nature chemical biology

ARTICLES

https://doi.org/10.1038/s41589-022-01072-w

Using fungible biosensors to evolve improved alkaloid biosyntheses

Simon d'Oelsnitz[1]✉, Wantae Kim[2], Nathaniel T. Burkholder[1], Kamyab Javanmardi[1], Ross Thyer[3], Yan Zhang[1], Hal S. Alper[2] and Andrew D. Ellington[1]✉

A key bottleneck in the microbial production of therapeutic plant metabolites is identifying enzymes that can improve yield. The facile identification of genetically encoded biosensors can overcome this limitation and become part of a general method for engineering scaled production. We have developed a combined screening and selection approach that quickly refines the affinities and specificities of generalist transcription factors; using RamR as a starting point, we evolve highly specific (>100-fold preference) and sensitive (half-maximum effective concentration (EC_{50}) < 30 μM) biosensors for the alkaloids tetrahydropapaverine, papaverine, glaucine, rotundine and noscapine. High-resolution structures reveal multiple evolutionary avenues for the malleable effector-binding site and the creation of new pockets for different chemical moieties. These sensors further enabled the evolution of a streamlined pathway for tetrahydropapaverine, a precursor to four modern pharmaceuticals, collapsing multiple methylation steps into a single evolved enzyme. Our methods for evolving biosensors enable the rapid engineering of pathways for therapeutic alkaloids.

Microbes have been extensively engineered for commercial-scale production of therapeutic plant metabolites, yielding many benefits over traditional plant-cultivation methods, such as reduced water and land use and faster and more reliable production cycles, and higher purity of target metabolites. Microbial fermentation is currently used for the commercial production of artemisinic acid, the immediate precursor to the antimalarial drug artemisinin, and is in development for scaled-up production of cannabinoids, opiates and tropane alkaloids[1–5]. However, scaling production typically requires several years and hundreds of person years to complete[6] and is largely bottlenecked by reliance on low-throughput analytical methods for assessing strain and pathway performance[7]. We believe that prokaryotic transcriptional regulators can be readily repurposed as biosensors to directly report on compound production and pathway performance in living cells[8,9], but, because methods for generating specific biosensors are lacking, there are virtually no extant biosensors for most plant metabolites. Directed evolution is potentially a starting point for the generation of new biosensor specificities but to date has proven quite limited, yielding improvements in responsiveness only to known effectors or close analogs thereof[10–13].

To overcome this limitation, we sought to exploit a key insight from natural selection, that a protein's substrate promiscuity correlates with its malleability[14]. Thus, by starting with biosensors that have expansive substrate scopes, it should be possible to create biosensors for virtually any compound. In particular, prokaryotic multidrug-resistance regulators, typically studied as mediators of broad-spectrum antibiotic resistance, have large substrate-binding pockets and are known to recognize a raft of structurally diverse lipophilic molecules via nonspecific interactions[15]. Early studies suggest that they may also be highly malleable; notably, just a single point mutation enabled one of these regulators, TtgR, to adopt substantial affinity for the noncognate ligand resveratrol[16]. We hypothesize that these regulators can serve as 'fungible biosensors' that may be readily changeable to adapt to new molecules, pathways and applications.

Using a new directed evolution circuit architecture that relies on both screening and selection, we can seamlessly filter sensor libraries of over 10^5 members into only a few high-performing variants in under 1 week. As proof, we start with a single multidrug-resistance regulator, RamR from *Salmonella typhimurium*, and evolve it to sensitively and specifically recognize five diverse therapeutic alkaloids. The high-resolution structures of these sensors reveal how the malleable effector-binding site can learn to specifically interact with entirely new ligands in wildly different ways. Ultimately, to demonstrate the utility of these sensors as a tool for metabolic engineering, we apply one sensor to engineer a multifunctional plant alkaloid methyltransferase capable of biosynthesizing tetrahydropapaverine (THP), an immediate precursor to four modern pharmaceuticals.

Results

Identifying a benzylisoquinoline alkaloid-responsive multidrug-resistance regulator. We have focused on generating sensors for benzylisoquinoline alkaloids (BIAs) as they (1) are rich in therapeutic activity, (2) have largely resolved biosynthetic pathways and (3) are the subject of ongoing academic and commercial efforts[3,4]. We reasoned that the lipophilic nature of alkaloids might lead multidrug-resistance regulators to display a basal affinity for these compounds. Therefore, we initially targeted five structurally diverse BIAs: THP, papaverine (PAP), rotundine (ROTU), glaucine (GLAU) and noscapine (NOS). These compounds are all therapeutically relevant and commercially available and belong to the structurally distinct benzylisoquinoline (THP and PAP), protoberberine, aporphine and phthalideisoquinoline BIA families, respectively (Fig. 1a and Supplementary Fig. 1). Furthermore, the complete microbial biosyntheses of NOS and ROTU have recently been reported[17,18].

[1]Department of Molecular Biosciences, University of Texas at Austin, Austin, TX, USA. [2]McKetta Department of Chemical Engineering, University of Texas at Austin, Austin, TX, USA. [3]Department of Chemical and Biomolecular Engineering, Rice University, Houston, TX, USA. ✉e-mail: simonsnitz@gmail.com; ellingtonlab@gmail.com

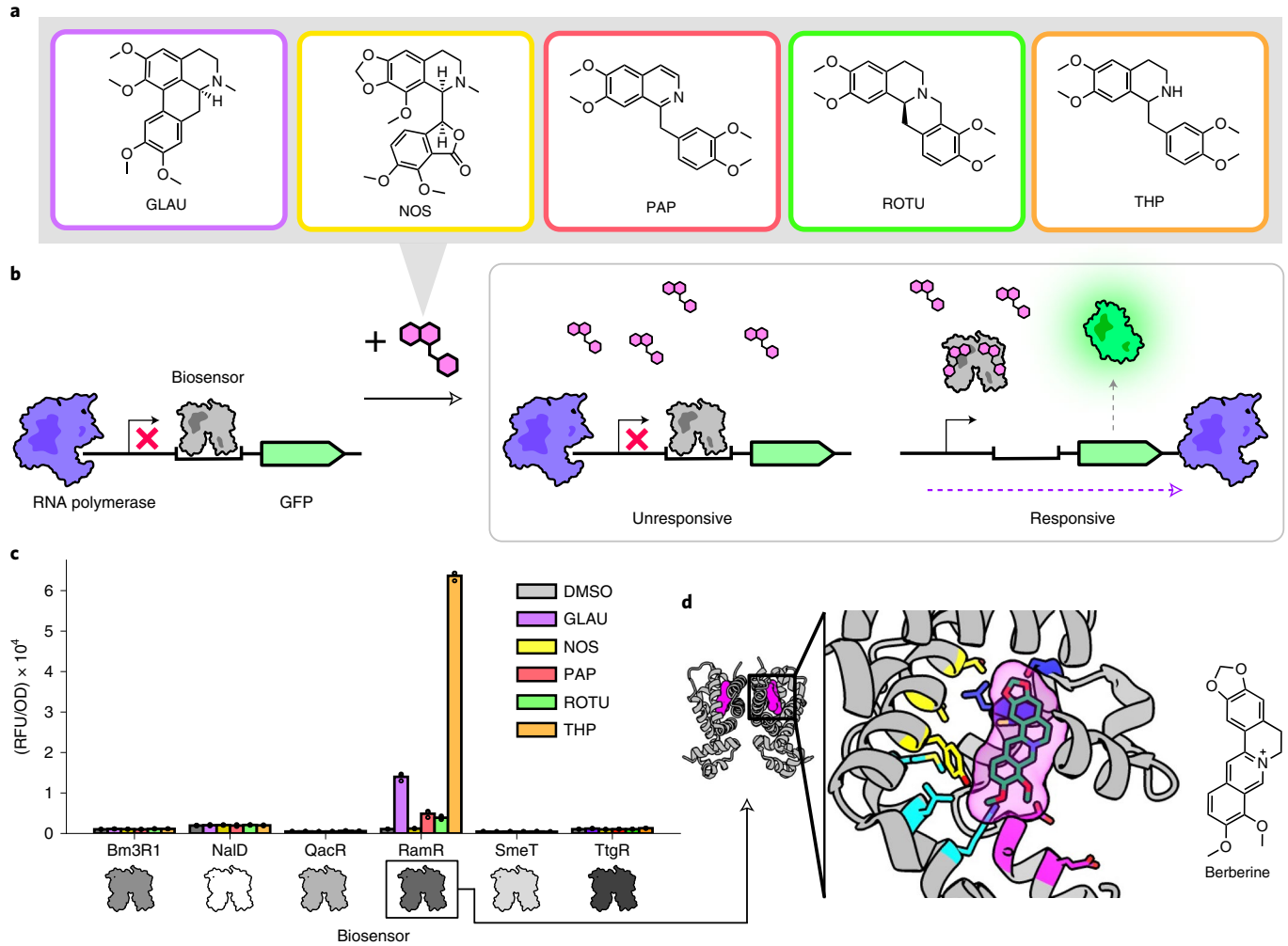


Fig. 1 | Screening identifies a biosensor responsive to BIAs. a, Chemical structures of the five BIAs used in the screen. **b**, Schematic of the genetic circuit used for screening the responsiveness of candidate sensors to target BIAs. **c**, Fluorescence response of six biosensors to all five BIAs. Ligand concentrations used for induction are indicated as follows: GLAU, 1 mM; NOS, 100 μM; PAP, 500 μM; ROTU, 250 μM; THP, 1 mM. Individual fluorescent measurement values for each condition performed in biological triplicate are shown. Data are displayed as mean ± s.e.m. RFU, relative fluorescence unit; OD, optical density. **d**, The global structure (left) and ligand-binding pocket (right) of RamR in complex with berberine (PDB 3VW2). Colored residues were targeted for mutagenesis.

To identify a template biosensor with some degree of BIA affinity, we assayed the responsiveness of six well-characterized multidrug-resistance regulators (QacR, TtgR, RamR, SmeT, NalD and Bm3R1) to the target BIAs. These sensors were chosen because their structures had largely been solved, they regulated multidrug-efflux pumps with corresponding structural features such as large hydrophobic binding pockets with flexible aromatic residues and they had been reported to respond to structurally distinct compounds. Regulators were constitutively expressed from one plasmid (pReg) that was cotransformed with another plasmid bearing the regulator's cognate promoter upstream of superfolder GFP (sfGFP) (pGFP). Sequences of promoters for *qacR* and *ttgR* were obtained from the literature[16,19], while promoters for the remainder were designed by either placing the sensor's operator downstream from a medium-strength promoter (*bm3R1*) or by modifying the −35 or −10 regions of the sensor's native promoter toward the *Escherichia coli* consensus (*nalD*, *smeT*, *ramR*)[20,21] (Supplementary Fig. 2). This design strategy was successful, as each regulator could readily repress transcription as measured by fluorescence (Supplementary Fig. 2). This also allowed further screening with otherwise unknown effectors for the semispecific transcription factors, and RamR from *S. typhimurium* was in fact found to be moderately responsive to many target BIAs (Fig. 1c). In addition, the structure of RamR had already been solved in complex with berberine (Protein Data Bank (PDB) 3VW2), an alkaloid related to our target ligands[22]. These conjoined informatics and experimental efforts thus quickly led to a quite rational approach to library design for directed evolution to improve affinity: five libraries encompassing five separate helices facing the ligand-binding pocket were created in which three residues in each library were site saturated using the degenerate NNS codon (Fig. 1d). Each library contained ~32,000 unique genotypes (32 × 32 × 32), and three to four libraries were pooled before selection, meaning that ~100,000–130,000 unique genotypes were assessed per round of evolution. In addition, error-prone libraries of the entire coding sequence were generated that had an average of two mutations per gene.

