (2020).
38. A. Pause, B. Peterson, G. Schaffar, R. Stearman, R. D. Klausner, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9533–9538 (1999).
39. B. Sandrock, J. M. Egly, *J. Biol. Chem.* **276**, 35328–35333 (2001).
40. A. S. Dixon *et al*., *ACS Chem. Biol.* **11**, 400–408 (2016).
41. Y.-C. Kwon, M. C. Jewett, *Sci. Rep.* **5**, 8663 (2015).
42. J. R. Porter, C. I. Stains, B. W. Jester, I. Ghosh, *J. Am. Chem. Soc.* **130**, 6488–6497 (2008).
43. S. L. Maude *et al*., *N. Engl. J. Med.* **378**, 439–448 (2018).
44. S. S. Neelapu *et al*., *N. Engl. J. Med.* **377**, 2531–2544 (2017).
45. J. A. Fraietta *et al*., *Nat. Med.* **24**, 563–571 (2018).
46. C. H. June, R. S. O'Connor, O. U. Kawalekar, S. Ghassemi, M. C. Milone, *Science* **359**, 1361–1365 (2018).
47. A. H. Long *et al*., *Nat. Med.* **21**, 581–590 (2015).
48. E. J. Wherry, M. Kurachi, *Nat. Rev. Immunol.* **15**, 486–499 (2015).
49. E. J. Wherry *et al*., *Immunity* **27**, 670–684 (2007).
50. K. E. Pauken *et al*., *Science* **354**, 1160–1165 (2016).
51. M. S. Wilken *et al*., Regulatory DNA keyholes enable specific and persistent multi-gene expression programs in primary T cells without genome modification. bioRxiv 2020.02.19.956730 [Preprint]. 20 February 2020; https://doi.org/10.1101/2020.02.19.956730.
52. Z. Chen *et al.*, Code used to simulate the cooperative binding system for: De novo design of protein logic gates, Zenodo (2020); https://doi.org/10.5281/zenodo.3697264.

ACKNOWLEDGMENTS

We thank A. Ng, T. Nguyen, D. Younger, M. Xie, and B. Groves for assistance with yeast experiments; C. Chow for assistance with protein purification; W. Lim for discussions on protein-binding cooperativity; M. Elowitz, R. Schulman, and N. Woodall for feedback on the manuscript; S. Bermeo, M. Lajoie, R. Langan, and A. Ljubetič for useful discussions; and S. Pennington for making the media for Y2H assays. **Funding:** This work was supported by the Howard Hughes Medical Institute (D.B.), the generosity of Eric and Wendy Schmidt by recommendation of the Schmidt Futures program (D.B. and Z.C.), IPD-WA State funding Y5/07-5568 (D.B.), NIH BTRR Yeast Resource Grant Y8-12/61-3650 (Z.C. and R.D.K.), Bruce and Jeannie Nordstrom/Patty and Jimmy Barrier Gift for the Institute for Protein Design (Z.C. and R.D.K.), Spark ABCA4/63-3819 (Z.C.), NIH P41 grant GM103533 (R.D.K.), Open Philanthropy (D.B. and B.I.M.W.), EMBO/80-7223 (B.I.M.W.), Burroughs Wellcome Fund Career Award at the Scientific Interface (S.E.B.), the Army Research Office W911NF-18-1-0200 (M.C.J.), the Air Force Research Laboratory Center of Excellence grant FA8650-15-2-5518 (M.C.J), the David and Lucile Packard Foundation (M.C.J.), and the Camille Dreyfus Teacher-Scholar Program (M.C.J.). A.H. was supported by the Department of Defense (DOD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program. T cell engineering work was supported by a charitable contribution from GlaxoSmithKline to the Altius Institute for Biomedical Sciences (J.R.P., H.L., M.W., C.C., J.A.S.) and NHLBI grant 7R33HL120752 (J.A.S). V.H.W. was supported by NIH grant P41 GM128577 and Ohio Eminent Scholar funds. H.E.-S. was supported by the Defense Advanced Research Projects Agency, contract no. HR0011-16-2-0045, and is a Chan-Zuckerberg investigator. SAXS data were collected at the Advanced Light Source (ALS) at LBNL, supported by the following grants from the NIH: P30 GM124169-01, ALS-ENABLE P30 GM124169, and S10OD018483; NCI SBDR (CA92584); and DOE-BER IDAT (DE-AC02-05CH11231). **Author contributions:** Z.C., S.E.B., and D.B. conceived of the idea; Z.C. and D.B. designed the research; Z.C. and D.B. wrote the manuscript with help from R.D.K and S.E.B.; Z.C. and R.D.K. performed biophysical characterizations with help from B.I.M.W.; Z.C. performed Y2H experiments; A.H. performed NanoBiT experiments and analyzed data with M.C.J.; F.B., M.J., and Z.L.V. performed nMS experiments and analyzed data with V.H.W.; J.R.P., M.W., and C.C. performed and analyzed data from T cell experiments with transcriptional repressors, supervised by S.G. and J.A.S.; G.D. performed the yeast titration assay and analyzed data with H.E.-S.; all authors discussed the results and commented on the manuscript. **Competing interests:** Z.C, S.E.B., and D.B. are coinventors on patent application PCT/US19/59654 that incorporates discoveries described in this manuscript. D.B. is a cofounder of and holds equity in Lyell Immunopharma and Sana Biotechnology. J.P., M.W., C.C., S.G., and J.A.S. are co-inventors on U.S. patent application 62/937,011 that includes discoveries described in this manuscript. J.P. H.L., and S.E.B. hold equity in Lyell Immunopharma. **Data and materials availability:** Raw data from nMS experiments have been deposited to http://files.ipd.uw.edu/pub/de_novo_logic_2019/190522_native_ms_raw.zip. Code used to simulate the cooperative binding system has been deposited at Zenodo (*52*). Plasmids used in this study are available upon request from the corresponding author.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6486/78/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S7
Tables S1 to S8
References (*53*–*71*)

5 June 2019; accepted 5 March 2020
10.1126/science.aay2790

Downloaded from http://science.sciencemag.org/ on April 4, 2020

Science

De novo design of protein logic gates

Zibo Chen, Ryan D. Kibler, Andrew Hunt, Florian Busch, Jocelynn Pearl, Mengxuan Jia, Zachary L. VanAernum, Basile I. M. Wicky, Galen Dods, Hanna Liao, Matthew S. Wilken, Christie Ciarlo, Shon Green, Hana El-Samad, John Stamatoyannopoulos, Vicki H. Wysocki, Michael C. Jewett, Scott E. Boyken and David Baker

Science **368** (6486), 78-84.
DOI: 10.1126/science.aay2790

Designer gates

Signaling in cells can occur through protein-protein interactions. Chen *et al.* describe the design of logic gates that can regulate protein association. The gates were built from small, designed proteins that all have a similar structure but where one module can be designed to interact specifically with another module. Using monomers and covalently connected monomers as inputs and encoding specificity through designed hydrogen-bond networks allowed the construction of two-input or three-input gates based on competitive binding. The modular control elements were used to regulate the association of elements of transcription machinery and split enzymes in vitro and in yeast cells.
Science, this issue p. 78

ARTICLE TOOLS http://science.sciencemag.org/content/368/6486/78

SUPPLEMENTARY MATERIALS http://science.sciencemag.org/content/suppl/2020/04/01/368.6486.78.DC1

REFERENCES This article cites 69 articles, 19 of which you can access for free
http://science.sciencemag.org/content/368/6486/78#BIBL

PERMISSIONS http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works

Downloaded from http://science.sciencemag.org/ on April 4, 2020