Accelerated Engineering of Optogenetic Tools with a High-throughput Microwell Array Platform

Michael Rappleye

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Andre Berndt, Chair

Michael R. Bruchas

Azadeh Yazdan-Shahmorad

Program Authorized to Offer Degree:

Bioengineering

University of Washington

Abstract

Accelerated Engineering of Optogenetic Tools with a High-throughput Microwell Array Platform

Michael Rappleye

Chair of the Supervisory Committee:

Andre Berndt

Department of Bioengineering

Optogenetics has shifted from a cutting-edge technology to a mainstay experimental technique. At its core, optogenetics is the use of light sensitive proteins to report or control biological activity. The unique spatiotemporal precision afforded by optogenetic tools has propelled the technique to the forefront of neuroscience, where it is used to investigate the complexities of the mammalian nervous system. Several technologies need to be integrated to implement an optogenetic experiment, coalescing around light delivery to a light sensitive transgene. Typically, optogenetic experiments are limited by the performance of the protein-based, light-sensitive tools. Constructing and optimizing optogenetic tools is laborious and resource intensive, prohibiting extensive tool development. Here, we present a high-throughput optogenetic tool pipeline that screens thousands of protein variants each day in mammalian cells. To showcase the pipeline, dubbed Opto-MASS, we improved a widely adopted dopamine

sensor's response to 100 nM DA >6-fold, and demonstrated the sensor could detect dopamine transients *in vivo*. We improved a mu opioid receptor derived sensor > 4-fold and detected morphine administration *in vivo* in the mammalian nervous system. To further showcase Opto-MASS, we expanded the applications to screen for ligand selectivity by making three biophysical measurements of thousands of sensors nearly simultaneously. Opto-MASS directly links the measured phenotype to the underlying genotype, reducing the time from gene identification to tool application. Opto-MASS is poised to shift the optogenetic tool engineering landscape by rapidly constructing high performance optogenetic actuators and sensors that can be quickly shared with the circuit neuroscience community.

TABLE OF CONTENTS

Some figures throughout this dissertation were created with biorender.com.

LIST OF FIGURES

Some figures throughout this dissertation were created with biorender.com

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge my family and friends for all their support over the years. **Emily** and **Erin**, I wouldn't have made it this far without your unwavering support and willingness to pick up the phone whenever I call. To my mom and dad, thank you for providing me the opportunities, support, and space to grow into the adult and scientist I aspire to be.

To all my friends, I would not have created all the happy memories of Seattle without you. Justin and Sarah, you are great scientists and even better friends. I look forward to watching and participating in your progression through graduate school and beyond as scientists. Ian, Audrey, Kathleen, Elisa, Samantha, Saffie, Sarah, and Caitlin, navigating the uncertainty of the early years of graduate school was made all the better. **Kristian** and **Ross**, thank you for the long days in the mountains, long talks, and immeasurable support and helping me enjoy the things that matter in life. I would also like to thank all of my friends from the Washington Alpine Club.

Andre, thank you for creating a work environment in which I felt respected and enjoyed coming to lab every day. It made all the difference. Also, thank you for all the coffee.

From the mountain tops surrounding Seattle, to long hours side by side at a microscope, you have all provided immeasurable support and I would not be here without you. Thank you.

“For even the very wise cannot see all ends.”

-J.R.R. Tolkien, *The Fellowship of the Ring*

Chapter 1. GENETICALLY ENCODED OPTOGENETIC TOOLS

ABSTRACT

Optogenetics has become a mainstay technique for neuroscience due to its spatiotemporal and genetic precision. Optogenetics combines light sensitive proteins, expressed in the central nervous system, with *in vivo* microscopy techniques, enabling an all-optical observation or control of neuron signaling events. Optogenetic sensors, called genetically encoded fluorescent indicators (GEFIs), are single fluorophore proteins that change in fluorescence upon detection of a physiological signal. While technological advances in genetic expression and *in vivo* light delivery enhance the capabilities of optogenetics, engineering high performance sensors is a significant bottle neck for the field, limiting the scope of optogenetics. Here, we summarize general sensor function, sensor scaffold construction strategies, and optimization approaches that have been used to construct sensors that have been used *in vivo*. We identify several pitfalls engineers should avoid when screening GEFIs and propose advantages to using novel *in silico* and structural methods in parallel with high-throughput screening methods to optimize GEFIs. As optogenetics is a combination of multiple technologies, several approaches will feed and inform each other respectively and develop high performance GEFIs to advance our understanding of biology.

1.1 INTRODUCTION

At its core, optogenetics is an experimental technique that expresses light-sensitive, protein-based tools in biological systems. During an experiment, scientists use light stimulation to measure or control physiological activity. Optogenetics is unusually precise because tools are expressed in genetically defined subpopulations of cells and light stimulation is temporally precise. Furthermore, optogenetic tools can be targeted to subcellular compartments, and light stimulation spatially and temporally targeted. The spatio-temporal precision of optogenetics positions the technique to be applied to complex biological systems.

The mammalian nervous system is a complex network and structure. The brain drives human emotion, interaction, and perception. The nervous system is comprised of interconnected neural circuits that communicate through the firing of action potentials and the release of neurotransmitters, hormones, lipids, peptides, and other signaling molecules. These chemical releases shape brain activity and are responsible for human cognition, pain, emotion, socialization, reward, and other aspects of the human experience. To fully understand neural circuits, their function, and pathophysiology, scientists need to be able to link neuron firing and neurotransmitter release during animal behavior or emotion. The brain is difficult to study because most techniques lack the spatiotemporal resolution to link animal behavior to neural circuit activity, or the technique cannot be performed in freely behaving animals. For example, *ex vivo* slice work has spatially and temporally precise measurements, however it cannot link measurements to animal behavior. Functional MRI or EEG measurements lack genetic precision. However, one technique, optogenetics, pairs genetic tools with light-based techniques to control and observe neuron firing and neurotransmitter release *in vivo* [1], [2].

Broadly, optogenetic tools are divided into two categories, actuators and sensors. Optogenetic actuators use light to control neuron firing in specific cell populations (**Figure 1.1A)**. For more information about the development of optogenetic actuators, refer to Chapter 5 & Chapter 6. The second type of tool, optogenetic sensors, use light to report on neural activity or chemical release in specific neuron populations (**Figure 1.1B)**. Both sets of tools are genetically encoded, meaning the biological instructions to assemble the protein-based tools are encoded in DNA or RNA molecules and packaged into adeno associated virus (AAVs) or lentivirus particles. The AAV particles are injected into the mammalian central nervous system, where the genetic

material is translated into light sensitive proteins. Cell-type specific gene promoters in the AAV or lentiviral particles drive tool expression to defined subpopulations of neurons. The genetic promoters enable precise control or readout of neuron subtype activity *in vivo*. Additionally, genetic tools can be paired with transgenic Cre-recombinase organisms, that further define or restrict the neural subpopulations that optogenetic tools are expressed in.

During viral injection surgeries, light controlling hardware like optic fibers, cranial imaging windows, or head-mounted miniature microscopes are integrated with the animals for light delivery and recording. Time-locked control of neural firing or time-locked observations of cell activity can be measured and linked to animal behavior.

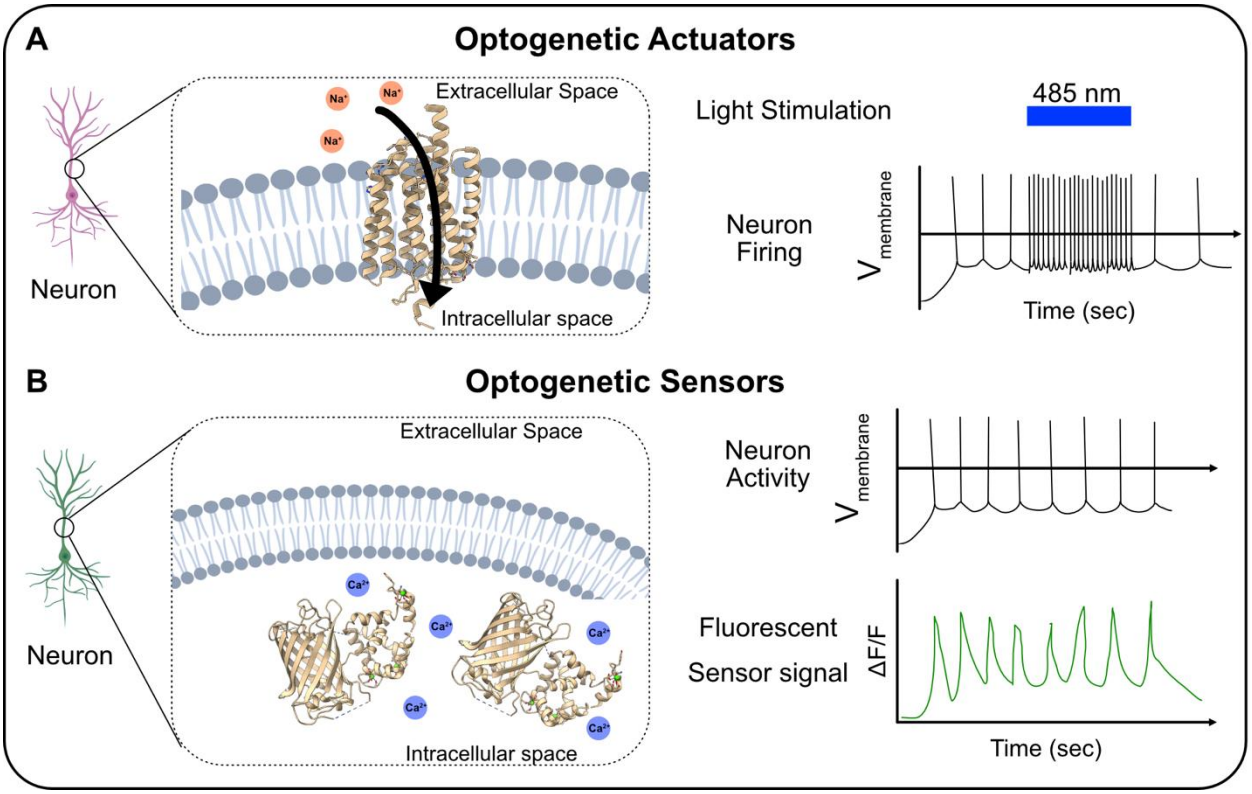

Figure 1.1. Optogenetics is centered around light sensitive proteins

A). Optogenetic actuators control neuronal firing through light activated ion channels. Ion channels can be excitatory, like Channelrhodopsin2 or ChRmine that increase neuron firing potential, and thus action potentials, as shown [3], [4]. Inhibitory tools, such as iC++, GtACRs, or KCRs inhibit neuron firing upon light stimulation [5]–[7] For a detailed overview of light controlled ion channels and their engineering, see Chapter 6. PDB: 6EID. B) Optogenetic sensors emit fluorescent light that is modulated by ligand signal. Pictured here is an idealized Ca^{2+} sensor. Observe the precisely time locked fluorescent signal with neuronal firing. Sensors can be membrane bound and detect neurotransmitter release or confined to the cytosol to detect intracellular signaling events [8]–[10]. Sensors are all optical readouts of neuronal activity with a single excitation and emission wavelength. PDB: 3EK4.

Optogenetics is the combination of multiple fields; optical methods for *in vivo* light delivery and observation; animal models and interpreting animal behavior; genetics and physiology; and light sensitive, protein-based tools. Advances in optics and image processing allow more precise light delivery and interpretation of larger data sets. Single-cell mRNA sequencing has identified new genetically defined populations of neurons, which enable more precise optogenetic tool expression. While technological advances push optogenetics forwards, light sensitive proteins are the foundation of the optogenetic approach.

Optogenetic sensors, called Genetically Encoded Fluorescent Indicators (GEFIs), are tracked *in vivo* and their changes in fluorescence intensity are directly correlated to their target ligand's concentration. GEFIs contain a single fluorophore and are measured through changes in fluorescence intensity of that fluorophore. These tools are remarkably powerful; their fast kinetics and spatial and genetic constraints enable them to link the activity of specific neurochemicals in specific neurons during animal behavior. However, engineering, bright, fast, ligand selective GEFIs that are compatible with mammalian physiology takes years, if not decades, of work. For example, the widely adopted Ca^{2+} sensor, GCaMP, took over 10 years from initial design to extensive *in vivo* mammalian use. Modern protein science needs to be applied to GEFI engineering because there remain hundreds of molecules without high performance sensors. Recent advancements in protein science, such as predictive structural modeling, advances in Cryo-EM and mutational techniques like deep mutational scanning provide guidance to engineer GEFIs. However, there remains remarkable gap in the available high-performance sensors for the remaining ligands neuroscientists wish to detect in the brain. Despite the incredible advancements the field of protein-engineering has made, GEFI development is still slow and resource intensive.

1.2 Optogenetic Tools Enable Cell-Type Specific Readouts of Biological Activity

Genetically encoded fluorescent indicators are protein-based, single-fluorophore molecules that change in fluorescence upon detection of a physiological event (**Figure 1.2**). One of the first GEFIs developed was a calcium sensor, called GCaMP [11]–[13]. When a neuron fires, calcium floods into the intracellular space, dramatically raising the concentration of free calcium to 100s of nanomolar [14], [15]. Ca^{2+} signaling is biologically ubiquitous, ensuring broad adoption of high-performance tools by the scientific community, and Ca^{2+} undergoes large concentration

fluctuations during signaling events, making it easy to detect. The importance of studying Ca^{2+} signaling was evident with the acceptance of fluorescent calcium sensitive dyes by the research community. However, fluorescent dyes are limited because dyes washout and cannot be targeted to specific cell types. Ideally, biologists would be able to link a genetically defined cell to its role in animal behavior. The cell-type specificity problem can be solved by fluorescent proteins, as their genetically encoded nature enables the genetic restriction of their expression.

The most widely adopted fluorescent protein is green fluorescent protein (GFP), discovered and cloned in 1962 and 1999, respectively [16], [17]. GFP is a bright, monomeric protein that absorbs blue light (474 nm) and emits green light (505 nm). However, once GFP is expressed in a cell, GFP does not toggle it's fluorescence like a calcium sensitive dye. GFP's chromophore is protected from environmental modulation by a barrel of eleven β-strands [18]. Roger Tsien and colleagues circularly permuted GFP (cpGFP) to enable modulation of the chromophore [13]. When a protein is circularly permuted, the primary DNA sequence is rearranged, shifting the start and stop locations of translation to a new location in the proteins DNA sequence (**Figure 1.2**). A functioning sensor was generated by grafting the cpGFP domain into calcium binding domains (M13 and calmodulin) with peptide linkers (**Figure 1.2**) [13]. Other research groups understood the usefulness of genetically encoded calcium indicators (GECI), and developed a family of tools, called GCaMP, using similar Ca^{2+} binding domains and fluorescent proteins [11].

Incremental improvements to GCaMP's expression and dynamic range led to GCaMP detecting neural activity in *Drosophila*. Using two-photon imaging, Ca^{2+} sensors were used to record neuron firing during odor detection in *Drosophila* [19]. Even further enhancements in GCaMP expression levels, signal to noise ratio, and ligand sensitivity enabled the detection of calcium signals in higher ordered organisms such as mice [2]. GCaMP presented itself as a powerful tool, enabling scientists to observe and directly link a neurons activity to an organism's behavior.

With these early experiments, the neuroscience field collectively identified how transformative high-performance biosensors are for understanding the brain. Researchers embarked on the task of engineering better variants with increased dynamic range, baseline fluorescence, affinity, and kinetics. High-performance calcium sensors would eventually be capable of monitoring a variety of calcium signaling events in higher order model organisms, such as sub-threshold calcium release, burst firing, and release of calcium in neuronal spines [20], [21].

However, it took almost a decade of improvements before GCaMP saw widespread adoption by neuroscientists.

The coalescing of better *in vivo* imaging technologies; improved genetic expression techniques; and dramatic improvements in GCaMP performance has led to GECI's becoming the *de facto* tool to monitor neuronal signaling events in freely behaving animals. While biosensors can use multiple fluorophores for precise detection and readout of ligands, such as FRET and BRET, we will focus on the engineering and development of single fluorophore biosensors as their single excitation and emission wavelength lends themselves as an effective tool for *in vivo* experimental preparations. They are effective because single fluorophore excitation and emission is simpler to execute experimentally, and they can be paired with spectrally orthogonal optogenetic sensors or actuators for dual ligand detection and neural circuit control [22].

The engineering principles for constructing calcium sensors can be applied to make indicators for a variety of ligands. By grafting the cpGFP domain into proteins that bind different molecules, researchers expanded the repertoire of single fluorophore sensors to include cAMP, maltose, nicotine, serotonin, glutamate, GABA, zinc, acetylcholine, and norepinephrine [8], [23]–[29]. This list is not exhaustive, demonstrating the widespread adoption of GEFIs in experimental preparations and their usefulness in answering scientific questions. Each of the GEFIs development required rigorous screening of mutants and *in vitro* biophysical characterization of the selected mutants prior to adoption by the scientific community.

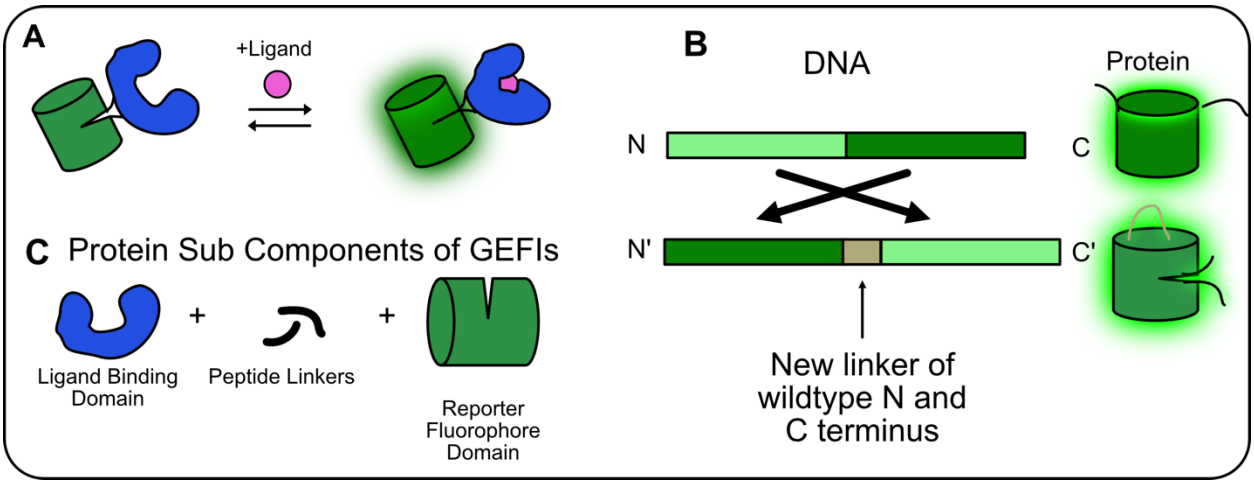

Figure 1.2. GEFI components and circular permutation of proteins

A). Upon ligand binding, the population of GEFI proteins shifts from a low fluorescence state to a higher fluorescent state.

B) Circular permutation of a protein occurs when the primary DNA sequence encoding the protein is rearranged. This alteration changes the sequence of peptide assembly during translation, however the overall three-dimensional structure and folding of the protein remain the same. New allosteric switches and sites can be introduced with circular permutation.

C) GEFIs are comprised of three components; 1) a domain that undergoes conformational change upon ligand binding and is amenable to the insertion of the cpGFP domain, 2) the reporter fluorophore domain, and 3) peptide linkers to physically link the reporter fluorophore and ligand binding domain.

1.3 Design Considerations and the Mechanism of Action of Genetically Encoded Fluorescent Indicator

1.3.1 *Protein Domain Considerations*

All single fluorophore genetically encoded fluorescent indicators (GEFIs) are comprised of three parts, 1) a single fluorophore domain inserted into 2) a ligand binding protein with 3) short peptide linkers connecting the sensing and reporter domains. When engineers construct GEFIs, they typically construct them by combining pre-existing protein domains together instead of *de novo* protein design [30]. When constructing GEFIs, engineers must consider the biophysical properties of the proteins they are combining and consider what constraints the experimental preparation places on the sensor. Protein characteristics such as ease of expression, folding kinetics, genetic payload constraints in AAVs, ligand concentration profile, stimulation, and observation source access all play a role in protein domain choice (**Table 1.1**).

When choosing a fluorophore, above all else it must generate enough signal to be detected above background fluorescence. Ideally, the fluorescent signal changes linearly with ligand concentrations in the physiologically relevant range. Red-shifted fluorophores can be used to image sensors deeper in tissues because their simulation and emission wavelengths can penetrate further into tissue with less energy and scattering [31]. However red fluorophores can present unique challenges such as protein aggregation, broader excitation and emission spectra, dimness, off-spectra excitation, and cytotoxicity [32]. Newer red fluorophores have been engineered to overcome some of the aggregation and off spectra excitation issues [33]. cpGFP is the typical fluorophore of choice, due to its high thermostability, rapid folding kinetics, brightness, high dynamic range, significant knowledge base of GFP's structure-function relationship [10], [32],

[34], [35]. While adequate for most sensors, engineers could further optimize the cpGFP domain by mutating residues in the β-barrel that alter folding to bias the folding pathway away from aggregation and misfolded intermediates. Mutations to the β-barrel could increase the stability of the H-bonding network on the exterior of the protein and enhance the β-barrel's thermostability– shielding the chromophore from twisting or hydration events that reduce sensor brightness [35]–[37].

While selecting ligand binding domains, engineers need to accommodate performance demands such as selectivity, affinity, and cross species compatible expression [38], [39] (**Table 1.1**). Ideally, sensing domains selectively bind the ligand of interest with an affinity that matches the physiological range [40]. Upon ligand binding, binding domains undergo large conformational changes that can be used to allosterically modulate the chromophore in the cpGFP. High resolution protein structures, of both ligand bound and unbound states of the protein provide immense guidance on GEFI engineering. Researchers have found success inserting cpGFP into a variety of domains, such as EPAC domains for detecting cAMP, periplasmic binding domains for a variety of small molecule ligands, voltage sensitive domains for GEVIs, M13 and calmodulin domains for calcium sensors, and more recently, G-protein coupled receptors (GPCRs) for the detection of a variety of ligands [8], [9], [26], [33], [41]–[49]. The use of GPCRs as ligand binding domains is significant because GPCRs are important mediators in a host of physiological processes, so they bind to many biologically relevant molecules and present a modular approach to engineering GEFIs.

Importantly, ligand binding domains should not interfere with endogenous signaling pathways. Typically, GEFIs and fluorophores must be overexpressed to achieve enough properly folded molecules for detection and functionality [40]. If their catalytic or allosteric domains are still active, they will alter biochemical signaling pathways in the model organism being tested. The ligand binding domain should have a well-defined biochemical signaling pathway or be biochemically orthogonal to the signaling pathway than the one being experimentally investigated. Well defined biochemical pathways can be tested for off target stimulation by the GEFIs.

Significantly, membrane bound GEFIs such as GPCR based sensors, like the GRAB or 'Light' families of sensors should be thoroughly vetted for their lack of signal amplification. GPCRs evolved as signal amplifiers, requiring only small amounts of ligand to induce many intracellular signaling events. Overexpression of slightly functional GPCRs on the membrane

could dramatically change the intracellular environment from signaling, leading to a significant alteration of the host organism. These GEFI sourced host organism perturbations would confound any conclusions drawn from the experiment. GEFI engineers can use several strategies to sidestep these problems; through steric hinderance of G-protein or β-arrestin coupling by the cpGFP domain (in the case of GPCRs), removal of catalytic domains, or mutating key residues in catalytic sites to remove their functionality [8], [29], [41], [50], [51].

The final component of GEFIs are the peptide linkers. There are two classes of peptide linkers in GEFIs, one, the link that connects the old C and N termini of the fluorophore and 2) the peptide linkers fusing the cpFP and the sensing domain. The first peptide linker is generally considered solved for the field, and little to no mutations are performed on the linker. Despite this, there may be dramatic ways to enhance GEFIs performance as this linker may directly modulate the thermostability of the reporter domain, which has a direct and significant impact on a GEFIs performance [52]. The benefit of mutations to the cpGFP old termini linker needs to be explored by the field. Generally, engineers should investigate shortening or lengthening the peptide linker to enhance stability of the β-barrel of the fluorophore. Mutating the amino acids in the linker (GGTGGS) may alter the hydrogen bonding network on the exterior of the β-barrel, enhancing the stability of the β-barrel and increasing the chromophores brightness through reducing undesired torque or hydration of the chromophore [37].

The second set of peptide linkers, between the sensing domain and ligand binding domain, are critical elements to change GEFI performance. Their role in the protein is transmute the ligand dependent conformational changes from the sensing domain to the reporter fluorophore domain. Linker length and composition are critical to the thermostability of the protein. The linkers must generate an energy landscape in which the protein can transition between the apo and bound state relatively easily, but a high enough activation energy between states the majority of the protein population is in the bright state only in the presence of the ligand. Researchers have a near impossible task of predicting protein folding dynamics and structure function relationships of amino acids in the peptide linkers. Due to the computational and predictive hurdle this presents, protein engineers typically test hundreds of linker variants to generate useful sensors.

Table 1.1. Protein domain considerations when engineering genetically encoded fluorescent indicators.

Sensor Component	Function	Desired Characteristics	Special Engineering Considerations
Sensing domain	Detect and bind ligand of interest	High ligand selectivity; physiologically relevant ligand affinity; large, ligand dependent conformational changes	Removal of catalytic sites or coupling sites for downstream signaling pathways; Ligand buffering; On/off kinetics High resolution structures aids protein engineering
Reporter domain	Change in fluorescence emission upon ligand detection	Thermostable, bright, high quantum yield upon ligand binding, narrow excitation and emission band	Red fluorophores are prone to aggregation, off spectra excitation and emission, dimness; red fluorophores can be imaged in deeper tissues; pH sensitivity
Peptide linkers	Allosteric linkage between sensing domain and reporter fluorophore	Short to increase thermostability	Rigidity of linkers can benefit sensor, thermal stability, may coordinate phenol to phenolate equilibrium of chromophore [39]

1.3.2 *GEFI Functional Mechanism: Chromophore Modulation*

Genetically Encoded Fluorescent Indicators (GEFIs) function by the fluorophore getting brighter upon ligand binding (**Figure 1.3**). To achieve this, the cpGFP's chromophore must switch between a dim and a bright state. The GFP chromophore is generated by a chemical reaction between S65, T66, and G67 (GFP numbering) that requires molecular oxygen [53]. The chromophore can be protonated or unprotonated at the phenol. When the chromophore is unprotonated and in a planar form, it absorbs light at approximately 475 nm and emits light at approximately 508 nm (**Figure 1.3**). However, in the protonated state, the chromophore absorbs and emits significantly less light, and the peak absorption is shifted to approximately 395 nm, with an emission of 503 nm[2], [53], [54].

In a GEFI, the chromophore is allosterically modulated between the two states in a ligand dependent manner. During ligand binding, an amino acid side chain is positioned closer to coordinate the phenolate state of the chromophore, generating fluorescence in the standard GFP wavelengths. The side chains can be from amino acids from different parts of the protein; the linkers, the sensing domain, or the chromophore itself, as determined through crystal structures of various sensors [39], [55]. Due to the complex molecular concert that must occur for GEFIs to function, it comes as no surprise initial GEFI scaffolds function poorly. If domain insertions yield

proper expression and some ligand dependent modulation of fluorescence, engineers should move ahead and optimize that scaffold. GEFI's are proteins, and thus the underlying genetic code can be mutated to change the protein's performance.

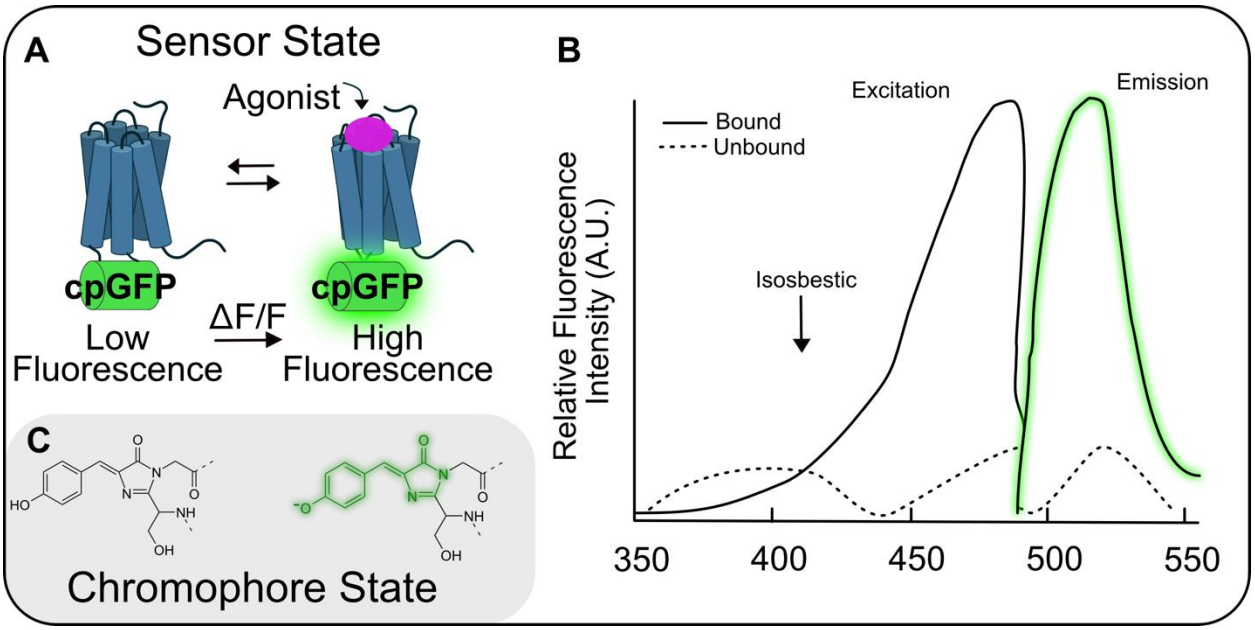

Figure 1.3. GEFI state and chromophore modulation

A). Proteins occupy several states, shifting between states with a thermodynamic equilibrium determined by their energy landscape. A functional GEFI, in the unbound state has a low fluorescence state. Upon ligand binding, the ligand binding domain is stabilized in a different conformation, and the peptide linkers transfer this conformational change to the cpGFP domain. The fluorophore is stabilized in a light emitting conformation.

B) The light absorption and emission are low in the unbound state of the receptor. Upon ligand binding, the chromophore absorbs and emits more light. Ideally, there is an isosbestic wavelength, where chromophore excitation and emission are constant regardless of ligand concentration. The isosbestic wavelength can be used to control for motion artifacts in fluorescence *in vivo*.

C) The chromophore phenol group is allosterically modulated between a protonated and unprotonated state by ligand binding. Figure inspired by [2], [56].

1.4 Genetically Encoded Fluorescent Indicators Can Be Improved Through Mutational Screening

An underlying tenant of biology is the reciprocal nature of the structure-function relationship of proteins, cellular structure, and organisms. Scientists can exploit the structure function relationship of proteins and mutate the primary amino acid sequence of GEFIs to alter

their performance. For example, the initial versions of GCaMP, and most sensors, did not generate enough signal fidelity to detect low frequency firing rates or low concentration ligand release *in vivo*[2]. Thus, researchers mutated the underlying DNA encoding the protein to enhance the sensors performance when expressed. Researchers can use several different techniques to improve sensor performance. For example, increasing expression levels of the protein, increasing thermostability, kinetics, and allosteric coupling of the sensing and reporting domains are all viable approaches to increasing GEFI performance [38]. These approaches usually rely on the physical *in vitro* testing of sensor performance because *in silico* models currently do not translate well into mammalian systems and require testing *in vitro*. When testing sensor variants, researchers must balance resource investment and the physiological relevance of the assay (**Figure 1.4**). Screens with more physiological relevance to the mammalian nervous system usually require more resources to execute and reduce the number of GEFI variants screened. Screening fewer variants reduces the chance of identifying the most optimal sensor. Less physiologically relevant systems require less resources to screen a greater number of variants, however the 'hits' may not translate well into mammalian host systems. Thus, when choosing screening platforms, researchers must balance physiological relevancy, resources, and the number of variants they can reasonably screen.

High resource screens, such as the neuron-based platform used to engineer the later generations of GCaMP, can overcome limitations presented by lower resource screens [57]. Initially, researchers improved GCaMP expression by adding the RSET expression cassette to the amino terminus of the sensor [13], [20], [58]. Increasing expression levels yields improved performance, as more molecules can absorb and emit light. Early efforts to screen and identify high dynamic range GCaMP were carried out in HEK293 or HeLa cells [13], [20], [59]. Additionally, they tested purified protein lysates of GCaMP2 to GCaMP3 and researchers found performance improvements in purified lysates did not transfer to cell-based performance. However, the fast Ca^{2+} signaling dynamics of neurons and sub compartments were not recapitulated in HEK293 or HeLa cells, and simpler systems such as bacterial lysates did not translate well [14].

To develop tools capable of detecting Ca^{2+} transients in dendritic spines, sensor variants need to be screened in an assay with equally fast calcium transients. To effectively discriminate between mutant's biophysical properties, the underlying calcium dynamics of the screening system

must be fast enough that the limiting factor is the GEFI's performance, not the physiology of the underlying assays.

To overcome this issue, researchers at Howard Hughes Medical Institute (HHMI) saw the need to develop a platform to screen calcium sensor variants. They constructed a neuron-based screening pipeline in which mutants were expressed in cultured primary rat neurons and functionally screened with a confocal microscope [57]. They targeted their mutations from crystal structures of the Ca^{2+} bound and unbound of GCaMP2 [55]. Beneficial mutations were then added together to find synergistic mutations and lead to an optimal sensor [57]. The platform has yielded several high performance GECIs because it effectively recapitulates the *in vivo* physiological Ca^{2+} signals while simultaneously screening sensors for kinetics, neuronal expression levels, protein folding, and baseline fluorescence [10], [60].

Despite the platform's success, the significant resource investment precludes the large-scale adoption of the platform to develop other sensors for *in vivo* optogenetic use. The GECI engineering pipeline presented a problem; the high access barrier due to the complexity of primary neuron cell culture technique precludes it from widespread adoption. On the other end of the resource investment spectrum, researchers optimized cAMP sensors in bacteria. cAMP is an important secondary messenger, altering ion channel function, gene regulation, and metabolism. Researchers leveraged the ability to screen a much greater number of variants in a less physiologically relevant organism to identify a high-performance sensor. A library of 30,000 cAMP sensor mutants was screened in bacteria. Sensors were assessed for their baseline fluorescence and response to bath additions of cAMP [26]. The sensors responded weakly to bath additions of cAMP, demonstrating that low resource investment into screening platforms can make functional screening difficult. Researchers still needed to test sensor functionality in higher order tissue systems of identified 'hits', adding in extra validation steps.

The field of genetically encoded voltage indicators (GEVIs) has seen significant investment in sensor development. Ed Boyden's group engineered a pipeline in which HEK293T cells expressing variants of a GEVI were screened along multiple performance characteristics to identify high performance variants [46]. Their pipeline was used to screen only 2,500 or 10,000 HEK293T cells expressing GEVIs. Initially, cells were screened for baseline fluorescence and membrane localization, not sensor signal or functionality. After the initial screening, external bath electrodes stimulated membrane voltage fluctuations in only 48 or 216 library variants in the first

and second pass through the library, respectively. Additionally, non-excitable HEK293T cells were used to screen the library, and the technique used to induce membrane voltage changes could lead to inaccurate voltage fluctuations across the membrane. Inaccurate stimulation techniques could bias selected sensors from the pipeline.

The field of sensor engineering has approached the problem with a variety of resource investment approaches. Researchers can invest resources upfront to make highly relevant specialty pipelines and guide mutation selection with precisely resolved crystal structures. On the other end of the resource investment spectrum, researchers simply expressed cAMP sensors bacterial cells and directly added the ligand. While a greater number of variants were functionally screened, the assay could not adequately discriminate the sensors performance for *in vivo* mammalian use. The pipeline used to engineer the GEVIs Archon1 and Archon2 lies in the middle of the resource investment spectrum, with cheap and scalable transfection reagents into mammalian cells, yet functionally screened low numbers of variants with an imprecise membrane voltage fluctuation technique.

The ideal pipeline would functionally screen a large variant library in as close of possible testing conditions to use case scenario with minimal resource investment (**Figure 1.4**). We propose to engineer a pipeline that uses traditional lipofectamine style reagents to express GPCR-based sensors in mammalian cells with cheap, scalable parts that do not preclude low resource labs from applying the technique.

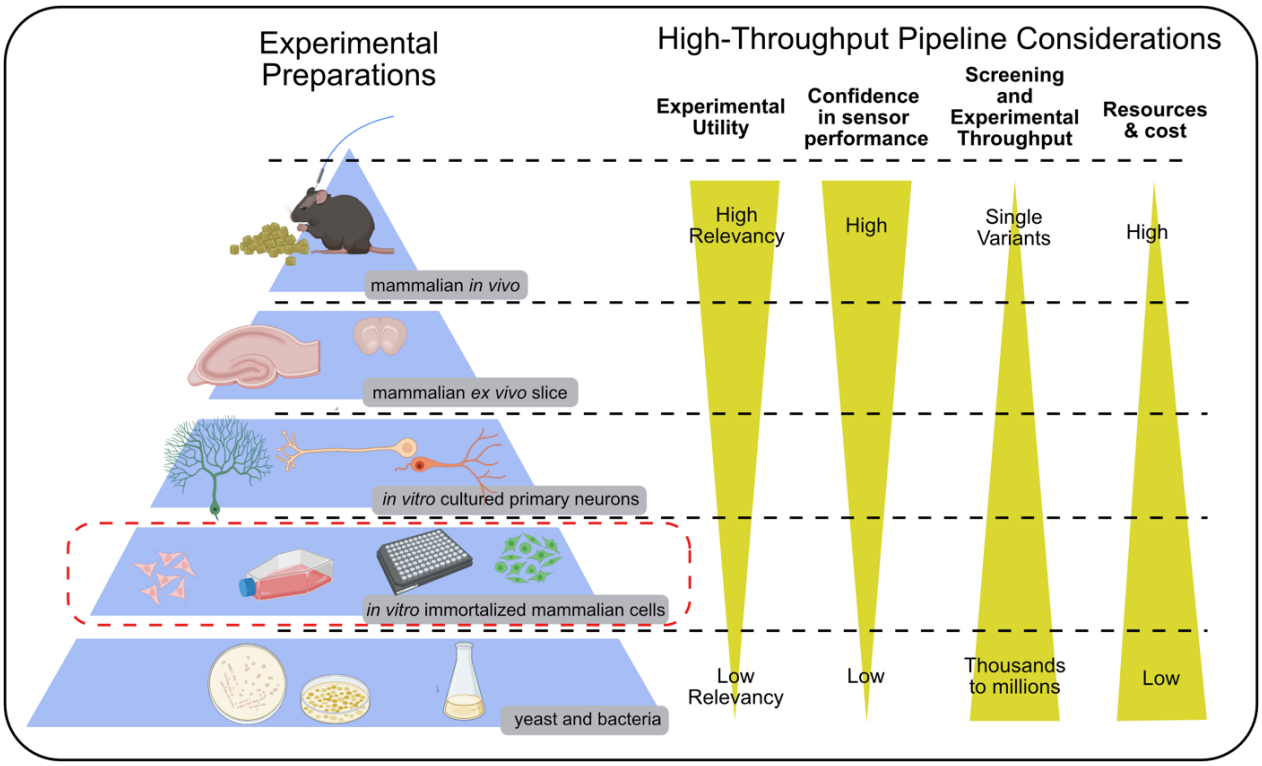

Figure 1.4. Experimental considerations when engineering optogenetic GEFIs for neuroscience

Engineers and researchers need to consider resource investment and experimental relevancy when designing platforms to screen GEFI variants. The ideal platform strikes a balance of resource investment, difficulty of execution, throughput, simplicity, and physiological relevancy to the mammalian nervous system. Less complex experimental preparations like yeast and *E. coli* are less physiologically relevant but are easier to screen, handle, and can have extremely high throughput. Less relevant experimental preparations also lend themselves to less confidence in sensor performance in the mammalian central nervous system. To overcome this, most researchers screen GEFI variants in less complex experimental preparations, and stepwise increase the experimental complexity and concomitantly the experimental relevancy by taking the highest performing variants at each stage and progress to the next step in complexity and relevancy. The apex of GEFI engineering is demonstration of expression and functionality in freely behaving animals, linking physiological signals to animal behavior with cell-type specificity and signal kinetics in the sub-second timescale with high fidelity.

1.5 Future Directions of Optogenetics: Combining Computational, Structural and High-Throughput Screening Methods

Studying the brain requires precise experimental preparations due to the functional, structural and genetic complexity of the brain. The optogenetic approach provides the necessary

spatial and temporal precision in genetically defined neuron subtypes needed to understand neural circuits. At the heart of this is high performance GEFIs. High performance sensors enable an all-optical readout of neural signaling events in freely behaving animals. GEFIs function by allosteric modulation of a cpGFP chromophore. GEFIs are an exercise in engineering protein allostery in fused protein domains, a task requiring screening of many mutants.

Engineers must consider resource investment when engineering GEFIs. Higher resource investment into biologically complex screens, such as cultured neurons, reduces the number of variants that are screened. However, variant 'hits' identified in biologically similar *in vitro* platforms have a higher chance of transferring *in vivo*. On the other end of the spectrum, screening a greater number of proteins with less relevant host systems and screening may yield to hits that do not translate well into the mammalian nervous system. While engineers refine *in vitro* assays, computational models can inform which residues to mutate and screen. Mutations identified with machine learning models are often synergistic, and performance benefits are greater than the sum of the individual mutations [61]. Both scientific approaches, computational modeling GEFI performance and high-throughput *in vitro* assays, can inform and provide data to refine each technique. Computational models mapping protein structure and function with mutations are benefited with high quality protein structure data and assay performance data. Protein structure data has become less of an experimental hurdle with several advancements.

High quality protein structures, particularly for membrane bound proteins, has been accelerated by the advancements in cryo-Electron Microscopy[62], [63]. GPCRs, an important pharmacological target on cell membranes, have been developed into a new class of GEFIs[64]. High quality structures of GPCRs in both the active, ligand bound state, and inactive form provide atomic level insights on structure function relationships. While Cryo-EM currently struggles with resolving ligands in binding pockets, structural insights of membrane bound proteins in ligand bound conformation better inform computational models and high-throughput assays of GEFI performance. The structural insights can be paired with computational efforts to model protein folding, such as AlphaFold2 [65]. As mining for new ligand binding domains yields new proteins, *in silico* models of protein folds will enable rapid construction of GEFIs to detect previously undetectable ligands or signaling events, such as activation of specific G_{α} subunits [66].

Advancements in protein expression, protein discovery, structural methods, *in silico* protein structure prediction and *de novo* design all accelerate the GEFI engineering field towards

high performance sensors. Paired with advances in genetic tool delivery and *in vivo* imaging methods, the field of optogenetics will push the boundaries of our understanding of the mammalian nervous system. Science is often limited by hypotheses that tools are capable of testing, for example, until the advancement of head mounted 2-photon imaging, deeper tissue structures in brain could not be imaged *in vivo*. Until the sequencing of GFP cDNA, researchers were limited to using synthetic dyes to trace molecules *in vitro.* Technology is pushing our understanding of protein structure and function. The development of new detergents to crystalize GPCRs has enabled determination of several GPCRs. Analyzing these new inactive and active state structures has informed the construction of a new subset of GEFI's using GPCRs as ligand binding domains. These GEFIs are poised to provide dramatic insights on cell-to-cell signaling in the mammalian brain, as GPCRs bind a variety of intermolecular signaling molecules. We will develop next-generation GEFIs to decode the mammalian nervous system.

