Design and engineering of genetically encoded protein biosensors for small molecules

Alison C. Leonard[1] and Timothy A. Whitehead[1,*]

[1]Department of Chemical and Biological Engineering, University of Colorado Boulder, Boulder, CO, 80305, USA

*Correspondence to:
Timothy A. Whitehead: timothy.whitehead@colorado.edu
JSC Biotechnology Building

3415 Colorado Avenue, Boulder, CO 80305

Phone: +1 (303)-735-2145

ABSTRACT

Genetically encoded protein biosensors controlled by small organic molecules are valuable tools for many biotechnology applications, including control of cellular decisions in living cells. Here we review recent advances in protein biosensor design and engineering for binding to novel ligands. We categorize sensor architecture as either integrated or portable, where portable biosensors uncouple molecular recognition from signal transduction. Proposed advances to improve portable biosensor development include standardizing a limited set of protein scaffolds, and automating ligand compatibility screening and ligand-protein interface design.

INTRODUCTION

Biosensors which transduce an input signal into a measurable output response have been developed for a wide range of molecules, including proteins, nucleic acids, inorganics, and small organic molecules. Genetically encoded biosensors allow for the quantitative measurement of metabolites in real-time in living cells and organisms [1]. Furthermore, these sensors enable control of cellular processes impacting a wide swathe of biotechnology including agrochemical control of plant traits [2], threat detection by engineered living cells [3], and spatiotemporal control of cellular therapies [4]. One challenge for cellular control is that, historically, scientists were limited to a handful of ligand response systems, including inducible expression systems for tetracyclines [5], IPTG [6], rapamycin [7], and chemically-induced dimerization activators derived from plant hormone pathways [2,4]. Recent advances in computational design [8,9] and/or directed evolution [10] of novel biosensors promise a near-future where such sensors can be constructed for nearly any small organic molecule.

An ideal genetically encodable small-molecule biosensor must possess tunable ligand binding affinity along with exquisite selectivity, be expressed in the target system without disrupting native cellular processes, and can be ported across domains of life using a range of cellular signal transduction mechanisms for different outputs. Here we highlight the recent literature in the development of such protein-based genetically encoded biosensors. As multiple excellent reviews already cover sensors for specific end applications [11–13], our review will instead focus on how current architectures approach the design criteria for the ideal sensor. We will give our perspective on the needed advances in portable sensor design.

MOLECULAR RECOGNITION AND SIGNAL TRANSDUCTION ARCHITECTURE

Biosensors which sense and respond to small molecules must transduce a molecular recognition event to some measurable output [14]. We broadly categorize the sensor architecture, or the protein structures engineered to perform the biosensor function, as either integrated or portable (**Fig. 1**). The unifying principle of integrated biosensors is that each molecular recognition event is compatible with only one output signal such as transcriptional activation or repression, channel

activation, or catalysis [15–18]. Integrated biosensors usually, but not always, are composed of a single macromolecular structure. Portable biosensors, by contrast, consist of separate and interchangeable molecular recognition and signal transduction components. Thus, the same ligand binding architecture can be coupled to different downstream response mechanisms, including fluorescence, transcriptional activation, and functional reconstitution of split proteins [9,12,19]. Figure 1 illustrates examples of various integrated and portable sensor types with representative cartoons of the molecular architecture. Table 1 contains the set of sensors reviewed here that have been engineered for binding to a novel ligand.

Categories of integrated biosensors

Transcription factors in which binding at a given DNA sequence is activated (or de-repressed) by ligand binding are a common integrated biosensor approach, building on foundational inducible expression systems for tetracyclines and other small molecules [5–7]. Transcription factors can either be engineered for binding to non-native effectors [20,21] or can be screened from diverse environmental sources at the genome scale [22] for binding to a target ligand. In an example of the latter approach, Ruegg *et al.* repurposed the multidrug-binding repressor EilR to engineer the "Jungle Express" ligand-inducible gene expression system compatible with diverse proteobacteria [17]. Engineered transcription factors are a common approach for novel metabolite sensors in bacteria, including work by Taylor *et al.* using computational protein design and directed evolution to engineer the allosteric transcription factor LacI for sensitivity to four saccharides not metabolized by *E. coli* [23], Rondon *et al.* to engineer caffeine activation into the PurR antirepressor [24•], d'Oelsnitz *et al.* using directed evolution to evolve the TetR-family regulator CamR for sensitivity to various bicyclic monoterpenes [25], and Rottinghaus *et al.* to engineer TyrR to distinguish phenylalanine and tyrosine [26]. However medium μM to mM sensor sensitivities [23,24•] limit existing engineered transcription factors largely to microbial synthetic biology and metabolic engineering applications [27–29]. Additionally, in cases where the specificity of allosteric transcription factor sensors for their target ligand has

been comprehensively assessed, sensors may demonstrate responsiveness to multiple inducer ligands [30].

Another important class of integrated biosensors are ligand-responsive eukaryotic membrane proteins, where repurposing existing binding pockets leads to either agonist or antagonist activity. The Designer Receptors Exclusively Activated by Designer Drugs (DREADD) biosensor technology, pioneered by Brian Roth's laboratory, is built on G-protein coupled receptors [31]. Recent notable DREADD sensors include the development of novel agonist ligands to activate hM3Dq and HM4D(Gi) neuronal DREADDs [32,33]. Ligand-gated ion channels are a relatively untapped protein with neuronal implications [34]. Magnus *et al.* directly activated neurons and modulated animal behavior in response to a drug by reengineering ligand-gated ion channels with a nicotinic acetylcholine receptor ligand binding domain to bind varenicline, commonly known as the smoking cessation aid Chantix [16].

