

Engineered biosensors from dimeric ligand-binding domains

Benjamin Jester, Christine E. Tinberg, Matthew S. Rich, David Baker, and Stanley Fields

ACS Synth. Biol., **Just Accepted Manuscript** • DOI: 10.1021/acssynbio.8b00242 • Publication Date (Web): 11 Sep 2018

Downloaded from <http://pubs.acs.org> on September 13, 2018

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.



Engineered biosensors from dimeric ligand-binding domains

Benjamin W. Jester,^{1,2} Christine E. Tinberg,³ Matthew S. Rich,² David Baker^{1,3} and Stanley Fields^{1,2,4}

¹Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195

²Department of Genome Sciences, University of Washington, Seattle, WA 98195

³Department of Biochemistry, University of Washington, Seattle, WA 98195

⁴Department of Medicine, University of Washington, Seattle, WA 98195

Current addresses:

Biologics Group, Lumen Bioscience, Seattle, WA 98103 (BWJ)

Department of Biology, University of Utah, Salt Lake City, UT 84112 (MSR)

Amgen Discovery Research, Amgen South San Francisco, South San Francisco, CA 94080 (CET)

Correspondence to fields@uw.edu, tel 206 616-4522, ORCID 0000-0001-5504-5925 (S.F.)

Abstract

Biosensors are important components of many synthetic biology and metabolic engineering applications. Here, we report a second generation of *Saccharomyces cerevisiae* digoxigenin and progesterone biosensors based on destabilized dimeric ligand-binding domains that undergo ligand-induced stabilization. The biosensors, comprising one ligand-binding domain monomer fused to a DNA-binding domain and another fused to a transcriptional activation domain, activate reporter gene expression in response to steroid binding and receptor dimerization. The introduction of a destabilizing mutation to the dimer interface increased biosensor dynamic range by an order of magnitude. Computational redesign of the dimer interface and functional selections were used to create heterodimeric pairs with further improved dynamic range. A heterodimeric biosensor built from the digoxigenin and progesterone ligand-binding domains functioned as a synthetic “AND”-gate, with 20-fold stronger response to the two ligands in combination than to either one alone. We also identified mutations that increase the sensitivity or selectivity of the biosensors to chemically similar ligands. These dimerizing biosensors provide additional flexibility for the construction of logic gates and other applications.

KEYWORDS: biosensor, protein engineering, ligand-binding domain, logic gate, transcription factor, directed evolution, next generation sequencing

1
2
3 Biosensors that detect small molecules can facilitate the engineering of microbial
4
5 strains that synthesize useful bioproducts,¹⁻² especially when combined with recent
6
7 technologies that can generate hundreds of thousands of combinations of variant
8
9 enzymes and promoters.³⁻⁴ Such biosensors allow the identification of those strains
10
11 producing the highest concentrations of desired products. They also allow the
12
13 engineering of cells that respond to exogenous chemicals, often by changing gene
14
15 expression in these cells.
16
17
18

19 Strategies to construct biosensors include RNA-based riboswitches,⁵ FRET
20
21 reporters,⁶ and allosteric transcription factors.⁷ These types of biosensors are often
22
23 limited in sensitivity and can require iterative rounds of improvement to obtain ligand
24
25 binding of sufficient affinity and specificity for use. An alternative approach exploits
26
27 natural or computationally designed ligand-binding domains that are destabilized by
28
29 mutation to be rapidly degraded in the absence of ligand.⁸ Stability is rescued by
30
31 binding to a small molecule ligand. Recent advances in protein engineering have
32
33 yielded *de novo* designed ligand-binding domains for several small molecules, including
34
35 digoxigenin,⁹ fentanyl,¹⁰ and 17 α -progesterone.¹¹ This approach provides a path for the
36
37 construction of a biosensor from a computationally designed or natural ligand-binding
38
39 domain recognizing a small molecule.
40
41
42
43

44 Here, we report a second generation of conditionally stable eukaryotic
45
46 biosensors that require both ligand binding and protein dimerization for activity.
47
48 Synthetic, dimeric transcription factors have been reported to respond to isopentenyl
49
50 diphosphate in yeast,¹² but their design requires a conformational switch upon ligand
51
52 binding. The biosensors we describe require only that the ligand-binding domain be
53
54
55
56
57
58
59
60

1
2
3 dimeric. The biosensors consist of two fusion proteins: a ligand-binding domain fused to
4 a DNA-binding domain; and a ligand-binding domain fused to a transcriptional activation
5 domain (Figure 1a). An active transcription factor is reconstituted upon small molecule
6 binding, which stabilizes the biosensor and increases its concentration in the cell.
7
8 Because each of the two ligand-binding domains binds its own ligand, this strategy also
9 affords the opportunity to create biosensors that require two different ligands to become
10 fully active. To test this concept, we started with the computationally-designed
11 digoxigenin-binding protein DIG10.3⁹ as a homodimeric ligand-binding domain. Relative
12 to the first-generation biosensors,⁸ the background activity of our second-generation
13 dimeric biosensors was greatly reduced because the constituents were split into
14 different proteins. Computational redesign of the dimer interface and ligand-binding
15 pocket allowed us to create a pair of orthogonal heterodimers that respond either to the
16 small molecule digoxigenin or to progesterone. We further improved these biosensors
17 using directed evolution and deep sequencing to identify mutations that favor
18 heterodimerization and small molecule binding selectivity. Finally, we demonstrate that
19 in combination, these approaches can result in an improved “AND”-gate and a pair of
20 orthogonal biosensors that can selectively activate independent reporters.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

44 Results

45 Dimeric biosensors are stabilized by ligand binding

46
47 We sought to extend the functionality of ligand-stabilized biosensors through the
48 use of homodimeric proteins. Whereas the first-generation biosensors consisted of a
49 ligand-binding domain fused to a DNA-binding domain and transcriptional activation
50
51
52
53
54
55
56
57
58
59
60

1
2
3 domain in a single polypeptide, these second-generation biosensors comprise two
4
5 separate polypeptide chains: a ligand-binding domain fused to a DNA-binding domain,
6
7 and a ligand-binding domain fused to an activation domain (Figure 1a). With this
8
9 approach, both fusion proteins should be unstable and rapidly degrade in the absence
10
11 of ligand. The ligand-binding domains become stabilized by binding to their cognate
12
13 ligands, reconstituting a functional transcription factor that activates the expression of a
14
15 reporter gene.
16
17

18
19 With first-generation biosensors based on the computationally-designed dimeric
20
21 digoxigenin-binding protein DIG10.3,⁹ many of the conditionally destabilizing mutations
22
23 identified by random mutagenesis and selection mapped to its homodimer interface,⁸
24
25 suggesting that dimerization plays an important role in this protein's stability. We asked
26
27 whether we could exploit ligand-stabilized dimerization of DIG10.3 to reconstitute an
28
29 active transcription factor. We fused the DIG10.3 monomer (termed DIG0) to a Gal4
30
31 DNA-binding domain and to a VP16 activation domain and transformed a yeast strain,
32
33 PF22, containing a firefly luciferase reporter gene driven by the *GAL1* promoter, with
34
35 expression plasmids for both constructs. Neither protein alone activated reporter
36
37 expression, but when Gal4-DIG0 and DIG0-VP16 were co-expressed, the addition of 50
38
39 μ M digoxigenin induced a ~20-fold increase in luciferase activity. In comparison to the
40
41 single-component sensor Gal4-DIG0-VP16, the dimeric biosensor had a wider dynamic
42
43 range but generated less than 1% of the reporter activity (Supporting Figure S1a). To
44
45 determine whether destabilizing mutations improved the dynamic range of the
46
47 biosensor, we incorporated a mutation (E83V, creating DIG1) that improved ligand-
48
49 induced stabilization in the single polypeptide biosensor context.⁸ DIG1-VP16 paired
50
51
52
53
54
55
56
57
58
59
60

1
2
3 with Gal4-DIG0 improved activation to ~400-fold (Figure 1b). Because E83 lies at the
4
5 dimer interface, this improvement may be driven by a reduction in the self-association of
6
7 DIG1-VP16, making more of it available to interact with Gal4-DIG0. A similar effect was
8
9 not observed when DIG1 was fused to the Gal4 DNA-binding domain or was used in
10
11 both the Gal4 and VP16 fusions (Supporting Figure S1b). Destabilization of the fusion
12
13 containing the Gal4 domain may impair DNA binding, while destabilizing mutations in
14
15 containing the Gal4 domain may impair DNA binding, while destabilizing mutations in
16
17 both sides of the dimer interface may prevent dimerization.
18

19 We next examined whether mutations to the ligand-binding pocket of DIG10.3
20
21 could make the dimers sensitive to an alternative ligand. A series of three tyrosine to
22
23 phenylalanine mutations (Y34F, Y99F, Y101F, designated PRO0) in this pocket
24
25 increase the affinity of DIG10.3 for progesterone, while decreasing it to digoxigenin.⁹
26
27 When the PRO0 domain was individually fused to a DNA-binding domain (Gal4-PRO0)
28
29 and to an activation domain (PRO0-VP16), 1 μ M progesterone induced a ~200-fold
30
31 increase in reporter activity (Figure 1b).
32
33
34

35 As the two ligand-binding domains must bind ligand and dimerize to function as a
36
37 biosensor in this context, we considered whether a simple “AND”-gate could be
38
39 generated by a combination of progesterone and digoxigenin ligand-binding domains.
40
41 The mutations in PRO0 do not fall on the dimer interface, and thus this domain should
42
43 be capable of dimerizing with either DIG0 or DIG1. We paired Gal4-PRO0 with DIG1-
44
45 VP16 and found that luciferase activity was 20-fold greater when both ligands were
46
47 present in the media than when either alone was present (Figure 1c).
48
49
50
51
52
53
54
55
56
57
58
59
60

Biosensor activity is improved by computational redesign of the ligand-binding domain dimer interface

To further improve biosensor function and provide greater control over dimerization, we redesigned the dimer interface to favor heterodimerization between the DNA-binding domain and activation domain fusions over homodimerization of two copies of the same fusion protein. We hypothesized that the dynamic range of the initial dimers was limited by non-productive homodimers reducing the amount of functional complex. Additionally, orthogonal heterodimer interfaces would allow the use of multiple DIG-based biosensors (*e.g.* digoxigenin-specific and progesterone-specific heterodimers) in the same cell without crosstalk between the biosensors.