Circuit design for biosensor evolution. While biosensors are typically evolved by screening sensor libraries via fluorescence-activated cell sorting in the absence and presence of the target ligand[10–13],

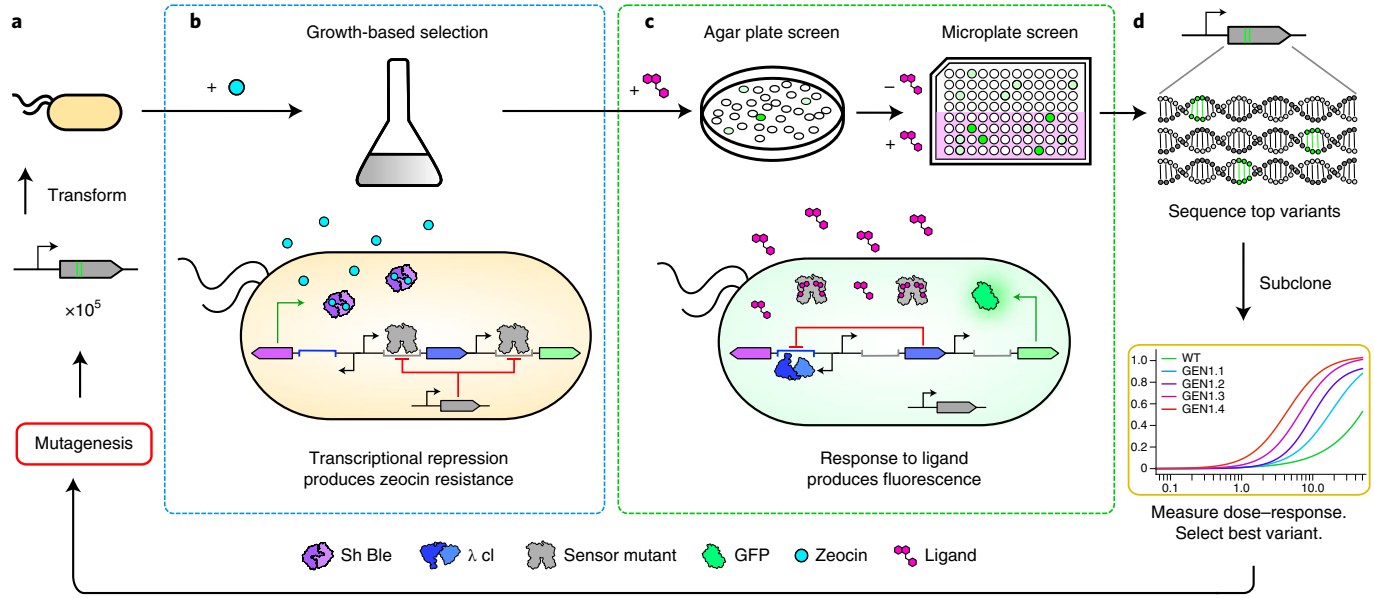


Fig. 2 | The SELIS approach for biosensor evolution. a, Libraries are generated and transformed into *E. coli* cells. **b**, Cells containing the sensor library are cultured in the presence of zeocin. Transcriptional repression by sensor variants prevents the expression of λ cI, which enables the expression of Sh Ble and confers zeocin resistance. Cells containing sensor variants that are unable to repress are eliminated from the population. Adding nontarget ligands at this stage enables counterselection for specificity. **c**, Binding of the sensor variant to the target ligand relieves repression of GFP expression, producing fluorescence. Cultures are plated on an LB agar plate containing the target ligand, and highly fluorescent colonies are cultured overnight. Subsequently, cultures from each picked colony are split and grown either with or without the target ligand. **d**, Variants that display high signal-to-noise ratios are sequenced, subcloned and rephenotyped with a wider range of ligand concentrations. The top-performing variant is then used for the next cycle of evolution. WT, wild type. In the example plot, the *x* axis represents ligand concentration and the *y* axis represents relative fluorescence units.

enrichment for strongly repressing sensor variants can be difficult via sorting due to inadequate filtering resolution and false positives that arise from dead or inactive cells. Selections, however, may enable higher-resolution filtering by enriching strongly repressing sensor variants in live, active cells through exponential amplification. Recent work using both growth-based and polymerase-based selections have already demonstrated these benefits for rapid biosensor evolution[23,24]. In turn, simpler fluorescence-based screens are already well suited to distinguish highly responsive sensors. Given these considerations, we designed a new directed evolution circuit architecture, specifically tailored for sensor evolution, that leverages the advantages of both selections and screens: seamless enrichment of ligand-inducible sensors (SELIS).

To selectively remove sensors with a reduced ability to repress transcription in the absence of the target ligand and variants that were responsive to nontarget ligands, we implemented an inverter circuit involving the λ cI repressor that would lead to expression of the zeocin-resistance protein encoded by the *Sh ble* gene only in the absence of ligand (Fig. 2b). *Sh ble* was chosen for its noncatalytic mechanism of action, enabling a more linear application of selection stringency than would be the case for other antibiotic-resistance elements[25]. Trial selections indeed showed enrichment for functionally repressing RamR variants in a zeocin-dependent manner (Supplementary Fig. 3).

Following this selection, we screened for variants that were more responsive to the target ligand by linking sensor output to the expression of GFP (Fig. 2c). Liquid cultures grown in the presence of zeocin (subjected to the negative selection) were plated directly onto solid medium containing the target ligand. Highly fluorescent clones were isolated and rephenotyped in liquid medium in both the presence and absence of the target ligand to better determine the signal-to-noise ratio for each chosen sensor variant. The stringency of the positive screen could be readily tuned by altering the amount of the target ligand applied on the solid medium.

Ultimately, variants with low background and a high signal-to-noise ratio were sequenced, and unique variants were more fully characterized using a wider range of ligand concentrations (Fig. 2d). The highest-performing biosensor variant was then used as a template for the next round of selection and screening. A library containing ~10^5 variants can be deconvoluted to yield phenotype and genotype data for high-performing clones in only 1 week, without the need for specialized equipment. The SELIS methodology should be broadly applicable for the evolution of virtually any prokaryotic ligand-inducible repressor and possibly even prokaryotic activators as well, and has already been applied to alter the specificity of another prokaryotic repressor, CamR[26].

Evolving RamR for BIAs. While multidrug-resistance regulators are known to recognize structurally diverse ligands, the ability to hone their inherently broad effector specificity has been rarely explored[16]. To demonstrate the utility and speed of SELIS, wild-type RamR was used as a starting point, and four rounds of evolution were performed for the five BIAs (THP, PAP, ROTU, GLAU and NOS) for a total of 20 RamR sensor generations. As library positions became fixed, site-saturation mutagenesis was performed at nonfixed positions to create new libraries for the next round of evolution (Supplementary Fig. 4). Following the first round of directed evolution, the strength of the promoter expressing the RamR variant and the concentration of the target BIA were reduced to increase the selection stringency for repression and ligand responsiveness, respectively (Supplementary Table 1). After the second round of evolution, nontarget BIAs at 100 µM were added during the growth-based selection to eliminate less-specific sensor variants.

Over the course of four generations of evolution, discrete evolutionary lineages became highly sensitive to their cognate BIA

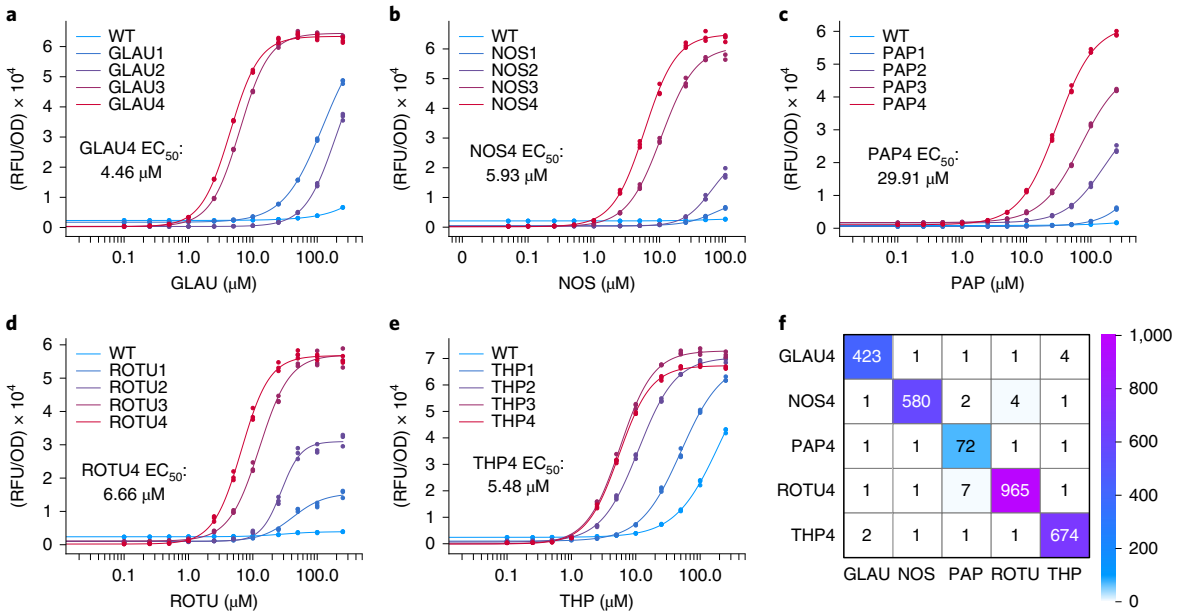


Fig. 3 | Evolution of highly specific BIA sensors from a generalist template. a–e, Transfer functions for all four generations of RamR variants with five different BIAs. The maximum ligand concentration was chosen based on the compound's solubility limit in 1% DMSO (100 μM for NOS and 250 μM for all other BIAs). Individual fluorescent measurements for each condition performed in biological triplicate are shown. **f**, Orthogonality matrix of the final evolved sensors. Fold response is shown for all BIAs for the GLAU4, NOS4, PAP4, ROTU4 and THP4 sensors. The indicated BIAs were applied at 100 μM in all conditions. Measurements for each condition represent the average of three biological replicates.

