

# UC Irvine

## UC Irvine Previously Published Works

### Title

Evolving Small-Molecule Biosensors with Improved Performance and Reprogrammed Ligand Preference Using OrthoRep

### Permalink

<https://escholarship.org/uc/item/1cj2407v>

### Journal

ACS Synthetic Biology, 10(10)

### ISSN

2161-5063

### Authors

Javanpour, Alex A  
Liu, Chang C

### Publication Date

2021-10-15

### DOI

10.1021/acssynbio.1c00316

Peer reviewed



# HHS Public Access

Author manuscript

ACS Synth Biol. Author manuscript; available in PMC 2022 October 15.

Published in final edited form as:

ACS Synth Biol. 2021 October 15; 10(10): 2705–2714. doi:10.1021/acssynbio.1c00316.

## Evolving small-molecule biosensors with improved performance and reprogrammed ligand preference using OrthoRep

Alex A. Javanpour<sup>1,2</sup>, Chang C. Liu<sup>1,2,3,4,\*</sup>

<sup>1</sup>Department of Biomedical Engineering, University of California, Irvine, CA 92697, USA.

<sup>2</sup>Center for Synthetic Biology, University of California, Irvine, CA 92697, USA.

<sup>3</sup>Department of Chemistry, University of California, Irvine, CA 92697, USA.

<sup>4</sup>Department of Molecular Biology & Biochemistry, University of California, Irvine, CA 92697, USA.

### Abstract

Genetically-encoded biosensors are valuable for the optimization of small-molecule biosynthesis pathways because they transduce the production of small-molecule ligands into a readout compatible with high-throughput screening or selection *in vivo*. However, engineering biosensors with appropriate response functions and ligand preferences remains challenging. Here, we show that the continuous hypermutation system, OrthoRep, can be effectively applied to evolve biosensors with high dynamic range, reprogrammed activity towards desired non-cognate ligands, and proper operational range for coupling to biosynthetic pathways. In particular, we encoded the allosteric transcriptional factor, BenM, on OrthoRep such that propagation of host yeast cells resulted in BenM's rapid and continuous diversification. When these cells were subjected to cycles of culturing and sorting on BenM activity in the presence and absence of its cognate ligand, muconic acid, or the non-cognate ligand, adipic acid, we obtained multiple BenM variants that respond to their corresponding ligands. These biosensors outperform previously-engineered BenM-based biosensors by achieving substantially greater dynamic range (up to ~180-fold induction) and broadened operational range. Expression of select BenM variants in the presence of a muconic acid biosynthetic pathway demonstrated sensitive biosensor activation without saturating response, which should enable pathway and host engineering for higher production of muconic and adipic acids. Given the streamlined manner in which high-performance and versatile biosensors were evolved using OrthoRep, this study provides a template for generating custom biosensors for metabolic pathway engineering and other biotechnology goals.

---

\*To whom correspondence should be addressed. Tel: +1 949-824-8534; ccl@uci.edu.

#### Author Contributions

A.A.J. and C.C.L. were responsible for the conception of this study and experimental design. A.A.J. carried out all experiments. A.A.J. and C.C.L. analyzed data and wrote the manuscript.

#### Conflict of Interest Statement

C.C.L. is a co-founder of K2 Biotechnologies, Inc., which focuses on the use of continuous evolution technologies applied to antibody engineering.

#### Supporting Information

Descriptions of key plasmids (Table S1) and strains (Table S2) used in this study; mean background fluorescence measurements of CCM and AA biosensors (Figure S1).

## Keywords

Biosensors; directed evolution; OrthoRep; metabolic engineering; muconic acid; adipic acid

---

## Introduction

Microbial biosynthesis of small-molecules is an attractive alternative to traditional chemical synthesis for a wide array of therapeutically and commercially relevant compounds. For example, yeast have been utilized for the biosynthetic production of the anti-malarial precursor, artemisinic acid;<sup>1</sup> a range of opioids, such as thebaine and hydrocodone;<sup>2</sup> perfume agents, such as santalene;<sup>3</sup> biofuels;<sup>4</sup> and oleochemicals.<sup>5</sup> However, engineering microbes to efficiently produce a target small-molecule often requires substantial modification of heterologous enzymes and host cell metabolism, necessitating search through a vast space of designs. Although large libraries of pathway and host modifications can be readily constructed using modern genetic engineering techniques, finding the most efficient producers is typically low throughput: agar plate screens on colonies, utilizing a visual output or growth complementation, are limited to small library sizes of up to 100,000; and microtiter plate-based screening approaches coupled analytical techniques such as mass spectrometry and liquid/gas chromatography are limited to libraries of up to 10,000.<sup>6,7</sup>

To convert the identification, engineering, and evolution of metabolic pathways into a high throughput exercise, synthetic biologists have focused on the generation of custom genetically-encoded biosensors that can specifically report on the production levels of desired molecules *in vivo*.<sup>8-13</sup> Properly engineered, such biosensors transduce the amount of a target small-molecule into a biological output, such as the strength of reporter gene expression. These outputs can then be used for sorting, as in the case of fluorescence reporters, or selection, as in the case of a selectable gene, forming the crucial link between an arbitrary molecule's production and the high-throughput enrichment of pathways and host modifications that efficiently produce the molecule. The generation of custom biosensors has therefore emerged as an important area of research within the field of metabolic engineering.

A powerful class of potential custom biosensors are allosteric transcription factors (aTFs), which have an effector-binding domain responsible for recognizing cognate ligands and a DNA binding domain that modulates target gene transcription when the cognate ligand is bound.<sup>14,15</sup> Because of their modular architecture and abundance, natural aTFs are attractive starting points for engineering custom biosensor properties, such as regulatory response, regulatory logic, and specificity for non-cognate ligands.<sup>11,12</sup> Recently, Jensen, Keasling, and colleagues demonstrated that the largest family of transcriptional regulators found in prokaryotes, the LysR-type transcriptional regulators (LTTRs), are both portable into workhorse bioproduction hosts such as *Saccharomyces cerevisiae* and highly engineerable for custom biosensor behaviors.<sup>16,17</sup> For example, they showed that the homotetrameric prokaryotic LTTR, BenM, can activate transcription of genes of interest in *S. cerevisiae* when BenM's cognate ligand, *cis,cis*-muconic acid (CCM), is present. This was done by engineering a common yeast promoter, CYC1p, to contain BenM's DNA operator

sequence, *BenO*. This hybrid CYC1p-*BenO* promoter was then used to control GFP expression, allowing for a series of directed evolution campaigns that yielded new BenM variants through fluorescence activated cell sorting (FACS). In these efforts, the Jensen Lab generated BenM variants with increased dynamic and operational range for CCM, inverted regulatory logic, as well as reprogrammed specificity for adipic acid (AA). For example, the most prominent BenM variants identified in Snoek *et al.* displayed a 15-fold increase in output level, a 40-fold shift in operational range, and complete inversion-of-function from ligand-induced activation to repression when compared to the parental BenM biosensor.<sup>17</sup> However, there is need for additional BenM biosensors with expanded dynamic range, operational range, and higher specificity given the diversity of metabolic engineering contexts in which BenMs may be applied.

