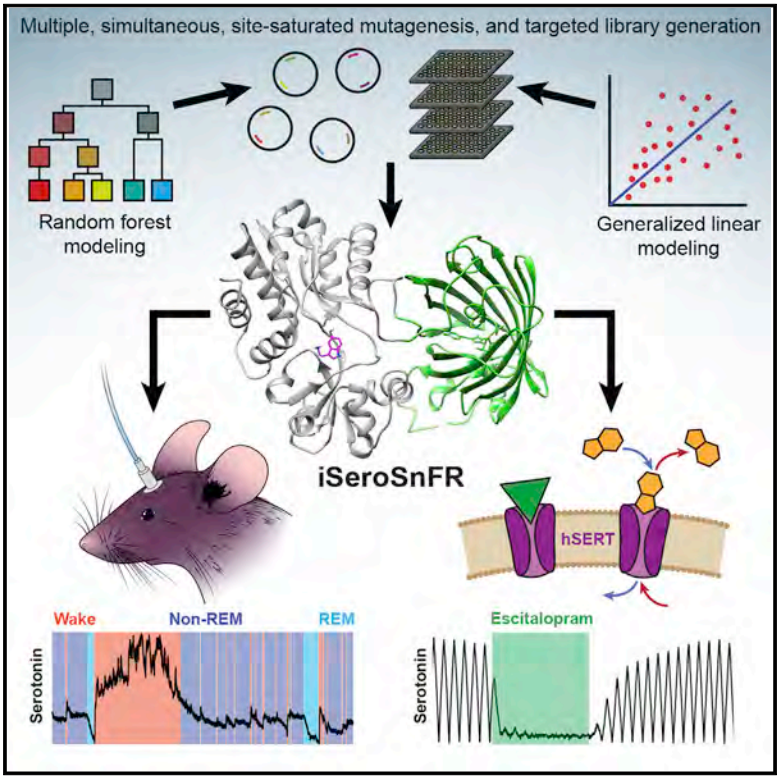
Cell

Directed Evolution of a Selective and Sensitive Serotonin Sensor via Machine Learning

Graphical Abstract



Highlights

- Developed a machine learning approach for rapid binding-pocket redesign
- Engineered a high dynamic range, sensitive, selective sensor for 5-HT: iSeroSnFR
- Demonstrated the use of iSeroSnFR for fiber photometry in awake behaving mice
- Developed a clinically relevant assay using iSeroSnFR for 5-HT transporter function

Authors

Elizabeth K. Unger, Jacob P. Keller, Michael Altermatt, ..., Viviana Gradinaru, Loren L. Looger, Lin Tian

Correspondence

loogerl@hhmi.org (L.L.L.), lintian@ucdavis.edu (L.T.)

In Brief

Machine-learning-guided binding-pocket redesign enables engineering of genetically encoded sensor for serotonin that detects serotonin release in freely behaving mice and is used for the development of an assay for serotonin transporter function and modulation by drugs.

Unger et al., 2020, Cell *183*, 1986–2002
December 23, 2020
https://doi.org/10.1016/j.cell.2020.11.040

CellPress

CellPress

Cell

Resource

Directed Evolution of a Selective and Sensitive Serotonin Sensor via Machine Learning

Elizabeth K. Unger,[1,12] Jacob P. Keller,[2,9,12] Michael Altermatt,[3,12] Ruqiang Liang,[1] Aya Matsui,[5] Chunyang Dong,[1] Olivia J. Hon,[6] Zi Yao,[4] Junqing Sun,[1] Samba Banala,[2] Meghan E. Flanigan,[6] David A. Jaffe,[1] Samantha Hartanto,[1] Jane Carlen,[1] Grace O. Mizuno,[1,10] Phillip M. Borden,[2,11] Amol V. Shivange,[3] Lindsay P. Cameron,[1] Steffen Sinning,[7] Suzanne M. Underhill,[8] David E. Olson,[1] Susan G. Amara,[8] Duncan Temple Lang,[1] Gary Rudnick,[7] Jonathan S. Marvin,[2] Luke D. Lavis,[2] Henry A. Lester,[3] Veronica A. Alvarez,[5] Andrew J. Fisher,[1] Jennifer A. Prescher,[4] Thomas L. Kash,[6] Vladimir Yarov-Yarovoy,[1] Viviana Gradinaru,[3] Loren L. Looger,[2,*] and Lin Tian[1,13,*]

[1]Departments of Biochemistry and Molecular Medicine, Chemistry, Statistics, Molecular and Cellular Biology, and Physiology and Membrane Biology, the Center for Neuroscience, and Graduate Programs in Molecular, Cellular, and Integrative Physiology, Biochemistry, Molecular, Cellular and Developmental Biology and Neuroscience, University of California, Davis, Davis, CA 95616, USA

[2]Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20174, USA

[3]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

[4]Department of Chemistry, University of California, Irvine, Irvine, CA 92697, USA

[5]Laboratory on Neurobiology of Compulsive Behaviors, National Institute on Alcohol Abuse and Alcoholism, NIH, Bethesda, MD 20892, USA

[6]Bowles Center for Alcohol Studies, Department of Pharmacology, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA

[7]Department of Pharmacology, Yale University School of Medicine, New Haven, CT 06520, USA

[8]Laboratory of Molecular and Cellular Neurobiology, National Institute on Mental Health, NIH, Bethesda, MD 20892, USA

[9]Present address: Department of Pharmacology and Molecular Therapeutics, Uniformed Services University, Bethesda MD 20814, USA

[10]Present address: Seven Biosciences, 630 Pena Ave., Suite 400, Davis, CA 95618, USA

[11]Present address: LifeEDIT, Research Triangle Park, NC 27709, USA

[12]These authors contributed equally

[13]Lead Contact

*Correspondence: loogerl@hhmi.org (L.L.L.), lintian@ucdavis.edu (L.T.)

https://doi.org/10.1016/j.cell.2020.11.040

SUMMARY

Serotonin plays a central role in cognition and is the target of most pharmaceuticals for psychiatric disorders. Existing drugs have limited efficacy; creation of improved versions will require better understanding of serotonergic circuitry, which has been hampered by our inability to monitor serotonin release and transport with high spatial and temporal resolution. We developed and applied a binding-pocket redesign strategy, guided by machine learning, to create a high-performance, soluble, fluorescent serotonin sensor (iSeroSnFR), enabling optical detection of millisecond-scale serotonin transients. We demonstrate that iSeroSnFR can be used to detect serotonin release in freely behaving mice during fear conditioning, social interaction, and sleep/wake transitions. We also developed a robust assay of serotonin transporter function and modulation by drugs. We expect that both machine-learning-guided binding-pocket redesign and iSeroSnFR will have broad utility for the development of other sensors and *in vitro* and *in vivo* serotonin detection, respectively.

INTRODUCTION

Serotonergic systems profoundly modulate diverse behaviors (Berger et al., 2009; Charnay and Léger, 2010). Serotonin (5-HT) dysregulation has been implicated in mental disorders, including depression and anxiety (Belmaker and Agam, 2008; Calhoon and Tye, 2015). Most antidepressants target some aspect of the serotonergic system; selective serotonin reuptake inhibitors (SSRIs) specifically target the 5-HT transporter (SERT) (Bos et al., 2012; Cipriani et al., 2018). Despite the critical importance of 5-HT, development of novel and more effective therapies has been challenging due to poor understanding of 5-HT dynamics, specifically the inability to measure 5-HT with high spatiotemporal resolution. Existing methods for measuring 5-HT, including microdialysis and fast-scan cyclic voltammetry (FSCV) (Abdalla et al., 2017; Jaquins-Gerstl and Michael, 2015; Peñalva et al., 2003; Schultz and Kennedy, 2008), as well as transporter assays relying primarily on radiolabeled 5-HT or analogs (Clarke and Khalid, 2015), lack the spatial or temporal resolution to adequately probe 5-HT dynamics and targetability.

A genetically encoded 5-HT sensor could potentially overcome these technical challenges (for review, see Broussard

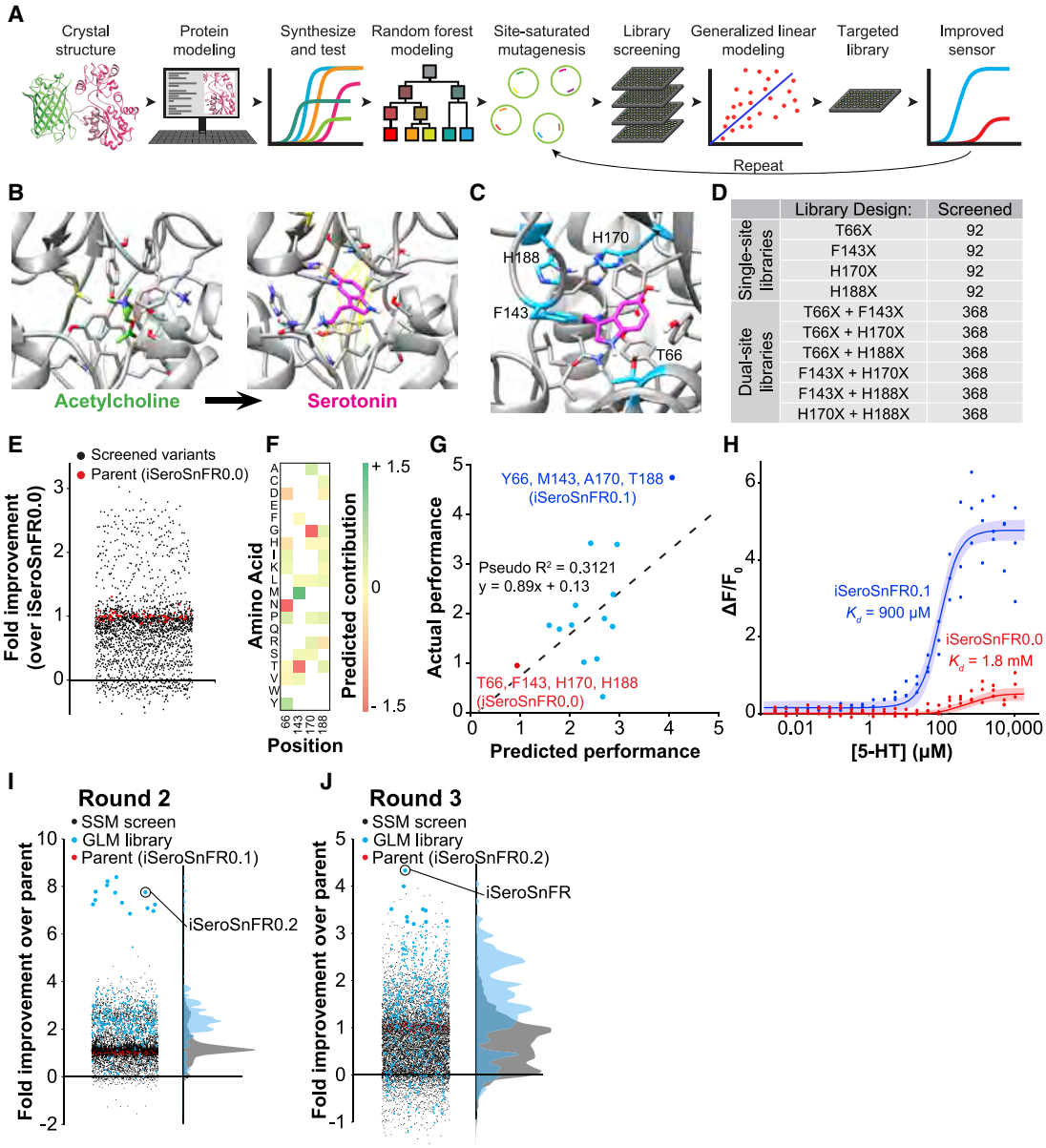


Figure 1. Using Machine Learning to Evolve Binding Proteins

(A) Overview of machine learning method.

(B) Schematic showing conversion of an acetylcholine (ACh)-binding protein to a serotonin (5-HT)-binding protein. ACh and 5-HT were docked into the binding pocket of AChSnFR0.6 using Rosetta. Statistical modeling was performed on these models, and promising mutations were synthesized and tested (see Figure S2 and Table S3). iSeroSnFR0.0 was chosen as a starting point for statistical modeling. Positions 66, 143, 170, and 188 were selected for further mutation.

(C) The binding pocket of iSeroSnFR0.0 was simulated using Rosetta and 5-HT (magenta) docked. Top-ranked positions are labeled (cyan).

(D) Table of DNA libraries created, and number of variants screened from each library.

(E) DNA libraries were generated, transformed into bacteria, grown, and lysed. Lysate was then screened with 10 mM 5-HT and compared to the parent sensor (iSeroSnFR0.0).

(F). Heatmap of the contribution of each mutation at each position screened, as predicted by the generalized linear model (GLM) (for additional information see Table S3).

(G) Combinations of mutations predicted to be better than the parent (iSeroSnFR0.0) were synthesized and tested as purified protein with 10 mM 5-HT. Dashed line represents the linear regression.

(H) Protein from iSeroSnFR0.0 and the top variant (iSeroSnFR0.1) was purified and tested against multiple concentrations of 5-HT. Fits were determined using the Hill equation. Shaded area denotes 95% confidence interval. The apparent K_d is defined as the concentration of the ligand producing 50% of the maximum fluorescence change.

(legend continued on next page)

et al., 2014; Lin and Schnitzer, 2016; Looger and Griesbeck, 2012). We and others have recently produced single-fluorescent protein (FP)-based sensors for neurotransmitters and neuromodulators based on either microbial periplasmic binding proteins (PBPs) or G-protein-coupled-receptors (GPCRs). Combined with modern microscopy, these sensors enable direct and specific measurements of diverse neurotransmitters and neuromodulators including glutamate (Marvin et al., 2013), GABA (Marvin et al., 2019), ATP (Lobas et al., 2019), dopamine (Patriarchi et al., 2018; Sun et al., 2018), acetylcholine (ACh) (Borden et al., 2020; Jing et al., 2020), and norepinephrine (Feng et al., 2019) with the necessary resolution for use in behaving animals (Corre et al., 2018; Madisen et al., 2015).

Although GPCR-based sensors can yield sensitive indicators, their response to pharmaceutical manipulations makes them problematic for use in any studies involving drug administration. Furthermore, it is difficult to target GPCR-based sensors to intracellular locations (e.g., to study 5-HT transport). PBP-based sensors are soluble and can therefore readily be targeted to subcellular locations, are amenable to high-throughput screening in bacteria, and easily allow detailed characterization in purified protein (Marvin et al., 2011). Naturally evolved PBPs typically bind few, if any, drugs targeting host proteins. In addition, microbial PBPs are bio-orthogonal to pathways in model organisms, promising minimal interference with endogenous signaling, a particular concern following long-term expression. Importantly, ligand binding in PBPs induces large conformational changes, resulting in very large dynamic ranges (Marvin et al., 2011, 2013, 2018). However, with no annotated PBPs for 5-HT, we opted to redesign the binding pocket of an existing PBP-based sensor to selectively bind 5-HT.

There are several complementary approaches available for binding-site redesign (Baker, 2019; Khoury et al., 2014). Site-saturated mutagenesis (SSM) combined with rational design can optimize sensor properties such as brightness, dynamic range, kinetics, and affinity (Cobb et al., 2013; Packer and Liu, 2015). However, SSM, while sufficient to convert our ACh sensor to one binding the cholinergic agonist nicotine (Shivange et al., 2019), cannot plausibly cover sufficient sequence space to radically change binding specificity to the structurally very different 5-HT. Computational design (Rosetta) can successfully predict proteins with high affinity and selectivity for a given small molecule (Richter et al., 2011), but has never been used for sensors. Recently, machine-learning approaches of varying complexity have been applied to optimize protein function (Bedbrook et al., 2019; Biswas et al., 2018; Ehren et al., 2008; Liao et al., 2007; Saito and Nakatsuji, 2001; Wu et al., 2019). Our approach combined computational design and machine learning to guide an SSM pipeline. Using this method, we developed the first PBP-based 5-HT sensor (iSeroSnFR) by redesigning the binding pocket of the ACh sensor iAChSnFR0.6 (Borden et al., 2020). iSeroSnFR contains 19 mutations relative to iAChSnFR0.6; these conferred >5,000-fold improvement in 5-HT affinity while eliminating choline and ACh binding. We demonstrate that iSeroSnFR enables imaging of 5-HT dynamics in brain slices and freely moving mice. In addition, we highlight the clinical relevance of iSeroSnFR for pharmacological assays.

RESULTS

Sensor Design

Before redesigning a binding pocket for structurally disparate molecules, we established our machine learning pipeline on structurally similar molecules (Figure S1; Table S1; Data S1; STAR Methods). As a starting scaffold for our sensor, we chose an early version of iAChSnFR, based on a choline-binding PBP, OpuBC, from the hyperthermophile *Thermoanaerobacter sp. X513* (Miller et al., 2011). In addition, this variant displayed detectable binding to 5-HT (apparent K_d >1 mM) (Figure S2A), making it a good starting template. We used a multi-stage pipeline to iteratively improve 5-HT binding and sensor response (Figure 1A). We first performed computational binding-pocket redesign using Rosetta, then iterative rounds of SSM guided by machine learning.

Step I: Computational Binding-Pocket Redesign

To perform computational modeling, we started with the structures of open, ligand-free iAChSnFR0.6 (6URU) and closed, choline-bound *Bacillus subtilis* OpuBC (3R6U), and created a model of the closed, ACh-bound form of iAChSnFR using Rosetta (Figure 1B). 5-HT conformers (rotation of the β-aminoethyl and hydroxy moieties) were generated using Open Eye Omega (Hawkins et al., 2010), and docked into the closed-iAChSnFR model using RosettaLigand (Bender et al., 2016; Davis and Baker, 2009) (Figures 1B and 1C). Next, Rosetta protein redesign (Taylor et al., 2016; Tinberg et al., 2013) was used to optimize the 5-HT binding pocket. In total, 250,000 variants were predicted and ranked based on computed ligand interaction (STAR Methods). The top 18 predicted variants were synthesized, purified, and examined for fluorescence response to 5-HT and other ligands including ACh (Figures S2B and S2C). Among the 18 variants, variant 7 showed the largest fluorescence response to 10 mM 5-HT ($\Delta F/F_0$ = 87% ± 20%) with no ACh response ($\Delta F/F_0$ = −4% ± 1%), representing an 18-fold improvement in 5-HT selectivity (Figure S2C). This mutated variant, named iSeroSnFR0.0 (Table S2), was then selected for further optimization.

Step II: Random Forest Modeling

We next optimized iSeroSnFR0.0 with SSM to improve 5-HT affinity. We used a random forest (RF) model to estimate the importance of each position interrogated by computational design (Table S2). We took the four highest-ranked positions from RF (66 > 170 > 143 > 188) (Figure 1D) and performed SSM at each site separately and in pairs (using degenerate NNK codons). We

(I) Raincloud plot where iSeroSnFR0.1 was used as the parent for a second round of screening followed by GLM analysis. A small library (32 possible combinations) was generated based on the GLM results and screened (cyan), which led to the discovery of iSeroSnFR0.2.
(J) Raincloud plot similar to (I), but using iSeroSnFR0.2 as the parent. This GLM-guided targeted library (96 possible combinations) was created and screened, leading to the discovery of iSeroSnFR.
See also Figures S1 and S2 and Data S1.

screened a total of 2,576 variants, including 92 from each single-site library and 368 from each dual-site library, for fluorescence response to 5-HT (10 mM). The library size was determined using the TopLib online library calculator (Nov, 2012). Of the screened variants, ~100 variants showed an improved response (~2- to 3-fold) compared to iSeroSnFR0.0 (Figure 1E). Subsequent analysis showed that top-performing variants frequently contained mutations at multiple positions, and the ordered contribution of each position (66 > 143 > 170 > 188) (Table S2) to the fluorescence response was nearly identical to that predicted by RF (66 > 170 > 143 > 188) (Figure 1F). No single mutation drastically improved 5-HT affinity, but combinations of mutations were frequently better than single mutations (Figure S2D). These results suggest that RF effectively predicts important positions contributing to sensor response, and simultaneous, beneficial contributions of multiple residues are essential to large-scale improvements.

Step III: Generalized Linear Model

Because single mutations offered only small improvements, whereas combinations gave much better results, mutations were clearly not additive. For example, we found a top-performing variant containing T66Y/H170A with 140% improved response, whereas T66Y and H170A alone showed only 40% and 10% improvements, respectively (Figure S2D). We next applied GLM to our dataset. This allowed us to identify individual mutations from our libraries that contribute to synergistic interactions—allowing us to design small targeted libraries.

GLM predicted that several amino acid mutations at each position would be beneficial (Figure 1F; Table S3), with 66Y, 66P, 143M, 170A, 188G, 188P, and 188T showing the strongest positive predictions. Given the beneficial T66Y/H170A mutant that we had already identified, we decided to synthesize variants combining those with GLM-predicted amino acid residues at positions 143 and 188. Out of 13 variants synthesized, 12 showed larger fluorescence response to 10 mM 5-HT than iSeroSnFR0.0 (Figure 1G) (GLM was moderately predictive: pseudo-R^2 = 0.31). One variant displayed 4.5-fold improved response over iSeroSnFR0.0, well above any SSM-screened variant. In purified protein, this variant (T66Y/F143M/H170A/H188T, named iSeroSnFR0.1) showed >9-fold increased fluorescence response $(\Delta F/F_0)_{max}$ relative to iSeroSnFR0.0 ($(\Delta F/F_0)_{max}$ = 480% ± 14% versus 50% ± 4%), and 2-fold increased 5-HT affinity (900 ± 110 μM versus 1.8 ± 0.5 mM) (Figure 1H).

In light of the substantial improvements from a single round of GLM-guided mutagenesis, we performed two more rounds of screening followed by GLM prediction, recruiting additional positions predicted by RF, and added others based on prior experience optimizing biosensors (e.g., linkers connecting the cpGFP to the binding protein, the interface between cpGFP and binding protein, and random mutations from previous rounds of screening) (Table S2). These subsequent rounds were screened at progressively lower 5-HT concentrations (round 2: 500 μM and round 3: 50 μM) to enrich for variants with tighter affinity. After each round, the top-performing variants were sequenced and re-analyzed using GLM to create a focused library; in total, ~13,000 variants were tested, of which 244 were sequenced. The best variant from round 2 (Figure 1I) came from the GLM-driven focused library, had 8-fold improved response over its parent iSeroSnFR0.1, and was named iSeroSnFR0.2. The best variant from round 3 (Figure 1J) was 5-fold improved over its parent iSeroSnFR0.2 (Figure 2A; Table S2). We named this final version iSeroSnFR, which contains 19 mutations relative to iAChSnFR0.6 (Figure 2A, PDB: 6PER); In purified protein, this variant exhibits 310 ± 30 μM affinity for 5-HT, and 800% $(\Delta F/F_0)_{max}$ (Figure 2B).

Our data show that GLM is highly effective at identifying beneficial mutations (Figures 1I and 1J): the average performance of GLM variants was significantly higher than SSM variants (1st, 2nd, and 3rd rounds: $p < 0.001$, Wilcoxon rank-sum test) and more importantly, the top variants in each round (including the top 15 variants from round 2) were from the GLM-inspired library as opposed to the SSM library. Thus, our machine-learning-guided mutagenesis pipeline can improve protein-ligand binding selectivity and affinity even for target molecules structurally distant from cognate ones, while still maintaining sensor function.

In Vitro *Characterization*

iSeroSnFR is highly specific for 5-HT over a wide array of endogenous molecules and drugs (Figures 2C and S3). Only tryptamine and dopamine showed detectable responses, but with 8- and 16-times weaker affinity (apparent K_d = 2.4 mM for tryptamine and 5.6 mM for dopamine), respectively. Other endogenous molecules showed marginal responses, with very low or negative responses, and/or titrations that did not fit a single-site binding isotherm.

To determine the *in situ* affinity in mammalian (HEK293T) cells, we cloned iSeroSnFR into the pMinDisplay expression vector (pDisplay lacking the HA epitope tag) (Marvin et al., 2013) such that iSeroSnFR would be displayed on the outer cell surface; we also cloned a version targeted to the post-synapse using full-length neuroligin (Nlgn) (STAR Methods). We observed robust membrane localization (Figure 2D) and clear response to 1.6 μM 5-HT (~50% $\Delta F/F_0$) (Figures 2E and S4A). The *in situ* affinity of the sensor on HEK293T cells was similar (390 ± 110 μM) to that in purified protein. Fortuitously, the fluorescence response of iSeroSnFR on HEK293T cells was significantly increased ($[\Delta F/F_0]_{max}$ = 1,700%). To characterize the utility of iSeroSnFR for physiologically relevant concentrations, we focused on the range from high-pM (338 pM) to mid-μM (60 μM) (Figures 2F and S4B). We observed small but reliable responses to all concentrations tested ($\Delta F/F_0$ = 16.8% ± 1.9% at 338 pM; similar responses up to 246 nM 5-HT) (Figure S4B). We quantified the ability of iSeroSnFR to faithfully report 5-HT by comparing 5-HT responses to Hank's balanced salt solution (HBSS) buffer responses (Figure S4C) with two metrics from signal detection theory: the receiver operating characteristic (ROC) (Figure S4D) and the discriminability index (d') (Figure S4E). ROC analysis indicates that 5-HT responses show essentially perfect discrimination (area under the curve = 0.990) (Figure S4D, bottom) of true-positives from false-positives, whereas buffer responses have no power (Figure S4D, top). Similarly, d' analysis shows that 5-HT responses have very strong separation from system noise ($d' > 3$) (Figure S4E, bottom), but buffer responses show essentially no separation (Figure S4E, top). Results on the surface of dissociated

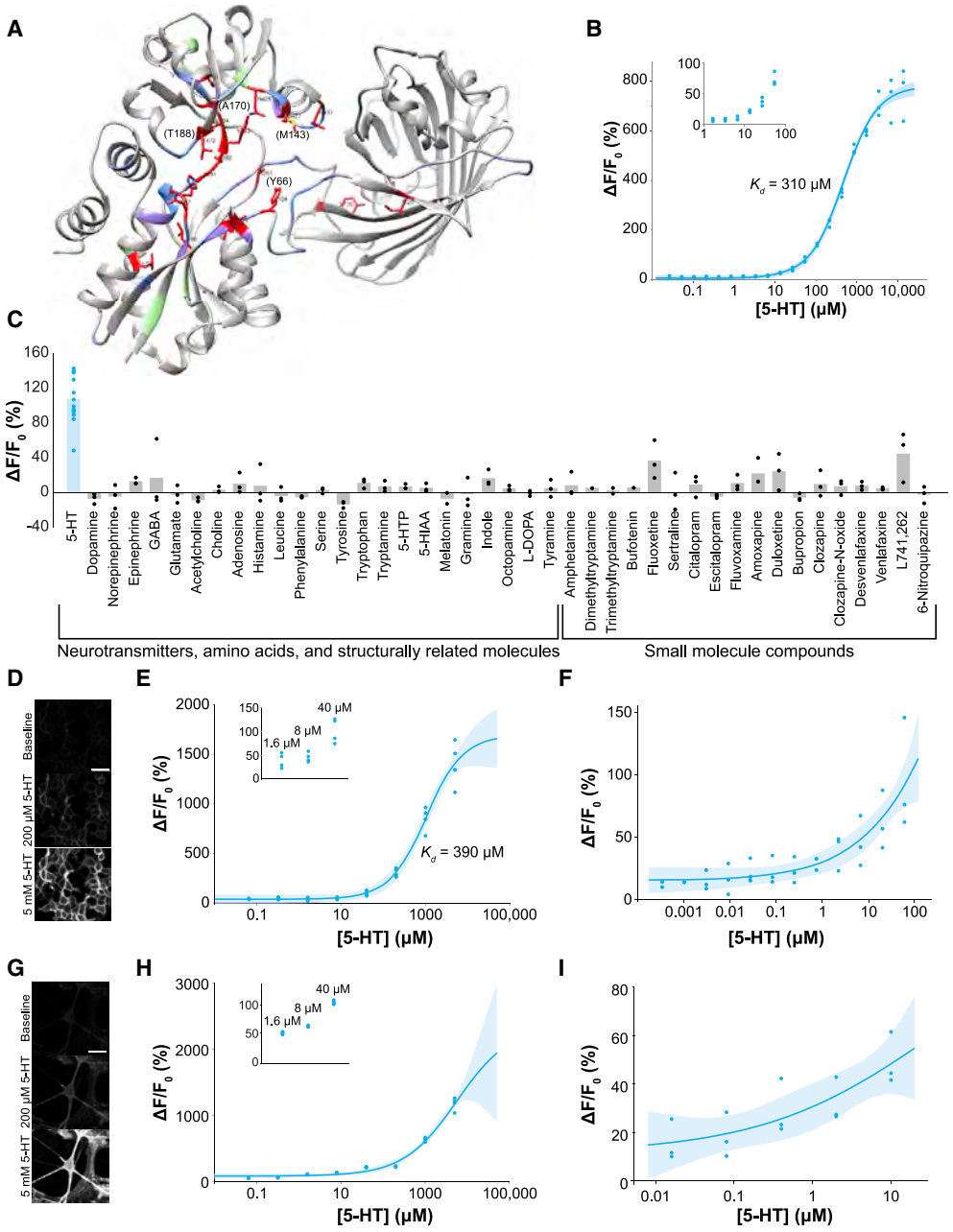


Figure 2. Affinity and Specificity of the Sensor

(A) Crystal structure of unliganded iSeroSnFR (PDB: 6PER). Mutations in iSeroSnFR relative to iAChSnFR0.6 are mapped onto the crystal structure (red). Positions interrogated by site-saturated mutagenesis (but not mutated in iSeroSnFR) are displayed in blue, mutations interrogated by Rosetta, but not SSM, in purple, and positions that were randomly mutated, in green.