(Supplementary Figs. 5–9). Despite having a barely detectable response to most target BIAs initially, four of the five final RamR variants had an EC_{50} value under 7 μM, highlighting the plasticity of this biosensor scaffold (Fig. 3a–e). Notably, the detectable concentration range for the final NOS biosensor is well within the level that is reported to be produced de novo in yeast[17]. In addition, the background signal was also reduced to less than 40% of the wild-type RamR signal for four of the five final biosensors (Extended Data Fig. 1). Low background signals generally correlated with an increased signal-to-noise ratio and a better limit of detection for the ligand.

What was most surprising was that, despite starting from the same generalist template, all five final biosensor variants proved to be extremely specific for their matching BIA, displaying >100-fold preference for their cognate BIA over all other noncognate BIAs at solubility-limiting concentrations (100 μM) of each compound (Fig. 3f, Extended Data Fig. 2 and Supplementary Fig. 10). High specificity is crucial for sensors used in strain engineering to avoid false positives arising from cross-reactivity with noncognate ligands, particularly biosynthetic precursors.

Structures reveal molecular adaptations to alkaloids. As both the ligand sensitivity and specificity of RamR could be dramatically transformed throughout evolution, with sensors accumulating from nine to 13 mutations, we sought to better understand the molecular adaptations by solving the crystal structures of the best-evolved sensors. We solved the structures of four of the five evolved sensors in complex with their cognate BIA: PAP4 with PAP (1.6 Å), ROTU4 with ROTU (1.6 Å), GLAU4 with GLAU (2.0 Å) and NOS4 with NOS (2.2 Å) (Supplementary Table 2). The overall folding and dimerization of the evolved variants was highly identical to those of wild-type RamR (Fig. 4a). A strong positive electron density was consistently detected at the binding site for each molecule in the asymmetric unit, which perfectly fit with the BIA chemical structures (Fig. 4b and Supplementary Fig. 11).

BIAs contain heterocycle isoquinoline and benzyl group moieties, with a wide variety of intervening connectivities distinguishing individual compounds. The configuration of each ligand complexed with RamR variants revealed that the isoquinoline component was always 'fixed' beneath Phe155 with π–π stacking interactions, with further recognition by a hydrophobic pocket created by mutations in residues 70, 85, 133 and 134 (Fig. 4c). For example, C134 is consistently changed to leucine to form a hydrophobic interaction with the isoquinoline ring. Similarly, M70 was frequently mutated into a shorter hydrophobic residue (leucine or isoleucine) to reinforce hydrophobic interactions with the BIA ligand. Mutations are not restricted to direct contact with the ligand: the L133I substitution supports other changes at position 85 (PAP4, T85M or L133I; ROTU4, T85I or L133I), making room for bulkier mutations at T85 with higher hydrophobicity. Identification of this common binding pattern and key residues involved in BIA recognition may facilitate structure-guided engineering of further sensors for morphinans and other therapeutic alkaloids.

Most surprisingly, despite the structural similarities among BIA ligands, each BIA biosensor employs a unique mode of recognition to accommodate heteroatoms and the benzyl ring moiety. This is easy to see, given that the binding sites for the benzyl moieties of PAP and NOS extend in entirely different directions (Fig. 4d) and occupy different, newly evolved pockets. Examining the unique benzyl pockets in greater detail, NOS extends into a side pocket for its specificity. The H135Y substitution assists the accommodation of the dimethoxybenzyl moiety by forming a pseudo π–π interaction and participating in the hydrogen bond network associated with the ester group of NOS. The mutation of E120 and D124 into highly flexible arginine residues creates an electrophilic network with H135Y and D152 (Fig. 5a). Unique accommodations are observed for the benzyl moieties of other ligands as well. A Y92G substitution opens a new binding cavity in PAP4, allowing occupancy by the dimethoxybenzyl group of PAP (Fig. 5b). K63Y and L156Y substitutions in ROTU4, along with Y92, form a triple-tyrosine

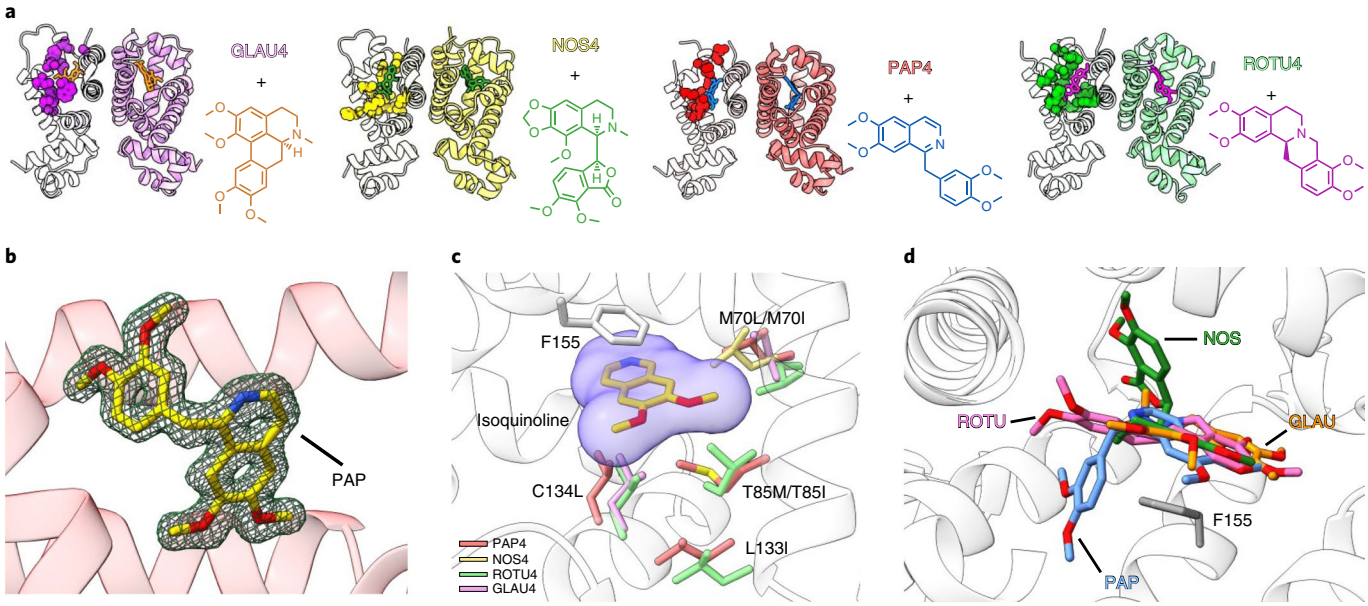


Fig. 4 | Crystal structures of evolved biosensors bound to cognate BIAs. a, Overall structures of the four evolved RamR variants in ribbon diagram. The specific ligands for each variant are shown in stick models with the binding site for one of the monomers shown in the space-filling model to highlight the binding pockets. **b**, Omit $F_o - F_c$ map (contoured at 3.0 σ) shown as a green mesh superimposed on the stick model of the PAP molecule (carbon atoms in yellow, oxygen atoms in red and nitrogen atom in blue). **c**, Superimposed structures of the complexes with the side chains of residues 70, 85, 133 and 134 in stick models with the color scheme as PAP4 (salmon), NOS4 (yellow), ROTU4 (green) and GLAU4 (violet). The isoquinoline ring part of all four ligands (yellow stick) is shown as a space-filling model with isoquinoline shown in stick model (color scheme is identical to that in **b**). The side chain of F155 π–π stacking with the isoquinoline ring is shown in stick model and is colored gray. **d**, Superimposed BIA structures from all BIA-specific RamR variants shown relative to each other and the F155 residue of RamR (ligand color scheme is identical to that in **a**).

'hydrophobic cage' that traps the dimethoxybenzyl group of ROTU (Fig. 5c). The L66W and Y92W substitutions in GLAU4 create a large tryptophan sandwich motif that pins the hydrophobic GLAU fused rings (Fig. 5d). Similarly, the nitrogen atom of PAP is coordinated by the K63R substitution in PAP4 (strongly anchored by the adjacent A123D substitution (Fig. 5b)), while, in ROTU4, the K63Y and L156Y substitutions coordinate two ordered water molecules to interact with the nitrogen atom of ROTU (Fig. 5c), and the native D152 residue of GLAU4 interacts with GLAU's nitrogen atom (Fig. 5d). Indeed, although all alkaloids exhibit similar orientation toward the nitrogen atom (Supplementary Fig. 12), each RamR variant employs a unique adaptation to stabilize it (Fig. 5a–d). These structural data highlight the inherent plasticity of the RamR protein to rapidly evolve new ligand specificity and stabilize completely divergent modes of binding, suggesting that it may be a 'privileged template' for biosensor engineering.

Evolved biosensors accelerate plant enzyme engineering. Numerous BIAs have been recognized for their therapeutic value as modern pharmaceuticals. Specifically, THP, which is naturally produced in plants, is used as an immediate precursor to four Food and Drug Administration-approved drugs: atracurium, cisatracurium, mivacurium and PAP[27]. In plants, THP is thought to be produced from the common intermediate norcoclaurine via an oxidase and four separate *O*-methyltransferases (OMTs)[28]. However, attempts to ferment THP by coexpressing these four methyltransferases have proven unsuccessful. As some BIA methyltransferases have promiscuous activities[29], we reasoned that it might be possible to use our biosensor to evolve a methyltransferase that could uniquely methylate all four positions, thereby reducing a four-enzyme pathway into a single multifunctional enzyme.

To create an abbreviated THP-biosynthetic circuit suitable for screening and eventually directed evolution, we sought to identify an optimal 'one-step' enzyme by screening several promiscuous methyltransferases. To aid this effort, as well as subsequent evolution, we developed and tuned a THP reporter plasmid (pThpR) (Supplementary Fig. 13) that relies on the most sensitive THP biosensor (THP4.1), with an EC_{50} under 2 μM (Supplementary Fig. 9). Because THP4.1 was somewhat responsive to a semi-methylated substrate (norreticuline) but not to the unmethylated substrate, this circuit should report on general methylation activity, likely favoring more completely methylated derivatives (Supplementary Fig. 13). To screen methyltransferase variants, *E. coli* cells cotransformed with pThpR and a plasmid expressing the methyltransferase were cultured in the presence of the substrate norlaudanosoline (NOR) for 24 h, and fluorescence was subsequently measured. Upon screening, one methyltransferase, GfOMT1 from *Glaucium flavum*, produced the highest fluorescent signal and was chosen as the starting point for evolution (Supplementary Fig. 14).