In this study, we sought to evolve high-performance BenM-based biosensors using the *in vivo* continuous hypermutation system, OrthoRep. By encoding BenM on OrthoRep, straightforward cycles of culturing and FACS in the presence and absence of CCM or AA yielded BenM variants with desired behaviors. Because BenM continuously diversifies when encoded on OrthoRep, we were able to implement many cycles of evolutionary improvement in a straightforward manner, resulting in superior multi-mutation BenM variants with high dynamic range, broad operational range, and improved ligand preference compared to literature benchmarks. This contrasts with past biosensor directed evolution strategies that carried out cycles of FACS enrichment on a static library of BenMs, restricting the improvements accessible.<sup>16,17</sup> While further rounds of mutagenesis and FACS enrichment were feasible in these previous studies, our approach obviates the need for and inefficiencies of *ex vivo* diversification strategies and subsequent library transformation. Combinatorial examination of mutations found in our evolved BenM variants showed that all substitutions characterized improved functionality in at least one context, validating the utility of OrthoRep-driven evolution cycles. Finally, we expressed our evolved BenM variants in the presence of a CCM production pathway to show operational sensing for the goal of evolving CCM and AA metabolic pathways. We argue that the demonstrated effectiveness of OrthoRep-driven BenM evolution combined with the inherent scalability of OrthoRep-based evolution experiments will advance future biosensor engineering campaigns, as the scope of desired biosensors, both in specific molecules sensed and performance across different operational ranges, is vast.

## Results and Discussion

### Directed evolution of BenM with OrthoRep

We sought to demonstrate OrthoRep-driven biosensor evolution by 1) increasing the ability of BenM to sense CCM and 2) reprogramming BenM to respond to AA. To establish the BenM evolution strain, we started with a CEN.PK113-5A *S. cerevisiae* strain previously engineered to express GFP under the control of the hybrid CYC1p-*BenO* promoter, which is activated by BenM in the presence of CCM (MeLS0138). We fully deleted the *URA3* and *TRP1* genes in this parent strain and inserted OrthoRep components by protoplast fusion and transformation following established pipelines.<sup>18-20</sup> The resulting BenM evolution strain includes a recombinant orthogonal p1 plasmid that encodes WT *BenM* along with

a selectable marker (*URA3*) and a nuclear (*CEN6/ARS4*) plasmid that encodes the error-prone orthogonal DNA polymerase, TP-DNAP1–4-2, along with a selectable marker (*TRP1*) (Figure 1A). This evolution strain uses the error-prone orthogonal DNAP to continuously replicate BenM at a high mutation rate of  $10^{-5}$  substitutions per base (spb).<sup>19</sup>

To evolve BenM, we first passaged the evolution strain for ~150 generations to accumulate BenM mutational diversity before introducing selection. Then, FACS-based evolution cycles were initiated. Each cycle consisted of 1) culturing cells in the presence of CCM or AA, 2) a positive sort where the top 0.5% most fluorescent cells were collected, 3) culturing cells in the absence of CCM or AA, and 4) a negative sort to remove any BenM variants that constitutively activate GFP expression (Figure 1B). During culturing steps, BenM autonomously diversifies, allowing cycles of this process to improve BenM activity over time. Over the span of 11 total cycles, with the concentration of ligand increasing from ~7 mM to ~14 mM at cycle 7 to further broaden operational range, the overall behavior of each population adapted as desired: the average fluorescence of the population increased when cultured in the presence of CCM or AA and remained at background levels in the absence of CCM and AA. BenM alleles present in the cultures after cycle 11 were then isolated *via* PCR from population p1 DNA and integrated into the genome of a fresh yeast reporter strain as a single copy (*ura3 0::REV1p-BenM-tTDHI*) through transformation. The reporter strain, which does not include OrthoRep components, encodes GFP under the control of the hybrid CYC1p-BenO promoter. This allowed us to measure BenM activity in a context where BenM is encoded on the yeast genome, replicating how it would be applied as a biosensor for metabolic engineering and biotechnology goals. Ninety-six clones were randomly picked and tested for their ability to respond to ~14 mM CCM or AA, and the top 7 clones with highest dynamic range, defined as the fold-induction of GFP relative to the condition with no ligand present, were sequenced (Figure 1C).

The top 7 BenM variants evolved to sense CCM (CCM-1–7) contained between 6 and 9 mutations including a core set of five substitutions shared among all variants (V88A, A115V, L174S, E223G, and Y257H) (Figure 1C). Additional substitutions that some CCM biosensors contained include D213G, which is structurally located in the dimer-dimer interface of BenM;<sup>21,22</sup> and a common substitution, A28T, which was hypothesized in a prior structural study to enhance BenM's interaction with the phosphate backbone of DNA.<sup>23</sup> Specifically, it was suggested that replacing A28 with T or S could provide a hydroxyl to enhance BenM's interaction with the phosphate backbone of DNA. Interestingly, other LTTRs contain a T or S at the position homologous to A28.<sup>23</sup> The top 7 BenM variants evolved to sense AA (AA-1–7) contained 5 to 8 substitutions including a core set of four substitutions shared among all variants (D76G, P199T, P201S, and Y293C) (Figure 1C). One of these substitutions, P201S, is at a position involved in the dimer-dimer interface of BenM and was also uncovered in previous work evolving BenM to sense AA; it likely functions to change ligand preference, as P201 forms an interaction with CCM.<sup>17,22</sup> Another substitution, Y293C, occurs at a position previously shown to be critical for binding of a secondary effector, benzoate.<sup>24,25</sup> Specifically, it's been demonstrated that at this secondary effector site, a hydrophobic binding pocket, the carboxyl group of benzoate forms hydrogen bonds to the side-chain hydroxyl of Y293.<sup>24</sup> Interestingly, mutation at Y293 (Y293F) abolishes BenM's benzoate activity.<sup>25</sup> It is unclear,

however, whether this substitution allows AA to bind to this secondary effector site or if it modulates the activity of AA in the primary effector site. Two other substitutions, D76G and G44D, were notable because many LTTRs have a charged or polar amino acid at D76, and G44 is highly conserved.<sup>23</sup> Although D76G and G44D could be neutral or deleterious hitchhiker mutations, we found this was not the case (see below). Overall, substitutions were found throughout BenM, highlighting the wide range of interdomain substitutions that can modulate allosteric regulation of BenM (Figure 1D).

### Evolved CCM and AA biosensors exhibit high dynamic range, operational range, and ligand preference

To fully characterize the performance of biosensors CCM-1–7 and AA-1–7, we obtained *in vivo* ligand response curves using the fluorescence of GFP driven by CYC1p-BenO as the readout (Figure 2A,B). Strains were grown in a range of ligand concentrations up to saturation in biological triplicate, and the fold-induction of fluorescence relative to the uninduced (*i.e.* no ligand) condition was determined for all concentrations (Figures 2C,D and S1). For comparison, WT BenM and a set of benchmark BenM-based biosensors evolved in previous studies were tested alongside CCM-1–7 and AA-1–7. These benchmark biosensors include TM (H110R, F211V, Y286N) and MP17\_D08 (A230V, F253S, Y286N, Y293H), which were previously evolved to sense CCM with improved dynamic range and expanded operational range, respectively,<sup>16,17</sup> and TiSNO120 (A130D, A153G, P201S, E287V), which was previously evolved for sensing AA.<sup>17</sup>

