VEGAS as a Platform for Facile Directed Evolution in Mammalian Cells

Justin G. English,[1,5,*] Reid H.J. Olsen,[1] Katherine Lansu,[1] Michael Patel,[2] Karoline White,[3] Adam S. Cockrell,[4] Darshan Singh,[1] Ryan T. Strachan,[1] Daniel Wacker,[1] and Bryan L. Roth[1,*]
[1]Department of Pharmacology, University of North Carolina, Chapel Hill, NC 27514, USA
[2]Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC 27514, USA
[3]Department of Biology, University of North Carolina, Chapel Hill, NC 27514, USA
[4]Department of Epidemiology, University of North Carolina, Chapel Hill, NC 27514, USA
[5]Lead Contact
*Correspondence: jgenglis@email.unc.edu (J.G.E.), bryan_roth@med.unc.edu (B.L.R.)
https://doi.org/10.1016/j.cell.2019.05.051

SUMMARY

Directed evolution, artificial selection toward designed objectives, is routinely used to develop new molecular tools and therapeutics. Successful directed molecular evolution campaigns repeatedly test diverse sequences with a designed selective pressure. Unicellular organisms and their viral pathogens are exceptional for this purpose and have been used for decades. However, many desirable targets of directed evolution perform poorly or unnaturally in unicellular backgrounds. Here, we present a system for facile directed evolution in mammalian cells. Using the RNA alphavirus Sindbis as a vector for heredity and diversity, we achieved 24-h selection cycles surpassing 10^{-3} mutations per base. Selection is achieved through genetically actuated sequences internal to the host cell, thus the system's name: viral evolution of genetically actuating sequences, or "VEGAS." Using VEGAS, we evolve transcription factors, GPCRs, and allosteric nanobodies toward functional signaling endpoints each in less than 1 weeks' time.

INTRODUCTION

Spontaneous genetic mutations diversify traits among a population of organisms while selective pressure culls diverse populations. This enrichment of ultimately advantageous traits is a process known as evolution by means of natural selection (Darwin and Bynum, 2009; Wallace, 1855, 1871). Humans can accelerate the development of organisms with desirable traits by guiding evolution through artificial selection. This technique can be traced back to the earliest agricultural crops (Diamond, 2002; Wright et al., 2005) and domesticated animals (Driscoll et al., 2009). As a biomedical laboratory tool, artificial selection has been instrumental in understanding myriad processes ranging from the cell cycle (Hartwell et al., 1970) to bacterial antibiotic resistance (Albert et al., 2005; Baym et al., 2016; Toprak et al., 2011). Artificial selection of targeted DNA sequences, rather than whole organisms, is called directed evolution (Arnold, 1998; Chen and Arnold, 1993). Directed evolution has been used to create novel binding proteins (Hanes and Plückthun, 1997; Xu et al., 2002), enzymes (Chen and Arnold, 1993; Kuchner and Arnold, 1997), chemogenetic tools (Armbruster et al., 2007), and fluorescent reporters (Campbell et al., 2002; Crameri et al., 1996) with broad scientific and industrial utility.

Directed evolution approaches typically use peptide display or microorganisms to screen large-scale DNA libraries that encode mutant proteins. "Hits" from these systems are isolated, mutagenized, and rescreened in an interrupted or iterative fashion. Iterative systems minimize evolution cycle time and omit user-biased "winner" selection by combining mutagenesis, selection, and heredity in parallel. Iterative systems have been improved using uninterrupted facile (McMahon et al., 2018) and continuous methods (Badran and Liu, 2015; Carlson et al., 2014; Esvelt et al., 2011). Although both methods have produced excellent results, these systems have been developed outside the context of the mammalian cell signaling environment. Consequently, incompatibility of function when transferring evolved products from unicellular to mammalian systems frequently occurs (see Armbruster et al., 2007 for examples), wherein additional rounds of selection and focused mutagenesis must be performed. Additionally, the currently available directed evolution systems omit classes of proteins that are usually incompatible with non-mammalian host systems—including G-protein coupled receptors (GPCRs).

GPCRs comprise one of the largest protein families in the human genome with >900 unique protein coding genes (Fredriksson et al., 2003; Wacker et al., 2017a). GPCRs represent the largest class of druggable targets and are known to regulate most biological processes (Hauser et al., 2017). Despite their importance, GPCRs are largely omitted from directed evolution studies due to their functional incompatibilities with non-mammalian systems (although, see Armbruster et al., 2007; Sarkar et al., 2008; Schütz et al., 2016). GPCRs are seven-transmembrane receptors that transduce extracellular signals into biological responses via heterotrimeric G proteins and β-arrestins (Gilman, 1987; Pierce et al., 2002). GPCR signal transduction is accomplished via a network of interacting molecular switches (Wacker et al., 2017a), yielding an isomerizing landscape of conformations that evoke unique cellular signaling cascades

(De Lean et al., 1980; Kobilka and Deupi, 2007; Onaran and Costa, 2009). GPCR-directed pharmaceuticals stabilize subsets of this signaling landscape leading to stabilization of active (e.g., agonism) or inactive (Neubig et al., 2003; Wacker et al., 2017a) states. A directed evolution system capable of targeting these states, and the signaling pathways downstream of such targets, could provide key insights necessary to advance cell signaling biology and drug development.

Here, we present a system for the viral evolution of genetically actuating sequences, which we dub "VEGAS," using a facile directed evolution platform in mammalian cells. Using the RNA alphavirus Sindbis for parallel mutagenesis, selection, and heredity, we demonstrate the robust, directed, and functional evolution of both GPCRs and allosteric GPCR intrabodies in mammalian cell culture in less than 1 week.

RESULTS

Sindbis Virus for Directed Evolution in Mammalian Cell Culture

Mammalian cell-based directed evolution has had many successes (Armbruster et al., 2007; Berman et al., 2018; Buchholz et al., 1998; Chan et al., 2017; Hess et al., 2016; Maheshri et al., 2006), albeit using time consuming, costly, and specialized screening platforms. The use of engineered viruses has advanced the field (Berman et al., 2018), serving as vectors for library storage, delivery, and heredity; the use of viruses, however, has been limited to conventional iterative systems involving panning, "winner-picking," and *ex vivo* mutagenesis. Such iterative directed evolution approaches sabotage the powerful evolutionary principles at play in competitive genetic populations (Huston, 1979). Here, we aimed to develop a mammalian directed evolution system where viral mutagenesis, selection, and heredity could operate simultaneously.

We required a mutagenic virus that could replicate freely at titers sufficient for constant reinfection of naive cells in culture, and for this, we focused on obligate RNA viruses, the most mutagenic viral class (Drake and Holland, 1999). Due to concerns related to laboratory safety and utility only a handful of RNA viruses are feasible for routine use. Of those available, we focused our efforts on the Alphavirus Sindbis, from the *Togaviridae* family (Strauss et al., 1984; Xiong et al., 1989). Sindbis virus is a single-stranded RNA virus encoding an RNA-dependent RNA replicase targeted to the viral genome by *cis*-acting, conserved 5-3′ sequences (Frolov et al., 2001). These sequences are required to initiate replication and RNA templates, even those from related viral families, cannot be replicated by the Sindbis virus replicase, resulting in high selectivity between the replicase and the Sindbis virus genome (Frolov et al., 2001), which functions simultaneously as a replication template and coding strand for viral protein translation. Sindbis virus has been engineered as a transgene delivery vector (Agapov et al., 1998; Schlesinger, 1993; Strauss and Strauss, 1994; Xiong et al., 1989); here, we further engineered Sindbis virus to control the packaging process using mammalian expression vectors.

We first determined that Sindbis virus can be continuously packaged in mammalian cell culture using an expression plasmid encoding the Sindbis virus structural genome (Figure 1A, also refer to Data S1 [VEGAS Supplemental Resource] and Data S2 [VEGAS Extended Protocol] for additional details). Transgenic Sindbis virus plasmid harboring green fluorescent protein (pTSin-EGFP) (STAR Methods) was packaged and titered at 5.45×10^{11} genomes/mL (Figure 1B) as determined by qRT-PCR targeting the Sindbis virus packaging signal sequence (STAR Methods). This initial titer was applied to 1×10^7 cells transfected with pCMV-SSG (Sindbis structural genome) (STAR Methods) at an MOI (multiplicity of infection) of 1. Harvesting and subsequent analysis of the culture media from these cells revealed high viral titer production, with 6.64×10^8 genomes/mL produced after 4 h and 5.57×10^{10} genomes/mL produced after 24 h (Figure 1B). The 24 h sample from round 1 was transferred to naive cells transfected with or without pCMV-SSG at an MOI = 1. After 24 h, pCMV-SSG transfected cells produced 6.37×10^8 genomes/mL while untransfected control cells produced 2.6×10^5 genomes/mL (Figure 1B). Fluorescent imaging of the infected culture over time confirmed passage of the EGFP transgene (Figure S1A). Transgene expression is rapid, with EGFP detectable in as few as 4 h post infection. These experiments demonstrate that Sindbis virus can be used for sustained transgene packaging in mammalian cell culture using a plasmid-borne structural genome.

RNA viruses, such as Sindbis, are highly mutagenic, with no known proof-reading capability. Approximations of RNA virus mutation rates range from 10^{-5} to 10^{-3} mutations per base replicated (Drake et al., 1998; Morley and Turner, 2017; Sanjuán et al., 2010; Schnell et al., 1996; Strauss and Strauss, 1994). As no prior study quantified the genetic stability of a non-essential transgenic gene during Sindbis virus replication, we next determined the mutation frequency of our directed evolution system. We initiated packaging of pTSin-EGFP in pCMV-SSG transfected cells and collected supernatant after 3, 6, 12, 24, and 36 h. The EGFP transgenic segment, as well as the vector template and initially packaged RNA, was amplified and sequenced using an Illumina NextSeq500. The sequences were aligned (STAR Methods) and quantified for positional sequence integrity of EGFP (Figures S1B and S1C). A significant ($p < 0.0001$) (STAR Methods) time-dependent increase in average mutation frequency was observed when comparing the 0HR (initial RNA) sample versus the 6, 12, 24, and 36 h samples (Figure 1C). The number of observed insertions and deletions also increased with time (Figure S1D). Nucleotide substitution rates were relatively even, with the exception of a modest ($p < 0.05$) C > G preference in samples 6–36 h (Figure 1D), and events occured proportionally across read lengths (Figures S1E–S1G). Linear regression analysis of mutation frequency versus time (Figure 1E) yielded an estimate of $1.0 \times 10^{-4} \pm 3.7 \times 10^{-5}$ mutations base^{-1}/h—~1 mutation per 1,000 bases replicated or >10^9 total mutations per day at the observed viral production rates. This high mutation rate, coupled to accumulating insertions and deletions infrequently accessible to rational design platforms, makes Sindbis virus an ideal candidate for developing a mammalian-directed evolution platform.

Directed Evolution of Transcription Factors with Sindbis Virus

To yield a robust directed evolution platform that leverages the replicative and mutagenic potential of Sindbis virus, artificial

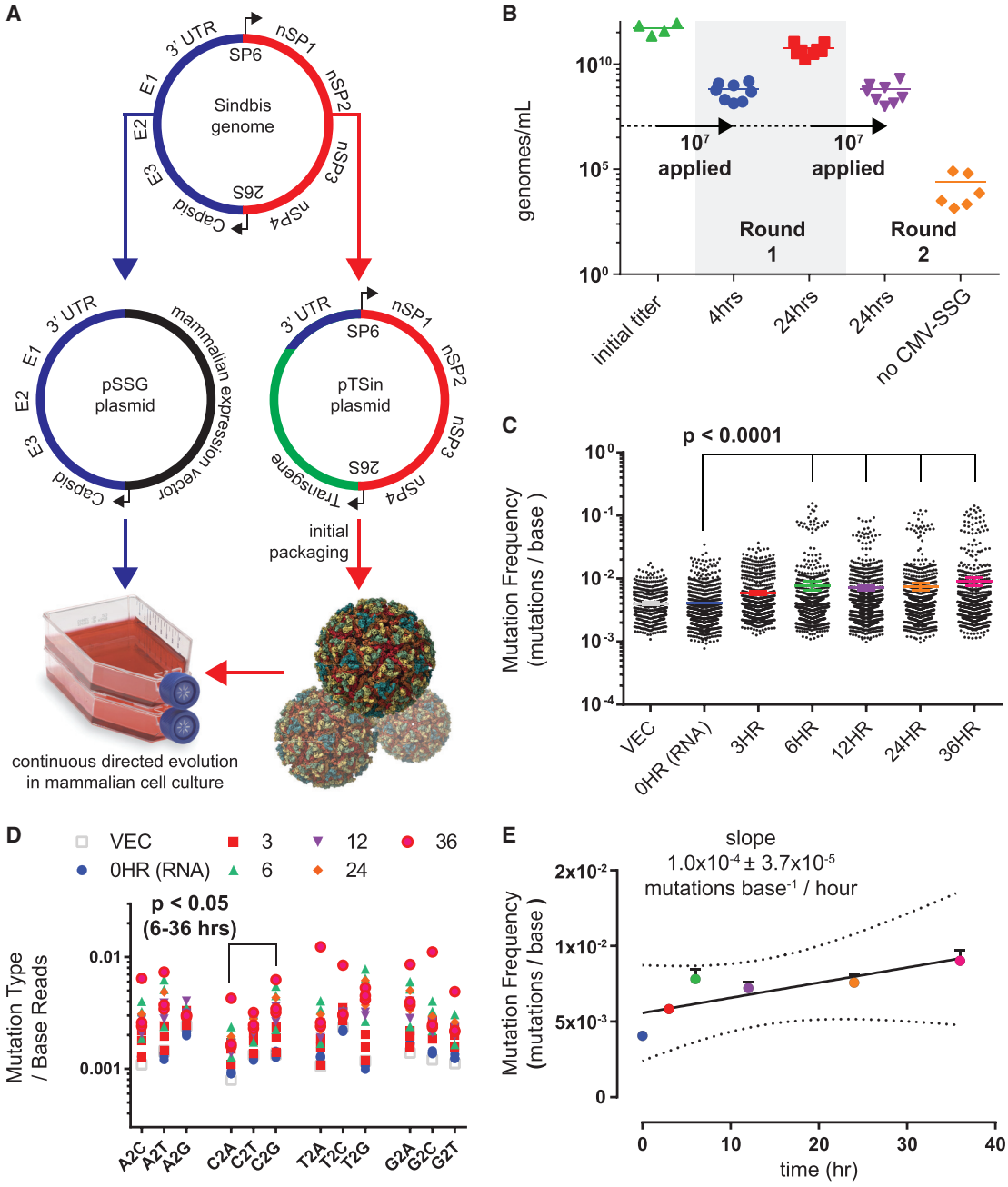


Figure 1. Sindbis Virus for Facile Directed Evolution in Mammalian Cell Culture

Development of Sindbis virus for facile, mutagenic viral propagation in mammalian cell culture.

(A) Design of plasmids used for facile directed evolution with Sindbis virus. Artificial Sindbis genome; Girdwood, MF459683.1. pSSG plasmid; capsid, E3, E2, E1, and 3′ UTR moved to a mammalian expression vector. pTSin plasmid; the structural genome elements of the artificial Sindbis genome replaced by any transgene sequence (pTSin). Propagation and selection can then be performed in mammalian cell culture using pTSin packaged virus applied to cells transfected with pSSG.

(B) qRT-PCR quantification of Sindbis virus production from cell culture. Data are represented as mean of individual biological replicates, N > 3.

(C) Mutations observed from Illumina paired-end sequencing of Sindbis packaged EGFP transgene over time. Mutation frequency is plotted as mutations observed per read at each nucleotide position across the transgene. Data are plotted for each individual replicate (N = 3; 24HR and VECTOR, N = 2) around mean ± 95% confidence interval.

(D) Base changes observed from sequencing of Sindbis packaged EGFP transgene over time. A, adenine, T, thymine, G, guanine, C, cytosine. Statistical comparison tested within base groups between each time point.

(E) Calculation of Sindbis mutation rate from sequencing of Sindbis packaged EGFP transgene over time. Data are represented as mean ± SEM and as linear regression, dotted line highlights the 99% confidence interval band.

See also Figure S1 and Data S1 and S2.

selective pressure must be applied. Each Sindbis viral particle requires 240 copies of each of the structural proteins E1, E2, and capsid to form a functional viral particle that can mature and propagate (Tang et al., 2011), and without this envelope, the virus is unable to mature and propagate. By engineering restrictions on structural genome transcription, we developed a system to apply selective pressure on transgenic Sindbis virus. As proof of concept for this method, we placed the Sindbis virus structural genome under control of the tetracycline operator sequence (pTETO$_7$-SSG) (Das et al., 2004; Gossen et al., 1995; Orth et al., 2000; STAR Methods) and packaged transgenic Sindbis virus with tetracycline transactivator (pTSin-tTA) (Gossen and Bujard, 1992; STAR Methods). We infected cells ± TETO$_7$-SSG with viral pTSin-tTA or pTSin-EGFP and then treated cells with either the tTA inhibitor doxycycline (DOX, 1 μM) or vehicle at the time of infection. Virus was packaged at 3.53 × 10^{10} genomes/mL in the vehicle + TETO$_7$-SSG cell line, while <10^6 genomes/mL were detected for all other conditions (Figure S2A).

Using the TS-tTA system, we sought to benchmark the capabilities of VEGAS by evolving tTA to be DOX-insensitive. To accomplish this, we packaged TS-tTA virus under non-selective conditions (R0) and exposed it to constant rounds of selection using increasing concentrations of DOX (Figure 2A). Seven selection rounds, encompassing just 7 days of evolution, produced a large number of full-length tTA sequences (see Table S1). By round 6, a consensus sequence dominated the observed coding sequence pool that was carried through to round 7. This consensus sequence, dubbed "R7," was completely resistant to DOX (Figure 2B). R7 possessed twenty-two coding mutations spanning all functional domains of the protein (Figure 2C). We had predicted that mutations directly involved in ligand interaction (Figures 2C and 2D; Kisker et al., 1995; Orth et al., 1999a, 1999b) would be enriched in the final consensus. To our surprise, none of these residues were mutated in R7. Instead, mutations accumulated primarily adjacent to key interacting residues for each functional domain (Figures 2C and S2B), many of which have been previously identified to reduce the effect of DOX on TETR-TETO interaction (Berens et al., 1992; Hecht et al., 1993; Müller et al., 1995; Orth et al., 1998; Scholz et al., 2004; Schubert et al., 2001; Smith and Bertrand, 1988; Urlinger et al., 2000; Wissmann et al., 1991; see Figure 2C and Table S2 for details). In addition, a cluster of negatively charged residues comprising helices 8 and 9 residing over the ligand binding pocket spanning Q149-H179 (Figure 2D) were converted to primarily positively charged residues (Figure 2C; Table S2). Among varying bacterial species, the net charge, but not specific residues, of this loop is conserved, and this conserved charge landscape has been proposed to attract the tetracycline-Mg^{2+} inducer to the ligand binding pocket (Orth et al., 1998). The mutations observed in R7 increase the net charge of this loop by +3.19, concentrated near the ligand entry tunnel. This gain in local charge presumably repels the positively charged DOX-Mg^{2+}.

Interestingly, in addition to augmenting the peptide sequence through directed evolution, our analysis of the nucleotide sequences from each round revealed codon usage optimizations as well (Figures S2C and S2D). Non-synonymous mutations acquired through tTA evolution converted rarely used codon sequences for BHK21, derived from *Mesocricetus auratus*, to the more frequently used GAC (D, +13%), GAG (E, +29%), AAG (K, +43%), TTG (L, 30%), and CAG (Q, +19%).

Augmenting TETR ligand sensitivity has been attempted previously using mammalian-directed evolution (Das et al., 2004), wherein 2 mutations were identified in 114 days. Our evolution of tTA generated an order of magnitude more functional mutations in 7 days thereby illustrating how our Sindbis virus system can be used for successful directed evolution of a transcription factor in mammalian cell culture. Key to the evolutionary component of this method is the actuation of a genetically encoded circuit to unlock expression of the Sindbis structural proteins, capsid, E1, and E2. Consequently, we have named this Sindbis virus system viral evolution of genetically actuating sequences, or "VEGAS."

VEGAS for the Evolution of GPCRs

With VEGAS in hand to perform directed evolution in mammalian cells, we focused our efforts on GPCRs, a superfamily of transmembrane receptors with substantial pharmacological and physiological importance (Hauser et al., 2017; Wacker et al., 2017a). Critical to the GPCR field is the mapping of interacting residues associated with the transition from an inactive to active receptor. Mapping these motifs can provide anchor points for homology modeling, evolutionary sequence analysis, and ligand design (Fan et al., 2009; Michino et al., 2015; Roth et al., 2017). Even among the best studied receptors, using extensive mutation campaigns and high resolution crystal structures of inactive and active receptor conformations, the field has struggled to consistently identify key residues involved in state transition (Dror et al., 2011; Huang et al., 2015a; Latorraca et al., 2017). Class A GPCRs possess conserved binding pockets and trigger motif residues involved in the inactive to active state transition. However, many class A GPCRs lack conservation within these motifs, a disproportionate number of which are classified as understudied or orphan receptors (Figure S3A). Here, we used VEGAS to identify previously unknown constitutively active mutations for the understudied receptor MRGPRX2; our approach demonstrates how VEGAS can illuminate the complex conformational changes involved in GPCR activation even in the absence of structural information.