Error-prone libraries of GfOMT1 with an average of two mutations per gene were cotransformed with pThpR, and highly fluorescent colonies were screened on solid medium supplemented with NOR (Fig. 6a). The resulting enzyme variants were subcloned and rephenotyped (Supplementary Fig. 15), and the best-performing variant was then used as the template for the next round of directed evolution. Over five rounds of screening, an OMT variant with seven substitutions (GEN5) produced a sixfold and 47-fold increase in fluorescent signal using pThpR when cultured with 100 μM or 10 μM NOR, respectively, when compared to wild-type GfOMT1 (Fig. 6b,c).

Throughout OMT evolution, later generations successively methylate more positions on NOR (Fig. 6d). Notably, the W22L substitution that occurred in GEN2 is thought to expand the substrate-binding pocket and enable production of the trimethylated intermediate, whereas the I258V substitution from GEN3 is adjacent to the expected *S*-adenosylmethionine-binding site and enables complete methylation to THP (Supplementary Fig. 16). The GEN3, GEN4 and GEN5 variants proved fully capable of

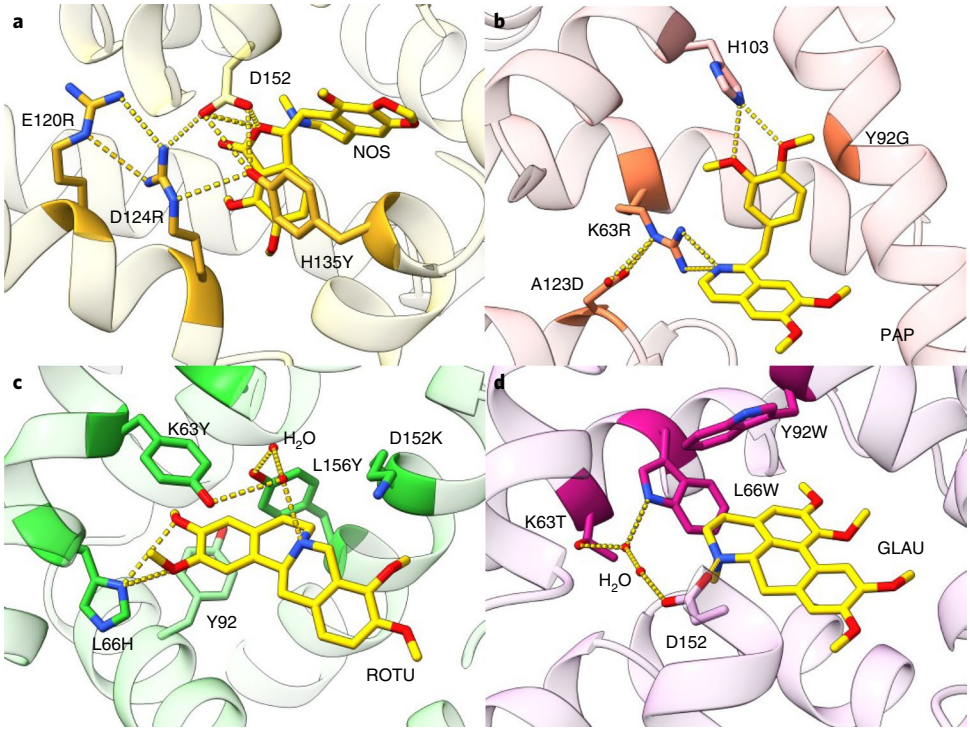


Fig. 5 | Unique molecular adaptations confer alkaloid specificity. a–d, Structures of NOS4 (**a**), PAP4 (**b**), ROTU4 (**c**) and GLAU4 (**d**) in complex with their cognate BIAs (shown in stick models with carbon atoms colored yellow, oxygen atoms in red and nitrogen in blue). Residues involved in specific interactions with the cognate ligand are displayed in stick models and labeled with carbon atoms in yellow, oxygen atoms in red and nitrogen atoms in blue.

making THP, with a product yield of 0.21 mg l^{-1}, 1.48 mg l^{-1} and 0.89 mg l^{-1}, respectively, from 2.88 mg l^{-1} of supplemented NOR (Supplementary Fig. 17). Importantly, the GEN4 mutant produced THP as its major product, and only trace amounts of the substrate NOR remained after fermentation (Fig. 6c). Interestingly, the GEN4 variant produced THP more selectively and efficiently than GEN5, despite both producing a similar fluorescent response with the pThpR reporter (Fig. 6c). This may be due to a higher accumulation of semi-methylated intermediates produced by the GEN5 enzyme.

Discussion

Using a new directed evolution circuit architecture, SELIS, we demonstrate that the TetR family member RamR can be readily adapted to specifically and sensitively recognize therapeutic alkaloids for which no other extant biosensors exist. Even though the wild-type sensor only poorly recognized a small subset of the target alkaloids, we were able to evolve biosensors for a wide variety of structures, most of which had EC_{50} values under 7 µM and >100-fold specificities for their cognate ligands, making them practical for screening even low-flux pathways during strain development.

While previous work has often attempted to identify biosensors that recognize ligands closely related to a target[10–13], we hoped to more broadly span chemical space by attempting to leverage an insight from natural selection: start with a generalist and evolve to a specialist[14]. This approach not only expands the chemical search space but also accelerates protein evolution by circumventing the exploration of generalist intermediates on the way to new specializations[30,31], an approach that is similar to the directed evolution of new specificities for highly promiscuous enzymes[32,33]. Pairing this generalist-first approach with our new selection-screening circuit architecture and focused mutagenesis libraries afforded massive improvements in biosensor performance with otherwise structurally diverse alkaloids (for example, the response to 100 µM NOS leaped from twofold to 509-fold after just three rounds of evolution (Extended Data Fig. 2)). Structural data for evolved RamR variants reveal that the RamR effector-binding site is extraordinarily malleable. Directed evolution of the generalist led to the same effector-binding site being able to accommodate both fixed interconnections between the rings (GLAU and ROTU) and flexible interconnections (NOS, PAP and THP). Common binding motifs for the isoquinoline ring abutted divergent benzyl-binding pockets, while still allowing for extremely fine discriminations (PAP and THP differ by only two hydrogen atoms).

One of the newly evolved biosensors (THP4.1) was introduced into a genetic circuit for the further evolution of a biosynthetic pathway for the therapeutic alkaloid THP, in a manner akin to prior biosensor-linked enzyme screens[11,16,34]. By identifying and broadening the substrate specificity of GfOMT1, we streamlined a more complex set of methylations for the production of THP to a single enzyme, with yield increasing from undetectable to 1.48 mg l^{-1}, providing an excellent starting point for additional scaling, especially as upstream biosynthetic machinery to enable de novo THP biosynthesis is already well established[35]. In particular, increasing the supply of the early intermediate L-DOPA and the cofactor *S*-adenosylmethionine should greatly improve yield[36,37].

These efforts should have immediate practical benefits, as THP is a precursor to several modern pharmaceuticals, including atracurium and cisatracurium, used as first-line drugs in the management of patients with coronavirus disease 2019, which have been at risk of experienced shortages due to increased demand[38]. Another drug synthesized using THP, PAP, is used for surgery on blood vessels. A shortage of this drug during 2016 in the US forced surgeons to trial other medications unapproved for these surgeries[39]. The combined sequence and structural data can also now inform intelligent library design for identifying additional biosensors for ligands bearing the isoquinoline moiety or even related groups, such as the quinoline and indole moieties abundant in natural and synthetic pharmaceuticals[40], potentially