Nearly all variants evolved in this study outperformed benchmark biosensors in their dynamic range and have comparable or broadened operational ranges that span multiple orders of magnitude (Figure 2A–D). For example, CCM-4 had a dynamic range of ~180-fold and CCM-6 had a dynamic range of ~135-fold with a graded response over at least ~0.02–14 mM CCM. This contrasts with TM, with a dynamic range of ~29-fold and a response that saturates at ~2 mM CCM, and MP17\_D08, with a dynamic range of ~13-fold and a response that saturates at ~1 mM CCM. Likewise, AA-5 had a dynamic range of ~180-fold and AA-3 had a dynamic range of ~150-fold with a graded response over at least ~1.5–112 mM AA. This contrasts with TiSNO120 with a dynamic range of ~30-fold and a response that saturates at ~3.5 mM AA. Notably, the large operational range of our AA biosensors, with several capable of discriminating up to ~110 mM AA, resulted from evolution campaigns that only used up to ~14 mM AA. This suggests that continued evolution of our AA biosensors may yield variants that are sensitive to even lower concentrations of AA, although these would have limited value for metabolic pathway evolution, because existing production pathways can already yield AA above the minimum concentrations our AA biosensors can detect.<sup>26</sup>

Several BenM variants we evolved also exhibited high preference for CCM over AA or vice-versa compared to benchmark biosensors. For example, CCM-4 had no detectable activity for AA, in contrast to TM and MP17\_D08 (Figure 2E). Likewise, AA-5 and AA-6 responded poorly to the non-cognate ligand, CCM, such that the saturation response for AA was ~24- and ~10-fold higher than the saturation response for CCM, respectively. This contrasts with TiSNO120 where the saturation response for AA is ~5-fold higher than

the saturation response for CCM (Figure 2F). Notably, we did not perform any negative selection against non-cognate ligands in our evolution campaigns. Therefore, observed improvements in ligand preference were likely the byproduct of the high activities evolved in the presence of cognate ligand, as high activity may be associated with strong ligand binding that results in ligand specialization.<sup>27–31</sup> Taken together, our OrthoRep-evolved BenMs form a new set of *in vivo* biosensors for CCM and AA distinguished by their diverse and exceptional performance. This should translate to versatility in their future application.

### Analysis of biosensor substitutions reveals complex contributions to function

We next sought to determine how individual and collections of substitutions in our evolved CCM and AA biosensors contribute to their functions. Because the evolved CCM-1–7 and AA-1–7 BenM variants each contain 4–9 substitutions, it was impractical to analyze all possible variant combinations. Instead, we divided substitutions into two categories: those shared by (almost) all alleles (common) and those that were unique to only some alleles (secondary). For CCM-1–7, 5 substitutions were shared among all variants and 6 substitutions were shared among 6 of the 7 variants, so we chose to include 6 substitutions as the common set: Y67C, V88A, A115V, L174S, E223G, and Y257H or CAVSGH for short. Likewise for AA-1–7, 4 substitutions were shared among all variants, defining the common set: D76G, P199T, P201S, and Y293C or GTSC for short. We then carried out a series of studies focusing primarily on the CAVSGH and GTSC common sets and, in select cases, the combination of common with secondary substitutions. Yeast reporter strains with GFP driven by CYC1p-BenO were engineered to encode genomically-integrated biosensor variants (*ura3 0::REV1p-BenM\_Variant-tTDH1*) containing various combinations of substitutions and GFP fluorescence was measured for all strains in the presence of no, medium, or high levels of CCM (0 mM, 0.88 mM or 14.08 mM) and AA (0 mM, 7.04 mM, or 112.60 mM) (Figure 3).

For CCM biosensors, we tested how each substitution in the common set contributes to function as an individual substitution in BenM or if removed from the common set. We found that a set of substitutions (C, A, V) increased biosensor response in the presence of CCM but did so in a way that also increased background activity in the absence of ligand as well as activity in the presence of the non-cognate ligand, AA (Figure 3A,B), implying that these individual substitutions can generically enhance BenM activity. However, the common substitutions S, G, and H individually resulted in activity comparable to WT, indicating that these substitutions were modulating BenM activity only in the context of other substitutions (Figure 3A). The subtractive analysis on the CAVSGH common set revealed more nuanced outcomes. CAVSGH had the desired biosensor behavior of high dynamic range for CCM (Figure 3A), including near-zero response in the absence of CCM and minimal sensitivity for the non-cognate ligand, AA (Figure 3B). However, when S, G, or H were individually removed, activity in the absence of CCM substantially increased, which was not the case when either A or V was removed (Figure 3A). This implies that S, G, and H enforce CCM-dependent activity and substantially improve dynamic range, whereas A and V generically increase BenM activity in the context of the other common substitutions. Interestingly, removal of C from CAVSGH resulted in an inactive BenM. This combined with the observation that C alone increased generic BenM activity (Figure 3A,B) implies



that C interacts with at least one other substitution in the common set in an epistatic manner. It is notable that the active biosensor, CCM-3 (Figures 1C and 2A), does not contain C but does contain AVSGH, suggesting that the secondary substitutions in CCM-3, namely A29T or D213G, compensate for the loss of C in the CAVSGH common set. Although we did not characterize the contributions of all secondary substitutions individually, we can conclude that the secondary substitutions in CCM-1–6 contribute additional function to the biosensing properties of the common set. This follows from the fact that CCM-7 is identical to the CAVSGH common set (Figure 1C) and CCM-1–6, containing secondary substitutions, all have greater dynamic range, higher response at saturating CCM concentrations, and larger operational range than CCM-7 (Figure 2A).

For AA biosensors, we comprehensively tested all 16 combinations of substitutions found in the GTSC common set (Figure 3C,D). We found that all individual substitutions in the GTSC common set are important for increased AA biosensor activity (Figure 3D). However, G and C showed activity increases in the presence of both CCM (Figure 3C) and AA (Figure 3D) whereas S and (to a lesser extent) T showed pronounced activity increases only in the presence of AA, suggesting that T and S are common substitutions responsible for changing the ligand preference of BenM to AA. When combined, GTSC exhibited high activity in the presence of AA but also appreciable activity in the absence of any ligand and in the presence of the non-cognate ligand, CCM (Figure 3C). This high background activity and/or low ligand selectivity was observed in all combinations of G, T, S, and C except the TS variant, providing further evidence that T and S are primarily responsible for reprogramming BenM's ligand preference. Likewise, high background and low ligand selectivity was exaggerated in the GC variant, supporting the conclusion that G and C increase the activity of BenM in a generic manner. Overall, we observed that the functional contributions of substitutions in the GTSC common set could be qualitatively explained by an additive model wherein the activity of BenMs in the presence of AA largely increased with the number of common substitutions and the background activity of BenMs was lowered by T and/or S, which also lowered activity in the presence of AA to a small extent in rare cases.

Secondary substitutions found in the AA biosensors were important for further reducing the background activity of the GTSC common set in the absence of ligand (Figure 3E,F). For example, G44D, T82I, I188N, S284P, and A294V, improved the dynamic range of GTSC by reducing background without substantially changing response in the presence of AA (Figure 3F). This suggests a model of evolution where there are a small number of large-effect substitutions that increase the induced activity of BenM, and a more diverse number of small-effect background-lowering substitutions accumulated throughout the mutational trajectories. This model is consistent with the higher stringency of positive selection steps versus negative selection steps during our evolution campaign. Taken together, these data reveal the subtleties of how OrthoRep-driven evolution achieved a diverse set of high-performance BenM biosensors that respond to CCM and AA and provide an additional collection of BenM variants with intermediate activities that may be useful in biosensor applications.