Using the Rosetta software suite, we redesigned the interface of the DIG0 dimer based on a “charge swap” or “knob-in-hole” model (Figure 2a). The charge swap design introduced the mutations R116E in DIG1 (to generate DIG1a) and E83R in its partner DIG0 (to generate DIG1b). For this design, self-association should be limited by unfavorable charge-charge repulsion. The knob-in-hole design introduced a hydrophobic pocket in DIG0 (I102A, to generate DIG2b) that could be complemented with bulkier side chains in DIG1 (I102M and V104F, to generate DIG2a). For this design, homodimerization should be disfavored by steric clashes and cavities in the dimer interface. Although included in the original designs, the E83V mutation was removed after early testing revealed poor reporter activation.

Biosensors built from both designs showed a >500-fold increase in reporter activation in the presence of digoxigenin (Figure 2b). When we assessed the ability of each of the DIG10.3 mutants to interact with other variants, the heterodimer designs

1
2
3 were largely orthogonal (Figure 2c). All of the charge swap and knob-in-hole fusion
4
5 proteins potently activated reporter expression when paired with their complement and
6
7 showed a weaker ability to activate reporter expression with other mutants or copies of
8
9 themselves. In contrast, the unmutated interface of DIG0 paired well with any other
10
11 mutant. The progesterone-binding mutations in DIG10.3 could be incorporated into both
12
13 interface redesigns to create improved progesterone biosensors (Supporting Figure S2).
14
15
16
17
18

19 **Selection of mutations that impart improved heterodimerization or ligand** 20 21 **selectivity** 22

23
24 Although the designed heterodimers represented a substantial improvement, we
25
26 sought to use directed evolution to identify additional heterodimer pairs and increase the
27
28 selectivity of the biosensors for highly similar steroid compounds. Toward this goal, we
29
30 devised a selection scheme that paired a single interface mutant of the DIG-VP16
31
32 fusion with a library of random Gal4-DIG mutants generated by error-prone PCR. The
33
34 library was selected for beneficial mutations using activation of a *HIS3* reporter gene in
35
36 either the presence or absence of ligand. Because Gal4-DIG0 interacted with most
37
38 interface mutants and these dimers responded to ligands similar to digoxigenin, we
39
40 repeated the selection by varying the fixed interface mutant fused to VP16 and the
41
42 ligand (Figure 3a). By comparing the results of each selection, we could identify
43
44 mutations that either: 1) generally improved dimerization and were enriched under all
45
46 conditions; 2) improved the specificity of heterodimerization and were enriched only with
47
48 a single fixed interface mutant; or 3) improved ligand binding selectivity and were
49
50 enriched differently between ligands. This analysis was aided by the use of deep
51
52
53
54
55
56
57
58
59
60

sequencing of the His⁺ colonies in bulk and comparison of the relative enrichment of mutations under each regime.

We tested 20 different selection conditions. Four fixed interface mutants, DIG1 (E83V), DIG5 (I81F), DIG1a (R116E), and DIG2a (I102M, V104F), fused to VP16 were each paired with the error-prone DIG10.3 library fused to the Gal4 domain. DIG1 was chosen because it dimerized well with DIG0, and DIG5, an interface mutant previously identified as destabilizing,⁸ was chosen to enable selection of mutants complementary to a different part of the dimer interface. DIG1a and DIG2a represented additional interface mutations. Gal4-DIG1b and Gal4-DIG2b were doped into the library as controls, as they require two or more base pair changes and would be unlikely to arise by error-prone PCR. The transformants were plated on solid media with one of five ligand conditions: DMSO (no-ligand control); digoxigenin, at two different concentrations (1 μ M or 10 μ M) to assess the effect of stringency; digitoxigenin (1 μ M), as a highly similar ligand that differs from digoxigenin by only a single hydroxyl group; and progesterone (1 μ M), to identify more selective mutants than PRO0. Enrichment scores were calculated relative to the frequency of mutations in the unselected library.

When we averaged the enrichment scores for each position of DIG10.3 regardless of selection condition, we found that mutations to the overwhelming majority of positions in DIG10.3 were deleterious, but that a handful were enriched (Supporting Figure S3). Mapped to the structure of the DIG10.3 dimer, the enriched mutations were distributed across both the surface and the ligand-binding pocket of the protein. Several mutations were strongly selected with digoxigenin regardless of the fixed interface mutant used in the selection. Of these commonly enriched mutations, mutations to P37

1
2
3 were the most frequent. In the presence of digoxigenin, the addition of a P37S mutation
4 to Gal4-DIG0 and pairing of this mutant with DIG1-VP16 roughly doubled the amount of
5 reporter expression (Figure 3b). Because mutations to this position were so widely
6 selected, we surmised that this benefit was the result of stabilization of the ligand-
7 binding domain. Several other mutations enriched in all of the digoxigenin selections
8 also demonstrated a modest increase in activation by digoxigenin (Supporting Figure
9 S4a). The increase in activation conferred by P37S boosted the sensitivity of both of the
10 heterodimer interfaces (Supporting Figure S4b). To confirm the hypothesis that these
11 mutations stabilized DIG10.3, we tested them in the context of a single component
12 biosensor (Supporting Figure S4c). In contrast with a destabilizing mutation (E83V), the
13 addition of P37S or G95D to a Gal4-DIG0-VP16 construct more than doubled the
14 activity of the biosensors in the absence of ligand. Because this sensor is not
15 dependent upon dimerization, increases in its activity are likely dependent solely on
16 stabilization by these mutations.

17
18 To identify selected mutations that might contribute to an orthogonal heterodimer
19 interface, we compared the relative enrichment scores for positions located at the dimer
20 interface (Figure 3c). Only in the selections using the DIG1a and DIG2a mutants did we
21 recover uniquely enriched mutations at interface positions. Two of these enriched
22 mutations, E83R (*i.e.* DIG1b) and I102A (*i.e.* DIG2b), represented complements created
23 during the computational redesign of the heterodimers. Two similar amino acid
24 changes, E83K and I102V respectively, were also enriched. When paired with DIG1a-
25 VP16, both E83R and E83K mutations in Gal4-DIG performed better than without the
26 E83 mutation (Figure 3d). All of the interface-specific mutations performed better with

1
2
3 their fixed interface partner than did Gal4-DIG0 (Supporting Figure S5), suggesting that
4
5 these mutations were under positive selective pressure.
6

7
8 We also identified two unexpected groups of mutations. The first consisted of
9
10 mutations to the flexible linker between Gal4 and DIG (enumerated in Supporting Figure
11
12 S6) that arose as the result of errors in the Gibson assembly of the libraries. When
13
14 linker mutations in Gal4-DIG0 were tested with DIG1-VP16, they increased biosensor
15
16 activation with digoxigenin by about 40-50% (Supporting Figure S6). This result
17
18 suggests that further improvements to the biosensors might be made through linker
19
20 optimization. The second group of unexpected mutations consisted of stop codons at
21
22 positions 41 and 42, which were enriched with three ligands (Supporting Figure S7a).
23
24 These premature stop codons might be revealing a cryptic activation domain, as
25
26 mutations to acidic amino acids, which are commonly found in activation domains, also
27
28 tended to be enriched in the same selections. A construct of Gal4-DIG0 truncated at
29
30 position 41 was insensitive to the addition of digoxigenin and was a more potent
31
32 activator of luciferase than was full-length DIG0 (Supporting Figure S7b), consistent
33
34 with the truncation revealing a ligand-independent activation domain.
35
36
37
38
39

40 Mutations that conferred ligand selectivity were identified by comparing the
41
42 enrichment scores among selections with digoxigenin, digitoxigenin, and progesterone.
43
44 Three positions in the ligand-binding pocket of DIG10.3 had mutations that were
45
46 selectively enriched (Figure 4a). F58I strongly enriched in nearly all selections with
47
48 digitoxigenin or digoxigenin, but not with progesterone. To determine if this mutant
49
50 enhanced the selectivity of DIG10.3 to bind digoxigenin and digitoxigenin over
51
52 progesterone, we tested it in the context of the single component Gal4-DIG0-VP16
53
54
55
56
57
58
59
60