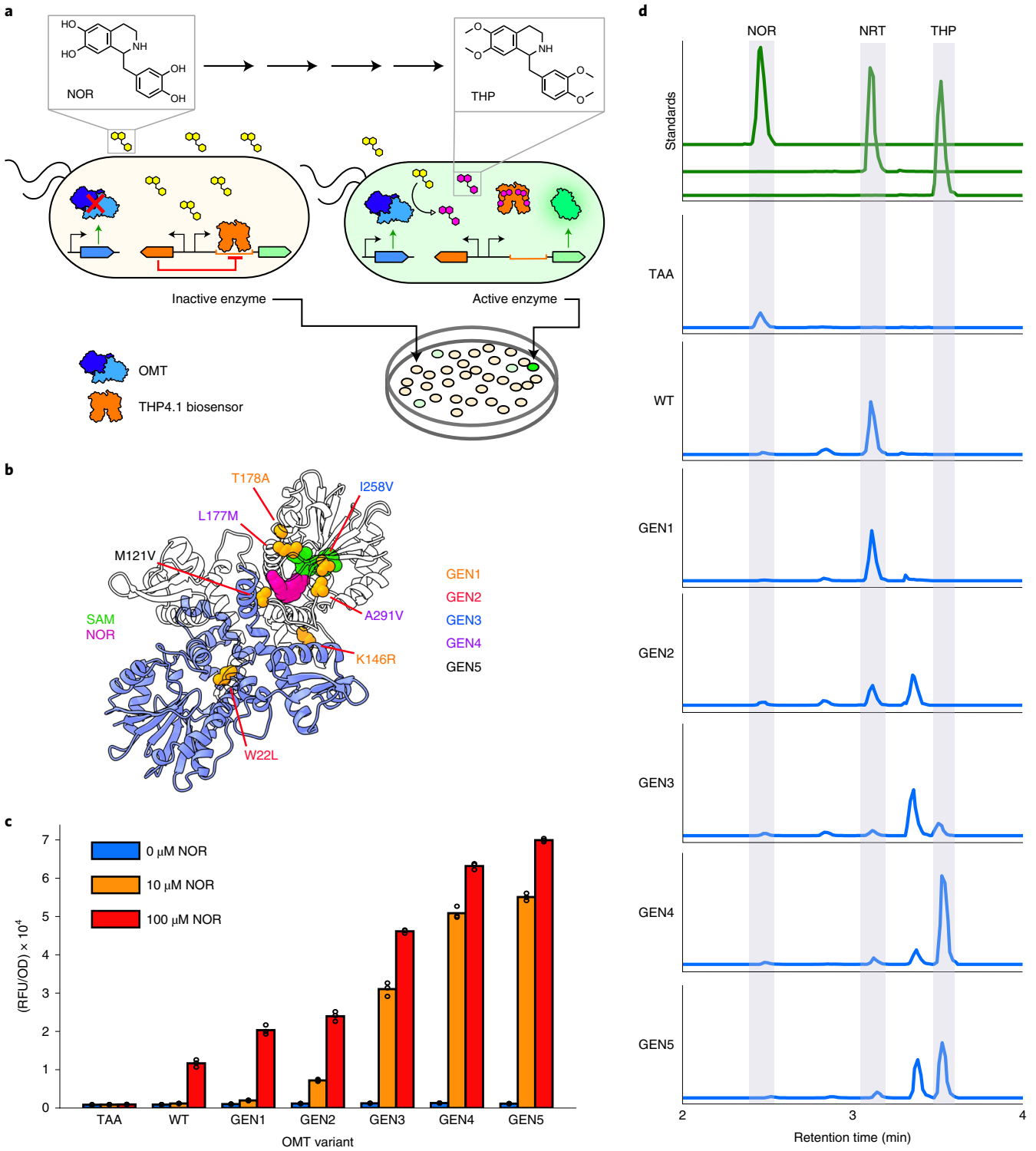


Fig. 6 | Evolved biosensor enables a new THP-biosynthetic pathway. a, Genetic circuit for OMT evolution. A plasmid expressing the OMT variant (blue) is cotransformed with a plasmid expressing GFP regulated by a THP-responsive biosensor (THP4.1 in orange). OMT libraries are plated with NOR, highly fluorescent colonies are picked and characterized, and the top variant is used for the following round of evolution. **b**, Homology structure of the template OMT (GfOMT1) with mutations in evolved variants shown in orange and labeled. The substrate NOR and the cofactor *S*-adenosylmethionine (SAM) are shown in pink and green, respectively. **c**, Fluorescence response from THP4.1 reporter plasmid-transformed cells cultured with either 0 μM, 10 μM or 100 μM NOR and an empty plasmid (TAA) control, GfOMT1 (wild type) or evolved OMT variants. Individual fluorescent measurement values for each condition performed in biological triplicate are displayed. Error bars represent mean ± s.e.m. **d**, Representative ion extracted chromatograms of strains expressing engineered OMT variants or controls, grown in the presence of 10 μM NOR. Norreticuline (NRT) has the identical exact mass of the expected double-methylated product nororientaline[29]. All LC–MS chromatograms were selected for the theoretical *m/z* values and retention times of the respective compounds of interest (Supplementary Table 3). Experimental chromatograms (TAA–GEN5) are scaled equally, while chromatograms for standards (NOR, NRT and THP) are not, for clarity. A small background peak for THP was subtracted; full chromatograms can be found in Supplementary Fig. 18.

providing engineered sensors that could accelerate ongoing efforts to improve morphinan-producing strains[3].

While previous efforts have largely focused on improving either sensitivity, specificity or background individually[10–13], we show that, by combining screening and selection, all of these properties can be improved concurrently, starting from well-chosen libraries that quickly converge on optimal outcomes (as evidenced by identifying multiple clones with identical protein sequences but different codons). Desired biosensor performance outcomes can be finely and rationally tuned during the implementation of SELIS in a manner that is more amenable to rapid prototyping and high-throughput identification than alternative, flow cytometry-based methods. For example, in one study[16], the bottom 80% of the sensor library population with the lowest fluorescence was sorted for the negative screen, while, in another[13], the bottom 5% of the population was sorted for the negative screen. Rather than relying on arbitrary gating stringencies, we leverage cellular growth to exponentially enrich living cells bearing highly active repressor variants. The fact that biosensors with intermediate ligand sensitivities and specificities can be manually chosen at intermediate SELIS rounds should yield analytical circuitry that can progressively keep pace with gradually improving product titers during strain development.

As biocatalyst-engineering projects become increasingly ambitious, via reconstituting long pathways in microbial hosts[41] or evolving enzyme cascades for pharmaceutical synthesis[42], there will likely be an increased reliance on high-throughput screening capabilities. We envisage that our approach to biosensor development will provide a chemical measurement platform integral to these future endeavors. Beyond their utility in high-throughput screening, biosensors have been used in dynamic regulatory schemes to improve production strain fitness and extend productivity lifetime[43,44]. Engineered sensors can also be paired with recently described genetic circuitry to reduce the limit of detection or improve the signal-to-noise ratio[45,46]. Furthermore, because TetR can be employed as a repressor in a wide range of medically and industrially relevant hosts, such as yeasts[47], plants[48] and mammalian cells[49], via a simple 'road-blocking' mechanism, the prospect now exists for evolving useful regulatory elements in bacteria and then directly adapting them to eukaryotic systems.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41589-022-01072-w.

Received: 8 November 2021; Accepted: 26 May 2022;
Published online: 7 July 2022

References

1. Ro, D.-K. et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006).
2. Luo, X. et al. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* **567**, 123–126 (2019).
3. Galanie, S., Thodey, K., Trenchard, I. J., Interrante, M. F. & Smolke, C. D. Complete biosynthesis of opioids in yeast. *Science* **349**, 1095–1100 (2015).
4. Nakagawa, A. et al. Total biosynthesis of opiates by stepwise fermentation using engineered *Escherichia coli*. *Nat. Commun.* **7**, 10390 (2016).
5. Srinivasan, P. & Smolke, C. D. Biosynthesis of medicinal tropane alkaloids in yeast. *Nature* **585**, 614–619 (2020).
6. Hodgman, C. E. & Jewett, M. C. Cell-free synthetic biology: thinking outside the cell. *Metab. Eng.* **14**, 261–269 (2012).
7. Rienzo, M. et al. High-throughput screening for high-efficiency small-molecule biosynthesis. *Metab. Eng.* **63**, 102–125 (2021).
8. Mitchler, M. M., Garcia, J. M., Montero, N. E. & Williams, G. J. Transcription factor-based biosensors: a molecular-guided approach for natural product engineering. *Curr. Opin. Biotechnol.* **69**, 172–181 (2021).
9. Hossain, G. S., Saini, M., Miyake, R., Ling, H. & Chang, M. W. Genetic biosensor design for natural product biosynthesis in microorganisms. *Trends Biotechnol.* **38**, 797–810 (2020).
10. Della Corte, D. et al. Engineering and application of a biosensor with focused ligand specificity. *Nat. Commun.* **11**, 4851 (2020).
11. Yao, J. et al. Developing a highly efficient hydroxytyrosol whole-cell catalyst by de-bottlenecking rate-limiting steps. *Nat. Commun.* **11**, 1515 (2020).
12. Snoek, T. et al. Evolution-guided engineering of small-molecule biosensors. *Nucleic Acids Res.* **48**, e3 (2020).
13. Machado, L., F. M., Currin, A. & Dixon, N. Directed evolution of the PcaV allosteric transcription factor to generate a biosensor for aromatic aldehydes. *J. Biol. Eng.* **13**, 91 (2019).
14. Trudeau, D. L. & Tawfik, D. S. Protein engineers turned evolutionists—the quest for the optimal starting point. *Curr. Opin. Biotechnol.* **60**, 46–52 (2019).
15. Wade, H. MD recognition by MDR gene regulators. *Curr. Opin. Struct. Biol.* **20**, 489–496 (2010).
16. Xiong, D. et al. Improving key enzyme activity in phenylpropanoid pathway with a designed biosensor. *Metab. Eng.* **40**, 115–123 (2017).
17. Li, Y. et al. Complete biosynthesis of noscapine and halogenated alkaloids in yeast. *Proc. Natl Acad. Sci. USA* **115**, E3922–E3931 (2018).
18. Valentic, T. R., Payne, J. T. & Smolke, C. D. Structure-guided engineering of a scoulerine 9-*O*-methyltransferase enables the biosynthesis of tetrahydropalmatrubine and tetrahydropalmatine in yeast. *ACS Catal.* **10**, 4497–4509 (2020).
19. Stanton, B. C. et al. Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat. Chem. Biol.* **10**, 99–105 (2014).
20. Morita, Y., Cao, L., Gould, V. C., Avison, M. B. & Poole, K. *nalD* encodes a second repressor of the *mexAB–oprM* multidrug efflux operon of *Pseudomonas aeruginosa*. *J. Bacteriol.* **188**, 8649–8654 (2006).
21. Sánchez, P., Alonso, A. & Martinez, J. L. Cloning and characterization of SmeT, a repressor of the *Stenotrophomonas maltophilia* multidrug efflux pump SmeDEF. *Antimicrob. Agents Chemother.* **46**, 3386–3393 (2002).
22. Yamasaki, S. et al. The crystal structure of multidrug-resistance regulator RamR with multiple drugs. *Nat. Commun.* **4**, 2078 (2013).
23. Meyer, A. J., Segall-Shapiro, T. H., Glassey, E., Zhang, J. & Voigt, C. A. *Escherichia coli* 'Marionette' strains with 12 highly optimized small-molecule sensors. *Nat. Chem. Biol.* **15**, 196–204 (2019).
24. Tominaga, M., Nozaki, K., Umeno, D., Ishii, J. & Kondo, A. Robust and flexible platform for directed evolution of yeast genetic switches. *Nat. Commun.* **12**, 1846 (2021).
25. Gatignol, A., Durand, H. & Tiraby, G. Bleomycin resistance conferred by a drug-binding protein. *FEBS Lett.* **230**, 171–175 (1988).
26. d'Oelsnitz, S., Nguyen, V., Alper, H. S. & Ellington, A. D. Evolving a generalist biosensor for bicyclic monoterpenes. *ACS Synth. Biol.* **11**, 265–272 (2022).
27. Ehrenworth, A. M. & Peralta-Yahya, P. Accelerating the semisynthesis of alkaloid-based drugs through metabolic engineering. *Nat. Chem. Biol.* **13**, 249–258 (2017).
28. Agarwal, P. et al. 3′*O*-methyltransferase, Ps3′OMT, from opium poppy: involvement in papaverine biosynthesis. *Plant Cell Rep.* **38**, 1235–1248 (2019).
29. Chang, L., Hagel, J. M. & Facchini, P. J. Isolation and characterization of *O*-methyltransferases involved in the biosynthesis of glaucine in *Glaucium flavum*. *Plant Physiol.* **169**, 1127–1140 (2015).
30. Collins, C. H., Arnold, F. H. & Leadbetter, J. R. Directed evolution of *Vibrio fischeri* LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones. *Mol. Microbiol.* **55**, 712–723 (2005).
31. Collins, C. H., Leadbetter, J. R. & Arnold, F. H. Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR. *Nat. Biotechnol.* **24**, 708–712 (2006).
32. Renata, H., Wang, Z. J. & Arnold, F. H. Expanding the enzyme universe: accessing non-natural reactions by mechanism-guided directed evolution. *Angew. Chem. Int. Ed. Engl.* **54**, 3351–3367 (2015).
33. Goldsmith, M. & Tawfik, D. S. Directed enzyme evolution: beyond the low-hanging fruit. *Curr. Opin. Struct. Biol.* **22**, 406–412 (2012).
34. Wu, J. et al. Design and application of a lactulose biosensor. *Sci. Rep.* **7**, 45994 (2017).
35. Nakagawa, A. et al. A bacterial platform for fermentative production of plant alkaloids. *Nat. Commun.* **2**, 326 (2011).
36. Wei, T., Cheng, B.-Y. & Liu, J.-Z. Genome engineering *Escherichia coli* for L-DOPA overproduction from glucose. *Sci. Rep.* **6**, 30080 (2016).
37. Chen, H., Wang, Z., Cai, H. & Zhou, C. Progress in the microbial production of *S*-adenosyl-L-methionine. *World J. Microbiol. Biotechnol.* **32**, 153 (2016).
38. Bohand, X., Jordan, D. & Dubois, F. Managing the risk of shortages and medication errors with curares during the COVID-19 pandemic: a hospital pharmacy experience. *Eur. J. Hosp. Pharm.* https://doi.org/10.1136/ejhpharm-2020-002605 (2021).
39. Ricci, J. A. et al. Comparing the outcomes of different agents to treat vasospasm at microsurgical anastomosis during the papaverine shortage. *Plast. Reconstr. Surg.* **138**, 401e–408e (2016).