## Evolved BenM variants show broad operational range in presence of pathway enzymes producing CCM

To show that our evolved BenM variants are applicable in the metabolic engineering of CCM and AA production pathways *in vivo*, we tested the performance of our biosensors in the presence of an established biosynthetic pathway for the production of CCM in *S. cerevisiae*.<sup>16</sup> The goal of this experiment was to demonstrate that our evolved biosensors 1) could properly report on the biosynthesis of CCM or AA and 2) possess an operational range that extends wide enough such that improvements in the biosynthesis pathway could be robustly detected in a high-throughput screen or selection. With these criteria in mind, we chose to test a subset of our CCM (CCM-3, CCM-4, and CCM-6) and AA (AA-3, AA-5, and AA-6) biosensors that exhibit high dynamic ranges, diverse operational ranges, and preference for their cognate ligand (Figures 1C and 2).

We genomically integrated the selected BenM variants (*ura3 0::REV1p-BenM\_Variant-tTDH1*) into a strain containing 1) GFP driven by the hybrid CYC1p-BenO promoter and 2) a CCM biosynthesis pathway consisting of codon-optimized AroZ from *Podospora anserina* (*PaAroZ*), CatA from *Candida albicans* (*CaCatA*), and AroY from *Klebsiella pneumonia* (*KpAroY*). As previously described, this particular combination of *PaAroZ*, *CaCatA*, and *KpAroY* results in the efficient production of CCM in *S. cerevisiae*.<sup>16</sup> While the addition of an enoate reductase (*ERBC* from *Bacillus coagulans*) can reduce CCM into AA, allowing us to also test our AA biosensors in the context of AA bioproduction, a multi-stage fermentation strategy utilizing a bioreactor in anaerobic conditions is required to produce active enoate reductase and consequent AA.<sup>26</sup> For practical reasons, we decided to forgo direct production of AA, instead only supplying AA exogenously, which would still allow us to assess the operational range and ligand preference of AA biosensors in the presence of the upstream CCM pathway.

As shown in Figure 4A, CCM-3, CCM-4, and CCM-6 were all activated in the presence of the CCM biosynthetic pathway. They also all demonstrated the capacity for substantial further activation for the detection of improvements in CCM biosynthesis. For example, CCM-4 shows an additional ~6-fold activation in response to 7.04 mM exogenous CCM and an additional ~16.5-fold activation in response to 14.08 mM exogenous CCM. The benchmark TM biosensor is also capable of detecting endogenously biosynthesized CCM without saturating response, but only weakly senses the addition of exogenous CCM, while the benchmark MP17\_D08 biosensor and WT BenM are both already saturated by the amount of endogenously biosynthesized CCM (Figure 4A). Compared to the benchmarks, our evolved CCM biosensors should therefore be more versatile in their application to the engineering of CCM biosynthetic pathways.

A similar set of behaviors was found for our AA biosensors. As shown in Figure 4B, variants AA-3, AA-5, and AA-6 all exhibited significant further activation upon addition of AA in the presence of biosynthesized CCM, up to ~60-fold, ~65-fold, and ~19-fold for AA-3, AA-5, and AA-6, respectively. Additionally, even when 14.08 mM of exogenous CCM was used to supplement endogenously biosynthesized CCM, AA-3, AA-5, and AA-6 responded at levels well below their corresponding responses in the presence of 14.08 mM AA. This suggests that AA-3, AA-5, and AA-6 discriminate AA over high amounts

of CCM, which should enable their application in engineering AA biosynthesis pathways that produce CCM as an intermediate. For comparison, the benchmark AA biosensor, TiSNO120, already exhibits saturation at AA concentrations below 14.08 mM. Overall, these data suggest that the BenM variants evolved in this study will be useful in the further optimization of CCM and AA production pathways in yeast.

## Conclusion

We have shown here the successful generation of high performance *in vivo* aTF biosensors for CCM and AA using a directed evolution process where a parent biosensor is encoded on OrthoRep for continuous hypermutation *in vivo*. The biosensors we report have substantially improved dynamic range (up to ~180-fold in several cases), operational range, and ligand preference compared to literature precedents, which should expand the application space for CCM and AA biosensing, especially in the context of metabolic engineering. Indeed, we have demonstrated that our evolved biosensors successfully coupled to an existing CCM production pathway while retaining capacity for an additional ~16-fold activation with higher concentrations of CCM. Recently, Jensen *et al.* showed that OrthoRep could be used to drive the evolution of a biosynthetic enzyme in the CCM production pathway, using the MP17\_08 BenM biosensor to guide selection for increased CCM titers.<sup>32</sup> Since the BenM variants from this study have higher dynamic range than MP17\_08 and are more effective at detecting high CCM and AA concentrations, they should be immediately applicable in the further evolution of CCM pathway productivity, as well as AA biosynthesis pathways that use an additional enzyme to convert CCM to AA. In the long run, we envision a general process where OrthoRep is first used to evolve aTF biosensors against desired small-molecules and then used to drive the evolution small-molecule production pathways with the biosensor to guide selection.

We attribute the high performance of our biosensors to the OrthoRep-driven evolution process through which they were generated. Specifically, OrthoRep effected the continuous hypermutation of BenM, allowing us to easily carry out many cycles of diversification and selection, because the process of evolution required only passaging of cells and FACS-based selection in the presence and absence of desired ligands. In this manner, our evolution campaigns were able to reach complex multi-substitution evolutionary outcomes that encompassed a variety of desired biosensor behaviors. While manually-intensive directed evolution cycles of error-prone PCR on BenM followed by cloning, transformation into yeast, and FACS selection may also be able to yield high performance BenM biosensors, the streamlined nature of biosensor evolution with OrthoRep is well matched to the demands of the broader biosensor engineering field. There are many small-molecules beyond CCM and AA for which effective aTF biosensors are lacking and for each distinct small-molecule, multiple concentration ranges need to be sensed. OrthoRep-driven evolution may prove uniquely scalable to match such breadth of target functions in the biosensor engineering space. Indeed, it may be possible to encode libraries of biosensor parents onto OrthoRep, such as those containing all natural LTTR effector-binding domains, in order to maximize the chance of finding initial biosensor activity against any small-molecule. Subsequent cycles of OrthoRep-driven evolution would automatically improve towards high dynamic range and hone towards desired operational ranges as demonstrated already with BenM.

These directions will guide future efforts in our lab and others, with the ultimate goal of having effective *in vivo* biosensors for all small-molecules of interest.

## Methods

### DNA plasmid construction

Plasmids used in this study are listed in Table S1. *E. coli* strain Top10 was used for all the DNA cloning steps. All primers used in this study were purchased from IDT. All enzymes for PCR and cloning were obtained from NEB. All plasmids were cloned via Gibson assembly.<sup>33</sup> Cloning of sgRNAs was performed as described by Ryan et al.<sup>34</sup>

### Yeast strains and Media/chemicals

All yeast strains used in this study are listed in Table S2. The auxotrophic selection marker used on p1 (*URA3*) was first fully deleted from the genome via CRISPR/Cas9<sup>34</sup> before integrating genetic cassettes onto p1 or p1/p2 transfer through protoplast fusion. The auxotrophic selection marker for the nuclear plasmid containing TP-DNAP1-4-2 (*TRP1*) was also fully deleted, in this case after protoplast fusion. Spacer sequences were designed using Yeaststriction v0.1.<sup>35</sup> Protoplast fusion was used to transfer OrthoRep into strain TrayBP-A7 and was performed as described before.<sup>20</sup> Yeast strains were grown in standard media including YPD (10 g/L Bacto Yeast Extract; 20 g/L Bacto Peptone; 20 g/L Dextrose) and appropriate synthetic drop-out media (6.7 g/L Yeast Nitrogen Base w/o Amino Acids (US Biological); 2 g/L Drop-out Mix Synthetic minus the appropriate nutrients w/o Yeast Nitrogen Base (US Biological); 20 g/L Dextrose). CCM (Sigma #15992) and AA (Sigma #A26357) were dissolved directly into SC media. Afterwards, the pH was adjusted to 4.0 and the media filter-sterilized. CCM was found to be insoluble above 2 g/L (14.08 mM).

