Designing Switchable Opioid Peptides for Interrogating the Effects of Cell Type-Specific Opioid Receptor Activation

by

Lequn Geng

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in The University of Michigan
2022

Doctoral Committee:

Assistant Professor Wenjing Wang, Chair
Assistant Professor Kristin Koutmou
Professor Anna Mapp
Professor Alan Smrcka

Lequn Geng

lequn@umich.edu

ORCID iD: 0000-0003-3838-6191

Dedication

To my parents, Suhua Li and Jun Geng.

Acknowledgements

I thank Dr. Wenjing Wang, my thesis advisor, for providing me with an excellent research environment, exciting scientific ideas to work on, and the best support I could ever imagine to get as a Ph.D. student. It has been three blessful years working with Dr. Wang, and I could not have asked for more.

I would also like to thank the Wang Lab members, especially Jiaqi Shen, Kayla Kroning, and Yemima Butler, for their substantial help and collaborations with my research projects.

Dr. William Birdsong and Dr. Peng Li, as well as their lab members, have been excellent collaborators. I thank their work on animal experiments and their suggestions on the tool development. Dr. William Birdsong and Dr. Manojkumar Puthenveedu generously provided research materials.

My dissertation committee, Dr. Wenjing Wang, Dr. Kristin Koutmou, Dr. Anna Mapp, and Dr. Alan Smrcka provided valuable advice and critical feedbacks on my research projects and this thesis. I thank their dedication to nurturing younger generations of scientists.

At the University of Michigan, Dr. Robert Kennedy, Dr. Zhan Chen, Dr. Ryan Bailey, and Dr. Xudong Fan have also been great advisors and contributed to where I am in my Ph.D. journey today. Dr. Ning Fang at Xiamen University, China, has been a great mentor to me since 2014.

Outside of academia, these persons have been great company and have influenced me in many positive ways: Yujie Zhang, Yaoyu Guan, Feng Xie, Jintao Jiang, and Guoming Gao. I am truly blessed to be with them in the past and in the years to come.

Table of Contents

List of Figures

List of Abbreviations

3D	three-dimensional
6TM	6-transmembrane
7TM	7-transmembrane
AAV	adeno-associated virus
ADF	abuse-deterrent formulations
APEX2	an engineered soybean ascorbate peroxidase
*As*LOV2	second LOV domain of *Avena sativa* phototropin 1
BBB	blood-brain barrier
CapC	chemically-activated protein domain controlling the C-terminus of an effector
CapN	chemically-activated protein domain controlling the N-terminus of an effector
CAPs	CapC and CapN
CCK2	cholecystokinin receptor 2
CD4	cluster of differentiation 4
Cre-*lox*P	Cre recombinase and locus of x-over, P1
CRISPR	clustered regularly interspaced short palindromic repeats
C-terminus	carboxyl terminus
cyclic AMP	cyclic adenosine monophosphate
DADLE	[D-Ala^2, D-Leu^5]-enkephalin
DAMGO	[D-Ala^2, N-MePhe^4, Gly-ol]-enkephalin
Dmt	2',6'-dimethyltyrosine
DNA	deoxyribonucleic acid
DOR	delta-opioid receptor

DREADDs	designer receptors exclusively activated by designer drugs
EGFP	evolved green fluorescent protein
ER	endoplasmic reticulum
FDA	U.S. Food and Drug Administration
GPCR	G protein-coupled receptors
HEK 293T	human embryonic kidney 293T cells
IBNtxA	iodobenzoylnaltrexamide
IRES	internal ribosome entry site
KO	knock-out
KOR	kappa-opioid receptor
LHA	lateral hypothalamic area
LOV domain	light, oxygen, voltage sensing domain
MOR	mu-opioid receptor
M-PROBE	cell membrane tethered opioid peptide for activating opioid receptors
mRNA	messenger ribonucleic acid
NAM	negative allosteric modulator
Nb	nanobody
NK1	neurokinin 1 receptor
NOP	noticeptin receptor
N-terminus	amino terminus
OR	opioid receptors
P2A	procine 2A self-cleaving peptide
PAM	positive allosteric modulator
PDYN	prodynorphin
PENK	proenkephalin
POMC	proopiomelanocortin
SAM	silent allosteric modulator
SPARK	Specific Protein Association tool giving transcriptional Readout with rapid Kinetics
TF	transcription factor

Tic	tetrahydroisoquinoline-3-carboxylic acid
TM	transmembrane domain
TEVcs	tobacco etch virus protease cleavage site
TEVp	tobacco etch virus protease
Tre	tetracycline response element
VP16	herpes simplex virus protein vmw65

List of Appendices

Abstract

Opioids are the most effective analgesics clinically available for treating severe pain, but they also cause adverse side effects such as addiction and respiratory suppression. Researchers and pharmaceutical companies have been trying to separate the desired outcomes of opioids from the unwanted ones, but this goal is yet to be achieved. This is in part due to the lack of understanding of how opioid receptors in different cell types and brain regions are linked to different opioid-induced effects.

This thesis describes the engineering of a novel genetic tool termed M-PROBE. M-PROBE is aimed at activating opioid receptors in the animal brain with cell type specificity, and thus enabling the study of correlation between cell type and behavior. M-PROBE is a multi-module protein encoded by designer DNA sequences. It has three major components: an opioid peptide for activating opioid receptors, a protein switch for controlling the activity of the opioid peptide, and a transmembrane domain for displaying M-PROBE on the cell outer membrane. To allow versatile applications of M-PROBE, two types of protein switches were designed, one of which is controlled by light (photoswitchable), and the other by a small molecule (chemically-activated).

All components of M-PROBE were first engineered through rational design. For the opioid peptide, met-enkephalin was identified as the optimal candidate based on a G-protein recruitment assay. For the protein switches, we selected the second light, oxygen, voltage sensing domain from oat as the starting point for the light switch, and an FK-506 binding protein previously engineered in-house as the starting point for the chemical switch. After re-engineering both proteins using circular permutation, they were confirmed to be capable of controlling a seven-amino-acid peptide,

SsrA, in a binding assay. Lastly, an extracellular domain-truncated human CD4 protein was used as the transmembrane domain.

After rational design, both the light switch and the chemical switch were improved using yeast surface-based directed evolution. I demonstrate that both directed revolution campaigns were highly effective in expanding the dynamic range (between the "closed" state and the "open" state) of the protein switches, and thus the theoretical overall performance of M-PROBE.

Finally, the chemically-activated M-PROBE was verified using both G-protein recruitment assay and a downstream secondary messenger (cyclic AMP) assay.

The engineering of M-PROBE has four main implications. First, M-PROBE is the first tool capable of activating any endogenous G protein-coupled receptors (GPCRs) with cell type specificity. This will expand our understanding of the endogenous GPCRs by providing a novel means of studying the correlation between cell type and behavior. The design principles of M-PROBE can be readily translated to other peptide GPCRs. Second, our re-engineered light and chemical switches have applications far beyond controlling opioid peptides. Since both switches can be applied to a range of short peptides, they open the door for studying many other cellular processes. Third, our directed evolution platform for engineering the light switch is the first one capable of labeling two states of a protein simultaneously. This design can be expanded to engineering other dual-state proteins. Lastly, M-PROBE itself may serve as a prototype of gene therapy that might solve the long-standing challenge of separating the analgesic effects of opioid from the adverse effects.

Chapter 1 Introduction to the Endogenous Opioid System and Current Methods for Studying It

1.1 The Opioid Epidemic

In an article[1] published in the Journal of the American Medical Association in 1915, David Macht wrote, "if the entire materia medica at our disposal were limited to the choice and use of only one drug, I am sure that a great many, if not the majority, of us would choose opium".

The juice from the opium poppy, *Papaver somniferum*, has been used medically for thousands of years, with the earliest literature dating back to the Greek and Latin period.[1] Opium is most known for its pain-relieving effect. In fact, even today, opium and its synthetic derivatives (termed "opioids") are the most effect analgesics for moderate-to-severe pain, and thus are often prescribed clinically for cancer and post-surgical patients.[2]

However, the use of opioids carries significant side effects. Opioids are known to be addictive, and opioid overdose can lead to respiratory suppression, eventually leading to death. It is estimated that one in every four patients who receive long-term opioid therapy in a primary care setting is subject to opioid addiction.[3] In 2019, more than 10 million Americans, or about 4% of the US adult population, reported misusing opioids.[4]

The death toll from the side effects of opioids is heavy. In the 1990s, prescription of opioids saw an increase in the US, leading to the first wave of opioid overdose deaths.[5] This wave was

followed by a rapid increase in heroin-induced overdose deaths[6] since 2010, and then by synthetic opioid (especially fentanyl)-induced deaths[7] since 2013. Together, these three waves contributed to about 500,000 deaths from 1999 to 2019.[8] In 2019 alone, opioid accounted for over 70% of the 70,630 drug overdose deaths in the US.[9] On October 26, 2017, the US Department of Health and Human Services declared the opioid epidemic to be a public health emergency[10], which is still in effect at the time of writing of this thesis.

1.2 Various Approaches to Combat the Opioid Epidemic

To date, many different methods have been proposed and implemented aiming at combating the opioid epidemic. Here I summarize strategies that are either currently in use or under active development. Designing novel opioid drugs is not included here, and will be more extensively discussed in the following sub-chapter.

One approach to reduce opioid misuse is abuse-deterrent opioid formulations (ADFs). These formulations make opioid abuse more difficult or less rewarding by targeting the expected routes of abuse. For example, a physical deterrent may form a viscous gel when the opioid tablet or capsule is crushed, preventing abuse by injection; a chemical deterrent may release an opioid antagonist when the opioid drug is manipulated.[11] The extended-release opioid formulations[12] may also be broadly considered as ADFs. They have a slower rate of entering the central nervous system and thus may have a lower abuse liability.[12] However, currently available ADFs do not deter taking multiple opioid tablets or capsules, the most common form of opioid abuse. In addition, ADFs are significantly more expensive than regular opioid formulations, making it a difficult economic

choice for patients and insurance companies.[11] Finally, although ADFs can deter opioid abuse using certain methods, the deterrent mechanisms can often be circumvented, and none of them has true abuse-deterrent properties.[13]

Another potential solution to opioid addiction is to eliminate patients' access to opioids through the oral route. Instead, opioids are administered through a slow-release, injectable formulation[14] or by a feedback-loop opioid delivery device that can detect pain and release the appropriate amount of opioid automatically[13]. Extended-release opioid injections have the caveat of not being able to be adjusted according to the level of pain, and their pain control effect has been found to be lesser than systematic opioid administrations.[13] In contrast, automatic opioid delivery device can adapt to the pain level, but they require surgery for implantation and removal, and their design and manufacturing is challenging.[13]

Opioid replacement therapy represents another strategy to solve opioid addiction. Also known as opioid pharmacotherapy or medication-assisted therapies, this approach uses drugs such as methadone (opioid agonist), buprenorphine (opioid partial agonist), naloxone (opioid antagonist), and naltrexone (opioid antagonist) which have safer profiles than many opioids, especially illicit drugs.[11] Naloxone, in particular, is also frequently distributed to first responders and potential bystanders to reverse opioid overdose.[15] Although such approaches can reduce the harm of opioid addiction, they are used only *after* patients have developed opioid use disorder, and therefore serve as a remedy rather than a fundamental solution to the opioid epidemic.

Other strategies to tackle the opioid epidemic include developing opioid vaccines or monoclonal antibodies that bind to commonly misused opioids[13], small molecules that target recognition, motivation, and drug reward circuits[16–18], and biomedical devices that can detect, treat, and prevent addiction[19]. These methods are either currently under development, or have their own

limitations such as being prohibitively expensive.[13] Clinically, there are additional strategies to manage opioid-induced side effects, such as dose reduction, opioid rotation, using alternative route of administration, and symptomatic treatment of adverse effects.[20] However, these strategies are aimed at reducing side effects such as constipation, nausea, sedation, and pruritus, but are not necessarily effective in reducing the key issue of opioid addiction. In addition, like all the approaches mentioned above, they do not fundamentally change the fact that opioids are addictive.

1.3 Drug Development Efforts to Decouple Analgesic Effect and Side Effects of Opioids

The most attractive and long-term solution to the opioid epidemic is arguably to develop novel opioids that do not have side effects in the first place. It is no secret that in the past one to two decades, both academic researchers and pharmaceutical companies have been working on minimizing opioid-induced side effects, especially addition and respiratory suppression. This sub-chapter will provide an overview of developments on this front.

Biased Ligands

Biased ligands are perhaps the most sought after approach so far in reducing opioid-induced side effects. When opioid receptors are activated by opioid drugs, they recruit both β-arrestin2 and G-protein, which then have different downstream pathways. A biased ligand causes the receptors to preferentially recruit β-arrestin2 or G-protein, and thus activates one downstream

pathway more than the other. Developing opioid biased ligands then have the potential to cause certain physiological outcomes while avoiding the others.

Most opioid biased ligands are developed for the mu-opioid receptor (MOR, one of the three types of opioid receptors) because a highly cited study[21] in the 1990s showed that MOR is the major type of opioid receptor responsible for both the analgesic effect and side effects of morphine, one of the most commonly prescribed opioids in clinical settings. From 1999 to 2005, neuroscientist Laura Bohn, first with Robert Lefkowitz at Duke University, then at the Ohio State University, published a series of papers[22–25] finding that mice lacking β-arrestin2 got stronger and longer lasting analgesic effect from morphine, and that β-arrestin2 knock-out mice had diminished morphine-induced side effects including constipation and respiratory suppression. These findings suggest that opioids that are biased towards the G-protein signaling pathway might have safer profiles, and sparked the interest to develop such opioids.

In 2009, Herkinorin was reported to be a G-protein biased ligand at MOR[26]. However, it was later found to be non-biased[27], and the true effect of this drug remains unclear. In 2013, Trevena, a company co-founded by Robert Lefkowitz and his former postdoctoral scholar Jonathan Violin, developed TRV130 (also known as oliceridine, marketed as "Olinvo") as a G-protein biased ligand with analgesic effect and reduced respiratory suppression effect than morphine[28]. However, the clinical trials were unsuccessful, and FDA rejected this drug in 2018[29]. The company also developed TRV734[30], which remains an investigational product at the time of writing of this thesis. In 2016, PZM21 was discovered to be a G-protein biased ligand through structure-based docking computation[27], and was claimed to have analgesic effect on pain perception while having little respiratory suppression or morphine-like reinforcement side effects. However, another group found that PZM21 has a limited bias factor, if present at all[31]. In the same year, the biosynthesized

compound mitragynine and its analogs were also reported to be a G-protein based ligand for MOR[32,33], but it is also an antagonist at the delta-opioid receptor (DOR)[33], making it complicated to interpret the true effect of this drug on MOR. In 2017, a series of compounds with prefix "SR-" were reported as G-protein biased ligands[34] by Laura Bohn. These compounds might be the first reported agonists to have statistically significant G-protein bias and reduced respiratory suppression effect.[34] The original paper claimed that the more G-protein biased, the less respiratory suppression. However, these compounds are not degraded in the brain for hours, raising questions regarding their pharmacokinetics and how their effects came into being.[35] Around the same period, other reported MOR biased ligands include endomorphins[34,36–38] (arrestin biased), fentanyl (contradicting evidence on arrestin biased or G-protein biased), MEB-1166[39], MEB-1170[39], and a series of PZM21 analogues[40].

Until now, no drug has been approved for their biased agonism at MOR. Some of these drugs (including herkinorin, TRV130, and PZM21) are questioned to be false positives and their "biased agonism" may be actually due to their low efficacy.[28,35] It remains obscure whether the ligand profiles of the biased ligands can actually be explained by biased agonist, or if they are due to partial agonism. In addition, as demonstrated above, many published biased ligands saw contradicting conclusions across studies. The ground on which biased agonists are based is also not without question. In 2019, Andrea Kliewer reported that the mice with mutated MOR that cannot recruit β-arrestin2 still experienced respiratory suppression after taking morphine and fentanyl, maybe even to a larger extent.[41] In 2020, a replicate[42] of the 2005 β-arrestin2 knockout experiment[25] in three independent labs found no difference between knockout mice and wild-type mice. Together, these data suggest that the G-protein pathway is also involved in respiratory suppression. Although there is compelling evidence that opioids with a large bias for the G-protein

pathway demonstrate reduced constipation and respiratory suppression[28,34], there is equally strong evidence supporting the involvement of G-protein pathway in respiratory suppression[31,41,42], possibly via regulating neuronal potassium channels[43,44] or voltage-gated calcium channels[45].

In summary, while biased agonism has been extensively studied and its potential in guiding new drug development heavily explored, the notion that G-protein pathway is responsible for analgesic effect and β-arrestin2 pathway is responsible for all other side effects is likely an oversimplification[41]. There is much more to be studied on what else lead to the various effects of opioids.

Allosteric Modulators

Allosteric modulators represent another possible approach for reducing opioid-related side effects. Allosteric modulators do not have intrinsic agonist or antagonist efficacy, but they can modulate the effects of orthosteric agonists (which binds to the main binding site) by binding to a non-orthosteric site on the receptors. Compared to other drugs that are designed to bind to the orthosteric sites, allosteric modulators have two main attractions: first, they often bind to non-conserved regions of a receptor, allowing selectivity among GPCR subtypes that cannot be readily differentiated by orthosteric agonists; second, they can enhance the effects of endogenous agonists while maintaining the spatial fidelity of native signaling. Several allosteric modulators at GPCRs have shown preclinical promise in neurodegenerative, psychiatric, and neurobehavioral diseases[46], and allosteric modulators at opioid receptors have also been explored for more than a decade.

There are three main types of allosteric modulators: positive allosteric modulators (PAMs), negative allosteric modulators (NAMs), and silent allosteric modulators (SAMs). PAMs can enhance the binding affinity and/or efficacy of an orthosteric agonist; NAMs can inhibit the

binding affinity and/or efficacy of an orthosteric agonist; SAMs do not affect the binding affinity and/or efficacy of an orthosteric agonist, but can serve as a competitive antagonist at the allosteric site, preventing PAMs and NAMs from binding.

In 2013, two compounds were identified to be MOR PAMs through a high-throughput screening approach with a β-arrestin2 recruitment assay.[47] They represent the first reported PAMs at any opioid receptor types. One of the compounds, BMS-986122, showed potential as a proof-of-concept drug for developing MOR NAMs with pain management utilities and improved side effect profiles[47], although no data related to the actual physiological effects of opioids were shown at the time of publication. In 2015, the same team applied computation-guided structural optimization to the two previously discovered compounds and discovered a new MOR PAM termed MS1.[48] MS1 can be more readily synthesized compared with its predecessors, but its effects were also only evaluated in vitro using cellular assays. Later, in 2021, the group showed the first in vivo data[49] ever published for any MOR PAM. Consistent with the hypothesis that MOR PAMs have therapeutic values by promoting the effect of endogenous opioid peptides and retaining their native signaling patterns, the authors discovered that BMS-986122 has antinociceptive effects in mice only when MOR is occupied by endogenous opioid peptides or exogenous opioids.[49] More interestingly, systematic injection of this compound at a working dose showed less constipation, reward, and respiratory suppression effects when compared with the same dose of morphine[49]. Around the same time, it was also discovered that two derivatives of the molecule MS1 can enhance the in vivo antinociceptive effects of several exogenous opioids with reduced withdrawal, tolerance, and respiratory suppression effects.[50]

Besides the BMS and M1 series of compounds, other MOR PAMs discovered in the past decade include the natural compound ignavine[51] and the neuropeptide oxytocin[52]. Both compounds

have been shown to enhance MOR signaling in vitro or in vivo, but their side effects have not been measured in vivo. In addition, oxytocin is an endogenous neuropeptide involved in mammalian labor and lactation[53], and has also been found to be a KOR PAM[54]. This poor selectivity may limit its potential as a MOR PAM when compared with synthesized compounds such as BMS-986122 and MS1.

Together, allosteric modulators, especially MOR PAMs with good selectivity, are a promising class of compounds that can enhance analgesic effects of opioids while having improved side effect profiles. Future researchers and clinicians will have to answer whether the effect of using allosteric modulators alone is anywhere comparable to using clinical opioids. Alternatively, to prove the clinical value of allosteric modulators, their side effects when used in combination with exogenous opioid drugs need to be evaluated.[49] Developing new allosteric modulators is challenging because these compounds are probe dependent and need to be tested with numerous orthosteric compounds in binding and functional assays.[55] However, this feature of allosteric modulators also offers the opportunity to tailor their activity towards specific opioid peptides or signaling pathways, allowing more specificity for drug development.[49]

Bi- and Multi-functional Opioid Receptor Ligands

The development of bifunctional opioid receptor ligands is based on the discovery that biological activity of drugs at a single receptor is often insufficient, and ligands with multiple activities may see better effects.[56,57] Bifunctional drugs still interact with receptors in a monovalent manner similar to conventional drugs, but they are capable of interacting with two or more targets. They are often made by connecting two pharmacophores together directly or via a linker, or by designing pharmacophores with overlapping or integrated functions.[58] Bifunctional drugs often

have improved potency due to synergistic effects, and compared with drug cocktails, they typically have more predictable pharmacokinetic and pharmacodynamic relationships, as well as lower risk for drug-drug interactions.[58]

Bi- or multi-functional ligands at opioid receptors typically combine agonist profiles at one opioid receptor with agonist or antagonist profiles at other opioid or non-opioid receptors, such as DOR, kappa-opioid receptor (KOR), nociceptin receptor (NOP), cholecystokinin receptor 2 (CCK2), and neurokinin 1 receptor (NK1).

One early direction is to develop drugs with MOR agonist / DOR antagonist profile. This is based on the findings[59–61] that selective blockage of DOR with antagonists can reduce morphine-induced tolerance and dependence. The first compound[62] with such profile was reported in 1992. The tetrapeptide amide Tyr-Tic-Phe-Phe-NH2 (TIPP-NH_2, Tic = tetrahydroisoquinoline-3-carboxylic acid) was found to have modest potency at MOR and strong antagonist activity at DOR.[62] Later development by the same group resulted in a more stable compound Dmt-Tic Ψ[CH_2NH]Phe-Phe-NH_2 (DIPP-NH_2[Ψ], Dmt = 2',6'-dimethyltyrosine)[63], which was a potent, balanced MOR agonist / DOR antagonist with sub-nanomolar binding affinities to both receptors. In 2000, the tetrapeptide H-Dmt-D-Arg-Phe-Lys-NH_2 ([Dmt^1]DALDA)[64] was found to be a highly potent MOR agonist. It also showed properties as a norepinephrine uptake inhibitor[65] and a releaser of endogenous opioid peptides acting at KOR and DOR[66]. Together, it showed enhanced analgesic effects in animals[65] and is considered a promising drug candidate for treating pain in the spinal level. While this drug has a low propensity to produce respiratory suppression[65], it does cause a strong tolerance effect[67,68].

Other bi- or multi-functional drugs were developed based on similar principles as the MOR agonist / DOR antagonist drugs. For example, the multifunctional MOR agonist / DOR agonist /

CCK-2 antagonist drug H-Tyr-D-Phe-Gly-D-Trp-NMeNle-Asp-Phe-NH2[69] was developed based on the findings that CCK2 receptor antagonists can reverse tolerance effects in mice[70] and attenuate morphine dependence[71]. DOR agonist / MOR agonist / NK1 antagonist drugs such as two octapeptides[72] and a heptapeptide[73] were developed based on observations that NK1 receptor antagonist can prevent and reverse morphine-induced tolerance[74], and the subsequent rationale that combining MOR agonist and NK1 antagonist activities could improve pain treatment while eliminating adverse side effects including tolerance and dependence[75]. Similarly, drugs with MOR agonist / KOR agonist profile may have the potential as non-addicting analgesics as the dysphoric effects by KOR agonists may counteract the euphoric effects of MOR agonists.[76] Drugs with MOR agonist / NOP agonist profiles[77] may also reduce opioid-induced side effects as NOP receptors can modulate opioid analgesia[78,79] and decrease opioid-induced dopamine release in rewards pathways[80,81].

The possibility of developing bi- and multi-functional drugs at opioid and other receptors are almost endless. For example, besides the interplay of opioid and non-opioid receptors mentioned above, interactions have been described between the opioid system and the serotonergic system[82,83], the dopaminergic system[84–86], and the cannabinoid system[87–90]. This grants great opportunities for new drug development, but also brings the challenge of understanding the real effect of bi- and multi-functional drugs. As a matter of fact, the endogenous neuronal systems controlling pain, rewards, and addiction are highly intermingled, and a new drug designed to target two or more receptors will often have unpredictable effects.

Most of the opioid-related bi- and multi-functional drugs developed to date were established through binding and activity assays in vitro, with a few exceptions[91] entering clinical trials, though with unfavorable outcomes. The true therapeutic value of this class of drugs is yet to

be established in the clinic. Future designers of bi- and multi-functional drugs may benefit from recent crystal structures of relevant GPCRs as well as computational drug design strategies[76] that are unavailable to early researchers. In addition, they will have to address the common problems of poor bioavailability, blood-brain barrier penetrability, and stability of bifunctional drugs, especially for the peptide-base ones[58].

Drugs Targeting MOR Isoforms

The existence of MOR isoforms, or mRNA splice variants, provide another possible avenue for drug development, although this has been less pursued than the other drug development strategies. The first suggestion of multiple MOR types came from early studies[92,93] in the 1970s and 1980s, where two MOR binding states with different affinities were discovered. Blocking of the higher affinity MOR type with antagonists changed the analgesic effect[94] of morphine in animals but did not alter morphine-induced respiratory suppression[94] or constipation[95,96]. It was later found that although morphine induces analgesia at both spinal and supraspinal sites through acting on MOR[97], the antagonist naloxonazine can only block the supraspinal analgesia in mice, but not the spinal analgesia[98,99]. In addition, the relative analgesic strengths of various MOR agonists were found to vary significantly between different mouse strains.[100] For example, morphine is a strong analgesic for BALB/c and CD-1 mice, but has little effect on CXBK mice. In contrast, for the CXBK strain, other MOR agonists such as fentanyl, methadone, and heroin all retained normal analgesic effect.[101] Taken together, these studies suggested the existence of multiple isoforms of MOR, which could explain the differential effects of opioids on physiological responses, animal strains, and different levels of pain.

In 1993, MOR was first cloned from rat.[102] Later studies identified only one *OPRM1* gene[103–106] but at least two additional mRNA variants[107,108], indicating alternative splicing of pre-mRNA. The group of Gavril Pasternak and Ying-Xian Pan at the Memorial Sloan Kettering Cancer Center in New York then extensively studied this phenomenon for over a decade and contributed to the discovery of a total of 33 splice variants in mouse, 17 in rat, and 19 in human.[109] These variants are generated from common alternative splicing patterns including exon skipping, alternative 5' and 3' splicing, intron retention, mutually exclusive exons, and alternative promoter and poly(A), and can be divided into three main categories: full-length 7-transmembrane (7TM) C-terminal splice variants, truncated 6-transmembrane (6TM) variants, and single transmembrane domain (TM) variants.[109] 7TM variants share all amino acid sequences from the N-terminus up until the C-terminal tail, but they have distinct C-terminal tails as a result of 3' splicing; 6TM variants are generated by 5' splicing and an alternative exon 11 promotor; single transmembrane variants are generated by exon skipping or insertion.

Importantly, the mRNA expression of the MOR splice variants was found to vary in different central nervous system locations as well as in different cellular locations.[110–113] For example, in the dorsal horn of the spinal cord, cells express either the MOR-1 or the MOR-1C splice variant, but not both. In this region, MOR-1 distribute equally in pre- and post-synaptic sites, but MOR-1C is exclusively found presynaptically.

The functions of different MOR splice variants also differ. For the 7TM variants with different C-terminal tails, opioid agonists show various efficacies. For example, morphine is a full agonist at the human MOR-1A variant, but partial agonist at human MOR-1B1, MOR-1B3, and MOR-1B5 variants.[114] Beta-endorphin is a full agonist for human MOR-1B5, but a partial agonist for human MOR-1A.[114] Evidence also showed that different 7TM variants have different G-protein

coupling and adenylyl cyclase inhibition activities[114], as well as different levels of morphine-induced receptor phosphorylation and internalization[115] in cell culture. Animal models also support the various roles of 7TM variants. Using antisense oligonucleotides to temporally knock down certain exons, it was found that exons have divergent roles in morphine-induced analgesia.[116] C-terminal truncated mouse models also showed the involvement of C-terminus in morphine-induced tolerance, reward, dependence, and desensitization[117–119], although these models did not necessarily reveal the role of any specific 7TM variant.

The single TM variant was first identified in cell culture and found to have a high mRNA level compared to the 7TM MOR-1 mRNA. A later study[120] showed that it acts as a molecular chaperone for the 7TM MOR-1. Through heterodimerization with MOR-1 in the endoplasmic reticulum (ER), single TM MOR facilitates MOR-1 folding and increases its expression. Downregulating a single TM variant in vivo led to reduced MOR-1 mRNA and protein expression as well as diminished morphine analgesia.[120] However, there is no evidence to date that the single TM variants interact with any opioids or have any downstream signaling.

The 6TM variants also have distinct functions from the 7TM variants. In mouse models that lack the 6TM variants, morphine and methadone showed normal analgesia, but the analgesic effect of heroin and fentanyl was greatly diminished.[121] In addition, although morphine-induced analgesia[121,122], reward[122], and respiratory suppression[122] were not affected in the 6TM knock-out (KO) or downregulated mice, morphine-induced hyperalgesia[122,123] and hyperlocomotion[122] were diminished. The 6TM KO mice also developed morphine tolerance more slowly than wild type mice[122]. Together, these lines of evidence indicate a role of the 6TM variants in mediating the analgesic effects of a sub-class of opioids as well as certain non-analgesic effects of morphine.

The only opioid ligand developed to date that targets a specific MOR isoform is iodobenzoylnaltrexamide (IBNtxA). First synthesized[124] in 2011, IBNtxA was later found to act on the 6TM MOR variants: 6TM MOR KO mouse strain lost IBNtxA analgesia[125], but a triple KO mouse strain that expresses no DOR, KOR, or 7TM MOR had normal IBNtxA analgesia[126]. In addition, in mice expressing only the 6TM MOR variant, IBNtxA had analgesic effect, but not morphine, fentanyl, or buprenorphine.[127,128] Taken together, these results suggest that the 6TM MOR variants are both necessary and sufficient for IBNtxA analgesia.

The drug profile of IBNtxA on animals are promising, as it was found to suppress pain while having no respiratory suppression, dependence, or reward effects.[125,126] The original study in 2011 suggested IBNtxA as a prototypical molecule for drug development, but not a drug for clinical development itself.[126] At at the time of writing of this thesis, there is yet any attempt to carry IBNtxA or similar molecules into the clinical phase, possibly due to its complex and still obscure opioid receptor pharmacology in vivo.[129]

Overall, MOR isoforms represent another layer of complexity of the endogenous opioid system and a unique therapeutic opportunity. Although it has been almost three decades since the discovery of the first MOR isoforms[107,108], only one drug[124,126] that targets a specific isoform category has been discovered, and none has entered the clinical trial. Due to the high similarities among variants[109], localization or functional study of any single variant remains challenging. Similarly, developing drugs that target a specific isoform is difficult as many isoforms share the exact same extracellular sequences. It is also unknown what percentage of MORs in human or animals are isoforms other than MOR-1[109], so the effect and value of any isoform-specific drug remain to be explored. Our understanding of the MOR isoforms and currently available drug design strategies seem insufficient in effectively targeting these isoforms in a clinical relevant

manner. Nonetheless, the current knowledge in MOR isoforms provides important insights into how opioid system works, and should be taken into consideration when evaluating any new opioid drugs.

1.4 Additional Insights of the Opioid System are Needed to Develop Opioids with Minimal Side Effects

The unsuccessful attempts so far to develop opioids with minimal side effects highlight the challenging nature of this quest. As a matter of fact, we currently only have limited understanding of what causes the various side effects of opioids. Biased agonism, allosteric modulation, bifunctional drugs, and opioid receptor isoforms each offer an opportunity for therapeutic development, but an ultimate and effective solution likely requires a more thorough understanding of the origin of each opioid-related physiological consequence.

Building upon the knowledge already mentioned in sub-chapter 1.2, this sub-chapter briefly summarizes what we currently know about the endogenous opioid system and highlights current techniques available to study this system and its effects. Briefly, receptor types and localization, receptor functions, and neuronal circuits responsible for specific physiology have been well-studied using a variety of currently available techniques. However, it calls for new techniques to reveal details such as the role of endogenous opioid receptors in each neuronal circuit, the role of intracellular opioid receptors, functions of opioid receptor heteromers, as well as the function of each MOR isoform in vivo.

Current Knowledge on the Endogenous Opioid System

The opioid receptors form arguably the most complex GPCR system in the brain. Three types of opioid receptors are generally considered to exist: MOR, KOR, and DOR. Some also consider the NOP receptor to be a fourth type of opioid receptor, although typical opioid ligands do not activate NOP, and NOP's ligand, nociceptin/orphanin FQ, does not activate any of the three main types of opioid receptors. In general, all three types of opioid receptors and their endogenous ligands can be found throughout the nervous system[130], having distinct but overlapping localizations. The endogenous opioid system is responsible for reducing responses to pain and stress, as well as for influencing reward and mood, among many other physiological and behavioral functions.[131] Each type of opioid receptor carries different but related functions, and most of the endogenous opioid ligands have cross reactivities at two or more opioid receptor types.

MOR was first cloned[102] in 1993, and was soon found to be the main opioid receptor for opioid analgesia. In addition to its major analgesic effect, MOR also have effects of reward, addiction, respiratory suppression, constipation, tolerance, dependence, and withdrawal. MOR is the sole receptor responsible for the analgesic effects and aversive effects of morphine[21], and is a target for other clinical or illicit opioids such as heroin, fentanyl, methadone, and oxycodone[132].

In contrast to MOR's euphoric effects, KOR is known for its aversive, or "anti-reward" effects[133,134]. First cloned[135] in 1993, KOR activation induces dysphoria, deteriorate mood, and psychotomimetic effects[136,137]. KOR is tightly related to stress[138] and has a role in withdrawal symptoms[139], addiction[140,141], and loss of motivation[142]. DOR was first cloned[143] in 1992. It is not essential for opioid reward, but mediates mood and learning. DOR normally helps alleviate negative mood associated with withdrawal[144,145], and regulate learning and memory[146,147].

Endogenous opioid ligands are also highly complex. Most if not all opioid peptides are derived from three precursors, proopiomelanocortin (POMC), proenkephalin (PENK), and prodynorphin (PDYN). However, none of these opioid peptide families are responsible for activating a single type of opioid receptor. Some peptides have certain selectivity towards one type of opioid receptor, but some have high affinity at all three types of opioid receptors.[148] POMC, PENK, and PDYN produce a total of ~30 different opioid peptides through specific cleavage, and most of these peptides have cross-reactivity at two or more opioid receptors, as summarized in **Table 1-1**.

Table 1-1 List of endogenous opioid peptides by precursor proteins, with amino acid sequence, number of amino acids, and K_i values (in nM) at three opioid receptors.

Unless specified, K_i values are from reference [148].

Peptide or precursor	Sequence (# of amino acids)	K_i at MOR (nM)	K_i at MOR (nM)	K_i at MOR (nM)
Proenkephalin peptides				
Leu-Enkephalin	YGGFL (5)	6.19 ± 1.00	0.37 ± 0.04	93.65 ± 13.39
Met-Enkephalin	YGGFM (5)	1.80 ± 0.15 5.9 ± 0.9[149] 9.5 ± 0.54[150,151]	0.45 ± 0.03 0.91 ± 0.07[150,151]	47.44 ± 11.99 > 1000[150,151]
Met-Enkephalin-Arg-Phe	YGGFMRF (7)	0.37 ± 0.12	0.57 ± 0.04	0.34 ± 0.11
Met-Enkephalin-Arg-Gly-Leu	YGGFMRGL (8)	1.40 ± 0.35	0.52 ± 0.18	0.48 ± 0.14
Metorphamide (Met-Enkephalin-Arg-Arg-Val-NH2)	YGGFMRRV**-NH2** (8)	0.08 ± 0.02 0.12 ± 0.02[151]	1.05 ± 0.05 2.65 ± 0.23[151]	0.04 ± 0.01 0.25 ± 0.03[151]
Bam 12	YGGFMRRVGRPE (12)	1.05 ± 0.20	2.62 ± 1.08	1.67 ± 0.49
Bam 18	YGGFMRRVGRPEWWMDYQ (18)	1.95 ± 0.54	5.15 ± 0.48	5.04 ± 2.22
Bam 22	YGGFMRRVGRPEWWMDYQ KRYG (22)	0.86 ± 0.09	2.39 ± 0.23	1.95 ± 0.27
Peptide E	YGGFMRRVGRPEWWMDYQ KRYGGFL (25)	1.04 ± 0.30	2.80 ± 0.67	1.55 ± 0.33
Peptide F	YGGFMKKMDELYPLEVEE EANGGEVLGKRYGGFM (34)	55.57 ± 10.26	6.58 ± 1.11	>1000

Prodynorphin peptides				
Dynorphin A 1-6	YGGFLR (6)	1.88 ± 0.07	1.04 ± 0.23	2.88 ± 0.47
Dynorphin A 1-7	YGGFLRR (7)	1.80 ± 0.38	2.71 ± 0.75	0.21 ± 0.06
Dynorphin A 1-8	YGGFLRRI (8)	2.56 ± 0.19	1.17 ± 0.20	0.45 ± 0.28
Dynorphin A 1-9	YGGFLRRIR (9)	2.62 ± 1.12	1.47 ± 0.58	0.13 ± 0.06
Dynorphin A 1-13	YGGFLRRIRPKLK (13)	4.14 ± 1.65	5.56 ± 1.93	0.23 ± 0.06
Dynorphin A 1-17 / Dynorphin A	YGGFLRRIRPKLKWDNQ (17)	2.64 ± 0.75 2.0 ± 0.5[149] 0.73[150]	1.29 ± 0.23 2.4[150]	0.37 ± 0.05 0.12[150]
Dynorphin A 1-32 / Dynorphin AB / Big Dynorphin	YGGFLRRIRPKLKWDNQK RYGGFLRRQFKVVT (32)	6.87 ± 0.09	3.34 ± 1.05	3.08 ± 0.82
Dynorphin B 1-13	YGGFLRRQFKVVT (13)	2.19 ± 0.48	3.48 ± 0.70	0.45 ± 0.12
Dynorphin B 1-29 / Leumorphin	YGGFLRRQFKVVTRSQQD PNAYYGGLFNV (29)	4.03 ± 1.60	3.89 ± 0.80	0.48 ± 0.05
α-neo Endorphin	YGGFLRKYPK (10)	1.14 ± 0.16	1.17 ± 0.13	0.13 ± 0.03
β-neo Endorphin	YGGFLRKYP (9)	5.34 ± 1.20	1.14 ± 0.37	0.48 ± 0.12
Proopiomelanocortin peptides				
β-Endorphin 1-27	YGGFMTSEKSQTPLVTLF KNAIIKNAY (27)	5.31 ± 0.29	6.17 ± 0.13	39.82 ± 5.34
β-Endorphin 1-31 / β-Endorphin	YGGFMTSEKSQTPLVTLF KNAIIKNAYKKGE (31)	3.73 ± 0.76 2.1[150] 4.4 ± 0.41[149]	5.02 ± 0.21 2.4	32.70 ± 6.00 96
Peptides with unknown precursor				
Endomorphin-1	YPWF**-NH2** (4)	0.36 ± 0.08[149] 0.67 ± 0.2[152] at Mu1 3.2 ± 0.5[152] Mu2	> 1000[149] > 500[152]	> 1000[149] > 500[152]
Endomorphin-2	YPFF**-NH2** (4)	0.69 ± 0.16[149] 0.43 ± 0.05[152] at Mu1 4.0 ± 0.05[152] at Mu2	> 1000[149] > 500[152]	> 1000[149] > 500[152]
Tyr-MIF-1 (Tyr-Pro-Leu-Gly-NH2)	YPLG**-NH2** (4)	987.8[153]	>1000[153]	>1000[153]
Tyr-W-MIF-1 (Tyr-Pro-Trp-Gly-NH2)	YPWG**-NH2** (4)	70.9[153]	>1000[153]	>1000[153]

Knockout Techniques Studied the Loss of Function of Opioid Receptors and Pathways

Knockout is one of the first class of techniques used by researchers to reveal the specific function of opioid receptors and their pathways. In a seminal paper[21] published in 1996, Matthes

and colleagues established the role of MOR in morphine analgesia and side effects. In this highly cited paper (over 1800 citations as of March 2022), the authors found that in MOR KO mice, both the analgesic and side effects of morphine were lost.[21] Other opioid receptor KO mice also revealed the role of these receptors. For example, conditional KO of KOR from mesolimbic dopaminergic neurons reduced conditional place aversion to KOR agonists[154], revealing the role of KOR in fine-tuning dopamine tone; DOR KO mice show intact morphine self-administration and depressive-like phenotype[155,156], indicating that DOR is not necessary for opioid reward but may be related to reducing negative mood.

In addition, KO mice can also be used to study downstream pathways of the opioid system. For example, as mentioned in sub-chapter 1.2, a series of important studies[22–25] around the year of 2000 established the role of β-arrestin2 in morphine-induced side effects, as β-arrestin2 KO mice did not develop constipation and respiratory suppression in these studies (contradicting evidence was found later by other groups).

Although knock-out is a well-established technique to study whether a receptor or other protein is necessary for certain opioid effects, it does have several limitations. First, new KO studies require the generation of new mouse lines, which can be limited by currently available techniques and is time-consuming, though the recent development of CRISPR[157] should relieve this constraint. Second, KO techniques, especially in the early days, may not completely remove all target receptors and proteins, causing possible false positive results. Third and most importantly, KO does not allow the study of receptor activation effects. While it can establish whether a receptor is *necessary* for certain physiological outcomes, it does not reveal whether activation of this receptor alone is *sufficient* for that outcome; that is, the causal effect of receptor activation cannot be studied.

Knock-in Techniques Revealed Localization of Opioid Receptors

Opioid receptors are present throughout the brain, and studying their localization both on the system level and on the cellular level is challenging. In the early years, researchers have been using radioligand binding assays[158,159] to identify opioid binding sites. However, this assay has low sensitivity due to the low abundance of opioid receptors. Most of the radioligands also have cross reactivity at multiple opioid receptors, resulting in a lack of specificity. Another common technique was to detect opioid receptors in brain homogenates. By using homogenates from different brain regions, binding difference between different brain structures were established.[158,160,161] While brain homogenate studies can reveal opioid receptor distribution in a quantitative way, they cannot obtain higher temporal resolution.

In contrast, knock-in technique can be specific, quantitative, and can study opioid receptor localization at both the brain structure level and the cellular level. For example, by using homologous recombination in knock-in mouse, DOR-EGFP (evolved green fluorescent protein) and MOR-mCherry can be expressed following the native expression pattern of endogenous DOR and MOR, respectively.[162] Expression of such fluorescent receptors also did not alter animal behaviors or opioid effects.[163] The fluorescent tags readily allowed the localization and quantification of both DOR and MOR throughout the mouse brain, and the overlap of the two colors enabled identification of possible DOR/MOR interaction and heterodimerization sites.[162] In addition, subcellular localization of opioid receptors became possible. It was found that under basal conditions, DOR mainly localized at the plasma membrane while MOR has a large intracellular pool.[162] Knock-in techniques are excellent in revealing receptor localizations, but itself cannot be used to study receptor functions. For functional assays, either removing (knocking-out) the receptors or activating the receptors is required.

Microinjection Techniques Linked Brain Regions to Opioid Effects

Due to the omnipresence of opioid receptors throughout the nervous system, it has been a topic of interest to study the role of receptors in different brain regions or cells. While the overall effects of opioids have long been known through global administration, the causal relationship between brain regions and opioid effects is a more recent discovery. The predominant technique for studying brain region – opioid effect relationship is microinjection. By delivering opioids to different brain regions using stereotaxic injection, opioid effects on animals can be revealed. For example, by injecting a panel of opioids into the periaqueductal gray region of the rat brain, the differential effects of morphine, fentanyl, oxycodone, methadone, and buprenorphine were established.[164] It was found that the first three opioids had analgesic effect while the last two did not show this effect[164], suggesting alternate MOR binding sites for methadone and buprenorphine in rat. Similarly, injecting [D-Ala2, N-MePhe4, Gly-ol]-enkephalin (DAMGO) or morphine into the raphe magnus region of the brain caused a decrease in both respiratory rate and heart rate in mice, indicating the possible role of raphe magnus in opioid-induced respiratory suppression.[165]

Microinjection is a powerful technique for studying brain region-specific opioid effects, but it also has several limitations. For example, there is no way to control the diffusion of opioid molecules once they are injected into the brain. Since the brain has a fluidic environment, there is no guarantee that the observed effects are produced solely by the targeted brain structure. In addition, even in the same brain structure, there can be multiple types of neurons and glial cells that may all express opioid receptors. Microinjection cannot differentiate opioid receptors on different cell types or in different neuronal circuits if they are in close proximity. However, it is very well possible that opioid receptors on different cell types or neuronal circuits have distinct roles.

Genetic Tools Revealed the Role of Neuronal Circuits on Physiological Consequences

The recent advancement in genetic techniques has allowed the study of opioid-related effects on the level of cell type or neuronal circuit for the first time. With cell type-specific promoters[166] or the Cre-*lox*P (Cre recombinase and locus of x-over, P1) system[167,168], genetically encoded tools can be expressed in specific cells or neuronal circuits without affecting nearby, untargeted neuronal population. By designing genetic tools to activate or inhibit neurons, the function of such cells and neuronal circuits can be studied at an unprecedented spatial resolution.

The most widely used genetic tool for this purpose is Designer Receptors Exclusively Activated by Designer Drugs (DREADDs). First designed[169] in 2007, DREADDs are mutated GPCRs that do not recognize the endogenous ligand, but recognize a synthetic molecule, clozapine *N*-oxide (CNO). The orthogonality of DREADDs can be achieved by introducing as little as one or two point mutations into conventional GPCRs.[169,170] Since its first introduction, DREADDs are soon highly recognized by the neuroscience community[171], and there is a range of excitatory[169] and inhibitory[169,170] DREADDs available for use today.

The use of DREADDs has been demonstrated in a range of neuroscience studies[171] including the study of opioid-related effects and neuronal circuits. For example, an inhibitory DREADD[170] has been used to study motivated behavior in rats[172]. It was found that inhibiting putative dopamine neurons in the midbrain eliminated tonic firing, and selective inhibition of ventral subiculum neurons projecting to nucleus accumbens shell decreased context-induced Fos protein expression.[172] DREADDs has also been used to study neuronal circuits involving reward[173], addiction[174], analgesia[175], and breathing[176], in which the endogenous opioid system may also be involved.

Although DREADDs provide unprecedented temporal resolution to study endogenous cells and neuronal circuits, they do have two caveats. First, the ligand for most DREADDs, CNO, was found to be not entirely orthogonal, as it can be reverse-metabolized into clozapine and produce stimulus effects in animals[177], possibly complicating the interpretation of experiment results. Second, DREADDs are an artificial system designed to study neuronal circuits in an orthogonal manner. While having the advantage of not interfering with the endogenous receptors and ligands, they also lose the potential to study the endogenous system. The overexpression of DREADDs does not mimic the true expression pattern of any endogenous receptors. Therefore, it is unknown whether the effects observed from the reward, analgesia, and addiction circuits are relevant to the opioid system.

In summary, DREADDs are an attractive tool to identify the effects of activating or inhibiting previously understudied neuronal circuits, but to establish the role of opioid receptors in these circuits, a different class of techniques will be needed.

1.5 Knowledge Gap that this Thesis Aims to Fill

Since the late 20th century, out knowledge about the endogenous opioid system has been rapidly expanding: three types of opioid receptors and more than 20 endogenous opioid ligands have been identified; the overall role and localization of each type of opioid receptor have been revealed; the effects of activating opioid receptors in certain brain regions have been established; the specific roles of each neuronal circuit in many opioid-related effects are also being gradually revealed. In addition to these studies on the system level, many fine details about the endogenous

opioid system and its downstream effects have also been discovered: biased signaling, intracellular opioid receptors, splice variants, heteromers, and allosteric modulation to name a few.

However, the development of a novel opioid drug with analgesic effects but not adverse side effects is yet successful. To continue understanding the endogenous opioid system, new classes of tools are needed, and we aim to fill this gap by developing genetic tools through protein engineering. One important missing piece from our understanding of the opioid system is the causal relationship between neuronal circuits and opioid effects. Although microinjection has revealed the link between brain region and opioid effects, it does not have the capability of studying each individual circuitry. In contrast, although DREADDs are excellent genetic tools capable of studying neuronal circuits, they are artificial proteins and do not represent the opioid receptors or ligands.

In this thesis, we aim to bridge the advantage of opioid microinjection and DREADDs by developing a new class of genetic tools that are capable of activating endogenous opioid receptors in specific neuronal circuits. These successful development of these tools will have three indications. First, this will provide a first glimpse into the role of individual neuronal circuits in opioid effects. We aim to provide neuroscientists with new technologies to dissect the function of each opioid receptor-expressing neuronal circuit. Second, creating such tools requires novel protein engineering methods and designs. Through our engineering efforts, we aim to provide protein engineers with new platforms and methodologies. We also aim to engineer tools with a wide range of utility beyond our prototypical purpose in the hope of benefiting a larger audience. Lastly, since our designs can interact with the endogenous opioid receptors, they themselves may serve as a starting point for creating a new class of opioid drugs. This is foreseeably a multi-decade

effort, but we hope our creation of these new tools will spark new innovations in drug design and development.

1.6 Dissertation Overview

This dissertation describes the engineering of a class of genetic tools termed M-PROBE. M-PROBE aims to provide a means of activating opioid receptors, especially MOR, in a cell type-specific or neuronal circuit-specific manner for the study of causal relationships between receptor activation and opioid effects. As will later be shown in **Figure 2-1**, M-PROBE has three components: an opioid peptide for activating opioid receptors, a protein switch for controlling the activity of the opioid peptide, and a transmembrane domain for displaying the opioid peptide on the outer cell membrane. To provide versatile control of M-PROBE, two types of protein switches were designed: a photoswitch and a chemoswitch, controlled by light and a small chemical molecule, respectively.

Chapter 2 describes the general design of M-PROBE and the engineering of a photoswitch: a protein that activates the opioid peptide in light but deactivates the peptide in the dark. The general design includes selection of an opioid peptide, a transmembrane domain, linkers, and a surface trafficking signal peptide for M-PROBE. I describe the rationale behind each selection and the supporting experimental results. In the second part, I describe the engineering of a photoswitch. I first explain the rational design of this photoswitch and characterization with different cellular assays. I then describe how directed evolution was applied to adjust the activity of the photoswitch. At the time of completion for this dissertation, the photoswitch for M-PROBE

has been proved to work on two short peptides, but I have not yet obtained evidence of it controlling an opioid peptide. The general utility of this photoswitch on non-opioid peptide and proteins will also be discussed.

Chapter 3 describes the engineering of a chemoswitch for controlling opioid peptides. Unlike the photoswitch, this protein switches its conformation upon binding of a small molecule. This molecule is then used as a control mechanism for activating a basally inactive opioid peptide. I describe the rational design and extensive directed evolution efforts to improve the chemoswitch. We have proved the utility of this chemoswitch on three peptides as well as on an opioid peptide. The general applicability of this chemoswitch will be described, as well as how it was used to control opioid peptides.

Chapter 4 will provide a summary of the main contributions of this dissertation, and suggest future directions for protein engineers and neuroscientists based on the tools described in this thesis.

Both the photoswitch (**Chapter 2**) and the chemoswitch (**Chapter 3**) were developed as a team with Jiaqi Shen, Ph.D. student in the Wenjing Wang Lab at the University of Michigan. Additional collaborators are mentioned at the beginning of each chapter.

Chapter 2 Developing Photoswitchable, Cell Membrane-Tethered Opioid Peptides[1]

2.1 Overview of M-PROBE

The overall design of M-PROBE is shown in **Figure 2-1**. A single chain genetic construct encodes a three-component protein with linkers connecting each component. The entire construct is displayed on the outer cell membrane by a transmembrane domain, allowing accessing the opioid receptors from the extracellular side. The N-terminus of the construct is an opioid peptide. Right after the opioid peptide is an engineered protein switch whose activity can be controlled by either light or a small molecule. This protein switch is essential because it allows temporal control

[1] Partially reproduced from: Geng, L., Shen, J., & Wang, W. Circularly permuted AsLOV2 as an optogenetic module for engineering photoswitchable peptides. *Chem. Commun.* **57,** 8051-8054 (2021), with permission from the Royal Society of Chemistry. Also related to: Geng, L., Kroning, K. & Wang, W. SPARK: A Transcriptional Assay for Recording Protein-Protein Interactions in a Defined Time Window. *Curr. Protoc.* **1(7),** e190 (2021). Experiments in this chapter are in collaboration with Jiaqi Shen and Kayla Kroning, Ph.D. students in the Wenjing Wang Lab, University of Michigan.

of the opioid peptide activity, correlating the time of receptor activation to the time of physiological effect. It also reduces desensitization and internalization due to prolonged receptor activation.

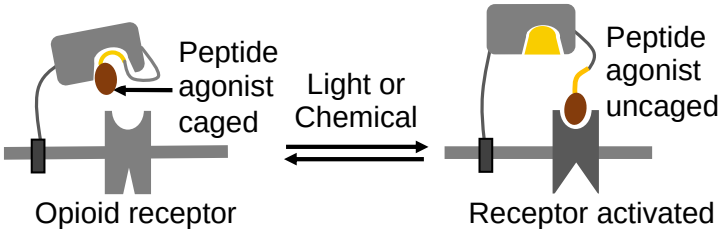


Figure 2-1 Overall design of M-PROBE. M-PROBE is made of a transmembrane domain, a photo- or chemical-switch, and an opioid peptide agonist controlled by the protein switch.

The major engineering effort for M-PROBE was on the protein switch, as neither the photoswitch nor the chemoswitch was readily available prior to our research. We chose light and chemical as the controlling mechanisms because they provide unique and complemental utilities: light is one of the fastest control mechanisms and allows reversibility, which is well-suited for studying the immediate opioid effects on the order of minutes or less; small molecules can be globally administered into animals and control opioid receptor activity in a large brain area or volume, which is well-suited to study opioid effects with a longer term or involve multiple brain regions. Together, both mechanism should provide complemental and versatile means of control for a range of neurological studies.

2.2 Design and Optimization of M-PROBE Opioid Peptide, Transmembrane Domain, Linkers, and Surface Trafficking Signal Peptide

Design of M-PROBE without a Protein Switch

The initial M-PROBE design involved an opioid peptide displayed on the outer cell membrane without a protein switch (**Figure 2-2**). This design is termed M-PROBE1.0. Besides the reasons that a protein switch was unavailable at that time, this design tested the following four pre-requisites for M-PROBE. First, a transmembrane protein must be capable of displaying M-PROBE on the outer cell membrane. Second, the opioid peptide must retain its activity even when its C-terminus is fused to a protein. Third, M-PROBE may be displayed with or without a surface trafficking signal peptide. If such peptide is used, it must be cleaved right before the N-terminus of the opioid peptide, because structural studies[178] showed that the N-terminal residue of opioid peptides are essential for entering the opioid receptor binding pocket. Lastly, the length of the linker must be appropriate to allow the opioid peptide to enter the receptor, without restricting the reach of the peptide or projecting it too far away from the cell surface.

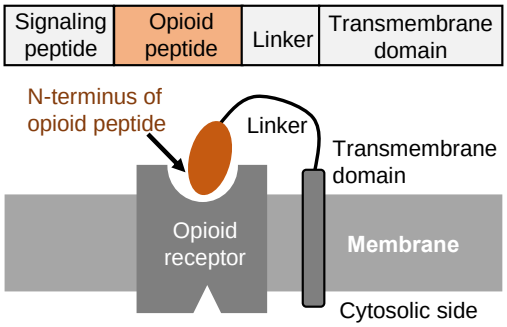


Figure 2-2 Design of M-PROBE1.0, an initial version of M-PROBE with only a membrane-tethered opioid peptide and no control mechanism. The surface trafficking signal peptide is not shown on the cartoon because it is designed to self-cleave after its C-terminus.

A Transcriptional Assay as M-PROBE1.0 Activity Readout

To test the activity of M-PROBE1.0, I created a transcriptional readout assay in human embryonic kidney (HEK) 293T cells together with Kayla Kroning, Ph.D. student in the Wenjing Wang Lab at the University of Michigan. This assay, originally termed Specific Protein Association tool giving transcriptional Readout with rapid Kinetics (SPARK, **Figure 2-3**), was based on a previously described system[179,180] developed by Wenjing Wang, my Ph.D. advisor.

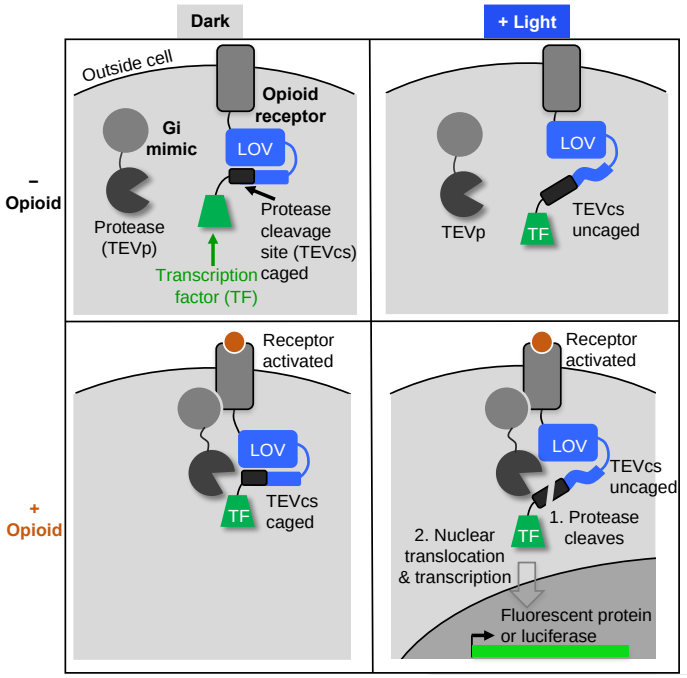


Figure 2-3 A light- and opioid-gated transcriptional reporter for detecting the activation of opioid receptors. A long-lasting transcriptional signal is generated after ~12 hours only when there was concurrent presence of light and opioid for the minimum required length of time. Opioid receptor activation leads to recruitment of Gi-mimic nanobody to the receptor, and light leads to uncaging of TEVcs. LOV, *As*LOV2 domain from common oat. TF, transcription factor.

SPARK is a light and ligand double-gated system. In the basal state (**Figure 2-3**, upper left), the transcription factor is inactive because it is tethered to the plasma membrane as part of a single construct of the artificially-expressed opioid receptor. Activation of the opioid receptor by opioids (**Figure 2-3**, lower left) leads to recruitment of a G_i-protein mimic, nanobody 39 (Nb39)[181], to the

receptor. This brings the tobacco etch virus protease (TEVp) into close proximity of the TEVp cleavage site (TEVcs). However, cleavage does not happen in the dark state because TEVcs is sterically blocked by a photoswitch termed a light, oxygen, and voltage sensing domain (LOV). When there is light (**Figure 2-3**, lower right), the TEVcs is uncaged from the photoswitch and becomes accessible by the TEVp, resulting in cleavage of TEVcs and release of the transcription factor into the nucleus. This eventually leads to the expression of an artificial reporter gene, such as the mCitrine fluorescent protein.

Using free, untethered opioids, we established this system for detecting the activation of all three types of opioid receptors (**Figure 2-4**). Typically, an appropriate amount of in-house prepared lentiviruses encoding the three genetic components (the opioid receptor – LOV – TF component, the TEVp component, and the reporter gene component) are used to transduce HEK 293T cells. About 2 days after lentiviral transduction, cells are stimulated with saturating concentrations (typically 10 μM) of agonist under light for 10-30 minutes. The agonists are then washed out, and cells are incubated overnight to allow sufficient time for gene transcription. The next day, fluorescent microscopy images are taken from a Nikon confocal microscope. When mCitrine is used as the reporter gene, we typically use mCherry as an expression marker, which is co-expressed with the TEVp component via a 2A self-cleaving peptide (specifically, P2A) in a single DNA construct. The other components do not have an expression marker.

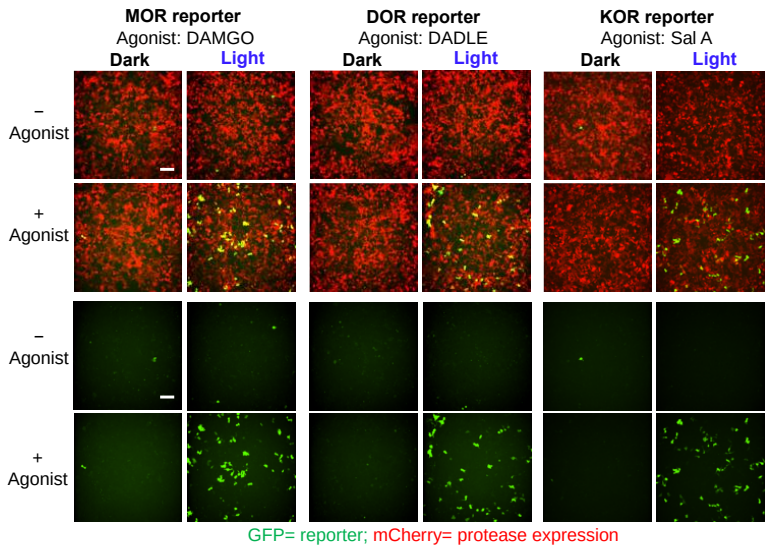


Figure 2-4 Reprehensive results from the transcription reporters for opioid receptor activation. Similar assays were developed for all three types of opioid receptors. All genetic constructs were transduced into HEK 293T cells with lentiviruses. Cells were stimulated with light and/or 10 μM of the indicated drug for 20 minutes, and imaged 12 hours after stimulation. MOR, mu-opioid receptor. DOR, delta-opioid receptor. KOR, kappa-opioid receptor. DAMGO, [D-Ala^2, N-MePhe^4, Gly-ol]-enkephalin, an MOR agonist. DADLE, [D-Ala^2, D-Leu^5]-enkephalin, a DOR agonist. Sal A, salvinorin A, a KOR agonist. Scale bars, 100 μm.

As shown in **Figure 2-4**, the experimental results were consistent with the design. Transcription of mCitrine only happened at the co-occurrence of light and ligand. Absence of either component led to significantly lower mCitrine expression. Note that the activation efficiency (defined by the ratio of mCitrine-expressing to mCherry-expressing cell count) was rather low in **Figure 2-4**. As a three-component system, SPARK was found to be highly expression-level dependent, and any variation in lentivirus activity or protein expression could lead to altered activation efficiency. In addition, we found that the affinity between Nb39 to activated opioid receptors was also a reason of generally low activation efficiency. When compared with the dopamine receptor D1 (DRD1) – β arrestin2 system, the opioid receptor – Nb39 system consistently produced less activation. To further characterize the opioid SPARK system, I applied a panel of opioids on MOR and quantified the activation efficiency (**Appendix 1**). Cell numbers

were counted using the Nikon image analysis software, and the ratio of mCitrine-to-mCherry-expressing cell numbers was calculated from 10 fields of view for each agonist condition. As shown by **Appendix 1**, with a full agonist such as fentanyl, an activation efficiency of ~40% can be achieved under the most optimized experimental condition. With less potent opioids such as leu-enkephalin and morphine, only ~20% activation efficiency can be obtained, and the activation efficiency is in general consistent with the potency of these opioids. With the ideal experimental conditions, the signal-to-noise (S/N) ratio of "+ drug" and "– drug" conditions under light was generally above 20, and the S/N of "+ light" and "– light" conditions in the presence of drug was generally above 10, showing good response to both ligand and light. Importantly, although activation efficiency varied from experiment to experiment, we found it to be consistent within the same experiment. In experiments where the activation efficiency was lower, the background signal was also lower. Therefore, comparison between experiment conditions should be valid as long as all data were collected from the same set of experiments.

The SPARK system has two major advantages for the purpose of testing M-PROBE1.0. First, SPARK provides a transcriptional, cumulative readout. Once opioid receptors are activated above a certain threshold, the reporter gene can be expressed for at least a day, allowing measurement of receptor activation at a later time. This is especially advantageous for M-PROBE1.0 because this version of M-PROBE is always "on", and measuring its effects using real-time assays is likely difficult. Second, SPARK provides an adjustable time window that allows fine-tuning the signal magnitude. Using free opioid agonists, we found 20 minutes to be the optimal time for the best signal-to-noise ratio: shorter time led to lowered signal with no significant reduction in background, and longer time led to higher background with no apparent increase in signal. This time frame is longer than other SPARK assays[180] such as the ones for the dopamine

D1 receptor or the β_2-adrenergic receptor (which typically take 1-10 min), likely due to the lower affinity between the TEVp and the GPCR components. If the opioid peptide activity on M-PROBE1.0 changes due to C-terminal fusion, we could adjust the light stimulation time to bring the signal into the proper range.

Screening a Panel of Opioid Peptides from Three Precursors for M-PROBE1.0

With the SPARK assay, I aimed to identify the best opioid peptide to use on M-PROBE for MOR. MOR was chosen because it is the most relevant opioid receptor for pain modulation as well as perhaps the most well-studied opioid receptor. As listed in **Table 1-1** (clickable link), there are more than 20 opioid peptides that can potentially activate MOR, with various K_i values. With the exception of endomorphins, which have unknown precursors[182], most of these peptides are from three precursors, POMC, PENK, and PDYN. In general, PENK peptides have lower K_i values and higher affinity at MOR, followed by PDYN and POMC peptides, but most peptides from all three precursors have cross-reactivity at MOR.

I first selected several representative opioid peptides from each precursor and tested them using MOR SPARK. All experimental procedures were the same as the regular SPARK except that an additional lentivirus encoding M-PROBE1.0 was used to transduce M-PROBE1.0 into HEK 293T cells. For this test, I used the amino sequence `MKTIIALSYIFCLVFA` as a self-cleaving surface trafficking signal peptide[178]. A human CD4 protein with most of the extracellular domain truncated (`LPTWSTPVQPMALIVLGGVAGLLLFIGLGIFFCVRCRHRRR`) was used as the transmembrane domain, and a flexible 17-amino-acid linker made up with glycine and serine was used to connect CD4 with the opioid peptide.

As shown by **Figure 2-5**, many of the opioid peptide tested showed activity from the SPARK assay. Consistent with the general high affinity of free PENK peptides towards MOR, these peptides also showed higher MOR activation than POMC or PDYN peptides when tethered to the cell membrane. Interestingly, endomorphin-1 and endomorphin-2, the only endogenous opioid peptides with high selectivity at MOR, showed little to no MOR activation when tethered to the cell surface. This is possibly due to the loss of the native C-terminal amidation from fusion with another protein.

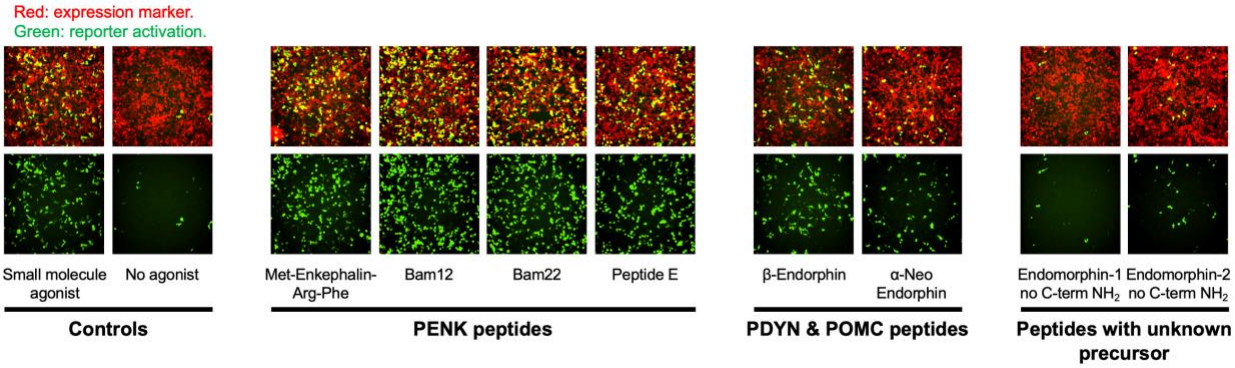


Figure 2-5 Screening opioid peptides mostly from three precursors for activating MOR, using the MOR activation transcription reporter assay. "Small molecule agonist" refers to DAMGO ([D-Ala^2, N-MePhe^4, Gly-ol]-enkephalin). PENK, proenkephalin. PDYN, prodynorphin. POMC, proopiomelanocortin.

Together with immunostaining data for M-PROBE (**Figure 2-6**), these results confirmed the following: First, M-PROBE can be displayed on the outer cell membrane with apparent membrane pattern. Second, the surface trafficking signal peptide we chose is indeed self-cleavable after its C-terminus, allowing the N-terminus of the opioid peptide to enter the MOR binding pocket. Third, opioid peptides that are tethered to the cell surface can still activate MOR and at least recruit a G_i-mimic protein. However, C-terminally amidated opioid peptides may lose their activity in this geometry. Lastly, opioid peptides from the PENK precursor generally cause higher

activation at MOR, although peptide with POMC and PDYN origins can also activate MOR to a lesser extent.

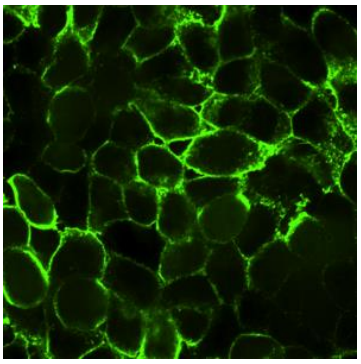


Figure 2-6 Representative immunostaining results for M-PROBE. HEK 293T cells were transduced with M-PROBE DNA using lentivirus. Two days after viral transduction, cells were cross-linked by formaldehyde and membrane-depleted by methanol. Expression pattern of M-PROBE was detected by immunostaining against an intracellular HA epitope tag using a rabbit anti-HA tag primary antibody followed by a goat anti-rabbit secondary antibody conjugated with Alexa Fluor 488.

I then moved on to investigate the shortest possible peptide sequence for satisfactory MOR activation. This is because shorter peptides are generally easier to be blocked by a protein switch. According to previous engineering methods[183,184] using a photoswitch, serial truncation is the best method to determine the necessary residues in an effector peptide or protein. By removing unimportant residues from the fusion site with the protein switch, key functional residues will generally be better "blocked" by the protein.

I started with the 12-amino-acid Bam12 peptide, because it shares five to eight N-terminal residues with shorter PENK peptides (**Table 1-1)** and its activity on MOR is similar to longer PENK peptides such as Bam22 and Peptide E (**Figure 2-5**). Eight serial truncations were made on Bam12, and peptides from Bam12(1-4) to Bam12 were tested in the same round of MOR SPARK assay. It was found that Bam12(1-5), or met-enkephalin, retained its activity towards MOR, but Bam12(1-4) lost most of its activity (**Figure 2-7**). The published K_i values[148] of Bam12 and met-

enkephalin at MOR are 1.05 ± 0.20 nM and 1.80 ± 0.15 nM, respectively, and our experimental results were consistent with these values. The loss of activity of Bam12(1-4) at MOR was encouraging because this means we can control the activity of met-enkephalin simply by blocking and releasing the last residue, methionine, using our protein switch. The length of 5 amino acids is also ideal to control, as manipulating longer peptides using protein switches[185–198] has been frequently demonstrated in previous literature.

Based on our data, we decided to use met-enkephalin (amino acid sequence YGGFM) as the prototypical opioid peptide for M-PROBE. The affinity of met-enkephalin is highest at DOR, followed by MOR (~5-fold lower) and KOR (~25-fold lower).[148] According to published K_i values of known opioid peptides and the loss of activity of tethered endomorphins at MOR, we were unable to identify a better candidate for M-PROBE. However, we believe the successful engineering of met-enkephalin M-PROBE should still be of tremendous interest to the opioid research community as well as protein engineers.

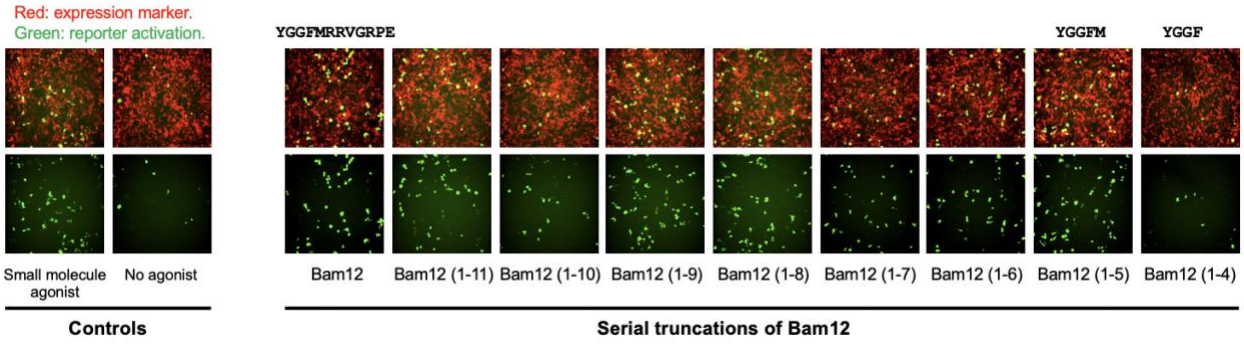


Figure 2-7 Serial truncations from Bam12 identified the minimal sequence required to activate the mu-opioid receptor, YGGFM. "Small molecule agonist" refers to DAMGO ([D-Ala^2, N-MePhe^4, Gly-ol]-enkephalin).

2.3 Rational Design of a Photoswitch

At the same time of screening for an ideal opioid peptide to use for M-PROBE, I aimed to identify or develop a photoswitch that can control the opioid peptide. The idea of using light to control protein activity and cellular processes was from optogenetics, broadly defined as the technique of photo-controlling engineered, light-sensitive proteins. Optogenetics gained its popularity among neuroscientist starting from 2007, primarily as a way to control neuronal activities through engineered ion channels or pumps. However, more recent development extended the field of optogenetics into controlling protein and peptide activities in general. Light is one of the fastest lever for temporal control and is well-suited for studying the immediately effects of opioid receptor activation. A search of available literature failed to identify a protein capable of controlling an opioid peptide in our proposed geometry of M-PROBE. Therefore, I settled to engineer such a protein based on existing studies.

This sub-chapter will give a brief overview of the field of optogenetics to showcase the use of light in biological and neuroscience research. I will then introduce the second light, oxygen, and voltage sensing domain from *Avena sativa* (*As*LOV2), which is the starting point of my engineering. Finally, I will explain our engineering strategy and demonstrate the successful rational design of an re-engineered *As*LOV2.

Introduction to Optogenetics

The foundation of optogenetics was laid in the late 1990s to early 2000s, and the term first appeared in literature[199] in 2006. Optogenetics broadly refers to the combined use of light ("opto-") and genetic targeting ("-genetics"). Optogenetic tools can be classified into sensors, which emits

light to report cellular or physiological changes, and actuators, which receives light and cause cellular or physiological changes.[200] For the purpose of this introduction, I will only introduce optogenetics actuators because M-PROBE itself is an actuator. The definition of "genetics" in "optogenetics" can be subject to interpretations. For the purpose of this introduction, only fully genetically tools will be considered. Partially genetically encoded tools that require externally introduced synthetic molecules will not be discussed.

The first optogenetic tools were developed in 2002 and 2003 by Gero Miesenböck. In 2002, Miesenböck used the phototransduction mechanisms of arrenstin2, rhodopsin, and G_q protein in *Drosophila melanogaster* to develop a method for optically depolarizing groups of genetically designated neurons.[201] The next year, the same group developed a way to directly use light to depolarize ion channels.[202] The temporal resolution of ion channel depolarization was improved by over 1000-fold[203] by Edward Boyden and Karl Deisseroth in 2005, and the first light-controlled ion channels with practical utilities[204,205], which are now still widely used, were published by the same groups in 2007. Around the same time, other light-controlled receptors using GPCRs[206] or chimeric rhodopsin-GPCRs termed Opto-XRs[207–209] were also developed.

To date, the most commonly known optogenetics tools to neuroscientists are still the channelrhodopsin- and halorhodopsin-based tools[204,205] for neuronal activation and silencing, respectively. However, the optogenetics field has far been extended to include non-receptor, soluble light-controlled proteins. These developments allowed the versatile use of light to control a variety of cellular processes both in neurons and in other cells.

One strategy is to utilize naturally found or artificially created light-induced protein dimerization pairs to control cellular processed by translocation. Examples of homodimerizers include the smallest light, oxygen, voltage (LOV) domain Vivid[210,211] from *Neurospora crassa*,

the plant photoreceptor protein UVR8[212], the naturally-occurring transcriptional regulator EL222[213–215], and a LOV domain from *Vaucheria frigida*[216]. Heterodimerization pairs include phytochrome B with phytochrome interaction factor 3[217–220], engineered "pMag" and "nMag" proteins derived from the light-dependent homodimerizer Vivid[221], and many notable natural protein pairs from *Arabidopsis thaliana*, including GIGANTEA with the LOV-bearing FKF1 protein[222], UVR8 with COP1[223], and the basic helix-loop-helix protein CIB1 with cryptochrome 2[224].

Complemental to light-induced dimerization, light-induced dissociation has also been investigated. Notable examples include the Zdark or Zdk protein which binds to *As*LOV2 in light but dissociates in dark[225,226], the engineered Dronpa proteins which associate with one another in light but dissociates in dark[227], and the natural LOV domain from *Rhodobacter sphaeroides*[228]. These light-controlled dissociation tools have been used to control a variety of biological entities, most notably, enzyme activities[227,228].

While light-dependent association or dissociation can be used to study many processes by controlling protein translocation and interactions, a more direct way to control protein or peptide activity by light is to "cage" and "uncage" the effector site (residues that confer functions to these proteins and peptides) or allosteric sites using a photoswitchable protein, by fusing or inserting the corresponding sites to the photoswitch. This approach has gained tremendous success since its first published example[183] in 2008.

In its essence, this approach utilizes light-sensitive proteins that exist in two states, an open or "on" state, and a closed or "off" state. Typically, these proteins belong to the Per-Arnt-Sim (PAS) superfamily, a group of sensor proteins.[229] PAS domains share a conserved core consisting of antiparallel β-sheets and several α-helices and loops. Signaling originates in the β-strand and

propagates to a terminal α-helix. The most common light-sensing PAS domains are LOV domains, which contain a cofactor, flavin mononucleotide, at their core. Upon absorbing light energy, FMN causes structural changes within the protein core which propagates and disrupts a terminal α-helix. Typically, the α-helix is bound with the protein core in the dark state, and unwinds from the core in the light state. By fusing proteins or peptides with this terminal helix, their functions are blocked in dark and released in light. The most successful photoswitch is probably the *As*LOV2 domain from the common oat. This domain will be introduced in more details in the following text.

*As*LOV2 Domain as a Commonly Used Photoswitch for Controlling Protein and Peptide Functions

*As*LOV2 is the second LOV domain from *Avene sativa* phototropin 1, and belongs to the superfamily of PAS domains. Similar to most other LOV domains, *As*LOV2 has a flavin mononucleotide (FMN) cofactor buried at the core of the protein. Light irradiation at 480 nm causes the formation of a covalent bond between FMN and cystein450, leading to a structural change in the protein core.[230] This change is propagated to the C-terminus of the protein, disrupting the hydrophobic interactions and hydrogen bonding between the PAS core and the C-terminal α-helix termed Jα-helix, and causing the Jα-helix to unwind from the PAS core[230,231].

As one of the most well-studied LOV domains, *As*LOV2 has been used for engineering photoswitchable optogenetic tools for over a decade. In 2008, an *As*LOV2-controlled transcription factor was created by fusing *As*LOV2 to the N-terminus of the *E. coli* transcription factor trp repressor[183], whose N-terminus is critical in DNA binding. The authors proposed an "allostery", or "shared helix" mechanism, by which the Jα-helix in *As*LOV2 forms a shared α-helix with the N-terminal helix of the trp repressor, and this shared helix acts as a lever for blocking or releasing

the trp repressor.[183] Under the same mechanism, several other proteins have been engineered to be photoswitchable, such the restriction enzyme PvuII[232], the apoptosis-executing domain from caspase-7[233], and the endogenous formin mDia[234]. Through these photoswitchable proteins, controlling of a variety of cellular processes have been demonstrated, including DNA cleavage activity[232], apoptosis[233], and serum response factor activity[234].

Another mechanism of using *As*LOV2 is steric blocking. Several proteins and virtually all peptides are controlled by this mechanism. Examples include a the Rac1 GTPase[235], the Stim1 transmembrane protein capable of activating the Orai1 calcium channel[184], a viral potassium channel[236], and numerous peptides with lengths from 4 to 38 amino acids[185–198,237,238].

Rational Design of a Circularly-Permuted AsLOV2

We first considered existing photoswitches for the possibility of controlling met-enkephalin in M-PROBE. For such proteins to work, it has to fuse with the C-terminus of met-enkephalin so that the met-enkephalin N-terminus can enter the MOR binding pocket. As indicated by MOR crystal structures[178], the N-terminal tyrosine residue is critical for opioids to bind to MOR. However, to our knowledge, although a variety of photoswitches[239,240] have been used to cage dozens of proteins and peptides, none has an obvious capability of fusing with a peptide's C-terminus.

We then decided to use the protein engineering strategy of circular permutation to create a new photoswitch for M-PROBE. Circular permutation was first discovered in 1979 as a natural phenomenon between two proteins, concanavalin A and favin.[241] It was found that the sequence and three-dimensional (3D) structures of both proteins were highly similar, except that the N- and C-termini were located on different positions of the protein.[241] Later studies found that this is due

to the post-translational modification of concanavalin A, which causes cleavage and an unusual protein ligation.[242] Soon after the discovery of this phenomenon, David Goldenberg and Thomas Creighton created the first engineered circularly permutated protein in 1983 by first joining the N- and C-termini with ligation ("circular"), followed by introducing a new termini at a different position ("permutation").[243] Since then, circular permutation has become a useful technique for protein designers to change the connectivity of proteins while retaining the overall 3D structure[244,245], creating new protein functions.

We reasoned that circular permutation would be an ideal technique to engineer our proposed light switch because we need to change the connectivity of the photoswitch while retaining its light-sensing functions. We picked *As*LOV2 as the engineering starting point for this new photoswitch for several reasons. First, the crystal structure of *As*LOV2 in both dark and light states are available[246], and the N- and C-termini of *As*LOV2 are in close proximity (**Figure 2-8**, left), allowing simple circular permutation design without disrupting the overall protein structure. Second, as illustrated by the preceding texts, *As*LOV2 is the most commonly used photoswitchable protein and its mechanisms are relatively well-understood. Lastly, previous studies have identified many mutations in *As*LOV2 that allow fine-tuning of its kinetics and dynamic range[179,180,247,248], which could guide us in improving our new photoswitch.

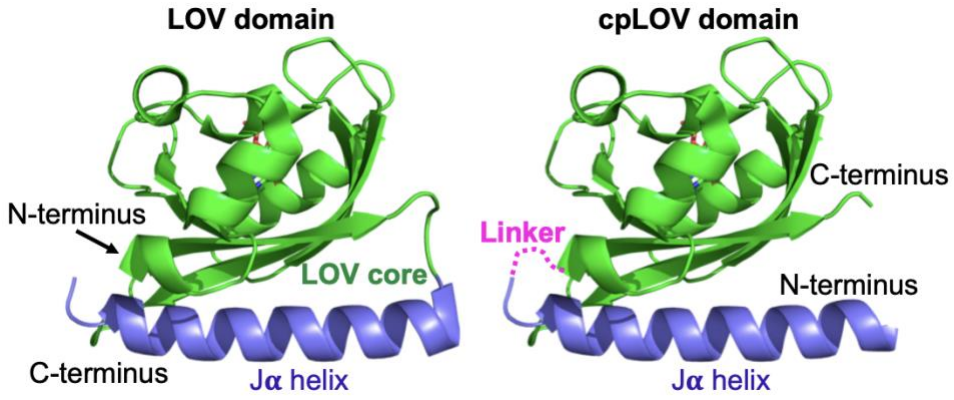


Figure 2-8 Crystal structure of the *As*LOV2 domain (PDB: 2V1A) and design of the circularly-permutated *As*LOV2, cpLOV. cpLOV was designed by first connecting the N- and C-termini on *As*LOV2 followed by creating a new opening to the N-terminus of the Jα-helix. Linker sequence: GSGS.

To engineer circularly permutated *As*LOV2 (cpLOV), we started from an *As*LOV2 variant, hLOV1[180], which contains 15 mutations from the wild-type *As*LOV2 and has been shown to have superior caging in the dark state. We first connected the hLOV1 termini with a flexible linker. Based on the crystal structure[246] (**Figure 2-8**), we reasoned that a four-amino-acid (GSGS) linker is sufficient to connect the original N- and C-termini of the LOV domain. We then introduced a new opening at the original "hinge region" connecting Jα-helix to the PAS core. After creating this circular protein, we made a new opening between amino acids L520 and H521, as H521 is the first helical residue on the original Jα-helix.

To check whether the light-induced conformational change of Jα-helix could still take place in cpLOV, we first used cpLOV to cage the seven-amino-acid TEV protease cleavage site, as this assay uses an easy-to-measure protease cleavage as the readout and has been previously demonstrated[179]. Screening of a series of fusion sites one-amino-acid-at-a-time gave promising results as one of the cpLOV candidates showed light dependence (**Appendix 2**). However, in this experiment, the dynamic range was extremely small due to the lack of protease cleavage even in the light state. Increasing TEV protease concentration, using TEVp variants with better catalytic

efficiency[249] or creating more space for TEVp to access TEVcs did not show significant improvement. We hypothesized that this could be due to the geometry of how TEVcs was caged by cpLOV. Since the last C-terminal residue (the P1' position) on TEVcs is most critical for TEVp cleavage[250], and this residue is the first one to fuse to cpLOV, it might have been overly blocked. In addition, hLOV1 is a high-caging LOV variant originally designed to minimize background leakage.[179,180] Although low background is also desired for M-PROBE in order to reduce receptor desensitization and internalization, the caging efficiency of this LOV variant might have been too high.

We then sought to use a different assay that could produce a more sensitive readout. We used cpLOV to cage a heptapeptide, SsrA, and used a yeast surface-based binding assay with SspB protein (**Figure 2-9**) to evaluate its accessibility in the dark and light states. In dark, SsrA is caged and does not bind to SspB; in light, SsrA is released and binds to SspB. This assay is highly sensitive because SspB is fused to an engineered peroxidase termed the APEX2 protein[251], which, in the presence of biotin-phenol and hydrogen peroxide, generates short-lived biotin-phenoxyl radicals and covalently tags nearby proteins with biotin molecules. These molecules can then be recognized by streptavidin-fluorophore conjugate. Due to the nature of enzymatic reactions and enrichment by the biotin-streptavidin interactions, this labeling process generates large signal amplification. We used yeast surface display to express individual SsrA-cpLOV fusion proteins because this allows future directed evolution[252,253] to improve cpLOV. We screened 10 fusion sites along the Jα-helix (**Figure 2-10**), from V529 to Q538, as it has been shown that the fusion sites on the Jα-helix affect the light-dependence of effector peptides, but no single fusion site is best for all peptides[185,189,194]. Out of the ten constructs tested (raw data in **Appendix 3**), three (cpLOV(b),

cpLOV(i), cpLOV(j)) showed statistically significant light-dark difference (**Figure 2-11**). Among them, cpLOV(i) showed comparable light-to-dark ratio as hLOV1.

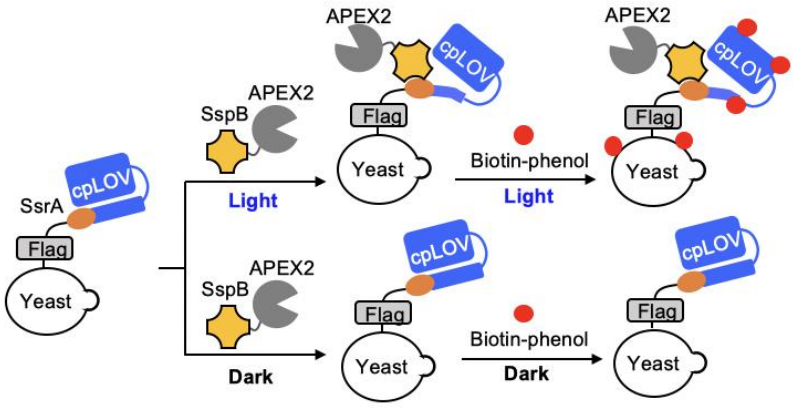


Figure 2-9 Yeast-surface-based assay for assessing the accessibility of cpLOV-caged SsrA. SspB is a binder to SsrA. APEX2 is an engineered soybean ascorbate peroxidase which can label nearby proteins with biotin molecules in the presence of biotin-phenol and hydrogen peroxide. When caged, SsrA cannot bind to SspB. When uncaged, SsrA binds to SspB-APEX2, and the corresponding yeast cell is labeled with biotin molecule.

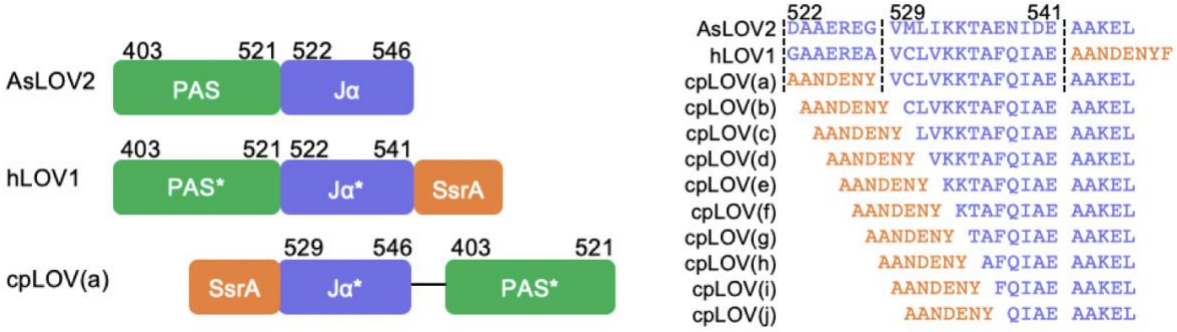


Figure 2-10 Rational design of cpLOV. We created ten SsrA-cpLOV fusion protein to screen for the best fusion site. *As*LOV2 is the wild type photoswitchable protein. hLOV1 is a high-caging variant. Numbers indicate the amino acid position on the original *As*LOV2 protein.

These data show that after circular permutation, cpLOV can still fold properly, recruit the FMN cofactor, and retain light-induced structural change. In addition, its performance, as measured by the dynamic range or S/N, can be similar to hLOV1. To our knowledge, this is the

first demonstration of a LOV domain with an α-helix docking and undocking from the N-terminus of the PAS core. Our results also highlight the importance of fusion site screening when using cpLOV. Similar to findings from previous studies, the fusion position plays a crucial role in engineering photoswitchable peptides, and the optimal fusion site must be for each individual peptide through site screening.

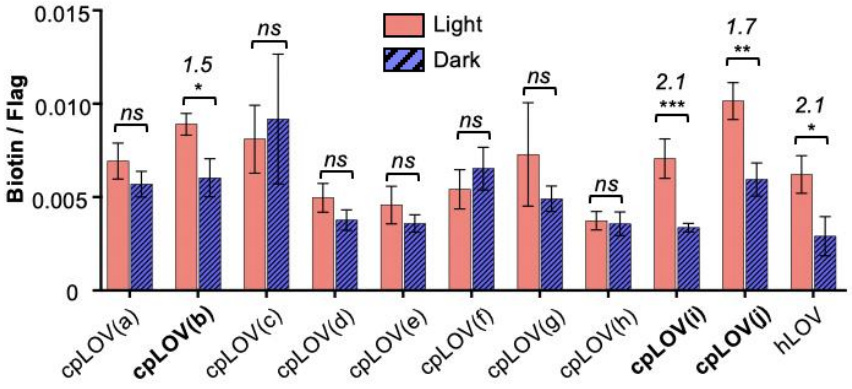


Figure 2-11 Quantification of the ten cpLOV fusion constructs and comparison with hLOV1. We used bar plot to show biotin/FLAG signal ratio for different constructs under light and dark conditions. Biotin and FLAG signals were measured by flow cytometry with raw data as shown in **Appendix 3**. Only FLAG-positive cells were included. The bars in the plot indicate the mean of the ratio and the error bars indicate the standard error of the mean. Numbers are the ratio of the mean biotin/FLAG signal ratio between light and dark conditions. P values were determined by unpaired two-tailed t-test. *P < 0.05; **P < 0.01; ***P < 0.001; ns, not significant.

2.4 cpLOV Can Replace the Photoswitch in Existing Optogenetic Tools

After we establish the functionality of cpLOV on the yeast surface, we asked how the performance of this new protein compares to that of the original hLOV1 in established optogenetic tools such as SPARK. This additional test aims to evaluate cpLOV in mammalian-cell-based, biologically-relevant applications. Since the mammalian cell environment is different from a more

solution-like yeast surface environment, we need to confirm its function in this regard. I cloned the best TEVcs-caging cpLOV into the transmembrane domain for the DRD1 SPARK assay to replace the original LOV domain. As shown by **Figure 2-12**, the new design is the same as a conventional SPARK system except that the connectivity of TEVcs and LOV domain is switched.

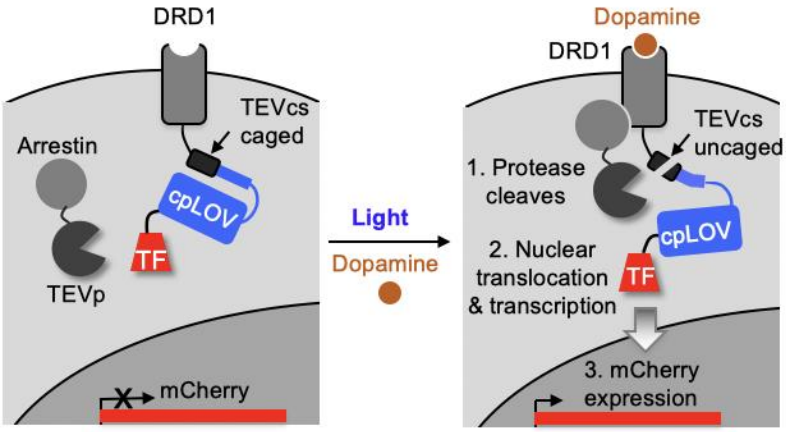


Figure 2-12 A DRD1 transcriptional SPARK assay where hLOV1 was replaced by cpLOV. Signal is only produced when there is concurrent presence of light and dopamine. DRD1, dopamine receptor D1. TEVp, tobacco etch virus protease. TEVcs, tobacco etch virus protease cleavage site, with amino acid sequence `ENLYFQM`. TF, transcription factor.

Using mCherry fluorescent protein as the reporter, we imaged four SPARK conditions using a confocal microscopy and quantified by the sum intensity of the reporter protein. We found that with some adjustment for experimental conditions, the results (**Figure 2-13**) were highly similar to the original SPARK assay, with a light-to-dark S/N (under drug conditions) of 8.4-fold. This experiment also revealed additional information about cpLOV for our later engineering. For example, for the original SPARK assay where the high-caging hLOV1 was used, both a regular TEVp[180] and a TEVp with improved catalytic efficiency[249] can be engineered onto the TEVp component. In that case, improved TEVp provided much higher signal and overall improved the SPARK assay.[249] However, for our assay, this TEVp led to high background and low S/N,

indicating that the dark-state caging capability of cpLOV is inferior to hLOV1. This is consistent with the our flow-cytometry-based quantifications (**Figure 2-11**) on the yeast surface, where hLOV1 showed lower background leakage than most cpLOV variants. This demonstrates that our cpLOV does have a different dynamic range than other LOV domains, and its performance in a given assay must be tested individually. To fully replace the function of hLOV1 from the SPARK assay, cpLOV may need to be further evolved for higher caging efficiency in the dark state.

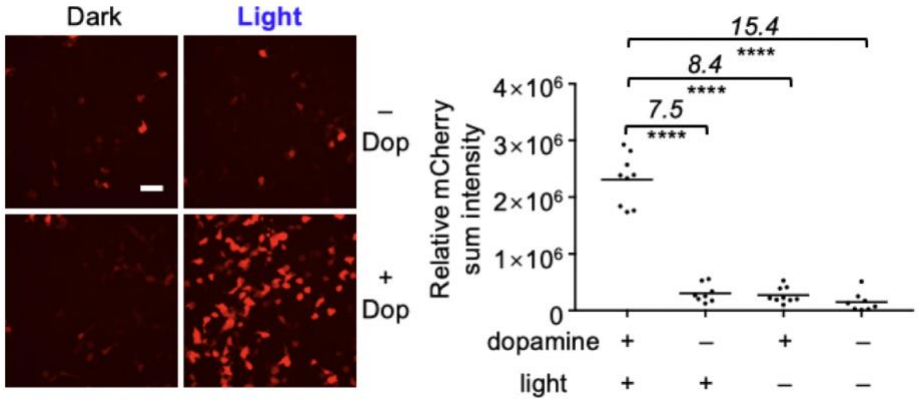


Figure 2-13 Representative confocal microscope fluorescent images (left) and quantifications (right) of the transcriptional assay shown in **Figure 2-12**. Cells were stimulated with light and/or 100 μM of dopamine for 10 minutes. The dot plot shows the relative mCherry sum intensity in each image. Eight to ten images were analyzed for each condition. The values above the dots indicate the ratio of total intensity between two conditions. P values were determined by Wilcoxon-Mann-Whitney test. ****P < 0.0001. Scale bar, 50 μm.

2.5 cpLOV Can be Used to Tune Existing Optogenetic Tools

Besides creating a LOV domain for M-PROBE, one of our original goals was to engineer new protein switch modules that can bring new engineering capabilities to researchers. The successful creation of the first LOV domain that fuses with the effector's C-terminus opened the

door for interesting engineering possibilities, such as a dual-caging strategy where a protein or peptide is flanked by a protein switch on both termini. We wanted to pursue this direction because high background activity is a long-standing challenge in the field of using light-sensing proteins[254]. Insufficient caging often leads to unwanted signal or effects, particularly when LOV domains are used to control highly sensitive proteins or when expressed for an extended period of time (e.g. in vivo for days to weeks). Many previous engineering efforts have focused on reducing the leakage of LOV domain by rational mutagenesis[255] or directed evolution[180,248], but this issue remains a challenge. We envisioned that dual caging might be a general approach for improving dark-state caging for many optogenetic tools and decided to pursue the first proof-of-concept study by ourselves.

We asked whether we can tune the dynamic range of the SPARK assay by the dual caging strategy so that this neuroscience tool can be adapted to a larger range of protein-protein interactions with various affinities. We used the established DRD1 SPARK (**Figure 2-14**, left) as our prototype and designed a dual-caged TEVcs (**Figure 2-14**, right) for the transcription factor component. Initially, we used a combination of hLOV1 and cpLOV, as hLOV1 has been shown to provide optimal caging for this assay. However, this design resulted in significantly reduced signal. This is presumably because hLOV1 was designed specifically for TEVcs through directed evolution[179,180], and the peptide is deeply embedded in hLOV1. Introducing an additional layer of caging to TEVcs in an already optimized system might have diminished the accessibility of TEVcs to TEVp. Therefore, we took a step back and replaced hLOV1 with a more general photoswitch, eLOV[179] (**Figure 2-14**, right), which is also derived from *As*LOV2 but with fewer mutations from the native protein and a less truncated Jα-helix.

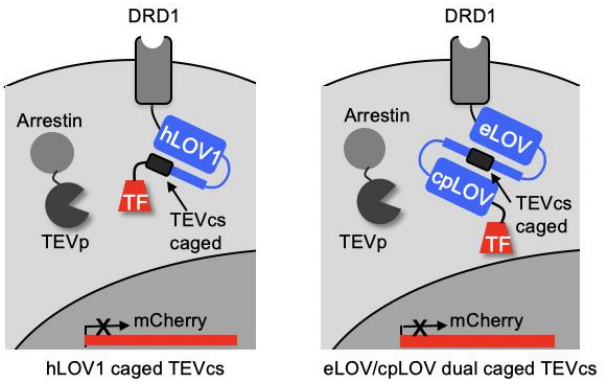


Figure 2-14 A dual-caged SPARK transcriptional assay for DRD1. eLOV and cpLOV were used in place of hLOV1. DRD1, dopamine receptor D1. TEVp, tobacco etch virus protease. TEVcs, tobacco etch virus protease cleavage site, with amino acid sequence `ENLYFQM`. TF, transcription factor.

Comparison between the original DRD1 SPARK and the new eLOV/cpLOV-dual-caging SPARK demonstrated the particularly strong caging efficiency for the latter strategy. As shown by **Figure 2-15**, the assay background under the "– drug" condition was reduced by an impressive 15-fold while signal was only reduced by 7-fold. This showed that dual-caging can effectively shift the light-dependent dynamic range of existing optogenetic tools. Specifically, by combining cpLOV with other photoswitchable proteins, the dynamic range of optogenetic tools can be shifted to the lower end. In principle, this strategy can be generally applied to photoswitchable peptides and proteins as long as they do not require a free N- or C-terminus to function. We expect optogenetics tools that suffer from high background to particularly benefit from our design.