MRGPRX2 is a primate-exclusive GPCR recently identified as an atypical opioid-recognition receptor (Lansu et al., 2017). MRGPRX2 possesses limited homology to other opiate receptors (see Table S3) and minimal conservation of classic interacting residues. The curious composition of MRGPRX2 hampers *de novo* prediction of functional motifs. We therefore used VEGAS to develop constitutively active mutations (CAMs) of MRGPRX2. As VEGAS requires an activity-coupled transcriptional response to gate selection, we screened MRGPRX2 activated by (+)-morphine, across a panel of transcription factor reporters driving luciferase (luc2P, FLuc) expression (Figure S3B). The serum response element (SRE) and serum response factor (SRF) minimal promoters gave 5-fold responses 4 h post-ligand addition and persisted for 24 h. We chose SRE for its lower total basal signal in unstimulated conditions and replaced luc2P with the Sindbis virus structural

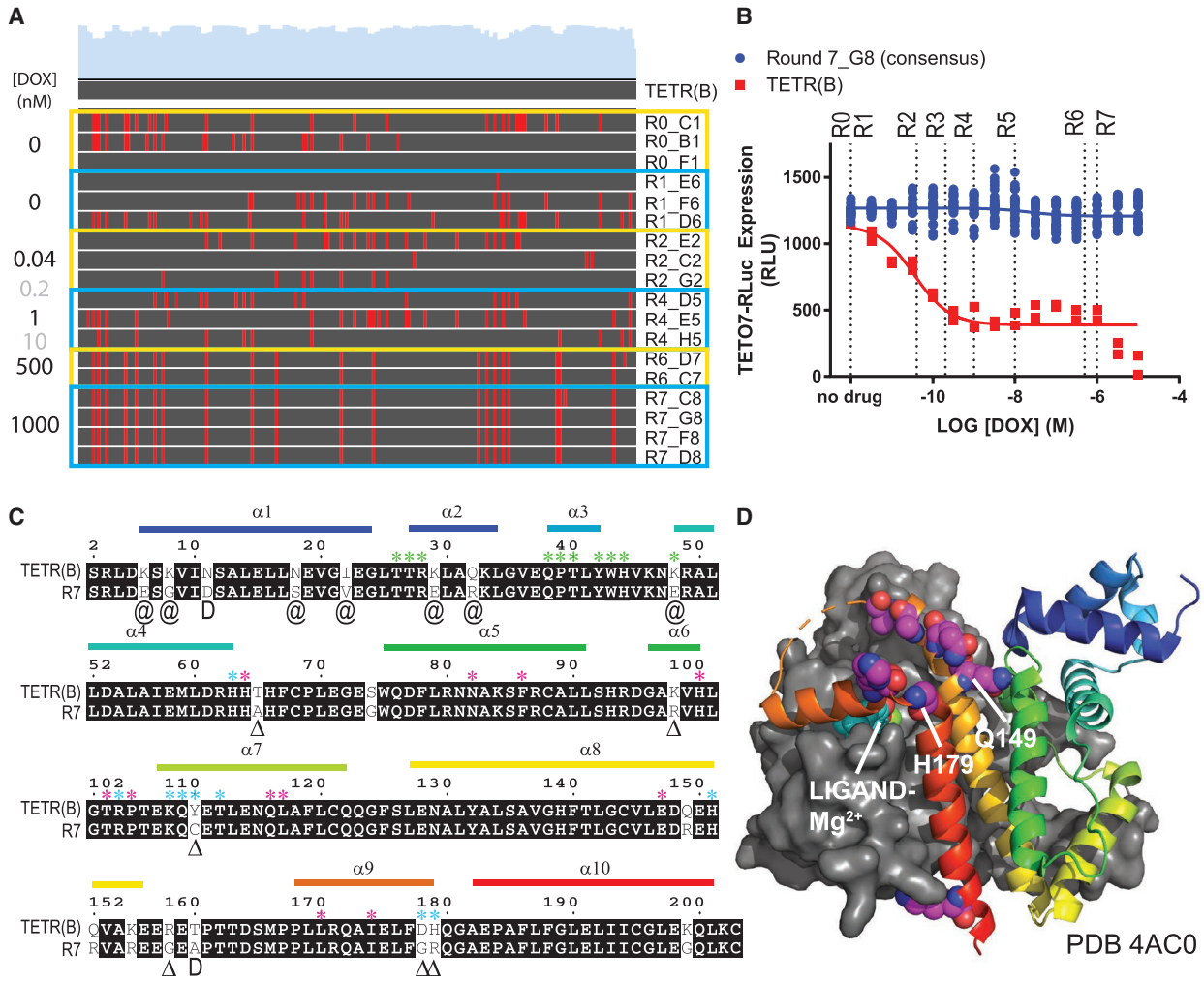


Figure 2. Directed Evolution of Transcription Factors with Sindbis

Sindbis was used as a directed evolution platform to generate a doxycycline (DOX)-resistant variant of the transcription factor tTA.

(A) Nucleotide sequence alignment of TETR clones isolated from each round of selection to the wild type TETR(B) sequence. Each round is outlined in yellow or blue, applied concentration of DOX to the left, name of individual clones to the right. Gray DOX values indicate no clones were isolated from the round. Red lines in the alignment denote a sequence mismatch from wild type.

(B) TETO7-Rluc reporter assay with increasing concentrations of DOX. Dotted lines are selection round DOX concentrations, for reference. Data are represented as mean ± SEM of individual biological replicates.

(C) Peptide sequence alignment of TETR(B) and the R7 consensus. Matching residues are shaded, mutations are unshaded. Alpha helices (α) are labeled and color coded to match with palettes in (D) and S2B. Exact residue (Δ), position (@), or subtype (D) substitutions previously published to enhance tTA activity in the presence of DOX as per Table S2. Residues (*) with direct involvement in DNA binding (green), ligand binding (magenta), and ligand entry (cyan) as per Orth et al. (2000) and Schubert et al. (2004).

(D) Crystal structure PDB: 4AC0 of TETR(B) in complex with minocycline-$Mg2^{+}$. Helix 8-9 ligand enclosure spanning Q149-H179 is displayed with spheres highlighting the residues for mutations Q149R, Q152R, K155R, R158G, T160A (no density), D178G, and H197R observed in R7.

See also Figure S2, Tables S1 and S2, and Data S1 and S2.

genome (SRE-SSG). Cells transfected with SRE-SSG were infected with transgenic Sindbis virus harboring MRGPRX2 (pTSin-MRGPRX2) and selected with diminishing amounts of (+)-morphine over 3 days (Figure 3). Resultant viral genomes were isolated, and their MRGPRX2 transgenes were tested in subsequent assays.

We presumed these clones would be CAMs, however, none of the isolated mutants mapped to classic sites of constitutive activity modulation identified in other GPCRs (Figure S3C).

We screened each mutant for activity in SRE-luc2P (Figure 3A), β-arrestin recruitment (Figure 3B), and phosphoinositide (PI) hydrolysis (Figure 3C) functional assays. We also quantified surface receptor expression via ELISA to ensure proper trafficking and expression (Figure S3D). For SRE and TANGO assays, basal activity across the variants increased at each evolutionary generation. TANGO basal activity reached 100% of wild type (+)-morphine stimulation for three independent mutants: L210P, Y67H+L210P+V265A, and L42P+D306G. PI hydrolysis,

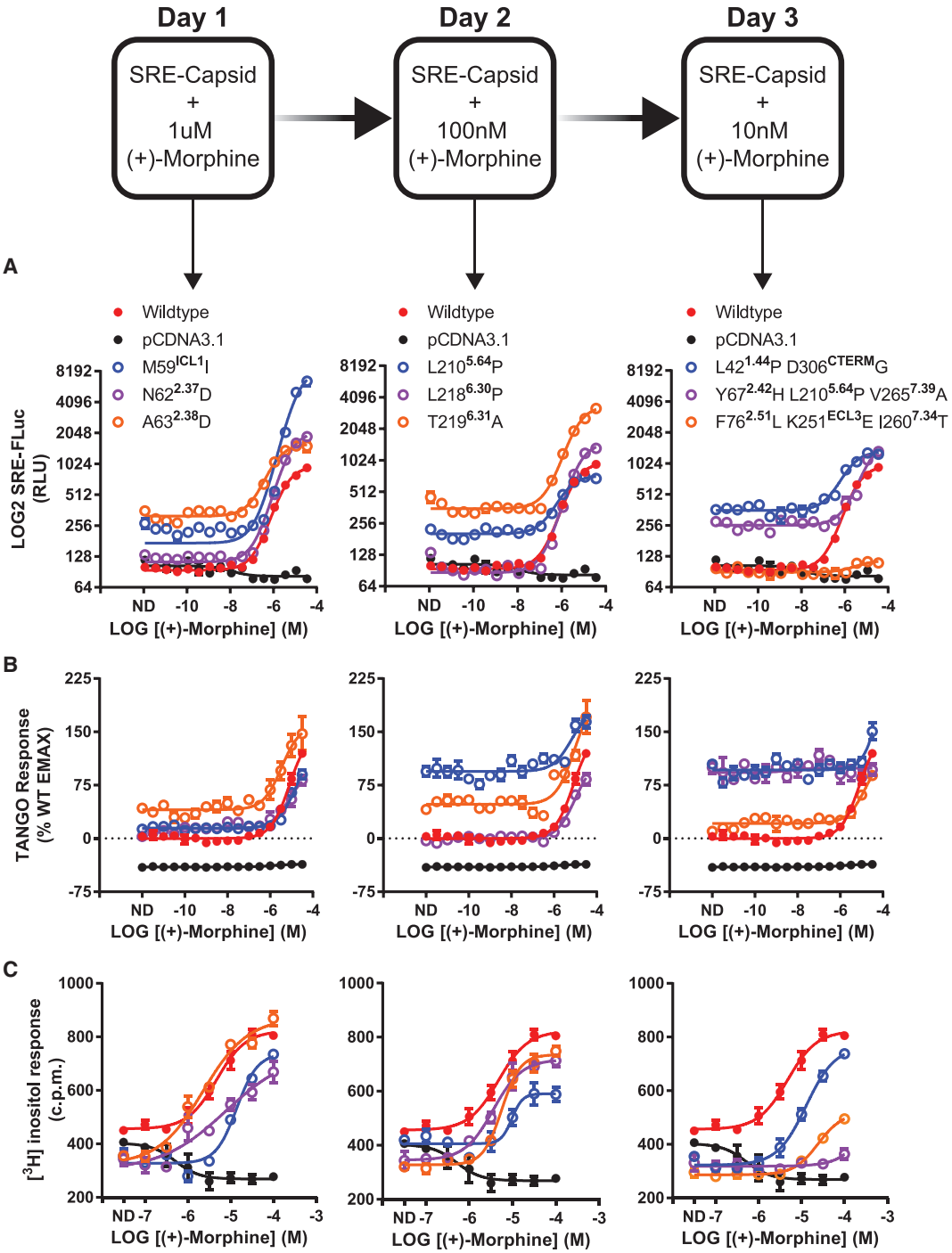


Figure 3. VEGAS for the Evolution of GPCRs

Using VEGAS multiple constitutively active mutants of the GPCR MRGPRX2 were produced in 3 days through application of decreasing concentrations of the MRGPRX2 agonist (+)-morphine. Mutations acquired in each round were tested functionally. Mutations are listed with their receptor residue position and Ballesteros-Weinstein annotation.

(A) Serum response element (SRE) reporter assay. FLuc production equates to relative receptor activation. ND, no drug. Data are represented as mean ± SEM, N = 3.

(B) TANGO reporter assay. RLuc production equates to receptor-mediated β-arrestin2 activation. ND, no drug. Data are represented as mean ± SEM, N = 3.

(C) Phosphoinositide hydrolysis assay. Accumulation of [^{3}H] inositol equates to receptor-mediated $G\alpha_q$ activation. ND, no drug. Data are represented as mean ± SEM, N = 3.

See also Figure S3, Table S3, and Data S1 and S2.

a proxy for $G\alpha_q$ activity, detected decreased basal activity, ligand potency, and efficacy for all mutants. Decreases in maximum agonist-induced $G\alpha_q$ activity correlated with increases in constitutive TANGO and SRE activity.

VEGAS for the Evolution of Active-State Nanobodies

GPCR ligands stabilize signal-state-specific receptor conformations (Kobilka and Deupi, 2007; Onaran and Costa, 2009; Strachan et al., 2014), and the development of novel ligands is enhanced by signal-state-specific GPCR crystal structures (Che et al., 2018; Manglik et al., 2016; Wang et al., 2017). Nanobodies, genetically encodable antigen recognition domains from dromedaries (Muyldermans et al., 2001), can be used to obtain these stabilized active state structures (Che et al., 2018; Manglik et al., 2016; Rasmussen et al., 2007; Staus et al., 2016). The nanobodies developed in these studies mimic the $G\alpha$ protein, displacing it. More desirable would be a nanobody that stabilizes the complex between GPCR and its transducer $G\alpha$ protein. These nanobodies would be allosteric modulators capable of enhancing GPCR-$G\alpha$ coupling. Using VEGAS, we can create allosteric nanobodies for multiple GPCR-$G\alpha$ pairings in less than a week.

To create GPCR nanobodies using VEGAS, we first generated a GPCR-targeted nanobody library by immunizing a llama against the serotonin 2A ($5\text{-}HT_{2A}$) GPCR bound to the high-affinity agonist lysergic acid diethylamide (LSD). We isolated single-domain antibodies from the peripheral blood mononuclear cells of the immunized llama through amplification of the variable region "V_HH" of IgG (Pardon et al., 2014; STAR Methods). The V_HH amplicon was used to generate a cDNA library of $\sim 1 \times 10^7$ colonies, which was subsequently packaged in Sindbis virus. This library was then used to evolve intracellular targeting nanobodies against $5\text{-}HT_{2A}$, as well as the dopamine-D2 (DRD2) and pH-sensing GPR68 (Huang et al., 2015b) receptors. Each of these receptors couples canonically to a different $G\alpha$ protein (2A, $G\alpha_q$; D2, $G\alpha_i$, 68, $G\alpha_s$). Developing nanobodies toward each receptor serves to demonstrate the broad applicability of the VEGAS system.

To evolve active state-stabilizing nanobodies for each GPCR, we first screened each receptor for transcription factor coupling (Figures S4A–S4F). All three receptors were determined to activate SRE with varying efficacy and we therefore chose to develop nanobodies that engage the SRE-signaling state for each receptor. Cells transfected with SRE-SSG and a GPCR were infected with an MOI = 1 of the viral nanobody library. To select for SRE activating nanobodies, $5\text{-}HT_{2A}$ and DRD2 cultures were incubated in the absence of ligand, while GPR68 was incubated at its inactive pH 8 (Huang et al., 2015b). Day 1 viral particles were harvested, the selection was repeated, and individual nanobody clones were isolated from the day 1 and 2 titers and sequenced. Clones with $N \geq 2$ identity in the subcloned population were selected and screened for GPCR-dependent SRE activation (Figure 4B). Each evolution series produced nanobodies capable of SRE activation only in the presence of the intended GPCR target, with the exception of VGS-Nb1 that appears to constitutively activate SRE even under GPCR-free conditions. To determine if the nanobodies obtained using VEGAS came directly from the originating library, or were evolved, we deep-sequenced the clonal library using a NextSeq500 (STAR Methods). Stringent end-to-end alignment of the entire sequence pool (total reads equaled 20 million) was assessed over 3 independent score cut-offs. Reads aligning to the VEGAS-derived nanobodies are displayed in Figure 4A. Each VEGAS-derived nanobody possessed sequences that were not detected within or outside the complementarity-determining regions (CDRs). In addition, we compared reference nanobodies (REF_Nbs) cloned from the parent library and VEGAS isolates to the amino acid frequency distribution of 1,346 deposited *Llama glama* VhH sequences from >50 animals (Table S4). Five positions with >99% sequence conservation across populations were conserved in the REF_Nb sequences, but were mutated in the VEGAS evolved sequences (Figure 4C). Both sequence analysis methods demonstrate that the VEGAS-derived nanobodies were not original to the library but evolved from that initial pool of nanobody cDNAs.

Positive Allosteric Modulation of GPCRs by VEGAS-Evolved Nanobodies

Using VEGAS, we produced 8 nanobodies targeted against 3 independent GPCRs in less than 1 week. Here, we interrogate their physical and molecular interactions with each target and provide a detailed characterization of the mechanism of VGS-Nb2, a positive allosteric modulator of the $5\text{-}HT_{2A}$ serotonin receptor.

First, we established whether the VEGAS-evolved nanobodies directly associated with their intended GPCR targets via bioluminescence resonance energy transfer (BRET). For BRET, GPCR-RLuc fusions and increasing concentrations of mVenus nanobody (mVenus-Nb) fusion proteins were co-transfected in to HEK293T cells. We observed a strong association between mVenus-VGS-Nb2 and $5\text{-}HT_{2A}$-RLuc, but no association to the closely related serotonin 2B ($5\text{-}HT_{2B}$) receptor (Figure 5A). Addition of the agonist serotonin (5-HT) at 1 μM or above (Figures 5A and S5A) had no effect on VGS-Nb2 association to either $5\text{-}HT_{2A}$ or $5\text{-}HT_{2B}$. We also observed association between mVenus-VGS-Nb6 and GPR68-RLuc at pH 8, but no association of the nanobody to $5\text{-}HT_{2A}$, the protein used to develop the initial nanobody library for directed evolution (Figure 5B). GPR68 activity increases with increasing pH (Huang et al., 2015b), we therefore stimulated our BRET assay with a pH 6 buffer and observed an increased association between mVenus-VGS-Nb6 and GPR68-RLuc (Figure 5B). Low, non-specific interaction of the DRD2-targeted nanobodies VGS-Nb7 and VGS-Nb8 was also observed (Figures S5C and S5D). However, VGS-Nb7 and VGS-Nb8 both increase SRE activity in the presence of DRD2 (Figure S5J) through an unknown mechanism.

The serotonin 2A (*HTR2A*, $5\text{-}HT_{2A}$) receptor is a GPCR of significant importance to mental health, disease, pharmacology, and homeostatic biology (McCorvy and Roth, 2015). Structures of $5\text{-}HT_{2A}$ and closely related 5-HT2-family receptors $5\text{-}HT_{2B}$ (Wacker et al., 2013) and $5\text{-}HT_{2C}$ (Peng et al., 2018) have yet to be obtained for their active states. Using VEGAS we have identified a nanobody that binds active $5\text{-}HT_{2A}$, but not $5\text{-}HT_{2B}$. We therefore further characterized the $5\text{-}HT_{2A}$ nanobody VGS-Nb2.