1
2
3 biosensor (Figure 4b). Activation of this F58I mutant biosensor by digoxigenin and
4
5 digitoxigenin did not change, but the relative activation by progesterone was reduced
6
7 significantly. Mutations to Y99 and Y101 were selectively enriched with progesterone
8
9 and digitoxigenin, but not digoxigenin. Mutations at both of these positions to
10
11 phenylalanine increase the relative affinity for progesterone by eliminating H-bonds to
12
13 digoxigenin.⁹ To test the effect of mutations that were enriched with progesterone but
14
15 not digoxigenin, we assayed F101C and F99N in the context of a single component
16
17 Gal4-PRO0-VP16 biosensor (Figure 4c). Although these mutations reduced the
18
19 absolute activity of this biosensor with progesterone by 7% and 25%, respectively,
20
21 activation by digoxigenin was reduced by more than 50% with either mutation. Despite
22
23 being enriched in a background of DIG0, both mutants were less sensitive to
24
25 digitoxigenin when added to PRO0. Combining both mutations in the same Gal4-PRO0-
26
27 VP16 scaffold resulted in a non-functional biosensor, likely due to collapse of the ligand-
28
29 binding pocket.
30
31
32
33
34

35 The mutations conferring improved selectivity for digoxigenin and progesterone
36
37 provided an opportunity to enhance the selectivity of the dimeric biosensors. When the
38
39 mutations used to create the charge swap heterodimer interface were combined with
40
41 the digoxigenin and progesterone “AND”-gate pair, background activation by the single
42
43 ligands increased substantially, reducing the relative activation by the addition of both
44
45 ligands (Figure 4d). We surmised that the heterodimerizing mutations in this context
46
47 inadvertently affected the selectivity and affinity of the ligand-binding domains for each
48
49 other, making the dimers more sensitive to the presence of a single ligand. When F58I
50
51
52
53
54
55
56
57
58
59
60

and F101C mutations were added to digoxigenin and progesterone ligand-binding domains, respectively, the functional charge swap “AND”-gate was restored.

Optimizing biosensor activity

We considered that the response of the dimeric biosensors might be further improved by adjusting the expression level of each component. We chose four constitutive promoters (*REV1*, *RNR1*, *SSB1*, and *TEF2*) with a range of expression levels¹³⁻¹⁴ that spanned two orders of magnitude (Supporting Figure S8). Using the knob-in-hole heterodimer biosensor, we found that increasing expression of both fusion proteins yielded an increase in biosensor activity (Supporting Figure S9). However, the use of the strongest promoter, *TEF2*, to drive expression of the activation domain fusion was less effective, likely due to toxicity associated with VP16.¹⁵ Increased expression of both components resulted in the most potent activation at lower levels of digoxigenin and a better than 1000-fold activation of luciferase activity with 50 μ M ligand.

With two orthogonal dimeric pairs that respond to different ligands, it became possible to measure the activation of separate reporter genes by the addition of each ligand. We used a knob-in-hole heterodimeric digoxigenin biosensor to drive a firefly luciferase reporter and a charge swap heterodimer progesterone biosensor to drive a *Renilla* luciferase reporter (Figure 5a). For the progesterone biosensor, we replaced Gal4 with a LexA DNA-binding domain and used a modified *GAL1* promoter containing LexA binding sites to drive expression of the *Renilla* luciferase reporter. Each pair of sensor components was cloned onto a single plasmid and transformed into the dual-reporter strain PGFLR14. When single-plasmid biosensors for digoxigenin and

progesterone were present in the same cell, they performed at least as well as their two plasmid counterparts (Figure 5b). The addition of mutations that enhance the selectivity of ligand binding would likely further improve the response of both sensors to the target ligand and improve the separation between activation by one ligand or the other.

Discussion

We show here that biosensors with a dynamic range spanning more than three orders of magnitude can be built from a dimeric ligand-binding domain. Compared to the previous single component transcription factors, the dimers drive expression of less reporter gene activity but have substantially improved dynamic ranges (Supporting Figure S10). Because these proteins must bind ligand, accumulate, and transduce their increased level into the activation of a reporter gene, these biosensors generally have slower kinetics than other types of biosensors, but their simplicity, ease of design, and tunability make them a desirable alternative. Biosensors based on heterodimeric ligand-binding domains may also facilitate new designs for logic-gate construction, an increasingly important utility in microbe engineering. Biosensors constituting an “AND”-gate could be used to create a checkpoint in a biosynthetic pathway that is contingent upon the presence of two metabolic precursors. “AND”-gates built from other homodimeric ligand-binding domains will require the design and engineering of a heterodimeric interface and a new ligand specificity in one of the proteins. The reliance on directed evolution to select for altered ligand-binding domains will likely restrict

1
2
3 binding specificities to ligands that are similar to the original substrate. However, as
4
5 computational protein design technologies improve, it may be possible to engineer *ab*
6
7 *initio* a pair of dimerizing ligand-binding domains that recognize more chemically diverse
8
9 molecules. Furthermore, the demonstration that these biosensors function in yeast
10
11 suggests that they will do so in other eukaryotes, similar to the single fusion
12
13 biosensors.⁸
14
15

16
17 The use of deep sequencing to characterize the output from the different
18
19 selection conditions provides advantageous features. The prevalent selection of
20
21 stabilizing mutations was unsurprising, given that such mutations are the simplest route
22
23 for unstable ligand-binding domains to increase in concentration and survive the
24
25 selection. Because these stabilizing mutations were so heavily enriched, it would have
26
27 been difficult to identify mutations conferring ligand selectivity by sequencing individual
28
29 colonies. Similarly, mutations in the linker between DIG and the DNA-binding domain
30
31 were poorly represented in the library and would not have been easily identified by
32
33 sequencing a handful of individual clones.
34
35
36

37
38 While the directed evolution scheme worked well to identify stabilizing and
39
40 ligand-specific mutations, it was less successful in selecting for complementary
41
42 heterodimer interfaces, e.g. of DIG1 and DIG5. Conversely, the positive enrichment for
43
44 the counterparts of DIG1a and DIG2a illustrates the power of selection to isolate
45
46 complementary interface mutants when a counterpart is available. Saturation
47
48 mutagenesis of the 10 amino acids that form the core of the dimer interface might yield
49
50 more novel complementary interface pairs.
51
52
53
54
55
56
57
58
59
60

Beyond the scope of biosensors, ligand-dependent dimers provide a simple framework for modulating biological functions. For example, ligand-stabilized dimers could see use in much the same way that rapamycin and its analogues have been used to chemically regulate protein-protein interactions and intracellular protein localization.¹⁶ With destabilized ligand-binding domains, dimerization is likely driven entirely by changes in protein concentration upon ligand binding and not by a conformational change. This feature could provide either a benefit or a liability depending on the application. As with transcription factor-based biosensors, the slower rate of activation and destabilization of the fused partners could be problematic, but they may provide a solution for systems suffering from high background activity. Wu *et al.*¹⁷ have demonstrated the importance of these tools to medicine by using an engineered “AND”-gate to activate cancer-killing CAR T-cells. Further development of ligand stabilized dimers for a variety of synthetic biology applications will significantly expand the toolkit for controlling biological interactions.

Materials and Methods

Yeast methods

YPD and synthetic dropout media were prepared as in Dunham *et al.*¹⁸ Yeast were transformed by the Gietz method.¹⁹

Yeast strains

1
2
3 *S. cerevisiae* strains used were derived from the yeast strain PJ74. This strain
4
5 was created by deleting the efflux pump *PDR5* in the yeast two-hybrid strain PJ69-4a.²⁰
6
7 A clean deletion of the open-reading frame was created by the 50:50 method.²¹
8
9

10 A firefly luciferase reporter strain was built from the *S. cerevisiae* strain PJ74. A
11
12 *GAL1* promoter for firefly luciferase was cloned via Gibson assembly into a pUG6
13
14 plasmid,²² adjacent to a NatMX cassette. The reporter construct and neighboring
15
16 antibiotic resistance gene were amplified by PCR with primers containing 50 bp of
17
18 sequence homologous to the *ura3* locus. Linear PCR product was transformed into the
19
20 PJ74 strain and selected for integration on YPD agar plates containing 100 µg/mL
21
22 nourseothricin. Genomic DNA for colony PCR was isolated by boiling cells from a single
23
24 colony in 20 mM NaOH for 10 min, pelleting cell debris by centrifugation, and using the
25
26 supernatant as template in a PCR reaction. Integration was confirmed by Sanger
27
28 sequencing of the of the genomic locus PCR amplicon. The resulting strain was named
29
30 PF22.
31
32
33
34

35 The dual luciferase *S. cerevisiae* strain was built by adding a *Renilla* luciferase
36
37 reporter to the firefly luciferase reporter strain, PF22. A construct encoding a *Renilla*
38
39 luciferase ORF driven by a modified *GAL1* promoter was cloned into a pUG6 plasmid
40
41 adjacent to a KanMX cassette. The modified *GAL1* promoter contained LexA DNA
42
43 binding sites in place of the native Gal4 DNA binding sequences. The *Renilla* reporter
44
45 construct and KanMX cassette were amplified by PCR from the plasmid using primers
46
47 containing 50 bp of homology to the YBR032W locus. The purified PCR product was
48
49 transformed into yeast, and transformants were plated on YPD agar with 200 µg/mL
50
51 G418 and 100 µg/mL nourseothricin. Crude genomic DNA was isolated from single
52
53
54
55
56
57
58
59
60

colonies as above and used to confirm faithful integration by PCR and Sanger sequencing. The resulting strain was name PGFLR14.