### Yeast Transformation

All transformations were performed via the high efficiency Gietz method.<sup>36</sup> p1 integrations were performed as before.<sup>18,20</sup> For p1 integrations, 2–4 µg of plasmid was linearized by digestion with ScaI, which generated blunt ends containing homologous regions to p1. For *CEN6/ARS4* nuclear plasmid transformations, roughly 100–500 ng of plasmid was transformed. Combinatorial variants were synthesized as gBlocks (IDT) and transformed directly into yeast. Transformants were selected on the appropriate selective solid SC media. Plates were grown at 30°C for 2 days for nuclear transformations and 4–5 days for p1 integrations.

### FACS-based evolution

Initial drift was performed by growing cells in 150 mL of SC-UW (pH 5.8) at 30 °C (250 r.p.m.) with daily 1:128 passages. After 21 passages (~150 generations), FACS-based screening began, consisting of alternating positive and negative sorts. After culture saturation (~20 hrs growth), cells were harvested and washed with a buffer containing HEPES-buffered saline (HBS; 20 mM HEPES pH 7.5, 150 mM NaCl) containing 0.2% maltose. After resuspending in buffer, FACS was performed with a Sony SH800 using a 70 µm Sony Sorting Chip. For positive sorting, cells were grown in SC-UW media (pH 4.0) containing 7.04 mM (rounds 1–6) or 14.08 mM (rounds 7–14) ligand (CCM or AA),

whereas for negative sorting, cells were grown in SC-UW media (pH 4.0) without any ligand. Roughly 50,000,000 events and 20,000,000 events were measured for positive and negative sorting, respectively, for each evolution condition. Fluorescence was measured using a 488 nm laser and the top 0.5% fluorescence cells were recovered for positive sorts, while cells exhibiting background fluorescence were recovered for negative sorts. Cells were recovered in 40 mL of SC-UW (pH 4.0) until saturation (30 °C, 250 r.p.m.) and were subsequently diluted 1:100 into 40 mL of appropriate media, as described above, for the next round of selection.

### Flow cytometry

All measurements to characterize biosensor responses were taken by flow cytometry. Strains were streaked onto solid media and single colonies picked into 400  $\mu$ L of SC or SC-UW media (pH 4.0). After 24 hrs growth (30 °C, 750 r.p.m.), cultures were diluted 1:100 into control (SC or SC-UW, pH 4.0) or induction media (SC or SC-UW + CCM/AA, pH 4.0). After 21 hrs growth (or 16 hrs for CCM-pathway containing strains), cells were diluted into 0.9% NaCl and measured on an Attune NxT Flow Cytometer (Life Technologies). Fluorescence of GFP was measured for 20,000 events, and the mean fluorescence for each population was determined. Fold induction was calculated by dividing mean fluorescence of the induced condition (for a given concentration of ligand) by the mean fluorescence of the uninduced condition. To fit the titration curves, a non-linear least squares regression using a four-parameter model was performed in GraphPad PRISM. Each concentration measured corresponds to the mean of 2 or 3 biological replicates.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgements

We thank M. Jensen (Technical University of Denmark) and E. Jensen (Technical University of Denmark) for insightful discussions and the gift of the BenM gene, reporter strain MeLS0138, and CCM production strain ST4245. We thank M. Jensen for helpful edits and comments on our manuscript. This work was supported by the National Institutes of Health [1DP2GM119163 to C.C.L., R35GM136297 to C.C.L.].

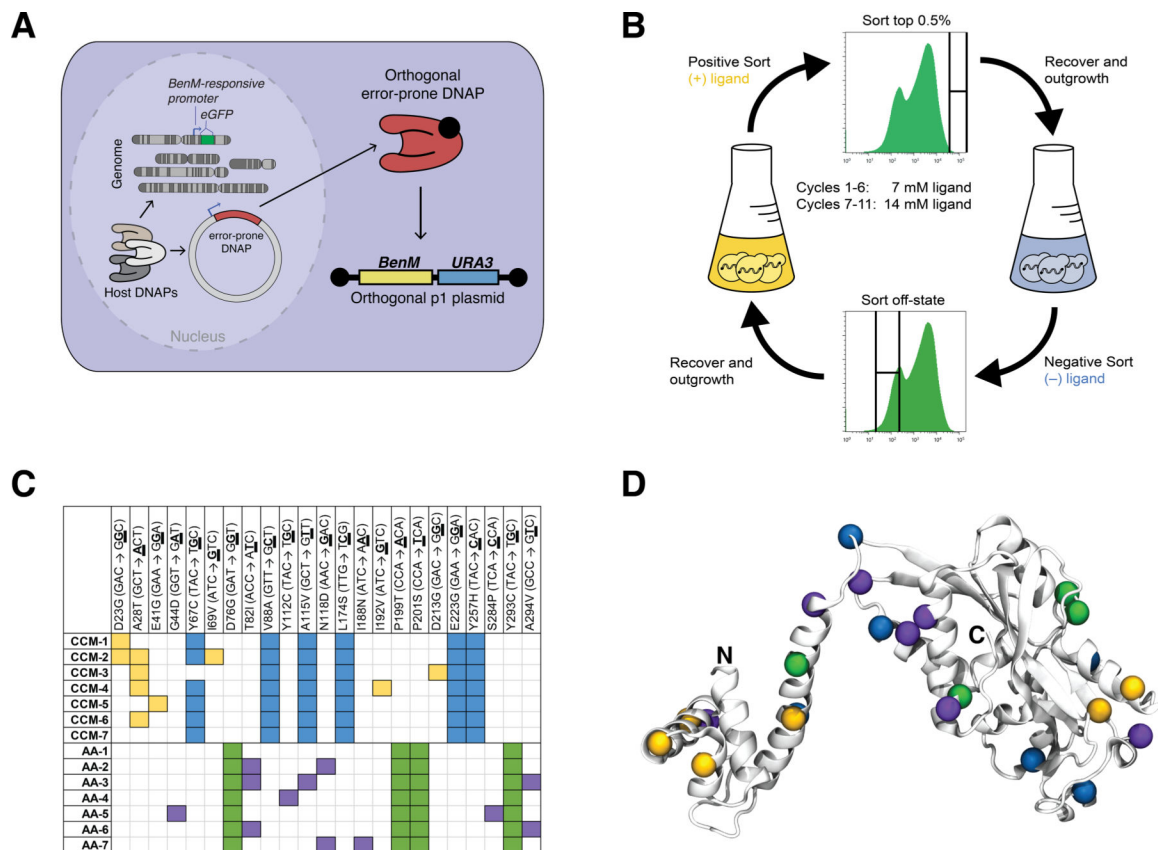
### References

- (1). Ro DK, Paradise EM, Quellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Chang MCY, Withers ST, Shiba Y, Sarpong R, and Keasling JD (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440, 940–943. [PubMed: 16612385]
- (2). Galanie S, Thodey K, Trenchard IJ, Interrante MF, and Smolke CD (2015) Complete biosynthesis of opioids in yeast. *Science* 349, 1095–1100. [PubMed: 26272907]
- (3). Scalcinati G, Partow S, Siewers V, Schalk M, Daviet L, and Nielsen J (2012) Combined metabolic engineering of precursor and co-factor supply to increase  $\alpha$ -santalene production by *Saccharomyces cerevisiae*. *Microb. Cell Fact.* 11, 117. [PubMed: 22938570]
- (4). Buijs NA, Siewers V, and Nielsen J (2013) Advanced biofuel production by the yeast *saccharomyces cerevisiae*. *Curr. Opin. Chem. Biol.* 17, 480–488. [PubMed: 23628723]
- (5). Zhou YJ, Buijs NA, Zhu Z, Qin J, Siewers V, and Nielsen J (2016) Production of fatty acid-derived oleochemicals and biofuels by synthetic yeast cell factories. *Nat. Commun.* 7, 11709. [PubMed: 27222209]