We first confirmed the interaction between $5\text{-}HT_{2A}$ and VGS-Nb2 via coimmunoprecipitation (coIP), as analyzed by western

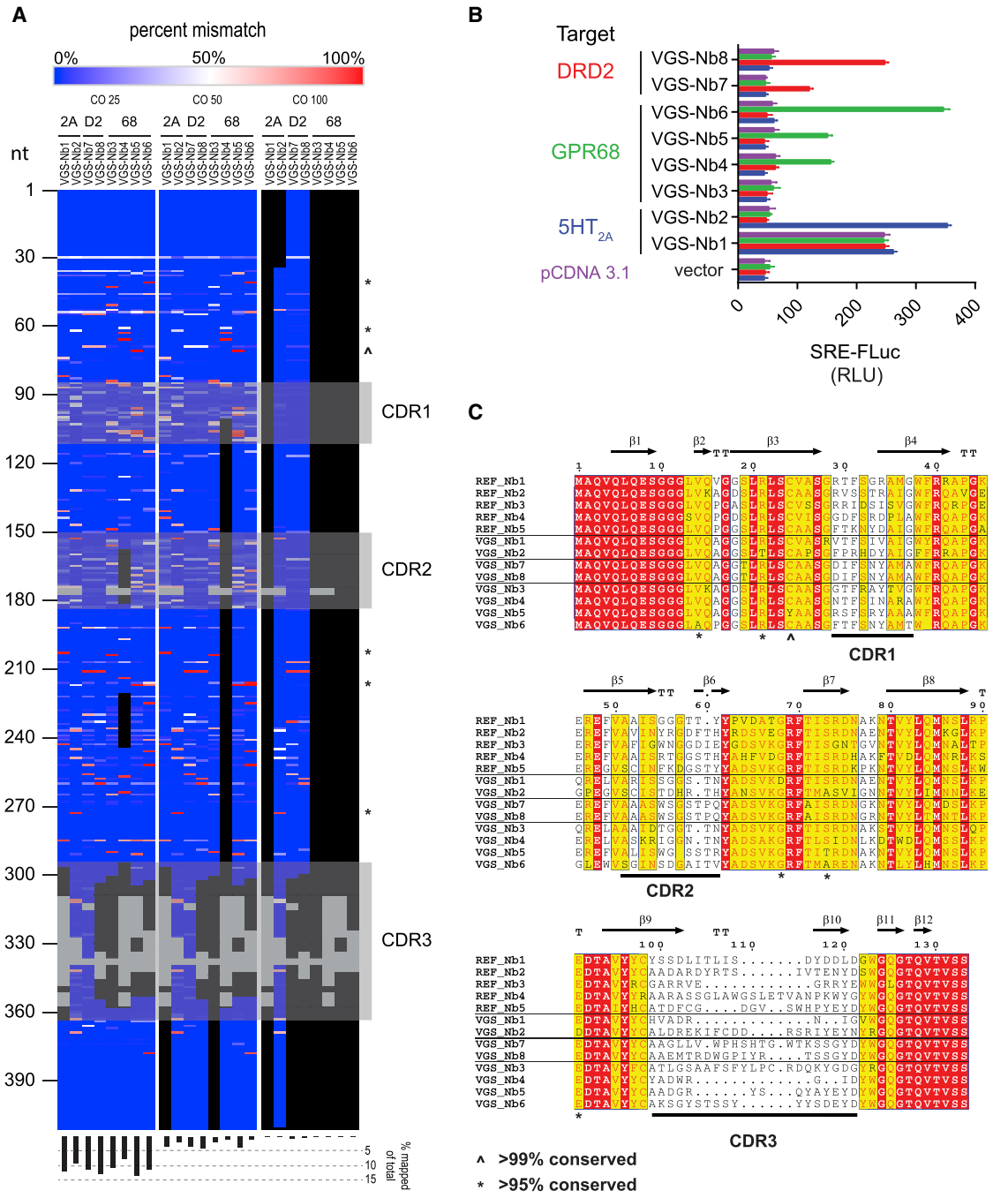


Figure 4. VEGAS for Evolution of Active-State Nanobodies

VEGAS was used to develop nanobodies that selectively activate diverse GPCR targets from a single cDNA library.

(A) Deep sequencing of the nanobody cDNA library used for VEGAS. 20 million reads were aligned to VEGAS-derived clones and plotted as % mismatch. Data was analyzed with score cut-offs (CO) of 25, 50, and 100 (STAR Methods). Grey blocks are gaps in alignment as per (C). Black blocks are regions with mapped reads <2,000 counts (<0.0001%). Symbols ^ and * mirror those on (C). Bottom histogram, percent total mapped reads for each alignment.

(B) Serum-response element (SRE) reporter assay. Nanobody:receptor:reporter transfection ratio of 5:1:1. FLuc production equates to relative receptor activation. Data are represented as mean ± SEM, N = 3.

(C) Amino acid sequence alignment of library (REF_NB#) and VEGAS-derived clones. Shading: 100% (red), >75% (yellow), <75% (white). Variations identified in VEGAS, but not reference sequence, derived clones at positions of high genetic conservation (see Table S4) are annotated, ^ >99% conserved, * >95% conserved. Nanobody secondary structure annotated above, retrieved from PDB 3P0G, chain B. Complementarity determining regions (CDRs) annotated below. β, beta sheet. TT, strict β-turn.

See also Figure S4 and Data S1 and S2.

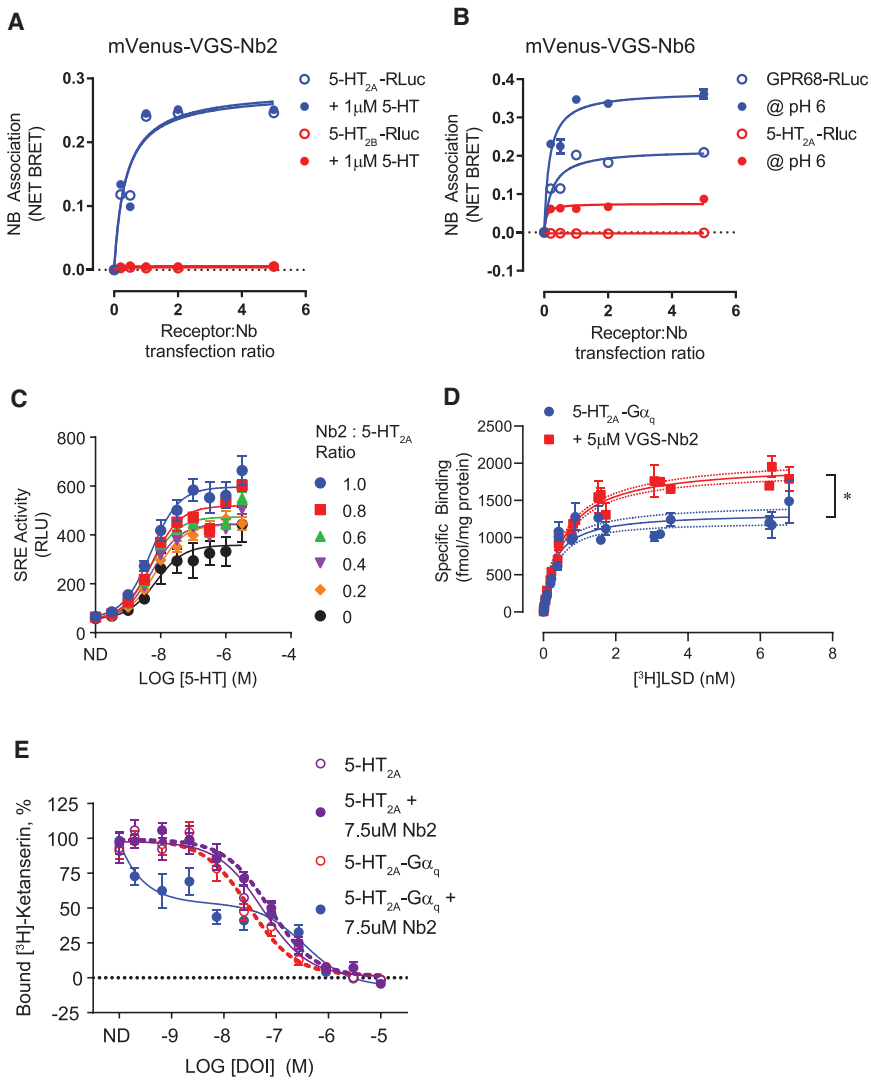


Figure 5. Positive Allosteric Modulation of GPCRs by VEGAS-Evolved Nanobodies

VEGAS-derived nanobodies were tested for direct association and allosteric modulation of their targets.

(A) Bioluminescence resonance energy transfer (BRET) association assay between 5-HT_{2A}-RLuc or 5-HT_{2B}-RLuc and mVenus-VGS-Nb2 at increasing transfection ratios of nanobody. Data are represented as mean ± SEM, N = 3. Symbols for 5-HT_{2B}-Rluc data underlie those for the +1 μM 5-HT data.

(B) Bioluminescence resonance energy transfer (BRET) association assay between GPR68-RLuc or 5-HT_{2A}-RLuc and mVenus-VGS-Nb6. Data are represented as mean ± SEM, N = 3.

(C) Serum response element (SRE) reporter assay. RLuc production equates to relative receptor activation. Data are represented as mean ± SEM, N = 3.

(D) Saturation radioligand binding assay. 5-HT_{2A}-$G\alpha_q$ membrane treated with vehicle or 5 μM VGS-Nb2. 5-HT_{2A}; K_d = 0.30 nM, B_{max} = 1,333 fmol/mg. 5-HT_{2A}+VGS-Nb2; K_d = 0.566 nM, B_{max} = 1,993 fmol/mg. Data are represented as mean ± SEM, N = 3, *p < 0.05.

(E) Competitive radioligand binding assay. 5-HT_{2A} and 5-HT_{2A}-$G\alpha_q$ membrane treated with vehicle or 7.5 μM VGS-Nb2 labeled with 1 nM [^{3}H]ketanserin and increasing concentrations of DOI. Data are represented as total-count normalized means ± SEM, N = 3.

See also Figure S5, Table S5, and Data S1 and S2.

blot (Figure S5B) and mass spectrometry (Figure S5E). These assays confirmed a stable interaction between the receptor and nanobody. These assays were performed in the absence of ligand, confirming our previous BRET observation that the interaction is ligand-independent. This is consistent with our directed evolution selection paradigm, which required a nanobody capable inducing receptor activity in the absence of ligand.

5-HT_{2A} couples to $G\alpha_q$ and β-arrestin to transduce its signal in cells (Wacker et al., 2017b). We tested whether VGS-Nb2 association to 5-HT_{2A} was Gα-dependent using $G\alpha_{q\Delta/11\Delta/s\Delta}$ knockout cell lines (Alvarez-Curto et al., 2016) in BRET recruitment assays (Figure S5A), coIP by western blot (Figure S5B), and mass spectrometry (Figure S5E). In all three studies, no appreciable difference in VGS-Nb2 association was detected. Notably, as with all epistasis experiments, knockout cell lines frequently adapt to gene loss by augmenting signaling pathways (Duncan et al., 2012; Luttrell et al., 2018). However, from the additional proteins identified in HEK-T and $G\alpha_{q\Delta/11\Delta/s\Delta}$ cells (Table S5), no canonical GPCR transducers were identified. We concluded from this analysis that VGS-Nb2 does not stabilize, or lock, a transducer-coupled state. This aligns with the evolved purpose of this nanobody to act as a positive allosteric modulator (PAM) of 5-HT_{2A} rather than inhibit transducer cycling.

To further validate VGS-Nb2 as a 5-HT_{2A} PAM, we first assessed its ability to positively modulate SRE signaling downstream of 5-HT_{2A}. At Nb ratios demonstrated to bind <50% of 5-HT_{2A} (Figure 5A), VGS-Nb2 increased the agonist-mediated SRE response by up to 2-fold (Figure 5C). This SRE signal could originate from $G\alpha_q$ and/or β-arrestin pathways. We assessed the effect of the nanobody using calcium and arrestin recruitment assays, respectively. VGS-Nb2 allosterically enhanced 5-HT_{2A} calcium release (Figure S5F), a $G\alpha_q$-mediated signal response. Conversely, VGS-Nb2 diminished mVenus-β-arrestin2 recruitment to the 5-HT_{2A}-RLuc fusion protein as a function of time (Figure S5G) and agonist concentration (Figures S5H and S5I). From these experiments, we hypothesized that VGS-Nb2 stabilizes the high-affinity $G\alpha_q$-coupled state of the receptor.

Unliganded receptors are rarely found in their active, or high-affinity, conformational state (Manglik et al., 2015). However, radiolabeled ligands can be used to probe and quantify high-affinity receptor sites. The number of these sites increases when allosteric effectors, such as $G\alpha_q$ or nanobodies, are bound

to the receptor (Che et al., 2018; Staus et al., 2016; Strachan et al., 2014). To test whether VGS-Nb2 stabilizes the 5-HT_{2A} $G\alpha_q$-coupled active state, as predicted from our functional data, we first employed radioligand saturation binding using the partial agonist [^{3}H]LSD (Figure 5D). Membranes from cells transfected with 5-HT_{2A} fused to its transducer $G\alpha_q$ (5-HT_{2A}-Gq) were incubated with increasing concentrations of [^{3}H]LSD ± 5 μM purified VGS-Nb2. As shown in Figure 5D, VGS-Nb2 increased labeled agonist binding sites by 50%. This increase in high-affinity agonist binding sites was additionally confirmed through competitive radioligand binding wherein 5-HT_{2A} and 5-HT_{2A}-$G\alpha_q$ membranes were incubated with ± 7.5 μM purified VGS-Nb2 (Figure 5E). In competition with the radiolabeled 5-HT_{2A} antagonist [^{3}H]ketanserin, the selective agonist DOI bound the 5-HT_{2A} receptor with a half maximal inhibitory concentration (IC_{50}) equal to 550 nM. Neither a local excess of $G\alpha_q$ (DOI IC_{50} = 307 nM, p = 0.1264) nor the addition of purified VGS-Nb2 (DOI IC_{50} = 746 nM, p = 0.3957) significantly affected DOI binding. However, in the presence of both $G\alpha_q$ and VGS-Nb2, ~50% of the available ligand binding sites were stabilized in the high-affinity conformation that bound DOI with an IC_{50} = 0.15 nM.

We have therefore demonstrated the directed evolution of multiple functionally distinct nanobody sequences against GPCRs using VEGAS. Of these, we have characterized VGS-Nb2 as a $G\alpha_q$-dependent positive allosteric modulator of 5-HT_{2A}.

DISCUSSION

Here, we demonstrate the development of a system for facile directed evolution in mammalian cells: VEGAS. Leveraging the alpha virus Sindbis as a vector for heredity, mutagenesis, and selection, we succeeded in evolving context-dependent functions for three independent classes of proteins. Our evolution targets were primary (tTA), secondary (GPCR), and tertiary (Nbs) interactors to downstream outputs, demonstrating the ability of VEGAS to provide tools at multiple levels of cell signaling. Our primary system evolved tTA to engage with $TETO_7$ in the presence of >1 μM DOX. Our secondary system evolved the GPCR MRGPRX2 to constitutively activate the serum response element via endogenous signaling pathways. Our tertiary system evolved nanobodies to selectively activate GPCRs, which in turn activated the serum response element via endogenous signaling pathways. Together, these applications showcase the ease and power of the VEGAS system as a tool for enabling directed evolution campaigns across a broad range of potential mammalian applications.

Directed evolution allows genetic sequences to evolve under selective pressure in an appropriate context. Through this process we are able to guide solutions to otherwise intractable biological problems (Hammer et al., 2017; Kan et al., 2016; Matsumoto et al., 2015; Shapiro et al., 2010). However, powerful systems for directed evolution in a mammalian cell context have lagged behind unicellular systems. VEGAS offers three major advantages for the directed evolution of biomedical tools and therapies.

First, VEGAS evolves within the signaling framework of the host cell. Signaling proteins never operate in isolation, but as interacting heteromeric complexes, to transfer information through the cell (Garrington and Johnson, 1999; Pawson and Scott, 1997; Purvis and Lahav, 2013; Varnaitė and MacNeill, 2016). The timing, location, and kinetics of these interactions is critical to performance and cannot be easily replicated in non-native environments. In addition, we can take advantage of the negative feedback (Amit et al., 2007; Behar et al., 2007; English et al., 2015; Ferrell, 2002; Howell et al., 2012; Subramaniam et al., 1989) mechanisms built in to endogenous signaling pathways to encode viral selection—as was done for both MRGPRX2- and nanobody-directed evolution in this study.

Second, VEGAS is wholly dependent on the host cell for transgene maturation (Garcia-Moreno et al., 2013). Directed evolution performance can falter when transferring tools evolved in one context to another (Armbruster et al., 2007). This may be a consequence of improper trafficking, failed compartmentalization, incorrect protein maturation, or absence of non-native co-factors. With VEGAS, mammalian translation is a requirement of the evolved product.

Third, VEGAS selection is constant and highly mutagenic, enabling it to overcome many of the pitfalls inherent to complex fitness landscapes (Romero and Arnold, 2009; Tracewell and Arnold, 2009). To avoid dead-ends and early fitness bias, directed evolution systems must sample toward saturation whenever possible. This directed evolution paradigm helps to maintain diversity by preserving even poor performing early evolution variants, which may ultimately rise to the highest fitness peaks. This is achieved with VEGAS, in part, because each host cell operates as a closed system. This allows evolved solutions derived in each cell to compete in the subsequent rounds—even when vastly superior solutions may have arisen elsewhere within the same selection cycle.

There are many potential applications of the VEGAS system. Sindbis virus has a transgene packaging capacity of >6 kb (Huang and Summers, 1991), placing few limits on the potential targets for directed evolution. High value targets would include: Cas9 variants evolved to better engage endogenous sequences (Doudna and Charpentier, 2014; Kleinstiver et al., 2015; Lee et al., 2018), fluorescent protein variants evolved for maturity time, photostability, brightness, or wavelength specificity in human tissue (Drobizhev et al., 2011; Piatkevich et al., 2017; Shaner et al., 2004, 2008), or designer receptors exclusively activated by designer drugs (DREADDs) for the chemogenetic control of cell signaling (Armbruster et al., 2007; Roth, 2016).

There also remains significant room for improvement. VEGAS has performed well with positive genetic selection and, as a synthetic biology system, can be engineered to perform AND and OR logic gated selection paradigms. It cannot at this time generate exclusive or NOT logic gated selection paradigms. Developing a dominant negative selective pressure for VEGAS will make these modes possible. In addition, a method to tune the speed of replication would be advantageous for developing slower phase gene circuits including those coupled to the cell cycle, metabolism, or other slow-maturing signal classes. Adapting VEGAS for mammalian cell bioreactors would also allow it to be applied as a continuous system (Badran and Liu, 2015). Here, we add VEGAS to the growing toolbox of synthetic

biology, filling an essential need for facile directed evolution in a mammalian context.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- METHOD DETAILS
  - Molecular Biology & Plasmid Construction
  - General Cell Culture
  - Sindbis Virus Production
  - Sustained Passage of Sindbis for Directed Evolution and Transgene Isolation
  - RNA Deep Sequencing
  - Quantification of Viral RNA via qRT-PCR
  - tTA Reporter Assay
  - Transcription Factor Reporter Primary Screen
  - SRE Reporter Assay
  - TANGO β-arrestin recruitment assay
  - Calcium flux assay
  - Bioluminescence resonance energy transfer association assay
  - Phosphoinositide hydrolysis assay
  - Surface expression enzyme-linked immunosorbent assay
  - Nanobody Production
  - Saturation and competitive radioligand binding assays
  - Co-Immunoprecipitation Analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2019.05.051.

ACKNOWLEDGMENTS

We thank our colleague Gary Johnson for providing resources for RNA-seq analysis. We also thank our colleague Mark Heise for providing the initial Sindbis virus vectors used in this study. This work was supported by NIH (1U01MH105892-01, 5U24DK116195-02, RO1MH112205, R37DA045657) and a grant from the University of North Carolina Lineberger Comprehensive Cancer Center to B.L.R.

AUTHOR CONTRIBUTIONS

Conceptualization, J.G.E.; Methodology, A.S.C., B.L.R., D.W., J.G.E., R.H.J.O., and R.T.S.; Software, D.S. and J.G.E.; Validation, J.G.E., K.L., R.H.J.O., and R.T.S.; Formal Analysis, J.G.E. and D.S.; Investigation, J.G.E., K.L., R.H.J.O., and R.T.S.; Resources, D.W., J.G.E., K.W., and M.P.; Data Curation, D.S. and J.G.E.; Writing – Original Draft, J.G.E.; Writing – Review & Editing, A.S.C., B.L.R., D.W., J.G.E., K.L., R.H.J.O., and R.T.S.; Visualization, J.G.E.; Supervision, B.L.R. and J.G.E.; Project Administration, B.L.R. and J.G.E.; Funding Acquisition, B.L.R. and J.G.E.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 19, 2018
Revised: February 6, 2019
Accepted: May 23, 2019
Published: July 4, 2019; corrected online: July 19, 2019

REFERENCES

Agapov, E.V., Frolov, I., Lindenbach, B.D., Prágai, B.M., Schlesinger, S., and Rice, C.M. (1998). Noncytopathic Sindbis virus RNA vectors for heterologous gene expression. Proc. Natl. Acad. Sci. USA *95*, 12989–12994.

Albert, T.J., Dailidiene, D., Dailide, G., Norton, J.E., Kalia, A., Richmond, T.A., Molla, M., Singh, J., Green, R.D., and Berg, D.E. (2005). Mutation discovery in bacterial genomes: metronidazole resistance in *Helicobacter pylori*. Nat. Methods *2*, 951–953.

Alvarez-Curto, E., Inoue, A., Jenkins, L., Raihan, S.Z., Prihandoko, R., Tobin, A.B., and Milligan, G. (2016). Targeted Elimination of G Proteins and Arrestins Defines Their Specific Contributions to Both Intensity and Duration of G Protein-coupled Receptor Signaling. J. Biol. Chem. *291*, 27147–27159.

Amit, I., Citri, A., Shay, T., Lu, Y., Katz, M., Zhang, F., Tarcic, G., Siwak, D., Lahad, J., Jacob-Hirsch, J., et al. (2007). A module of negative feedback regulators defines growth factor signaling. Nat. Genet. *39*, 503–512.

Armbruster, B.N., Li, X., Pausch, M.H., Herlitze, S., and Roth, B.L. (2007). Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. Proc. Natl. Acad. Sci. USA *104*, 5163–5168.

Arnold, F.H. (1998). Design by Directed Evolution. Acc. Chem. Res. *31*, 125–131.

Badran, A.H., and Liu, D.R. (2015). In vivo continuous directed evolution. Curr. Opin. Chem. Biol. *24*, 1–10.

Baym, M., Lieberman, T.D., Kelsic, E.D., Chait, R., Gross, R., Yelin, I., and Kishony, R. (2016). Spatiotemporal microbial evolution on antibiotic landscapes. Science *353*, 1147–1151.