(B) Purified iSeroSnFR binding to 5-HT.

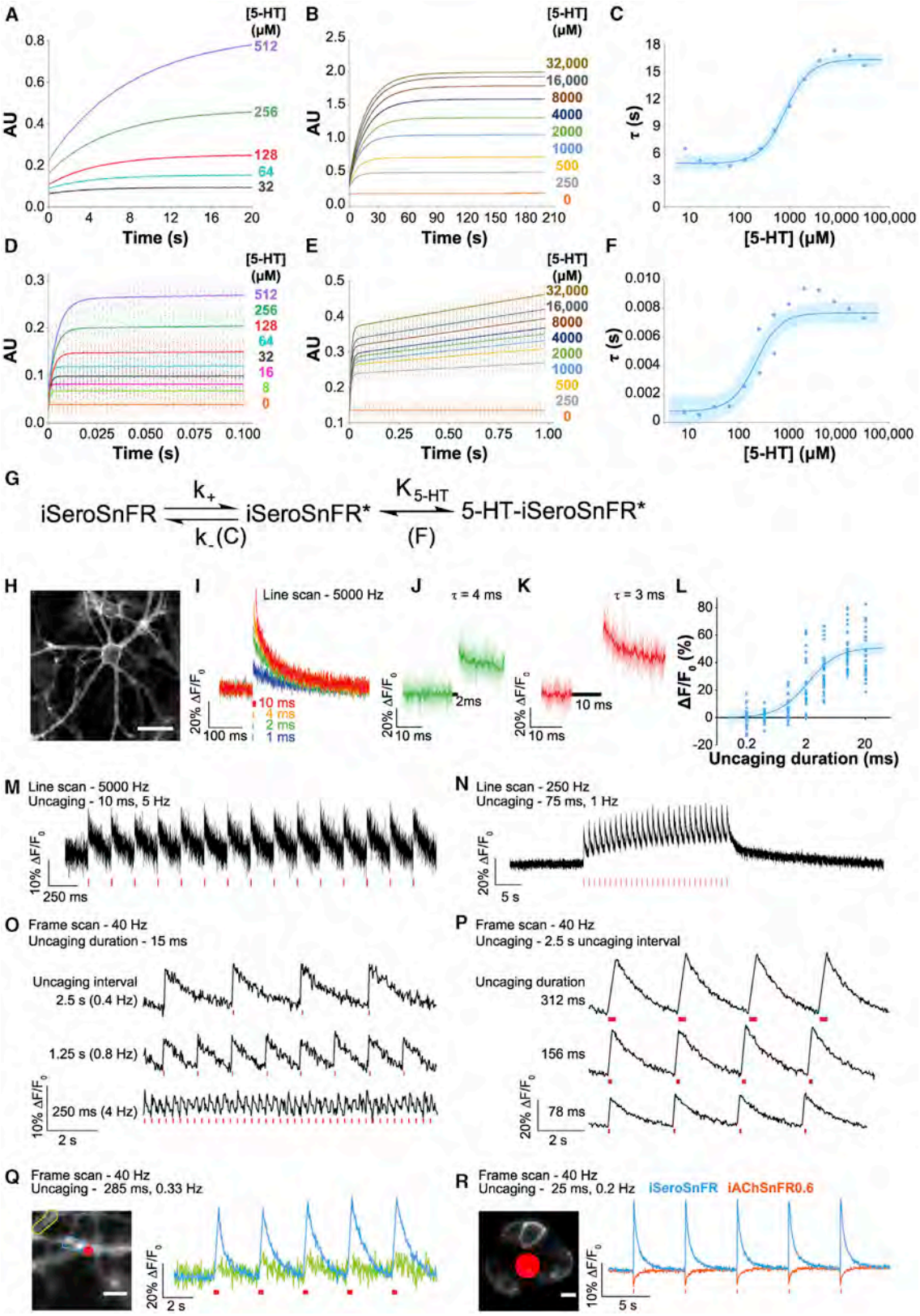
(C) Purified iSeroSnFR binding to multiple ligands. Due to differential compound solubility, the values displayed match the following concentrations: octopamine, L-phenylalanine, 80 μM; 5-HTP, 85 μM; sertraline, 110 μM; L-DOPA, tyramine, escitalopram, citalopram, amoxapine, 125 μM; all other compounds were tested at either 100 or 105 μM. For the full concentration curve for each compound, see Figure S3.

(D–F) Response of membrane-displayed iSeroSnFR in HEK293T cells. Representative images (D), and dose-response curves for higher concentrations (E) and lower concentrations (F). n = 3–4.

(G–I) Response of membrane-displayed iSeroSnFR-PDGFR in cultured neurons. (G–I) Representative images (G), and dose-response curve for higher concentrations (H) and lower concentrations (I). n = 3–4. For raw traces, see Figure S4.

(B, E, F, H, and I) Fits were determined using the Hill euqation. Shaded area denotes 95% confidence interval. Scale bars represent 50 μm. Insets show magnifications of the points at low concentrations.

See also Figures S3 and S4.



(legend on next page)

hippocampal neurons were similar to those on HEK293 cells ($\Delta F/F_0$ = 15.8% ± 2.0% at 16 nM 5-HT) (Figures 2G–2I and S4F).

We next sought to assess sensor kinetics. Using stopped-flow fluorescence, we found that the sensor reached saturation after ~10–40 s (Figures 3A and 3B). With a single exponential fit, $\tau_{apparent}$ was between 5 and 18 s (Figure 3C). We noted, however, that there appeared to be a very fast and robust initial rise (Figures 3D and 3E). We therefore fit a double-exponential curve with two phases of activation: a fast initial rise (Figure 3F), and a slower rise to saturation. Our double-exponential fit the data much more precisely and showed that the two phases occur with a rise τ of 0.5–10 ms (fast) and 5–18 s (slow). We postulate that ligand binding rapidly activates fluorescence of the sensor (Figure 3G, $K_{5\text{-HT}}$), but that the sensor in purified protein exists in an "inactive" (iSeroSnFR) and an "active" (iSeroSnFR*) conformation, and the transition between these states is slow (Figure 3G, k_+).

We then validated the kinetics of iSeroSnFR by using light-evoked release of 5-HT from a caged version (PA-N-5HT). Single pulses (1 ms) of uncaged 5-HT were robustly detected on the surface of cultured neurons (Figures 3H and 3I), and the amplitude of fluorescence changes increased with of uncaging pulse duration (Figures 3I–3L). Decay to baseline featured a fast component (τ_{decay} ~4 ms at 10 ms pulse duration) followed by a slower component (τ_{decay} ~150 ms at 10 ms pulse duration) (Figures 3J and 3K). Diffusion of uncaged 5-HT away from the imaging location likely explains the slower component. To determine how iSeroSnFR responds to repeated 5-HT bursts, we used various combinations of uncaging pulse durations and frequencies. One-photon line-scans at 5,000 Hz or 250 Hz, and frame-scans at 40 Hz revealed robust, reproducible fluorescence increases to trains of uncaging pulses at various frequencies over prolonged imaging (line scans: 10 ms uncaging pulses at 5 Hz, Figure 3M, and 75 ms uncaging pulses at 1 Hz, Figure 3N; frame scan: 15 ms uncaging pulses at 0.4, 0.8, and 4 Hz, Figure 3O). Fluorescence response amplitude increased with uncaging pulse duration under frame-scan (Figure 3P). However, when line-scan mode was used, the amplitude decreased with increased pulse frequency (Figures 3O and S5A), possibly due to accumulation of 5-HT in the bath and resulting sensor saturation. It is also possible that this results from the slow phase of the sensor response, as described above. 5-HT uncaged at one dendrite produced a negligible response in dendrites ~10 μm away (Figure 3Q). Although 5-HT uncaging readily activated iSeroSnFR, it failed to activate iAChSnFR0.6 (Figure 3R), indicating that the fluorescence response is specific to release of 5-HT, and not artefactual.

Imaging Endogenous Release of 5-HT in Brain Slice

Serotonergic fibers from the raphe nuclei innervate many brain regions, with dense projections found in medial prefrontal cortex (mPFC), orbitofrontal cortex (OFC), striatum, bed nucleus of the stria terminalis (BNST), and basolateral amygdala (BLA) (Bang et al., 2012; Belmer et al., 2017; Zhou et al., 2015). To demonstrate iSeroSnFR's utility in mouse brain slices, we first examined the expression level of iSeroSnFR in striatum and in mPFC. We cloned iSeroSnFR into regular and Cre-dependent adeno-associated virus (AAV) backbones, producing AAV2/9.*CAG*.iSeroSnFR.NIgn and AAV2/9.*CAG*.FLEx.iSeroSnFR.PDGFR. We injected AAV2/9.*CAG*.iSeroSnFR.NIgn into dorsal striatum of C57BL/6J mice (Figure 4A), and a mix of AAV2/9.*CAG*.FLEx.iSeroSnFR.PDGFR and AAV2/9.*CMV*.Cre (9:1 ratio) into mPFC of *Rosa*$^{Ai14/+}$ (Cre-dependent tdTomato expression) (Madisen et al., 2010) (Figure S5B). After 2 weeks recovery, iSeroSnFR expression was obvious in dorsal striatum as evidenced by widespread green fluorescence (Figure 4B) and in mPFC as evidenced by green fluorescence on membranes and processes of red, Cre-positive neurons (Figure S5C). We demonstrated robust functionality by imaging iSeroFnFR with two-photon frame-scan imaging while applying exogenous 5-HT via microinjection (Figures S5D–S5F), or by imaging iSeroSnFR with one-photon photometry and applying exogenous 5-HT via bath application (Figure S5G).

We next imaged iSeroSnFR responses to electrically triggered release in mouse brain slices using one-photon photometry. Brief electrical stimuli triggered 5-HT release, readily detected by iSeroSnFR (1 pulse: $\Delta F/F_0$ = 0.71% ± 0.12%); response amplitudes correlated with the number of electrical stimuli ($\Delta F/F_0$ = 1.1% ± 0.1% [2 pulses], 2.0% ± 0.4% [5 pulses], 2.8% ± 0.5% [10 pulses], 4.1% ± 0.7% [20 pulses], and 5.8% ± 0.9% [40 pulses], n = 7) (Figures 4C and 4D). Surprisingly, neither amplitude nor decay kinetics of fluorescence transients was affected by bath application of cocaine (10 μM), citalopram (1 μM), or

Figure 3. Sensor Kinetics

(A–F) iSeroSnFR was purified and tested in a stopped-flow apparatus, with increasing concentrations of 5-HT.

(A and B) Average traces showing full time courses for low (A) and high (B) concentrations.

(D and E) Magnification of the first 100 ms and 1 s, respectively, of the data in (A) and (B), respectively, with double exponential fits shown.

(C and F) The τ (1/rate constant) for each concentration for the slow phase (C) and the fast phase (F) was fit using the Hill equation. Shaded area denotes 95% confidence interval.

(G) Model of iSeroSnFR function showing two rate-limiting steps: isomerization between the inactive and active states, followed by binding of serotonin, for full fluorescence activation. n = 16–18 trials for each concentration.

(H–R) Primary cultured neurons (H–Q) and HEK cells (R) were exposed to 200 μM caged-5-HT (PA-N-5-HT) and uncaged using 405 nm laser stimulation.

(H) Representative image.

(I) 5-HT was uncaged as noted. Traces represent a 9-trial average for each replicate. Biological replicates = 3. There was no image acquisition during 405 nm laser stimulation.

(J and K) Data expanded from (I). Faded lines depict raw traces, dark lines represent average traces.

(L) Data from (I) was plotted and fit with using the Hill equation. Shaded area denotes 95% confidence interval.

(M–R) 5-HT was uncaged as noted. Red lines represent uncaging epochs.

(M and N) Line scans (128 × 1 pixel) at 5 kHz (M) and 250 Hz (N).

(O–R) Frame scans (128 × 128 pixel) at 40 Hz. For raw traces, see Figure S5. (Q) Red dot represents uncaging spot. Blue and green traces represent data from regions of interest outlined on the image (~2 μm and ~20 μm from the uncaged region). (R) Response of membrane-displayed iSeroSnFR and iAChSnFR0.6 to serotonin uncaging on HEK293T cells. Scale bars represent 50 μm (H) and 10 μm (Q and R).

See also Figure S5.

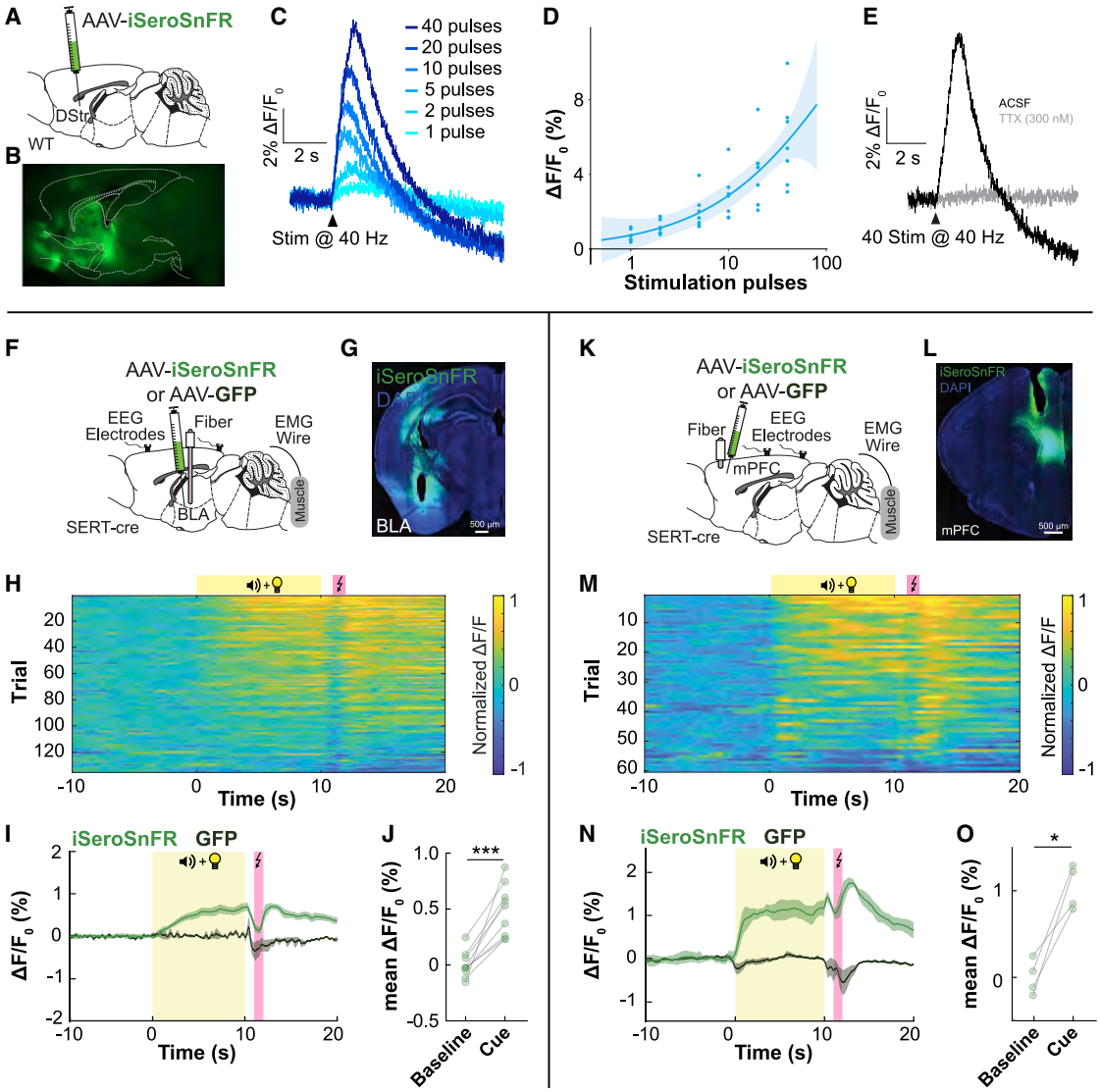


Figure 4. Detection of Electrically or Behaviorally Triggered Release of 5-HT

(A–D) WT mice were injected with AAV2/9.*CAG*.iSeroSnFR.Nlgn into the dorsal striatum (DStr). Slices were prepared (300 μm) and imaged using one-photon photometry.

(A and B) Schematic (A) and representative (B) image of sensor injection and expression.

(C) Slices were stimulated using a monopolar saline-filled glass electrode (0.5 ms; 50 μA) at the frequencies noted.

(D) Data from (C) fitted using the Hill equation. Shaded area represents 95% confidence interval.

(E) Tetrodotoxin (TTX; 300 nM) was added to the perfusion solution, and the slice was stimulated at 40 Hz for 1 s (n = 11 slices from 3 mice). For more information, see Figure S5.

(F–O) Fiber-photometry recording of 5-HT release in response to fear-conditioning in BLA (F–J) and mPFC (K–O). Mice were injected with either AAV2/9.*CAG*-iSeroSnFR.Nlgn or AAV2/5.*CAG*-GFP (as a negative control) followed by optical fiber implantation into BLA (F and G) or mPFC (K and L). Yellow box indicates unconditioned stimulus (tone + house lights); pink box illustrates foot shock. Single-trial traces (H and M) or average trace across all trials (I and N). Shaded area represents SEM. Statistical comparison was made based on the average fluorescence for the 10 s before the cue onset (baseline) and the 10 s of cue presentation (J and O). N =15 trials/animal, n = 9 $BLA^{iSeroSnFR}$, 3 BLA^{GFP}, 4 $mPFC^{iSeroSnFR}$, and 3 $mPFC^{GFP}$). *p < 0.05, ***p < 0.001, paired Student's t test.

See also Figure S5 and S6.

reserpine (1 μM) (Figure S5H). Importantly, however, tetrodotoxin (TTX, 300 nM) (Figure 4E) completely abolished electrically evoked fluorescence transients, indicating that signals are indeed action potential-dependent.

Imaging Endogenous Release of 5-HT in Freely Behaving Mice

We then examined the utility of iSeroSnFR in reporting physiologically relevant 5-HT dynamics in freely behaving mice across

multiple brain regions (Figure 4F-O). Serotonergic DRN neurons respond to both rewarding and aversive stimuli, but little is known about the dynamics of 5-HT release in regions receiving inputs from DRN (Li et al., 2016; Ren et al., 2018). Multiple lines of evidence highlight 5-HT as a critical neuromodulator of fear learning (Bocchio et al., 2016; Burghardt et al., 2007; Marcinkiewcz et al., 2016). Therefore we injected AAV2/9.*CAG*.iSeroSnFR.NIgn or AAV2/9.*CAG*.GFP into BLA (Figures 4F and 4G) and mPFC (Figures 4K and 4L), followed by fiber photometry during a fear learning paradigm. Mice were trained for 15 trials (Curzon et al., 2009), during which a cue period (10 s tone and light) was followed by a 1-s pause, and then a 1-s footshock (Figures 4H–4J and 4M–4O). We observed a rise in 5-HT during the cue period that was immediate in mPFC but slower in BLA, followed by a dip during the footshock and further rise afterward that decayed slowly in both brain regions, which was reliable across individual trials (Figures 4H and 4M). In control GFP-expressing mice, smaller deflections were associated with cue onset and foot shock, likely due to motion artifacts (Figures 4I and 4N).

We next imaged 5-HT release in response to social interaction in multisite recordings in three brain regions simultaneously: OFC, BNST, and BLA, which are suggested to mediate this behavior in different ways (Kiser et al., 2012). We injected AAV2/9.*CAG*.iSeroSnFR.NIgn into all three regions (Figures S6A–S6D) followed by multi-site fiber photometry during introduction of a same-sex intruder (Figure S6E). Wheel running has been shown to buffer behavioral consequences of stress in rodents, with concomitant changes in presynaptic 5-HT receptor and transporter expression in DRN (Greenwood et al., 2005). To assess this, we allowed mice access to a functional running wheel for 6 weeks prior to social interaction testing (Figure S6E). We found that iSeroSnFR reliably detected changes in 5-HT release during intruder presentation (Figures S6F–S6K; both in red "run" and blue "locked" trials), consistent with reports that $DRN^{5\text{-}HT}$ cells are activated by social interaction (Li et al., 2016). This increase was similar across all three regions, reaching statistical significance in OFC and BLA (Figures S6I–S6K). Free access to the wheel significantly increased 5-HT signal peak during social interaction, but only in OFC (Figures S6F–S6K; red "run" versus blue "locked"). Taken together, these data suggest that iSeroSnFR permits robust recording of 5-HT release triggered by fear conditioning and social interaction and behavioral modulation by wheel-running in multiple brain regions.

Imaging 5-HT Dynamics in Sleep-Wakefulness Cycles

We next tested the utility of iSeroSnFR in reporting 5-HT dynamics over long time intervals in sleep-wakefulness cycles (Portas et al., 2000). Electroencephalographic (EEG) and electromyographic (EMG) electrodes were implanted for simultaneous polysomnographic (sleep) recordings (Figures 5A and S7A). iSeroSnFR tracked 5-HT fluctuations across sleep-wake cycles over prolonged times, correlating well with EEG/EMG signals (Figures 5B and S7B). When variable durations of sleep-wake states were normalized, iSeroSnFR signal amplitude was highest in wakefulness, decreased during non-rapid eye movement (NREM, "slow-wave") and was lowest during rapid eye movement (REM, "paradoxical") sleep (Figures 5C and S7C). We then examined 5-HT dynamics during state transitions. Fluorescence signals rose sharply at waking, peaked during wakefulness, and decreased immediately after sleep onset, further decreasing during REM sleep. The largest signal increase occurred during the REM-to-wakefulness transition, and the largest decrease was from non-rapid eye movement (NREM)-to-REM (Figures 5D and S7D), clearly visible in individual trials. In control mice expressing GFP, no behavioral state-dependent changes in fluorescence were observed. Similar patterns were observed in both mPFC and BLA, however, in BLA, we observed more prominent fluorescence decreases during wakefulness-to-NREM transitions.

SERT Function and Pharmacology in Cell Culture

Given the clinical significance of drugs affecting 5-HT signaling and transport, we developed an iSeroSnFR-based assay to functionally characterize hSERT and drugs modulating it. We previously developed the oscillating stimulus transporter assay (OSTA) using intracellularly expressed fluorescent sensors for glutamate (Keller and Looger, 2016) and glucose (Keller et al., 2019). Because iSeroSnFR is intracellularly targetable, and does not respond to most clinical drugs, it is ideal for use with OSTA to assess hSERT-mediated transport (Figure 6A).

We first confirmed reported ionic requirements for hSERT function (reviewed in Rudnick and Sandtner, 2019). Extracellular sodium (10 or 160 mM) and chloride (10 or 160 mM) were both required for 5-HT uptake, while high extracellular potassium (150 mM) in the absence of sodium and 5-HT drove efflux, without dependence on chloride (Figures 6B and 6C). We next measured hSERT's K_m for sodium in 5-HT uptake by subjecting cells to a smoothly varying ramp of external sodium concentration, interleaved with high sodium (150 mM) bouts, verified using the dye sulforhodamine 101 (200 mM). (Figure 6D). Uptake rates in individual cells were plotted as a function of $[Na^+]$ and fitted using the Hill equation, yielding an average K_m of 6.5 ± 0.1 mM and a Hill coefficient of 1.9 ± 0.1 (Figures 6E and 6F), similar to published values (Koldsø et al., 2011; Sneddon, 1969). These results demonstrate that intracellularly expressed iSeroSnFR can precisely quantify transporter biophysical properties.