Chapter 2. DEVELOPING A HIGH-THROUGHPUT PIPELINE TO OPTIMIZE GENETICALLY ENCODED FLUORESCENT INDICATORS[1]

ABSTRACT

Fluorescent sensor proteins are instrumental for detecting biological signals *in vivo* with high temporal accuracy and cell-type specificity. However, engineering sensors with physiological ligand sensitivity and selectivity is difficult because their performance is measured through individual mutagenesis *in vitro* to assess their performance. The vast mutational landscape proteins constitute is hindrance to sensor development, particularly for sensors that require screening in mammalian host systems. Here, we developed a novel high-throughput engineering platform that functionally tests thousands of variants nearly simultaneously in HEK293T cells. We showcase the capabilities of our platform, called Optogenetic Microwell Array Screening System (Opto-MASS), by engineering a monoamine *in vivo* capable optogenetic sensor within weeks. With our platform, we screened over 13,000 mammalian cells expressing a dopamine sensor library to identify an improved variant. The improved variant, called dMASS^{3A}, has a 1.6-fold improved response saturating conditions of dopamine compared to the parent scaffold *in vitro*. The increase in dynamic range comes with no loss of molecular selectivity or baseline brightness. dMASS^{3A} was expressed *in vivo* in Vgat-cre mouse dorsal medial striatum and nucleus accumbens. Using fiber photometry, dMASS^{3A} detected sucrose concentration dependent dopamine transients *in vivo* in the rat nucleus accumbens. Opto-MASS addresses the need for improved methods to construct optogenetic sensors. Traditional techniques screen sensors one-by-one, Opto-MASS presents a method to rapidly constructing *in vivo* capable optogenetic sensors by functional screening of thousands of sensor variants expressed in mammalian cells. We showcased our platforms

[1] This chapter has text directly from and adapted from: Rappleye M, Gordon-Fennell A, Zamorano C A, Castro D C, Matarasso A K, Stine C, Wait S J, Lee J D, Siebart J, Suko A, Smith N, Muster J, Matreyek K A, Fowler D.M , Stuber G D, Bruchas M R, and Berndt A, "Opto-MASS: A high-throughput protein engineering platform for genetically encoded fluorescent sensors enabling all optical *in vivo* detection of monoamines and opioids" *Pre-print, BioRxiv:* https://doi.org/10.1101/2022.06.01.494241

versatility by optimizing monoamine and opioid sensors that are *in vivo* capable, highlighting the ability to optimize optogenetic sensors for neurotransmitters with diverse physiological roles.

2.1 Introduction: Optogenetic Microwell Array Screening System (Opto-MASS) bridges resource investment and screening relevancy

Genetically encoded fluorescent indicators (GEFIs) are protein-based sensors that increase in fluorescence intensity upon target ligand binding [13]. The basic engineering principle combines a ligand selective binding domain with a fluorescent reporter protein and tunes their connection by mutating the amino acids linking the two domains. Recently, a new subset of GEFIs were constructed by grafting a circularly permuted fluorophore (cpGFP) into the third intracellular loop of dopamine G-protein Coupled Receptors (GPCRs) to engineer dopamine sensors (DA) [9], [50](**Figure 2.1A**). As with most GEFIs, several hundred sensor variants were screened to optimize signal amplitudes and dopamine detection. However, the screened mutations represent a small fraction of the 160,000 variants that constitute the mutational landscape of the 4 targeted residues (20^4). It is likely that better sensors could be identified within the remaining sequence space if more variants could be screened. The grafting principle has been demonstrated to work on a host of GPCRs, expanding the available sensors to include acetylcholine, serotonin, norepinephrine, and orexin [28], [29], [67], [68]. But similarly, protein engineering bottle necks testing to only a few hundred variants during the process of optimizing ligand sensitivity and signal output.

Traditionally, engineering fluorescent biosensors requires a multistep, resource intensive process. Researchers generate individual mutations in plasmid DNA by PCR, purify them from *E.coli* one-by-one, and express the variants in a heterologous expression system for analysis. Membrane-bound GPCR-based sensors require testing in mammalian host cells such as HEK293 cultures because yeast and bacteria cells have difficulty expressing a diversity of fully functional GPCRs at their membranes [69]. On the other hand, to test constructs in mammalian cell cultures, researchers transfect individual plasmids into cells seeded in multi-well plates (24-384 wells), limiting throughput. The fluorescent output of sensor variants is then tested upon ligand application, often under saturating conditions to elicit maximum responses. The mutation and screening process must be repeated potentially hundreds of times, as performance is notoriously difficult to predict in these highly dynamic fusion proteins. The field needs to address the significant gap presented by resource intensive techniques currently used to engineer sensors for

the wealth of GPCRs in mammalian physiology because current advances preclude the development of sensors for the majority of GPCRs.

Here, to address this gap, we present a high throughput platform to rapidly construct genetically encoded fluorescent indicators. Furthermore, the mammalian host cells are engineered to express one single variant while using commercial transfection reagents. We physically separate the individual sensor expressing cells into single wells in a microwell array. Taken together, we functionally screen hundreds of cells simultaneously, resulting in thousands of tested cells per day. We quickly rank the sensor's phenotypes in real-time using fluorescence microscopy and automated image analysis. When considering the engineering of GEFI optimization platforms, we must design the platform for broad application, ease of adoption, and throughput. Protein engineering pipelines using mammalian cells have been applied to the optimization of genetically encoded voltage and calcium indicators [46], [57]. However, the functional screening of sensors was either limited in throughput or did not provide signal readouts under dynamic conditions. To showcase the broad applicability of our pipeline we optimized an *in vivo* capable monoamine sensor and engineered a neuropeptide sensor capable of all optical, *in vivo* detection of exogenous opioids. The Opto-MASS is an accessible platform that can rapidly optimize *in vivo* capable biosensors to detect a variety of ligand types. Notably, the platform addresses a gap in the field by providing a higher throughput platform to optimize GPCR-based biosensors to *in vivo* capabilities and increase the number of scientific questions that can be answered by these highly precise tools.

2.2 RESULTS

2.2.1 *Design, Engineering, and Optimization of Opto-MASS*

We envisioned a method that functionally screens thousands of optogenetic sensor variants each day to significantly increase engineering throughput which is often limited to a few hundred variants in several months. New, improved sensors could then be used as a scaffold to iteratively mutate and improve upon with the platform (**Figure 2.1B**). We identified four necessary features to achieve this goal. One, a single step library generation strategy to make a large, unbiased library of sensor variants in DNA plasmids. Two, a mammalian expression system wherein one single plasmid is expressed per cell, while maintaining high transfection efficiency. Three, the ability to readout functional, dynamic signals from hundreds of cells simultaneously under physiological

conditions. Four, recovery of genetic content encoding high-performing variants from individual cells.

To achieve these goals, we put the 'landing pad' HEK 293T TetBxb1BFP cells at the center of our platform [70]. They enable facile expression of a single variant per cell in mammalian cells. The sensor expressing cells are screened in customized PDMS microwell arrays placed in 24 well cell culture plates (**Figure 2.1**). The microwell arrays physically separate the cells to enable an easy, functional readout of fluorescent signals from hundreds of cells simultaneously. The cells on each array are ranked based on ligand-dependent fluorescence changes in real time. We physically recover the cells by aspirating them with a glass micropipette controlled by a micromanipulator. Next, we perform single cell RT-PCR on the recovered cell to identify the sensor encoding gene. We validate the recovered gene's phenotype by biophysical characterization in cultured HEK293 cell populations. After assessing the sensor's performance, we could iterate the process and use the recovered variant as a scaffold for the next library or package the sensor into a virus for *in vivo* experiments.

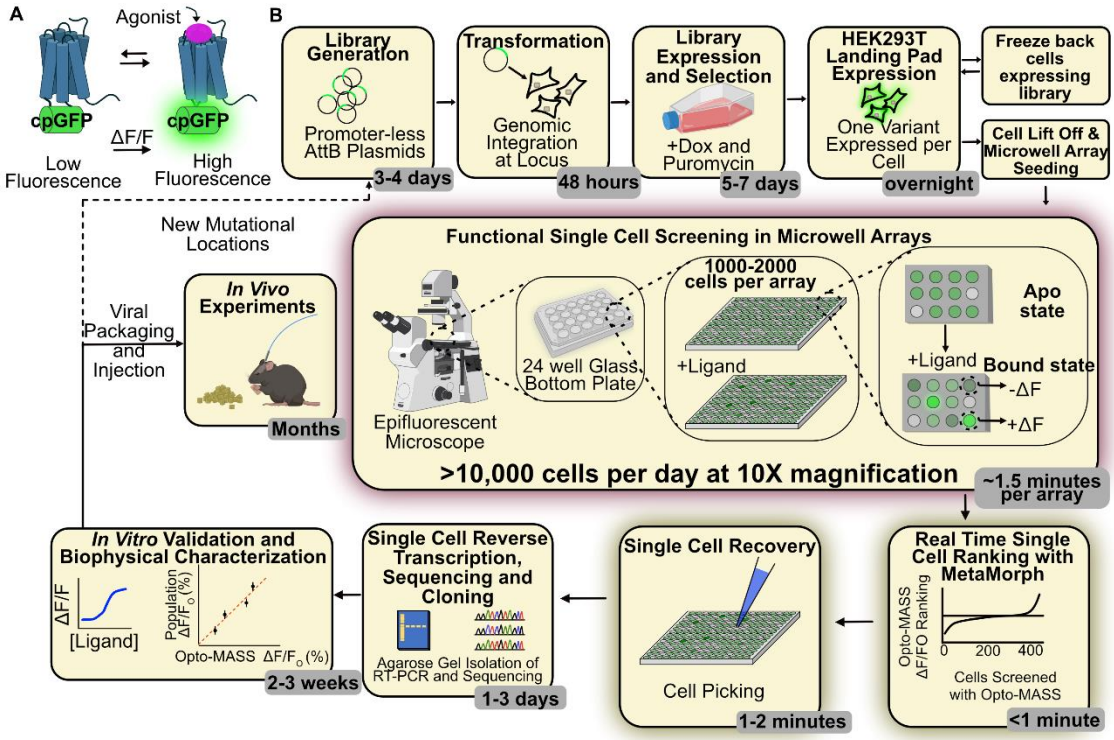

Figure 2.1. The Optogenetic Microwell Array Screening System: overview

A) A new class of GEFIs uses GPCRs as ligand binding domains. Upon ligand binding, conformational changes in the receptor domain allosterically modulate the fluorophore of the inserted cpGFP domain and increase its fluorescence.

B) Researchers need a facile, rapid, and scalable system to build optogenetic sensors. Opto-MASS combines microwell array screening technology and mammalian genetic expression systems to rapidly screen thousands of cells/variants of sensors on a platform that can complete sensor engineering in less than a month. First, a randomized mutational library of the sensor is constructed in landing pad compatible plasmids. The library is then recombined in the HEK293T "landing pad" cell genome into a single locus. After doxycycline induction and puromycin selection for 5-7 days, the library can be frozen back for later screening or seeded onto microwell arrays. The Opto-MASS microwell arrays are screened on an inverted fluorescent microscope, and sensors are ranked nearly instantaneously in each field of view. After identifying the highest performing cells, a glass micropipette is used to aspirate them physically. Next, we use RT-PCR to identify the underlying sensor gene. The recovered gene is then transfected into HEK293WT cell cultures to characterize the sensor's biophysical phenotype in detail on a population level. If the sensor retains the desired characteristics, we can package it into viral vectors for in vivo experiments or use the sensor as a template for a new round of library screening.

2.2.2 *HEK293T TetBxb1BFP landing pad mammalian expression system.*

The landing pad expression system has several key features that make it an amenable solution to our design goals. The plasmids encoding for sensor variants contain no constitutive mammalian promoter. Instead, upstream of our genes of interest, an AttB recombination site enables irreversible recombination with a AttP site inserted into the genome. We cloned a tricistronic gene downstream of the AttB recombination site on our sensor library plasmids. The three genes encode for: a green fluorescent sensor, a mCherry (control for image analysis), and a puromycin resistance gene (selection of recombined cells) and are each separated by self-cleaving P2A sequences (**Figure 2.2A**). Thus, after Bxb1 mediated recombination into the engineered locus, each protein functions independently. A Tet inducible promoter drives the expression of the tricistronic cassette, allowing the tuning of genetic expression (4-10 μg/mL doxycycline). As a result, while Fugene6 transfection reagents can introduce more than one plasmid per cell, only one plasmid can recombine into the genomic landing pad, thus only a single variant is expressed per cell **(Figure 2.2B)**.

2.2.3 *Microwell array design and cell seeding*

The design goal for the microwell array was to image thousands of physically separated cells within a single field of view for easy analysis of sensor function and physical recovery of sensor variants. The microwell arrays were designed to accommodate a single HEK293T cell. The physical separation provided by the microwells expedites the automated analysis of fluorescent signals from individual cells (**Figure 2.1**). We fabricated a silicon master mold from a 100mm silicon wafer, using deep reactive ion etching (DRIE) to etch away the wafer and reveal a negative mold of the arrays (University Wafer, **Figure 2.2C**). We varied well diameter (20-100 μm), distance (5-10 μm) and depth (20-50 μm) for initial prototyping. Our goal was to reach maximum well density in the field of view of our camera (Photonics Prime 95B 2048x2048 pixels at 11 μm per pixel) while holding only one cell per well. The array size is 2.5 x 2.5 mm to match the field of view of our camera at 5X magnification. Optimal well parameters were found at 35 μm diameter, 6.75 μm distance and 35 μm depth resulting in 3600 wells per array. This enables the observation of a maximum of 3600 wells at 5X magnification or 900 wells at 10X. Each 100mm silicon wafer carries 76 microarrays (**Figure 2.2C**). We use polydimethylsiloxane (PDMS, Sylgard 184) to cast

our Opto-MASS arrays by standard soft lithography techniques. (**Figure 2.2D**). We chose PDMS due to its optical clarity, low cost, and low cytotoxicity on our experimental timescales (12-16 hours). After curing the PDMS on the master mold, the PDMS slab was nanostamped with BSA to reduce cell adhesion outside of the microwells during initial cell seeding (See **Methods 2.4.1**).

After removing dust and debris from the back of the PDMS slab, individual arrays were cut from the PDMS slab and placed in the wells of a 24 well, glass-bottom dish (CellVis, NC0397150, Fisher Scientific). We optimized cell suspension seeding conditions for the HEK293T landing pad cells to maximize well occupancy while reducing the occurrence of multiple cells in a microwell (**Figure 2.2E-H**). For our library screening experiments, we chose to seed 80,000 HEK 293T cells at a concentration of $0.5x10^6$ cells/mL per array. This results in single cell occupied wells at about 30%-50% array occupancy, routinely providing up to 1800 observable cells at 5X or 300-500 cells at 10X magnification.

2.2.4 *Automated Image Acquisition and Analysis Protocol*

During functional screening of the libraries, an efficient way to track and rank the cells response to the addition of the ligand is needed. Our microscope and imaging set up was operated using MetaMorph imaging software (Molecular Devices). The control fluorophore, mCherry, was imaged before and after stimulation to remove any cells that moved into or out of the field of view during stimulation (**Figure 2.3A-C**). During stimulation, the cells were imaged continually under GFP wavelengths (EX: 474/27 nm, EM: 520/35 nm) and the ligand of interest was added to the bath to screen for sensor functionality (**Figure 2.3D**). mCherry was used to define regions of interest for analysis in MetaMorph. Regions of Interest (ROIs) were then measured for size and excluded if they were too large or too small for typical cells (**Figure 2.3** and **Figure 2.4**).

The ROIs were transferred to the pre stimulation mCherry image to measure mean grayscale values as a control for fluorophore expression. The ROIs were then transferred to an image stack of the stimulation period. The ROI's average GFP grayscale value (i.e., fluorescence intensity) was measured and exported to Excel for offline analysis. Next, the stimulation image stack was split into two different stacks, the pre- and post-stimulation stacks, with the average fluorescence intensity projection taken for both stacks. The resulting images were then divided and multiplied by 1000, so that ROIs that increased in fluorescence had a value higher than 1000, and those that decreased had a value lower than 1000 (**Figure 2.3**). We dubbed the ratio value the

Opto-MASS Ranked Ratio for each ROI and it was used to identify ROIs that had the greatest increase in fluorescence in the field of view. The Ranked Ratio was found to be an effective measure of fluorescence change (**Figure 2.3**).

We don't include images close to the ligand addition event because there may be motion artifacts that alter the fluorescence calculation (**Figure 2.3**). After identifying the ten highest ranked ROIs, the ten ROIs were moved to the live field of view for physical recovery and a duplicated stack of the stimulation for manual review prior to recovery. Cells are recovered under bright field. See **Figure 2.4** for detailed task execution during screening.

2.2.5 *Opto-MASS screens functionally diverse sensor populations*

For initial testing, we aimed to demonstrate that the platform could correctly identify signals from sensors with different but known signal amplitudes under similar ligand concentrations. For this purpose, we cloned the dopamine (DA) sensors dLight1.2, dLight1.3a, and dLight1.3b into landing pad plasmids [9]. We generated isogenic populations of the sensors by stably integrating the different dopamine sensors into separate landing pad cell populations, inducing their expression with doxycycline (10 μg/mL) and selecting with puromycin (0.75-1μg/mL) (**Figure 2.2I**). We seeded the isogenic populations onto different arrays (i.e. each cell population expressing one sensor on separate arrays) to test for sensor functionality (**Figure 2.2I and J**).

We expected that the dopamine dependent fluorescent changes at saturating DA concentrations (100 μM) should create differentiable responses between these sensors. We achieved this goal after extensive protocol optimization. The key step was to seed cells and let them recover on the arrays overnight in doxycycline supplemented media prior to screening. Using the optimized protocol, we could distinguish the different dopamine sensor populations based on their fluorescent output. Importantly, we yielded a reduced coefficient-of-variance for the sensor populations on the arrays to demonstrate the signal readout was representative of the underlying sensors performance (**Figure 2.2K**). The significantly improved CoV of approximately 10% $\Delta F/F_o$ provides a narrow window for potential outliers and cell-to-cell variability (**Figure 2.2K**)[57].

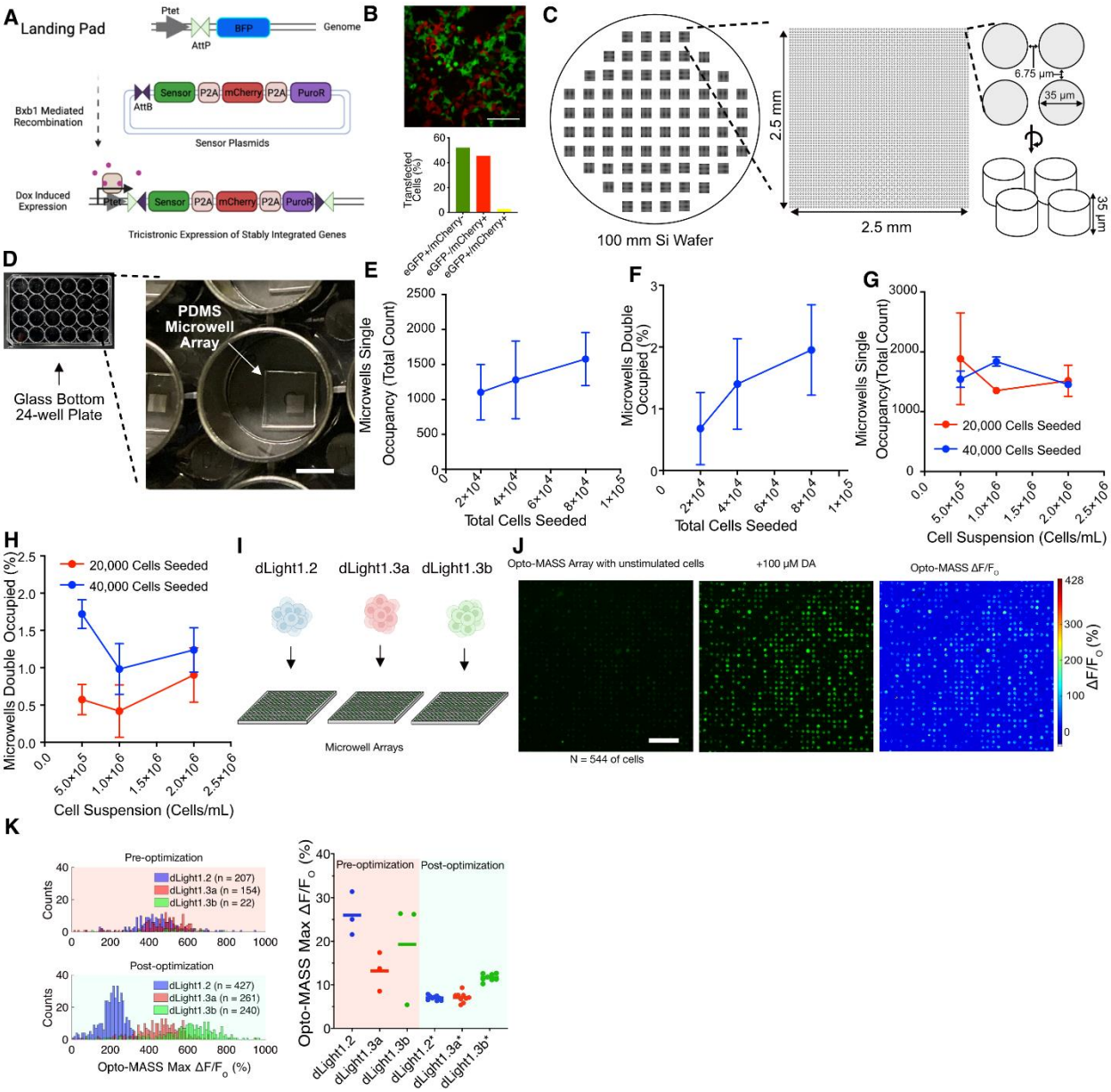

Figure 2.2. Design and engineering of the Optogenetic Microwell Array Screening System

A) Schematic of the genomic landing pad in the HEK293T landing pad cells (HEK 293T TetBxb1BFP) pre and post sensor plasmid integration.

B) Representative false-color image of HEK293T TetBxb1BFP cells transfected with a mix of mCherry and eGFP plasmids (upper panel 100 µm scale bar). Summary data for the field of view for the fluorophore expression (lower panel). The majority of cells express only one fluorophore.

C) Silicon master mold of the microwell arrays. The Si wafer was etched using Deep Reactive Ion Etching.

D) Polydimethylsiloxane (PDMS) microwell arrays are cut out from the master mold after curing and BSA nanostamping and are placed 'wells up' in the bottom of a 24 well glass bottom plate for easy screening of thousands of variants in a day.

E-H) HEK293T landing pad cells expressing dopamine sensors are seeded onto microwell arrays at varying total amounts to determine most optimal seeding density. Total cells seeded was altered to determine if it

increased well occupancy (E and F), or cell suspension concentration was altered to determine effects on well occupancy (G and H).

I) For initial testing, we seeded three distinct dopamine sensors with known signal amplitudes onto separate arrays.

J) Representative stimulation of an isogenic population stimulation of landing pad cells expressing dLight1.3b and stimulated with 100 μM dopamine on the Opto-MASS. N = 544 cells in field of view. 200 μm scale bar.

J) Array stimulation optimization. Cell seeding density, cell handling, microscope focusing, and array fabrication, were optimized to reduce the coefficient of variation in arrays stimulation of three different dopamine sensors. Saturating concentrations of dopamine were used. Average CoV pre-optimization 19.52% (n = 9 microwell arrays) to 8.49% (n = 26 microwell arrays).

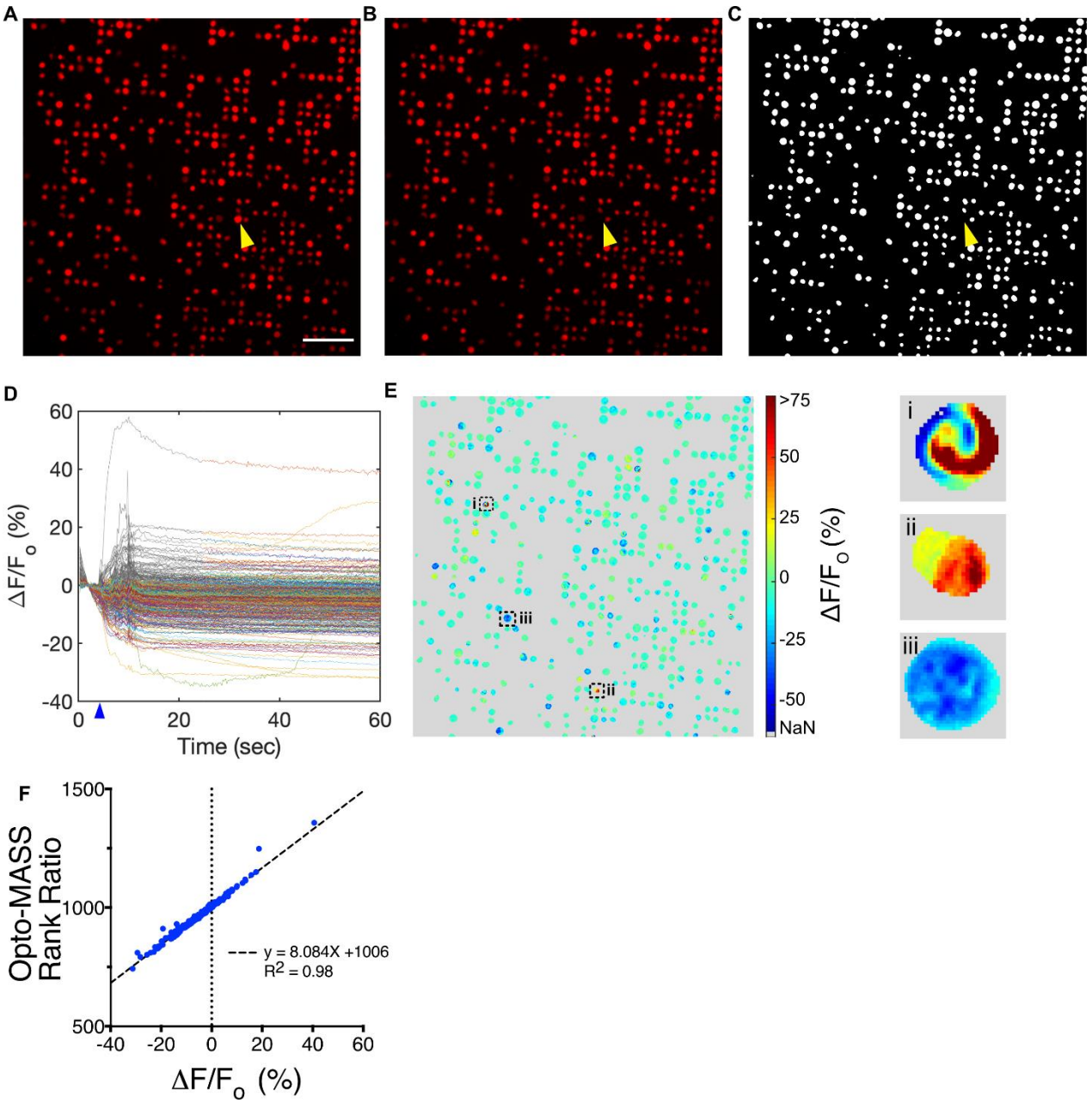

Figure 2.3. A rapid analysis algorithm to identify high performance sensors

A) Example false color image of pre-stimulation control fluorophore imaging. Yellow arrow denotes a ROI that moved during stimulation. Scale bar 200 µm.

B) Post-stimulation imaging of the control fluorophore post stimulation of the same array in A. Yellow denotes the absence of the ROI highlighted in A.

C) Binary mask generated from pre and post stimulation images in A and B. Note the lack of ROI by the yellow arrow.

D) GFP fluorescence changes of the ROIs from C over time due to 500 nM dopamine addition. Blue arrow denotes approximate time of DA addition.

E) Heat map of pixel changes in mask area (C) due to ligand addition. Note that areas outside the mask were assigned a grey value. Inset images are larger views of areas i, ii, and iii. Note the differences in individual pixels in the ROIs, and the resulting ROI increasing or decrease in fluorescence.
F) The Opto-MASS rank ratio generated by the division of the baseline and post stimulation average intensity projections correlates with the post-hoc calculation of fluorescence change. Note, average fluorescence change for the last 35 seconds is used as change in fluorescence percentage.

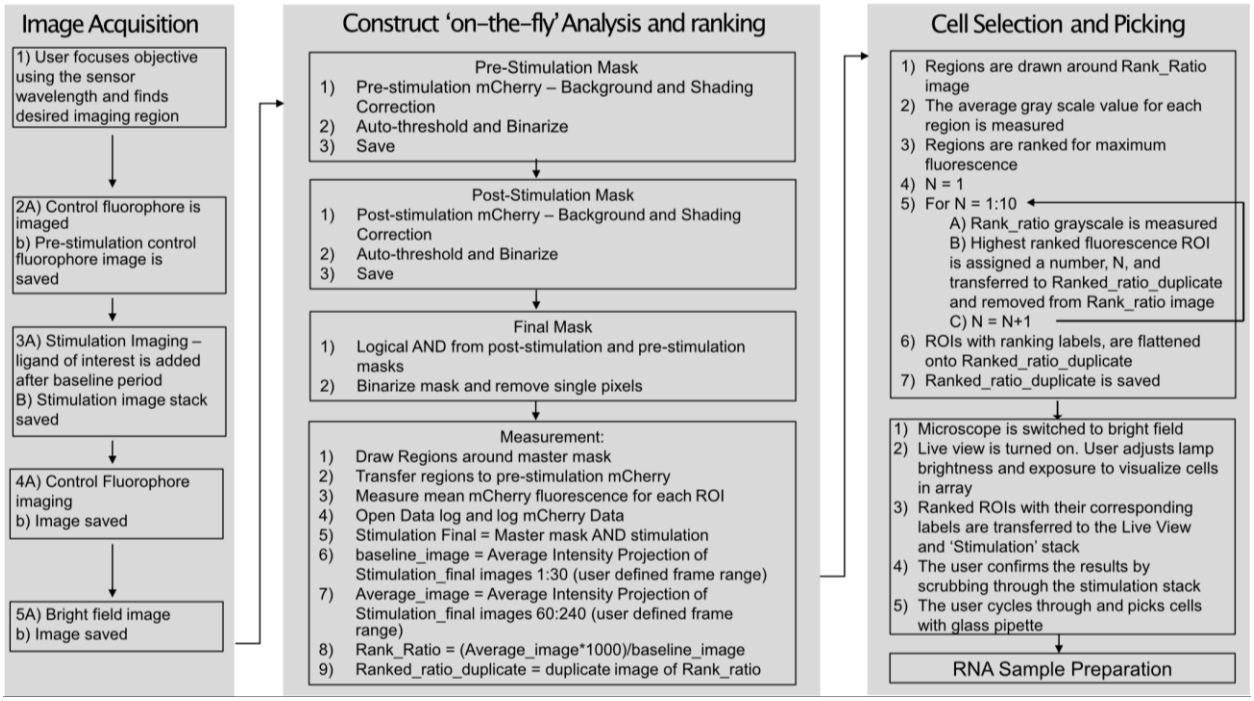

Figure 2.4.: Detailed high throughput screening task list.

The major imaging tasks for high throughput imaging and screening can be divided into three major tasks. The commands are executed in MetaMorph journals to control various aspects of the microscope.

2.2.6 *Using Opto-MASS's enhanced screening capabilities to identify a high-performance monoamine sensor with* in vivo *capabilities*

Next, we demonstrated that we could identify variants with optimized signal amplitude and ligand sensitivity variants from a large mutational library of an existing sensor framework. We chose the dopamine sensor dLight1.1 for this purpose because of its reliable signal generation and utility of dopamine sensors to the field of neuroscience. We targeted the four residues flanking either side of the fluorophore cpGFP inserted into the human D1 dopamine receptor (**Figure 2.5A**). These residues play a critical role in coupling fluorophore brightness to ligand-dependent changes

in the receptor domain, but only 585 variants were tested of the 20^4 possible variants [9], [39]. We built a library of randomized mutations at these sites by incorporating DNA primers with degenerate codons (IDT) in a single PCR step and subsequent Gibson Assembly (NEB) (**Figure 2.5**A, **Figure 2.6**, > 200,000 *E.coli* transformants). Sampling DNA sequences from 23 colonies revealed a relatively even distribution of four nucleic acids at the targeted sites (**Figure 2.5B**). We transfected the randomized plasmid library into a population of 250K cells (cultured in a well of a 6-well plate). The landing pad cells had been transfected with plasmids encoding for a nuclear localized Bxb1 recombinase twenty-four hours prior (Fugene6, Promega). We drove library expression and selection by adding doxycycline (10 μg/mL) and puromycin (1 μg/mL), respectively, after 24h. After 2-4 days of selection, cells were combined and passaged into a T25 flask. Cells cultures were expanded for a total of 5-7 days, until selection was complete. Next, we lifted the cells with Trypsin (0.05%) and EDTA from the flasks and seeded cells onto microarrays placed in glass bottom 24-well plates. We let the cells recover in incubators for 12-16 hours overnight prior to screening (37°C, 5% CO_2). Plates were imaged under epifluorescence using the GFP channel (474 nm excitation, 520 nm emission, 500 millisecond (ms) exposure) for dopamine signals and the mCherry (578 and 641 nm) channel as negative controls. Cells on each array were stimulated by dopamine application near physiological (500 nM) to saturating conditions (10 μM) via an automated syringe pump at consistent time points (**Figure 2.5C and D**). The simultaneous increase of fluorescent signals from the cells demonstrates that diffusion of the dopamine was immediate (**Figure 2.3D**). We could faithfully rank cells in each field of view in real time using customized MetaMorph scripts (**Figure 2.3 & Figure 2.4**). In the initial and following sessions we recovered the highest-ranking cells from the arrays by using glass micropipettes connected to a syringe for aspiration and controlled by a micromanipulator. Each cell was placed into separate microcentrifuge tubes containing Tris/EDTA buffer and the reducing agent dithiothreitol (DTT 2.44 mM) to prevent RNA degradation by RNAse.

We recovered the sensor encoding gene from each cell by RT-PCR (SSIV First Strand Synthesis, ThermoFisher) using a gene specific primer. PCR amplified cDNA was recovered from agarose gels following electrophoresis and identified by routine Sanger Sequencing (GeneWiz). The recovered sensor sequences were cloned into mammalian expression vectors (pC_DNA3.1, CMV promoter) for subsequent transfection (Lipofectamine 3000, ThermoFisher) and biophysical characterization into HEK293WT cell cultures.

Importantly, the high-throughput capabilities of the platform allowed us to identify highly dynamic variants at lower ligand concentrations instead of being biased towards maximum brightness under saturating conditions. In one trial we screened ~13,000 cells/variants at a low 500 nM dopamine stimulation in a single session to find more sensitive variants (**Figure 2.5C**). During this session, 30.3% of cells decreased in brightness upon ligand addition ($\Delta F/F_o < -5\%$), 59.6% had no fluorescence response ($-5\% < \Delta F/F_o < 5\%$), and 10.1% increased in fluorescence ($\Delta F/F_o > 5\%$).

We observed several high performing variants (**Figure 2.5D**) and recovered cell 3A to pursue further characterization in HEK293 cell populations. We dubbed the recovered variant dMASS^{3A} because it was a dopamine indicator identified using the Opto-MASS pipeline. All the targeted linker sites were altered compared to dLight1.1 in dMASS^{3A} (**Figure 2.5E**). When screened with epifluorescent microscopy, dMASS^{3A} had a lower K_D compared to the parent construct, dLight1.1 (K_D = 323 nM and 625 nM, respectively**, Figure 2.5F**). dMASS^{3A} had 1.6-fold greater response than dLight1.1 at 10 μM DA ($p < 0.0001$, Student's T-test, unpaired, **Figure 2.5G**). The baseline fluorescence of dMASS^{3A} was found to be close to the parent construct, dLight1.1 ($p = 0.0734$, unpaired Student's t-test, **Figure 2.5H**).

The increased fluorescence output came at no apparent loss of molecular selectivity of the sensor (**Figure 2.5I**). To align our measurements with previous studies, we tested dMASS^{3A} on a confocal microscope for low concentrations of dopamine (**Figure 2.5J-K** >6-fold, $p<0.0001$, unpaired two-tailed t-test). Here, dMASS^{3A} also outperforms the parent construct, dLight1.1, demonstrating sensors can be enhanced in specific ways by targeting Opto-MASS screening conditions towards the desired sensor characteristics.

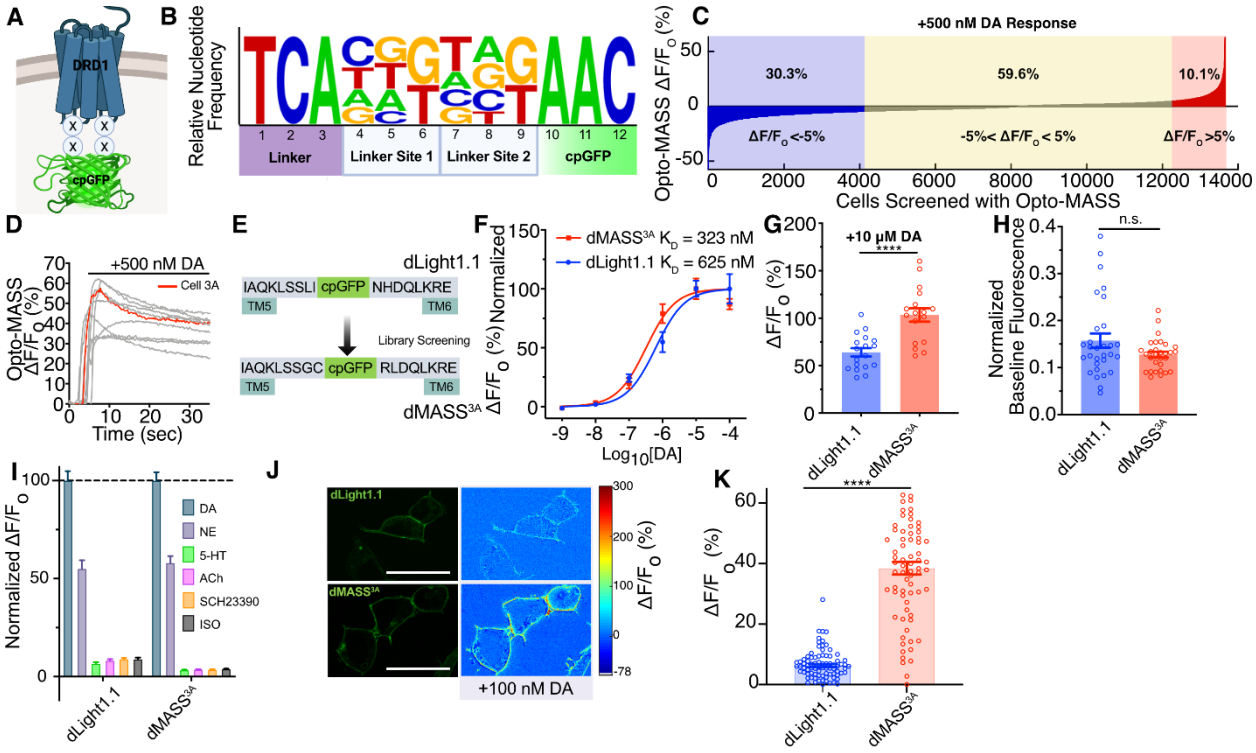

Figure 2.5: Opto-MASS engineers a high-performance DA sensor and *in vitro* characterization

A) Schematic of dopamine sensor library in the cell membrane with site saturated mutation locations denoted by X's. The four residues are located in the linker regions between the cpGFP and GPCR DRD1 domains and were targeted by site-saturated mutagenesis.

B) Summary of 23 sampled sequencing results from the c-terminal linker reveals near equal distribution of all possible nucleotides at the targeted sites.

C) Aggregate responses of dopamine sensor library. 13,656 cells screened. 30.3% of the screened cells had a DA dependent decrease in fluorescence ($\Delta F/F_o$ < -5%) 59.6% had no change in fluorescence and 10.1% had an increase in fluorescence ($\Delta F/F_o$ > 5%).

D) Exemplary responses from selected cells from the screen. Cell 3A, isolated to become variant dMASS^{3A}, highlighted in red.

E) Mutational changes post screening to linker regions in the dopamine sensor.

F) Apparent affinity of parent scaffold and Opto-MASS improved sensor under epifluorescence microscopy. (Single site binding n = 3 wells, 6 cells/well).

G) dLight1.1 and dMASS^{3A} response to 10 μM DA. ($p < 0.0001$, unpaired t-test, n = 5-6 cells/well, 3 wells).

H) Baseline fluorescence normalized to C-terminal tagged mRuby3. ($p = 0.0734$, unpaired Student's t-test).

I) Pharmacological selectivity of dLight1.1 and dMASS^{3A} to 10 μM of compounds, normalized to 10 μM DA stimulation(dLight1.1 100 ± 4.85%, dMASS^{3A} 100±4.29), NE is norepinephrine(dLight1.1 55.06 ±4.31 dMASS^{3A} 58.00±3.46%) , 5-HT is 5-hydroxytryptamine (dLight1.1 6.61± 0.69% dMASS^{3A} 3.32± 0.20%), ACh is acetylcholine (dLight1.1 8.22±0.58%, dMASS^{3A} 3.61 ±0.14%), SCH23390 (D_1 receptor antagonist dLight1.1 8.87±0.58%, dMASS^{3A} 3.58±0.16%), Iso is isoproterenol (dLight1.1 8.89±0.83%, dMASS^{3A} 3.80±0.21%). N = 40 cells. All plots mean ± SEM.

J) Representative confocal images of HEK293 cells expressing dopamine sensors and their membrane bound response to 100 nM dopamine bath addition. 30 μm scale.

K) dLight1.1 and dMASS^{3A} response to 100 nM dopamine imaged on a confocal. (>6-fold, $p<0.0001$, unpaired two-tailed t-test, n = 3 wells)

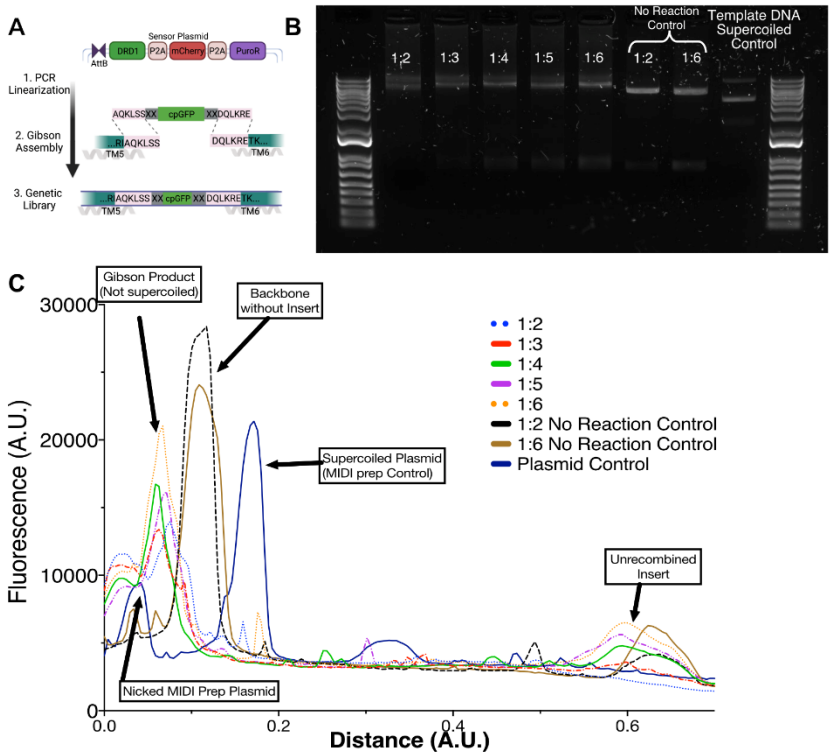

Figure 2.6: Optimization of sensor library generation

A) Schematic of dopamine sensor library construction using Gibson Assembly. (HiFi DNA Assembly Master Mix, NEB # E5520.

B) 1% agarose TAE gel. Lanes, from left to right: backbone:insert ratios: 1:2, 1:3, 1:4, 1:5, 1:6; 1:2 and 1:6 no reaction control and supercoiled DNA of the parent scaffold used for the library. The assembled plasmid travels slower than the linear backbone through the agarose due to not being supercoiled (see lane 1:2 Gibson Product). The molar amount of backbone added to each Gibson Assembly reaction was held constant. 1:6 molar ratio was used for library construction.

C) Fluorescence image analysis of the lanes, averaged over 30 pixels wide. Image analysis completed in FIJI.