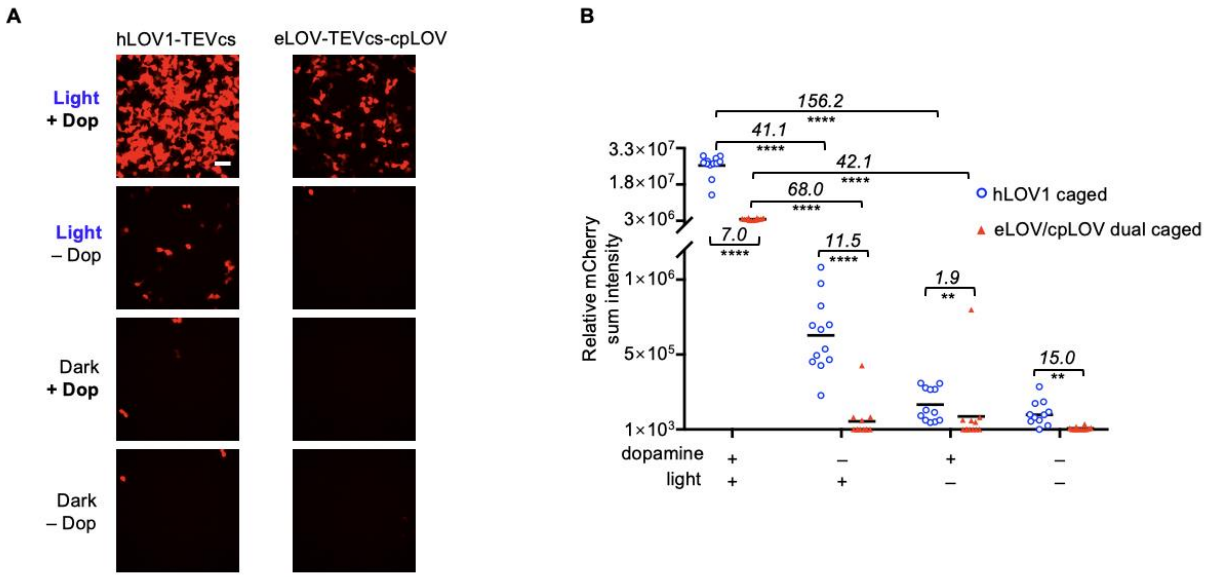


Figure 2-15 Representative confocal microscope fluorescent images (A) and quantifications (B) of the transcriptional assay shown in **Figure 2-14**. Cells were stimulated with light and/or 100 μM of dopamine for 10 minutes. The dot plot shows the relative mCherry sum intensity in each image. Twelve images were analyzed for each condition. The values above the dots indicate the ratio of total intensity between two conditions. P values were determined by Wilcoxon-Mann-Whitney test. ****P < 0.0001. Scale bar, 50 μm.

Chapter 3 Developing a Chemically-Activated Protein Switch for Opioid Peptides and Short Peptides in General[2]

3.1 Introduction to Chemogenetics and Chemically-Activated Protein Switches

Classes of Chemogenetic Tools

Similar to the lexicology of "optogenetics", chemogenetics refers to the use of a chemical entity ("chemo-"), often an organic small molecule, to control genetically-encoded proteins ("-genetics"). Chemogenetic tools allow remote control of cellular processes using an often orthogonal lever, enabling the study of cellular activities, physiology, and behavior with user-defined temporal resolution. Just as optogenetic tools encompass a wide range of engineering methods and working mechanisms (discussed under **Chapter 2.3**), chemogenetic tools can also be divided into several main categories.

To neuroscientists, the most prominent chemogenetics tools are DREADDs (discussed under **Chapter 1.4**), which are used to control neuronal activities by orthogonal small molecules.

[2] Partially from: Shen, J., Geng, L., Li, X., Emery, C., Kroning, K., Shingles, G., Lee, K., Heyden, M., Li, P., Wang, W. A general method for chemogenetic control of peptide function. *In review.* Rational design and directed evolution of CapN and CapC, as well as cell culture applications were in collaboration with Jiaqi Shen, Ph.D. student in the Wenjing Wang Lab, University of Michigan. Animal experiments were work from Xingyu Li and Catherine Emery from the Peng Li Lab at the University of Michigan. Simulation experiment was work from Matthias Heyden at Arizona State University. Kerry Lee, my undergraduate mentee at the University of Michigan, involved in the design of the GloSensor assay and performed the experiment.

The earliest DREADD-like technology can perhaps date back to 1991, when a β_2-adrenergic receptor was engineered only to bind to a non-natural ligand, 1-(3',4'-dihydroxyphenyl)-3-methyl-L-butanone, although with a low potency.[256] In 1998-1999, the first RASSLs (receptor activated solely by synthetic ligand) were designed[257] and soon demonstrated in the very first published chemogenetic study[258]. RASSLs were later used, most notably, to reveal the cells and receptors involved in the taste of sweet[259], unami[259], and bitter[260]. However, RASSLs are limited by the activity of the RASSL ligands on endogenous receptors. In 2007, the first Designer Receptors Exclusively Activated by Designer Drugs (DREADDs) were developed. DREADDs ligands are generally considered orthogonal to the endogenous system, representing another advancement from RASSLs. To date, DREADDs are the most used and successful chemogenetics tools in the neuroscience field, and readers can refer to **Chapter 1.4** for a more detailed description.

Though popular, the GPCR-based RASSLs and DREADDs only represent the tip of the iceberg of available chemogenetic tools and engineering strategies. For example, certain ion channels can also be made sensitive to a chemical ligand or thermal energy. These ligand-gated ion channels can be used in conceptually similar ways as RASSLs and DREADDs, though they often suffer from cross-reactivity with endogenous ligands and vice versa[261].

Chemically-induced dimerization or dissociation is another aspect of chemogenetic tools, and they are conceptually similar to the light-controlled counterparts (reviewed under **Chapter 2.3**). One seminal example is based on the 12-kilodalton FK-506 binding protein (FKBP12)[262,263] which is a ubiquitous eukaryotic immunophilin protein with prolyl isomerase activities. By creating a homodimer of the FKBP12 binder, FK506 (which results in "FK1012"), homodimerization of FKBP12 can be achieved.[264] A later study[265] identified another protein, FRB (FKBP12-rapamycin-binding domain), that can bind to FKBP-12 in the presence of an

immunosuppressant drug, rapamycin, allowing chemically-controlled heterodimerization. These tools, together with other conceptually similar examples, have been used to enable "chemically induced proximity" for the study of biology and medicine.[266] Chemically-induced dissociation is another engineering possibility and may be seen as a reverse process of chemically-induced dimerization, though the development seems more recent. A notable example was a "chemically disrupted proximity" design published in 2019 by Dustin Maly, where the hepatitis C virus protease, NS3a, binds to a peptide in the basal state and dissociates from the peptide in the presence of a small molecule called ANR.[267] Similar to chemically induced association systems, this chemically induced dissociate system can be used to control a variety of cellular processes.[267]

The last category of chemogenetic designs is similar to the idea of photoswitchable proteins, where an effector peptide or protein itself is made "switchable" by insertion or direct fusion with a chemoswitchable protein. An early example of insertion embedded a truncated FKBP12 into kinases, which allowed allosteric activation by an FKBP12 binder molecule.[268] Later, the same group designed a protein made of part of FKBP12 and part of FRB, which can also be inserted into allosteric sites in kinases to enable control by rapamycin.[269] More recently, the insertion strategy has been applied beyond kinases and FKBP12-related proteins. A circularly permuted bacterial dihydrofolate reductase was inserted into sites in nanobodies that are essential for their affinity to the binding targets.[270] This allowed chemogenetic control of nanobody activities. In a different design strategy, the effector protein can also be "inserted" into chemical switches. In 2017, Dustin Maly designed a chemically controlled guanine nucleotide exchange factor by sandwiching it between a BH3 peptide and a BLC-xL protein.[271] BH3 interacts with BLC-xL and occludes the effector's active site in the absence of a small molecule. Addition of the small molecule disrupts BH3/BLC-xL interaction and restores the effector's function.

Directly fusing the effector and the chemoswitch is similar to the insertion strategy except that the effector is now at one of the termini of the chemoswich. To our knowledge, the only example before our design was a ligand-induced degradation (LID) system[272] developed by Thomas Wandless in 2011. By screening FKBP12 variants with degron-like sequences on the C-terminus and *de novo* engineering using synthetic oligonucleotide libraries, the authors created a FKBP12 protein where an extended C-terminus both serves as a degron and a binder to an active site in FKBP12. In the basal state, the degron is inactive due to steric blocking. A small molecule, shield-1, however, can bind to the same active site with higher affinity and displace the C-terminal tail from the site, restoring the degron's function.

Comparison between Chemogenetics and Optogenetics

In this dissertation, chemogenetics and optogenetics are both pursued as means of controlling M-PROBE. Building on the brief overview in **Chapter 1.6**, this sub-chapter further compares chemogenetics with optogenetics and elaborates on the pros and cons of both approaches. Readers should see why each type of tool has its unique advantages in M-PROBE.

Chemogenetic and optogenetic tools share the "-genetic" feature but differ in their control mechanisms. As both can be fully genetically encoded, they both have the capability of targeting specific cell type, neuronal circuits, or neuronal projections, as discussed in **Chapter 1.4**. The differences between "chemo-" and "opto-" control make these tools suited for different purposes, and researchers should choose based on the desired spatiotemporal resolution, among other features.

Chemogenetic and optogenetic tools have the same spatial resolution during expression, similar to all other genetically encoded tools. Using cell-type-specific promoters or the Cre-*lox*P

system, genetic tools can be precisely targeted to specific cells and neuronal circuits. In addition, they can be targeted to specific subcellular locations using a variety of localization signals[273,274]. However, the spatial resolution for chemogenetic and optogenetic tools during stimulation is different. In mice, ligands for chemogenetic tools are often administered globally, allowing the control of proteins in many organs and locations at the same time, although local administration of small molecules are also possible. Most small molecules will inevitably diffuse in the animal, making confined stimulation difficult if such genetic tools were expressed in undesired regions. On the other hand, this allows the use of chemogenetic tools in deep tissues or regions hard to reach physically. In contrast, light often can only illuminate a small body or brain area in mice, especially free-moving ones. For non-transparent animals, light have poor tissue penetration.[274] For example, in the mouse brain, light can penetrate for ~2 mm. This is well-suited for local activation of optogenetic tools, but is a hurdle for using light in a large body area.

The temporal resolution of chemo- and opto-genetic tools also differ. Light provides almost unsurpassable temporal resolution as it can be easily controlled on the scale of millisecond or less. Often, the temporal resolution of optogenetic tools depends on how fast these tools can be turned on and revert back, rather than how fast light itself can be turned on and off. Many optogenetic tools are auto-switchable: they are turned on in the presence of light, and will turn off automatically in the absence of light. For these tools, the temporal resolution depends on the on-off kinetics of the photoswitchable protein. Other tools[275], often not fully genetically encoded, require one wavelength to be turned on and another to be turned off. Chemogenetics tools are often controlled on a much longer scale, on the order of minutes to hours and more. The temporal resolution of chemogenetics tools depends on how fast the ligands get degraded or cleared out from the animal of interest, and differs from one tool to another. As a reference, one study found DREADD ligand

CNO to be effective in the mouse brain from 15 minutes post-injection to 30-60 minutes post-injection.[276] Therefore, for studies that require temporal resolution on the minute or second level, optogenetic tools are more well-suited. For studies that require longer stimulation, both chemogenetic tools and optogenetics tools can be used, although prolonged light stimulation in animal organs may cause phototoxicity or organ damage, especially if the optogenetic tool requires laser stimulation.

Spatiotemporal resolution is not the only factor to consider when choosing between chemogenetic and optogenetic tools. Another common aspect of concern is how to access these tools after they are expressed in the brain. Neuroscience is arguably the field where chemogenetic and optogenetic tools are most used, and gaining control to the genetic tools in the presence of animal skull and blood-brain barrier (BBB) is not always easy. For optogenetic tools, a common technique is to implant light-emitting devices such as optical fiber or wearable light-emitting diode. The invasiveness of light delivery and possible hindrance on animal mobility and comfort needs to be considered when designing such experiments. Routine injection techniques are often used to control chemogenetics tools, and they are far less invasive than implanting light-emitting devices on the animal body or skull. As a result, they are less likely to interfere with the animal behavior. However, once drug is injected, there is less control of when it is degraded or cleared out, or how much can cross the BBB. Testing different drug concentrations and routes of administrations may be necessary in order to get the best result. In addition, many compounds do not readily cross the BBB. Cautions must be taken even if a compound has been previously shown to cross the BBB, especially if there is only few literature. For example, in our own experience, the ligand shield-1 does not cross the BBB in mice very well, even though it has been reported so[277].

Lastly, the invasiveness and toxicity of chemogenetic and optogenetic tools need to be considered. As mentioned, the need to implant optical fiber or light-emitting devices is a notable source of invasiveness when using optogenetics. However, light is generally considered orthogonal to the endogenous environment and does not cause noticeable interference with the normal physiology of animals. A potential concern besides invasiveness is phototoxicity. It is well-known that prolonged exposure to strong light can damage cells and tissues. This is particularly a concern for tools that require laser stimulation, such as the neuronal stimulation or inhibition tools[204,205] based on channelrhodopsin and halorhodopsin. However, most photoswitchable proteins such as LOV domain can be controlled with much weaker light (such as intensity similar to room light), and phototoxicity is generally not a concern. Toxicity of chemogenetic ligands depend on the molecule. For tool designers, low toxicity and orthogonality are desired features of the chemogenetic ligand. However, the effect of these ligands on animals are not always obvious or known. For example, the most popular DREADD ligand, CNO, was generally thought to be orthogonal since DREADD was first introduced[169] in 2007. It was more than a decade later when CNO was found to be reverse-metabolized into a molecule that produces stimulus effect in animals.[177] For chemogenetic ligands, comprehensive evaluation of their on-target and off-target effects, pharmacokinetics, and pharmacodynamics[276], preferably at the time of their first development, is desirable. Such studies will provide an early evaluation of these drugs and guide tool users to the proper selection of chemogenetic and optogenetic tools. In the lack of such studies, however, researchers will need to evaluate the pros and cons of both classes of tools using their best judgment and domain knowledge.

3.2 Overall Engineering Strategy for the Chemically-Activated Protein Switch in M-PROBE

We studied the various engineering strategies to design chemogenetic tools in order to identify the best strategy for making a chemoswitch for M-PROBE. We are particularly interested in the published single-chain chemoswitches because that is the desired configuration of M-PROBE. For these switches, both insertion and fusion strategies can render a protein or peptide chemoswitchable. Based on published literature, we believe that insertion requires more engineering efforts because there are numerous possible insertion sites. In addition, insertion can be less predictable. As demonstrated by inserting circularly permuted bacterial dihydrofolate reductase into several nanobodies[270], for different effector proteins, the effect of insertion can even be opposite: while addition of the small molecule increases the activity of some nanobodies, the exact same small molecule can decrease the activities of other nanobodies.[270] In contrast, a direct fusion strategy is more similar to what we wanted to achieve in M-PROBE, especially given that the N-terminus of the M-PROBE opioid peptide has to be free from fusion.

We started our rational design using LID[272] because this is the closest example we identified, and it has several key proven advantages. First, the functionality of LID has been proven on a 19-amino-acid peptide, and we hypothesized that a similar caging mechanism may apply to other peptides. Second, LID uses FKBP12 with an F36V mutation, and its ligand, shield-1, has sub-nanomolar affinity towards FKBP12(F36V), which is 1000-fold higher than that for the wild-type FKBP.[278] This will minimize disturbing the function of endogenous FKBPs. Lastly, shield-1 is membrane permeant and has been previously used in live animals.[279] This is important for the final application of M-PROBE in vivo. At the time before we started engineering, a study using

shield-1 in live animals[279] showed that shield-1 can cross BBB. However, in our hand, using the tools we later developed, we were unable to observe this effect using even the most sensitive assay we had, despite extensive animal experiment efforts. Nonetheless, shield-1 was very cell-membrane-permeant and can reach other animal organs except for the brain.

We proposed a two-step engineering strategy to adapt LID to control opioid peptides in the M-PROBE geometry. First, we will test and re-engineer LID so that it will not only be able to control a degron, but also any short peptide fused to its C-terminus. This is possible because certain residues on LID must interact with the FKBP12(F36V) active site. We hypothesized that the caging mechanism of LID can be general: by fusing any short peptide after the residues responsible for FKBP12(F36V) binding, these peptides should be pulled into the active site and blocked by FKBP12(F36V). Ideally, not all amino acids that form the degron are required for FKBP12(F36V) binding, so that we can diminish the degron function of LID and convert it into a general chemoswitch. The key engineering focus of this step is to identify and/or improve the FKBP12(F36V) binding sequence. We named the resulting protein as CapN, for "chemically-activated protein domain caging the N-terminus of the effector".

Our second step was to re-engineer CapN into "CapC". As previously mentioned in **Chapter 2.3**, the N-terminus of opioid peptides must be free in order to activate opioid receptors. Any protein switch can therefore only control opioid peptides from their C-terminus. We planned to use circular permutation similar to how we engineered cpLOV. The challenge for engineering cpCapN, or CapC, is that the N- and C-termini of FKBP12(F36V) are not in immediate proximity, and a much longer linker is required. We proposed to first use the strongest FKBP12(F36V) binder developed from CapN to mitigate the loss of binding caused by a long, flexible linker, then improve the key residues on this linker.

We were aware that this would be a multi-step, long engineering process with no guaranteed success, and we envisioned several importance of this process and possible exit points. The innovations and potential values we would bring to protein engineers and tool users are as follows:

1. Before our design, there was no chemogenetic protein switch with general applicability. Unlike the *As*LOV2 domain which has been applied to dozens of proteins and peptides (reviewed in **Chapter 2.3**), most chemoswitchable designs are effector-specific. Engineering each chemoswitchable entity often requires extensive effort, preventing these tools from being widely used and adapted by biologists and neuroscientists. The successful engineering of either CapN or CapC, or both, will offer a generally applicable chemoswitchable protein for the first time, expanding the toolbox of chemogenetics.

2. CapN and CapC will allow controlling effectors from two different terminus. As demonstrated by many previous photoswitchable engineering efforts (reviewed in **Chapter 2.3**), it often critical to fuse switchable protein domains close to the key residues on the effectors. CapN and CapC will offer the versatility of controlling an effector from different termini, increasing the chance of successful engineering of these tools.

3. Similar to the tandem use of *As*LOV2 and cpLOV as demonstrated in **Chapter 2.5**, CapN and CapC can be used in the same construct. This will likely offer a different dynamic range (between the "open" state and the "closed" state) than using CapN or CapC alone, and can be particularly useful for applications that require minimized background.

Overall, both using CapN and CapC on their own and using them together will offer new protein engineering capabilities that are previously unavailable, and we expect these tools to be

appreciated by both the protein engineering community as well as tools users including biologists and neuroscientists.

3.3 Rational Design and Directed Evolution Created a Chemically-Activated Protein Switch, CapN

As mentioned in **Chapter 3.2**, our first step was to engineer CapN, which is capable of controlling a peptide by fusing to the effector's N-terminus.

Rational Design of CapN

The design of CapN was inspired by the LID system[272] where an added 19-amino-acid sequence to the C-terminus of FKBP12(F36V) serves as a degron and a shield-1-displaceable binder to the FKBP12(F36V) active site at the same time. The original LID paper did not explicitly differentiate the degron sequences from linker sequences on these 19 amino acids, or suggest which residues from the 19 amino acids bind to the FKBP12(F36V) active site. However, it did show that the last 5 amino acids, `RRRGN`, is a degron. Based on this information and the hydrophobic active site (**Figure 3-1**) from the FKBP12 crystal structure[280], we hypothesized that the degron `RRRGN` was brought to the FKBP12(F36V) active site by the hydrophobic residues immediate to the N-terminus of the degron, as underlined here: `TRGVEEVAEGVVLLRRRGN`. In contrast, the more N-terminal residues serve as the linker from the FKBP12(F36V) C-terminus to the 19 amino acids.

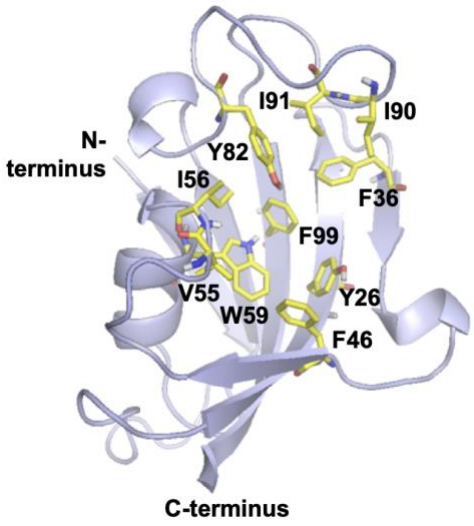


Figure 3-1 Hydrophobic binding pocket of FKBP12 shown by its crystal structure (PDB:1FAP). The hydrophobic residues around the active site are shown in yellow and stick representation.

Based on these hypotheses, we attempted to use LID to cage two short peptides that have been previously shown to be controllable by the AsLOV2 domain and for which we have an established readout. We replaced the degron (RRRGN) with a heptapeptide, TEVcs (ENLYFQG)[250], or an octapeptide, modified SsrA (AANDENYF)[248] (**Figure 3-2**). DNA encoding the LID-peptide fusion protein was transformed into the EBY100 strain of *Saccharomyces cerevisiae* (baker's yeast) and displayed on the yeast surface[252]. This allows a solution-like environment easy for measuring peptide activities as well as enables further yeast-based directed evolution[253] of the protein.

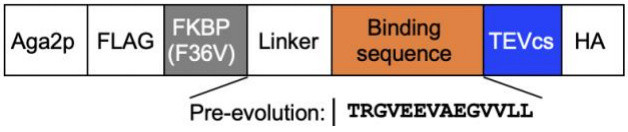


Figure 3-2 Geometry of CapN-caged SsrA. The original linker and binding sequence (TRGVEEVAEGVVLL) from LID was fused with TEVcs or SsrA to the C-terminus of FKBP12(F36V).

The accessibility of TEVcs was evaluated using a protease cleavage assay (**Figure 3-3**, top) where an epitope tag, HA, is removed from the yeast surface if the TEVcs is cleaved by the protease. TEVcs accessibility is therefore inversely correlated with the HA abundance, measurable by anti-HA antibody staining (i.e. higher HA signal, less CapN opening). In comparison, the accessibility of SsrA was measured by a binding assay with SspB-APEX2 (**Figure 3-3**, bottom), similar to how cpLOV-caged SsrA was measured as shown in **Chapter 2.3**. Briefly, when SsrA is "uncaged" and binds with SspB-APEX2, the APEX2 enzyme labels nearby proteins with biotin. Fluorophore-linked streptavidin can then be used to measure the abundance of biotin on the yeast, and the signal is positively correlated to the SsrA accessibility (i.e. lower fluorescence, less CapN opening). Yeast cells for both assays can be handled in a solution-like assay where they are simply incubated with purified TEV protease or SspB-APEX2 protein. After labeling, they were then measured by flow cytometry, showing both the expression level of the CapN construct and the corresponding signal. In the initial test, we did not observe shield-1 dependence for either TEVcs or SsrA, suggesting that LID is ineffective in introducing steric blocking to peptides beyond the original degradation sequence.

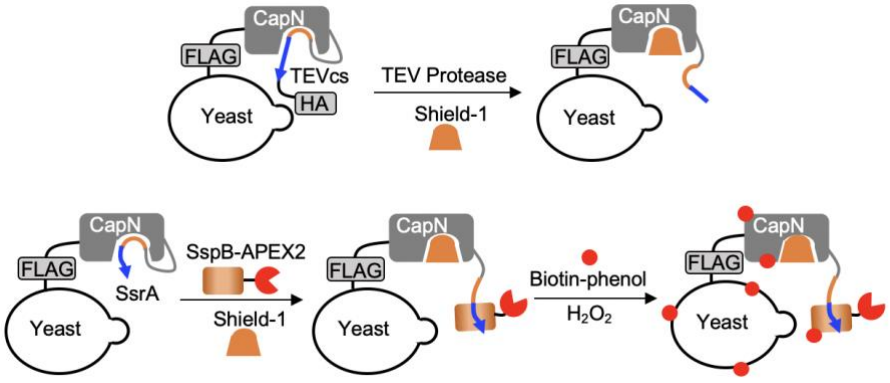


Figure 3-3 Yeast-surface-based labeling assays for measuring the accessibility of CapN-caged TEVcs (ENLYFQG, top) and SsrA (AANDENYF, bottom). TEVcs or SsrA was displayed on the yeast surface by fusing to the yeast Aga2p protein. Accessibility of TEVcs was evaluated by protease cleavage. FLAG and HA signals indicate protein expression level and TEVcs cleavage, respectively. Accessibility of SsrA was evaluated by binding to its binding partner, SspB. APEX2 labels protein within close proximity with biotin-phenol. FLAG and biotin signals indicate protein expression level and SsrA-SspB association, respectively. APEX2 is an engineered ascorbate peroxidase. FLAG and HA are epitope tags.

Directed Evolution of CapN

Based on our hypothesis that the N-terminal residues immediately before the effector peptide are most critical for binding to FKBP12(F36V), we focused on improving these residues. Using the TEVcs cleavage design with the yeast surface display platform, we designed four libraries with six to nine random amino acids in the place of the original six amino acids to the N-terminus of TEVcs (**Figure 3-4**, top). The added amino acids ensured that there was sufficient length for interacting with the protein switch, in case the degron sequence (RRRGN) in the original LID also contributed to FKBP12(F36V) binding. This targeted mutagenesis approach (as opposed to random mutagenesis) allowed more coverage of the residues of interest. Theoretically, each random residue has 20 possibilities (any of the natural amino acid), and the higher limit of the number of variants for a yeast library is on the order of 10^7. Therefore, six random amino acids,

which have 6.4×10^7 combinatorial possibilities, are the higher number that can possibly be fully covered by a single yeast library. For our libraries with 7, 8, or 9 random amino acids, we expect them to sample some variations but not to fully cover all possibilities. We expected this strategy to be more effective than fully randomizing the entire FKBP12(F36V) construct or the 14 amino acids on the C-terminus because they are the residues with the highest possibility of reaching the FKBP12(F36V) active site.

We used flow-cytometry-based fluorescence-activated cell sorting (FACS) to measure and sort yeast cells. FACS has the advantage of displaying each individual cell as a dot on an x-y plot, where the x-axis is normally showing the protein expression level and the y-axis is normally showing the signal. In this particular directed evolution campaign, the expression of CapN-TEVcs fusion protein was measured by a mouse anti-FLAG-tag immunoglobulin G (IgG) primary antibody followed by a goat anti-mouse-IgG secondary antibody conjugated with Alexa Fluor 568, and the signal was measured by a rabbit anti-HA-tag IgG primary antibody followed by a goat anti-rabbit-IgG secondary antibody. Being able to detect both expression and signal is advantageous because we can select cells not solely based on the signal intensity, but based on the signal-to-expression ratio. The latter method is more reliable because yeast cells vary vastly on the their protein expression level, with measured differences spanning over two to three orders of magnitude.

We used an alternating selection strategy which is common for FACS sorting. This strategy uses two types of selections termed positive selection and negative selection. For the purpose of this thesis, we define positive selection as the round of labeling and selection where our protein switches are “on” (e.g. when shield-1 is added to CapN-expressing yeast cells), and negative selection as the round where our protein switches are “off”. FACS displays each yeast cell as an

individual dot on an x-y plot (i.e. signal vs expression plot) and allows users to "gate" the desired population by literally drawing an enclosed shape on the plot shown on a computer screen. The instrument then physically sorts any yeast cells that fall into the gate by increasing the voltage between two metal plates when the charged liquid droplet containing that cell reaches the appropriate position, allowing it to be collected by a collection tube set on the side of the stream. Typically, for each round of selection, both positively and negatively labeled yeasts in two separate tubes are prepared, and the user decides which type of selection to execute and how to draw the gate based on a number of factors. This is arguably the most difficult and critical part of FACS-based directed evolution, as gating can easily make or break a directed evolution, and it is heavily dependent on personal experience and preferences. Based on my experience with multiple directed evolution campaigns, in a rough descending order of importance, these factors include: the current difference between positive and negative cells, the difference between the current and the previous rounds of cells, the desired "direction" of evolution, the current estimated number of variants, room on the FACS plot for drawing the gate, the distribution of the cell population and the number of "outliers" (i.e., possibilities for further improvement), as well as the need to remove or retain expected interfering cells during the first round. For the directed evolution of CapN, we combined the four libraries after the first round of selection because they were all designed to achieve the same outcome and we only needed one improved variant regardless of its number of amino acids.

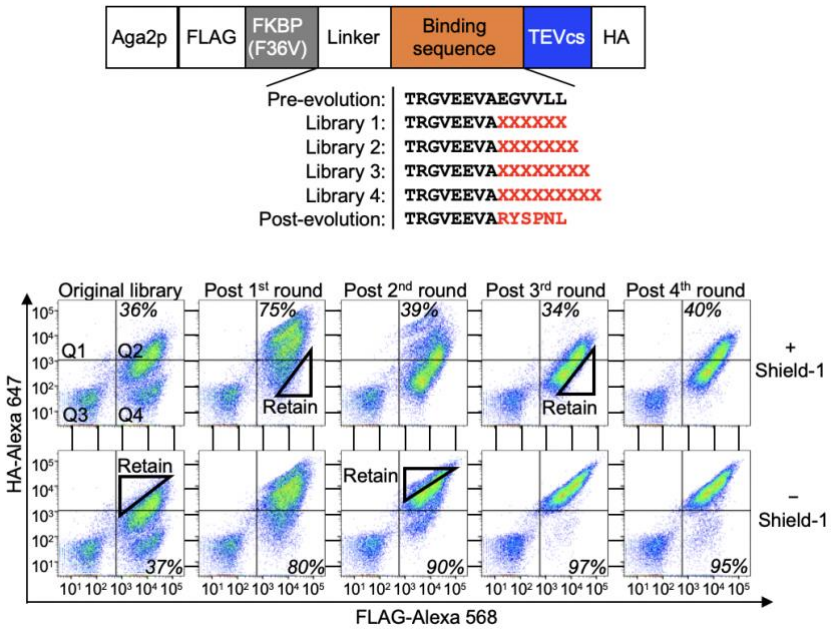


Figure 3-4 Library design and selection scheme for CapN-caged TEVcs. For directed evolution (Libraries 1-4), the last six amino acids of the binding sequence was mutated to six to nine random amino acids. The post-evolution sequence shown is the final CapN used for the rest of this study. Aga2p is the yeast protein for displaying CapN on the yeast surface. "X" indicates any of the twenty amino acids. Amino acids that are different from the original LID sequence are highlighted in red.

For the initial CapN libraries ("Original library" as shown on **Figure 3-4**, lower), there was no difference between the "+ shield-1" and "– shield-1" conditions (**Figure 3-4**, lower). For cells with good protein expression (high FLAG signal; Q2 and Q4 on the FACS plots), there was a vast amount of variation in signal, spanning over at least three orders of magnitude. This indicated the successful construction of the libraries as there were a considerable number of variants and room for selection. In addition, this also implies that our target residues were likely indeed important in binding with FKBP12(F36V), as varying these residues caused a difference in signal intensity.

We then performed one round of negative selection by gating for cells with high signal-to-expression ratio in the absence of shield-1 (**Figure 3-4**, lower). The gate was drawn in such a way that the slope of the lower boundary of the gate was roughly the same as the slope of the top of the

cell population (i.e. gate and population shape are parallel), and the gate gently touched a sparse region of the cells. This is because in general, directed evolution is most effective when the cells that are just a little off the main population are selected, so that there is some evolutionary pressure but not a high risk of selecting mostly false positive outliers. We performed negative selection because we needed to remove cells expressing proteins with early stop codons. Since all CapN libraries were completely randomized, the DNA encoding the variants can unexpectedly have one of the three stop codons (`TAA`, `TAG`, `TGA`) in the place of an amino acid. For our CapN geometry, these variants do not express TEVcs or HA, and would always have a low signal-to-expression ratio regardless of the shield-1 condition. Such cells likely had shown up in the lower right population in Q4 of the upper plot for the "Original library". If we had performed positive selection instead for this round, these cells would have interfered with the selection of true positive variants by misleading our gating, because theoretically, they should have had lower signal than that of the true positives on the FACS plot, and in a positive selection round for CapN, gating is mainly based on the position of the low signal cells.

After this round, the libraries ("Post 1st round" as shown on **Figure 3-4**, lower) showed a small but noticeable difference between "+ shield-1" and "– shield-1" conditions. Overall, post-1st-round cells under both conditions showed higher signal than the original libraries. This showed that the selection pressure from the previous round of gating worked, by retaining cells with a "tighter" CapN which protected TEVcs from being cleaved, and thus having a higher HA signal. Notice that even under the "+ shield-1" condition, the population generally showed higher HA signal than the previous round, which is opposite to what we eventually aimed to evolve for. However, this was well expected and is one caveat of using alternating selection. For each round, it is most common that populations under both experimental conditions evolve in the same

direction. However, if the selection is truly effective, after several (typically two to five) rounds, the populations will eventually diverge as the working variants start to be the dominating group in the library. We used position selection for this round because of two reasons. First, in the alternating selection strategy, the selection is often done in a truly alternating way unless the previous round of selection does not show much improvement. Since we performed negative selection in the first round and the population showed an obvious shift towards higher signal, it is most likely that positive selection was needed for this round. If we had performed another round of negative selection, it is likely that we have had selected variants that always show high signal regardless of the presence of shield-1, which can never open up. Second, based on the cell distribution on the FACS plot, there were more variants available for positive selection than negative selection. The upper population in the "– shield-1" condition showed a tight and clean distribution, indicating lack of variety. In comparison, the lower population for the "+ shield-1" condition showed a range of signal with differences of about two orders of magnitude.

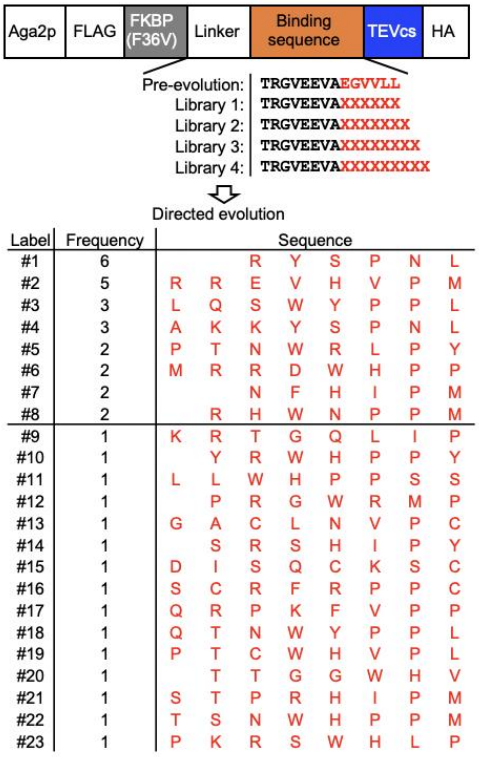
After this round of selection, the libraries, for the first time, showed a clear difference between "+ shield-1" and "– shield-1" conditions ("Post 2nd round" as shown on **Figure 3-4**, lower). This indicated that many cells selected from the previous round can not only open their CapN in the presence of shield-1, but retained good caging in the absence of shield-1. At this point, CapN showed divergent signal under different shield-1 conditions, and can thus be considered functional. However, there were still a large number of variants as the yeast population was visibly spread out over the FACS plots. We decided to perform another round of selection to further enrich the working clones, so that when individual yeast DNA is isolated from the libraries, we would have less characterization to do. For this round, we alternated from the previous round and performed negative selection.

The resulting "Post 3rd round" libraries (**Figure 3-4**, lower) had tighter yeast populations than the previous round, indicating removal of some less functional variants. In addition, the overall position of the population on the FACS plot did not shift much from the previous round. This showed that the performance of the current CapN libraries had stabilized.

We performed a final round of alternating selection, and the resulting population ("Post 4th round" as shown on **Figure 3-4**, lower) was almost identical to the pre-selection one. This indicated that the selection has come to an end and further selection would very unlikely introduce significant changes to CapN. The tight cell distribution implied that the current libraries either had enriched very few working variants, or were made up with variants of extremely similar performance.

Characterization of Post-Evolution CapN and Comparison with Pre-Evolution CapN

After four rounds of directed evolution, CapN libraries showed obvious improvement from the original libraries. We extracted the plasmids from the post-4th-round libraries and sequenced forty individual clones. Twenty-three distinct sequences were identified from sequencing (**Figure 3-5**, left), showing that the post-4th-round libraries had a relatively large number of variants despite the their tight distribution on the FACS plot. Most sequences were rich in hydrophobic amino acids, which was in agreement with FKBP's hydrophobic active site (**Figure 3-1**).



Label	Frequency	Sequence							
#1	6			R	Y	S	P	N	L
#2	5	R	R	E	V	H	V	P	M
#3	3	L	Q	S	W	Y	P	P	L
#4	3	A	K	K	Y	S	P	N	L
#5	2	P	T	N	W	R	L	P	Y
#6	2	M	R	R	D	W	H	P	P
#7	2			N	F	H	I	P	M
#8	2		R	H	W	N	P	P	M
#9	1	K	R	T	G	Q	L	I	P
#10	1		Y	R	W	H	P	P	Y
#11	1	L	L	W	H	P	P	S	S
#12	1		P	R	G	W	R	M	P
#13	1	G	A	C	L	N	V	P	C
#14	1		S	R	S	H	I	P	Y
#15	1	D	I	S	Q	C	K	S	C
#16	1	S	C	R	F	R	P	P	C
#17	1	Q	R	P	K	F	V	P	P
#18	1	Q	T	N	W	Y	P	P	L
#19	1	P	T	C	W	H	V	P	L
#20	1		T	T	G	G	W	H	V
#21	1	S	T	P	R	H	I	P	M
#22	1	T	S	N	W	H	P	P	M
#23	1	P	K	R	S	W	H	L	P

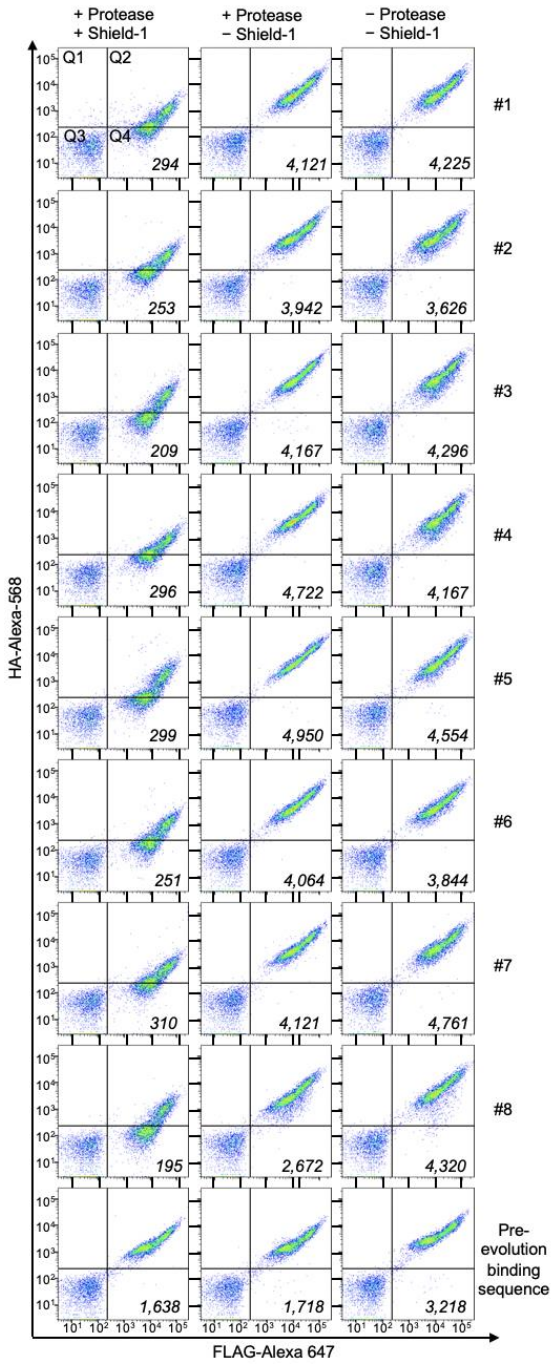


Figure 3-5 Sequences (left) and flow cytometry characterization (right) of post-evolution CapN individual clones. Clone #1 is the final CapN selected. Numbers in the flow cytometry plots are median HA intensity of FLAG-positive cells (Q2 + Q4).

We characterized eight clones that appeared more than once from sequencing and found that they had highly similar performance as characterized by flow cytometry (**Figure 3-5**, right).

This was expected as all clones must fall into the post-4th-round library range, and that range was pretty small as shown by the "Post 4th round" FACS plots from the lower panel of **Figure 3-4**. We selected the most enriched clone with a C-terminus ending in `RYSPNL` as the final post-evolution CapN. When compared with the pre-evolution CapN, the post-evolution CapN showed both tighter caging in the absence of shield-1 and larger opening in the presence of shield-1 (**Figure 3-5**, left). As a result, the dynamic range was greatly improved.

Although CapN was evolved using TEVcs, we expected that the caging mechanism should also apply to other short peptides. When `RYSPNL` interacts with the FKBP12(F36V) active site, it should bring any peptide that is fused to its C-terminus into the FKBP12(F36V) active site, thus blocking its function in the absence of shield-1. To test this hypothesis, we used the same post-evolution CapN to cage SsrA. With virtually no optimization, we observed lower background, higher signal, and a greatly enhanced dynamic range (**Figure 3-5**, right), showing that CapN is a truly generally applicable protein switch to short peptides.

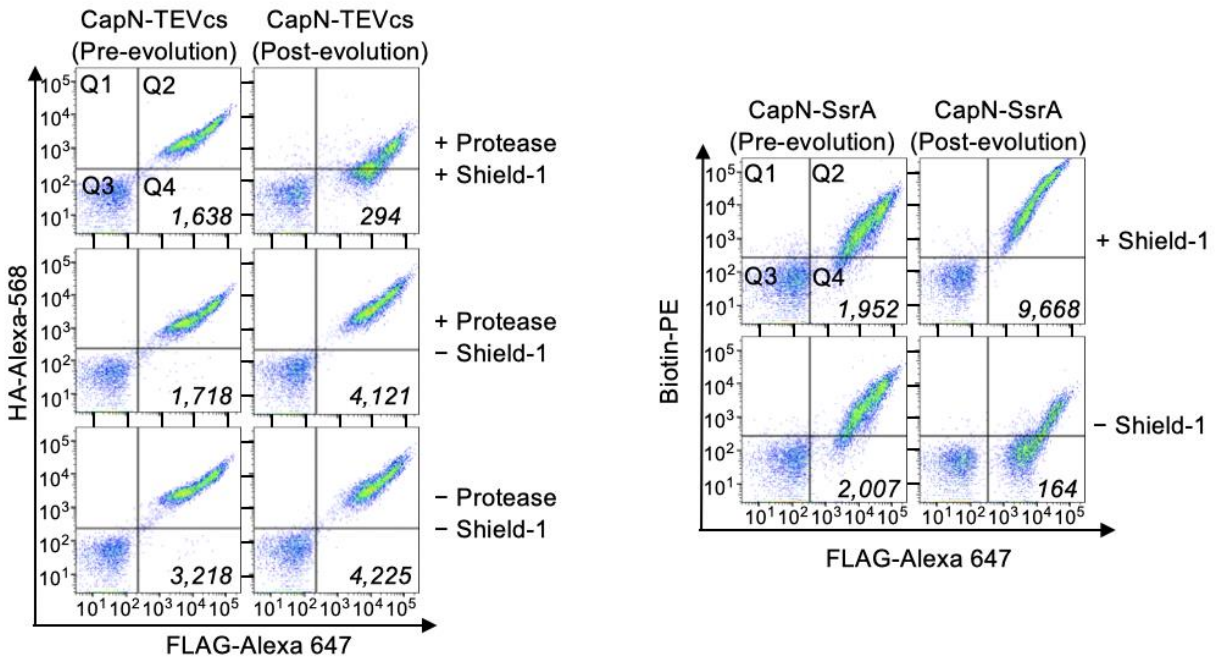


Figure 3-6 Comparison of pre- and post-evolution CapN in caging TEVcs (left) and SsrA (right). In the left plot, values are median HA intensity of FLAG-positive cells (Q2 + Q4). In the right plot, values are median biotin intensity of FLAG-positive cells (Q2 + Q4).

We next measured the dose response of CapN to shield-1, as this is an important factor to consider when using CapN in cell culture or in vivo. Using CapN-caged SsrA, we measured the median biotin signal for yeasts when they were exposed to 2-5000 nM of shield-1, and found that the half maximum response was reached at 53 nM (**Figure 3-7**). The signal started to show up at ~10 nM, and reached the maximum at ~100 nM. Although this number is higher than the single-digit nanomolar affinity[281] of DREADD-binding clozapine at several DREADDs, the 10-100 nM range is still generally considered high binding affinity.

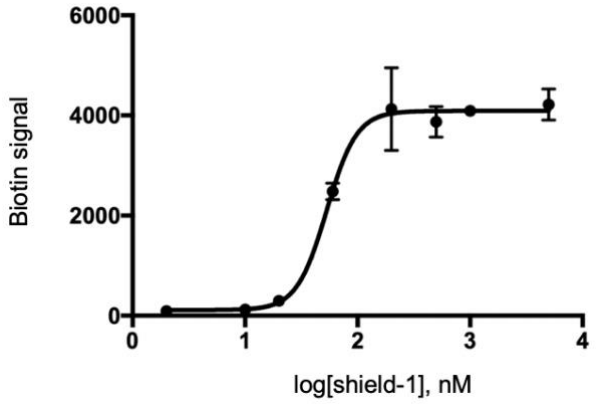


Figure 3-7 CapN dose response curve to shield-1. FACS was used to analyze CapN-caged SsrA on yeast surface treated with different concentrations of shield-1. Three technical replicates were performed for each condition. The median biotin signal was plotted against shield-1 concentration. Half maximum response was observed at 53 nM. Errors bars are standard error of the mean.

We also assessed whether the binding between shield-1 and CapN is reversible, as this is another key aspect to consider when using CapN. With yeasts expressing CapN-SsrA, we first incubated the cells with shield-1 for 10 minutes. Then, shield-1 was washed away, and yeasts were incubated with SspB-APEX2 fusion protein immediately, 1 hour, or 12 hours after shield-1 was washed out. Following SspB-APEX2 incubation, APEX labeling was immediately performed, and any bound SspB-APEX2 will label the yeast surface with biotin molecules. As shown by **Figure 3-8**, even 12 hours after removing shield-1 from the yeast media, SsrA was still able to strongly recruit SspB-APEX2, indicating the "on" state of CapN. Therefore, CapN is effectively an irreversible switch that can only be turned on, but not off. Based on this observation, CapN should be used for studies where a one-way "on" switch is desired, but not for studies where reversibility is expected.

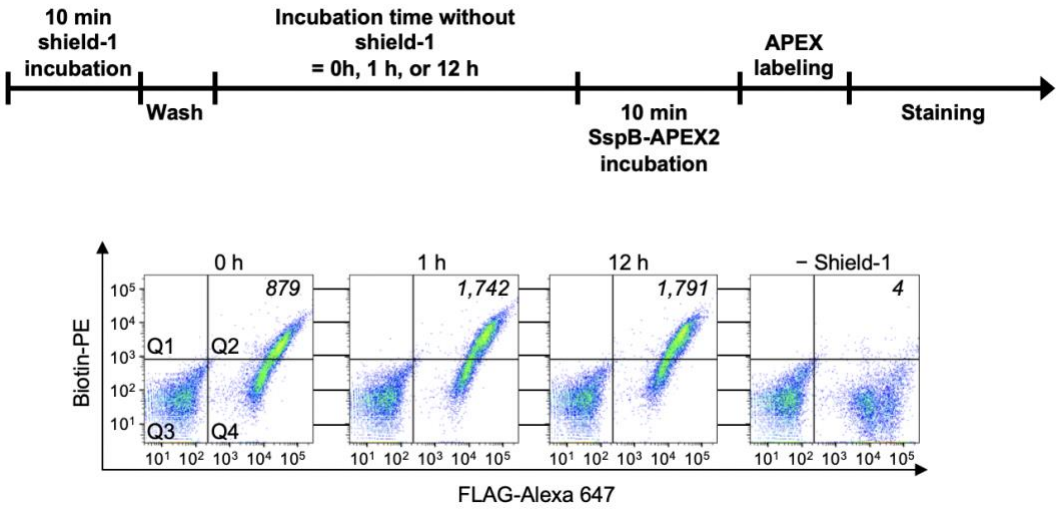


Figure 3-8 Schematics (top) and results (bottom) for assessing the reversibility of CapN. Yeast cells were incubated with shield-1 for 10 min, followed by washing to remove excess shield-1. Yeast cells were then incubated at room temperature for 0-12 h before the accessibility of SsrA was evaluated using SspB-APEX2 and biotin-phenol labeling as shown. Values in the FACS plots are median biotin intensity of FLAG-positive cells (Q2 + Q4).

3.4 Re-engineering CapN into CapC for Controlling Opioid Peptides

Now that we've achieved our first step of the two-step engineering strategy (discussed in **Chapter 3.2**), we aimed to re-engineer CapN into an analogous protein, CapC, for caging peptides from their C-terminus. As previously mentioned (in **Chapter 2.3**), opioid peptides must have a free N-terminus in order to interact with the opioid receptor binding pocket. Therefore, only CapC can be used for the purpose of M-PROBE.

In addition, we expect CapC to have an impact beyond its prototypical use in M-PROBE. The reasoning is two-fold. First, as demonstrated by numerous previous studies, caging of peptide is most effective when the key residues on a peptide are closely bound to the protein switch. Therefore, we expect peptides with a crucial or functional C-terminal sequence to be better caged

by CapC rather than CapN. Second, CapN and CapC can be used in tandem to further increase the caging efficiency and reduce "leaking" in the absence of shield-1, similar to how *As*LOV2 and cpLOV can be used together to tune existing tools as demonstrated in **Chapter 2.5**. This is especially critical for applications in live animals where the protein is expressed for a long time, often weeks to months, before an experiment is performed.

Rational Design of CapC

Similar to how we engineered cpLOV from *As*LOV2, we applied a circular permutation strategy to re-engineer CapN into CapC. We reasoned that circular permutation could work because presumably, the post-evolution CapN binding sequence, `RYSPNL`, works by interacting with the FKBP12(F36V) active site, regardless of its connectivity with FKBP12(F36V). A challenge of this engineering is the relatively long distance between the FKBP12(F36V) N-terminus and the active site (**Figure 3-1**). In fact, the active site and N-terminus are on located on the two opposite faces of FKBP12(F36V), which is perhaps the most challenging case for the circular permutation strategy to work. However, since `RYSPNL` was specifically evolved to bind tightly with the FKBP12(F36V) active site, we expected it to still bind even when the linker connecting it to the active site is changed.

To test this design strategy, we fused the putative binding sequence `RYSPNL` to the N-terminus of FKBP12(F36V) with a flexible glycine- and serine-rich linker (**Figure 3-9**, right). In contrast to the 8-amino-acid linker for CapN (**Figure 3-9**, left), we designed a 17-amino-acid linker based on our estimation of the distance between the N-terminus of FKBP12(F36V) to its active site. We fused SsrA to the N-terminus of the binding sequence, and measured its shield-1 dependence using the same assay as shown in the lower panel of **Figure 3-3**. SsrA was chosen

because we CapN showed a larger dynamic range (difference between "+ shield-1" and "– shield-1" conditions) in caging SsrA than caging TEVcs, and this could facilitate selection during directed evolution.

Figure 3-9 Schematics showing the connectivity of CapN and CapC. The putative binding sequence from the post-evolution CapN, RYSPNL, was kept as a starting point for CapC.

This initial CapC construct showed promising shield-1 dependence (**Figure 3-10**). Without any optimization, there was already a visible difference between the "+ shield-1" and "– shield-1" conditions on the FACS plot. However, the dynamic range was small compared to the post-evolution CapN-caged SsrA (**Figure 3-6**). The difference might be explained by the unoptimized binding sequence or linker in this new geometry. Nevertheless, this showed the binding sequence RYSPNL still had the potential to interact with the FKBP12(F36V) active site, and further directed evolution can be expected to improve its performance.

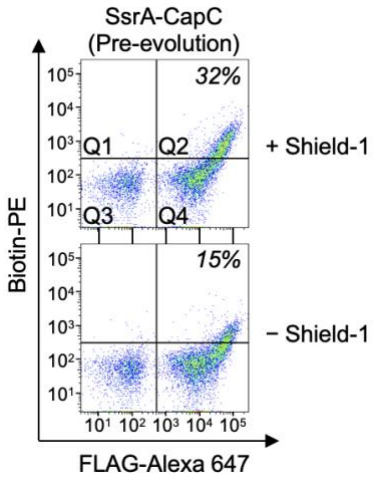


Figure 3-10 Initial characterization of CapC-caged SsrA on the yeast surface. Labeling methods were the same as those in the lower panel of **Figure 3-3**. Numbers are the percentage of cells in Q2 over (Q2 + Q4).

Directed Evolution of CapC Improved Dynamic Range

Based on the data as shown in **Figure 3-10**, we applied a directed evolution strategy similar to that used for CapN (as shown in **Chapter 3.3**) to improve the dynamic range of CapC. To generate libraries, we applied site-saturated, targeted mutagenesis to two critical regions (**Figure 3-11**) simultaneously. One region is the residues immediate to the C-terminus of SsrA, because these residues are likely the most critical ones in pulling the effector peptide into the FKBP12(F36V) active site. Another region is the three to six amino acids right after the binding sequence, because they might be important in positioning the binding sequence and SsrA towards the FKBP12(F36V) active site. Between these two regions, we kept the binding sequence `PNL` because they form a hydrophobic surface which we believed might be important in interacting with the FKBP12(F36V) active site. Through a later collaboration with Dr. Matthias Heyden at Arizona State University, we found that the hydrophobic leucine residue showed a high tendency to interact with the FKBP12(F36V) active site in an all-atom molecular dynamics simulation (**Appendix 4**).

We created three libraries with various levels of mutations in these two regions. Each of the three library has six random residues because this produces the highest number of combinatorial possibilities that a single yeast library can possibly fully cover, as discussed in more detail in **Chapter 3.3**.

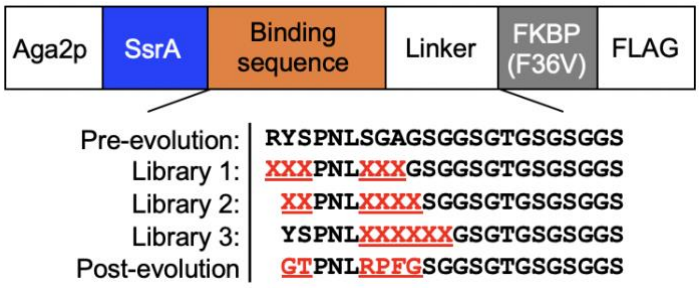


Figure 3-11 Design of CapC libraries. For directed evolution (Libraries 1-3), amino acids highlighted in red were mutated randomly into any of the twenty amino acids. The post-evolution sequence is the final CapC used for the rest of this study. Aga2p is the yeast protein for displaying CapC on the yeast surface. "X" indicates any of the twenty amino acids. Amino acids that are different from the post-evolution CapN sequence are highlighted in red

APEX labeling was used for CapC libraries. Briefly, yeasts were incubated with purified SspB-APEX2 protein, washed, and resuspended in 1% bovine serum albumin, which prevents nonspecific APEX labeling by quenching any radicals before they reach a neighboring protein. The enzymatic labeling process was started with biotin-phenol in a hydrogen peroxide environment. The reaction was allowed to run for two minutes before it was quenched by a mixture of sodium ascorbate and Trolox. After APEX labeling, yeast cells were first labeled with streptavidin-phycoerythrin (PE) and mouse anti-FLAG-tag IgG primary antibody, then with goat anti-mouse-IgG secondary antibody conjugated with Alexa Fluor 647. Yeast cells were analyzed and sorted using FACS, and all subsequent procedures were similar to those for the CapN directed evolution, which have been described in more details in **Chapter 3.3**. After three CapC libraries

were created, they were combined and selected according to the scheme as shown in **Figure 3-12**. The combined library size was determined to be around 1×10^7.

For the initial library (shown as "Combined libraries" in **Figure 3-12**), the difference between "+ shield-1" and "– shield-1" conditions was diminished. This was because the original template (shown as "SsrA-CapC (Pre-Evolution)" in **Figure 3-12**) had just a little shield-1 dependence, and the mutations introduced into the libraries masked this difference by increasing the signal variance. This demonstrated the successful construction of the libraries. For this round, there was plenty of room and variations for both positive and negative selections. We decided to perform a negative selection so that we could retain the CapC variants with good caging in the absence of shield-1. This decision was made based on our goal of having an M-PROBE with minimal background, as it will be expressed in animal for days to weeks. However, if we had performed positive selection for this round, the directed evolution might have worked equally well. The first-round negative selection was possible because for this particular library construct (**Figure 3-11**), the expression marker, FLAG tag, was strategically placed at the very C-terminus. For cells that had a stop codon in the place of the FLAG tag, they would have shown up in Q3 rather than Q4, and would therefore not be selected in either positive or negative selection. For the selection, we collected 1.5×10^4 cells out of 5×10^6 cells.

Our first round of selection was highly effective. As shown by the plots of post-1st-round cells in **Figure 3-12**, the background was largely diminished and was already lower than the pre-evolution template. For the second round of selection, a positive selection was the best choice because (1) there was much more room on the FACS plot for positive selection than negative selection, and (2) the overall population was low on the FACS plot, and further negative selection may cause the loss of high-shield-1-responsive variants. For this round, we retained only 3×10^3

cells out of 3 × 10^6 cells. To our surprise, the post-2nd-round population showed a very large enhancement in signal while keeping a relatively low background. When compared with any previous round or the original template, the signal was higher. In addition, both the "+ shield-1" and "– shield-1" populations had a tight distribution, showing a reduction in the number of variants.

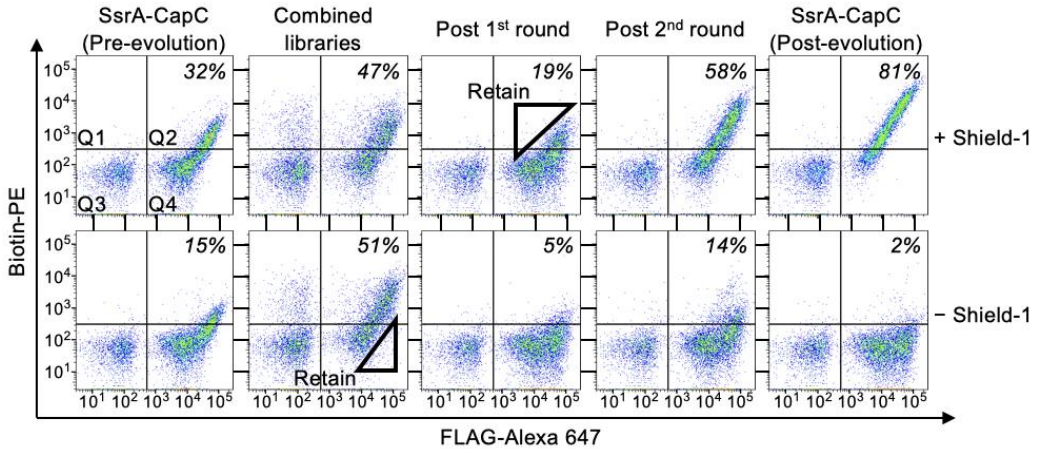


Figure 3-12 Selection scheme for CapC directed evolution. Numbers are the percentage of cells in Q2 over (Q2 + Q4).

We decided that the post-2nd-round libraries were already good enough given the signal-to-noise ratio, and sequenced twenty individual clones (**Figure 3-13**). Eighteen distinct sequences were identified and individually characterized by flow cytometry. Many of the clones showed appreciable dynamic range, and we picked the one with the highest dynamic range (#18 in **Figure 3-13**, panel B) as the final post-evolution CapC (binding and linker sequences: GTPNLRPFGSGGSGTGSGSGGS).

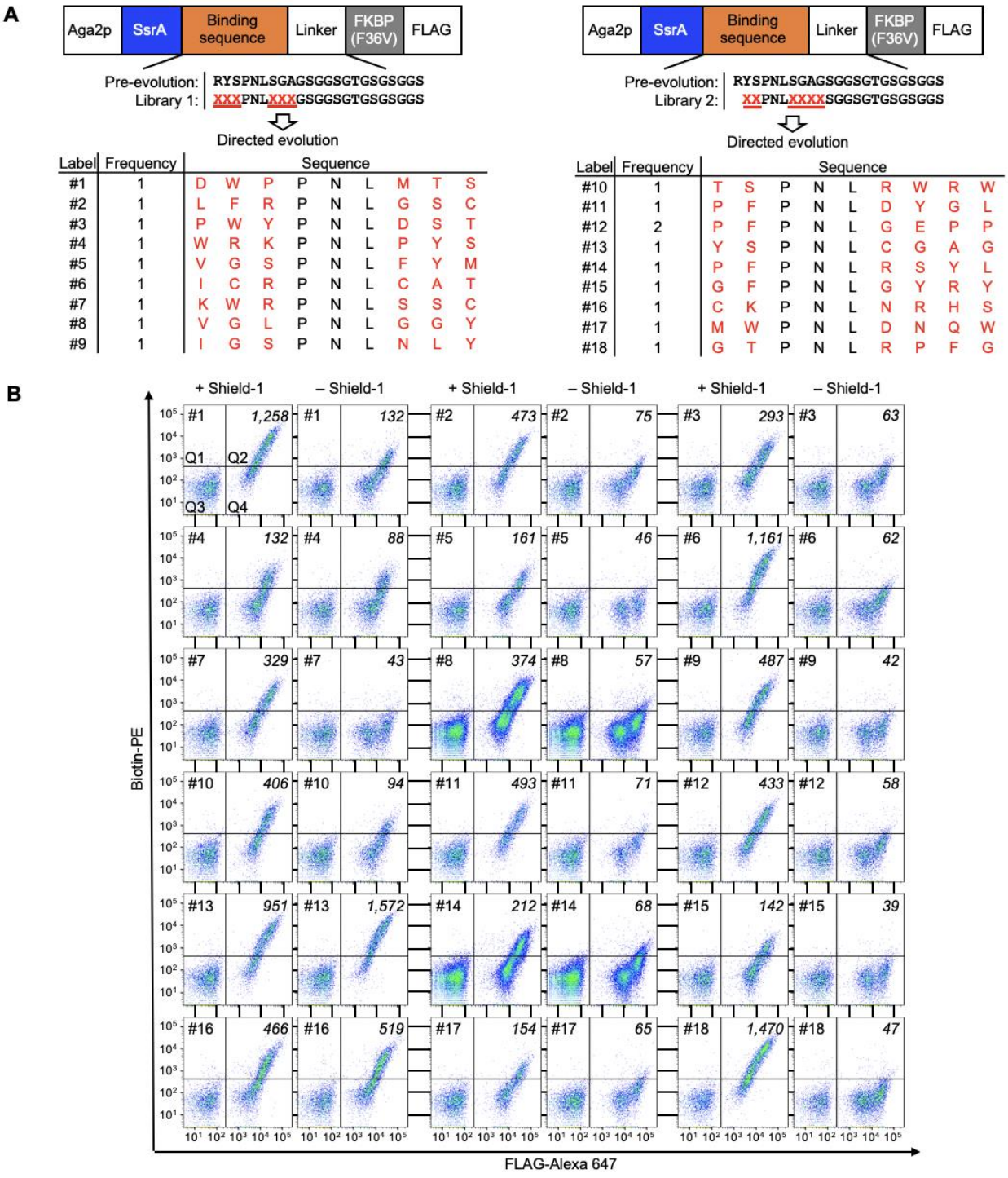


Figure 3-13 Sequences (A) and flow cytometry characterization (B) of post-evolution CapC individual clones. Clone #18 is the final CapC selected. Numbers are median HA intensity of FLAG-positive (Q2 + Q4) cells.

3.5 Tandem Use of CapN and CapC can Reduce Background

One of our goals of developing CapC is to allow its tandem use with CapN in a similar way that *As*LOV2 and cpLOV can work together. This dual caging strategy might solve the long-standing challenge that a single switchable protein often has insufficient caging and suffers from unwanted leakage, especially when expressed in vivo for an extended period of time. To test whether this design could work, we selected TEVcs over SsrA as the caged peptide. This is because SsrA can already be so well-caged by CapC alone (**Figure 3-12**, **Figure 3-13**), and there was not much room for further background reduction. In comparison, by using a stronger TEV protease incubation conditions (such as longer incubation time and higher TEV protease concentration), we can push the CapN and CapC systems to their limits and test whether dual caging could further reduce background.

As shown by **Figure 3-14**, under a strong TEV protease condition, both CapN- and CapC-caged TEVcs had significant leakage in the absence of shield-1. However, when CapN and CapC were used together on TEVcs, leakage was largely diminished. Importantly, the dual-caged TEVcs can still be opened up by shield-1, though to a much lesser extent than the single-caged TEVcs. This is consistent with the dual-caging strategy involving *As*LOV2 and cpLOV, where both the background and the signal were reduced (**Figure 2-15**). Therefore, dual-caging strategy can be used to shift the dynamic range of the switchable tools. The results on the yeast surface supported our initial hypothesis that using both CapN and CapC simultaneously can enhance caging efficiency and reduce unwanted leaking. For the rest of this thesis, I will use the term "CAPs" to refer to the tandem use of CapN and CapC.

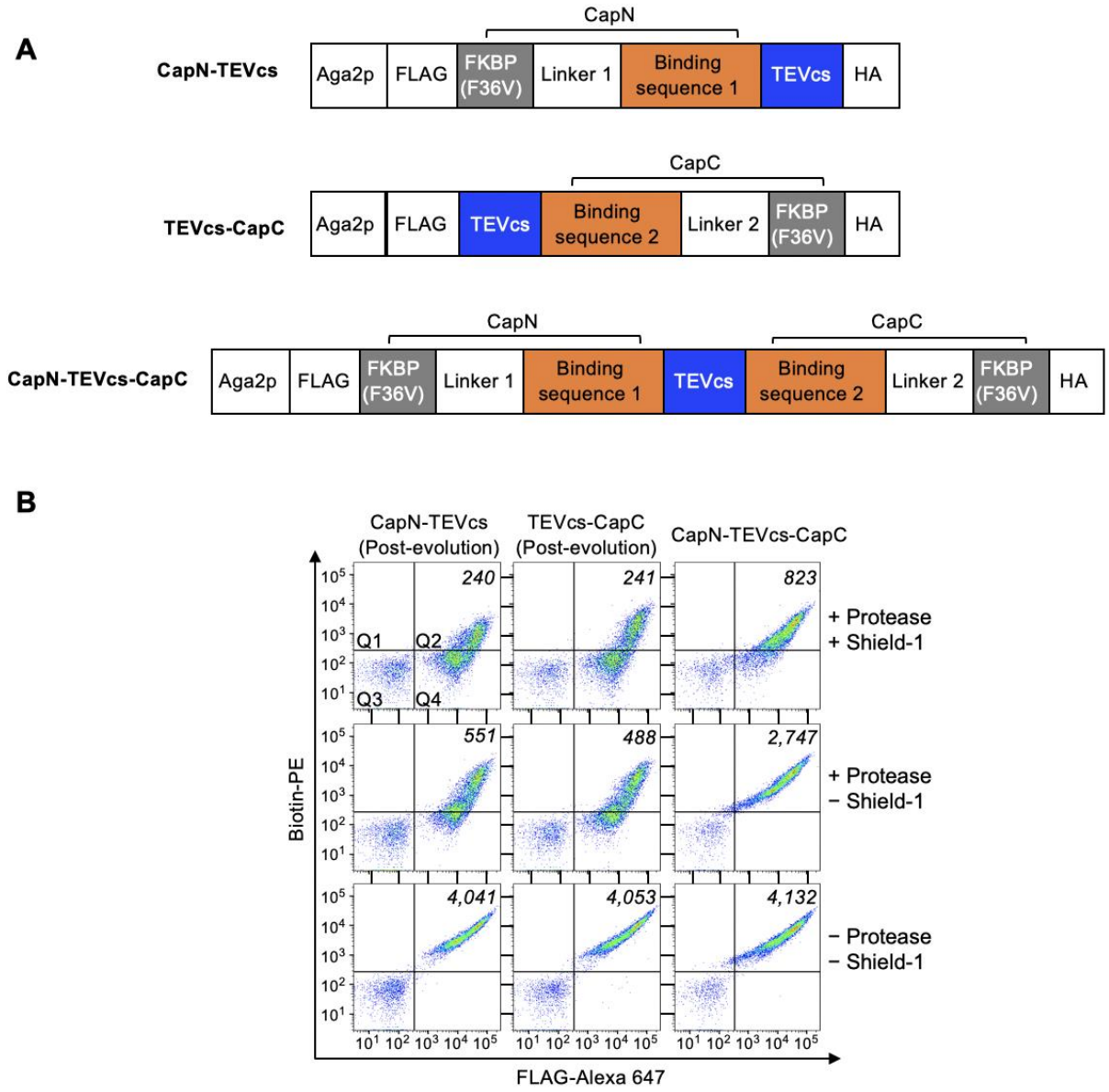


Figure 3-14 Comparison of single CapN, single CapC, and tandem CAPs in caging TEVcs. **(A)** Schematics of the three constructs tested. CapN-TEVcs-CapC is the combined use of both post-evolution CAPs. Aga2p is the yeast protein for displaying constructs on the yeast surface. TEVcs, TEV protease cleavage site (ENLYFQ/G, cleaved between Q and G). FLAG and HA are epitope tags. **(B)** FACS plots of the three constructs shown in (A). Values are median biotin intensity of cells in Q2 and Q4.

3.6 Demonstrating General Applicability of CAPs at Various Short Peptides in vitro and in vivo

To further establish the use of CAPs, we aimed to demonstrate our tools in the mammalian cellular environment both in intro and in vivo, and to extend the effector to other peptides which neither CapN nor CapC was evolved for. To this end, we designed several protein translocation and gene transcription assays in HEK 293T cell culture, neuronal culture, as well as mouse brain and liver. In this sub-chapter, I show the successful control of three peptides in the mammalian cellular environment in vitro and in vivo, including TEVcs, SsrA, and a nuclear localization signal.

CAPs can Control Protein Translocation to Plasma Membrane in HEK 293T Cells

Protein transportation to different subcellular compartments is closely linked to their function, and controlled localization to the plasma membrane has been used to manipulate cellular processes such as phagocytosis[282] and calcium influx[184]. To demonstrate the use of CAPs for protein translocation, we designed a shield-1 induced membrane localization system (**Figure 3-15** A). In this design, dual-caged SsrA peptide is localized on the plasma membrane of HEK 293T cells by fusing to a transmembrane domain. The interacting partner of SsrA, SspB protein, is fused to the protein of interest (POI). As an example, we used EGFP as the POI. Without shield-1, no binding between SsrA and SspB should occur, and EGFP should be found throughout the cell. Addition of shield-1 should uncage SsrA from CAPs and allow it to bind with SspB-EGFP, translocating EGFP to the membrane. As shown by **Figure 3-15** B, under no shield-1 condition, EGFP was localized in the cytosol and nucleus with no apparent membrane pattern as predicted. Upon adding shield-1, we observed EGFP translocation to the plasma membrane within seconds.

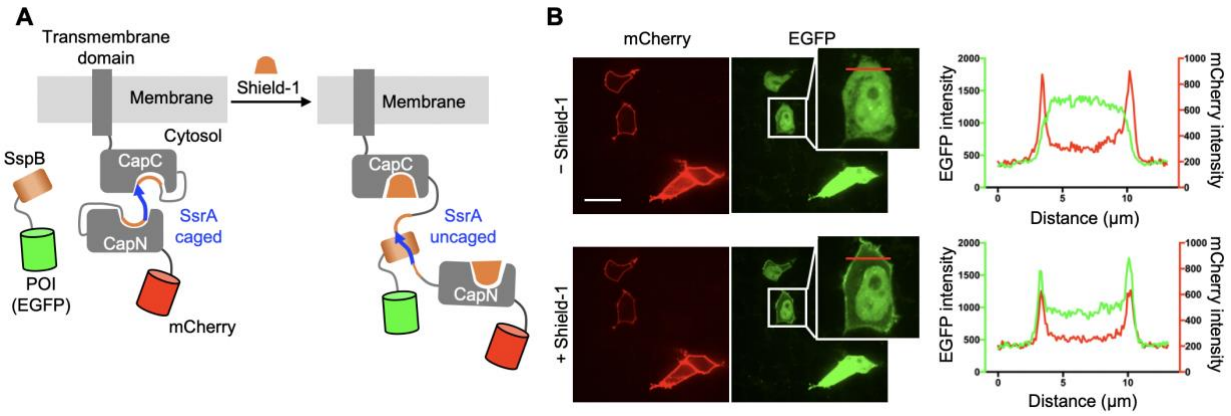


Figure 3-15 Schematics (A) and results (B) for CAPs-controlled protein translocation to the plasma membrane. Protein of interest (EGFP as the example) fused to SspB was translocated to the plasma membrane when SsrA was uncaged from CAPs. mCherry was used as a membrane protein marker. Transmembrane domain is CAAX. b, The right panel shows the intensity profiles of mCherry and EGFP along the red line in the images. POI, protein of interest. Scale bar, 20 μm.

This showed that shield-1 can easily permeabilize the cell membrane and open up CAPs on the order of seconds, making the CAPs-shield-1 system useful even for experiments that require a fast temporal control. Importantly, this also demonstrated that CAPs retained its functionality when directly transferred from yeast surface to mammalian cell culture, and there was no need of re-optimization. In principle, by directing CAPs-SsrA to different subcellular compartments with appropriate localization signals, POI could be recruited to various locations of interest in a shield-1-dependent manner.

CAPs can Delocalize Proteins from the Plasma Membrane to the Cytosol in HEK 293T Cells

In the second design, we aimed to achieve the reverse process of membrane localization: to remove the POI from the plasma membrane. This could be useful for temporally-controlled perturbation of proteins that function when localized on the plasma membrane. We anchored

CAPs-caged TEVcs to the plasma membrane, followed by the POI (**Figure 3-16** A). Under no shield-1 condition, TEVcs should be inaccessible to the co-expressed TEV protease, and POI should remain localized on the plasma membrane. With shield-1, TEVcs should be uncaged and cleaved by TEV protease, releasing POI from the cell membrane. **Figure 3-16** B shows that prior to shield-1 addition, the example POI, EGFP, was almost exclusively bound to the plasma membrane. Adding shield-1 depleted the EGFP from the plasma membrane and significantly increased its presence in the cytoplasm.

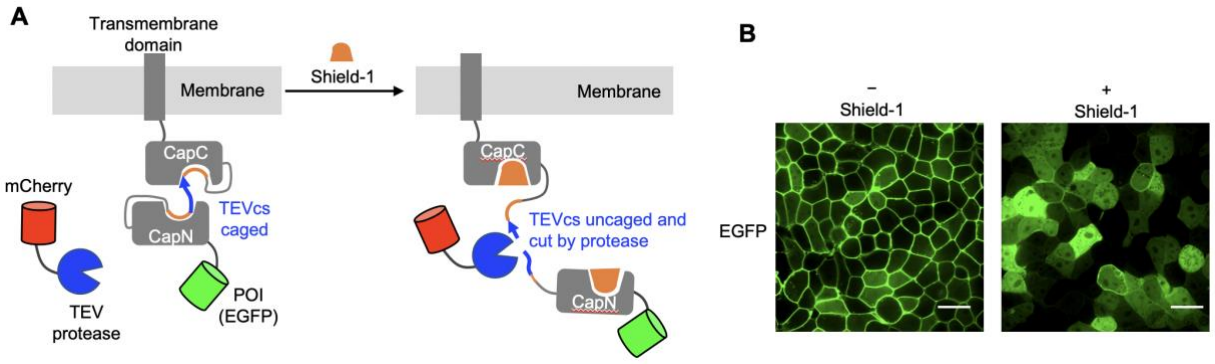


Figure 3-16 Schematics (A) and representative fluorescence microscopy images (B) for CAPs-controlled membrane depletion of a protein of interest. Protease cleavage of TEVcs allows POI (EGFP as example) to be removed from the plasma membrane. mCherry was used as a protease expression marker. Transmembrane domain was CAAX. TEVcs, TEV protease cleavage site (ENLYFQ/M, cleaved between Q and M). POI, protein of interest. Scale bar, 20 μm.

Therefore, this system can delocalize POI from its membrane location to cytoplasm location in a shield-1-dependent manner. Similar to CAPs-caged SsrA, when directly transferred from the yeast surface to mammalian cell culture, CAPs were still capable of caging TEVcs efficiently. The only modification needed was to change the P1' position (the last amino acid) on TEVcs from the canonical glycine to a less active methionine to reduce unwanted cleavage, as

TEV protease is present for much longer time in this assay (2-3 days) than in the yeast surface assay (3 hours).

CAPs can Control the Nuclear-Cytosolic Distribution of Proteins in HEK 293T Cells

The nucleocytoplasmic distribution of many eukaryotic proteins is a common determinant of their functions.[283] Previously, nuclear localization signal peptides (NLS) has been controlled by light through customized engineering with the AsLOV2 domain.[189,190] To test if NLS can be controlled by CAPs in a shield-1-dependent manner, we designed a single-chain construct where POI is expressed as a fusion protein with CAPs-caged NLS (**Figure 3-17** A). In the no shield-1 condition, NLS should be sterically blocked, and POI should be found throughout the cell. Addition of shield-1 should open up CAPs and expose NLS to endogenous importins, bringing more POI to the nucleus. Due to the strength of the NLS, we added a weak nuclear export signal, PKIt[189], to the construct to reduce the nuclear localization of the protein in the basal state. **Figure 3-17** B shows that the example POI, EGFP, was found both in the cytosol and the nucleus when there was no shield-1. Upon addition of shield-1, EGFP was depleted from the cytosol and preferentially localized in the nucleus with a statistically different cytosol-to-whole-cell ratio than that in the basal state.

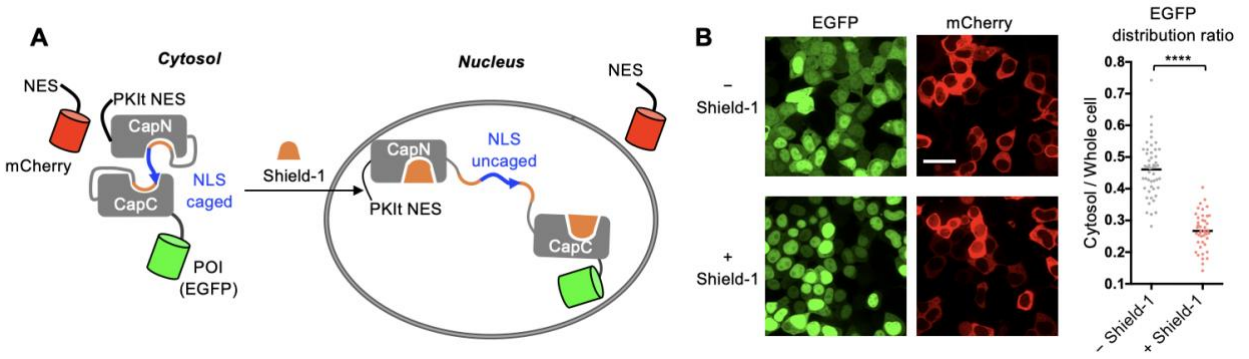


Figure 3-17 Schematics (A) and results (B) for CAPs controlled nuclear localization by caging a nuclear localization signal peptide. Shield-1-dependent uncaging of NLS brought POI (EGFP as example) from cytosol to nucleus. mCherry was used to indicate the cytosol. NLS, nuclear localization signal peptide (PKKKRKV). POI, protein of interest. NES, nuclear export signal peptide (LQLPPLERLTLD). PKIt NES, truncated cAMP-dependent protein kinase inhibitor alpha (PKIt) NES (LALKLAGLDI). The right panel shows the quantification of EGFP total intensity distribution. The ratio was calculated by the EGFP total intensity in cytosol to that in whole cell. The cytosol and whole cell was determined by mCherry. The center lines indicated mean values of the ratio. P value was determined by unpaired two-tailed t-tests. ****P < 0.0001. Scale bar, 20 μm.

This study demonstrates the general applicability of CAPs: they not only can cage TEVcs and SsrA which they were initially optimized with, but can be transferred to other peptides such as NLS. However, there was some nuclear localization pattern in the absence of shield-1 (**Figure 3-17** B), suggesting that the NLS is not completely blocked by CAPs. This is possibly because the residues in the middle of the NLS sequence are also important for importin recognition and nuclear localization, while CAPs are best at caging residues at the N- and C-termini of a peptide. To further eliminate nuclear pattern in the absence of shield-1, NLS sequences with a weaker strength[189] could be used. We also noticed that under the "+ shield-1" condition, there was a weak cytosolic localization pattern in cells with a high level of EGFP expression. This was possibly due to the equilibrium of nucleus-cytosol distribution. The current system is therefore not suited to activate proteins that will gain their functions in the nucleus. Instead, it can be used to study the function of cytosolic proteins by depleting them from the cytosol.

CAPs can Control Gene Transcription in HEK 293T Cell Culture and Neuronal Culture

Gene transcription is a ubiquitous process in all living organisms, and temporally-controlled gene transcription provides useful means for studying the functional roles of a specific protein in living cells and animals. To demonstrate the utility of CAPs in controlling gene transcription, we designed a split transcription factor based on the two-hybrid system[284] (**Figure 3-18** A). In this design, CAPs-caged SsrA is fused to a DNA-binding domain (DBD), and SspB to a transcription-activation domain, VP16. In the presence of shield-1, SspB-VP16 will be recruited to DBD-SsrA and initiate transcription of the reporter gene.

We first tested this system in HEK 293T cells using the Gal4 DBD and the UAS-mCherry reporter gene. With the eight-amino-acid SsrA peptide (**Figure 3-18** B) used in previous experiments, a ~4-fold and ~25-fold shield-1-induced increase in mCherry intensity was observed for CapN and CAPs caged SsrA, respectively (**Figure 3-18** C). As expected, the tandem use of CAPs showed both lower background and higher dynamic range than using CapN alone. This was consistent with our prior observation of the CAPs-caged TEVcs on yeast surface (**Figure 3-14**). We then sought to further lower the background because transcriptional systems are highly sensitive and are often practically useful only when undesired transcription is reduced to the minimum. It has been previously shown that the amino acid sequence of SsrA is tunable[248,285], so we adjusted the SsrA sequence from both the N- and C-termini (**Figure 3-18** B) to look for more efficient caging. The best result was from a seven-amino-acid SsrA sequence (`AANDENY`) (**Figure 3-18** D), which showed 156-fold shield-1 dependent reporter gene expression change, with two-fold lower background and six-fold higher signal-to-noise ratio (**Figure 3-18** C) than the original sequence tested. In this set of experiments, we found that the expression level of SspB-VP16, shown by EGFP signal, was positively correlated to the amount of shield-1 added and reporter gene

expression. Conditions with higher shield-1 concentration and higher level of reporter gene expression also showed higher EGFP signal, although presumably EGFP expression should be consistent across all conditions. Since shield-1 added to the EGFP-SspB-VP16 alone did not increase the level of EGFP, it suggests that the SsrA-SspB interaction stabilized the SspB protein, and non-interacting SspB was degraded.

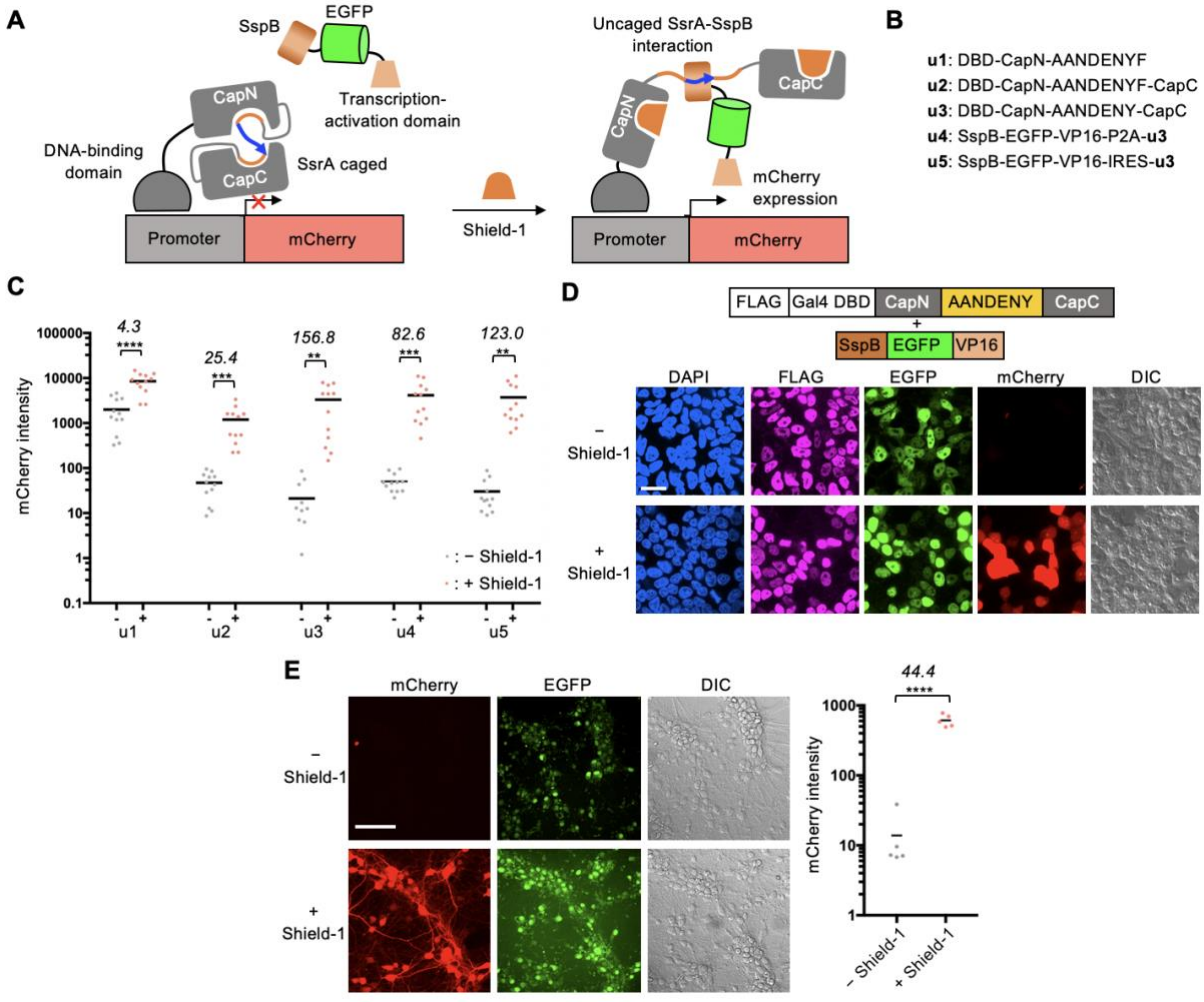


Figure 3-18 CAPs-controlled gene transcription. **(A)** Schematics of shield-1-induced gene transcription. Uncaging of SsrA reconstitutes the split transcription factor and results in reporter gene (mCherry as example) expression. EGFP is used as an expression marker for SspB and transcription-activation domain. Transcription-activation domain is VP16 for all following experiments. DNA-binding domain is specified under each experiment. **(B)** Summary of main constructs tested. Amino acid sequences of SsrA are highlighted. DBD, DNA-binding domain. P2A, a self-cleaving peptide. IRES, internal ribosome entry site. **(C)** Quantification of mCherry expression level for constructs shown in (B). Numbers are the ratio of

mean mCherry intensity of "+ shield-1" to that of "− shield-1" conditions for each construct. The center lines indicate mean values of mCherry intensity. For this experiment, Gal4 was used as DBD, and UAS-mCherry was used as reporter gene. n = 12 for all conditions. **(D)** Representative fluorescence microscopy image of HEK 293T cells expressing the best-performing non-single-component construct, u3. Same DBD and reporter gene as in (C) Scale bar, 20 μm. FLAG is an epitope tag. **(E)** Representative fluorescence microscopy images of rat cortical neurons expressing the single-component construct, u4, and quantification of mCherry expression level. For this experiment, TetR was used as DBD, and TRE-mCherry was used as reporter gene. The number on the plot is the ratio of mean mCherry intensity of "+ shield-1" to that of the "– shield-1" conditions. The center lines indicate mean values of mCherry intensity. n = 5 for both conditions. Scale bar, 100 μm. *P* values were determined by unpaired two-tailed *t*-tests. $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$.

To apply this chemical-dependent transcriptional system in neuroscience studies, we next tested it in cultured neurons. To enable homogeneous expression of DNA in these stringent experiments, we made single viral constructs that express both the DNA-binding and the transcription-activation domains via the self-cleaving peptide P2A or the internal ribosome entry site (IRES) (**Figure 3-18** B). Upon adding shield-1, the P2A and IRES constructs showed 83- and 123-fold increase in mCherry expression, respectively, which are comparable to the two-component system (**Figure 3-18** C). The P2A construct was then used to introduce the TetR DBD and VP16 transcription-activation domain into cultured rat cortical neurons through adeno-associated viruses (AAV), together with another AAV encoding the TRE-mCherry reporter gene. Shield-1 induced 44-fold increase in mCherry reporter gene expression compared to the no shield-1 condition (**Figure 3-18** E), showing that this system could work robustly in cultured neurons.

CAPs can Control Gene Transcription in Living Animals

To enable chemical-dependent peptide regulation in living organisms, we next tested the CAPs system in mouse brain and liver. For the application in brain, we performed stereotactic injection of AAV encoding shield-1-dependent gene regulation constructs into the lateral

hypothalamic area (LHA). Seven days after viral delivery, aquashield-1 (a water-soluble analogue of the shield-1 molecule) was locally administered into LHA. Forty-eight hours after the treatment with aquashield-1 or saline, mice were euthanized and perfused and the brain tissues were processed for analysis (**Figure 3-19** A). In the saline treated control brains ("− aquashield-1"), there were only a few sparse neurons with mCherry expression throughout the entire LHA region (**Figure 3-19** B). In contrast, mCherry was observed in a large cluster of LHA neurons in the aquashield-1 treated brains ("+ aquashield-1"), which is more than 16-fold than the control (**Figure 3-19** B), suggesting a shield-1 dependent gene expression.

To test the application of our chemical-dependent gene regulation system beyond neuronal tissues, we next injected AAVs encoding the CAPs constructs into mouse liver. On the seventh and eighth day after viral delivery, aquashield-1 or saline was administered via intraperitoneal injection for a total of two times (**Figure 3-19** C). Two days after the first aquashield-1 injection, the liver tissues were harvested for analysis. In the control mice injected with saline, limited cells express mCherry in the whole liver, whereas mCherry expressing cells were greatly increased in the liver from the aquashield-1 treated animals (**Figure 3-19** D), demonstrating that systematic injection of aquashield-1 can control CAPs in mouse liver tissues. This result showed the unique advantage of using a small molecule to control peptide activity in comparison to light. Light is confined by its illumination area and is difficult to be applied globally in living animals. In contract, injection of a small molecule brings it to the whole body of the animal, allowing global control of peptide activities in various organs in the body.

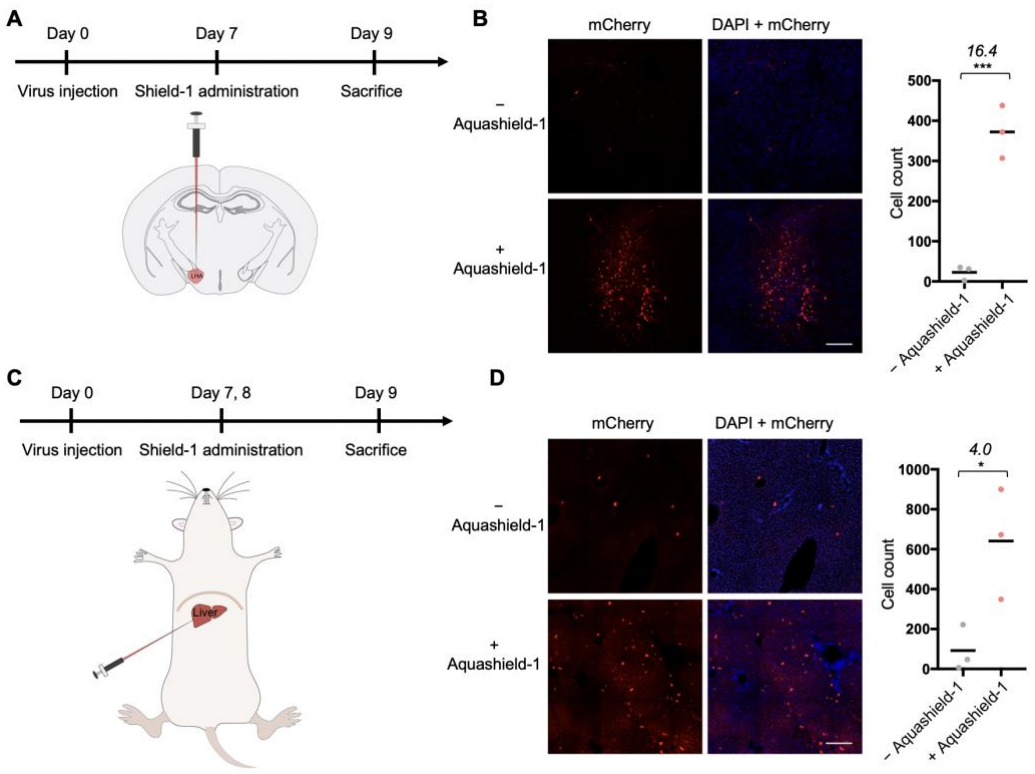


Figure 3-19 CAPs-controlled gene transcription in mouse brain and liver. **(A)** Timeline for the aquashield-1-induced transgene expression in mouse brain. Aquashield-1 was locally administered to mice (1 μL, 1 mM). **(B)** Representative fluorescence microscopy images of brain sections of the lateral hypothalamic area and quantification of total number of cells expressing mCherry. Numbers on the plot are the ratio of mean cell count of "+ aquashield-1" to that of the "− aquashield-1" conditions. The center lines indicate mean values of cell count. n = 3 for both conditions. Scale bar, 200 μm. **(C)** Timeline for the aquashield-1-induced transgene expression in mouse liver. Aquashield-1 is administered to mice via two intraperitoneal (IP) injections (40 mg/kg) with 24 hours apart. **(D)** Representative fluorescence microscopy images of liver sections from injection site and quantification of total number of cells expressing mCherry. Numbers on the plot are the ratio of mean cell count of "+ aquashield-1" to that of the "− aquashield-1" conditions. The center lines indicate mean values of cell count. n = 3 for both conditions. Scale bar, 200 μm. *P* values were determined by unpaired two-tailed *t*-tests. **P* < 0.1; ****P* < 0.001.

Taken together, these results demonstrate that the aquashield-1-induced gene regulation works in multiple organs in living organisms, including brain and liver. In addition, aquashield-1 can be readily administered through intraperitoneal injection to activate gene transcription in animal tissue of interest. This is advantageous over light-induced gene transcription systems as it provides a global control and introduces less disturbance to the animals.

3.7 CapC can Control Opioid Peptide Met-Enkephalin

Besides establishing CapN and CapC as general protein switches for controlling short peptides such as TEVcs, SsrA, and NLS, our other goal was to apply CapC in M-PROBE for controlling opioid peptides. As shown in **Chapter 2.2**, the five-amino-acid met-enkephalin is an ideal effector for photo- or chemical-control due to its short length and activity towards MOR. With CapC, we aimed to control met-enkephalin in a shield-1-dependent manner.

Because we expected CapC-caged met-enkephalin to have very low background and reduced signal compared to the non-caged met-enkephalin, we adapted two highly sensitive assays for measuring CapC activation. The first assay (**Figure 3-20**) was based on a luciferase termed NanoLuc[286] from the deep sea shrimp, *Oplophorus gracilirostris*. Luciferases are light-producing enzymes found from firefly, sea pansy, and shrimp, among other species. They catalyze the conversion of their substrates into luminescence signals. Among these enzymes, the NanoLuc luciferase is one of the most sensitive[287] and has been used to detect protein-protein interactions in a split protein configuration[288]. We developed a split NanoLuc assay for detecting opioid receptor activation by fusing one split half to the opioid receptor and the other half to a G_i-mimic nanobody, Nb44[289]. Opioid receptor activation leads to recruitment of Nb44 to the receptor and reconstitution of the split NanoLuc into a functioning luciferase. Our second assay was a circularly permutated firefly luciferase whose activity has been engineered to be dependent on the cyclic adenosine monophosphate (cAMP) concentration.[290] By expressing this real-time sensor termed GloSensor in cells, they produce higher luminescence when the cAMP level is high and lower luminescence when the cAMP level is low. The GloSensor assay also complements the split NanoLuc assay in that it measures a key downstream event rather than a binding between G protein and the receptor.

Ideally, M-PROBE activation will not only cause G protein recruitment, but also downstream signaling events.

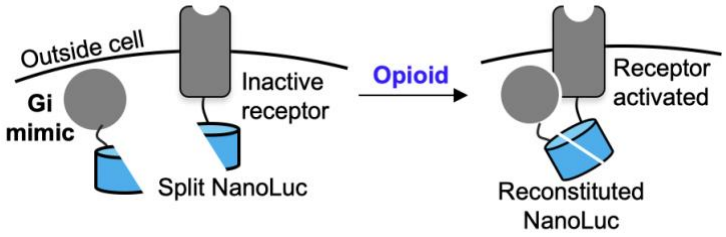


Figure 3-20 Schematics for a split NanoLuc assay for detecting opioid receptor activation.

We first tested CapC-caged met-enkephalin by expressing it as a separate construct (**Figure 3-21**, left. Sequence was YGGFMPNLRPFGSGGSGTGSGSGGS with binder sequence underlined) on the cell surface on top of the constructs that were already expressed in the split NanoLuc assay (one construct expressing an opioid receptor fused to a split NanoLuc half and another construct expressing a G protein mimic fused to another split NanoLuc half). This was our ultimate goal of using M-PROBE because it can target endogenous opioid receptors. However, this test did not result in any detectable signal both in the "+ shield-1" and "– shield-1" conditions.

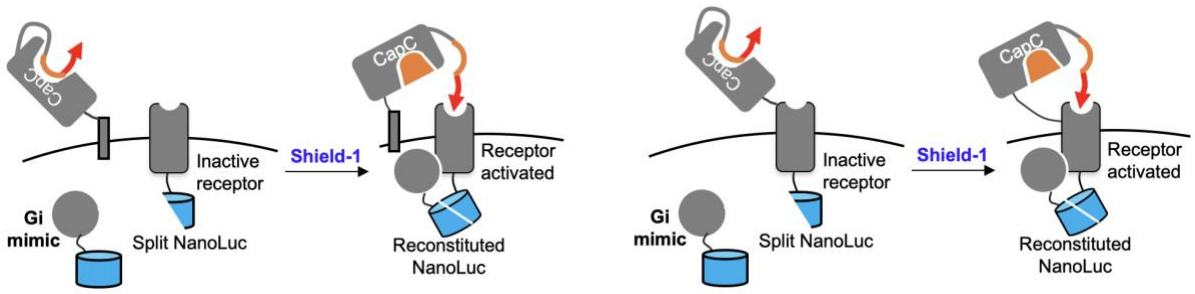


Figure 3-21 Schematics of using a split NanoLuc assay for detecting the activity of CapC-caged met-enkephalin in a two-chain design (left) and a single-chain design (right).

This result was important as it showed that the activity of CapC-caged met-enkephalin was low. There were many possible reasons why this assay did not show signal. First, the opioid peptide might have been too deeply embedded into FKBP12(F36V), and cannot regain its activity even after it was released by sheild-1. This might have been caused by the `PNL` sequence right after the opioid peptide. Second, the geometry of the CapC-caged met-enkephalin construct might have prevented met-enkephalin from binding or activating the opioid receptors. We could not predict how FKBP12(F36V), a large protein compared with met-enkephalin, affect the geometry of met-enkephalin. However, we thought this was unlikely because the linker between FKBP12(F36V) and met-enkephalin was pretty long, and the two membrane-bound proteins can move freely on the cell surface with many interaction possibility at various distances. Another possible reason why this experiment showed no signal was that the expression levels of the proteins were not optimized. In this three-component system, protein expression for all components can vary, and as we had previous seen in SPARK, protein expression level can greatly influence tool performance. It was possible that the M-PROBE expression was too low in comparison to the opioid receptor construct. Lastly, M-PROBE and the opioid receptor might not have had the opportunity to interact due to the two-construct design. For them to interact, both constructs have to come into the appropriate distance. Since the cell outer member is natively cluttered with a variety of proteins and other entities, M-PROBE and opioid receptor might not have had the chance to interact properly.

To troubleshoot this issue and prove that CapC can indeed cage met-enkephalin, we stepped back and simplified the design, as shown in the right panel of **Figure 3-21**. Rather than expressing M-PROBE and opioid receptors separately as we would expect in the final application

of M-PROBE, we fused CapC-caged met-enkephalin directly onto the N-terminus of the opioid receptor. This design solved the problem of low interaction possibilities, and at the same time, ensured that M-PROBE and opioid receptors were expressed in a fixed ratio. Given the proximity of met-enkephalin to the opioid receptor, we expected this design to give much higher signal.

As shown by **Figure 3-22**, shield-1 dependence (between "Shield-1" and "No drug" conditions) was observed for this experimental design. This was the first time that we showed it was possible to control the activity of an opioid receptor through a chemically-gated membrane-tethered opioid peptide. In the same experiment, we also designed controls to provide more information. The "Naloxone" condition represented the basal state in the absence of any receptor activation. In comparison, although the "No drug" state had no external drug molecules added, it did have an CapC-caged met-enkephalin on the cell surface, and its signal would come from the basal activity (i.e. background or leakage) of that M-PROBE. The "Loperamide" condition showed the highest possible signal as loperamide is a full agonist at MOR. As can be seen from the graph, there was a considerable amount of background leakage from M-PROBE. Nonetheless, leakage itself did not cause maximum receptor activation. Both shield-1 and loperamide could bring the signal higher by either uncaging an opioid peptide (met-enkephalin) from CapC or directly acting on MOR. The high background for this experiment was not completely unexpected. Since M-PROBE was directly linked to MOR in this design, it was prone to significant basal activity in the absence of shield-1. It was encouraging to see that CapC was still able to cage met-enkephalin to some extent despite its close proximity to MOR, and that the shield-1-induced met-enkephalin release caused more receptor activation.

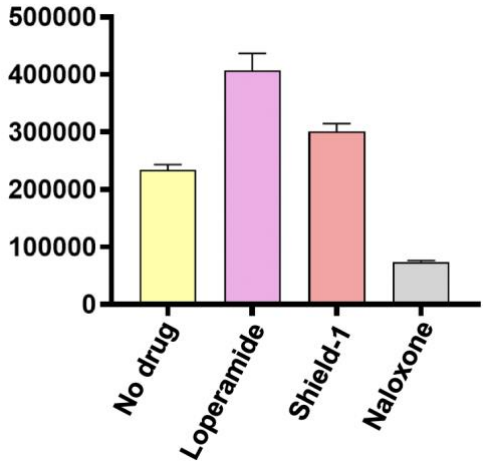


Figure 3-22 Quantification of the split NanoLuc assay for the single-chain M-PROBE-MOR design (**Figure 3-21**, right panel). Y-axis is luminescence signal in relative luminescence unit. X-axis shows the different experimental conditions.

