A general strategy to construct small molecule biosensors

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:40046433

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
A general strategy to construct small molecule biosensors

A dissertation presented

by

Justin Feng

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biological and Biomedical Sciences

Harvard University

Cambridge, Massachusetts

April 2017
A general strategy to construct small molecule biosensors

Abstract

Biosensors for small molecules can be used in applications that range from metabolic engineering to orthogonal control of transcription. Despite their broad utility, it remains a longstanding challenge to rapidly create new small molecule biosensors that are robust and specific. Here, we produce biosensors based on a ligand-binding domain (LBD) by using a method that, in principle, can be applied to any target molecule. The LBD may be fused to any protein-based reporter such as a fluorescent protein or a transcriptional activator and is destabilized by mutation such that the fusion accumulates only in cells containing the target ligand.

We illustrate the power of this method by developing biosensors for digoxin and progesterone in *Saccharomyces cerevisiae* and demonstrating portability to both mammalian cell culture and multicellular plants. Addition of ligand to yeast, mammalian, or plant cells expressing a biosensor activates transcription with a dynamic range of up to ~100-fold. We use the biosensors to improve the biotransformation of pregnenolone to progesterone in yeast by creating a novel selection to enable directed evolution of a metabolic pathway and to regulate Cas9 activity in mammalian cells.

We further propose to expand the biosensor development and deployment to new systems such as prokaryotes and cell-free lysates. Lastly, we propose an antibody
fragment based scaffold that can have its target response readily switched with minimal effort to arbitrary molecules while retaining high levels of affinity and specificity.
Table of contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title page</td>
<td>i</td>
</tr>
<tr>
<td>Copyright</td>
<td>ii</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Table of contents</td>
<td>v</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>vi</td>
</tr>
<tr>
<td>Chapter 1 – Background and motivation</td>
<td>1</td>
</tr>
<tr>
<td>Screens for metabolic engineering</td>
<td>1</td>
</tr>
<tr>
<td>In vivo biosensors</td>
<td>3</td>
</tr>
<tr>
<td>Conditionally destabilized proteins</td>
<td>5</td>
</tr>
<tr>
<td>Computationally designed small molecule binding proteins</td>
<td>5</td>
</tr>
<tr>
<td>Chapter 2 – Construction and characterization of conditionally destabilized LBDs</td>
<td>7</td>
</tr>
<tr>
<td>Attributions</td>
<td>7</td>
</tr>
<tr>
<td>Fluorescent biosensors built from engineered LBDs</td>
<td>7</td>
</tr>
<tr>
<td>TF-biosensors amplify ligand-dependent responses</td>
<td>12</td>
</tr>
<tr>
<td>TF-biosensors are tunable and modular</td>
<td>19</td>
</tr>
<tr>
<td>Chapter 3 – Application of conditionally destabilized LBD biosensors</td>
<td>22</td>
</tr>
<tr>
<td>Attributions</td>
<td>22</td>
</tr>
<tr>
<td>Progesterone and hydrocortisone production in Saccharomyces cerevisiae</td>
<td>22</td>
</tr>
<tr>
<td>TF-biosensors enable a selection in yeast to improve the bioproduction of a small molecule</td>
<td>23</td>
</tr>
<tr>
<td>Yeast-derived biosensors port directly to mammalian cells and can be used to tightly regulate CRISPR/Cas9 genome editing</td>
<td>30</td>
</tr>
<tr>
<td>Environmental detection in the plant Arabidopsis thaliana</td>
<td>33</td>
</tr>
<tr>
<td>Discussion</td>
<td>38</td>
</tr>
<tr>
<td>Chapter 4 – Future directions</td>
<td>44</td>
</tr>
<tr>
<td>Motivation</td>
<td>44</td>
</tr>
<tr>
<td>Bacterial hosts for biosensor development and deployment</td>
<td>44</td>
</tr>
<tr>
<td>Cell-free biosensors</td>
<td>49</td>
</tr>
<tr>
<td>Antibody-based scaffolds for generalizable biosensor construction</td>
<td>51</td>
</tr>
<tr>
<td>Chapter 4 – Final thoughts</td>
<td>56</td>
</tr>
</tbody>
</table>
Chapter 5 – Materials and Methods ................................................................. 59

References.................................................................................................................. 78
Acknowledgements

First and foremost, I would like to thank my parents Liching and Bill. They have supported me in countless ways, always encouraged me to do my best, and have sacrificed much to afford me every opportunity they could provide. Without them I would be nowhere close to the person I am today.

I would like to thank the doctors, nurses, and staff at the BIDMC Division of Hematology-Oncology, in particular Dr. Robin Joyce and Dr. Alexandra Bailey.

I’ve had the great fortune of working with scientific collaborators and mentors that have almost exclusively been both talented and kind, with there being far too many to name. Some that I am indebted to include, but are not limited to:

- Dave Steakley and Jasper Rine for teaching me scientific discipline and rigor and the power of yeast genetics while I was an undergraduate at UC Berkeley
- Jeff Kim, Oliver Liu, Quinn Mitrovitch, and Chris Paddon for helping me explore and dream about the potential of synthetic biology during my time at Amyris
- George Church for encouraging me to dream bigger while providing an amazing lab with many wonderful people to collaborate with, learn from, and spend time with
- Many other current and former labmates: Raj Chari, James DiCarlo, Alex Chavez, Dan Goodman, Noah Taylor, Gleb Kuznetsov, Alex Garruss, Chris Gregg, Marc Lajoie, Matthieu Landon, and Michael Napolitano
- Dan Mandell, with whom I have worked with for the entirety of my graduate career, for being an incredible mentor, collaborator, and friend.

Lastly, I’d like to thank my friends who have been with me along the way – Gray Chadwick, Ammar Rizvi, Alix Chan, Su Vora, Sami Farhi, Nikit Patel, and many more. Science and graduate school have been a defining part of my life, but it’s the time that I’ve spent with the amazing people and all our great moments together that I’ll truly remember!
Chapter 1 – Background and motivation

Biosensors capable of sensing and responding to small molecules \textit{in vivo} have wide-ranging applications in biological research and biotechnology, including metabolic pathway regulation (Zhang et al., 2012), biosynthetic pathway optimization (Raman et al., 2014; Tang and Cirino, 2011), metabolite concentration measurement and imaging (Paige et al., 2012), environmental toxin detection (Gil et al., 2000), and small molecule-triggered therapeutic response (Ye et al., 2012). Despite such broad utility, no single strategy for \textit{in vivo} biosensor construction has proven sufficiently generalizable to gain widespread use. Current methods typically couple binding to system-specific signals, and use natural protein (Tang et al., 2013) or nucleic acid aptamer (Yang et al., 2013) binding domains which limits the scope of small molecules that can be detected. A general solution to small molecule biosensing should be adaptable to arbitrary small molecules and responses.

Screens for metabolic engineering

Production of useful chemical compounds by recombinant microbes dates back to the 1979 introduction of human insulin-producing \textit{E. coli} (Goeddel et al., 1979). Since then, metabolic engineering has emerged as an intensely studied and growing field. With the advancing capabilities to sequence and synthesize DNA (Buermans and den Dunnen, 2014; Kosuri and Church, 2014), our ability to discover, design, and create new biological phenotypes has increased drastically as well. The range of molecules
produced in microorganisms has vastly expanded, and this type of biological manufacturing holds great promise for enabling large-scale production of industrially relevant compounds ranging from fuels to pharmaceuticals to cosmetics.

Despite this potential, there have been few instances of successfully using metabolic engineering to produce small molecules at a scale and price competitive with alternative processes. One of the most notable achievements is the recent development of an *Saccharomyces cerevisiae* strain capable of producing the anti-malarial drug precursor artemisinic acid with a titer of 25 g/L (Paddon et al., 2013). While this is a landmark achievement in synthetic biology, it took the better part of a decade, tens of millions of dollars, and an extraordinary amount of man hours to develop. The enormity of this effort demonstrates the need for further technological development to realize the promise of metabolic engineering.

A major hurdle to increasing titers to industrially relevant levels is the inability to rapidly screen and select for molecules of interest. Production strains are often optimized by making large libraries of unique variants through random or rationally designed changes to regulatory and coding regions of metabolic pathways. It is very difficult to predict which sets of changes will be beneficial rather than neutral or detrimental. Because of this, each clone needs to be assayed for biosynthetic efficiency.

Currently, the standard methods for assaying small molecule production are gas and liquid chromatography and mass spectrometry (Choi and Antoniewicz, 2011).
While these methods are both quantitative and are able to detect a wide range of molecules, they exhibit very low throughput, with a maximum sampling size of $\sim 10^4$ (Dietrich et al., 2010). This means that only a small number of mutants can be screened, leaving a large portion of the sequence landscape unexplored.

There has been a great deal of work on creating fluorometric assays for higher throughput screens. One method is the addition of an exogenous chemical to the production culture that responds to the molecules of interest. An example is the Nile red dye, which fluoresces in the presence of hydrophobic molecules (Greenspan et al., 1985) and may allow for colorimetric detection of chemicals with long hydrocarbon chains such as terpenes. These assays usually exhibit low sensitivity but can be screened in microtiter plates with a throughput of $\sim 10^5$ samples (Dietrich et al., 2010).

**In vivo biosensors**

In addition to chemically induced fluorometric assays, a number of *in vivo* biosensor strategies have been developed for various applications (Bernhardt et al., 2007; Lütke-Eversloh and Stephanopoulos, 2007; Mohn et al., 2006). These typically utilize a specific detectable intermediate, such as lycopene (Alper et al., 2005), or develop a biosensor specifically for the molecule of interest. Such biosensors have been made using RNA aptamers and proteins with allosteric responses (Michener et al., 2012). *In vivo* biosensors present many advantages, including the ability to avoid costly extraction of the molecule of interest. Depending on the output response, they may also
be used in fluorescently activated cell sorting (FACS) or in growth complementation selections, both of which provide many orders of magnitude higher throughput than microtiter plates.

To date, *in vivo* biosensors have been developed and applied in several contexts, but no single strategy has achieved widespread use (Dietrich et al., 2010; Michener et al., 2012). The dominant problem is that for each application, a novel biosensor must be developed that can link detection of the target molecule to an assayable response.

Methods for finding RNA aptamers *in vitro* have been used for many years (Ellington and Szostak, 1990). Despite this, moving RNA aptamer devices into cells has proven difficult in some cases, and RNA based responses are limited (Michener et al., 2012). Protein-based biosensors are genetically encodable and can be linked to a wide array of cellular activities. One especially appealing application would be to link the sensor response to growth complementation, either through antibiotic resistance or through remediation of an auxotrophy. Creating a synthetic selection in this way would allow for the screening of vastly larger libraries (Rogers et al., 2016).

The ideal biosensor development platform would have three main qualities: the ability to be robustly generalizable to arbitrary targets, the ability to easily tune the biosensor response to different molecule concentrations, and the ability to be linked biological responses to enable high throughput screens and selections.
**Conditionally destabilized proteins**

A promising approach to biosensor design in eukaryotes uses conditionally stable ligand-binding domains (Banaszynski et al., 2006). In the absence of a cognate ligand, these proteins are degraded by the ubiquitin proteasome system (Egeler et al., 2011). Binding of ligand stabilizes the LBD and prevents degradation. Fusing the destabilized LBD to a protein of interest, such as an enzyme, fluorescent protein, or transcription factor, renders the fusion conditionally stable and allows for sensor response at the protein or DNA level. Naturally occurring LBDs can be engineered for conditional stabilization (Banaszynski et al., 2006; Iwamoto et al., 2010; Miyazaki et al., 2012), making it possible in principle to convert any LBD into a biosensor for a target ligand.