Behar, M., Hao, N., Dohlman, H.G., and Elston, T.C. (2007). Mathematical and computational analysis of adaptation via feedback inhibition in signal transduction pathways. Biophys. J. *93*, 806–821.

Berens, C., Altschmied, L., and Hillen, W. (1992). The role of the N terminus in Tet repressor for tet operator binding determined by a mutational analysis. J. Biol. Chem. *267*, 1945–1952.

Berman, C.M., Papa, L.J., 3rd, Hendel, S.J., Moore, C.L., Suen, P.H., Weickhardt, A.F., Doan, N.-D., Kumar, C.M., Uil, T.G., Butty, V.L., et al. (2018). An Adaptable Platform for Directed Evolution in Human Cells. J. Am. Chem. Soc. *140*, 18093–18103.

Bourdon, D.M., Wing, M.R., Edwards, E.B., Sondek, J., and Harden, T.K. (2006). Quantification of Isozyme-Specific Activation of Phospholipase C-β2 by Rac GTPases and Phospholipase C-E by Rho GTPases in an Intact Cell Assay System. Methods Enzymol. *406*, 489–499.

Bredenbeek, P.J., Frolov, I., Rice, C.M., and Schlesinger, S. (1993). Sindbis virus expression vectors: packaging of RNA replicons by using defective helper RNAs. J. Virol. *67*, 6439–6446.

Buchholz, F., Angrand, P.-O., and Stewart, A.F. (1998). Improved properties of FLP recombinase evolved by cycling mutagenesis. Nat. Biotechnol. *16*, 657–662.

Campbell, R.E., Tour, O., Palmer, A.E., Steinbach, P.A., Baird, G.S., Zacharias, D.A., and Tsien, R.Y. (2002). A monomeric red fluorescent protein. Proc. Natl. Acad. Sci. USA *99*, 7877–7882.

Carlson, J.C., Badran, A.H., Guggiana-Nilo, D.A., and Liu, D.R. (2014). Negative selection and stringency modulation in phage-assisted continuous evolution. Nat. Chem. Biol. *10*, 216–222.

Chan, K.Y., Jang, M.J., Yoo, B.B., Greenbaum, A., Ravi, N., Wu, W.-L., Sánchez-Guardado, L., Lois, C., Mazmanian, S.K., Deverman, B.E., and Gradinaru, V. (2017). Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. Nat. Neurosci. *20*, 1172–1179.

Che, T., Majumdar, S., Zaidi, S.A., Ondachi, P., McCorvy, J.D., Wang, S., Mosier, P.D., Uprety, R., Vardy, E., Krumm, B.E., et al. (2018). Structure of the Nanobody-Stabilized Active State of the Kappa Opioid Receptor. Cell *172*, 55–67.

Chen, K., and Arnold, F.H. (1993). Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. Proc. Natl. Acad. Sci. USA *90*, 5618–5622.

Crameri, A., Whitehorn, E.A., Tate, E., and Stemmer, W.P.C. (1996). Improved green fluorescent protein by molecular evolution using DNA shuffling. Nat. Biotechnol. *14*, 315–319.

Darwin, C., and Bynum, W.F. (2009). On the Origin of Species by Means of Natural Selection: Or, the Preservation of Favored Races in the Struggle for Life (Penguin Classics).

Das, A.T., Zhou, X., Vink, M., Klaver, B., Verhoef, K., Marzio, G., and Berkhout, B. (2004). Viral evolution as a tool to improve the tetracycline-regulated gene expression system. J. Biol. Chem. *279*, 18776–18782.

De Lean, A., Stadel, J.M., and Lefkowitz, R.J. (1980). A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled β-adrenergic receptor. J. Biol. Chem. *255*, 7108–7117.

Di Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J.-M., Taly, J.-F., and Notredame, C. (2011). T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. Nucleic Acids Res. *39*, W13-7.

Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication. Nature *418*, 700–707.

Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096.

Drake, J.W., and Holland, J.J. (1999). Mutation rates among RNA viruses. Proc. Natl. Acad. Sci. USA *96*, 13910–13913.

Drake, J.W., Charlesworth, B., Charlesworth, D., and Crow, J.F. (1998). Rates of spontaneous mutation. Genetics *148*, 1667–1686.

Driscoll, C.A., Macdonald, D.W., and O'Brien, S.J. (2009). From wild animals to domestic pets, an evolutionary view of domestication. Proc. Natl. Acad. Sci. USA *106* (*Suppl 1*), 9971–9978.

Drobizhev, M., Makarov, N.S., Tillo, S.E., Hughes, T.E., and Rebane, A. (2011). Two-photon absorption properties of fluorescent proteins. Nat. Methods *8*, 393–399.

Dror, R.O., Arlow, D.H., Maragakis, P., Mildorf, T.J., Pan, A.C., Xu, H., Borhani, D.W., and Shaw, D.E. (2011). Activation mechanism of the β2-adrenergic receptor. Proc. Natl. Acad. Sci. USA *108*, 18684–18689.

Duncan, J.S., Whittle, M.C., Nakamura, K., Abell, A.N., Midland, A.A., Zawistowski, J.S., Johnson, N.L., Granger, D.A., Jordan, N.V., Darr, D.B., et al. (2012). Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer. Cell *149*, 307–321.

English, J.G., Shellhammer, J.P., Malahe, M., McCarter, P.C., Elston, T.C., and Dohlman, H.G. (2015). MAPK feedback encodes a switch and timer for tunable stress adaptation in yeast. Sci. Signal. *8*, ra5.

Esvelt, K.M., Carlson, J.C., and Liu, D.R. (2011). A system for the continuous directed evolution of biomolecules. Nature *472*, 499–503.

Fan, H., Irwin, J.J., Webb, B.M., Klebe, G., Shoichet, B.K., and Sali, A. (2009). Molecular docking screens using comparative models of proteins. J. Chem. Inf. Model. *49*, 2512–2527.

Ferrell, J.E., Jr. (2002). Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. Curr. Opin. Cell Biol. *14*, 140–148.

Filizola, M., and Devi, L.A. (2013). Grand opening of structure-guided design for novel opioids. Trends Pharmacol. Sci. *34*, 6–12.

Fredriksson, R., Lagerström, M.C., Lundin, L.-G., and Schiöth, H.B. (2003). The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol. Pharmacol. *63*, 1256–1272.

Frolov, I., Hardy, R., and Rice, C.M. (2001). Cis-acting RNA elements at the 5′ end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. RNA *7*, 1638–1651.

Garcia-Moreno, M., Sanz, M.A., Pelletier, J., and Carrasco, L. (2013). Requirements for eIF4A and eIF2 during translation of Sindbis virus subgenomic mRNA in vertebrate and invertebrate host cells. Cell. Microbiol. *15*, 823–840.

Garrington, T.P., and Johnson, G.L. (1999). Organization and regulation of mitogen-activated protein kinase signaling pathways. Curr. Opin. Cell Biol. *11*, 211–218.

Gilman, A.G. (1987). G proteins: transducers of receptor-generated signals. Annu. Rev. Biochem. *56*, 615–649.

Gossen, M., and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. USA *89*, 5547–5551.

Gossen, M., Freundlieb, S., Bender, G., Müller, G., Hillen, W., and Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. Science *268*, 1766–1769.

Hammer, S.C., Kubik, G., Watkins, E., Huang, S., Minges, H., and Arnold, F.H. (2017). Anti-Markovnikov alkene oxidation by metal-oxo-mediated enzyme catalysis. Science *358*, 215–218.

Hanes, J., and Plückthun, A. (1997). In vitro selection and evolution of functional proteins by using ribosome display. Proc. Natl. Acad. Sci. USA *94*, 4937–4942.

Hartwell, L.H., Culotti, J., and Reid, B. (1970). Genetic control of the cell-division cycle in yeast. I. Detection of mutants. Proc. Natl. Acad. Sci. USA *66*, 352–359.

Hauser, A.S., Attwood, M.M., Rask-Andersen, M., Schiöth, H.B., and Gloriam, D.E. (2017). Trends in GPCR drug discovery: new agents, targets and indications. Nat. Rev. Drug Discov. *16*, 829–842.

Hecht, B., Müller, G., and Hillen, W. (1993). Noninducible Tet repressor mutations map from the operator binding motif to the C terminus. J. Bacteriol. *175*, 1206–1210.

Hess, G.T., Frésard, L., Han, K., Lee, C.H., Li, A., Cimprich, K.A., Montgomery, S.B., and Bassik, M.C. (2016). Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. Nat. Methods *13*, 1036–1042.

Howell, A.S., Jin, M., Wu, C.-F., Zyla, T.R., Elston, T.C., and Lew, D.J. (2012). Negative feedback enhances robustness in the yeast polarity establishment circuit. Cell *149*, 322–333.

Huang, M.J., and Summers, J. (1991). Infection initiated by the RNA pregenome of a DNA virus. J. Virol. *65*, 5435–5439.

Huang, X.-P., Setola, V., Yadav, P.N., Allen, J.A., Rogan, S.C., Hanson, B.J., Revankar, C., Robers, M., Doucette, C., and Roth, B.L. (2009). Parallel functional activity profiling reveals valvulopathogens are potent 5-hydroxytryptamine(2B) receptor agonists: implications for drug safety assessment. Mol. Pharmacol. *76*, 710–722.

Huang, W., Manglik, A., Venkatakrishnan, A.J., Laeremans, T., Feinberg, E.N., Sanborn, A.L., Kato, H.E., Livingston, K.E., Thorsen, T.S., Kling, R.C., et al. (2015a). Structural insights into μ-opioid receptor activation. Nature *524*, 315–321.

Huang, X.-P., Karpiak, J., Kroeze, W.K., Zhu, H., Chen, X., Moy, S.S., Saddoris, K.A., Nikolova, V.D., Farrell, M.S., Wang, S., et al. (2015b). Allosteric ligands for the pharmacologically dark receptors GPR68 and GPR65. Nature *527*, 477–483.

Huston, M. (1979). A General Hypothesis of Species Diversity. Am. Nat. *113*, 81–101.

Isberg, V., de Graaf, C., Bortolato, A., Cherezov, V., Katritch, V., Marshall, F.H., Mordalski, S., Pin, J.-P., Stevens, R.C., Vriend, G., and Gloriam, D.E. (2015). Generic GPCR residue numbers - aligning topology maps while minding the gaps. Trends Pharmacol. Sci. *36*, 22–31.

Kan, S.B.J., Lewis, R.D., Chen, K., and Arnold, F.H. (2016). Directed evolution of cytochrome c for carbon-silicon bond formation: Bringing silicon to life. Science *354*, 1048–1051.

Katritch, V., Fenalti, G., Abola, E.E., Roth, B.L., Cherezov, V., and Stevens, R.C. (2014). Allosteric sodium in class A GPCR signaling. Trends Biochem. Sci. *39*, 233–244.

Kisker, C., Hinrichs, W., Tovar, K., Hillen, W., and Saenger, W. (1995). The complex formed between Tet repressor and tetracycline-Mg2+ reveals mechanism of antibiotic resistance. J. Mol. Biol. *247*, 260–280.

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V.V., Nguyen, N.T., Zheng, Z., Gonzales, A.P.W., Li, Z., Peterson, R.T., Yeh, J.-R.J., et al. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature *523*, 481–485.

Kobilka, B.K., and Deupi, X. (2007). Conformational complexity of G-protein-coupled receptors. Trends Pharmacol. Sci. *28*, 397–406.

Kroeze, W.K., Sassano, M.F., Huang, X.-P., Lansu, K., McCorvy, J.D., Giguère, P.M., Sciaky, N., and Roth, B.L. (2015). PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. Nat. Struct. Mol. Biol. *22*, 362–369.

Kuchner, O., and Arnold, F.H. (1997). Directed evolution of enzyme catalysts. Trends Biotechnol. *15*, 523–530.

Lansu, K., Karpiak, J., Liu, J., Huang, X.-P., McCorvy, J.D., Kroeze, W.K., Che, T., Nagase, H., Carroll, F.I., Jin, J., et al. (2017). *In silico* design of novel probes for the atypical opioid receptor MRGPRX2. Nat. Chem. Biol. *13*, 529–536.

Latorraca, N.R., Venkatakrishnan, A.J., and Dror, R.O. (2017). GPCR Dynamics: Structures in Motion. Chem. Rev. *117*, 139–155.

Lee, J.K., Jeong, E., Lee, J., Jung, M., Shin, E., Kim, Y.H., Lee, K., Jung, I., Kim, D., Kim, S., and Kim, J.S. (2018). Directed evolution of CRISPR-Cas9 to increase its specificity. Nat. Commun. *9*, 3048.

Luttrell, L.M., Wang, J., Plouffe, B., Smith, J.S., Yamani, L., Kaur, S., Jean-Charles, P.-Y., Gauthier, C., Lee, M.-H., Pani, B., et al. (2018). Manifold roles of β-arrestins in GPCR signaling elucidated with siRNA and CRISPR/Cas9. Sci. Signal *11*, eaat7650.

Maheshri, N., Koerber, J.T., Kaspar, B.K., and Schaffer, D.V. (2006). Directed evolution of adeno-associated virus yields enhanced gene delivery vectors. Nat. Biotechnol. *24*, 198–204.

Manglik, A., Kim, T.H., Masureel, M., Altenbach, C., Yang, Z., Hilger, D., Lerch, M.T., Kobilka, T.S., Thian, F.S., Hubbell, W.L., et al. (2015). Structural insights into the dynamic process of β2-adrenergic receptor signaling. Cell *161*, 1101–1111.

Manglik, A., Lin, H., Aryal, D.K., McCorvy, J.D., Dengler, D., Corder, G., Levit, A., Kling, R.C., Bernat, V., Hübner, H., et al. (2016). Structure-based discovery of opioid analgesics with reduced side effects. Nature *537*, 185–190.

Matsumoto, Y., Chen, R., Anikeeva, P., and Jasanoff, A. (2015). Engineering intracellular biomineralization and biosensing by a magnetic protein. Nat. Commun. *6*, 8721.

McCorvy, J.D., and Roth, B.L. (2015). Structure and function of serotonin G protein-coupled receptors. Pharmacol. Ther. *150*, 129–142.

McMahon, C., Baier, A.S., Pascolutti, R., Wegrecki, M., Zheng, S., Ong, J.X., Erlandson, S.C., Hilger, D., Rasmussen, S.G.F., Ring, A.M., et al. (2018). Yeast surface display platform for rapid discovery of conformationally selective nanobodies. Nat. Struct. Mol. Biol. *25*, 289–296.

Michino, M., Beuming, T., Donthamsetti, P., Newman, A.H., Javitch, J.A., and Shi, L. (2015). What can crystal structures of aminergic receptors tell us about designing subtype-selective ligands? Pharmacol. Rev. *67*, 198–213.

Morley, V.J., and Turner, P.E. (2017). Dynamics of molecular evolution in RNA virus populations depend on sudden versus gradual environmental change. Evolution *71*, 872–883.

Müller, G., Hecht, B., Helbl, V., Hinrichs, W., Saenger, W., and Hillen, W. (1995). Characterization of non-inducible Tet repressor mutants suggests conformational changes necessary for induction. Nat. Struct. Biol. *2*, 693–703.

Muyldermans, S., Cambillau, C., and Wyns, L. (2001). Recognition of antigens by single-domain antibody fragments: the superfluous luxury of paired domains. Trends Biochem. Sci. *26*, 230–235.

Nakamura, Y., Gojobori, T., and Ikemura, T. (2000). Codon usage tabulated from international DNA sequence databases: status for the year 2000. Nucleic Acids Res. *28*, 292.

Neubig, R.R., Spedding, M., Kenakin, T., and Christopoulos, A.; International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification (2003). International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII. Update on terms and symbols in quantitative pharmacology. Pharmacol. Rev. *55*, 597–606.

Onaran, H.O., and Costa, T. (2009). Allosteric coupling and conformational fluctuations in proteins. Curr. Protein Pept. Sci. *10*, 110–115.

Orth, P., Cordes, F., Schnappinger, D., Hillen, W., Saenger, W., and Hinrichs, W. (1998). Conformational changes of the Tet repressor induced by tetracycline trapping. J. Mol. Biol. *279*, 439–447.

Orth, P., Schnappinger, D., Sum, P.E., Ellestad, G.A., Hillen, W., Saenger, W., and Hinrichs, W. (1999a). Crystal structure of the tet repressor in complex with a novel tetracycline, 9-(N,N-dimethylglycylamido)- 6-demethyl-6-deoxy-tetracycline. J. Mol. Biol. *285*, 455–461.

Orth, P., Saenger, W., and Hinrichs, W. (1999b). Tetracycline-chelated Mg2+ ion initiates helix unwinding in Tet repressor induction. Biochemistry *38*, 191–198.

Orth, P., Schnappinger, D., Hillen, W., Saenger, W., and Hinrichs, W. (2000). Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. Nat. Struct. Biol. *7*, 215–219.

Pándy-Szekeres, G., Munk, C., Tsonkov, T.M., Mordalski, S., Harpsøe, K., Hauser, A.S., Bojarski, A.J., and Gloriam, D.E. (2018). GPCRdb in 2018: adding GPCR structure models and ligands. Nucleic Acids Res. *46* (D1), D440–D446.

Pardon, E., Laeremans, T., Triest, S., Rasmussen, S.G.F., Wohlkönig, A., Ruf, A., Muyldermans, S., Hol, W.G.J., Kobilka, B.K., and Steyaert, J. (2014). A general protocol for the generation of Nanobodies for structural biology. Nat. Protoc. *9*, 674–693.

Pawson, T., and Scott, J.D. (1997). Signaling through scaffold, anchoring, and adaptor proteins. Science *278*, 2075–2080.

Peng, Y., McCorvy, J.D., Harpsøe, K., Lansu, K., Yuan, S., Popov, P., Qu, L., Pu, M., Che, T., Nikolajsen, L.F., et al. (2018). 5-HT_{2C} Receptor Structures Reveal the Structural Basis of GPCR Polypharmacology. Cell *172*, 719–730.

Piatkevich, K.D., Suk, H.-J., Kodandaramaiah, S.B., Yoshida, F., DeGennaro, E.M., Drobizhev, M., Hughes, T.E., Desimone, R., Boyden, E.S., and Verkhusha, V.V. (2017). Near-Infrared Fluorescent Proteins Engineered from Bacterial Phytochromes in Neuroimaging. Biophys. J. *113*, 2299–2309.

Pierce, K.L., Premont, R.T., and Lefkowitz, R.J. (2002). Seven-transmembrane receptors. Nat. Rev. Mol. Cell Biol. *3*, 639–650.

Purvis, J.E., and Lahav, G. (2013). Encoding and decoding cellular information through signaling dynamics. Cell *152*, 945–956.

Rasmussen, S.G.F., Choi, H.-J., Rosenbaum, D.M., Kobilka, T.S., Thian, F.S., Edwards, P.C., Burghammer, M., Ratnala, V.R.P., Sanishvili, R., Fischetti, R.F., et al. (2007). Crystal structure of the human β_2 adrenergic G-protein-coupled receptor. Nature *450*, 383–387.

Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. *42*, W320-4.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.

Romero, P.A., and Arnold, F.H. (2009). Exploring protein fitness landscapes by directed evolution. Nat. Rev. Mol. Cell Biol. *10*, 866–876.

Roth, B.L. (2016). DREADDs for Neuroscientists. Neuron *89*, 683–694.

Roth, B.L., Irwin, J.J., and Shoichet, B.K. (2017). Discovery of new GPCR ligands to illuminate new biology. Nat. Chem. Biol. *13*, 1143–1151.

Sane, J., Kurkela, S., Levanov, L., Nikkari, S., Vaheri, A., and Vapalahti, O. (2012). Development and evaluation of a real-time RT-PCR assay for Sindbis virus detection. J. Virol. Methods *179*, 185–188.

Sanjuán, R., Nebot, M.R., Chirico, N., Mansky, L.M., and Belshaw, R. (2010). Viral mutation rates. J. Virol. *84*, 9733–9748.

Sarkar, C.A., Dodevski, I., Kenig, M., Dudli, S., Mohr, A., Hermans, E., and Plückthun, A. (2008). Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. Proc. Natl. Acad. Sci. USA *105*, 14808–14813.

Schlesinger, S. (1993). Alphaviruses–vectors for the expression of heterologous genes. Trends Biotechnol. *11*, 18–22.

Schnell, M.J., Buonocore, L., Whitt, M.A., and Rose, J.K. (1996). The minimal conserved transcription stop-start signal promotes stable expression of a foreign gene in vesicular stomatitis virus. J. Virol. *70*, 2318–2323.

Scholz, O., Henssler, E.-M., Bail, J., Schubert, P., Bogdanska-Urbaniak, J., Sopp, S., Reich, M., Wisshak, S., Köstner, M., Bertram, R., and Hillen, W. (2004). Activity reversal of Tet repressor caused by single amino acid exchanges. Mol. Microbiol. *53*, 777–789.