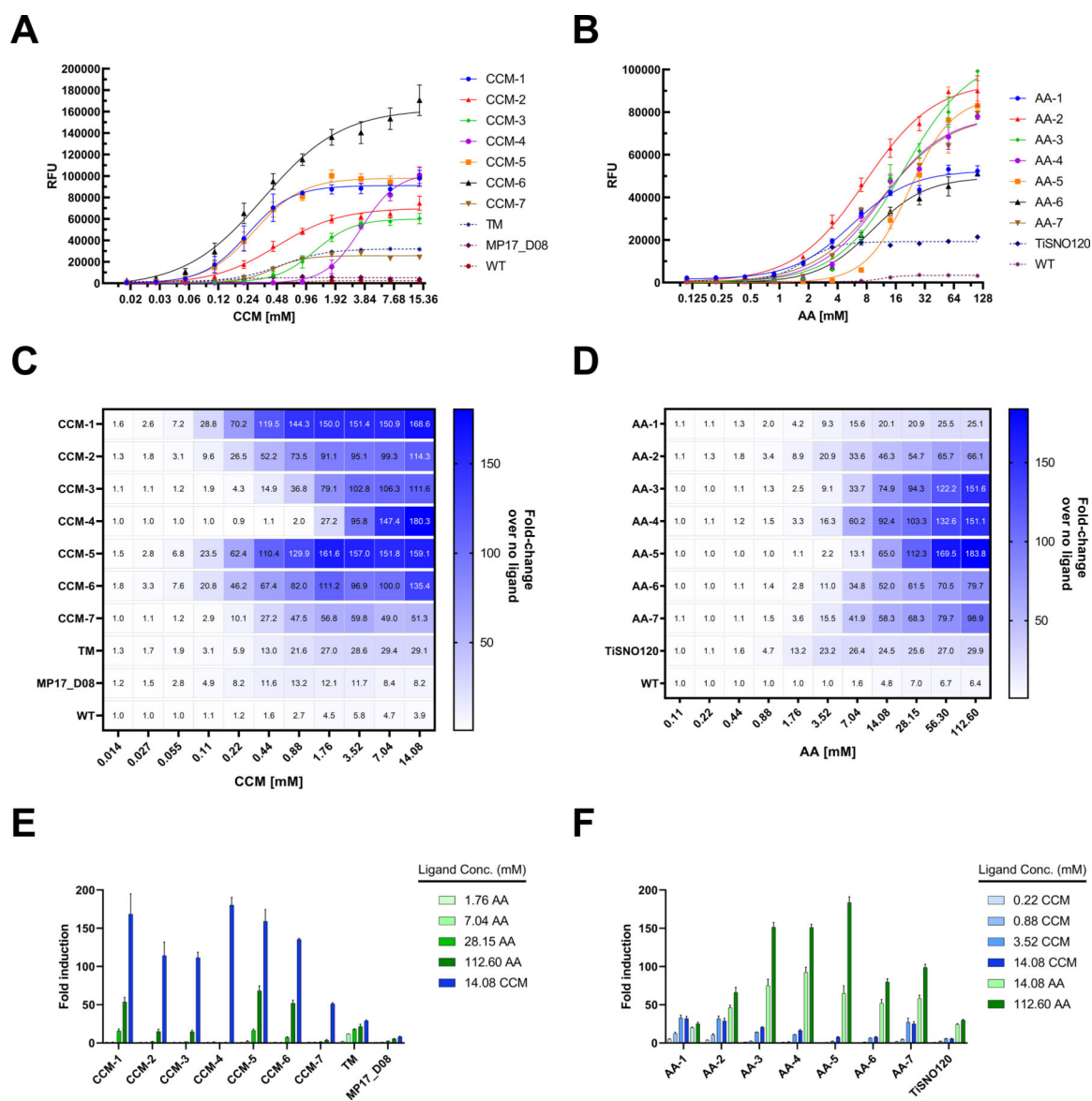
- (6). Leemhuis H, Kelly RM, and Dijkhuizen L (2009) Directed evolution of enzymes: Library screening strategies. *IUBMB Life* 61, 222–228. [PubMed: 19180668]
- (7). Xiao H, Bao Z, and Zhao H (2015) High throughput screening and selection methods for directed enzyme evolution. *Ind. Eng. Chem. Res.* 54, 4011–4020. [PubMed: 26074668]
- (8). Rogers JK, Taylor ND, and Church GM (2016) Biosensor-based engineering of biosynthetic pathways. *Curr. Opin. Biotechnol.* 42, 84–91. [PubMed: 26998575]
- (9). Raman S, Rogers JK, Taylor ND, and Church GM (2014) Evolution-guided optimization of biosynthetic pathways. *Proc. Natl. Acad. Sci. U. S. A.* 111, 17803–17808. [PubMed: 25453111]
- (10). Zhang J, Jensen MK, and Keasling JD (2015) Development of biosensors and their application in metabolic engineering. *Curr. Opin. Chem. Biol.* 28, 1–8. [PubMed: 26056948]
- (11). Liu Y, Liu Y, and Wang M (2017) Design, optimization and application of small molecule biosensor in metabolic engineering. *Front. Microbiol.* 8, 2012. [PubMed: 29089935]
- (12). Mahr R, and Frunzke J (2016) Transcription factor-based biosensors in biotechnology: current state and future prospects. *Appl. Microbiol. Biotechnol.* 100, 79–90. [PubMed: 26521244]
- (13). Dietrich JA, Shis DL, Alikhani A, and Keasling JD (2013) Transcription factor-based screens and synthetic selections for microbial small-molecule biosynthesis. *ACS Synth. Biol.* 2, 47–58. [PubMed: 23656325]
- (14). Cuthbertson L, and Nodwell JR (2013) The TetR Family of Regulators. *Microbiol. Mol. Biol. Rev.* 77, 440–475. [PubMed: 24006471]
- (15). Maddocks SE, and Oyston PCF (2008) Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* 154, 3609–3623. [PubMed: 19047729]
- (16). Skjoedt ML, Snoek T, Kildegaard KR, Arsovska D, Eichenberger M, Goedecke TJ, Rajkumar AS, Zhang J, Kristensen M, Lehka BJ, Siedler S, Borodina I, Jensen MK, and Keasling JD (2016) Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast. *Nat. Chem. Biol.* 12, 951–958. [PubMed: 27642864]
- (17). Snoek T, Chaberski EK, Ambri F, Kol S, Bjørn SP, Pang B, Barajas JF, Welner DH, Jensen MK, and Keasling JD (2020) Evolution-guided engineering of small-molecule biosensors. *Nucleic Acids Res.* 48, e3. [PubMed: 31777933]
- (18). Ravikumar A, Arrieta A, and Liu CC (2014) An orthogonal DNA replication system in yeast. *Nat. Chem. Biol.* 10, 175–177. [PubMed: 24487693]
- (19). Ravikumar A, Arzumanyan GA, Obadi MKA, Javanpour AA, and Liu CC (2018) Scalable, Continuous Evolution of Genes at Mutation Rates above Genomic Error Thresholds. *Cell* 175, 1946–1957. [PubMed: 30415839]
- (20). Javanpour AA, and Liu CC (2019) Genetic Compatibility and Extensibility of Orthogonal Replication. *ACS Synth. Biol.* 8, 1249–1256. [PubMed: 31095905]
- (21). Ruangprasert A, Craven SH, Neidle EL, and Momany C (2010) Full-Length Structures of BenM and Two Variants Reveal Different Oligomerization Schemes for LysR-Type Transcriptional Regulators. *J. Mol. Biol.* 404, 568–586. [PubMed: 20932977]
- (22). Ezezika OC, Haddad S, Neidle EL, and Momany C (2007) Oligomerization of BenM, a LysR-type transcriptional regulator: Structural basis for the aggregation of proteins in this family. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 63, 361–368.
- (23). Alanazi AM, Neidle EL, and Momany C (2013) The DNA-binding domain of BenM reveals the structural basis for the recognition of a T-N11-A sequence motif by LysR-type transcriptional regulators. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 69, 1995–2007. [PubMed: 24100318]
- (24). Ezezika OC, Haddad S, Clark TJ, Neidle EL, and Momany C (2007) Distinct Effector-binding Sites Enable Synergistic Transcriptional Activation by BenM, a LysR-type Regulator. *J. Mol. Biol.* 367, 616–629. [PubMed: 17291527]
- (25). Craven SH, Ezezika OC, Haddad S, Hall RA, Momany C, and Neidle EL (2009) Inducer responses of BenM, a LysR-type transcriptional regulator from *Acinetobacter baylyi* ADP1. *Mol. Microbiol.* 72, 881–894. [PubMed: 19400783]
- (26). Raj K, Partow S, Correia K, Khusnutdinova AN, Yakunin AF, and Mahadevan R (2018) Biocatalytic production of adipic acid from glucose using engineered *Saccharomyces cerevisiae*. *Metab. Eng. Commun.* 6, 28–32. [PubMed: 29487800]

- (27). Lundin E, Näsval J, and Andersson DI (2020) Mutational Pathways and Trade-Offs Between HisA and TrpF Functions: Implications for Evolution via Gene Duplication and Divergence. *Front. Microbiol* 11, 2516.
- (28). Stiffler MA, Hekstra DR, and Ranganathan R (2015) Evolvability as a Function of Purifying Selection in TEM-1  $\beta$ -Lactamase. *Cell* 160, 882–892. [PubMed: 25723163]
- (29). Galvão TC, and De Lorenzo V (2006) Transcriptional regulators à la carte: Engineering new effector specificities in bacterial regulatory proteins. *Curr. Opin. Biotechnol.* 17, 34–42. [PubMed: 16359854]
- (30). Shehata L, Maurer DP, Wec AZ, Lilov A, Champney E, Sun T, Archambault K, Burnina I, Lynaugh H, Zhi X, Xu Y, and Walker LM (2019) Affinity Maturation Enhances Antibody Specificity but Compromises Conformational Stability. *Cell Rep.* 28, 3300–3308.e4. [PubMed: 31553901]
- (31). Collins CH, Leadbetter JR, and Arnold FH (2006) Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR. *Nat. Biotechnol.* 24, 708–712. [PubMed: 16715074]
- (32). Jensen ED, Ambri F, Bendtsen MB, Javanpour AA, Liu CC, Jensen MK, and Keasling JD (2021) Integrating continuous hypermutation with high-throughput screening for optimization of *cis,cis*-muconic acid production in yeast. *Microb. Biotechnol Epub* Mar 1, 2021. DOI: 10.1111/1751-7915.13774.
- (33). Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, and Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. [PubMed: 19363495]
- (34). Ryan OW, and Cate JHD (2014) Multiplex engineering of industrial yeast genomes using CRISPRm, in *Methods in Enzymology* (Doudna JA, and Sontheimer EJ, Eds.), pp 473–489. Academic Press.
- (35). Mans R, van Rossum HM, Wijsman M, Backx A, Kuijpers NGA, van den Broek M, Daran-Lapujade P, Pronk JT, van Maris AJA, and Daran JMG (2015) CRISPR/Cas9: A molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 15, 1–15.
- (36). Gietz RD, and Schiestl RH (2007) Quick and easy yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* 2, 35–37. [PubMed: 17401335]