**Computationally designed small molecule binding proteins**

Designed LBDs would be useful when a natural binder possessing high affinity and specificity does not exist or is impractical due to protein instability, subcellular localization, or interference with cellular processes. The Baker lab at the University of Washington recently described a general framework for computational design of LBDs with high affinity and substrate selectivity (Tinberg et al., 2013). First, the authors modeled an optimized binding pocket for the steroid digoxigenin (DIG). They then scanned Protein Data Bank (PDB) for protein scaffolds that could accommodate the binding pocket geometries, finding that PDB code 1Z1S, a protein of unknown function,
served as a suitable scaffold. Their in silico design for a DIG binding protein yielded a protein with ~10 μM affinity. Two rounds of mutagenesis further improved affinity to 653 pM. In addition to creating the high affinity DIG binding protein (DIG10.3), the authors showed that key mutations in the binding pocket can switch binding specificity from DIG to other steroid compounds. For example, mutations Y34F/Y99F/Y101F improved affinity to <17 nM from 243 nM for progesterone and while weakening affinity for DIG to ~580 nM from 653 pM. These results show great promise for developing new, physiologically orthogonal binding proteins for ligands that do not have a suitable natural partner.

Conditionally destabilized biosensors

We decided to use conditionally destabilized proteins as a platform for developing new small molecule biosensors due to the perceived generality and modularity of this strategy. For our initial test, we chose to use the DIG10.3 binding protein to create steroid biosensors. We chose DIG10.3 for its high binding affinity, ability to rationally alter target binding specificities, and crystallography data available to conduct post hoc analyses on any discovered biosensor mutations.
Chapter 2 – Construction and characterization of conditionally destabilized LBDs

Attributions

The majority of this chapter was previously published in the journal *eLife* (Feng et al., 2015) and does not require explicit permission to reproduce here. The work was done in collaboration with the Baker Lab at University of Washington, the Fields Lab and the University of Washington, and the Medford Lab at Colorado State University.

**Fluorescent biosensors built from engineered LBDs**

LBDs intended for biosensor development should recognize their targets with high affinity and specificity. To develop conditionally destabilized LBD biosensors (Figure 2.1a), we began with the computationally-designed binding domain DIG10.3 (Tinberg et al., 2013), hereafter DIG0, which binds the plant steroid glycoside digoxin and its aglycone digoxigenin with picomolar affinities. Introduction of three rationally-designed binding site mutations into DIG0 resulted in a progesterone binder (PRO0) with nanomolar affinity (Tinberg et al., 2013). We constructed genetic fusions of DIG0 and PRO0 to a yeast-enhanced GFP (yEGFP, LBD-biosensors DIG0-GFP and PRO0-GFP) and constitutively expressed them in *S. cerevisiae*. The fusions showed little change in fluorescence in response to digoxin or progesterone, respectively (Figure 2.1b,c, Figure 2.2). Work by Wandless and co-workers has shown that mutagenesis of LBDs can be used to identify variants that are stable only in the presence of a target ligand (Banaszynski et al., 2006). We randomly mutagenized the LBDs of DIG0-GFP and PRO0-GFP...
GFP by error-prone PCR and subjected libraries of $10^5$ integrants to multiple rounds of FACS, sorting alternately for high fluorescence in the presence of the ligand and low fluorescence in its absence. We isolated LBD variants having greater than 5-fold activation by cognate ligand (Figure 2.1b,c, Figure 2.2). By making additional variants that contain only one of the up to four mutations found in the progesterone biosensors, we showed that some mutations are additive, while others predominately contribute to sensitivity (Figure 2.3a). Many of the conditionally-destabilizing mutations identified in DIG$_0$ involve residues that participate in key dimer interface interactions (Figure 2.1d). The conditionally-destabilizing mutations of PRO$_0$ are located throughout the protein (Figure 2.3b-d); the DIG$_0$ interface mutations also rendered PRO$_0$-GFP conditionally stable on binding progesterone (Figure 2.3e).
Figure 2.1. A general method for construction of biosensors for small molecules. (a) Modular biosensor construction from a conditionally destabilized LBD and a genetically fused reporter. The reporter is degraded in the absence but not in the presence of the target small molecule. (b) yEGFP fluorescence of digoxin LBD-GFP biosensors upon addition of 250 µM digoxin or DMSO vehicle. (c) yEGFP fluorescence of progesterone LBD-GFP biosensors upon addition of 50 µM progesterone or DMSO vehicle. (d) Positions of conditionally destabilizing mutations of DIG0 mapped to the crystal structure of the digoxin LBD (PDB ID 4J9A). Residues are shown as colored spheres and key interactions highlighted in insets. In b-c, fold activation is shown above brackets, (-) indicates cells lacking biosensor constructs, and error bars represent the standard error of the mean (s.e.m.) of three biological replicates.
Figure 2.2. Population responses to cognate ligand for cells bearing LBD-biosensors. 
(a) yEGFP fluorescence of DIG\textsubscript{0}-GFP upon addition of 250 µM digoxin or DMSO vehicle. 
(b) yEGFP fluorescence of DIG\textsubscript{1}-GFP upon addition of 250 µM digoxin or DMSO vehicle. 
(c) yEGFP fluorescence of PRO\textsubscript{0}-GFP upon addition of 50 µM progesterone or DMSO vehicle. 
(d) yEGFP fluorescence of PRO\textsubscript{1}-GFP upon addition of 50 µM progesterone or DMSO vehicle. Data for each condition are presented for a single representative biological replicate with 20,000 events measured by flow cytometry.
Figure 2.3. Characterization of mutations conferring progesterone-dependent stability. (a) Single-mutant deconvolutions of mutations conferring progesterone sensitivity. The parental biosensor appears in the leftmost column of each panel. (b-d) Positions of mutations in PRO₁ (b), PRO₂ (c), and PRO₃ (d) are mapped to the crystal structure of the digoxin LBD (PDB ID 4J9A) and are shown in colored spheres. (e) Fold activation of PRO₀-GFP biosensors with digoxin biosensor mutations upon addition of 50 µM progesterone. In a and e, error bars represent s.e.m. of three biological replicates.
**TF-biosensors amplify ligand-dependent responses**

To improve the dynamic range and utility of the biosensors, we built conditionally-stable LBD-transcription factor fusions (TF-biosensors) by placing an LBD between an N-terminal DNA-binding domain (DBD) and a C-terminal transcriptional activation domain (TAD, Figure 2.4a). The use of TFs serves to amplify biosensor response and allows for ligand-dependent control of gene expression (Beerli et al., 2000; Louvion et al., 1993; Shoulders et al., 2013). Our initial constructs used the DBD of Gal4, the destabilized LBD mutant DIG1 (E83V), and either VP16 or VP64 as a TAD to drive the expression of yEGFP under control of a GAL1 promoter. The dynamic range of TF-biosensor activity was maximal when the biosensor was expressed using a weak promoter and weak activation domain, because of lower yEGFP expression in the absence of ligand (Figure 2.5a,b).

We chose Gal4-DIG1-VP16 (hereafter G-DIG1-V) for further TF-biosensor development because it has both a large dynamic range and maximal activation by ligand. A FACS-based screen of an error-prone PCR library of G-DIG0-V, G-DIG1-V, and G-DIG2-V variants identified mutations L77F and R60S in the Gal4 dimer interface (hereafter G.L77F, G.R60S) that further increased TF-biosensor response by lowering background activity in the absence of ligand (Figure 2.4b and Figure 2.5c). Although these Gal4 mutations were identified by screening libraries of digoxin-dependent TF-biosensors, they also increased progesterone-dependent activation of the G-PRO-V
series of biosensors, indicating a shared mechanism of conditional stability in both systems (Figure 2.5d). Combining mutations in Gal4 and DIG0 or PRO0 led to activations of up to 60-fold by cognate ligand, a ten-fold improvement over the most responsive LBD-biosensors (Figure 2.4c,d, Figure 2.6a) and a dynamic range that has been challenging to achieve with stability-based biosensors in yeast (Rakhit et al., 2011). The TF-biosensors were also rapidly activated, showing a five-fold increase in signal after one hour of incubation with ligand and full activation after ~14 hours (Figure 2.4e,f, Figure 2.6). In contrast to the LBD-biosensors, the TF-biosensors exhibited a broad range of fluorescence levels across single cells, as well as a population of nonfluorescent cells in the presence of ligand (Figure 2.6). We used FACS to isolate cells from the nonfluorescent population and found those cells to be inviable, possibly indicating plasmid loss or toxicity from biosensor activation.

Upon withdrawal of ligand, strains expressing TF-biosensors rapidly exhibited reduction in signal, reaching half of their maximum yEGFP fluorescence after approximately 5 hours and nearly undetectable fluorescence after 10-15 hours (Figure 2.4g,h). The response of the TF-biosensors to the withdrawal of ligand is likely much faster than observed by fluorescence, as the reduction in fluorescence signal is dependent on both the degradation of the TF-biosensors as well as the degradation and dilution of previously expressed yEGFP.
**Figure 2.4. Ligand-dependent transcriptional activation.** (a) TF-biosensor construction from a conditionally destabilized LBD, a DNA binding domain, and a transcriptional activation (TAD) domain. (b) Positions of conditionally destabilizing mutations of Gal4 mapped to a computational model of Gal4-DIG homodimer. Residues are shown as colored spheres and key interactions are highlighted in insets. The TAD is not shown. (c) Concentration dependence of response to digoxin for digoxin TF-biosensors driving yEGFP expression. (d) Concentration dependence of response to progesterone for progesterone TF-biosensors driving yEGFP expression. (e) Time dependence of response to 250 µM digoxin for digoxin TF-biosensors. (f) Time dependence of response to 50 µM progesterone for progesterone TF-biosensors. (g) Time-dependent response to withdrawal of 250 µM digoxin for digoxin TF-biosensors. (h) Time-dependent response to withdrawal of 50 µM progesterone for progesterone TF-biosensors. In c-f, (-) indicates cells lacking biosensor plasmids and error bars represent s.e.m. of three biological replicates. Marker symbols in e and g are the same as in c. Marker symbols in f and h are the same as in d.
Figure 2.4 (Continued)
**Figure 2.5. Improvements to TF-biosensor response.** Digoxin-dependent expression of yEGFP by G-DIG₁-V TF-biosensors either (a) containing VP64 or VP16 as the TAD and expressed from a CYC₁ promoter or (b) containing a VP16 TAD and expressed from a CYC₁, ADH₁, or TEF₁ promoter. (c) Individual mutations identified in a FACS analysis of an error-prone PCR library of G-DIG-V biosensors were tested for their effect on biosensor function using digoxigenin. Transformants were analyzed in an yEGFP yeast reporter strain containing a deletion of PDR₅ (PyE₁₄). Improvements in fold activation relative to parental sequences were localized to mutations in Gal4. (d) R₆₀S and L₇₇F mutations found in Gal4 were introduced into G-DIG₁-V, G-DIG₂-V, and G-PRO₁-V. In each case, the Gal4 mutations had the effect of lowering the amount of luciferase activity in the absence of the relevant ligand. In a-d, error bars represent s.e.m. of three biological replicates.
**Figure 2.5 (Continued)**

(a) Graph showing mean yEFGP fluorescence for different TAD combinations: VP64 and VP16. The bars indicate the activation domain strength with DMSO and 100 µM digoxin.

(b) Graph showing mean yEFGP fluorescence for different promoters: CYC1p, ADH1p, and TEF1p. The bars indicate the promoter strength with DMSO and 100 µM digoxin.

