

PROTEIN DESIGN

Computational design of a modular protein sense-response system

Anum A. Glasgow^{1*}, Yao-Ming Huang^{1*†}, Daniel J. Mandell^{1,2*‡}, Michael Thompson¹, Ryan Ritterston^{1§}, Amanda L. Loshbaugh^{1,3}, Jenna Pellegrino^{1,3}, Cody Krivacic^{1,4}, Roland A. Pache^{1||}, Kyle A. Barlow^{1,2¶}, Noah Ollikainen^{1,2}, Deborah Jeon¹, Mark J. S. Kelly⁵, James S. Fraser^{1,3,6}, Tanja Kortemme^{1,2,3,4,6,7#}

Sensing and responding to signals is a fundamental ability of living systems, but despite substantial progress in the computational design of new protein structures, there is no general approach for engineering arbitrary new protein sensors. Here, we describe a generalizable computational strategy for designing sensor-actuator proteins by building binding sites de novo into heterodimeric protein-protein interfaces and coupling ligand sensing to modular actuation through split reporters. Using this approach, we designed protein sensors that respond to farnesyl pyrophosphate, a metabolic intermediate in the production of valuable compounds. The sensors are functional in vitro and in cells, and the crystal structure of the engineered binding site closely matches the design model. Our computational design strategy opens broad avenues to link biological outputs to new signals.

In the past two decades, computational protein design has created diverse new protein structures spanning helical (1–5), alpha-beta (6–8), and beta-sheet (9, 10) folds. By contrast, our ability to computationally design arbitrary protein functions de novo lags far behind, with relatively few examples that often require screening of many variants (11, 12). One unsolved challenge is the de novo design of small-molecule sensor-actuators in which ligand binding by a protein directly controls changes in downstream functions, a key aspect of cellular signal transduction (13).

Sensing and responding to a small-molecule signal requires both recognition of the target and linking target recognition to an output response. Exciting progress has been made with the design of proteins recognizing new ligands (10, 11, 14–16). A general solution to the second problem, coupling ligand recognition to diverse output responses, has remained challenging. Existing approaches have used a ligand that fluoresces upon binding (10), engineered the sensor components to be unstable and hence inactive in the absence of the ligand (14, 17), or repurposed an allosteric transcription factor (18). Each of these strategies places constraints on the input signals or output responses that can be used.

Here, we describe a computational strategy to engineer protein complexes that can sense a small molecule and respond directly using different biological outputs, creating modular

sensor-actuator systems. Unlike previous work (10, 11, 14, 15) that reengineered existing binding sites or placed ligands into preformed cavities, we build small-molecule binding sites de novo into heterodimeric protein-protein interfaces to create new and programmable chemically induced dimerization systems (CIDs). This strategy is inspired by naturally occurring and reengineered CID systems (19) that have been widely used but are limited to a small number of existing or similar input molecules. We aimed to design synthetic CIDs that could similarly link binding of a small molecule to modular cellular responses through genetically encoded fusions of each sensor monomer to a split reporter (Fig. 1A) but respond to new, user-defined inputs.

To demonstrate this strategy, we chose farnesyl pyrophosphate (FPP) as the target ligand. FPP is an attractive target because it is a toxic intermediate in a commonly engineered biosynthesis pathway for the production of valuable terpenoid compounds (20). Sensors for FPP could be used, for example, to optimize pathway enzymes or, when linked to appropriate outputs, to regulate pathway gene expression in response to changes in metabolite concentrations (21). Our computational strategy (Fig. 1B and supplementary materials and methods) proceeds in four main steps: (i) defining the geometries of minimal FPP-binding sites composed of three to four side chains (termed “motif residues”) that form key interactions with the target ligand; (ii) modeling these

geometries into a dataset of heterodimeric protein-protein interfaces (termed “scaffolds”) and computationally screening for coarsely compatible scaffolds (22); (iii) optimizing the binding sites in these scaffolds using flexible backbone design methods previously used to predict ligand-binding specificities (23–25) but not tested in the de novo design of binding sites (“reshaping”); and (iv) ranking individual designs for experimental testing according to several design metrics, including ligand-binding energy predicted using the Rosetta force field (26) and ligand burial.

Starting with five FPP-binding-site geometries and up to 3462 heterodimeric scaffolds, we selected the highest-ranked designs across three engineered scaffolds for experimental testing (Fig. 1B and supplementary materials and methods): (i) the FKBP-FRB complex originally responsive to rapamycin (27) (one design), (ii) a complex of the bacterial proteins RapF and ComA (28) (four designs), and (iii) an engineered complex of maltose-binding protein (MBP) and an ankyrin repeat (AR) protein (29) (four designs) (Fig. 2A, table S1, and fig. S1). The ligand was placed into the rapamycin site in FKBP-FRB, but the binding sites in the other two complexes were modeled de novo.