2.2.7 *dMASS^{3A} detects dopamine signals* in vivo

Due to the broad utility of GPCR-based sensors in neuroscience, we aimed to demonstrate that sensors engineered by our pipeline are compatible with neuronal *in vivo* and *in vitro* detection methods. Here, we validated dMASS^{3A} *in-vivo* within the dorsal medial striatum (DMS) and nucleus accumbens (NAc) of mice during consummatory behavior using fiber photometry (**Figure 2.7A**). Dopamine signaling in the DMS and NAc are important modulators of operant and consummatory behavior[71], [72], but the precise dopamine signaling dynamics in these structures during free consumption of different reward magnitudes remains unclear. To investigate this, we recorded dopamine release using dMASS^{3A} using fiber photometry during limited windows of free-access consumption of 5 concentrations of sucrose (**Figure 2.7A**). We trained head-fixed mice to lick and consume sucrose during 100 trials of 3s access (**Figure 2.7B-D**). Mice exhibited consumption of sucrose that was dependent on the concentration of sucrose which was manifested as more licking for high concentrations of sucrose compared to low concentrations (**Figure 2.7C-D**; $F_{4,20} = 114.63$, P = 1.72e-13). dMASS^{3A} was strongly expressed in the NAc and DMS (**Figure 2.7E**) and showed clear dynamics during behavior (**Figure 2.7F-H**). In the DMS, dMASS^{3A} signals showed strong, transient increases in response to the onset of the access period that did not scale with the concentration of sucrose (**Figure 2.7G&H**). The transient increase in dMASS^{3A} fluorescence in the DMS coincides with the onset of licking behavior implies that dopamine release in the DMS may initiate but not sustain motor actions of consumption [73], [74]. On the contrary, in the NAc, dMASS^{3A} signals showed a two-component response with an initial rise at the onset of the access period and a secondary rise or drop in signal during the middle of the access period (**Figure 2.7G**). The mean and peak dMASS^{3A} fluorescence in the NAc showed clear scaling with the concentration of sucrose with higher fluorescence during consumption of higher concentrations of sucrose and decreases in fluorescence during access periods with lower concentrations of sucrose (**Figure 2.7G&H**). The change in dMASS^{3A} fluorescence in the NAc does not directly track licking behavior, as the dMASS^{3A} response begins to return to baseline before licking returns to 0 (**Figure 2.7C &G**). Instead, dMASS^{3A} response in the NAc may represent an initial detection response and a secondary value signal related to the concentration of sucrose during the access period[75].

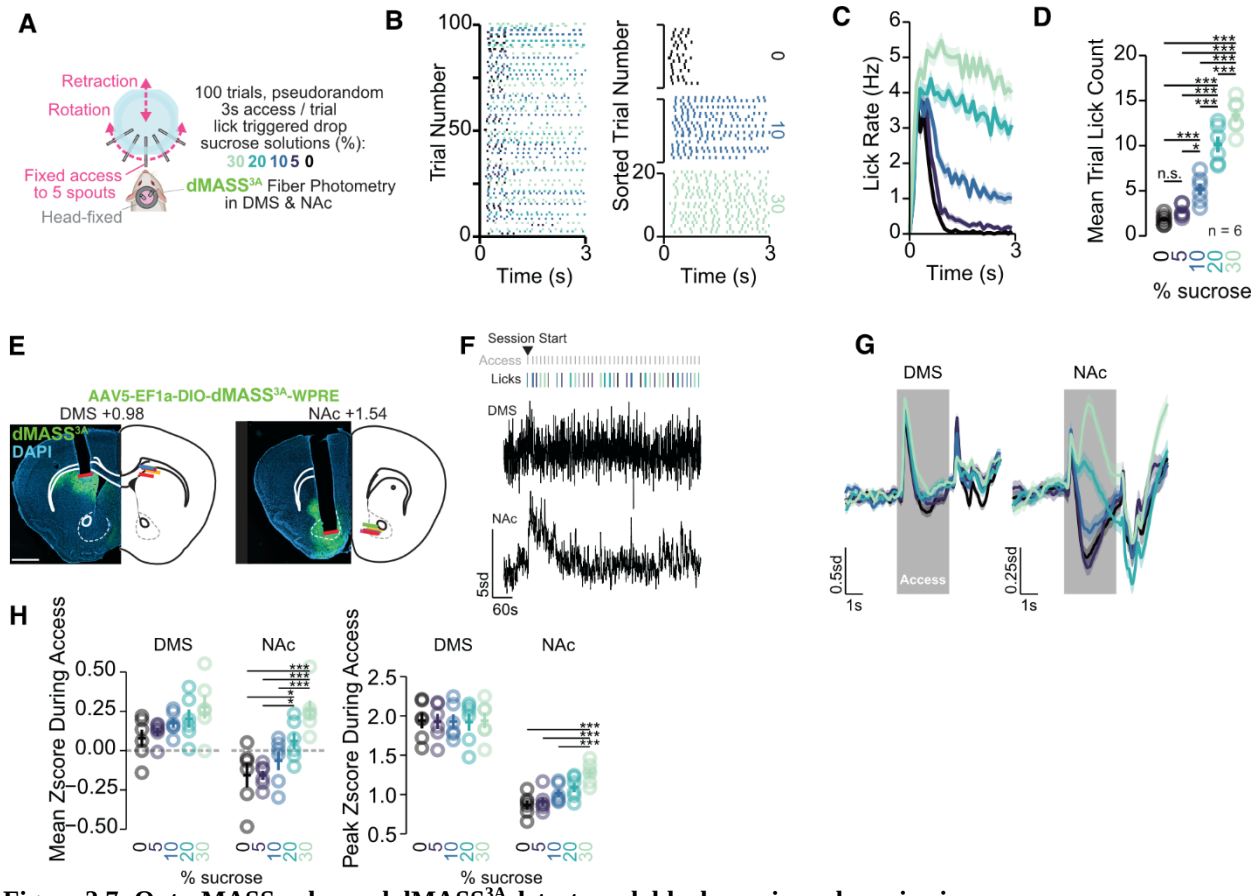

Figure 2.7: Opto-MASS enhanced dMASS^{3A} detects scalable dopamine release *in vivo*

A) Training RPE paradigm for pseudorandom access to sucrose solutions of varying concentrations. Over 3 sessions, mice were given 100 trials of 3s access to 5 concentrations of sucrose (colors in subsequent plots indicate concentration of sucrose).

B) Consummatory licks during 3s access to each concentration of sucrose displayed as mean binned licking over time (C) and mean total lick count during access for each concentration (D).

E) Histological images of dMASS^{3A} expression and fiber tip placement in the dorsal medial striatum and Nucleus Accumbens.

F) Representative fiber photometry trace of dMASS^{3A} fluorescence simultaneously recorded in the DMS and NAc. Grey lines indicate 3s access periods and Lick lines indicate licks for each concentration.

G) Mean fluorescence traces in the DMS (left) and NAc (right) over time during access to each concentration (ribbon depicts SEM).

H) Mean (left) and peak (right) fluorescence during the access period shows clear scaling in the NAc but not in the DMS. (2-way repeated measures ANOVA, pair-wise Tukey HSD post hoc. *: P<0.05; **: P<0.01; ***:P<0.001).

2.3 DISCUSSION

Our goal was to construct a high-throughput engineering pipeline for GPCR-based sensors that overcomes the resource constraints of traditional techniques. We used our pipeline to functionally test 13,656 cells expressing a dopamine biosensor library in mammalian cells and identified the high-performance biosensor, dMASS^{3A}, that has a >6-fold increase in response to 100 nM DA compared to the parent scaffold. 30.3% of the screened cells had a DA dependent decrease in fluorescence ($\Delta F/F_o < -5\%$), 59.6% had no change in fluorescence and 10.1% had an increase in fluorescence ($\Delta F/F_o > 5\%$). We validated our identified sensor, dMASS^{3A} *in vivo*. dMASS^{3A} folded, expressed, and trafficked in mouse NAc and DMS, and recorded behavior-linked changes in DA during reward prediction error testing paradigms.

The Opto-MASS platform had a low coefficient of variance when imaging isogenic populations of cells and could rapidly measure and assess a sensor variants biophysical performance. Our pipeline provides several advantages over the field's current methods. First, we harness a mammalian expression system that ensures only a single variant from our library is expressed in each cell. By expressing our libraries in mammalian cells, we immediately overcome protein folding and expression challenges that arise from screening libraries in *E.coli* or yeast. Additionally, biophysical characteristics of fluorescent sensors in protein lysates, *E.coli*, or yeast may not transfer to mammalian cells [2]. Consequently, each variant that is screened passes the hurdle of mammalian transgene expression, protein folding dynamics, and cell surface trafficking. Additionally, using doxycycline inducible expression system ensures a controllable expression level for our constructs. We can use commercial transfection reagents to induce stable integration, removing the need for difficult viral packaging or dilution with 'dummy' plasmids with calcium phosphate transfection.

Second, due to time and resource restrictions researchers usually functionally screen only hundreds or low thousands of variants for each sensor. Opto-MASS functionally screens thousands of variants a day. While still short of statistically covering the possible mutation space, Opto-MASS is orders of magnitudes faster than alternative *in vitro* methods. Thus, the unbiased approach is a step towards more efficiently covering the mutational space and identifying optimized sensor variants.

Third, the platform can be expanded to engineer other light-based optogenetic tools. Red shifted optogenetic sensors are particularly difficult to develop due to their aggregation, low signal to noise ratio, and off spectrum activation. Opto-MASS is modular, and red-shifted GEFIs can be engineered after cloning in cpmRuby or cpmApple and imaging under 560 nm excitation. Opto-MASS can also be expanded to engineer proteins that modulate physiology under light stimulation, or optogenetic actuators. If the actuators physiological effect can be measured with a GEFI, the optimization of the actuator just requires a spectrally orthogonal GEFI to measure phenotypic differences in the actuator library.

The sensor dMASS^{3A} is optimized to detect low concentration detection of DA *in vivo*. While dMASS^{3A} demonstrated no loss of molecular selectivity compared to the parent construct, high concentration NE release *in vivo* may mask low concentration DA release. Additionally, rapid neural firing has been demonstrated to lower intracellular pH, and the pH sensitivity of the sensor has not been adequately explored. The residues flanking the cpGFP domain suggest they coordinate the phenol to phenolate state transition of the GFP chromophore, so the sensor may be particularly sensitive to changes in intracellular pH [39]. Rapid reductions of dMASS^{3A} fluorescence signal *in vivo* may be due to pH decreases or rapid reuptake of DA. Furthermore, our *in vivo* experiments highlighted the dose dependent nature of DA release in the NAc, and large releases of DA in the DMS. The experiments are a remarkable demonstration of dMASS^{3A}'s ability to detect dopamine release *in vivo*. Further *in vivo* experiments should express the sensor in regions of the brain known to have lower, tonic DA releases, such as the medial prefrontal cortex (mPFC). Two-photon imaging of dMASS^{3A} in the mPFC could highlight the sensor's ability to detect low concentration releases of DA in the mammalian nervous system.

Despite the advances Opto-MASS and the improved tool dMASS^{3A} present to the field, there is room for optimizations and further applications to the system. First, our gene recovery rate was around 30-40%, and further improvement would make the identification of high-performance genes easier. The throughput of our technique is approximately 10,000 HEK293T cells per day. So far, we tested ~26,000 dopamine sensor variants during pipeline optimization and library screening, which is less than 2% of the 20^4 possible mutations. Other fluorescence measurement techniques, such as flow cytometry can screen an order of magnitude more mammalian cells in a day. However, flow cytometry takes a single, instantaneous fluorescence intensity measurement in time, and cannot make measurements like cell surface expression and kinetics. While these

biophysical characteristics were not used to screen sensors during the dMASS^{3A} screening, computation and screening thresholds for these elements could easily be implemented.

Other pipelines have been developed to engineer genetically encoded biosensors in mammalian cells. Opto-MASS provides a direct link of the sensor functional phenotype and genotype. Other pipelines can provide that direct phenotype-genotype linkage but cannot functionally screen thousands of sensors [46], or can functionally screen thousands of cells but only provide mutational enrichment scores [76] . Opto-MASS provides both, functional screening of thousands of biosensors and a direct phenotype-genotype linkage. However, protein engineering strategies that look at changes in variant enrichment at a population level, like deep mutational scanning, may reveal mutational combinations not identified with the direct phenotype-to-genotype strategy, or may miss them all together. While Opto-MASS did not screen sensors for biophysical characteristics such as kinetics and membrane localization, sensors should be assessed on kinetics and cell surface expression as these are critical performance metrics. Some pipelines screened sensors on these biophysical characteristics [46]. We chose to not make these types of measurements with Opto-MASS, as we determined that sensors that traffic to the membrane poorly would not survive antibiotic selection or have high enough dynamic range for selection on Opto-MASS.

We used degenerate NNK codons to introduce mutations into our peptide linkers. NNK nucleotides bias the library to the amino acids represented by the frequency they appear in the NNK codon space. Site saturated mutagenesis techniques that incorporate amino acids at equal ratios may increase the chance that a low probability linker identity is measured and observed on Opto-MASS arrays. Additionally, the sensor library was cultured in HEK293T cells for seven to ten days at high expression levels, under mammalian antibiotic selection. Sensor mutants that are translated, folded, and trafficked to the cell surface with lower metabolic cost may be overrepresented in the screened population as their host cells divided more rapidly. While this genetic change due to cell division may in theory be beneficial, mature neurons do not divide and cell division mediated selective pressure may represent a selective force on the sensor characteristics that is not relevant to the desired use case in the central nervous system. Variant frequency changes due to length in culture could be sidestepped by using FACS to identify recombined HEK293T TetBxb1BFP cells approximately 48 hours after library expression.

Opto-MASS presents the first high-throughput genetically encoded biosensor pipeline that functionally screens thousands of variants in mammalian cells and provides a direct phenotype-genotype linkage. Opto-MASS also provides several beneficial aspects. After transfection and expansion, we can freeze the library expressing cells in liquid N_2 and recover them quickly before additional screening sessions. This effectively requires only one library generation, landing pad cell transfection, selection and expansion step each for each library. In contrast, traditional testing methods would require months if not years of preparations, measuring and analysis and significantly more resources for DNA purification and cell cultures. Opto-MASS has demonstrated its ability to rapidly measure the biophysical characteristics of thousands of monoamine sensors in multiplex and identify high performance sensors. The enhanced throughput of the Opto-MASS system presents the next chapter in the field of optogenetic tool engineering, opening new avenues as protein engineers will not be slowed down by traditional techniques which require significant resource investment to develop high quality sensors.

2.4 METHODS

2.4.1 *Microwell Array Fabrication*

The master mold (negative) of the microwell arrays was fabricated using a ICP4-SPTS-DSi using deep reactive ion etching (DRIE) to etch silicon wafers. To prepare the wafer for DRIE, AZ1512 photoresist (MicroChemicals) was spun onto 100 mm silicon wafers (University Wafer) with a two-step process. First, wafers were spun at 500 RPM for 5 sec to spread the photoresist, with a final thirty second spin step at 2500 RPM. After incubation at 100°C for 60 seconds on a hotplate, the wafer was exposed with a chrome on borosilicate glass mask and developed with AZ340 developer (MicroChemicals).

Polydimethylsiloxane (PDMS) (Sylgard 184, Corning) was used to construct the microwell arrays. The elastomer and the curing agent were thoroughly mixed at a ratio 10:1 (w/w), poured onto the silicon wafer and placed in a desiccator to remove air bubbles. After the air bubbles were removed, the wafer was moved to a 55-75° C incubator to cure for several hours.

After curing, the PDMS was removed using razorblades and forceps. The PDMS was plasma treated for sixty seconds and then firmly pressed onto a dried bovine serum albumen (BSA) layer to nanostamp a layer of BSA onto the surface (FischerSci Cat# BP1600). The dried BSA layer was made by incubating a 2% w/v solution of BSA in PBS in a Petri dish for approximately 30 minutes. After, the dish was rinsed 3X with PBS and left to dry. After pressing the plasma charged PDMS into the BSA, debris were removed from back of the PDMS using tape. The PDMS microwell arrays were then cut out using a scalpel and placed upright into a glass bottom 24 well dish.

2.4.2 *Building Genetic Libraries for Screening on the Platform*

Gibson Assembly was used to build the genetic libraries of sensors. The insert and backbone were PCR amplified using Q5 polymerase (New England Biolabs (NEB) M0515) and degenerate codons (IDT) were introduced with primers during PCR amplification of the cpGFP insert. To subclone the cpGFP moiety with two flanking mutational regions, we used primers to amplify the cpGFP domain out of a mammalian expression plasmid containing GCaMP6f. 1 µL of DpnI was used to digest PCR templates. PCR products were isolated from a 1% agarose gel stained with SyberSafe (Invitrogen Cat #S33102) and New England Biolabs (NEB) Monarch DNA

Gel Extraction Kit (Cat # T1020L). After gel isolation, the insert and backbone were assembled using NEB HiFi DNA Assembly (NEB Cat #: E2621). A total of 0.2 pmol of DNA was used in the Gibson Assembly, with a 6:1 molar ratio of insert (cpGFP moiety) to vector and incubated at 50° C for 60 minutes.

The assembly was cleaned up with NEB PCR cleanup kit and double eluted from the column with 10 µL of prewarmed water. We then transformed 33 µl of electrocompetent cells (NEB Cat #C3020K) (2000 V, τ = 5ms) in ice cold cuvettes (1 mm gap) with 2 µL of the elution. Immediately after pulsing the cells, 967 µL of prewarmed SOC was added to the cuvettes. After 1 hour recovery at 37° C and 240 RPM in a 15 mL recovery tube, a dilution of the recovery media was plated on an agar plate and grown overnight to estimate library size. The remaining recovery media was added to 125 mL of Luria Broth with ampicillin and grown overnight at 37° C and 240 RPM. The 15 mL recovery tube was rinsed several times with fresh LB to ensure all transformants were added to the large overnight culture.

After overnight growth, the agar plate was counted to estimate total transformants and colonies were randomly selected for library sampling. Library plasmids were isolated using Machery-Nagel NucleoBond Xtra Midi EF kit (Ref # 740420.50). The resulting plasmid prep was used to generate stable integrated cell lines with the landing pad cell line. The dopamine sensor library used dLight1.1 and the MOR sensor from reference [9].

2.4.3 *Library Transfection into Landing Pad Cells and Cell Seeding on Microwell Arrays*

After validation the library was correctly assembled in transformed *E.coli* through Sanger Sequencing selected colonies, HEK293T landing pad cells were stably recombined with our library using a double transfection protocol. Landing pad cells were maintained in standard growth media supplemented with 1-2 µg/mL doxycycline. The day of transfection, the cells were gently lifted off the growth substrate using 0.05% Trypsin/EDTA (Invitrogen Cat # 25300120). Lift off was stopped by adding growth media, approximately 250,000 cells per well were seeded into 6 well dishes and final culture volume was 2 mL. The DNA transfection reagents were prepared by incubating 3 µg of plasmid DNA encoding the Bxb1 recombinase and 6 µL of Fugene6 reagent (Promega Cat # E2693) in 300 µL of Opti-MEM for 15 minutes and then added to the cell suspension. After 24 hours incubation, cells were lifted off the growth substrate and centrifuged at 500 RCF for 5 minutes. After seeding at 250,000 cells/well in a 6 well dish, the cells were

transfected a second time using the same protocol with library plasmids. For each round of library screening, five wells of the six well plate were transfected with the genetic library and combined after puromycin selection for screening.

The nano-stamped PDMS microwell arrays were then briefly plasma treated to charge the inner wells again. Quickly after plasma treatment, standard growth media was added to the wells, and they were placed in a desiccator to remove bubbles in the microwells. The plates were briefly returned to the tissue culture incubator to raise the temperature of the media and balance the pH. Next, the landing pad cells were lifted from the growth substrate with 0.05% Trypsin/EDTA. Once a single cell suspension was achieved, the cells were counted and then 40K cells were seeded at a concentration of $0.5X10^6$ onto the arrays. Cells were slowly pipetted above the array with a micropipette. The cells were returned to the incubator for 10 minutes. After 10 minutes, the 24-well plates were then placed in a centrifuge and spun down at 100 RCF for 5 minutes. The arrays were then rinsed with growth media several times to remove cells not in microwells and cell debris. The final rinse is with DMEM/10% FCS supplemented with doxycycline at the concentration used during selection and half selection puromycin concentration. The cells were then returned to the incubator overnight.

2.4.4 *Library Screening and Cell Selection*

The morning of cell selection experiments, the arrays were washed twice with standard growth media, and then once with imaging Tyrode's supplemented with GlutaMax (Gibco Ref: 35050-1), sodium pyruvate (GIBCO Ref: 11360-070) and MEM Non-Essential Amino Acids (Gibco Ref: 11140-050). A MetaMorph Journal was used to control the microscope during the imaging sequence. In brief, during image capture stage, the arrays were imaged for the control fluorophore and then were imaged at GFP excitation/emission continually for one minute. Ligand was added by automatic pump or hand to the bath. The control fluorophore was imaged and a bright field image. The images were then analyzed using MetaMorph.

ROIs with the greatest response were identified and added to the live field of view and the image stack of the stimulation. To define the greatest responding ROI's, the average intensity projection of baseline and stimulation images are divided. The image numbers that make up the 'baseline' and 'stimulation' are dependent on the time of ligand addition. The baseline images are typically defined as the first ten images of the stimulation time course. The proportion of images

from the stimulation stack post stimulation are selected to be a brief period after the ligand addition to the end of the stack.

The user can verify the selected ROIs prior to cell picking using glass micropipettes. The micropipette tips were then transferred to 200 µL PCR tubes with 5 µL of a TE/DTT buffer and immediately placed on dry ice for the remainder of the screening session. Positive pressure was applied during tip breakage into the solution.

After the screening session is over, the samples were processed to covert the mRNA of the sensor into cDNA using an adapted protocol of the ThermoFischer SuperScript IV Reverse Transcriptase (SSIV RT) protocol. 2 µL of the cDNA product was then amplified in a 25 µL reaction using Q5 polymerase. After DNA cleanup, the PCR product was Sanger Sequenced to check for contamination and cloned into a pCMV backbone using Gibson Assembly to validate the genes performance. After transformation of the Gibson Assembly into chemically competent cells, the clones were grown up in 5 mL Luria Broth cultures with (100 µg/mL) ampicillin and plasmid DNA was isolated using the Machery Nagel Endotoxin Free Miniprep kit. The plasmids were transfected into HEK293 WT cultures in plastic 24 well dishes to validate the genes performance.

2.4.5 *Selected Gene Recovery: Reverse Transcriptase Reaction*

Single cell recovery tubes were prepared by diluting 5 µL of 0.1 M DTT into 200 µL of TE buffer. 5 µL of the TE/DTT buffer was added to each PCR tube. After a single cell was deposited into the tube, the tubes were incubated on dry ice for the remainder of the library screening session. After library screening, the tubes were removed from the dry ice and placed on wet ice. Each tube was processed with reagents from SuperScript IV First Strand Synthesis kit (Invitrogen Cat# 18091050). To each tube, 0.5 µL of 0.1 M DTT and 0.5 µL of RNAse Inhibitor was added. The samples were then placed on dry ice for five minutes, and then moved back to wet ice. Next, 0.5 µL of the following was added to each tube, DI H_20, a 10 mM dNTP mix, and 2 µM primer.

Next, the primers were annealed to the mRNA by incubating the samples at 95°C for 30 seconds, 4° C 1 minute, 65° C 5 minutes. The samples were then returned to wet ice. Next, 2 µL of SSIV RT 5X Master Mix was added to each sample. The samples were pipetted up and down thoroughly. Finally, 0.5 µL of the SSIV RT Enzyme was added to each tube and the samples were

pipetted up and down thoroughly. The reverse transcriptase reaction was carried out in the following manner the samples were incubated at 53°C for 10 minutes, and then 80° C for 10 minutes to inactivate the reverse transcriptase. Next, 0.5 µL of RNAseH was added to each tube and the samples were incubated for 20 minutes at 37° C to remove any mRNA from the cDNA. Samples were stored at -20° C prior to PCR amplification of cDNA with Q5 (New England Biolabs) or SuperFiII (ThermoFischer).

2.4.6 *HEK293 Maintenance and Transfection for* in vitro *imaging assays*

HEK293 cells were cultured on tissue culture treated plastic in DMEM supplemented with 10% Fetal bovine serum (FBS) and 1% penicillin and streptomycin at 37° C with a 5% CO2 atmosphere. One day prior to transfection, cells were lifted off the growth substrate with 0.05% Trypsin/EDTA. The cells were then seeded into 24 well tissue culture plates. Cells were grown to 70-80% and then transfected. During transfection, growth media was replaced with fresh 250 µL of media. The DNA transfection reagents were prepared using the standard protocol. In brief, per well of transfection, 25 µL Opti-MEM, 1 µg of DNA and 1.5 µL of P3000 were mixed. After five minutes of equilibration, the Opti-MEM/DNA/P3000 mix were added to a tube containing 25 µL Opti-MEM. And 1.5 µL of Lipofectamine. The DNA/P3000/Lipofectamine were incubated for approximately 15 minutes at room temperature before addition to the wells. After incubation for 3-4 hours, the transfection media was removed, and fresh media was added. Reactions were scaled for different wells according to the manufacturer's directions.

Cells were imaged with an sCMOS camera (Photometrics Prime95B) on an epifluorescent microscope (Leica DMI8) using a 20X objective (Leica HCX PL FLUOTAR L 20x/0.40 NA CORR) forty-eight hours after transfection. A Lumencor Light Engine LED and Semrock Filters were used for fluorescence imaging.

Prior to imaging, cells were rinsed once with Tyrode's (125 mM NaCl, 2 mM KCl, 2 mM $CaCl_2$, 2 mM Mg Cl_2 30 mM Glucose and 25 mM HEPES). Cells were imaged in Tyrode's solution that was supplemented with GlutaMax (Gibco Ref: 35050-1), sodium pyruvate (GIBCO Ref: 11360-070) and MEM Non-Essential Amino Acids (Gibco Ref: 11140-050). Bath additions of ligands were done by hand for validation experiments and the volume added was always equivalent to the pre-addition bath volume. Ligands were prepared in a Tyrode's solution.

For confocal images of fluorescence responses HEK293 cells were plated onto poly-l-lysine (Cultrex 3438-100-01) coated glass bottom plates and imaged in Tyrode's solution with a Nikon A1R microscope with an 40X oil objective (CFI Plan Fluor NA 1.30) at room temperature (≈23°C). A 488 nm laser was used for GFP and sensor imaging and a 561 nm laser was used for red fluorophore imaging. Ligands were hand pipetted into the bath.

Fluorescence change graphs were generated by taking the average intensity projection of five frames prior to ligand addition, five frames after ligand addition in FIJI (NIH). The resulting images were then divided in MATLAB (2019a) to determine pixel-by-pixel fluorescence change, and a color scale was overlaid.

2.4.7 *Affinity Curves*

HEK293WT cells were transfected with Lipofectamine 3000 reagents for the dopamine affinity curves. In brief, HEK293WT cells were seeded onto poly-L-lysine coated 96 well glass bottom plates and expanded to 70-80% confluency and transfected. After transfection, the cells were incubated for 24-48 hours prior to imaging.

The day of imaging, cells were rinsed and imaged in 50 µL of supplemented Tyrode's. During imaging, 150 µL of a dopamine and Tyrode's solution was hand added to the bath during imaging. Cells were imaged with a 63X air objective.

Cells were analyzed in FIJI (NIH). Five to six cells from each well were hand circled from background subtracted image stacks (Rolling ball, 100 pixels). The average fluorescence intensity was measured for each cell and exported to Excel for analysis. Analyzed data was imported into GraphPad Prism 8 to calculate EC50 values using the nonlinear fit function and Least Squares fit.

2.4.8 *Calculation of $\Delta F/F_o$ and Opto-MASS Array Rank Ratio*

The change in fluorescence was measured by hand circling regions of interest (ROIs) around background subtracted images in FIJI. The ROIs were measured for mean grey value over time. The measurements were imported into Excel, where Equation 1 was used to calculate the change in fluorescence:

$$\Delta F/F_o\ (\%) = \frac{(F_i - F_{baseline})}{F_{baseline}} \times 100 \qquad (2.1)$$

Where F_i is the ROIs mean fluorescence for a frame and $F_{baseline}$ is the average fluorescence for the first five or six frames after imaging began. When peak fluorescence is shown, the maximum fluorescence is identified in a time course, and five frames surrounding maximum fluorescence are averaged to account for noise in the fluorescence measurement.

To assess a ROIs ligand dependent fluorescence response, the Opto-MASS Array Rank Ratio is calculated to compare to other ROIs in the field of view. The Opto-MASS Array Rank ratio is calculated by the following equation:

$$\text{Array Rank Ratio} = \frac{\text{AIP}_{\text{stimulation}}}{\text{AIP}_{\text{baseline}}} \times 1000 \quad (2.2)$$

Where $AIP_{Stimulation}$ is the average intensity projection of the imaging period after ligand stimulation, and $AIP_{baseline}$ is the average intensity projection of the time period before ligand stimulation and Array Rank Ratio is a unitless number used to assess the ligand dependent fluorescent response of an ROI. The imaging period $AIP_{stimulation}$ start time is defined several frames after the ligand addition to remove any motion or kinetic artifacts.

2.4.9 *Molecular Cloning*

Unless explicitly stated, DNA constructs were cloned with either Q5 Polymerase, Platinum SuperFi II, Site Directed Mutagenesis, Gibson Assembly, *In Vitro Assembly*, or standard restriction enzyme cloning. PCR products were verified on 1% agarose gels stained with SyberSafe and cleaned up with NEB Monarch PCR Clean Up kits. The pCAG NLS HA Bxb1 plasmid was gifted from the Fowler lab (Addgene #51271), the AttB puromycin plasmid used to generate libraries was gifted from the Fowler lab and the pCMV dLight1.1 plasmid was sourced from Addgene (#111053), and rat MOR plasmid was a gift from the Bruchas Lab.

2.4.10 In vivo *fiber photometry dMASS^{3A}*

Under isoflurane anesthesia (5-2%), 6 heterozygous Vgat-cre mice (9 weeks old, 3m and 3f) were injected with 400nL (Nanoject III, 1nL / minute) of AAV5-DIO-dMASS^{3A} into the dorsolateral striatum (AP: 1.25; ML: +/-1.6; DV: -2.3; angle: 10) and the contralateral nucleus accumbens (AP: 1.7; ML: +/-1.5; DV: -4.5 angle: 10) with hemispheres balanced across subjects. We then lowered 6mm optic fibers (Doric: 400µm, 0.37 NA, 1.25 zirconia ferrule) 0.1mm above

each injection target and fixed the fibers to a headring and the mouse's skull using super glue and dental cement.

Following 1 week of recovery and return to presurgical body weight, mice were food restricted to 90% of their baseline body weight for 5 days prior to behavioral sessions. Mice were habituated to handing and the head-fixation apparatus for 2 days before being head-fixed. The head-fixed behavioral apparatus consisted of a custom, 3d printed head fixation stage and 5x multi-spout that was able to rotate and retract using micro servos (Tower Pro SG92R). Each multi-spout was attached to an independent solenoid (Parker) which were calibrated before the experiment to ensure they delivered ~1.5 µL per delivery. Licks were detected on each spout using a capacitive touch sensor (Adafruit MPR121). An Arduino Mega was used to control hardware and record the timing of behavioral and hardware events.

Mice were first trained to consume sucrose from a metal lickspout in a single 10-minute free-access session with a single spout in the extended position that mice could freely lick for 30% sucrose. Free consumption was achieved by delivering a drop sucrose via solenoid opening (~1.5 µL over ~15ms / delivery) immediately following each lick. Next, mice were trained on the forced-choice free-access multi-spout assay. Each session consisted of 100 trials of 3s access to 1 of 5 concentrations of sucrose (0, 5, 10, 20, 30%) with a random inter-trial interval of 11-16s drawn from a uniform distribution. During each access period, the spout extended forward, and the mouse could lick for sucrose for 3s, then the spout was retracted, and the multi-spout head was immediately rotated to the spout of the next trial. Mice were trained over 11 sessions and were then recorded over 3 sessions.

We recorded dMASS^{3A} fluorescence in the NAc and DMS simultaneously by connecting each fiber to patch cables (Doric: 400µm, 0.37 NA, 1.25 zirconia ferrule) coupled to a 5-port mini cube (Doric) and integrated fiber photometry system (RZ10X, Tucker-Davis Technologies). We used 465nm light modulated at 331 Hz for measuring dMASS^{3A} fluorescence and 405nm light modulated at 209 Hz for measuring autofluorescence. Light emission was collected using the same fiber and was measured using a photosensor (Lux). During collection, signals were low pass filtered at 6Hz and demodulated. Excitation power for both wavelengths was set to 30 µw. The timing of hardware and behavioral events were recorded using TTL inputs to the fiber photometry system.

Fiber photometry signals for dMASS^{3A} were post processed using custom Python and R scripts. The 405nm channel was inspected for abrupt changes in signal power that could be attributed to motion, but none were observed (probably because the mice were headfixed throughout recording). As result, the 405nm channel was not used to correct the dMASS^{3A} signal and was discarded from further analysis. We corrected for photobleaching for each session by fitting and subtracting a 4th degree polynomial. We normalized the fluorescent signal by taking a z-score using the mean and standard deviation of the signal throughout the entire session. The signal was then smoothed using a 100ms moving average and then down sampled to 20Hz. Next, we used behavioral time stamps to extract peri-event time histograms centered on the access period and then baseline corrected by subtracting the mean signal during the 3s prior the onset of access.

Chapter 3. GENETICALLY ENCODED SENSORS FOR DETECTION OF OPIOID NEUROPEPTIDES AND OPIOIDS[2]

ABSTRACT

The mu opioid receptor (MOR) was one of the first pharmacologically targeted receptors in human history. The MOR is a GPCR that mediates the pain-relieving actions of opioids and is involved with a host of physiological roles such as maternal bonding, social interaction, feeding, addiction, and respiratory regulation. Scientists currently lack tools to measure endogenous opioid release *in vivo* in freely behaving animals in a cell specific manner. GPCR-based optogenetic sensors have recently been applied to making biosensors with high spatiotemporal precision. However, initial MOR sensor scaffolds do not have a large enough dynamic range and signal to noise ratio for *in vivo* applications. Due to the unpredictable nature of protein folding, many hundreds to thousands of sensor variants must be tested to identify high performance sensors. Traditional techniques to validate sensor variants take months, if not years. Previously, we reported on a high throughput pipeline, Opto-MASS, that rapidly constructed an optimized dopamine sensor by screening more than ten thousand mammalian cells expressing a dopamine sensor library. Here, we apply Opto-MASS to develop an opioid sensor. The pipeline identified the improved variant μMASS^{2A} after screening over 23,000 HEK293T cells with 1 micromolar [D-Ala^2, N-MePhe^4, Gly-ol]-enkephalin. μMASS^{2A} has a ~4.6-fold and ~3.8-fold greater response over the parent scaffold to 500 nM and saturating concentrations of DAMGO. μMASS^{2A} binds the endogenous enkephalin, methionine-enkephalin with high affinity. μMASS^{2A} does not internalize during *in vitro* tests and detects morphine administration *in vivo* in rodent models. μMASS^{2A} represents the next step towards developing highly sensitive optogenetic sensors for the *in vivo* detection of neuropeptides.

[2] This chapter has text directly from and adapted from: Rappleye M, Gordon-Fennell A, Zamorano C A, Castro D C, Matarasso A K, Stine C, Wait S J, Lee J D, Siebart J, Suko A, Smith N, Muster J, Matreyek K A, Fowler D.M , Stuber G D, Bruchas M R, Berndt A, "Opto-MASS: A high-throughput protein engineering platform for genetically encoded fluorescent sensors enabling all optical *in vivo* detection of monoamines and opioids" *Pre-print, BioRxiv:* https://doi.org/10.1101/2022.06.01.494241

3.1 INTRODUCTION

Opioids are among the oldest drugs in human history [77]. Opioids have been used for medicinal and recreational purposes since the ancient Sumerians and the alkaloids from the opium poppy have a significant role in human history, shaping modern medicine, ancient and modern culture, wars, and geopolitics through time. Scientists have long sought to understand the link between exogenous opioids, endogenous opioids and the opioid receptors and their roles in mammalian physiology. There are four known opioid receptors, mu, kappa, delta, and nociceptin. All four of the opioid receptors are G-protein coupled receptors (GPCRs), which are membrane bound receptors with seven transmembrane regions that modulate the intracellular environment through G-protein signaling. There is high sequence homology between the four receptors, with the DOR, KOR, and MOR sharing greater than 70% sequence homology in the TM regions [78].

All GPCRs are important pharmacological targets, with complex intracellular signaling cascades that contribute to the regulation of almost every aspect of mammalian physiology. The mu opioid receptor (MOR) contributes to the regulation of a significant amount of human emotion, appetite, addiction, socialization, maternal-infant bonding, intestinal mobility, immune cell activation, and other critical functions [79]–[84]. The MOR is encoded by the *OPRM1* gene and maintains the canonical seven transmembrane domain (TM) topology found in class A GPCRs and couples the inhibitor G-protein, $G\alpha_{i/o}$ [85]. Researchers first cloned the receptor in 1993 [86]–[88], and proceeded to generate gene knockout mice. $MOR^{-/-}$ mice lose the pain-relieving properties of morphine, the place preference mediated by morphine, and evidence of morphine dependence [89]. While understanding the role of all four opioid receptors is important to understanding mammalian biology, the MOR mediates the pharmacological effects of classic exogenous opioids like morphine, so understanding the MOR's function and ligand interactions in the central nervous system is critical. The physiological and societal effects of the MOR and its ligands has placed it at the forefront of scientific research. Understanding MOR's involvement with cognition *in vivo* will broaden our understanding of the regulation of physiological processes such as addiction, socialization, maternal-infant bonding, and pain-relief.

Researchers want to link opioid signaling in neural circuits to animal behavior *in vivo*. Currently, there are no tools that have the requisite cell-type specificity, kinetics, or dynamic range to detect opioid release *in vivo*. Microdialysis has been implemented with optogenetics to release

opioid peptides, however, microdialysis lacks the cell-type specificity for the site of opioid action and kinetics to directly link animal behavior to opioid peptide signaling [90]. Genetic approaches, such as a nanobody sensor, can be targeted to specific genetic cell subtypes, however nanobody sensors require total internal reflection fluorescence (TIRF) microscopy [91]. TIRF currently cannot be used *in vivo,* so nanobody sensors cannot link animal behavior and opioid activity. A single chain nanobody-fluorophore opioid sensor, named MSPOTIT, has been developed [92]–[94]. However, the biosensor lacks the pharmacological selectivity of the endogenous MOR. DOR selective agonists, such as DADLE, activate the sensor. The sensor's off-kinetics are on the order of days, hindering its ability to link animal behavior to opioid release in neural circuits as endogenous signals occur on faster timescales than days. Additionally, the biosensor has not been demonstrated to express *in vivo,* and improved variants needed tissue fixation for measurement benefits to become apparent.

A GPCR-based GEFI has been constructed out of the rat MOR[9]. However, it lacks the dynamic range and cell surface expression needed for *in vivo* applications. The sensors expression levels, brightness, folding kinetics and dynamic range need to be improved for *in vivo* use. Traditional engineering techniques would dedicate significant resources and time to optimize the sensors characteristics above the threshold needed for *in vivo* work. Here, the Opto-MASS pipeline can be applied to quickly screen thousands of sensor variants and identify a more optimal sensor capable of *in vivo* ligand detection.

Using the Opto-MASS pipeline, we screened over 23,000 mammalian cells expressing a genetic library of MOR sensors and isolated an improved variant with a ~4.6-fold and ~3.8-fold greater response over the parent scaffold to 500 nM and saturating concentrations of [D-Ala^2, N-MePhe^4, Gly-ol]-enkephalin (DAMGO), respectively. The improved variant is dubbed μMASS^{2A}. We fully characterized the biophysical properties of the sensor and its loss of function control variant *in vitro* and demonstrated its ability to detect opioid ligands *in vivo*. We demonstrate Opto-MASS can construct and optimize a neuropeptide-based sensor capable of *in vivo* opioid detection. The engineering of μMASS^{2A}, a neuropeptide sensor, expands the sensor space Opto-MASS is capable of improvement from monoamines to more sensitive biosensors required for neuropeptide detection.

3.2 RESULTS

3.2.1 *Opto-MASS engineers a neuropeptide sensor capable of* in vivo *detection of opioids*

For the final validation of the capabilities of our pipeline, we chose to engineer a GPCR-based sensor framework that currently lacks *in vivo* detection capabilities. We selected to optimize a sensor prototype called mLight which is based on the Mu-opioid GPCR (MOR)[9]. Neuropeptides such as endogenous opioids are hypothesized to function through low concentration volume transmission instead of fast, high concentration synaptic transmission like monoamine neurotransmitters [95]. The low concentrations make *in vivo* neuropeptide detection more difficult. Endogenous opioid peptides include endorphins, enkephalins, dynorphins and nociceptin which help regulate motivation, stress, reward, gastrointestinal mobility, hedonic homeostasis, feeding and other behaviors through the opioid receptors [96]–[99]. The opioid peptides bind to different subtypes of opioid receptors (mu, delta, kappa, nociceptin) with varying affinity but are rarely exclusive to just one target receptor [100]. Current techniques to monitor opioid peptide release either lack cell-type specificity, the kinetics to link signaling events with animal behavior or are incompatible with current *in vivo* imaging technologies [90], [92], [101]. Previously, researchers have inserted the cpGFP moiety into the ICL3 of the mu opioid receptor (MOR) to make a prototype opioid GEFI mLight, however it suffered from poor dynamic range and cell surface expression, precluding it from *in vivo* use [9].

First, we validated that the published cpGFP domain insertion location in mLight was the most optimal (**Figure 3.1A-B**). We then enhanced membrane trafficking in the landing pad cell system by adding membrane trafficking and ER export sequences (TS-ER) to the sensor scaffold (**Figure 3.1**)[102]. Next, to increase the allosteric coupling between the two domains, we targeted mutations to four residues within the linkers between MOR and cpGFP (**Figure 3.2**, **Figure 3.1**). We screened >23,000 cells/variants at 1 μM [D-Ala^2, N-MePhe^4, Gly-ol]-enkephalin (DAMGO), a synthetic enkephalin with high selective for the MOR. We recovered variant 2A from the library for testing in HEK293 cell populations and dubbed the variant μMASS^{2A}. μMASS^{2A} had significantly better responses to 500 nM (~4.6 fold) and saturating concentrations of DAMGO (~3.8 fold) than the parent construct mLight (**Figure 3.2C and D**). Similar to the native MOR, the sensor could detect several types of opioid peptides, such as the enkephalins, dynorphins and beta-endorphins with differing apparent affinities (**Figure 3.2E**, DAMGO K_D = 243 nM, Methionine-

enkephalin K_D = 99 nM, leucine-enkephalin K_D = 637 nM, β-endorphin K_D = 1176 nM, Dynorphin A K_D = 1125 nM [100]).

The exogenous opioid agonists morphine and fentanyl (at 1 μM) activated μMASS^{2A} at lower levels compared to Met-Enk (26.4±3.64% and 4.63±1.23% respectively) (**Figure 3.2F**). We demonstrated that the sensor activity is reversible by applying the opioid receptor antagonist naloxone, which could abolish fluorescence signal and be used as an important pharmacological control for *in vivo* experiments (**Figure 3.2G**) [103]. Next, we engineered a loss of function (LF) variant by mutating Asp147$^{3.32}$Gly (See Appendix B for information on Ballesteros-Weinstein Numbering). Asp147$^{3.32}$ is hypothesized to coordinate a key intermolecular bond with the primary amine in the canonical opioid signaling motif [64], [78], [104]. The μMASS^{LF} response signal to 10 μM of various opioid peptides was significantly lower (**Figure 3.2H**) and the μMASS^{LF} had a similar baseline fluorescence compared to the parent μMASS^{2A} (**Figure 3.2I** $p = 0.90$, unpaired t-test, Welch's correction). μMASS^{LF} lacks any membrane bound fluorescence changes when imaged under confocal microscopy (**Figure 3.2J**). μMASS^{2A} did not have any significant ligand dependent internalization in HEK293WT cells when exposed to 10 μM DAMGO and imaged over an hour at 37°C in comparison to a C-terminally tagged rat MOR (**Figure 3.1E-G**).