(c) Table listing various mutations and their effects on mean yEFGP fluorescence. The table compares DMSO and 50 µM digoxigenin conditions.

(d) Graph showing mean luciferase activity for different combinations of LBD mutations and VP16 mutations. The bars indicate the activity with DMSO, 50 µM digoxin, and 1 µM progesterone.
Figure 2.6. Population responses to cognate ligand for cells bearing TF-biosensors. (a) yEGFP fluorescence from G-DIG₀-V activation upon addition of digoxin or DMSO vehicle (left) and yEGFP fluorescence from G₁₇₇₋DIG₁₋V activation upon addition of digoxin or DMSO vehicle (right). (b) yEGFP fluorescence from G-PRO₀-V activation upon addition of progesterone or DMSO vehicle (left) and yEGFP fluorescence from G₁₇₇₋PRO₁₋V activation upon addition of progesterone or DMSO vehicle (right). Data for each condition are presented for a single representative biological replicate with 20,000 events measured by flow cytometry.
**TF-biosensors are tunable and modular**

An attractive feature of the TF-biosensors is that the constituent parts – the DBD/promoter pair, the LBD, the TAD, the reporter, and the yeast strain – are modular, such that the system can be modified for additional applications. To demonstrate tunability, we replaced the DBD of G-DIGi-V with the bacterial repressor LexA and replaced the Gal4 DNA-binding sites in the GAL1 promoter with those for LexA. LexA-based TF-biosensors with DIGi and a weak TAD (B42) showed a strong response to digoxin (nearly 40-fold) only when the promoter driving reporter expression contained LexA-binding sites (Figure 2.7a). These results demonstrate that the biosensors can function with different combinations of DBDs and TADs, which could produce diverse behaviors and permit their use in eukaryotes requiring different promoters. Furthermore, the reporter gene can be swapped with an auxotrophic marker gene to enable growth selections. The TF-biosensors drove expression of the HIS3 reporter most effectively when steroid was added to the growth media, as assessed by growth of a histidine auxotrophic strain in media lacking histidine (Figure 2.7b,d). Fusion of the Mata2 degron to the biosensor improved dynamic range by reducing growth of yeast in the absence of ligand. Finally, the yeast strain could be modified to improve biosensor sensitivity toward target ligands by deletion of the gene for a multidrug efflux pump (Ernst et al., 2005), thereby increasing ligand retention (Figure 2.7c-d).
Figure 2.7. Tuning TF-biosensors for different contexts. (a) The TAD and DBD of the TF-biosensor and the corresponding binding site for the DBD in the reporter promoter can be swapped for a different application. Expression of a plasmid-borne luciferase reporter was driven by TF-biosensors containing either a LexA or Gal4 DBD and either a VP16 or B42 TAD. Promoters for the reporter contained DNA-binding sites for either Gal4 or LexA. (b) TF-biosensors were transformed into the yeast strain PJ69-4a and tested for growth on –his minimal media containing 1mM 3-aminotriazole (3-AT) and the indicated steroid. To determine the effect of including an additional destabilization domain, the degron from Mata2 was cloned into one of four positions. (c) G-DIG1-V biosensor response to digoxigenin in yEGFP reporter strain PyE1 either with or without a deletion to the ORF of PDR5. (d) Ligand and TF-biosensor dependent growth on –his media in yeast strains containing deleted ORFs for efflux-related transcription factors (PDR1 and PDR3) or ABC transporter proteins (YOR1, PDR5, SNQ2). In a and c, error bars represent s.e.m. of three biological replicates.
Figure 2.7 (Continued)
Chapter 3 – Application of conditionally destabilized LBD biosensors

Attributions

The majority of this chapter was previously published in the journal *eLife* (Feng et al., 2015) and does not require explicit permission to reproduce here. The work was done in collaboration with the Baker Lab at University of Washington, the Fields Lab and the University of Washington, and the Medford Lab at Colorado State University.

**Progesterone and hydrocortisone production in *Saccharomyces cerevisiae***

Progesterone is a steroid hormone that regulates many high-level functions including reproduction and nervous system activity (Rupprecht, 2003). It is used medicinally to balance hormone levels, such as in patients undergoing menopausal symptoms. Complete synthesis of progesterone from a simple carbon source using a microbial host would possibly allow for more economical production of progesterone and other steroid based drugs.

There has been previous work on using *S. cerevisiae* as a host for the biosynthetic production of progesterone. In 1998, Duport and colleagues showed that the addition of five genes into yeast diverts flux through the endogenous yeast ergosterol pathway to create pregnenolone and progesterone (Duport et al., 1998).

This work was followed by a demonstration of the total biosynthesis of the progesterone derivative hydrocortisone in yeast by incorporating the remaining heterologous mammalian hydrocortisone pathway into *S. cerevisiae* (Szczebara et al.,
Hydrocortisone is a glucocorticoid that is used as an anti-inflammatory or immunosuppressive drug, as well as a starting point for synthesis of other steroid compounds. The final strain, UCYI, produces titers of 11.5 mg/L, far below typical optimized production titers of grams per liter. Metabolite analysis showed non-detectable amounts of the intermediate progesterone, indicating that there is a bottleneck at or upstream of progesterone production.

**TF-biosensors enable a selection in yeast to improve the bioproduction of a small molecule**

Improving bioproduction requires the ability to detect how modifications to the regulation and composition of production pathways affect product titers. Current product detection methods such as mass spectrometry or colorimetric assays are low-throughput and are not scalable or generalizable. LBD- and TF-biosensors could be coupled to fluorescent reporters to enable high throughput library screening or to selectable genes to permit rapid evolution of biosynthetic pathways (Chou and Keasling, 2013; Dietrich et al., 2010; Tang and Cirino, 2011). Yeast-based platforms have been developed for the biosynthesis of pharmaceutically relevant steroids, such as progesterone and hydrocortisone (Duport et al., 1998; Szczebara et al., 2003). A key step in the production of both steroids is the conversion of pregnenolone to progesterone by the enzyme 3β-hydroxysteroid dehydrogenase (3β-HSD). We aimed to use a progesterone biosensor to detect and improve this transformation. An important feature
of biosensors intended for pathway engineering is their ability to detect a product with minimal activation by substrate or other related chemicals. TF-biosensors built from PRO1 showed the greatest dynamic range and selectivity for progesterone over pregnenolone when driving yEGFP expression or when coupled to a HIS3 reporter assay (Figure 3.1a,b and Figure 3.2a-d). We observed moderate activation to high concentrations of digitoxigenin and nearly no activation to the other steroids. We investigated whether this sensor could be used to detect the in vivo conversion of pregnenolone to progesterone by episomally-expressed 3β-HSD (Figure 3.1c). Using Gl771-PRO1-V driving a yEGFP reporter, we could detect progesterone production, with biosensor response greatest when 3β-HSD was expressed from a high copy number plasmid and from a strong promoter (Figure 3.1d).

We then sought to use the biosensor to improve this enzymatic transformation. To select for improved progesterone production, we required a growth assay in which wild-type 3β-HSD could no longer complement histidine auxotrophy when the yeast were grown on plates supplemented with pregnenolone. To this end, the selection stringency was tuned by adding the His3 inhibitor 3-aminotriazole (Figure 3.2e). We mutagenized the 3β-HSD coding sequence using error-prone PCR and screened colonies that survived the HIS3 selection for their yEGFP activation by pregnenolone. By transforming evolved 3β-HSD mutations into a fresh host background, we showed that the mutations in the enzyme, and not off-target plasmid or host escape mutations,
were responsible for increased biosensor response (Figure 3.1e). Two of the mutants, 3\(\beta\)-HSD N139D and 3\(\beta\)-HSD F67Y, were assayed for progesterone production using gas chromatography and mass spectrometry and were found to produce two-fold more progesterone per OD than cells bearing the wild-type enzyme (Figure 3.1f).
Figure 3.1. Application of biosensors to metabolic engineering in yeast. (a) Fold activation of Gl77F-PRO1-V by a panel of steroids in yEGFP reporter strain PyE1. Progesterone became toxic at levels of 100 µM and above, leading to substantial cell death. Beta-estradiol and hydrocortisone were not soluble in yeast growth media at levels above 25 µM. (b) Growth of degron-G-PRO1-V in HIS3 reporter strain PJ69-4a is stimulated by progesterone but not pregnenolone. (c) Schematic for directed evolution of 3β-HSD using TF-biosensors for conversion of pregnenolone to progesterone. (d) Fold activation of Gl77F-PRO1-V by a panel of plasmids expressing wild-type 3β-HSD under varying promoter strengths in yEGFP reporter strain PyE1 when incubated in 50 µM pregnenolone. Data for plasmids containing CEN/ARS and 2 µ (2 micron) origins are shown. (e) Fold activation of Gl77F-PRO1-V by a panel of evolved 3β-HSD mutants expressed under the TDH3 promoter on a CEN/ARS plasmid and incubated in 50 µM pregnenolone. (f) Progesterone titer in 1 OD of cells produced by strains expressing 3β-HSD mutants. In a and d-f, data are presented as mean ± s.e.m. of three biological replicates. In d and e, (-) indicates cells lacking 3β-HSD. * indicates significance with a threshold of p < .05 using 2-tailed Student’s t-test.
Figure 3.1 (Continued)

(a) Graph of yEGFP fold activation with various compounds.

(b) Diagram showing ligand stabilization and TF biosensor.

(c) Chemical structures of pregnenolone and progesterone.

(d) Bar graph of yEGFP fold activation with different promoters.

(e) Bar graph of yEGFP fold activation with different mutants.

(f) Bar graph of progesterone concentration with different mutants.

Legend:
- 1 μM
- 10 μM
- 25 μM
- 100 μM
Figure 3.2. Specificity of PRO biosensors enables selection for auxotrophy complementation. Specificity for progesterone (PRO) over digoxigenin (DIG), digoxin (DGX), digitoxigenin (DTX), pregnenolone (PRE), β-estradiol (B-EST), and hydrocortisone (HYD) for (a) G-PRO0-V (b) G-PRO1-V (c) G-PRO2-V and (d) G-PRO3-V. Progesterone became toxic at levels of 100 µM and above, leading to substantial cell death. Beta-estradiol and hydrocortisone were not soluble in yeast growth media at levels above 25 µM. (e) Growth response of yeast strain PyE1 transformed with 3β-HSD on CEN/ARS plasmids under various promoters and plated on SC –his (and –ura –leu for plasmid maintenance) containing titrations of 3-AT and either 0.5% DMSO (upper panels) or 50 µM pregnenolone (lower panels). In a-d, error bars represent s.e.m. of three biological replicates.
Figure 3.2 (Continued)
Yeast-derived biosensors port directly to mammalian cells and can be used to tightly regulate CRISPR/Cas9 genome editing