40. Welsch, M. E., Snyder, S. A. & Stockwell, B. R. Privileged scaffolds for library design and drug discovery. *Curr. Opin. Chem. Biol.* **14**, 347–361 (2010).
41. Cravens, A., Payne, J. & Smolke, C. D. Synthetic biology strategies for microbial biosynthesis of plant natural products. *Nat. Commun.* **10**, 2142 (2019).
42. Huffman, M. A. et al. Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science* **366**, 1255–1259 (2019).
43. Hartline, C. J., Schmitz, A. C., Han, Y. & Zhang, F. Dynamic control in metabolic engineering: theories, tools, and applications. *Metab. Eng.* **63**, 126–140 (2021).
44. Rugbjerg, P., Sarup-Lytzen, K., Nagy, M. & Sommer, M. O. A. Synthetic addiction extends the productive life time of engineered *Escherichia coli* populations. *Proc. Natl Acad. Sci. USA* **115**, 2347–2352 (2018).
45. Wan, X. et al. Cascaded amplifying circuits enable ultrasensitive cellular sensors for toxic metals. *Nat. Chem. Biol.* **15**, 540–548 (2019).
46. Ho, J. M. L., Miller, C. A., Parks, S. E., Mattia, J. R. & Bennett, M. R. A suppressor tRNA-mediated feedforward loop eliminates leaky gene expression in bacteria. *Nucleic Acids Res.* **49**, e25 (2021).
47. Nagahashi, S. et al. Regulation by tetracycline of gene expression in *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* **255**, 372–375 (1997).
48. Gatz, C. & Quail, P. H. Tn10-encoded tet repressor can regulate an operator-containing plant promoter. *Proc. Natl Acad. Sci. USA* **85**, 1394–1397 (1988).
49. Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl Acad. Sci. USA* **89**, 5547–5551 (1992).

Methods

Strains, plasmids and media. *E. coli* DH10B (New England Biolabs) was used for all routine cloning and directed evolution. All biosensor systems were characterized in *E. coli* DH10B. *E. coli* BL21 DE3 (New England Biolabs) was used for protein expression. LB Miller (LB) medium (BD) was used for routine cloning, fluorescence assays, directed evolution and orthogonality assays unless specifically noted. Terrific broth (Thermo Fisher Scientific, 22711022) was used for protein purification. LB with 1.5% agar (BD) plates were used for routine cloning and directed evolution. The plasmids described in this work were constructed using Gibson assembly and standard molecular biology techniques. Synthetic genes, obtained as gBlocks, and primers were purchased from IDT. Relevant plasmid sequences are provided in Supplementary Table 4.

Benzylisoquinoline alkaloids. Cells were induced with the following chemicals: NOR (HDH Pharma, 29030), THP (Tokyo Chemical, N0918), PAP (MP Biomedicals, 190261), GLAU (Carbosynth, FG137572), ROTU (Alfa Aesar, J63328), NOS (Aldrich, 363960-5G), NRT (Selena Chem, CSC000735172).

Chemical transformation. For routine transformations, strains were made competent for chemical transformation. Five milliliters of an overnight culture of DH10B cells was subcultured into 500 ml LB medium and grown at 37 °C and 250 r.p.m. for 3 h. Cultures were centrifuged (3,500*g*, 4 °C, 10 min), and pellets were washed with 70 ml chemical competence buffer (10% glycerol, 100 mM $CaCl_2$) and centrifuged again (3,500*g*, 4 °C, 10 min). The resulting pellets were resuspended in 20 ml chemical competence buffer. After 30 min on ice, cells were divided into 250-µl aliquots and flash frozen in liquid nitrogen. Competent cells were stored at −80 °C until use.

Promoter design and biosensor response assay. Promoters for *ttgR* and *qacR* were derived from sequences in the literature[16,19]. For the *ramR* promoter, sequences for a region 60 bp upstream of the known operator sequence as well as the operator itself were extracted from the *S. typhimurium* genome (WP_000113609.1). NalD and SmeT are homologs of TtgR, therefore modifications from the *pttgr* promoter were made to match the sequence of the *nalD* operator[20] and the *smeT* operator[21]. For Pbm3r1, the known *bm3r1* operator[19] was placed immediately after the −10 region of a synthetic medium-strength promoter. All promoter sequences are listed in Supplementary Fig. 2. The pReg and pGFP equivalents for each regulator were cotransformed into DH10B cells, which were plated on LB agar plates with appropriate antibiotics. Three separate colonies were picked for each transformation and were grown overnight. The following day, 20 µl of each culture was then used to inoculate six separate wells in a 2-ml 96-deep-well plate (Corning, P-DW-20-C-S) sealed with an AeraSeal film (Excel Scientific) containing 900 µl LB medium, one for each test ligand and a solvent control. After 2 h of growth at 37 °C, cultures were induced with 100 µl LB medium containing either 10 µl DMSO or 100 µl LB medium containing one of the five target BIAs dissolved in 10 µl DMSO. Cultures were grown for an additional 4 h at 37 °C and 250 r.p.m. and subsequently centrifuged (3,500*g*, 4 °C, 10 min). Supernatant was removed, and cell pellets were resuspended in 1 ml PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4). One hundred microliters of the cell resuspension for each condition was transferred to a 96-well microtiter plate (Corning, 3904), from which the fluorescence (excitation, 485 nm; emission, 509 nm) and absorbance (600 nm) were measured using the Tecan Infinite M1000 plate reader.

RamR library design and construction. Five semi-rational libraries were designed, each targeting three inward-facing residues on one of five helices of the RamR ligand-binding pocket (Fig. 1d and Supplementary Fig. 4). Random mutagenesis was performed by amplifying the *ramR* gene using *Taq* polymerase in a buffer with biased nucleotide compositions. The average number of mutations introduced during random mutagenesis was calculated by sequencing 30 library members chosen at random. Libraries were generated using overlap PCR with redundant NNS codons using AccuPrime Pfx (Thermo Fisher, 12344024) and cloned into pReg. *E. coli* DH10B bearing pSelis was transformed with the resulting library. Transformation efficiency always exceeded 10^6 for each round of selection, indicating several fold coverage of the library. Transformed cells were grown in LB medium overnight at 37 °C with carbenicillin and chloramphenicol.

Directed evolution of RamR biosensors. Cell culture (20 µl) bearing the sensor library was seeded into 5 ml fresh LB containing appropriate antibiotics, 100 µg ml^{-1} zeocin (Thermo Fisher, R25001) and 100 µM nontarget BIAs (for rounds three and four) and grown at 37 °C for 7 h. Following incubation, 0.5 µl of culture was diluted into 1 ml LB medium, from which 100 µl was further diluted into 900 µl LB medium. Three hundred microliters of this mixture was then plated across three LB agar plates containing carbenicillin, chloramphenicol and the target BIA dissolved in DMSO. Plates were incubated overnight at 37 °C. The following day, the brightest colonies were picked and grown overnight in 1 ml LB medium containing appropriate antibiotics in a 96-deep-well plate sealed with an AeraSeal film at 37 °C. A glycerol stock of cells containing pSelis and pReg encoding the parental RamR variant was also inoculated into 5 ml LB for overnight growth.

The following day, 20 µl of each culture was used to inoculate two separate wells in a new 96-deep-well plate containing 900 µl LB medium. Additionally, eight separate wells containing 1 ml LB medium were inoculated with 20 µl of the overnight culture expressing the parental RamR variant. A typical arrangement would have 44 unique clones on the top half of the plate, duplicates of those clones on the bottom half of the plate and the right-most column occupied by cells harboring the parental *ramR* variant. After 2 h of growth at 37 °C, the top half of the 96-well plate was induced with 100 µl LB medium containing 10 µl DMSO, whereas the bottom half of the plate was induced with 100 µl LB medium containing the target BIA dissolved in 10 µl DMSO. The concentration of BIA used for induction is typically the same concentration used in the LB agar plate for screening during that particular round of evolution. Cultures were grown for an additional 4 h at 37 °C and 250 r.p.m. and subsequently centrifuged (3,500*g*, 4 °C, 10 min). Supernatant was removed, and cell pellets were resuspended in 1 ml PBS. One hundred microliters of the cell resuspension for each condition was transferred to a 96-well microtiter plate, from which the fluorescence (excitation, 485 nm; emission, 509 nm) and absorbance (600 nm) were measured using the Tecan Infinite M1000 plate reader. Clones with the highest signal-to-noise ratio (generally the top 5–10% of the screened clones) were then sequenced and subcloned into a fresh pReg vector.

For sensor variant validation, the subcloned pReg vectors expressing the sensor variants were transformed into DH10B cells expressing pGFP. These cultures were then assayed, as described in Promoter design and biosensor response assay, using eight different concentrations of the target BIA. Sensor variants that displayed a combination of low background, a reduced EC_{50} for the target BIA and a high signal-to-noise ratio were used as templates for the next round of evolution. A description of key parameters tuned throughout evolution and their impact on phenotype is provided in Supplementary Fig. 19.

Dose–response measurements. Glycerol stocks (20% glycerol) of strains containing the plasmids of interest were inoculated into 1 ml LB medium and grown overnight at 37 °C. Twenty microliters of overnight culture was seeded into 900 µl LB medium containing ampicillin and chloramphenicol in a 2-ml 96-deep-well plate sealed with an AeraSeal film. Following growth at 37 °C and 250 r.p.m. for 2 h, cultures were induced with 100 µl of an LB medium solution containing appropriate antibiotics and the inducer molecule dissolved in 10 µl DMSO. Cultures were grown for an additional 4 h at 37 °C and 250 r.p.m. and subsequently centrifuged (3,500*g*, 4 °C, 10 min). Supernatant was removed, and cell pellets were resuspended in 1 ml PBS. The cell resuspension (100 µl) for each condition was transferred to a 96-well microtiter plate, from which the fluorescence (excitation, 485 nm; emission, 509 nm) and absorbance (600 nm) were measured using the Tecan Infinite M1000 plate reader. Population distributions of cells induced with each cognate effector were homogeneous (Supplementary Fig. 20), and BIA induction did not appear to affect apparent cell viability (Supplementary Fig. 21).