Plasmid constructs

A list of plasmids can be found in the Supporting Information (Table S1). DIG10.3 and the destabilized mutants were cloned as DNA-binding domain and transcriptional activation domain fusions by Gibson Assembly.²³ Briefly, PCR fragments for each insert were generated using Phusion High-Fidelity DNA Polymerase (NEB, Waltham, MA) under standard PCR conditions. Amplification primers were ordered from Integrated DNA Technologies and contained 25-30 base overlaps with a neighboring insert or the plasmid backbone. Recipient vectors p415CYC and p416CYC²⁴ were linearized by digestion with BamHI and Sall (restriction enzymes from NEB, Waltham, MA) and purified by DNA Clean & Concentrator kit (Zymo Research, Irvine, CA). DNA-binding domain fusions were cloned into p415CYC, and transcriptional activation domain fusions were cloned into p416CYC. Plasmids were co-transformed into *S. cerevisiae* and were plated on the synthetic dropout media lacking leucine and uracil (SD -leu -ura).

To express different levels of the biosensor components, promoters were cloned into expression plasmids by replacing the *CYC1* promoter in p415CYC and p416CYC plasmids via Gibson assembly. Plasmids were linearized by digestion with SacI and Sall. The *TEF2* promoter was amplified from a p414TEF plasmid²⁴ by PCR. Promoter sequences for *REV1*, *RNR1*, and *SSB1* were amplified from genomic DNA isolated from the PJ69-4a strain. Yeast cells were lysed by boiling in 20 mM NaOH for 10 min,

1
2
3 pelleted by centrifugation, and supernatant was used as template for PCR with Phusion
4
5 polymerase (35 cycles of 98°C for 10 sec, 62°C for 15 sec, 72°C for 45 sec). Assembly
6
7 reactions included both promoter and biosensor component amplicons. DNA-binding
8
9 domain fusions were cloned into a p415 plasmid and transcriptional activation domain
10
11 fusions were cloned into a p416 plasmid.
12
13

14
15 To create plasmids capable of expressing both biosensor components, PCR
16
17 products encoding a complete transcriptional unit (*i.e.* promoter, ORF, and terminator)
18
19 were amplified and cloned by Gibson assembly. Recipient plasmids containing the co-
20
21 expressed biosensor component were digested with *Sall*. Amplification primers were
22
23 designed to insert constructs such that the two ORFs were oriented in parallel, with the
24
25 terminator of the first transcriptional unit next to the promoter of the second. Dimers
26
27 using Gal4 as the DNA-binding domain were cloned into p415, and LexA DNA-binding
28
29 domain constructs into p416.
30
31
32
33
34

35 **Mutagenesis and selection**

36
37 A table of mutations tested and their functions can be found in the Supporting
38
39 Information (Table S2). Mutations were introduced into the DIG domain by Gibson
40
41 Assembly using mutagenic primers. Pairs of primers containing mutant codons and 30
42
43 bp of complementarity were designed as forward and reverse primers from the site of
44
45 mutagenesis and used in Phusion polymerase PCR reactions with primers annealing to
46
47 the 3' and 5' end of each gene respectively. The resulting gene fragments containing
48
49 altered codons were then reassembled in Gibson reactions into either p415CYC or
50
51 p416CYC.
52
53
54
55
56
57
58
59
60

Random mutations were introduced to DIG10.3 with a GeneMorph II Random Mutagenesis Kit (Agilent, Santa Clara, CA). Using primers immediately flanking DIG10.3, mutagenic PCR was performed according to manufacturer's protocol (95°C 30 sec, 68°C 30 sec, 72°C 1 min, 30 cycles) using 160 ng of a p415CYC plasmid containing a Gal4-DIG0 insert as a template. The resulting PCR product was purified and cloned via Gibson Assembly into a p415CYC plasmid containing Gal4 and a multiple cloning site (linearized with HindIII and Sall). The plasmid library was transformed into ElectroMAX DH10B cells (Thermo Fisher Scientific, Waltham, MA) via electroporation per manufacturer's protocol. Transformants were allowed to recover in 1 mL SOC with shaking at 37°C for 1 hr. A 1:1,000 dilution of cells was plated on LB agar containing ampicillin and was found to be comprised of roughly 5×10^6 members. The library was expanded by inoculating the remainder of the transformed cells into 25 mL LB with ampicillin and grown for 18 hr at 30°C. Plasmid library was purified from culture using five Qiaprep Miniprep columns (Qiagen, Hilden, Germany). Sanger sequencing of 10 colonies revealed 2-9 bp changes per DIG, with an average mutation rate of 4 bp changes per clone.

For the selection of Gal4-DIG mutants, yeast cells were first transformed with an interface variant of DIG-VP16, followed by transformation of the mutant library. All selections were performed with the PJ74 strain; this strain also includes a *GAL1* promoter-driven *HIS3* reporter gene at the *LYS2* locus. A single plasmid encoding a DIG-VP16 interface variant was transformed into yeast and plated on SD -ura agar plates. A single colony was picked into 5 mL SD -ura liquid media and grown for 16 hr at 30°C. The cultures were back diluted to an $OD_{600} = 0.1-0.2$ in 50 mL and grown on a

shaker at 225 rpm until the cultures at double at least twice (5-7 hr). Cells were pelleted, washed and transformed with the Gal4-DIG library on a p415CYC plasmid. The transformation was performed as five samples, with each using 500 ng of Gal4-DIG library, supplemented with 0.5 ng of plasmids encoding Gal4-DIG1b and Gal4-DIG2b, and 2×10^8 yeast. The five samples were pooled, mixed, split, and plated in equal volumes on different selective media. Each 14 cm agar plate received roughly 2×10^8 cells and was prepared from SD -ura -leu -his +1 mM 3-aminotriazole, containing either DMSO, 1 μ M digoxigenin, 10 μ M digoxigenin, 1 μ M digitoxigenin, or 1 μ M progesterone (1% final DMSO). Plates were grown for seven days at 30°C before all cells were collected from each plate by gentle scraping. Plasmid was isolated from the selected cells using a Zymoprep Yeast Plasmid Miniprep kit (Zymo Research, Irvine, CA). The transformation and selection was performed once for each of the five interface variants, with the 1 μ M digoxigenin selections performed in triplicate.

Luciferase assays

Single colonies of luciferase reporter strains transformed with biosensor constructs were inoculated into 200 μ L of SD -leu -ura +2% glucose in a clear Costar 96-well culture plate. Replicates of three to five colonies were assayed for each transformant. Cultures were incubated at 30°C for 6 hours, and then the OD₆₀₀ was measured with a Synergy H1 plate reader (BioTek, Winooski, VT). To capture cultures in mid-log growth after overnight incubation, each culture was back-diluted to an OD₆₀₀ = 0.1-0.15 and then serially diluted 1:3 with fresh media three times. Plates were sealed with a Breathe-Easy membrane (Sigma-Aldrich, St. Louis, MO) and incubated at 30°C

for about 16 hours. The density of each culture was measured, and samples at an $OD_{600} = 0.3-0.5$ were used for subsequent inoculations. Each culture was then diluted into a new 96-well plate with fresh media to an $OD_{600} = 0.1$ in 198 μL and induced with 2 μL of 100X ligand in DMSO (1% final DMSO). Culture plates were sealed and incubated at 30°C until the cultures had gone through 3 to 4 doublings (6-7 hr).

To measure firefly luciferase activity, cultures were resuspended by pipetting, and a final OD_{600} was measured for normalizing reporter activity. 5X passive lysis buffer (125 mM Tris-phosphate, pH 7.8; 10 mM DTT; 10 mM EDTA; 50% glycerol, 10% Triton X-100, store at -20°C) was warmed to room temperature and 25 μL was added to wells of a new 96-well plate. 100 μL of induced culture was added to each well and gently mixed by pipetting, carefully avoiding the formation of bubbles. Plates were incubated at room temperature for 20 min.

Firefly luciferase activity was measured by mixing equal volumes of lysed cultures with a firefly luciferase assay reagent (25 mM Tris, pH 7.8; 15 mM MgSO_4 ; 3 mM ATP; 2 mM EDTA; 1 mM DTT; 250 μM coenzyme A; 1 mM D-luciferin). 100 μL of lysed culture and 100 μL of firefly luciferase assay reagent were gently mixed in a white 96-well LUMITRAC 200 plate (Greiner Bio-one, Kremsmünster, Austria) and incubated at room temperature for 5 minutes. Luciferase activity was measured on a Synergy H1 plate reader with a 1 second integration time. The activity of each sample was normalized by dividing the raw luminescence by the final OD_{600} . The average normalized luminescence and standard deviation were calculated from each set of biological replicates.