Yeast is an attractive platform for engineering *in vivo* biosensors because of its rapid doubling time and tractable genetics. If yeast-derived biosensors function in more complex eukaryotes, the design-build-test cycle in those organisms could be rapidly accelerated. We first assessed the portability of yeast TF-biosensors to mammalian cells. Single constructs containing digoxin and progesterone TF-biosensors with the greatest dynamic ranges (without codon optimization) were stably integrated into human K562 cells using PiggyBac transposition. We characterized the dynamics of the TF-biosensors in human cells by dose response and time course assays similar to the yeast experiments (Figure 3.3a-d). As with yeast, the human cells demonstrated greater sensitivity to digoxin, with fluorescence activation increasing up to 100 nM of cognate ligand for digoxin biosensors and 1 mM for progesterone biosensors. We observed >100-fold activation for the most sensitive progesterone biosensor G_{L77F}-PRO_1-V. The increase in mammalian dynamic range over yeast may arise from more aggressive degradation of destabilized biosensors or greater accumulation of target-stabilized biosensors or reporters resulting from larger cell sizes and slower doubling times. The time course data show that fluorescence increased four-fold within four hours of target introduction and rose logarithmically for 24-48 hours.
We next assessed whether these biosensors could drive more complex mammalian phenotypes. The CRISPR/Cas9 system has proved to be an invaluable tool for genome editing (DiCarlo et al., 2013; Gratz et al., 2013; Hwang et al., 2013; Mali et al., 2013a). Despite the high programmability and specificity of Cas9-mediated gene editing achieved to date, unchecked Cas9 activity can lead to off-target mutations and cytotoxicity (Fu et al., 2013; Mali et al., 2013b; Pattanayak et al., 2013). Further, it may be desirable to tightly regulate Cas9 activity such that gene editing occurs only under defined conditions. To facilitate inducible gene editing, we fused human codon-optimized versions of the DIG3 and PRO1 LBDs to the N-terminus of Cas9 from *Streptococcus pyogenes*. We integrated this construct into a reporter cell line containing an EGFP variant with a premature stop codon that renders it non-functional. Upon separate stable integration of the DIG-Cas9 and PRO-Cas9 fusions, we transfected a guide RNA targeting the premature stop codon as well as a donor oligonucleotide containing the sequence to restore EGFP activity *via* homologous recombination. After a 48-hour incubation period, we observed an ~18-fold increase in GFP positive cells with digoxigenin relative to the mock control (Figure 3.3e).
biological replicates. Data are presented as mean ± s.e.m. of three
restored upon transfection of a guide RNA and donor oligonucleotide with matching
sequence in the presence of active Cas9. Figure 3.3. Activation of biosensors in mammalian cells and regulation of CRISPR/Cas9 activity. (a) Concentration dependence of response to digoxin for constructs containing digoxin TF-biosensors and Gal4 UAS-E1b-EGFP reporter individually integrated into K562 cells. G<sub>R60S,L77F</sub>-PRO<sub>V</sub> serves as a digoxin insensitive control. (b) Concentration dependence of response to progesterone for constructs containing progesterone TF-biosensors and Gal4 UAS-E1b-EGFP reporter individually integrated into K562 cells. G<sub>R60S</sub>-DIG<sub>V</sub> serves as a progesterone insensitive control. (c) Time dependence of response to 100 nM digoxin for constructs containing digoxin TF-biosensors and Gal4 UAS-E1b-EGFP reporter individually integrated into K562 cells. G<sub>R60S,L77F</sub>-PRO<sub>V</sub> serves as a digoxin insensitive control. (d) Time dependence of response to 25 µM progesterone for constructs containing progesterone TF-biosensors and Gal4 UAS-E1b-EGFP reporter individually integrated into K562 cells. G<sub>R60S</sub>-DIG<sub>V</sub> serves as a progesterone insensitive control. (e) DIG<sub>V</sub> and PRO<sub>V</sub> fused to the N-terminus of S. pyogenes Cas9 were integrated into a K562 cell line containing a broken EGFP. EGFP function is restored upon transfection of a guide RNA and donor oligonucleotide with matching sequence in the presence of active Cas9. Data are presented as mean ± s.e.m. of three biological replicates.
Environmental detection in the plant *Arabidopsis thaliana*

To assess generalizability of these sensors to multicellular organisms, we engineered G-DIG1-V to function as an environmental biosensor in plants. The DIG1 sequence was codon optimized for expression in *Arabidopsis thaliana*. We tested biosensor fusions to two different degrons, Matα2 from yeast and DREB2a from *Arabidopsis* (Sakuma et al., 2006), and we used the VP16 and VP64 variants as the TAD. We initially tested the G-DIG1-TAD variants with a transient expression assay using *Arabidopsis* protoplasts and a reporter gene consisting of firefly luciferase under the control of a Gal4-activated plant promoter (pUAS::Luc). The biosensor containing the Matα2 degron and VP16 TAD showed the highest fold activation of luciferase in the presence of digoxigenin (Figure 3.4a). We next inserted the genes encoding G-DIG1-V-Matα2 and the Gal4-activated pUAS::Luc into a plant transformation vector and stably transformed them into *Arabidopsis* plants. Primary transgenic plants were screened *in vivo* for digoxigenin-dependent luciferase production, and responsive plants were allowed to set seed for further testing. Second generation transgenic plants (T1, heterozygous) were tested for digoxin- or digoxigenin-dependent induction of luciferase expression. After 42 hours, we observed 30-50 fold induction of luciferase activity in digoxin-treated plants compared to the uninduced control (Figure 3.5). Both digoxin and digoxigenin were capable of inducing the biosensor. Digoxigenin-dependent luciferase induction was observed in multiple independent transgenic T1
lines (Figure 3.4b), and a rising dose response to digoxigenin was observed in the transgenic plants (Figure 3.4c). The specificity of the digoxigenin biosensor in plants parallels that in yeast cells (Figure 3.4d).
Figure 3.4. Characterization of DIG biosensor in plants. (a) Test of DIG variants engineered for plant function in Arabidopsis protoplasts. Two transcriptional activation domains, VP16 (V) and VP64 (VP64), as well as two degrons, yeast MATα2 and Arabidopsis DREB2a, were added to DTF-1 (G-DIG1), and the proteins were constitutively expressed from the CaMV35S promoter. The Gal4-activated pUAS promoter controls expression of a luciferase reporter. Transformed protoplasts were incubated with digoxigenin at 0, 100 µM, and 500 µM for 16 hours. (b) Digoxigenin-dependent activation of luciferase expression in three independent transgenic Arabidopsis lines. Plants were incubated in the absence (Control) or presence (DIG) of 100 µM digoxigenin for 42 hours and imaged. Quantification of luciferase expression is presented as mean relative luciferase units ± s.d. of ten plants. (c) Specificity of luciferase activation in transgenic Arabidopsis plants. All inducers were tested at 100 µM concentration. DIG, digoxigenin; DIGT, digitoxigenin; β-EST, β-Estradiol. Data are presented as mean fold activation relative to the control ± s.e.m. of ten technical replicates. (d) Digoxigenin dose response curve in transgenic Arabidopsis plants. Data are presented as mean fold induction relative to the control ± s.e.m. of ten biological replicates.
Figure 3.4 (Continued)

![Graph and images showing relative luciferase units and fold changes with different concentrations of digoxigenin, digitoxigenin, and digoxigenin-estradio.](image)
Figure 3.5. Application of biosensors in plants. (a) Activation of luciferase expression in transgenic *Arabidopsis* plants containing the G-DIG1-V biosensor in the absence (left) or presence (right) of 100 µM digoxin. Luciferase expression levels are false colored according to scale to the right. Relative luciferase units corresponding to 1 min of image pixel integration (to avoid saturating pixels) are shown above each individual plant. (b) Brightfield image of plants shown in a.
Discussion

*In vivo* biosensors for small molecules enable the regulation and detection of cellular responses to endogenous metabolites and exogenous chemicals. Here we show that LBDs can be conditionally destabilized to create biosensors that function in yeast, mammalian cells, and plants, and we demonstrate the use of these biosensors in metabolic engineering and genome editing applications. While this method requires a high-affinity ligand-binding domain as a starting point, nearly all small molecules of interest have a natural protein interactor. Furthermore, the use of *de novo* designed binders opens the possibility of generating biosensors for ligands with unsuitable or unknown binding proteins. By incorporating standard mutagenesis and screening, our method constitutes a simple platform for sensor development that can be applied to many areas of biotechnology. These sensors act either at the level of post-translational control over protein function or at the level of transcription (Figure 3.6a), and they can be tuned by altering any of their components (Figure 3.6b), or by modifying efflux of the target ligand in the host organism. These tunable features should make the biosensors useful in many different cellular and environmental contexts.

Our results suggest a general mechanism of conditional stabilization for LBDs, allowing the rational development of biosensors for other targets. Furthermore, the portability of the mutations we identified suggests a structural basis for conferring conditional stability to the DIG₀ LBD scaffold. Both the DIG₀ LBD and Gal4 are
homodimers, and the majority of the conditionally-stabilizing mutations are located at the dimer interfaces. A computational model of the Gal4-DIG$_0$ complex indicates that the orientation of the two domains allows a homodimeric fusion to form (Figure 2.4b). These results suggest an allosteric interplay between ligand binding and dimer formation: weakening of the dimer interface, in either the DIG$_0$ or the Gal4 domain, is compensated by ligand binding. This LBD scaffold is derived from a member of the nuclear transport factor 2 family, a fold class that typically has a large dimer interface ($\sim$1200 Å$^2$) that facilitates the large and open ligand-binding site ($\sim$600 Å$^2$). These protein folds are well suited for de novo design of other LBDs because of their large binding pocket and natural substrate diversity (Todd et al., 2002). Exploiting dimer interfaces to modulate stability without impairing ligand binding may be a general mechanism to confer conditional stability on LBDs. This possibility is supported by the observation that interface mutations in DIG$_0$ and Gal4 conferring digoxin-dependent stability lead to progesterone-dependent stability in a progesterone biosensor (Figure 2.3e and Figure 2.5d).

A longstanding challenge in metabolic engineering is to rapidly detect and control how changes to the regulation and composition of biosynthetic pathways affect product titers. Transcriptional control by a product or intermediate (Raman et al., 2014; Tang and Cirino, 2011; Zhang et al., 2012) and directed evolution of constituent pathway elements (Agresti et al., 2010; Alper et al., 2005; Dietrich et al., 2013) have
emerged as promising strategies towards this goal. These approaches require high selectivity against intermediates (Zhang and Keasling, 2011), a feature demonstrated here that can be explicitly considered during the computational design and screening process. Our method allows biosensors to be generated that are highly selective for a small molecule, facilitating a simple directed evolution strategy without requiring prior structural or bioinformatic knowledge about the targeted enzyme(s) or pathway(s). Because the biosensors are TF-based, sophisticated systems of optimizing metabolic output, such as dynamic control of gene expression (Zhang et al., 2012) and feedback-regulated genome evolution (Chou and Keasling, 2013), are possible.

A reliance on the general principles of protein stability and ligand binding allows the development of biosensors that function in any organism with similar protein quality control machinery. Here, we engineered biosensors based on a designed scaffold derived from a bacterial protein. These biosensors required only minimal modifications to retain high levels of sensitivity when developed in yeast and deployed across mammalian and plant species, demonstrating unprecedented portability for biosensors. In some cases, these biosensors showed a greater dynamic range in mammalian cells relative to yeast, possibly due to larger cell volume and variations in protein degradation machinery. Further work using biosensors based on multiple different LBD scaffolds and introduced into diverse organisms will allow us to better understand the principles by which biosensor variability across hosts arises.
Small molecule biosensors with the modularity incorporated here enable diverse cellular responses to a variety of exogenous and endogenous signals (Banaszynski et al., 2008). Gene editing is an area that requires particularly tight coupling of cell response to activation signals. The CRISPR/Cas9 system provides a facile and robust genome-editing platform, but it can result in off-target genetic changes (Fu et al., 2013; Mali et al., 2013b; Pattanayak et al., 2013). Proposed solutions include optimizing guide RNA sequences (Cho et al., 2014; Fu et al., 2014), building chimeric Cas9 fusions requiring the presence of two Cas9 molecules in close proximity (Guilinger et al., 2014; Mali et al., 2013b; Ran et al., 2013; Tsai et al., 2014), and regulating Cas9 activity by chemical or light-based inducers (Dow et al., 2015; Polstein and Gersbach, 2015; Zetsche et al., 2015). While small molecule inducers including doxycycline and rapamycin have been used, these molecules may confer leaky expression and cytotoxicity (Xie et al., 2008). Thus, an expanded chemical repertoire is needed for tightly regulated gene editing and gene therapy applications. By exploiting the low background of the LBD-biosensors, we produced biosensor-Cas9 fusions with tightly controlled activation (Figure 3.3e). This switch-like control over CRISPR/Cas9 activity could reduce background activity and off-target editing, a critical feature for safer gene therapies (Mandal et al., 2014; Schwank et al., 2013; Wu et al., 2013).