Orthogonality assays. For each evolutionary lineage (for example, wild type, THP1, THP2, THP3, THP4) all regulators were expressed from the pReg plasmid using the same promoter, which is P114-RBS(*riboJ*), P114-RBS(*riboJ*), P103-RBS(*elvJ*), P114-RBS(*riboJ*) and P103-RBS(*riboJ*) for the GLAU, NOS, PAP, ROTU and THP lineages, respectively. These plasmids were cotransformed with a plasmid encoding pGFP, and, the following day, three individual colonies were picked into LB and grown overnight. Fluorescence assays were performed as described in Dose–response measurements above, but either 100 µM of each BIA in 1% DMSO or 1% DMSO itself was used for induction. For post-screening sensor variant comparison (Supplementary Figs. 5–9e,f), only one biological replicate was taken. The fluorescence of cells containing both the pReg plasmid expressing the sensor variant (see Supplementary Table 1 for the promoters used for each variant) and the pSELIS plasmid was assayed with 100 µM of each BIA in 1% DMSO.

Protein purification. Coding sequences for RamR variants were cloned into an ampicillin-resistance pUC plasmid with a T7 RNA polymerase promoter driving expression of the gene of interest with an encoded N-terminal His6-3C tag. Plasmids were transformed into electrocompetent BL21 DE3 cells, and single transformants were grown to saturation in LB supplemented with 1,000 µg ml^{-1} carbenicillin. Cultures were diluted 1/250 in terrific broth supplemented with antibiotics in baffled flasks and incubated at 37 °C with agitation (250 r.p.m.) until they reached midlog phase. Protein expression was induced by adding IPTG to achieve a final concentration of 0.5 mM. For PAP4 only, PAP was also added during IPTG induction to reach a final concentration of 100 µM. Cells were cultured for 18 h at 18 °C. Cells were collected by centrifugation at 8,000*g* for 10 min, and the cell pellets were resuspended in 25 ml wash buffer (50 mM K_2HPO_4, 300 mM NaCl and 10% glycerol at pH 8.0) with a protease inhibitor cocktail (cOmplete, Mini, EDTA-free, Roche) and lysozyme (0.5 mg ml^{-1}). Cells were incubated for 20 min at 4 °C with gentle agitation and lysed by sonication (Model 500, Fisher Scientific). The lysate was repeatedly clarified by centrifugation (35,000*g* for 30 min), and protein was recovered by immobilized metal ion affinity chromatography using Ni–NTA resin and gravity flow columns. The eluate was concentrated and dialyzed, with 3C protease added to the dialysis cassette, into the appropriate buffer

followed by purification to apparent homogeneity by size-exclusion fast protein liquid chromatography. All RamR variants were dialyzed into 20 mM Tris (pH 8.0), 200 mM NaCl and 3 mM DTT.

X-ray crystallography. To form co-crystals of RamR variants in complex with individual ligands, 1 mM substrate was added to 10 mg ml^{-1} purified protein and incubated overnight at 4 °C except for the PAP4 protein, which already formed a complex with PAP during the protein expression step. Rod-shaped co-crystals were grown using the sitting-drop vapor-diffusion method at room temperature for PAP4, ROTU, GLAU4 and NOS4 in conditions with 0.1 M MES (pH 6.0–7.5), 14–23% PEG 3350, 0.2 M ammonium sulfate and 0.1 M sodium chloride. Individual crystals were flash frozen directly in liquid nitrogen after a brief incubation with a reservoir solution supplemented with 25% (vol/vol) glycerol. X-ray diffraction data were collected at the BL 5.0.1 beamline in ALS (Berkeley, CA). X-ray diffraction was processed to resolutions of 1.6 Å, 1.6 Å, 2.0 Å and 2.2 Å for PAP4 with PAP, ROTU4 with ROTU, GLAU4 with GLAU and NOS4 with NOS using HKL2000. In the Phenix software, phases were obtained by molecular replacement using a previously solved RamR wild-type structure as the initial search model (PDB 3VVX)[50]. Notably, all data collected for ROTU4 ($n > 8$) exhibited the crystal defect of twinning, creating issues for identifying a structural solution. Using the package Xtriagle from the Phenix suite, we identified a twin operator of (-h, k, -k-l) that was successfully used for subsequent refinement. The molecular replacement solutions for each structure were iteratively built using Coot with the Phenix refine package[51]. The refinement of ROTU4 was finalized using REFMAC. The quality of the final refined structures was evaluated by MolProbity[52]. The final statistics for data collection and structure determination are shown in Supplementary Table 2.

Biosensor-linked *O*-methyltransferase activity assay. All OMTs were expressed with the P114-RBS(*riboJ*) promoter–RBS on the pReg plasmid backbone (no regulator present). Cells were cotransformed with both the OMT plasmid and the THP reporter plasmid and plated on an LB agar plate containing appropriate antibiotics. Three individual colonies from each transformation were picked into LB and grown overnight. Resulting cultures were diluted 50-fold into 1 ml LB medium containing NOR and 1 mg ml^{-1} ascorbic acid in a 96-deep-well plate and were grown at 30 °C for 18 h. Subsequently, the fluorescence of cultures was measured in the same manner as previously described in Dose–response measurement above.

***O*-methyltransferase evolution.** The OMT-coding region of the GfOMT1 expression vector was targeted for random mutagenesis, introducing an average of two mutations relative to the template. Libraries were assembled using Gibson assembly and were subsequently transformed into chemically competent cells containing the THP reporter plasmid. In parallel, the template OMT plasmid was transformed into the same cells for downstream comparative analysis. Cultures were then plated on an LB agar plate containing NOR. Highly fluorescent colonies were picked and grown overnight with appropriate antibiotics. The following day, cultures were grown with NOR and 1 mg ml^{-1} ascorbic acid at 30 °C for 18 h. Fluorescence of the resulting cultures was measured as described above, and variants that were significantly more fluorescent than the template OMT control were sequenced. Unique top-performing OMT variants were then subcloned into the pReg vector and transformed into cells bearing the THP reporter plasmid. Three individual colonies from each transformation were subcultured into fresh LB medium, and the fluorescence of each culture in the presence of NOR (refer to Supplementary Fig. 15) was subsequently measured as described above. The top-performing OMT variant was then used as the template for the following round of evolution.

Homology model. The homology model of the GEN5 GfOMT1 variant was constructed using SWISS-MODEL (https://swissmodel.expasy.org/). PDB 5ICE was used as a template for structure generation.

Liquid chromatography–mass spectrometry. Cells containing the plasmid expressing each OMT variant with the P114-RBS(*RiboJ*) promoter were transformed and plated onto an LB agar plate containing appropriate antibiotics. The following day, three colonies from each plate were cultured overnight in LB and subsequently diluted 50-fold into 1 ml LB containing 10 µM NOR and 1 mg ml^{-1} ascorbic acid. These cultures were grown for 18 h at 30 °C and centrifuged at 16,000*g* for 1 min, and the resulting supernatant was filtered using a 0.2-µm filter. The samples were then diluted 1:100-fold into filter-sterilized water before LC–MS analysis. Samples were analyzed using an Agilent 6546 Q-TOF LC–MS with a dual Agilent Jet Stream electrospray ionization source in positive mode. Chromatographic separations were obtained under gradient conditions by injecting 5 µl onto an Agilent RRHD Eclipse Plus C18 column (50 × 2.1 mm, 1.8-µm particle size) with an Agilent ZORBAX Eclipse Plus C18 narrow-bore guard column (12.5 × 2.1 mm, 5-µm particle size) on an Agilent 1260 Infinity II liquid chromatography system. The mobile phase consisted of eluent A (water with 0.1% formic acid) and eluent B (acetonitrile). The gradient was as follows: 5% B to 95% B from 0 to 6 min (0.3 ml min^{-1}), hold at 95% B from 6 to 7 min (0.3 ml min^{-1}), 95% B to 5% B from 7 to 7.1 min (0.3 ml min^{-1}) and hold at 5% B from 7.1 to 9 min (0.45 ml min^{-1}). The sample tray and column compartment were set to 7 °C and 30 °C, respectively. The fragmentor was set to 180 V. Q-TOF data were processed using the Agilent MassHunter Qualitative Analysis software.

To create the chromatograms shown in Supplementary Fig. 18, signal count from the EIC within a window ±0.05 min relative to the retention time for the corresponding alkaloid (Supplementary Table 3) was extracted for each scan (m/z ratios of the following: 288.1230, 302.1387, 316.1543, 330.1700 and 344.1856) and added together. Because a substantial background peak was detected around the expected retention time for THP, the signal count in this window of time was removed for all samples that had a count at or below this background level. The resulting chromatograms were then used to create Fig. 5d.

Statistical analysis and reproducibility. All data in the text are displayed as mean ± s.e.m. unless specifically indicated. Bar graphs, fluorescence and growth curves, dose–response functions and orthogonality matrices were all plotted in Python 3.6.9 using Matplotlib and seaborn. Dose–response curves and EC_{50} values were estimated by fitting to the Hill equation $y = d + (a - d)x^b(c^b + x^b)^{-1}$ (where y = output signal, b = Hill coefficient, x = ligand concentration, d = background signal, a = maximum signal and c = EC_{50}), with the scipy.optimize.curve_fit library in Python.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The coordinates for the complex structures have been deposited in the PDB: RamR in complex with berberine, PDB 3VW2; PAP4 in complex with PAP, PDB 7N53; ROTU4 in complex with ROTU, PDB 7N4W; NOS4 in complex with NOS, PDB 7N4Z; GLAU4 in complex with GLAU, PDB 7N54. Protein sequence information was retrieved from the NCBI database: RamR, 3VVX_A; TtgR, WP_010952495.1; QacR, WP_001807342.1; SmeT, WP_014648459.1; NalD, WP_003092152.1; Bm3R1, WP_013083972.1; GfOMT1, AKO60152.1. Plasmid sequences relevant to this study (Supplementary Notes 1 and 2) were deposited in Addgene. Source data are provided with this paper.

Code availability

Code used to generate bar plots, dose–response functions and orthogonality matrices presented in this text is accessible at https://github.com/simonsnitz/plotting.

References

50. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).
51. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Struct. Biol.* **60**, 2126–2132 (2004).
52. Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).

Acknowledgements

Funding from DARPA Soils (HR00111920019 to A.D.E.), Welch (F-1654 to A.D.E.) and AFSOR (FA9550-14-1-0089 to H.S.A. and A.D.E.) is acknowledged. This work is partially supported by grants from the National Institutes of Health (R01GM104896 and R01GM125882 to Y.Z.). We thank K.J. Blake in the Chemistry Department at the University of Texas at Austin for performing LC–MS analysis and S. Kar for his thoughtful advice on selection circuit design.