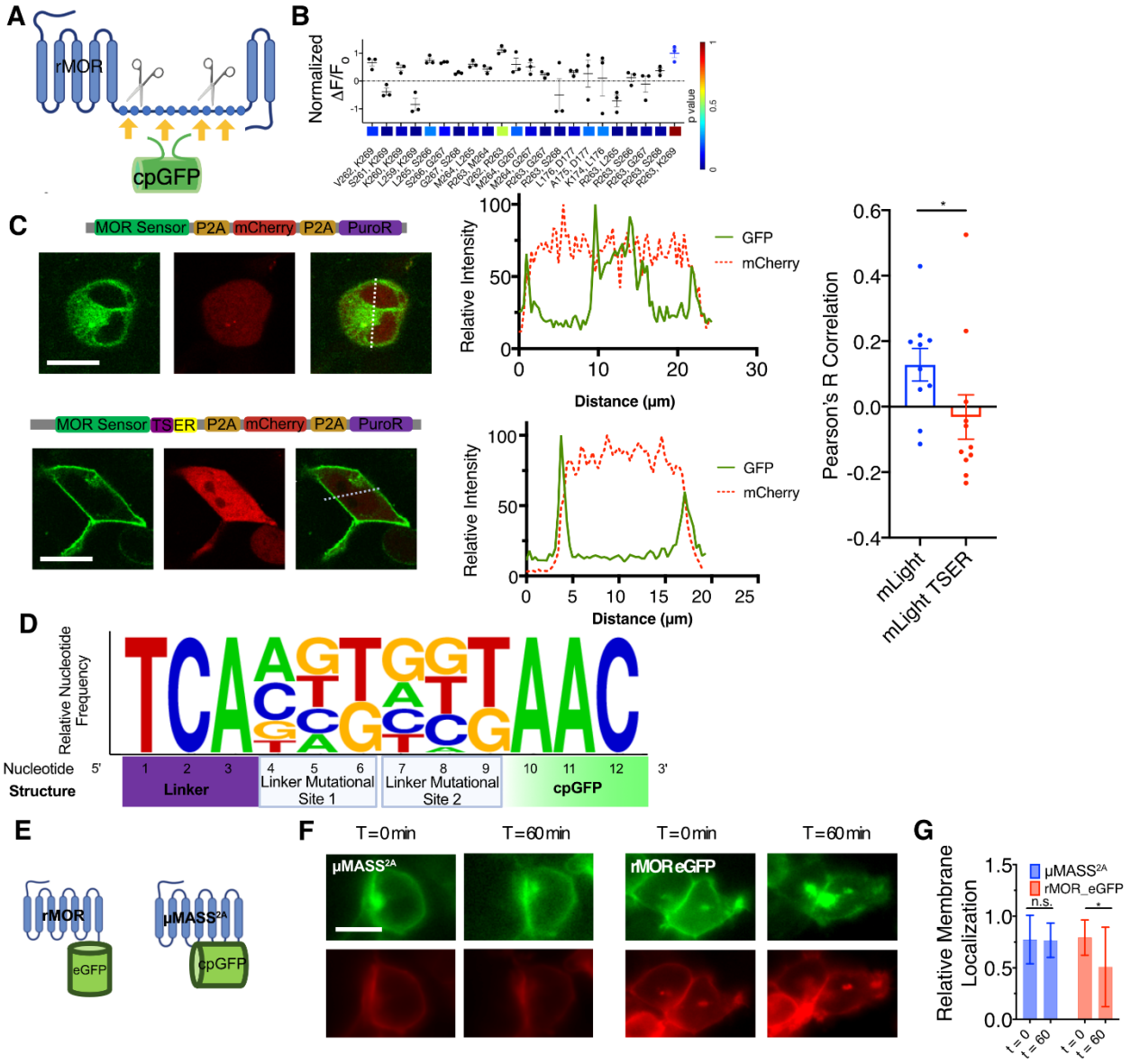

Figure 3.1: Design, engineering, and characterization of μMASS libraries and extended biophysical characterization of improved variants

A) Schematic showing the insertion of the cpGFP reporter domain into the rat mu-opioid receptor. Panel inspired by [50].
B) cpGFP domain insertion screening results. Residues denote the position of the N- and C-terminal MOR residues next to the cpGFP linker. The original insertion of cpGFP in mLight between R263 and K269 yielded the highest signal amplitude. Three wells, ANOVA, multiple comparisons.
C) Membrane trafficking signals (TS-ER) were added to the MOR sensor to reduce cytotoxicity from aggregation in the cytosol in the HEK293T landing pad cell line. Cells were imaged on a 40X confocal microscope and analyzed for low correlation with the cytosolic mCherry fluorophore. Summary data n = 2 wells, 5 cells/well. Mann-Whitney Test, p =0.0357.
D) Relative nucleotide frequency in eighteen selected colonies from the μMASS library generation. Logo creation using [48], [49].
E) rMOR-EGFP and μMASS^{2A} internalization experimental set up.

F) Representative epifluorescent images of HEK293 cells expressing membrane-localized control fluorophores (red, mRuby-CaaX) and either rMOR tagged with eGFP or green fluorescent μMASS^{2A}. Images were taken before adding 10 μM DAMGO and 60 min after. 10 μm scale bar.
G) Summary data for the relative membrane localization of GFP fluorescence before and after one hour of 10 μM DAMGO addition. rMOR internalizes while μMASS^{2A} remains localized in the plasma membrane. N = 2-3 wells, 4-5 cells/well. Paired t-test, n.s. p>0.05, p=0.035 rMOR_eGFP.

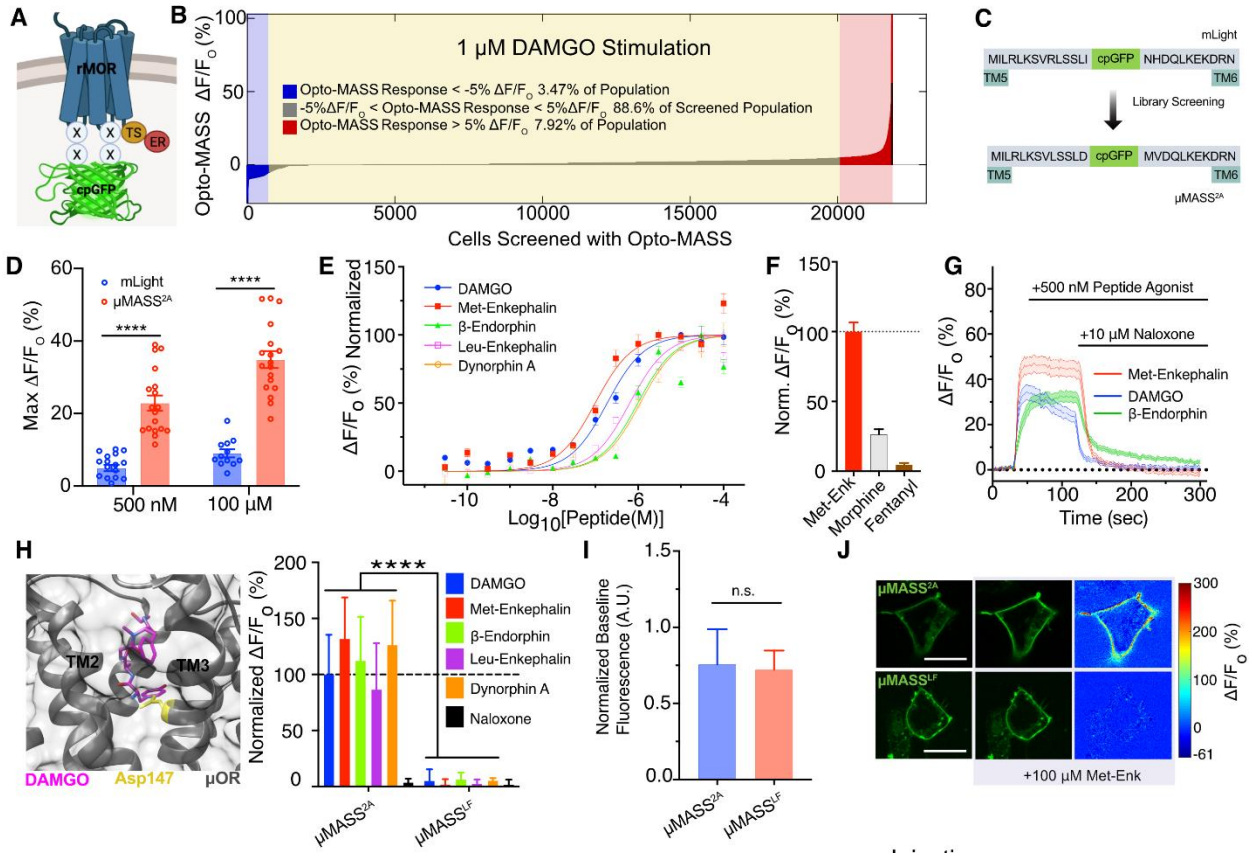

Figure 3.2: Engineering, design and *in vitro* characterization of an opioid biosensor using Opto-MASS

A) cpGFP receptor insertion into rat mu-opioid receptor (MOR) and locations of linker residues targeted by site saturated mutation for generating a MOR sensor library. We added the membrane trafficking signals TS-ER to the c-terminus.
B) Aggregated sensor library response using Opto-MASS. 21,839 cells were screened over two days, resulting in 3.47% with a ligand dependent decrease in fluorescence ($\Delta F/F_o$ < -5%), 88.6% with no change in fluorescence (-5% < $\Delta F/F_o$ < 5%), and 7.92% with an increase in fluorescence ($\Delta F/F_o$ > 5%).
C) Mutational changes in the selected variant μMASS^{2A} compared to the parent scaffold mLight.
D) mLight and μMASS^{2A} response to 500 nM (mLight 5.00±0.87%; μMASS^{2A}: 22.9±2.1%) and 100 μM DAMGO (mLight 9.06±1.11, μMASS^{2A} 34.31±3.11). P <0.0001, unpaired t-test n = 12-18 cells, 2-3 wells.
E) Apparent affinity curves of μMASS^{2A} for common opioid peptides and DAMGO. K_D as follows: DAMGO 243 nM, Methionine-enkephalin 99 nM, leucine-enkephalin 637 nM, β-endorphin 1176 nM, Dynorphin A 1125 nM. n = 18 cells/3wells each.
F) μMASS^{2A}'s normalized response to 1 μM endogenous and exogenous opioids, morphine, and fentanyl. (n = 19-21 cells, three wells).
G) μMASS^{2A} responses are reversible by addition of the opioid receptor antagonist naloxone. N = 27-28 cells/3 wells.

H) Left: The mutated residue for generating a loss of function variant from μMASS^{2A} is highlighted in yellow (PDB ID: 6DDF). Right: Comparison of μMASS^{2A} and μMASS^{LF} (loss-of-function) ligand responses which were normalized to 10 μM DAMGO (Unpaired t-test, **** p<0.0001, n=3 wells, 6-9 cells per well).
I) μMASS^{2A} and μMASS^{LF} do not have significantly different baseline fluorescence. The fluorescence was normalized to a C terminally tagged mRuby3 (p = 0.90 unpaired t-test with Welch's correction, n = 16-27 cells/2-3 wells, mean ± SEM).
J) Representative confocal images of μMASS^{2A} and μMASS^{LF} expressed in HEK293 cells and their responses to 100 μM met-enkephalin.

3.2.2 *In vivo opioid recording with μMASS^{2A}*

Having developed an optimized GEFI for MOR, we next validated μMASS^{2A} *in vivo* in a brain site that is highly enriched in MORs, the nucleus accumbens (NAc). MORs in the NAc have long been shown to dramatically modulate motivated behaviors, including food consumption, social bonding, and drug seeking[96][105]–[107]. Although an important site for opioid reward, the temporal characteristics of opioid signaling *in vivo* have largely remained a mystery due to the inability to track it over subsecond timescales[90], [108]. Therefore, to determine whether μMASS^{2A} could be used to monitor ongoing MOR activity, we used a combination of fiber photometry and pharmacological agonism (**Figure 3.3K, M**). First, mice were habituated to a chamber in which they could freely move around. During this period, μMASS^{2A} fluorescence was recorded to establish a relative baseline. μMASS^{2A} showed strong expression in nucleus accumbens (**Figure 3.3**). After 10 minutes, mice were injected with either saline or 1mg or 10mg of morphine and allowed to continue exploring the chamber. Behavioral and fluorescent activity was recorded for a total duration of 2 hours. Overall, injections of 1mg and 10mg of morphine produced an increase in μMASS^{2A} fluorescence (**Figure 3.3**) relative to saline test days ($F_{(8,32}$= 1.7, P= 0.131). During the first 10 minutes of the test day, μMASS^{2A} activity was similar across drug conditions (t_{1mg}= 0.998, p_{1mg}= 0.06, t_{10mg}= 0.904, p_{10mg}= 0.40). The lower dose produced a more rapid response within 60 minutes of morphine injection (t_{1mg}= 3.07, p_{1mg}= 0.01) which persisted until the end of the two-hour period (t_{1mg}= 3.78, p_{1mg}= 0.001), although both doses were similar in magnitude by 60 minutes (t_{1mg}= 3.07, p_{1mg}= 0.01, t_{10mg}= 2.96, p_{10mg}= 0.01). Although 1mg and 10mg doses of morphine produce different maximal fluorescent responses, it should be noted that morphine only activates in μMASS^{2A} at ≈25% of the capacity that met-enkephalin can produce.

However, despite this ceiling, it should be appreciated that even the low dose of morphine was detectible by the biosensor, highlighting its sensitivity *in vivo*.

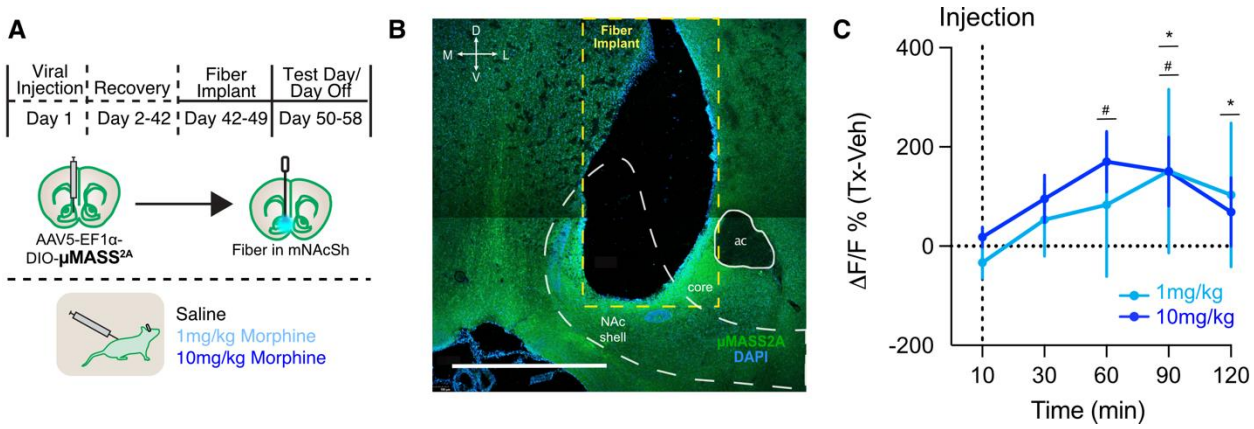

Figure 3.3: *In vivo* recording of opioid administration with μMASS^{2A}

A) Schematic of in vivo pharmacology μMASS^{2A} experiments. μMASS^{2A} was expressed in nucleus accumbens of Penk-cre mice. Fiber photometry recordings were obtained after injecting mice with saline, 1mg/kg morphine, or 10mg/kg morphine.

B) Histological image of μMASS^{2A} expression and fiber tip placement in the nucleus accumbens, 1 mm scale.

C) Mean fluorescence (circles) and SEM (bars) at 10, 30, 60, 90, and 120 minute time points after systemic injections of saline (black), 1mg/kg morphine (light blue) or 10mg/kg morphine (dark blue). * indicates statistical significance between 1mg morphine and saline. # indicates statistical significance between 10mg morphine and saline.

3.3 Discussion and Future Directions

Here, we used a high-throughput protein engineering pipeline to identify an *in vivo* capable opioid biosensor, μMASS^{2A}, that has a significantly better response to 500 nM (~4.6 fold) DAMGO than the parent construct and can detect morphine administration *in vivo*. μMASS^{2A} was selected after screening >23,000 HEK293T cells expressing sensor variants with Opto-MASS. The Opto-MASS platform has previously demonstrated it could improve *in vivo* monoamine sensors in mammalian cells (Chapter 2). Here, we used Opto-MASS to make an opioid sensor. Endogenous neuropeptides signal through a notably different signaling profile. Neuropeptides are thought to signal with long range, low concentration releases and classic neurotransmitters are thought to have large, high concentration releases that are spatially constrained at neuron synapses [109]–[111] . With the construction of μMASS^{2A}, we demonstrated Opto-MASS can optimize sensors to detect ligands that signal through a distinctly different signaling mechanism than classic neurotransmitters.

We screened our μMASS library with the peptide mimetic, DAMGO, which may have biased μMASS^{2A} binding affinity for enkephalins over endorphins. The binding affinity profile of endogenous opioid peptides for μMASS^{2A} does not bracket with the endogenous ligand affinity profile of the MOR. We anticipated the ligand affinity profile for μMASS^{2A} to bracket the MOR ligand affinity profile, similar to previously published sensors that bracket their receptors ligand affinity profiles [9], [33], [67]. Additionally, μMASS^{2A} responds poorly to exogenous opioids, which have a different overall molecular structure than endogenous opioid peptides. The change in μMASS^{2A}'s ligand affinity profile may have arisen due to screening the library with DAMGO, a biased ligand. We chose DAMGO due to its chemical stability. There is *in vitro* assay evidence DAMGO biases the MOR similarly to Met-enkephalin, but differently β-endorphin [100], [112], [113]. While β-endorphin contains the structure of Met-enkephalin within its peptide sequence, evidence suggests each peptide has a unique role in opioid signaling in mammals. The differences in biasing of cell signaling pathways may explain μMASS^{2A}'s lower affinity for β-endorphin. To bias the receptors, ligands must stabilize the receptor in different conformations upon binding[114], [115]. Screening the library with a biased ligand may lead to the selection of a sensor

that preferentially detects ligands with a similar bias. As seen with the μMASS^{2A} sensor, screening with biased ligands may lead to selecting a variant with an altered ligand affinity profile.

μMASS^{2A} has rapid on and off kinetics *in vitro,* demonstrating once its dynamic range is large enough to robustly detect endogenous opioid release, the temporal fidelity of the signal will be high. μMASS^{2A} is one of the first, if not the first, MOR-based sensors shown to be functional *in vivo.* It sensitively binds and detects μ-opioids and does not have ligand dependent internalization in HEK293T cells. μMASS^{2A} can detect dose dependent morphine administration *in vivo*. However, it has not demonstrated robust detection of endogenous opioid release *in vivo.* Despite the limitations to detecting exogenous opioids and higher molecular weight opioid peptides, such as dynorphins and endorphins, μMASS^{2A} presents a significant innovation to the opioid biology field. In the future, screening of the library should include screening with the ligand the sensor is desired to detect *in vivo.*

So far, we tested ≈24,000 mu-opioid sensor variants during pipeline optimization and library screening, which is less than 2% of the 20^4 possible mutations in each case. However, this represents an order of magnitude higher throughput compared to previous studies. Importantly, this number was sufficient to identify already significantly improved variants. The biased agonism and MOR pleiotropy demonstrates the need to develop Opto-MASS to screen for ligand selectivity. While μMASS^{2A} currently cannot detect endogenous opioid peptide release *in vivo*, it presents a significant step forward to detecting the pharmacokinetics of morphinan opioid compounds *in vivo*.

3.4 METHODS

3.4.1 *Library Construction, Mammalian Cell Culture and HEK293T Imaging*

HEK293 cells were cultured on tissue culture treated plastic at 37° C with a 5% CO2 atmosphere. One day prior to transfection, cells were lifted off the growth substrate with 0.05% Trypsin/EDTA. The cells were then seeded into 24 well tissue culture plates. Cells were grown to 70-80% and then transfected. During transfection, growth media was replaced with fresh 250 µL of media. The DNA transfection reagents were prepared using the standard protocol. In brief, per well of transfection, 25 µL Opti-MEM, 1 µg of DNA and 1.5 µL of P3000 were mixed. After five minutes of equilibration, the Opti-MEM/DNA/P3000 mix were added to a tube containing 25 µL Opti-MEM. And 1.5 µL of Lipofectamine. The DNA/P3000/Lipofectamine were incubated for approximately 15 minutes at room temperature before addition to the wells. After incubation for 3-4 hours, the transfection media was removed, and fresh media was added on. Reactions were scaled for different wells according to the manufacturer's directions.

3.4.2 *cpGFP Domain insertion validation in rMOR*

cpGFP was subcloned into various insertion sites in a pcDNA3.1 vector expressing rat MOR. Cloning was performed using primers from IDT and HiFi DNA Assembly (New England Biolabs). Domain insertions were transfected into HEK293 cells and fluorescence response to 10 µM DAMGO was imaged with a 20X air objective on a Leica DMI8 microscope. Fluorescence response was measured by ROI in FIJI. Data analysis done in MATLAB R2019b.

3.4.3 *µMASS Library Construction, Trafficking and Screening in HEK 293T TetBxb1BFP cells*

Prior to µMASS library construction, we increased cell surface membrane trafficking by cloning IgK leader chain to the N terminus of the rMOR receptor, and $KIR_{2.1}$ trafficking and ER export signals to the C terminus. Cloning was performed with insertional mutagenesis with long primers using NEB Site Directed Mutagenesis kit. The IgK leader chain was removed from

recovered variants prior to *in vitro* characterization. HEK 293T TetBxb1BFP cells stably expressing sensor variants (8 μg/mL doxycycline) were plated on poly-l-lysine coated glass coverslips and imaged on a Nikon A1R scanning confocal using a 40X oil objective. Intensity profiles of red and green fluorescence was measured using FIJI (NIH). Intensity profile correlation coefficients were calculated by normalizing red and green fluorescence intensity profiles and calculating the Pearson's R-correlation coefficient. Normalization and Pearson's Correlation Coefficient were calculated in Excel.

μMASS library construction was performed as detailed in 2.4.2 and **Figure 2.6**. μMASS library expression is particularly doxycycline sensitive, as the rMOR gene is slightly toxic to HEK293T cells. Library screening was performed with the same imaging conditions as the dMASS library, with the notable exception of ligands being hand added to the bath due to the dead volume with a syringe pump system and reagent expenses.

3.4.4 *Peptide Affinity Curves*

Peptides were stored desiccated at 4°C in the dark until reconstitution. Peptides were brought to room temperature before being reconstituted at a concentration of 500 μM with distilled water. DAMGO was reconstituted to a concentration of 1 millimolar. B-endorphin was reconstituted with 10% DMSO added to the water. Peptides were vortexed and centrifuged and allowed to solubilize for at least one hour prior to use. Aliquots were stored in the dark at -20°C. For peptide affinity curves imaging, landing pad cells expressing the μMASS^{2A} sensor were imaged under epifluorescent microscopy with a 63X air objective. Cells were seeded at 25,000 cells per well in a 96 well dish and imaged the next day after overnight growth in 8 μg/mL doxycycline. A 63X air objective was used to image the cells continuously for sixty seconds. 150 μL of peptide solution was hand pipetted into a bath of 50 μL bath.

Cells were analyzed in FIJI (NIH). Five to six cells from each well were hand circled from background subtracted image stacks (Rolling ball, 100 pixels). The average fluorescence intensity was measured for each cell and exported to Excel for analysis. Analyzed data was imported into GraphPad Prism 8 to calculate EC50 values using the nonlinear fit function and Least Squares fit.

3.4.5 *In Vitro internalization ssay*

12 mm glass coverslips were coated with poly-l-lysine. After incubation in poly-l-lysine for approximately one to two hours at room temperature, or overnight at 4°C, and rinsed three times with 1X PBS. HEK293 cells were seeded onto the coverslips and grown to 70-80% confluency. The HEK293 cells were transfected using the Lipofectamine 3000 reagents, with an increased total amount of DNA (1500 ng per well and a 1:1 molar split between the control fluorophore plasmids and sensor expression plasmids. See Section 2.4.6 for detailed transfection instructions). The first contains the pCMV mRuby-CaaX and the other contains a pCMV rMOR eGFP or the pCMV μMASS^{2A} plasmid.

Forty-eight hours after transfection, we imaged the coverslips on an epifluorescence microscope with a 40X oil objective at 37°C. Cells were imaged for one hour, during which they were sampled 60 times. An mCherry image and an eGFP image were taken at each timepoint. Images were analyzed in FIJI to collect intensity profiles, analyzed in Excel, and plotted in GraphPad Prism (Ver 8.4.2).

3.4.6 In vivo *fiber photometry μMASS2^{A}*

For μMASS^{2A} experiments, Penk-cre mice (9 weeks old, 2m and 3f) were injected with 200nL (Hamilton, 100nL/minute) of AAV5-EF1a-DIO- μMASS^{2A} into nucleus accumbens (AP: 1.7; ML: +/-1.0; DV: -4.4) and implanted with a 5mm optic fiber (Doric: 400μm, 0.48 NA, 5mm brass ferrule) 0.1mm above each injection target. Fibers were fixed to the skull with Metabond.

Following 6 weeks of recovery, mice were habituated to the test chamber (25cm x 25cm x 25cm) and allowed to explore for 30 minutes. Near the end of the habituation day, mice were systemically injected with saline (i.p.) to habituate them to the injection procedure. 48 hours later, mice were again placed into the test chamber and allowed to explore for 10 minutes. After 10 minutes, mice were injected with saline or 1mg/kg or 10mg/kg morphine and returned to the test chamber for the remaining duration of the two-hour test. Behavioral videos and photometry recordings were collected for the entire two-hour test.

At the conclusion of the experiment, mice were transcranial perfused with 20mL of PBS and 20 mL of 4% PFA. Skulls were removed and post fixed for 24h before the brain was removed and post fixed for an additional 24h. Brains were frozen at -20°C and then sectioned at 40μm on a

cryostat (Leica). Sections were collected in PBS then mounted onto glass slides and cover slipped with fluoroshield with DAPI (Sigma). Sections were imaged at 5x magnification under an epifluorescence microscope (Zeiss ApoTome2) using Zen (Blue Edition, Zeiss) or at 10x magnification under a confocal microscope (Olympus Fluoview FV3000) using FV31S-SW. The location of optic fibers was determined by mapping fiber position onto a mouse histological atlas (The Mouse Brain, Paxinos 2001).

For μMASS^{2A} experiments, we recorded μMASS^{2A} fluorescence in the NAc by connecting an optic fiber to the implanted fiber using a ferrule sleeve (Doric, catalog no. ZR_2.5). Two light-emitting diodes (LEDs) were used to excite μMASS^{2A}. A 531-Hz sinusoidal LED light (Thorlabs, LED light, catalog no. M470F3; LED driver, catalog no. DC4104) was bandpass filtered (470 ± 20 nm, Doric, catalog no. FMC4) to excite μMASS^{2A} and evoke μ opioid-dependent emission. Laser intensity for the 470-nm wavelength band was measured at the tip of the optic fiber and adjusted to 50 μW before each day of recording. μMASS^{2A} fluorescence traveled through the same optic fiber before being bandpass filtered (525 ± 25 nm, Doric, catalog no. FMC4), transduced by a femtowatt silicon photoreceiver (Newport, catalog no. 2151) and recorded by a real-time processor (TDT, catalog no. RZ5P). The timing of injection was recorded using behavioral video recordings. The envelopes of the 531-Hz signal were extracted in real time by the TDT program Synapse at a sampling rate of 1,017.25 Hz.

For μMASS^{2A} experiments, fiber photometry signals were post processed using custom MATLAB scripts. We corrected for photobleaching for each session by fitting our 470 fluorescent signals to a 4-term polynomial function. We normalized the fluorescent signal by subtracting the vehicle treated fluorescent signal from the treatment group signal from. Next, we used behavioral video recordings to extract the time of injection and baseline corrected by subtracting the mean signal during the 9.5 minutes before injection. To calculate the fluorescence at the desired timepoints, we averaged a range of the raw, baseline corrected, decay adjusted 470 fluorescence from +/- 1 minute from the desired timepoint (ex. average fluorescence from 9 to 11 minutes for 10-minute timepoint) for each animal and treatment group.

Chapter 4. EXPANDING OPTO-MASS TO SCREENING FOR LIGAND SELECTIVE SENSORS

ABSTRACT

G-protein Coupled Receptor derived biosensors have revolutionized the field of optogenetics, such as the 'Light', 'GRAB', and MASS families of sensors. GPCR-based biosensors can be used to detect a significant space of neurochemical signaling in the brain because GPCRs bind a diverse and large group of molecules. Despite their immense impact and utility, the sensor's ligand binding profile is constrained to the GPCRs native affinity profile. This presents a significant problem for detecting ligands that bind to promiscuous GPCRs, such as the opioid peptides. The sensitive, GPCR-based sensor μMASS^{2A} binds Met-enkephalin with high affinity ($EC50_{met}$ = 99 nM) and β-endorphin with low affinity ($EC50_{\beta\text{-endorphin}}$ = 1176 nM). Research suggests that Met-enkephalin and β-endorphin have distinct physiological roles. Researchers need β-endorphin selective sensors to investigate the physiological role of β-endorphin in freely behaving animals. If μMASS^{2A}'s affinity for β-endorphin and Met-enkephalin were swapped, β-endorphin signals could be measured in isolation *in vivo*. However, ligand binding selectivity is bestowed by multiple amino acid locations in the ligand binding pocket, necessitating the screening of large, complex mutational libraries. However, traditional techniques to screen for ligand selectivity are prohibitively laborious and resource intensive. Here, we expand the Opto-MASS platform to the high-throughput screening of ligand selective sensors to screen thousands of proteins along multiple biophysical characteristics in mammalian cells. To showcase the platform and address the need for the β-endorphin selective biosensor, we screened a library targeting N127$^{2.63}$, K303$^{6.58}$, W318$^{7.35}$, and E229$^{5.35}$ in the rat mu opioid receptor. We measured approximately 13,000 mammalian cells along two performance axes for β-endorphin selectivity over structurally similar Met-enkephalin and screened for pharmacological reversibility in the presence of a nonselective opioid competitive antagonist. These efforts highlight the expanded applicability of Opto-MASS and its ability to quantify two biophysical characteristics and screen a third. The rapid, multiparameter, multiplex quantification of biosensors with Opto-MASS pushes the boundaries of GEFI engineering.

4.1 INTRODUCTION: THE NEED FOR LIGAND SELECTIVITY

Ideal biosensors selectively and sensitively detect ligands with fine spatio-temporal precision. The Opto-MASS engineered opioid peptide sensor, μMASS^{2A}, has varying affinities for endogenous opioids. The sensor μMASS^{2A} uses the rat μ opioid receptor (rMOR) as the ligand binding domain. The endogenous rMOR binds a variety of opioid peptides [100]. Opioid peptides are released from dense core vesicles and diffuse over large distances in the mammalian brain, interacting with many neurons. The long-distance diffusion and signaling of opioid peptides complicates *in vivo* opioid receptor-ligand studies and their role in physiological behavior, as distal release sites cannot be easily mapped with retro- or anterograde transynaptic viruses like direct synaptic junctions[116], [117]. Genetic knockout and animal behavioral models, receptor structure studies, and gene expression patterns suggest that each opioid peptide and receptor have different physiological roles. However, experiments using genetic knockout models or microdialysis cannot make precise measurements required for neural circuit mapping. Genetically encoded fluorescent indicators can link opioid peptide release and animal behavior in neural circuits. The endogenous ligand affinity profile of the opioid receptors and tools derived from them, such as μMASS^{2A}, necessitates the engineering of ligand selectivity to make precision tools. Receptor-ligand binding interactions are complex, and many mutational variants need to be screened to identify ligand selective sensors. Opto-MASS can identify high performance sensors and their response to one ligand. We broadened the screening capabilities of Opto-MASS to identify high performance, ligand selective sensors through multiple ligand stimulation periods. The high performance, ligand selective sensors will benefit the study of neuropeptide signaling, circuit neuroscience and biology.

4.1.1 *Opioid Peptide Signaling System: Diverse Ligands and Receptor Interactions*

There are four different opioid receptors, mu, kappa, delta, and Opioid Receptor-Like 1(ORL1). ORL1 was classified as an opioid receptor after molecular cloning confirmed high sequence homology with the other three opioid receptors. Interestingly, ORL1's cognate ligand, nociceptin, despite being similar in structure to dynorphin A, does not interact with a high affinity to the three 'classic' opioid receptors. The endogenous opioid peptides and morphinan compounds do not interact with ORL1[118] . Due to this lack of pharmacological crosstalk, ORL1 and its cognate ligands will be omitted from this discussion. Understanding the structural differences of

opioid receptors and their resulting interactions with ligands can guide residue identification for ligand selectivity engineering.

The remaining three opioid receptors are believed to have distinct roles in physiology because are encoded by three different genes; *OPRM1* for μ (MOR), *OPRD1* for δ (DOR), and *OPRK1* for κ (KOR) [86]–[88], [119]–[122]. Scientists discovered three opioid peptide groups arising from three different precursor genes; enkephalins originate from the *proenkephalin* gene (*PENK),* β-endorphin peptides originate from the *proopiomelanocortin* gene (*POMC),* and dynorphin originates from *prodynorphin* (*PDYN)* [123]–[126]. Each peptide product of opioid precursor genes undergoes posttranslational modification by peptidases. Peptides can undergo further modification such as acetylation or amination, further diversifying opioid peptides (**Table 4.2**).

The major peptide products of *PENK* are the heptapeptide, Met-enkephalin and Leu-enkephalin. While longer opioid peptides, such as dynorphin and β-endorphin contain enkephalin sequences, evidence points to longer peptides rarely being cleaved into an enkephalin and the remaining peptide due to the absence of the non-enkephalin remaining peptide in tissue extracts[100].

Table 4.2. Major opioid peptide precursor mRNA and peptide products from mice that bind to opioid receptors. Note: the peptide precursor genes can encode and produce other notable peptide hormones, for example *POMC* produces β-melanocyte-stimulating hormone in human *POMC.*

Opioid Peptide Group	Precursor Gene	Peptides from gene	Peptide Sequence
Enkephalins	Proenkephalin (*PENK)*	6 Met-enkephalin, 1 leucine enkephalin	YGGFM / YGGFL
Endorphins	Proopiomelanocortin (*POMC)*	β-endorphin (1-31, 1-27, and 1-26) all with or without N terminal acetylation	β-endorphin$_{1-31}$: YGGFMTSEKSQTPLVTLFKN AIIKNAYKKGE
Dynorphins	Prodynorphin (*PDYN*)	α-neoendorphin, Dynorphin A$_{1-13/17,}$ Dynorphin A$_{1-8}$	Dynorphin A$_{1-13}$: YGGFLRRIRPKLK
Nociceptin/orphanin FQ	Pronociceptin (*PNOC*)	Nociceptin/Orphanin FQ	FGGFTGARKSARKLANQ

The hypothesis of one opioid peptide family binds to one opioid receptor arose from the discovery of three receptors and three opioid peptide genes. However, evidence suggests this is not true. Each opioid peptide interacts with all the opioid receptors, but with varying affinities. In fact, opioid receptor expression may overlap solely with a peptide with lower affinity for that

receptor, such as KOR and the *PENK* gene in some cases, alluding to different role for each opioid peptide at each opioid receptor subtype[127], [128]. Gene knockout studies of structurally similar ligands, such as the enkephalins and β-endorphin show each peptide having different physiological roles in behavior. For example, *PENK* and *POMC* knockout mice have shown enkephalin to modulate opioid dependent changes in sucrose consumption, but not β-endorphin [96], [129]. Additionally, stress induced ethanol consumption is changed in β-endorphin knockout mice, but not *PENK* knockout mice[130]. These findings suggest distinct physiological roles for opioid peptides despite their structural similarity. For researchers to investigate the distinct physiological role of each peptide, they need fast, peptide selective tools that can be targeted to genetically defined cell types in freely behaving animals.

4.1.2 *Structural Relationships of Opioid Ligands*

There is structural diversity of the opioid ligands that bind the orthosteric pocket of the MOR. Broadly, there are two categories, endogenous opioid peptides, and exogenous small molecule opioids (**Figure 4.1**). The three endogenous ligand classes — endorphins, enkephalins and dynorphins — all contain the amino terminal sequence of Tyr-Gly-Gly-Phe with diverging sequences progressing along the peptide (**Table 4.2**). In the Tyr-Gly-Gly-Phe motif, the tyrosine's hydroxyl group is separated by six carbon-carbon bonds to an amine in the peptide. This hydroxyl, six C-C bonds to amine is a common structural motif in exogenous and endogenous opioid ligands (**Figure 4.1**). The common opioid structural motif is the molecular 'message' in the opioid ligand address-message hypothesis [131], [132]. Nociceptin contains Phe-Gly-Gly-Phe at its amino terminus. The N terminal Phe dramatically reduces its binding affinity for DOR, MOR and KOR, and will be excluded from this discussion.

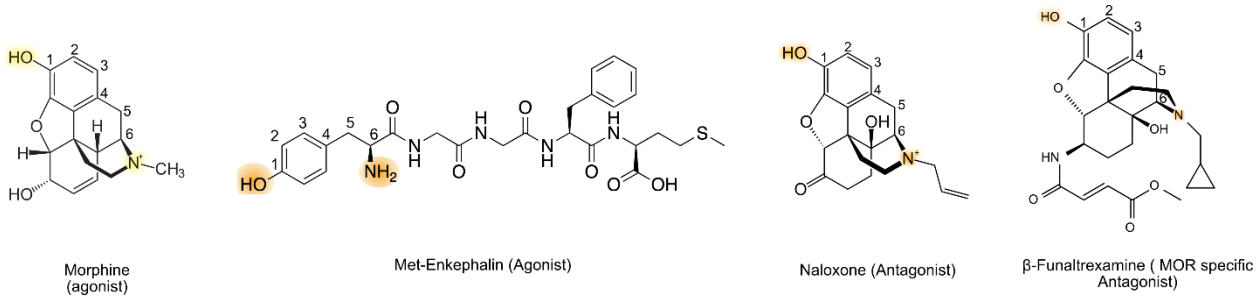

Figure 4.1: Functionally diverse opioid ligands have common structural motifs
Opioid ligands, both agonists and antagonists contain common structural features, notably the hydroxyl to amine, separated by six carbon-carbon bonds, fitting with the opioid ligand address message hypothesis [131].

Exogenous opioids include small molecule alkaloids derived from the plant *Papeaver Somniferum*, such as morphine, codeine, and thebaine. Morphine is a classical MOR agonist and is hypothesized to be one of the first alkaloids isolated from a plant for medicinal purposes [77]. Morphine can be used as a molecular scaffold to synthesize many other opiates. However, there are completely synthetic opiates, such as fentanyl, first synthesized by Janssen Pharmaceuticals in 1960, that do not require a morphinan structure for synthesis. Notably, the μMASS^{2A} sensor responds poorly to exogenous opioids but well to endogenous opioids. We hypothesize this is a consequence of screening the library with the peptide mimetic, DAMGO, and biasing our selection to a short, Met-enkephalin similar peptide.

Generally, all endogenous opioid peptides tested generate signal in the μMASS^{2A} sensor, albeit with different affinities. To adequately discriminate between specific opioid release, genetically encoded tools must not generate large signals to off-target peptides. To engineer ligand selective sensors, we will build mutational libraries targeting regions of the MOR known to mediate ligand selectivity. After expressing the library of sensors in HEK293T TetBxb1BFP cells, we will screen our libraries with an altered ligand addition protocol. The altered ligand addition protocol will enable multiple biophysical measurements of a variant's ligand selectivity. After gene recovery we will validate the sensors biophysical characteristics *in vitro* prior to advancing to *in vivo* experiments.

4.1.3 *Structural Determinants of Opioid Receptor Ligand Selectivity*

There is considerable economic, societal, and geopolitical interest in understanding opioid receptor-ligand interactions for the construction of nonaddictive, pain-relieving drugs. These interests, along with other scientific interests in opioid signaling in animal behavior, have motivated the scientific field to generate high resolution, experimentally determined structures of the four opioid receptors with bound agonists and antagonists[64], [78], [103], [133], [134]. Understanding the differences in ligand-residue interaction from experimentally determined structures can inform residue identification to engineer ligand selective, opioid receptor derived biosensors. The nociceptin/orphanin FQ receptor binds nociceptin, which is structurally different than the other opioid peptides and does not interact significantly with DOR, KOR or MOR, so the NOR will be omitted from most of the discussion.

The family of opioid receptors share remarkable similarity among the three canonical opioid receptors, mu, delta, and kappa, with >70% sequence similarity in the TM regions [78]. The canonical selectivity of the receptors to endogenous ligands is as follows; MOR to β-endorphin, KOR to dynorphins, and DOR to the enkephalins. However, all the endogenous peptides can signal across the opioid receptor family, as all the canonical opioid peptides contain the Tyr-Gly-Gly-Phe peptide sequence. The one receptor, one peptide hypothesis is an oversimplification of opioid peptide signaling. Due to the high structural homology of the opioid receptor TM and opioid peptides, it comes as no surprise the ligand binding pocket is highly homologous between DOR, KOR and MOR structures. The opioid 'molecular message' motif binds with high affinity to this region (**Figure 4.1**) [131].

The ligand-receptor subtype selectivity comes from other structural elements in the ligands and a selectivity filter near the extracellular side of the receptor [135]. Notably, as you move along the opioid receptors from the intracellular space to the extracellular space, the diversity of the OR ligand binding pockets increases (**Figure 4.2** Ref: [135]). The opioid receptor diversity spatially mirrors the increase in opioid peptide diversity; as you move further along the primary sequence of the ligand, away from the highly conserved structures that bind deep in the receptor binding pocket (opioid 'message'), the ligand diversity increases (opioid 'address'). Structural studies revealed agreement in the binding of opioid ligands 'message' deep in the ligand binding pocket, and opioid 'addresses' interacting with receptor residues closer to the surface [64]. The organization of structural diversity increasing from intracellular to extracellular space in GPCRs has been observed in the literature. This structural diversity gradient matches with the role of GPCRs in biology. GPCRs mediate diverse signals from different extracellular environments to more evolutionarily conserved signaling mechanisms internally (e.g. lipid signaling, gene regulation, Ca^{2+}, cAMP) [136], [137].

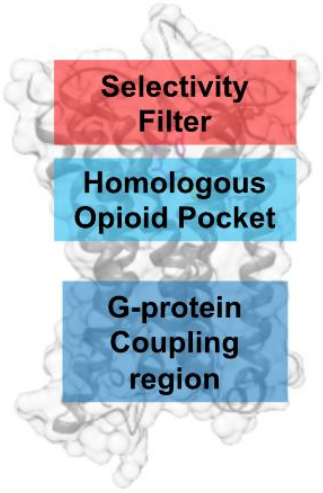

Figure 4.2: Structural determinants of opioid receptor subtype selectivity
Receptor level summary of structural organization of opioid receptors. PDB ID: 6DDF

4.1.4 *Summary: The Need for Ligand Selectivity*

The opioid receptor-ligand system exemplifies the need for ligand selective biosensors. The genetic distribution of opioid receptors and the opioid peptides suggests unique biological roles for the receptors and the endogenous opioid ligands. Behavioral assays with genetic knockout mice support the hypothesis of complex ligand receptor signaling, and unique ligand behavior. The structural motifs of the four opioid receptors, stratification of receptor and peptide expression in the brain and complex post-translational peptide processing creates a swath of ligand-receptor interactions. Current tools, like genetic knockout assays, do not have the resolution to disentangle peptide-receptor interactions in neural circuits. Current measurement techniques such as microdialysis lack the spatio-temporal resolution and cell-type specificity required to link animal behavior and ligand interactions in neural circuits. Existing genetically encoded tools either are incompatible with *in vivo* imaging technologies or do not have the spatio-temporal resolution to link opioids to animal behavior. To disentangle ligand-receptor interactions and animal behavior, fast, cell-type specific tools with ligand selectivity need to be developed. Genetically encoded tools such as μMASS^{2A} provide an easy solution if ligand selectivity can be engineered into the biosensor.