Schubert, P., Schnappinger, D., Pfleiderer, K., and Hillen, W. (2001). Identification of a stability determinant on the edge of the Tet repressor four-helix bundle dimerization motif. Biochemistry *40*, 3257–3263.

Schubert, P., Pfleiderer, K., and Hillen, W. (2004). Tet repressor residues indirectly recognizing anhydrotetracycline. Eur. J. Biochem. *271*, 2144–2152.

Schütz, M., Schöppe, J., Sedlák, E., Hillenbrand, M., Nagy-Davidescu, G., Ehrenmann, J., Klenk, C., Egloff, P., Kummer, L., and Plückthun, A. (2016). Directed evolution of G protein-coupled receptors in yeast for higher functional production in eukaryotic expression hosts. Sci. Rep. *6*, 21508.

Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N.G., Palmer, A.E., and Tsien, R.Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat. Biotechnol. *22*, 1567–1572.

Shaner, N.C., Lin, M.Z., McKeown, M.R., Steinbach, P.A., Hazelwood, K.L., Davidson, M.W., and Tsien, R.Y. (2008). Improving the photostability of bright monomeric orange and red fluorescent proteins. Nat. Methods *5*, 545–551.

Shapiro, M.G., Westmeyer, G.G., Romero, P.A., Szablowski, J.O., Küster, B., Shah, A., Otey, C.R., Langer, R., Arnold, F.H., and Jasanoff, A. (2010). Directed evolution of a magnetic resonance imaging contrast agent for noninvasive imaging of dopamine. Nat. Biotechnol. *28*, 264–270.

Smith, L.D., and Bertrand, K.P. (1988). Mutations in the Tn10 tet repressor that interfere with induction. Location of the tetracycline-binding domain. J. Mol. Biol. *203*, 949–959.

Smith, T.F., and Waterman, M.S. (1981). Identification of common molecular subsequences. J. Mol. Biol. *147*, 195–197.

Staus, D.P., Wingler, L.M., Strachan, R.T., Rasmussen, S.G.F., Pardon, E., Ahn, S., Steyaert, J., Kobilka, B.K., and Lefkowitz, R.J. (2014). Regulation of β2-adrenergic receptor function by conformationally selective single-domain intrabodies. Mol. Pharmacol. *85*, 472–481.

Staus, D.P., Strachan, R.T., Manglik, A., Pani, B., Kahsai, A.W., Kim, T.H., Wingler, L.M., Ahn, S., Chatterjee, A., Masoudi, A., et al. (2016). Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. Nature *535*, 448–452.

Strachan, R.T., Sun, J.P., Rominger, D.H., Violin, J.D., Ahn, S., Rojas Bie Thomsen, A., Zhu, X., Kleist, A., Costa, T., and Lefkowitz, R.J. (2014). Divergent transducer-specific molecular efficacies generate biased agonism at a G protein-coupled receptor (GPCR). J. Biol. Chem. *289*, 14211–14224.

Strauss, J.H., and Strauss, E.G. (1994). The alphaviruses: gene expression, replication, and evolution. Microbiol. Rev. *58*, 491–562.

Strauss, E.G., Rice, C.M., and Strauss, J.H. (1984). Complete nucleotide sequence of the genomic RNA of Sindbis virus. Virology *133*, 92–110.

Subramaniam, M., Schmidt, L.J., Crutchfield, C.E., 3rd, and Getz, M.J. (1989). Negative regulation of serum-responsive enhancer elements. Nature *340*, 64–66.

Tang, J., Jose, J., Chipman, P., Zhang, W., Kuhn, R.J., and Baker, T.S. (2011). Molecular links between the E2 envelope glycoprotein and nucleocapsid core in Sindbis virus. J. Mol. Biol. *414*, 442–459.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192.

Toprak, E., Veres, A., Michel, J.-B., Chait, R., Hartl, D.L., and Kishony, R. (2011). Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. Nat. Genet. *44*, 101–105.

Tracewell, C.A., and Arnold, F.H. (2009). Directed enzyme evolution: climbing fitness peaks one amino acid at a time. Curr. Opin. Chem. Biol. *13*, 3–9.

Urlinger, S., Baron, U., Thellmann, M., Hasan, M.T., Bujard, H., and Hillen, W. (2000). Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. Proc. Natl. Acad. Sci. USA *97*, 7963–7968.

van der Kant, R., and Vriend, G. (2014). Alpha-bulges in G protein-coupled receptors. Int. J. Mol. Sci. *15*, 7841–7864.

Varnaitė, R., and MacNeill, S.A. (2016). Meet the neighbors: Mapping local protein interactomes by proximity-dependent labeling with BioID. Proteomics *16*, 2503–2518.

Volkers, G., and Hinrichs, W. (2012). 4AC0: TETR(B) in complex with minocycline and magnesium. https://www.rcsb.org/structure/4AC0.

Wacker, D., Wang, C., Katritch, V., Han, G.W., Huang, X.-P., Vardy, E., McCorvy, J.D., Jiang, Y., Chu, M., Siu, F.Y., et al. (2013). Structural features for functional selectivity at serotonin receptors. Science *340*, 615–619.

Wacker, D., Stevens, R.C., and Roth, B.L. (2017a). How Ligands Illuminate GPCR Molecular Pharmacology. Cell *170*, 414–427.

Wacker, D., Wang, S., McCorvy, J.D., Betz, R.M., Venkatakrishnan, A.J., Levit, A., Lansu, K., Schools, Z.L., Che, T., Nichols, D.E., et al. (2017b). Crystal Structure of an LSD-Bound Human Serotonin Receptor. Cell *168*, 377–389.

Wallace, A.R. (1855). XVIII.—On the law which has regulated the introduction of new species. Ann. Mag. Nat. Hist. *16*, 184–196.

Wallace, A.R. (1871). Contributions to the Theory of Natural Selection (Macmillan).

Wang, S., Wacker, D., Levit, A., Che, T., Betz, R.M., McCorvy, J.D., Venkatakrishnan, A.J., Huang, X.-P., Dror, R.O., Shoichet, B.K., and Roth, B.L. (2017). D_4 dopamine receptor high-resolution structures enable the discovery of selective agonists. Science *358*, 381–386.

Wissmann, A., Baumeister, R., Müller, G., Hecht, B., Helbl, V., Pfleiderer, K., and Hillen, W. (1991). Amino acids determining operator binding specificity in the helix-turn-helix motif of Tn10 Tet repressor. EMBO J. *10*, 4145–4152.

Wright, S.I., Bi, I.V., Schroeder, S.G., Yamasaki, M., Doebley, J.F., McMullen, M.D., and Gaut, B.S. (2005). The effects of artificial selection on the maize genome. Science *308*, 1310–1314.

Xiong, C., Levis, R., Shen, P., Schlesinger, S., Rice, C.M., and Huang, H.V. (1989). Sindbis virus: an efficient, broad host range vector for gene expression in animal cells. Science *243*, 1188–1191.

Xu, L., Aha, P., Gu, K., Kuimelis, R.G., Kurz, M., Lam, T., Lim, A.C., Liu, H., Lohse, P.A., Sun, L., et al. (2002). Directed evolution of high-affinity antibody mimics using mRNA display. Chem. Biol. *9*, 933–942.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Donkey polyclonal anti-rabbit IgG HRP	Jackson ImmunoResearch	711-035-152; RRID: AB_10015282
Horse polyclonal anti-mouse IgG HRP	Cell Signaling	7076S; RRID: AB_330924
Mouse monoclonal anti-FLAG M2-Peroxidase (HRP)	Sigma-Aldrich	A8592; RRID: AB_439702
Mouse polyclonal anti-FLAG-M2	Sigma-Aldrich	F1804; RRID: AB_262044
Rabbit polyclonal anti-GFP	Novus Biologicals	NB600-308; RRID: AB_10003058
Bacterial and Virus Strains		
One Shot Stbl3 Chemically Competent *E. coli*	Thermofisher	C737303
Chemicals, Peptides, and Recombinant Proteins		
(+)-Morphine base	NIDA Drug Supply	9300-012
[^{3}H]Ketanserine	PerkinElmer	NET791025
[^{3}H]-myo-inositol	PerkinElmer	NET1177001MC
[N-Methyl-^{3}H]-Lysergic Acid Diethylamide ([3H]-LSD)	PerkinElmer	NET638250UC
1-(4-iodo-2,5-dimethoxyphenyl)propan-2-amine HCl (DOI)	Tocris	2643
1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC)	Avanti Polar Lipids	850355C
10x Hank's Buffered Saline Solution (HBSS)	Life Technologies	14065-056
3x FLAG Peptide	Sigma-Aldrich	F4799
4% paraformaldehyde	Fisher	AAJ19943K2
4-(2-Aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF)	Sigma-Aldrich	A8456
Aprotinin	Sigma-Aldrich	A1153
Bovine Serum Albumin (BSA), fatty-acid free	Akron Biotech	AK8909
Bright-Glo	Promega	E2620
Carbenicillin	Gold Bio	C-103-25
cholesteryl hemisuccinate (CHS)	Sigma-Aldrich	C6512
Coelenterazine h	Promega	S2011
Decyl Maltose Neopentyl Glycol (DMNG)	Anatrace	NG322
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich	276855
Dopamine HCl	Tocris	3584
Doxycycline HCl (DOX)	Sigma-Aldrich	D3447
D-Phosphate Buffered Saline (D-PBS), Ca^{2+}/Mg^{2+} free	ThermoFisher	14190144
E-64	Sigma-Aldrich	E3132
GppNHp	Abcam	ab146659
Hygromycin B	KSE	98-923
imidazole	Sigma-Aldrich	I5513
iodoacetamide	Sigma-Aldrich	I6125
Leupeptin	Sigma-Aldrich	L2884
Lipid A	Sigma-Aldrich	L5399
lysergic acid diethylamide	synthetic	Wacker et al., 2017b
Methiothepin mesylate salt	Sigma-Aldrich	M149
n-dodecyl-β-D-maltopyranoside (DDM)	Anatrace	D319
n-Octyl-β-D-Glucopyranoside	Anatrace	O311
nuclease-free water (H_2O)	NEB	B1500
Penicillin/Streptomycin	ThermoFisher	15140122
Phenoxybenzamine HCl	Sigma-Aldrich	B019

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
poly-L-lysine	Sigma-Aldrich	P2636
Puromycin	Gemini	400-128P
Serotonin HCl (5-HT)	Sigma-Aldrich	H9523
Trypsin	VWR	45000-660
Tryptose Phosphate Broth (TPB)	ThermoFisher	T8159
UltraPure Phenol:Chloroform:Isoamyl Alcohol (25:25:1, w/v)	ThermoFisher	15593031
Versene	ThermoFisher	15040066
Critical Commercial Assays		
Agencourt AMPure magnetic beads	Beckman Coulter	A63881
Bac-to-Bac Baculovirus Expression System	Invitrogen	10359016
Bio-Beads SM2 Resin	BioRad	1523920
Bioruptor Pico sonication device	Diagenode	B01060010
Cellfectin II reagent	ThermoFisher	10362100
Clarity Western ECL Substrate	BioRad	1708370
KAPA Hyperprep kit	Roche	KK8500
LDS Gel Loading Buffer	ThermoFisher	NP0007
MagMax Viral RNA Isolation Kit	ThermoFisher	AM1939
Magnetic FLAG-M2 Beads	Sigma-Aldrich	M8823
mMessage mMachine *in vitro* RNA transcription kit	ThermoFisher	AM1340
NEBuilder HiFi DNA Assembly Master Mix	NEB	E2621
Neon Transfection Kit	ThermoFisher	MPK10096
Neon Transfection System	ThermoFisher	MPK5000
NuPage 4-12% Bis-Tris Protein Gels	ThermoFisher	NP0322
PD MiniTrap G-25 Columns	GE Life Sciences	28918007
PowerPrep HP Plasmid Maxiprep system	Origene	NP100010
Precision Plus Protein Dual Color Standard	BioRad	1610374
PreScission Protease	Genscript	Z02799
PrimeSTAR Max DNA polymerase	Takara Bio	R045
QIAprep spin miniprep kits	QIAGEN	27104
QIAquick Gel Extraction Kit	QIAGEN	28115
RNA binding beads (YSI)	PerkinElmer	RPNQ0013
Sartorius Vivaspin 20 Centrifugal Concentrators 100kDa MWCO	Cole-Parmer	VS2002
Sartorius Vivaspin 500 Centrifugal Concentrators 100kDa MWCO	Cole-Parmer	VS0141
SuperScript IV One-Step RT-PCR System	Invitrogen	12594025
SuperSignal ELISA Pico Chemiluminescent Substrate	Sigma-Aldrich	37069
TALON Metal Affinity Resin	Takara	635653
TaqMan Fast Virus 1-Step Master Mix	ThermoFisher	4444432
TransIT-2020 Transfection Reagent	VWR	MIR5400
Deposited Data		
Raw and analyzed data	This paper	GEO: GSE123269
Experimental Models: Cell Lines		
Hamster: BHK21	ATCC	CCL-10
Human: HEK293T	ATCC	CRL-3216
Human: HEK-$G\alpha_{11\Delta/s\Delta/olf\Delta/12\Delta/13\Delta}$	gift	Asuka Inoue, Tohoku University
Human: HEK-P	gift	Asuka Inoue, Tohoku University
Human: HTLA	gift	Richard Axel, Columbia University
Insect: Sf9 cells, suspension in ESF 921 1M cells/mL	Expression Systems	94-001S

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oligonucleotides		
See Table S6	N/A	N/A
Recombinant DNA		
See Table S6	N/A	N/A
Other		
0.45uM media filters	EMD Millipore	SCGP00525
10 μg/mL Carbenicillin supplemented LB agar plates	Teknova	L1010
24-well plates	Sigma-Aldrich	CLS3527
384-well white plates	Black Dog	781098
5% normal goat serum	Vector Laboratories	S-1000
96-well white plates	Black Dog	655098
BamHI-HF	NEB	R3136
Canted neck culture flasks	Sigma-Aldrich	CLS430641U
ClaI	NEB	R0197
CutSmart Buffer	NEB	B7204
Dulbecco's Modified Eagle Media (DMEM)	VWR	45000
Fetal Bovine Serum (FBS)	VWR	89510-186
Filtermat A, GF/C	PerkinElmer	1450-421
Immobilon PVDF membranes	Sigma-Aldrich	IPSN07852
LB Broth	ThermoFisher	10855001
MeltiLex Solid Scintillant, for Microbeta	PerkinElmer	1450-441
MEM-α with nucleosides	ThermoFisher	32571036
myo-inositol-free DMEM	Caisson Labs	DML13
NheI-HF	NEB	R3131
NotI-HF	NEB	R3189
PET-A, FLEX, 96-well clear sample plates	PerkinElmer	1450-401C
RNAsin	Promega	N2111
Sf-900 II SFM media	ThermoFisher	10902096
T7 DNA Ligase	NEB	M0318
XbaI	NEB	R0145

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Justin English (jgenglis@email.unc.edu).

METHOD DETAILS

Molecular Biology & Plasmid Construction

All standard plasmids were constructed via PCR amplification of the desired amplicons using PrimeSTAR Max DNA polymerase (Takara Bio, #R045) and primers (Table S6, Eton Biosciences). Ligation of backbones and amplicons was performed using NEBuilder HiFi DNA Assembly Master Mix (NEB, #E2621). Clones were isolated by transformation of ligated DNA to One Shot Stbl3 Chemically Competent E. Coli (Thermofisher, #C737303) and selected on LB agar plates supplemented with 100μg/mL carbenicillin (Teknova L1010). Individual colonies were grown shaking at 37°C overnight in liquid LB broth (ThermoFisher, 10855001) supplemented with 100 μg/mL carbenicillin (GoldBio, C-103-25). Plasmids were purified with QIAprep Spin Miniprep Kits (QIAGEN, #27104) or PowerPrep HP Plasmid Maxiprep Systems (OriGene, #NP100010), dependent on downstream application. For construction of viral-sequence containing vectors see specific methods sections. All constructs were designed and confirmed via Sanger sequencing alignment (Eton Biosciences) using Benchling (https://Benchling.com). The list of plasmids used in this study can be found on Table S6, those necessary to perform VEGAS directed evolution have been made available at Addgene.org.

General Cell Culture

All cells were grown in a humidified 37°C incubator with 5% CO_2 using media supplemented with 100 I.U./mL penicillin and 100 μg/mL streptomycin (ThermoFisher, #15140122), unless otherwise indicated. The human cell lines HEK293T (ATCC, #CRL-3216), HTLA (kindly provided by Dr. Richard Axel, Columbia Univ.), HEK-G$\alpha_{q/11/s\Delta}$, and HEK parental (HEKp, both kindly provided by Asuka Inoue, Tohoku Univ.) were maintained in DMEM (VWR, #45000) containing 10% fetal bovine serum (FBS, VWR, #89510-186). HTLA cells were additionally supplemented with 5 mg/mL Puromycin (Gemini, 400-128P) and 100 mg/mL Hygromycin B (KSE, 98-923). The hamster cell line BHK21 (ATCC, #CCL-10) was maintained in MEM-α with nucleosides (ThermoFisher, #32571036) containing 5% FBS, 100 I.U./mL penicillin, 100ug/mL streptomycin, and 10% Tryptose Phosphate Broth (ThermoFisher, #T8159).

Sindbis Virus Production

Construct Design

pTSin (pTransgenicSindbis, Table S6), a transgene-free variant of pSinRep5 (kindly provided by Mark Heise, UNC Chapel Hill) (see Bredenbeek et al., 1993), was used as the base plasmid for the construction of all transgenic Sindbis virus packaging experiments. Each transgene of interest was subcloned to pTSin via PCR amplification adding 5′-NotI & 3′-ClaI cut sites or double NotI sites. The amplicon and pTSin were both digested overnight at 37°C with NotI-HF (NEB, #R3189) and ClaI (NEB, #R0197). The digested DNA fragments were purified using the QIAquick Gel Extraction Kit (QIAGEN, #28115) and ligated with T7 DNA ligase (NEB, #M0318). Clones were selected and confirmed as described previously.

Preparation of mRNA

Our preparation of RNA for Sindbis virus packaging is modified from previous methods (Bredenbeek et al., 1993). pTSin plasmids containing transgene, as well as pSinHelper and pSinCapsid (both kindly provided by Mark Heise, UNC Chapel Hill), must be linearized before converting them to mRNA for viral packaging. Linearization was accomplished by mixing 3μg of plasmid with 2 μL XbaI (NEB, #R0145), 8 μL CutSmart Buffer (NEB, #B7204), raised to 80 μL with nuclease-free water (hereafter H_2O, NEB, #B1500) and incubated at 37°C for 24 hr. The linearized DNA was extracted by adding 20 μL H_2O to the digestion mix, followed by 100 μL UltraPure Phenol:Chloroform:Isoamyl Alcohol (25:24:1 w/v, ThermoFisher, #15593031). The mixture was vortexed for 15 s, centrifuged at 13,000 x g for 5 min, and the top aqueous phase was transferred to a new tube. The extraction was then repeated. The extracted DNA was then treated with 10μL 3M sodium acetate (pH 5.2), mixed by inversion, and precipitated with 220μL 100% ethanol. The precipitant was kept at −80°C for a minimum of 20 min (or held indefinitely for future use). The DNA:ethanol mixture was then centrifuged at 13,000 x g for 10 min at 4°C. The liquid was gently aspirated from the pellet and the retained pellet was washed with 300μL 75% ice-cold ethanol prior to centrifugation at 13,000 x g for 3 min at 4°C. The wash was aspirated and the pellet was resuspended in 10μL RNAsin (Promega, #N2111) treated H_2O (1:10, v:v). DNA was then immediately *in vitro* transcribed to mRNA.

Conversion of linearized pTSin transgene plasmids to mRNA was performed using the Ambion mMessage mMachine *in vitro* mRNA transcription kit (Fisher, #AM1340). Reagents were added in the following order at room temperature: 1μL reaction buffer, 5μL CAP/dNTP mix, 1μL GTP, 10μL linearized pTSin vector, and 0.7μL enzyme. The reaction was incubated at 37°C for 1 hr and then used immediately or stored at −80°C until needed.

Packaging to Sindbis Viral Particles

To package pTSin in to viral particles pTSin, pSinHelper, and pSinCapsid mRNA must be electroporated in to BHK21 cells. One day prior to electroporation BHK21 cells were split to canted-neck culture flasks seeded at $5x10^5$ cells/dish. One flask is prepared per desired electroporation. On the day of electroporation cells were washed with 10mL Ca^{2+}/Mg^{2+} free DPBS (ThermoFisher, #14190144), disassociated with trypsin (VWR, #45000-660), rinsed with 7mL ice-cold DPBS, and held on ice. Cells were centrifuged at 500 x g for 5 min at 4°C. Cell pellet was resuspended in 10mL ice-cold DPBS and centrifuged at 500 x g for 5 min 4°C. The DPBS wash was repeated 2 additional times. Before the final wash, cells were counted by hemocytometer. The washed cell pellet was then resuspended in ice cold Neon E1 electroporation buffer (ThermoFisher, #MPK10096) to $5x10^6$ cells/mL and aliquoted 115μL for each transgene to be packaged. To each cell aliquot 10μL of pTSin, pSinHelper, and pSinCapsid mRNA were added, pipette mixed, and electroporated with a Neon Transfection System (ThermoFisher, #MPK5000) set to 1400V, 10 width, 3 pulses. Electroporated cells were incubated at room temperature for 10 minutes and then applied to 25mL pre-incubated serum-free MEM-α media (ThermoFisher, #32571036) supplemented with 10% tryptose phosphate broth and pen/strep in canted neck culture flasks (Sigma-Aldrich, #CLS430641U). After 24 hr media was decanted, 0.45μm filtered (EMD Millipore, #SCGP00525), and stored at 4°C for no longer than 1 month for use in downstream applications.