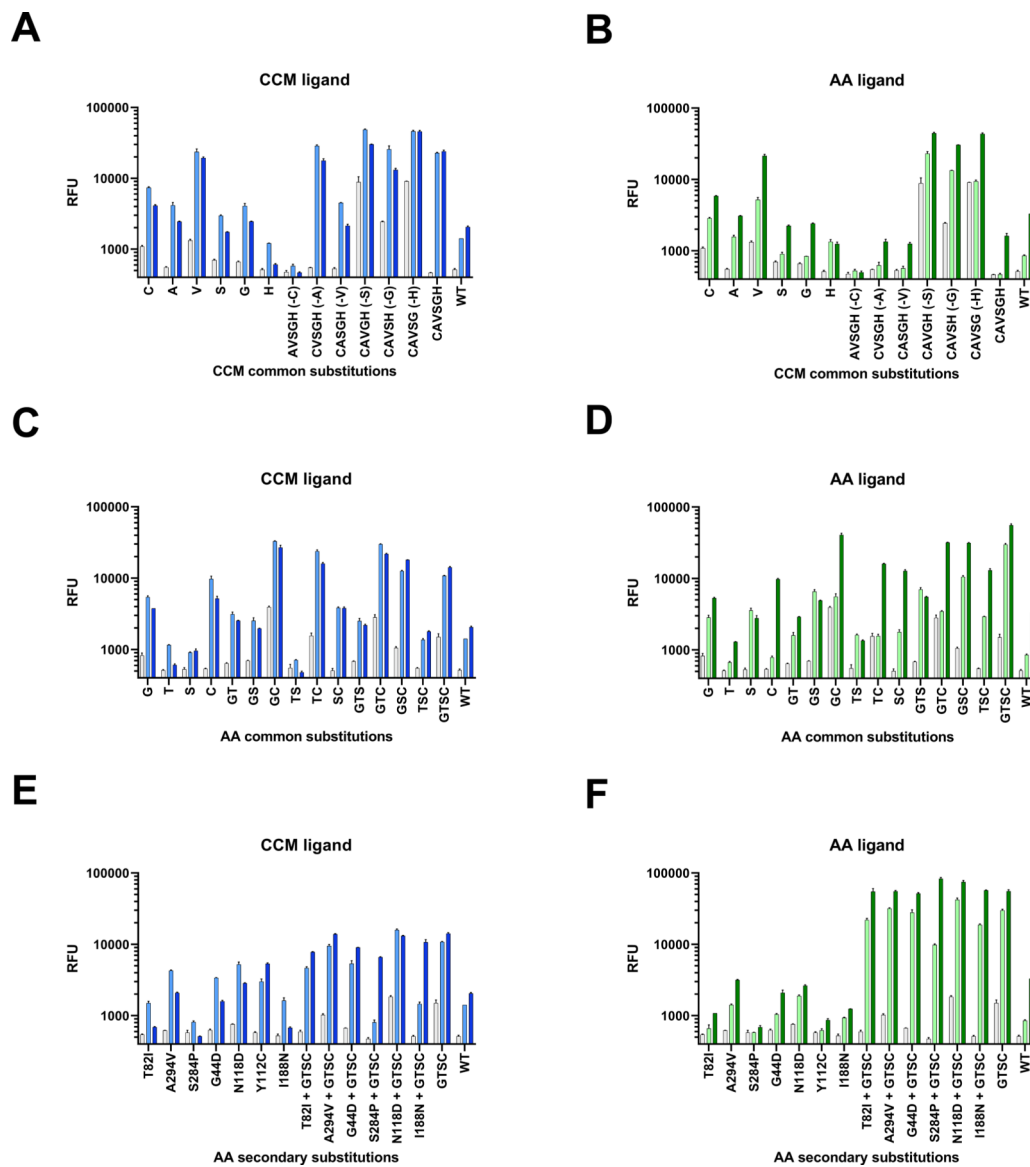




**Figure 1.** Overview of directed evolution of BenM and list of evolved BenM variants. (A) OrthoRep evolution strain. The accessory plasmid p2 is not shown for simplicity. (B) Overview of FACS-based evolution pipeline. (C) Tables showing substitutions in the top 7 CCM and AA hits. Common substitutions are highlighted blue or green (CCM or AA, respectively), and secondary substitutions are highlighted yellow or purple (CCM or AA, respectively). (D) Crystal structure of BenM (PDB: 3K1N) with substitutions in (C) shown.



**Figure 2.** Response function and ligand preference measurements of evolved BenM variants. (A and B) Titration curves of top 7 (A) CCM and (B) AA hits. (C and D) Heat map of titration data displaying mean fold-induction over the uninduced control (0 mM ligand, see Figure S1) for (C) CCM and (D) AA. (E and F) Titration of off-target ligand to measure ligand preference of (E) CCM and (F) AA BenM variants. For all panels, each concentration was measured in biological triplicates, and the mean  $\pm$  one standard deviation (error bars) is shown.



**Figure 3.** Combinatorial reconstruction of variants observed in top 7 CCM and AA hits. (A) All 6 CCM common substitutions as single or quintuple mutants induced with 0 mM (first bar in each set), 0.88 mM (second bar in each set), or 14.08 mM (third bar in each set) CCM. The common substitution that was removed to yield a given quintuple mutant is listed in parenthesis. (B) Same as (A), but induced with 0 mM, 7.04 mM, or 112.60 mM AA. (C) All combinations of the 4 AA common substitutions induced with 0 mM (first bar in each set), 0.88 mM (second bar in each set), or 14.08 mM (third bar in each set) CCM. (D) Same as (C), but induced with 0 mM, 7.04 mM, or 112.60 mM AA. (E) Secondary AA substitutions induced with 0 mM (first bar in each set), 0.88 mM (second bar in each set), or 14.08 mM (third bar in each set) CCM. (F) Same as (E), but induced with 0 mM, 7.04 mM, or 112.60 mM AA. Common substitution abbreviations for (A and B): C, Y67C; A, V88A; V, A115V; S, L174S; G, E223G; H, Y257H. Common substitution abbreviations for (C to F): G, D76G;

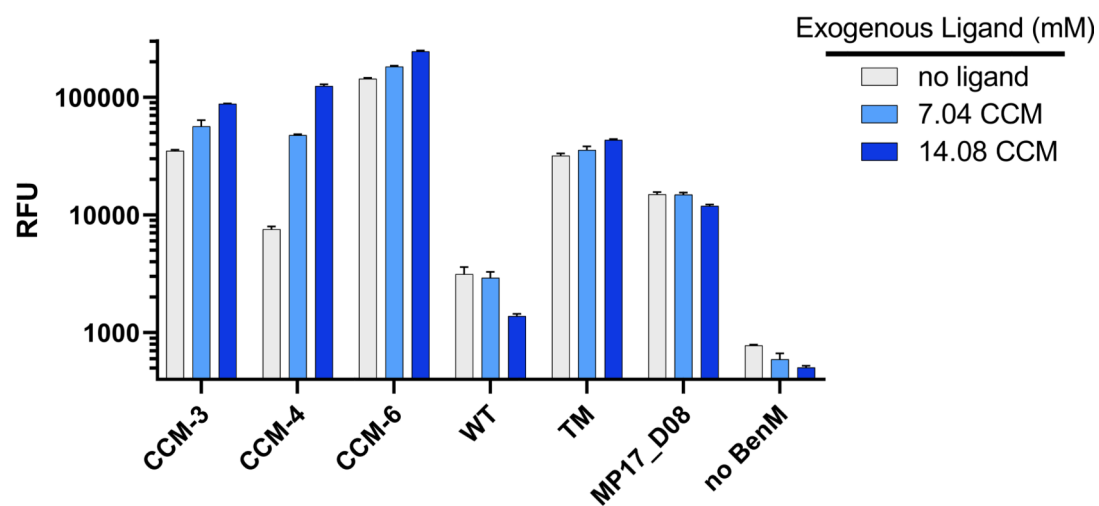
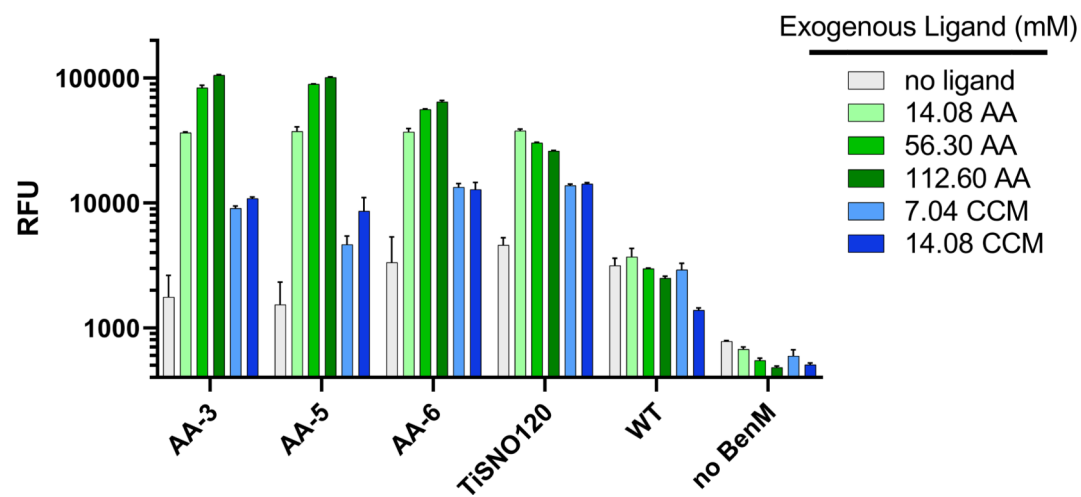
T, P199T; S, P201S; C, Y293C. Each concentration was measured in biological duplicates, and the mean  $\pm$  one standard deviation (error bars) is shown.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**A****B****Figure 4.**

Mean fluorescence measurements of (A) CCM and (B) AA biosensor variants in strains producing endogenous CCM. Exogenous ligand was added to the media alongside cultures without exogenous ligand. Each concentration was measured in biological triplicates, and the mean  $\pm$  one standard deviation (error bars) is shown.