To test these computationally designed FPP sensors, we genetically fused the engineered sensor proteins to a well-studied split reporter, the enzyme murine dihydrofolate reductase [mDHFR (30); Fig. 2B and supplementary materials, appendix 1] and expressed the fusion constructs in *Escherichia coli*. We reasoned that functional sensors should exhibit increased growth through FPP-driven dimerization of the sensor proteins and resulting complementation of functional mDHFR under conditions in which endogenous *E. coli* DHFR was specifically inhibited by trimethoprim. Because FPP does not efficiently enter *E. coli*, we added its metabolic precursor, mevalonate, to the growth medium and coexpressed an engineered pathway of five enzymes (20) (Fig. 2B, pMBIS) to produce FPP from mevalonate in the cells. We then monitored sensor function as change in growth in the presence or absence of mevalonate under otherwise identical conditions (Fig. 2B and supplementary materials and methods). In the following, we denote designs by their scaffold (S1, S2, and S3), design generation (1, 2, and 3), and successive letter (A, B, C, etc.) (table S1 and fig. S1).

¹Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA, USA. ²Bioinformatics Graduate Program, University of California San Francisco, San Francisco, CA, USA. ³Biophysics Graduate Program, University of California San Francisco, San Francisco, CA, USA. ⁴UC Berkeley–UCSF Graduate Program in Bioengineering, University of California San Francisco, San Francisco, CA, USA. ⁵Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, CA, USA. ⁶Quantitative Biosciences Institute, University of California San Francisco, San Francisco, CA, USA. ⁷Chan Zuckerberg Biohub, San Francisco, CA, USA.

*These authors contributed equally to this work. †Present address: 23andMe, Inc., Therapeutics, South San Francisco, CA, USA. ‡Present address: GRO Biosciences, Boston, MA, USA. §Present address: Gryphon Scientific, LLC, Takoma Park, MD, USA. ||Present address: Novozymes A/S, Biologisk Vej 2, 2800 Kgs. Lyngby, Denmark. ¶Present address: Adimab LLC, Palo Alto, CA, USA.

#Corresponding author. Email: tanjakortemme@gmail.com

Although seven of the nine selected designs showed only a small signal (S2-1A, B, C, D, S3-1A, B) or no signal (S1-1A), two designs (S3-1C, D) displayed a robust signal response to FPP (Fig. 2C and fig. S2). Both designs resulted from the AR-MBP scaffold (Fig. 2A, S3). For this scaffold, we also generated two libraries: library 1 based on our ensemble design predictions (Fig. 2A and table S2) and library 2 using error-prone polymerase chain reaction (epPCR) starting from design S3-1C. After an initial growth-based selection and subsequent plate-based screens in the presence and absence of FPP (supplementary materials and methods, fig. S3), we identified 36 hits from which we confirmed 27 FPP-responsive sequences by individual growth assays (figs. S4 and S5). One of the most active designs identified across both libraries (S3-2A) was a variant of design S3-1C with two additional mutations distal from the designed FPP-binding site introduced by epPCR. This variant displayed essentially equal activity as the original S3-1C design

when tested under identical conditions (Fig. 2C, table S1, and fig. S2). These results show that library screening or epPCR was not necessary to identify functional sensors; instead, we obtained functional sensors directly by computational design. However, library 1 provided additional active sequences from the sequence tolerance predicted in the ensemble design simulations (Fig. 2A, table S2, and fig. S4).

To further characterize the identified best design, S3-2A (table S1), we performed single-site saturation mutagenesis at (SSM) 11 positions (table S3). We tested the resulting mutants with the growth-based split mDHFR reporter in the presence and absence of FPP under more stringent conditions by increasing the trimethoprim concentration (fig. S6). Whereas at most positions the originally designed amino acid (Fig. 2A, design S3-1C) appeared to be optimal under these conditions, we saw considerable improvements for mutations at two positions, R194A (Fig. 2A, design S3-2B) and R194A/L85G (Fig. 2A, design S3-2C).

These two designs displayed increasing responses to mevalonate at higher trimethoprim concentrations (Fig. 2D). For the most active design, S3-2C, we confirmed that the sensor signal was dependent on the expression of the sensor proteins [Fig. 2E, -IPTG (isopropyl- β -D-thiogalactopyranoside)] and the metabolic pathway that converts added mevalonate to FPP (Fig. 2E, pMBIS). To test for specificity for FPP, we confirmed that the sensor signal was absent when preventing the accumulation of FPP either by inactivating the fifth enzyme in the pathway by a single point mutation (Fig. 2, B and E, *ispA* R116A) or by adding a sixth enzyme that converts FPP to amorpha-diene (Fig. 2, B and E, pB5K). To test whether the sensor signal was dependent on the original four motif side chains, we mutated each individually to alanine and observed decreased sensitivity to the presence of mevalonate for three of the four motif side chains (Fig. 2F, L89, F133, and R145, but not W114). Finally, we tested whether the sensor signal of design