The dual luciferase assay was adapted from a published protocol²⁵ using a Dual-Glo Luciferase Assay (Promega, Madison, WI). The dual luciferase reporter strain was transformed with plasmids, cultured, and incubated with ligand as with the firefly luciferase assay. After cultures were treated with ligand on 96-well plates and grown for 3-4 doublings, cells were lysed by gently mixing 100 μ L of culture with 25 μ L of 5X Passive Lysis Buffer (Promega, Madison, WI). 50 μ L of this mixture was combined with 50 μ L of the Glo luciferase reagent on a white 96-well LUMITRAC plate and incubated at room temperature for 20 min. Firefly luciferase activity was measured on a Synergy H1 plate reader with a 1 sec integration time. 50 μ L of Stop and Glo reagent was then added to each well, and *Renilla* luciferase activity was measured after the plate was incubated at room temperature for 20 min. The final OD₆₀₀ of each culture prior to measurement of luciferase activity was used to normalize raw luminescence, and average luciferase activities were calculated from 3-5 biological replicates.

Computational redesign

The homodimer interface of computationally designed DIG binder DIG10.2 was redesigned to render it heterodimeric using the Rosetta software suite. Although all computational analysis was performed on DIG10.2 because of its higher resolution crystal structure, redesigned interfaces should translate to the higher affinity variant DIG10.3 used in this study.

DIG10.2 homodimer (PDB ID 4J8T, chains A and B) was relaxed with coordinate restraints to prepare it for design calculations.²⁶ A model of the conditionally unstable first-generation biosensor DIG10.2 E83V/ DIG10.2 heterodimer was then generated by

mutating E83 on chain A to Val and then repacking and minimizing sidechains. This model was used as an input for fixed backbone interface redesign in RosettaScripts.²⁷ Several different design protocols were used, including Monte Carlo interface design calculations optimizing Rosetta total score,²⁸ greedy optimization²⁹ calculations optimizing interface $\Delta\Delta G$ and total score, or Monte Carlo design followed by greedy optimization of the interface core, then short range electrostatic interactions at the interface edge, then long range electrostatic interactions. In all calculations, native residue identity was up-weighted and the amino acids Pro, Cys, Gly, and His were disallowed if non-native. Designs having good interface metrics (Rosetta $\Delta\Delta G$, total score, Holes score,³⁰ and solvent-accessible surface area) were manually inspected and several were selected for further evaluation.

For each selected design, global protein docking using FastRelax refinement³¹ was used to validate the designed heterodimeric interface and evaluate both corresponding undesired homodimeric interfaces. Docking was run on the distributed computing platform Rosetta@home. A successful design was one in which the desired heterodimer had a funnel-shaped docking profile,³² in which Rosetta score is more favorable for docked configurations having low RMSD to the starting design model, and both corresponding homodimers either had non-funnel-shaped docking profiles or had funnel-shaped docking profiles but scored poorly relative to heterodimer. Two designs, the “charge swap” design (chain A: E83V, R116E; chain B: E83R), and the “knob-in-hole” design (chain A: E83V, I102M, V104F; chain B: I102A), were experimentally evaluated.

Sequencing

To prepare samples for high throughput sequencing, amplicons encoding the selected DIG from the Gal4-DIG plasmids were generated using qPCR. PCR reactions were prepared using Phusion polymerase with 1 μ L selected plasmid library, 500 nM each primer, 1X GC Phusion Buffer, 1X SYBR green, and supplemented with an additional 1 mM $MgCl_2$. A control reaction was performed using 50 ng of the naïve Gal4-DIG plasmid library. The forward and reverse amplification primers were designed to anneal 100 bp upstream and downstream of the DIG sequence, with the forward primer annealing to Gal4 in order to reduce capture of the DIG mutant fused to VP16. Samples were run on a MiniOpticon Real-Time PCR System (Bio-Rad, Hercules, CA) and cycled (98°C 10 sec, 72°C 30 sec) until the RFU approached 0.7. The amplification of all selections was performed for 9-17 cycles before being placed on ice. PCR products were cleaned up using a DNA Clean and Concentrator Kit (Zymo Research, Irvine, CA).

Amplicons were prepared for high throughput sequencing by tagmentation. Using a Nextera DNA Library Preparation kit (Illumina, San Diego, CA), 50-100 ng of each library amplicon was fragmented according to manufacturer's protocol. Each tagmented product was purified by a DNA Clean and Concentrator Kit and used in a limited cycle PCR reaction (72°C 3 min, 98°C 30 sec, 7 cycles: 98°C 10 sec, 63°C 30 sec, 72°C 3 min) according to manufacturer's protocol. Uniquely indexed primers were used for each library in order to multiplex sequencing of the libraries. Products were again purified by a DNA Clean and Concentrator Kit and run on a PAGE gel; all libraries showed a smear of sizes ranging from about 150-700 bp. 2 nM of each library was

loaded into a NextSeq 500/550 Mid Output 300 cycle kit (Illumina, San Diego, CA), yielding a sequencing depth of about 90,000-300,000 reads per selected library.

Sequencing results from the library selections were analyzed as a pileup of amino acid changes for each position within DIG. Despite each Gal4-DIG clone containing three amino changes on average, for the analysis we made the assumption that each mutation was selected in isolation. Briefly, sequencing reads were trimmed of their Nextera adapters with Trim Galore!

(https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and overlapping forward and reverse reads were merged using PEAR.³³ Merged and unmerged reads were aligned to a reference DIG sequence using Bowtie 2.³⁴ Using a custom python script, each DNA sequence was translated and used to create a pileup of amino acid counts at each position within DIG. The resulting file contained a tab delimited table indicating the number of times a given amino acid was observed at each position. A complete set of scripts and detailed protocol for file analysis is available online (https://github.com/msr2009/PROTEIN_PILEUP).

Mutational enrichment scores were calculated by comparing the frequency of each mutation appearing in selection relative to its frequency in the naïve mutant library. The frequency of each mutation in the naïve Gal4-DIG library was calculated by dividing the number of counts for a given amino acid change by the sum of counts for all amino acids at that position. A pseudo-count of 0.5 was added to each amino acid count in order to ensure a non-zero frequency. A mutation frequency was similarly calculated from the selected libraries. An enrichment score for each mutation was calculated as the $\log_2$ of the ratio the mutant frequency under selection and the mutant frequency in

the naïve library. In order to filter out low confidence enrichment scores, we required that the number of sequencing counts for any given mutation represent at least 0.1% of the total sequencing reads at a given position in DIG. No score was calculated for the native amino acid. For the selections performed in triplicate (1 μ M digoxigenin), an average enrichment score was determined.

Supporting Information

All Supporting Information is available free of charge at the ACS Synthetic Biology website (<https://pubs.acs.org/journal/asbcd6>) and includes Tables S1, S2 and Figures S1-S10.

Author Contributions

B.W.J. constructed plasmids, built yeast strains, performed reporter assays and selections, and analyzed data. C.E.T computationally designed initial heterodimer pairs. M.S.R. conceived sequencing analysis pipeline and wrote scripts. B.W.J., C.E.T., M.S.R., D.B., and, S.F. wrote the manuscript.

Acknowledgments

We thank Vatsan Raman and Maitreya Dunham for comments on the manuscript. This work was supported by grant P41 GM103533 (to S.F. and D.B.). S.F. and D.B. are investigators of the Howard Hughes Medical Institute.

References

1. Zhang, J.; Jensen, M. K.; Keasling, J. D., Development of biosensors and their application in metabolic engineering. *Curr Opin Chem Biol* **2015**, *28*, 1-8.
2. Rogers, J. K.; Taylor, N. D.; Church, G. M., Biosensor-based engineering of biosynthetic pathways. *Curr Opin Biotechnol* **2016**, *42*, 84-91.
3. Garst, A. D.; Bassalo, M. C.; Pines, G.; Lynch, S. A.; Halweg-Edwards, A. L.; Liu, R.; Liang, L.; Wang, Z.; Zeitoun, R.; Alexander, W. G.; Gill, R. T., Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nat Biotechnol* **2017**, *35* (1), 48-55.
4. Wang, H. H.; Isaacs, F. J.; Carr, P. A.; Sun, Z. Z.; Xu, G.; Forest, C. R.; Church, G. M., Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **2009**, *460* (7257), 894-898.
5. Hallberg, Z. F.; Su, Y.; Kitto, R. Z.; Hammond, M. C., Engineering and In Vivo Applications of Riboswitches. *Annu Rev Biochem* **2017**, *86*, 515-539.
6. Zadran, S.; Standley, S.; Wong, K.; Otiniano, E.; Amighi, A.; Baudry, M., Fluorescence resonance energy transfer (FRET)-based biosensors: visualizing cellular dynamics and bioenergetics. *Appl Microbiol Biotechnol* **2012**, *96* (4), 895-902.
7. Dietrich, J. A.; Shis, D. L.; Alikhani, A.; Keasling, J. D., Transcription factor-based screens and synthetic selections for microbial small-molecule biosynthesis. *ACS Synth Biol* **2013**, *2* (1), 47-58.
8. Feng, J.; Jester, B. W.; Tinberg, C. E.; Mandell, D. J.; Antunes, M. S.; Chari, R.; Morey, K. J.; Rios, X.; Medford, J. I.; Church, G. M.; Fields, S.; Baker, D., A general strategy to construct small molecule biosensors in eukaryotes. *eLife* **2015**, *4*.