We next used this technique to investigate drugs targeting hSERT. Drugs were added to both influx and efflux buffers to ensure consistent exposure. As expected, 5-HT transport was reduced by the competitive inhibitor cocaine (10 μM) (Figure 6G). Similarly, the SSRI escitalopram rapidly diminished 5-HT transport by >50% at 500 nM and abolished it at 1 μM (Figures 6H and 6I). The tricyclic antidepressant clomipramine (100 nM) also inhibited 5-HT transport, but with slower on- and off-rates than escitalopram (Wang et al., 2013) (Figure 6J). A newer antidepressant, vilazodone, caused transport function to rapidly disappear, and despite washouts of >20 minutes, did not return to baseline (Figure 6K). Because vilazodone's reported K_i (0.5 nM) (Heinrich et al., 2004) is comparable to that of clomipramine (0.3 nM) (Tatsumi et al., 1997), this result was unexpected and requires further investigation. We next measured longitudinal effects of stimulation of the protein kinase C pathway through addition of phorbol 12-myristate 13-acetate (PMA), which decreases 5-HT transport (400 nM) (Bermingham and Blakely, 2016). PMA induced a slow decrement in transport rate over ~40 minutes, consistent with its reported effects (Figure 6L)

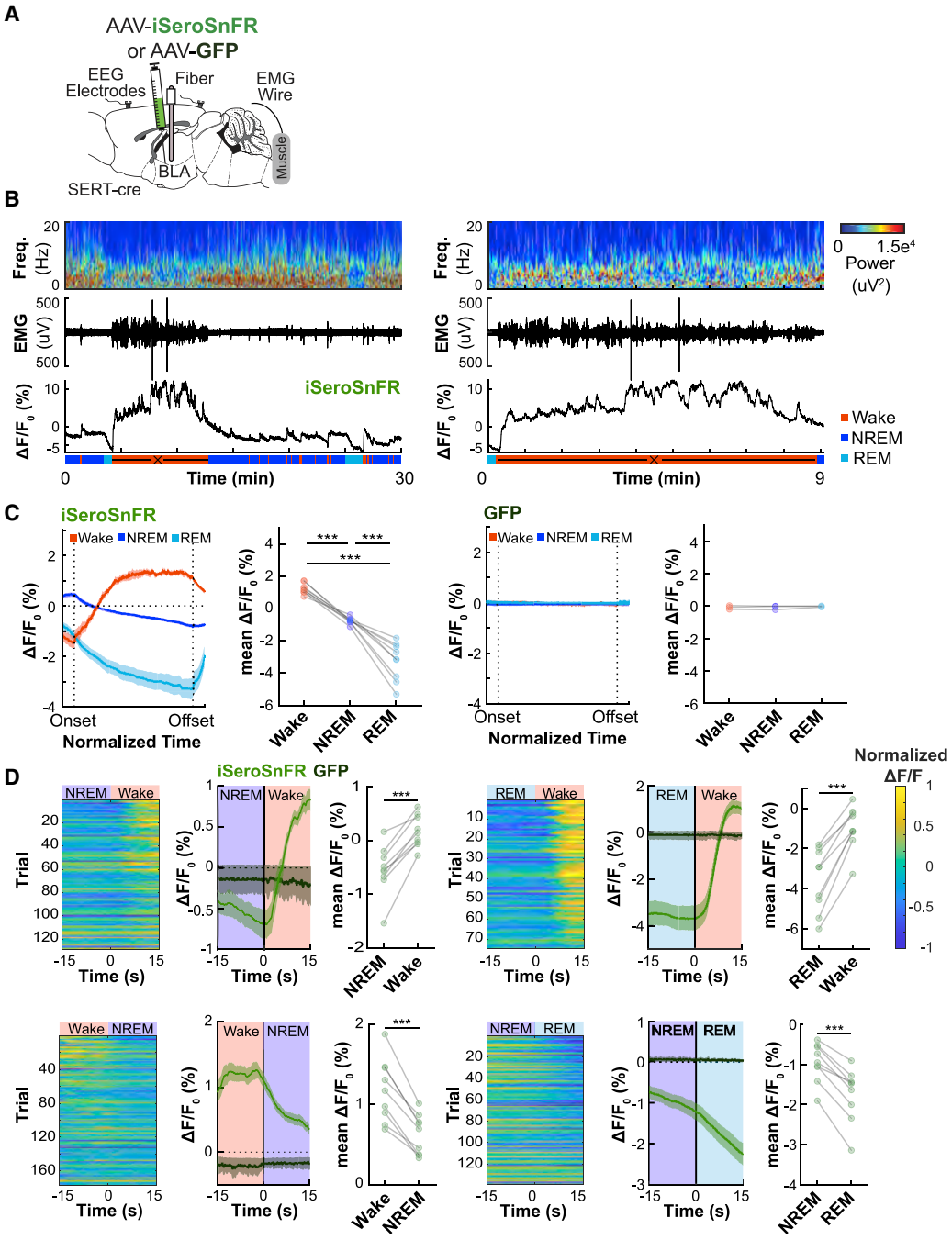


Figure 5. Detection of 5-HT Release during Sleep-Wakefulness Cycles in BLA

(A) Mice were injected with either AAV2/9.*CAG*-iSeroSnFR.Nlgn or AAV2/5.*CAG*-GFP (as a negative control), and an optical fiber was implanted into basolateral amygdala (BLA). EEG screw electrodes and EMG wires were implanted to classify sleep-wake states.

(B) Representative BLA^{iSeroSnFR} EEG spectrograms, EMG, and fiber photometry traces over time across sleep-wake cycles (left) and walking episode (right).

(C) Temporal dynamics of iSeroSnFR (left) and GFP (right) activity during waking, NREM, and REM episodes (data from B, left) across time, normalized from onset to offset. Statistical comparisons of fluorescence levels were performed on the last 10% of data within each behavioral state (one-way ANOVA with Bonferroni correction).

(D) Single-trial or averaged fluorescence change across all trials of iSeroSnFR from NREM to wake, REM to wake, wake to NREM, and NREM to REM transitions. Statistical comparisons of changes in fluorescence using and BLA^{iSeroSnFR} were made based on the average fluorescence over 15 s before and 15 s after the behavioral state transition. N = 15 trials/animal, n = 3 BLA^{GFP}, and n = 9 BLA^{iSeroSnFR}. Data represent mean ± SEM. ***$p < 0.001$, paired Student's t test.

See also Figure S7.

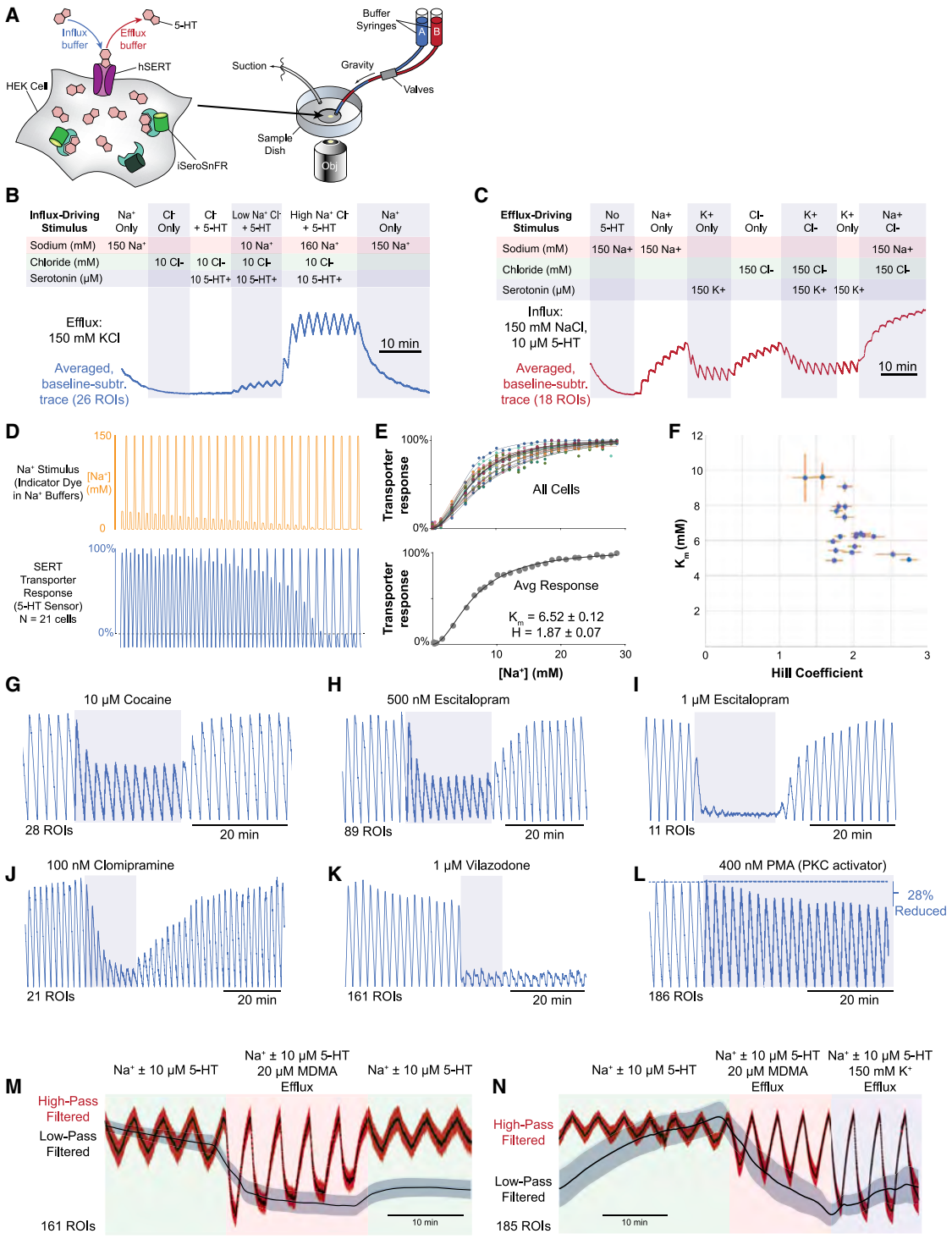


Figure 6. Application to Monitoring hSERT Transport

(A) Schematic of the oscillating stimulus transporter assay (OSTA).

(B) Confirmation of ionic requirements of hSERT for 5-HT uptake: influx only occurs in the presence of 5-HT, Na^+, and Cl^-.

(C) Ions driving hSERT-mediated 5-HT efflux: rates of efflux increase with K^+ and are insensitive to Cl^-.

(D–F) Sodium dependence of 5-HT transport through hSERT.

(D) Cells were subjected to a decreasing linear gradient of Na^+ in the influx buffer, with standardized "fiducial" bouts at saturating sodium concentration interleaved. Red fluorescence (sulforhodamine 101 at 200 nM in fiducials) was used as a readout for the sodium stimulus (top panel), and green fluorescence (iSeroSnFR) reflected changes in cytosolic [5-HT] due to hSERT function (bottom panel).

(legend continued on next page)

(Bermingham and Blakely, 2016). Since the decrement was relatively slow, it is likely that it represents extraction of transporters from cell membranes, also consistent with literature (Bermingham and Blakely, 2016).

Finally, we examined pharmacologically driven efflux of 5-HT, which has been proposed to underlie the psychoactive effects of psychedelics such as 3,4-methyl enedioxy methamphetamine (MDMA, "ecstasy"; acting mainly via SERT) (Rudnick and Sandtner, 2019). We omitted KCl from the efflux buffer, so oscillations were instead driven solely by presence or absence of 5-HT (0 or 10 μM). 20 μM MDMA was then added to the 5-HT-free buffer to probe its capacity for driving 5-HT efflux. We observed a 2-fold enhancement in efflux in the presence of 20 μM MDMA (Figure 6M). The ratio of efflux magnitude was about 1:2:3 for Na^+-, MDMA-, and K^+-driven effluxes, respectively (Figure 6N). The smaller magnitude of efflux driven by 5-HT removal alone results from the ability of extracellular Na^+ to stabilize hSERT in an outward-facing conformation, thereby slowing down reverse transport (Zhang et al., 2016). MDMA-driven efflux probably reflects a different process in which hSERT catalyzes exchange of intracellular 5-HT with extracellular MDMA, without utilizing the entire transport cycle (Rudnick and Wall, 1992). As a whole, these experiments show that iSeroSnFR coupled with OSTA is a powerful tool for rapid, precise pharmacological characterization.

DISCUSSION

Use of Machine Learning to Accelerate Directed Evolution

Our combined Rosetta and machine-learning-guided directed evolution approach was quite effective. After just one round of each, we screened fewer than 2,600 variants, but made dramatic improvements to the sensor's affinity, specificity, and fluorescence response. After just two more rounds of machine-learning-guided directed evolution, we had screened a total of ~16,000 variants, interrogated more than 60 different protein scaffold positions, and introduced 19 mutations into our final sensor, increasing its 5-HT affinity by >5,000-fold, abolishing choline/ACh binding, and increasing fluorescence response by 3-fold compared to the starting scaffold, iAChSnFR0.6. The efficacy of the design cycle diminished somewhat with each round. However, these results validate the rankings of the initial RF model, which predicted high impacts for the top 4 positions and progressively lower impacts for each successive position. Nonetheless, in each round, the best variant came from the GLM-predicted library; in rounds 1 and 2, these variants were nearly twice as good as the best-performing variant from the SSM library. Because RF uses an ensemble method, we could identify the most important features (protein positions) and provide guidance for subsequent stages. GLM is a good classifier even for small (<10,000) training datasets and is ideal for first-pass data analysis (Yang et al., 2019a). Our entire dataset was used for training (no test data), because our models were used only to guide one round of library design at a time.

Headroom remains in the iSeroSnFR scaffold, particularly for higher-affinity versions for better detection of sparse signaling events. It is likely that the machine-learning-guided approach detailed here can produce further gains, although each round thus far produced diminishing returns. A high-resolution structure of ligand-bound iSeroSnFR could reinvigorate this process, but we have been as yet unable to obtain such a structure. Alternatively, the addition of more biophysical parameters to the model, or more advanced ML models such as universal transformers (Dehghani et al., 2019), Bayesian optimization (Yang et al., 2019b), or neural networks (Kato et al., 2005), could extract sequence/function relationships that we missed. On a related note, it will be broadly useful for the field to somehow incorporate ML-gleaned insights back into the biophysical potential functions underlying structure-based computational protein design. Regardless, the method described here was highly effective for our purposes, and we expect that it can be used to engineer other PBP-based neuromodulator sensors, and more generally to tackle other challenging protein-engineering tasks.

Molecular Properties of iSeroSnFR

iSeroSnFR is highly selective for 5-HT over other endogenous signaling molecules and drugs (Figures 2C and S3). The physiological concentration of 5-HT has been estimated using FSCV to be in the high-nM to low-μM range (Bunin and Wightman, 1998; Bunin et al., 1998). The functional affinity of iSeroSnFR is in the high-μM range, much larger than endogenous monoamine receptors. However, its very large fluorescence response $(\Delta F/F_0)_{max}$, which is 3- to 30-fold larger than that of GPCR-based sensors (Jing et al., 2019; Mizuno et al., 2019; Wan et al., 2020), means that low concentrations are still readily detectible. On mammalian cells, iSeroSnFR reliably detected concentrations as low as ~350 pM (Figures 2F and S4B). Indeed, this high dynamic range and weaker affinity offers the additional advantages of fast kinetics (Helassa et al., 2018; Marvin et al., 2018) and low 5-HT buffering potential. Furthermore, the bio-orthogonal nature of the sensor reduces concerns of interference with endogenous receptors due to sensor overexpression. We have shown that the kinetics of iSeroSnFR are in the low millisecond range, both by stopped-flow (Figures 3D–3F) and by uncaging (Figures 3I–3K). We calculated both slow and fast components of the on-kinetics in purified protein, also observed with iGABASnFR and iGluSnFR (Marvin et al., 2019). Further investigation is needed to determine the mechanism underlying the apparently two-component on-kinetics. Nonetheless, we were able to demonstrate robust *in vivo* detection of millisecond-level release events. We believe that two-component on-kinetics could potentially serve as an advantage, in that the sensor could rapidly report lower concentration events, and then slowly read out larger changes in baseline occurring over several seconds. A higher-affinity sensor would quickly max out after low

(E) Fits of single-cell (top) and grouped (bottom) 5-HT transport responses to the stimulus shown in (D).
(F) Scatterplot of fitted parameters for individual cells: Hill coefficient versus K_m.
(G–L) hSERT-mediated 5-HT transport responses to various pharmacological agents (as indicated) under a standardized stimulus.
(M and N) MDMA-mediated 5-HT efflux: MDMA at 20 μM in the efflux buffer significantly increased the rate of 5-HT efflux compared to Na^+ alone. (N) In the last epoch of the experiment, K^+ was substituted for Na^+ in the efflux buffer for comparison to other experiments.

concentration events, and miss larger changes in baseline. However, the sensor is not without its limitations, and low concentration events might still be missed. Future engineering efforts will focus on producing a complementary sensor with higher affinity such that both high and low concentration events can be detected.

Imaging Endogenous 5-HT Release *Ex Vivo* and *In Vivo*

With iSeroSnFR, we replicated several salient biological results using photometry, with much higher temporal resolution than previously possible. We detected electrically evoked release, which was blocked by TTX. Surprisingly, we could not perturb observed fluorescent transients using blockers for SERT such as cocaine or citalopram, nor by the VMAT inhibitor reserpine. We expected cocaine and citalopram to prolong the 5-HT transients measured in slice by inhibiting 5-HT reuptake, as we showed *in vitro* (Figure 6G). Given the fast off-kinetics (Figures 3K and 3L), it is unlikely that sensor unbinding dominates the response waveform and obscures drug effects. Alternatively, in the brain slice preparation, released 5-HT potentially diffuses away, with SERT playing a minimal role in observed waveforms. Although possible, this is unlikely, given previous results using FSCV (Matsui and Alvarez, 2018). However, other groups have seen similar negative results, e.g., the inability of the SERT blocker fluoxetine to affect electrical stimulation-evoked 5-HT release amplitude and kinetics in the mouse dorsal raphe nucleus (Bunin and Wightman, 1998). It is also unclear why inhibition of VMAT with reserpine failed to block the transients by depleting 5-HT storage in synaptic vesicles. It is possible that reserpine only slowly depletes 5-HT from vesicles in brain slice. We also noticed that the apparent affinity in slice is right-shifted from that in dissociated neuronal culture. This is not unexpected, given the added confounds of tissue penetration for bath applied 5-HT and dilution of micropipetted 5-HT in the slice environment, which do not exist in monolayered cultured cells. Reassuringly, the successful *in vivo* detection of endogenously released 5-HT suggests that this reflects an artifact of the slice preparation in this context and not a true weakening of affinity following long-term expression. Future experiments will focus on direct comparison with FSCV and microdialysis to further characterize the properties—and deficiencies—of the sensor.

Our *in vivo* studies very clearly show functionality of the sensor across multiple brain regions and multiple behaviors, across both short and long time windows. iSeroSnFR signals are consistent with 5-HT microdialysis studies in fear conditioning in both BLA and mPFC (Ferrés-Coy et al., 2013; Forster et al., 2006; Fujino et al., 2002; Hashimoto et al., 1999; Kawahara et al., 1993; Rueter and Jacobs, 1996; Yokoyama et al., 2005), but offer much improved temporal resolution, indicating that iSeroSnFR is a viable alternative to these techniques. iSeroSnFR revealed obvious 5-HT increases in mPFC and BLA during cued fear conditioning, consistent with previous studies (Bauer, 2015). OFC responds robustly to social interaction (Figure S6I), and it is clear that serotonergic input to frontal cortex and other brain regions is complex and heterogeneous. The high sensitivity and millisecond-level temporal resolution of iSeroSnFR will be instrumental in working out such details. Furthermore, the causes and consequences of endogenous 5-HT release, and how closely optogenetically triggered release mimics it, remain uncertain (Correia et al., 2017; Garcia-Garcia et al., 2018; Marcinkiewcz et al., 2016; Ohmura et al., 2014; Seo et al., 2019). *In vivo* recording in target brain regions using iSeroSnFR will provide new opportunities to study 5-HT modulation by electrical, opto/chemogenetic, or behavioral manipulations, 5-HT interaction with other neurotransmitters such as glutamate or dopamine, and 5-HT function in the context of various behavioral states.

For initial *in vivo* demonstration of iSeroSnFR, we opted for fiber photometry, a simple and widely used technique. We replicated published FSCV results as a proof-of-principle, in multiple behavioral paradigms, across multiple brain regions. This technique, however, does not demonstrate the full potential of optical sensors, because essentially all spatial resolution is lost. *In vivo* two-photon imaging is the gold standard for high-resolution tracking of neural signaling in intact animals. iSeroSnFR is completely compatible with two-photon imaging, as we have demonstrated in brain slice (Figures S5B–S5F). Its spectral and biophysical properties are similar to that of iGluSnFR and GCaMP, routinely used in *in vivo* two-photon imaging. We thus anticipate no sensor-specific challenges for *in vivo* two-photon iSeroSnFR imaging in diverse preparations.

Finally, while we have demonstrated the uses of iSeroSnFR for several neuroscience applications, the vast majority of 5-HT is produced in the gut, and indeed 5-HT receptors exist throughout the body. Studies of 5-HT in peripheral systems would be exciting and informative, and we expect that iSeroSnFR could be adapted for these purposes.

Screening 5-HT Transporter Modulators

The soluble nature of iSeroSnFR provides an additional advantage over the use of integral-membrane GPCR-based scaffolds. The excellent dynamic range of the sensor, combined with its robust cytoplasmic expression, allowed us to easily monitor time-resolved kinetics of 5-HT influx and efflux, and how transport is distinctly modulated by various drugs, including two drugs of abuse and three antidepressants. GPCR-based sensors would be incompatible with this application, because they cannot be expressed cytoplasmically, and GPCRs respond directly to many of the very drugs such an assay is designed to characterize. iSeroSnFR, on the other hand, is far more specific, allowing these critical pharmacological experiments. Given the clinical significance of hSERT and its mutants in psychological disorders (Jarrett et al., 2007; Margoob and Mushtaq, 2011; Ozaki et al., 2003; Wankerl et al., 2014), this assay will greatly facilitate serotonergic drug discovery by identifying and quantifying the underlying properties of drug-hSERT interactions, paving the way for screening patient-specific drugs and doses for SSRIs or other drugs.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2020.11.040.

ACKNOWLEDGMENTS

We would like to thank Drs. Liqun Luo (Stanford University) and Jing Ren (MRC) for their critical reading and feedback. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the NIH (P30-GM124165). The Pilatus 6M detector on the 24-ID-C beamline is funded by an NIH-ORIP HEI grant (S10-RR029205). This research used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract DE-AC02-06CH11357. This work was supported by funding to L.T. (BRAIN Initiative U01NS090604, U01NS013522, and DP2MH107056 from NIH), to E.K.U. (Mistletoe Foundation Research Fellowship), and to G.O.M. (ARCS Scholarship), as well as by the Howard Hughes Medical Institute. V.G. is a Heritage Principal Investigator supported by the Heritage Medical Research Institute, the NIH (BRAIN RF1MH117069), the Center for Molecular and Cellular Neuroscience of the Chen Institute, and the Beckman Institute for CLARITY, Optogenetics and Vector Engineering Research. V.A.A. is funded by Intramural Programs of NIAAA and NINDS ZIA-AA000421 and Innovation Award from NIH-DDIR. T.L.K. is funded by the NIH (R01AA019454, P60AA011605, and U24AA025475), and O.J.H. is funded by the NIH (5T32NS007431-20).

AUTHOR CONTRIBUTIONS

E.K.U., J.P.K., M.A., L.L.L., and L.T. conceived of and designed the study. E.K.U. designed the machine-learning method, screened and optimized sensors, and characterized them in purified protein, mammalian cells, cultured neurons, and brain slice, with significant contribution from C.D., D.A.J., and J.S. J.P.K., S.S., and G.R. designed OSTA and stopped-flow experiments, and J.P.K. performed them. M.A. and V.G. designed and performed *in vivo* fiber photometry and EEG/EMG recording in BLA and mPFC in fear learning and sleep/wake cycles. O.J.H., M.E.F., and T.L.K. designed and performed *in vivo* fiber photometry experiments in BLA, OFC, and BNST during social interaction. R.L. and V.Y.-Y. designed and performed computational Rosetta modeling. Z.Y. and J.A.P. provided luciferase experimental data for establishing machine-learning methods. J.C. and D.T.L. provided significant insight for the machine-learning methods. J.S. characterized the sensor in acute slice using two-photon imaging. A.M. and V.A.A. designed and performed photometry imaging in acute slice. S.H. and A.J.F. performed crystallography. J.S.M., P.M.B., A.V.S., H.A.L., and L.L.L. provided iAChSnFR0.6 and performed preliminary experiments on serotonin binding. S.B. and L.D.L. synthesized caged serotonin. G.O.M. provided dissociated neuronal cultures. L.P.C. and D.E.O. produced chemical reagents. S.M.U. and S.G.A. provided SSRIs and guidance in cell-assay design. E.K.U., J.P.K., L.L.L., and L.T. wrote the manuscript with significant input from other authors.

DECLARATION OF INTERESTS

L.T. and G.O.M. are co-founders of Seven Biosciences. D.E.O. is a founder of Delix.

Received: November 23, 2019
Revised: June 22, 2020
Accepted: November 20, 2020
Published: December 16, 2020

REFERENCES

Abdalla, A., Atcherley, C.W., Pathirathna, P., Samaranayake, S., Qiang, B., Peña, E., Morgan, S.L., Heien, M.L., and Hashemi, P. (2017). In Vivo Ambient Serotonin Measurements at Carbon-Fiber Microelectrodes. Anal. Chem. *89*, 9703–9711.

Baker, D. (2019). What has de novo protein design taught us about protein folding and biophysics? Protein Sci. *28*, 678–683.

Bang, S.J., Jensen, P., Dymecki, S.M., and Commons, K.G. (2012). Projections and interconnections of genetically defined serotonin neurons in mice. Eur. J. Neurosci. *35*, 85–96.

Bauer, E.P. (2015). Serotonin in fear conditioning processes. Behav. Brain Res. *277*, 68–77.

Bedbrook, C.N., Yang, K.K., Robinson, J.E., Gradinaru, V., and Arnold, F.H. (2019). Machine learning-guided channelrhodopsin engineering enables minimally-invasive optogenetics. bioRxiv. https://doi.org/10.1101/565606.

Belmaker, R.H., and Agam, G. (2008). Major depressive disorder. N. Engl. J. Med. *358*, 55–68.

Belmer, A., Klenowski, P.M., Patkar, O.L., and Bartlett, S.E. (2017). Mapping the connectivity of serotonin transporter immunoreactive axons to excitatory and inhibitory neurochemical synapses in the mouse limbic brain. Brain Struct. Funct. *222*, 1297–1314.

Bender, B.J., Cisneros, A., 3rd, Duran, A.M., Finn, J.A., Fu, D., Lokits, A.D., Mueller, B.K., Sangha, A.K., Sauer, M.F., Sevy, A.M., et al. (2016). Protocols for Molecular Modeling with Rosetta3 and RosettaScripts. Biochemistry *55*, 4748–4763.

Berger, M., Gray, J.A., and Roth, B.L. (2009). The expanded biology of serotonin. Annu. Rev. Med. *60*, 355–366.

Bermingham, D.P., and Blakely, R.D. (2016). Kinase-dependent Regulation of Monoamine Neurotransmitter Transporters. Pharmacol. Rev. *68*, 888–953.

Bessette, P.H., Mena, M.A., Nguyen, A.W., and Daugherty, P.S. (2003). Construction of designed protein libraries using gene assembly mutagenesis. Methods Mol. Biol. *231*, 29–37.