Our biosensor design approach should have numerous applications in agriculture. For example, biosensors could be developed to enable plants to monitor the
environment for pollutants, toxins or dangerous compounds. Coupling biosensors with a phytoremediation trait could enable plants to both sense a contaminant and activate a bioremediation gene circuit. When paired with an agronomic or biofuel trait, such biosensors could serve as triggers for bioproduction. In the transgenic *Arabidopsis* plants, we observed ligand-dependent activation in all cells, tissues and organs examined (Figure 3.5).

The technology introduced here operates at either the transcriptional or post-translational level. These biosensors can be developed in yeast and readily transferred with minimal modification to other eukaryotic cell types, where they retain a high level of sensitivity (Figure 3.6c). The generality of our approach arises from the universality of the transcriptional activation and protein degradation machinery across eukaryotes, together with the modularity and tunability of the constituent parts. These biosensors should find broad application, including improving metabolically engineered pathway flux and product titers, exerting ligand-dependent control over genome editing, and detecting exogenous small molecules or endogenous metabolites.
Figure 3.6. Schematic of biosensor platform. (a) Biosensors for small molecules are modularly constructed by replacing the LBD with proteins possessing altered substrate preferences. (b) Activity of the biosensor can be tuned by 1) introducing destabilizing mutations (red Xs), 2) adding a degron, 3) altering the strength of the TAD or DNA binding affinity of the TF, 4) changes in the number of TF binding sites or sequences, and 5) titrating 3-aminotriazole, an inhibitor of His3. (c) Yeast provide a genetically tractable chassis for biosensor development prior to implementation in more complex eukaryotes, such as mammalian cells and plants.
Chapter 4 – Future directions

Motivation

Our biosensor development strategy operates on the general principles of protein folding and ligand binding. Through the experiments in porting our digoxin and progesterone biosensors to K562 human cell culture lines and *Arabidopsis thaliana*, as well through previous work by Wandless and colleagues (Cho et al., 2013), it is evident that the idea of using conditionally destabilized ligand binding domains to sense and respond to small molecules is applicable across many different contexts.

Despite this, there remains a great deal of room for expanding generalizability and broadening accessibility in both the scope of systems that LBD biosensors can be deployed and in the range of molecules that can be targeted for biosensor development. As we look forward towards addressing the current limitations of biosensor development, the two main foci are expanding the systems in which biosensors can be engineered and deployed, as well as creating generalizable scaffolds that can be developed into biosensors for any number of new small molecule targets for which no biosensor currently exists.

**Bacterial hosts for biosensor development and deployment**

In addition to eukaryotic fungal species such as *Saccharomyces cerevisiae*, prokaryotic organisms like bacteria are also popular hosts for bioproduction and thus present an additional need for biosensor-driven optimization of enzymes and
biosynthetic pathways (Tang and Cirino, 2011; Zhang et al., 2012). Prokaryotes lack the conserved ubiquitin-proteasome system (UPS) that is expected to be critical for protein stability-based biosensor function across eukaryotes (Egeler et al., 2011). Despite this, bacterial species do possess a protein quality control system consisting of multiple proteases (Dougan et al., 2002). In addition, there is evidence that destabilized proteins fused to a GFP reporter are targeted to inclusion bodies, which prevents GFP from folding and fluorescing (Drew et al., 2005).

As in our eukaryotic biosensors, the prokaryotic biosensors can be constructed so that the destabilized LBD-reporter fusion is degraded or targeted to an inclusion body in the absence of the stabilizing target ligand, preventing the reporter protein from carrying out its function. In the presence of stabilizing target ligand, the fusion is stabilized and the reporter protein carries out its function. If conditional stability can be demonstrated in bacteria, hosts such as *E. coli* can serve as chassis for both the development and deployment of protein stability-based sensors.

The reporter protein may be a fluorescent protein, transcription factor, enzyme, signaling protein, or other functional protein. It has been previously found that using transcription factors as reporters for biosensor stability amplifies the signal relative to direct fusion to a fluorescent reporter (Figure 2.4). To enable discovery and use of biosensors in *E. coli* with maximum dynamic range, in addition to our previously described direct GFP-LBD fusions, the LBDs can be fused to a panel of transcription
factors including LexA, LacI, 933W, and cl from Lambda phage. These E. coli transcription factors primarily function as repressors by binding to their cognate DNA promoter site and blocking transcription. In these cases, biosensor instability in the absence of ligand will cause degradation or aggregation of the construct, activating transcription of the reporter gene. Stability in the presence of the target ligand will abrogate expression of the reporter.

**Preliminary results**

Preliminary data for the repressor system spanning multiple biosensor expression levels as well as multiple operator binding site strengths for each repressor has demonstrated dose-responsive repression of GFP expression using our ported digoxin and progesterone biosensors in E. coli (Figure 4.1). The sensors only exhibited conditional stability in the N-terminal repressor fusion, indicating a possible dependency on an initial unfolded translational product to reduce signal response. We observe high levels of ligand-induced stabilization, indicating room for improvement of background signal ablation through E. coli specific screens for destabilizing mutations.

To obtain bacterial TF biosensors that activate, rather than repress, gene expression in the presence of the target small molecule, the LBDs can be genetically fused to bacteriophage RNA polymerases from T7, T3, and SP6. Sigma factors have also been shown to act as orthogonal, sequence-specific transcriptional activators and can also be fused LBD biosensors (Rhodius et al., 2014). Expression levels for these new TF
biosensors may have to be optimized, but based on the preliminary results using repressor TFs and our previous experience with transcriptional activators, it is anticipated that these will yield higher levels of dynamic range while also enabling arbitrary transcriptional outputs in response to small molecule concentration in bacteria. Once the activation system has been developed, new libraries of the LBD can be screened for *E. coli* specific mutations.
Figure 4.1. Ligand-dependent transcriptional repression in *E. coli*. (a) Concentration dependence of response to digoxin for digoxin 933W cl repressor-biosensors regulating GFP expression. (b) Concentration dependence of response to progesterone for progesterone 933W cl repressor-biosensors regulating GFP expression. Error bars represent s.e.m. of three biological replicates.
Cell-free biosensors

This work was done in conjunction with Kalia Firester, an undergraduate at Harvard College.

All in vivo approaches to biosensors require that the targets are membrane permeable and non-toxic. Further, biosensor library size is limited by the transformation efficiency of the host organism. In vitro biosensors remove the requirements of target membrane permeability and non-toxicity, and in vitro libraries can surpass the tractable size of in vivo DNA libraries by many orders of magnitude. Using a cell-free system would also allow for precise control over reaction environments such as redox or pH levels in cell lysates. In vitro transcription and translation can be carried out by a mixture of purified components or by whole cell lysate from cells including bacteria, yeast, human, wheat germ and rabbit reticulocytes (Stech and Kubick, 2015). Cell lysates can also be stored and transported in a freeze-dried paper based system, enabling use as an extremely low cost and portable environmental detector or point of care diagnostic (Pardee et al., 2016). Potential limitations to using cell lysates for engineering and deploying conditionally destabilized biosensors include energy depletion, reduced protein quality control machinery capacity, and low transcription and translation levels. Previous work has demonstrated rabbit reticulocyte lysates to be a suitable system for combined protein translation and degradation (Ristriani et al., 2009).
To port the existing biosensors into a cell-free system, previously characterized lysates from eukaryotic organisms such as yeast, rabbit reticulocyte, HeLa, and wheat germ can be used. Endogenous transcription and translational machinery will be tested using previously designed yeast and human promoters. The use of exogenously added RNA polymerases for both the GFP-direct fusion and TF-biosensor constructs will be tested by expressing them under the promoters for T7, T3, and SP6 RNA polymerases and supplying the corresponding purified RNA polymerase to the cell-free system.

The reporters in the TF-biosensors use the VP16 transcriptional activator to recruit endogenous transcriptional machinery to the Gal4 DNA binding site, thereby driving expression of an output signal gene downstream from the Gal4 binding site. In a cell-free system, it is possible that the endogenous translational machinery is not sufficiently recruited to detectably express the reporter gene using the existing Gal4/VP16 system. To address this possibility, alternative transcription factor systems should be tested in parallel. The LBDs can be genetically fused directly to the T7, T3, and SP6 RNA polymerases. The reporter gene can be placed under the cognate promoter for each polymerase, and the performance of each polymerase/promoter pair will be quantified in terms of characteristics like dynamic range, dose response, and kinetics of activation and deactivation. The RNA polymerase used to express the LBD fusion may be orthogonal to the RNA polymerase used to express the reporter gene.
To screen for new biosensors, libraries of variant LBD DNA can be expressed *in vitro* using transcription and translation machinery from the lysate, or optionally supplemented with additional polymerases, ribosomes or other transcriptional or translational machinery. Emulsions can be combined with FACS or microfluidic devices to sort in a similar fashion to our yeast biosensor enrichment strategy.

The DNA sequence of successful biosensors can be determined by multiple methods, including mRNA display as follows: the mRNA sequence encoding each LBD variant is fused to the LBD protein by covalent linkage. The mRNA of LBD variants passing the screen is reverse transcribed to produce cDNA, and then the cDNA can be read by techniques such as Sanger or next-generation sequencing. Functional biosensors can be enriched by using pull-downs for standard sequence tags fused to the biosensor constructs.

**Antibody-based scaffolds for generalizable biosensor construction**

For some small molecule targets a known LBD may not exist, may interfere with cellular processes, or may bind the target with insufficient affinity. In the same vein, despite the promise of using computational design to create designer LBDs for biosensor use, engineering LBDs *de novo* remains a high-level challenge requiring expertise that is not easily accessible to the broader scientific community. We developed this technology as a generalizable method for creating new biosensors, but the results from the initial paper (Feng et al., 2015) serve mainly as a proof of principle rather than
a true demonstration of ligand generalizability. The ultimate tool for biosensor development would be a conditionally destabilized scaffold (or set of scaffolds) that can have its binding target readily switched with minimal effort to arbitrary molecules while demonstrating high levels of affinity and specificity. Achieving this goal would be the most important way to enable accessibility of this technology to many more applications by lowering the barrier of entry.

One promising class of proteins that has potential to be used as generalized biosensor scaffolds are antibodies. Monoclonal antibodies (mAbs) or fragments including FAbs, scFvs, or nanobodies can be engineered to tightly bind target molecules. mAbs and mAb fragments can be expressed in mammalian cells, plants and yeast. They may also be expressed in *E. coli* by periplasmic secretion, by oxidizing the cytosol, or by replacing disulfide forming cysteines with other standard, noncanonical, or nonstandard amino acids. Antibodies are particularly interesting because of their ability to be engineered to bind a vast number of ligands. Ligand-binding loop sequences may be known *a priori*, may be determined from antibodies raised by immunization in animals or tissue culture, or by may be discovered by screening from phage display, yeast display, or *in vitro* methods. The complementary determining regions (CDRs) of antibodies may be identified through sequencing, structural analysis, and homologies, and transferred between scaffolds to effectively swap binding specificities between scaffolds (Nelson et al., 2010). Due to their ability to bind a wide
multitude of molecules in a modular manner combined with the numerous technologies
to develop new CDRs for desired targets, a conditionally destabilized antibody would
be a truly generalizable biosensor scaffold.