Author contributions

S.d.O. designed the experiments and performed biosensor evolution and characterization. S.d.O. and R.T. performed protein purification. Enzyme evolution was carried out by S.d.O. and K.J., and X-ray crystallography was conducted by W.K. and N.T.B. The manuscript was written by S.d.O. with support from A.D.E., R.T., Y.Z. and H.S.A. S.d.O., A.D.E. and H.S.A. supervised all aspects of the study.

Competing interests

S.d.O., K.J., R.T. and A.D.E. have filed two patent applications on materials described in this text. R.T. and A.D.E have equity in GRO Biosciences, a company developing protein therapeutics. The other authors declare no conflict of interest.

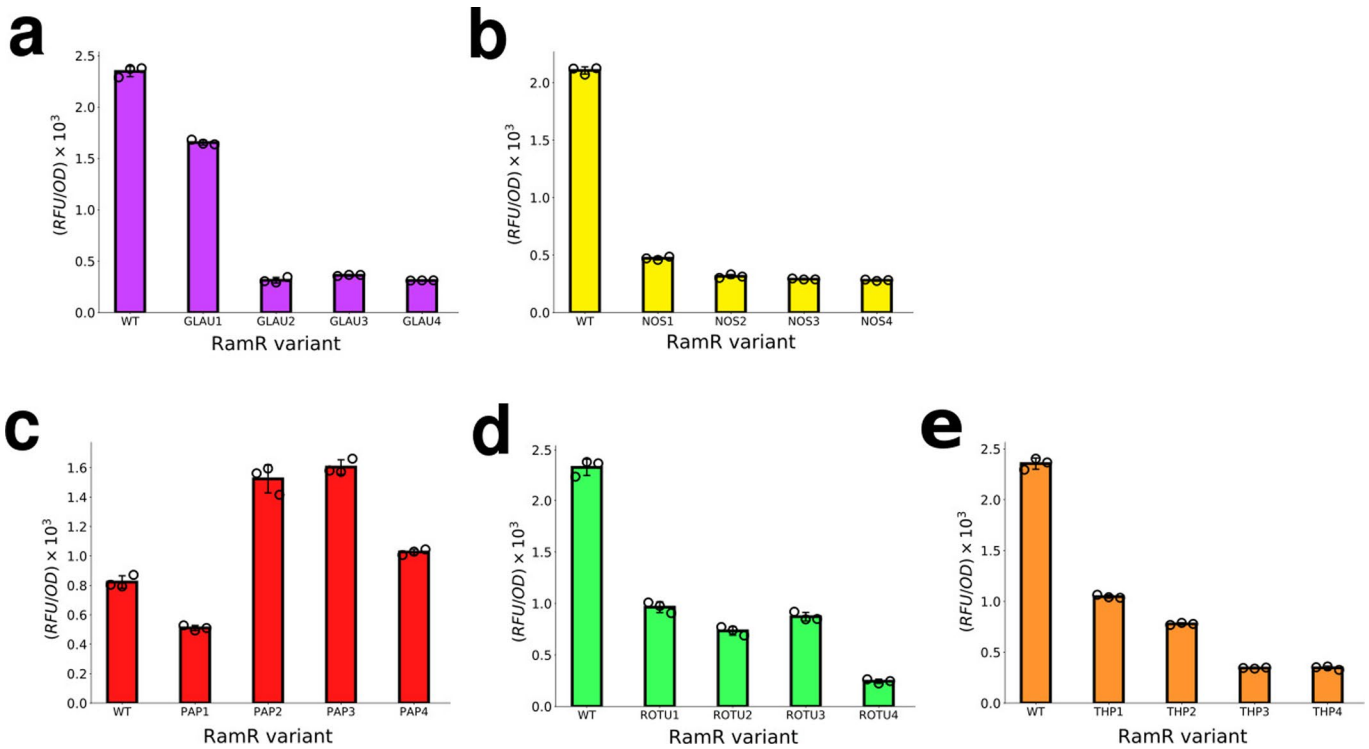
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41589-022-01072-w.

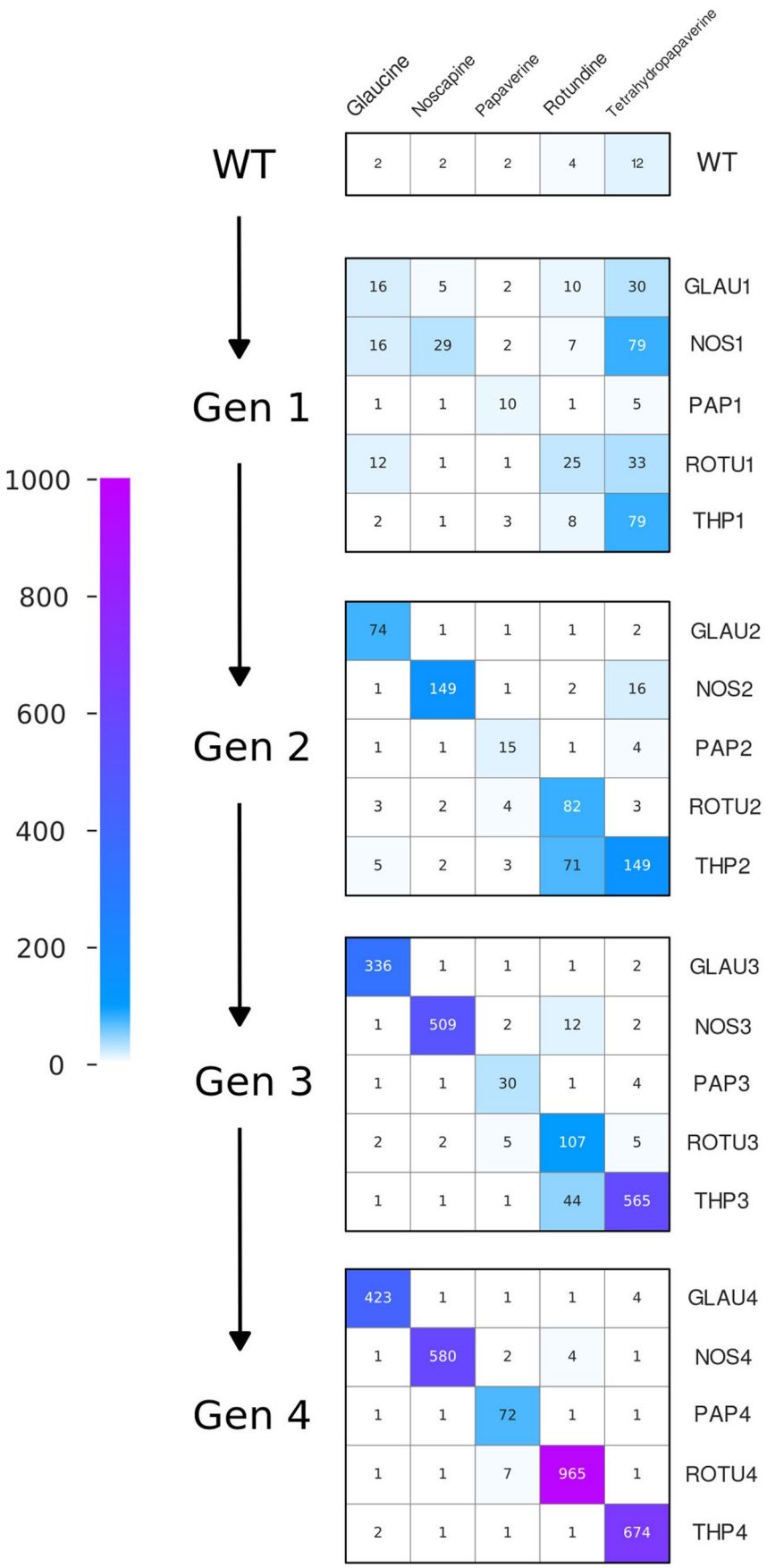
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41589-022-01072-w.

Correspondence and requests for materials should be addressed to Simon d'Oelsnitz or Andrew D. Ellington.

Peer review information *Nature Chemical Biology* thanks Michael Jensen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



Extended Data Fig. 1 | Background fluorescence measurements for all RamR generations. The same promoter was used to express variants from each evolutionary trajectory (*see* methods). Individual fluorescent measurements for each condition performed in biological triplicate are shown alongside error bars that represent the SE ± the mean.



Extended Data Fig. 2 | Cross reactivity of all evolved sensors. Fold-response is shown for all BIAs for the native RamR protein, the first, second, third, and fourth generations from top to bottom, respectively. 100 μM of the indicated BIA was applied in all conditions. Measurements for each condition represent the average of three biological replicates.

nature portfolio

Corresponding author(s): Simon d'Oelsnitz and Andy Ellington

Last updated by author(s): Apr 16, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Fluorescence and absorbance data was collected using the iControl 3.4.2.0 software associated with the Tecan Infinite M1000 instrument. X-ray diffraction data were collected at BL 5.0.1 beamline in ALS (Berkeley, CA). Mass spectrometry data was collected using commercial code associated with the Agilent 6546 Q-TOF LC/MS instrument.
Data analysis	All data in the manuscript are displayed as mean ± s.e.m. unless specifically indicated. Bar graphs, fluorescence/growth curves, dose response functions, and orthogonality matrices were all plotted in Python 3.6.9 using Matplotlib v3.5.1 and Seaborn v0.11.2 . Dose response curves and EC50 values were estimated by fitting to a hill equation with the optimize.curve_fit function of the SciPy v1.8.0 library in Python.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The coordinates for the complex structures have been deposited in the Protein Data Bank. Relevant structure descriptions and associated PDB codes are listed below.
RamR in complex with berberine. PDB: 3VW2

PAP4 in complex with papaverine. PDB: 7N53
ROTU4 in complex with rotundine. PDB: 7N4W
NOS4 in complex with noscapine. PDB: 7N4Z
GLAU4 in complex with glaucine. PDB: 7N54

Protein sequence information was retrieved from the NCBI database. Accession codes of proteins relevant to this study are listed below.
RamR: 3VVX_A
TtgR: WP_010952495.1
QacR: WP_001807342.1
SmeT: WP_014648459.1
NalD: WP_003092152.1
Bm3R1: WP_013083972.1
GfOMT1: AKO60152.1

Plasmids relevant to this study (see Supplementary Note 1 and 2) were deposited in Addgene.
Raw data is provided as source data files.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was performed. A sample size of three was chosen for all biological assays, unless otherwise specified, since this is standard in the field.
Data exclusions	No data was excluded.
Replication	Experiments were replicated once, and all attempts at replication were successful.
Randomization	Samples were allocated into experimental groups via genotype. Measurements were taken from random replicates of identical genotypes.
Blinding	The investigators were not blinded. Blinding during data collection was not required since the conditions were well controlled. Blinding during data analysis was not possible since the differences between samples under different conditions were visually apparent. Blinding was also not necessary since the results are quantitative and didn't require subjective interpretation. Blinding is not commonly used in the field.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging