nature computational science

Article

https://doi.org/10.1038/s43588-024-00611-w

Machine learning-guided engineering of genetically encoded fluorescent calcium indicators

Received: 6 July 2023

Accepted: 15 February 2024

Published online: 21 March 2024

Check for updates

Sarah J. Wait[1,2], Marc Expòsit[1,3,13], Sophia Lin[4,5,13], Michael Rappleye[2,6,7], Justin Daho Lee[1,2], Samuel A. Colby[1], Lily Torp[2,6], Anthony Asencio[2,6], Annette Smith[2], Michael Regnier[2,6], Farid Moussavi-Harami[2,8,9], David Baker[3,10,11], Christina K. Kim[4,5] & Andre Berndt[1,2,6,12]

Here we used machine learning to engineer genetically encoded fluorescent indicators, protein-based sensors critical for real-time monitoring of biological activity. We used machine learning to predict the outcomes of sensor mutagenesis by analyzing established libraries that link sensor sequences to functions. Using the GCaMP calcium indicator as a scaffold, we developed an ensemble of three regression models trained on experimentally derived GCaMP mutation libraries. The trained ensemble performed an in silico functional screen on 1,423 novel, uncharacterized GCaMP variants. As a result, we identified the ensemble-derived GCaMP (eGCaMP) variants, eGCaMP and eGCaMP^{+}, which achieve both faster kinetics and larger $\Delta F/F_0$ responses upon stimulation than previously published fast variants. Furthermore, we identified a combinatorial mutation with extraordinary dynamic range, eGCaMP^{2+}, which outperforms the tested sixth-, seventh- and eighth-generation GCaMPs. These findings demonstrate the value of machine learning as a tool to facilitate the efficient engineering of proteins for desired biophysical characteristics.

Genetically encoded fluorescent indicators (GEFIs) are protein-based sensors that allosterically fuse fluorescent reporters to ligand-binding domains. Ligand binding induces changes in fluorescence intensity, enabling the monitoring of biological compounds such as intracellular second messengers or neuromodulators in freely moving animals[1]. GEFIs have become essential tools in neuroscience, with sensors for calcium, dopamine, norepinephrine, endocannabinoids and opioids, among others[2–11]. To match each sensor's characteristics, like dynamic range or kinetics, with experimental needs, GEFIs require extensive engineering. Current methods, such as trial-and-error mutagenesis, often require substantial time and resource commitments. Recently, machine learning (ML) algorithms have shown proficiency in engineering enzymes, fluorescent proteins and optogenetic tools[12–17]. In this study, we developed an ML approach to predict the biophysical

[1]Molecular Engineering and Sciences Institute, University of Washington, Seattle, WA, USA. [2]Institute of Stem Cell and Regenerative Medicine, University of Washington, Seattle, WA, USA. [3]Institute for Protein Design, University of Washington, Seattle, WA, USA. [4]Center for Neuroscience, University of California, Davis, Davis, CA, USA. [5]Department of Neurology, University of California, Davis, Davis, CA, USA. [6]Department of Bioengineering, University of Washington, Seattle, WA, USA. [7]Institute of Pharmacology and Toxicology, University of Zürich, Zurich, Switzerland. [8]Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA, USA. [9]Division of Cardiology, University of Washington, Seattle, WA, USA. [10]Department of Biochemistry, University of Washington, Seattle, WA, USA. [11]Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA. [12]Center for Neurobiology of Addiction, Pain, and Emotion, University of Washington, Seattle, WA, USA. [13]These authors contributed equally: Marc Expòsit, Sophia Lin. e-mail: berndtuw@uw.edu

https://doi.org/10.1038/s43588-024-00611-w

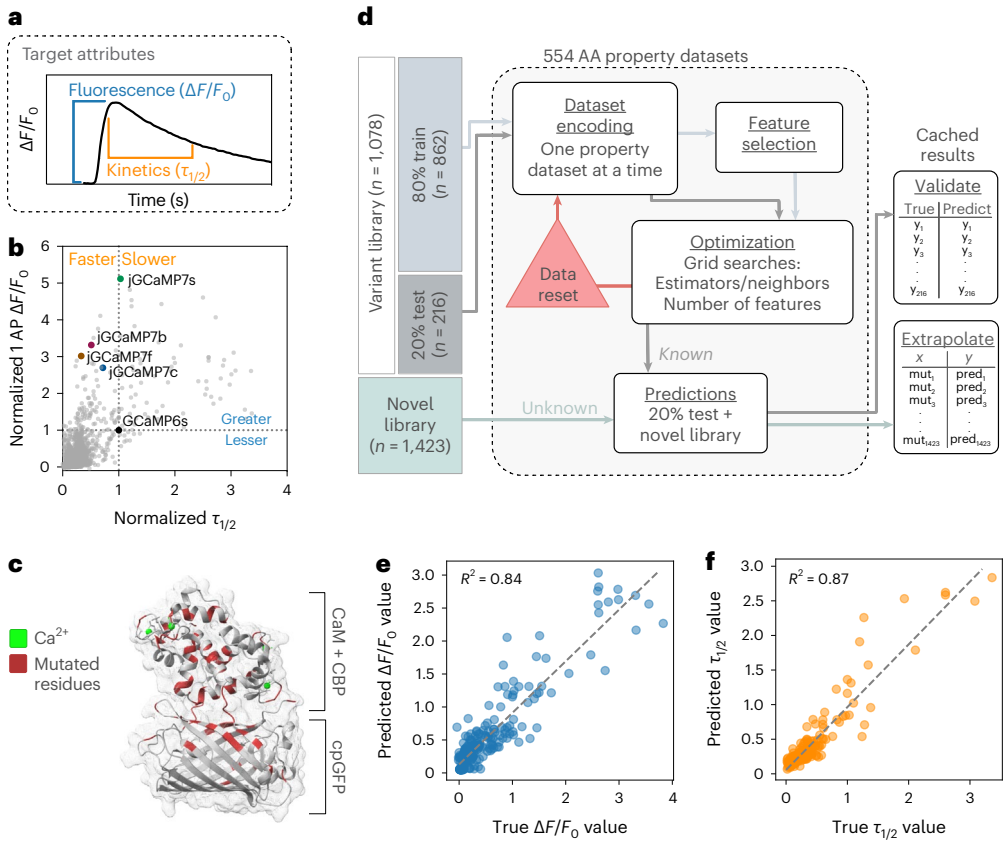


Fig. 1 | Description of variant library, computational approach and ensemble cross-validation. a, Description of the biophysical attributes of the GCaMP sensor targeted for engineering: fluorescence change ($\Delta F/F_0$) and kinetics ($\tau_{1/2}$). **b**, Scatter plot depicts the 1AP $\Delta F/F_0$ by the $\tau_{1/2}$ for each of the 1,078 variants in the variant library[4,5]. Each value was normalized to GCaMP6s as 1.0 for 1AP $\Delta F/F_0$ and $\tau_{1/2}$. Published variants are indicated with colored dots and text labels. **c**, Crystal structure of GCaMP3–D380Y (Protein Data Bank (PDB): 3SG3, gray) with 75 residues (red) in which mutation information exists in the variant library[4,5]. These 75 residues indicate the positions used to form the novel library. Brackets indicate the GCaMP domains CaM, CBP and cpGFP. **d**, Overview of model training schema. The variant library[4,5] was split randomly into an 80% training set and a 20% testing set. The data were encoded using the AAindex property datasets. The train set underwent feature selection before being optimized using a grid search of key hyperparameters for each model. The optimized model was used to form predictions on the 20% test set and the novel library. The final test set and novel library predictions were cached for downstream analysis. **e**, The scatter contains the true $\Delta F/F_0$ value by the predicted $\Delta F/F_0$ value made by the ensemble for each variant in the test set. The dotted line depicts $R^2 = 1.0$. The R^2 value denotes the coefficient of determination of the scatter data. **f**, The scatter contains the true $\tau_{1/2}$ value by the predicted $\tau_{1/2}$ value made by the ensemble for each variant in the test set. The dotted line depicts $R^2 = 1.0$. The R^2 value denotes the coefficient of determination of the scatter data.

characteristics of previously untested mutations. We selected the calcium indicator GCaMP as a protein sensor scaffold to develop this platform. GCaMP is a chimeric protein that consists of circularly permuted GFP (cpGFP) fused to calmodulin (CaM) and calmodulin-binding peptide (CBP). GCaMP sensors have been widely adopted in neuroscience research and have undergone several generations of improvements to optimize their capabilities[2–6,18,19]. Thus, datasets from the functional characterization of more than a thousand mutants are publicly available[4,5]. Using these data, we developed a stacked ML ensemble that predicted the functional characteristics of untested GCaMP variants. Using mutations proposed by the ensemble, we identified variants that accelerate the off-rate kinetics and increase the change in fluorescence over baseline ($\Delta F/F_0$) upon activation. We demonstrated that ML ensembles can effectively learn from complex mutational datasets and that we can harness their predictive power to guide protein engineering efforts.

Results

Description of ensemble development

Our training data consisted of 1,078 characterized mutants derived from cultured neuron screening[4,5,20]. Within the library, we focused on the change in fluorescence over baseline ($\Delta F/F_0$) in response to one action potential (AP; 1AP $\Delta F/F_0$) and decay kinetics of the sensor signal ($\tau_{1/2}$, decay half-time after 10 APs; Fig. 1a). When normalized to the sixth generation GCaMP sensor, GCaMP6s[4] (1AP $\Delta F/F_0$ and $\tau_{1/2} = 1.0$), we can see a broad distribution of variant capabilities and mutation locations within the GCaMP structure (Fig. 1b,c and Supplementary Fig. 1a). We found the sequence similarity is not deterministic for $\Delta F/F_0$ or kinetics, as seen by the variability in mutation impact regardless of GCaMP generation (Supplementary Fig. 1b,d).

Before training, the variants in the library were randomly assigned to training and testing sets at an 80/20 ratio for cross-validation, where the mean values between the training and test sets were not significantly different (Supplementary Fig. 1c,e). We tested three methods of encoding the mutation dataset: one-hot encoding, label encoding and functional encoding with amino acid property datasets found on AAindex[21]. AAindex comprises 554 complete matrices describing different amino acid properties, such as size, polarity and hydrophobicity. Encoding with AAindex property matrices improved the cross-validation R^2 by an average of 20% over one-hot encoded or label-encoded libraries (Supplementary Fig. 2c). We tested the 554 AAindex property datasets to determine which properties led to the largest R^2 values during cross-validation. The predictions from the top-five performing datasets were used to form the final ensemble's predictions (Fig. 1d and Supplementary Fig. 2a). We found that amino acid property datasets that described hydrophobicity were commonly associated with higher-performing predictive capabilities in the $\Delta F/F_0$ model (Supplementary Fig. 2b–d; Supplementary

https://doi.org/10.1038/s43588-024-00611-w

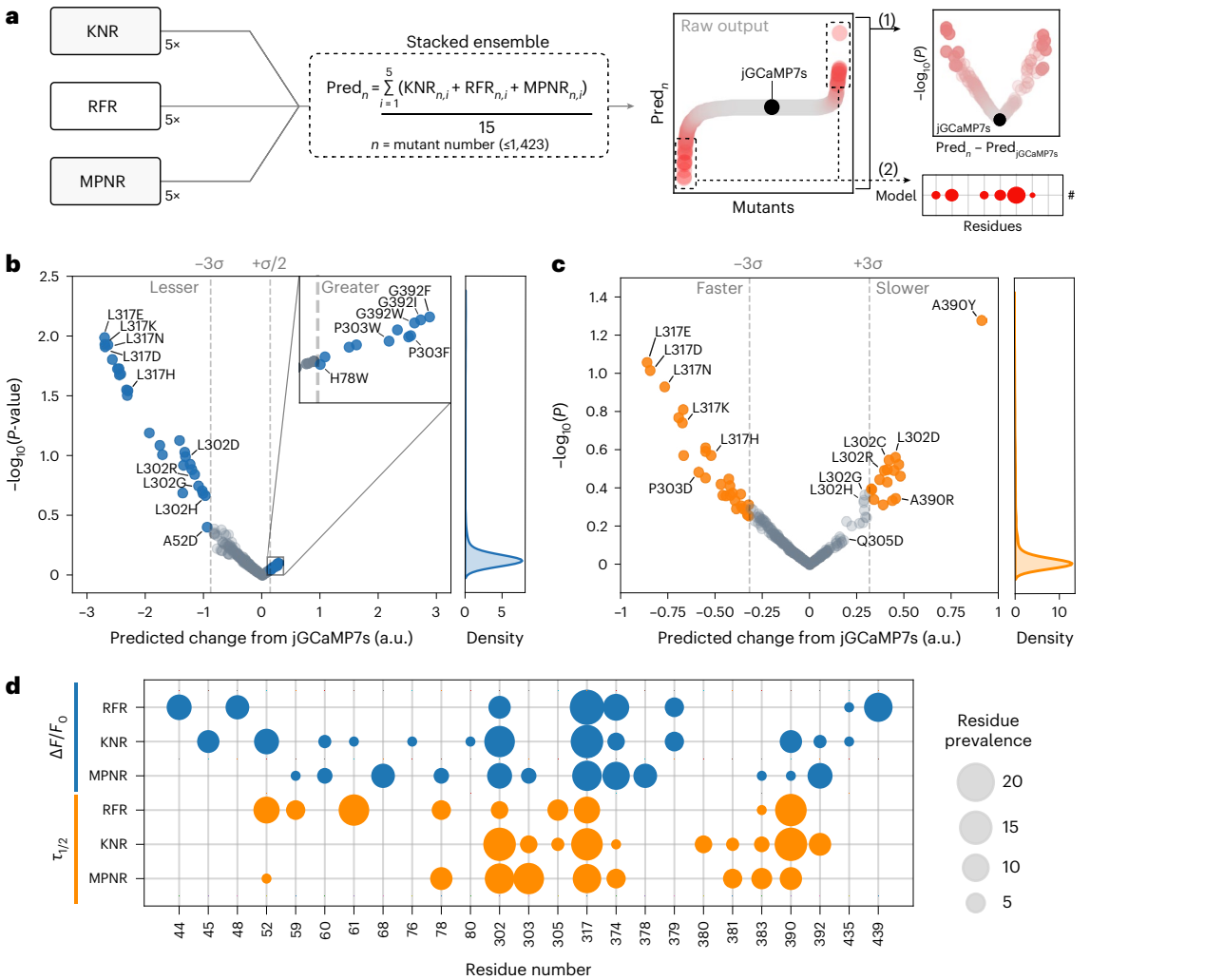


Fig. 2 | In vitro verification of ensemble predictions. a, Brief description of prediction analysis. The stacked ensemble predictions were formed by averaging the predictions from the 15 contributor models for each variant ($Pred_n$) in the novel library. The raw output is the prediction ($Pred_n$) for each mutant, with a prediction for jGCaMP7s as a benchmark. (1) The volcano plots were formed by subtracting the benchmark jGCaMP7s prediction from the variant prediction (x axis) and P values were derived by performing an unpaired t-test between the 15 predictions for variant$_n$ and the 15 predictions for jGCaMP7s. (2) The bubble plot indicates the number of times a given residue appears in the top 2.5% and bottom 2.5% of predictions. **b**, Volcano plots depicting the ensemble's prediction for a given mutation change in fluorescent response from jGCaMP7s (x axis) and the $\log_{10}P$ of the given prediction. P values were calculated using a two-sided unpaired t-test on ensemble predictions (15 models) for jGCaMP7s and the given mutation. Right: Kernel density estimation depicts the spread of $\log_{10}P$ obtained. Dotted lines are included at indicated σ values. **c**, Volcano plots depicting the ensemble's prediction for given mutations change kinetic capability from jGCaMP7s (x axis) and the $\log_{10}P$ of the given prediction. P values were derived using a two-sided unpaired t-test on ensemble prediction (15 models) for jGCaMP7s and given mutation. Right: Kernel density estimation depicts the spread of $\log_{10}P$ obtained. Dotted lines are included at indicated standard deviations (σ). **d**, Bubble plot depicting the number of times each residue (x axis) appeared in the top 2.5% and bottom 2.5% of predicted values for each regressor that constitute the $\Delta F/F_0$ ensemble and the kinetics ($\tau_{1/2}$) ensemble.

Table 1). In comparison, amino acid property datasets associated with protein folding and energetics were common amongst the higher performing predictive capabilities in the kinetics model (Supplementary Fig. 2b–d and Supplementary Table 2).

To improve prediction capabilities, we performed a stacked ensemble comprising a random forest regressor (RFR), K-neighbors regressor (KNR) and multi-layer perceptron network regressor (MPNR)[22,23]. The ensemble's predictions for each mutation are the average response from the 15 models (5 amino acid properties × 3 regressor types). During cross-validation, the ensembles for $\Delta F/F_0$ and kinetics achieved R^2 values greater than 0.80 for predictions made on the test dataset (Fig. 1e,f). The $\Delta F/F_0$ ensemble achieved a higher R^2 value than any models contributing to the prediction, demonstrating the beneficial collaborative effect of ensembling (Supplementary Fig. 2c).

Analysis of ensemble predictions

We utilized the trained ensembles to predict a novel library's $\Delta F/F_0$ and kinetics. This library substituted each of the 75 positions previously mutated in the dataset with the remaining 19 amino acids in the seventh-generation GECI jGCaMP7s (Fig. 1c). After removing redundancies, the library contained 1,423 untested variants. We calculated the 'predicted change from jGCaMP7s' and performed an unpaired t-test between the 15 predictions made for each mutant (1 from each contributor model) and the 15 predictions made for jGCaMP7s. Using this method, we identified mutations that were predicted to substantially affect the $\Delta F/F_0$ and the kinetics of jGCaMP7s (Fig. 2a–c). In our model training, the jGCaMP7s sequence was purposely withheld. Nevertheless, the prediction ranked jGCaMP7s within the top 15% of variants for a large $\Delta F/F_0$ response. Consequently, the ensemble predicted most variants, such as L317E, L317K, L317N, L317D and L317H (<−2.2 a.u.),

https://doi.org/10.1038/s43588-024-00611-w

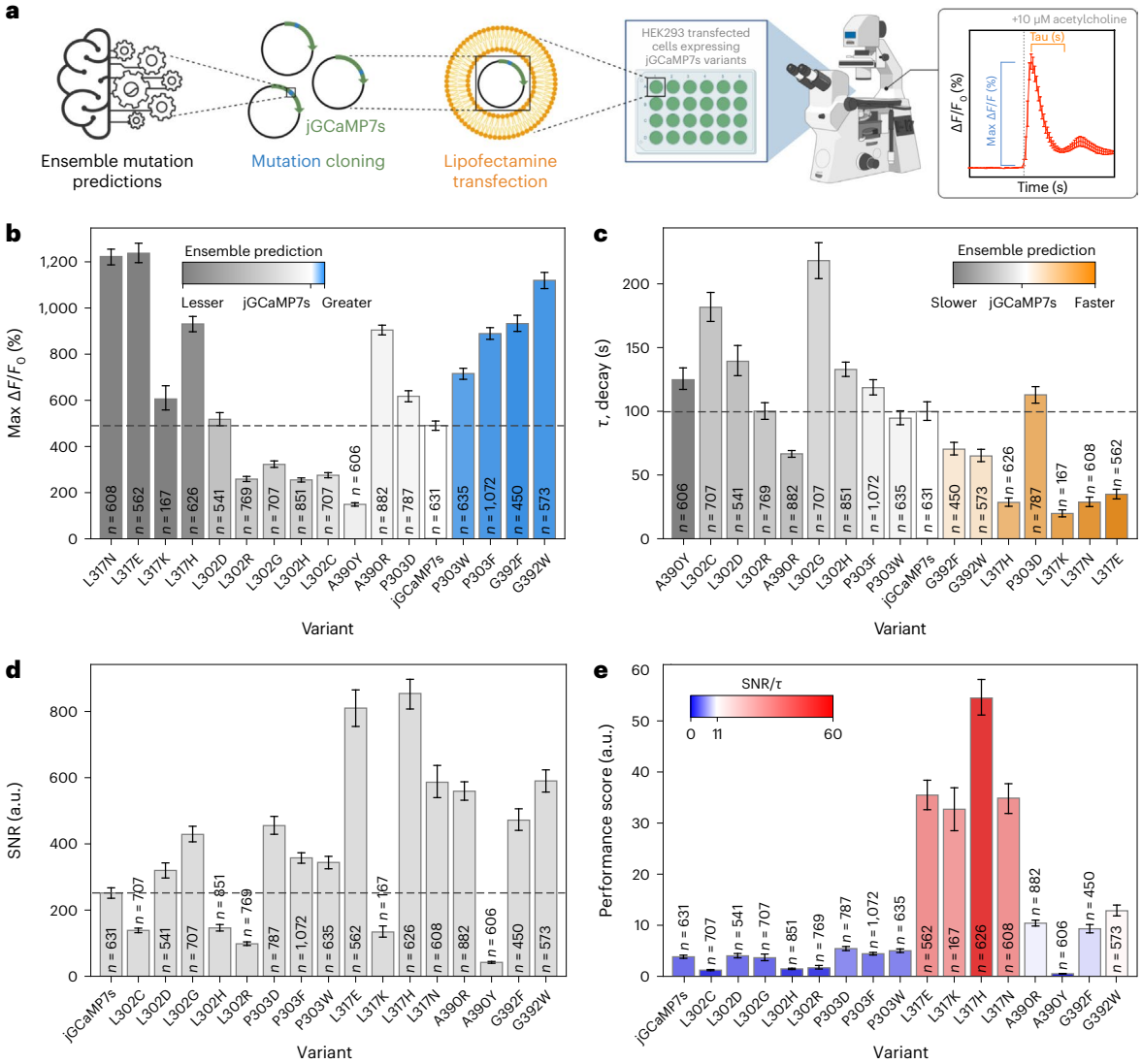


Fig. 3 | Gq/IP$_3$ assay in HEK293 cells to validate ensemble predictions. **a**, Mutation predictions isolated from the ensemble are used as the basis for downstream variant analysis. Variants of interest are cloned into the jGCaMP7s backbone, then transfected into HEK293 cells using lipofectamine transfection. After 48 hours post-transfection, cells are time-course imaged using an epifluorescent microscope. Panel a made using Biorender. **b**, Maximum fluorescence responses (equation (1)) were obtained from each variant of jGCaMP7s expressed in HEK293 cells and stimulated with 10 μM acetylcholine. Heat mapping indicates the ensemble's prediction. Mutations are sorted in order of the ensemble's predicted performance. The dotted line depicts mean performance of the base construct, jGCaMP7s. **c**, Decay values (τ; equation (4)) obtained from each variant of jGCaMP7s expressed in HEK293 cells and stimulated with 10 μM acetylcholine. Heat mapping indicates the ensemble's prediction. Mutations are sorted in order of the ensemble's predicted performance. The dotted line depicts mean performance of the base construct, jGCaMP7s. Visual representations of the qualifications in **b** and **c** are found on the representative response trace. **d**, SNR (equation (2)) of each mutant of jGCaMP7s expressed in HEK293 cells and stimulated with 10 μM acetylcholine. Mutations are sorted in ascending order based on residue number and final residue composition. The dotted line depicts mean performance of the base construct, jGCaMP7s. **e**, Performance score, SNR/τ (equations (2) and (4)), obtained from each mutant of jGCaMP7s expressed in HEK293 cells and stimulated with 10 μM acetylcholine. Heat mapping highlights the highest-scoring mutants or those with high $\Delta F/F_0$ (%) responses and fast decay speeds. Mutations are sorted in ascending order based on residue number and final residue composition. (n, number of cells quantified; bars show mean; error bars show 95% bootstrapped CI.)

to have a decreased $\Delta F/F_0$, while variants such as G392F, G392I and G392W were all predicted to have an increased (>0.25 a.u.) response (Fig. 2b). L317E, L317D, L317N and L317K were all predicted to decay faster (<−0.6 a.u.) than jGCaMP7s, while variants such as A390Y, L302D and L302C were predicted to decay slower (>0.3 a.u.; Fig. 2c). All these variants fell outside 99.7% (±3 standard deviations (σ)) of $-\log_{10}P$, except for large $\Delta F/F_0$ predictions, indicating that the 15 contributing models displayed confidence in the effect of the mutation (±3σ, $\Delta F/F_0$: 0.612; kinetics: 0.242; Fig. 2b,c).

Next, we identified the residues whose mutations had the strongest positive or negative impact on $\Delta F/F_0$ and kinetics. We isolated the top and bottom 2.5% of the ranked predictions and counted the times each residue appeared (Fig. 2a(ii)). We designated these as 'impactful residues,' as these residue positions were predicted to alter protein function substantially. We found that 22% and 18% of the impactful mutations in the $\Delta F/F_0$ and kinetics libraries were at L317 (Fig. 2d), despite only 1.3% of variants in the novel library harboring an L317 mutation. Similarly, L302 predictions accounted for 14% and 16% of the impactful mutations of the $\Delta F/F_0$ and kinetics libraries (Fig. 2d). L317 is located on the interface between CaM and CBP, and L302 is on the linker between CaM and cpGFP (Supplementary Fig. 3a–c). By contrast, residue A390 was 4.5 times more impactful in the kinetics

https://doi.org/10.1038/s43588-024-00611-w

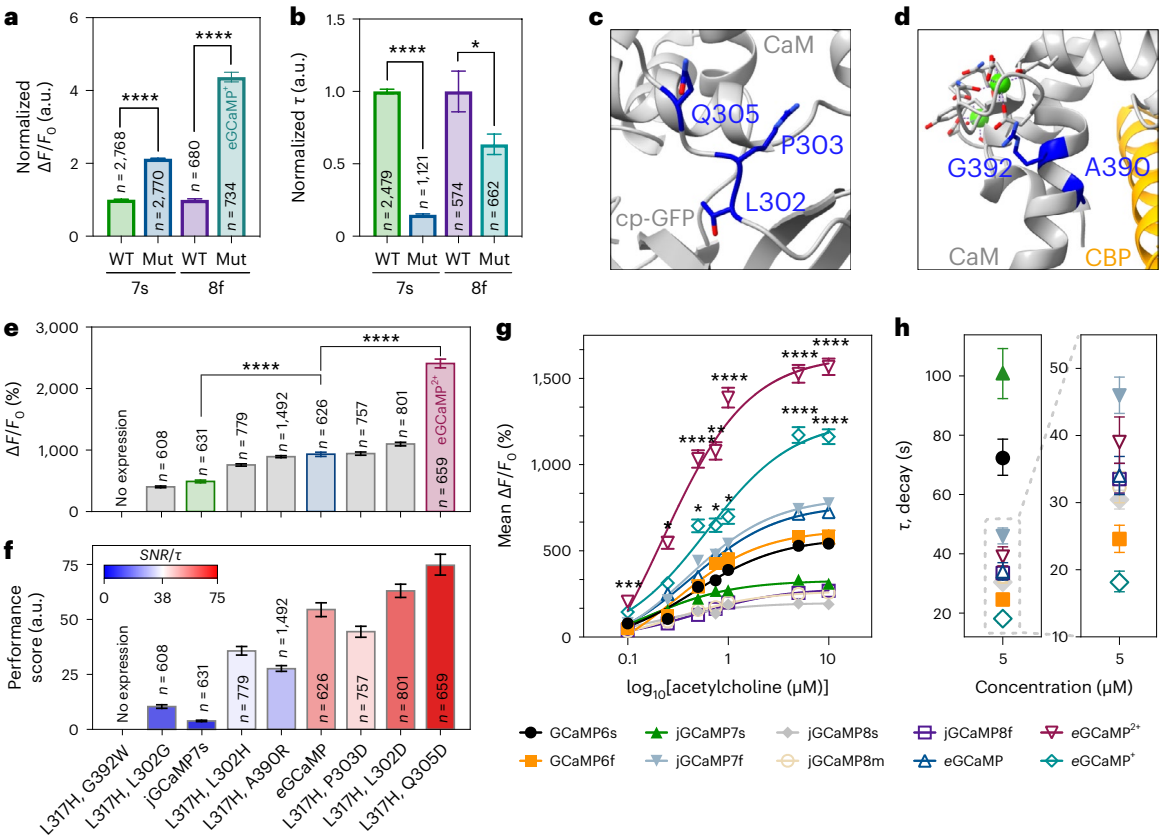


Fig. 4 | Identification of eGCaMP$^+$ and eGCaMP^{2+} in HEK293 cells. **a**, Fluorescent responses ($\Delta F/F_0$) upon 10 μM acetylcholine stimulation. WT indicates jGCaMP7s (7s) or jGCaMP8f (8f). Mut indicates L317H in jGCaMP7s or A289H in jGCaMP8f. Data are normalized to WT (*n*, number of cells; error bars are mean ± s.e.m.; ****$P < 0.0001$ (unpaired *t*-test, two-tailed)). **b**, Decay kinetics (τ) upon 10 μM acetylcholine. Mutation values (Mut) are normalized to WT (*n* = number of cells; error bars are mean ± s.e.m.; *$P = 0.0161$, ****$P = <0.0001$ (unpaired *t*-test, two-tailed)). **c**, Crystal structure of GCaMP3–D380Y (PDB: 3SG3, gray) with Q305 and linker residues P303 and L302 colored in dark blue, CaM and fluorescent reporter cpGFP. **d**, A390 and G392 are colored dark blue. Bound Ca^{2+} (green spheres) in the EF-Hand motifs and the CBP (orange) are included. **e**, $\Delta F/F_0$ of combinatorial mutations of jGCaMP7s upon 10 μM acetylcholine, sorted in order of performance and identified on the *x* axis of **f**. (*n*, number of cells; bars depict mean; error bars show bootstrapped 95% CI; ****$P < 0.0001$ (unpaired *t*-test, two-tailed)). **f**, Performance score (ratio of SNR/τ) for combinatorial jGCaMP7s mutations upon 10 μM acetylcholine. Mutations are sorted by $\Delta F/F_0$ performance. (*n*, number of cells; bars depict mean; error bars show bootstrapped 95% CI[49]). **g**, $\Delta F/F_0$ of GCaMP variants stimulated by different acetylcholine concentrations (*x* axis). eGCaMP is jGCaMP7s L317H, eGCaMP^{2+} is jGCaMP7s L317H/Q305D, and eGCaMP$^+$ is jGCaMP8f A289H. Error bars are mean ± s.e.m. Solid line depicts nonlinear data fits. (*$P < 0.05$; **$P = 0.0010$; ***$P = 0.0007$; ****$P < 0.0001$, unpaired *t*-test between variants and jGCaMP7f, two-tailed). **h**, Kinetic decay (τ; equation (4)) of the indicated variants stimulated with 5 μM acetylcholine. Plotted points indicate the mean τ for each variant to the indicated stimuli, and error bars are mean ± s.e.m. Four independent biological replicates per concentration + construct (**g**,**h**).

predictions than in the $\Delta F/F_0$ predictions. Like L317, A390 is located on the interface between CaM and CBP but on the opposing side (Supplementary Fig. 3d). Impactful residues also tended to cluster. For instance, the kinetics library displays 38% prediction prevalence surrounding residue clusters Y380, R381, R383 and L302, P303, Q305. The prevalence of these residues is 2.38-fold higher in kinetics than the $\Delta F/F_0$ predictions. These residues are located close to each other in the residue linker and CaM (Supplementary Fig. 3e). Within the $\Delta F/F_0$ predictions, residue clusters N44, K45, H48, V52 and M374, M378, K379 displayed 31% prediction prevalence, 3.9-fold higher than the kinetics library. Interestingly, we observed that all these residues face inward toward one another, suggesting they may be involved in interactions essential for $\Delta F/F_0$ (Supplementary Fig. 3f).

In vitro performance of ensemble predictions

We benchmarked 17 predicted mutations in vitro by stimulating human embryonic kidney cells (HEK293) cells with acetylcholine[2,3,24–26] (Fig. 3a). The ideal configuration would be to evaluate them in the same manner as the training data. However, owing to the lower throughput of cultured neuron screens, we first performed an intermediate acetylcholine assay step in HEK293 cells. We found the acetylcholine assay approximated variant performances accurately before cultured neuron assays (Supplementary Fig. 4a–f).

We identified four mutations (P303W, P303F, G392F and G392W) that displayed their predicted increase in $\Delta F/F_0$ as well as five mutations (A390Y, L302C, L302H, L302G and L302R) that displayed the predicted decrease compared with jGCaMP7s (Fig. 3b and Supplementary Table 3). The overall accuracy (equation (6)) of the $\Delta F/F_0$ model is 0.56 (Supplementary Fig. 5c and Supplementary Table 4). The score is largely affected by L317 mutations, which are predicted to decrease $\Delta F/F_0$ but display the opposite in vitro. Within the training data, the GCaMP variants that contained a 317E/H/K/N mutation had decreased $\Delta F/F_0$ compared with jGCaMP7s, an association in which the ensemble learned (Supplementary Fig. 6a). However, each previously characterized variant that contained a mutation at residue 317 also contained an alanine at residue 52 (Supplementary Table 6). When we tested the L317H variant in jGCaMP7f, which contains A52, we observed the loss of $\Delta F/F_0$ capabilities that the model predicted (Supplementary Fig. 6b). The mutations that changed kinetics largely aligned with the ensemble predictions, with an accuracy score of 0.75 (Fig. 3c, Supplementary Fig. 5).

Variants P303D, L317E, L317H, L317K, L317N, G392F and G392W were predicted to accelerate decay kinetics. Of these variants, 85%

https://doi.org/10.1038/s43588-024-00611-w

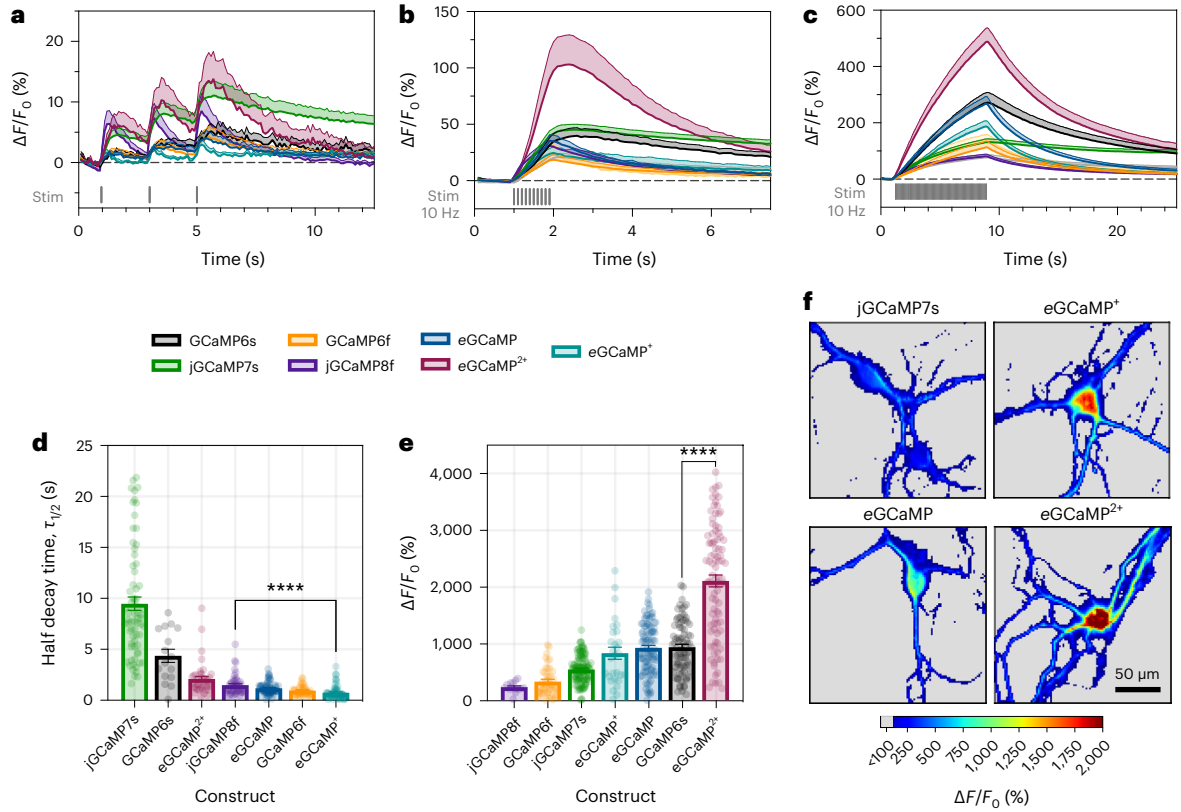


Fig. 5 | eGCaMP, eGCaMP$^+$ and eGCaMP^{2+} $\Delta F/F_0$ and kinetics characteristics in primary neurons. a, $\Delta F/F_0$ (%) recordings of each variant to one AP stimuli applied at 0.5 Hz over 6 seconds (lines depict mean, shading depicts s.e.m.). The applied stimuli are shown in gray. **b**, $\Delta F/F_0$ (%) recordings of each variant to 10 AP stimuli applied at 10 Hz over 1 second (lines depict mean, shading depicts s.e.m.). The applied stimuli are shown in gray. **c**, $\Delta F/F_0$ (%) recordings of each variant to 80 AP stimuli applied at 10 Hz over 8 seconds (lines depict mean, shading depicts s.e.m.). The applied stimulus is shown in gray. **d**, Half decay time values after 10 AP stimuli, scatter depicts neurons quantified. (bars depict mean; error bars show s.e.m.; $^*P = 0.045$ (unpaired t-test, two-tailed), four biologically independent samples per construct per concentration examined over four independent experiments)). **e**, Maximum $\Delta F/F_0$ (%) achieved after stimulation with 40 mM KCl. Error bars are mean ± s.e.m.; $^{****}P < 0.0001$ (unpaired t-test, two-tailed), 6 biologically independent replicates per construct over >2 independent experiments). **f**, Representative images of maximal fluorescence response to the 40 mM KCl stimulation variant indicated in **e**. Heat mapping displays $\Delta F/F_0$ (%) achieved by each pixel. Scale bar, 50 μm.

showed shorter decay times than jGCaMP7s, with L317K displaying a decay time that was fivefold faster than jGCaMP7s (Supplementary Table 5). Additionally, 71% of the variants predicted to decrease decay (L302C, L302D, L302G, L302H, L302R, A390R and A390Y) demonstrated the predicted behavior, with L302G exhibiting a decay time 2.18-fold longer than jGCaMP7s. Residue L317 is known to be involved in extensive hydrophobic interactions between CaM and CBP[27]. Thus, each mutation at L317 may destabilize the CaM and CBP interactions, accelerate kinetics and alter $\Delta F/F_0$ responses.

Several variants with large $\Delta F/F_0$ maintained a signal-to-noise ratio (SNR, equation (2)) 1.5-fold larger than jGCaMP7s (Fig. 3d and Supplementary Table 7). We created a performance score by dividing the SNR by the tau value of the decay (equations (2) and (4)) to highlight variants that combine both characteristics (Fig. 3e). L317H had the highest performance score, 14.23-fold greater than jGCaMP7s (Supplementary Table 8). Hence, we selected the jGCaMP7s L317H variant for in-depth characterization and named it 'ensemble-GCaMP' (eGCaMP). These results demonstrate that the ensemble could effectively predict enhanced sensor function while substantially reducing the experimental burden to identify variants with desirable biophysical characteristics.

Identification of eGCaMP$^+$ and eGCaMP^{2+}

We introduced the equivalent 317H mutation into the eighth-generation GCaMP, jGCaMP8f[6], testing if the beneficial effects alter divergent GCaMP iterations (Supplementary Fig. 7a). jGCaMP8f A289H improved the $\Delta F/F_0$ response fourfold over jGCaMP8f (Fig. 4a) with 36% faster decay (Fig. 4b). The fast decay kinetics and large $\Delta F/F_0$ responses provide a promising variant we named 'ensemble-GCaMP+' (eGCaMP$^+$).

Next, we tested a select combination of additional mutations on eGCaMP. We chose variants L302D, P303D, A390R and G392W for their increased $\Delta F/F_0$ in vitro (Fig. 3b). Other mutants were selected based on their locations. L302 and P303 are key residues in the linker between cpGFP and CaM[3,28] (Fig. 4c). Residue G392 forms a hydrogen bond with residue G398, which lies in one of the EF-hand domains and has been previously observed to influence the Ca^{2+} affinity[3,27] (Supplementary Fig. 3d), and A390 lies on the interaction face between CaM and CBP (Fig. 4d). We tested Q305 due to its proximity to the linker residues (Fig. 4c), hydrogen bonding interactions with Y380 (Supplementary Fig. 3e), and prevalence in the impactful residues for kinetics (Fig. 2d). All combinations, except for L317H/G392W, led to functional proteins (Fig. 4e,f and Supplementary Fig. 7b). On average, all variants exhibited decay times fivefold faster than jGCaMP7s and 50% displayed equal or improved $\Delta F/F_0$ responses to those of eGCaMP (Supplementary Fig. 7b and Supplementary Table 9). We observed the largest $\Delta F/F_0$ in the L317H/Q305D, named eGCaMP^{2+}, with an almost 2.5-fold increase in $\Delta F/F_0$ over eGCaMP and a fivefold increase over jGCaMP7s (Fig. 4e and Supplementary Table 10). The variant also achieved the highest performance score, a 1.36-fold increase over eGCaMP (Fig. 4f and Supplementary Fig. 7b and Supplementary Tables 11 and 12). The good dynamic range may result from intraprotein interactions within CaM. One possible explanation is that the decreased R-group length in the

Q305D mutation requires a more substantial conformational change to form the hydrogen bond with residue Y380 (Supplementary Fig. 3e). The resulting conformational change may have downstream effects on both the cpGFP/CaM linker (Fig. 4a) and R381, which faces inward toward the chromophore (Supplementary Fig. 3e). The dramatic effects of this mutation suggest a collaborative role between the cpGFP/CaM linker and the inward loop of CaM in stabilizing the phenol/phenolate transition of the chromophore[29–31].

We benchmarked the biophysical and photophysical properties of eGCaMP, eGCaMP^{2+}, and eGCaMP^{+} against published variants, including widely used constructs such as GCaMP6s, GCaMP6f, jGCaMP7s, jGCaMP7f, jGCaMP8s, jGCaMP8m and jGCaMP8f[4–6]. The excitation and emission spectra of the eGCaMP variants remained unchanged from the previously published GCaMPs, with excitation peaks at ~495 nm and emission peaks at ~515 nmv (ref. 5 and Extended Data Fig. 1a–d). We found that eGCaMP, eGCaMP^{+} and eGCaMP^{2+} had lower baseline fluorescence than GCaMP6s, jGCaMP7s and jGCaMP8f (Fig. 4f and Supplementary Fig. 8a–d). The three ensemble variants demonstrated good $\Delta F/F_0$ responses and SNRs in the acetylcholine assays (Fig. 4g and Supplementary Fig. 7c). At every tested concentration, they maintained a larger $\Delta F/F_0$ than all previous variants (Fig. 4g and Supplementary Tables 13–19). For example, eGCaMP^{2+} achieved 2.5-fold greater $\Delta F/F_0$ values at 0.1 µM acetylcholine than the highest-performing previous variant, with decay times comparable to jGCaMP7f (Fig. 4g,h and Supplementary Table 18). Additionally, the decay time of eGCaMP^{+} was the fastest of all tested variants (46% faster than jGCaMP8f), while the maximum $\Delta F/F_0$, was second only to eGCaMP^{2+} (Fig. 4g,h and Supplementary Table 20). eGCaMP achieved a $\Delta F/F_0$ close to jGCaMP7f but with a 26% faster decay (Fig. 4g,h and Supplementary Tables 13 and 20). Using purified proteins, we found that the eGCaMP and eGCaMP^{2+} variants achieved similar dissociation constants (K_d) to those published for jGCaMP8f[6] (Supplementary Table 21). eGCaMP^{+} displayed a K_d shift to the micromolar range, consistent with previously published studies finding a tradeoff between sensitivity and kinetics[4–6] (Supplementary Table 21). The eGCaMPs had slightly diminished extinction coefficients compared with GCaMP6f but displayed larger quantum yields (Supplementary Table 21).

eGCaMP, eGCaMP^{+} and eGCaMP^{2+} performance in primary neurons

We benchmarked the eGCaMPs against previous variants in cultured primary neurons stimulated by extracellular electrical fields to evoke APs[4,5,32]. eGCaMP^{2+} displayed a $\Delta F/F_0$ of 10.1% in response to 1 AP, similar to jGCaMP8f (Fig. 5a and Supplementary Table 22). At 10 AP, jGCaMP8f saturated quickly, while eGCaMP^{2+} achieved a 2.34-fold larger response than jGCaMP7s. At 80 AP, eGCaMP^{2+} achieved a 1.82-fold larger $\Delta F/F_0$ than GCaMP6s (Fig. 5b,c and Supplementary Tables 23 and 24). The average $\Delta F/F_0$ of eGCaMP^{2+} full saturation by 40 mM KCl was 1938%, which is twofold larger than GCaMP6s (Fig. 5e and Supplementary Table 25). While the KCl saturation responses were quantified in the cell body, the proximal projections in eGCaMP^{2+} still maintained >1,000% $\Delta F/F_0$ increases (Fig. 5f). At 80 AP trains, both eGCaMP and eGCaMP^{+} achieved higher $\Delta F/F_0$ response amplitudes than the previously published fast variants GCaMP6f and jGCaMP8f (Fig. 5c and Supplementary Table 24). These results are compounded by both eGCaMP and eGCaMP^{+} achieving 10 AP half decay times ($\tau_{1/2}$) of 1.17 s and 0.74 s for each variant, respectively, which is faster than jGCaMP8f's, whose 10 AP half decay time was 1.49 s (Fig. 5d and Supplementary Table 26). Furthermore, eGCaMP decayed eightfold faster than jGCaMP7s, as well as a diminished response to 1AP stimulus, highlighting the ability of the ensemble to correctly predict the single point mutation's functional effect (Fig. 5a,d and Supplementary Tables 22 and 26).

eGCaMP^{+} and eGCaMP^{2+} performance in vivo

Next, we benchmarked eGCaMP^{2+} and eGCaMP^{+} in vivo, against GCaMP6f. We injected each variant of Cre-dependent GCaMP virus in the medial prefrontal cortex (mPFC), and a retrograde Cre virus in the nucleus accumbens (NAc; Extended Data Fig. 2a). This labeled a relatively sparse population of mPFC to NAc projections neurons with the GCaMP sensor. An optical fiber was implanted above the mPFC to measure the GCaMP fluorescence signal in response to brief foot shocks, which has been previously shown to elicit responses in these neurons[33]. Histology images showed qualitatively similar GCaMP expression in mPFC cell bodies and axons in NAc across all groups of mice (Extended Data Fig. 2b). All three GCaMP variants exhibited a time-locked increase in fluorescence during the foot shock, followed by a slow decay in the sensor fluorescence (Extended Data Fig. 2c). We calculated the mean response to the foot shock for each sensor and found that eGCaMP^{2+} exhibited a larger change in response compared with GCaMP6f and eGCaMP^{+} (Extended Data Fig. 2d), similar to our results in culture. We also calculated the mean sensor decay response 3 to 4 s after the foot shock (which is dependent on the off kinetics of the sensor; a faster sensor will return to baseline after the shock, while a slower sensor will still maintain a higher signal after the shock). We found that the eGCaMP^{+} decay response was smaller than that of eGCaMP^{2+}, supporting our previous findings the eGCaMP^{+} has the fastest off kinetics among our sensors (Extended Data Fig. 2e).

Discussion

Incorporating ML into our engineering pipeline enabled us to efficiently identify new GCaMP variants with enhanced $\Delta F/F_0$ responses and faster decay kinetics. We achieved good predictive performance in the cross-validation phase by using an ensemble of three regressor models, encoding our dataset with amino acid characteristics, and focusing solely on sequence inputs for learning. These predictive capabilities translated to the in vitro space, where many in silico predicted characteristics accurately reflected the mutant's true performance. As a result of these engineering efforts, we identified three new GcaMP constructs: eGCaMP, eGCaMP^{+} and eGCaMP^{2+}.

We made several critical design decisions when developing our approaches, such as our encoding method, chosen models, ensemble and exploring only single-point mutations. We chose to encode the sequence with biophysical properties underlying amino acids in each position to facilitate the formation of meaningful learning patterns. The selection of five amino acid datasets was made semi-arbitrarily, as it provided better learning capabilities over other common encoding methods without dramatically impacting computational demands, processing time and storage requirements.

Ensembling ML models (considering the input from multiple models) is preferable over single-model predictions, as no singular model is perfectly optimized to perform all tasks[34]. The three selected models have diverse learning strategies and make different assumptions about the data, which is important when ensembling. Decision tree learning methods, such as RFRs, are computationally efficient models well suited for small training libraries, making them a strong foundation within our ensemble's learning[35]. The KNR[36] similarity metric can capture the variability between the performances of nearly identical sequences and highlights residues whose mutation led to large differences in the targeted sensor characteristics. MPNRs are deep-learning models capable of extracting high-level features from the data, making them useful for identifying key residues or properties that lead to the observed biophysical response[35]. Adding more models improved our predictive capability, as demonstrated by the increased R^2 of the ensemble compared with the sole contributor's performance.

One of the major hurdles of protein engineering is the susceptibility of proteins to experience epistasis, in which combinations of mutations non-additively influence the phenotypic characteristics[37]. Though the mutation library we worked with had more than a thousand well-characterized variants, the large number of mutated residues renders the dimensionality incredibly large. As such, we felt that the risk of epistasis upon combinatorial mutation was too great

and that the relatively limited size of the library in comparison to its dimensionality rendered this application better suited to single-point mutation testing.

While the incorporation of ML substantially improved our engineering capabilities, there were several limitations within our study. Within our analysis of model results, we opted to perform a high-throughput acetylcholine assay step within HEK293 cells to approximate sensor performance before transitioning promising variants to neuron culture screens. Ideally, the model predictions should be validated against data acquired in the same manner as the training data. The HEK293 assay approximates sensor capabilities but does not fully reflect the predictions made by the model. Importantly, these differences did not compromise the ensemble's ability to guide our engineering efforts; however, validating the ML-ensemble prediction using the same host system as the training data could offer a more complete understanding of the model's predictive capabilities.

Although structural insights guided the engineering of previously published GCaMPs, we developed the ensemble pipeline to be structure-agnostic. This was crucial, as we aim to engineer GEFIs without relying on molecular structures. Owing to the exclusion of structure information, extrapolation outside of the observed sequence space may be difficult. Our approach is best suited for exploration within a sequence space with only minor variations from the training data. However, one could incorporate spatial information from crystal structures or structure predictions in the ensemble's learning to aid extrapolation in future studies.

The approach we employ here allowed the majority of the screening to occur in silico, which reduced the experimental burden while achieving substantial improvements. We selected variants for in vitro testing based on their predicted performance for $\Delta F/F_0$ or off-kinetics. Thus, the selected variants displayed compensation within favorable characteristics, such as a lower baseline fluorescence, which was not a prediction criterion. The lower baseline did not impact the performance of eGCaMPs in neuron cultures or in vivo fiber photometry. Hence, it would be an acceptable tradeoff in many use scenarios. As a consideration for future studies, metrics for favorable characteristics could be included in ensemble training to preserve them within the final variants.

Lastly, our training dataset was more biased toward influential residues, which were chosen through crystal structure analysis and previous trial and error. As a result, highly mutated positions came to the forefront of final predictions; however, this does not mean that they are not influential or that the mutations that the ensemble suggests cannot be exploited further. Likewise, it did not preclude less explored residues from being chosen as influential in sensor performance. As ML becomes more prevalent, several considerations for data acquisition may help generate better suited, unbiased datasets. First, sequence space and dimensionality must be well defined. Smaller dimensionality offers more in-depth analysis and comprehension of combinatorial mutations. Larger numbers of residue positions will span a much greater sequence space but limit the study to small iterations from the starting sequence. The data should have equal numbers of mutations per residue in their characterization to avoid biases that may arise due to unbalanced prevalence. Furthermore, identifying 'loss-of-function' mutations is as vital to training as 'gain-of-function' mutations. The use case of iterative model training, in which the user is informed by ML and then retrains the model with additional information, is an ideal application. However, testing only promising variants should be avoided, as this may introduce bias into the dataset during retraining.

The ML ensemble used in this study has demonstrated a good capacity to guide fluorescent biosensor engineering. The ensemble's predictions helped identify variants with large $\Delta F/F_0$ values and fast decay kinetics, while highlighting clusters of impactful residues for each biophysical property, which may be further exploited by mutation-library-based high-throughput screening. These findings illustrate the ensemble's ability to guide engineering efforts and improve experimental efficiency. Moreover, since our model's learning is based solely on the sequence–function relationship and all contributor model optimization is unbiased, the final ensemble platform can be broadly applied to any genotype-to-phenotype mutation library. Applying this ML platform to mutation studies of proteins with quantifiable output characteristics, including other protein sensors, has the potential to accelerate the engineering of these proteins.

Methods

Data preprocessing

The Chen and Dana studies provide a functional characterization of more than 1,000 GCaMP variants that span the GCaMP6 and jGCaMP7 iterations[4,5,20]. Each study normalized the results to base constructs for data such as $\Delta F/F_0$ response (equation (1)) to stimuli of 1 AP, 3 AP, 10 AP and 160 AP, and decay half-time after 10 AP. To cross-compare mutation libraries, we re-normalized the Chen et al. 2013 dataset such that GCaMP6s was 1.0 for all metrics. Each variant was incorporated into a single dataframe that comprised 453 columns: one column containing the primary key variant identifier, 451 columns corresponding to the sequence of each GCaMP variant, and the final column containing each variant's empirically derived performance.

$$\Delta F/F_0 = \frac{(F - F_0)}{F_0} \times 100 \quad (1)$$

Within the variant library used for model training, the independent variable consists of the sequence of each mutation. The dependent variable is the $\Delta F/F_0$ response (1AP $\Delta F/F_0$) or kinetics capability ($\tau_{1/2}$). To encode our independent variable data, we attempted several different methods: label encoding, one-hot encoding or by adding functional information. Within our label encoding, we randomly assigned an integer value to each amino acid and replaced each residue label in the GCaMP sequence with the dummy label. For one-hot encoding, the full extent of possible residues at each position is considered in a Boolean manner (20 amino acids × 450 residue positions). To perform encoding with functional data, we developed a dictionary of amino acid properties by web scraping the AAindex database[21]. AAindex consists of matrices that each describe a different amino acid property, such as size, polarity or hydrophobicity. The general shape and composition of each one of the property datasets is a list of 20 float type values, in which the order is linked to the amino acid, and the float type value is a quantitative value that is dependent on the property in question. We used the 554 complete property datasets to formulate an unbiased model training paradigm in two steps. To perform the encoding, we replace each amino acid in the sequence with the corresponding value from the property dataset. The final variant library used in model training consisted of the fully encoded GCaMP sequence and the variants' empirically derived performance capability.

Generation of the novel variant library

To generate a library of unknown sequences, we performed a single-point saturation of the jGCaMP7s sequence at 75 residue locations. These 75 residues correspond to the 75 residues that contain mutagenesis information in the variant library. The outcome was a novel point-saturation-mutation library that contained 1,423 untested variants (77 redundant with previously tested variants).

Ensemble training

The models that we developed were from the pip installable package Scikit Learn in Python 3.8.5 to develop an RFR, KNR and an MPNR and developed in Jupyter Notebooks and Google Colab. The models were trained on the encoded sequence of each variant linked to their empirically derived performance capability. The performance capabilities correspond to their $\Delta F/F_0$ response to 1 AP or half decay time after 10 AP. The data were split into train/test sets at a ratio of 80:20 with a random seed of 42 for downstream optimization efforts. We performed the

'SelectKBest' feature selection function found in Scikit Learn. Optimization of the model was done by grid-search hyperparameter tuning, unique to each model type. We used the coefficient of determination (R^2) and mean squared error to track the fit of each model. This process was repeated over each of the 554 datasets for the three models (~1,662 times). Each model's top five performing property datasets (that led to the highest R^2 values) were advanced to generate predictions on the novel variant library. Each contributor model (5 amino acid properties × 3 regressor models) forms predictions independently, and the final predictions are the average response from each contributor model for each target attribute ($\Delta F/F_0$ response 1AP $\Delta F/F_0$ or kinetics capability $\tau_{1/2}$). The predicted values returned by the ensemble are numeric values originating from a normalized library, making the predictions unitless.

PCA clustering

Each feature within the data was first scaled using Sklearn's StandardScaler. We passed the scaled data into Sklearn's PCA function with no defined number of components. We chose the optimum number of components by finding where the explained variance of the PCA of the data passed 0.8. We reinitialized the PCA with the determined number of principal components and fit the function with the standardized data. We then used the principal component space coordinates to find the ideal number of clusters for *K*-means clustering. We determined the ideal number of clusters by using the 'elbow method' on the within cluster sum of square. After finding the clusters, we labeled each input to their *K*-means-defined cluster.

Molecular cloning

Predicted mutations were reflected into the cytomegalovirus-jGCaMP7s backbone (Addgene ID: 104463) using point-mutation primers (Integrated DNA Technologies) and PCR amplification with either Q5-polymerase (New England Biolabs; M0492L) or Superfi-II polymerase (Invitrogen; 12368010). Amplification of the DNA fragment was verified with agarose gel electrophoresis. Blunt-end DNA circularization was achieved with Kinase, Ligase, and DpnI enzyme treatment (New England Biolabs: E0554S). Circularized DNA was transformed into competent *Escherichia coli* cells (DH5α or TOP10) and grown on agar plates that contain either ampicillin or kanamycin selection antibiotic (50 μg ml^{-1}). Upon colony formation, single colonies were picked and grown in 5 ml cultures containing LB Broth (Fisher BioReagents; BP9723-2) and selection antibiotic (ampicillin or kanamycin; 50 μg ml^{-1}) overnight (37 °C, 230 rpm). DNA was isolated using DNA prep kits (Machery Nagel; 740490.250). Sanger sequencing (Genewiz) of the isolated plasmid DNA was used to confirm the presence of the intended mutation.

Genes encoding the GCaMP variants were cloned into a CAG-driven backbone, pCAG-Archon1-KGC-EGFP-ER2-WPRE (Addgene; 108423), using Gibson assembly (New England Biolabs; E2621L). All subsequences were verified with Sanger sequencing (Genewiz).

Acetylcholine assays

Human embryonic kidney cells (HEK293; ATCC Ref: CRL-1573, passages 3 to 25) were cultured in Dulbecco's Modified Eagle Medium + GlutaMAX (Gibco; 10569-010) supplemented with 10% fetal bovine serum (Biowest; S1620). When cultures reached 85% confluency, the cultures were seeded at 100,000 cells per well or 50,000 cells per well in 24-well or 48-well plates, respectively. At 24 hours after cell seeding, the cells were transfected using Lipofectamine3000 (Invitrogen; L3000015) at 1,000 ng of DNA per well of a 24-well plate, according to the manufacturer's instructions. HEK293 cells were authenticated by ATCC before shipping with short tandem repeat profiling following ISO 9001 standards. At 48 hours post-transfection, the plates were prepared for imaging by washing and then replacing culturing media volume with imaging solution (Tyrode's pH = 7.33; 125 mM NaCl, 2 mM KCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 30 mM dextrose, 25 mM HEPES (triple supplemented with 1% GlutaMAX (Gibco; 35050-1), 1% sodium pyruvate (Gibco; 11360-070) and 1% MEM non-essential amino acids (Gibco; 11140-050)). Crystalline power acetylcholine chloride (Alfa Aesar; L02168.14) was resuspended into imaging solution (Tyrode's pH = 7.33; 125 mM NaCl, 2 mM KCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 30 mM dextrose, 25 mM HEPES) into 2× the desired final concentration. During imaging, 1:1 volumes of the acetylcholine-tyrodes imaging solution were hand-pipetted into the bath volume to bring the final acetylcholine concentration to the desired concentration. Imaging was performed on an sCMOS camera (Photometrics Prime95B) on an epifluorescent microscope (Leica DMI8) using a ×20 objective (Leica HCX PL FLUOTAR L 20x/0.40 NA CORR) controlled by Metamorph software. A Lumencor Light Engine LED and Semrock Filters (Excitation: FF01-474-27; Emission: FF01-520/35) were used for fluorescence imaging.

Analysis of fluorescence assays

Analysis of HEK293 cell fluorescence imaging data was done by FUSE, a custom cloud-based semi-automated time series fluorescence data analysis platform written in Python. First, the cell segmentation quality of the selected Cellpose[38] model was manually verified. For the segmentation of cells expressing cytosolic fluorescent indicators, model 'cyto' was selected as our base model. If the selected Cellpose model was low performing, we further trained the Cellpose model using the Cellpose 2.0 human-in-the-loop system[39]. Using an 'optimized' segmentation model, fluorescence time-series data are extracted for each region of interest. Using the raw fluorescence data, percentage fluorescence change from the baseline ($\Delta F/F_0$) over time was calculated using equation (1). The SNR was calculated using equation (2).

$$\mathrm{SNR} = \frac{(F_{\max} - F_0)}{\text{standard deviation}(F_0)} \tag{2}$$

The exponential decay constant (λ) was calculated using equation (3), where $F(t)$ is the change in fluorescence at a time (t) after the maximum fluorescence (F_0) was achieved. Importantly, F_0 was normalized to 1.0, such that $F(t)$ depicts the change in fluorescence over time, t.

$$F(t) = F_0 e^{-\lambda t} \tag{3}$$

The exponential time constant (τ) was isolated by using the known reciprocal relationship of λ and τ.

$$\tau = \frac{1}{\lambda} \tag{4}$$

The dynamic range (DR) was defined as the ratio of the maximum fluorescence intensity to the baseline fluorescence intensity (equation (5)). All $\Delta F/F_0$, SNR, τ and DR values were quantified using a custom python script.

$$\mathrm{DR} = \frac{F_{\max}}{F_0} \tag{5}$$

To calculate the accuracy of our model, we classified kinetics predictions into variants that are either faster or slower than jGCaMP7s, and $\Delta F/F_0$ predictions into variants containing a larger or smaller $\Delta F/F_0$ than jGCaMP7s. To evaluate our model's performance, we computed an accuracy score (equation (6)) using the empirical data, which is the ratio of sum of the true positives (TP) and true negatives (TN) to the total number of predictions.

$$\text{Accuracy score} = \frac{\mathrm{TP} + \mathrm{TN}}{n_{\text{predictions}}} \tag{6}$$

To calculate the precision of our model, we classified kinetics predictions into variants that are either faster or slower than jGCaMP7s, and $\Delta F/F$ predictions into variants containing a larger or smaller $\Delta F/F_0$ than jGCaMP7s. To evaluate our model's performance, we calculated the precision (equation (7)) of our models using the empirical data,

which is the ratio of number of TP over the number of TP and false positives (FP).

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (7)$$

Optical properties of purified proteins

Proteins were purified by large-scale nickel column protein purification and SEC purification[40,41]. Purified protein isolates were diluted to 10 µM in 30 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS), 100 mM KCl, pH 7.2 with either 10 mM ethylene glycol-bis(β-aminoethyl ether)-*N*,*N*,*N*′,*N*′-tetraacetic acid (EGTA) (low Ca^{2+}) or 10 mM CaEGTA buffers (high Ca^{2+}) (Invitrogen; C3008MP). Protein absorbance spectra were recorded for each condition using a UV-vis spectrophotometer (NanoDrop 2000/2000c Spectrophotometers; Thermo Scientific). Fluorescence emission and excitation spectra for each condition were measured with a spectrum-capable plate reader (SpectraMax M5; Molecular Devices).

For calcium titrations, GCaMP protein was first diluted (0.5 µM protein) in triplicate in high-Ca^{2+} or low-Ca^{2+} buffers. These two solutions were mixed in various ratios to give 11 different free calcium concentrations (Invitrogen; C3008MP). GCaMP fluorescence (excitation 485 nm; emission 535 nm) was measured using a SpectraMax M5 (Molecular Devices). Calcium titration curves were fit (Prism; GraphPad) to sigmoidal binding functions, and the Hill coefficient and K_d for Ca^{2+} binding for the GCaMP variants were extracted.

The absorbance under saturating conditions was measured using 2 µM protein diluted into high-Ca^{2+} buffer at 500 nm (DU800 spectrophotometer; Beckman Coulter). The chromophore concentration was measured from the absorbance (447 nM) of protein denatured by 1 M NaOH (extinction coefficient 44,000 M^{-1} cm^{-1}). The extinction coefficient was calculated using Beer's law, where the absorbance was that of the saturated protein at 500 nm, and the concentration was extracted using the absorbance of the denatured protein.

Quantum yield measurements were measured at 460 nm light using an integrating-sphere spectrometer (Hamamatsu) for 0.3 µM protein diluted in high-Ca^{2+} buffer.

Isolation of cortical neurons

Tissue culture plates with 24 wells were coated with matrigel (mixed 1:20 in cold-phosphate-buffered saline (PBS), Corning; 356231) solution and incubated at 4 °C overnight before use. Sterile dissection tools were used to isolate cortical brain tissue from P0 rat pups (male and female, Sprague Dawley, Envigo). Tissue was minced until 1 mm pieces remained, then lysed in equilibrated (37 °C, 5% CO_2) enzyme (20 U ml^{-1} Papain (Worthington Biochemical; LK003176) in 5 ml of EBSS (Sigma; E3024)) solution for 30 minutes at 37 °C, 5% CO_2 humidified incubator. Lysed cells were centrifuged at 200*g* for 5 minutes at room temperature, and the supernatant was removed before cells were resuspended in 3 ml of EBSS (Sigma; E3024). Cells were triturated 24× with a pulled Pasteur pipette in EBSS until homogenous. EBSS was added until the sample volume reached 10 ml before spinning at 0.7*g* for 5 minutes at room temperature. Supernatant was removed, and enzymatic dissociation was stopped by resuspending cells in 5 ml EBSS (Sigma; E3024) + final concentration of 10 mM HEPES buffer (Fisher; BP299-100) + trypsin inhibitor soybean (1 mg ml^{-1} in EBSS at a final concentration of 0.2%; Sigma, T9253) + 60 µl of fetal bovine serum (Biowest; S1620) + 30 µl 100 U ml^{-1} DNase1 (Sigma;11284932001). Cells were washed twice by spinning at 0.7*g* for 5 minutes at room temperature, removing supernatant and resuspending in 10 ml of Neuronal Basal Media (Invitrogen; 10888022) supplemented with B27 (Invitrogen; 17504044) and glutamine (Invitrogen; 35050061) (NBA++). After final wash, spin and supernatant removal, cells were resuspended in 10 ml NBA++ before counting. Just before neurons were plated, matrigel was aspirated from the wells. Neurons were plated on the prepared culture plates at desired seeding density. At 24 hours after plating, 1 µM AraC (Sigma; C6645) was added to the NBA++ growth medium to prevent the growth of glial cells. Plates were incubated at 37 °C and 5% CO_2 and maintained by exchanging half of the media volume for each well with fresh, warmed Neuronal Basal Media (Invitrogen; 10888022) supplemented with B27 (Invitrogen; 17504044) and glutamine (Invitrogen; 35050061) every three days[42,43].

Calcium phosphate transfection of primary cortical neurons

Isolated primary cortical neurons were transfected using the calcium phosphate transfection kit from Sigma Aldrich (Sigma-Aldrich; CAPHOS-1KT). Half of the neuron media was changed 24 hours before transfection, saving the removed conditioned media to add to the neurons after transfection. Reagents were mixed in a ratio of 3 µl $CaCl_2$:24.5 µl H_2O:1,000 ng DNA before being added dropwise to bubbled 2× HEPES buffered saline (30 µl). The final solution was vortexed for 4 seconds and left undisturbed for 20 minutes. The solution was added dropwise to each well of neurons in a 24-well plate and shaken to distribute equally. Neurons were left to incubate for 1 hour at 37 °C with 5% CO_2. The cells were rinsed twice with HBSS before adding the conditioned medium removed on the previous day and mixed with half-fresh medium.

Electrical field stimulation

On the day of imaging, ~24–36 hours post-transfection, cells were washed once with imaging solution and then transferred to E-Stim Tyrode's (pH = 7.33; 150 mM NaCl, 4 mM KCl, 3 mM $CaCl_2$, 1 mM $MgCl_2$, 10 mM Dextrose, 10 mM HEPES)[32]. A custom wire holding piece was designed to fit into 48-well plates with silver wires 10 mm apart. 100 mA pulses, with a 3 ms pulse width, were administered at either 0.5 Hz or 10 Hz frequency using a pulse generator (Warner Instruments; SIU-102B), triggered with Sutter Instruments Integrated Patch Amplifier with Patch Panel, time-locked using Igor Pro 8. Imaging was performed with a digital camera (Hamamatsu ORCA-Flash4.0; C11440) at 100 ms exposure attached to an epifluorescent microscope (Leica DM IL). The light was generated using a SOLA Light Engine (Lumencor; SOLA SE 5-LCR-SB) with a 488 nm wavelength filter lens. Bulk fluorescence traces were acquired using FIJI imaging software with background subtraction (rolling = 50 stack) and hand-drawn regions of interest. The baseline was defined as the first 50 measurements before the event trigger. Maximum $\Delta F/F_0$ was determined by finding the maximum value within each $\Delta F/F_0$ trace. The $\Delta F/F_0$ traces were then normalized to have a maximum value of 1.0 by dividing each value in the trace by the maximum $\Delta F/F_0$. The half-decay time was recorded as the time in which the normalized $\Delta F/F_0$ passed 0.5. Final traces and plots were created using Prism9.

Potassium chloride assays

On the day of imaging, ~24–36 hours post-transfection, cells were washed once with imaging solution, then replaced with imaging solution (Tyrode's pH = 7.33; 125 mM NaCl, 2 mM KCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 30 mM dextrose, 25 mM HEPES (triple supplemented with 1% GlutaMAX (Gibco; 35050–1), 1% sodium pyruvate (Gibco; 11360-070), and 1% MEM non-essential amino acids (Gibco; 11140-050)). Powdered potassium chloride (Sigma; P9541-500G) was diluted in double distilled water to a concentration of 2 M. This solution was then diluted to 80 mM in imaging solution (Tyrode's pH = 7.33; 125 mM NaCl, 2 mM KCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 30 mM dextrose, 25 mM HEPES). During imaging, 1:1 volumes of KCl solution were hand-pipetted into the bath to bring the final KCl concentration to the desired concentration. Imaging was performed on an sCMOS camera (Photometrics Prime95B) on an epifluorescent microscope (Leica DMI8) using a ×20 objective (Leica HCX PL FLUOTAR L 20x/0.40 NA CORR). A Lumencor Light Engine LED, and Semrock filters (excitation: FF01-474-27; emission: FF01-620/35) were used for fluorescence imaging. Bulk fluorescence traces were acquired using FIJI imaging software with background subtraction (rolling = 50 stack) and hand-drawn regions of interest. The baseline was

defined as the first 30 measurements before KCl addition. Maximum $\Delta F/F_0$ values were obtained using a custom Python script. Final traces were plotted in Prism9.

Animals
Male and female C57BL/6 J mice (6–7 weeks old) were obtained from the Jackson Laboratory and maintained on a 12 h reverse light–dark cycle (lights on at 21:00) at 22 °C, group-housed with same-sex cage mates and given ad libitum access to food and water. Mice were left undisturbed for 1 week following arrival before the start of testing. All experiments occurred in the dark cycle. All experiments were conducted in accordance with the UC Davis Institutional Animal Care and Use Committee.

Stereotaxic surgery
Mice were anesthetized under 1.5–2% isoflurane and placed in a stereotaxic apparatus (RWD) on a heat pad. Three different adeno-associated virus (AAV) cre-dependent GCaMP variants were tested: AAV5-Syn-FLEX-GCaMP6f (Addgene 100834; final titer 1.1×10^{13} genomic copies per ml) AAV1-EF1a-DIO-eGCaMP^{+} (Fred Hutch Virus Core; final titer 1.25×10^{12} IU ml^{-1}); or AAV1-EF1a-DIO-eGCaMP^{2+} (Fred Hutch Virus Core; final titer 6.80×10^{11} IU ml^{-1}) (IU, units of infectious particles). AAV cre-dependent GCaMP variant (1 µl) was infused into the mPFC (M/L: −0.35, A/P: 1.98, D/V: −2.25 mm relative to bregma), and 500 nl of retroAAV-Syn-Cre (Addgene 105553; final titer 9.50×10^{12} GC ml^{-1}) was infused into the nucleus accumbens (NAc; M/L: −0.35, A/P: +1.25, D/V: −4.6 mm). Injections were performed at a rate of 150 nl min^{-1} using a Hamilton syringe controlled by an injection pump (World Precision Instruments). The virus was allowed to diffuse for 5 min before withdrawing the needle. Chronically implantable fibers (RWD; 400 µm core, 0.37NA, 1.25 mm ceramic ferrule) were implanted above the mPFC injection site (M/L: −0.35, A/P: 1.98, D/V: −1.5 mm) to allow for blue light delivery and fluorescence signal recording. Recordings began 4 weeks after surgery to allow sufficient time for viral expression.

Fiber photometry recording
Fiber photometry recordings were performed using RWD's Tricolor Multi Channel Fiber Photometry System. Briefly, 470 nm and 410 nm light pulses were alternately delivered through a 400 µm patchcord (0.57NA; Doric Lenses) connected to an optical fiber implanted above the PFC. Fluorescence was recorded with a cMOS sensor using RWD software at a frequency of 20 Hz. The 410 nm wavelength excitation light represents the isosbestic wavelength of the sensor, which allows us to get a control signal that shows non-Ca^{2+} related signal changes that could contribute to the measured Ca^{2+}-dependent signal. The 410 nm signal was linearly scaled to best fit the 470 nm signal using least-squares regression. The motion-corrected 470 nm signal was obtained by subtracting the 410 nm signal from the 470 nm signal[44]. The corrected 470 nm trace was then z-scored for further analysis.

Shock delivery
During the fiber photometry recording, mice were given a 2 s, 1.0-mA foot shock twice, separated by at least 60 seconds. Shocks were delivered using a behavior box with a built-in shock floor (Med Associates). The time of shock delivery was synchronized to the fiber photometry recording using transistor-to-transistor logic time stamps.

Fiber photometry analysis
Data analysis was performed using MATLAB (MathWorks v2020b). The 410 nm trace was linearly scaled to the 470 nm trace and subtracted for each recording. The corrected 470 nm trace was then z-scored for further analysis. To calculate the mean shock response, the mean trace from $t = 1$ to 2 s after the shock onset was calculated, and then the mean baseline trace from $t = -2$ to 0 s before the shock was subtracted from that. To calculate the mean decay after the shock, the mean trace from $t = 3$ to 4 s after the shock onset was calculated, and then the mean baseline trace from $t = -2$ to 0 s before the shock was subtracted from that. No animals were excluded from analysis.

Histology
Mice were anesthetized under 5% isofluorane and perfused with 20 ml cold PBS, followed by 20 ml of cold 4% paraformaldehyde. Brains were extracted and post-fixed overnight in paraformaldehyde before being transferred to PBS. Brains were sliced on a vibratome (Leica) to a thickness of 60 µm. For immunostaining, brain slices were first washed in PBS with 0.3% Triton-X then blocked for 60 min in PBS with 0.3% Triton-X and 5% normal donkey serum. Slices were stained overnight with anti-GFP-AlexaFluor488 antibody (1:1,000 in blocking solution, Life Technologies A-21311) at 4 °C. Histology images were captured using a Keyence BZ-X180 fluorescence microscope, with an 80 W halide lamp and PlanApo 10 ×0.45 NA air objective. GFP fluorescence was visualized using the commercially provided GFP set excitation/emission filters. Images were processed using ImageJ (Fiji).

Material requests
Plasmids for eGCaMP^{+} and eGCaMP^{2+} can be obtained directly from Addgene for mammalian expression or subcloning encoded in pCAG backbones (201147, 201148) and virus production for CRE-dependent expression encoded in pAAV-EF1a-DIO backbones (201149, 201150).

Ethics statement
All animal procedures performed at the University of Washington were approved by the University of Washington's Animal Use Committee (protocol 4422-01) and follow the National Institutes of Health and the Association for Assessment and Accreditation of Laboratory Animal Care International guidelines. All experiments at the University of California, Davis and University of Washington were conducted in accordance with the Institutional Animal Care and Use Committee.

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

Data availability
All of the datasets generated within this study are available on figshare at https://doi.org/10.6084/M9.FIGSHARE.23750682.V1 (ref. 45). We included the Chen[4] and Dana[5] datasets used to run our model and an amino acid property matrix derived from AAindex[30] in the Supplementary Data. The GCaMP crystal structure used in this paper is accessible online (https://www.rcsb.org/structure/3sg3), GCaMP3–D380Y (RCSB: 3SG3) and in the Supplementary Data. Source data are provided with this paper.

Code availability
The source code is available for download from GitHub at https://doi.org/10.5281/ZENODO.8179256 (ref. 46) and CodeOcean at https://doi.org/10.24433/CO.0624159.v1 (refs. 47,48). Custom Python scripts are available from figshare[45].

References
1. Baird, G. S., Zacharias, D. A. & Tsien, R. Y. Circular permutation and receptor insertion within green fluorescent proteins. *Proc. Natl Acad. Sci. USA* **96**, 11241–11246 (1999).
2. Tian, L. et al. Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nat. Methods* **6**, 875–881 (2009).
3. Akerboom, J. et al. Optimization of a GCaMP calcium indicator for neural activity imaging. *J. Neurosci.* **32**, 13819–13840 (2012).
4. Chen, T.-W. et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).

5. Dana, H. et al. High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nat. Methods* **16**, 649–657 (2019).
6. Zhang, Y. et al. Fast and sensitive GCaMP calcium indicators for imaging neural populations. *Nature* **615**, 884–891 (2023).
7. Patriarchi, T. et al. Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science* **360**, eaat4422 (2018).
8. Sun, F. et al. Next-generation GRAB sensors for monitoring dopaminergic activity in vivo. *Nat. Methods* **17**, 1156–1166 (2020).
9. Feng, J. et al. A genetically encoded fluorescent sensor for rapid and specific in vivo detection of norepinephrine. *Neuron* **102**, 745–761.e8 (2019).
10. Dong, A. et al. A fluorescent sensor for spatiotemporally resolved imaging of endocannabinoid dynamics in vivo. *Nat. Biotechnol.* **40**, 787–798 (2022).
11. Rappleye, M. et al. Optogenetic microwell array screening system: a high-throughput engineering platform for genetically encoded fluorescent indicators. *ACS Sens.* **8**, 4233–4244 (2023).
12. Saito, Y. et al. Machine-learning-guided library design cycle for directed evolution of enzymes: the effects of training data composition on sequence space exploration. *ACS Catal.* **11**, 14615–14624 (2021).
13. Romero, P. A. & Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* **10**, 866–876 (2009).
14. Wu, Z., Kan, S. B. J., Lewis, R. D., Wittmann, B. J. & Arnold, F. H. Machine learning-assisted directed protein evolution with combinatorial libraries. *Proc. Natl Acad. Sci. USA.* **116**, 8852–8858 (2019).
15. Saito, Y. et al. Machine-learning-guided mutagenesis for directed evolution of fluorescent proteins. *ACS Synth. Biol.* **7**, 2014–2022 (2018).
16. Bedbrook, C. N. et al. Machine learning-guided channelrhodopsin engineering enables minimally invasive optogenetics. *Nat. Methods* **16**, 1176–1184 (2019).
17. Unger, E. K. et al. Directed evolution of a selective and sensitive serotonin sensor via machine learning. *Cell* **183**, 1986–2002.e26 (2020).
18. Tian, L., Akerboom, J., Schreiter, E. R. & Looger, L. L. in *Progress in Brain Research* Vol. 196 (eds Knöpfel, T. & Boyden, E. S.) 79–94 (Elsevier, 2012).
19. Nakai, J., Ohkura, M. & Imoto, K. A high signal-to-noise Ca^{2+} probe composed of a single green fluorescent protein. *Nat. Biotechnol.* **19**, 137–141 (2001).
20. Wardill, T. J. et al. A neuron-based screening platform for optimizing genetically-encoded calcium indicators. *PLoS ONE* **8**, e77728 (2013).
21. Kawashima, S. et al. AAindex: amino acid index database, progress report 2008. *Nucleic Acids Res.* **36**, D202–D205 (2008).
22. Dong, X., Yu, Z., Cao, W., Shi, Y. & Ma, Q. A survey on ensemble learning. *Front. Comput. Sci.* **14**, 241–258 (2020).
23. Zhou, Z.-H. in *Machine Learning* (ed. Zhou, Z.-H.) 181–210 (Springer, 2021).
24. Yang, Y. et al. Improved calcium sensor GCaMP-X overcomes the calcium channel perturbations induced by the calmodulin in GCaMP. *Nat. Commun.* **9**, 1504 (2018).
25. Song, Z., Wang, Y., Zhang, F., Yao, F. & Sun, C. Calcium signaling pathways: key pathways in the regulation of obesity. *Int. J. Mol. Sci.* **20**, 2768 (2019).
26. Nausch, B., Heppner, T. J. & Nelson, M. T. Nerve-released acetylcholine contracts urinary bladder smooth muscle by inducing action potentials independently of IP_3-mediated calcium release. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **299**, R878–R888 (2010).
27. Ding, J., Luo, A. F., Hu, L., Wang, D. & Shao, F. Structural basis of the ultrasensitive calcium indicator GCaMP6. *Sci. China Life Sci.* **57**, 269–274 (2014).
28. Souslova, E. A. et al. Single fluorescent protein-based Ca^{2+} sensors with increased dynamic range. *BMC Biotechnol.* **7**, 37 (2007).
29. Akerboom, J. et al. Crystal structures of the GCaMP calcium sensor reveal the mechanism of fluorescence signal change and aid rational design. *J. Biol. Chem.* **284**, 6455–6464 (2009).
30. Barnett, L. M., Hughes, T. E. & Drobizhev, M. Deciphering the molecular mechanism responsible for GCaMP6m's Ca^{2+}-dependent change in fluorescence. *PLoS ONE* **12**, e0170934 (2017).
31. Nasu, Y., Shen, Y., Kramer, L. & Campbell, R. E. Structure- and mechanism-guided design of single fluorescent protein-based biosensors. *Nat. Chem. Biol.* **17**, 509–518 (2021).
32. Fenno, L. E. et al. Comprehensive dual- and triple-feature intersectional single-vector delivery of diverse functional payloads to cells of behaving mammals. *Neuron* **107**, 836–853.e11 (2020).
33. Kim, C. K. et al. Molecular and circuit-dynamical identification of top-down neural mechanisms for restraint of reward seeking. *Cell* **170**, 1013–1027.e14 (2017).
34. Wolpert, D. H. & Macready, W. G. No free lunch theorems for optimization. *IEEE Trans. Evol. Comput.* **1**, 67–82 (1997).
35. Yang, K. K., Wu, Z. & Arnold, F. H. Machine-learning-guided directed evolution for protein engineering. *Nat. Methods* **16**, 687–694 (2019).
36. Yao, Z. & Ruzzo, W. L. A regression-based K nearest neighbor algorithm for gene function prediction from heterogeneous data. *BMC Bioinf.* **7**, S11 (2006).
37. Starr, T. N. & Thornton, J. W. Epistasis in protein evolution. *Protein Sci.* **25**, 1204–1218 (2016).
38. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).
39. Pachitariu, M. & Stringer, C. Cellpose 2.0: how to train your own model. *Nat. Methods* **19**, 1634–1641 (2022).
40. Klima, J. C. et al. Incorporation of sensing modalities into de novo designed fluorescence-activating proteins. *Nat. Commun.* **12**, 856 (2021).
41. Klima, J. C. et al. Bacterial expression and protein purification of mini-fluorescence-activating proteins. Preprint at *Protocol Exchange* https://doi.org/10.21203/rs.3.pex-1077/v1 (2021).
42. Catapano, L. A., Arnold, M. W., Perez, F. A. & Macklis, J. D. Specific neurotrophic factors support the survival of cortical projection neurons at distinct stages of development. *J. Neurosci.* **21**, 8863–8872 (2001).
43. Martin, D. L. Synthesis and release of neuroactive substances by glial cells. *Glia* **5**, 81–94 (1992).
44. Kim, C. K. et al. Simultaneous fast measurement of circuit dynamics at multiple sites across the mammalian brain. *Nat. Methods* **13**, 325–328 (2016).
45. Wait, S. J. et al. Machine learning ensemble directed engineering of genetically encoded fluorescent calcium indicators. *figshare* https://doi.org/10.6084/M9.FIGSHARE.23750682.V1 (2023).
46. sarahwaity/ProteiML: v0.1.1. *Zenodo* https://doi.org/10.5281/ZENODO.8179256 (2023).
47. Wait, S. J. A. B. ProteiML. *CodeOcean* https://doi.org/10.24433/CO.0624159.v1 (2024).
48. BerndtLab—overview. *GitHub* https://github.com/BerndtLab (2024).
49. Dragicevic, P. in *Modern Statistical Methods for HCI* (eds Robertson, J. & Kaptein, M.) 291–330 (Springer, 2016).

Acknowledgements
S.J.W. was supported by the National Science Foundation (DGE-2140004) and the Herbold Foundation. M.E. was supported by 'La Caixa' foundation and the Rafael del Pino Foundation. J.D.L. was supported by 1F31DA056121-01A1. S.L. and C.K.K. were supported by the Burroughs Wellcome Fund (CASI 1019469) and the Searle Scholars Program (SSP-2022-107). A.B. was supported by the Brain Research Foundation, University of Washington (UW) Royalty Research Fund, UW Innovation Pilot Award, National Institute of General Medical Sciences (R01 GM139850-01), National Institute of Mental Health (RF1MH130391), National Institute of Neurological Disorders and Stroke (U01NS128537), National Institute on Drug Abuse (R21DA051193, P30 DA048736 01 Pilot) and the McKnight Foundation's Technological Innovations in Neuroscience Award. The research received additional support from the UW Center of Excellence in Neurobiology of Addiction, Pain, and Emotion Center and Institute for Stem Cell and Regenerative Medicine Shared Equipment. We thank H. R. Daniels (UC Davis) for assistance with histology for in vivo data. We would like to thank M. B. Colby for extensive input and guidance throughout this process.

Author contributions
S.J.W. conceived and designed the experiments, performed the experiments, analyzed the data, contributed materials and analysis tools, and wrote the paper. S.L. conceived and designed the experiments, performed the experiments, analyzed the data and wrote the paper. M.E., M.Ra. and J.D.L. analyzed the data and contributed materials/analysis tools. S.A.C. analyzed the data. A.S. contributed materials/analysis tools. A.A. performed the experiments, analyzed the data and contributed materials/analysis tools. L.T. performed the experiments and analyzed the data. M.Re., F.M.-H. and D.B. contributed materials/analysis tools. A.B. and C.K.K. conceived and designed the experiments, analyzed the data, contributed materials/analysis tools and wrote the paper.

Competing interests
The authors declare no competing interests.

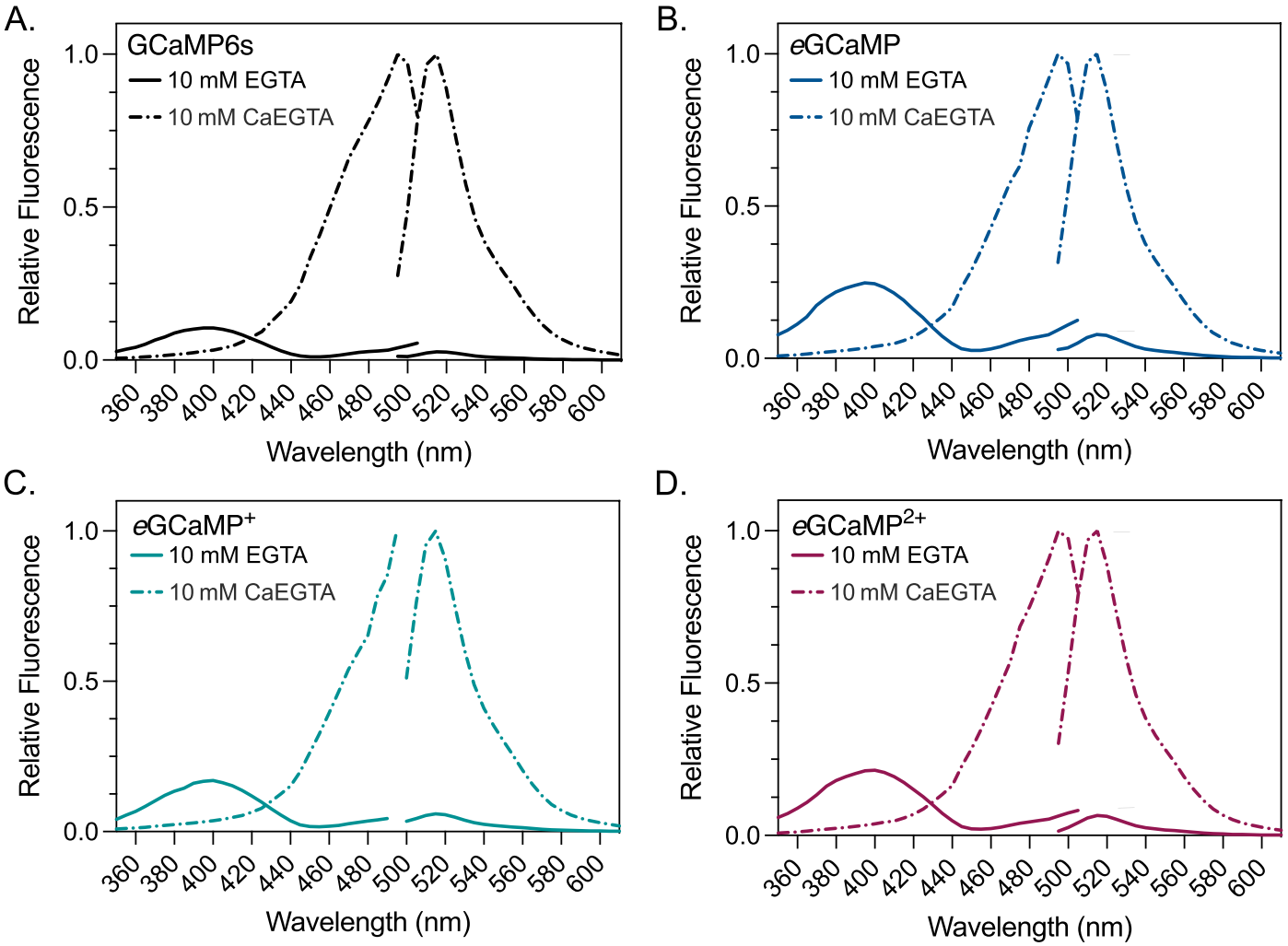
Additional information
Extended data is available for this paper at https://doi.org/10.1038/s43588-024-00611-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43588-024-00611-w.

Correspondence and requests for materials should be addressed to Andre Berndt.

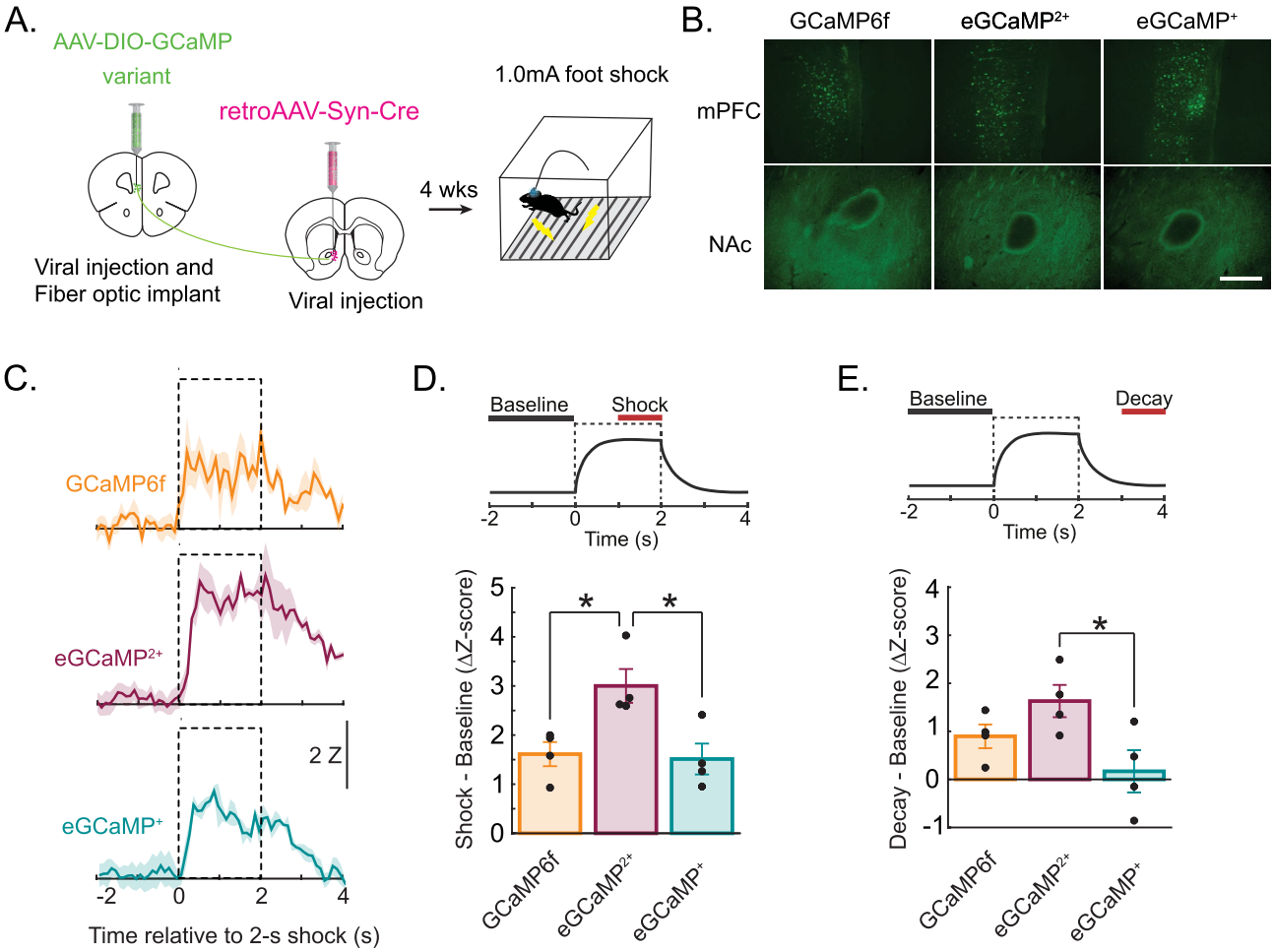
Peer review information *Nature Computational Science* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available. Primary Handling Editor: Jie Pan, in collaboration with the *Nature Computational Science* team.

https://doi.org/10.1038/s43588-024-00611-w



Extended Data Fig. 1 | Excitation and Emission Spectra of eGCaMP Sensors. **A**. Purified GCaMP6s protein diluted into buffer containing either 10 mM EGTA or 10 mM CaEGTA. Emission spectra were calculated using a fixed excitation at 450 nm and excitation spectra were calculated using a fixed emission at 520 nm. **B**. Purified *e*GCaMP protein diluted into buffer containing either 10 mM EGTA or 10 mM CaEGTA. Emission spectra were calculated using a fixed excitation at 450 nm and excitation spectra were calculated using a fixed emission at 520 nm. **C**. Purified *e*GCaMP^{+} protein diluted into buffer containing either 10 mM EGTA or 10 mM CaEGTA. Emission spectra were calculated using a fixed excitation at 450 nm and excitation spectra were calculated using a fixed emission at 520 nm. **D**. Purified *e*GCaMP^{2+} protein diluted into buffer containing either 10 mM EGTA or 10 mM CaEGTA. Emission spectra were calculated using a fixed excitation at 450 nm and excitation spectra were calculated using a fixed emission at 520 nm.

https://doi.org/10.1038/s43588-024-00611-w



Extended Data Fig. 2 | In Vivo Performance of eGCaMP^{+} and eGCaMP^{2+} expressed in mouse mPFC. A. Experimental timeline. Mice were injected with an AAV-Cre dependent-GCaMP variant in the mPFC and a retroAAV-Syn-Cre was injected in NAc. An optic fiber was implanted above the mPFC to allow for light delivery and fluorescence recording. **B.** Representative fluorescence images of GCaMP expression in mPFC and NAc (stained with anti-GFP-Alexafluor488). Scale bar, 130 µm. **C.** Mean Z-scored fluorescence changes in response to a foot shock (n = 4 total shock trials, collected from 2 mice for each GCaMP variant, Line depicts mean, shading depicts SEM). **D.** Comparison of the mean shock response between the three GCaMP variants. Top: schematic of how the shock response was calculated (*see methods*). Bottom: Mean change in Z-scored fluorescence response to shock (n = 4 total shock trials, collected from 2 mice for each GCaMP version). P-values were calculated using a One-way ANOVA followed by Tukey's multiple comparisons in panels (D) and (E): *P < 0.05. All data show mean +/− SEM. **E.** Comparison of the mean decay to shock between the three GCaMP variants. Top: schematic of how the decay to shock was calculated (see methods). Bottom: Mean change in Z-scored fluorescence decay to shock (n = 4 total shock trials, collected from 2 mice for each GCaMP version). P-values were calculated using a One-way ANOVA followed by Tukey's multiple comparisons in panels (D) and (E): *P < 0.05. All data show mean +/- SEM.

nature portfolio

Corresponding author(s): Andre Berndt

Last updated by author(s): Feb 14, 2024

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	Metamorph
Data analysis	ImageJ, MATLAB, Google Colab, Jupyter Notebooks, CellPose, https://codeocean.com/capsule/9982751/tree, https://github.com/sarahwaity/ProteiML, https://github.com/BerndtLab, https://doi.org/10.1101/2023.04.13.536801

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

Data

Policy information about availability of data

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

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

All of the datasets generated within this study are available on figshare(ref 48). The files include data used for main figures, supplemental figures, and supplementary tables. We included the Chen (ref 4) and Dana (ref 5) datasets used to run our model and an amino acid property matrix derived from AAINDEX (ref

32) in the Supplementary Data. The GCaMP crystal structure used in this manuscript is accessible online (https://www.rcsb.org/structure/3sg3), GCaMP3 D380Y (RCSB: 3SG3) and in the Supplementary Data.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size	Samples were taken at least until we reached statistical significance, and typically beyond. Most data sets include hundreds of samples for each average.
Data exclusions	Size exclusion for automated cell segmentation: For cell segmentation, average sizes for cells were calculated. Regions of interest that were outside of one SD mean for the size of cells were excluded.
Replication	Each in vitro data point contains samples from at least 3 biological replicates, that is cell cultures wells. No technical replicates were takes and no outliers removed. The in vivo data contains data from 2 biological replicates (mice) under repeated time-locked stimulation, which is standard in the field.
Randomization	The study was conducted with large numbers of samples and variants by individual researchers who required complete oversight from the preparation of experiments, to conduct, and analysis. We minimized research bias by including large sample sizes and an unbiased, automated image analysis algorithm. No outliers were removed and data was analyzed independently.
Blinding	Test were not blinded because we had only one person taken measurments. However, data analysis was unbiased by an automated image analysis pipeline including a cell segmentation process (CellPose) which automatically identified all sensor expressing cells and extracted change in fluorescence from them.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used	anti-GFP-AlexaFluor488 (Life Technology cat #A-21311)
Validation	Manufacturer validated AB using western blots against positive and negative controls.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	HEK293, ATCC: CRL-1573
Authentication	HEK293 cell were authenticated by ATCC prior to shipping with STR profiling following ISO 9001 and ISO/IEC 17025 quality standards. We discard cells after 25 passages and start new cultures from P2 or P3 frozen stocks.
Mycoplasma contamination	HEK293 cells were tested negative for mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	Commonly misidentified lines were not used

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	6–7-week-old male and female C57BL/6J mice (JAX) were used for in vivo studies. Cultured neurons were dissected from P0 rat pubs (female and male, Sprague Dawley, Envigo).
Wild animals	N/A
Reporting on sex	Male and female mice were used.
Field-collected samples	N/A
Ethics oversight	All experiments were conducted in accordance with UC Davis's and University of Washington's Institutional Animal Care and Use Committees.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☒	☐	Public health
☒	☐	National security
☒	☐	Crops and/or livestock
☒	☐	Ecosystems
☒	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☒	☐	Demonstrate how to render a vaccine ineffective
☒	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☒	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☒	☐	Increase transmissibility of a pathogen
☒	☐	Alter the host range of a pathogen
☒	☐	Enable evasion of diagnostic/detection modalities
☒	☐	Enable the weaponization of a biological agent or toxin
☒	☐	Any other potentially harmful combination of experiments and agents