Next, we sought to confirm that M-PROBE can still induce the downstream signaling effects as free opioids. This is an important step to show that the function of met-enkephalin was well-preserved in this membrane-tethered geometry. We expressed the same single chain met-enkephalin-CapC-MOR on the membrane of HEK 293T cells, and co-transfected cells with the cAMP-dependent luciferase. After overnight protein expression, we first stimulated HEK 293T cells with forskolin to artificially enhance the cAMP concentration in the cells. This was because MOR is an inhibitory GPCR, and its activation leads to decreased cAMP level. Starting from a high cAMP concentration would magnify the effect of MOR activation. As shown by **Figure 3-23**, upon forskolin stimulation, cells showed a significant enhancement in luminescence signal. We then treated cells with four different conditions: no drug, shield-1, MOR agonist DAMGO, or MOR antagonist naloxone. Cells with no treatment did not show any change in their luminescence level. In comparison, cells treated with a high concentration (10 micromolar) of DAMGO saw a drastic decrease (~80%) in luminescence. The luminescence level of shield-1-treated cells fell in the middle of these two conditions, and had a ~50% decrease. Interestingly, treating cells with a

high concentration (10 micromolar) of naloxone caused a jump in luminescence signal. The increase was even larger than treating cells with 1 micromolar of forskolin. Hypothetically, naloxone should have no effect on the luminescence level because it only binds to MOR but does not induce any downstream signaling. A reasonable explanation to naloxone-induced signal increase is that M-PROBE has been constantly activating MOR (and thus suppressing cAMP level and luminescence signal) in its basal state. High concentration of naloxone displaced the met-enkephalin in M-PROBE and stopped its MOR binding, causing cAMP level to recover.

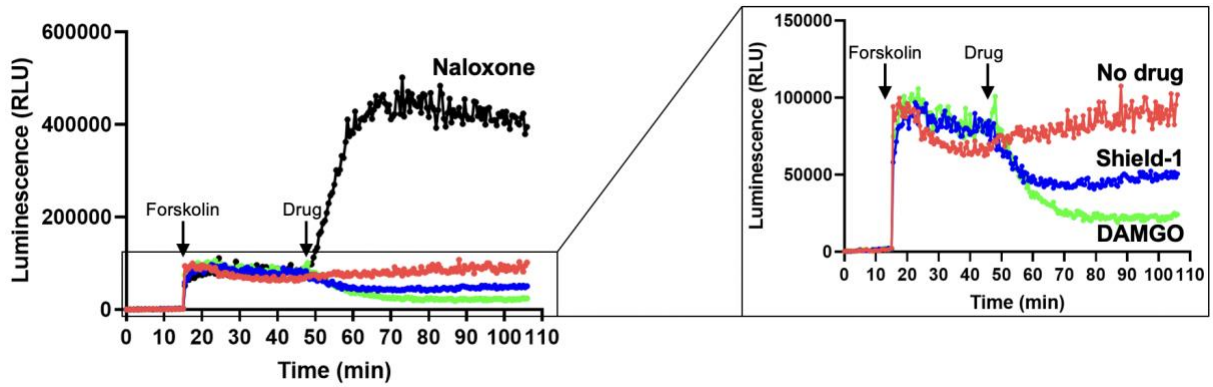


Figure 3-23 Activity of CapC-caged met-enkephalin as detected by the GloSensor cAMP assay. RLU, relative luminescence unit.

These data were consistent with the results (**Figure 3-22**) from the split NanoLuc assay, where the effect of shield-1 was stronger than the "no drug" condition but weaker than the agonist condition. In addition, both assays showed that M-PROBE had a considerable amount of leakage when shield-1 was not present. Nonetheless, the shield-1-induced M-PROBE activation was clearly shown in both assays.

In summary, by using a binding assay and a downstream cAMP assay, we proved that CapC can effectively control met-enkephalin to activate MOR in a drug-dependent manner,

although further optimization is needed to reduce its background leakage. This is the first prototype of a functioning M-PROBE, and theoretically, it can be used to target endogenous opioid receptors in specific cell population or neuronal circuits in the animal brain.

Chapter 4 Concluding Remarks and Future Directions

This thesis demonstrated the engineering of three genetic tools: (1) a cell-membrane-tethered opioid peptide that can activate opioid receptors, (2) a photoswitchable protein that allows reversible control of peptide activities by light, and (3) a pair of generally applicable chemoswitches that can be either individually or tandemly used to control activities of short peptides. The first tool opened the door for genetically targeting and activating endogenous opioid receptors in specific cell types or neuronal circuits for the first time; the second tool added a one-of-a-kind C-terminus-fusing photoswitch to the already diverse toolbox of optogenetics; the third tool presented two complemental chemoswitches at once, representing the first examples of chemically-controlled proteins with true generalizability.

When the tools are used separately, they offer new engineering modalities for creating novel optogenetic and chemogenetic proteins, as shown in **Chapters 2.4**, **2.5**, **3.5**, and **3.6**. When the tools are combined, they offer unprecedented possibility of activating opioid receptors with cell-type-specificity and temporal control at the same time, as demonstrated in **Chapter 3.7**. Future protein engineers, biologists, and neuroscientists will find our tools useful in creating new synthetic biology, interrogating the nuances of the endogenous opioid system and signaling, and possibly designing new gene therapies to fill the inadequacy of the currently available opioid medications. In this chapter, I will summarize the major findings and results from this thesis, recap the novelties of our protein designs, and attempt to propose possible future directions of tool improvement and applications.

4.1 Summary of Results and Impact

The results from this thesis have been shown in detail in **Chapter 2** and **Chapter 3**. This sub-chapter will briefly summarize the results and highlight how their novelties are of importance to different research communities.

Summary of Research Results

The main findings and results of this thesis can be summarized as below:

First, we established the first cell membrane-tethered opioid peptide for activating the opioid receptors. We demonstrated that membrane-anchored opioid peptides can be properly trafficked to the cell surface by a self-cleaving signal peptide, activate opioid receptors, and cause the receptors to recruit a G-protein mimic.

Second, we developed the first LOV-domain-based photoswitchable protein, cpLOV, that can control the effectors by fusing to their C-terminus. We showed that this unique geometry can not only substitute the conventional N-terminal fusion geometry, but can also tune the dynamic range of existing optogenetic tools.

Third, we designed the first chemically dependent protein switches, CapN and CapC, that have general applicability in controlling short peptides. We demonstrated that, unlike their predecessors, CapN and CapC are generalizable towards multiple short peptides, including a protease cleavage site, a protein interacting partner peptide, and a nuclear localization signal peptide.

Fourth, we demonstrated a “dual-caging” strategy for using switchable proteins for the first time. This strategy was made possible by our design of cpLOV and CapC, which complement

*As*LOV2 and CapN. We showed that both the light-controlled and the chemical-controlled pairs reduced background compared to using a single switchable protein.

Lastly, we demonstrated, for the first time, that we can control the activity of a membrane-tethered opioid peptide remotely using a bio-orthogonal small molecule. We further showed that the enhanced opioid peptide activity activated the opioid receptors, causing both G-protein-mimic recruitment and a change in the cellular cAMP level.

Novelty and Expected Impact in the Fields of Protein Engineering, Opioid, and GPCR Research

We expect that the results and novelty from our research will benefit a wide range of research community, especially protein engineers, opioid researchers, and other peptide GPCR researchers.

These results are significant to the protein engineering community because of the following reasons for a couple reasons. First, our cpLOV design is the first switchable LOV that can fuse to the C-terminus of the peptides or proteins under control (another group published a similar cpLOV[291] around the same time of our publication). Although *As*LOV2 is the most widely used genetic photoswitch and have been extensively applied to more than 30 peptides and proteins during the past decade (reviewed in **Chapter 2.3**), it was never engineered to control the C-terminus of a peptide. The new modality of cpLOV enabled more flexible design of photoswitchable protein, doubling the possibilities of optogenetic designs. In the past, for effectors that can obviously be better controlled through their C-terminus, researchers must identify ways to circumvent the need of controlling N-terminal fusions by truncations, insertions, or using novel geometries.[236] With cpLOV, future engineering of such tools will be much easier.

Another reason why protein engineers could benefit from our tools is that we offered two generally applicable chemoswitches to the protein engineering community. Both CapN and CapC are first examples of a drug-dependent protein with a broadly applicable mechanism. Before their invention, designing chemically controlled tools is limited by the need to individually optimize each chemoswitch with the target effector. Only certain laboratories had the full protein engineering capable to design and optimize these tools. CapN and CapC opened more possibilities for researchers to design and apply chemogenetic tools without doing tedious engineering.

Lastly, our tool is appealing to the protein engineering community because we, for the first time, demonstrated a working dual-caging strategy. We not only showed that the natural *As*LOV2 can be used together with the designed cpLOV, but also proved that dual protein switches can be designed from the ground up, as shown by our CapN and CapC. This should stimulate new protein engineering ideas that make use of the extremely low background of dual-caged peptides and proteins. Dual caging is also an effective strategy for re-purposing existing optogenetics or chemogenetic tools. As we have demonstrated in **Chapters 2.5** and **3.5**, dual caging introduced a significant shift in the dynamic range of the current tools. Before us, extensive mutational studies[247,255,292] or directed evolution campaigns[179,248] were required if one wanted to adjust the tool kinetics and dynamic range. Dual caging provids a new and easier possibility of achieving the same results.

Our results are also valuable to the opioid community for the following three reasons. First, for the past several decades, researchers and pharmaceutical companies have been trying to gain understanding of the endogenous opioid system and develop novel opioids with minimized side effects. Although there have been considerable progress towards this goal (reviewed in **Chapters 1.3** and **1.4**), we still do not have the knowledge on the exact functions of opioid receptors in each

cell type and neuronal circuit in the nervous system, and there is yet any drug that can truly eliminate opioid-related side effects while retaining analgesic effects. Our tools brought the theoretical possibility of studying opioid receptors in each individual cell type or neuronal circuit for the first time, opening up possibilities of new insights into the opioid system. Since M-PROBE is fully genetically encoded, it is not limited by adjacency of brain structures or opioid receptors, representing an advancement from the conventional microinjection techniques (reviewed in **Chapter 1.4**).

Second, M-PROBE can not only target cell-membrane-bound opioid receptors as it was originally designed, but can also be used to study intracellular opioid receptors. It is now well-recognized that opioid receptors exist intracellularly[289,293], and the amount of intracellular MOR may even be more than that on the cell surface[162]. Currently, to specifically study intracellular opioid receptors, researchers need to use cell-permeant opioid agonists coupled with non-cell-permeant opioid antagonists. This process introduces additional complexities to data interpretation and suffers from non-specific receptor activation or inhibition. As a membrane-bound genetic tool, M-PROBE can theoretically be targeted to intracellular organelle membranes as well, such as the surface of the Golgi and the mitochondria. By using specific localization signals[273,274], M-PROBE can have minimal expression on the cell outer membrane, circumventing the need to block the membrane-bound receptors.

Third, In a prototypical design (**Chapter 3.7**), we demonstrated for the first time that membrane-tethered opioid peptides can be controlled remotely by a bio-orthogonal small molecule, shield-1. Theoretically, the opioid peptides should also be controllable by light, using the cpLOV domain that we developed (**Chapter 2**). Temporal control is a key consideration when using new tools to study the opioid system as it allows researchers to precisely correlate behavioral

or physiological effects with the time of specific receptor activation. The control mechanism we demonstrated in this thesis should be of great interest to the opioid community as these mechanisms should be transferrable to in vivo studies.

Besides protein engineers and opioid researchers, our tools should also be of interest to researchers studying other GPCRs, especially peptide GPCRs (i.e. GPCRs that have a peptide ligand). GPCRs are the largest class of cell surface receptors and represent the largest protein family encoded by the human genome.[294] With over 800 members[295], they are the most successful family of therapeutic targets, and about one third to half of all the drugs on the market target one GPCR or another. About half of the GPCRs are olfactory GPCRs related to the sense of smell and taste. Among the other ~400 members, at least 118 recognize endogenous peptide or protein ligands.[295] Our study highlighted the possibility of targeting a peptide GPCR (MOR) with cell-type- and neuronal-circuit-specificity, which will allow studying the exact functions of these receptors in cells and circuits of interest. This approach is conceptually generalizable and is not limited to opioid peptides. For other peptide GPCRs with complex localizations and functions, we expect that similar tools can be designed. In addition, for non-peptide GPCRs, if there is an urgent need to use fully genetically-encoded M-PROBE-like tools to study their functions, it might be possible to design and evolve peptides that can act on these GPCRs. Such efforts will require skilled protein engineers to perform directed evolution or computational designs, but the payback can be well worth it.

4.2 Future Directions

The design of our tools should bring forward numerous opportunities for further optimization and applications. This sub-chapter will propose both immediate and long-term directions that future researchers can pursue.

Immediate tool improvements and applications

One apparent direction is to apply cpLOV to control opioid peptides such as met-enkephalin. In the course of this dissertation, we investigated the CapC-caged met-enkephalin with the split NanoLuc and GloSensor assays (**Chapter 3.7**) before attempting the cpLOV-caged met-enkephalin. CapC features a very large dynamic range (over 100-fold difference between the "open" and "closed" states for caging SsrA, **Chapter 3.4**) and low background, but even in this case, there was considerable amount of leakage and only a small drug-dependent signal increase when caging met-enkephalin, as measured by both a G protein binding assay and a cAMP assay (**Chapter 3.7**). Since the cpLOV dynamic range was much smaller on the yeast surface (**Chapter 2.3**), we did not attempt to test cpLOV-caged met-enkephalin.

To tune the dynamic range of cpLOV, directed evolution is most likely needed. Since *As*LOV2 has a relatively well-studied structure-function relationship, targeted mutagenesis can be easily applied. Several regions of interest include the Jα-helix residues that directly or partially face the PAS domain, the "hinge" region connecting Jα-helix with the PAS core, regions around the N-terminal turn-helix-turn, and regions on the β-sheets that form hydrophobic interactions with the Jα-helix. Among these possibilities, the hinge region should be given particular interest, because in previous directed evolution efforts on *As*LOV2, two separate groups using different directed evolution strategies both identified key mutations in the hinge region that were thought to be important to the Jα-helix docking to the PAS core[179,248]. Since our cpLOV uses a GSGS linker

as the hinge, it likely served only as a linker, without strengthening or weakening the Jα-helix docking. We expect that four to six site-saturated mutations would cover most of the possibilities for this region and result in cpLOV variants with a different dynamic range.

The CAPs system may also be further investigated, especially for its exact caging mechanisms. Although CAPs were designed to be generally applicable (**Chapter 3.2**) and did show this characteristic (**Chapter 3.6** and **3.7**), it was a surprise to see that the best CapC for caging SsrA had the sequence of GTPNLRPFG, as glycine and threonine mostly serve in non-functional links with no hydrophobic interaction tendencies. This is in contradiction to the putative hydrophobic active site of FKBP (**Figure 3-1**) and to our molecular modeling (**Appendix 4**) that hydrophobic interactions were important for FKBP interaction. In addition, CapC-caged met-enkephalin saw a significant amount of background leakage (**Chapter 3.7**). We hypothesized that the directed evolution results of CapC might have been impacted by the effector peptide, SsrA. The SsrA sequence used for CapC directed evolution was AANDENYF. The last two residues, tyrosine and phenylalanine, are two of the most hydrophobic and bulky amino acids. It is possible that these residues themselves bound to the FKBP12(F36V) active site. As a result, the first two amino acids on the designed FKBP12(F36V) "binder" became glycine and threonine, two of the smallest amino acids, to accommodate binding of tyrosine and phenylalanine. Should this be the case, the general applicability of CapC should be re-evaluated, and this protein may be re-evolved using TEVcs (ENLYFQG) as the effector peptide.

Nevertheless, we showed that both cpLOV and CAPs are functional switchable proteins that can enable chemical-control and photo-control for peptides, respectively. There are endless possibilities of using cpLOV and CAPs, either in a single-cage design or in a dual-cage design. By controlling peptide activities using our new switchable proteins, researchers will more readily gain

control of signal transduction[296], cell motility[235], and intracellular potassium level[236] among others. Special attention should be paid to peptides and proteins with a critical or functional C-terminus. Prior to cpLOV and CapC, it has been challenging to control such peptides and proteins because the effector site is far away from the fusion site. Although this challenge has been circumvented in the past using insertion strategies or alternative control mechanisms, we provided a more direct way of addressing this issue in this thesis. In addition, we brought opportunities for engineering ultra-low background genetic tools by the combined use of two photoswitches (*As*LOV2 and cpLOV) or chemoswitches (CapN and CapC), as demonstrated in **Chapters 2.5** and **3.5**.

The last immediate application is to use CapC-caged met-enkephalin in animal studies. Despite its leakage and low drug-dependence, interesting biological discoveries can still be made, as this is the first tool that can achieve cell-type-specific or neuronal-circuit-specific opioid receptor activation. We expect that this version of the M-PROBE can be used in both behavioral studies or brain slice physiology studies. Brain circuits with particular interest are those involved in pain, breathing, and reward.

Long term directions

Besides immediate applications, the tools we developed should also be able to serve as basic building blocks for new optogenetic and chemogenetic designs. Here I describe three possible long-term directions.

First, using CapC-caged opioid peptide, there is a possibility to design a novel class of closed-loop cell calcium level stabilizer (**Figure 4-1**). Calcium signaling regulates numerous cellular processes and pathways, and many neurological disorders, such as epilepsy[297] and anxiety disorders[298], involve abnormal calcium levels or calcium channels. Closed-loop medical devices

have become available in recent years, with FDA issuing a draft guidance[299] at the end of 2021. These devices typically require invasive physical implantation, and it would be interesting to design gene therapies that have a built-in feedback control. As shown by **Figure 4-1**, by expressing two genetic constructs on the same cell, it is possible to achieve a calcium-level-dependent cellular activity inhibitor. In this design, the proximity of the two proteins are controlled by a pair of calcium-dependent binder: calmodulin and the MK2 peptide, which have been previously used[179] in engineering calcium-dependent tools. High calcium level brings two proteins together, and when CapC is active, the tethered opioid peptide activates a KOR-based DREADD[170], causing inhibitory effects. When calcium level is low, the opioid peptide is further away from the KOR DREADD, reverting to the non-inhibitory state. If successful, this tool would automatically "detect" high calcium levels in the cells and autoinhibit such abnormalities, bringing potential new solution to epilepsy and beyond.

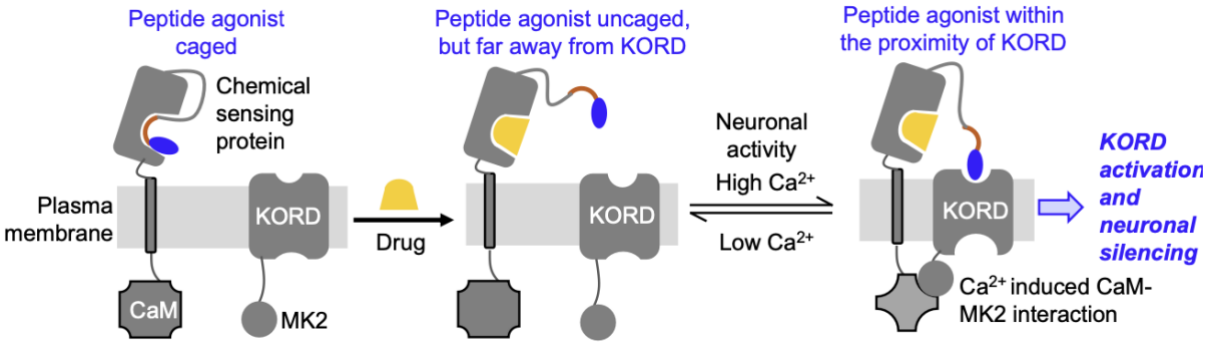


Figure 4-1 Design of a closed-loop cellular calcium level stabilizer based on the chemically controlled M-PROBE. KORD, KOR DREADD. CaM, calmodulin. MK2, a CaM binding peptide.

Another long-term direction is to develop a single-chain M-PROBE fused to the N-terminus of MOR as a new DREADD-like tool. DREADDs are one the most widely used chemogenetic tools for manipulating specific neuronal or non-neuronal cell populations. A main

disadvantage of DREADDs in that the most popular DREADD ligand, CNO, is not as selective as it was thought, and can be reverse-metabolized into a compound that can interfere with normal physiological functions.[300] By converting our single-chain M-PROBE MOR (**Figure 3-21**, right) into an orthogonal opioid receptor, we could circumvent the need to use the traditional DREADD ligands. This will also enable multiplexed control of cellular activities, since our design uses a different small molecule than DREADDs.

Lastly, our designs of CapC- or cpLOV-caged opioid peptides can be a prototype for gene therapy themselves. Over two decades of research and drug development efforts have not identified a novel opioid with analgesic effects and minimized side effects at the same time. As gene therapy is gaining popularity, we believe it would be reasonable to consider gene therapy targeting pain (after careful ethics and safety reviews), especially for patients with a terminal illness. If it is possible to differentiate the opioid analgesic functions from other effects by targeting specific cell types or neuronal circuits, the tools shown in this thesis would offer a direct prototype.

Appendices

Appendix 1. Further Characterization and Quantification of the SPARK Assay for Measuring MOR Activation

Related to **Chapter 2.2**

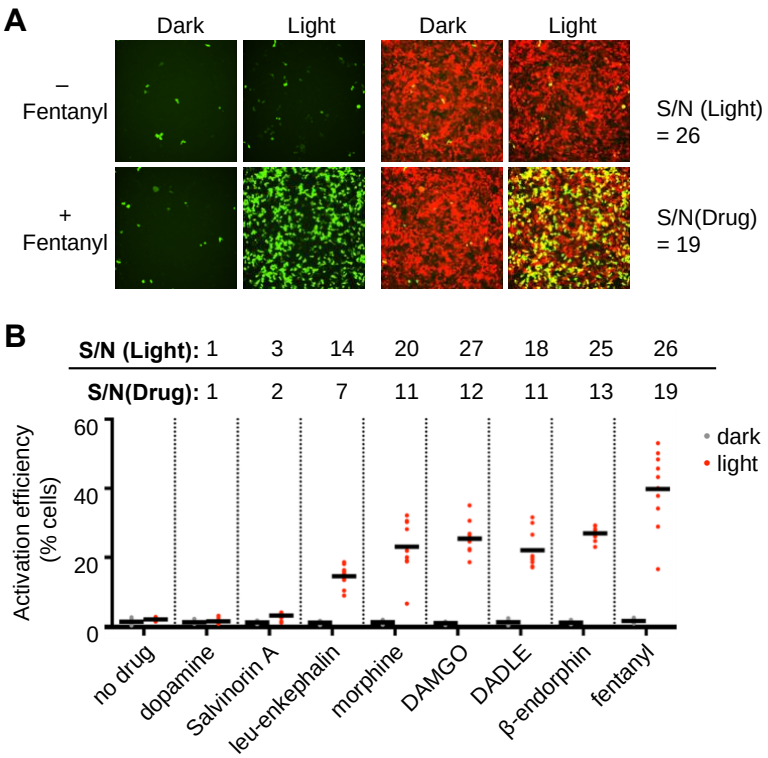


Figure A-1 Representative MOR SPARK assay results using fentanyl as the agonist (A) and quantification of MOR SPARK assay results for a panel of opioids. Cells were stimulated with light and opioids for 20 minutes. Green is reporter activation; red is protease expression marker. Activation efficiency was quantified by the ratio of the number of green-expressing cells to that of the red-expressing cells. Cells were counted using a built-in Nikon analysis software. S/N, signal-to-noise ratio.

Appendix 2. Testing the Pre-Evolution cpLOV in Caging TEVcs Using a Yeast Surface Assay

Related to **Chapter 2.3**

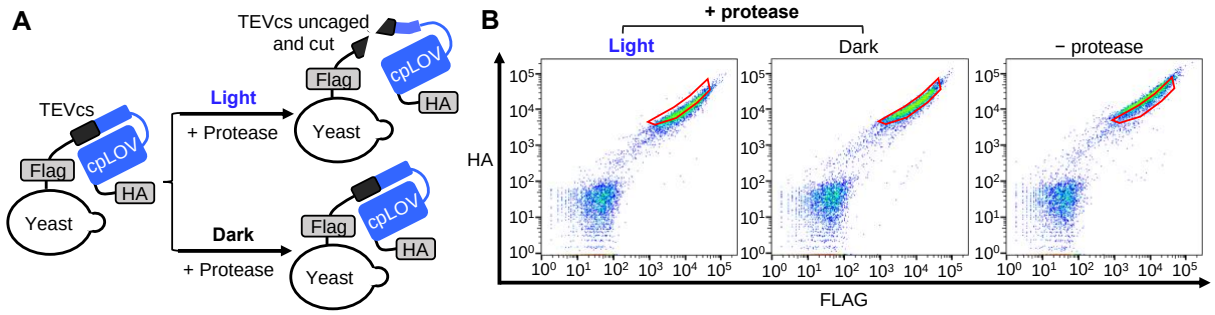


Figure A-2 Testing of pre-evolution cpLOV in caging TEVcs on the yeast surface. The encircled area on the FACS plots represent yeast cells treated with protease in the dark. The yeast population in the light condition shifted downward, showing protease cleavage.

Appendix 3. Raw Flow Cytometry Data for the Rationally Designed SsrA-cpLOV Constructs Tested

Related to **Chapter 2.3**

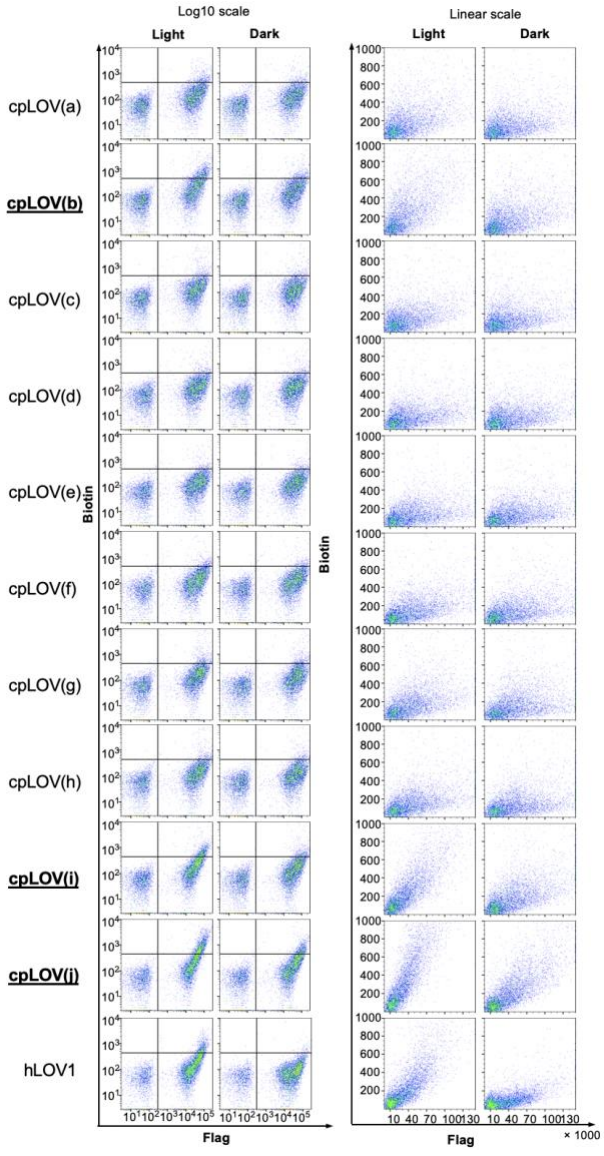


Figure A-3 Raw flow cytometry data for cpLOV-caged SsrA on the yeast surface.

Appendix 4. All-Atom Molecular Dynamics Simulations for the Binding Between the Putative CAPs Binder Sequences to the FKBP12(F36V) Active Site

Related to **Chapter 3.4**

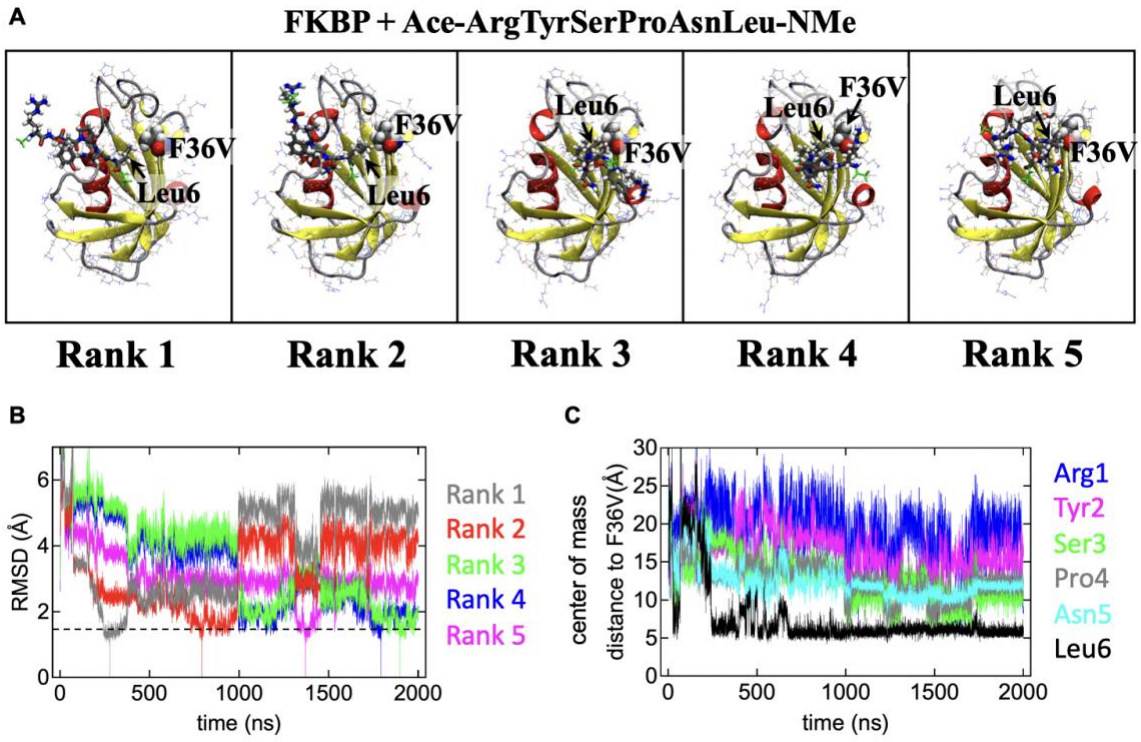


Figure A-4 A two microsecond all-atom molecular dynamics simulations for the binding between the putative CAPs binder sequences (a capped ArgTyrSerProAsnLeu peptide) to the FKBP12(F36V) active site, in 150 mM of buffer. **(A)** The central configurations for the top 5 clusters (Rank 1-5) obtained from RMSD clustering indicate direct interactions between Leu6 of the peptide (shown in a "licorice" representation; cap residues are shown in green, other atoms in CPK colors with gray carbons) and the F36V binding site of FKBP (shown as van-der-Waals spheres). The secondary structure of the FKBP protein is shown in a cartoon representation with red α-helices and yellow β-sheets. **(B)** RMSD time traces with respect to the structures shown in a indicate the longevity of the respective conformations within the simulations. RMSD of 0 indicate the simulation time points corresponding to the structures in (A). A horizontal dashed line indicates the 1.5 Å cutoff used for clustering. **(C)** Time traces of the center of mass distances between each individual sidechain of the peptide and the sidechain of the F36V binding site indicate a persistent proximity of Leu6 to the binding site for a large fraction of the simulation

trajectory (distances of 5-6 Å). Fractions of the simulation trajectory with close proximity of Leu6 to the F36V binding site include all configurations associated with the top 5 clusters shown in (A).

References

1. Macht, D. I. The history of opium and some of its preparations and alkaloids. *J. Am. Med. Assoc.* **LXIV**, 477–481 (1915).
2. Centers for Disease Control and Prevention. Prescription Opioids. https://www.cdc.gov/opioids/basics/prescribed.html (2021).
3. Fleming, M. F., Balousek, S. L., Klessig, C. L., Mundt, M. P. & Brown, D. D. Substance use disorders in a primary care sample receiving daily opioid therapy. *J. Pain* **8**, 573–582 (2007).
4. U.S. Department of Health and Human Services. *Key Substance Use and Mental Health Indicators in the United States: Results from the 2019 National Survey on Drug Use and Health*. (2019).
5. Centers for Disease Control and Prevention. *Vital Signs: Overdoses of Prescription Opioid Pain Relievers --- United States, 1999--2008*. 60(43);1487-1492 (2011).
6. Centers for Disease Control and Prevention. *Increases in Heroin Overdose Deaths — 28 States, 2010 to 2012*. 63(39);849-854 (2014).
7. Centers for Disease Control and Prevention. *Deaths Involving Fentanyl, Fentanyl Analogs, and U-47700 — 10 States, July–December 2016*. 66(43);1197-1202 (2017).
8. CDC WONDER. https://wonder.cdc.gov/.
9. Centers for Disease Control and Prevention. Understanding the Opioid Overdose Epidemic. https://www.cdc.gov/opioids/basics/epidemic.html (2022).

10. U.S. Health and Human Services. HHS Acting Secretary Declares Public Health Emergency to Address National Opioid Crisis. https://public3.pagefreezer.com/browse/HHS.gov/31-12-2020T08:51/https://www.hhs.gov/about/news/2017/10/26/hhs-acting-secretary-declares-public-health-emergency-address-national-opioid-crisis.html.

11. Skolnick, P. The Opioid Epidemic: Crisis and Solutions. *Annu. Rev. Pharmacol. Toxicol.* **58**, 143–159 (2018).

12. Webster, L. *et al.* Human Abuse Potential of the New Opioid Analgesic Molecule NKTR-181 Compared with Oxycodone. *Pain Med. Off. J. Am. Acad. Pain Med.* **19**, 307–318 (2018).

13. Park, K. & Otte, A. Prevention of Opioid Abuse and Treatment of Opioid Addiction: Current Status and Future Possibilities. *Annu. Rev. Biomed. Eng.* **21**, 61–84 (2019).