Another class of integrated biosensors are ligand-responsive proteins fused to fluorescent proteins (reviewed comprehensively here [35]). Usually such sensors are constructed from natural ligand binding domains [36]. A notable example for an engineered sensor was recently described in which Herud-Sikimić *et al.* mutated the receptor TrpR, which natively recognizes the chemically similar tryptophan, to recognize the plant hormone auxin [37•]. TrpR is a homodimer and undergoes a conformational change upon ligand binding; fusion of different fluorescent proteins yields a FRET-based biosensor with a signal to noise ratio of approximately two. Engineered TrpR is one example of the gray area in the integrated-portable sensor dichotomy as the ligand-induced conformational change could be leveraged for the reconstitution of split proteins.

The above are examples where proteins with existing ligand binding pockets are redesigned or engineered for new chemical specificity. In contrast, the Karanicolas lab has developed a distinct 'bottoms-up' approach called chemical rescue of structure [38] which can, in principle, be applied to any protein. The central idea is that a protein structural feature removed by mutation can be replaced by a small molecule ligand. In the absence of ligand the protein is non-

functional, but binding of the ligand restores protein structure and hence function. Recently, this technology was deployed for the chemogenetic control of antibodies [39•]. Advantages of this technology include the applicability to many different proteins, integration of biosensing without changing the size of the gene (important for gene therapy applications), and the relative simplicity of constructing new sensors. Potential disadvantages include the substantial instability of the non ligand-bound protein [40], the relatively weak affinity for ligands, the ability to activate, but not repress, proteins with ligands, and the relatively limited palettes of activating chemicals.

Categories of portable biosensors

Portable biosensors decouple ligand binding from downstream signal transduction (**Fig 1**), which allows one to mix and match a given ligand binding module with different signal transduction mechanisms. The major categories of portable biosensors are chemically-induced dimerization (CID) modules [41,42] and protease or degron regulable domains. CID modules are composed of two separate proteins that form a dimer in the presence of a given ligand. Such chemically-induced proximity can be exploited for transcriptional and post-translational control of protein function.

The first category of CID modules are known as "molecular glues" in which a ligand can bind both CID proteins separately and independently. In this case the ligand serves as a glue for the two proteins [41]. Kang *et al*. described a method called COMBINES-CID involving phage display in which engineered nanobodies sandwich a known ligand, succeeding in developing a sensor for cannabidiol [43]. Cao *et al.* developed a mathematical model to define these molecular glues by comparing the function of this cannabidiol sensor to auxin and known immunomodulatory drugs [44•]. Guo *et al.* used phage display to identify an antibody mimetic compatible with a starting methotrexate-binding anchor domain, creating a novel molecular glue that sandwiches methotrexate [45]. Glasgow *et. al.* computationally designed *de novo* ligand

binding sites into a novel heterodimeric CID protein-protein interface [9]. While molecular glues can achieve robust CID with a well-defined mathematical model, with the potential to engineer novel binding sites into dimerization interfaces, a major disadvantage is that ligand binding must be engineered into two distinct protein surfaces, creating a two-fold design problem.

Molecular glues have also been utilized for the spatiotemporal control of chimeric antigen receptor T (CAR-T) cell therapies [4,46,47•]. In a clever inversion of CID, Giordano-Attianese *et al.* computationally designed a protein interface in which ligand binding directly inhibits dimerization. This enables the STOP-CAR system, in which the chimeric antigen receptor disassembles in the presence of ligand, disrupting function [47•].

A distinct CID module called a "molecular ratchet" [48•] is exemplified by the abscisic acid (ABA) hormone perception system in plants. In this system, ABA binding triggers a conformational change in the soluble receptor protein (PYR1 and other PYR1-like proteins (PYLs)) which recruits a partner protein for heterodimerization. Key to the interaction is a coordinated water with hydrogen bonds between the ligand, the receptor backbone, and a tryptophan indole nitrogen in the partner protein. Ligand binding results in a conformational change that closes this tryptophan latch [49]. The singular advantage of this architecture relative to molecular glues is that, given latch closure with a suitable ligand H-bond acceptor, there is a complete separation from ligand binding and partner protein recruitment, simplifying the engineering challenge as only one protein needs to be engineered. The Cutler group first showed that the PYR1 binding pocket could be repurposed to sense and respond to different agrochemicals, most notably mandipropamid [2]. The resulting $PYR1^{Mandi}$ sensor, when expressed in transgenic plants, could control water use in response to mandipropamid [2]. Ziegler *et al*. showed that $PYR1^{Mandi}$ could be ported into mammalian cells and zebrafish embryos for targeted protein localization in response to mandipropamid [50]. Recently, Beltran *et al.* showed that PYR1 can be repurposed to create biosensors for a diverse range of

compounds [51••]. Remarkably, sensors for 21 of out 38 compounds screened could be identified from a mutational library of just 37,778 mutants. Directed evolution resulted in seven low nanomolar and highly specific biosensors characterized by quantifying the off-target response to chemically similar ligands. This relatively high specificity is notable compared to allosteric transcription factor systems with known promiscuity [30]. These sensors show portability between multiple output systems, including ELISA-like immunoassays, eukaryotic transcriptional activation, and post-translational control of protein function [51••]. Steiner *et al.* have also developed an analytical mathematical model for molecular ratchets and employed this to predict sensor behavior in different transduction mechanisms [48•]. Combined, these examples demonstrate the pliability of CID modules to recognize diverse ligands as well as control signal transduction systems across domains of life.