Biswas, S., Kuznetsov, G., Ogden, P.J., Conway, N.J., Adams, R.P., and Church, G.M. (2018). Toward machine-guided design of proteins. bioRxiv. https://doi.org/10.1101/337154.

Bocchio, M., McHugh, S.B., Bannerman, D.M., Sharp, T., and Capogna, M. (2016). Serotonin, Amygdala and Fear: Assembling the Puzzle. Front. Neural Circuits *10*, 24.

Borden, P.M., Zhang, P., Shivange, A.V., Marvin, J.S., Cichon, J., Dan, C., Podgorski, K., Figueiredo, A., Novak, O., Tanimoto, M., et al. (2020). A fast genetically encoded fluorescent sensor for faithful in vivo acetylcholine detection in mice, fish, worms and flies. bioRxiv. https://doi.org/10.1101/2020.02.07.939504.

Bos, N., Dreier, S., Jørgensen, C.G., Nielsen, J., Guerrieri, F.J., and d'Ettorre, P. (2012). Learning and perceptual similarity among cuticular hydrocarbons in ants. J. Insect Physiol. *58*, 138–146.

Breiman, L. (2001). Breiman and Cutler's Random Forests for Classification and Regression. Mach. Learn. *45*, 5–32.

Broussard, G.J., Liang, R., and Tian, L. (2014). Monitoring activity in neural circuits with genetically encoded indicators. Front. Mol. Neurosci. *7*, 97.

Broussard, G.J., Unger, E.K., Liang, R., McGrew, B.P., and Tian, L. (2018). Imaging glutamate with genetically encoded fluorescent sensors. In Biochemical Approaches for Glutamatergic Neurotransmission, S. Parrot and L. Denoroy, eds. (Springer New York), pp. 117–153.

Bryksin, A.V., and Matsumura, I. (2010). Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. Biotechniques *48*, 463–465.

Bunin, M.A., and Wightman, R.M. (1998). Quantitative evaluation of 5-hydroxytryptamine (serotonin) neuronal release and uptake: an investigation of extrasynaptic transmission. J. Neurosci. *18*, 4854–4860.

Bunin, M.A., Prioleau, C., Mailman, R.B., and Wightman, R.M. (1998). Release and uptake rates of 5-hydroxytryptamine in the dorsal raphe and substantia nigra reticulata of the rat brain. J. Neurochem. *70*, 1077–1087.

Burghardt, N.S., Bush, D.E.A., McEwen, B.S., and LeDoux, J.E. (2007). Acute selective serotonin reuptake inhibitors increase conditioned fear expression: blockade with a 5-HT(2C) receptor antagonist. Biol. Psychiatry *62*, 1111–1118.

Calhoon, G.G., and Tye, K.M. (2015). Resolving the neural circuits of anxiety. Nat. Neurosci. *18*, 1394–1404.

Charnay, Y., and Léger, L. (2010). Brain serotonergic circuitries. Dialogues Clin. Neurosci. *12*, 471–487.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Cho, J.R., Treweek, J.B., Robinson, J.E., Xiao, C., Bremner, L.R., Greenbaum, A., and Gradinaru, V. (2017). Dorsal Raphe Dopamine Neurons Modulate Arousal and Promote Wakefulness by Salient Stimuli. Neuron *94*, 1205–1219.

Cipriani, A., Furukawa, T.A., Salanti, G., Chaimani, A., Atkinson, L.Z., Ogawa, Y., Leucht, S., Ruhe, H.G., Turner, E.H., Higgins, J.P.T., et al. (2018). Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. Lancet *391*, 1357–1366.

Clarke, R.J., and Khalid, M.A.A. (2015). Pumps, Channels and Transporters: Methods of Functional Analysis (John Wiley & Sons).

Cobb, R.E., Chao, R., and Zhao, H. (2013). Directed Evolution: Past, Present and Future. AIChE J. *59*, 1432–1440.

Corre, J., van Zessen, R., and Pascoli, V. (2018). Dopamine neurons projecting to medial shell of the nucleus accumbens drive heroin reinforcement. eLife *7*, e39945.

Correia, P.A., Lottem, E., Banerjee, D., Machado, A.S., Carey, M.R., and Mainen, Z.F. (2017). Transient inhibition and long-term facilitation of locomotion by phasic optogenetic activation of serotonin neurons. eLife *6*, e20975.

Curzon, P., Rustay, N.R., and Browman, K.E. (2009). Cued and contextual fear conditioning for rodents. In Methods of Behavior Analysis in Neuroscience, J.J. Buccafusco, ed. (CRC Press/Taylor & Francis).

Davis, I.W., and Baker, D. (2009). RosettaLigand docking with full ligand and receptor flexibility. J. Mol. Biol. *385*, 381–392.

Dehghani, M., Gouws, S., Vinyals, O., Uszkoreit, J., and Kaiser, L. (2019). Universal Transformers (ICLR).

Dobson, A.J., Barnett, A.G., and Barnett, A.G. (2008). An Introduction to Generalized Linear Models (Chapman and Hall/CRC).

Ehren, J., Govindarajan, S., Morón, B., Minshull, J., and Khosla, C. (2008). Protein engineering of improved prolyl endopeptidases for celiac sprue therapy. Protein Eng. Des. Sel. *21*, 699–707.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? Acta Crystallogr. D Biol. Crystallogr. *69*, 1204–1214.

Feng, J., Zhang, C., Lischinsky, J.E., Jing, M., Zhou, J., Wang, H., Zhang, Y., Dong, A., Wu, Z., Wu, H., et al. (2019). A Genetically Encoded Fluorescent Sensor for Rapid and Specific In Vivo Detection of Norepinephrine. Neuron *102*, 745–761.

Ferrés-Coy, A., Pilar-Cuellar, F., Vidal, R., Paz, V., Masana, M., Cortés, R., Carmona, M.C., Campa, L., Pazos, A., Montefeltro, A., et al. (2013). RNAi-mediated serotonin transporter suppression rapidly increases serotonergic neurotransmission and hippocampal neurogenesis. Transl. Psychiatry *3*, e211.

Forster, G.L., Feng, N., Watt, M.J., Korzan, W.J., Mouw, N.J., Summers, C.H., and Renner, K.J. (2006). Corticotropin-releasing factor in the dorsal raphe elicits temporally distinct serotonergic responses in the limbic system in relation to fear behavior. Neuroscience *141*, 1047–1055.

Fujino, K., Yoshitake, T., Inoue, O., Ibii, N., Kehr, J., Ishida, J., Nohta, H., and Yamaguchi, M. (2002). Increased serotonin release in mice frontal cortex and hippocampus induced by acute physiological stressors. Neurosci. Lett. *320*, 91–95.

Garcia-Garcia, A.L., Canetta, S., Stujenske, J.M., Burghardt, N.S., Ansorge, M.S., Dranovsky, A., and Leonardo, E.D. (2018). Serotonin inputs to the dorsal BNST modulate anxiety in a 5-HT_{1A} receptor-dependent manner. Mol. Psychiatry *23*, 1990–1997.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Gong, S., Zheng, C., Doughty, M.L., Losos, K., Didkovsky, N., Schambra, U.B., Nowak, N.J., Joyner, A., Leblanc, G., Hatten, M.E., et al. (2003). A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. Nature *425*, 917–925.

Greenwood, B.N., Foley, T.E., Day, H.E.W., Burhans, D., Brooks, L., Campeau, S., and Fleshner, M. (2005). Wheel running alters serotonin (5-HT) transporter, 5-HT1A, 5-HT1B, and alpha 1b-adrenergic receptor mRNA in the rat raphe nuclei. Biol. Psychiatry *57*, 559–568.

Hashimoto, S., Inoue, T., and Koyama, T. (1999). Effects of conditioned fear stress on serotonin neurotransmission and freezing behavior in rats. Eur. J. Pharmacol. *378*, 23–30.

Hawkins, P.C.D., Skillman, A.G., Warren, G.L., Ellingson, B.A., and Stahl, M.T. (2010). Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. J. Chem. Inf. Model. *50*, 572–584.

Heinrich, T., Böttcher, H., Gericke, R., Bartoszyk, G.D., Anzali, S., Seyfried, C.A., Greiner, H.E., and Van Amsterdam, C. (2004). Synthesis and structure–activity relationship in a class of indolebutylpiperazines as dual 5-HT(1A) receptor agonists and serotonin reuptake inhibitors. J. Med. Chem. *47*, 4684–4692.

Helassa, N., Dürst, C.D., Coates, C., Kerruth, S., Arif, U., Schulze, C., Wiegert, J.S., Geeves, M., Oertner, T.G., and Török, K. (2018). Ultrafast glutamate sensors resolve high-frequency release at Schaffer collateral synapses. Proc. Natl. Acad. Sci. USA *115*, 5594–5599.

Huang, R., Fang, P., and Kay, B.K. (2012). Improvements to the Kunkel mutagenesis protocol for constructing primary and secondary phage-display libraries. Methods *58*, 10–17.

Jaquins-Gerstl, A., and Michael, A.C. (2015). A review of the effects of FSCV and microdialysis measurements on dopamine release in the surrounding tissue. Analyst (Lond.) *140*, 3696–3708.

Jarrett, M.E., Kohen, R., Cain, K.C., Burr, R.L., Poppe, A., Navaja, G.P., and Heitkemper, M.M. (2007). Relationship of SERT polymorphisms to depressive and anxiety symptoms in irritable bowel syndrome. Biol. Res. Nurs. *9*, 161–169.

Jing, M., Zhang, Y., Wang, H., and Li, Y. (2019). G-protein-coupled receptor-based sensors for imaging neurochemicals with high sensitivity and specificity. J. Neurochem. *151*, 279–288.

Jing, M., Li, Y., Zeng, J., Huang, P., Skirzewski, M., Kljakic, O., Peng, W., Qian, T., Tan, K., Zou, J., et al. (2020). An optimized acetylcholine sensor for monitoring in vivo cholinergic activity. Nat. Methods *17*, 1139–1146.

Jones, K.A., Porterfield, W.B., Rathbun, C.M., McCutcheon, D.C., Paley, M.A., and Prescher, J.A. (2017). Orthogonal Luciferase-Luciferin Pairs for Bioluminescence Imaging. J. Am. Chem. Soc. *139*, 2351–2358.

Kabsch, W. (2010). XDS. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Kato, R., Nakano, H., Konishi, H., Kato, K., Koga, Y., Yamane, T., Kobayashi, T., and Honda, H. (2005). Novel strategy for protein exploration: high-throughput screening assisted with fuzzy neural network. J. Mol. Biol. *351*, 683–692.

Kawahara, H., Yoshida, M., Yokoo, H., Nishi, M., and Tanaka, M. (1993). Psychological stress increases serotonin release in the rat amygdala and prefrontal cortex assessed by in vivo microdialysis. Neurosci. Lett. *162*, 81–84.

Keller, J.P., and Looger, L.L. (2016). The Oscillating Stimulus Transporter Assay, OSTA: Quantitative Functional Imaging of Transporter Protein Activity in Time and Frequency Domains. Mol. Cell *64*, 199–212.

Keller, J.P., Marvin, J.S., Lacin, H., Lemon, W.C., Shea, J., Kim, S., Lee, R.T., Koyama, M., Keller, P.J., and Looger, L.L. (2019). In vivo glucose imaging in multiple model organisms with an engineered single-wavelength sensor. bioRxiv. https://doi.org/10.1101/571422.

Khoury, G.A., Smadbeck, J., Kieslich, C.A., and Floudas, C.A. (2014). Protein folding and de novo protein design for biotechnological applications. Trends Biotechnol. *32*, 99–109.

Kiser, D., Steemers, B., Branchi, I., and Homberg, J.R. (2012). The reciprocal interaction between serotonin and social behaviour. Neurosci Biobehav Rev *36*, 786–798.

Koldsø, H., Noer, P., Grouleff, J., Autzen, H.E., Sinning, S., and Schiøtt, B. (2011). Unbiased simulations reveal the inward-facing conformation of the human serotonin transporter and Na(+) ion release. PLoS Comput. Biol. *7*, e1002246.

Kunkel, T.A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA *82*, 488–492.

Li, Y., Zhong, W., Wang, D., Feng, Q., Liu, Z., Zhou, J., Jia, C., Hu, F., Zeng, J., Guo, Q., et al. (2016). Serotonin neurons in the dorsal raphe nucleus encode reward signals. Nat. Commun. *7*, 10503.

Liao, J., Warmuth, M.K., Govindarajan, S., Ness, J.E., Wang, R.P., Gustafsson, C., and Minshull, J. (2007). Engineering proteinase K using machine learning and synthetic genes. BMC Biotechnol. *7*, 16.

Liaw, A., and Wiener, M. (2002). Classification and Regression by randomForest. https://rdrr.io/cran/randomForest/.

Lin, M.Z., and Schnitzer, M.J. (2016). Genetically encoded indicators of neuronal activity. Nat. Neurosci. *19*, 1142–1153.

Lobas, M.A., Tao, R., Nagai, J., Kronschläger, M.T., Borden, P.M., Marvin, J.S., Looger, L.L., and Khakh, B.S. (2019). A genetically encoded single-wavelength sensor for imaging cytosolic and cell surface ATP. Nat. Commun. *10*, 711.

Looger, L.L., and Griesbeck, O. (2012). Genetically encoded neural activity indicators. Curr. Opin. Neurobiol. *22*, 18–23.

Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., et al. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. *13*, 133–140.

Madisen, L., Garner, A.R., Shimaoka, D., Chuong, A.S., Klapoetke, N.C., Li, L., van der Bourg, A., Niino, Y., Egolf, L., Monetti, C., et al. (2015). Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. Neuron *85*, 942–958.

Marcinkiewcz, C.A., Mazzone, C.M., D'Agostino, G., Halladay, L.R., Hardaway, J.A., DiBerto, J.F., Navarro, M., Burnham, N., Cristiano, C., Dorrier, C.E., et al. (2016). Serotonin engages an anxiety and fear-promoting circuit in the extended amygdala. Nature *537*, 97–101.

Margoob, M.A., and Mushtaq, D. (2011). Serotonin transporter gene polymorphism and psychiatric disorders: is there a link? Indian J. Psychiatry *53*, 289–299.

Marvin, J.S., Schreiter, E.R., Echevarría, I.M., and Looger, L.L. (2011). A genetically encoded, high-signal-to-noise maltose sensor. Proteins *79*, 3025–3036.

Marvin, J.S., Borghuis, B.G., Tian, L., Cichon, J., Harnett, M.T., Akerboom, J., Gordus, A., Renninger, S.L., Chen, T.-W., Bargmann, C.I., et al. (2013). An optimized fluorescent probe for visualizing glutamate neurotransmission. Nat. Methods *10*, 162–170.

Marvin, J.S., Scholl, B., Wilson, D.E., Podgorski, K., Kazemipour, A., Müller, J.A., Schoch, S., Quiroz, F.J.U., Rebola, N., Bao, H., et al. (2018). Stability, affinity, and chromatic variants of the glutamate sensor iGluSnFR. Nat. Methods *15*, 936–939.

Marvin, J.S., Shimoda, Y., Magloire, V., Leite, M., Kawashima, T., Jensen, T.P., Kolb, I., Knott, E.L., Novak, O., Podgorski, K., et al. (2019). A genetically encoded fluorescent sensor for in vivo imaging of GABA. Nat. Methods *16*, 763–770.

Matsui, A., and Alvarez, V.A. (2018). Cocaine Inhibition of Synaptic Transmission in the Ventral Pallidum Is Pathway-Specific and Mediated by Serotonin. Cell Rep. *23*, 3852–3863.

McCoy, A.J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr. D Biol. Crystallogr. *63*, 32–41.

Miller, D.A., Suen, G., Bruce, D., Copeland, A., Cheng, J.-F., Detter, C., Goodwin, L.A., Han, C.S., Hauser, L.J., Land, M.L., et al. (2011). Complete genome sequence of the cellulose-degrading bacterium Cellulosilyticum lentocellum. J. Bacteriol. *193*, 2357–2358.

Mizuno, G.O., Unger, E.K., and Tian, L. (2019). Real-time monitoring of neuromodulators in behaving animals using genetically-encoded indicators. In Compendium on In Vivo Monitoring in Real-Time Molecular Neuroscience, G.S. Wilson and A.C. Michael, eds. (World Scientific), pp. 1–18.

Nov, Y. (2012). When second best is good enough: another probabilistic look at saturation mutagenesis. Appl. Environ. Microbiol. *78*, 258–262.

Ohmura, Y., Tanaka, K.F., Tsunematsu, T., Yamanaka, A., and Yoshioka, M. (2014). Optogenetic activation of serotonergic neurons enhances anxiety-like behaviour in mice. Int. J. Neuropsychopharmacol. *17*, 1777–1783.

Ollikainen, N., de Jong, R.M., and Kortemme, T. (2015). Coupling Protein Side-Chain and Backbone Flexibility Improves the Re-design of Protein-Ligand Specificity. PLoS Comput. Biol. *11*, e1004335.

Ozaki, N., Goldman, D., Kaye, W.H., Plotnicov, K., Greenberg, B.D., Lappalainen, J., Rudnick, G., and Murphy, D.L. (2003). Serotonin transporter missense mutation associated with a complex neuropsychiatric phenotype. Mol. Psychiatry *8*, 933–936.

Packer, M.S., and Liu, D.R. (2015). Methods for the directed evolution of proteins. Nat. Rev. Genet. *16*, 379–394.

Patriarchi, T., Cho, J.R., Merten, K., Howe, M.W., Marley, A., Xiong, W.-H., Folk, R.W., Broussard, G.J., Liang, R., Jang, M.J., et al. (2018). Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. Science *360*, eaat4422.

Peñalva, R.G., Lancel, M., Flachskamm, C., Reul, J.M.H.M., Holsboer, F., and Linthorst, A.C.E. (2003). Effect of sleep and sleep deprivation on serotonergic neurotransmission in the hippocampus: a combined in vivo microdialysis/EEG study in rats. Eur. J. Neurosci. *17*, 1896–1906.

Portas, C.M., Bjorvatn, B., and Ursin, R. (2000). Serotonin and the sleep/wake cycle: special emphasis on microdialysis studies. Prog. Neurobiol. *60*, 13–35.

Quan, J., and Tian, J. (2011). Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries. Nat. Protoc. *6*, 242–251.

Rathbun, C.M., Porterfield, W.B., Jones, K.A., Sagoe, M.J., Reyes, M.R., Hua, C.T., and Prescher, J.A. (2017). Parallel Screening for Rapid Identification of Orthogonal Bioluminescent Tools. ACS Cent. Sci. *3*, 1254–1261.

Ren, J., Friedmann, D., Xiong, J., Liu, C.D., Ferguson, B.R., Weerakkody, T., DeLoach, K.E., Ran, C., Pun, A., Sun, Y., et al. (2018). Anatomically Defined and Functionally Distinct Dorsal Raphe Serotonin Sub-systems. Cell *175*, 472–487.

Richter, F., Leaver-Fay, A., Khare, S.D., Bjelic, S., and Baker, D. (2011). De novo enzyme design using Rosetta3. PLoS ONE *6*, e19230.

Ritz, C., Baty, F., Streibig, J.C., and Gerhard, D. (2015). Dose-Response Analysis Using R. PLoS ONE *10*, e0146021.

Rouillard, J.-M., Lee, W., Truan, G., Gao, X., Zhou, X., and Gulari, E. (2004). Gene2Oligo: oligonucleotide design for in vitro gene synthesis. Nucleic Acids Res. *32*, W176-80.

Rudnick, G., and Sandtner, W. (2019). Serotonin transport in the 21st century. J. Gen. Physiol. *151*, 1248–1264.

Rudnick, G., and Wall, S.C. (1992). p-Chloroamphetamine induces serotonin release through serotonin transporters. Biochemistry *31*, 6710–6718.

Rueter, L.E., and Jacobs, B.L. (1996). A microdialysis examination of serotonin release in the rat forebrain induced by behavioral/environmental manipulations. Brain Res. *739*, 57–69.

Saito, T., and Nakatsuji, N. (2001). Efficient gene transfer into the embryonic mouse brain using in vivo electroporation. Dev. Biol. *240*, 237–246.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Schultz, K.N., and Kennedy, R.T. (2008). Time-resolved microdialysis for in vivo neurochemical measurements and other applications. Annu. Rev. Anal. Chem. (Palo Alto, Calif.) *1*, 627–661.

Seo, C., Guru, A., Jin, M., Ito, B., Sleezer, B.J., Ho, Y.-Y., Wang, E., Boada, C., Krupa, N.A., Kullakanda, D.S., et al. (2019). Intense threat switches dorsal raphe serotonin neurons to a paradoxical operational mode. Science *363*, 538–542.

Shivange, A.V., Borden, P.M., Muthusamy, A.K., Nichols, A.L., Bera, K., Bao, H., Bishara, I., Jeon, J., Mulcahy, M.J., Cohen, B., et al. (2019). Determining the pharmacokinetics of nicotinic drugs in the endoplasmic reticulum using biosensors. J. Gen. Physiol. *151*, 738–757.

Sneddon, J.M. (1969). Sodium-dependent accumulation of 5-hydroxytryptamine by rat blood platelets. Br. J. Pharmacol. *37*, 680–688.

Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. Protein Expr. Purif. *41*, 207–234.

Sun, F., Zeng, J., Jing, M., Zhou, J., Feng, J., Owen, S.F., Luo, Y., Li, F., Wang, H., Yamaguchi, T., et al. (2018). A Genetically Encoded Fluorescent Sensor Enables Rapid and Specific Detection of Dopamine in Flies, Fish, and Mice. Cell *174*, 481–496.

Tatsumi, M., Groshan, K., Blakely, R.D., and Richelson, E. (1997). Pharmacological profile of antidepressants and related compounds at human monoamine transporters. Eur. J. Pharmacol. *340*, 249–258.

Taylor, N.D., Garruss, A.S., Moretti, R., Chan, S., Arbing, M.A., Cascio, D., Rogers, J.K., Isaacs, F.J., Kosuri, S., Baker, D., et al. (2016). Engineering an allosteric transcription factor to respond to new ligands. Nat. Methods *13*, 177–183.

R Development Core Team (2013). R: A language and environment for statistical computing (R Foundation for Statistical Computing).

Tinberg, C.E., Khare, S.D., Dou, J., Doyle, L., Nelson, J.W., Schena, A., Jankowski, W., Kalodimos, C.G., Johnsson, K., Stoddard, B.L., and Baker, D. (2013). Computational design of ligand-binding proteins with high affinity and selectivity. Nature *501*, 212–216.

Wan, J., Peng, W., Li, X., Qian, T., Song, K., Zeng, J., Deng, F., Hao, S., Feng, J., Zhang, P., et al. (2020). A genetically encoded GRAB sensor for measuring serotonin dynamics in vivo. bioRxiv. https://doi.org/10.1101/2020.02.24.962282.

Wang, H., Goehring, A., Wang, K.H., Penmatsa, A., Ressler, R., and Gouaux, E. (2013). Structural basis for action by diverse antidepressants on biogenic amine transporters. Nature *503*, 141–145.

Wankerl, M., Miller, R., Kirschbaum, C., Hennig, J., Stalder, T., and Alexander, N. (2014). Effects of genetic and early environmental risk factors for depression on serotonin transporter expression and methylation profiles. Transl. Psychiatry *4*, e402.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer).

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the Tidyverse. J. Open Source Softw. *4*, 1686.

Williams, G. (2011). Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery (Springer-Verlag).

Wu, Z., Kan, S.B.J., Lewis, R.D., Wittmann, B.J., and Arnold, F.H. (2019). Machine learning-assisted directed protein evolution with combinatorial libraries. Proc. Natl. Acad. Sci. USA *116*, 8852–8858.

Yang, K.K., Wu, Z., and Arnold, F.H. (2019a). Machine-learning-guided directed evolution for protein engineering. Nat. Methods *16*, 687–694.

Yang, K.K., Chen, Y., Lee, A., and Yue, Y. (2019b). Batched Stochastic Bayesian Optimization via Combinatorial Constraints Design. arXiv, arXiv:1904.08102.

Yokoyama, M., Suzuki, E., Sato, T., Maruta, S., Watanabe, S., and Miyaoka, H. (2005). Amygdalic levels of dopamine and serotonin rise upon exposure to conditioned fear stress without elevation of glutamate. Neurosci. Lett. *379*, 37–41.

Zhang, Y.-W., Turk, B.E., and Rudnick, G. (2016). Control of serotonin transporter phosphorylation by conformational state. Proc. Natl. Acad. Sci. USA *113*, E2776–E2783.

Zhou, X., Lu, J.Y.-F., Darling, R.D., Simpson, K.L., Zhu, X., Wang, F., Yu, L., Sun, X., Merzenich, M.M., and Lin, R.C.S. (2015). Behavioral training reverses global cortical network dysfunction induced by perinatal antidepressant exposure. Proc. Natl. Acad. Sci. USA *112*, 2233–2238.