Framework mutations to antibody products may provide generalized
destabilized scaffolds, such that sensors to new targets can be produced simply by
grafting ligand recognizing elements such as variable loops onto destabilized antibodies
or antibody fragments (Figure 4.2a,b). Recent work by the Cepko lab has demonstrated
such modularity in mammalian cells expressing conditionally destabilized nanobodies
dependent on binding to different fluorescent proteins (Tang et al., 2016). Similarly,
destabilizing mutations may be ported from an scFv biosensor onto other scFv scaffolds
to render them conditionally destabilized (Figure 4.2c).

Work spanning over a decade by the Rabbitts lab has yielded an scFv scaffold
that exhibits disulfide-free stability in the yeast cytosol (Tanaka et al., 2003; Zhang and
Rabbitts, 2014). This promises to be a suitable starting point for further evolution of
antibody fragments in yeast, both for finding cytosol-stable scaffolds, as well as for
converting into a conditionally destabilized scaffold. To screen for new biosensors, the
stability of the published disulfide free scaffold (Tanaka and Rabbitts, 2008) should be
optimized by combinatorially testing fusions to our Gal4-VP16 system and using
ePCR mutagenesis to find orientations and mutations that exhibit higher activation of
our yEGFP reporter. Causal mutations can be deconvoluted by testing point mutants
singly or in combinations, with the brightest (and presumably most stable) versions being used as starting points for our biosensor engineering work flow. After conditionally destabilizing mutations are found using our standard biosensor enrichment workflow, modularity of the resulting mutations will be tested by porting new CDRs into the conditionally destabilized scaffold, and porting destabilizing mutations into new scaffolds.
Figure 4.2. Schematic of generalizable antibody-based biosensor platform. (a) Antibody fragments such as scFvs can be screened for conditionally destabilizing mutations in the scaffold region (red X’s). (b) CDRs may be swapped into a conditionally destabilized antibody fragment, altering its binding specificity and target response. (c) Conditionally destabilizing mutations may be copied into new antibody fragments to render them responsive to their binding target.
Chapter 4 – Final thoughts

The work presented in this thesis advances our capabilities to sense and respond to molecules in cells by leveraging the principles of protein folding and endogenous cellular protein quality control machinery. We have demonstrated the utility of these conditionally destabilized biosensors by applying them to wide ranging uses such as bioproduction optimization, post-translational control of enzymatic activity, and to act as environmental sensors.

The most immediately impactful use of our biosensor technology is the ability to enable directed evolution of bioproduction pathways. Previous methods of biosensor development have often been useful for very specific contexts such as sensing a particular metabolite or allowing for a small number of chemical inducers to be used to control cellular activity. It has remained a challenge to rapidly produce multiple novel biosensors, particularly for ligands that have been less well studied and do not have well characterized protein interactors. For metabolic engineering to thrive and compete commercially as a strategy to produce chemicals, the throughput at which mutants and chemicals can be screened must be improved drastically, and the ability to more rapidly generate robust biosensors is a key step in that direction.

Another promising application of the conditionally destabilized biosensors is their use as cheap, portable, and easily renewable environmental or diagnostic sensors. While we have not yet ported our existing sensors into an *in vitro* system, the required
components have been previously shown to work in cell-free lysates (Ristriani et al., 2009). If our biosensors could be demonstrated to work in cellular lysates, a paper based system would be an ideal way to deploy them (Pardee et al., 2014). The signal noise, response time, and reproducibility would all need to be carefully characterized to find the proper use cases. It is unlikely they would replace traditional medical diagnostics such as chromatography and mass spectrometry, but for chemical detection where extreme precision is not required and cost and portability are paramount, our sensors could be a useful alternative. The modularity in assay response could also enable colorimetric or other easily observable outputs, negating the need for any expensive equipment.

While the work presented here has made progress in developing a generalizable strategy for sensor development, there remains a great deal of room for improvement. For our technology to gain wide usage and to have a demonstrable effect on researchers outside of our immediate subfield, the barriers to engineering a robust biosensor must be further lowered. The current requirements for a high-affinity ligand binder for each novel target greatly limits the target space. I believe the proposed antibody fragment based generalizable scaffold holds a great deal of promise for having a truly transformative effect on the field of biosensor engineering. We will continue to pursue this idea and test the limits of modularity of the binding regions and destabilizing mutations in antibody fragment biosensors.
While it remains to be seen which of the expansions to our biosensor development platform are able to be implemented successfully, the utility and portability of the sensors we have developed demonstrate a substantial step forward in our ability to sense and respond to small molecules.
Chapter 5 – Materials and Methods

Culture and growth conditions. Biological replicates are defined as samples inoculated from distinct colonies. Growth media consisted of YPAD (10 g/L yeast extract, 20 g/L peptone, 40 mg/L adenine sulfate, 20 g/L glucose) and SD media (1.7 g/L yeast nitrogen base without amino acids, 5 g/L ammonium sulfate, 20 g/L glucose and the appropriate amount of dropout base with amino acids [Clontech]). The following selective agents were used when indicated: G418 (285 mg/L), pen/strep (100 U/mL penicillin and 100 µg/mL streptomycin).

LBD-yEGFP library construction. The DIG10.3 sequence (Tinberg et al., 2013) was cloned by Gibson assembly (Gibson et al., 2009) into a pUC19 plasmid containing yeast enhanced GFP (yEGFP, UniProt ID B6UPG7) and a KanMX6 cassette flanked by 1000 and 500 bp upstream and downstream homology to the HO locus. The DIG10.3 sequence was randomized by error-prone PCR using a Genemorph II kit from Agilent Technologies. An aliquot containing 100 ng of target DNA (423 bp out of a 7.4 kb plasmid) was mixed with 5 µL of 10X Mutazyme buffer, 1 µL of 40 mM dNTPS, 1.5 µL of 20 µM forward and reverse primer containing 90 bp overlap with the pUC19 plasmid (oJF70 and oJF71), and 1 µL of Mutazyme polymerase in 50 µL. The reaction mixture was subject to 30 cycles with Tm of 60 °C and extension time of 1 min. Vector backbone was amplified using Q5 polymerase (NEB) with oJF76 and oJF77 primers with Tm of 65 °C and extension time of 350 s. Both PCR products were isolated by 1.5% agarose gel
electrophoresis and the randomized target was inserted as a genetic fusion to yEGFP by Gibson assembly (Gibson et al., 2009). Assemblies were pooled, washed by ethanol precipitation, and resuspended in 50 µL of dH2O, which was drop dialyzed (Millipore) and electroporated into E. cloni supreme cells (Lucigen). Sanger sequencing of 16 colonies showed a mutation rate of 0-7 mutations/kb. The library was expanded in culture and maxiprepped (Qiagen) to 500 µg/µl aliquots. 16 µg of library was drop dialyzed and electrotransformed into yeast strain Y7092 for homologous recombination into the HO locus. Integrants were selected by growth on YPAD solid media containing G418 followed by outgrowth in YPAD liquid media containing G418.

**LBD-yEGFP library selections.** Libraries of DIG0-yEGFP and PRO0-yEGFP integrated into yeast strain Y7092 were subject to three rounds of fluorescence activated sorting in a BD FACSaria IIu. For the first round, cells were grown overnight to an OD₆₀₀ of ~1.0 in YPAD containing steroid (500 µM digoxigenin or 50 µM progesterone), and cells showing the top 5% of fluorescence activation were collected and expanded overnight to an OD₆₀₀ of ~1.0 in YPAD lacking steroid. In the second sort, cells displaying the lowest ~3% fluorescence activation were collected. Cells passing the second round were passaged overnight in YPAD containing steroid to an OD₆₀₀ of ~1.0 and sorted once more for the upper 5% of fluorescence activation. The sorted libraries were expanded in YPAD liquid culture and plated on solid YPAD media. Ninety-six colonies from each library were clonally isolated and grown overnight in deep well plates containing 500
µL of YPAD. Candidates were diluted 1:50 into two deep well plates with SD-complete media: one plate supplemented with steroid and the other with DMSO vehicle. Cells were grown for another 4 h, and then diluted 1:3 into microtitre plates of 250 µL of the same media. Candidates were screened by analytical flow cytometry on a BD LSRFortessa cell analyzer. The forward scatter, side scatter, and yEGFP fluorescence (530 nm band pass filter) were recorded for a minimum of 20,000 events. FlowJo X software was used to analyze the flow cytometry data. The fold activation was calculated by normalizing mean yEGFP fluorescence activation for each steroid to the mean yEGFP fluorescence in the DMSO only control. Highest induction candidates were subject to Sanger sequencing with primers flanking the LBD sequence.

**TF-biosensor reporter plasmid construction and integration.** Reporter genes were cloned into the integrative plasmid pUG6 or the CEN plasmid pRS414 using the Gibson method (Gibson et al., 2009). Each reporter (either yEGFP or firefly luciferase) was cloned to include a 5’ GAL1 promoter (*S. cerevisiae* GAL1 ORF bases (-455)-(5)) and a 3’ CYC1 terminator. For integration, linearized PCR cassettes containing both the reporter and an adjacent KanMX antibiotic resistance cassette were generated using primers containing 50 bp flanking sequences of homology to the *URA3* locus. Integrative PCR product was transformed into the yeast strain PJ69-4a using the Gietz method (Gietz and Schiestl, 2007) to generate integrated reporter strains.
**G-DIG/PRO-V plasmid construction.** G-DIG/PRO-V fusion constructs were prepared using the Gibson method (Gibson et al., 2009). Constructs were cloned into the plasmid p416CYC (p16C). Gal4 (residues 1-93, UniProt ID P04386), DIG10.3 (Tinberg et al., 2013), and VP16 (residues 363-490, UniProt ID P06492) PCR products for were amplified from their respective templates using Phusion high-fidelity polymerase (NEB, Waltham, MA) and standard PCR conditions (98 °C 10 s, 60 °C 20 s, 72 °C 30 s; 30 cycles). The 8-residue linker sequence GGSGGSGGG was used between Gal4 and DIG10.3. PCR primers were purchased from Integrated DNA technologies and contained 24-30 5’ bases of homology to either neighboring fragments or plasmid. Clones containing an N-terminal degron were similarly cloned fusing residues 1-67 of Mata2 (UniProt ID P0CY08) to the 5'-end of G-DIG-V. Plasmids were transformed into yeast using the Gietz method (Gietz and Schiestl, 2007), with transformants being plated on synthetic complete media lacking uracil (SD -ura).

**G-DIG-V mutant construction.** Mutations were introduced into DIG10.3/pETCON (Tinberg et al., 2013) or the appropriate G-DIG/PRO-V construct using Kunkel mutagenesis (Kunkel, 1985). Oligos were ordered from Integrated DNA Technologies, Inc. For mutants constructed in pETCON/DIG10.3, the mutagenized DIG10.3 gene was amplified by 30 cycles of PCR (98 °C 10 s, 61 °C 30 s, 72 °C 15 s) using Phusion high-fidelity polymerase (NEB, Waltham, MA) and 5’- and 3’- primers having homologous overlap with the DIG10.3-flanking regions in p16C-G-DIG-VP64.
Genes were inserted into p16C-Gal4-(HE)-VP16 by Gibson assembly (Gibson et al., 2009) using vector digested with HindIII and EcoRI-HF.