Other receptor ligand systems with some pharmacological cross talk, such as the dopamine receptors, would also benefit from ligand selectivity. High concentration releases of norepinephrine will bind dopamine sensors, like dLight1.1 or GRAB$_{DA}$, due to the molecular similarity of norepinephrine and dopamine[9], [33] . For example, if dLight1.1 or GRAB$_{DA}$ were expressed in the frontal cortex, high concentration norepinephrine release would generate a similar signal to the low concentration dopamine releases researchers would anticipate to detect in the frontal cortex. Receptor-ligand binding interactions are complex. High-throughput screening of libraries would be needed to engineer ligand selectivity into highly dynamic proteins such as GPCR-based sensors like μMASS^{2A}. Thus, engineering Opto-MASS to screen the ligand selectivity of sensors would accelerate the development of high-performance tools. Multiparameter, multiplexed screening of genetic libraries in mammalian cells would expand Opto-MASS's ability to make high quality, precision optogenetic tools for more exact *in vivo* experiments.

4.2 RESULTS

4.2.1 *Expanding Opto-MASS to multiparameter multiplexed screening*

We expanded the engineering capabilities of Opto-MASS to ligand selective sensors. Non-selective sensors generate a fluorescent signal when they bind an off-target ligand (**Figure 4.3B**). Selective (on-target) sensors need to generate a ligand selective signal, even in the presence of the off-target ligand. Ideally, an on-target sensor still has its fluorescence signal abolished by a competitive antagonist (**Figure 4.3C&D**). To enable the identification of ligand selective sensors, we need a technique that rapidly interrogates a variant's ability to detect a ligand of interest in the presence of an off-target ligand (**Figure 4.3**).

We determined that sequential ligand addition of will enable screening for ligand selectivity. By adding ligands in sequence, we can quantify any off-target response first, and then ensure the sensor responds to on-target ligand in presence of the off-target ligand. To meet this goal, we have four imaging periods during our stimulation paradigm. We measure the biophysical properties of each sensor during the four periods. First, we image a baseline period. Second, we add the off-target ligand to the bath. Third, the ligand of interest is added to the bath and finally a control antagonist is added to bath (**Figure 4.3D**). With these four imaging periods, we can determine if our sensor generates an off-target signal above threshold (periods 1 and 2). We also can measure the response to our on-target ligand in the presence of the off-target ligand (Period 3). Finally, the addition of the control antagonist ensures selected variants have a pharmacological control for *in vivo* experiments (Period 4, **Figure 4.3**).

To demonstrate Opto-MASS's ability to screen sensors for ligand selectivity, we will screen for opioid selective sensors. More than one opioid precursor gene can be expressed in neurons and neuropeptides diffuse over long distances, so opioid sensors may interact with many opioid peptide types originating from different sources. Additionally, once we add the ligand of interest to the bath, the presence of the off-target ligand ensures the on-target ligand can still bind. The final drug added to the bath, the antagonist naloxone, will be added to the bath to ensure that the mutations that endow the ligand selectivity does not remove the possibility of using naloxone as a pharmacological control for *in vivo* experiments.

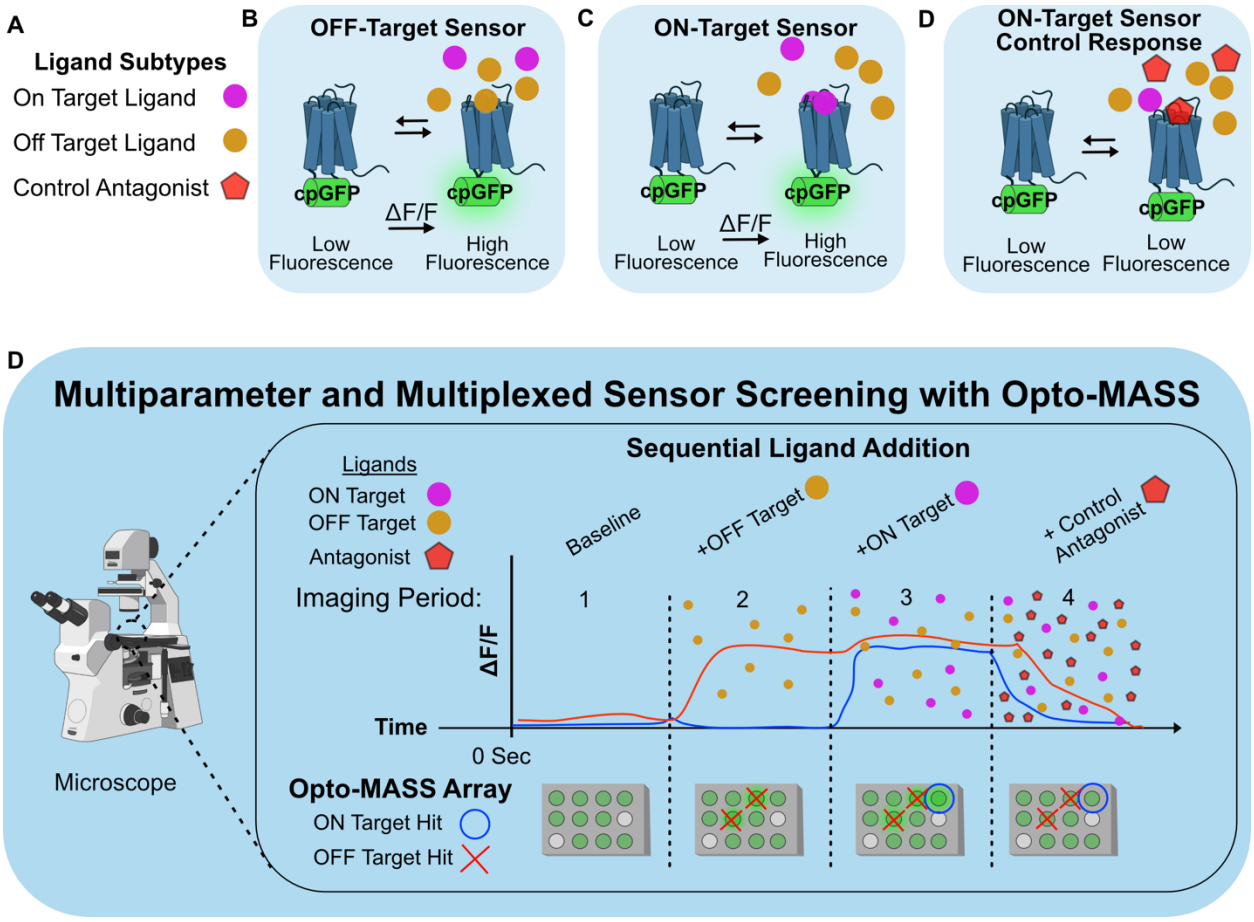

Figure 4.3: Multiparameter and multiplexed sensor screening with Opto-MASS

A) Sensors encounter several types of ligands when expressed *in vivo*. These include off-target and on-target ligands that may bind and generate fluorescent signals.

B) An off-target sensor binds a ligand the sensor should not detect and generates a signal, even in the presence of the on-target ligand.

C) ON-target sensors selectively bind the ligand of interest, even in the presence of the off-target ligand.

D) To identify on-target sensors, we will mutate ligand binding pockets of sensors and express them in mammalian cells. After library expression and cell seeding onto Opto-MASS arrays, we will sequentially add ligands to the bath. First, the off-target ligands are added, to identify sensor variants with off-target responses. Next, the desired ligand is added to the bath, identifying sensors that respond to the on-target ligand. Finally, a pharmacological control, to abolish fluorescence response is added to the bath. This is an important control for *in vivo* experiments to account for motion artifacts.

To validate our new ligand selectivity screening method and image analysis algorithms, and demonstrate we can link phenotype to genotype, we built a control sensor based on μMASS^{2A} with a different affinity profile for opioid peptides. μMASS^{2A} is the rat μ opioid receptor (rMOR) with a cpGFP domain insertion into the third intracellular loop (ICL3, see Chapter 3). μMASS^{2A} responds to DAMGO, a MOR selective opioid peptide mimetic with a high affinity. However, other opioid receptors, such as the DOR, have low affinity for DAMGO. The exclusion of DAMGO is predominantly encoded by residues in extracellular loop 1 (ECL1) in DOR [138]–

[140]. We hypothesized that grafting DOR's ECL1 onto our μMASS^{2A} sensor should decrease the sensor's affinity for the MOR selective ligand **(Figure 4.4A).** This DAMGO insensitive μMASS^{2A} sensor, called μMASS^{2AID} (for Insensitive to DAMGO) is an adequate control for validating new pipeline features because it has an approximately one order of magnitude lower affinity for DAMGO (μMASS^{2A} $EC50_{DAMGO}$ = 296 nM, μMASS^{2AID} $EC50_{DAMGO}$ = 2094 nM **Figure 4.4B)** and in culture has a significantly different response to 300 nM DAMGO compared to the parent sensor (≈ 2.59-fold difference, p<0.0001, Welch's t-test, two tailed) with no difference in response to 1 μM Met-Enk (p =0.1396, Welch's t-test, two tailed, **Figure 4.4C**). Both sensors are reversible, and +10μM of the competitive antagonist naloxone reduced their signal in the presence of 1300 nM peptide agonist (**Figure 4.4C).**

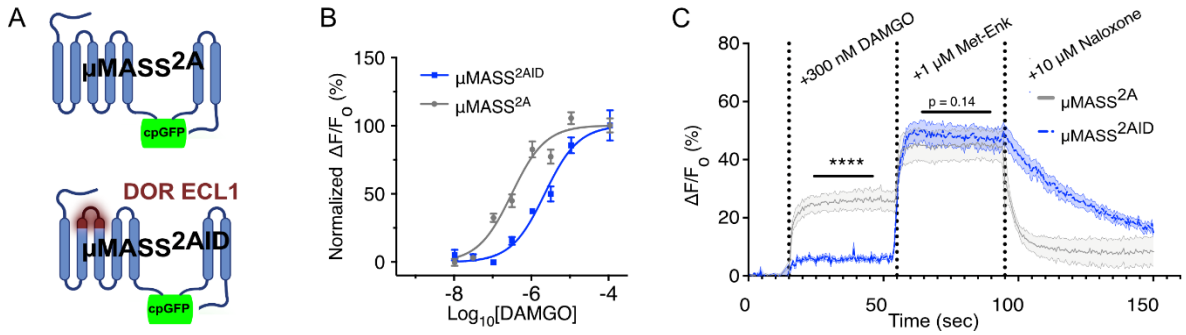

Figure 4.4: Grafting ECL1 onto μMASS^{2A} to make DAMGO insensitive sensor μMASS^{2AID}

A) Schematic of the grafting of extracellular Loop 1 (ECL1) of the rat DOR onto μMASS^{2A} to make the μMASS^{2A} Insensitive to DAMGO sensor (μMASS^{2AID}).

B) Normalized affinity curves to DAMGO in HEK293T TetBxb1BFP cells stably expressing μMASS^{2A} or μMASS^{2AID}. Data was collected on a confocal microscope. μMASS^{2A} $EC50_{DAMGO}$ = 296 nM, μMASS^{2AID} $EC50_{DAMGO}$ = 2094 nM, n = 3 wells, single site binding affinity curve.

C) μMASS^{2A} or μMASS^{2AID} average response to sequential ligand addition of DAMGO, Met-Enk and Naloxone in culture. μMASS^{2A} or μMASS^{2AID} are stably expressed in HEK293T TetBxb1BFP cells. Mean ± SEM plotted. 300 nM DAMGO response, p <0.0001. 1 μM Met-Enk response, p = 0.1396. Welch's t-test. n = 3 wells/23-31 ROIs.

Next, we seeded isogenic populations of each sensor variant onto arrays and screened them to see if the different signals were retained (**Figure 4.5**). We used this stimulation data to validate additional functionality we added to the 'on-the-fly' analysis to screen for ligand selectivity (**Figure 4.5B**). Using the values from three array stimulations of the control sensor, μMASS^{2A}, we can determine a threshold to omit any off-target sensors that respond greater than that threshold. Setting a threshold to exclude off-target sensors is determined by Equation (4. 1),

$$\omega = \mu_{OffTargetLigand} - 1.5\sigma_{off\ target\ ligand} \qquad (4.\ 1)$$

Where ω is the Opto-MASS Array Rank Ratio used to determine a subthreshold response to the ligand, μ is the control sensors mean population response to the off-target ligand, and σ is the standard deviation of the control sensors population response to the off-target ligand. Using Equation (4. 1) we can omit most off-target sensors. We can then use another threshold to identify sensors that respond to the on-target ligand but didn't respond to the off-target stimulation (**Figure 4.5B&C**). In an example of two arrays from this experiment, we can see that this thresholding system removes 93.8% of the off-target sensors and has only one false positive. However, it omits 143 ROIs of the 'on target' population that are responding below the threshold for the on-target ligand, leaving only 37% as a possible identification as an 'on-target' sensor. The average µMASS^{2A} Array Rank Ratio response to 1 µM Met-Enk was calculated to be 1207.16 with a standard deviation of 98.6. If we lower the 'on-target' threshold to one standard deviation below the average control sensor response (µMASS^{2A}) to 1108.4, 76.21% of the 'on-target' population would be identified as an 'on-target' hit. The threshold for determining an 'on-target' hit is dependent on the threshold the experimenter wants to set. Lowering the threshold for identifying an on-target hit does not increase the chance of a false positive for this example data set.

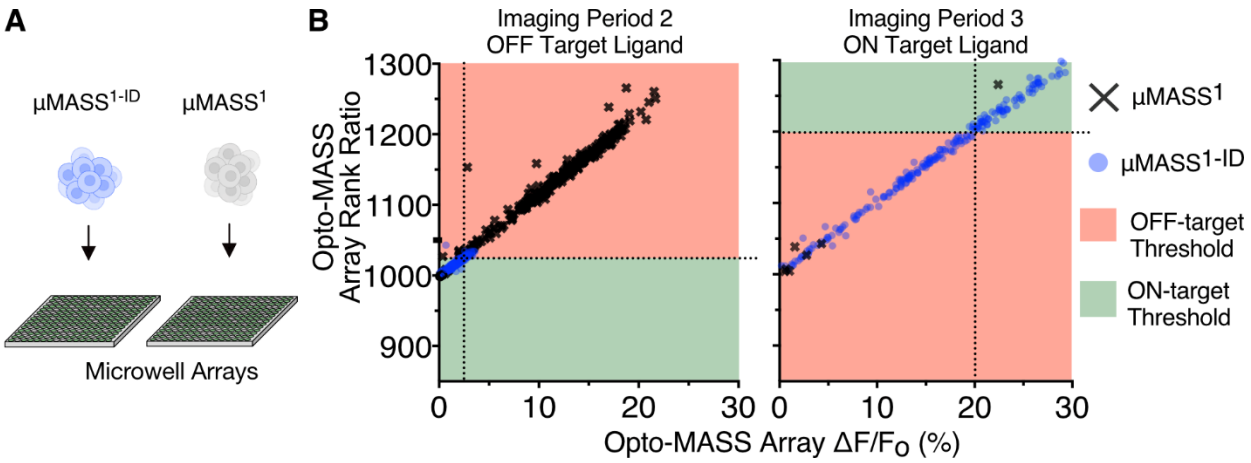

Figure 4.5: Example screening of two isogenic opioid sensor populations validates rapid algorithm

A) HEK293T Bxb1TetBFP recombined with one of two opioid sensors, µMASS^{2AID} or µMASS^{2A}, are seeded onto Opto-MASS arrays for screening.

B) Example: During off-target ligand addition, cells with an Opto-MASS Array Rank Ratio of above 1022 (mean response of control – 1.5 STD) are removed from analysis during on-target addition. 93.8% of the off-target sensor (µMASS^{2A}) are removed from analysis and ranking in the next section (17 remaining). Of the 17 remaining OFF-target cells, only one is above the threshold for 20% for fluorescence ranking. 82.9% of the on-target cell population passes off-target thresholding of 1022, and 67 are above the 1200 threshold for an on-target hit, for a sensitivity of 98.5%. The ON-Target ligand threshold can be lowered to increase the number of available cells.

4.2.2 *Harnessing Structural Determinants of Opioid Receptor Subtype Selectivity to Engineer Opioid Peptide Selective Sensors*

The selectivity of the opioid receptors is bestowed by residues closer to the extracellular space in the ligand binding pocket or by residues in the extracellular loops [141], [142]. Extracellular loops have been demonstrated to comprise 'selectivity filters' to alter the affinity of opioid peptide subclasses from the three different opioid receptors [140]–[142]. For example, the MOR selective ligand DAMGO is excluded from DOR's binding pocket by ECL1[138], [139], and KOR's binding pocket by four specific residues in the ECL3 of KOR[142]. With this in mind, we targeted residues that are divergent among the three canonical opioid receptors, because we hypothesize we can tune μMASS^{2A} for β-endorphin selectivity over Met-enkephalin by targeting these residues.

Structural studies and opioid receptor ECL grafting studies aid in the identification of structural determinants of OR subtype selectivity. To generate a library that biases our sensor to β-endorphin over enkephalins, we targeted residues that are divergent among the opioid receptors, close to bound ligands, and far from the highly homologous opioid binding core. Residues that are divergent between receptor subtypes may mean the residues contribute to ligand selectivity. In the first libraries, we want to target residues that are close enough to interact with bound ligands. We hypothesize residues that are divergent and within 4 Å of the ligand may contribute to ligand selectivity, and mutations may enhance differences between Met-enkephalin and β–endorphin binding. The active state structure of DAMGO-bound μOR-G_i and inactive-bound structure provided insights on identifying these residues [64]. The residues are as follows; N127$^{2.63}$, K303$^{6.58}$, W318$^{7.35}$, E229$^{5.35}$ [78], [135] (**Figure 4.6A**).

A library targeting all four residues was constructed with Gibson Assembly and stably recombined and expressed in HEK293T TetBxb1BFP cells (See Methods 2.4.2 & 2.4.3). The library was screened with a sequential addition of three ligands: 300 nM Met-enkephalin, 1 μM β-endorphin, and 10 μM naloxone to test for β-endorphin selectivity and sensor reversibility. ROIs were analyzed for their response to the first two ligand additions. ROIs that responded below the threshold for 300 nM Met-enkephalin were analyzed for their response to 1 μM β-endorphin. Cells not responding to Met-enkephalin but identified to respond above threshold for β-endorphin would

be recovered. The threshold values to determine β-endorphin selectivity were determined the day of by screening an isogenic population of μMASS^{2A} expressing cells. Three Opto-MASS arrays seeded with μMASS^{2A} cells were stimulated with 300 nM Met-Enk, 1 μM β-endorphin, and 10 μM naloxone. The Met-Enk responses of the three arrays were averaged and used to calculate the first selectivity threshold using Equation (4. 1). The average response to β-endorphin for μMASS^{2A} was 1153.44. We lowered the threshold for β-endorphin reactivity to 1100 to be more generous and increase the chances of identifying an 'on-target' hit. After experimental thresholds were determined, 12,946 library expressing cells were screened.

Zero cells of the 12,946 had the desired response (**Figure 4.6B & Table 4.3**). Twelve cells from the 12,946 cells functionally screened were initially flagged as β-endorphin selective, however analysis of the stimulation videos revealed the positive β-endorphin signal was a cell or debris moving into the ROI during β-endorphin stimulation, after a nonresponse during the Met-Enk stimulation. Approximately 98% of the library had a response below both agonist thresholds. **Table 4.3** summarizes library population distributions into each threshold in comparison to the control μMASS^{2A} population. Due to the lack of cells having the desired biophysical characteristics, no cells were recovered.

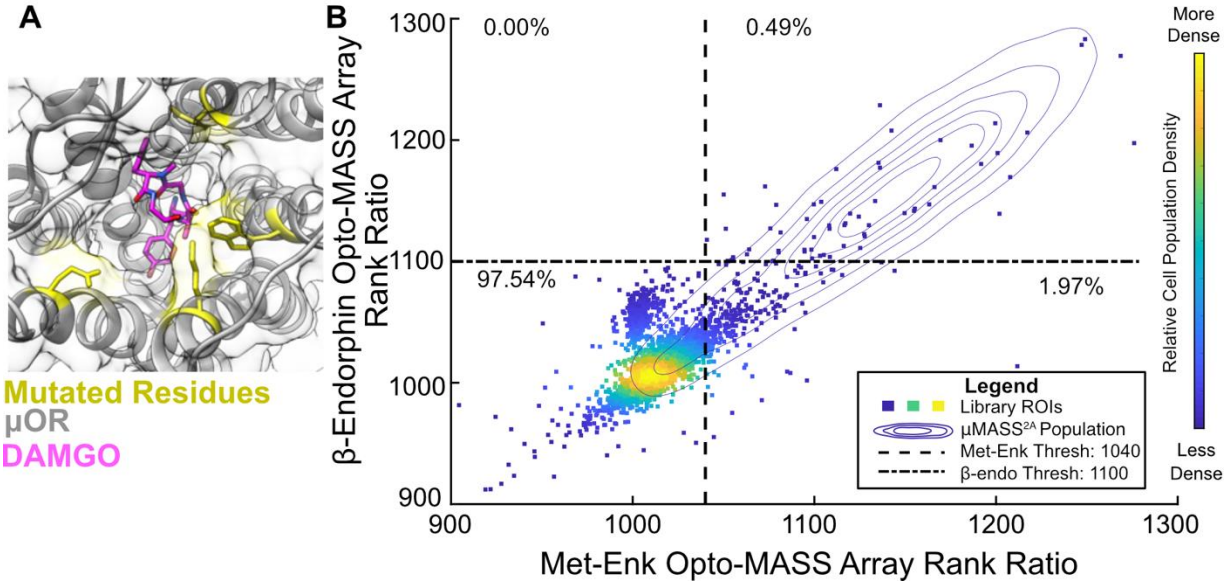

Figure 4.6: Multiparameter, multiplexed Opto-MASS screening reveals rMOR N127$^{2.63}$, K303$^{6.58}$, W318$^{7.35}$, and E229$^{5.35}$ do not bestow β-endorphin selectivity

A) Residues N127$^{2.63}$, K303$^{6.58}$, W318$^{7.35}$, and E229$^{5.35}$ were selected for mutation because residues in those positions are divergent among three opioid receptors. In the inactive crystal structures K303$^{6.58}$, W318$^{7.35}$, and E229$^{5.35}$ are within 4 Å of the bound antagonist β-funaltrexamine in crystalized MOR. The mutated residues for the library are colored yellow. PDB code: 6DDE.

B) B-endorphin selective library screening results with control μMASS^{2A} sensor population overlaid. With the Opto-MASS platform, 12,946 cells expressing the selectivity library were screened, and are plotted as colored squares. Plotted thresholds (dotted lines) were experimentally determined the day of by screening an isogenic μMASS^{2A} population. Prior to screening, three arrays seeded with isogenic populations of were screened to determine thresholds. ROIs that had a response below 1040 when stimulated with 100 nM Met-Enk and above 1100 when stimulated with 1 μM β-endorphin were selected for pick up. No cells met the threshold criteria in the library, 12 false positives from the library were removed due to imaging artifacts (cells or debris moving into wells during stimulation) and are not plotted. Inset percentages are percentage of selectivity library meeting quadrant thresholds.

Table 4.3. Summary of cell populations from library screening. 1121 HEK293T TetBxb1BFP cells stably recombined with μMASS^{2A} were screened on three arrays to establish thresholds. 12,946 HEK293T TetBxb1BFP cells were screened on 39 arrays. Twelve cells met threshold criteria, however all β-endorphin signals above the 1100 threshold were generated by debris moving during ligand stimulation. The dark grey column is the desired threshold response.

Cell Population	> **1040 Met-Enk** > **1100 β-endo** (non-selective)	**> 1040 Met-Enk < 1100 β-endo** (Met-Enk selective)	< **1040 Met-Enk < 1100 β-endo** (no response)	**< 1040 Met-Enk > 1100 β-endo** (β-endo selective)
μMASS^{2A}	75.71%	16.95%	7.23%	0.0089%
μMASS^{2A} β-endorphin Selectivity Library	0.49%	1.97%	97.54%	0.00%

4.3 Discussion and Future Directions

4.3.1 *Library Limitations*

Ligand selective sensors enable detection of different subtypes of opioid peptides. Genetic knockout studies suggest that opioid subtypes have distinct physiological roles, necessitating the development of tools to dissect the action of peptide release in neural circuits. Currently, there are not publicly available GEFIs capable of *in vivo* detection of selective endogenous opioid peptide release. We have demonstrated, with our control sensors μMASS^{2AID} and μMASS^{2A}, that we can set thresholds that rapidly identify the sensors with selective ligand response profiles. In one day, we screened almost 13,000 sensors on our pipeline and measured multiple biophysical characteristics, nearly simultaneously, of all 13,000 sensors.

To engineer ligand selectivity of our μMASS^{2A} sensor, we targeted residues that are divergent among the μ, κ, and δ opioid receptors and are within 4 Å of a bound ligand. Currently, of the approximately 13,000 cells expressing this library, 0.00% met the threshold to be identified

as β-endorphin selective over Met-enkephalin. The residues identified for mutation, $N127^{2.63}$, $K303^{6.58}$, $W318^{7.35}$, and $E229^{5.35}$, were chosen because they are within 4 Å of a bound ligand and are divergent between the MOR, DOR, and KOR, and are not buried in the ligand binding pocket. The chosen residues are within 4 Å of a covalently bound antagonist in an MOR crystal structure. The receptor was stabilized with a T4 lysozyme insertion in the ICL3. There are many limitations to interpreting this structure, as the GPCR is in the inactive state during antagonist binding. GPCRs have several states and fluctuate between the states. The binding of ligands, coupling of G-proteins, phosphorylation by GPCR kinases, all alter the receptor population states and may shift the population to stabilizing in certain conformations. The residues may be closer or further from bound ligands during the active state conformation, so they may play less of a role in agonist binding.

Recently, there was an experimentally solved structure of the active state MOR with G_i coupled and DAMGO bound[64]. However, the structure was experimentally solved with Cryo-EM, which cannot resolve ligands with the many degrees of rotational freedom DAMGO has. To overcome the limitation of Cryo-EM, the authors computationally modeled the position of DAMGO with molecular dynamics. Further Opto-MASS experiments should avoid over interpreting this structure and the interactions between ligand and receptor. The structure should not be the sole source of data to inform the selection of the next mutational library, but it does provide crucial insights about differences in residue location relative to an agonist or antagonist. This may inform which residues to mutate in the next rounds of library generation.

While significantly more cells need to be screened for complete statistical coverage of the libraries mutational space, it may be advantageous from a resource perspective to target other residues. To target other residues, we would need to construct and express a new library. A potential reason there are no β-endorphin selective receptor variants is that we mutated residues that play a significant role in peptide binding, not peptide selectivity. Our mutational selection may have been too close to the highly homologous core of the ligand binding pocket, and we should target residues further from the core that bestow ligand selectivity. We can target residues in the second and third intracellular loop, as these loops have been demonstrated in chimeric and mutational studies to bestow aspects of opioid peptide ligand affinity, and these ECL are close to the ligand binding pocket. Targeting these residues, Thr218 (ECL2), Asp216 (ECL2), E310 (ECL3), and Thr307 (ECL3), may alter the selectivity of the hypothesized ligand filter [138],

[141][142]. Interestingly, another residue, $Q^{2.60}$, is conserved among all the opioid receptor subtypes but only comes within interaction distance of a ligand (4Å) in the structure of DAMGO bound MOR. If mutating the previous residues do not yield any change in selectivity, targeting $Q^{2.60}$, as it may mediate some aspect unique to peptide derived conformational changes. Should mutations to residues in the ECLs fail to bestow ligand selectivity, structural modeling of biased ligands should be considered to identify residues that may be playing a key role in mediating peptide subtype conformational changes [143], [144]. Additionally, several single nucleotide polymorphisms have been identified that alter glycosylation of the MOR and β-endorphin affinity. Targeting this glycosylation may alter the electrostatics surrounding the receptor ligand binding pocket and enhance β-endorphin affinity [145].

We will avoid mutating residues deep in the ligand binding pocket. Residues deeper in the pocket tend to be hydrophobic and mutations may alter protein folding and trafficking to the membrane. These residues are highly conserved between the opioid receptors and are likely critical to defining the high affinity binding pocket for the opioid ligand motifs (**Figure 4.1**). Additionally, these residues may be critical for ligand dependent conformational changes of the receptor. These residues include $D^{3.32}$, $D^{3.47}$, $Y^{3.33}$, $M^{3.36}$, $W^{6.48}$, $Y^{7.43}$ [135], [144]. Further improvements to the library generation process could incorporate mutation schemes that ensure a statistically equal chance of each amino acid in the library. We chose not to implement a library generation scheme like that, for ease of library generation and testing the translation efficiency of different codons. However, statistically rare amino acid permutations may not have been measured in our cell population, and these rare variants may encode for β-endorphin selectivity. Our library generation techniques typically limit our number of *E.coli* transformants to a range of 200,000 to 300,000, dramatically under sampling the sequence space of our libraries. The throughput of our library screening in mammalian cells is approximately 10,000-13,000 cells per day. A smaller sequence space that still samples the chemical identities of the twenty amino acids across all 4 mutated sites, like NDT codons instead of NNK codons, could be implemented due to the throughput screening numbers [146]. It is feasible with Opto-MASS screening throughput to gain 95% statistical coverage of four amino acid locations (62,118 transformants) over five to six days of screening.

4.3.2 *Multiplex and Multi-parameter screening*

GPCR-based genetically encoded fluorescent indicators were a pivoting point for detecting biological signals *in vivo*. GPCRs mediate a significant amount of mammalian cell-to-cell signaling and are highly druggable. GEFIs that provide insight into GPCR signaling, activation, or ligand-receptor interactions *in vivo* provide particularly vital information about neural circuit function. While some GPCRs are highly selective to binding one endogenous ligand, some, like the opioid receptors, are not. Opioid receptors can signal with a variety of endogenous ligands. Each opioid ligand is hypothesized to have a unique physiological role, and researchers need tools to isolate endogenous peptide signaling to investigate those roles. Opioid ligand-receptor interactions are of great interest to the structural biology, pharmacology, and neuroscience fields. A clear understanding of receptor subtype binding of ligands is important to develop GEFIs engineered for ligand selective.

Traditionally, researchers would mutate rationally or computationally selected residues in the ligand binding pocket and screen the GEFIs for ligand selective functionality one-by-one. Researchers would need to perform several ligand addition experiments for each mutant as well. We leveraged the ability of Opto-MASS to screen large number of GEFI variants nearly simultaneously and expanded the screening capabilities to screen sensors for ligand selectivity (**Figure 4.3 & Figure 4.6**). Currently, the platform enables engineers to rapidly identify variants that do not respond to the off-target ligand but do respond to the on-target ligand in the presence of the off-target ligand. Three sequential additions of ligand and the resulting biophysical analysis means each Opto-MASS array screening takes several minutes in comparison to the single minute for stimulation and screening in Chapter 2 and Chapter 3. The reality of the multi-ligand screening protocol is the execution takes longer, limiting the total throughput of the pipeline. Faster delivery of the ligands with microfluidics or automated pumps may reduce the time needed to screen, increasing the Opto-MASS throughput once again.

The nature of multi-ligand addition also generates significantly more noise in fluorescent signal readout. There are more events of fluorescent cells and debris moving through the bath and altering fluorescence measurements in comparison to single ligand stimulation. These noise generating events may be reduced with more precise delivery of ligands, or moving the PDMS arrays out of the tissue culture wells they are seeded and recovered overnight in. The source of the contaminating events is HEK293T cells loosely attached to the glass substrate outside of the

PDMS arrays. Changing the location of the PDMS arrays to clean wells with no off array HEK293T cells will reduce these contaminating artifacts.

Additionally, there is a small population of HEK293T cells that did positively respond to β-endorphin without exceeding the response threshold to Met-enkephalin. The appearance of this population is evident in the post hoc analysis (**Figure 4.6 & Figure 4.7**). The population is sparse but present. This population may contain a high level of contaminating events; however it may contain variants with a slight binding preference to β-endorphin. During post-hoc analysis, the population was revealed to be an imaging artifact (**Figure 4.7**). Despite the fact this particular population, in this particular instance, is the result of a contaminating event does not exclude the possibility that a subthreshold population of interest may happen during the next library screening. During screening, this population cannot be visualized by the researcher using Opto-MASS, as it is below the target threshold. Ideally, the functionality to make a real time plot of screened cells would enable the identification of unique populations not bounded within the original thresholds and should be added to the Opto-MASS pipeline. If this population of interest represented true underlying mutations, these cells could be isolated, validated, and combined to make β-endorphin selective GEFIs, as mutations to GEFI binding pockets typically have a synergistic effect on sensor performance [61].

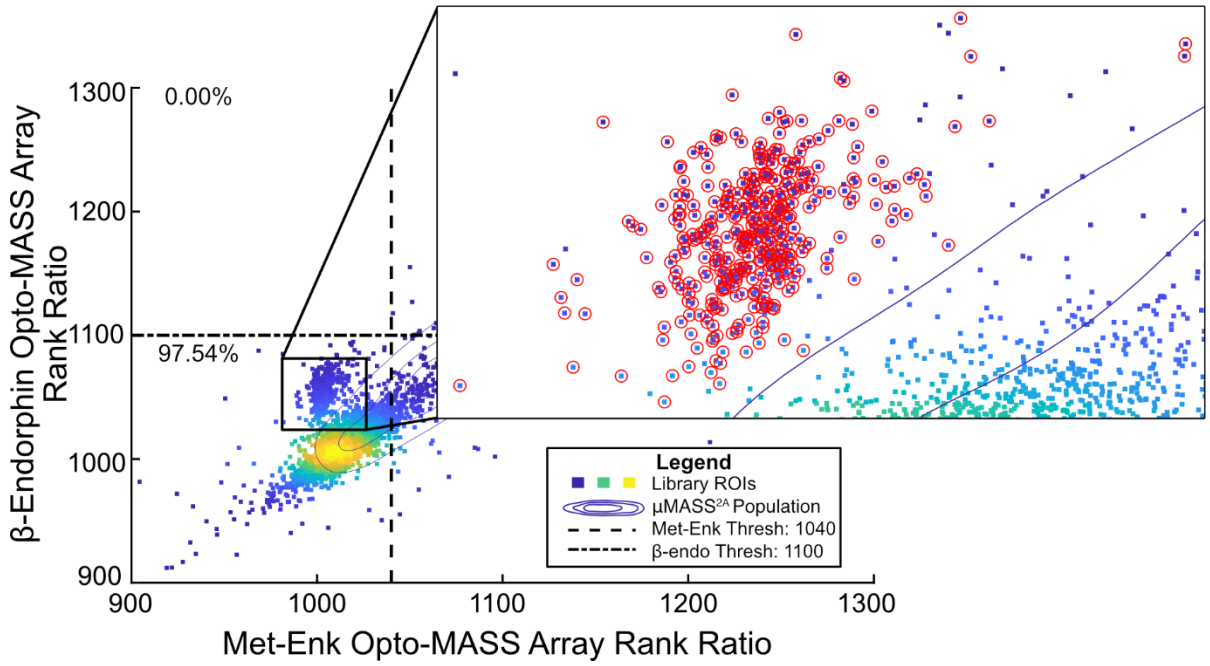

Figure 4.7: Stimulation artifacts generate false populations of interest

During ligand stimulation, the addition of more liquid can disturb cells and cause movement. Occasionally, these movements are above the focal plane, and create fluorescence artifacts in which fluorescence level increases for the whole stimulation imaging period. A population of interest was thought to have arisen, from Opto-MASS Array 9. However, the video has an imaging artifact. All the red circled cells are from the same stimulation with the same artifact. The population of interest is a false positive.

Different experimental parameters may highlight the new screening capabilities of Opto-MASS. GEFI scaffolds with greater dynamic range may be better suited for initial ligand selectivity screens. μMASS^{2A} has low overall dynamic range, and currently is not suitable for extensive use for *in vivo* detection of endogenous peptide release (See Chapter 3.3 *Discussion*). Higher dynamic range sensors have a higher signal to noise ratio, reducing the chance of an ROI selection due to an artifact. Additionally, initial proof of concept screens may want to target ligands that are structurally more divergent, β-endorphin contains the structure of Met-enkephalin within itself, complicating identifying residues that bestow ligand selectivity. Screening between Met-enkephalin and Leu-enkephalin, or endogenous peptides over exogenous morphinan compounds may be better experimental avenues to pursue as the ligands are more structurally diverse. Additionally, computational models of ligand specific interactions could better inform residue selection.

Multi-addition, ligand selectivity screening with Opto-MASS opens a new frontier for the platform. The Opto-MASS platform will be the first platform to screen sensor variants along functional performance axes for multiple ligands in rapid succession. Other pipelines measure

sensor biophysical characteristics that are not the sensor's function, but relevant to sensor performance [46]. Additionally, we chose to recover individual cells with the desired phenotype, leaving valuable information about ligand-receptor binding interactions of non-selected cells unknown. Implementation of DNA barcoding to each variant, and *in situ* sequencing of those barcodes could provide a rich data set to inform models about receptor-ligand binding and sensor function [147]. Despite the technical difficulties presented by the chosen ligands and receptors, Opto-MASS has demonstrated the ability to screen for ligand selective sensors, expanding the boundaries of GEFIs that can be engineered with Opto-MASS. Opto-MASS continues to demonstrate the ability to push the forefront of optogenetics sensor engineering.

4.4 METHODS

4.4.1 *Molecular Cloning and Cell Culture*

Unless otherwise stated, DNA plasmids were constructed using site directed mutagenesis, Gibson Assembly, or standard restriction enzyme cloning. HEK293 were cultured with DMEM (Invitrogen Cat# 10569044) supplemented with 10% Fetal Bovine Serum (Invitrogen cat# 26140079), 1% penicillin/streptomycin (Invitrogen cat# 15140122) at 37° C with a 5% CO_2 atmosphere on tissue culture treated plastic. Cells were passaged with Trypsin/EDTA. HEK293T TetBxb1BFP cells were cultured in media supplemented with 1-2 μg/mL doxycycline instead of pen/strep.

4.4.2 *DAMGO Affinity Curves*

HEK293T TetBxb1BFP clone 4 cells were stably recombined with either a plasmid encoding μMASS^{2A} and a control fluorophore or μMASS^{2AID} with a control fluorophore using the standard transfection and recombination protocol outlined in section 2.4.3. The cells were plated onto poly-L-lysine coated 96 well glass bottom plates, cultured overnight at 37° C with 5% CO_2 atmosphere and supplemented with 4 ug/mL doxycycline. The day of imaging, the cells were rinsed twice with Tyrode's solution and imaged with a 40X oil objective and a 488 nm laser on a Nikon A1R confocal microscope. Images were analyzed for fluorescence change using FIJI (1.53c, National Institutes of Health). Data was normalized to maximum response and plotted using GraphPad Prism (ver 8.4.2).

4.4.3 *Selectivity Library Construction*

Site saturated mutagenesis was introduced to four sites in the receptor domain of μMASS^{2A} using Gibson Assembly. In brief, two dsDNA insertions and a backbone vector were PCR amplified with ssDNA primers containing degenerate codons using SuperFi Platinum II. After PCR amplification, the reactions were incubated with 2 μL of FastDigest DpnI for 30 minutes at 37° C. dsDNA PCR products were cleaned up using Machery Nagel PCR clean up kits. The inserts were assembled into the backbone using NEB HiFi DNA Assembly at a molar ratio of 9:9:2 (insert:insert:backbone, 0.2 pmol total DNA) in a 10.2 μL total volume reaction. The DNA

assembly was cleaned and eluted with 10 μL of water. 2 μL of the product was transformed into 66 μL of NEB 10-β cells (New England Biolabs: 101417-922). After recovery in 1 mL of prewarmed SOC, the *E. coli* were rotated at 240 RPM in a 37° C atmosphere. The library was diluted and grown in Luria Broth with ampicillin (50 μg/mL) overnight at 37° C at 240 RPM. The plasmid DNA was recovered with a Machery Nagel endotoxin free DNA prep kit. Fifteen colonies from the dilution plate were selected at random for sequencing.

4.4.4 *Library Screening Stimulation*

HEK293T TetBxb1BFP cells stably expressing the ligand selectivity library were seeded onto PDMS microwell arrays as previously described in 2.4.3. Notably, this library required only 3 μg/mL doxycycline for stable expression. After recovery overnight, the arrays were rinsed 3X with prewarmed, triple supplemented Tyrode's solution and moved to the microscope. The control fluorophore was imaged before and after sensor imaging and stimulation with 578/21 nm excitation and 641/75 nm emission (200 ms, 10 A.U.). During stimulation, the arrays were constantly illuminated with blue light (Semrock GFP filters; excitation 474/27 nm, emission 520/35 nm) for 150 seconds. At ten seconds, fifty, and one hundred seconds, ligands were added to the bath by hand. The first ligand is the undesired agonist, the second ligand added is the desired ligand, and the final ligand is the control antagonist. One final bright field image was taken. All images are automatically saved and then immediately analyzed using the screening algorithm in Section 4.4.5. Plots in **Figure 4.6** and **Figure 4.7** generated using the dscatter function from [148] .

4.4.5 *Library Screening Algorithm*

The images were analyzed with an updated algorithm to identify the highest performing cell that does not detect the undesired ligand. Masks are generated by a logical AND of the pre and post stimulation control fluorophore images. Next, to generate the array rank ratio of the stimulation image, average intensity projections are taken of the baseline imaging, the first ligand, second ligand, and third ligand imaging windows. The Opto-MASS array rank ratios of the undesired ligand, desired ligand response, and control antagonist were calculated using Equation (2.2). Integrated Morphometry Analysis was used to measure the average fluorescence of the Array Rank Ratios. ROIs were defined by the mask generated by the control fluorophore and the fluorescence intensity for each ROIs Opto-MASS Array Rank Ratio were measured. ROIs with

too great of fluorescence change to the undesired ligand were removed, determined by a threshold calculated using Equation (4. 1). The greatest responding ROIs in the desired ligand were ranked for fluorescence response. ROIs were transferred to the stimulation imaging stack and the live field of view for user confirmation prior to cell recovery. For detailed imaging and image analysis information see **Figure 4.8.**

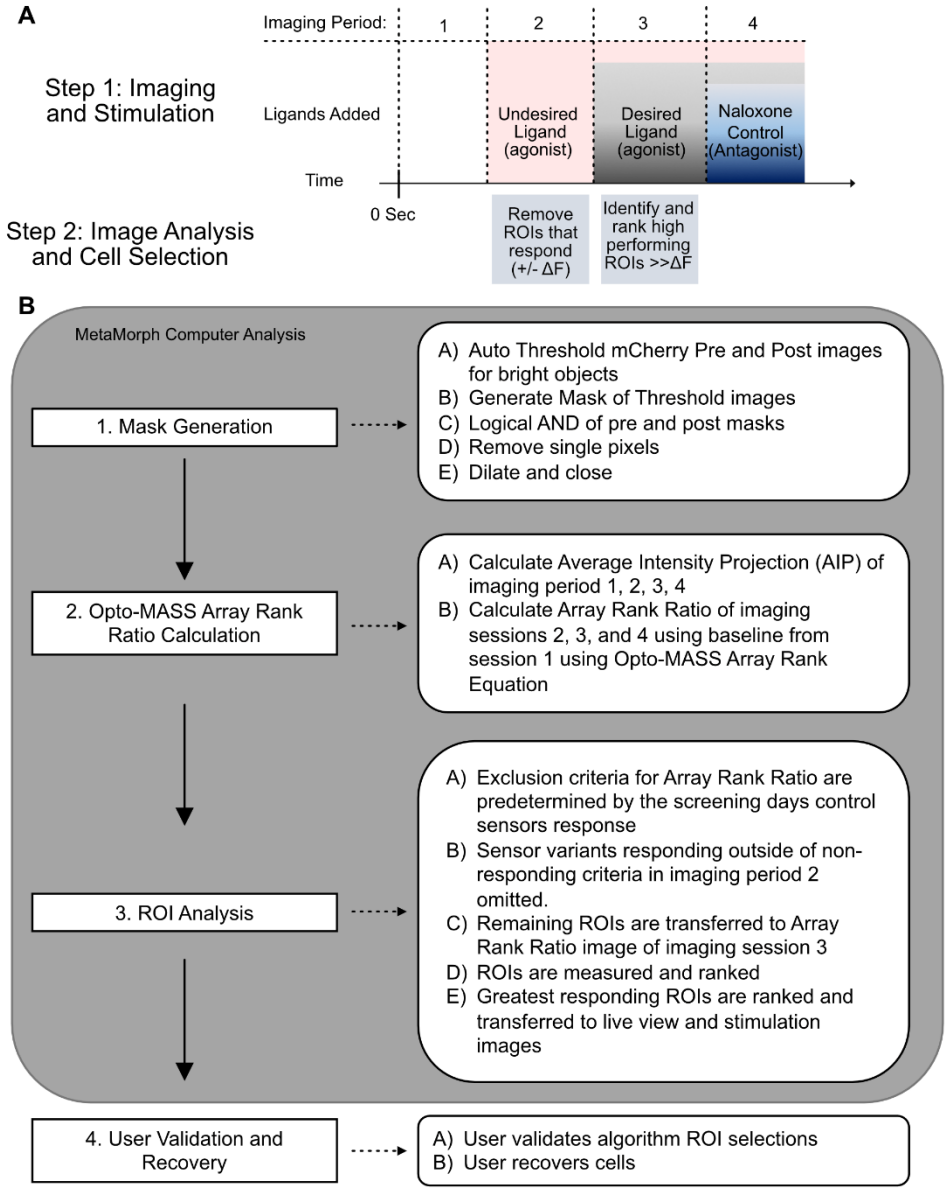

Figure 4.8: Overview of Opto-MASS multiplexed multiparameter screening

A) Overview of the multistep library stimulation. Opto-MASS arrays are continuously imaged for four imaging periods. Period one is to determine sensors baseline fluorescence. Period two determines the sensors response to the undesired ligand, period three is the sensors response to the desired ligand in the presence of the undesired ligand. Period four is a control period, showing the sensor can be turned off with a pharmacological control.