Zhuang, X., Masson, J., Gingrich, J.A., Rayport, S., and Hen, R. (2005). Targeted gene expression in dopamine and serotonin neurons of the mouse brain. J. Neurosci. Methods *143*, 27–32.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Chicken polyclonal anti-GFP	Aves Labs	GFP-1020; RRID: AB_2307313
goat anti-chicken IgY H&L Alexa Fluor 488	Abcam	RRID: AB_150169
Bacterial and Virus Strains		
Top10	Fisher	C404003
BL21 (DE3)	NEB	C25271
LOBSTR-BL21 (DE3)	Kerafast	EC1001
CJ236	NEB	E4141
AAV-*CAG*-iSeroSnFR-NIgn	In house	N/A
AAV-*CAG*.FLEx.iSeroSnFR-PDGFR	In house	N/A
AAV-*CAG*.FLEx.iSeroSnFRNull-PDGFR	In house	N/A
AAV1-*CMV*-PI-Cre-rBG	Penn Vector Core	AV-1-PV1090
AAV5-*CMV*-HI-eGFP-Cre.WPRE.SV40	Penn Vector Core	AV-5-PV2004
Biological Samples		
Hippocampal dissociated neuronal culture	In house	N/A
Chemicals, Peptides, and Recombinant Proteins		
Serotonin	Enzo	ALX-550-328-G001
Glutamate	Sigma	G1251
GABA	Sigma	A5835
Choline	Fisher	AC110290010
Acetylcholine	Sigma	A2661-100G
Tyrosine	Sigma	T3754-50G
L-DOPA	Sigma	378850
Dopamine	Sigma	H8502-25G
Norepinephrine	Sigma	74480
Epinephrine	Sigma	E4642-5G
L-Tryptophan	Sigma	T0254-25G
5-Hydroxytryptophan	Sigma	H9772-250MG
Tryptamine	Sigma	246557-5G
5-Hydroxy indole acetic acid	Sigma	H8876-500MG
L-Phenylalanine	Sigma	P2126-100G
Histamine	Sigma	H7250-10G
L-Leucine	Sigma	E811-25G
L-Serine	Sigma	S4250-5G
Adenosine	Sigma	A9251
Tyramine	Sigma	T2879-1G
Octopamine	Sigma	O0250-1G
Melatonin	Calbiochem	444300
Gramine	Sigma	G10806-25G
Indole	Sigma	W259306-SAMPLE-K
Amphetamine	Sigma	A5880-1G
Bufotenine	Gift from DEO	N/A
Trimethyltryptamine	Gift from DEO	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Dimethyltryptamine	Gift from DEO	N/A
L-741,626	Tocris	1003
Fluoxetine	Tocris	0927
Fluoxamine	Tocris	1033
Sertraline	Tocris	2395
Citalopram (racemic)	Tocris	5763
Escitalopram	Tocris	4796
Duloxetine	Tocris	4798
Amoxapine	Sigma	A129-100MG
Bupropion	Tocris	2831
Clozapine	Tocris	0444
Clozapine-N-Oxide	Tocris	6329
6-Nitroqupazine	Sigma	Q109-100MG
Venlafaxine	Tocris	2917
Desvenlafaxine	Sigma	D2069-50MG
PA-N caged serotonin	Made by LDLand SB	N/A
NPEC caged serotonin	Tocris	3991
MDMA	Gift from DEO	N/A
TCEP	Sigma	C4706-2G
Jeffamine	Hampton	HR2-597
Citalopram	Abcam	ab120133
Tetrodotoxin	Tocris	1078
Buprenorphine	MWI	O56163
Carprofen	Henry Schein	024751
Ketamine	Henry Schein	010177
Xylazine	Henry Schein	033198
Critical Commercial Assays		
DNA miniprep kit	QIAGEN	27104
PCR purification kit	QIAGEN	28104
ZymoPURE Plasmid Miniprep Kit	Zymo	D4208T
MCSG-2 crystallization screen	Anatrance	MCSG-2
Endo-free Plasmid Maxi kit	QIAGEN	12362
jetPrime	Polypus	114-75
AAVance concentration reagent	SBI	AAV100A-1
Effectine	QIAGEN	301425
Deposited Data		
pRSET-iSeroSnFR	This paper	Addgene 128483
pMinDisplay-iSeroSnFR	This paper	Addgene 128484
pAAV-CAG-iSeroSnFR-Nlgn	This paper	Addgene 128485
pAAV-CAG.FLEx.iSeroSnFR	This paper	Addgene 128486
pMinDisplay-iSeroSnFR-EnhancedExport	This paper	Addgene 128941
pAAV-CAG.FLEx.iSeroSnFR-EnhancedExport	This paper	Addgene 129180
iSeroSnFR open conformation crystal structure	This paper	PDB: 6PER
Sequence of iSeroSnFR	This paper	Genebank:MN083270

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Cell Lines		
HEK293T	ATCC	1573
AAV293	Agilent	240073
Experimental Models: Organisms/Strains		
Timed-pregnant rats	Charles River	N/A
C57BL/6J	Jackson Laboratory	000664
SERT-Cre	MMRRC	017260-UCD RRID: MMRRC_017260-UCD
SERT-Cre	Jackson Laboratory	Zhuang et al., 2005; RRID: IMSR_JAX:014554
Ai14	Jackson Laboratory	Madisen et al., 2010; 007914
Recombinant DNA		
pRSET-A	Invitrogen	V35120
pMinDisplay	Invitrogen	V66020
pAAV	Addgene	98932
Software and Algorithms		
Rosetta	Rosetta Commons	https://www.rosettacommons.org/
liganddock	N/A	https://www.rosettacommons.org/manuals/rosetta3_user_guide/app_ligand_docking.html
Backrub	N/A	https://www.rosettacommons.org/manuals/archive/rosetta3.4_user_guide/d1/d09/backrub.html
R (RStudio)	R-project	https://www.r-project.org/ - https://rstudio.com/
randomForest	N/A	https://www.rdocumentation.org/packages/randomForest/versions/4.6-14/topics/randomForest
Glm	N/A	https://www.rdocumentation.org/packages/stats/versions/3.5.3/topics/glm
MATLAB R2017b	Mathworks	RRID: SCR_001622
EEGLAB plug-in	Swartz Center for Comp. Neuro.	https://sccn.ucsd.edu/eeglab/index.php RRID: SCR_007292
SleepSign 3.0	Kissei Comtec	http://www.sleepsign.com
Ethovision 14	Noldus	https://www.noldus.com/ethovision-xt/new
Video Freeze Software	Med Associates	https://www.med-associates.com/product/video-fear-conditioning/
ScanImage (SI 2016b)	Vidrio Technologies	http://scanimage.vidriotechnologies.com/display/SIH/ScanImage+Home
OMEGA	OpenEye Scientific	https://www.eyesopen.com/
CCP4	Collaborative Computational Project No. 4	https://www.ccp4.ac.uk
Living Image	PerkinElmer	https://www.perkinelmer.com/product/li-software-for-spectrum-1-seat-add-on-128113
OmniPlex System	Plexon	https://plexon.com/products/omniplex-d-neural-data-acquisition-system-1/
ABET II	Lafayette Instruments	https://lafayetteneuroscience.com/products/abetii-operant-ctrl-software

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
pCLAMP10	Molecular Devices	https://mdc.custhelp.com/app/answers/detail/a_id/18779/~/axon%E2%84%A2pclamp%E2%84%A2-10-electrophysiology-data-acquisition-%26-analysis-software-download
AxoGraph X	N/A	https://axograph.com/
Other		
Mantis	Formulatrix	N/A
Fluorescent Platereader	Tecan	Infinite M200Pro
IVIS Lumina	Xenogen	N/A
Nanodrop	Thermo	2000
Small animal Stereotax	David Kopf Instruments	*In vivo* fiber photometry experiments 1900
2-photon microscope	Scientifica	N/A
Ti:Sapphire laser	Coherent	Chameleon ultra
Axopatch-200B amplifier	Molecular Devices	200B
Vibrating Microtome	Leica	1200
Mosquito	TTP Labtech	HTS

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed and will be fulfilled by lead contact, Lin Tian, lintian@ucdavis.edu

Materials Availability

The following plasmids have been made available through Addgene:

128483	pRSET-iSeroSnFR
128484	pMinDisplay-iSeroSnFR
128485	pAAV-CAG.iSeroSnFR-NIgn
128486	pAAV-CAG.FLEx.iSeroSnFR-PDGFR
128941	pMinDisplay-iSeroSnFR-EnhancedExport
129180	pAAV-CAG.FLEx.iSeroSnFR-PDGFR-EnhancedExport

AAV viruses are available through the Canadian Neurophotonics Centre.

Data and Code Availability

The full sequence of iSeroSnFR, and has been deposited in GenBank:MN083270. The crystal structure has been deposited in the Protein Data Bank:6PER). Code has been provided directly in this STAR Methods section. Code updates will be provided at https://github.com/ekunger/iSeroSnFR.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice and Rats

The experimental procedures performed on animals followed the guidelines of the Animal Care and Use Committees at the University of California Davis, the California Institute of Technology, the University of North Carolina, Chapel Hill, and the National Institute on Drug Abuse. Mice and rats were housed in a barrier facility with a 12:12 hr light:dark cycle, and food and water were available *ad libitum*.

Animals were group housed by sex wherever possible. *Rosa*$^{Ai14/Ai14}$ (Madisen et al., 2010) mice were maintained as homozygous breeding pairs or crossed to wild-type C57BL/6J (Jax # 000664). *SERT*$^{Cre/+}$ (Zhuang et al., 2005) mice were bred as male *SERT*$^{Cre/+}$ x female wild-type C57BL/6J (Jax # 000664) as recommended by the depositing lab. Both male and female animals were used in experiments and were between the age of 2 and 9 months. For *in vivo* mouse experiments, *SERT*-Cre (Zhuang et al., 2005; (Gong et al., 2003)) were crossed to wild-type C57BL/6J (Jax # 000664). Both wild-type mice and mice heterozygous for the Cre transgene were used for experiments. Following surgery, subjects were singly housed for at least five days in a residence room for full recovery before animals were moved to a behavioral room. Animals were habituated to the new light/dark cycle and behavioral room for at least 2 weeks before experiments were started. Rats were ordered as timed-pregnant females and sacrificed shortly after arrival in the facility.

METHOD DETAILS

Methods specific to luciferase

Development of a machine learning-guided library design paradigm (methods specific to luciferase)

Before pursuing the ambitious goal of drastically shifting the specificity of a choline-binding protein to 5-HT, we first established our machine learning approach on a more modestly modified ligand/binding pocket pair. We selected firefly luciferase, an enzyme that has been engineered for multicomponent imaging (Jones et al., 2017). We synthesized a brominated analog (4'Br-Luc) of the cognate substrate -luciferin (-Luc), and mutated the enzyme active site to accommodate the substitution (Figure S1A).

We first chose random forest (RF) (Liaw and Wiener, 2002) because it is tolerant of missing data points and capable of extracting information from relatively small training datasets, which are a practicality of most protein mutagenesis studies. Importantly, a strength of RF is "feature importance" detection, giving us the ability to prioritize protein positions for SSM (Figure 1A). Second, we chose generalized linear modeling (GLM, Gaussian regression) (Dobson et al., 2008) because it is a better classifier, while still tolerating small training datasets, and its linear nature provides precise, interpretable predictions about the effects of individual mutations that we could easily translate into new variants and combinations thereof (Figure 1A).

To build the models, we used sequence information from 222 published and unpublished luciferase variants as our input variables, and the experimentally determined preference for -Luc versus 4'Br-Luc as the output variable. Each position within the protein was an independent variable with a categorical distribution (amino acid), and the preference for -Luc versus 4'Br-Luc (fold-preference as calculated by the amplitude of photon output in the presence of -Luc versus 4'Br-Luc) was our target with a continuous numerical distribution. We then compared the importance of each position using the RF model, and the contribution of each mutation using the GLM (identity link), followed by experimental validation of these predictions. We found that RF was effective at predicting positions, and GLM made correct predictions, but only at positions highly ranked by RF. We therefore decided that for future sensor design, we would use the two models in a sequential fashion: first we would use the RF to rank the positions for SSM, and then we would analyze the results of our SSM libraries with the GLM to identify the best individual mutations at each position.

The top residues predicted by the RF model were positions 218, 347, 316 and 247, which were frequently mutated to accommodate other 4'-substituted analogs in a previous screen (Jones et al., 2017). The GLM predicted S347G, L342A, S347A, Y255N, R218A, F247Y, and G316S as the top mutations (see table below). To experimentally validate these predictions, we created a small library covering both high-ranked and lowranked positions predicted by RF, all of which had very statistically significant GLM predictions of improved 4'Br–Luc specificity (Figure 7B). We tested 276 variants from this library for their photon output with D-Luc and 4'Br–Luc and sequenced the top 10% favoring 4'Br–Luc (~1000x increased binding compared to the starting variant), as well as the bottom 10% favoring D-Luc (Figure 7C). The number of mutations per variant was not significantly different between the two populations, indicating a similar overall mutation rate (Figure 7D). However, the incidence of specific mutations was different between the two populations, and when compared to the input data (Figure 7E). For example, at position 347, which was highly ranked by both RF and GLM, the incidence of the glycine mutation compared to the native serine was 26:1 (G:S) in the top 10% of variants but 5:19 in the bottom ones (Figure 7E). Conversely, at position 342, which was not ranked highly by RF, a leucine-to-alanine mutation was predicted by the GLM to heavily bias the preference toward 4'Br–Luc. This mutation appeared more frequently in the bottom variants (L:A, 18:9 – top 10% versus 3:21 – bottom 10%, Figure 7E), demonstrating a correct prediction by the RF model, and an incorrect prediction for the GLM. In general, the RF more linearly predicted the preference (pseudo-R^2 = 0.45, Figure 7F), than GLM (pseudo-R^2 = 0.10, Figure 7G).

Construction of luciferase library and mutants

DNA inserts for the luciferase shuffle library and point mutants were generated as described by Jones et al. (2017). Two sections of the luciferase gene (*pgl4-luc2*), denoted R1 and R2, were targeted for gene assembly. The R1 region comprises amino acids 199-275 and was assembled with primers R1-F0 to R1-F235 (forward primers, Table M1) and R1-R0 to R1-R119 (reverse primers). The R2 region comprises amino acids 275-347 and was assembled with primers R2-F0 to R2-F264 (forward, Table M2) and R2-R0 to R2-R228 (reverse). The gene assembly primers were designed using gene2oligo (Rouillard et al., 2004) to identify desirable T_m values. To assemble the genes encoding point mutants, primers containing the codon(s) of interest were used in place of the primers coding

for the wild-type sequence. For the shuffle library inserts, mutagenic primers were mixed with wild-type primers in a 1:1 molar ratio. Libraries were assembled as described by Bessette et al. (2003). All PCR reactions were performed using Q5 High-Fidelity DNA polymerase.

Library DNA was inserted into linearized template vector pET28-R1del-mRFP as described by Rathbun et al. (2017). Library inserts were assembled with the linearized pET vectors using Gibson assembly (Gibson et al., 2009). Gibson assembly master mixes were prepared following the Prather recipe on https://www.openwetware.org/wiki/Gibson_Assembly, with all materials purchased from New England BioLabs. For each assembly, 50 ng of *Dpn*I digested, linearized vector was combined with insert (5:1 insert:vector ratio) and added to 10 mL of master mix. The mixtures were incubated at 50°C for 60 min, then 2 mL was transformed into chemically competent cells (TOP10 *E.* coli, Thermo-Fisher). Transformants were plated on square, agar plates containing kanamycin. Cells were plated to exceed 3X the library size. Cells were scrapped of the plates, combined, and pelleted. DNA was isolated using ZymoPURE Plasmid Miniprep Kit and saved for screening.

Primer lists

All primers were purchased from Integrated DNA Technologies, Inc. (San Diego, CA) and are written in the 5′→3′ direction. Upper case letters denote bases coding for the luciferase gene. Lower case letters denote bases added to ensure similar melting temperatures (T_m) for all primers. Bases highlighted in red denote sites targeted for mutagenesis.

Forward Primers	
R1-F0	GGATCCACCGGATTGCCCAAGGGCGTAGCCCTACC
R1-F35	GCACCGCACCGCTTGTGTCCGATTCAGTCATGCCC
R1-F70	GCGACCCCATCTTCGGCAACCAGATCATCCCCGACA
R1-F106	CCGCTATCCTCAGCGTGGTGCCATTTCACCACGGC
R1-F141	TTCGGCATGTTCACCACGCTGGGCTACTTGATCTGCG
R1-F178	GCTTTCGGGTCGTGCTCATGTACCGCTTCGAGGAG
R1-F213	GAGCTATTCTTGCGAAGCTTcggccctgtacagtcgctg
R1-F235	cagcgactgtacagggc
Reverse Primers	
R1-R0	GGGCAATCCGGTGGATCC
R1-R18	ACAAGCGGTGCGGTGCGGTAGGGCTACGCCCTT
R1-R51	GCCGAAGATGGGGTCGCGGGCATGACTGAATCGGAC
R1-R87	CCACGCTGAGGATAGCGGTGTCGGGGATGATCTGGTT
R1-R124	CGTGGTGAACATGCCGAAGCCGTGGTGAAATGGCA
R1-R159	AGCACGACCCGAAAGCCGCAGATCAAGTAGCCCAG
R1-R194	cgAAGCTTCGCAAGAATAGCTCCTCCTCGAAGCGGTACATG

Region 1 (R1, wild-type *luc2* primers). Lower case letters denote non-luciferase sequences included for maintaining proper T_m and do not appear in the final assembled gene.

Forward Primers	
R2-F0	acacatCGAGGAGGAGCTATTCTTGCGAAGCTTGCA
R2-F36	AGACTATAAGATTCAATCTGCCCTGCTGGTGCCCACAC
R2-F74	TATTTAGCTTCTTCGCTAAGAGCACTCTCATCGACAAGTACGAC
R2-F117	CTAAGCAACTTGCACGAGATCGCCAGCGGCGGG
R2-F150	GCGCCGCTCAGCAAGGAGGTAGGTGAGGCC
R2-F181	GTGGCCAAACGCTTCCACCTACCAGGCATCCG
R2-F213	CCAGGGCTACGGCCTGACAGAAACAACTAGTGCCA
R2-F248	TTCTGATCACCCCCgtcgcctccggaggaga
R2-F264	tctcctccggaggcg
Reverse Primers	
R2-R0	TAGCTCCTCCTCGatgtgt
R2-R19	GGGCAGATTGAATCTTATAGTCTTGCAAGCTTCGCAAGAA
R2-R59	GCTCTTAGCGAAGAAGCTAAATAGTGTGGGCACCAGCA
R2-R97	TCGTGCAAGTTGCTTAGGTCGTACTTGTCGATGAGAGT
R2-R135	TTGCTGAGCGGCGCCCCGCCGCTGGCGATC
R2-R165	GGAAGCGTTTGGCCACGGCCTCACCTACCTCC
R2-R197	AGGCCGTAGCCCTGGCGGATGCCTGGTAGGT
R2-R228	acGGGGGTGATCAGAATGGCACTAGTTGTTTCTGTC

Region 2 (R2, wild-type *luc2* primers). Lower case letters denote non-luciferase sequences included for maintaining proper T_m and do not appear in the final assembled gene.

Forward Primers	
R1-F35_A	GCACCGCACCGCTTGTGTCGCCTTCAGTCATGCCC
R1-F141_Y	TACGGCATGTTCACCACGCTGGGCTACTTGATCTGCG
R1-F141_N	TTCGGCATGTTCACCACGCTGGGCAATTTGATCTGCG
R1-F141_Y+N	TACGGCATGTTCACCACGCTGGGCAATTTGATCTGCG
R2-F117_S	CTAAGCAACTTGCACGAGATCGCCAGCGGCTCT
R2-F213_A	CCAGGGCTACGGCGCCACAGAAACAACTAGTGCCA
R2-F213_G	CCAGGGCTACGGCCTGACAGAAACAACTGGGGCCA
R2-F213_A+G	CCAGGGCTACGGCGCCACAGAAACAACTGGGGCCA
Reverse Primers	
R1-R51_A	GCCGAAGATGGGGTCGCGGGCATGACTGAAGGCGAC
R1-R124_Y	CGTGGTGAACATGCCGTAGCCGTGGTGAAATGGCA
R1-R159_N	AGCACGACCCGAAAGCCGCAGATCAAATTGCCCAG
R2-R135_S	TTGCTGAGCGGCGCAGAGCCGCTGGCGATC
R2-R197_A	GCGCCGTAGCCCTGGCGGATGCCTGGTAGGT
R2-R228_G	acGGGGGTGATCAGAATGGCACTAGTTGTTTCTGTG

Primers used to construct luciferase point mutants and shuffle library. The bases highlighted in red denote sites targeted for mutagenesis.

R1 insert amplifaction primers	
R1-insert-Fwd	GACAAAACCATCGCCCTGATCATGAACAGTAGT GGATCCACCGGATTGCCCAA
R1-insert-Rev	CACCAGCAGGGCAGATTGAATCTTATAGTCTTGCAAGCTTCGCAAGA ATAGCTCCTCCTC
R2 insert amplifaction primers	
R2-insert-Fwd	CGAGGAGGAGCTATTCTTGCG
R2-insert-Rev	CTTCGGGGGTGATCAGAAT
Vector amplifcation primers	
R1-vector-Rev	GCAAGACTATAAGATTCAATCTGCCCTGCTGGTG
R2-vector-Fwd	ATTCTGATCACCCCCGAAG

Primers used to amplify inserts for the R1 or R2 regions, along with the pET vector backbone.

General bioluminescence imaging protocol

All analyses were performed with bacterial cell lysates in black 96-well plates (Greiner Bio One). Plates containing luminescent reagents were imaged in a light-proof chamber with an IVIS Lumina (Xenogen) CCD camera chilled to −90°C. The stage was kept at 37°C during the imaging session, and the camera was controlled using Living Image software. Exposure times were set to 60 s, and data binning levels were set to medium. Regions of interests were selected for quantification, and total flux values were analyzed using Living Image software. All data were exported to Microsoft Excel for further analyses.

Lysate screening of combination library

The shuffle library was screened following the protocol previously described in Jones et al. (2017), with modifications. Library DNA was transformed into chemically competent BL21-DE3 *E. coli* cells. Transformants were plated on square, agar plates containing kanamycin. Colonies were picked and expanded in LB-Kan media in a 96-well deep well plate (500 mL/well). The plate was incubated at 37°C overnight. An aliquot of the overnight culture (4 mL) was used to inoculate 400 mL of auto-induction LB media (Studier, 2005), and the cells were incubated at 30°C with shaking (200 rpm) for 24 h. The remaining starter cultures were stored at 4°C and used for subsequent plasmid recovery and sequencing analysis. The cells were pelleted and lysed in lysis buffer (600 mL). Cell lysate (90 mL) was added to 96-well black plates, followed by 10X luciferin and ATP solution (10 mL, 100 mM luciferin and 1 mM ATP final concentration) in phosphate buffer (250 mM phosphate buffer, pH = 7.8). The plate was then imaged as described above.

Methods specific to iSeroSnFR

Rosetta Modeling

Rotamers for 5-HT were generated with OMEGA from OpenEye Scientific. The closed conformation PDB structure of *Thermoanaerobacter spX513* OpuBC was combined with the open structure of iAChSnFR0.4 (provided by PMB). We used rosetta_script in Rosetta 3 (2015.19.57819) for docking with protein-ligand interface design instructed by design.xml (below). Residues within 6Å of the ligand, as well as those within 8Å and facing toward the ligand, were mutated. Residues within 10Å, as well as those within 12Å and facing the ligand, were repacked. We restricted aromatic residues (Trp, His, Tyr, and Phe) to only be mutated into other aromatic residues. We designed 48,000 models and filtered them with total_score for the top 5% and then picked the top 10 models according to the interface_delta_X score. These 10 models were further re-designed with a protocol to couple protein side-chain and backbone

flexibility from the Kortemme lab (Ollikainen et al., 2015). Shape complementarity scores for the top redesigns were calculated with Sc routine in CCP4. Top designs were chosen for synthesis and characterization.

Rosetta design was performed using the following scripts:

Design.options

```
-in
  -file
  -s SRO_X513.pdb
  -extra_res_fa SRO.params
  -packing
  -ex1
  -ex2
  -linmem_ig 10
  -parser
  -protocol design.xml
  -overwrite
  -mistakes
  -restore_pre_talaris_2013_behavior true
```

Backrub.options

```
-in
  -path
  -database ./rosetta/main/database/
  -file
  -extra_res_fa SRO.params
  -packing
  -ex1
  -ex2
  -extrachi_cutoff 0
  -resfile X513_6A.res
  -out
  -nstruct 20
  -coupled_moves
  -mc_kt 0.6
  -ntrials 1000
  -initial_repack false
  -ligand_mode true
  -ligand_weight 2.0
  -run
  -ignore_zero_occupancy false
```

Backrub.bash

```
./main/source/bin/coupled_moves.linuxgccrelease -database ./main/database/ protocols.backrub.BackrubMover -extra_res_fa SRO.params -ex1 -ex2 -extrachi_cutoff 0 -nstruct 20 -coupled_moves::mc_kt 0.6 -coupled_moves::ntrials 1000 -coupled_moves::initial_repack false -coupled_moves::ligand_mode true -coupled_moves::ligand_weight 2.0 -mute all -coupled_moves::fix_backbone false -coupled_moves::bias_sampling true -coupled_moves::bump_check true -s 14_SRO_X513_0280.pdb -resfile 14_SRO_X513_0280.res
```

Design.xml

```
< ROSETTASCRIPTS >
  < SCOREFXNS >
  < ligand_soft_rep weights = ligand_soft_rep >
  < /ligand_soft_rep >
  < hard_rep weights = ligand >
  < /hard_rep >
  < /SCOREFXNS >
  < TASKOPERATIONS >
```

< DetectProteinLigandInterface name = design_interface cut1 = 6.0 cut2 = 8.0 cut3 = 10.0 cut4 = 12.0 design = 1 resfile = ”resfile.txt”/ >
< /TASKOPERATIONS >
< LIGAND_AREAS >
< inhibitor_dock_sc chain = X cutoff = 6.0 add_nbr_radius = true all_atom_mode = false/ >
< inhibitor_final_sc chain = X cutoff = 6.0 add_nbr_radius = true all_atom_mode = false/ >
< inhibitor_final_bb chain = X cutoff = 7.0 add_nbr_radius = false all_atom_mode = true alpha_restraints = 0.3/ >
< /LIGAND_AREAS >
< INTERFACE_BUILDERS >
< side_chain_for_docking ligand_areas = inhibitor_dock_sc/ >
< side_chain_for_final ligand_areas = inhibitor_final_sc/ >
< backbone ligand_areas = inhibitor_final_bb extension_window = 3/ >
< /INTERFACE_BUILDERS >
< MOVEMAP_BUILDERS >
< docking sc_interface = side_chain_for_docking minimize_water = false/ >
< final sc_interface = side_chain_for_final bb_interface = backbone minimize_water = false/ >
< /MOVEMAP_BUILDERS >
< MOVERS >
< FavorNativeResidue name = favor_native bonus = 1.2/ >
< ddG name = calculateDDG jump = 1 per_residue_ddg = 1 scorefxn = hard_rep/ >
< PackRotamersMover name = designinterface scorefxn = hard_rep task_operations = design_interface/ >
< Translate name = translate chain = X distribution = uniform angstroms = 3 cycles = 50/ >
< Rotate name = rotate chain = X distribution = uniform degrees = 360 cycles = 50/ >
< SlideTogether name = slide_together chains = X/ >
< HighResDocker name = high_res_docker cycles = 12 repack_every_Nth = 3 scorefxn = ligand_soft_rep movemap_builder = docking/ >
< FinalMinimizer name = final scorefxn = hard_rep movemap_builder = final/ >
< InterfaceScoreCalculator name = add_scores chains = X scorefxn = hard_rep/ >
< /MOVERS >
< PROTOCOLS >
< Add mover_name = translate/ >
< Add mover_name = rotate/ >
< Add mover_name = slide_together/ >
< Add mover_name = favor_native/ >
< Add mover_name = designinterface/ >
< Add mover_name = high_res_docker/ >
< Add mover_name = calculateDDG/ >
< Add mover_name = final/ >
< Add mover_name = add_scores/ >
< /PROTOCOLS >
< /ROSETTASCRIPTS >

Cloning

The bacterial expression vector pRSET-A was used for library screening, which contains a His_6-tag for purification. The mammalian expression vector pMinDisplay, which contains a myc-tag, an IgK-leader secretion tag, a Kozak sequence, and a PDGFR transmembrane domain for tethering to the membrane, was used for transfection into HEK293 cells. The viral expression vector pAAV was used for infection of HEK293 cells, primary neurons, mouse brain slices, and *in vivo* fiber photometry. All vectors were cloned with a combination of QuikChange, circular polymerase extension cloning (Quan and Tian, 2011), overlap extension cloning (Bryksin and Matsumura, 2010), digestion/ligation, or Kunkel mutagenesis (Kunkel, 1985). Due to issues arising from double-stranded break repair machinery in the bacteria, which drastically reduced the cloning efficiency into FLEx backbones, sequences inserted into pAAV containing FLEx sites were first subcloned into a custom shuttle vector, then cut with SacI and HindIII and ligated into a matching custom pAAV.