**G-PRO-V mutant construction.** The gene for DIG10.3 Y34F/Y99F/Y101F was amplified from the appropriate DIG10.3/pETCON (Tinberg et al., 2013) construct by 30 cycles of PCR (98 °C 10 s, 59 °C 30 s, 72 °C 15 s) using Phusion high-fidelity polymerase (NEB, Waltham, MA) and 5′- and 3′- primers having homologous overlap with the DIG10.3-flanking regions in p16C-G-DIG-VP64 (DIG_fwd and DIG_rev). Genes were inserted into p16C-GDVP16 by Gibson assembly (Gibson et al., 2009) using p16C-Gal4-(HE)-VP16 vector digested with HindIII and EcoRI-HF.

**G-DIG-V error-prone library construction.** A randomized G-DIG-V library was constructed by error-prone PCR using a Genemorph II kit from Agilent Technologies. An aliquot containing 20 ng p16C GDVP16, 20 ng p16C GDVP16 E83V, and 20 ng p16C Y36H was mixed with 5 µL of 10X Mutazyme buffer, 1 µL of 40 mM dNTPS, 1.5 µL of 20 µM forward and reverse primer containing 37- and 42-bp overlap with the p16C vector for homologous recombination, respectively (GDV_ePCR_fwd and GDV_ePRC_rev), and 1 µL of Mutazyme polymerase in 50 µL. The reaction mixture was subjected to 30 cycles of PCR (95 °C 30 s, 61 °C 30 s, 72 °C 80 s). Template plasmid was digested by adding 1 µL of DpnI to the reaction mixture and incubating for 3 hr at
37 °C. Resulting PCR product was purified using a Quiagen PCR cleanup kit, and a second round of PCR was used to amplify enough DNA for transformation. Gene product was amplified by combining 100 ng of mutated template DNA with 2.5 µL of 10 µM primers (GDV_ePCR_fwd and GDV_ePRC_rev), 10 µL of 5X Phusion buffer HF, 1.5 µL of DMSO, and 1 µL of Phusion high-fidelity polymerase (NEB, Waltham, MA) in 50 µL. Product was assembled by 30 cycles of PCR (98 °C 10 s, 65 °C 30 s, 72 °C 35 s).

Following confirmation of a single band at the correct molecular weight by 1% agarose gel electrophoresis, the PCR product was purified using a Quaigen PCR cleanup kit and eluted in ddH₂O. Yeast strain PyE1 ΔPDR5 was transformed with 9 µg of amplified PCR library and 3 µg of p16C Gal4-(HE)-VP16 triply digested with SalI-HF, BamHI-HR, and EcoRI-HF using the method of Benatuil (Benatuil et al., 2010), yielding ~10⁶ transformants. Following transformation, cells were grown in 150 mL of SD -ura media. Sanger sequencing of 12 individual colonies revealed an error rate of ~1-6 mutations per gene.

**G-DIG-V library selections.** An error-prone library of G-DIG₀/DIG₁/DIG₂/-V transformed into yeast strain PyE1 ΔPDR5 was subjected to three rounds of cell sorting using a Cytopeia (BD Influx) fluorescence activated cell sorter. For the first round, cells displaying high fluorescence in the presence of digoxin (on-state) were collected. Transformed cells were pelleted by centrifugation (4 min, 4000 rpm) and resuspended to a final OD₆₀₀ of 0.1 in 50 mL of SD -ura media, pen/step antibiotics, and 5 µM digoxin.
prepared as a 100 mM solution in DMSO. The library was incubated at 30 °C for 9 h and then sorted. Cells displaying the highest fluorescent values in the GFP channel were collected (1,747,058 cells collected of 32,067,013 analyzed; 5.5%), grown up at 30 °C in SD-ura, and passaged twice before the next sort. For the second round of sorting, cells displaying low fluorescence in the absence of digoxin (off-state) were collected. Cells were pelleted by centrifugation (4 min, 4000 rpm) and resuspended to a final OD$_{600}$ of 0.1 in 50 mL of SD-ura media supplemented with pen/strep antibiotics. The library was incubated at 30 °C for 8 h and then sorted. Cells displaying low fluorescent values in the GFP channel were collected (1,849,137 cells collected of 22,290,327 analyzed; 11.1%), grown up at 30 °C in SD-ura, and passaged twice before the next sort. For the last sorting round, cells displaying high fluorescence in the presence of digoxin (on-state) were collected. Cells were prepared as for the first sort. Cells displaying the highest fluorescent values in the GFP channel were collected (359,485 cells collected of 31,615,121 analyzed; 1.1%). After the third sort, a portion of cells were plated and grown at 30 °C. Plasmids from 12 individual colonies were harvested using a Zymoprep Yeast miniprep II kit (Zymo Research Corporation, Irvine, CA) and the gene was amplified by 30 cycles of PCR (98 °C 10 s, 52 °C 30 s, 72 °C 40 s) using Phusion high-fidelity polymerase (NEB, Waltham, MA) with the T3 and T7 primers. Sanger sequencing (Genewiz, Inc., South Plainfield, NJ) was used to sequence each clone in the forward (T3) and reverse (T7) directions.
**G-DIG-V error-prone library mutation screens.** Of twelve sequenced clones from the library sorts, two showed significantly improved (>2-fold) response to DIG over the input clones (clone 3 and clone 6). Clone 3 contains the following mutations: Gal4_T44T (silent), Gal4_L77F, DIG10.3_E5D, DIG10.3_E83V, DIG10.3_R108R (silent), DIG10.3_L128P, DIG10.3_I137N, DIG10.3_S143G, and VP16_A44T. Clone 6 contains the following mutations: Gal4_R60S, Gal4_L84L (silent), VP16_G17G (silent), VP16_L48V, and VP16_H98H (silent). To identify which mutations led to the observed changes in DIG response, variants of these clones with no silent mutations and each individual point mutant were constructed using Kunkel mutagenesis (Kunkel, 1985). Oligos were ordered from Integrated DNA Technologies, Inc. Sequence-confirmed plasmids were transformed into PyE1 ΔPDR5f and plated onto selective SD-ura media. Individual colonies were inoculated into liquid media, grown at 30 °C, and passaged once. Cells were pelleted by centrifugation (4 min, 1700 x g) and resuspended to a final OD_{660} of 0.1 in 1 mL of SD-ura media supplemented 50 µM DIG prepared as a 100 mM solution in DMSO. Following a 6 hr incubation at 30 °C, cells were pelleted, resuspended in 200 µL of PBS, and cellular fluorescence was measured on an Accuri C6 flow cytometer using a 488 nm laser for excitation and a 575 nm band pass filter for emission. FlowJo software version 7.6 was used to analyze the flow cytometry data. Data are given as the mean yEGFP fluorescence of the single yeast population in the absence of DIG (off-state) and the mean yEGFP fluorescence of the higher fluorescing yeast population in the presence...
of DIG (on-state).

**Computational model of Gal4-DIG.** A model of the Gal4-DIG10.3 fusion was built using Rosetta Remodel (Huang et al., 2011) to assess whether the linker between Gal4 and the DIG LBD, which are both dimers, would allow for the formation of a dimer in the fusion construct. In the simulation, the Gal4 dimer was held fixed while the relative orientation of the DIG LBD monomers were sampled symmetrically using fragment insertion in the linker region. Constraints were added across the DIG LBD dimer interface to facilitate sampling. The lowest energy model satisfied the dimer constraints, indicating that a homodimer configuration of the fusion is possible.

**TF-biosensor titration assays in yeast.** Yeast strain PyE1 transformed with p16C plasmids containing G-LBD-V variants were inoculated from colonies into SD–ura media supplemented and grown at 30 °C overnight (16 h). 10 µL of the culture was resuspended into 490 µL of separately prepared media each containing a steroid of interest (SD–ura media supplemented the steroid of interest and DMSO to a final concentration of 1% DMSO). Resuspended cultures were then incubated at 30°C for 8 hours. 125 µL of incubated culture was resuspended into 150 µL of fresh SD–ura media supplemented with the steroid of interest and DMSO to a final concentration of 1%. These cultures were then assayed by analytical flow cytometry on a BD LSRFortessa using a 488 nm laser for excitation. The forward scatter, side scatter, and yEGFP
fluorescence (530 nm band pass filter) were recorded for a minimum of 20,000 events. FlowJo X software was used to analyze the flow cytometry data. The fold activation was calculated by normalizing mean yEGFP fluorescence activation for each steroid to the mean yEGFP fluorescence in the DMSO only control. G-PROe-V was assayed on a separate day from the other TF biosensors under identical conditions.

TF-biosensor kinetic assays in yeast. Yeast strain PyE1 was transformed with p16C plasmids containing G-LBD-V variants were inoculated from colonies into SD –ura media and grown at 30 °C overnight (16 h). 5 µL of each strain was diluted into 490 µL of SD –ura media in 2.2 mL plates. Cells were incubated at 30 °C for 8 hours. 5 µL of steroid was then added for a final concentration of 250 µM digoxin or 50 µM progesterone. For each time point, strains were diluted 1:3 into microtitre plates of 250 µL of the same media. Strains were screened by analytical flow cytometry on a BD LSRFortessa cell analyzer. The forward scatter, side scatter, and yEGFP fluorescence (530 nm band pass filter) were recorded for a minimum of 20,000 events. FlowJo X software was used to analyze the flow cytometry data. The fold activation was calculated by normalizing mean yEGFP fluorescence activation for each time point to the mean yEGFP fluorescence at T = 0 h.

Luciferase reporter assay. Yeast strains containing either a plasmid-borne or integrated luciferase reporter were transformed with p16C plasmids encoding TF-biosensors.
Transformants were grown in triplicate overnight at 30 °C in SD –ura media containing 2% glucose in sterile glass test tubes on a roller drum. After ~16 hours of growth, OD$_{600}$ of each sample was measured and cultures were back diluted to OD$_{600}$ = 0.2 in fresh SD –ura media containing steroid dissolved in DMSO or a DMSO control (1% DMSO final). Cultures were grown at 30 °C on roller drum for 8 hrs prior to taking readings.

Measurement of luciferase activity was adapted from a previously reported protocol (Leskinen et al., 2003). 100 uL of each culture was transferred to a 96-well white NUNC plate. 100 uL of 2 mM D-luciferin in 0.1 M sodium citrate (pH 4.5) was added to each well of the plate and luminescence was measured on a Victor 3V after 5 minutes.

**Yeast deletion strain creation.** Genomic deletions were introduced into the yeast strains PJ69-4a and PyE1 using the 50:50 method (Horecka and Davis, 2013). Briefly, forward and reverse primers were used to amplify an URA3 cassette by PCR. These primers generated a product containing two 50 bp sequences homologous to the 5’ and 3’ ends of the ORF at one end and a single 50 bp sequence homologous to the middle of the ORF at the other end. PCR products were transformed into yeast using the Gietz method (Gietz and Schiestl, 2007) and integrants were selected on SD –ura plates. After integration at the correct locus was confirmed by a PCR screen, single integrants were grown for 2 days in YEP containing 2.5% ethanol and 2% glycerol. Each culture was plated on synthetic complete plates containing 5-fluoroorotic acid. Colonies were screened for deletion of the ORF and elimination of the Ura3 cassette by PCR and
confirmed by DNA sequencing.