- 1
2
3 9. Tinberg, C. E.; Khare, S. D.; Dou, J.; Doyle, L.; Nelson, J. W.; Schena, A.;
4 Jankowski, W.; Kalodimos, C. G.; Johnsson, K.; Stoddard, B. L.; Baker, D.,
5 Computational design of ligand-binding proteins with high affinity and selectivity. *Nature*
6 **2013**, *501* (7466), 212-216.
7
8
9
- 10
11 10. Bick, M. J.; Greisen, P. J.; Morey, K. J.; Antunes, M. S.; La, D.; Sankaran, B.;
12 Reymond, L.; Johnsson, K.; Medford, J. I.; Baker, D., Computational design of
13 environmental sensors for the potent opioid fentanyl. *eLife* **2017**, *6*.
14
15
16
- 17 11. Dou, J.; Doyle, L.; Jr Greisen, P.; Schena, A.; Park, H.; Johnsson, K.; Stoddard,
18 B. L.; Baker, D., Sampling and energy evaluation challenges in ligand binding protein
19 design. *Protein Sci* **2017**, *26* (12), 2426-2437.
20
21
22
- 23 12. Chou, H. H.; Keasling, J. D., Programming adaptive control to evolve increased
24 metabolite production. *Nat Commun* **2013**, *4*, 2595.
25
26
27
- 28 13. Lee, M. E.; DeLoache, W. C.; Cervantes, B.; Dueber, J. E., A Highly
29 Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth Biol* **2015**, *4*
30 (9), 975-86.
31
32
33
- 34 14. Peng, B.; Williams, T. C.; Henry, M.; Nielsen, L. K.; Vickers, C. E., Controlling
35 heterologous gene expression in yeast cell factories on different carbon substrates and
36 across the diauxic shift: a comparison of yeast promoter activities. *Microb Cell Fact*
37 **2015**, *14*, 91.
38
39
40
- 41 15. Gill, G.; Ptashne, M., Negative effect of the transcriptional activator GAL4. *Nature*
42 **1988**, *334* (6184), 721-4.
43
44
45
- 46 16. Putyrski, M.; Schultz, C., Protein translocation as a tool: The current rapamycin
47 story. *FEBS Lett* **2012**, *586* (15), 2097-105.
48
49
50
51
52
53
54
55
56
57
58
59
60

17. Wu, C. Y.; Roybal, K. T.; Puchner, E. M.; Onuffer, J.; Lim, W. A., Remote control of therapeutic T cells through a small molecule-gated chimeric receptor. *Science* **2015**, *350* (6258), aab4077.
18. Dunham, M. J.; Gartenberg, M. R.; Brown, G. W., *Methods in yeast genetics and genomics: a CSHL course manual*. 2015 ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, 2015; p 256.
19. Gietz, R. D.; Schiestl, R. H., High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* **2007**, *2* (1), 31-4.
20. James, P.; Halladay, J.; Craig, E. A., Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **1996**, *144* (4), 1425-36.
21. Horecka, J.; Davis, R. W., The 50:50 method for PCR-based seamless genome editing in yeast. *Yeast* **2014**, *31* (3), 103-12.
22. Guldener, U.; Heck, S.; Fielder, T.; Beinhauer, J.; Hegemann, J. H., A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res* **1996**, *24* (13), 2519-24.
23. Gibson, D. G.; Young, L.; Chuang, R. Y.; Venter, J. C.; Hutchison, C. A., 3rd; Smith, H. O., Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **2009**, *6* (5), 343-5.
24. Mumberg, D.; Muller, R.; Funk, M., Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **1995**, *156* (1), 119-22.

25. Merritt, G. H.; Naemi, W. R.; Mugnier, P.; Webb, H. M.; Tuite, M. F.; von der Haar, T., Decoding accuracy in eRF1 mutants and its correlation with pleiotropic quantitative traits in yeast. *Nucleic Acids Res* **2010**, *38* (16), 5479-92.
26. Nivon, L. G.; Moretti, R.; Baker, D., A Pareto-optimal refinement method for protein design scaffolds. *PLoS One* **2013**, *8* (4), e59004.
27. Fleishman, S. J.; Leaver-Fay, A.; Corn, J. E.; Strauch, E. M.; Khare, S. D.; Koga, N.; Ashworth, J.; Murphy, P.; Richter, F.; Lemmon, G.; Meiler, J.; Baker, D., RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS One* **2011**, *6* (6), e20161.
28. Fleishman, S. J.; Whitehead, T. A.; Ekiert, D. C.; Dreyfus, C.; Corn, J. E.; Strauch, E. M.; Wilson, I. A.; Baker, D., Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science* **2011**, *332* (6031), 816-21.
29. Nivon, L. G.; Bjelic, S.; King, C.; Baker, D., Automating human intuition for protein design. *Proteins* **2014**, *82* (5), 858-66.
30. Sheffler, W.; Baker, D., RosettaHoles2: a volumetric packing measure for protein structure refinement and validation. *Protein Sci* **2010**, *19* (10), 1991-5.
31. Berger, S.; Procko, E.; Margineantu, D.; Lee, E. F.; Shen, B. W.; Zelter, A.; Silva, D. A.; Chawla, K.; Herold, M. J.; Garnier, J. M.; Johnson, R.; MacCoss, M. J.; Lessene, G.; Davis, T. N.; Stayton, P. S.; Stoddard, B. L.; Fairlie, W. D.; Hockenbery, D. M.; Baker, D., Computationally designed high specificity inhibitors delineate the roles of BCL2 family proteins in cancer. *eLife* **2016**, *5*.

- 1
2
3 32. Chaudhury, S.; Berrondo, M.; Weitzner, B. D.; Muthu, P.; Bergman, H.; Gray, J.
4 J., Benchmarking and analysis of protein docking performance in Rosetta v3.2. *PLoS*
5 *One* **2011**, 6 (8), e22477.
6
7
8
9
10 33. Zhang, J.; Kobert, K.; Flouri, T.; Stamatakis, A., PEAR: a fast and accurate
11 Illumina Paired-End reAd mergeR. *Bioinformatics* **2014**, 30 (5), 614-20.
12
13
14 34. Langmead, B.; Salzberg, S. L., Fast gapped-read alignment with Bowtie 2. *Nat*
15 *Methods* **2012**, 9 (4), 357-9.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

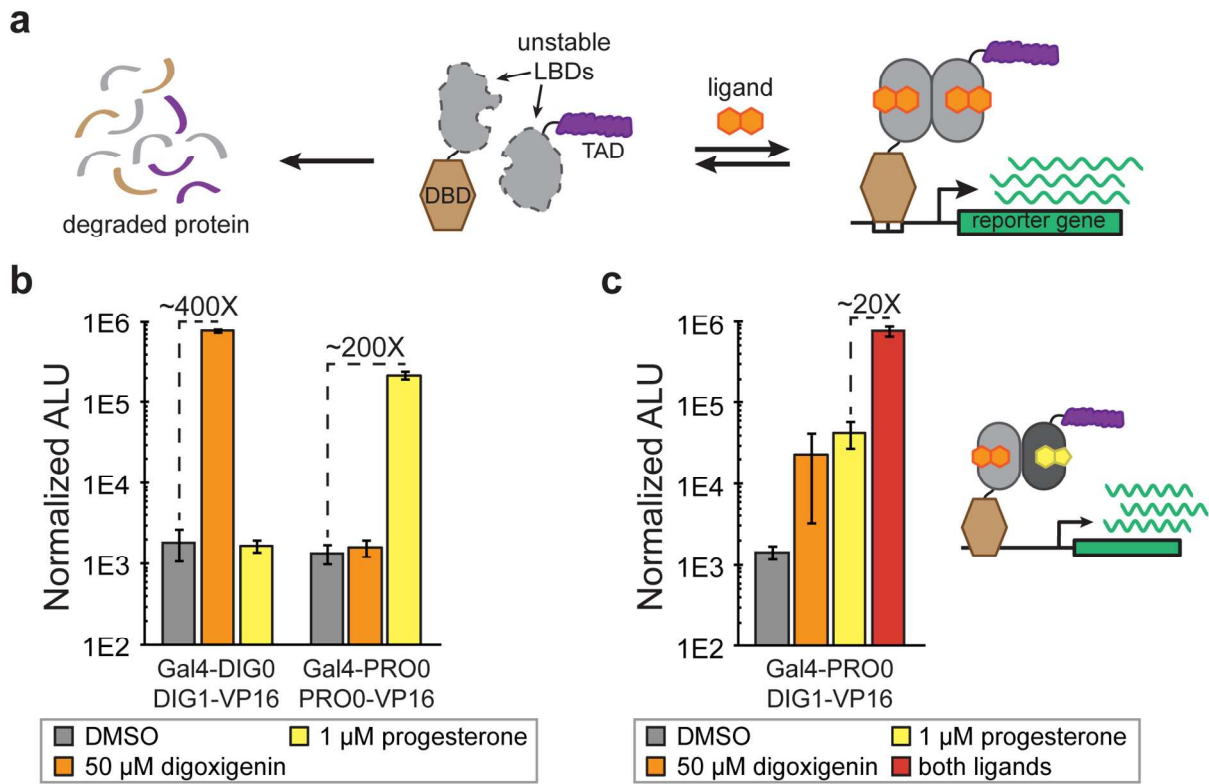


Figure 1: Validation of ligand-dependent stabilization of dimeric biosensors. **a.** Cartoon illustration of the strategy to create biosensors from unstable dimeric ligand-binding domains. In the absence of ligand, the DNA-binding domain (DBD) and transcriptional activation domain (TAD) fused to ligand binding domain (LBD) monomers are degraded and inactive. Addition of ligand stabilizes both proteins, facilitating dimerization and reporter activation. **b.** Activation of a luciferase reporter with a digoxigenin-dependent dimer (Gal4-DIG0 & DIG1-VP16) and a progesterone-dependent dimer (Gal4-PRO0 & PRO0-VP16). Fold activations are shown relative to no ligand (DMSO) **c.** Reporter activation driven by a simple “AND”-gate using Gal4-PRO0 and DIG1-VP16. Stabilization of two different ligand-binding domains, PRO0 by progesterone and DIG1 by digoxigenin, are required for dimerization. Fold activation denotes increase in reporter expression when both ligands are present over just progesterone.