Library generation

Libraries were constructed by Kunkel mutagenesis (Kunkel, 1985). Single stranded dU-DNA was generated by transforming the plasmid to CJ236 cells and infecting with M13K07 bacteriophage. Chloramphenicol was used initially to ensure an F-pilus$^+$ host; ampicillin was used to ensure the presence of the plasmid, and kanamycin was used to ensure bacteriophage infection. Single stranded dU-DNA was harvested using a kit according to the manufacturer’s instructions. Phosphorylated primers were designed

in the reverse direction with either a specific mutation or a degenerate codon (NNK), and at least 10 bp on either side of the mutation. An additional primer was included in each reaction to destroy a cut site (either NheI or XbaI) in order to allow easy enrichment for mutated sequences (Huang et al., 2012). Primers were annealed to the sequence by raising the temp to 95°C and lowering it slowly to 25°C (2°/min) using a thermocycler. Gaps were filled in and ligated using T7 polymerase and T4 ligase. DNA was transformed to Top10 cells and grown at 37°C overnight. DNA was extracted using a miniprep kit according to the manufacturer's instructions, and digested with either NheI or XbaI. DNA was purified using a PCR cleanup kit according to the manufacturer's instructions.

Library screening

Library size was determined using the TopLib online library calculator (Nov, 2012). DNA libraries were transformed to BL21(DE3) cells, plated on LB-agar with ampicillin, and allowed to grow at 37°C overnight. Colonies were picked at random into 2.3 mL deep-well 96-well plates and grown in 1 mL of autoinduction medium (2xYT + 0.5% glycerol, 0.5% glucose, 0.2% α-lactose, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 20 mM $MgSO_4$) for 8 h at 37°C, then for 24-36 h at 18°C, shaking at 250-300 rpm. 4 wells of each plate were designated for the parent protein, which was transformed and plated at the same time as the libraries. At the end of the growth period, 50 μL was removed and stored at −80°C in 25% glycerol for subsequent DNA recovery. Cultures were then pelleted by centrifugation and washed 3x with PBS. Pellets were dried briefly, then frozen at −80°C for 15 mins, and thawed. Pellets were resuspended in 100 μL of B-PER Complete reagent and incubated shaking at 30°C for 1 h. Cell debris was pelleted by centrifugation and the supernatant was transferred to a clean plate. 96-well plates were expanded into optically clear 384-well plates containing 45 μL PBS and 4 μL lysate, then read on a fluorescent plate reader. Following the initial read, 5-HT was prepared fresh and 1 μL was added to experimental wells, and 1 μL of PBS was added to control wells, and the plate was read again. The performance of each variant was calculated as the difference between the first and second read, divided by the first read, and normalized to the PBS well, then compared to the parent wells. Any variant showing considerable improvement over the parent was subsequently retested at multiple concentrations, then regrown from the glycerol stock, DNA extracted by miniprep, and sequenced.

Protein purification

Plasmids were regrown from glycerol stock or transformed to LOBSTR-BL21(DE3) cells. For small yields, cells were grown in 5 mL of autoinduction medium at 37°C for 8 h, then 18°C for 24-36 h, shaking at 250-300 rpm. Then cultures were pelleted by centrifugation and the supernatant removed. Pellets were subjected to 3 rounds of freeze-thaw, first at −80°C, followed by 37°C waterbath. Then pellets were resuspended in 500 μL of B-PER Complete and incubated at 30°C for 1 h, shaking at 250-300 rpm. Cell debris was pelleted by centrifugation and supernatant was transferred to a 96-well deep-well plate. Protein was purified using the His-tag with prepacked NiNTA tips (IMAC20) connected to an electronic multichannel with a preset protein purification protocol. Briefly, NiNTA beads were washed with 10 mM imidazole, then the lysate was bound to the NiNTA beads, and washed 8X with 25 mM imidazole, then eluted with 250 mM imidazole.

For large yields, plasmids were first grown in 5 mL of autoinduction medium at 37°C, then expanded to 100 mL and grown at 18°C for 36-48 h, shaking at 250-300 rpm. Then cultures were pelleted by successive rounds of centrifugation followed by supernatant removal. Pellets were subjected to 3 rounds of freeze-thaw cycles as above, then resuspended in 10 mL of B-PER Complete and incubated at 30°C for 1 h, shaking at 250-300 rpm. Then NiNTA beads were washed with 10 mM imidazole in PBS and incubated with the lysate in 10 mM imidazole on a rotator at 4°C for 18 h. Beads were then washed with 25 mM imidazole in PBS 6 times (600-1000 mL total wash volume), then eluted with 250 mM imidazole in PBS. Protein was then dialyzed in 1X PBS (specificity and affinity assays) or 0.1X TBS (crystallization), and subsequently concentrated using concentrator columns (10 kDa MWCO).

Specificity screening

Purified protein was measured by nanodrop, and diluted with 1X PBS to reach a final concentration of 100 nM. 40 μL was added to a 384-well plate, and read on a fluorescent plate reader. Small molecule neurotransmitters, drugs and other compounds were dissolved in either 1X PBS, ascorbate buffer (to reduce oxidation: Na-(L) Ascorbate 137 mM, KCl 5.3 mM, $CaCl_2$ 952 μM, $MgSO_47H_2O$ 833 μM, $MgCl_26H2O$ 1 mM, Na_2PO_42H2O 423 μM, KH_2PO_4 345 μM, glucose 5.6 mM, $NaHCO_3$ 4.2 mM), 10% ethanol, or 10% DMSO. pH was carefully maintained at 7.4 for all experiments. Then 10 μL of either vehicle or compound was added to the protein and read again on the plate reader. Due to maximum solubility issues, not all compounds were tested at the same concentration.

Crystallization and data collection

Crystals were grown by hanging-drop vapor diffusion at room temperature (21°C) using iSeroSnFR (20 mg/mL, mixed with 20 mM 5-HT and 10 mM TCEP) and the commercially available crystallization screen MCSG 2. Crystals grew in condition G11, in a drop containing 200 nL protein and 200 nL reservoir solution (0.1 M HEPES:NaOH pH 7.0, 30% Jeffamine ED-2001 pH 7.0). Crystals were harvested and cryoprotected in reservoir solution containing 30% ethylene glycol before flash-cooling in liquid nitrogen. All data were collected at APS beamline 24-ID-C. Diffraction data were processed with XDS (Kabsch, 2010) and scaled with Aimless (Evans and Murshudov, 2013). iSeroSnFR crystallized in the orthorhombic space group *I*222 with unit cell parameters a = 80.67 Å, b = 99.26 Å, c = 150.56 Å. The structure of iSeroSnFR was determined by molecular replacement with PHASER (McCoy, 2007) using the GFP domain (residues 177-402) of the similarly-constructed maltose sensor (PDB: 3OSQ) and the soluble binding domain of

the BilE ABC transporter from *Listeria monocytogenes* (PDB: 4Z7E) as search models. Manual model building was done with the program Coot (Emsley and Cowtan, 2004). The resulting structure was refined with PHENIX to a final R_{factor} = 19.21% and R_{free} = 23.84%. Data collection refinement statistics are listed below.

Synchrotron (Beamline)	APS (24-ID-C)
Wavelength (Å)	0.97920
Space Group	*I*222
Unit Cell Parameters (Å)	*a* = 80.67, *b* = 99.26, *c* = 150.56
Resolution Range (Å)	85-2.10 (2.16-2.10)
No. observed reflections	107,170 (8,844)
No. unique reflections	34,796 (2,849)
Completeness (%)	98.1 (68.2)
I/σ (I)	11.4 (2.0)
R_{merge} [a] (%)	7.0 (63.8)
$CC_{1/2}$	99.8 (68.2)
Monomers per ASU	1
Matthew's Coefficient (Å^3/Da)	2.40
Solvent Content (%)	48.81
Refinement Statistics	
No. of reflections (F > 0)	34,793
R_{factor} [b] (%)	18.83
R_{free} [b] (%)	23.46
RMS bond length (Å)	0.008
RMS bond angle (°)	1.077
Coordinate Error[c] (Å)	0.21
Average B Factors	
Protein (Å^2)	43.6
Water (Å^2)	40.0
Ethylene Glycols (Å^2)	46.4
Ramachandran Plot Statistics[c]	
Favored (%)	489 (96.0%)
Allowed (%)	17 (3.4%)
Outliers (%)	3 (0.6%)
No. of atoms	
Protein	4,023
Ethylene Glycols	20 (5 molecules)
Waters	226
PDBID	**6PER**

[a] $R_{merge} = [\sum_h \sum_i |I_h - I_{hi}| / \sum_h \sum_i I_{hi}]$ where I_h is the mean of I_{hi} observations of reflection *h*. Numbers in parenthesis represent highest resolution shell.
[b] *R*-Factor and [c] $R_{free} = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$ x 100 for 95% of recorded data (*R*-Factor) or 5% data (R_{free})
[c] Ramachandran plot statistics from MolProbity (Chen et al., 2010)

Protein Sequences

Insertion of cpsfGFP is after residue 106, but residues 105-108 were mutated previously as part of the linker regions. Amino acid numbering matches normal N-to-C numbering of each domain. The first 29 amino acids of X513 OpuBC are an endogenous secretion

signal that was omitted, and the first D of X513 OpuBC corresponds to 1 throughout the rest of the manuscript as it was the first amino acid available in the initial Rosetta modeling.

IgGκ signal peptide
X513 OpuBC 30-104
Inserted linker
sfGFP 147-238 circular permutation linker sfGFP 1-146
Inserted linker
X513 OpuBC 109-305
Myc epitope
PDGFR transmembrane domain 513-561

METDTLLLWVLLLWVPGSTGDRSANDTVVVGSINHTEQIIVANMLAEMIEAHTDLKVVRKLNLGGVNVNFEAIKR
GGANNGIDIYLEYVGYGLVDILGYPEPNVYIIADKQKNGIKANFKIRYNVEDGSVQLADHYQQNTPIGDGPVLLP
DNHYLSTQSVLSKDPNEKRDHMVLLEFVTAAGITLGMDELYKGGTGGSMSKGEELFTGVVPILVELDGGVNGHKF
SVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKD
DGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFPPPGATDPEGAYETVKKEYKRKWNIVWLKPLG
FNASYVLAVKDELAKQYNLKTFSDLAKISDKLILGANMMFLENPDGYPGLQKLYNFKFKHTKSMDAGIPYTAIDN
NEVQVIDATATDGLLVSHKLKILEDDKAFFPPYYAAPIIRQDVLDKHPELKDVLNKLANQISLEEMQKLNYKRDG
EGQDPAKVAKEFLKEKGLILQVDEQKLISEEDLNAVGQDTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILI
MLWQKKPR*

Neuroligin1 signal peptide
X513 OpuBC 30-104
Inserted linker
sfGFP 147-238 circular permutation linker sfGFP 1-146
Inserted linker
X513 OpuBC 109-305
Neuroligin exrtracellular predicted transmembrane intracellular

MALPRCMWPNYVWRAMMACVVHRGSGAPLTLCLLGCLLQTFHVLSQKLEFANDTVVVGSINHTEQIIVANMLAEM
IEAHTDLKVVRKLNLGGVNVNFEAIKRGGANNGIDIYLEYVGYGLVDILGYPEPNVYIIADKQKNGIKANFKIRY
NVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRDHMVLLEFVTAAGITLGMDELYKGGTGGS
MSKGEELFTGVVPILVELDGGVNGHKFSVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYP
DHMKQHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFPPPGA
TDPEGAYETVKKEYKRKWNIVWLKPLGFNASYVLAVKDELAKQYNLKTFSDLAKISDKLILGANMMFLENPDGYP
GLQKLYNFKFKHTKSMDAGIPYTAIDNNEVQVIDATATDGLLVSHKLKILEDDKAFFPPYYAAPIIRQDVLDKHP
ELKDVLNKLANQISLEEMQKLNYKRDGEGQDPAKVAKEFLKEKGLIGSTLELVPHLHNLNDISQYTSTTTKVPST
DITLRPTRKNSTPVTSAFPTAKQDDPKQQPSPFSVDQRDYSTELSVTIAVGASLLFLNILAFAALYYKKDKRRHD
VHRRCSPQRTTTNDLTHAPEEEIMSLQMKHTDLDHECESIHPHEVVLRTACPPDYTLAMRRSPDDIPLMTPNTIT
MIPNTIPGIQPLHTFNTFTGGQNNTLPHPHPHPHSHSTTRVGGSGG*

Tissue Culture

HEK293T cells were grown in DMEM, supplemented with fetal bovine serum (FBS) and Penicillin-Streptomycin (Pen Strep). Cells were transfected with Effectene per the manufacturer's instructions. Prior to imaging, cells were washed with Hank's Balanced Salt Solution (HBSS) supplemented with 2 mM $MgCl_2$ and 2 mM $CaCl_2$. All images were collected in HBSS + Mg + Ca.

Dissociated hippocampal neuronal culture

Primary hippocampal neuronal cultures were prepared as previously described. Briefly, E18 rat pups were decapitated, and the brains were dissected into ice-cold neural dissection solution (NDS, 10 mM HEPES in HBSS, pH 7.4). Hippocampi were removed, enzymatically digested with 0.25% trypsin, washed with pre-warmed plating medium (PM, Minimal Essential Medium supplemented with 10% FBS and 100 U/mL Pen Strep) and then mechanically digested by trituration. Cells were plated on 35 mm MatTek glass bottom dishes previously coated with a mixture of 20 μM poly-L-ornithine and 50 ng/mL laminin. Cells were kept at 37°C, 5% CO_2 in PM for ~24 h and then in Neurobasal medium supplemented with 2% B27, and one third medium exchanges were performed every 3-4 days. On the seventh day *in vitro* (DIV), cells were infected with AAV2/1- or AAV2/9-*CAG*-iSeroSnFR-Nlgn or AAV2/1- or AAV2/9-*CAG*.FLEx.iSeroSnFR-PDGFR and AAV2/1-*CMV*-Cre and transferred to feeding medium (Neurobasal medium supplemented with 2% B27, 1% GlutaMAX, 1 μg/mL Gentamicin, and 10 μM FuDR).

Dose-response curves

Prior to imaging, all cultures were washed and incubated in HBSS containing 2 mM $MgCl_2$ and 2 mM $CaCl_2$ to minimize background due to fluorescent culture medium components. All data collected for the same experiment were collected with identical imaging parameters. Cultures were imaged on a laser scanning confocal microscope (Zeiss 710). Dishes were perfused using a gravity-fed inlet connecting 8 channels via a commutator and controlled by a ValveBank controller, and a pump-controlled outlet. 5-HT and 5-HIAA were made fresh. Washes and test concentrations were perfused onto the cultures at a rate of ~1 mL/minute.

Stopped-flow

iSeroSnFR protein was purified as above and kinetics were determined by mixing equal volumes of 0.2 μM protein with varying concentrations of 5-HT in PBS in an Applied Photophysics SX20 stopped-flow fluorimeter with 490 nm LED excitation and 510 nm long pass filter. Data were collected at 37°C.

Uncaging

Prior to imaging, all cultures were washed and incubated in HBSS containing 2 mM $MgCl_2$ and 2 mM $CaCl_2$ to minimize background due to fluorescent culture medium components. All data collected for the same experiment were collected with identical imaging parameters. MatTek dishes containing primary neurons were washed 3X with HBSS, then 1 mL of HBSS with 200 μM PA-N-5HT or NPEC-5HT was added to the dish. Dishes were imaged on a laser scanning confocal (Zeiss 880). The 488 nm laser was used for imaging at 0.5% power. The 405 nm laser was used for uncaging at 100% power. Bidirectional line scans (1 × 128 px) were performed at 5000 Hz (200 μs/scan). Frame scans (128 × 128 px) were performed at 40 Hz (25 ms/scan). The uncaging area was selected using the Zen software's bleaching function, then setting the initial delay, the bleaching period, and the inter-bleach interval.

Virus production

Virus was produced as previously described (Broussard et al., 2018). Briefly, transfer plasmid, capsid plasmid and helper plasmid were grown in a large culture volume and purified using an endo-free maxiprep kit, per the manufacturer's instructions. AAV293 cells were grown in DMEM + FBS + Pen Strep, as described above, and transfected with all three plasmids (17 μg transfer, 14.5 μg helper, 8.4 μg capsid) using jetPrime reagent per the manufacturer's instructions. After 72h incubation, medium was collected and cleared by centrifugation (2000 xg, 5 min) and filtered through a 0.22 μm filter. 4°C AAVance concentration reagent was added and incubated at 4°C for 18-96h. Virus was pelleted by centrifugation (1500 xg, 30 min, 4°C), washed by resuspending in 4°C culture medium, and pelleted again (1500 xg, 3 min). Virus was finally resuspended in 4°C PBS. Genomic titers were determined by qPCR within the WPRE region.

Surgical procedures for slice experiments

Viral injections

Injection procedures were performed as previously described (Broussard et al., 2018). Briefly, animals were anesthetized using 0.5%–2.5% isoflurane and mounted on a stereotaxic apparatus (900). For injections into the mPFC, a small craniotomy (1-2 mm diameter) was performed around Bregma to expose the mPFC. Virus injection was performed using a glass pipette beveled at 45° with a 15–20-μm opening and back-filled with mineral oil. A fitted plunger controlled by a hydraulic manipulator was inserted into the pipette and used to load and inject the viral solution. Either AAV2/9-*CAG*-iSeroSnFR-Nlgn or AAV2/9-*CAG*.FLEx.iSeroSnFR-PDGFR and AAV2/1-*CMV*-Cre was injected into $SERT^{Cre/+}$; $Rosa^{Ai14/+}$ mice in 3 to different locations on each side, ~250 μm apart, and 500-1500 μm deep such that a large area of the mPFC would be labeled. For injections into the striatum, AAV2/9-CAG-iSeroSnFR-Nlgn was injected bilaterally into striatum and nucleus accumbens using a Nanoject II or III. Virus was injected at AP +1.2; ML ± 1.0; DV −4.6, −4.1, −3.6, −3.1 mm. 250 nL was injected at each depth for a total of 1000nL/side. Mice were allowed to recover > 2 weeks to allow for sensor expression.

Brain slices for 2-photon imaging

Brain slices for two-photon imaging were prepared as previously described (Broussard et al., 2018). Briefly, animals were sacrificed, and the brain was dissected out and placed into ice-cold cutting solution (73 mM NaCl, 2.5 mM KCl, 2 mM $MgCl_2x6H_2O$, 1.25 mM

$NaH_2PO_4xH_2O$, 25 mM $NaHCO_3$, 25 mM glucose, 0.5 mM $CaCl_2$, 75mM sucrose). 300 μm slices were cut using a vibrating microtome. Slices were allowed to recover for 45 mins at 37°C in cutting solution, before being transferred to artificial cerebrospinal fluid (aCSF) (125 mM NaCl, 2.5 mM KCl, 1 mM $MgCl_2x6H_2O$, 1.25 mM $NaH_2PO_4xH_2O$, 25 mM $NaHCO_3$, 10 mM glucose, 2 mM $CaCl_2$). Imaging was carried out on a custom-built 2-photon microscope. ScanImage version 5 software was used to control the microscope and collect the images. Slices were scanned at 30 Hz using 920 nm light. Ligands were applied by microinjection through a pulled glass pipette placed on the surface of the slice within the imaging area. Pressure was applied manually, and ligand expulsion was confirmed by slight deformation of the tissue.

Brain slices for 1-photon photometry

Brain slices for photometry imaging were prepared as follows. Mice were anesthetized with isoflurane and brains were removed and placed in warm (33°C) aCSF containing the following (in mM): 124 NaCl, 2.5 KCl, 1.3 $MgCl_2$, 2.5 $CaCl_2$, 1.0 NaH_2PO_4, 26.2 $NaHCO_3$, 20 glucose, 0.4 ascorbate, and 3 kynurenic acid. Sagittal slices (230 μm) were prepared in warm aCSF using a Leica vibratome. Slices were incubated in warm (33°C) 95%O_2/5%CO_2 oxygenated aCSF containing kynurenic acid (3 mM) for at least 30 min and moved to room temperature (22°C–24°C) until used. Photometric recordings were made from slices containing striatum and transferred to a recording chamber maintained at 34°C, and aCSF perfused at the rate of 2 mL/min.

5-HT transients were detected under a 40x water immersion objective on an upright fluorescent microscope connected with a photomultiplier tube (PMT). Regions of interests (ROIs) were excited using Solis LED and excitation/emission lights were filtered with a dichroic. ROIs were excited for 20 s every 2 minutes using TTL signals to control the LED. Monopolar saline-filled glass electrodes (~4 MΩ, stimulation: 0.5 ms, 50 μA) were used to evoke 5-HT release with various pulses and frequencies. The current produced by the PMT was digitized by Axopatch-200B amplifier and sampled with pClamp 10 software. Data were analyzed with AxoGraphX software. 5-HT transients were calculated as the ratio of the peak amplitude transient (ΔF) to the baseline value (F_0), defined as $\Delta F/F_0$.

Surgical procedures for *in vivo* mouse experiments

General

At the beginning of surgery mice were anesthetized with 5% isoflurane for induction and later 1% isoflurane was used for maintenance. After induction of anesthesia, ketoprofen (5 mg/kg) and buprenorphine SR (1 mg/kg) were subcutaneously injected. The mouse was mounted on a stereotaxic frame. During surgery, body temperature was maintained with a heating pad. Before a sterile scalpel was used to make an incision, the hair covering the skin above the skull was removed, and the skin was subsequently sterilized with chlorhexidine. To have consistent horizontal alignment of the skull, bregma and lambda were leveled to be on the same z axis while two points on the surface of the skull 2.5 mm to either side of lambda were used to level the skull with regard to the y axis. Following viral injection, optical fiber and EEG/EMG implantation, bupivacaine (up to 1 mg/kg) was applied subcutaneously for local analgesia before the wound was closed. Tissue adhesive was applied to the closed wound. For the duration of the post-operative care period, mice were provided with ibuprofen (30 mg/kg) in the drinking water.

Viral injection

To inject virus and implant optical fibers for optogenetic and fiber photometry experiments, craniotomy holes were made over the mPFC (anterior-posterior (AP) axis: 1.3 mm; medio-lateral (ML) axis: 0.4 mm) and BLA (AP: −1.42 mm; ML: 3.0 mm). Virus was injected using a blunt 35 (mPFC) or 33-gauge (BLA) microinjection needle within a 10 μL microsyringe. The injection needle was lowered into the mPFC (dorsoventral (DV): −2.3 mm) or BLA (DV: −5.6 mm) and 0.5 μL of AAV was infused per site at a rate of 100 nL per min. Injection volume was controlled by a microsyringe pump, which was connected to a controller. Following injection, the virus was allowed to diffuse into the tissue for an additional 10 min before the needle was withdrawn over approximately 10 min.

Optical fiber implantation

After viral injection, optical fibers were mounted in a stereotaxic holder and inserted into tissue targeting the mPFC (AP: 1.3 mm; ML: 0.4 mm; DV: −1.8 mm; MFC_400/430-0.48_3mm_ZF1.25_FLT) or the BLA (AP: −1.42 mm; ML: −3.0 mm; DV: −5.1 mm; MFC_400/430-0.48_7mm_ZF1.25_FLT), or the DRN (with 20 degrees angle AP: −4.3 mm; ML: 1.1 mm; DV: −2.85 mm; MFC_400/430-0.57_5mm_ZF2.5(G)_FLT). First, a layer of adhesive cement was applied to the surface of the skull around the optical fiber followed by a layer of dental cement to secure the optical fiber.

EEG and EMG implantation

A drill bit was used to make three craniotomy holes (for reference and ground: AP: −3.2 mm; ML: −2.8 mm; for first EEG channel: AP: 1.7 mm; ML: 0.8 mm; for second EEG channel: AP: −1.3 mm; ML: 1.3 mm) into which 0.10" electrodes with wire lead were inserted. Similar to optical fiber implantation, electrodes were covered with adhesive cement to secure them to the skull. Lead wires from the screw electrodes were connected to an EEG/EMG headmount (8201, Pinnacle Technology Inc.). To improve electrical conduction, silver epoxy was spread over the connections between lead wires and headmount. EMG wires from the headmount were inserted into the trapezius muscles. Finally, dental cement was used to cover all wires and their connections to the headmount.

Histology

Perfusion

Animals were euthanized with 100 mg/kg euthasol i.p. injection followed by trans-cardial perfusion with ice-cold 1x phosphate buffered saline (PBS) and subsequent perfusion with ice-cold 4% paraformaldehyde (PFA) in 1x PBS. After extraction of mouse brains, samples were post-fixed in 4% PFA at 4°C overnight on a shaker. Mouse brains were cryo-protected by immersion in 30% sucrose in 1x PBS solution for > 2 days, before samples are embedded in O.C.T. Compound and frozen in dry ice for > 1 h. Samples were then transferred to a −80°C freezer for long-term storage or directly sliced into 50 μm sections on a cryostat (Leica Biosystems).

Immunohistochemistry

Before staining, brain sections were washed 1-3 times with 1x PBS to remove O.C.T. Compound. Afterward, slices were incubated overnight at 4°C on a shaker in a 1x PBS solution containing 0.1% Triton X-100, 10% normal goat serum (NGS) or normal donkey serum (NDS) and primary antibody chicken polyclonal anti-GFP (1:500). This was followed by a washing step in which slices were washed 3-4 times for 10-15 min each in 1x PBS. Brain slices were then incubated at 4°C overnight or RT for 2 h on a shaker in a 1x PBS solution containing 0.1% Triton X-100, 10% NGS/DGS and secondary antibody goat anti-chicken IgY H&L Alexa Fluor 488 (1:500), again followed by washing with 1x PBS. To image the slices, they were mounted on glass microscope slides. DAPI-containing mounting media and slices were covered with a glass. Images were taken on a Zeiss LSM 880 confocal microscope. Regions where fibers were misplaced of expression of iSeroSnFR or GFP was insufficient were excluded from other analysis.

Fiber photometry

Fiber photometry in mPFC, BLA and DRN was performed as in a previous study (Cho et al., 2017). A 490 nm LED operated at 211 Hz and a 405 nm LED operated at 531 Hz were used for iSeroSnFR excitation or for an isosbestic wavelength, respectively. The two LEDs were controlled by a custom-written software (provided by Drs. Karl Deisseroth and Tom Davidson at Stanford University) and processor. Two filters were used to clean the excitation lights in the LEDs. Dichroic mirrors guided the light into a patch cord connected to an implanted optical fiber using a ferrule. The excitation wavelength light power at the end of the patch cord was measured with a power meter and set to 70 μW. iSeroSnFR emission signals were passed back through the patch cord, dichroic mirror, an emission filter and then a focusing lens until it reached the femtowatt photoreceiver. Signals from the photoreceiver were sent to an RX8-2 processor using a BNC cable and split into two signals corresponding to each LED excitation wavelength based on modulation frequency. These signals were then digitized at a 382 Hz sampling rate followed by low-pass filtering with a 15 Hz threshold. To synchronize iSeroSnFR fluorescence with EEG/EMG recordings, TTL pulses from a TTL pulse generator were fed into an additional digital input channel using a BNC cable.

For multisite fiber photometry experiments (OFC-BNST-BLA), mice were habituated to patch cord tethering for 2 days prior to the start of behavior. For photometry recordings, 470nm light was delivered at 100-120uW and iSeroSnFR signals were recorded for the duration of the fear conditioning protocol using a Neurophotometrics (San Diego, CA) system. Background fluorescence measurements were taken by briefly turning off LEDs while mice were tethered, but prior to the start of behavioral testing. Signals were analyzed using a custom MATLAB script. Briefly, background fluorescence was subtracted for each ROI, and the recorded traces were low-pass filtered at 2Hz. Traces were then fit to a biexponential curve, and the fit was subtracted from the signal to correct for baseline drift. The first two minutes of baseline recordings were discarded to improve curve fitting. ΔF/F% was calculated as (100*(470 signal - fitted signal) / fitted signal). Traces were then z-scored.

Polysomnographic recordings

For fiber photometry experiments with concurrent polysomnographic recordings, mice were singly housed in clear Plexiglas cylindrical tubes (diameter = 15,'' height = 20'') with *ad libitum* food and water. Prior to the start of experiments, mice were habituated to the experimental setup for > 2 days. Animals were connected to a customized pre-amplifier (10x gain, 0.5 Hz high-pass filter, 10 Hz high-pass filter for EMG) attached to a commutator to allow unrestricted movement. EEG/EMG signals were digitized using the OmniPlex System with a sampling rate of 1 kHz. Plexon's proprietary file format (pl2) was converted to European data format (.edf) with MATLAB using the EEGLAB plug-in.

Sleep/wake analysis

Sleep-wake classification based on EEG/EMG signals was conducted using the SleepSign analysis software. Wake states were defined by desynchronized EEG and high EMG activity, NREM sleep by synchronized EEG with high power at < 4 Hz and low EMG activity, and REM sleep by desynchronized EEG with high power at theta frequencies (6-9 Hz) and very low EMG tone. Non-overlapping 5 s windows of EEG/EMG signals were assigned a behavioral state. Delta power, theta ratio and EMG integral were the parameters used for automatic scoring with arbitrarily chosen thresholds. Afterward, classification was manually corrected by an experimenter (M.A.) who was blind to expressed transgene.

Fear conditioning

Behavioral experiments were performed in an operant chamber located within a sound-attenuating box (Lafayette Instruments (mPFC/BLA), or Med Associates (DRN)). Tone (±light), and electric shock were programmed in and controlled by ABET II software

(Lafayette Instruments, mPFC/BLA), or Video Freeze Software (Med Associates Inc., DRN). Each experiment consisted of 15 individual trials. A single trial consisted of a 10 s long cue and a 1 s break followed by a 1 s electric footshock of 0.6 mA (mPFC/BLA) or a 27 s long cue and a 1.5 s shock (0.6 mA). The intervals between single trials were randomly chosen from a uniform distribution between 45 and 75 s (mPFC/BLA) or set to 2 mins (DRN). The cue included a house-light (mPFC/BLA) and 70 dB 5 kHz tone (all). Fear conditioning was performed after sleep experiments.

Running wheels

In running wheel experiments, C57BL/6J mice were singly housed starting one week after surgery, and were given continuous access to a functional or locked running wheel (Med Associates, St. Albans, VT) in the home cage. Wheel revolutions were counted wirelessly for the duration of the experiment. Behavioral testing began following 6 weeks of wheel running.

Free Social Interaction

Test mice were placed into an empty novel cage and allowed to freely explore. After a 2 min baseline period, a sex and age-matched social target (C57) was placed into the cage for a total test time of 12 min.

Oscillating Stimulus Transporter Assay

OSTA was performed as previously described (Keller and Looger, 2016). Briefly, HEK293 cells were cultured in DMEM, 10% FBS, 1% Penicillin/Streptomycin, and 1% L-glutamine and transfected using FuGENE (Promega) with plasmids encoding human serotonin transporter (hSERT, Addgene 15483) and p-cyto-iSeroSnFR, which omits the transmembrane anchor to induce cytosolic localization. Cells were seeded onto 35 mm coverslip bottomed dishes (MatTek) with 10-20 mm microwells to become 60%–80% confluent when measured. A gravity-fed, four-channel perfusion system delivered solutions at 2-3 mL/min, and wide-diameter solution reservoirs (2L bottles and large beakers) were used to minimize changes in height of solution surfaces and concomitant changes in perfusion rates. A short segment of Tygon tubing with an inner diameter of 1/16'' was added after the perfusion manifold to reduce the flow velocity which allowed for better cell stability. This outlet was positioned within 2-3 mm from the imaging area, and a suction tube at the edge of the dish or microwell removed the solutions. Buffer changes were estimated to be > 90% in 1 s and ~100% in 1.5 s. For all data in Figure 6, cells were imaged on a Zeiss 800 inverted laser scanning confocal microscope with excitation at 488 and 561 nm wavelengths, and emission at 495-550 and 570-650, respectively, and using an EC Plan-Neofluar 10x/0.30 M27 objective.

For ion substitution experiments (Figures 6B and 6C), substitutions were either equiosmolar or within 20 mOsm due to addition of salts. NMDG and gluconate were used as null substitutes for cations and anions, respectively. Each raw ROI trace was normalized by overall minimum and maximum fluorescence values, then averaged to produce the trace shown.

For quantification of sodium dependence of transport (Figures 6D–6F), perfusion buffers were oscillated between influx buffer (IB: 10 μM 5-HT, 150 mM NaCl, 10 glucose, 2 $CaCl_2$, 2 $MgCl_2$, 5 MOPS pH 7.5 with TRIS base) and efflux buffer (EB: same as IB, but without 5-HT and substituting KCl for NaCl to increase efflux and allow faster oscillation) for standardized ''fiducial,'' full-strength sodium bouts. Sulforhodamine-101 was added as a marker at 200 nM in IB, and the microscope's red fluorescence channel gain and laser power were adjusted to provide maximum dynamic range without saturated pixels. Once steady oscillations in fluorescence were attained under fiducial conditions, ''test'' bouts were interleaved between the fiducial bouts, in which a gradient mixer system gradually diluted IB with decreased-sodium IB (substituted with a dye-free NMDG-Cl version of IB). The plot of red fluorescence acted as a readout of sodium concentrations, and the fiducials remained very nearly unchanged, indicating the fidelity of the perfusion and gradient systems. For presentation in Figure 6, the red fluorescence trace was normalized by subtraction of a smoothed moving-window average with window size equal to the stimulus period, then divided by a smoothed moving window maximum with window size equal to two periods (to include always one fiducial). In the green (iSeroSnFR) channel, since heights of fiducial bouts changed gradually over time for unknown reasons, individual traces were corrected, similarly to the red channel, by normalizing the entire trace to smoothed moving window minimum and maximum functions. Residual oscillations of unknown cause, observed at zero sodium concentration, were subtracted out, and resulting peak heights of the test bouts were taken as readouts of the integrated rates of transport, assuming approximate linearity. These rates were expressed as percentages of transport observed during the fiducials, which was assumed to be saturated with regard to the effects of sodium. The data were then fit by the Hill equation in GraphPad Prizm, either individually or as a group.

For pharmacological data (Figures 6G–6L), the stimulus was the same IB/EB oscillation as in the fiducials described above, but with pharmacological agents at indicated concentrations superadded to both buffers in the middle epoch of the experiment. Oscillating signals representing transporter flux were isolated from tonic changes by a applying a high-pass frequency filter in the form of a two-sided boxcar filter with window length equal to the stimulus period. To convert this signal into a purely positive-going one, a smoothed moving-window minimum function with window length of one stimulus period was first subtracted, then the result was divided by the maximum of the entire experiment. While there is also presumably information contained in the discarded tonic signal, nothing appeared remarkable offhand, and the isolated oscillating signal gave the most straightforward readout of drug effects. Raw images are available for further inspection and alternative processing.

For experiments on MDMA (Figures 6M and 6N), the baseline stimulus was simply oscillation of IB ± 10 μM 5-HT, without using potassium for the efflux bouts as above. Although this required longer periods, it allowed more direct probing of MDMA's ability to cause efflux. In Figure 6N, the last segment of the trace shows, for reference, decreases of intracellular 5-HT in response to

substituting potassium for sodium, which might be used as a standard for future efflux experiments. These data were processed similarly to those above, with the exception that the minimum- and maximum functions were reversed to show more directly the downgoing magnitudes of 5-HT efflux.

Cocaine and MDMA were used according to Institutional Biosafety Committee and Drug Enforcement Agency protocols of Howard Hughes Medical Institute, Janelia Research Campus and National Institute of Mental Health.

QUANTIFICATION AND STATISTICAL ANALYSIS

Image analysis

Images were analyzed using Fiji (Schindelin et al., 2012). For each experiment, a region of interest (ROI) was selected manually and stored in the ROI manager. Then all images from that experiment were analyzed by calculating the average pixel intensity within the ROI for each frame. Timestamps were extracted using the LSM toolbox plugin. For each experiment, a baseline period was collected and averaged as the basal fluorescence. $\Delta F/F_0$ was calculated for each frame as (average pixel intensity – basal fluorescence)/basal fluorescence.

Fiber photometry

Fiber photometry data processing was conducted as previously reported (Cho et al., 2017). Briefly, a 4^{th} order Butterworth filter with zero-phase distortion was applied to low-pass filter (with a threshold of 2 Hz) the fluorescent signals from the 490 nm and 405 nm excitation wavelengths. The 405 nm signal was aligned to the 490 nm signal using a least-squares linear fit. $\Delta F/F_0$ values were computed as follows: (490 nm signal – fitted 405 nm signal)/(fitted 405 nm signal).

EEG spectrogram

Fiber photometry data processing was conducted as previously reported (Cho et al., 2017). In short, a Morlet wavelet time-frequency decomposition of EEG signals was performed and convolved with a complex Morlet wavelet having a Gaussian shape both in time and frequency domain around its central frequency.

Statistical methods

Student's t test and 2-way ANOVA were performed on *in vivo* fiber photometry data as noted using Excel or Graphpad Prism. All other statistics were performed using R (R Development Core Team, 2013), including the tidyr, tidyverse, and ggplot2 packages (Wickham, 2016; Wickham et al., 2019).

Statistical modeling was performed using the data mining GUI: Rattle (Williams, 2011), which includes the randomForest package (Breiman, 2001) and the glm function.

RandomForest was performed using the following call:

```
randomForest(formula = x5HT ~.,
data = crs$dataset[, c(crs$input, crs$target)],
ntree = 500, mtry = 14, importance = TRUE, replace = FALSE, na.action = randomForest::na.roughfix)
Type of random forest: regression
Number of trees: 500
No. of variables tried at each split: 14
Missing value imputation is active.
```

Generalized linear modeling was performed using the following call:

```
glm(formula = x5HT ~., family = Gaussian(identity), data = crs$dataset[,
c(crs$input, crs$target)])
```

Hill Equation fitting was performed using the drc package (Ritz et al., 2015), with the following code:

```
library(tidyr)
library(drc)
library(ggplot2)
#copy data to clipboard, concentrations are rows, trials are columns
read.excel < - function(header = TRUE, ...){read.table("clipboard," sep = "\t," header = header, ...)}
rawdata = read.excel()
#Tidy data
colnames(rawdata)[1] < - "concs"
longdata < - gather(rawdata, Trial, dFFs, -concs)
#Fit the Hill equation
mycurve < - drm(longdata$dFFs ~longdata$concs, fct = LL.4())
#Calculate 95% confidence intervals
minconc < - min(longdata$concs, na.rm = TRUE)
maxconc < - max(longdata$concs, na.rm = TRUE)
```

```
newdata < - expand.grid(concs = exp(seq(log(minconc/2), log(maxconc*2), length = maxconc*2)))
pm < - predict(mycurve, newdata = newdata, interval = ‘‘confidence’’)
newdata$p < - pm[,1]
newdata$pmin < - pm[,2]
newdata$pmax < - pm[,3]
#b = (-)Hill coefficient, c = min offset, d = max, e = Kd/EC50
summary(mycurve)
#for saving and opening in illustrator, use eps, max path = 30,000 points
ggplot(longdata, aes(x = concs, y = dFFs)) + geom_point(color = ’’blue’’) +
geom_line(data = newdata, aes(x = concs, y = p), color = ’’blue’’) +
scale_x_log10() +
xlab(‘‘[5HT] (uM)’’) + ylab(‘‘dF/F’’) +
geom_line(data = newdata, aes(x = concs, y = pmin)) +
geom_line(data = newdata, aes(x = concs, y = pmax)) +
ggtitle(‘‘Dose Response Curve’’)
ggsave(plot = last_plot(), file = ’’C:\\Users\\EKUnger\\Desktop\\5HTcurves.eps,’’ width = 14, height = 8)
```

Fisher’sexacttest was performed using the fisher.test package with the following code:

```
fisher.test(rbind(c(x1,x2, ...),c(y1,y2, ...)));
```

Wilcoxon rank-sum test was performed using the wilcox.test package with the following code:

```
wilcox.test(ssmmuts ~glmmuts, data = mydata)
```

ROC and d’ analysis was performed using the following code:

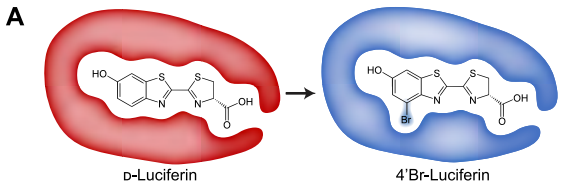
```
library(tidyverse)
library(pROC)
library(psycho)
read.excel < - function(header = TRUE, ...){read.table(‘‘clipboard,’’ sep = ’’\t,’’ header = header, ...)}
rawdata = read.excel() #read in data: cols = trials/ROIs, rows = time points
longdata < - gather(rawdata, ROI, dFFs, -Baseline.Signal) #tidy data
dffs < - longdata$dFFs
signoise < - longdata$Baseline.Signal #Baseline.Signal = classifier column, 0 = baseline, 1 = signal
plot(x = dffs, y = signoise) #frequency distribution
glm.fit = glm(signoise ~dffs, family = binomial) #establish classifier
lines(dffs, glm.fit$fitted.values) #check classifier fidelity
par(pty = ‘‘s’’) #plot constrained to square
roc(signoise, glm.fit$fitted.values,
plot = TRUE,
legacy.axes = TRUE,
xlab = ’’False Positive Rate,’’
ylab = ’’True Positive Rate,’’
print.auc = TRUE) #ROC plot
roc.info < - roc(signoise, glm.fit$fitted.values, legacy.axes = TRUE)
roc.df < - data.frame(
truepos = roc.info$sensitivities*100,
falsepos = (1-roc.info$specificities)*100,
misses = (1-roc.info$sensitivities)*100,
correj = roc.info$specificities*100,
thresholds = roc.info$thresholds) #data for d’ analysis
indices < - psycho::dprime(
n_hit = roc.df$truepos,
n_fa = roc.df$falsepos,
n_miss = roc.df$misses,
n_cr = roc.df$correj)
plot(x = roc.df$thresholds, y = indices[[1]], xlab = ‘‘Threshold,’’ ylab = ‘‘d’ score’’)
par(pty = ‘‘m’’) #plot fills space by default
```

Double-exponential fits were estimated using the nls function with the following code:

```
library(ggplot2)
library(tidyverse)
read.excel < - function(header = TRUE, ...){read.table(‘‘clipboard,’’ sep = ’’\t,’’ header = header, ...)}
rawdata = read.excel()
```

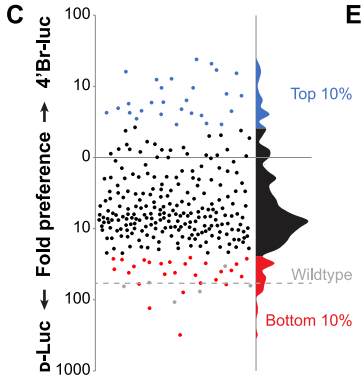
```
colnames(rawdata)[1] < - ‘‘timesec’’
longdata2 < - slice(rawdata, 1:8000) % > % gather(Trial,dFFs, -timesec)
longdata3 < - slice(rawdata, 1:51) % > % gather(Trial,dFFs, -timesec)
#Double exponential equation
#m1 + m2*(1 - exp(-m3*x)) + m4*(1 - exp(-m5*x))
# Where m1 is Y-intercept
# Where m2 is ΔF1 and m3 is k1 (Slow phase)
# Where m4 is ΔF2 and m5 is k2 (fast phase)
doublexp < - function(x,m1,m2,m3,m4,m5) m1 + m2*(1-exp(-m3*x)) + m4*(1-exp(-m5*x))
nlsfit < - nls(longdata2$dFFs ~doublexp(x = longdata2$timesec, m1,m2,m3,m4,m5),
data = longdata2, start = list(m1 = 1.8, m2 = 1.2, m3 = 0.05, m4 = 0.1, m5 = 2000),
nls.control(maxiter = 500, minFactor = 0.000001),
lower = c(0,0,0,0,0), algorithm = ‘‘port,’’ trace = TRUE)
summary(nlsfit))
m1 = coef(nlsfit)[[1]]
m2 = coef(nlsfit)[[2]]
m3 = coef(nlsfit)[[3]]
m4 = coef(nlsfit)[[4]]
m5 = coef(nlsfit)[[5]]
ggplot(longdata3, aes(x = timesec, y = dFFs)) +
geom_point(size = 0.1, alpha = 0.05) +
theme_light(base_size = 16) +
ylab(‘‘AU’’) + xlab(‘‘Time (s)’’) +
stat_function(geom = ‘‘path,’’ fun = function(x)
(m1 + m2*(1 - exp(-m3*x)) + m4*(1 - exp(-m5*x))))
```

Supplemental Figures



B

Position	RF rank	Native	Mutation	GLM
218	97	R	A	0.6817
247	63	F	Y	0.5138
255	39	Y	N	0.8542
316	69	G	S	0.5065
342	0	L	A	1.8681
347	86	S	G	1.8726



D

# Mutations	0	1	2	3	4	5	6
Top 10%	0	1	7	15	4	0	0
Bottom 10%	7	1	7	8	8	0	0

p = 0.2853

E

Position	Residue	Top 10%		Bottom 10%			Input data			
		Freq.	%	Freq.	%	p*	Freq.	%	p* (top)	p* (bot.)
218	**R**	22	81.5	8	25.8	6.6E-4	181	81.5	1	1.8E-6
	A	5	18.5	16	51.6		2	0.9		
	other	NA		NA			39	17.6		
247	**F**	9	33.3	12	38.7	0.2651	195	87.8	2.4E-9	3.4E-5
	Y	18	66.7	12	38.7		11	5.0		
	other	NA		NA			16	7.2		
255	**Y**	21	77.8	18	58.1	1	175	78.8	1	0.6106
	N	6	22.2	6	19.4		2	0.9		
	other	NA		NA			45	20.3		
316	**G**	15	55.6	12	38.7	0.7819	127	57.2	1	0.5223
	S	12	44.4	12	38.7		28	12.6		
	other	NA		NA			67	30.2		
342	**L**	18	66.7	3	9.7	1.5E-4	210	94.6	6.7E-5	<2E-16
	A	9	33.3	21	67.7		1	0.5		
	other	NA		NA			11	5.0		
347	**S**	1	3.7	19	61.3	2.6E-8	199	89.6	<2E-16	0.1661
	G	26	96.3	5	16.1		4	1.8		
	other	NA		NA			19	8.6		

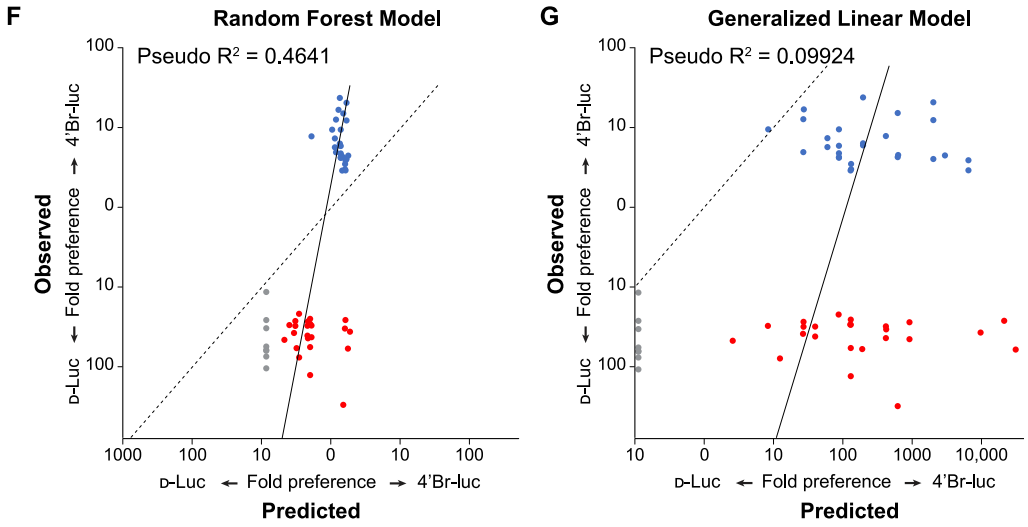
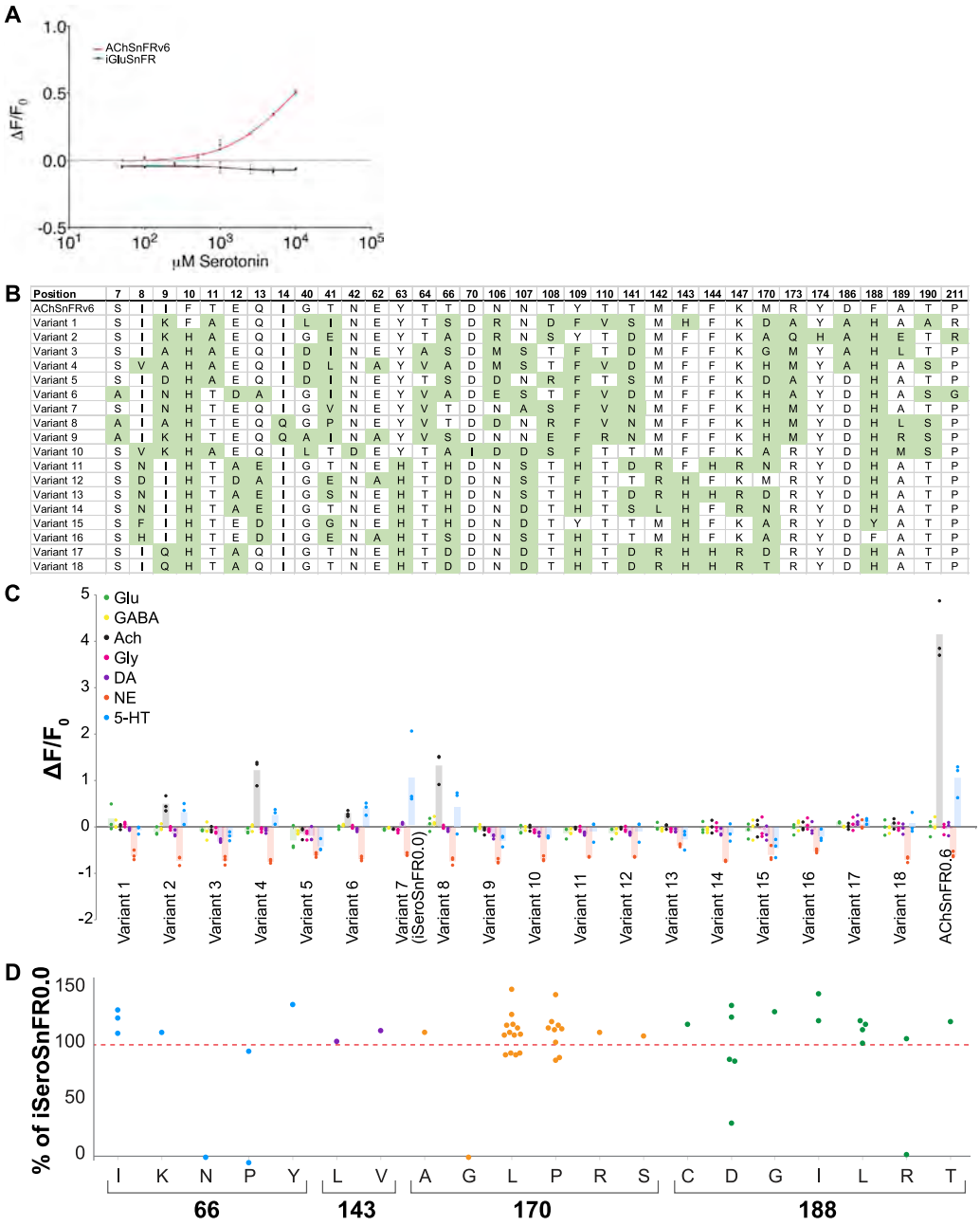


Figure S1. Development of a Machine-Learning-Guided Library Design Paradigm, Related to Figure 1

A. Schematic showing the evolution of substrate preference from D-luc to 4'Br-luc. B. Random forest (RF) modeling and generalized linear modeling (GLM) were performed on 222 variants (see Table S2). C. The combination library in (B) was generated and 276 variants were screened for their preference for D-luc or 4'Br-luc. The top and bottom 10% of variants were sequenced. D. Table showing no difference between the mutation rate of the top and bottom 10% of variants (Fisher's Exact test). Variants with no mutations were omitted from statistical analysis. E. Table showing the frequency of different mutations predicted by statistical modeling. *p values were calculated by Fisher's exact test, comparing the mutated amino acid(s) to the native amino acid, and the top 10% to either the bottom 10% or the input data as noted. F,G. Comparison between the RF (F) and GLM (G) prediction and the actual data.



Position	7	8	9	10	11	12	13	14	40	41	42	62	63	64	66	70	106	107	108	109	110	141	142	143	144	147	170	173	174	186	188	189	190	211
AChSnFRv6	S	I	I	F	T	E	Q	I	G	T	N	E	Y	T	T	D	N	N	T	Y	T	T	M	F	F	K	M	R	Y	D	F	A	T	P
Variant 1	S	I	K	F	A	E	Q	I	L	I	N	E	Y	T	S	D	R	N	D	F	V	S	M	H	F	K	D	A	Y	A	H	A	A	R
Variant 2	S	I	K	H	A	E	Q	I	G	E	N	E	Y	T	A	D	R	N	S	Y	T	D	M	F	F	K	A	Q	H	A	H	E	T	R
Variant 3	S	I	A	H	A	E	Q	I	D	I	N	E	Y	A	S	D	M	S	T	F	T	D	M	F	F	K	G	M	Y	A	H	L	T	P
Variant 4	S	V	A	H	A	E	Q	I	D	L	N	A	Y	V	A	D	M	S	T	F	V	D	M	F	F	K	H	M	Y	A	H	A	S	P
Variant 5	S	I	D	H	A	E	Q	I	D	I	N	E	Y	T	S	D	D	N	R	F	T	S	M	F	F	K	D	A	Y	D	H	A	T	P
Variant 6	A	I	N	H	T	D	A	I	G	I	N	E	Y	V	A	D	E	S	T	F	V	D	M	F	F	K	H	A	Y	D	H	A	S	G
Variant 7	S	I	N	H	T	E	Q	I	G	V	N	E	Y	V	T	D	N	A	S	F	V	N	M	F	F	K	H	M	Y	D	H	A	T	P
Variant 8	A	I	A	H	T	E	Q	Q	G	P	N	E	Y	V	T	D	D	N	R	F	V	N	M	F	F	K	H	M	Y	D	H	L	S	P
Variant 9	A	I	K	H	T	E	Q	Q	A	I	N	A	Y	V	S	D	N	N	E	F	R	N	M	F	F	K	H	M	Y	D	H	R	S	P
Variant 10	S	V	K	H	A	E	Q	I	L	T	D	E	Y	T	A	I	D	D	S	F	T	T	M	F	F	K	A	R	Y	D	H	M	S	P
Variant 11	S	N	I	H	T	A	E	I	G	T	N	E	H	T	H	D	N	S	T	H	T	D	R	F	H	R	N	R	Y	D	H	A	T	P
Variant 12	S	D	I	H	T	D	A	I	G	E	N	A	H	T	D	D	N	S	T	F	T	T	R	H	F	K	M	R	Y	D	H	A	T	P
Variant 13	S	N	I	H	T	A	E	I	G	S	N	E	H	T	H	D	N	S	T	H	T	D	R	H	H	R	D	R	Y	D	H	A	T	P
Variant 14	S	N	I	H	T	A	E	I	G	T	N	E	H	T	H	D	N	D	T	H	T	S	L	H	F	R	N	R	Y	D	H	A	T	P
Variant 15	S	F	I	H	T	E	D	I	G	G	N	E	H	T	H	D	N	D	T	Y	T	T	M	H	F	K	A	R	Y	D	Y	A	T	P
Variant 16	S	H	I	H	T	E	D	I	G	E	N	A	H	T	S	D	N	S	T	H	T	T	M	H	F	K	A	R	Y	D	F	A	T	P
Variant 17	S	I	Q	H	T	A	Q	I	G	T	N	E	H	T	D	D	N	D	T	H	T	D	R	H	H	R	D	R	Y	D	H	A	T	P
Variant 18	S	I	Q	H	T	A	Q	I	G	T	N	E	H	T	D	D	N	D	T	H	T	D	R	H	H	R	T	R	Y	D	H	A	T	P

Figure S2. iSeroSnFR Design, Related to Figure 1

A. Purified protein of iAChSnFR0.6 and iGluSnFR were tested against multiple concentrations of 5-HT. n = 3. Data represent mean ± s.e.m.s.e.m. (B). Table showing the top 18 variants predicted by Rosetta modeling. Mutations are highlighted in green, compared to iAChSnFR0.6. C. The top 18 variants were synthesized and tested as purified protein against multiple ligands. Protein concentration: 100 nM, ligand concentration: 10 mM. Shaded bars represent the mean. D. Single mutants from the first SSM screen were tested with 10 mM 5-HT and compared to iSeroSnFR0.0. Red dashed line indicates iSeroSnFR0.0 (set to 100%).

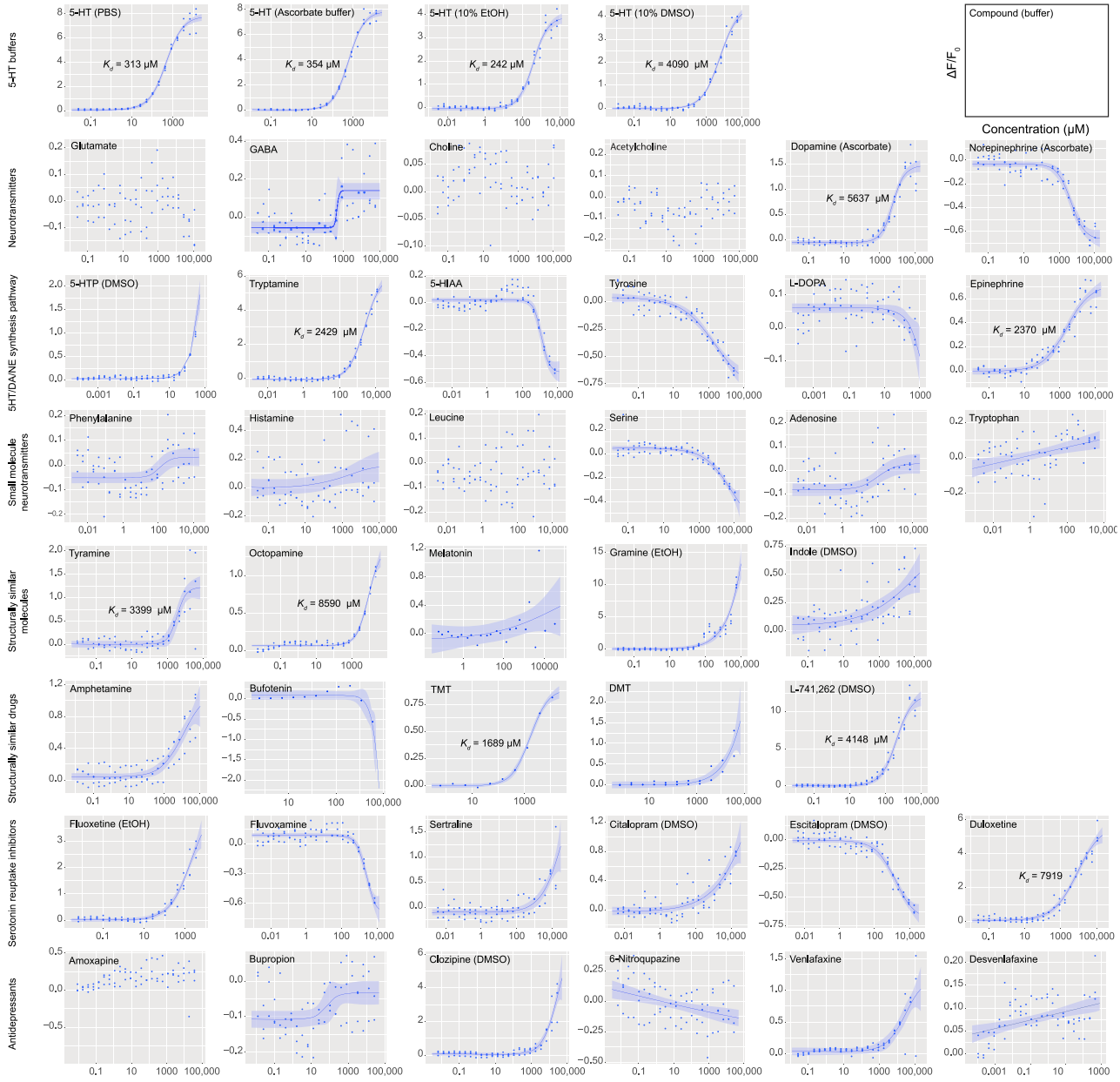


Figure S3. Specificity of iSeroSnFR, Related to Figure 2

iSeroSnFR was purified and tested at 100 nM against multiple ligands as noted. Some ligands were tested in alternative buffers as noted. If no buffer is listed, the ligand was dissolved in PBS. Lines were fitted and K_d's were determined using the drc package in R. Shaded areas represent the 95% confidence interval. If no line is present, convergence failed, and no fit was calculated.

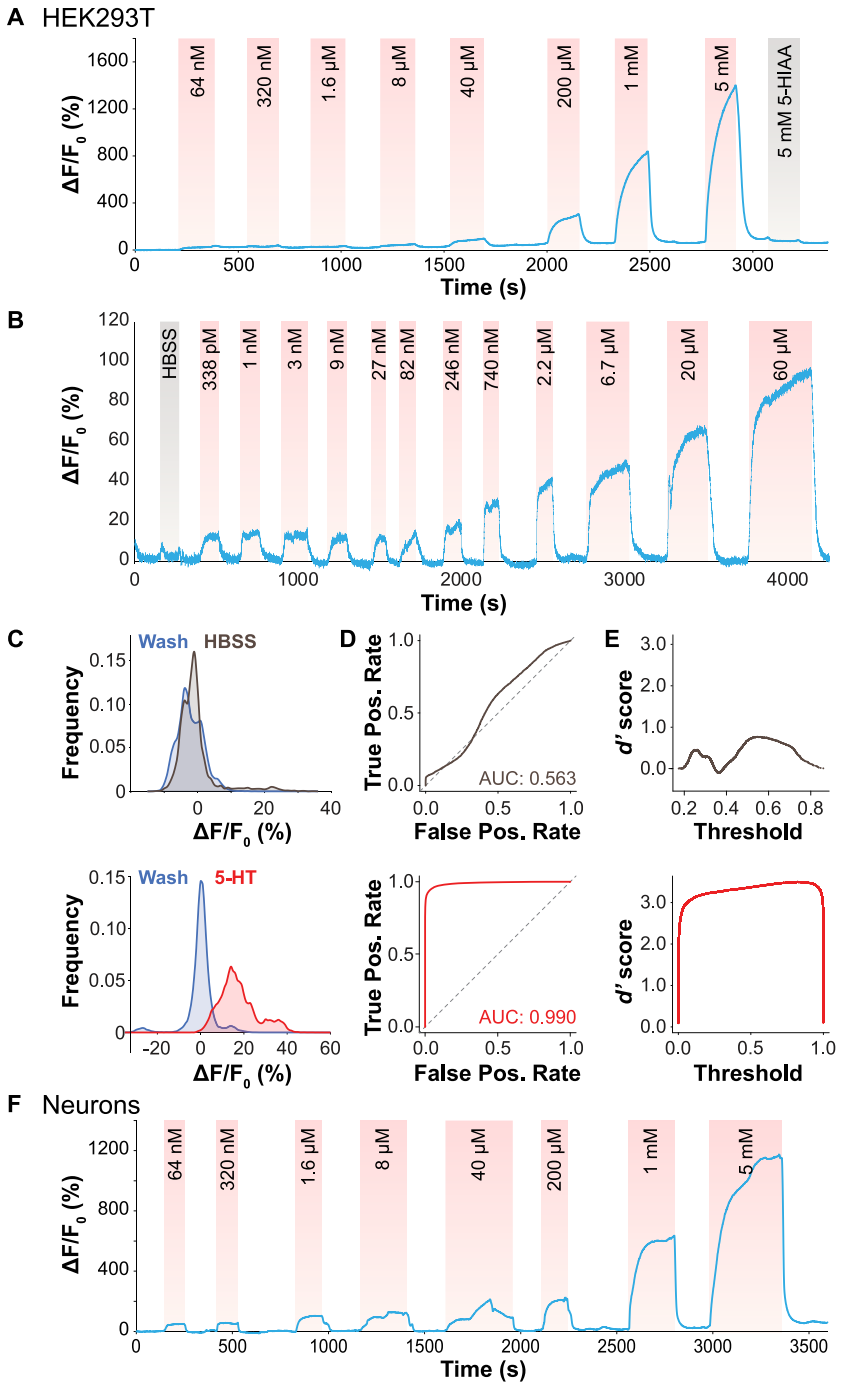


Figure S4. *In Situ* Titration, Related to Figure 2

Representative traces from HEK293T cells (A,B) and primary cultured E18 rat hippocampal neurons (F) for which the images and dose response curves are displayed in Figures 2D–2F (HEK cells) and Figures 2G and 2H (neurons). C-E. HBSS (top) and concentrations of 5-HT between 338 pM and 246 nM (bottom) were further analyzed for sensor sensitivity. C. Distribution of responses is shown. D. ROC analysis was performed where responses during the "wash" period were defined as a false positive, and responses during the "HBSS" or "5-HT" period were defined as a true positive. E. A *d'* score was calculated based on the ROC analysis. (A) and (F) were imaged at 1 Hz, (B) was imaged at 5 Hz. n = 3-4 dishes.

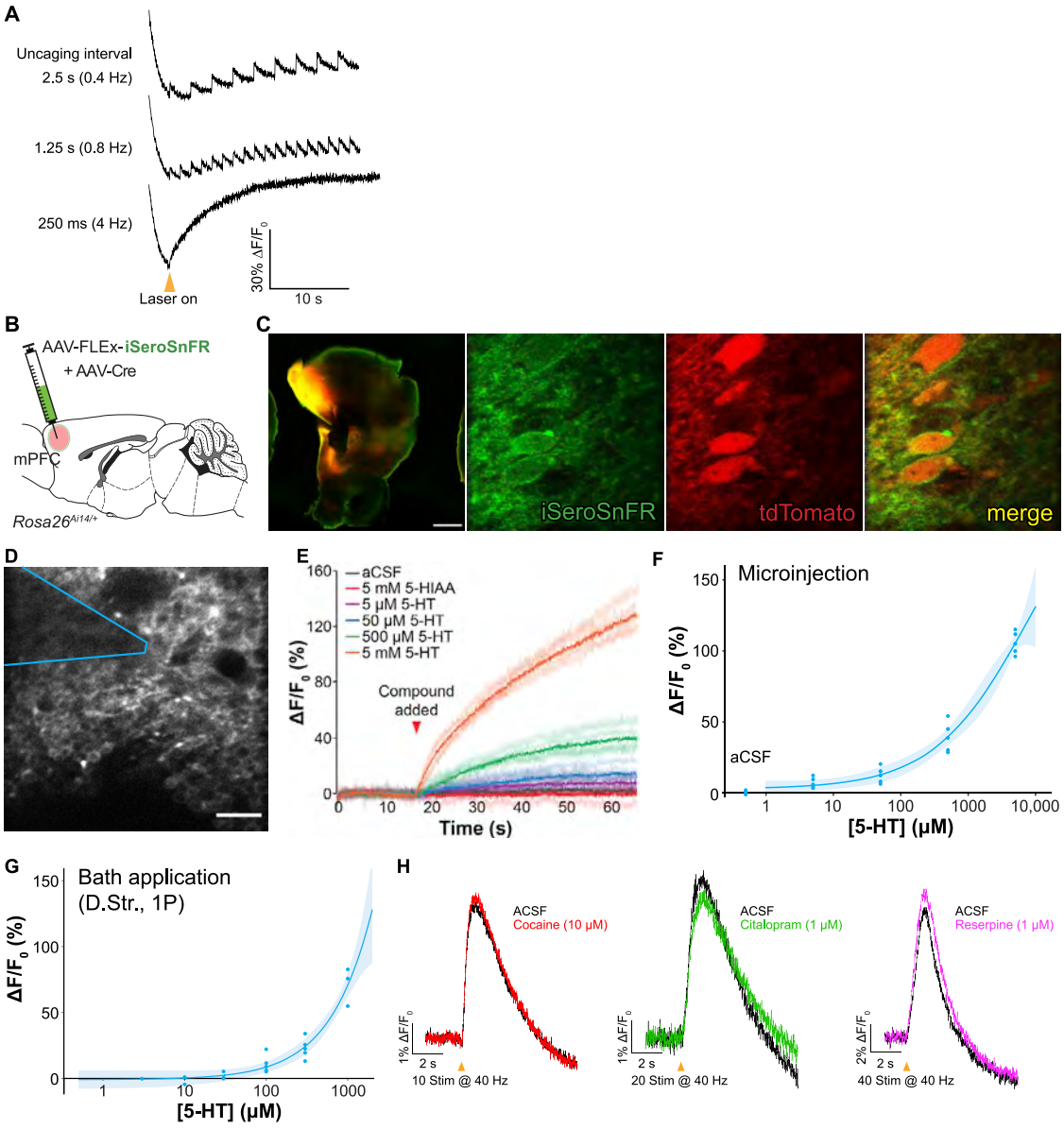


Figure S5. Kinetics and Two-Photon Characterization of iSeroSnFR Expression in mPFC, Related to Figures 3 and 4

A. Representative raw traces from Figure 3O, showing significant bleaching followed by gradual accumulation of 5-HT. Biological replicates = 3. B-F. AAV2/9.*CAG*.FLEx.iSeroSnFR.PDGFR combined with AAV2/1.*CMV*.Cre was injected into the medial prefrontal cortex (mPFC) of *Rosa*$^{Ai14/+}$ mice. Acute slices were prepared (300 μm) and imaged using a 2-photon microscope (920 nm). C,D. Representative images. Scale bars represent 500 μm (low mag.) and 10 μm (high mag.). E. Different concentrations of 5-HT (in aCSF) were microinjected using a pulled glass pipette positioned at the surface of the slice (outlined in blue in (D). 5-HT was expelled manually. (n = 5 slices from 3 mice). Saturated lines indicate mean, pale lines indicate raw traces. F. Average values from the period between 30 and 60 s in (E) were plotted and fitted using the Hill equation. Shaded area represents 95% confidence interval. (G). Mouse brain slices from Figures 4A–4D were perfused with different concentrations of 5-HT. (H). Mouse brain slices from Figures 4A–4D were perfused with either ACSF or the drug noted and stimulated with a monopolar saline-filled glass electrode (0.5 ms, 50 μA).

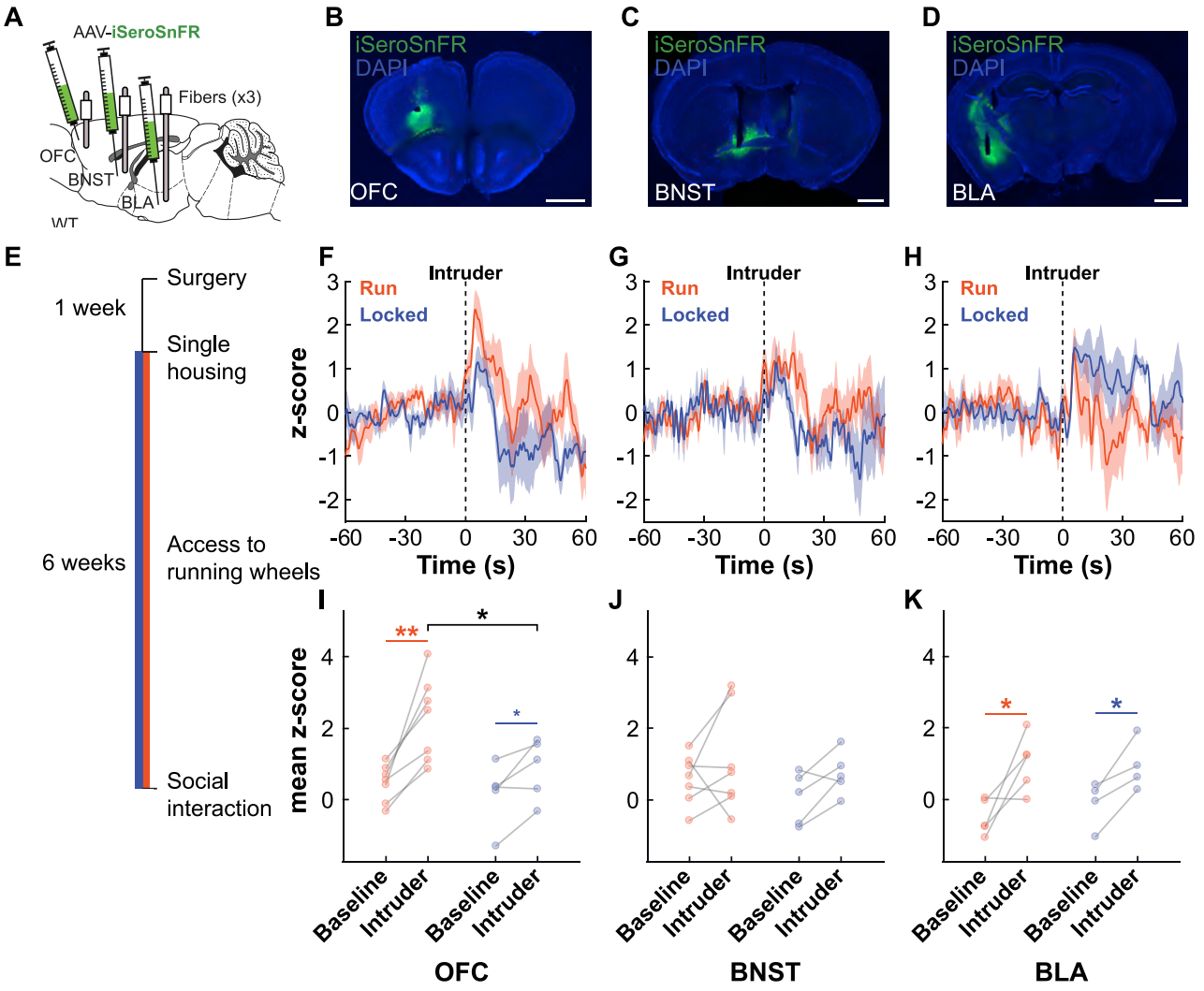


Figure S6. Imaging 5-HT Dynamics in Social Interaction, Related to Figure 4

A. WT C57BL/6J mice were injected with AAV.*CAG*.iSeroSnFR.Nlgn into orbitofrontal cortex (OFC), bed nucleus of the stria terminalis (BNST) and basolateral amygdala (BLA). For multi-site fiber photometry recordings, optical fibers were inserted targeting all three brain regions. B-D. Representative images of iSeroSnFR expression and fiber location. E. Graphical representation of the timeline of surgery and behavioral experiments. After 1 week of recovery animals were provided access to running wheels, but only half were functional (run) and the other half were fixed (locked). F-H. Average fluorescence changes before and after introduction of an intruder. I-K. Mean z-scores frome each animal from the −2-0 s prior to intruder presentation (baseline) and 4-6 s after intruder presentation (intruder). Single trials per animal; N = 4-7 mice. Data represent mean ± s.e.m. *p < 0.05, **p < 0.01, 2-way ANOVA. Scale bars represent 1 mm.

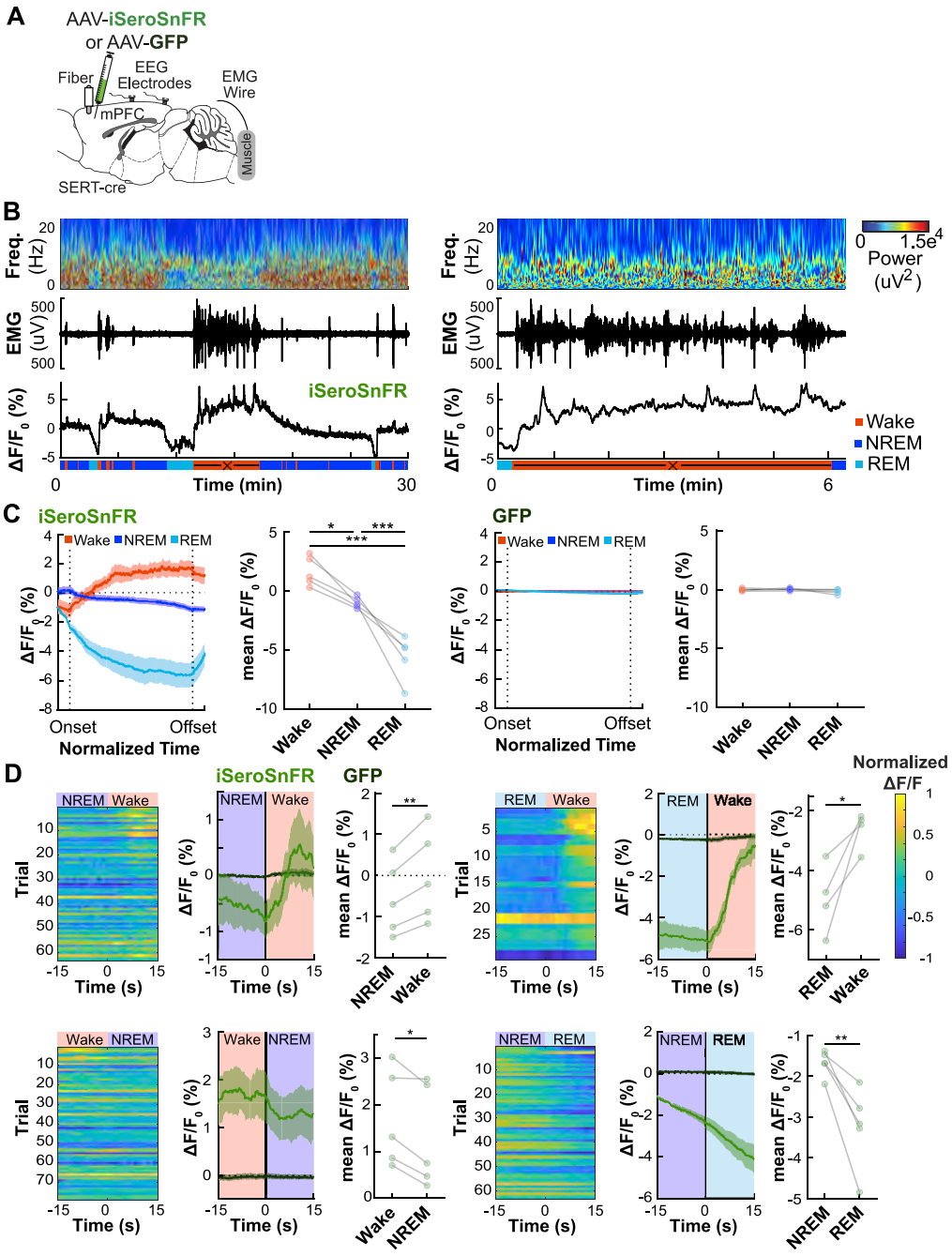


Figure S7. 5-HT Dynamics in Sleep-Wake Cycle in mPFC, Related to Figure 5

A. Mice were injected with either AAV2/9.*CAG*-iSeroSnFR.Nlgn or AAV2/5.*CAG*-GFP (as a negative control). For fiber photometry recordings, an optical fiber was inserted targeting mPFC. EEG screw electrodes and EMG wires were implanted to classify sleep-wake states. B. Representative mPFC^{iSeroSnFR} EEG spectrograms, EMG and fiber photometry traces over time across sleep-wake cycles (left) and walking episode (right). C. Temporal dynamics of mPFC^{iSeroSnFR} (left) and mPFC^{GFP} (right) activity during waking, NREM, and REM episodes within normalized time. Statistical comparisons of fluorescence levels were performed on the last 10% of data within each behavioral state (one-way ANOVA with Bonferroni correction). D. Fluorescence changes of iSeroSnFR and GFP across NREM to Wake, REM to Wake, Wake to NREM and NREM to REM transitions. Statistical comparisons of changes in fluorescence using mPFC^{iSeroSnFR} were made based on the average fluorescence over 15 s before and 15 s after the behavioral state transition. Data represent mean (dark lines) ± s.e.m. (shaded areas). Each row of the heatmap represents one instance of a state transition. n = 3 mPFC^{GFP}; n = 4 mPFC^{iSeroSnFR}. *p < 0.05, **p < 0.01, ***p < 0.001, paired Student's t test.