B) Detailed explanation of algorithm analysis used to rapidly identify desired constructs.

Chapter 5. EXPANDING OPTO-MASS APPLICATIONS TO OPTOGENETIC ACTUATORS

Abstract:

Optogenetic actuators use light to control specific biological signaling pathways in cells. In circuit neuroscience, tools that inhibit, or reduce neuronal firing have proven to be a powerful approach to determine a neuron subtype's role in neural circuits and animal behavior. Inhibitory opsins have been demonstrated as a useful way to inhibit neuron firing. While there are many opsins available in nature, researchers may wish to optimize certain biophysical characteristics of an opsin. Currently, opsin engineering is limited by low throughput methods of validating single opsin variants. Here, we expand the applications of the high-throughput platform, Opto-MASS, to opsin engineering. We demonstrate the platform can express parapinopsin (PPO), a bistable inhibitory opsin, with Pink Flamindo, a fluorescent biosensor for cAMP. We inhibit forskolin dependent cAMP production in mammalian cells by adapting the Opto-MASS screening capabilities. This demonstration of PPO activation by Opto-MASS expands Opto-MASS to the engineering of optogenetic actuators.

5.1 INTRODUCTION

Optogenetic tools are broadly divided into two types of tools. Optogenetic sensors pair light and genetic tools to observe and report on neuronal activity. Optogenetic actuators pair light and genetic tools to control biological signaling pathways. When applied to neuroscience, the most common optogenetic actuators modulate neural firing, either increasing or decreasing the firing rate. Optogenetic actuators can have several different approaches to altering neural firing. There are light-activated microbial opsins, that increase the conductance of ions across the membrane, or pump ions across the membrane, to alter the membrane firing potential. A second type of tool uses vertebrate opsins, which use light to trigger intracellular signaling pathways, such as G-proteins, to modulate a cell's intracellular signaling pathways.

Engineered and naturally discovered microbial opsins have been shown to effectively activate or suppress neural firing. Optogenetic actuators use light to control the activity of neurons. Channelrhodopsin (ChR), a light activated ion channel, initiated the widespread acceptance and adoption of optogenetics[1]. Light sensitivity, kinetics, spectral shifts, and conductance were all engineered into channelrhodopsins [3], [4], [149], [150]. Initially, light sensitive proteins were solely used to stimulate activity in neurons, but neural inhibition is a significant regulator of neural circuits and an important experimental variable to control. Inhibitory ion pumps and ion channels were soon engineered or identified from nature, expanding the suite of optogenetic tools to include tools to silence neural activity [151]–[154]. Light activated ion channels and ion pumps present a useful solution to neural inhibition, however there are limitations to using these tools for neural inhibition [154]–[156]. These limitations include compensatory firing, Cl^- conducting channels may raise resting membrane potential, Cl^- may be excitatory in early developmental stages of the mammalian nervous system, and pH sensitivity. To overcome some of the limitations presented by ion channels and pumps, protein engineers harnessed inhibitory GPCRs to modulate neural firing.

Recently, researcher identified a light sensitive, inhibitory opsin from lamprey, called parapinopsin[157]. Parapinopsin (PPO) has been demonstrated to be a useful optogenetic tool for presynaptic inhibition[158]. PPO functions by absorbing light predominantly in the ultraviolet range in the dim state. Upon transition to the light state, PPO's peak absorption shifts to a longer

wavelength in the amber range, approximately 500 nm. During light absorption, a molecule called retinal shifts between states and stabilizes PPO between inactive and active G-protein signaling states, thus toggling PPO 'on' and 'off'[157], [158].

PPO overcomes kinetic issues with other inhibitory GPCR tools like DREADDS, which are constrained by the kinetics of drug absorption and metabolism. PPO also does not have the compensatory firing issues that light activated inhibitory channels and pumps may have. Engineers have also built chimeras of the light sensitive portions of rhodopsin and structures that couple to inhibitory GPCRs, such as the Opto-XRs and Opto-MORs[159], [160]. However, these tools may not traffic well to the surface of neurons, and can signal through off target pathways. Researchers may desire to alter the function of PPO. Neuroscientists may have a need for a more sensitive PPO, or to shift the absorption spectra, or desire to shift the G-protein subtype the receptor preferentially couples. Traditional protein engineering techniques would require scientists to build, screen and analyze the biophysical characteristics of individual mutants. The protein engineering process would be particularly laborious, as the biochemical signaling pathways PPO uses to modulate neural function are only fully recapitulated in mammalian cells. Opsins are complex molecular machines, and typically bestowing new functions onto the protein requires mutations of multiple sites. Testing these combinatorial mutations would require a prohibitively large screen by traditional methods. Researchers need a high-throughput method to rapidly assess the function of PPO in mammalian cells.

5.2 Results: Expanding Opto-MASS to Optogenetic Actuators

The optogenetic tool platform, Opto-MASS, can engineer sensors for dynamic range and ligand selectivity. We highlighted the platforms versatility during screening for ligand selectivity, by screening thousands of cells along three performance axes (Chapter 4). Significantly, the platform has not been applied to the screening of optogenetic actuators. Optogenetic actuators use light to modulate cellular physiology. When developing sensors, the platform harnesses a sensitive readout of sensor fluorescence. Here, we propose to leverage that capability by using a fluorescent sensor to measure downstream modulation of the cAMP signaling pathway by PPO. A tricistronic genetic cassette was stably recombined into HEK293T TetBxb1BFP cells (**Figure 5.1A**). HEK293T cells stably expressing PPO and the red fluorescent cAMP biosensor, Pink Flamindo [41], were seeded into PDMS Opto-MASS arrays and imaged (**Figure 5.1**). We assigned the blue spectrum

(474 ± 27 nm), previously used to observe sensor signals, to the stimulation wavelength of the optogenetic actuator. We assigned the control fluorophore spectra (554 ± 23 nm) to a fluorescent sensor for downstream signaling modulated by the PPO gene (**Figure 5.1**). We can use biophysical measurements taken over several imaging periods to quantify PPO activation. HEK293T TetBxb1BFP cells exposed to blue light after +100 µM forskolin (FSK) stimulation had an average 1.6-fold or 2.9-fold reduction in cAMP signal.

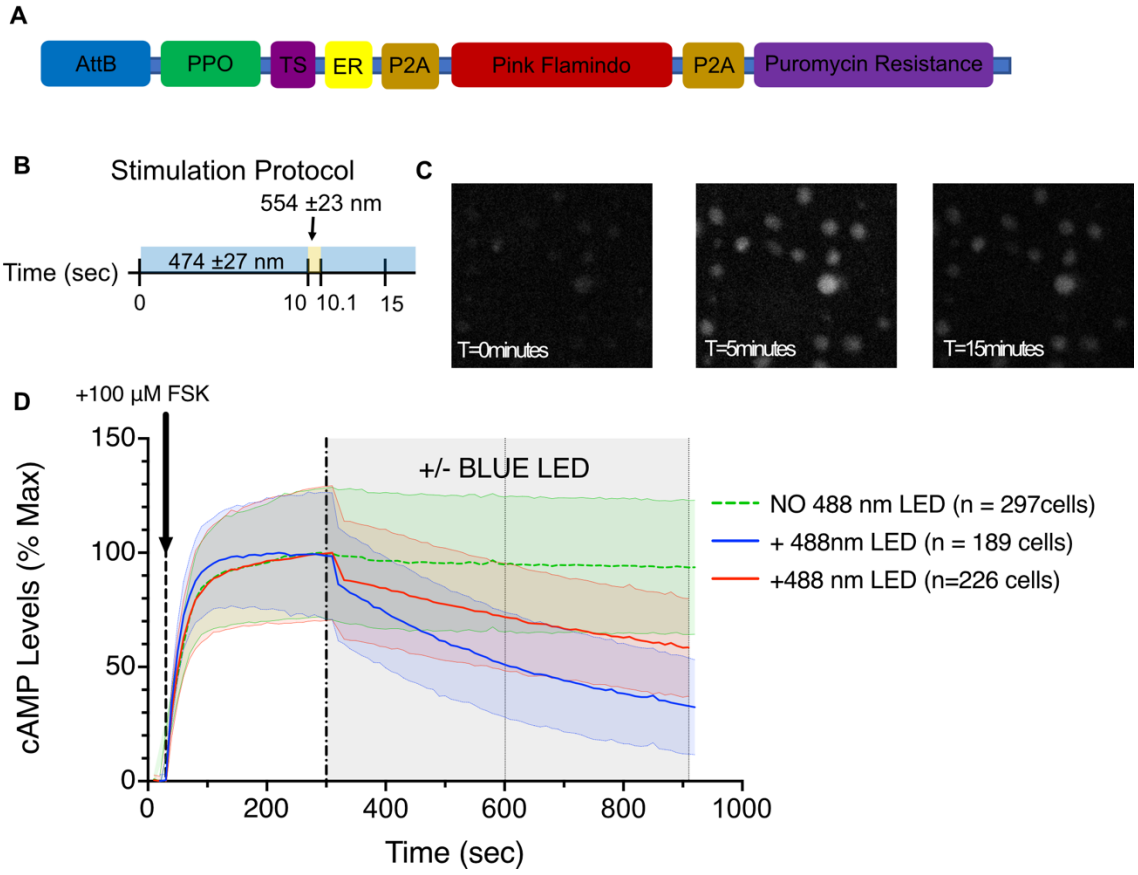

Figure 5.1: Opto-MASS can screen for parapinopsin functionality

A) Inhibitory optogenetic actuator, parapinopsin, is cloned into a HEK293T TetBxb1BFP compatible plasmid with membrane trafficking signals, red shifted cAMP sensor Pink Flamindo, and puromycin resistance gene.
B) PPO stimulation and cAMP observation require orthogonal wavelengths of light.
C) Representative images of PPO screening on Opto-MASS arrays.
D) Greater than 100 mammalian cells expressing the tricistronic cassette were screened on Opto-MASS arrays, demonstrating a blue light dependent decrease in Pink Flamindo fluorescence.

5.3 Discussion and Future Directions

Opto-MASS demonstrated it was able to screen optogenetic actuators, namely parapinopsin. cAMP production was stimulated by the broadly acting adenyl cyclase agonist, forskolin, and blue light stimulated PPO was shown to reduce the forskolin induced signal. Interestingly, the fold change reduction in cAMP signal was different in each array, suggesting a high biological variability in the stimulations. The measured biological variability could be due to temperature fluctuations that reduce PPO or AC enzymatic turnover, genetic expression levels, resting cAMP levels, or time spent in Tyrode's, not cell culture media. Tyrode's is missing bicarbonate, which is involved in maintaining mammalian homeostasis. Great care should be taken to account for variability in PDMS thickness, as light absorption by the PDMS will vary with thickness and thus alter the amount of light reaching the PPO molecules. Additionally, Pink Flamindo absorbs light in the blue range due to the choice of having a cpmApple fluorophore. The absorption of blue light may bleach the Pink Flamindo chromophore.

Significantly, missing from the data set is a genetic control. Mutation of Lys281$^{7.42}$Ala would remove the PPO ability to covalently bind retinal. This mutated PPO (PPO^{mut}) would serve as a powerful negative genetic control and could account for any blue light photobleaching of Pink Flamindo, and ensure the Opto-MASS pipeline is capable of directly measuring cellular phenotypes due to actuator genotype. To further improve Opto-MASS's actuator screening abilities, a more sensitive cAMP sensor could be implemented. Currently, there are green fluorescent cAMP sensors with better dynamic range or sensitivity [26], [161]. These tools could be implemented for a more precise readout of cAMP fluctuation. We chose to measure cAMP as we had fluorescent cAMP tools on hand. PPO couples to $G_{\beta\gamma,}$ which modulates downstream GIRK channels to reduce firing potential. GIRK channels are robust pathways to inhibit neurons, and in the future researchers may wish to modulate PPO's ability to effect GIRK channels. Recently, several genetically encoded fluorescent indicators to detect changes in K^+ have been published and could be implemented to screen PPO's $G_{\beta\gamma}$ coupling and downstream signaling efficiency [162], [163]. Additionally, if monitoring intracellular potassium levels proves to be too challenging due to the K^+ GEFI's sensitivity and dynamic range, membrane localization of a fluorescently tagged G_β could be implemented, as it has been demonstrated as a valid way to screen PPO signaling[158]. The PPO stimulation results on the Opto-MASS platform expand the space Opto-MASS is capable

of engineering and open the door to high throughput opsin engineering. Rapid optimization of opsins will lead to more precise optogenetic experimental preparations to clearly disentangle neural circuits.

5.4 METHODS

5.4.1 *PPO Cloning, Expression, Stimulation and Data Analysis*

The parapinopsin gene was a gift of the Bruchas laboratory and was subcloned into the AttB expression vector with two trafficking signals, TS ER, using standard molecular biology technique and Gibson assembly (NEB HiFi). The PPO constructs were stably recombined into HEK293T TetBxb1BFP cells using the Methods from 2.4.3. HEK293T cells expressing PPO were imaged in Opto-MASS microwells under 10X magnification and stimulated with blue light (474±24 nm) at power of 5.49 mW with a LumenCore Light engine and a Leica DMI8 epifluorescent microscope controlled by MetaMorph (Molecular Devices). Data analysis was performed in Excel and GraphPad Prism (Ver 8.4.2).

Chapter 6. STRUCTURAL BASIS FOR ION SELECTIVITY AND ENGINEERING IN CHANNELRHODOPSINS[3]

ABSTRACT

Channelrhodopsins have become an integral part of modern neuroscience approaches due to their ability to control neuronal activity in targeted cell populations. The recent determination of several channelrhodopsin X-ray structures now enables us to study their function with unprecedented molecular precision. We will discuss how these insights can guide the engineering of the ion conducting pathway to increase its selectivity for Cl^-, $Ca^{2+,}$ and K^+ ions and improve the overall conductance. Engineering such channelrhodopsins would further increase their utility in neuroscience research and beyond by controlling a wider range of physiological events. To thoroughly address this issue, we compare channelrhodopsin structures with structural features of voltage and ligand-gated K^+, Cl^- and Ca^{2+} channels and discuss how these could be implemented in channelrhodopsins.

[3] This chapter is previously published as a review: Rappleye M and Berndt A, "Structural basis for ion selectivity and engineering in channelrhodopsins," *Curr. Opin. in Struct. Biology*, Volume 57, 2019, Pages 176-184, https://doi.org/10.1016/j.sbi.2019.04.008.

6.1 INTRODUCTION

Channelrhodopsins are light-activated cation and anion channels from microorganisms such as *Chlamydomonas reinhardtii*. The majority of known channelrhodopsins are non-selective cation channels that conduct monovalent and divalent cations and protons [164]. Additionally, a family of engineered and naturally evolved channelrhodopsins is highly selective for Cl^- [165]–[167]. Channelrhodopsins were quickly adopted as tools for basic neuroscience research because the conducted ion species play major roles in exciting or inhibiting neuronal activity. Researchers can control ionic currents in targeted cell populations by direct light stimulation [168]. Since the early adoption of channelrhodopsins as research tools, several studies aimed to increase their utility by enhancing or changing biophysical properties.

Conventional channelrhodopsins conduct Na^+, K^+ and Ca^{2+} non-selectively which depolarize neurons and triggers action potentials. Exclusively Ca^{2+} or K^+-selective channelrhodopsins could be applied to specialized tasks such as the initiation of intracellular signaling cascades or hyperpolarization of neuron membranes. Furthermore, the unitary conductance of channelrhodopsins is about $10\text{-}10^4$ smaller compared to many neuronal ion channels (~40 fS for Channelrhodopsin-2) [169]. Increasing the ion transport rate could reduce the necessary amount of channelrhodopsins in cell membranes to trigger a physiological response to light.

Several recently published x-ray structures identified the position and molecular composition of the ion conducting pathway, generating confidence for targeted engineering approaches [170]–[175]. Molecular engineering has often been inspired by examples found in nature. Therefore, we will examine how similarities and differences between channelrhodopsins and ligand and voltage-gated ions channels can guide efforts to engineer channelrhodopsins. However, one missing piece for the thorough comprehension of channelrhodopsin function is an ion-conducting, open-state structure. Consequently, the ion binding sites have not been identified with absolute certainty and the precise mechanisms of selectivity and conductivity remain elusive. Hence, we will also discuss approaches that could reveal a channelrhodopsin open-state structure and thus provide a path towards rational engineering of channelrhodopsin ion pores.

6.2 THE ION-CONDUCTING PATHWAY IN CHANNELRHODOPSIN

Channelrhodopsins belong to the family of microbial opsins which share common structural features and biophysical properties. For example, all channelrhodopsins contain a seven α-helices transmembrane domain. The membrane domain encloses a covalently bound retinal molecule, which serves as the light-absorbing chromophore. It has been shown that Channelrhodopsin-2 (ChR2) can conduct large organic cations such as dimethyl-amine (2.6 Å radius). Following these results, the effective pore size was determined to be 6.4 Å, which indicates that ions can be transported in partially or fully hydrated states [12]. Furthermore, the permeability for ions is inversely correlated to their sizes, i.e., smaller ions such as Li^+ (0.9 Å radius) generate larger currents than Cs^+ (1.8 Å radius). Ca^{2+} and Mg^{2+} have higher binding probabilities than monovalent cations but lower transport rates, which significantly lowers current amplitudes at high Ca^{2+} and Mg^{2+} concentrations [176]. Protons have the highest binding probability and transport rates at acidic pH values, but their contribution to the current amplitudes is smaller under neutral pH conditions.

A chain of glutamate residues (E82/83, E90, E97, E101) in ChR2 has long been suggested to be a crucial element of the putative pore. The crystal structure of the channelrhodopsin hybrid C1C2 revealed that the homolog glutamate residues (E121/122, E129, E136, E140) are indeed located throughout transmembrane helix 2 (TM2) [170] (**Figure 6.1**). Although C1C2 forms a dimer, each subunit contains an independent ion pathway defined by four alpha-helices (TM1, 2, 3, 7). The extracellular vestibule (opening = 8 Å diameter) is slightly electronegative and closed at a constriction site called the central gate (**Figure 6.1, Figure 6.2**).

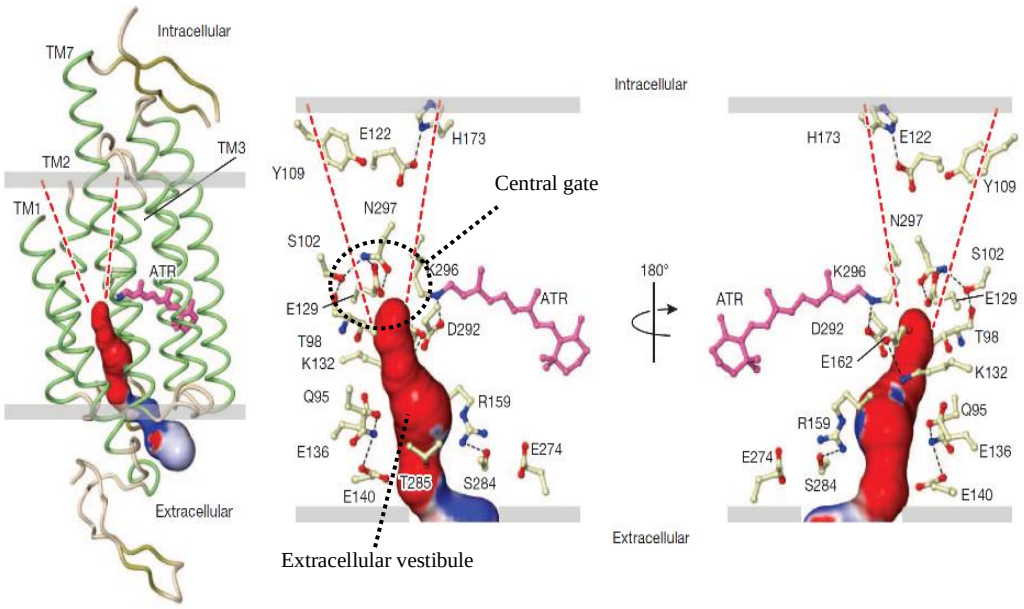

Figure 6.1. Closed-State x-ray structure of the cation-conducting channelrhodopsin C1C2 (PDB ID: 3UG9) The structure depicts the overall structure with seven transmembrane alpha-helices (TM). The pore is formed by TM1, TM2, TM3, TM7 and lined with acidic residues that generate an electronegative cavity (red). The extracellular vestibule extends to the central gate, near the light-absorbing chromophore retinal (magenta). The cytosolic site is constricted by S102, E129, and N297 near the center as well as Y109 at the putative pore entrance. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature, [177].

The cytoplasmic side is entirely occluded in this dark-state structure, by several highly polar residues (S102, E129, and N297) near the center of the pore and by Y109 near the entrance. Strikingly, replacing a majority of the acidic pore residues by neutral or basic residues converts C1C2 into the Cl^--selective channel iC++ [2][15]. Similarly, replacing the central E90 by glutamine in ChR2 generate chloride conductivity in the engineered variant ChloC [4]. Soon after the engineering of anion selectivity, a natural Cl^- conducting channelrhodopsin, GtACR1 (*Guillardia theta* Anion-Conducting Channelrhodopsin), was found by screening the genomic DNA of *Guillardia theta* [166]. Three crystal structures of GtACR1 and iC++ have since been published [179]–[181]. While the overall configuration of GtACR1 and iC++ is similar, the molecular composition of the ion conducting pathway stands in stark contrast to C1C2. For example, C1C2 has seven carboxylate residues in its ion-conducting pathway, and 14 on the intra- and extracellular surface, which potentially attract cations and repel negatively charged anions. In contrast, GtACR1 has only three acidic residues in its pore, while 12 of the carboxylate residues on its inner and outer surface are replaced by the basic residues arginine or lysine (**Figure 6.2**).

Although the protonation states of acidic and basic residues remain elusive, we can assume that the electrostatic potential of the ion conducting pathway and protein surface is a major determinant of ion selectivity in channelrhodopsins. As discussed below, we can find these principles in a majority of ligand and voltage-gated channels. However, there are some striking differences. For example, the ion pathways of many ligand and voltage-gated channels are comprised of 2, 4, 5 or 6 topologically quasi-identical protein subunits or domains which form pores with topological 2, 4, 5 or 6-fold symmetry, respectively [182]–[185]. These configurations are critical for the formation of highly selective filters in Ca^{2+} and K^{+} channels. In contrast, the monomeric channelrhodopsin pore has no symmetrical features.

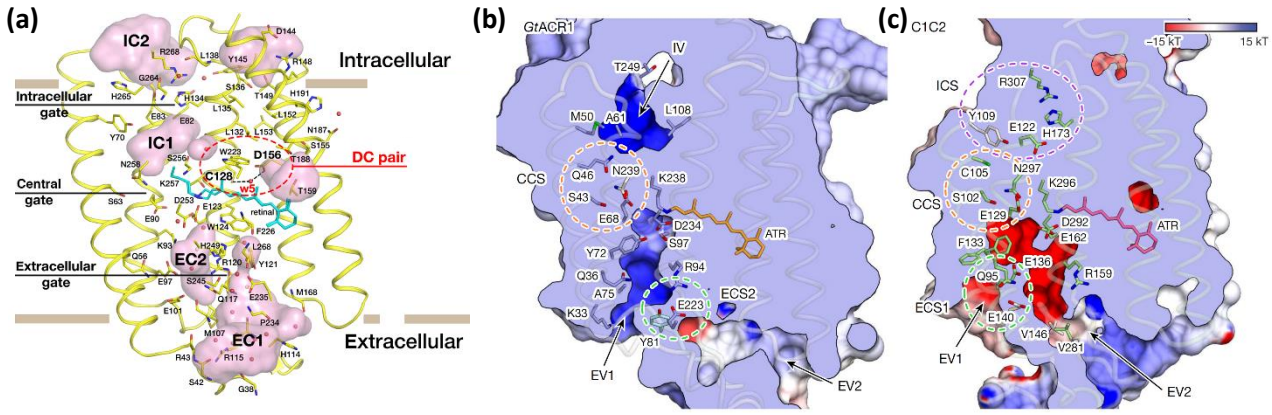

Figure 6.2. Comparison of channelrhodopsins with varying ion transport rates.

(a) Channelrhodopsin-2 (ChR2) has three constriction sites in its dark-state structure: Intracellular gate (cytosolic site), Central gate (center) and Extracellular gate (within the extracellular vestibule). IC1, IC2, EC1, EC2 depict intra- and extracellular vestibules respectively. (PDB ID: 6EID). From [11]. Reprinted with permission from AAAS. **(b)** The conductance rates of Cl^{-} selective GtACR1 are 25 times larger compared to ChR2. There is only one restriction site at the center of the pore (CCS, red circle) (PDB ID: 6CSM). The extracellular vestibule 1 (EV1) is open albeit forming an alternative pathway compared to C1C2 shown in (c). The constriction site ECS2 (green circle) separates EV1 from EV2. The lack of acidic residues generates a potentially positively charged, electrostatic environment within the pore (blue). **(c)** The ion pathway of C1C2 (PDB ID: 3UG9) is occluded at two sites, at the center (CCS, red circle) and at the cytosolic entrance (ICS, purple circle), which could contribute to larger photocurrents compared to ChR2. The excess of acidic residues in the C1C2 pore creates a negative electrostatic potential (red) which promotes cation transport. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature, [180].

An interesting finding in the channelrhodopsin structures is an apparent correlation between the number of constriction sites and the conductivity. For example, C1C2 generates larger

photocurrents than ChR2. ChR2 has three constrictions, one cytosolic, one extracellular and one in the center [186] (**Figure 6.2A**). C1C2 has two constrictions, at the center of the pore and at the cytosolic entrance, with the extracellular vestibule being open in the dark-state [170] (**Figure 6.1 & Figure 6.2C**). Furthermore, GtACR1 has a more than 10 times larger conductivity (~550 fS) compared to ChR2 and is only occluded at the center in the closed state [166] (**Figure 6.2B**). It is likely that channelrhodopsins evolved from light-gated proton pumps such as bacteriorhodopsin which do not form open pores [187]. Hence, extensive constriction sites could be remnants from protein predecessors. Nevertheless, upon light-activation, channelrhodopsins must open these barriers to enable ion flow. Channelrhodopsin cycle through open and closed intermediates under continuous illumination. While currently speculative, it is possible that more constriction sites result in a higher ratio of closed state intermediates under continuous illumination thereby lowering current amplitudes and conductance measured using stationary noise analysis [3] [6]. Additional structures of channelrhodopsins with large current amplitudes such as ChR2-T159C, or *Chloromonas oogama* channelrhodopsin could test this correlation [21] [22]. If true, future engineering approaches should aim to remove the extracellular and cytosolic constriction sites to increase conductivity.

6.3 Engineering Chloride Conducting Channelrhodopsins for Shunting Inhibition

Channelrhodopsins such as iC++ can conduct monovalent anions (Cl^-, I^-, Br^-) but not larger charged molecules like aspartate and glutamate [15]. Therefore, the dominant ion species in anion selective channelrhodopsins, expressed in vertebrate brains, is Cl^-. Consequently, these ion channels can mimic the function of neuronal $GABA_A$ receptors and suppress action potential generation through shunting inhibition [178]. Several families of ligand-gated ion channels also have members that are either cation or anion selective. For example, the nicotinic acetylcholine receptor (nAch) and the 5-HT receptor conduct cations while glycine and GABA receptors conduct anions [18] [23] [24]. All four of the channels are pentameric cys-loop channels that share a similar overall topology. However, the selectivity filter of the cation-conducting receptors contains a horizontal ring of 5 glutamates (1 from each subunit) whereas the anion-selective channels place alanine at these positions (**Figure 6.3A&B**). Mutations of glutamates to alanines increased the

selectivity for Cl^- in the nAch and 5-HT_{3A} receptors [190] [191]. Interestingly, the structure of a pentameric Cl^- channel from *Caenorhabditis elegans* revealed that the pore does not contain any potentially charged residues [192] (**Figure 6.3B**). Nevertheless, the electrostatic potential of the pore is positive due to oriented peptide dipoles in the transmembrane alpha helices. This feature can also be found in CLC Cl^- selective ion channels [193]. The nAch and 5-HT receptors reverse the dipole moment by the placement of glutamates within the selectivity filter, by creating a repulsive electrostatic barrier for negatively charged ions.

Another instance of topologically similar variants that are either cation or anion selective are Ca^{2+}-activated TMEM16 channels. For example, TMEM16F is a non-selective cation channel while TMEM16A is Cl^- selective. TMEM16A contains several basic residues in its putative pore that are positively charged when protonated [194] [195]. Conversely, the murine cation channel TMEM16F contains neutral glutamine at one of these positions. When this glutamine is replaced by lysine, the ion selectivity of TMEM16F is significantly biased toward anion permeation, indicating that electrostatic barriers play major roles for anion vs. cation selectivity [196]. Another example is the cystic fibrosis transmembrane conductance regulator (CFTR), which contains several basic residues within the pore region (**Figure 6.3C**) [182]. Thus, Cl^- selectivity in CFTR is most likely mediated by the positive electrostatic potential of the pore which attracts anions and repels cations. These examples demonstrate that electrostatic potentials are a major driver for anion vs. cation selectivity in a wide variety of structurally diverse ion channel families, including channelrhodopsins.

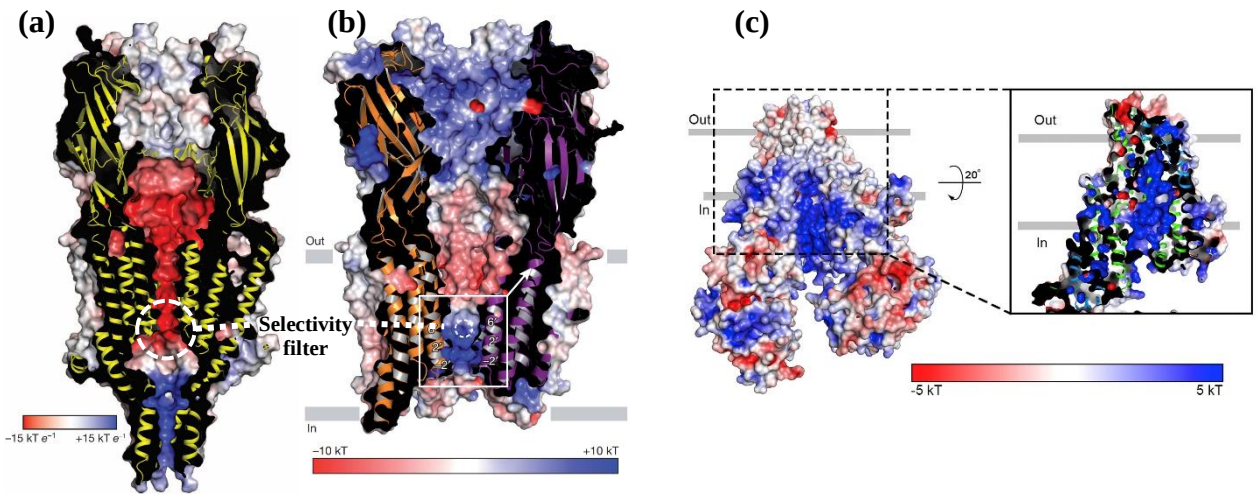

Figure 6.3. Electrostatic potentials within the pore regions of ligand-gated cation and anion channels.

(a) Cross section through the x-ray structure of the pentameric, mouse serotonin 5-HT3 receptor, a cation channel (PDB ID: 4PIR). The electrostatic potential of the selectivity filter (white circle) is electronegative (red) due to charged glutamates which orient the peptide dipoles within the pore region and facilitate the transport of cations. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature, [197]. **(b)** Cross section through the x-ray structure of the pentameric, glutamate-gated Cl^- channel GluCl-alpha from *C. elegans* (PDB ID: 3RIA). The selectivity filter is electropositive (blue, narrow bottom) albeit the lack of charged residues in this region. Instead, oriented peptide dipoles provide an electrostatic environment that facilitates the transport of anions. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature, [192]. **(c)** Cryo-EM structure of the anion selective Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) from zebrafish. The ion-conducting pathway is formed at the interface of the dimer. Lysine and arginine residues generate a potentially electropositive environment (blue) within the pore that attracts anions and repels cations. Note, CFTR does not have a single selectivity filter. Instead, anion selectivity is controlled by several ion binding sites along the pore. Reprinted from [16], with permission from Elsevier.

6.4 Potassium Selective Channelrhodopsins for Hyperpolarization of Neuron Membranes

Ultimately, ion gradients determine if ion channels depolarize or hyperpolarize membrane potentials. The shunting capabilities of Cl^- conducting channelrhodopsins are thus dependent on Cl^- gradients. Cl^- gradients are variable within neurons and do not always permit strong inhibition. For example, pre-synaptic terminals have higher intracellular Cl^- levels than dendrites and somata [198]. Consequently, Cl^- channels depolarize high intracellular Cl^- cell membranes from resting potentials and could induce neurotransmitter release through activation of voltage-gated Ca^{2+} channels [199]. On the other hand, K^+ gradients are relatively stable throughout neurons. Thus, K^+

mediated currents will cause membrane hyperpolarization and strong neuronal inhibition under most physiological conditions [200]. Hence, the engineering of K^+ selective channelrhodopsins would be another critical tool in the optogenetic arsenal.

Studying K^+ channel structures aims to solve an obvious conundrum. How can these channels selectively transport the larger K^+ ions (1.33 Å radius) over the smaller Na^+ ions (0.99 Å radius)? While the exact mechanisms are still debated, there is a consensus on several key factors. 1. K^+ channels share a common tetrameric pore structure which results in selectivity filters with four-fold symmetry [201] (**Figure 6.4A**). 2. The filter has four contiguous ion binding sites. 3. The hydration shell, surrounding K^+ ions is largely removed and replaced by main chain carbonyl oxygen atoms in the selectivity filter. The thermodynamic costs for removing the hydration shell are lower for K^+ than for Na^+, which is crucial for K^+ selectivity. 4. The selectivity filter is occupied by two or more contiguous K^+ ions. Cations move through the pore by electrostatic ion-ion repulsion (knock-on mechanisms) (**Figure 6.4B**) [34][35]. As demonstrated by molecular dynamics studies, the specific configuration of the selectivity filter favors K^+ transport, as Na^+ would not occupy binding sites in an optimal position for efficient knock-on transport driven by electrostatic repulsion. All available scientific evidence strongly suggests that the architecture of the selectivity filter is the overall governing factor for K^+ selectivity. For example, the nonselective, bacterial cation channel NaK can be converted into a K^+ channel when the number of ion binding sites in the selectivity filter is increased from 2 to 4 [204] [205] (**Figure 6.4A**).

The strict architectural requirements render the engineering of K^+ selective channelrhodopsins extremely difficult. Several features that do not exist in channelrhodopsins have to be implemented, such as a distinct selectivity filter with 4-fold symmetry and four contiguous ion binding sites that mimic the hydration shell of K^+ ions. However, channelrhodopsin pores are asymmetrical, ions are most likely not dehydrated during transport, and channelrhodopsins do not have a distinct selectivity filter. Also, the central gate is essential for the light-gated opening mechanism which excludes this region as a potential site for an ion selectivity filter. Instead, a potential engineering strategy could aim to elongate the pore-forming transmembrane alpha helices into the extra or intracellular space to allow more flexibility for positioning a selectivity filter. If successful, any potential K^+ selective channelrhodopsin has to be thoroughly analyzed and confirmed by electrophysiological ion substitution experiments with K^+.

Only electrophysiological ion substitution experiments can confirm that shifts in reversal potential originate from K^+ selectivity and not Cl^- [167], [206], [207].

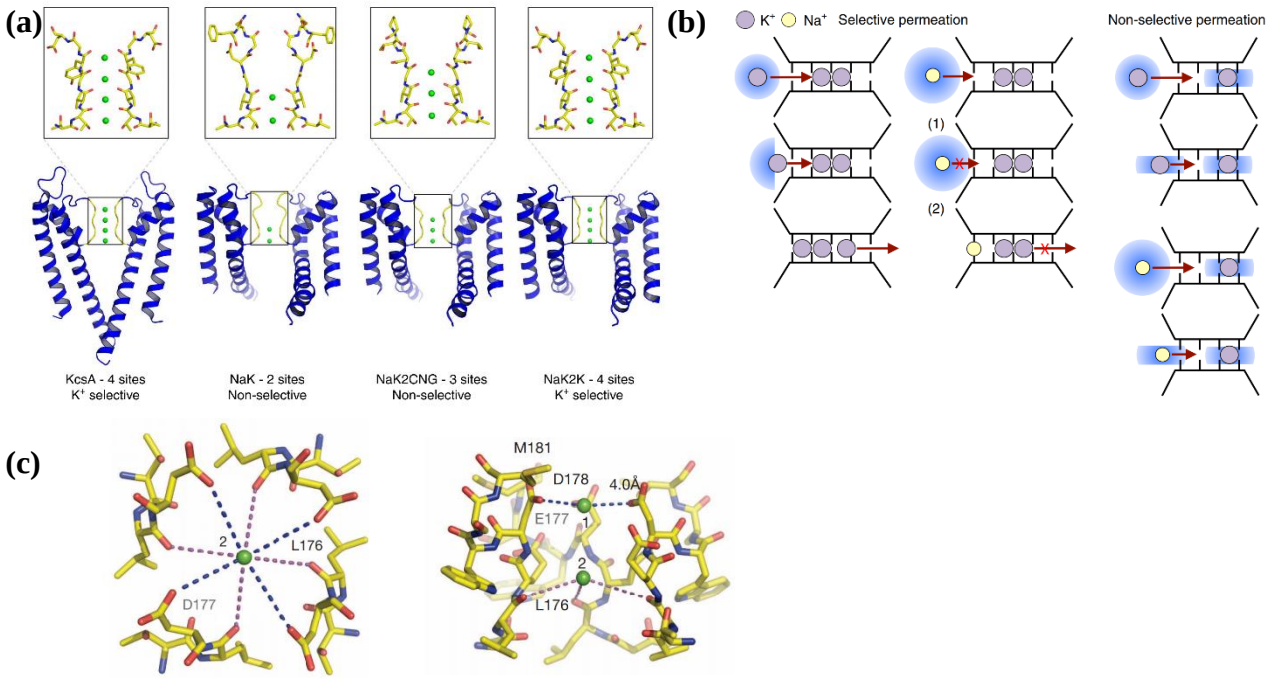

Figure 6.4. Structural features for ion selectivity in K^+ and Ca^{2+} ion channels.

(a) Left: structure of the selectivity filter in the native, K^+ conducting channel KcsA, including four bound, dehydrated K^+ ions (green) (PDB 1K4C). The selectivity filter provides four ion binding sites through main chain carbonyl oxygen atoms (red). 2nd left: The non-selective cation channel NaK provides only two ion binding sites (PDB ID: 3E8H). 2nd right: Increasing the binding sites of NaK from 2 to 3 in NaK2CNG is not sufficient to generate an exclusively K^+ selective pore (PDB ID: 3K03). Right: Further increase of the ion binding sites to 4 in the engineered channel NaK2K generates a highly K^+ selective filter region similar to KscA (PDB ID: 3OUF). Note, only two of the four pore domains are shown. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature, [37]. **(b)** Model of selective K^+ ion transport. Left/Middle: K^+ selective filters remove hydration shells (blue) which thermodynamically favors K^+ entry over Na^+. K^+ ions occupy 2-3 sites and are in direct contact which facilitates ion transport through electrostatic repulsion. Sodium ions do not bind in ideal positions for electrostatic repulsion. Right: Nonselective pores have fewer ion binding sites and only partially dehydrate ions, which results in equal binding probabilities for Na^+ and K^+. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature, [208]. **(c)** The structure of the engineered channel Ca_vAb shows that the selectivity filter orients Ca^{2+} (green) with carboxyl side groups as well as main chain carbonyl groups to accommodate divalent charges. Two hydrated Ca^{2+} move through the binding sites by mutual electrostatic repulsion. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature, [209]

6.5 CALCIUM SELECTIVE CHANNELRHODOPSINS FOR INDUCING SYNAPTIC VESICLE RELEASE

Native neurotransmitter release is mediated by Ca^{2+} flowing through highly selective voltage-gated Ca^{2+} channels into pre-synaptic terminals. An optimized Ca^{2+} selective channelrhodopsin (CaChR) could trigger pre-synaptic neurotransmitter release with a lower risk of generating backpropagating action potentials.

Efforts to engineer a Ca^{2+} selective ChR must overcome several challenges unique to Ca^{2+} conductivity; Na^+ ions are similar in size to Ca^{2+} (0.95 Å to 0.99 Å radii) and Na^+ is significantly more abundant than Ca^{2+} under physiological conditions. Nevertheless, nature has evolved multiple Ca^{2+}- selective ion channels. Most of them are hypothesized to harness the divalent nature of Ca^{2+} in their selectivity filters, such that the movement of ions through the filter is dependent on the divalent charge, effectively excluding Na^+ [209]. Many Ca^{2+} channels share the same tetrameric topology, with P-loops that harbor the selectivity filter, similar to K^+ channels (**Figure 6.4C**). Ca^{2+} selectivity in mammalian $Ca_v1.1$ is coupled to a planar ring of four glutamate residues (EEEE motif) positioned in pore loops between TM5 and 6 in each of the four domains [210]. The selectivity filter is hypothesized to function through selective binding and trapping of Ca^{2+} ions in the EEEE motif. When a second Ca^{2+} enters, it electrostatically pushes the EEEE motif trapped ion through the filter (knock-off mechanism) (**Figure 6.4C**) [211], [212]. Ca^{2+} ions are hypothesized to remain fully or partially hydrated as they permeate through the pore. Their divalent nature is critical to the putative function of the EEEE motif, as monovalent cations do not possess a large enough charge density to push Ca^{2+} out of the selectivity filter. Hence, any rational design effort to engineer CaChRs should aim to mimic the transmembrane domain topology seen in tetrameric P-loop channels. To mimic P-loop channels, CaChRs would require the generation of 2-3 non-dehydrating ion binding sites within the selectivity filter (1-2 with high Ca^{2+} selectivity). While still challenging, engineering a Ca^{2+} filter is more flexible in contrast to the longer and narrower K^+ filter.

Ca^{2+}-selective filters do not necessarily depend on carboxyl side chain groups, which generates this steric flexibility. In TRPV1 channels, main-chain carbonyl oxygen atoms of four glycine residues coordinate Ca^{2+} binding in the selectivity filter [213]. Furthermore, bacterial voltage-gated Na^+ channels such as NavAb also utilize EEEE motifs [214], [209]. To explain this

discrepancy, it has been proposed that the larger net charge of Ca^{2+} filters facilitates Ca^{2+} over Na^{+} binding. Electrostatics are affected by the residues within the vicinity of selectivity filters, [38],[215]. Thus, electrostatic simulation of ideal pore formations should precede and guide future engineering of CaChR, ideally combined with structure determination. To further enhance conductivity, additional efforts should be taken to generate electronegative vestibules at the entrance and exit of the pore. These elements are hypothesized to attract and bind Ca^{2+} for the 'knock-off' of Ca^{2+} through the filter into a hydrophobic region. Electronegative vestibules at the exit accelerate 'knocked-off' Ca^{2+}-ions through the hydrophobic pore region into the cytoplasm. All evidence suggests that these elements play a critical role in high-conductance Ca^{2+} channels [216], [217].