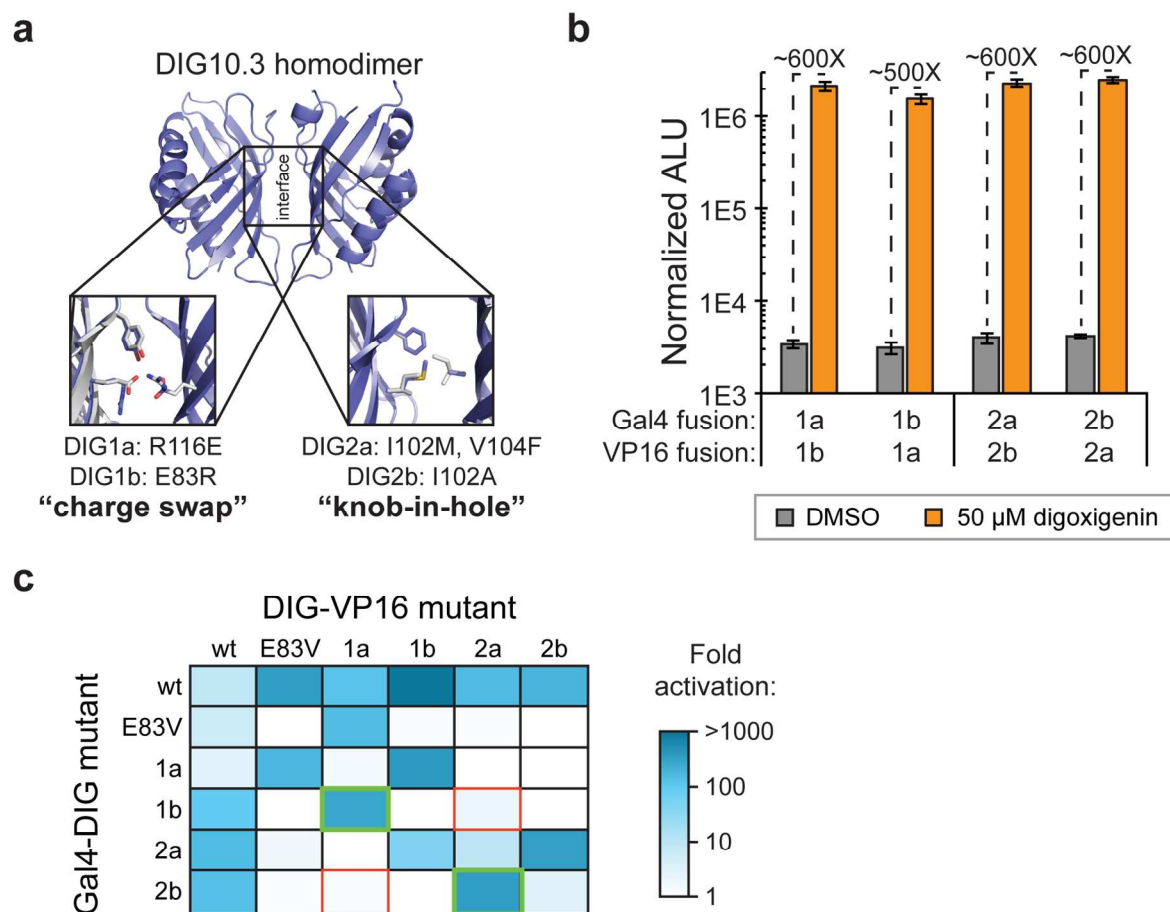


Figure 2: Engineering of dimer interfaces to create orthogonal ligand-binding domain pairs. **a.** Models of the “charge swap” and “knob-in-hole” interface mutants made from the dimer structure of DIG10.3 (PDB code: 4J9A). Both the native (white) and mutant (blue) residues are shown. **b.** Function of dimer interface mutants on dimeric biosensors. Each mutant pair of DNA-binding domain and activation domain fusions was tested with or without digoxigenin for activation of a luciferase reporter. **c.** Heatmap of fold activation of a luciferase reporter with all possible Gal4 and VP16 fusions to DIG mutants, tested with 50 μ M digoxigenin. The interaction between two heterodimer pairs chosen for further improvement are highlighted with green boxes; red boxes illustrate relevant cross reactivity.

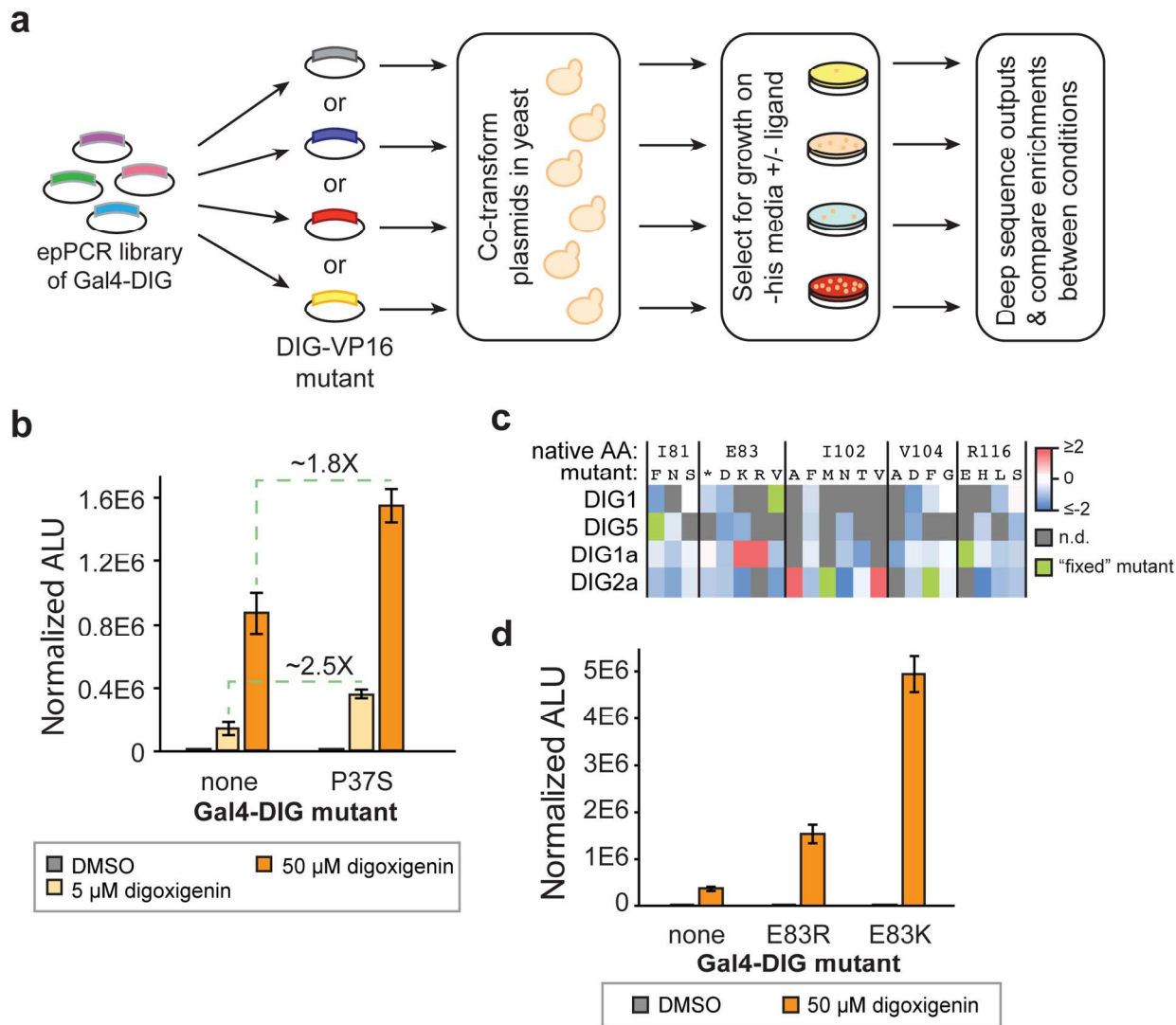


Figure 3: Selection for mutations to improve biosensor function. **a.** Selection scheme used to select for mutants in Gal4-DIG. An epPCR library of DIG mutants fused to Gal4 was co-transformed with each of four “fixed” interface mutants of DIG-VP16. Transformed cells were selected for growth on media lacking histidine and in the presence or absence of a ligand. Cells were collected, deep sequenced, and mutational enrichments compared to identify effects. **b.** Effect of P37S mutant of Gal4-DIG paired with DIG1-VP16 on luciferase activation with two concentrations of digoxigenin. Dotted green line represents fold increase in luciferase activity with P37S. **c.** Heatmap of