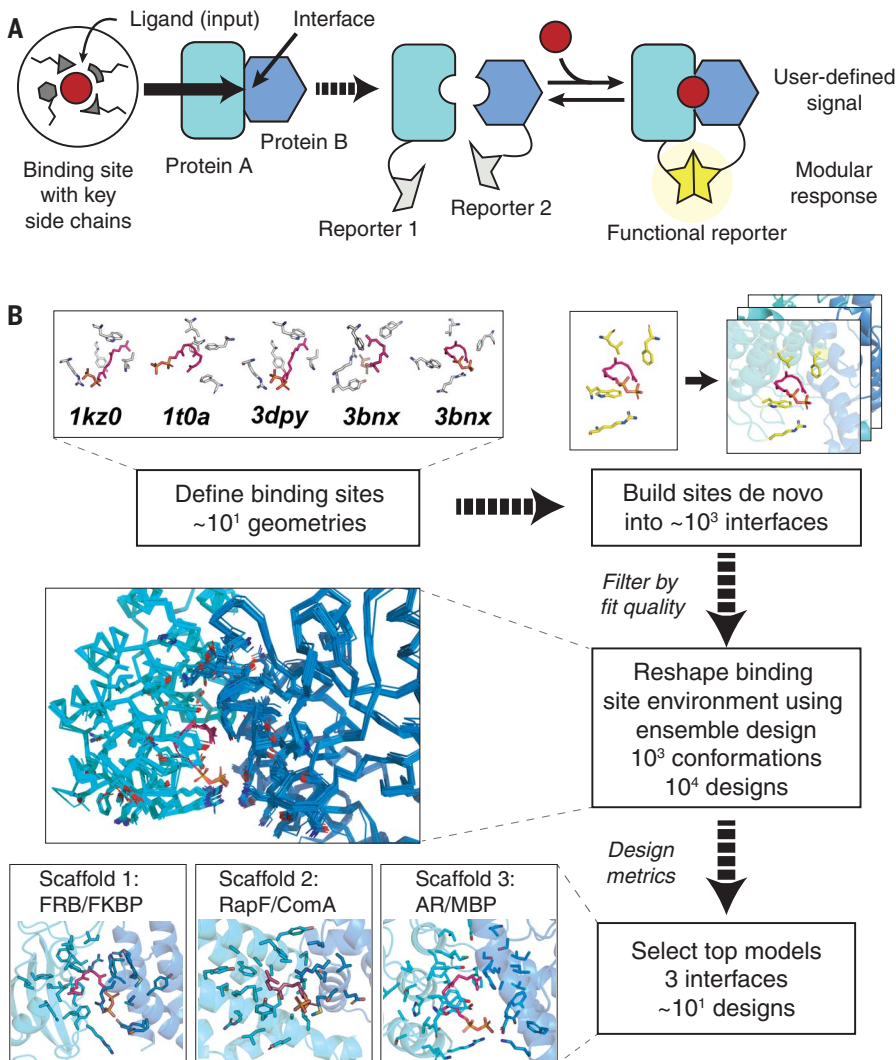


Fig. 1. Computational design. (A) Diagram of the design strategy. A small-molecule binding site is built de novo into protein-protein interfaces (left) to create synthetic CIDs (right). Linking the designed sensor proteins to split reporters yields modular CID systems in which different reporter outputs can be coupled to user-defined small-molecule input signals. (B) Steps in the design of a synthetic CID system sensing FPP. Top: Binding-site geometries with key interacting side chains selected from FPP-binding proteins (PDB codes indicated) are computationally modeled into a large number of protein-protein interfaces. Middle: Binding sites with feasible geometries are reshaped and optimized by flexible backbone design; shown is a conformational ensemble for a single sequence. Bottom: Top designs from three different scaffolds selected for experimental tests (Fig. 2).