Protease domains can also serve as a portable molecular recognition architecture, in which a protease domain links inhibitor-specific binding to activation or inhibition of downstream signal transduction. Tague *et al.* used the HCV NS3 protease domain as a ligand-inducible connection (LInC) controlled by protease inhibitor binding, in which ligand binding activates transcription by preserving the connection between modular DNA-binding and transcriptional activation domains [52]. The Pleiotropic Response Outputs from a Chemically-Inducible Single Receiver (PROCISiR) system described in Foight *et al.* utilizes a NS3a protease "reader" to incorporate multiple orthogonal ligand recognition systems for protease inhibitors with distinct output responses into the same protein architecture [53]. Small-molecule protease control using degron tags has also been a successful strategy for controlling CAR-T cells. Juillerat *et al.* and Sahillioglu *et al.* created CAR-T off-switches in which binding of an inhibitor blocks protease cleavage to remove a degron tag [54,55]. Conversely in Richman *et al*., small molecule binding displaces and exposes a degron leading to CAR-T degradation [56]. Jan *et al.* uses distinct on- and off-switches to control CAR-T function using ligand-induced degradation at a zinc finger degron tag (off-switch) and ligand-induced dimerization (on-switch) [57]. While ligand-

controlled protease domains are effective portable biosensors for many small FDA approved antivirals, this approach is limited to known protease inhibitors.

In pursuit of a more general protein degradation strategy applicable to engineered binding domains for novel ligands, Bick *et al*. embedded a computationally-designed fentanyl sensor in a conditionally stable ligand binding domain that is degraded in the absence of ligand [8]. This conditional sensor domain can then be attached to transcription factors enabling downstream expression of a desired reporter.

OPPORTUNITIES FOR IMPROVEMENTS TO PORTABLE BIOSENSORS

The development of portable biosensors enables use of a single molecular recognition architecture in many different downstream applications, reducing the need for redundant sensors for the same ligand. Advances in several key areas will improve the pipeline for facile development of sensors for any small molecule.

Standardization of diverse protein scaffolds for automated ligand pocket design

One open question is the accessible chemical space of existing portable biosensors - are there differences in the molecules they recognize relative to integrated sensors? To address this question, for the engineered sensors cited in this review (Table 1) we quantified ligand properties using PubChem. We find that ligands for existing engineered portable biosensors are more lipophilic (**Fig 2A**), have larger molecular weights (**Fig 2A**), and have more rotatable bonds (**Fig 2B**) than those of the integrated biosensors (p-values 3.3e-6, 2.7e-5, 4.2e-8 respectively). However, no statistically significant differences were observed for the amount of polar surface area or the number of hydrogen bond acceptors. We stress that these trends are preliminary and may be exaggerated by the small sample size of engineered sensors to date, the differences in screening platforms used, the choice of ligands to screen, or other experimental artifacts. Nevertheless, a given protein scaffold is likely to be compatible with only ligands with certain

characteristics, such as the need for a ligand hydrogen bond acceptor at a critical position in the PYR1 pocket [51••].

To reach the goal of portable biosensors for nearly any desired small organic molecule, the creation of more diverse and well-characterized scaffolds, coupled with automated *in silico* ligand docking and screening to identify the appropriate scaffold, may be required to enable a more diverse chemical space for ligand selection. Here, developing a cognate protein binding interface(s) is a critical step in sensor development. Molecular glue and molecular ratchet architectures provide different challenges and opportunities for the automated design of CID interfaces. Molecular glues require design on two independent binding interfaces, which increases the complexity of design but gives greater ligand selection flexibility. Molecular ratchets are advantageous in that the ligand binding and dimerization interfaces are distinct, which enables dimerization at ligand saturation and streamlines the design process in that only one protein must be designed. However, the choice of compatible ligands are constrained by the volume, geometrical requirements, and chemical functional group compatibility of the scaffold binding pocket. For protease inducible domains, one unexplored area is redesigning protease binding sites for competitive inhibition of novel ligands, but that may prove more difficult than repurposing existing binding pockets.

While new and diverse CID scaffolds are needed, a countervailing issue is that every new scaffold needs to be reconfigured for downstream signal transduction systems, such as split proteins, reducing the benefits of portability. For these reasons we propose that it is most efficient to develop automated ligand pocket design methods built on only a handful of well-understood protein scaffolds.

Choice of ligands for cellular control

Ligand choice for contemporary cellular control applications is determined in many ways by the availability of existing sense and response modules (e.g. doxycycline for the TetON and TetOFF systems). As the development of greater scaffold diversity expands the space of available ligands

for novel biosensors, ligand choice can be more finely tuned to the requirements of the cellular control application. Biological and economical requirements will vary between human, plant, and microbial applications, which can include the ligand tissue distribution, lack of interactions with other drugs or other off-target effects, a wide therapeutic window to avoid toxicity, regulatory approval, cell permeability, and ligand cost. Further research is needed for ways to automate this multi-objective optimization ligand selection process.

Shape and quality of transfer functions

In the context of portable biosensors, which have a variety of signal transduction mechanisms to enable cellular control for variegated contexts (**Fig 1**), a transfer function may be defined as the biological output (e.g. enzymatic activity, cell fate decisions) as a function of ligand concentration. Predictive and tunable control of living cells requires advances in three major areas.

First, predictive models should be developed and validated. Equilibrium closed form models exist for both molecular glues and molecular ratchets [48•,58], but extension of these models to biological outputs require further model building and experimental testing for the range of transduction mechanisms. Relatedly, a major contributor to the shape of the transfer function is the biosensor architecture itself. For example, the ternary binding equilibria of molecular glues lead to a bell-shaped transfer function where increasing ligand concentrations can reduce the amount of the ternary complex. By contrast, molecular ratchets have saturable kinetics with minimal cooperativity that are well modeled by the Hill equation [48•]. The shape of the transfer function has large implications for cellular control, particularly for ligands with tissue-specific heterogeneity in which small differences in ligand concentration can have an outsized impact on signal depending on the concentration. Advances are needed in system architectures to convert these signals to digital (all or none responses) or to configurations allowing adaptation (a transient response that returns to baseline values).

Finally, advances are needed in improving the predictive control of split proteins, particularly with respect to basal activity in the absence of ligand. Numerous split proteins have been developed in the last two decades for use with CID and other biosensor systems [19]. To improve split protein design and reduce background signal, Dagliyan *et. al.* and Dolberg *et. al.* have both developed computational split protein optimization workflows [59, 60••]. The Split Protein Optimization by Reconstitution Tuning (SPORT) workflow in Dolberg *et. al.* was validated by engineering an improved split tobacco etch virus protease for transcription factor activation with a low background and high signal induction [60••]. Future structure-based and/or machine learning approaches are anticipated for the predictive design of split proteins.

CONCLUSIONS

Genetically-encoded protein biosensors have been deployed for a diverse range of cellular control applications, demonstrating that the sensor architecture has a tangible impact on the sensor versatility and output signal. The decoupling of molecular recognition and signal transduction in portable biosensors enables the tailoring of sensor output to the application end goals. Advances in the predictive modeling of transfer function output and automated ligand design platforms will further improve biosensor capability, approaching the goal of protein biosensors for any small organic molecule.

Acknowledgements

We thank members of the Whitehead lab for careful reading and suggestions. This work was supported by National Science Foundation NSF Award #2128287 to T.A.W. and an NSF GRFP to A.C.L. A.C.L. also thanks the Interdisciplinary Quantitative Biology program at CU Boulder for ongoing support.

REFERENCES

1. Teng Y, Zhang J, Jiang T, Zou Y, Gong X, Yan Y: **Biosensor-enabled pathway optimization in metabolic engineering**. *Curr Opin Biotechnol* 2022, **75**:102696.

2. Park SY, Peterson FC, Mosquna A, Yao J, Volkman BF, Cutler SR: **Agrochemical control of plant water use using engineered abscisic acid receptors**. *Nature* 2015, **520**:545–548.

3. Shemer B, Shpigel E, Hazan C, Kabessa Y, Agranat AJ, Belkin S: **Detection of buried explosives with immobilized bacterial bioreporters**. *Microb Biotechnol* 2021, **14**:251–261.

4. Wu CY, Roybal KT, Puchner EM, Onuffer J, Lim WA: **Remote control of therapeutic T cells through a small molecule-gated chimeric receptor**. *Science* 2015, **350**:aab4077.

5. Gossen M, Bujard H: **Tight control of gene expression in mammalian cells by tetracycline-responsive promoters**. *Proc Natl Acad Sci U S A* 1992, **89**:5547–5551.

6. Studier FW: **Protein production by auto-induction in high density shaking cultures**. *Protein Expr Purif* 2005, **41**:207–234.

7. Liberles SD, Diver ST, Austin DJ, Schreiber SL: **Inducible gene expression and protein translocation using nontoxic ligands identified by a mammalian three-hybrid screen**. *Proc Natl Acad Sci U S A* 1997, **94**:7825–7830.

8. Bick MJ, Greisen PJ, Morey KJ, Antunes MS, La D, Sankaran B, Reymond L, Johnsson K, Medford JI, Baker D: **Computational design of environmental sensors for the potent opioid fentanyl**. *Elife* 2017, **6**.

9. Glasgow AA, Huang YM, Mandell DJ, Thompson M, Ritterson R, Loshbaugh AL, Pellegrino J, Krivacic C, Pache RA, Barlow KA, et al.: **Computational design of a modular protein sense/response system**. *Science* 2019, **366**:1024.

10. D'Ambrosio V, Pramanik S, Goroncy K, Jakočiūnas T, Schönauer D, Davari MD, Schwaneberg U, Keasling JD, Jensen MK: **Directed evolution of VanR biosensor specificity in yeast**. *Biotechnology Notes* 2020, **1**:9–15.

11. Sahillioglu AC, Schumacher TN: **Safety switches for adoptive cell therapy**. *Curr Opin Immunol* 2022, **74**:190–198.

12. Gangopadhyay SA, Cox KJ, Manna D, Lim D, Maji B, Zhou Q, Choudhary A: **Precision Control of CRISPR-Cas9 Using Small Molecules and Light**. *Biochemistry* 2019, **58**:234.

13. Lieb A, Weston M, Kullmann DM: **Designer receptor technology for the treatment of epilepsy**. *EBioMedicine* 2019, **43**:641.

14. Chen Z, Elowitz MB: **Programmable protein circuit design**. *Cell* 2021,

184:2284–2301.

15. Younger AKD, Dalvie NC, Rottinghaus AG, Leonard JN: **Engineering Modular Biosensors to Confer Metabolite-Responsive Regulation of Transcription**. *ACS Synth Biol* 2017, **6**:311–325.

16. Magnus CJ, Lee PH, Bonaventura J, Zemla R, Gomez JL, Ramirez MH, Hu X, Galvan A, Basu J, Michaelides M, et al.: **Ultrapotent chemogenetics for research and potential clinical applications**. *Science* 2019, **364**.

17. Ruegg TL, Pereira JH, Chen JC, DeGiovanni A, Novichkov P, Mutalik VK, Tomaleri GP, Singer SW, Hillson NJ, Simmons BA, et al.: **Jungle Express is a versatile repressor system for tight transcriptional control**. *Nat Commun* 2018, **9**.

18. Faccio G: **From Protein Features to Sensing Surfaces**. *Sensors* 2018, **18**.

19. Lim SA, Wells JA: **Chapter Twelve - Split enzymes: Design principles and strategy**. In *Methods in Enzymology*. Edited by Tawfik DS. Academic Press; 2020:275–296.

20. Galvão TC, de Lorenzo V: **Transcriptional regulators à la carte: engineering new effector specificities in bacterial regulatory proteins**. *Curr Opin Biotechnol* 2006, **17**:34–42.

21. Tang S-Y, Fazelinia H, Cirino PC: **AraC regulatory protein mutants with altered effector specificity**. *J Am Chem Soc* 2008, **130**:5267–5271.

22. Hanko EKR, Paiva AC, Jonczyk M, Abbott M, Minton NP, Malys N: **A genome-wide approach for identification and characterisation of metabolite-inducible systems**. *Nat Commun* 2020, **11**:1213.

23. Taylor ND, Garruss AS, Moretti R, Chan S, Arbing MA, Cascio D, Rogers JK, Isaacs FJ, Kosuri S, Baker D, et al.: **Engineering an allosteric transcription factor to respond to new ligands**. *Nat Methods* 2016, **13**:177.

24• Rondon R, Wilson CJ: **Engineering Alternate Ligand Recognition in the PurR Topology: A System of Novel Caffeine Biosensing Transcriptional Antirepressors**. *ACS Synth Biol* 2021, **10**:552–565.

•An example of allosteric transcription factor engineering, in which the PurR antirepressor was reengineered to respond to caffeine only. A system of caffeine antirepressors with altered DNA binding was then used to regulate a genetic circuit using NOR logic. This work builds on previous efforts to rewire LacI-type repressors to antirepressors.

25. d'Oelsnitz S, Nguyen V, Alper HS, Ellington AD: **Evolving a Generalist Biosensor for Bicyclic Monoterpenes**. *ACS Synth Biol* 2022, **11**:265–272.

26. Rottinghaus AG, Xi C, Amrofell MB, Yi H, Moon TS: **Engineering ligand-specific biosensors for aromatic amino acids and neurochemicals**. *Cell Systems* 2022, **13**:204–214.e4.

27. Hanko EKR, Sherlock G, Minton NP, Malys N: **Biosensor-informed engineering of Cupriavidus necator H16 for autotrophic D-mannitol production**. *Metab Eng* 2022, **72**:24–34.

28. Seok JY, Han YH, Yang JS, Yang J, Lim HG, Kim SG, Seo SW, Jung GY: **Synthetic biosensor accelerates evolution by rewiring carbon metabolism toward a specific metabolite**. *Cell Rep* 2021, **36**:109589.

29. Wu Y, Jameel A, Xing X-H, Zhang C: **Advanced strategies and tools to facilitate and streamline microbial adaptive laboratory evolution**. *Trends Biotechnol* 2022, **40**:38–59.

30. Nasr MA, Timmins LR, Martin VJJ, Kwan DH: **A versatile transcription factor biosensor system responsive to multiple aromatic and indole inducers.** *ACS Synth Biol* 2022, **11**:1692-1698.

31. Urban DJ, Roth BL: **DREADDs (designer receptors exclusively activated by designer drugs): chemogenetic tools with therapeutic utility**. *Annu Rev Pharmacol Toxicol* 2015, **55**:399–417.

32. Weston M, Kaserer T, Wu A, Mouravlev A, Carpenter JC, Snowball A, Knauss S, von Schimmelmann M, During MJ, Lignani G, et al.: **Olanzapine: A potent agonist at the hM4D(Gi) DREADD amenable to clinical translation of chemogenetics**. *Sci Adv* 2019, **5**:eaaw1567.

33. Thompson KJ, Khajehali E, Bradley SJ, Navarrete JS, Huang XP, Slocum S, Jin J, Liu J, Xiong Y, Olsen RHJ, et al.: **DREADD Agonist 21 Is an Effective Agonist for Muscarinic-Based DREADDs in Vitro and in Vivo**. *ACS Pharmacol Transl Sci* 2018, **1**:61–72.

34. Lieb A, Qiu Y, Dixon CL, Heller JP, Walker MC, Schorge S, Kullmann DM: **Biochemical autoregulatory gene therapy for focal epilepsy**. *Nat Med* 2018, **24**:1324.

35. Greenwald EC, Mehta S, Zhang J: **Genetically Encoded Fluorescent Biosensors Illuminate the Spatiotemporal Regulation of Signaling Networks**. *Chem Rev* 2018, **118**:11707–11794.

36. Dong A, He K, Dudok B, Farrell JS, Guan W, Liput DJ, Puhl HL, Cai R, Wang H, Duan J, et al.: **A fluorescent sensor for spatiotemporally resolved imaging of endocannabinoid dynamics in vivo**. *Nature Biotechnology 2021 40:5* 2021, **40**:787–798.

37• Herud-Sikimić O, Stiel AC, Kolb M, Shanmugaratnam S, Berendzen KW, Feldhaus C, Höcker B, Jürgens G: **A biosensor for the direct visualization of auxin**. *Nature* 2021, **592**:768.

•In this work the tryptophan receptor TrpR was engineered to instead recognize the related metabolite indole-3-acetic acid, known as the plant hormone auxin. This is a rare example of a fluorescent metabolite biosensor engineered for novel ligand specificity.

38. Deckert K, Budiardjo SJ, Brunner LC, Lovell S, Karanicolas J: **Designing allosteric control into enzymes by chemical rescue of structure**. *J Am Chem Soc* 2012, **134**:10055–10060.

39• Khowsathit J, Bazzoli A, Cheng H, Karanicolas J: **Computational Design of an Allosteric Antibody Switch by Deletion and Rescue of a Complex Structural Constellation**. *Cite This: ACS Cent Sci* 2020, **2020**:403.

•Applies the chemical rescue of structure approach of ligand-inducible control to antibodies. This work demonstrates that allosteric switches can be designed into complex molecules with sufficient structural knowledge.

40. Xia Y, DiPrimio N, Keppel TR, Vo B, Fraser K, Battaile KP, Egan C, Bystroff C, Lovell S, Weis DD, et al.: **The designability of protein switches by chemical rescue of structure: mechanisms of inactivation and reactivation**. *J Am Chem Soc* 2013, **135**:18840–18849.

41. Schreiber SL: **The Rise of Molecular Glues**. *Cell* 2021, **184**:3–9.

42. Stanton BZ, Chory EJ, Crabtree GR: **Chemically induced proximity in biology and medicine**. *Science* 2018, **359**.

43. Kang S, Davidsen K, Gomez-Castillo L, Jiang H, Fu X, Li Z, Liang Y, Jahn M, Moussa M, Dimaio F, et al.: **COMBINES-CID: An Efficient Method for de Novo Engineering of Highly Specific Chemically Induced Protein Dimerization Systems**. *J Am Chem Soc* 2019, **141**:10948–10952.

44• Cao S, Kang S, Mao H, Yao J, Gu L, Zheng N: **Defining molecular glues with a dual-nanobody cannabidiol sensor**. *Nat Commun* 2022, **13**:815.

•Reports the crystal structure of molecular glue CID interface for cannabidiol (CDB),

elegantly designed from engineered nanobodies. This work quantitatively compares their CDB sensor to auxin and immune-modulatory imide drug binding to provide useful insight into molecular glue binding dynamics with and without ligand.

45. Guo Z, Smutok O, Johnston WA, Walden P, Ungerer JPJ, Peat TS, Newman J, Parker J, Nebl T, Hepburn C, et al.: **Design of a methotrexate-controlled chemical dimerization system and its use in bio-electronic devices**. *Nat Commun* 2021, **12**:7137.

46. Duong MLT, Collinson-Pautz MR, Morschl E, Lu A, Szymanski SP, Zhang M, Brandt ME, Chang WC, Sharp KL, Toler SM, et al.: **Two-Dimensional Regulation of CAR-T Cell Therapy with Orthogonal Switches**. *Molecular therapy oncolytics* 2018, **12**:124–137.

47• Giordano-Attianese G, Gainza P, Gray-Gaillard E, Cribioli E, Shui S, Kim S, Kwak MJ, Vollers S, Corria Osorio ADJ, Reichenbach P, et al.: **A computationally designed chimeric antigen receptor provides a small-molecule safety switch for T-cell therapy**. *Nature Biotechnology 2020 38:4* 2020, **38**:426–432.

•This innovative CAR-T control mechanism combines computational design of a novel ligand binding interface with a unique CAR-T control system architecture.

48• Steiner PJ, Swift SD, Bedewitz M, Wheeldon I, Cutler SR, Nusinow DA, Whitehead TA: **A Closed Form Model for Molecular Ratchet-Type Chemically Induced Dimerization Modules**. *Biochemistry* 2022, doi:10.1021/acs.biochem.2c00172.

•This paper develops an equilibrium mathematical model for molecular ratchet-type CID modules. Transfer functions for a variety of signal transduction mechanisms are derived and compared with experimental results.

49. Melcher K, Ng LM, Zhou XE, Soon FF, Xu Y, Suino-Powell KM, Park SY, Weiner JJ, Fujii H, Chinnusamy V, et al.: **A gate–latch–lock mechanism for hormone signalling by abscisic acid receptors**. *Nature* 2009, **462**:602–608.

50. Ziegler MJ, Yserentant K, Dunsing V, Middel V, Gralak AJ, Pakari K, Bargstedt J, Kern C, Petrich A, Chiantia S, et al.: **Mandipropamid as a chemical inducer of proximity for in vivo applications**. *Nature Chemical Biology 2021 18:1* 2021, **18**:64–69.

51•• Beltrán J, Steiner PJ, Bedewitz M, Wei S, Peterson FC, Li Z, Hughes BE, Hartley Z, Robertson NR, Medina-Cucurella AV, et al.: **Rapid biosensor development using plant hormone receptors as reprogrammable scaffolds**. *Nat Biotechnol* 2022,

doi:10.1038/s41587-022-01364-5.

••This work demonstrates that the plant hormone receptor PYR1 can be engineered to accommodate a diverse set of ligands. Furthermore, this work demonstrates the portability of molecular ratchet-type sensors by coupling the CID system to both *in vitro* and *in vivo* sensor outputs with pM to μM sensitivities.

52. Tague EP, Dotson HL, Tunney SN, Sloas DC, Ngo JT: **Chemogenetic control of gene expression and cell signaling with antiviral drugs**. *Nat Methods* 2018, **15**:519.

53. Foight GW, Wang Z, Wei CT, Greisen P, Warner KM, Cunningham-Bryant D, Park K, Brunette TJ, Sheffler W, Baker D, et al.: **Multi-input chemical control of protein dimerization for programming graded cellular responses**. *Nat Biotechnol* 2019, **37**:1209–1216.

54. Juillerat A, Tkach D, Busser BW, Temburni S, Valton J, Duclert A, Poirot L, Depil S, Duchateau P: **Modulation of chimeric antigen receptor surface expression by a small molecule switch**. *BMC Biotechnol* 2019, **19**:44.

55. Sahillioglu AC, Toebes M, Apriamashvili G, Gomez R, Schumacher TN: **Crash-it switch enables reversible and dose-dependent control of tcr and car t-cell function**. *Cancer Immunology Research* 2021, **9**:999–1007.

56. Richman SA, Wang LC, Moon EK, Khire UR, Albelda SM, Milone MC: **Ligand-Induced Degradation of a CAR Permits Reversible Remote Control of CAR T Cell Activity In Vitro and In Vivo**. *Mol Ther* 2020, **28**:1600–1613.

57. Jan M, Scarfò I, Larson RC, Walker A, Schmidts A, Guirguis AA, Gasser JA, Słabicki M, Bouffard AA, Castano AP, et al.: **Reversible ON- and OFF-switch chimeric antigen receptors controlled by lenalidomide**. *Sci Transl Med* 2021, **13**.

58. Douglass EF Jr, Miller CJ, Sparer G, Shapiro H, Spiegel DA: **A comprehensive mathematical model for three-body binding equilibria**. *J Am Chem Soc* 2013, **135**:6092–6099.

59. Dagliyan O, Krokhotin A, Ozkan-Dagliyan I, Deiters A, Der CJ, Hahn KM, Dokholyan NV: **Computational design of chemogenetic and optogenetic split proteins**. *Nat Commun* 2018, **9**:4042.

60••Dolberg TB, Meger AT, Boucher JD, Corcoran WK, Schauer EE, Prybutok AN, Raman S, Leonard JN: **Computation-guided optimization of split protein systems**. *Nature Chemical Biology 2021 17:5* 2021, **17**:531–539.

••This paper addresses several problems in engineering split proteins for biosensor signal transduction systems, notably succeeding in reducing background signal for a split TEV protease. The open source computational pipeline can be applied to create split proteins for a variety of signal transduction applications.

FIGURES

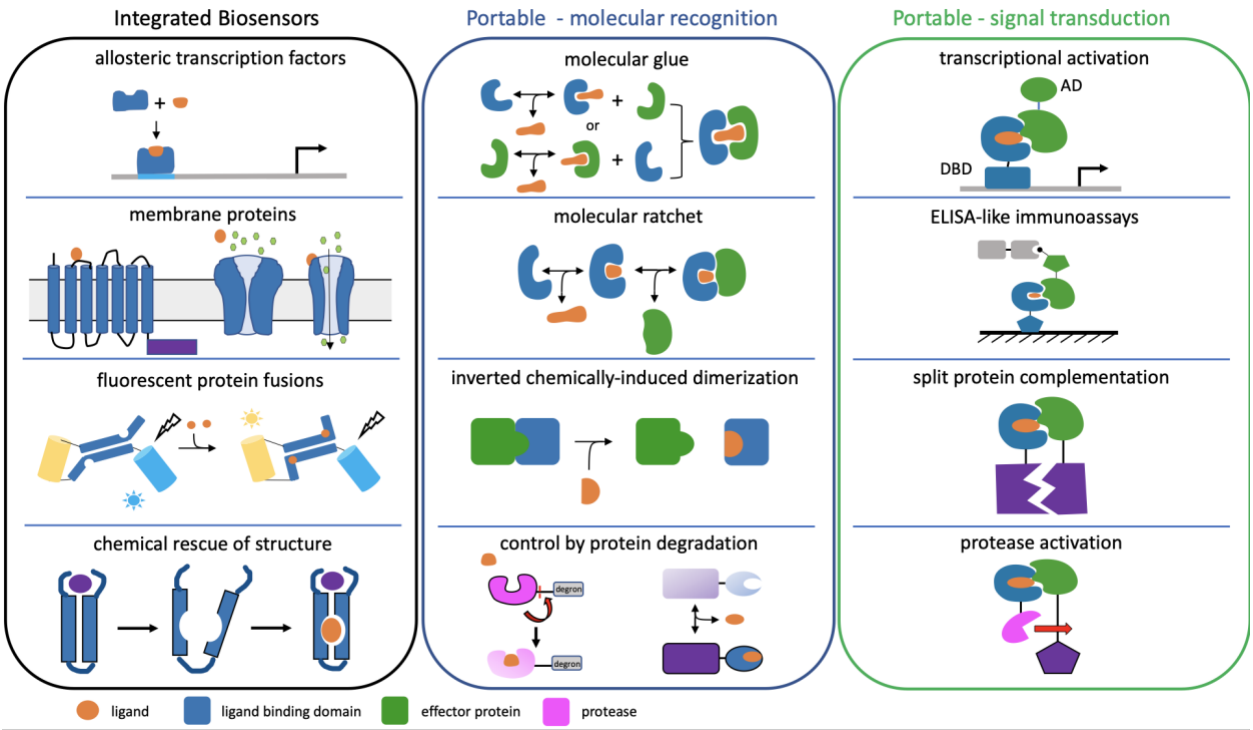

Figure 1. Molecular architecture and examples of integrated and portable biosensors. Examples of integrated biosensors include allosteric transcription factors [3,10,21,23,24,25], membrane proteins like GPCRs or ion channels [16,32,33], fluorescent protein fusions [37], and chemical rescue of structure [39]. Examples of portable molecular recognition mechanisms include molecular glues [9,43,45], molecular ratchets [2,51], inverted chemically-induced dimerization [47], and control by protein degradation [8].

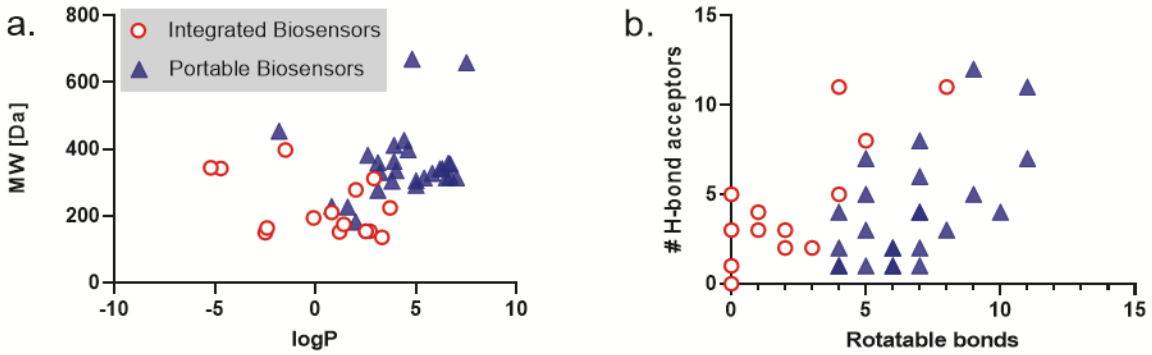

Figure 2. Accessible chemical space of existing engineered biosensors. Ligands for integrated and portable engineered biosensors occupy different portions of the accessible chemical space, shown by **a.** molecular weight vs octanol-water partition coefficient logP and **b.** number of hydrogen bond acceptors vs number of rotatable bonds. All plotted data is listed in table 1.

Table 1. List of sensors reviewed here that have been engineered for binding to a novel ligand.

Category	Ligand	Sensor Type	MW (Da)	logP	H-bond acceptors	Rotatable bonds	Ref
Integrated	Vanillin	Allosteric transcription factor	152.15	1.2	3	2	10
Integrated	D-Arabinose	Allosteric transcription factor	150.13	-2.5	5	0	21
Integrated	Fucose	Allosteric transcription factor	164.16	-2.4	5	4	23
	Gentiobiose		342.3	-4.7	11	4	
	Lactitol		344.31	-5.2	11	8	
	Sucralose		397.6	-1.5	8	5	
Integrated	Caffeine	Allosteric transcription factor	194.19	-0.1	3	0	24
Integrated	Borneol	Allosteric transcription factor	154.25	2.7	1	0	25
	Fenchol		154.25	2.5	1	0	
	Eucalyptol		154.25	2.5	1	0	
	Camphene		136.23	3.3	0	0	
Integrated	Varenicline	Membrane protein	211.26	0.8	3	0	16
Integrated	Olanzapine	Membrane protein	312.4	2.9	4	1	32
Integrated	DREADD agonist 21	Membrane protein	278.35	2	3	1	33
Integrated	Auxin	Fluorescent protein fusions	175.18	1.4	2	2	37
Integrated	6-(benzyloxy)indazole	Chemical rescue of structure	224.26	3.7	2	3	39
Portable	Farnesyl pyrophosphate	Molecular glue	382.33	2.6	3	11	9
Portable	Cannabidiol	Molecular glue	314.5	6.5	2	6	43
Portable	Methotrexate	Molecular glue	454.4	-1.8	5	9	45
Portable	Mandipropamid	Molecular Ratchet	411.9	3.9	1	10	2

Portable	Cannabidiloic acid	Molecular Ratchet	358.5	6.6	3	7	51
	delta9-THC		314.5	7	1	4	
	(+/-)CP 47,497		318.5	6.7	2	7	
	AB-PINACA		330.4	3.2	2	8	
	4F-MDMB-BUTINACA		363.4	3.9	1	9	
	JWH-167		305.4	5	0	7	
	JWH-030		291.4	5	0	6	
	JWH-016		341.4	6.2	0	5	
	JWH-007		355.5	6.7	0	6	
	JWH-018		341.4	6.3	0	6	
	JWH-072		313.4	5.4	0	4	
	JWH-015		327.4	5.8	0	4	
	JWH-193		398.5	4.6	0	5	
	(+/-)WIN-55212-2		426.5	4.4	0	4	
	Diazinon		304.35	3.8	0	7	
	Pirimiphos		277.28	3.1	2	5	
	Dimethoate		229.3	0.8	1	5	
	Chlorfenvinphos		359.6	3.1	0	7	
Portable	A1331852	Inverted chemical induced dimerization	658.8	7.5	2	7	47
	A1155463		669.8	4.8	2	11	

Portable	Fentanyl	Control by protein degradation	336.5	4	0	6	8