Sustained Passage of Sindbis for Directed Evolution and Transgene Isolation

Construct Design

For sustained passage and selection by directed evolution the Sindbis virus structural genome (pG100, SSG; 7662-11718 of full length Sindbis virus genome GenBank: MF459683.1, kindly provided by Mark Heise, UNC Chapel Hill) was subcloned using PCR/HiFi assembly as described above to mammalian expression plasmids possessing either CMV (pCDNA3.1), TETO7 (pTRE3G-BI-ZsGreen1), or SRE (pGL4.33) promoters (see Table S6 for source and mapping information).

Cell Culture and Selection

One day prior to selection $5x10^5$ BHK21 cells were plated in canted neck culture flasks (Sigma-Aldrich, #CLS430641U). The following day 10μg of the appropriate SSG plasmid was transfected in to the cells using the TransIT-2020 Transfection Reagent (VWR, MIR5400). The transfected cells were incubated for six hr before removing all media, rinsing the cells with PBS, and then applying the appropriate titer of virus diluted to 2mL in serum-free MEM-α media supplemented with 10% TPB, pen/strep, and either doxycycline HCl (Sigma-Aldrich, #D3447), (+)-morphine (NIDA Drug Supply, 9300-012), or vehicle. For GPR68-targeted selection, media was additionally supplemented with 2mg/L sodium bicarbonate and equilibrated to pH 8.0 in 5% CO_2 overnight prior to use. After incubating 1 hr with intermittent rocking 23mL of additional supplemented media was added and the cells were incubated for 24 hr. Following culturing and viral propagation the cell culture media is decanted, 0.45μm filtered, and stored at 4°C for no longer than 1 month.

Transgene Isolation

Using the MagMax Viral RNA Isolation Kit (ThermoFisher, AM1939) a 1mL aliquot of the viral stock was processed to collect a viral RNA sample. The purified RNA was maintained, precipitated in 100% EtOH at −80°C, until use in downstream applications unless noted otherwise. Isolating only translated transgenes from the viral pool is critical for the success of VEGAS. To isolate positively selected transgene sequences we used a forward primer annealing to the 26S promoter (26S-F, 5′-atctctacggtggtcctaaatagt-3′) alongside 8 reverse primers (pooled as "SinRev") annealing to the conserved RNA structural components of the viral 3′ UTR (See Table S6). Production of a cDNA library from the Sindbis virus RNA genome with SinRev and subsequent PCR amplification with 26S-F & SinRev were performed in series using the SuperScript IV One-Step RT-PCR System (Invitrogen, #12594025). Amplicons and pCDNA3.1 were digested with NheI-HF (NEB, #R3131) and BamHI-HF (NEB, #R3136), gel purified, and ligated together. Colonies were selected, cultured, plasmids were purified, and identified by sequencing as described above. Any clone appearing ≥ 2 times out of 12 clones was selected for functional screening.

RNA Deep Sequencing

Sample Processing

Samples were prepared as follows. For EGFP mutation analysis, the pTSin-EGFP construct was independently packaged via NEON electroporation into three separate cell populations as described above, using fresh mRNA prepared on the day. During packaging, viral samples were decanted and 0.45 μM filtered at 3, 6, 12, 24, and 36 hr from each of the three independently grown cell flasks. At each collection time point the cells were washed with PBS and returned to incubate with fresh media. Viral RNA for all time points and replicates was collected simultaneously using the MagMax Viral RNA Isolation Kit (ThermoFisher, AM1939) from a 1mL aliquot of each viral stock. For nanobody library analysis, the pTSin-Nb library mRNA used for subsequent directed evolution experiments was purified and used for sequencing.

The purified RNA samples, synthesized mRNA, and pTSin-EGFP vector template were immediately amplified as described above with the SuperScript IV One-Step RT-PCR System (Invitrogen, #12594025) using 26S-F and the SinRev primer pool.

For mutation rate analysis the EGFP transgene amplicons were processed using a KAPA HyperPlus kit (KAPABiosystems, KK8512) with a 30' digestion period followed by standard size selection with KAPA Pure Beads (KAPABiosystems, KK8001). The resulting fragments were barcoded with the Illumina compatible SeqCAP Adapater Kit A & B (KAPABiosystems, 7141530001 and 07141548001). Samples were normalized, pooled, and processed using a NextSeq NSQ 500/550 Hi Output KT v2.5 300 cycle kit (Illumina, 200249808) calibrated to acquire dual-indexed, 2 × 150 bp reads, yielding an average of 4.3×10^8 reads per sample.

For nanobody library analysis the amplicons were > 100bp size-selected using Agencourt AMPure magnetic beads (Beckman Coulter, #A63881). The purified DNA was diluted to 10ng/μL in 10mM Tris, 1mM EDTA, pH 7.5 sonication buffer and sheared to 150bp fragments on a Bioruptor Pico sonication device (Diagenode, #B01060010) using a 30"/30" cycle program for 30 cycles at 4°C. The sheared DNA was then used for library construction using a KAPA Hyperprep Kit (Roche, #KK8500) and Illumina TruSeq indexed adapters (IDT, HPLC purified) following the manufacturer's recommended protocol. Samples were normalized and processed using 12-plex, single-indexed, 1 × 150 bp Illumina NextSeq500 sequencing.

The EGFP experiment sequence data was initially processed for exact matching paired-end reads > 25bp in length. The longest perfect match from each read pair was isolated and the remaining non-matching sequences were omitted from further analysis. The truncated reads were aligned to sense and anti-sense EGFP and Sindbis genome reference sequences using a Smith-Waterman algorithm (Smith and Waterman, 1981) on the UNC Longleaf Linux Cluster (UNC ITS). Alignment score weighting was m3, x1, o5, e1 for correct match (m), mismatch penalty (x), insert/deletion penalty (o), and gap extension penalty (e). Reads were associated with their high scoring reference sequence. The sequences aligning to anti-sense EGFP were reverse-complemented. All EGFP aligning reads were then re-aligned to the EGFP reference sequence and a sequence where EGFP was excised (TGDEL). Sequences scoring best for the sense EGFP reference were then used to retrieve individual base position, deletion, and insertion counts for the EGFP transgene. No score cut-offs were implemented.

The nanobody experiment sequence data was aligned to the cloned nanobody reference sequence using a Smith-Waterman algorithm (Smith and Waterman, 1981) on the UNC Longleaf Linux Cluster (UNC ITS). Alignment score weighting for all nanobodies was m1, x5, o5, e5 for correct match (m), mismatch penalty (x), insert/deletion penalty (o), and gap extension penalty (e). Sequence reads containing the 3′ barcode and vector backbone were trimmed and the final alignments were designated with varying score cut-offs as described in Figure 4A.

Data were analyzed on Graphpad Prism 8. The mutation frequency data per time point (Figure 1C) statistics were calculated using One-Way ANOVA with Geisser-Greenhouse correction and Sidak multiple comparisons correction. Multiplicity adjusted P values for each comparison are $p < 0.0001$ for all significant comparisons, using family-wise significance and confidence level set to 99.9%. The mutation type per base reads data (Figure 1D) was plotted as individual replicates. The values represent the number of total substitutions divided by the number of total reads for the mutated base per replicate and time point. Statistics were calculated within each base substitution group using Two-Way ANOVA and Tukey's multiple comparisons correction with one family per column (substitution type). A minimum multiplicity adjusted P value ($\alpha = 0.05$) was applied for each comparison. The mutations base^{-1}/time (Figure 1E) data was plotted with a linear regression model for mean ± SD of N = 715 measures with 99% confidence band displayed. Data in Figures S1B–S1D are residuals of the data displayed in Figure 1 and no statistical analysis was performed. For Figures S1E–S1G the number of mutations observed per total reads at the indicated read length are plotted with a linear regression model beginning at the observed linear range of read length 50. 95% confidence bands are displayed for each time point.

Quantification of Viral RNA via qRT-PCR

Quantification of Sindbis viral genomic RNA by TaqMan qRT-PCR was performed as described previously (Sane et al., 2012) with modification. We designed a probe and primer pair specific for the packaging signal sequence spanning g.138-a.269 of nsP1 in pTSin; Probe: 5′-/5HEX/ATTTTGGAC/ZEN/ATAGGCAGCGCACC/3IABkFQ/-3′, Forward: 5′-GTTCCTACCACAGCGACG-3′, Reverse: 5′-GGTACTGGTGCTCGGAAAAC-3′ (IDT) (see Table S6). Sindbis virus containing media was diluted across 3 × 10-fold serial dilutions, 4 replicates each, mixed directly with TaqMan® Fast Virus 1-Step Master Mix (ThermoFisher, #4444432) and cycled on a BioRad CFX96 Touch RT-PCR machine (BioRad). Serially diluted (10^{-12}-10^{-6} M) *in vitro* transcribed Sindbis virus RNA reference samples were present on each sample plate. Standard curves were calculated on CFX Manager (BioRad) and used to calculate and report genomes/mL for each sample.

tTA Reporter Assay

The wild-type and R7_G8 tetracycline transactivator (tTA) sequences were subcloned in pCDNA3.1. The reporter was constructed by subcloning luc2p from SRE-luc2p (Promega, pGL4.33) over the ZsGreen1 CDS of pTRE3G-BI-ZsGreen1 (Takara, 631339) to create pTRE3G-BI-luc2p. The day before transfection, HEK293T cells were split to yield approximately 9x10^6 cells/15-cm plate next day. The following day cells were transfected with 10μg DNA per 15-cm, 5μg of pTRE3G-BI-luc2p and either 5μg of WT or R7 constructs using TransIT-2020. Six hours after transfection media and transfection reagents were removed, cells were washed with PBS, dissociated using versene (ThermoFisher, 15040066), centrifuged and resuspended in DMEM supplemented with 1% dialyzed FBS. Transfected cells were then plated onto poly-L-lysine-coated 384-well white clear bottom cell culture plates at a density of 10,000 cells/well in a total of 20μl. Doxycycline HCl solutions were prepared in plating media at 2 x and added to cells (20 μL per well) for overnight incubation. After 20-22 hr overnight incubation, media and drug solutions were removed from plates and 20μl per well of 1:20 diluted Bright-Glo reagent (Promega, E2620) was added per well. The plate was incubated for 20 min at room temperature in the dark before being counted using a luminescence counter. Results (relative luminescence units) were plotted as a function of drug concentration and analyzed using ''log(inhibitor) versus response'' in GraphPad Prism 8.0.

Transcription Factor Reporter Primary Screen

The MRGPRX2, DRD2, 5-HT2$_A$, and GPR68 constructs originated from their respective PRESTO-TANGO plasmids (Kroeze et al., 2015), from which the C-terminal V2Tail-TEV-tTA sequence was removed by PCR mutagenesis. The transcription factor reporter constructs were purchased from Promega (see Table S6). Two days before transfection, HEK293T cells were split to yield approximately 9x10^6 cells/15-cm plate in 2 days. On the day of transfection each well of a poly-L-lysine-coated 384-well white clear bottom cell culture plate received 30ng of DNA,15ng of one transcription factor reporter and 15ng of either GPCR or pCDNA3.1, pre-incubated with TransIT-2020. One plate was prepared for each time point, with 8 wells of GPCR and 8 wells of empty vector for each reporter construct. The prepared cells were then washed with PBS, dissociated using trypsin, centrifuged, and resuspended in DMEM supplemented with 1% dialyzed FBS. Cells were seeded in the DNA pre-loaded plates at 10,000 cells/well to a final volume of 40 μL and incubated for 24 hr. The following day single drug concentrations and vehicle (DMSO) were prepared in drug buffer (1 x HBSS, 20 mM HEPES, 0.1% BSA, 0.01% ascorbic acid, pH 7.4) at 3 x and added to cells (20μl per well) at the appropriate time point for each plate. For each reporter 4 replicates of drug treated and 4 replicates of vehicle were added for each transfection condition at each time point. After incubation, media and drug solutions were removed from plates and 20μl per well of BrightGlo reagent (purchased from Promega, after 1:20 dilution) was added per well. The plate was incubated for 20 min at room temperature in the dark before being counted using a luminescence counter. Results (relative luminescence units) were plotted as a function of time in GraphPad Prism 8.0.

SRE Reporter Assay

The GPCR constructs from the transcription factor reporter primary screen were used in these assays. Nanobody constructs were direct clones from Sindbis virus transgene isolation to pCDNA3.1. The day before transfection, HEK293T cells were split to yield approximately 9x10^6 cells/15-cm plate next day. The following day cells were transfected with 15 μg DNA per 15-cm dishes; 5 μg of SRE-luc2P (Promega, pGL4.33), 5 μg of GPCR, and an appropriate ratio of Nb and empty vector to 5ug. The next day, media

and transfection reagents were removed, cells were washed with PBS, dissociated using versene, centrifuged and resuspended in DMEM supplemented with 0.1% dialyzed FBS. Transfected cells were then plated onto poly-L-lysine-coated 384-well white clear bottom cell culture plates at a density of 10,000 cells/well in a total of 40μl. The cells were incubated for 12 hr and then developed for untreated assays or received drug stimulation. Drug solutions were prepared in drug buffer (1 x HBSS, 20 mM HEPES, 0.1% BSA, 0.01% ascorbic acid, pH 7.4) at 3 x and added to cells (20μl per well) for overnight incubation. After 6 hr incubation, media and drug solutions were removed from plates and 20 mL per well of BrightGlo reagent (purchased from Promega, after 1:20 dilution) was added per well. The plate was incubated for 20 min at room temperature in the dark before being counted using a luminescence counter. Results (relative luminescence units) were plotted as a bar graphs or as a function of drug concentration and analyzed using ''log(agonist) versus response (three parameters)'' in GraphPad Prism 8.0.

TANGO β-arrestin recruitment assay

The MRGPRX2 Tango construct, which contains the TEV cleavage site and the tetracycline transactivator (tTA) fused to the C terminus of the receptor, were designed and assays were performed as previously described (Kroeze et al., 2015; Lansu et al., 2017). Each VEGAS-derived MRGPRX2 mutant was subcloned over the wild-type sequence via HiFi assembly. HTLA cells expressing TEV fused- β-Arrestin2 and a tetracycline transactivator-driven luciferase (kindly provided by Dr. Richard Axel, Columbia Univ.) were grown in HTLA media (10% FBS DMEM containing 5 mg/mL Puromycin and 100 mg/mL Hygromycin B). The day before transfection, HTLA cells were split to yield approximately $9x10^6$ cells/15-cm plate next day. The following day cells were transfected with 10μg DNA per 15-cm with MRGPRX2 Tango, or one of the VEGAS derived mutants, using TransIT-2020. The next day, media and transfection reagents were removed, cells were washed with PBS, dissociated using versene, centrifuged and resuspended in DMEM supplemented with 1% dialyzed FBS. Transfected cells were then plated onto poly-L-lysine-coated 384-well white clear bottom cell culture plates at a density of 10,000 cells/well in a total of 40μl. The cells were incubated for 12 hr before receiving drug stimulation to allow for recovery and adherence to the plate. Drug solutions were prepared in drug buffer (1 x HBSS, 20 mM HEPES, 0.1% BSA, 0.01% ascorbic acid, pH 7.4) at 3 x and added to cells (20μl per well) for overnight incubation. Drug solutions used for the Tango assay were exactly the same as used for the SRE assay, which was conducted in parallel to the Tango assay. After 6 hr incubation, media and drug solutions were removed from plates and 20 mL per well of BrightGlo reagent (purchased from Promega, after 1:20 dilution) was added per well. The plate was incubated for 20 min at room temperature in the dark before being counted using a luminescence counter. Results (relative luminescence units) were plotted as a function of drug concentration, normalized to % wild-type stimulation, and analyzed using ''log(agonist) versus response (three parameters)'' in GraphPad Prism 8.0.

Calcium flux assay

Assays were designed and performed as previously described (Wacker et al., 2017b), using the same 5-HT2AR stable cell line created with the Flp-In 293 T-Rex Tetracycline inducible system (Invitrogen). The day before transfection, 5-HT2AR cells were split to yield approximately $9x10^6$ cells/15-cm plate next day. The following day cells were transfected with 8 μg DNA per 15-cm with adjusted ratios of empty vector (pCDNA3.1) or pCDNA3.1-VGS-Nb2 using TransIT-2020. Six hours after transfection, 1 μM final concentration of doxycycline HCl was applied to induce receptor expression. The next day, media and transfection reagents were removed, cells were washed with PBS, dissociated using versene, centrifuged and resuspended in DMEM supplemented with 0.1% dialyzed FBS. cells were seeded in 384-well poly-L-lysine plates at a density of 10,000 cells/well at least 16-24 hr before the calcium flux assay. On the day of the assay, the cells were washed in FLIPR buffer (1 x HBSS, 2.5 mM probenecid, and 20 mM HEPES, pH 7.4), pre-treated with the 1 μM alkylating agent phenoxybenzamine (Sigma-Aldrich, B019) for 30 min, washed again in FLIPR buffer and then incubated with 20μL/well Fluo-4 Direct dye (ThermoFisher, F10471) reconstituted in FLIPR buffer for 1 hr at 37°C. After dye loading, cells were placed in a FLIPRTETRA fluorescence imaging plate reader (Molecular Dynamics). Drug dilutions were prepared at 3 x final concentration in drug buffer (1 x HBSS, 20 mM HEPES, 0.1% BSA, 0.01% ascorbic acid, pH 7.4), aliquoted into 384-well plates. and placed in the FLIPRTETRA for drug stimulation. The fluidics module and plate reader of the FLIPRTETRA were programmed to read baseline fluorescence for 10 s (1 read/s), then 10μl of drug/well was added and read for 5 min (1 read/s). Fluorescence in each well was normalized to the average of the first 10 reads (i.e., baseline fluorescence). Then, the maximum-fold increase, which occurred within the first 60 s after drug addition, was determined and fold over baseline was plotted as a function of drug concentration. Data were analyzed using ''log(agonist) versus response (three parameters)'' in Graphpad Prism 8.0.

Bioluminescence resonance energy transfer association assay

Assays were designed and performed as previously described (Che et al., 2018), with additions. The day before transfection, HEK293T cells were split to yield approximately $9x10^6$ cells/15-cm plate next day. The following day cells were transfected with 12 μg DNA per 15-cm dishes; 2 μg of GPCR-RLuc DNA held constant and an adjusted ratio of empty vector (pCDNA3.1) or pCDNA3.1-mVenus-Nb using TransIT-2020. The next day, media and transfection reagents were removed, cells were washed with PBS, dissociated using versene, centrifuged and resuspended in DMEM supplemented with 1% dialyzed FBS. Transfected cells were plated in poly-lysine coated 96-well white clear bottom cell culture plates in plating media (DMEM + 1% dialyzed FBS) at a density of 40-50,000 cells in 200μl per well and incubated overnight. The next day, media was aspirated and cells were washed twice with 60μL of drug buffer (1 x HBSS, 20 mM HEPES, 0.1% BSA, 0.01% ascorbic acid, pH 7.4), then 60μL of drug buffer was added per well,

followed by 30μL of drug (3X) per well, and finally 10μL of the RLuc substrate, coelenterazine h (Promega, S2011, 5 mM final concentration). Plates were incubated for 5 min to allow for substrate diffusion, and then read for both luminescence at 485 nm and fluorescent eYFP emission at 530 nm for 1 s per well using a Mithras LB940 multimode microplate reader. The ratio of eYFP/RLuc was calculated per well and the net bioluminescence resonance energy transfer (BRET) ratio was calculated by subtracting the eYFP/RLuc per well from the eYFP/RLuc ratio in wells without mVenus-Nb present. The net BRET ratio was plotted as a function of nanobody concentration using Graphpad Prism 8 (Graphpad Software Inc., San Diego, CA).

Phosphoinositide hydrolysis assay

Phosphoinositide (PI) hydrolysis assays measuring inositol phosphates (IP) were performed using the scintillation proximity assay (Bourdon et al., 2006; Huang et al., 2009). The MRGPRX2 WT construct was developed from the TANGO system as described above and mutant constructs were directly subcloned as viral transgenes to pCDNA3.1. The day before transfection, HEK293T cells were split to yield approximately $9x10^6$ cells/15-cm plate next day. The following day cells were transfected with 5 μg DNA per 15-cm dish with TransIT-2020. On the day before the assay, transfected cells were seeded into 96-well poly-lysine coated plates at a density of 40-50,000 cells/well in 100μL inositol-free DMEM (Caisson Labs, DML13) containing 1% dialyzed FBS. After 6 hr, an additional 100μL of label media was added containing 1μCi/well (final concentration) of [^{3}H]-myo-inositol (PerkinElmer, NET1177001MC) in inositol-free DMEM containing 1% dialyzed FBS and plates were incubated overnight for 16-18 hr. The next day, label media was removed and cells were washed twice with 60μL of drug buffer (1 x HBSS, 20 mM HEPES, 0.1% BSA, 0.01% ascorbic acid, pH 7.4), then 60μL of drug buffer was added per well. Afterward, 30μL of drug (3X) was added per well and incubated at 37°C for 1 hr. To capture IP accumulation, lithium chloride (10μL/well, 15 mM final concentration) was added 30 min before lysis. The assay was terminated by replacement of the incubation medium with 40μL of 50 mM formic acid. After overnight incubation at 4°C, 10μL of lysates were added to 96-well flexible, clear microplates (PerkinElmer, 1450-401C) containing 75μL of 0.2 mg/well RNA binding yttrium silicate beads (PerkinElmer), and incubated for 1 hr on a shaker. Afterward, plates were centrifuged at 300 x g for 1 min, and radioactivity was measured using a Wallac MicroBeta Trilux plate reader (PerkinElmer). Data were plotted as counts per minute (CPM) as a function of drug concentration, and analyzed using ''log(agonist) versus response (three parameters)'' in GraphPad Prism 8.0.

Surface expression enzyme-linked immunosorbent assay

To confirm cell surface expression of MRGPRX2 and its mutants, immunohistochemistry was done using cells plated on 384-well plates, as described earlier, at 10,000 cells/well. Cells were fixed with 20μl/well 4% paraformaldehyde (Fisher, #AAJ19943K2) for 10 minutes at room temperature. After fixation, cells were washed twice with 40μl/well PBS. Blocking was performed with 20μl/well 5% normal goat serum (Vector Laboratories, #S-1000) in PBS for 30 minutes at room temperature. After blocking, 20μl/well monoclonal ANTI-FLAG M2-Peroxidase (HRP) antibody (Sigma-Aldrich, A8592) diluted 1/10,000 in PBS was added and incubated for 1 hour at room temperature. This was followed by two washes with 80μl/well PBS. Then, 20μl/well SuperSignal enzyme-linked immunosorbent assay (ELISA) Pico Chemiluminescent Substrate (Sigma-Aldrich, #37069) was added, and luminescence was counted using a MicroBeta Trilux luminescence counter. Data were plotted as relative luminescent units (RLU) in GraphPad Prism 8.0.

Nanobody Production

Construct Preparation

5-HT2AR proteoliposomes for immunization were prepared using a receptor construct with truncated intracellular loop 3 (icl3) and C terminus. The final construct lacking residues 278-304 and 404-471 was cloned into a modified pFastBac vector introducing a haemagglutinin (HA) signal sequence followed by a FLAG tag at the N terminus, and a PreScission protease site followed by a 10 × His tag at the C terminus.

Expression and purification of 5-HT2AR

High-titer recombinant baculovirus (> 10^9 viral particles per ml) was generated using the Bac-to-Bac Baculovirus Expression System (Invitrogen, 10359016). Recombinant baculovirus was obtained by transfecting ~5μg of recombinant bacmid into $5x10^5$ settled *Spodoptera frugiperda* cells (Sf9, Expression Systems, 94-001S) in a 24 well plate (Sigma-Aldrich, CLS3527) using 3μl Cellfectin II Reagent (ThermoFisher, 10362100). After 5-12 hr, media was exchanged for 1 mL Sf-900 II SFM media (ThermoFisher, 10902096) and incubated for 4-6 days at 27°C. P0 viral stock with ~10^9 virus particles per ml was harvested as the supernatant and used to generate high-titer baculovirus stock by infection of 40-1000 mls of Sf9 cells and incubation for several days. Expression of 5-$HT2_A$ was carried out by infection of Sf9 cells at a cell density of 2-3 × 10^6 cells/ml in ESF921 media (Expression Systems) with P1 virus at a MOI (multiplicity of infection) of 3-5. After 48 hr of expression at 27°C in the presence of 10 μM Methiothepin (Sigma-Aldrich, M149), cells were harvested by centrifugation, washed in PBS, and stored at −80°C until use. Cells were disrupted by thawing frozen cell pellets in a hypotonic buffer (10 mM HEPES, pH 7.5, 10 mM $MgCl_2$, 20 mM KCl and protease inhibitors 500μM AEBSF, 1μM E-64, 1μM Leupeptin, 150 nM Aprotinin). Membranes were purified by repeated centrifugation in a high osmolarity buffer containing 1.0 M NaCl, 10 mM HEPES, pH 7.5, 10 mM $MgCl_2$, 20 mM KCl, to remove soluble and membrane associated proteins. Purified membranes were directly flash-frozen in liquid nitrogen and stored at −80°C.

Purified membranes were resuspended in buffer containing 10 mM HEPES, pH 7.5, 10 mM $MgCl_2$, 20 mM KCl, 150 mM NaCl, 20μM LSD (synthesized in house, see Wacker et al., 2017b), and protease inhibitors before incubating at room temperature for

1 h. After 30 min incubation in the presence of 2 mg/ml iodoacetamide (Sigma, I6125), membranes were solubilized in 10 mM HEPES, pH 7.5, 150 mM NaCl, 1% (w/v) n-dodecyl-β-D-maltopyranoside (DDM, Anatrace, D310), 0.2% (w/v) cholesteryl hemisuccinate (CHS, Sigma, C6512), 20μM LSD, and protease inhibitors for 2 h at 4°C. Unsolubilized material was removed by centrifugation at 150,000 × g for 30 min, and 15 mM imidazole (Sigma, I5513) was added to the supernatant. Proteins were bound to TALON metal affinity resin (Takara, 635653) overnight at 4°C using approximately 750μl resin for protein purified from 1 L of cells. The resin was then washed with 10 column volumes (cv) of Wash Buffer I (50 mM HEPES, pH 7.5, 800 mM NaCl, 0.1% (w/v) DDM, 0.02% (w/v) CHS, 20 mM imidazole, 10% (v/v) glycerol, and 20μM LSD, followed by 10 cv of Wash Buffer II (25 mM HEPES, pH 7.5, 150 mM NaCl, 0.05% (w/v) DDM, 0.01% (w/v) CHS, 10% (v/v) glycerol, and 20μM LSD). Proteins were eluted in 2.5 cv of Wash Buffer II + 250 mM imidazole, concentrated in a 100 kDa molecular weight cut-off Vivaspin 20 concentrator (Cole-Parmer, VS2002) to 500μl, and imidazole was removed by desalting the protein over PD MiniTrap G-25 columns (GE Life Sciences, 28918007). The C-terminal 10 × His-tag was removed by addition of His-tagged PreScission protease (GenScript) and incubation overnight at 4°C. Protease, cleaved His-tag and uncleaved protein were removed by passing the suspension through equilibrated TALON IMAC resin (Clontech) and collecting the flow-through. 5-$HT2_A$ /LSD complexes were then concentrated to ~8.5 mg/ml with a 100 kDa molecular weight cut-off Vivaspin 500 centrifuge concentrator (Cole-Parmer, VS0141). Protein purity and monodispersity were tested by analytical size-exclusion chromatography.

100 μL of 5-$HT2_A$ concentrated to 8.5 mg/ml were mixed with 300 μL of 4 mg/ml 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC, Avanti Polar Lipids, 850355C) and cholesteryl hemi succinate (CHS) at a ratio of 9:1 (w/w) and 100 μL 5 mg/ml Lipid A (Sigma, L5399) both in 20 mM HEPES, pH 7.5, 100 mM NaCl, 1% (w/v) n-Octyl-β-D-Glucopyranoside (OG, Anatrace, O311). Lipids, detergents, and protein were allowed to equilibrate on ice for 90 min, before rapid dilution with 1 mL of 20 mM HEPES, pH 7.5, 100 mM NaCl to bring OG below its critical micelle concentration. To remove detergent and form liposomes the sample was dialyzed against 20 mM HEPES, pH 7.5, 100 mM NaCl for 24 hr, and subsequently incubated with Bio-Beads SM2 resin (BioRad, 1523920) for 4 hr.

Llama immunization was done by Capralogics using a first injection of 200 μg, followed by 5 additional injections of 100 μg of 5-$HT2_A$ proteoliposomes.

Library Preparation

A nanobody library was generated as described previously (Pardon et al., 2014). In brief, immunized llama blood was taken to isolate peripheral blood lymphocytes. RNA was purified from these lymphocytes and reverse transcribed by PCR to obtain cDNA. The resulting library was cloned into pTSin plasmid to a complexity of ~1 × 10^7 colonies.

Nanobody Purification

VGS-Nb2 was subcloned from its original pCDNA3.1 destination vector after VEGAS directed evolution to pMESy4 (kindly provided by Jan Steyaert, Vrije Universiteit, Brussels) and purified following steps 70-73 described in the previous protocol (Pardon et al., 2014). Nanobodies were concentrated, desalted (10 mM HEPES, 100 mM NaCl, and 10% Glycerol), and stored at 80°C for future use.

Saturation and competitive radioligand binding assays

Radioligand assays were performed in parallel utilizing the same membrane preparations, binding buffer (50 mM Tris, 10 mM $MgCl_2$, 0.1 mM EDTA, 0.1% BSA, 0.01% ascorbic acid, pH 7.4), and purified VGS-Nb2 nanobody. For saturation assays, 5-$HT2_A$-$G\alpha_q$ membrane (50 μL) treated with purified VGS-Nb2 (5μM final concentration) or vehicle were added to round-bottom 96-well plates. A range of [N-Methyl-^{3}H]-Lysergic Acid Diethylamide ([^{3}H]-LSD, PerkinElmer, #NET638250UC) concentrations up to 6.80nM was added (25 μL) along with either vehicle (25 μL) or unlabeled 10 mM 1-(4-iodo-2,5-dimethoxyphenyl)propan-2-amine HCl (25 μL, DOI HCl, Tocris, #2643) to determine non-specific binding. For competition assays, 5-$HT2_A$ and 5-$HT2_A$-$G\alpha_q$ membranes (50μL) treated with purified VGS-Nb2 (7.5μM final concentration) or vehicle were added to round-bottom 96-well plates. 5-$HT2_A$ membranes were additionally treated with 50mM GppNHp (Abcam, ab146659) to uncouple all Gα proteins. The antagonist radioligand [^{3}H]Ketanserine (PerkinElmer, #NET791025) was added (25μL) to 1nM final concentration alongside a range of unlabeled DOI concentrations (25μL). Plates were incubated for 1.5 hr at room temperature. Both saturation and competition binding plates were harvested immediately after incubation by vacuum filtration onto 0.3% polyethyleneimine pre-soaked 96-well filter mats (PerkinElmer, 1450-421) using a 96-well Filtermate harvester, followed by three washes of cold wash buffer (50 mM Tris, pH 7.4). Solid scintillant (PerkinElmerm 1450-441) was melted onto dried filters and radioactivity was counted using a Wallac Trilux MicroBeta counter (PerkinElmer). For analysis of saturation binding data non-specific counts were removed using "Remove Baseline and Column Math" and then analyzed using "One Site – Specific Binding" with asymmetrical 95% confidence intervals in Graphpad Prism 8.0. Statistics were performed as extra sum-of-squares F test (α, 0.05) for difference between best-fit values of each dataset. Competitive binding data was analyzed using the "One Site – Fit logIC50" model and compared for best fit versus a "Two sites – Fit logIC50" model in Graphpad Prism 8.0 using an extra sum-of-squares F test (α, 0.05). Each dataset was then normalized to the "Top" value of the best fit model and replotted with shared parameters. Within parameter comparisons of IC_{50} were performed with an extra sum-of-squares F test (α, 0.05).

Co-Immunoprecipitation Analysis

Co-Immunoprecipitation

The FLAG-5-$HT2_A$ and FLAG-DRD2 constructs were developed from their respective TANGO plasmids with the V2Tail-TEV-tTA C-terminal tag removed. The mVenus-VGS-Nb2 plasmid is the same described for use in BRET experiments. The day before

transfection, HEK293T and HEK $G\alpha_{q/11/s\Delta}$ cells were split to yield approximately $6x10^6$ cells/15-cm plate next day. The following day cells were transfected with 10 μg DNA per 15-cm; 8 μg GPCR construct and 2 μg nanobody using TransIT-2020. 6 hr after transfection media and transfection reagents were removed, cells were washed with PBS, and incubated in DMEM supplemented with 0.1% dialyzed FBS. 48 hours after transfection the cells were washed, lysed, and FLAG-GPCR was FLAG-immunoprecipitated as described previously (Staus et al., 2014). In brief, cells were scraped in to ice-cold lysis buffer [20 mM HEPES, 100mM NaCl, 0.5% decyl maltose-neopentyl glycol (DMNG, Anatrace, NG322)] and incubated with gentle rotation at 4°C for 1 hr. Insoluble material was separated by centrifugation at 14,000 x g for 15 minutes at 4°C. The soluble lysate was immunoprecipitated with magnetic FLAG-M2 beads (Sigma-Aldrich, M8823), washed with TBS-M (50mM Tris HCl, 150mM NaCl, 0.5% MNG, pH 7.4), and eluted with 3 μL of 5ug/μL 3x FLAG peptide (Sigma-Aldrich, #F4799) in TBS-M. Samples were then used immediately for downstream applications.

Western Blot

Co-immunoprecipitated fractions taken from solubilized cell lysate (L), unbound waste (U), bead wash (W), and final elution (E) were mixed with LDS gel loading buffer (Thermo-Fisher, #NP0007) containing 50mM fresh dithiothreitol and heated to 65°C for 5 minutes. Samples were then run on NuPage 4%–12% Bis-Tris Protein Gels (Thermo-Fisher, #NP0322) according to manufacturer's instructions. Samples were then transferred to Immobilon PVDF membranes for western blotting (Sigma-Aldrich, #IPSN07852) according to manufacturer's instructions. Protein-adhered membranes were blocked with TBS-T (50mM Tris, 150mM NaCl, 0.1% Tween 20) with 3% BSA for 1 hr followed by overnight incubation with primary antibodies to detect mVenus (Anti-GFP, rabbit, 1:1000 dilution, Novus Biologicals, #NB600-308) and FLAG (Anti-FLAG-M2, mouse, 1:1000, Sigma-Aldrich, F1804) at 4°C. The following day membranes were washed 4 × 15 min in TBST and probed for 1 hr with anti-rabbit IgG HRP (1:5000, Jackson ImmunoResearch, #711-035-152) for mVenus detection or anti-mouse IgG HRP (1:3000, Cell Signaling, #7076S) for FLAG detection. Blots were washed again 4x 15 min in TBST, mixed with Clarity Western ECL Substrate (BioRad, #1705061) and imaged on a ChemiDoc Touch Imaging System (BioRad, #1708370).

LC-MS/MS Analysis

Immunoprecipitated samples were run on SDS-PAGE in every other lane to avoid cross contamination and bands were excised from 150kDa to 15kDa as determined by Precision Plus Protein Dual Color Standard (BioRad, #1610374). The proteins were reduced, alkylated, and in-gel digested with trypsin overnight at 37°C. Peptides were extracted and dried via vacuum centrifugation. All peptide samples were stored at −80°C until further analysis.

The samples were analyzed by LC/MS/MS using an Easy nLC 1200 coupled to a QExactive HF mass spectrometer (ThermoScientific). Samples were injected onto an Easy Spray PepMap C18 column (75 μm id × 25 cm, 2 μm particle size) (Thermo Scientific) and separated over a 45 min method. The gradient for separation consisted of 5%–38% mobile phase B at a 250 nl/min flow rate, where mobile phase A was 0.1% formic acid in water and mobile phase B consisted of 0.1% formic acid in 80% acetonitrile. The QExactive HF was operated in data-dependent mode where the 15 most intense precursors were selected for subsequent fragmentation. Resolution for the precursor scan (m/z 350–2000) was set to 120,000 with a target value of 3×10^6 ions. MS/MS scans resolution was set to 15,000 with a target value of 1×10^5 ions, 100 ms max IT. The normalized collision energy was set to 27% for HCD. Dynamic exclusion was set to 30 s and precursors with unknown charge or a charge state of 1 and $\geq$ 8 were excluded.

Data were searched against a Uniprot Human database (containing ~22,000 protein sequences), appended with the mVenus-VGS-Nb2 protein sequences, using Sequest within Proteome Discoverer 2.1 (Thermo Scientific). The following parameters were used to identify tryptic peptides for protein identification: 10 ppm precursor ion mass tolerance; 0.02 Da product ion mass tolerance; up to two missed trypsin cleavage sites. Carbamidomethylation of Cys was set as a fixed modification, and oxidation of Met and acetylation of N terminus were set as variable modifications. Scaffold (version Scaffold_4.7.3) was used to validate MS/MS based peptide and protein identifications. Protein/peptide identifications were accepted if they could be established at greater than 95% probability by the Scaffold Local FDR algorithm.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analysis was performed using Prism 8 (GraphPad). The specific details of each analysis type are outlined in the experimental methods section, figures, and figure legends of the specific experiment. The number of biological experimental replicates per experiment (herein, N) is outlined in each figure legend and experimental methods section.

DATA AND CODE AVAILABILITY

All graphs and data plots were generated using Prism 8 (GraphPad) unless noted otherwise herein. Alignment of tTA nucleotide sequences (Figure 2A) was performed and graphically rendered using Benchling (Benchling.com). Alignment of tTA peptide sequences (Figure 2C) as well as VEGAS cloned Nb sequences (Figure S4B) was performed using T-COFFEE (Di Tommaso et al., 2011) and graphically rendered using ESPript 3.0 (Robert and Gouet, 2014). Analysis of codon usage frequency (Figure S2D) was plotted using Morpheus (https://software.broadinstitute.org/morpheus). The datasets for human, hamster, and Sindbis virus codon usage frequency were obtained in 2018 from the Codon Usage Database (Nakamura et al., 2000). Protein structure renderings (Figures 2D and S2B) were created using PyMol 2.2 (Schrödinger) from PDB 4AC0 (Volkers and Hinrichs, 2012) and PDB 1QPI

(Orth et al., 2000). Alignment of GPCR switch regions (Figure S3A) was performed using GPCRdb (Isberg et al., 2015; van der Kant and Vriend, 2014; Pándy-Szekeres et al., 2018) and graphically plotted using Morpheus. Snake plots (Figure S3D) were rendered using GPCRdb. Alignment of nanobody NGS reads was graphically rendered using the Integrative Genomics Viewer (Robinson et al., 2011; Thorvaldsdóttir et al., 2013).

Data are deposited on the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo) under the accession number GEO: GSE123269.

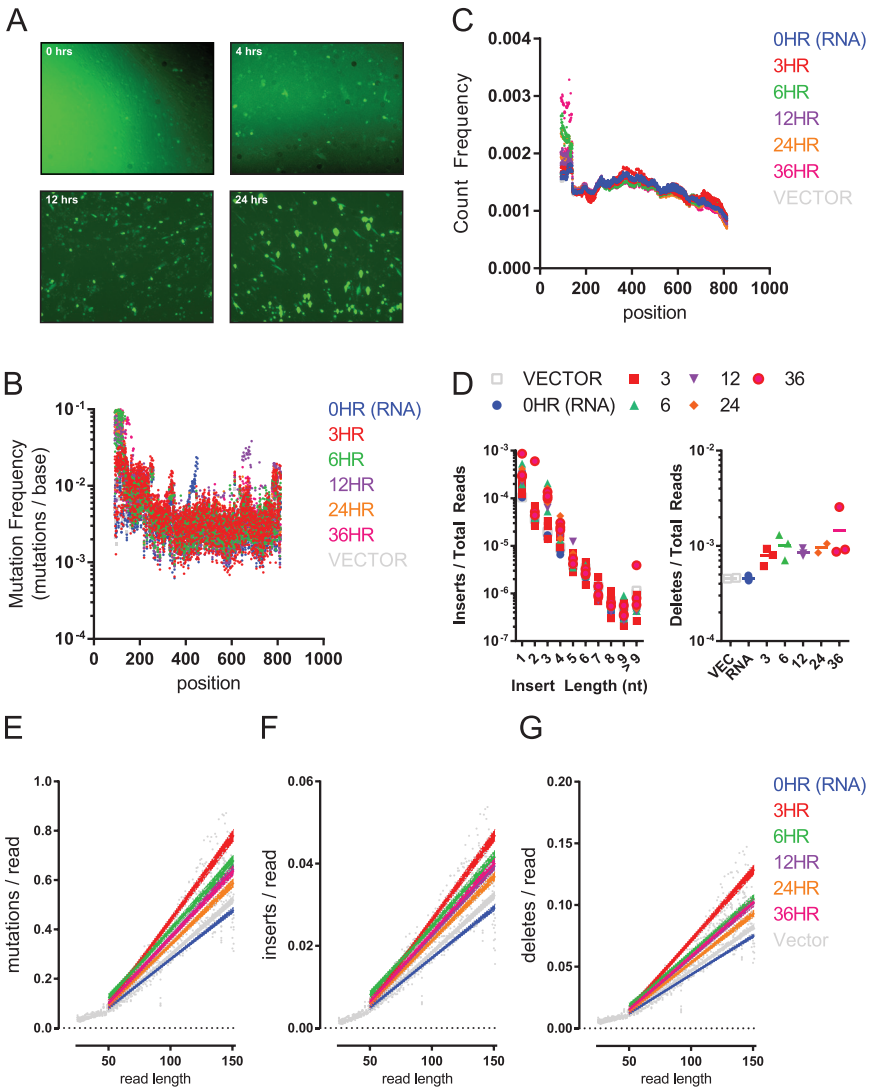


Figure S1. Sindbis Virus for Facile Directed Evolution in Mammalian Cell Culture, Related to Figure 1

(A) Epifluorescence microscope images of BHK-21 cells in culture 0, 4, 12, and 24 hr after application of pTSin-EGFP virus. (B) Mutations observed per base measured via Illumina paired-end sequencing at each position in the EGFP transgene over time. Data are plotted as individual replicates. (C) Read distribution from Illumina paired-end sequencing per position in the EGFP transgene over time. Data are plotted as individual replicates. (D) Insert number and size and deletion number observed from sequencing of Sindbis packaged EGFP transgene over time. Data are plotted as individual replicates. (E) Mutations observed per read length from sequencing of Sindbis packaged EGFP transgene over time. Data are represented individually and as linear regression, shading represents 95% confidence band. (F) Insertions observed per read length from sequencing of Sindbis packaged EGFP transgene over time. Data are represented individually and as linear regression, shading represents 95% confidence band. (E) Deletions observed per read length from sequencing of Sindbis packaged EGFP transgene over time. Data are represented individually and as linear regression, shading represents 95% confidence band.