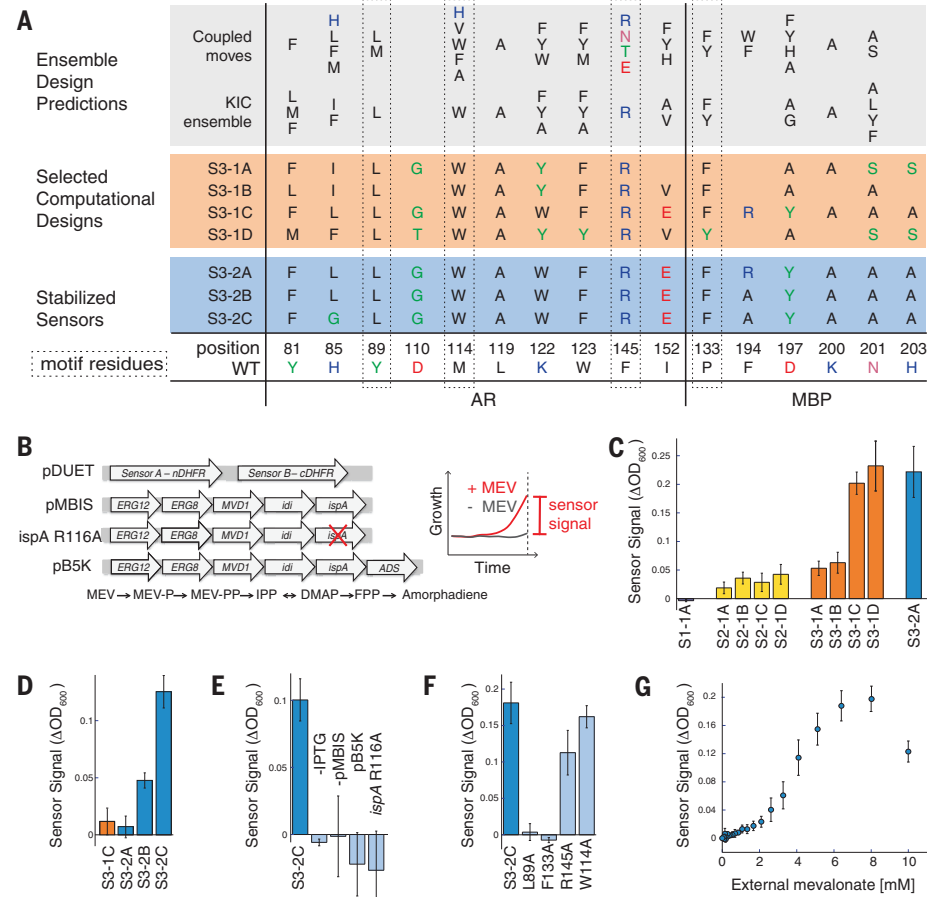
S3-2C was dependent on the concentration of FPP by titrating the extracellular concentration of the mevalonate precursor (Fig. 2G). Although the sensor signal initially increased with increasing mevalonate concentrations, as expected, the signal decreased at the highest mevalonate concentration tested. This behavior likely arises from FPP-mediated toxicity previously observed at this mevalonate concentration using the same FPP-biosynthesis pathway (20). We confirmed a consistent dependency of the sensor signal both on sensor expression (by varying the concentration of the inducer, IPTG) and on mevalonate concentration in the growth medium for seven of our designs (fig. S7, S3-1A, B, C, D, S3-2A, B, C). These results confirm that sensor function in *E. coli* is specific to FPP produced by an engineered pathway, dependent on key residues in the de novo designed binding site, dose dependent in *E. coli*, and sensitive to FPP concentrations in a relevant range (i.e., below the toxicity level).

To confirm biochemically that FPP increases the binding affinity of the AR-MBP complex as designed, we purified the designed AR and MBP proteins without attached reporters [see the supplementary materials and methods; these constructs contained several previously published mutations to stabilize AR (37), which when tested in the split mDHFR reporter assay led to active sensor S3-2D (table S1, fig. S8, and supplementary materials, appendix 2)]. We determined the apparent binding affinity of the designed AR and MBP proteins comprising the S3-2D sensor (Fig. 3A, table S1, and fig. S1) in the absence and presence of 200 μ M FPP using biolayer interferometry (Fig. 3B, fig. S9, and supplementary materials and methods). The presence of FPP led to a >100-fold stabilization of the interaction between the AR and MBP proteins comprising sensor S3-2D [dissociation constant (K_D) from >200 to 2.1 μ M, Fig. 3C; for comparison, the original AR-MBP scaffold had a K_D of 4.4 nM (29)]. Binding of FPP to the designed AR component of S3-2D

alone was weak, and binding of FPP to the designed MBP component of S3-2D alone was not detectable (Fig. 3D). These results confirm in vitro with purified components that design S3-2D functions as a CID system responding to FPP.

To determine whether FPP is recognized in the de novo engineered binding site as predicted by the design model, we determined a 2.2-Å resolution crystal structure of the ternary complex of FPP bound in the engineered AR-MBP interface (supplementary materials and methods; table S4). The crystal structure of the bound complex is in excellent overall agreement with the design model (Fig. 4, A to C). Despite twinning in the crystals, examining unbiased omit maps allowed modeling of unexplained density in the engineered binding site as FPP (Fig. 4B and fig. S10) and confirmed the side-chain conformations in the designed binding pocket (Fig. 4, C and D). Overall, in a 10-Å shell around FPP in the binding pocket, the C α root mean square

Fig. 2. Sensor function in bacteria. (A) Designed sequences at key positions for scaffold 3. Gray shading indicates preferred residues from flexible backbone reshaping by kinematic closure [KIC (23, 24)] or coupled moves (25). Orange shading indicates individual computational designs selected based on ligand burial (S3-1A), consensus (S3-1B), optimized ligand packing (S3-1C), and predicted ligand-binding score (S3-1D). Blue shading indicates sensors stabilized by two additional mutations from SSM (S3-2B and S3-2C also contained two mutations from epPCR that were not in the designed FPP-binding site; fig. S1). **(B)** Constructs (left; for details, see supplementary materials, appendix 1) used in the split mDHFR reporter assay (right). pDUET, sensor proteins linked to the split mDHFR reporter; pMBIS, engineered pathway of five enzymes to convert mevalonate (MEV) into FPP (20); *ispA* R116A, pMBIS containing R116A mutation in *ispA* that reduces catalytic activity 13-fold (37); pB5K, pMBIS with amorphadiene synthase (ADS) (20). Sensor signal is quantified as the change in optical density at wavelength 600 nm (OD_{600}) in the presence and absence of mevalonate. **(C)** Sensor signal in the split mDHFR assay for computational designs based on scaffold 1 (FKBP-FRB12, purple bar), scaffold 2 (RapF-ComA, yellow bars), and scaffold 3 (AR-MBP, orange bars). Sensor S3-2A (identified from library 2 with two mutations distal from the designed FPP-binding site; table S1), is shown for comparison (blue bar). **(D)** Improvement of sensor signal by stability-enhancing mutations in S3-2B and S3-2C at increased stringency [trimethoprim concentration 6 μ M versus 1 μ M in (C)]. **(E)** Dependence of S3-2C sensor signal on sensor expression (-IPTG) and FPP production (-pMBIS, pB5K, *ispA* R116A). **(F)** Dependence of the S3-2C sensor signal on motif residues. **(G)** Dependence of the S3-2C sensor signal on the concentration of the FPP precursor mevalonate added extracellularly. Error bars indicate standard deviation from at least four biological replicates and eight technical replicates for each biological replicate.



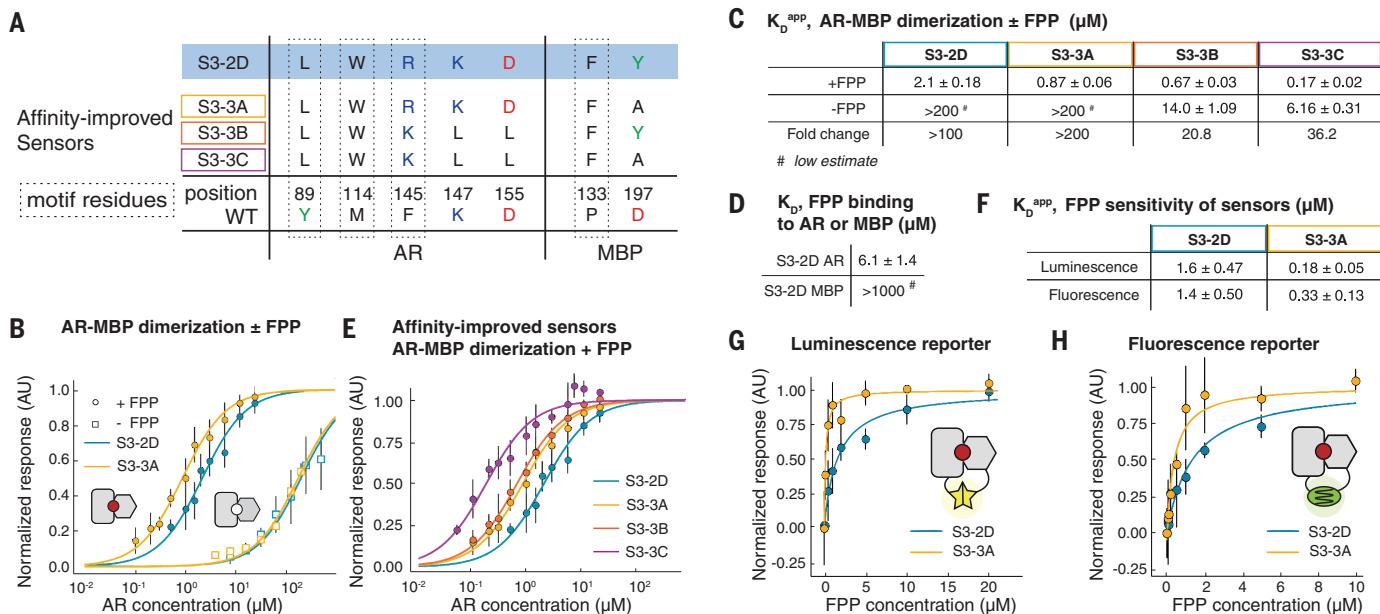


Fig. 3. In vitro sensor characterization and output modularity. (A) Sequence changes in sensor constructs tested in vitro. Motif residues are also shown. The starting construct, S3-2D (blue), is identical to S3-2C in the engineered FPP-binding site but contains additional previously published stabilizing mutations in AR (31) (shown in table S1). (B to H) In vitro binding measurements from biolayer interferometry (BLI) using purified protein (B to E) or FPP titrations with sensors expressed using in vitro TxTI (F to H). (B) Apparent AR interaction with immobilized MBP in the presence (closed circles) or absence (open squares) of 200 μM FPP, comparing designs S3-2D (blue) and S3-3A containing the Y197A

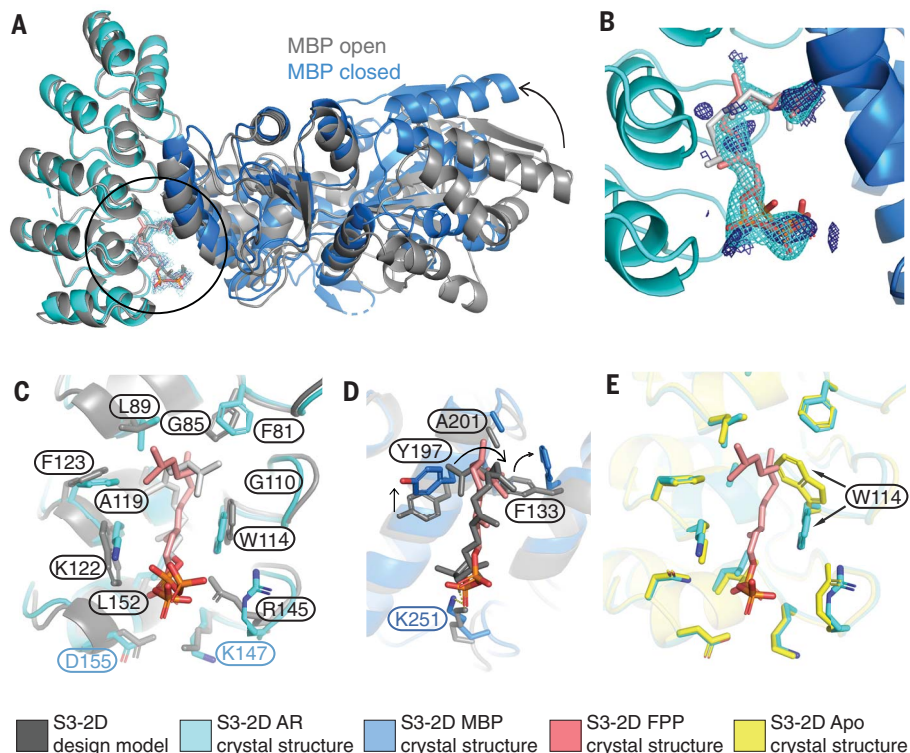
deviation (rmsd) between the model and the structure was 0.53 Å and the all-heavy-atom rmsd was 1.13 Å. Although crystals formed only in the presence of FPP, only one of the two complexes in the asymmetric unit contained FPP in the binding site (fig. S11). This behavior allowed us to compare *apo* and *holo* states of the complex. Most of the designed side chains are in identical conformations in the FPP-bound *holo* and FPP-minus *apo* states (Fig. 4E), suggesting favorable preorganization of the designed binding site. An exception is W114 on AR, which is partly disordered in the *apo* state (fig. S11), providing a potential explanation for why a W114A mutation is less detrimental for sensor activity (Fig. 2F) than expected based on the observed packing interactions between W114 and FPP in the *holo* state. A second slight deviation between the model and the crystal structure appeared to be caused by potential steric clashes of the engineered Y197 on MBP with the modeled FPP conformer, which led to rearrangements in the FPP structure and a rotamer change in designed residue F133 on MBP (Fig. 4D). Many of the original models from computational design favored a smaller alanine side chain at this position (Fig. 2A). These observations led to the prediction that a Y197A mutation might stabilize the ternary complex and, indeed, de-

sign S3-3A containing the Y197A mutation showed an increased (>200 fold) stabilization of the complex with FPP, with an apparent dissociation constant of 870 ± 60 nM for the designed AR and MBP proteins comprising sensor S3-3A in the presence of 200 μM FPP (Fig. 3, B and C). We also confirmed that design S3-3A (table S1) is active in *E. coli* (fig. S12). To further improve the design based on the crystal structure of design S3-2D, we used an additional round of flexible backbone design using the Rosetta coupled-moves method (25) starting from the FPP-bound crystal structure. These simulations suggested three additional mutations leading to design S3-3B: R145K, K147L, and D155L (Fig. 3A). These mutations, when combined with the Y197A mutation (design S3-3C), enhanced the apparent binding affinity of the designed AR and MBP proteins comprising sensor S3-3C in the presence of 200 μM FPP to 170 ± 20 nM (Fig. 3, C and E), which is within 40-fold of the original scaffold AR-MBP interaction affinity (29), but also strengthened the binding affinity of the protein-protein dimer in the absence of FPP to 6.2 ± 0.3 μM (fig. S13). The design simulations optimized sequences for stability of the ternary complex without also destabilizing the dimer in the absence of the small molecule. Methods integrating negative design (32) could be incor-

porated to improve the dynamic range of the system (supplementary text). A key advantage of our CID design strategy is the ability to link an engineered sensor, the input of which is specific to a user-defined small-molecule signal, to a modular output that can in principle be chosen from many available split reporters (Fig. 1A). To test this concept, we linked the engineered CID sensors S3-2D and S3-3A to two additional outputs: a dimerization-dependent fluorescent protein (33) and a split luciferase (34) (Fig. 3, G and H, and supplementary materials, appendix 3). We tested input-output responses using an in vitro transcription-translation system (TxTI) (35) in which FPP can be added at defined concentrations to the assay extract, in contrast to the cell-based split mDHF assay. The TxTI assay revealed a nanomolar FPP sensitivity (K_D^{app}) for our best sensor, S3-3A (Fig. 3F), that was essentially identical for both reporters (180 ± 50 and 330 ± 130 nM by luminescence and fluorescence detection, respectively; Fig. 3, G and H, and fig. S14), and additionally confirms the improvements in design S3-3A containing the Y197A mutation over design S3-2D (the K_D^{app} for S3-2D was 1.6 ± 0.5 and 1.4 ± 0.5 μM for the luminescence and fluorescence reporters, respectively; Fig. 3, F to H). These results show that our CID

Fig. 4. The S3-2D crystal structure closely matches the computational design model.

(A) Overlay of the design model (gray) with the crystal structure (designed AR, cyan; designed MBP, blue; FPP, pink) showing FPP binding in the computationally designed binding site at the AR-MBP interface (circle). The design crystallized in the closed MBP conformation, whereas MBP was in the open conformation in the original scaffold on which the model was based, leading to a difference in rigid-body orientation (arrow) of one lobe of MBP distal to the FPP-binding site. (B) FPP overlaid with $2mF_o - DF_c$ electron density map (1.2σ , cyan) and ligand $2mF_o - DF_c$ omit map (1.0σ , dark blue). Strong density peaks were present in both maps for the phosphates and several anchoring hydrophobic groups. (C) Open-book representation of the FPP-binding site on AR showing close match of designed side-chain conformations to the crystal structure. (D) Open-book representation of the FPP-binding site on MBP indicating a clash between the position of MBP Y197 in the crystal structure (blue) and the designed FPP orientation in the model (gray), causing slight rearrangements of FPP and F133 (arrows). (E) Alignment of the *holo* (cyan) and *apo* (yellow) structures of S3-2D showing overall agreement with the exception of the side chain of W114 (arrows). In (C) to (E), residues are labeled black when designed and cyan or blue when present in the original scaffold complex.



sensor design strategy is compatible with modular outputs.

The most critical feature of our approach is the ability to computationally design small-molecule binding sites de novo into protein-protein interfaces. A previous computational analysis suggested that the appearance of pockets around artificially generated protein-protein interfaces may be an intrinsic geometric feature of protein structure (36), lending support to the idea that our approach is extensible to many other small molecules and interfaces. The design method presented here thus introduces a generalizable way to create synthetic sensing systems with different outputs that can be used in diverse biological contexts to respond to user-specified molecular signals.

REFERENCES AND NOTES

- P. S. Huang *et al.*, *Science* **346**, 481–485 (2014).
- T. M. Jacobs *et al.*, *Science* **352**, 687–690 (2016).
- A. R. Thomson *et al.*, *Science* **346**, 485–488 (2014).
- R. B. Hill, D. P. Raleigh, A. Lombardi, W. F. DeGrado, *Acc. Chem. Res.* **33**, 745–754 (2000).
- P. B. Harbury, J. J. Plecs, B. Tidor, T. Alber, P. S. Kim, *Science* **282**, 1462–1467 (1998).
- G. J. Rocklin *et al.*, *Science* **357**, 168–175 (2017).
- N. Koga *et al.*, *Nature* **491**, 222–227 (2012).
- P. S. Huang *et al.*, *Nat. Chem. Biol.* **12**, 29–34 (2016).
- E. Marcos *et al.*, *Nat. Struct. Mol. Biol.* **25**, 1028–1034 (2018).
- J. Dou *et al.*, *Nature* **561**, 485–491 (2018).
- H. Lechner, N. Ferruz, B. Höcker, *Curr. Opin. Chem. Biol.* **47**, 67–76 (2018).
- J. Dou *et al.*, *Protein Sci.* **26**, 2426–2437 (2017).
- R. M. Gordley, L. J. Bugaj, W. A. Lim, *Curr. Opin. Struct. Biol.* **39**, 106–114 (2016).

- M. J. Bick *et al.*, *eLife* **6**, e28909 (2017).
- C. E. Tinberg *et al.*, *Nature* **501**, 212–216 (2013).
- N. F. Polizzi *et al.*, *Nat. Chem.* **9**, 1157–1164 (2017).
- J. Feng *et al.*, *eLife* **4**, e10606 (2015).
- N. D. Taylor *et al.*, *Nat. Methods* **13**, 177–183 (2016).
- D. M. Spencer, T. J. Wandless, S. L. Schreiber, G. R. Crabtree, *Science* **262**, 1019–1024 (1993).
- V. J. Martin, D. J. Pitera, S. T. Withers, J. D. Newman, J. D. Keasling, *Nat. Biotechnol.* **21**, 796–802 (2003).
- F. Zhang, J. Keasling, *Trends Microbiol.* **19**, 323–329 (2011).
- A. Zanghellini *et al.*, *Protein Sci.* **15**, 2785–2794 (2006).
- M. Babor, D. J. Mandell, T. Kortemme, *Protein Sci.* **20**, 1082–1089 (2011).
- D. J. Mandell, E. A. Coutsias, T. Kortemme, *Nat. Methods* **6**, 551–552 (2009).
- N. Ollikainen, R. M. de Jong, T. Kortemme, *PLOS Comput. Biol.* **11**, e1004335 (2015).
- R. F. Alford *et al.*, *J. Chem. Theory Comput.* **13**, 3031–3048 (2017).
- J. Choi, J. Chen, S. L. Schreiber, J. Clardy, *Science* **273**, 239–242 (1996).
- M. D. Baker, M. B. Neiditch, *PLOS Biol.* **9**, e1001226 (2011).
- H. K. Binz *et al.*, *Nat. Biotechnol.* **22**, 575–582 (2004).
- J. N. Pelletier, F. X. Campbell-Valois, S. W. Michnick, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12141–12146 (1998).
- M. A. Kramer, S. K. Wetzel, A. Plückthun, P. R. Mittl, M. G. Grütter, *J. Mol. Biol.* **404**, 381–391 (2010).
- J. A. Davey, R. A. Chica, *Protein Sci.* **21**, 1241–1252 (2012).
- S. C. Alford, Y. Ding, T. Simmen, R. E. Campbell, *ACS Synth. Biol.* **1**, 569–575 (2012).
- A. S. Dixon *et al.*, *ACS Chem. Biol.* **11**, 400–408 (2016).
- R. Marshall, V. Noireaux, *Methods Mol. Biol.* **1772**, 61–93 (2018).
- M. Gao, J. Skolnick, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 3784–3789 (2012).
- L. Song, C. D. Poulter, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3044–3048 (1994).

ACKNOWLEDGMENTS

We thank J. Keasling and F. Zhang for advice on FPP production in microbes and pathway constructs; E. de los Santos,

Z. Sun, V. Noireux, and R. Murray for TxTI advice and reagents; S. Alford and R. Campbell for dimerization-dependent fluorescent protein constructs; A. Anand, V. Ruiz, B. Adler, and A. Maxwell for contributions to computational database and characterization; S. O'Connor for developing a database for design models; and members of the Kortemme lab for discussion. **Funding:** This work was supported by a grant from the National Institutes of Health (NIH) (R01-GM110089) and a W.M.F. Keck Foundation Medical Research Award to T.K. We additionally acknowledge the following fellowships: NIH IRACDA and UC Chancellor's Postdoctoral Fellowships (A.A.G.), PhRMA Foundation Predoctoral Fellowship in Informatics (D.J.M.), NIH F32 Postdoctoral Fellowship (M.T.), and National Science Foundation Graduate Research Fellowships (J.P. and N.O.). **Author contributions:** D.J.M. and T.K. conceived the idea for the project; D.J.M. developed and performed the majority of the computational design with contributions from A.A.G., R.A.P., K.A.B., N.O., J.P., and T.K.; A.A.G. and Y.-M.H. designed the experimental approach and performed the majority of the experimental characterization with contributions from R.R., A.L.L., C.K., D.J., and M.J.S.K.; M.T. and J.S.F. determined the crystal structure; M.J.S.K., J.S.F., and T.K. provided guidance, mentorship, and resources; and A.A.G. and T.K. wrote the manuscript with contributions from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Coordinates and structure files have been deposited to the Protein Data Bank (PDB) with accession code 60B5. All other relevant data are available in the main text or the supplementary materials. Rosetta source code is available from rosettacommons.org. Upon publication, constructs will be made available by Addgene.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6468/1024/suppl/DC1
Materials and Methods
Supplementary Text
Tables S1 to S4
Figs. S1 to S14
References (38–73)

1 May 2019; accepted 7 October 2019
10.1126/science.aax8780

Computational design of a modular protein sense-response system

Anum A. Glasgow, Yao-Ming Huang, Daniel J. Mandell, Michael Thompson, Ryan Ritterson, Amanda L. Loshbaugh, Jenna Pellegrino, Cody Krivacic, Roland A. Pache, Kyle A. Barlow, Noah Ollikainen, Deborah Jeon, Mark J. S. Kelly, James S. Fraser and Tanja Kortemme

Science **366** (6468), 1024-1028.
DOI: 10.1126/science.aax8780

Sense and respond

Many signaling pathways start with cellular proteins sensing and responding to small molecules. Despite advances in protein design, creating a protein-based sense-and-respond system remains challenging. Glasgow *et al.* designed binding sites at the interface of protein heterodimers (see the Perspective by Chica). By fusing each monomer to one half of a split reporter, they linked ligand-driven dimerization to the reporter output. The computational design strategy provides a generalizable approach to create synthetic sensing systems with different outputs.

Science, this issue p. 1024; see also p. 952

ARTICLE TOOLS

<http://science.sciencemag.org/content/366/6468/1024>

SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2019/11/20/366.6468.1024.DC1>

RELATED CONTENT

<http://science.sciencemag.org/content/sci/366/6468/952.full>

REFERENCES

This article cites 72 articles, 12 of which you can access for free
<http://science.sciencemag.org/content/366/6468/1024#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 © 2019 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works