6.6 Future Directions and conclusions

Structural studies of highly selective Ca^{2+} and K^{+} ion channels suggest that a virtual symmetry of the pore and selectivity filters are critical elements for K^{+} and Ca^{2+} selectivity. Therefore, engineering a pseudo-symmetrical pore structure into channelrhodopsins might allow for the addition and manipulation of K^{+} and Ca^{2+} selectivity. A significant problem to overcome is the lack of an open-state structure of channelrhodopsin. Guided attempts to design ion specificity would be significantly accelerated if we could observe photo-induced cycling of the pore and gating mechanism through close-open-close states. However, traditional X-ray crystallography approaches encounter several issues. For example, crystallization of purified channelrhodopsins under light would result in a heterogeneous mix of open and closed-state intermediates with varying conformations, which would effectively prevent crystal formation. Generating nanobodies or antigen-binding (Fab) fragments that could bind and stabilize an open structure would be technically difficult because the "agonist," light, would need to be permanently applied within host animals [218] [219]. Even if possible, light is technically a fast-dissociating "agonist." Thus, even under constant illumination, channelrhodopsins molecules continue to cycle through open and closed states, which makes the accumulation of stable open-state intermediates difficult [220]. One alternative strategy could be to use variants that have an extended lifetime of the open state, such as the Stop-Function-Opsins (SFO) [49]. Dark-state protein crystals could be illuminated and immediately frozen in liquid nitrogen for subsequent structural x-ray analysis. This approach has

been successfully applied for bacteriorhodopsin and blue-light-activated LOV domains (light-oxygen-voltage) [50] [51]. However, this method could also destroy protein crystals due to tension induced by conformational changes. Furthermore, spectroscopic analysis of SFO variants showed complex photochemistry that could result in a heterogeneous mix of intermediates under extended illumination and interfere with the accumulation of a stable open state intermediate during cryo-trapping [52] [53].

On the other hand, serial femtosecond crystallography with X-ray free electron lasers (XFEL), would be capable of bypassing these challenges [54]. XFEL techniques have been used to observe crystal structures of membrane proteins, such as G-protein coupled receptors in a multitude of physical states [55]. Moreover, researchers were able to pair XFEL with timed light excitation of photoactivatable proteins in nano-crystals to determine light-induced conformational changes. The proof-of-concept has been demonstrated for photosystem II, fluorescent proteins, as well as for the light-gated proton pump Bacteriorhodopsin, which shares the main topological features with channelrhodopsins [223]–[226]. In these studies, protein crystals were grown in the dark and light-activated immediately before injection into the x-ray beam. Pairing this technique, dubbed time-resolved serial femtosecond crystallography (TR-SFX), with lipidic cubic phase crystallization methods, could be a feasible path for the observation of light-induced ion conduction pathways in channelrhodopsins.

In conclusion, we have analyzed how past engineering efforts generated Cl^- conducting ChRs by mimicking structural elements found in ligand-gated, Cl^- conducting ion channels. We also summarized key structural elements from Ca^{2+} and K^+ channels that contribute to their respective selectivity. Future engineering efforts are encouraged to apply these principles to generate Ca^{2+} and K^+ conductive channelrhodopsins for light-triggered control of physiological signaling mechanisms.

Chapter 7. CONCLUDING REMARKS

7.1 PROTEIN ENGINEERING FOR OPTOGENETICS

At its core, optogenetics is the application of light sensitive proteins in the mammalian nervous system. The combination of the spatiotemporal precision, paired with the genetic specificity provided by genetically encoded tools, makes optogenetics a method with unprecedented access to the mammalian nervous system.

Typically, the foundation of an optogenetic experiment requires researchers to express a protein or chimeric protein from other organisms in mammalian cells. These proteins typically are not immediately optimal for use in mammalian cells. They need to be mutated, have trafficking signals added, or have their expression levels increased or decreased to be compatible with the host organism's physiology. Thus begins the protein engineering. Mutations to proteins can bestow them with new catalytic activity, increase or decrease binding affinities, alter allostery between domains, or change folding dynamics. Despite advances of *in silico* protein folding predictions, there will always remain the need for *in vitro* validation prior to use in the mammalian nervous system.

In vitro tests of protein mutants can be resource and time intensive, driving the desire for high-throughput screening techniques. High-throughput screening techniques come with their own set of challenges. Experimental design must be carefully considered; if you make a mistake in your design, high-throughput experiments will repeat the mistake at great speed, much to your frustration. Despite the higher resource commitment to establish a high-throughput screen, the data generated by these screens can be used for multiple purposes. High-throughput screens with quality data generation can inform *in silico* models, enabling further refinement of computational models. Similarly, *in silico* models can identify non-intuitive, but useful mutation locations, and screen larger data sets than is reasonable for humans. The locations of site saturated mutagenesis can be informed by *in silico* models. The rich data set generated by high-throughput screens based off of these *in silico* informed locations can be fed back into the models, and can start a feedback loop to further refine both genetic libraries and computational models. Protein engineering will continue with advances in genetic expression techniques that enable easy

manipulation of higher order organisms, such as mammalian cells, and *in silico* models better informing mutational selections. Opto-MASS is a step towards true high throughput screening (> 1 million variants a day) in mammalian cells to provide rich data sets for computational models, and more thoroughly search a proteins sequence space.

7.2 Future Directions of Opto-MASS

When the Berndt lab set out to design and build the Opto-MASS pipeline, we envisioned a method to screen thousands of protein variants to identify highly desired sensor phenotypes. The pipeline would enable the easy recovery of identified cells for the sequencing of the underlying gene. The foundation of this implementation is the ability to rapidly screen thousands of mammalian cells because we want to use our end product in a mammalian expression system. The rapid screening and assessment of a protein's phenotype is considered a high-throughput technique. As with all high-throughput techniques, there should be a higher level of consideration and scrutiny during experimental design because resource commitments to execute a high-throughput screen are greater than typical, low throughput experiments. For the Opto-MASS pipeline, there are several biological requirements and resource limitations we needed to consider and engineer solutions for. The resulting pipeline has some unique characteristics as a direct result of these limitations. The pipeline and physical measurement of cells is quite easy to execute. Setting up experiments to run on Opto-MASS is not technically difficult; however experimental setup requires significant level of time investment for the researcher because there are many Opto-MASS specific consumables that must be built. While these initial steps require a time commitment, they are typically executed with a high success rate. However, the final step of gene recovery is technically and experimentally difficult.

A significant bottleneck to the usefulness of the pipeline is gene recovery. Our ability to directly link the genotype to measured phenotype is powerful characteristic of Opto-MASS. Other high-throughput techniques use large pools of data to identify enriched variants. These enriched variants are inferred to be the best genotype and must be validated. While there are several techniques and ways to engineer around the cell physical recovery problem, the Opto-MASS pipeline could have *in situ* barcode sequencing implemented to recover even more information such as kinetics, dynamic range and brightness of non-selected sensors. This would provide a highly informative dataset, linking multiple characteristics to many protein variants.

Prior to library recombination and expression, each variant would acquire a DNA barcode. Each variant in the library would be linked to a barcode using long form reads of high throughput sequencing. Currently, we observe and measure thousands of variants in each library but select individual cells to identify the gene. With *in situ* barcoding, we could treat each microwell like a minireactor, testing multiple performance axes for each variant, and after a gauntlet of biophysical tests such as kinetics, cell surface trafficking, ligand selectivity, *in situ* sequencing of the barcodes would link the measured biophysical phenotype to genotype [147].

The rich data sets provided by these experiments would link the sensor genotype to the sensor's biophysical characteristics like kinetics, dynamic range, ligand selectivity, and cell surface trafficking. These data sets could inform high quality *in silico* protein models and machine learning algorithms due to the detail and breadth of data they provide of a sensor's biophysical characteristics.

Our library construction currently generates a greater sequence space than is reasonable to screen. We use NNK degenerate codons, limiting the possible codon space to 32 out of 64. Adoption of NDT codons, which samples all the chemical identities encoded by the twenty canonical amino acids with a reduced set of twelve amino acids, could generate informed models of amino acid residue location and function. Opto-MASS screening throughput could reasonably cover the sequence space with 95% coverage [146]. Opto-MASS presents several avenues to pursue with the engineering of optogenetic tools. Opto-MASS can continue to be applied in its current form to optimize a host of GPCR derived sensors. There are over 700 GPCRs, and less than 100 published GPCR-based sensors. Opto-MASS could be used to engineer ligand selectivity for opioid receptor derived sensors and other endogenous receptors that bind many different opioid ligands and post-translationally modified opioid peptides. Optogenetic actuators, such as opsins could be engineered for spectral shifts or bias towards activating specific intracellular signaling pathways.

Enhanced optogenetic tools that are used to elucidate signaling mechanisms in higher order *in vivo* models such as non-human primates, could reveal therapeutic avenues not previously possible. For example, ligand selective opioid sensors could detect splice variants or post-translationally modified opioid peptide release that is unique to disease states. Advances in single cell RNA-seq can identify enzymes that may be responsible for these post-translational opioid peptide modifications. Small molecules pharmaceuticals targeting these enzymes to inhibit or

encourage their enzymatic activity could be developed as therapeutics, all as a result of identification with a ligand selective sensor. Opto-MASS could be used to engineer DREADDs that target specific intracellular signaling pathways. Paired with advances in genetic targeting that shorten neuron specific promoters, DREADDS could be genetically targeted to specific neurons and activated with orally administered, inert pharmaceuticals [227], [228]. When engineering therapeutics that modulate downstream signaling pathways, such as opsins or DREADDS, the differences in ion conductance in human neurons should be considered [229]. Opto-MASS uses a mammalian host system based on an immortalized human cell line, being slightly closer on the evolutionary tree to non-human primate experimental models than rats and mice. This is beneficial for validating protein folding and trafficking to cell surface membranes, and codon frequency considerations in non-human primate experiments. Generally, as mammal body size increases, neurons increase in size as well. The greater surface area of larger neurons will enable more molecules of sensor to occupy the membrane in membrane bound receptors, potentially increasing signal. However, to image fluorescent tools occupied in lower cortical layers may be difficult, as multiphoton imaging has limitations on imaging depth that may be exceeded by the larger tissues presented in non-human primate models.

Opto-MASS in its current form presents a step forward in the development of next-generation optogenetic tools. Opto-MASS rapidly assesses genetic libraries in mammalian cells, and directly links phenotype to genotype. This is unique to many high-throughput platforms. Initial screening in mammalian cells surpasses a large hurdle in the validation of genetic tools for the mammalian nervous system. Opto-MASS's current state will yield useful tools, however improvement of data collection with *in situ* barcode sequencing, enhanced gene recovery, and novel stimulation techniques will broaden the utility of the approach. As researchers implement additional techniques with Opto-MASS, it will continue to push the frontier of optogenetic protein engineering.

BIBLIOGRAPHY

[1] E. S. Boyden, F. Zhang, E. Bamberg, G. Nagel, and K. Deisseroth, "Millisecond-timescale, genetically targeted optical control of neural activity.," *Nat Neurosci*, vol. 8, no. 9, pp. 1263–1268, Sep. 2005, doi: 10.1038/nn1525.

[2] L. Tian *et al.*, "Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators," *Nat Methods*, vol. 6, no. 12, pp. 875–881, 2009, doi: 10.1038/nmeth.1398.

[3] G. Nagel *et al.*, "Channelrhodopsin-2, a directly light-gated cation-selective membrane channel," *Proc Natl Acad Sci U S A*, vol. 100, no. 24, pp. 13940–13945, Nov. 2003, doi: 10.1073/pnas.1936192100.

[4] N. C. Klapoetke *et al.*, "Independent optical excitation of distinct neural populations," *Nat Methods*, vol. 11, no. 3, pp. 338–346, 2014, doi: 10.1038/nmeth.2836.

[5] A. Berndt, "Structure-Guided Transformation, 420. doi:10.5061/dryad.9r0p6," vol. 420, 2014, doi: 10.5061/dryad.9r0p6.

[6] E. G. Govorunova, O. A. Sineshchekov, R. Janz, X. Liu, and J. L. Spudich, "Natural light-gated anion channels: A family of microbial rhodopsins for advanced optogenetics," *Science (1979)*, vol. 349, no. 6248, pp. 647–650, 2015, doi: 10.1126/science.aaa7484.

[7] E. G. Govorunova *et al.*, "Kalium channelrhodopsins are natural light-gated potassium channels that mediate optogenetic inhibition," *Nat Neurosci*, Jul. 2022, doi: 10.1038/s41593-022-01094-6.

[8] J. Wan, W. Peng, X. Li, T. Qian, K. Song, and J. Zeng, "A genetically encoded GRAB sensor for measuring serotonin dynamics in vivo," 2020, doi: 10.1101/2020.02.24.962282.

[9] T. Patriarchi *et al.*, "Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors," *Science (1979)*, vol. 360, no. 6396, Jun. 2018, doi: 10.1126/science.aat4422.

[10] T. W. Chen *et al.*, "Ultrasensitive fluorescent proteins for imaging neuronal activity," *Nature*, vol. 499, no. 7458, pp. 295–300, 2013, doi: 10.1038/nature12354.

[11] J. Nakai, M. Ohkura, and K. Imoto, "A high signal-to-noise Ca2+ probe composed of a single green fluorescent protein," *Nat Biotechnol*, vol. 19, no. 2, pp. 137–141, Feb. 2001, doi: 10.1038/84397.

[12] T. Nagai, A. Sawano, E. S. Park, and A. Miyawaki, "Circularly permuted green fluorescent proteins engineered to sense Ca $^{2+}$," *Proceedings of the National Academy of Sciences*, vol. 98, no. 6, pp. 3197–3202, Mar. 2001, doi: 10.1073/pnas.051636098.

[13] G. S. Baird, D. A. Zacharias, and R. Y. Tsien, "Circular permutation and receptor insertion within green fluorescent proteins," *Proc Natl Acad Sci U S A*, vol. 96, no. 20, pp. 11241–11246, 1999, doi: 10.1073/pnas.96.20.11241.

[14] B. L. Sabatini, T. G. , Oertner, and K. Svoboda, "The Life Cycle of Calcium Ions in Dendritic Spines," *Neuron*, vol. 33, pp. 439–452, Jan. 2002.

[15] R. D. Burgoyne, "Neuronal calcium sensor proteins: Generating diversity in neuronal Ca 2+ signalling," *Nature Reviews Neuroscience*, vol. 8, no. 3. pp. 182–193, Mar. 2007. doi: 10.1038/nrn2093.

[16] M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher, “Green fluorescent protein as a marker for gene expression,” *Science (1979)*, vol. 263, no. 5148, pp. 802–805, 1994, doi: 10.1126/science.8303295.

[17] D. C. Prasher, V. K. Eckenrode, W. W. Ward, F. G. Prendergast, and M. J. Cormier, “Primary structure of the Aequorea victoria green-fluorescent protein,” 1992.

[18] M. Ormö, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, and S. J. Remington, “Crystal Structure of the Aequorea victoria Green Fluorescent Protein,” *Science (1979)*, vol. 273, no. 5280, pp. 1392–1395, 1996, doi: 10.1126/science.273.5280.1392.

[19] J. W. Wang, A. M. Wong, J. Flores, L. B. Vosshall, and R. Axel, “Two-photon calcium imaging reveals an odor-evoked map of activity in the fly brain,” *Cell*, vol. 112, no. 2, pp. 271–282, 2003, doi: 10.1016/S0092-8674(03)00004-7.

[20] L. Tian *et al.*, “Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators,” *Nat Methods*, vol. 6, no. 12, pp. 875–881, 2009, doi: 10.1038/nmeth.1398.

[21] T.-W. Chen *et al.*, “Ultrasensitive fluorescent proteins for imaging neuronal activity,” *Nature*, vol. 499, no. 7458, pp. 295–300, 2013, doi: 10.1038/nature12354.

[22] J. Vierock, C. Grimm, N. Nitzan, and P. Hegemann, “Molecular determinants of proton selectivity and gating in the red-light activated channelrhodopsin Chrimson,” *Sci Rep*, vol. 7, no. 1, pp. 1–15, 2017, doi: 10.1038/s41598-017-09600-8.

[23] K. Bera *et al.*, “Biosensors Show the Pharmacokinetics of S-Ketamine in the Endoplasmic Reticulum,” *Front Cell Neurosci*, vol. 13, no. 6, 2019, doi: 10.3389/fncel.2019.00499.

[24] J. S. Marvin *et al.*, “Stability, affinity, and chromatic variants of the glutamate sensor iGluSnFR,” *Nat Methods*, vol. 15, no. 11, pp. 936–939, 2018, doi: 10.1038/s41592-018-0171-3.

[25] J. S. Marvin *et al.*, “A genetically encoded fluorescent sensor for in vivo imaging of GABA,” *Nat Methods*, vol. 16, no. 8, pp. 763–770, 2019, doi: 10.1038/s41592-019-0471-2.

[26] C. R. Hackley, E. O. Mazzoni, and J. Blau, “cAMPr: A single-wavelength fluorescent sensor for cyclic AMP,” *Sci Signal*, vol. 11, no. 520, 2018, doi: 10.1126/scisignal.aah3738.

[27] E. K. Unger *et al.*, “Directed Evolution of a Selective and Sensitive Serotonin Sensor via Machine Learning,” *Cell*, vol. 183, no. 7, pp. 1986-2002.e26, 2020, doi: 10.1016/j.cell.2020.11.040.

[28] M. Jing *et al.*, “A genetically encoded fluorescent acetylcholine indicator for in vitro and in vivo studies,” *Nat Biotechnol*, vol. 36, no. 8, pp. 726–737, 2018, doi: 10.1038/nbt.4184.

[29] J. Feng *et al.*, “A Genetically Encoded Fluorescent Sensor for Rapid and Specific In Vivo Detection of Norepinephrine,” *Neuron*, vol. 102, no. 4, pp. 745-761.e8, 2019, doi: 10.1016/j.neuron.2019.02.037.

[30] J. C. Klima *et al.*, “Incorporation of sensing modalities into de novo designed fluorescence-activating proteins,” *Nat Commun*, vol. 12, no. 1, Dec. 2021, doi: 10.1038/s41467-020-18911-w.

[31] R. Weissleder, “A clearer vision for in vivo imaging,” *Nat Biotechnol*, pp. 316–317, 2001.

[32] J. Akerboom *et al.*, “Genetically encoded calcium indicators for multi-color neural activity imaging and combination with optogenetics,” *Front Mol Neurosci*, no. FEB, Feb. 2013, doi: 10.3389/fnmol.2013.00002.

[33] F. Sun *et al.*, “Next-generation GRAB sensors for monitoring dopaminergic activity in vivo,” *Nat Methods*, vol. 17, no. 11, pp. 1156–1166, Nov. 2020, doi: 10.1038/s41592-020-00981-9.

[34] J. J. Ding, A. F. Luo, L. Y. Hu, D. C. Wang, and F. Shao, “Structural basis of the ultrasensitive calcium indicator GCaMP6,” *Sci China Life Sci*, vol. 57, no. 3, pp. 269–274, 2014, doi: 10.1007/s11427-013-4599-5.

[35] J. D. Pédelacq, S. Cabantous, T. Tran, T. C. Terwilliger, and G. S. Waldo, “Engineering and characterization of a superfolder green fluorescent protein,” *Nat Biotechnol*, vol. 24, no. 1, pp. 79–88, 2006, doi: 10.1038/nbt1172.

[36] B. T. Bajar *et al.*, “Improving brightness and photostability of green and red fluorescent proteins for live cell imaging and FRET reporting,” *Sci Rep*, vol. 6, no. 1, p. 20889, 2016, doi: 10.1038/srep20889.

[37] T. D. Craggs, “Green fluorescent protein: Structure, folding and chromophore maturation,” *Chem Soc Rev*, vol. 38, no. 10, pp. 2865–2875, 2009, doi: 10.1039/b903641p.

[38] J. A. Kaczmarski, J. A. Mitchell, M. A. Spence, V. Vongsouthi, and C. J. Jackson, “Structural and evolutionary approaches to the design and optimization of fluorescence-based small molecule biosensors,” *Curr Opin Struct Biol*, vol. 57, pp. 31–38, 2019, doi: 10.1016/j.sbi.2019.01.013.

[39] Y. Nasu, Y. Shen, L. Kramer, and R. E. Campbell, “Structure- and mechanism-guided design of single fluorescent protein-based biosensors,” *Nat Chem Biol*, vol. 17, no. May, 2021, doi: 10.1038/s41589-020-00718-x.

[40] M. A. Labouesse and T. Patriarchi, “A versatile GPCR toolkit to track in vivo neuromodulation : not a one-size- fi ts-all sensor,” *Neuropsychopharmacology*, no. October 2020, pp. 1–5, 2021, doi: 10.1038/s41386-021-00982-y.

[41] K. Harada *et al.*, “Red fluorescent protein-based cAMP indicator applicable to optogenetics and in vivo imaging,” *Sci Rep*, vol. 7, no. 1, pp. 1–9, 2017, doi: 10.1038/s41598-017-07820-6.

[42] D. Dimitrov *et al.*, “Engineering and characterization of an enhanced fluorescent protein voltage sensor,” *PLoS One*, vol. 2, no. 5, pp. 2–6, 2007, doi: 10.1371/journal.pone.0000440.

[43] A. Jung, J. E. Garcia, E. Kim, B.-J. Yoon, and B. J. Baker, “Linker length and fusion site composition improve the optical signal of genetically encoded fluorescent voltage sensors,” *Neurophotonics*, vol. 2, no. 2, p. 021012, 2015, doi: 10.1117/1.NPh.2.2.021012.

[44] F. St-Pierre, J. D. Marshall, Y. Yang, Y. Gong, M. J. Schnitzer, and M. Z. Lin, “High-fidelity optical reporting of neuronal electrical activity with an ultrafast fluorescent voltage sensor,” *Nat Neurosci*, vol. 17, no. 6, pp. 884–889, 2014, doi: 10.1038/nn.3709.

[45] Y. Gong *et al.*, “High-speed recording of neural spikes in awake mice and flies with a fluorescent voltage sensor,” *Science (1979)*, vol. 350, no. 6266, pp. 1361–1366, 2015, doi: 10.1126/science.aab0810.

[46] K. D. Piatkevich *et al.*, “A robotic multidimensional directed evolution approach applied to fluorescent voltage reporters,” *Nat Chem Biol*, vol. 14, no. April, p. 1, 2018, doi: 10.1038/s41589-018-0023-6.

[47] J. S. Marvin *et al.*, “An optimized fluorescent probe for visualizing glutamate neurotransmission,” *Nat Methods*, vol. 10, no. 2, pp. 162–170, 2013, doi: 10.1038/nmeth.2333.

[48] Y. Yang *et al.*, “Improved calcium sensor GCaMP-X overcomes the calcium channel perturbations induced by the calmodulin in GCaMP,” *Nat Commun*, vol. 9, no. 1, 2018, doi: 10.1038/s41467-018-03719-6.

[49] J. Akerboom *et al.*, “Optimization of a GCaMP Calcium Indicator for Neural Activity Imaging,” *Journal of Neuroscience*, vol. 32, no. 40, pp. 13819–13840, 2012, doi: 10.1523/JNEUROSCI.2601-12.2012.

[50] F. Sun *et al.*, “A genetically-encoded fluorescent sensor enables rapid and specific detection of dopamine in flies, fish, and mice,” *Cell*, vol. 174, no. 2, pp. 481–496, 2018, doi: 10.1101/332528.

[51] Y. Ohta, T. Furuta, T. Nagai, and K. Horikawa, “Red fluorescent cAMP indicator with increased affinity and expanded dynamic range,” *Sci Rep*, no. November 2017, pp. 1–9, 2018, doi: 10.1038/s41598-018-20251-1.

[52] Y. Yu and S. Lutz, “Circular permutation: A different way to engineer enzyme structure and function,” *Trends Biotechnol*, vol. 29, no. 1, pp. 18–25, 2011, doi: 10.1016/j.tibtech.2010.10.004.

[53] S. J. Remington, “Green fluorescent protein: A perspective,” *Protein Science*, vol. 20, no. 9, pp. 1509–1519, 2011, doi: 10.1002/pro.684.

[54] R. Y. Tsien, “the Green Fluorescent Protein,” *Annu Rev Biochem*, vol. 67, no. 1, pp. 509–544, 1998, doi: 10.1146/annurev.biochem.67.1.509.

[55] J. Akerboom *et al.*, “Crystal structures of the GCaMP calcium sensor reveal the mechanism of fluorescence signal change and aid rational design,” *Journal of Biological Chemistry*, vol. 284, no. 10, pp. 6455–6464, 2009, doi: 10.1074/jbc.M807657200.

[56] L. Ravotto, L. Duffet, X. Zhou, B. Weber, T. Patriarchi, and E. C. Carroll, “A Bright and Colorful Future for G-Protein Coupled Receptor Sensors,” *Front Cell Neurosci*, vol. 14, no. March, pp. 1–9, 2020, doi: 10.3389/fncel.2020.00067.

[57] T. J. Wardill *et al.*, “A Neuron-Based Screening Platform for Optimizing Genetically-Encoded Calcium Indicators,” *PLoS One*, vol. 8, no. 10, pp. 1–12, 2013, doi: 10.1371/journal.pone.0077728.

[58] Y. N. Tallini *et al.*, “Imaging cellular signals in the heart in vivo: Cardiac expression of the high-signal Ca 2 indicator GCaMP2,” 2006. [Online]. Available: www.pnas.orgcgidoi10.1073pnas.0509378103

[59] J. Nakai, M. Ohkura, and K. Imoto, “A high signal-to-noise ca2+ probe composed of a single green fluorescent protein,” *Nat Biotechnol*, vol. 19, no. 2, pp. 137–141, 2001, doi: 10.1038/84397.

[60] H. Dana *et al.*, “High-performance calcium sensors for imaging activity in neuronal populations and microcompartments,” *Nat Methods*, vol. 16, no. 7, pp. 649–657, Jul. 2019, doi: 10.1038/s41592-019-0435-6.

[61] E. K. Unger *et al.*, “Directed Evolution of a Selective and Sensitive Serotonin Sensor via Machine Learning,” *Cell*, vol. 183, no. 7, pp. 1986-2002.e26, Dec. 2020, doi: 10.1016/j.cell.2020.11.040.

[62] A. Rivera-Calzada and M. Carroni, “Editorial: Technical Advances in Cryo-Electron Microscopy,” *Frontiers in Molecular Biosciences*, vol. 6. Frontiers Media S.A., Aug. 22, 2019. doi: 10.3389/fmolb.2019.00072.

[63] Y. Cheng, “Membrane protein structural biology in the era of single particle cryo-EM,” *Current Opinion in Structural Biology*, vol. 52. Elsevier Ltd, pp. 58–63, Oct. 01, 2018. doi: 10.1016/j.sbi.2018.08.008.

[64] A. Koehl *et al.*, “Structure of the μ-opioid receptor-Gi protein complex,” *Nature*, vol. 558, no. 7711, pp. 547–552, 2018, doi: 10.1038/s41586-018-0219-7.
[65] J. Jumper *et al.*, “Highly accurate protein structure prediction with AlphaFold,” *Nature*, vol. 596, no. 7873, pp. 583–589, Aug. 2021, doi: 10.1038/s41586-021-03819-2.
[66] C. Bayly-Jones and J. C. Whisstock, “Mining folded proteomes in the era of accurate structure prediction,” *PLoS Comput Biol*, vol. 18, no. 3, Mar. 2022, doi: 10.1371/journal.pcbi.1009930.
[67] C. Dong *et al.*, “Psychedelic-inspired drug discovery using an engineered biosensor,” *Cell*, vol. 184, no. 10, pp. 2779-2792.e18, 2021, doi: 10.1016/j.cell.2021.03.043.
[68] L. Duffet *et al.*, “A genetically encoded sensor for in vivo imaging of orexin neuropeptides,” *Nat Methods*, vol. 19, no. 2, pp. 231–241, Feb. 2022, doi: 10.1038/s41592-021-01390-2.
[69] E. C. McCusker, S. E. Bane, M. A. O’Malley, and A. S. Robinson, “Heterologous GPCR expression: A bottleneck to obtaining crystal structures,” *Biotechnol Prog*, vol. 23, no. 3, pp. 540–547, 2007, doi: 10.1021/bp060349b.
[70] K. A. Matreyek, J. J. Stephany, and D. M. Fowler, “A platform for functional assessment of large variant libraries in mammalian cells,” *Nucleic Acids Res*, vol. 45, no. 11, pp. 1–12, 2017, doi: 10.1093/nar/gkx183.
[71] M. F. Roitman, G. D. Stuber, P. E. M. Phillips, R. M. Wightman, and R. M. Carelli, “Dopamine Operates as a Subsecond Modulator of Food Seeking,” *Journal of Neuroscience*, vol. 24, no. 6, pp. 1265–1271, 2004, doi: 10.1523/JNEUROSCI.3823-03.2004.
[72] R. D. Palmiter, “Dopamine signaling in the dorsal striatum is essential for motivated behaviors: Lessons from dopamine-deficient mice,” *Ann N Y Acad Sci*, vol. 1129, pp. 35–46, 2008, doi: 10.1196/annals.1417.003.
[73] M. W. Howe and D. A. Dombeck, “Rapid signalling in distinct dopaminergic axons during locomotion and reward,” *Nature*, vol. 535, no. 7613, pp. 505–510, 2016, doi: 10.1038/nature18942.
[74] T. N. Lerner *et al.*, “Intact-Brain Analyses Reveal Distinct Information Carried by SNc Dopamine Subcircuits,” *Cell*, vol. 162, no. 3, pp. 635–647, 2015, doi: 10.1016/j.cell.2015.07.014.
[75] W. Schultz, “Dopamine reward prediction-error signalling: A two-component response,” *Nat Rev Neurosci*, vol. 17, no. 3, pp. 183–195, 2016, doi: 10.1038/nrn.2015.26.
[76] H. Tian *et al.*, “All-optical electrophysiology with improved genetically encoded voltage indicators reveals interneuron network dynamics in vivo,” *BioRxiV*, doi: 10.1101/2021.11.22.469481.
[77] M. J. Brownstein, “A brief history of opiates, opioid peptides, and opioid receptors.,” *Proceedings of the National Academy of Sciences*, vol. 90, no. 12, pp. 5391–5393, Jun. 1993, doi: 10.1073/pnas.90.12.5391.
[78] S. Granier *et al.*, “Structure of the δ-opioid receptor bound to naltrindole,” *Nature*, vol. 485, no. 7398, pp. 400–404, 2012, doi: 10.1038/nature11111.
[79] J. P. Higham, C. S. Barr, C. L. Hoffman, T. M. Mandalaywala, K. J. Parker, and D. Maestripieri, “Mu-opioid Receptor (OPRM1) Variation, Oxytocin Levels and Maternal Attachment in Free-Ranging Rhesus Macaques Macaca mulatta,” *Behavioral Neuroscience*, vol. 125, no. 2, pp. 131–136, Apr. 2011, doi: 10.1037/a0022695.

[80] G. B. Stefano and R. Kream, “Endogenous opiates, opioids, and immune function: Evolutionary brokerage of defensive behaviors,” *Seminars in Cancer Biology*, vol. 18, no. 3. pp. 190–198, Jun. 2008. doi: 10.1016/j.semcancer.2007.12.001.

[81] R. J. Bodnar, “Endogenous opioids and feeding behavior: A 30-year historical perspective,” *Peptides*, vol. 25, no. 4. pp. 697–725, Apr. 2004. doi: 10.1016/j.peptides.2004.01.006.

[82] S. C. Ribeiro, S. E. Kennedy, Y. R. Smith, C. S. Stohler, and J. K. Zubieta, “Interface of physical and emotional stress regulation through the endogenous opioid system and μ-opioid receptors,” *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 29, no. 8. pp. 1264–1280, Dec. 2005. doi: 10.1016/j.pnpbp.2005.08.011.

[83] G. Drolet, É. C. Dumont, I. Gosselin, R. Kinkead, S. Laforest, and J.-F. Trottier, “Role of endogenous opioid system in the regulation of the stress response,” *Prog Neuropsychopharmacol Biol Psychiatry*, vol. 25, no. 4, pp. 729–741, May 2001, doi: 10.1016/S0278-5846(01)00161-0.

[84] P. Holzer, “Opioid receptors in the gastrointestinal tract,” *Regulatory Peptides*, vol. 155, no. 1–3. pp. 11–17, Jun. 05, 2009. doi: 10.1016/j.regpep.2009.03.012.

[85] P. S. Puttfarcken, L. L. Werling, and B. M. Cox, “Effects of chronic morphine exposure on opioid inhibition of adenylyl cyclase in 7315c cell membranes: a useful model for the study of tolerance at mu opioid receptors.,” *Mol Pharmacol*, vol. 33, no. 5, pp. 520–7, May 1988, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/2835651

[86] Y. Chen, A. Mestek, J. Liu, A. Hurley, and L. Yu, “Molecular cloning and functional expression of a mu opioid receptor from rat brain,” *Mol Pharmacol*, vol. 44, pp. 8–12, 1993, doi: 10.1016/0167-0115(94)90214-3.

[87] R. C. Thompson, A. Mansour, H. Akil, and S. J. Watson, “Cloning and pharmacological characterization of a rat μ opioid receptor,” *Neuron*, vol. 11, no. 5, pp. 903–913, 1993, doi: 10.1016/0896-6273(93)90120-G.

[88] J. B. Wang, Y. Imai, C. M. Eppler, P. Gregor, C. E. Spivak, and G. R. Uhl, “μ opiate receptor: cDNA cloning and expression,” *Proc Natl Acad Sci U S A*, vol. 90, no. 21, pp. 10230–10234, 1993, doi: 10.1073/pnas.90.21.10230.

[89] H. W. D. Matthes *et al.*, “Loss of morphine-induced analgesia, reward effect and withdrawal symptoms in mice lacking the mu opioid-receptor gene,” *Nature*, vol. 383, pp. 819–823, 1996, doi: https://doi.org/10.1038/383819a0.

[90] R. Al-Hasani *et al.*, “In vivo detection of optically-evoked opioid peptide release,” *Elife*, vol. 7, pp. 1–13, 2018, doi: 10.7554/eLife.36520.

[91] M. Stoeber *et al.*, “A genetically encoded biosensor reveals location bias of opioid drug action,” pp. 1–27, 2018.

[92] K. E. Kroning, M. Li, D. I. Petrescu, and W. Wang, “A genetically encoded sensor with improved fluorescence intensity for opioid detection at cellular resolution,” *Chemical Communications*, vol. 57, no. 81, pp. 10560–10563, 2021, doi: 10.1039/d1cc04524e.

[93] K. E. Kroning and W. Wang, “Designing a Single Protein-Chain Reporter for Opioid Detection at Cellular Resolution,” *Angewandte Chemie - International Edition*, vol. 60, no. 24, pp. 13358–13365, Jun. 2021, doi: 10.1002/anie.202101262.

[94] K. E. Kroning, M. Li, D. I. Petrescu, and W. Wang, “A genetically encoded sensor with improved fluorescence intensity for opioid detection at cellular resolution,” *Chemical Communications*, vol. 57, no. 81, pp. 10560–10563, Oct. 2021, doi: 10.1039/d1cc04524e.

[95] G. Buzsáki, C. Geisler, D. A. Henze, and X. J. Wang, “Interneuron Diversity series: Circuit complexity and axon wiring economy of cortical interneurons,” *Trends Neurosci*, vol. 27, no. 4, pp. 186–193, 2004, doi: 10.1016/j.tins.2004.02.007.

[96] D. C. Castro *et al.*, “An endogenous opioid circuit determines state-dependent reward consumption,” *Nature*, vol. 598, no. 7882, pp. 646–651, 2021, doi: 10.1038/s41586-021-04013-0.

[97] H. A. Shippenberg TS, “Differential effects of mu and kappa opioid systems on motivational processes.,” *NIDA Res Monogr.*, no. 75, pp. 563–66..

[98] B. B. Land, M. R. Bruchas, J. C. Lemos, M. Xu, E. J. Melief, and C. Chavkin, “The dysphoric component of stress is encoded by activation of the dynorphin κ-opioid system,” *Journal of Neuroscience*, vol. 28, no. 2, pp. 407–414, 2008, doi: 10.1523/JNEUROSCI.4458-07.2008.

[99] H. W. D. Matthes *et al.*, “Loss of morphine-induced analgesia, reward effect and withdrawal symptoms in mice lacking the μ-opioid-receptor gene,” *Nature*, vol. 383, no. 6603, pp. 822–823, 1996, doi: 10.1038/383819a0.

[100] L. D. Fricker, E. B. Margolis, I. Gomes, and L. A. Devi, “Five Decades of Research on Opioid Peptides: Current Knowledge and Unanswered Questions,” *Mol Pharmacol*, vol. 98, no. 2, pp. 96–108, 2020, doi: 10.1124/mol.120.119388.

[101] M. Stoeber *et al.*, “A Genetically Encoded Biosensor Reveals Location Bias of Opioid Drug Action,” *Neuron*, vol. 98, no. 5, pp. 963-976.e5, 2018, doi: 10.1016/j.neuron.2018.04.021.

[102] V. Gradinaru *et al.*, “Molecular and Cellular Approaches for Diversifying and Extending Optogenetics,” *Cell*, vol. 141, no. 1, pp. 154–165, 2010, doi: 10.1016/j.cell.2010.02.037.

[103] A. Manglik *et al.*, “Crystal structure of the μ-opioid receptor bound to a morphinan antagonist,” *Nature*, vol. 485, no. 7398, pp. 321–326, 2012, doi: 10.1038/nature10954.

[104] W. Huang *et al.*, “Structural insights into μ-opioid receptor activation,” *Nature*, vol. 524, no. 7565, pp. 315–321, 2015, doi: 10.1038/nature14886.

[105] J. M. Richard and H. L. Fields, “Mu-opioid receptor activation in the medial shell of nucleus accumbens promotes alcohol consumption, self-administration and cue-induced reinstatement,” *Neuropharmacology*, vol. 108, pp. 14–23, 2016, doi: 10.1016/j.neuropharm.2016.04.010.

[106] S. L. Resendez *et al.*, “μ-Opioid receptors within subregions of the striatum mediate pair bond formation through parallel yet distinct reward mechanisms,” *Journal of Neuroscience*, vol. 33, no. 21, pp. 9140–9149, 2013, doi: 10.1523/JNEUROSCI.4123-12.2013.

[107] V. P. Bakshi and A. E. Kelley, “Feeding induced by opioid stimulation of the ventral striatum: Role of opiate receptor subtypes,” *Journal of Pharmacology and Experimental Therapeutics*, vol. 265, no. 3, pp. 1253–1260, 1993.

[108] A. G. Difeliceantonio, O. S. Mabrouk, R. T. Kennedy, and K. C. Berridge, “Enkephalin surges in dorsal neostriatum as a signal to eat,” *Current Biology*, vol. 22, no. 20, pp. 1918–1924, 2012, doi: 10.1016/j.cub.2012.08.014.

[109] S. J. Smith *et al.*, “Single-cell transcriptomic evidence for dense intracortical neuropeptide networks,” *Elife*, vol. 8, no. 1, pp. 1–35, 2019, doi: 10.7554/eLife.47889.

[110] S. C. Baraban and M. K. Tallent, “Interneuron Diversity series: Interneuronal neuropeptides - Endogenous regulators of neuronal excitability,” *Trends Neurosci*, vol. 27, no. 3, pp. 135–142, 2004, doi: 10.1016/j.tins.2004.01.008.

[111] M. Ludwig and G. Leng, “Dendritic peptide release and peptide-dependent behaviours,” *Nature Reviews Neuroscience*, vol. 7, no. 2. pp. 126–136, Feb. 2006. doi: 10.1038/nrn1845.
[112] A. T. Ehrlich *et al.*, “Biased Signaling of the Mu Opioid Receptor Revealed in Native Neurons,” *iScience*, vol. 14, pp. 47–57, Apr. 2019, doi: 10.1016/j.isci.2019.03.011.
[113] T. Kenakin, “New concepts in pharmacological efficacy at 7TM receptors: IUPHAR Review 2,” *British Journal of Pharmacology*, vol. 168, no. 3. pp. 554–575, Feb. 2013. doi: 10.1111/j.1476-5381.2012.02223.x.
[114] J. de Neve, T. M. A. Barlow, D. Tourwé, F. Bihel, F. Simonin, and S. Ballet, “Comprehensive overview of biased pharmacology at the opioid receptors: biased ligands and bias factors,” *RSC Med Chem*, vol. 12, no. 6, pp. 828–870, 2021, doi: 10.1039/d1md00041a.
[115] J. J. Liu, R. Horst, V. Katritch, R. C. Stevens, and K. Wüthrich, “Biased signaling pathways in β2-adrenergic receptor characterized by 19F-NMR.,” *Science*, vol. 335, no. 6072, pp. 1106–10, Mar. 2012, doi: 10.1126/science.1215802.
[116] C. Saleeba, B. Dempsey, S. Le, A. Goodchild, and S. McMullan, “A student’s guide to neural circuit tracing,” *Frontiers in Neuroscience*, vol. 13, no. AUG. Frontiers Media S.A., 2019. doi: 10.3389/fnins.2019.00897.
[117] X. Xu *et al.*, “Viral Vectors for Neural Circuit Mapping and Recent Advances in Trans-synaptic Anterograde Tracers,” *Neuron*, vol. 107, no. 6. Cell Press, pp. 1029–1047, Sep. 23, 2020. doi: 10.1016/j.neuron.2020.07.010.
[118] C. Mollereau and L. Mouledous, “Tissue distribution of the opioid receptor-like (ORL1) receptor,” *Peptides (N.Y.)*, vol. 21, no. 7, pp. 907–917, Jul. 2000, doi: 10.1016/S0196-9781(00)00227-8.
[119] K. Yasuda *et al.*, “Cloning and functional comparison of kappa and delta opioid receptors from mouse brain.,” *Proceedings of the National Academy of Sciences*, vol. 90, no. 14, pp. 6736–6740, Jul. 1993, doi: 10.1073/pnas.90.14.6736.
[120] C. J. Evans, D. E. Keith, H. Morrison, K. Magendzo, and R. H. Edwards, “Cloning of a Delta Opioid Receptor by Functional Expression,” *Science (1979)*, vol. 258, no. 5090, pp. 1952–1955, Dec. 1992, doi: 10.1126/science.1335167.
[121] B. L. Kieffer, K. Befort, C. Gavériaux-Ruff, and C. G. Hirth, “The delta-opioid receptor: isolation of a cDNA by expression cloning and pharmacological characterization.,” *Proc Natl Acad Sci U S A*, vol. 91, no. 3, p. 1193, Feb. 1994, doi: 10.1073/pnas.91.3.1193-c.
[122] C. J. Evans, D. E. Keith, H. Morrison, K. Magendzo, and R. H. Edwards, “Cloning of a Delta Opioid Receptor by Functional Expression,” *Science (1979)*, vol. 258, no. 5090, pp. 1952–1955, Dec. 1992, doi: 10.1126/science.1335167.
[123] J. Hughes, T. W. Smith, H. W. Kosterlitz, L. A. Fothergill, B. A. Morgan, and H. R. Morris, “Identification of two related pentapeptides from the brain with potent opiate agonist activity,” *Nature*, vol. 258, no. 5536, pp. 577–579, 1975, doi: 10.1038/258577a0.
[124] C. H. LI and D. CHUNG, “Primary structure of human β-lipotropin,” *Nature*, vol. 260, no. 5552, pp. 622–624, Apr. 1976, doi: 10.1038/260622a0.
[125] L. H. Lazarus, N. Ling, and R. Guillemin, “beta-Lipotropin as a prohormone for the morphinomimetic peptides endorphins and enkephalins.,” *Proceedings of the National Academy of Sciences*, vol. 73, no. 6, pp. 2156–2159, Jun. 1976, doi: 10.1073/pnas.73.6.2156.

[126] A. Goldstein, S. Tachibana, L. I. Lowney, M. Hunkapiller, and L. Hood, “Dynorphin-(1-13), an extraordinarily potent opioid peptide,” *Proc Natl Acad Sci U S A*, vol. 76, no. 12, pp. 6666–6670, 1979, doi: 10.1073/pnas.76.12.6666.