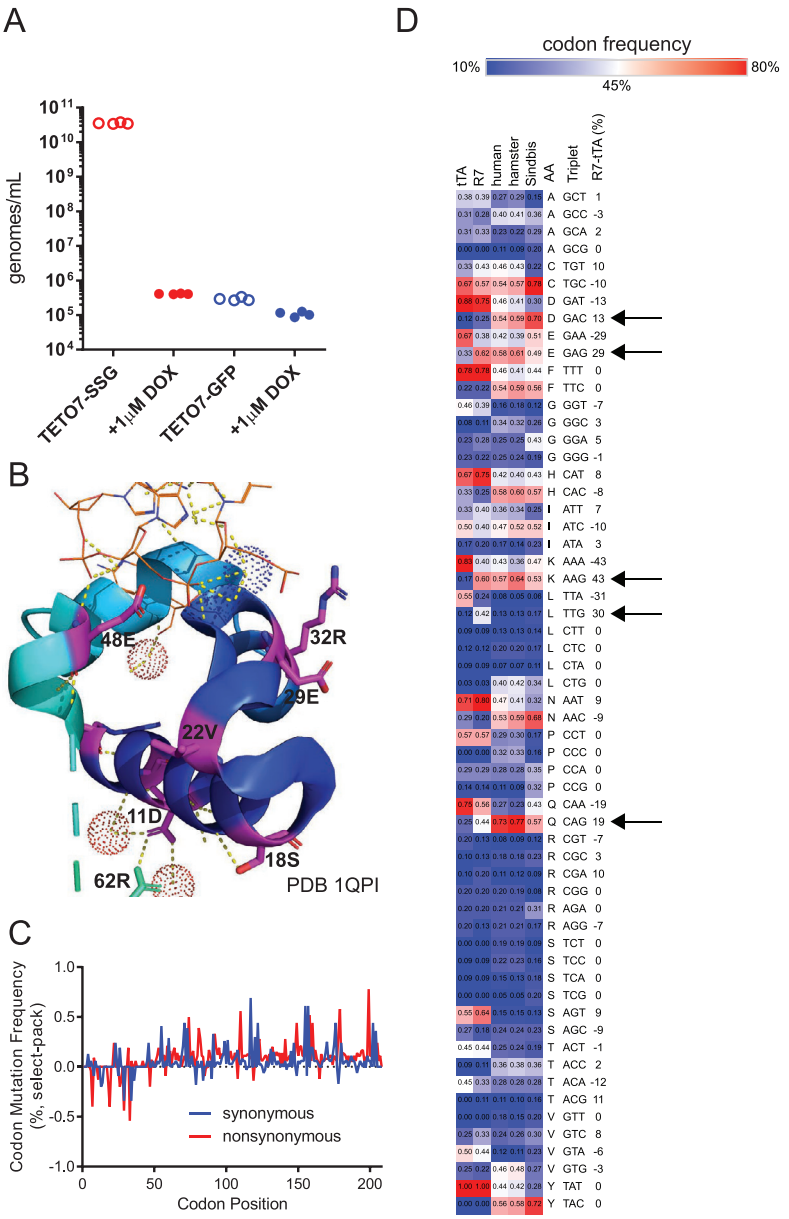


Figure S2. Directed Evolution of Transcription Factors with Sindbis, Related to Figure 2

(A) qRT-PCR quantification of Sindbis virus production from cell culture. BHK-21 cells transfected with TETO7-SSG (red) or TETO7-GFP (blue) were infected with pTSin-tTA and treated with 1 μM Doxycycline (filled circles) or vehicle (open circles). Data are plotted as individual replications. (B) Crystal structure PDB 1QPI of TETR in complex with TET operator DNA sequence. Residue interactions (yellow dash), waters (dot spheres), and DNA (orange) highlighted. R7 mutations modeled in purple and annotated with residue number and mutation. (C) Codon mutation frequency of all acquired TETR sequences versus initially packaged variants. Synonymous (blue) and nonsynonymous (red) mutations plotted at each position. (D) Codon usage frequencies (%) for wild-type tTA, R7, *Homo sapiens* (human), *Mesocricetus auratus* (hamster), and Sindbis plotted as a heatmap. The R7-tTA (%) column summarizes the % change in usage for a particular codon triplet between R7 and tTA. Arrows, > 10% conversion of R7 codon usage to those more frequently used in humans and hamsters.

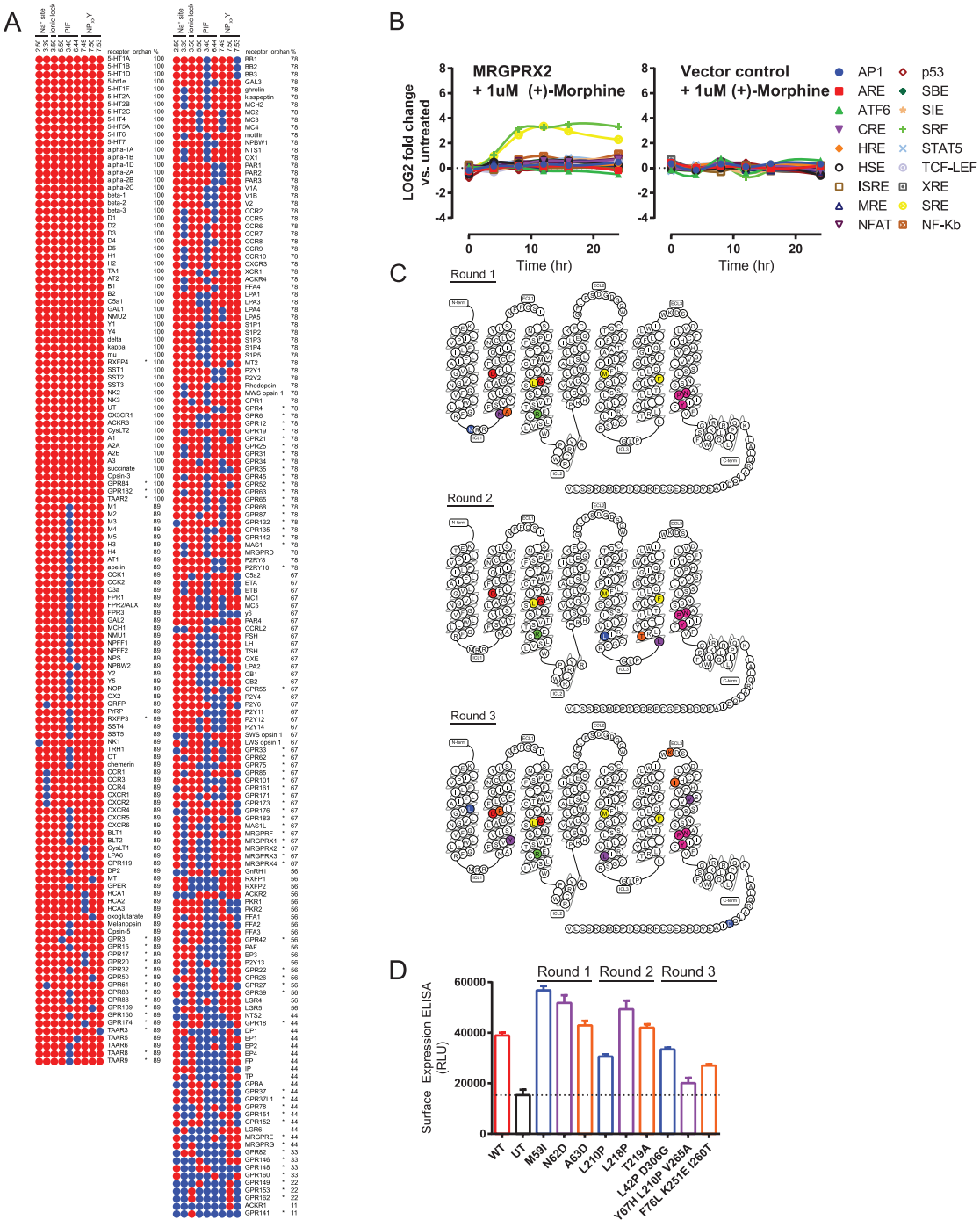


Figure S3. VEGAS for the Evolution of GPCRs, Related to Figure 3

(A) Residue identity alignment of all human Class A GPCRs across functional domains associated with transition from inactive to active receptor conformation. Ballesteros-Weinstein annotation of residue position was used for alignment. The receptor orphan classification (*) and conservation score (%) are listed. Red, conserved; blue, non-conserved. (B) Transcription factor reporter primary screen for MRGPRX2 mediated (+)-Morphine response. RLuc production equates to drug-mediated activation. Left, MRGPRX2 transfected cells. Right, empty vector transfected cells. Data are represented as mean ± SEM. (C) Snake plots of MRGPRX2 sequence and predicted secondary structure. Mutants from each round are marked in blue, purple, and orange. The locations of the Na^+ pocket (red), ionic lock (green), PIF (yellow), and NPXXY (pink) domains are also highlighted. (D) Surface expression enzyme-linked immunosorbent assay (ELISA) to quantify the total expression of MRGPRX2 (WT), untransfected cells (UT), and the MRGPRX2 mutants obtained from VEGAS. Data are represented as mean ± SEM, N = 3.

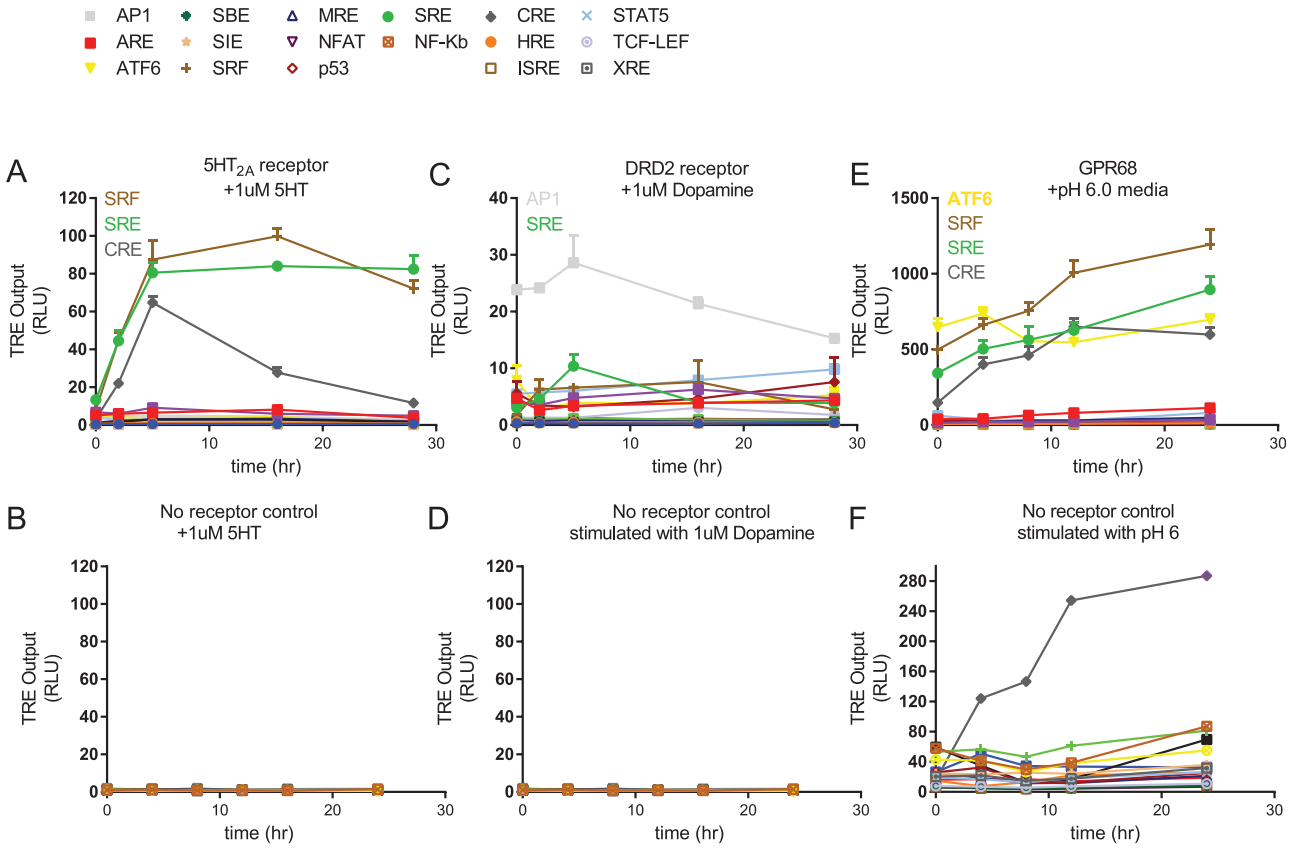
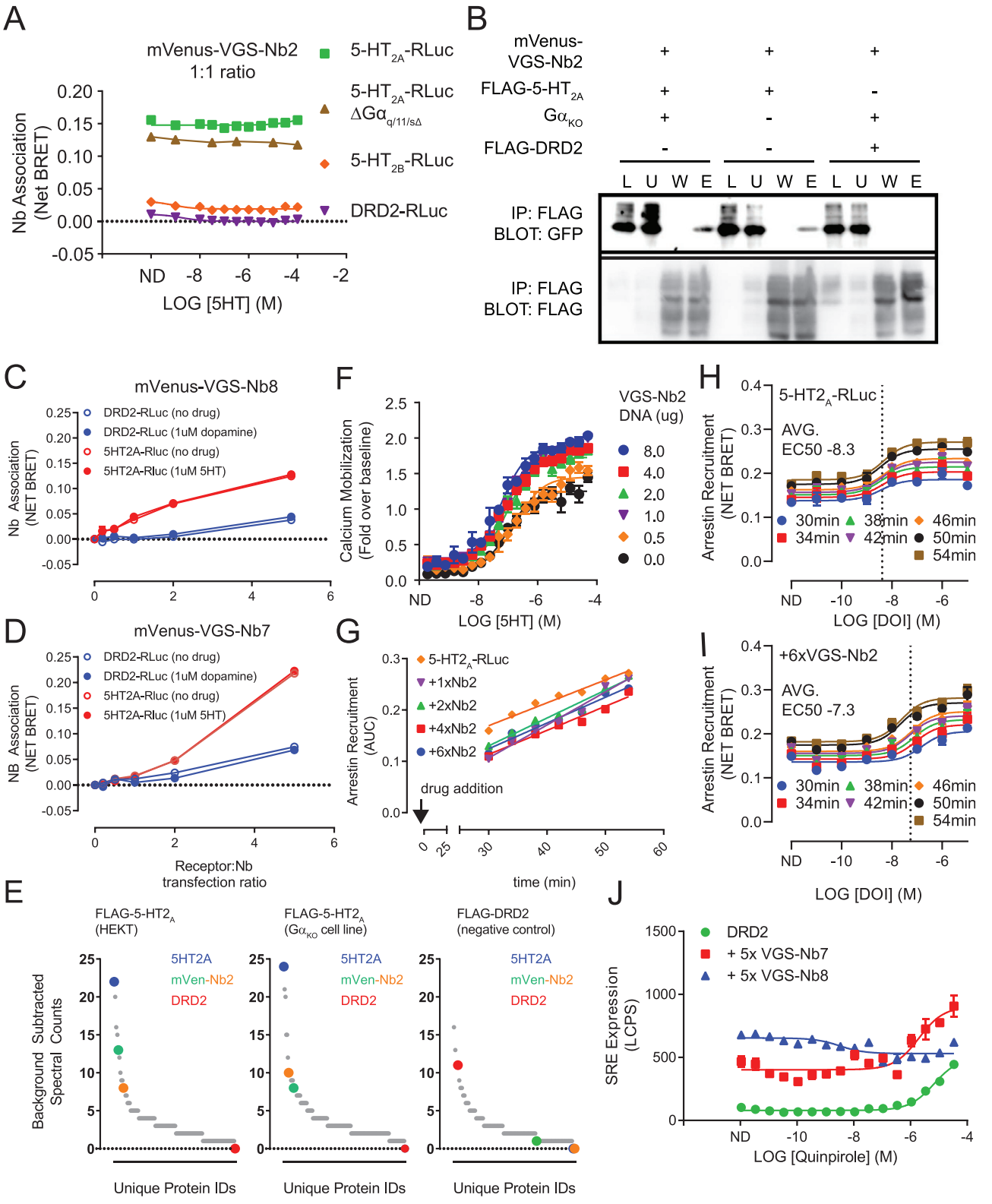


Figure S4. VEGAS for Evolution of Active-State Nanobodies, Related to Figure 4

VEGAS was used to develop nanobodies that selectively activate diverse GPCR targets from a single cDNA library. (A–F)Transcription factor reporter primary screen. Cells transfected with GPCRs (A,C,E) or pCDNA3.1 (B,D,F), and the indicated transcription response element (TRE) were stimulated. TRE-RLuc production equates to drug-mediated activation of the GPCR, quantified over time. Data are represented as mean ± SEM, N = 1.



(legend on next page)

Figure S5. Positive Allosteric Modulation of GPCRs by VEGAS-Evolved Nanobodies, Related to Figure 5

VEGAS derived nanobodies were tested for direct association and allosteric modulation of their targets. (A) Bioluminescence resonance energy transfer (BRET) association assay between GPCR-RLuc constructs and mVenus-VGS-Nb2 in the presence of increasing concentrations of 5HT. Data are represented as mean ± SEM, N = 3.

(B) Co-immunoprecipitation determination from FLAG pull-down by western blot analysis. Loaded (L), unbound (U), washed, (W), and eluted (E) fractions annotated. (C) Bioluminescence resonance energy transfer (BRET) association assay between DRD2-RLuc or 5HT2A-RLuc and mVenus-VGS-Nb8 at increasing transfection ratios of nanobody. Data are represented as mean ± SEM, N = 3 (D) Bioluminescence resonance energy transfer (BRET) association assay between DRD2-RLuc or 5HT2A-RLuc and mVenus-VGS-Nb7 at increasing transfection ratios of nanobody. Data are represented as mean ± SEM, N = 3. (E) Co-immunoprecipitation determination from FLAG pull-down by LC-MS/MS analysis. 5HT2A, blue; mVenus, green; VGS-Nb2, orange; DRD2, red; other peptides, gray. Data are represented as total background subtracted spectral counts, N = 2. See Table S5 for raw data. (F) Calcium flux assay. Calcium mobilization equates to relative GPCR-mediated $G\alpha_q$ activity. Data are represented as mean ± SEM, N = 3. (G–I) Bioluminescence resonance energy transfer (BRET) association assay between 5HT2A-RLuc and mVenus-βArrestin-2 in the presence of increasing concentrations of DOI. (H) receptor only, (I) VGS-Nb2 co-transfected, (G) linear regression of area under the curve datasets for G, H, and additional nanobody concentrations. Data are represented as mean ± SEM, N = 3. (J) Serum response element (SRE) reporter assay. DRD2 alone or co-transfected with 5x VGS-Nb7 or VGS-Nb8. RLuc production equates to relative receptor activation. Data are represented as mean ± SEM, N = 2.