1
2
3 enrichment scores for selected interface residues with each of the four “fixed” interface
4
5 mutants and 1 μ M digoxigenin. Mutants with no data are shown in gray, and the “fixed”
6
7 mutation used in the VP16 fusion for selection is denoted in green. **d.** Effect of two
8
9 Gal4-DIG mutants, E83 substituted for either lysine or arginine, when paired with
10
11 DIG1a-VP16. Each was tested for the activation of a luciferase reporter with
12
13
14 digoxigenin.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

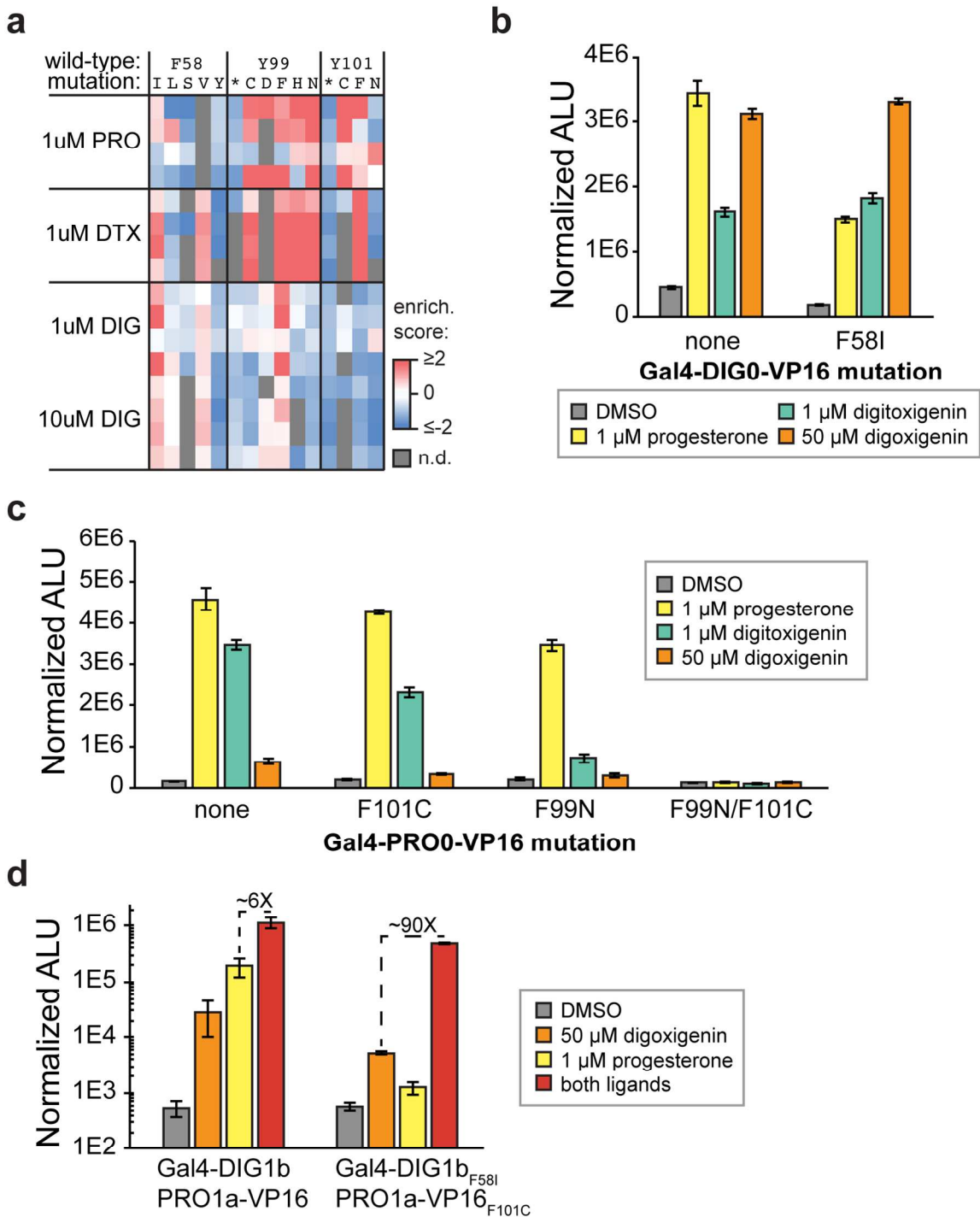


Figure 4: Identification of mutations affecting selectivity of ligand binding. **a.** Heatmap of enrichment scores for mutations to three ligand-binding domain binding pocket residues that exhibited unique profiles based on the ligand used in selection. **b.** Effect of a

1
2
3 digoxigenin-selective mutation, F58I, tested in the context of a single-component, Gal4-
4
5 DIG0-VP16 biosensor activating luciferase reporter expression. **c.** Effect of two
6
7 progesterone-selective mutants, F101C and F99N, tested in the context of a single-
8
9 component, Gal4-PRO0-VP16 biosensor activating luciferase reporter expression. **d.**
10
11 Effect of combining heterodimerizing “chargeswap” mutations and ligand -selective
12
13 binding pocket mutations in both components of an “AND”-gate.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

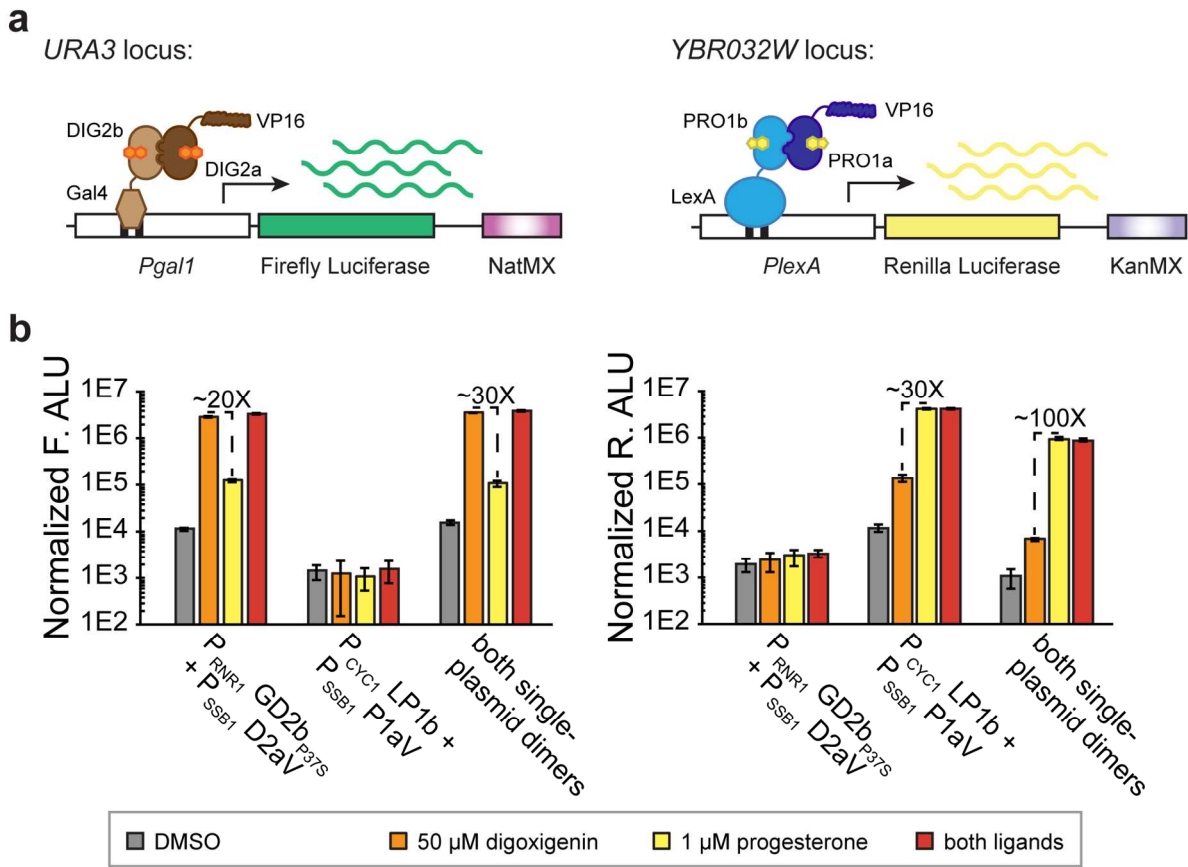


Figure 5: Use of two orthogonal biosensors to detect different ligands in the same cell.

a. Cartoon illustration of the design for orthogonal digoxigenin and progesterone biosensors. A Gal4-dependent firefly luciferase reporter is driven using the knob-in-hole heterodimeric digoxigenin biosensor. A LexA-dependent *Renilla* luciferase reporter is driven by a charge swap heterodimeric progesterone biosensor. **b.** Comparison of the ability for the digoxigenin and progesterone biosensors to drive expression of a firefly luciferase (Normalized F. ALU) and a *Renilla* luciferase (Normalized R. ALU) reporter when the biosensor components are present alone versus when both biosensors are present in the same cell.

For Table of Contents Use Only

Title: Engineered biosensors from dimeric ligand-binding domains

Authors: Benjamin W. Jester, Christine E. Tinberg, Matthew S. Rich, David Baker and Stanley Fields