[127] J. le Merrer, J. A. J. Becker, K. Befort, and B. L. Kieffer, “Reward Processing by the Opioid System in the Brain,” *Physiol Rev*, vol. 89, no. 4, pp. 1379–1412, Oct. 2009, doi: 10.1152/physrev.00005.2009.

[128] U. Arvidsson *et al.*, “The kappa-opioid receptor is primarily postsynaptic: combined immunohistochemical localization of the receptor and endogenous opioids.,” *Proceedings of the National Academy of Sciences*, vol. 92, no. 11, pp. 5062–5066, May 1995, doi: 10.1073/pnas.92.11.5062.

[129] M. D. Hayward, A. Schaich-Borg, J. E. Pintar, and M. J. Low, “Differential involvement of endogenous opioids in sucrose consumption and food reinforcement,” *Pharmacol Biochem Behav*, vol. 85, no. 3, pp. 601–611, Nov. 2006, doi: 10.1016/j.pbb.2006.10.015.

[130] I. Racz *et al.*, “The Opioid Peptides Enkephalin and β-Endorphin in Alcohol Dependence,” *Biol Psychiatry*, vol. 64, no. 11, pp. 989–997, Dec. 2008, doi: 10.1016/j.biopsych.2008.05.008.

[131] C. Chavkin and A. Goldstein, “Specific receptor for the opioid peptide dynorphin: structure--activity relationships.,” *Proceedings of the National Academy of Sciences*, vol. 78, no. 10, pp. 6543–6547, Oct. 1981, doi: 10.1073/pnas.78.10.6543.

[132] M. Filizola and L. A. , Devi, “How opioid drugs bind to receptors,” *Nature*, vol. 485, pp. 314–317, 2012.

[133] H. Wu *et al.*, “Structure of the human κ-opioid receptor in complex with JDTic,” *Nature*, vol. 485, no. 7398, pp. 327–332, May 2012, doi: 10.1038/nature10939.

[134] A. A. Thompson *et al.*, “Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic,” *Nature*, vol. 485, no. 7398, pp. 395–399, 2012, doi: 10.1038/nature11085.

[135] M. Filizola and L. A. Devi, “Grand opening of structure-guided design for novel opioids,” *Trends Pharmacol Sci*, vol. 34, no. 1, pp. 6–12, 2013, doi: 10.1016/j.tips.2012.10.002.

[136] T. Flock, A. S. Hauser, N. Lund, D. E. Gloriam, S. Balaji, and M. M. Babu, “Selectivity determinants of GPCR-G-protein binding,” *Nature*, vol. 545, no. 7654, pp. 317–322, 2017, doi: 10.1038/nature22070.

[137] A. J. Venkatakrishnan, X. Deupi, G. Lebon, C. G. Tate, G. F. Schertler, and M. Madan Babu, “Molecular signatures of G-protein-coupled receptors,” *Nature*, vol. 494, no. 7436, pp. 185–194, 2013, doi: 10.1038/nature11896.

[138] T. Onogi *et al.*, “DAMGO, a μ-opioid receptor selective agonist, distinguishes between μ- and δ-opioid receptors around their first extracellular loops,” *FEBS Lett*, vol. 357, no. 1, pp. 93–97, 1995, doi: 10.1016/0014-5793(94)01341-W.

[139] M. Minami *et al.*, “A single residue, Lys108, of the delta-opioid receptor prevents the mu-opioid-selective ligand [D-Ala2,N-MePhe4,Gly-ol5]enkephalin from binding to the delta-opioid receptor.,” *Mol Pharmacol*, vol. 50, no. 5, pp. 1413–22, Nov. 1996, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/8913373

[140] M. Minami *et al.*, “DAMGO, a μ-opioid receptor selective ligand, distinguishes between μ-and κ-opioid receptors at a different region from that for the distinction between μ- and δ-opioid receptors,” *FEBS Lett*, vol. 364, no. 1, pp. 23–27, 1995, doi: 10.1016/0014-5793(95)00340-F.

[141] T. G. Metzger and D. M. Ferguson, “On the role of extracellular loops of opioid receptors in conferring ligand selectivity,” *FEBS Lett*, vol. 375, no. 1–2, pp. 1–4, 1995, doi: 10.1016/0014-5793(95)01185-H.
[142] T. Seki, M. Minami, T. Nakagawa, Y. Ienaga, A. Morisada, and M. Satoh, “DAMGO recognizes four residues in the third extracellular loop to discriminate between μ- and κ-opioid receptors,” *Eur J Pharmacol*, vol. 350, no. 2–3, pp. 301–310, 1998, doi: 10.1016/S0014-2999(98)00240-4.
[143] Z. Zhao, T. Huang, and J. Li, “Molecular dynamics simulations to investigate how pzm21 affects the conformational state of the μ-opioid receptor upon activation,” *Int J Mol Sci*, vol. 21, no. 13, pp. 1–17, Jul. 2020, doi: 10.3390/ijms21134699.
[144] A. Ricarte, J. A. R. Dalton, and J. Giraldo, “Structural Assessment of Agonist Efficacy in the μ-Opioid Receptor: Morphine and Fentanyl Elicit Different Activation Patterns,” *J Chem Inf Model*, vol. 61, no. 3, pp. 1251–1274, 2021, doi: 10.1021/acs.jcim.0c00890.
[145] C. Bond *et al.*, “Single-nucleotide polymorphism in the human mu opioid receptor gene alters-endorphin binding and activity: Possible implications for opiate addiction,” 1998. [Online]. Available: www.pnas.org.
[146] M. T. Reetz, D. Kahakeaw, and R. Lohmer, “Addressing the numbers problem in directed evolution,” *ChemBioChem*, vol. 9, no. 11, pp. 1797–1804, Jul. 2008, doi: 10.1002/cbic.200800298.
[147] D. Feldman *et al.*, “Optical Pooled Screens in Human Cells,” *Cell*, vol. 179, no. 3, pp. 787-799.e17, Oct. 2019, doi: 10.1016/j.cell.2019.09.016.
[148] P. H. C. Eilers and J. J. Goeman, “Enhancing scatterplots with smoothed densities,” *Bioinformatics*, vol. 20, no. 5, pp. 623–628, Mar. 2004, doi: 10.1093/bioinformatics/btg454.
[149] F. Schneider, D. Gradmann, and P. Hegemann, “Ion selectivity and competition in channelrhodopsins,” *Biophys J*, vol. 105, no. 1, pp. 91–100, Jul. 2013, doi: 10.1016/j.bpj.2013.05.042.
[150] K. Oda *et al.*, “Crystal structure of the red light-activated channelrhodopsin Chrimson,” *Nat Commun*, vol. 9, no. 1, p. 3949, Sep. 2018, doi: 10.1038/s41467-018-06421-9.
[151] V. Gradinaru, K. R. Thompson, and K. Deisseroth, “eNpHR: A Natronomonas halorhodopsin enhanced for optogenetic applications,” *Brain Cell Biol*, vol. 36, no. 1–4, pp. 129–139, Aug. 2008, doi: 10.1007/s11068-008-9027-6.
[152] A. Berndt *et al.*, “Structural foundations of optogenetics: Determinants of channelrhodopsin ion selectivity,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 4, pp. 822–829, 2016, doi: 10.1073/pnas.1523341113.
[153] Y. S. Kim *et al.*, “Crystal structure of the natural anion-conducting channelrhodopsin GtACR1,” *Nature*, vol. 561, no. 7723, pp. 343–348, Sep. 2018, doi: 10.1038/s41586-018-0511-6.
[154] M. Mahn *et al.*, “High-efficiency optogenetic silencing with soma-targeted anion-conducting channelrhodopsins,” *Nat Commun*, vol. 9, no. 1, Dec. 2018, doi: 10.1038/s41467-018-06511-8.
[155] M. Mahn, M. Prigge, S. Ron, R. Levy, and O. Yizhar, “Biophysical constraints of optogenetic inhibition at presynaptic terminals,” *Nat Neurosci*, vol. 19, no. 4, pp. 554–556, Mar. 2016, doi: 10.1038/nn.4266.

[156] J. S. Wiegert, M. Mahn, M. Prigge, Y. Printz, and O. Yizhar, “Silencing Neurons: Tools, Applications, and Experimental Constraints,” *Neuron*, vol. 95, no. 3. Cell Press, pp. 504–529, Aug. 02, 2017. doi: 10.1016/j.neuron.2017.06.050.
[157] M. Koyanagi *et al.*, “Bistable UV pigment in the lamprey pineal,” 2004. [Online]. Available: www.pnas.orgcgidoi10.1073pnas.0400819101
[158] B. A. Copits *et al.*, “A photoswitchable GPCR-based opsin for presynaptic inhibition,” *Neuron*, vol. 109, no. 11, pp. 1791-1809.e11, Jun. 2021, doi: 10.1016/j.neuron.2021.04.026.
[159] R. D. Airan, K. R. Thompson, L. E. Fenno, H. Bernstein, and K. Deisseroth, “Temporally precise in vivo control of intracellular signalling.,” *Nature*, vol. 458, no. 7241, pp. 1025–9, 2009, doi: 10.1038/nature07926.
[160] E. R. Siuda *et al.*, “Spatiotemporal Control of Opioid Signaling and Behavior,” *Neuron*, vol. 86, no. 4, pp. 923–935, 2015, doi: 10.1016/j.neuron.2015.03.066.
[161] L. Wang *et al.*, “A high-performance genetically encoded fluorescent indicator for in vivo cAMP imaging,” *BioRxiv*, 2022, doi: 10.1101/2022.02.27.482140.
[162] C. C. Torres Cabán *et al.*, “Tuning the Sensitivity of Genetically Encoded Fluorescent Potassium Indicators through Structure-Guided and Genome Mining Strategies,” *ACS Sens*, vol. 7, no. 5, pp. 1336–1346, May 2022, doi: 10.1021/acssensors.1c02201.
[163] Y. Shen *et al.*, “Genetically encoded fluorescent indicators for imaging intracellular potassium ion concentration,” *Commun Biol*, vol. 2, no. 1, Dec. 2019, doi: 10.1038/s42003-018-0269-2.
[164] G. Nagel *et al.*, “Channelrhodopsin-2, a directly light-gated cation-selective membrane channel,” *Proc Natl Acad Sci U S A*, vol. 100, no. 24, pp. 13940–13945, Nov. 2003, doi: 10.1073/pnas.1936192100.
[165] A. Berndt, S. Y. Lee, C. Ramakrishnan, and K. Deisseroth, “Structure-guided transformation of channelrhodopsin into a light-activated chloride channel,” *Science (1979)*, vol. 344, no. 6182, pp. 420–424, 2014, doi: 10.1126/science.1252367.
[166] E. G. Govorunova, O. A. Sineshchekov, R. Janz, X. Liu, and J. L. Spudich, “Natural light-gated anion channels: A family of microbial rhodopsins for advanced optogenetics,” *Science (1979)*, vol. 349, no. 6248, pp. 647–650, 2015, doi: 10.1126/science.aaa7484.
[167] J. Wietek *et al.*, “Conversion of channelrhodopsin into a light-gated chloride channel,” *Science (1979)*, vol. 344, no. 6182, pp. 409–412, 2014, doi: 10.1126/science.1249375.
[168] E. S. Boyden, F. Zhang, E. Bamberg, G. Nagel, and K. Deisseroth, “Millisecond-timescale, genetically targeted optical control of neural activity.,” *Nat Neurosci*, vol. 8, no. 9, pp. 1263–1268, Sep. 2005, doi: 10.1038/nn1525.
[169] K. Feldbauer, D. Zimmermann, V. Pintschovius, J. Spitz, C. Bamann, and E. Bamberg, “Channelrhodopsin-2 is a leaky proton pump,” *Proc Natl Acad Sci U S A*, vol. 106, no. 30, pp. 12317–12322, Jul. 2009, doi: 10.1073/pnas.0905852106.
[170] H. E. Kato *et al.*, “Crystal structure of the channelrhodopsin light-gated cation channel,” *Nature*, vol. 482, no. 7385, pp. 369–374, 2012, doi: 10.1038/nature10870.
[171] Y. S. Kim *et al.*, “Crystal structure of the natural anion-conducting channelrhodopsin GtACR1,” *Nature*, vol. 561, no. 7723, pp. 343–348, Sep. 2018, doi: 10.1038/s41586-018-0511-6.
[172] H. E. Kato *et al.*, “Structural mechanisms of selectivity and gating in anion channelrhodopsins,” *Nature*, vol. 561, no. 7723, pp. 349–354, Sep. 2018, doi: 10.1038/s41586-018-0504-5.

[173] H. Li *et al.*, “Crystal structure of a natural light-gated anion channelrhodopsin,” *Elife*, vol. 8, p. e41741, Jan. 2019, doi: 10.7554/eLife.41741.
[174] O. Volkov *et al.*, “Structural insights into ion conduction by channelrhodopsin 2,” *Science (1979)*, vol. 358, no. 6366, p. eaan8862, Nov. 2017, doi: 10.1126/science.aan8862.
[175] K. Oda *et al.*, “Crystal structure of the red light-activated channelrhodopsin Chrimson,” *Nat Commun*, vol. 9, no. 1, p. 3949, Sep. 2018, doi: 10.1038/s41467-018-06421-9.
[176] F. Schneider, D. Gradmann, and P. Hegemann, “Ion selectivity and competition in channelrhodopsins,” *Biophys J*, vol. 105, no. 1, pp. 91–100, Jul. 2013, doi: 10.1016/j.bpj.2013.05.042.
[177] H. E. Kato *et al.*, “Crystal structure of the channelrhodopsin light-gated cation channel,” *Nature*, vol. 482, no. 7385, pp. 369–374, 2012, doi: 10.1038/nature10870.
[178] A. Berndt *et al.*, “Structural foundations of optogenetics: Determinants of channelrhodopsin ion selectivity,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 4, pp. 822–829, 2016, doi: 10.1073/pnas.1523341113.
[179] H. E. Kato *et al.*, “Structural mechanisms of selectivity and gating in anion channelrhodopsins,” *Nature*, vol. 561, no. 7723, pp. 349–354, Sep. 2018, doi: 10.1038/s41586-018-0504-5.
[180] Y. S. Kim *et al.*, “Crystal structure of the natural anion-conducting channelrhodopsin GtACR1,” *Nature*, vol. 561, no. 7723, pp. 343–348, Sep. 2018, doi: 10.1038/s41586-018-0511-6.
[181] H. Li *et al.*, “Crystal structure of a natural light-gated anion channelrhodopsin,” *Elife*, vol. 8, p. e41741, Jan. 2019, doi: 10.7554/eLife.41741.
[182] Z. Zhang and J. Chen, “Atomic Structure of the Cystic Fibrosis Transmembrane Conductance Regulator,” *Cell*, vol. 167, no. 6, pp. 1586-1597.e9, 2016, doi: 10.1016/j.cell.2016.11.014.
[183] A. Alam and Y. Jiang, “Structural studies of ion selectivity in tetrameric cation channels,” *J Gen Physiol*, vol. 137, no. 5, pp. 397–403, May 2011, doi: 10.1085/jgp.201010546.
[184] R. E. Hibbs and E. Gouaux, “Principles of activation and permeation in an anion-selective Cys-loop receptor,” *Nature*, vol. 474, p. 54, May 2011.
[185] X. Hou, L. Pedi, M. M. Diver, and S. B. Long, “Crystal Structure of the Calcium Release-Activated Calcium Channel Orai,” *Science (1979)*, vol. 338, pp. 659–664, 2012, doi: 10.1126/science.1159850.
[186] O. Volkov *et al.*, “Structural insights into ion conduction by channelrhodopsin 2,” *Science (1979)*, vol. 358, no. 6366, p. eaan8862, Nov. 2017, doi: 10.1126/science.aan8862.
[187] H. Luecke, B. Schobert, H. T. Richter, J. P. Cartailler, and J. K. Lanyi, “Structural changes in bacteriorhodopsin during ion transport at 2 angstrom resolution,” *Science (1979)*, vol. 286, no. 5438, pp. 255–260, 1999, doi: 10.1126/science.286.5438.255.
[188] A. Berndt *et al.*, “High-efficiency channelrhodopsins for fast neuronal stimulation at low light levels,” *Proceedings of the National Academy of Sciences*, vol. 108, no. 18, pp. 7595–7600, 2011, doi: 10.1073/pnas.1017210108.
[189] N. C. Klapoetke *et al.*, “Independent optical excitation of distinct neural populations,” *Nat Methods*, vol. 11, no. 3, pp. 338–346, 2014, doi: 10.1038/nmeth.2836.
[190] M. J. Gunthorpe and S. C. R. Lummis, “Conversion of the Ion Selectivity of the 5-HT3AReceptor from Cationic to Anionic Reveals a Conserved Feature of the Ligand-gated Ion Channel Superfamily ,” *Journal of Biological Chemistry* , vol. 276, no. 14, pp. 10977–10983, Apr. 2001, doi: 10.1074/jbc.M009575200 .

[191] J.-L. Galzi, A. Devillers-Thiery, N. Hussy, S. Bertrand, J.-P. Changeux, and D. Bertrand, “Mutations in the channel domain of a neuronal nicotinic receptor convert ion selectivity from cationic to anionic,” *Nature*, vol. 359, no. 6395, pp. 500–505, 1992, doi: 10.1038/359500a0.
[192] R. E. Hibbs and E. Gouaux, “Principles of activation and permeation in an anion-selective Cys-loop receptor,” *Nature*, vol. 474, p. 54, May 2011.
[193] E. Park, E. B. Campbell, and R. MacKinnon, “Structure of a CLC chloride ion channel by cryo-electron microscopy,” *Nature*, vol. 541, no. 7638, pp. 500–505, 2017, doi: 10.1038/nature20812.
[194] C. Paulino, V. Kalienkova, A. K. M. Lam, Y. Neldner, and R. Dutzler, “Activation mechanism of the calcium-activated chloride channel TMEM16A revealed by cryo-EM,” *Nature*, vol. 552, p. 421, Dec. 2017.
[195] S. Dang *et al.*, “Cryo-EM structures of the TMEM16A calcium-activated chloride channel,” *Nature*, vol. 552, no. 7685, pp. 426–429, Dec. 2017, doi: 10.1038/nature25024.
[196] H. Yang *et al.*, “TMEM16F Forms a Ca2+-Activated Cation Channel Required for Lipid Scrambling in Platelets during Blood Coagulation,” *Cell*, vol. 151, no. 1, pp. 111–122, 2012, doi: https://doi.org/10.1016/j.cell.2012.07.036.
[197] G. Hassaine *et al.*, “X-ray structure of the mouse serotonin 5-HT3 receptor,” *Nature*, vol. 512, p. 276, Aug. 2014.
[198] J. R. Pugh and C. E. Jahr, “Axonal GABAA Receptors Increase Cerebellar Granule Cell Excitability and Synaptic Activity,” *Journal of Neuroscience*, vol. 31, no. 2, pp. 565–574, 2011, doi: 10.1523/JNEUROSCI.4506-10.2011.
[199] M. Mahn, M. Prigge, S. Ron, R. Levy, and O. Yizhar, “Biophysical constraints of optogenetic inhibition at presynaptic terminals,” *Nat Neurosci*, vol. 19, p. 554, Mar. 2016.
[200] D. L. Fortin *et al.*, “Optogenetic photochemical control of designer K+ channels in mammalian neurons,” *J Neurophysiol*, vol. 106, no. 1, pp. 488–496, Jul. 2011, doi: 10.1152/jn.00251.2011.
[201] A. Alam and Y. Jiang, “Structural studies of ion selectivity in tetrameric cation channels,” *J Gen Physiol*, vol. 137, no. 5, pp. 397–403, May 2011, doi: 10.1085/jgp.201010546.
[202] D. A. Köpfer, C. Song, T. Gruene, G. M. Sheldrick, U. Zachariae, and B. L. de Groot, “Ion permeation in K⁺ channels occurs by direct Coulomb knock-on,” *Science (1979)*, vol. 346, no. 6207, pp. 352 LP – 355, Oct. 2014, doi: 10.1126/science.1254840.
[203] W. Kopec *et al.*, “Direct knock-on of desolvated ions governs strict ion selectivity in K+ channels,” *Nat Chem*, vol. 10, no. 8, pp. 813–820, 2018, doi: 10.1038/s41557-018-0105-9.
[204] M. G. Derebe, D. B. Sauer, W. Zeng, A. Alam, N. Shi, and Y. Jiang, “Tuning the ion selectivity of tetrameric cation channels by changing the number of ion binding sites,” *Proceedings of the National Academy of Sciences*, vol. 108, no. 2, pp. 598–602, 2011, doi: 10.1073/pnas.1013636108.
[205] D. B. Sauer, W. Zeng, J. Canty, Y. Lam, and Y. Jiang, “Sodium and potassium competition in potassium-selective and non-selective channels,” *Nat Commun*, vol. 4, p. 2721, Nov. 2013.
[206] A. Berndt, S. Y. Lee, C. Ramakrishnan, and K. Deisseroth, “Structure-guided transformation of channelrhodopsin into a light-activated chloride channel,” *Science (1979)*, vol. 344, no. 6182, pp. 420–424, 2014, doi: 10.1126/science.1252367.

[207] E. G. Govorunova, O. A. Sineshchekov, R. Janz, X. Liu, and J. L. Spudich, "Natural light-gated anion channels: A family of microbial rhodopsins for advanced optogenetics," *Science (1979)*, vol. 349, no. 6248, pp. 647–650, 2015, doi: 10.1126/science.aaa7484.
[208] W. Kopec *et al.*, "Direct knock-on of desolvated ions governs strict ion selectivity in K+ channels," *Nat Chem*, vol. 10, no. 8, pp. 813–820, 2018, doi: 10.1038/s41557-018-0105-9.
[209] L. Tang *et al.*, "Structural basis for Ca2+ selectivity of a voltage-gated calcium channel," *Nature*, vol. 505, no. 7481, pp. 56–61, Jan. 2014, doi: 10.1038/nature12775.
[210] J. Wu *et al.*, "Structure of the voltage-gated calcium channel Cav1.1 complex," *Science (1979)*, vol. 350, no. 6267, 2015, doi: 10.1126/science.aad2395.
[211] P. T. Ellinor, J. Yang, W. A. Sather, J. F. Zhang, and R. W. Tsien, "Ca2+ channel selectivity at a single locus for high-affinity Ca2+ interactions," *Neuron*, vol. 15, no. 5, pp. 1121–1132, 1995, doi: 10.1016/0896-6273(95)90100-0.
[212] B. Corry and S. H. Chung, "Mechanisms of valence selectivity in biological ion channels," *Cellular and Molecular Life Sciences*, vol. 63, no. 3, pp. 301–315, 2006, doi: 10.1007/s00018-005-5405-8.
[213] E. Cao, M. Liao, Y. Cheng, and D. Julius, "TRPV1 structures in distinct conformations reveal activation mechanisms," *Nature*, vol. 504, no. 7478, pp. 113–118, Dec. 2013, doi: 10.1038/nature12823.
[214] N. Chakrabarti, C. Ing, J. Payandeh, N. Zheng, W. A. Catterall, and R. Pomès, "Catalysis of Na+ permeation in the bacterial sodium channel Na(V)Ab," *Proc Natl Acad Sci U S A*, vol. 110, no. 28, pp. 11331–11336, Jul. 2013, doi: 10.1073/pnas.1309452110.
[215] I. K. Kaufman, D. G. Luchinsky, W. A. T. Gibby, P. V. E. McClintock, and R. S. Eisenberg, "Putative resolution of the EEEE selectivity paradox in L-type Ca2+and bacterial Na+biological ion channels," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2016, no. 5, p. 54027, 2016, doi: 10.1088/1742-5468/2016/05/054027.
[216] Z. Yan *et al.*, "Structure of the rabbit ryanodine receptor RyR1 at near-atomic resolution," *Nature*, vol. 517, no. 7532, pp. 50–55, 2015, doi: 10.1038/nature14063.
[217] J. Wu *et al.*, "Structure of the voltage-gated calcium channel Cav1.1 complex," *Science (1979)*, vol. 350, no. 6267, 2015, doi: 10.1126/science.aad2395.
[218] A. Manglik, B. K. Kobilka, and J. Steyaert, "Nanobodies to Study G Protein–Coupled Receptor Structure and Function," *Annu Rev Pharmacol Toxicol*, vol. 57, no. 1, pp. 19–37, Jan. 2017, doi: 10.1146/annurev-pharmtox-010716-104710.
[219] S. Maeda *et al.*, "Development of an antibody fragment that stabilizes GPCR/G-protein complexes," *Nat Commun*, vol. 9, no. 1, p. 3712, 2018, doi: 10.1038/s41467-018-06002-w.
[220] E. Ritter, K. Stehfest, A. Berndt, P. Hegemann, and F. J. Bartl, "Monitoring light-induced structural changes of channelrhodopsin-2 by UV-visible and Fourier transform infrared spectroscopy," *Journal of Biological Chemistry*, vol. 283, no. 50, pp. 35033–35041, 2008, doi: 10.1074/jbc.M806353200.
[221] H. N. Chapman *et al.*, "Femtosecond X-ray protein nanocrystallography," *Nature*, vol. 470, no. 7332, pp. 73–78, 2011, doi: 10.1038/nature09750.
[222] B. Stauch and V. Cherezov, "Serial Femtosecond Crystallography of G Protein–Coupled Receptors," *Annu Rev Biophys*, vol. 47, no. 1, pp. 377–397, May 2018, doi: 10.1146/annurev-biophys-070317-033239.

[223] M. J. Latimer *et al.*, “Simultaneous Femtosecond X-ray Spectroscopy and Diffraction of Photosystem II at Room Temperature,” *Science (1979)*, vol. 340, no. 6131, pp. 491–495, 2013, doi: 10.1126/science.1234273.

[224] N. Coquelle *et al.*, “Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography,” *Nat Chem*, vol. 10, no. 1, pp. 31–37, 2018, doi: 10.1038/NCHEM.2853.

[225] J. Tenboer *et al.*, “Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein,” *Science (1979)*, vol. 346, no. 6214, pp. 1242–1246, 2014, doi: 10.1126/science.1259357.

[226] R. Andersson *et al.*, “A three-dimensional movie of structural changes in bacteriorhodopsin,” *Science (1979)*, vol. 354, no. 6319, pp. 1552–1557, 2016, doi: 10.1126/science.aah3497.

[227] A. Lieb, M. Weston, and D. M. Kullmann, “Designer receptor technology for the treatment of epilepsy,” *EBioMedicine*, vol. 43. Elsevier B.V., pp. 641–649, May 01, 2019. doi: 10.1016/j.ebiom.2019.04.059.

[228] D. J. Urban and B. L. Roth, “DREADDs (designer receptors exclusively activated by designer drugs): Chemogenetic tools with therapeutic utility,” *Annual Review of Pharmacology and Toxicology*, vol. 55. Annual Reviews Inc., pp. 399–417, Jan. 06, 2015. doi: 10.1146/annurev-pharmtox-010814-124803.

[229] L. Beaulieu-Laroche *et al.*, “Allometric rules for mammalian cortical layer 5 neuron biophysics,” *Nature*, vol. 600, no. 7888, pp. 274–278, Dec. 2021, doi: 10.1038/s41586-021-04072-3.

[230] J. A. Ballesteros and H. Weinstein, “Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors,” 1995, pp. 366–428. doi: 10.1016/S1043-9471(05)80049-7.

[231] V. Isberg, B. Vroling, R. van der Kant, K. Li, G. Vriend, and D. Gloriam, “GPCRDB: An information system for G protein-coupled receptors,” *Nucleic Acids Res*, vol. 42, no. D1, Jan. 2014, doi: 10.1093/nar/gkt1255.

APPENDIX A: DNA AND PEPTIDE SEQUENCES OF SENSORS

Significant DNA and Peptide Sequences used in this study:

cpGFP
Linkers
Trafficking sequences
GPCR sequences

μMASS[2A]

atggacagcagcaccggcccagggaacaccagcgactgctcagacccctagctcaggcaagttgctccccagcacctggctcctggct
caacttgtcccacgttgatggcaaccagtccgatccatgcggtctgaaccgcaccgggcttggcgggaacgacagcctgtgccctcagac
cggcagcccttccatggtcacagccattaccatcatggccctctactctatcgtgtgtgtagtgggcctcttcggaaacttcctggtcatgtatg
tgattgtaagatacaccaaaatgaagactgccaccaacatctacattttcaaccttgctctggcagacgccttagcgaccagtacactgcccctt
tcagagtgtcaactacctgatgggaacatggcccttcggaaccatcctctgcaagatcgtgatctcaatagattactacaacatgttcaccagc
atattcaccctctgcaccatgagcgtggaccgctacattgctgtctgccacccaGTCaaagccctggatttccgtacccccccgaaatgcca
aaatcgtcaacgtctgcaactggatcctctcttctgccatcggtctgcctgtaatgttcatggcaaccacGaaatacaggcaggggtccatag
attgcaccctcacgttctcccacccaacctggtactgggagaacctgctcaaaatctgtgtctttatcttcgctttcatcatgccgatcctcatcat
cactgtgtgttacggcctgatgatcttacgactcaagagcgttcgcctgagctcactTGAtaacgtctatatcaaggccgacaagcagaag
aacggcatcaaggcgaacttcaagatccgccacaacatcgaggacggcggcgtgcagctcgcctaccactaccagcagaacaccccccat
cggcgacggccccgtgctgctgcccgacaaccactacctgagcgtgcagtccaaactttcgaaagaccccaacgagaagcgcgatcaca
tggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaagggcggtaccggagggagcatggtgagc
aagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaacggccacaagttcagcgtgtccggcg
agggtgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggcccaccctcgtg
accaccctgacctacggcgtgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaagtccgccatgcccgaaggc
tacatccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaa
ccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaacATGGTTgaccaactg
aaagaaaaggacaggaatctgcgcaggatcacccggatggtgctggtggtcgtggctgtatttatcgtctgctggacccccatccacatcta
cgtcatcatcaaagcgctgatcacgattccagaaaccacatttcagaccgtttcctggcacttctgcattgctttgggttacacgaacagctgc
ctgaatccagttctttacgccttcctggatgaaaacttcaagcgatgcttcagagagttctgcatcccaacctcgtccacgatcgaacagcaaa
actccactcgagtccgtcagaacactagggaacatccctccacggctaatacagtggatcgaactaaccaccagctagaaaatctggaggc
agaaactgctccattgcccGGCGGAGGTAAGAGCAGAATCACAAGCGAGGGCGAGTACATCCC
TCTGGACCAGATCGACATCAACGTGGGAGCACCAGCCGCCGCTTTCTGCTACGAGA
ACGAAGTCGGCtaa

MDSSTGPGNTSDCSDPLAQASCSPAPGSWLNLSHVDGNQSDPCGLNRTGLGGNDSLCP
QTGSPSMVTAITIMALYSIVCVVGLFGNFLVMYVIVRYTKMKTATNIYIFNLALADALAT
STLPFQSVNYLMGTWPFGTILCKIVISIDYYNMFTSIFTLCTMSVDRYIAVCHPVKALDFR
TPRNAKIVNVCNWILSSAIGLPVMFMATTKYRQGSIDCTLTFSHPTWYWENLLKICVFIF
AFIMPILIITVCYGLMILRLKSVRLSSLDNVYIKADKQKNGIKANFKIRHNIEDGGVQLAY

HYQQNTPIGDGPVLLPDNHYLSVQSKLSKDPNEKRDHMVLLEFVTAAGITLGMDELYK
GGTGGSMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGK
LPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYIQERTIFFKDDGNYKTRAE
VKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNMVDQLKEKDRNLRRITRMVLVVVAV
FIVCWTPIHIYVIIKALITIPETTFQTVSWHFCIALGYTNSCLNPVLYAFLDENFKRCFREFC
IPTSSTIEQQNSTRVRQNTREHPSTANTVDRTNHQLENLEAETAPLPGGGKSRITSEGEYI
PLDQIDINVGAPAAAFCYENEVG*

dMASS[3A]

atgaggactctgaacacctctgccatggacgggactgggctggtggtggagagggacttctctgttcgtatcctcactgcctgtttcctgtcg
ctgctcatcctgtccacgctcctggggaacacgctggtctgtgctgccgttatcaggttccgacacctgcggtccaaggtgaccaacttctttg
tcatctccttggctgtgtcagatctcttggtggccgtcctggtcatgccctggaaggcagtggctgagattgctggcttctggccctttgggtcc
ttctgtaacatctgggtggcctttgacatcatgtgctccactgcatccatcctcaacctctgtgtgatcagcgtgGACaggtattgggctatct
ccagccctTTCcggtatgagagaaagatgacccccaaggcagccttcatcctgatcagtgtggcatggaccttgtctgtactcatctccttc
atcccagtgcagctcagctggcacaaggcaaaacccacaagcccctctgatggaaatgccacttccctggctgagaccatagacaactgt
gactccagcctcagcaggacatatgccatctcatcctctgtaatcagcttttacatccctgtggccatcatgattgtcacctacaccaggatcta
caggattgctcagaaactgagctcaGGTTGTaacgtctatatcaaggccgacaagcagaagaacggcatcaaggcgaacttcaagat
ccgccacaacatcgaggacggcggcgtgcagctcgcctaccactaccagcagaacacccccatcggcgacggccccgtgctgctgccc
gacaaccactacctgagcgtgcagtccaaactttcgaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgcc
gccgggatcactctcggcatggacgagctgtacaagggcggtaccggagggagcatggtgagcaagggcgaggagctgttcaccggg
gtggtgcccatcctggtcgagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgagggtgagggcgatgccacctacg
gcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgcagt
gcttcagccgctaccccgaccacatgaagcagcacgacttcttcaagtccgccatgcccgaaggctacatccaggagcgcaccatcttctt
caaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatcgagctgaagggcatcga
cttcaaggaggacggcaacatcctggggcacaagctggagtacaacCGGTTGgaccaactgaaaagagaaactaaagtcctgaag
actctgtcggtgatcatgggtgtgtttgtgtgctgttggctacctttcttcatcttgaactgcattttgcccttctgtgggtctggggagacgcagc
ccttctgcattgattccaacacctttgacgtgtttgtgtggtttgggtgggctaattcatccttgaaccccatcatttatgcctttaatgctgattttcg
gaaggcattttcaaccctcttaggatgctacagactttgccctgcgacgaataatgccatagagacggtgagtatcaataacaatggggccg
cgatgttttccagccatcatgagccacgaggctccatctccaaggagtgcaatctggtttacctgatcccacatgctgtgggctcctctgagg
acctgaaaaaggaggaggcagctggcatcgccagacccttggagaagctgtccccagccctatcggtcatattggactatgacactgacgt
ctctctggagaagatccaacccatcacacaaaacggtcagcacccaacctga

dMASS[3A] Peptide Sequence:

MRTLNTSAMDGTGLVVERDFSVRILTACFLSLLILSTLLGNTLVCAAVIRFRHLRSKVTN
FFVISLAVSDLLVAVLVMPWKAVAEIAGFWPFGSFCNIWVAFDIMCSTASILNLCVISVD
RYWAISSPFRYERKMTPKAAFILISVAWTLSVLISFIPVQLSWHKAKPTSPSDGNATSLAE
TIDNCDSSLSRTYAISSSVISFYIPVAIMIVTYTRIYRIAQKLSSGCNVYIKADKQKNGIKA
NFKIRHNIEDGGVQLAYHYQQNTPIGDGPVLLPDNHYLSVQSKLSKDPNEKRDHMVLLE
FVTAAGITLGMDELYKGGTGGSMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEG
DATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYIQ
ERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNRLDQLKRETK
VLKTLSVIMGVFVCCWLPFFILNCILPFCGSGETQPFCIDSNTFDVFVWFGWANSSLNPII
YAFNADFRKAFSTLLGCYRLCPATNNAIETVSINNNGAAMFSSHHEPRGSISKECNLVYL
IPHAVGSSEDLKKEEAAGIARPLEKLSPALSVILDYDTDVSLEKIQPITQNGQHPT*

Parapinopsin: (Lys281 for PPO[mut])

MENLTSLDLLPNGEVPLMPRYGFTILAVIMAVFTIASLVLNSTVVIVTLRHRQLRHPLNFS
LVNLAVADLGVTVFGASLVVETNAVGYFNLGRVGCVIEGFAVAFFGIAALCTIAVIAVD
RFVVVCKPLGTLMFTRRHALLGIAWAWLWSFVWNTPPLFGWGSYELEGVRTSCAPDW
YSRDPANVSYITSYFAFCFAIPFLVIVVAYGRLMWTLHQVAKLGMGESGSTAKAEAQVS
RMVVVMVVAFLVCWLPYALFAMIVVTKPDVYIDPVIATLPMYLTKTSTVYNPIIYIFMN
RQFRDCAVPFLLCGRNPWAEPSSESATAASTSATSVTLASAPGQVSPSGGGKSRITSEGE
YIPLDQIDINVGAPAAAFCYENEVG

APPENDIX B: BALLESTEROS WEINSTEIN GPCR NUMBERING

The Ballesteros-Weinstein numbering technique for GPCRs was proposed in 1995[230]. Due to different length and mutations of the same GPCR between organism species, a number scheme needed to be developed to enable facile communication between biophysicists, biologists, pharmacology, and structural biologists that were discussing the same GPCR, but with slightly different lengths due to mutational insertions. Classic protein numbering mutational schemes, where the wild-type amino acid is first, followed by its number in the primary sequence, then it's mutation (e.g. A147Y) was inadequate, as there are commonly insertional mutations in GPCRs that shift these numbers. In Ballesteros-Weinstein numbering, each GPCR residue is assigned a three-digit number. The first number assigns which of the TM domains the residue is closest to, followed by a period and a two-digit number. The two-digit number denotes the number of residues between the identified residue and the most conserved residue in the respective TM domain. The most conserved residue in each TM domain is assigned the value of 50. For example, the Ballesteros-Weinstein number for residue D147 in the MOR, used to create the loss of function variant of μMASS^{2A}, is 3.32, meaning the residue is in the 3rd transmembrane domain, and 18 residues before the most highly conserved residue in transmembrane domain 3. Importantly, there may not be 50 amino acids between each highly conserved residue, the value of 50 is arbitrarily chosen as a reference point.

Of note, numbers below 50 correspond to residues approaching the extracellular surface of the receptor, numbers above 50 correspond to residues approaching the intracellular space. The most highly conserved residue for each TM domain, its amino acid identity, and the percentage of GPCRs with that amino acid in the location: N1.50: 98%, D2.50: 90%, R3.50: 95%, W4.50: 97%, P5.50: 78%, P6.50: 99%, P7.50: 88% [231].

APPENDIX C: IMPORTANT PRIMERS FOR THIS STUDY

Table 0.4. Important primers for this study.

Primer Name	Primer Sequence	Purpose of Primer
1543_dlight_RT2_R	tgcagttcaagatgaagaaagg	Reverse Transcription of dMASS library variants
1652-Bx30_2_RT_R	gcagtttctgcctccagatttt	Reverse Transcription of µMASS library variants
1530_5'UTR_11_F	tgtctggtcaaccaccgc	Recombination specific primer for cDNA amplification
1278-Bx6.1_lib_insF / 1642-Bx6_1_lib_insR	ttgctcagaaactgagctcaNNKNNKAACGTCTATATCAAGGCC gtttctcttttcagttggtcMNNMNNGTTGTACTCCAGCTTGTG	Library Generation dMASS (insert for Gibson)
1414_Bx30_1_lib_insF / 1641_Bx30_1_lib_insR	agagcgttcgcctgagctcaNNKNNKAACGTCTATATCAAGGCC tccttttctttcagttggtcMNNMNNGTTGTACTCCAGCTTGTG	Library Generation µMASS (insert for Gibson)
2533_MOR_SL_baF	gattccagaaaccacatttcagaccgtttccNNKcacttctgcattgctttgggttacac	Ligand Selectivity Library µMASS[2A]
2534_MOR_SL_Ins2R	gaaatgtggtttctggaatcgtgatcagcgcMNNgatgatgacgtagatgtggatggggg	
2535_MOR_SL_Ins2F	ccacccaacctggtactggNNKaacctgctcaaaatctgtgtctttatcttcgc	
2536_MOR_SL_ins1R	ccagtaccaggttgggtgggag	
2537_MOR_SL_ins1F	ctgccctttcagagtgtcNNKtacctgatgggaacatggcccttc	
2538_MOR_SL_baR	acactctgaaagggcagtgtactgg	
1351_RT-dLight_R	ttgggtgctgaccgtttt	R primer for cDNA amplification of dMASS library
1653_bx30_2_secondR	gatgatgacgtagatgtggatgg	R primer for cDNA amplification of µMASS library

VITA

Publications

Rappleye M *et al.,* "Opto-MASS: A high-throughput protein engineering platform for genetically encoded fluorescent sensors enabling all optical *in vivo* detection of monoamines and opioids" *Pre-print, BioRxiv:* https://doi.org/10.1101/2022.06.01.494241

Klima, J.C., Doyle, L.A., Lee, J.D., **Rappleye M** *et al.* "Incorporation of sensing modalities into de novo designed fluorescence-activating proteins." *Nat Communications* 12**,** 856 (2021) (PMID: 33558528). https://doi.org/10.1038/s41467-020-18911-w

Rappleye M, and Berndt A, "Structural Basis for Ion Selectivity and Engineering in Channelrhodopsins", *Current Opinion in Structural Biology,* 2019 Jun 4, 57:176-184.

Valera E, Berger J, Hassan U, Ghonge T, Liu J, **Rappleye M,** Winter J, Abboud D, Haidry Z, Healey R, Hung N, Leung N, Mansury N, Hasnain A, Lannon C, Price Z, White K, Bashir R, "A Microfluidic biochip platform for electrical quantification of proteins,", *Lab on a Chip,* 6 April 2018.

Taneja I, Reddy B, Damhorst G, Dave Zhao S, Hassan U, Price Z, Jensen T, Ghonge T, Patel M, Wachspress S, Winter J, **Rappleye M,** Smith G, Healey R, Ajmal M, Khan M, Patel J, Rawal H, Sarwar R, Soni S, Anwaruddin S, Davis B, Kumar J, White K, Bashir R, Zhu R, "Combining Biomarkers with EMR Data to Identify Patients in Different Phases of Sepsis," *Scientific Reports*, 7 September 2017.

Hassan U, Ghonge T, Reddy B, Patel M, **Rappleye M,** Taneja I, Tanna A, Healey R, Manusry N, Price Z, Jensen T, Berger J, Hasnain A, Flaugher E, Liu S, Davis B, Kumar J, White K, Bashir R, "A point-of-care microfluidic biochip for quantification of CD64 expression from whole blood for sepsis stratification," *Nature Communications*, 3 Jul, 2017.

SELECTED CONFERENCE PRESENTATIONS & TALKS

M. Rappleye, *et al.,* "A highly sensitive mu-opioid peptide sensor engineered with a massively parallel GPCR optogenetic tool building pipeline," Gordon Research Symposium, Optogenetic Approaches to Understanding Neural Circuits and Behavior, July 16-17, 2022 - *Talk*

M. Rappleye, et al., " A highly sensitive mu-opioid peptide sensor engineered with a massively parallel GPCR optogenetic tool building pipeline," Gordon Research Conference, Optogenetic Approaches to Understanding Neural Circuits and Behavior, July 17th-22nd 2022, Newry ME, - *Poster* – received prize

M. Rappleye, Sarah J. Wait, Jamison C. Siebart, Justin D. Lee, Amanda A. Nguyen, Daniel C. Castro, Azra Suko, Jennifer D. Deem, Michael R. Bruchas and Andre Berndt. "A highly sensitive mu-opioid peptide sensor optimized with a massively parallel, high throughput GPCR optogenetic tool building platform." Society for Neuroscience, November 13-16, 2021, Chicago, IL & Virtual

M. Rappleye, "High-throughput optimization of GPCR-based sensors"**,** February 22, 2021, UW Center of Excellence in Neurobiology of Addiction, Pain and Emotion, Seattle, WA & Virtual - *Talk*

M. Rappleye, J. Winter, M. Patel, P. Duarte Guevara, E. Flaugher, U. Hassan, B. Reddy, T. Jensen, Rashid Bashir, "Point-of-Care Biochip to Quantify Inflammatory Response by Measuring IL6 from Whole Blood," BMES 2016 Conference, Minneapolis, MN -*Talk*