**TF-biosensor specificity assays.** Yeast strains expressing the TF-biosensors and yEGFP reporter (either genetically fused or able to be transcriptionally activated by the TAD) were grown overnight at 30 °C in SD –ura media for 12 hours. Following overnight growth, cells were pelleted by centrifugation (5 min, 5250 rpm) and resuspended into 500 µL of SD –ura. 10 µL of the washed culture was resuspended into 490 µL of separately prepared media each containing a steroid of interest (SD –ura media supplemented with the steroid of interest and DMSO to a final concentration of 1% DMSO). Steroids were tested at a concentration of 100 µM digoxin, 50 µM progesterone, 250 µM pregnenolone, 100 µM digitoxigenin, 100 µM beta-estradiol, and 100 µM hydrocortisone. Stock solutions of steroids were prepared as a 50 mM solution in DMSO.

Resuspended cultures were then incubated at 30°C for 8 hours. 125 µL of incubated culture was resuspended into 150 µL of fresh SD –ura media supplemented the steroid of interest, and DMSO to a final concentration of 1%. These cultures were then assayed by analytical flow cytometry on a BD LSRFortessa using a 488 nm laser for excitation. The forward scatter, side scatter, and yEGFP fluorescence (530 nm band pass filter) were recorded for a minimum of 20,000 events. FlowJo X software was used to analyze the flow cytometry data. The fold induction was calculated by normalizing mean
yEGFP fluorescence activation for each steroid to the mean yEGFP fluorescence in the DMSO only control.

**Yeast spotting assays.** Yeast strain PJ69-4a transformed with p16C plasmids containing degron-G-DIG-V variants were first inoculated from colonies into SD -ura media and grown at 30 °C overnight (16 h). 1 mL of each culture was pelleted by centrifugation (3000 rcf, 2 min), resuspended in 1mL of fresh SD -ura and the OD\textsubscript{660} was measured. Each culture was then diluted in SD -ura media to an OD\textsubscript{660} = 0.2 and incubated at 30 °C for 4-6 hrs. 1 mL of each culture was pelleted and resuspended in sterile, distilled water and the OD\textsubscript{660} measured again. Each transformant was then diluted to an OD\textsubscript{660} = 0.1. Four 1/10 serial dilutions of each culture were prepared in sterile water (for a total of 5 solutions). 10 µL of each dilution was spotted in series onto several SD –ura –his agar plates containing 1mM 3-aminotriazole and the indicated steroid. Steroid solutions were added to agar from 200x steroid solutions in DMSO (0.5% DMSO final in plates).

**3β-HSD plasmid and library construction.** The 3β-HSD ORF was synthesized as double-stranded DNA (Integrated DNA Technologies, Inc.) and amplified using primers oJF325 and oJF326 using KAPA HiFi under standard PCR conditions and digested with BsmBI to create plasmid pJF57. 3β-HSD expression plasmids (pJF76 through pJF87) were generated by digesting plasmid pJF57 along with corresponding plasmids from the Yeast Cloning Toolkit (Lee et al., 2015) with Bsal and assembled.
using the Golden Gate Assembly method (Engler et al., 2008). The 3β-HSD sequence was randomized by error-prone PCR using a Genemorph II kit from Agilent Technologies. An aliquot containing 100 ng of target DNA was mixed with 5 µL of 10X Mutazyme buffer, 1 µL of 40 mM dNTPS, 1.5 µL of 20 µM forward and reverse primer containing 90-bp overlap with the 3β-HSD expression plasmids and 1 µL of Mutazyme polymerase in 50 µL. The reaction mixture was subject to 30 cycles with Tm of 60 °C and extension time of 1 min. Vector backbone was amplified using KAPA HiFi polymerase with oJF387 and oJF389 (pPAB1) or oJF387 and oJF389 (pPOP6) with Tm of 65 °C and extension time of 350 s. PCR products were isolated by 1.5% agarose gel electrophoresis and assembled using the Gibson method (Gibson et al., 2009). Assemblies were pooled, washed by ethanol precipitation, and resuspended in 50 µL of dH2O, which was drop dialyzed (Millipore) and electroporated into E. cloni supreme cells (Lucigen). Sanger sequencing of 16 colonies showed a mutation rate of 0-4 mutations/kb. The library was expanded in culture and maxiprepped (Qiagen) to 500 µg/µL aliquots. 16 µg of library was drop dialyzed and electrotransformed into yeast strain PyE1.

3β-HSD progesterone selections. PyE1 transformed with libraries of 3β-HSD were seeded into 5 mL of SD –ura –leu media supplemented and grown at 30 °C overnight (24 h). Cultures were measured for OD600, diluted to an OD600 of 0.0032, and 100 µL was plated onto SD –ura –leu –his plates supplemented 35 mM 3-AT and either 50 µM pregnenolone or 0.5% DMSO.
**Progesterone bioproduction and GC/MS analysis.** Production strains were inoculated from colonies into 5 mL SD–ura media and grown at 30 °C overnight (16 h). 1 mL of each culture was washed and resuspended into 50 mL of SD–ura with 250 µM of pregnenolone and grown at 30 °C for 76 h. OD$_{600}$ measurements were recorded for each culture before pelleting by centrifugation. Cells were lysed by glass bead disruption, and lysates and growth media were extracted separately with heptane. Extractions were analyzed by GC/MS.

**TF-biosensor EGFP assays in mammalian cells.** The K562 cell line was obtained from the ATCC. The cell line was not authenticated and not tested for mycoplasma contamination. For each TF-biosensor, 1 µg of the PiggyBac construct along with 400 ng of transposase were nucleofected into K562 cells using the Lonza Nucleofection system as per manufacturer settings. Two days post-transfection, cells underwent puromycin selection (2 µg/mL) for at least eight additional days to allow for unintegrated plasmid to dilute out and ensure that all cells contained the integrated construct. An aliquot of 100,000 cells of each integrated population were then cultured with 25 µM of progesterone, 1 µM of digoxigenin, or no small molecule. Forty-eight hours after small molecule addition, cells were analyzed by flow cytometry using a BD Biosciences Fortessa system. Mean EGFP fluorescence of the populations was compared.

**Construction of K562 cell lines.** The PiggyBac transposase system was employed to
integrate biosensor constructs into K562 cells. Vector PB713B-1 (Systems Biosciences) was used as a backbone. Briefly, this backbone was digested with NotI and HpaI and G-LBD-V, Gal4BS-E1b-EGFP (EGFP; enhanced GFP ref or UniProt ID A0A076FL24), and sEF1-Puromycin were cloned in. Gal4BS represents four copies of the binding sequence. For hCas9, the PiggyBac system was also employed, but the biosensors were directly fused to the N-terminus of Cas9 and were under control of the CAGGS promoter. Cas9 from S. pyogenes was used.

**TF-biosensor-Cas9 assays.** Construct integration was carried out as for the Cas9 experiments for EGFP assays, except that the constructs were integrated into K562 containing a broken EGFP reporter construct. Introduction of an engineered nuclease along with a donor oligonucleotide can correct the EGFP and produce fluorescent cells. Upon successful integration (~10 days after initial transfection), 500,000 cells were nucleofected with 500 ng of guide RNA (sgRNA) and 2 µg of donor oligonucleotide. Nucleofected cells were then collected with 200 µL of media and 50 µL aliquots were added to wells containing 950 µL of media. Each nucleofection was split into four separate wells containing 1 µM of digoxigenin, 25 µM of progesterone, or no small molecule. Forty-eight hours later, cells were analyzed using flow cytometry and the percentage of EGFP positive cells was determined.

**TF-biosensor assays in protoplasts.** Digoxin transcriptional activators were initially
tested in a transient expression assay using *Arabidopsis* protoplasts according previously described methods (Yoo et al., 2007), with some modifications. Briefly, protoplasts were prepared from 6-week old *Arabidopsis* leaves excised from plants grown in short days. Cellulase Onozuka R-10 and Macerozyme R-10 (Yakult Honsha, Inc., Japan) in buffered solution were used to remove the cell wall. After two washes in W5 solution, protoplasts were re-suspended in MMg solution at 2 x 10^5 cells/mL for transformation. Approximately 10^4 protoplasts were mixed with 5 µg of plasmid DNA and PEG4000 at a final concentration of 20%, and allowed to incubate at room temperature for 30 minutes. The transformation reaction was stopped by addition of 2 volumes of W5 solution, and after centrifugation, protoplasts were re-suspended in 200 µL of WI solution (at 5 x 10^5/mL) and plated in a 96-well plate. Digoxigenin (Sigma-Aldrich, St. Louis, MO) was added to the wells, and protoplasts were incubated overnight at room temperature in the dark, with slight shaking (40 rpm). For luciferase imaging, protoplasts were lysed using Passive Lysis Buffer (Promega, Madison, WI) and mixed with LARII substrate (Dual-Luciferase Reporter Assay System, Promega). Luciferase luminescence was collected by a Stanford Photonics XR/MEGA-10 ICCD Camera and quantified using Piper Control (v.2.6.17) software.

**Plant plasmid construction.** G-DIG\(^{i}\)-V was recoded to function as a ligand-dependent transcriptional activator in plants. Specifically, an *Arabidopsis thaliana* codon optimized protein degradation sequence from the yeast MAT\(\alpha\)2 gene was fused in frame in
between the Gal4 DBD and the DIG1 LBD. The resulting gene sequence was codon-optimized for optimal expression in *Arabidopsis thaliana* plants and cloned downstream of a plant-functional CaMV35S promoter to drive constitutive expression in plants, and upstream of the octopine synthase (*ocs*) transcriptional terminator sequence. To quantify the transcriptional activation function of DIG10.3, the luciferase gene from *Photinus pyralis* (firefly) was placed downstream of a synthetic plant promoter consisting of five tandem copies of a Gal4 Upstream Activating Sequence (UAS) fused to the minimal (-46) CaMV35S promoter sequence. Transcription of luciferase is terminated by the E9 terminator sequence. These sequences were cloned into a pJ204 plasmid and used for transient expression assays in Arabidopsis protoplasts.

**Construction of transgenic Arabidopsis plants.** After confirmation of function in transient tests, the digoxin biosensor genetic circuit was transferred to pCAMBIA 2300 and was stably transformed into *Arabidopsis thaliana* ecotype Columbia plants using a standard Agrobacterium floral dip method (Clough and Bent, 1998). Transgenic plants were selected in MS media (Murashige and Skoog, 1962) containing 100 mg/L kanamycin.

**TF-biosensor assays in transgenic plants.** Transgenic plants expressing the digoxin biosensor genetic circuit were tested for digoxigenin-induced luciferase expression by placing 14-16 day-old plants in liquid MS (- sucrose) media supplemented with 0.1 mM
digoxigenin in 24-well plates, and incubated in a growth chamber at 24 °C, 100 µE.m².s⁻¹ light. Luciferase expression was measured by imaging plants with a Stanford Photonics XR/MEGA-10 ICCD Camera, after spraying luciferin and dark adapting plants for 30 minutes. Luciferase expression was quantified using Piper Control (v.2.6.17) software. Plants from line KJM58-10 were used to test for specificity of induction by incubating plants, as described above, in 0.1 mM digoxigenin, 0.1 mM digitoxigenin, and 0.02 mM β-estradiol. All chemicals were obtained from Sigma-Aldrich (St. Louis, MO).
References


Genome Res. 24, 132–141.


Sakuma, Y., Maruyama, K., Qin, F., Osakabe, Y., Shinozaki, K., and Yamaguchi-


Tang, S.-Y., Qian, S., Akinterinwa, O., Frei, C.S., Gredell, J. a, and Cirino, P.C. (2013). Screening for enhanced triacetic acid lactone (TAL) production by recombinant
Escherichia coli expressing a designed TAL reporter. J. Am. Chem. Soc.