14. Angst, M. S. & Drover, D. R. Pharmacology of drugs formulated with DepoFoam: a sustained release drug delivery system for parenteral administration using multivesicular liposome technology. *Clin. Pharmacokinet.* **45**, 1153–1176 (2006).

15. Doe-Simkins, M., Walley, A. Y., Epstein, A. & Moyer, P. Saved by the Nose: Bystander-Administered Intranasal Naloxone Hydrochloride for Opioid Overdose. *Am. J. Public Health* **99**, 788–791 (2009).

16. Heidbreder, C. A. & Newman, A. H. Current perspectives on selective dopamine D3 receptor antagonists as pharmacotherapeutics for addictions and related disorders. *Ann. N. Y. Acad. Sci.* **1187**, 4–34 (2010).

17. Bubar, M. J. & Cunningham, K. A. Prospects for serotonin 5-HT2R pharmacotherapy in psychostimulant abuse. *Prog. Brain Res.* **172**, 319–346 (2008).

18. Higgins, G. A. & Fletcher, P. J. Therapeutic Potential of 5-HT2C Receptor Agonists for Addictive Disorders. *ACS Chem. Neurosci.* **6**, 1071–1088 (2015).

19. U.S. Food and Drug Administration. As part of efforts to combat opioid crisis, FDA launches innovation challenge to spur development of medical devices – including digital health and diagnostics – that target pain, addiction and diversion. *FDA* https://www.fda.gov/news-events/press-announcements/part-efforts-combat-opioid-crisis-fda-launches-innovation-challenge-spur-development-medical-devices (2018).

20. Rogers, E., Mehta, S., Shengelia, R. & Reid, M. C. Four Strategies for Managing Opioid-Induced Side Effects in Older Adults. *Clin. Geriatr.* **21**, (2013).

21. Matthes, H. W. *et al.* Loss of morphine-induced analgesia, reward effect and withdrawal symptoms in mice lacking the mu-opioid-receptor gene. *Nature* **383**, 819–823 (1996).

22. Bohn, L. M. *et al.* Enhanced morphine analgesia in mice lacking beta-arrestin 2. *Science* **286**, 2495–2498 (1999).

23. Bohn, L. M., Gainetdinov, R. R., Lin, F. T., Lefkowitz, R. J. & Caron, M. G. Mu-opioid receptor desensitization by beta-arrestin-2 determines morphine tolerance but not dependence. *Nature* **408**, 720–723 (2000).

24. Bohn, L. M., Lefkowitz, R. J. & Caron, M. G. Differential mechanisms of morphine antinociceptive tolerance revealed in (beta)arrestin-2 knock-out mice. *J. Neurosci. Off. J. Soc. Neurosci.* **22**, 10494–10500 (2002).

25. Raehal, K. M., Walker, J. K. L. & Bohn, L. M. Morphine side effects in beta-arrestin 2 knockout mice. *J. Pharmacol. Exp. Ther.* **314**, 1195–1201 (2005).

26. Groer, C. E. *et al.* An opioid agonist that does not induce mu-opioid receptor--arrestin interactions or receptor internalization. *Mol. Pharmacol.* **71**, 549–557 (2007).

27. Manglik, A. *et al.* Structure-based discovery of opioid analgesics with reduced side effects. *Nature* **537**, 185–190 (2016).

28. DeWire, S. M. *et al.* A G protein-biased ligand at the μ-opioid receptor is potently analgesic with reduced gastrointestinal and respiratory dysfunction compared with morphine. *J. Pharmacol. Exp. Ther.* **344**, 708–717 (2013).

29. U.S. Food and Drug Administration. FDA Briefing Document. https://www.fda.gov/media/121233/download (2018).

30. TRV734. *Trevena, Inc.* https://www.trevena.com/pipeline/trv734.

31. Hill, R. *et al.* The novel μ-opioid receptor agonist PZM21 depresses respiration and induces tolerance to antinociception. *Br. J. Pharmacol.* **175**, 2653–2661 (2018).

32. Kruegel, A. C. *et al.* Synthetic and Receptor Signaling Explorations of the Mitragyna Alkaloids: Mitragynine as an Atypical Molecular Framework for Opioid Receptor Modulators. *J. Am. Chem. Soc.* **138**, 6754–6764 (2016).

33. Váradi, A. *et al.* Mitragynine/Corynantheidine Pseudoindoxyls As Opioid Analgesics with Mu Agonism and Delta Antagonism, Which Do Not Recruit β-Arrestin-2. *J. Med. Chem.* **59**, 8381–8397 (2016).

34. Schmid, C. L. *et al.* Bias Factor and Therapeutic Window Correlate to Predict Safer Opioid Analgesics. *Cell* **171**, 1165-1175.e13 (2017).

35. Conibear, A. E. & Kelly, E. A Biased View of μ-Opioid Receptors? *Mol. Pharmacol.* **96**, 542–549 (2019).

36. Rivero, G. *et al.* Endomorphin-2: a biased agonist at the μ-opioid receptor. *Mol. Pharmacol.* **82**, 178–188 (2012).

37. Thompson, G. L. *et al.* Systematic analysis of factors influencing observations of biased agonism at the mu-opioid receptor. *Biochem. Pharmacol.* **113**, 70–87 (2016).

38. Burgueño, J. *et al.* A Complementary Scale of Biased Agonism for Agonists with Differing Maximal Responses. *Sci. Rep.* **7**, 15389 (2017).

39. Mebias Discovery. News. https://mebiasdiscovery.com/news/ (2017).

40. Ma, M. *et al.* Synthesis and Evaluation of Novel Biased μ-Opioid-Receptor (μOR) Agonists. *Mol. Basel Switz.* **24**, E259 (2019).

41. Kliewer, A. *et al.* Phosphorylation-deficient G-protein-biased μ-opioid receptors improve analgesia and diminish tolerance but worsen opioid side effects. *Nat. Commun.* **10**, 367 (2019).

42. Kliewer, A. *et al.* Morphine-induced respiratory depression is independent of β-arrestin2 signalling. *Br. J. Pharmacol.* **177**, 2923–2931 (2020).

43. Montandon, G. *et al.* G-protein-gated Inwardly Rectifying Potassium Channels Modulate Respiratory Depression by Opioids. *Anesthesiology* **124**, 641–650 (2016).

44. Liang, X., Yong, Z. & Su, R. Inhibition of protein kinase A and GIRK channel reverses fentanyl-induced respiratory depression. *Neurosci. Lett.* **677**, 14–18 (2018).

45. Seward, E., Hammond, C. & Henderson, G. Mu-opioid-receptor-mediated inhibition of the N-type calcium-channel current. *Proc. Biol. Sci.* **244**, 129–135 (1991).

46. Nickols, H. H. & Conn, P. J. Development of allosteric modulators of GPCRs for treatment of CNS disorders. *Neurobiol. Dis.* **61**, 55–71 (2014).

47. Burford, N. T. *et al.* Discovery of positive allosteric modulators and silent allosteric modulators of the μ-opioid receptor. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 10830–10835 (2013).

48. Bisignano, P. *et al.* Ligand-Based Discovery of a New Scaffold for Allosteric Modulation of the μ-Opioid Receptor. *J. Chem. Inf. Model.* **55**, 1836–1843 (2015).

49. Kandasamy, R. *et al.* Positive allosteric modulation of the mu-opioid receptor produces analgesia with reduced side effects. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2000017118 (2021).

50. Pryce, K. D. *et al.* A promising chemical series of positive allosteric modulators of the μ-opioid receptor that enhance the antinociceptive efficacy of opioids but not their adverse effects. *Neuropharmacology* **195**, 108673 (2021).

51. Ohbuchi, K. *et al.* Ignavine: a novel allosteric modulator of the μ opioid receptor. *Sci. Rep.* **6**, 31748 (2016).

52. Meguro, Y. *et al.* Neuropeptide oxytocin enhances μ opioid receptor signaling as a positive allosteric modulator. *J. Pharmacol. Sci.* **137**, 67–75 (2018).

53. Kiss, A. & Mikkelsen, J. D. Oxytocin--anatomy and functional assignments: a minireview. *Endocr. Regul.* **39**, 97–105 (2005).

54. Miyano, K. *et al.* Oxytocin Is a Positive Allosteric Modulator of κ-Opioid Receptors but Not δ-Opioid Receptors in the G Protein Signaling Pathway. *Cells* **10**, 2651 (2021).

55. Remesic, M., Hruby, V. J., Porreca, F. & Lee, Y. S. Recent Advances in the Realm of Allosteric Modulators for Opioid Receptors for Future Therapeutics. *ACS Chem. Neurosci.* **8**, 1147–1158 (2017).

56. Morphy, R. & Rankovic, Z. Designed Multiple Ligands. An Emerging Drug Discovery Paradigm. *J. Med. Chem.* **48**, 6523–6543 (2005).

57. Cavalli, A. *et al.* Multi-target-directed ligands to combat neurodegenerative diseases. *J. Med. Chem.* **51**, 347–372 (2008).

58. Schiller, P. W. Bi- or multifunctional peptide drugs. *Life Sci.* **86**, 598–603 (2010).

59. Abdelhamid, E. E., Sultana, M., Portoghese, P. S. & Takemori, A. E. Selective blockage of delta opioid receptors prevents the development of morphine tolerance and dependence in mice. *J. Pharmacol. Exp. Ther.* **258**, 299–303 (1991).

60. Fundytus, M. E., Schiller, P. W., Shapiro, M., Weltrowska, G. & Coderre, T. J. Attenuation of morphine tolerance and dependence with the highly selective delta-opioid receptor antagonist TIPP[psi]. *Eur. J. Pharmacol.* **286**, 105–108 (1995).

61. Hepburn, M. J., Little, P. J., Gingras, J. & Kuhn, C. M. Differential effects of naltrindole on morphine-induced tolerance and physical dependence in rats. *J. Pharmacol. Exp. Ther.* **281**, 1350–1356 (1997).

62. Schiller, P. W. *et al.* Differential stereochemical requirements of mu vs. delta opioid receptors for ligand binding and signal transduction: development of a class of potent and highly delta-selective peptide antagonists. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 11871–11875 (1992).

63. Schiller, P. W. *et al.* The Opioid μ Agonist/δ Antagonist DIPP-NH2[Ψ] Produces a Potent Analgesic Effect, No Physical Dependence, and Less Tolerance than Morphine in Rats. *J. Med. Chem.* **42**, 3520–3526 (1999).

64. Schiller, P. W. *et al.* Synthesis and in vitro opioid activity profiles of DALDA analogues. *Eur. J. Med. Chem.* **35**, 895–901 (2000).

65. Shimoyama, M., Shimoyama, N., Zhao, G. M., Schiller, P. W. & Szeto, H. H. Antinociceptive and respiratory effects of intrathecal H-Tyr-D-Arg-Phe-Lys-NH2 (DALDA) and [Dmt1] DALDA. *J. Pharmacol. Exp. Ther.* **297**, 364–371 (2001).

66. Szeto, H. H., Soong, Y., Wu, D., Qian, X. & Zhao, G.-M. Endogenous opioid peptides contribute to antinociceptive potency of intrathecal [Dmt1]DALDA. *J. Pharmacol. Exp. Ther.* **305**, 696–702 (2003).

67. Zhao, G.-M. *et al.* Profound Spinal Tolerance after Repeated Exposure to a Highly Selective μ-Opioid Peptide Agonist: Role of δ-Opioid Receptors. *J. Pharmacol. Exp. Ther.* **302**, 188–196 (2002).

68. Ben, Y., Smith, A. P., Schiller, P. W. & Lee, N. M. Tolerance develops in spinal cord, but not in brain with chronic [Dmt1]DALDA treatment. *Br. J. Pharmacol.* **143**, 987–993 (2004).

69. Hruby, V. J. *et al.* Design of novel peptide ligands which have opioid agonist activity and CCK antagonist activity for the treatment of pain. *Life Sci.* **73**, 699–704 (2003).

70. Hoffmann, O. & Wiesenfeld-Hallin, Z. The CCK-B receptor antagonist Cl 988 reverses tolerance to morphine in rats. *Neuroreport* **5**, 2565–2568 (1994).

71. Maldonado, R. *et al.* Inhibition of morphine withdrawal by the association of RB 101, an inhibitor of enkephalin catabolism, and the CCKB antagonist PD-134,308. *Br. J. Pharmacol.* **114**, 1031–1039 (1995).

72. Yamamoto, T. *et al.* Design, Synthesis, and Biological Evaluation of Novel Bifunctional C-Terminal-Modified Peptides for δ/μ Opioid Receptor Agonists and Neurokinin-1 Receptor Antagonists. *J. Med. Chem.* **50**, 2779–2786 (2007).

73. Yamamoto, T. *et al.* A Structure–Activity Relationship Study and Combinatorial Synthetic Approach of C-Terminal Modified Bifunctional Peptides That Are δ/μ Opioid Receptor Agonists and Neurokinin 1 Receptor Antagonists. *J. Med. Chem.* **51**, 1369–1376 (2008).

74. Powell, K. J., Quirion, R. & Jhamandas, K. Inhibition of neurokinin-1-substance P receptor and prostanoid activity prevents and reverses the development of morphine tolerance in vivo and the morphine-induced increase in CGRP expression in cultured dorsal root ganglion neurons. *Eur. J. Neurosci.* **18**, 1572–1583 (2003).

75. Kalso, E. Improving opioid effectiveness: from ideas to evidence. *Eur. J. Pain Lond. Engl.* **9**, 131–135 (2005).

76. Cunningham, C. W., Elballa, W. M. & Vold, S. U. Bifunctional opioid receptor ligands as novel analgesics. *Neuropharmacology* **151**, 195–207 (2019).

77. Zaveri, N., Jiang, F., Olsen, C., Polgar, W. & Toll, L. Designing Bifunctional NOP receptor-mu opioid receptor ligands from NOP receptor-selective scaffolds. Part I. *Bioorg. Med. Chem. Lett.* **23**, 3308–3313 (2013).

78. Tian, J.-H. *et al.* Bidirectional modulatory effect of orphanin FQ on morphine-induced analgesia: antagonism in brain and potentiation in spinal cord of the rat. *Br. J. Pharmacol.* **120**, 676–680 (1997).

79. Mogil, J. S. *et al.* Orphanin FQ is a functional anti-opioid peptide. *Neuroscience* **75**, 333–337 (1996).

80. Di Giannuario, A., Pieretti, S., Catalani, A. & Loizzo, A. Orphanin FQ reduces morphine-induced dopamine release in the nucleus accumbens: a microdialysis study in rats. *Neurosci. Lett.* **272**, 183–186 (1999).

81. Di Giannuario, A. & Pieretti, S. Nociceptin differentially affects morphine-induced dopamine release from the nucleus accumbens and nucleus caudate in rats. *Peptides* **21**, 1125–1130 (2000).

82. Hamdy, M. M., Elbadr, M. M. & Barakat, A. Fluoxetine uses in nociceptive pain management: a promising adjuvant to opioid analgesics. *Fundam. Clin. Pharmacol.* **32**, 532–546 (2018).

83. Haleem, D. J. Serotonin-1A receptor dependent modulation of pain and reward for improving therapy of chronic pain. *Pharmacol. Res.* **134**, 212–219 (2018).

84. Emery, M. A., Bates, M. L. S., Wellman, P. J. & Eitan, S. Differential effects of oxycodone, hydrocodone, and morphine on the responses of D2/D3 dopamine receptors. *Behav. Brain Res.* **284**, 37–41 (2015).

85. Dai, W.-L. *et al.* Blockade of neuronal dopamine D2 receptor attenuates morphine tolerance in mice spinal cord. *Sci. Rep.* **6**, 38746 (2016).

86. Rivera, A. *et al.* Dopamine D4 receptor stimulation prevents nigrostriatal dopamine pathway activation by morphine: relevance for drug addiction. *Addict. Biol.* **22**, 1232–1245 (2017).

87. Bushlin, I., Gupta, A., Stockton, S. D., Miller, L. K. & Devi, L. A. Dimerization with cannabinoid receptors allosterically modulates delta opioid receptor activity during neuropathic pain. *PloS One* **7**, e49789 (2012).

88. Grenald, S. A. *et al.* Synergistic attenuation of chronic pain using mu opioid and cannabinoid receptor 2 agonists. *Neuropharmacology* **116**, 59–70 (2017).

89. Maguire, D. R. & France, C. P. Reinforcing effects of opioid/cannabinoid mixtures in rhesus monkeys responding under a food/drug choice procedure. *Psychopharmacology (Berl.)* **235**, 2357–2365 (2018).

90. Maguire, D. R. & France, C. P. Antinociceptive effects of mixtures of mu opioid receptor agonists and cannabinoid receptor agonists in rats: Impact of drug and fixed-dose ratio. *Eur. J. Pharmacol.* **819**, 217–224 (2018).

91. Tris Pharma, Inc. *Efficacy, Safety, and Tolerability of Oral Cebranopadol Versus Morphine Sulfate PR in Subjects With Chronic Moderate to Severe Pain Related to Cancer.* https://clinicaltrials.gov/ct2/show/NCT01964378 (2021).

92. Pasternak, G. W. & Snyder, S. H. Identification of novel high affinity opiate receptor binding in rat brain. *Nature* **253**, 563–565 (1975).

93. Wolozin, B. L. & Pasternak, G. W. Classification of multiple morphine and enkephalin binding sites in the central nervous system. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 6181–6185 (1981).

94. Ling, G. S., Spiegel, K., Lockhart, S. H. & Pasternak, G. W. Separation of opioid analgesia from respiratory depression: evidence for different receptor mechanisms. *J. Pharmacol. Exp. Ther.* **232**, 149–155 (1985).

95. Heyman, J. S., Williams, C. L., Burks, T. F., Mosberg, H. I. & Porreca, F. Dissociation of opioid antinociception and central gastrointestinal propulsion in the mouse: studies with naloxonazine. *J. Pharmacol. Exp. Ther.* **245**, 238–243 (1988).

96. Paul, D. & Pasternak, G. W. Differential blockade by naloxonazine of two mu opiate actions: analgesia and inhibition of gastrointestinal transit. *Eur. J. Pharmacol.* **149**, 403–404 (1988).

97. Takemori, A. E., Larson, D. L. & Portoghese, P. S. The irreversible narcotic antagonistic and reversible agonistic properties of the fumaramate methyl ester derivative of naltrexone. *Eur. J. Pharmacol.* **70**, 445–451 (1981).

98. Paul, D., Bodnar, R. J., Gistrak, M. A. & Pasternak, G. W. Different mu receptor subtypes mediate spinal and supraspinal analgesia in mice. *Eur. J. Pharmacol.* **168**, 307–314 (1989).

99. Pick, C. G., Nejat, R. J. & Pasternak, G. W. Independent expression of two pharmacologically distinct supraspinal mu analgesic systems in genetically different mouse strains. *J. Pharmacol. Exp. Ther.* **265**, 166–171 (1993).

100. Mogil, J. S. The genetic mediation of individual differences in sensitivity to pain and its inhibition. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 7744–7751 (1999).

101. Snyder, S. H. & Pasternak, G. W. Historical review: Opioid receptors. *Trends Pharmacol. Sci.* **24**, 198–205 (2003).

102. Chen, Y., Mestek, A., Liu, J., Hurley, J. A. & Yu, L. Molecular cloning and functional expression of a mu-opioid receptor from rat brain. *Mol. Pharmacol.* **44**, 8–12 (1993).

103. Wang, J. B. *et al.* Human mu opiate receptor. cDNA and genomic clones, pharmacologic characterization and chromosomal assignment. *FEBS Lett.* **338**, 217–222 (1994).

104. Kozak, C. A., Filie, J., Adamson, M. C., Chen, Y. & Yu, L. Murine chromosomal location of the mu and kappa opioid receptor genes. *Genomics* **21**, 659–661 (1994).

105. Belknap, J. K. *et al.* Localization to chromosome 10 of a locus influencing morphine analgesia in crosses derived from C57BL/6 and DBA/2 strains. *Life Sci.* **57**, PL117-124 (1995).

106. Giros, B., Pohl, M., Rochelle, J. M. & Seldin, M. F. Chromosomal localization of opioid peptide and receptor genes in the mouse. *Life Sci.* **56**, PL369-375 (1995).

107. Bare, L. A., Mansson, E. & Yang, D. Expression of two variants of the human mu opioid receptor mRNA in SK-N-SH cells and human brain. *FEBS Lett.* **354**, 213–216 (1994).

108. Zimprich, A., Simon, T. & Höllt, V. Cloning and expression of an isoform of the rat μ opioid receptor (rMOR1B) which differs in agonist induced desensitization from rMOR1. *FEBS Lett.* (1995) doi:10.1016/0014-5793(95)00028-8.

109. Pasternak, G. W., Childers, S. R. & Pan, Y.-X. Emerging Insights into Mu Opioid Pharmacology. in *Substance Use Disorders: From Etiology to Treatment* (eds. Nader, M. A. & Hurd, Y. L.) 89–125 (Springer International Publishing, 2020). doi:10.1007/164_2019_270.

110. Abbadie, C., Gultekin, S. H. & Pasternak, G. W. Immunohistochemical localization of the carboxy terminus of the novel mu opioid receptor splice variant MOR-1C within the human spinal cord. *Neuroreport* **11**, 1953–1957 (2000).

111. Abbadie, C., Pan, Y., Drake, C. T. & Pasternak, G. W. Comparative immunohistochemical distributions of carboxy terminus epitopes from the mu-opioid receptor splice variants MOR-1D, MOR-1 and MOR-1C in the mouse and rat CNS. *Neuroscience* **100**, 141–153 (2000).

112. Abbadie, C., Pan, Y. X. & Pasternak, G. W. Differential distribution in rat brain of mu opioid receptor carboxy terminal splice variants MOR-1C-like and MOR-1-like immunoreactivity: evidence for region-specific processing. *J. Comp. Neurol.* **419**, 244–256 (2000).

113. Abbadie, C., Pasternak, G. W. & Aicher, S. A. Presynaptic localization of the carboxy-terminus epitopes of the mu opioid receptor splice variants MOR-1C and MOR-1D in the superficial laminae of the rat spinal cord. *Neuroscience* **106**, 833–842 (2001).

114. Pan, L. *et al.* Identification and characterization of six new alternatively spliced variants of the human mu opioid receptor gene, Oprm. *Neuroscience* **133**, 209–220 (2005).

115. Koch, T. *et al.* Carboxyl-terminal splicing of the rat mu opioid receptor modulates agonist-mediated internalization and receptor resensitization. *J. Biol. Chem.* **273**, 13652–13657 (1998).

116. Rossi, G. C., Pan, Y. X., Brown, G. P. & Pasternak, G. W. Antisense mapping the MOR-1 opioid receptor: evidence for alternative splicing and a novel morphine-6 beta-glucuronide receptor. *FEBS Lett.* **369**, 192–196 (1995).

117. Kest, B., Hopkins, E., Palmese, C. A., Adler, M. & Mogil, J. S. Genetic variation in morphine analgesic tolerance: a survey of 11 inbred mouse strains. *Pharmacol. Biochem. Behav.* **73**, 821–828 (2002).

118. Kest, B. *et al.* Naloxone-precipitated withdrawal jumping in 11 inbred mouse strains: evidence for common genetic mechanisms in acute and chronic morphine physical dependence. *Neuroscience* **115**, 463–469 (2002).

119. Klein, G. *et al.* A survey of acute and chronic heroin dependence in ten inbred mouse strains: evidence of genetic correlation with morphine dependence. *Pharmacol. Biochem. Behav.* **90**, 447–452 (2008).

120. Xu, J. *et al.* Stabilization of the μ-opioid receptor by truncated single transmembrane splice variants through a chaperone-like action. *J. Biol. Chem.* **288**, 21211–21227 (2013).

121. Pan, Y.-X. *et al.* Involvement of exon 11-associated variants of the mu opioid receptor MOR-1 in heroin, but not morphine, actions. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 4917–4922 (2009).

122. Marrone, G. F. *et al.* Genetic dissociation of morphine analgesia from hyperalgesia in mice. *Psychopharmacology (Berl.)* **234**, 1891–1900 (2017).

123. Oladosu, F. A. *et al.* Mu Opioid Splice Variant MOR-1K Contributes to the Development of Opioid-Induced Hyperalgesia. *PloS One* **10**, e0135711 (2015).

124. Majumdar, S. *et al.* Generation of novel radiolabeled opiates through site-selective iodination. *Bioorg. Med. Chem. Lett.* **21**, 4001–4004 (2011).

125. Wieskopf, J. S. *et al.* Broad-spectrum analgesic efficacy of IBNtxA is mediated by exon 11-associated splice variants of the mu-opioid receptor gene. *Pain* **155**, 2063–2070 (2014).

126. Majumdar, S. *et al.* Truncated G protein-coupled mu opioid receptor MOR-1 splice variants are targets for highly potent opioid analgesics lacking side effects. *Proc. Natl. Acad. Sci.* **108**, 19778–19783 (2011).

127. Lu, Z. *et al.* Mediation of opioid analgesia by a truncated 6-transmembrane GPCR. *J. Clin. Invest.* **125**, 2626–2630 (2015).

128. Lu, Z. *et al.* Truncated μ-Opioid Receptors With 6 Transmembrane Domains Are Essential for Opioid Analgesia. *Anesth. Analg.* **126**, 1050–1057 (2018).

129. Islam, A. *et al.* Abuse Liability, Anti-Nociceptive, and Discriminative Stimulus Properties of IBNtxA. *ACS Pharmacol. Transl. Sci.* **3**, 907–920 (2020).

130. Le Merrer, J., Becker, J. A. J., Befort, K. & Kieffer, B. L. Reward processing by the opioid system in the brain. *Physiol. Rev.* **89**, 1379–1412 (2009).

131. Bodnar, R. J. Endogenous opiates and behavior: 2017. *Peptides* **124**, 170223 (2020).

132. Darcq, E. & Kieffer, B. L. Opioid receptors: drivers to addiction? *Nat. Rev. Neurosci.* **19**, 499–514 (2018).

133. Koob, G. F. *et al.* Addiction as a stress surfeit disorder. *Neuropharmacology* **76 Pt B**, 370–382 (2014).

134. Koob, G. F. & Le Moal, M. Addiction and the brain antireward system. *Annu. Rev. Psychol.* **59**, 29–53 (2008).

135. Minami, M. *et al.* Cloning and expression of a cDNA for the rat kappa-opioid receptor. *FEBS Lett.* **329**, 291–295 (1993).

136. Pfeiffer, A., Brantl, V., Herz, A. & Emrich, H. M. Psychotomimesis mediated by kappa opiate receptors. *Science* **233**, 774–776 (1986).

137. Ranganathan, M. *et al.* Dose-related behavioral, subjective, endocrine, and psychophysiological effects of the κ opioid agonist Salvinorin A in humans. *Biol. Psychiatry* **72**, 871–879 (2012).

138. Crowley, N. A. & Kash, T. L. Kappa opioid receptor signaling in the brain: Circuitry and implications for treatment. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **62**, 51–60 (2015).

139. Lalanne, L. *et al.* Kappa opioid receptor antagonism and chronic antidepressant treatment have beneficial activities on social interactions and grooming deficits during heroin abstinence. *Addict. Biol.* **22**, 1010–1021 (2017).

140. Bruchas, M. R., Land, B. B. & Chavkin, C. The dynorphin/kappa opioid system as a modulator of stress-induced and pro-addictive behaviors. *Brain Res.* **1314**, 44–55 (2010).

141. Chavkin, C. & Koob, G. F. Dynorphin, Dysphoria, and Dependence: the Stress of Addiction. *Neuropsychopharmacology* **41**, 373–374 (2016).

142. Kreibich, A. *et al.* Presynaptic Inhibition of Diverse Afferents to the Locus Ceruleus by κ-Opiate Receptors: A Novel Mechanism for Regulating the Central Norepinephrine System. *J. Neurosci.* **28**, 6516–6525 (2008).

143. Kieffer, B. L., Befort, K., Gaveriaux-Ruff, C. & Hirth, C. G. The delta-opioid receptor: isolation of a cDNA by expression cloning and pharmacological characterization. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 12048–12052 (1992).

144. Pradhan, A. A., Befort, K., Nozaki, C., Gavériaux-Ruff, C. & Kieffer, B. L. The delta opioid receptor: an evolving target for the treatment of brain disorders. *Trends Pharmacol. Sci.* **32**, 581–590 (2011).

145. Gendron, L., Cahill, C. M., von Zastrow, M., Schiller, P. W. & Pineyro, G. Molecular Pharmacology of δ-Opioid Receptors. *Pharmacol. Rev.* **68**, 631–700 (2016).

146. Klenowski, P., Morgan, M. & Bartlett, S. E. The role of δ-opioid receptors in learning and memory underlying the development of addiction. *Br. J. Pharmacol.* **172**, 297–310 (2015).

147. Pellissier, L. P., Pujol, C. N., Becker, J. a. J. & Le Merrer, J. Delta Opioid Receptors: Learning and Motivation. *Handb. Exp. Pharmacol.* **247**, 227–260 (2018).

148. Mansour, A., Hoversten, M. T., Taylor, L. P., Watson, S. J. & Akil, H. The cloned mu, delta and kappa receptors and their endogenous ligands: evidence for two opioid peptide recognition cores. *Brain Res.* **700**, 89–98 (1995).

149. Zadina, J. E., Hackler, L., Ge, L. J. & Kastin, A. J. A potent and selective endogenous agonist for the mu-opiate receptor. *Nature* **386**, 499–502 (1997).

150. Corbett, A. D., Paterson, S. J., & Kosterlitz, H. W. Opioids I. in *Handbook of experimental pharmacology* (ed. Herz, A.) vol. 104/I 645–679 (Springer, New York, 1993).

151. Weber, E. *et al.* Metorphamide: isolation, structure, and biologic activity of an amidated opioid octapeptide from bovine brain. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 7362–7366 (1983).

152. Goldberg, I. E. *et al.* Pharmacological characterization of endomorphin-1 and endomorphin-2 in mouse brain. *J. Pharmacol. Exp. Ther.* **286**, 1007–1013 (1998).

153. Zadina, J. E., Kastin, A. J., Ge, L. J. & Hackler, L. Mu, delta, and kappa opiate receptor binding of Tyr-MIF-1 and of Tyr-W-MIF-1, its active fragments, and two potent analogs. *Life Sci.* **55**, PL461-466 (1994).

154. Chefer, V. I., Bäckman, C. M., Gigante, E. D. & Shippenberg, T. S. Kappa opioid receptors on dopaminergic neurons are necessary for kappa-mediated place aversion. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **38**, 2623–2631 (2013).

155. Charbogne, P., Kieffer, B. L. & Befort, K. 15 years of genetic approaches in vivo for addiction research: opioid receptor and peptide gene knockout in mouse models of drug abuse. *Neuropharmacology* **76**, 10.1016/j.neuropharm.2013.08.028 (2014).

156. Lutz, P.-E. & Kieffer, B. L. Opioid receptors: distinct roles in mood disorders. *Trends Neurosci.* **36**, 195–206 (2013).

157. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).

158. Terenius, L. Stereospecific interaction between narcotic analgesics and a synaptic plasm a membrane fraction of rat cerebral cortex. *Acta Pharmacol. Toxicol. (Copenh.)* **32**, 317–320 (1973).

159. Simon, E. J., Hiller, J. M. & Edelman, I. Stereospecific binding of the potent narcotic analgesic (3H) Etorphine to rat-brain homogenate. *Proc. Natl. Acad. Sci. U. S. A.* **70**, 1947–1949 (1973).

160. Kuhar, M. J., Pert, C. B. & Snyder, S. H. Regional distribution of opiate receptor binding in monkey and human brain. *Nature* **245**, 447–450 (1973).

161. Hiller, J. M., Pearson, J. & Simon, E. J. Distribution of stereospecific binding of the potent narcotic analgesic etorphine in the human brain: predominance in the limbic system. *Res. Commun. Chem. Pathol. Pharmacol.* **6**, 1052–1062 (1973).

162. Erbs, E. *et al.* A mu-delta opioid receptor brain atlas reveals neuronal co-occurrence in subcortical networks. *Brain Struct. Funct.* **220**, 677–702 (2015).

163. Scherrer, G. *et al.* Knockin mice expressing fluorescent delta-opioid receptors uncover G protein-coupled receptor dynamics in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 9691–9696 (2006).

164. Morgan, M. M., Reid, R. A., Stormann, T. M. & Lautermilch, N. J. Opioid selective antinociception following microinjection into the periaqueductal gray of the rat. *J. Pain* **15**, 1102–1109 (2014).

165. Hellman, K. M., Mendelson, S. J., Mendez-Duarte, M. A., Russell, J. L. & Mason, P. Opioid microinjection into raphe magnus modulates cardiorespiratory function in mice and rats. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **297**, R1400-1408 (2009).

166. Zheng, C. & Baum, B. J. Evaluation of promoters for use in tissue-specific gene delivery. *Methods Mol. Biol. Clifton NJ* **434**, 205–219 (2008).

167. O'Neil, K. T., Hoess, R. H. & DeGrado, W. F. Design of DNA-binding peptides based on the leucine zipper motif. *Science* **249**, 774–778 (1990).

168. Sauer, B. & Henderson, N. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 5166–5170 (1988).

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

170. Vardy, E. *et al.* A New DREADD Facilitates the Multiplexed Chemogenetic Interrogation of Behavior. *Neuron* **86**, 936–946 (2015).

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

172. Marchant, N. J. *et al.* Behavioral and Physiological Effects of a Novel Kappa-Opioid Receptor-Based DREADD in Rats. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **41**, 402–409 (2016).

173. Ben-Shaanan, T. L. *et al.* Activation of the reward system boosts innate and adaptive immunity. *Nat. Med.* **22**, 940–944 (2016).

174. Ferguson, S. M. & Neumaier, J. F. Using DREADDs to investigate addiction behaviors. *Curr. Opin. Behav. Sci.* **2**, 69–72 (2015).

175. Saloman, J. L. *et al.* Gi-DREADD Expression in Peripheral Nerves Produces Ligand-Dependent Analgesia, as well as Ligand-Independent Functional Changes in Sensory Neurons. *J. Neurosci. Off. J. Soc. Neurosci.* **36**, 10769–10781 (2016).

176. Amorim, M. R. *et al.* The Effect of DREADD Activation of Leptin Receptor Positive Neurons in the Nucleus of the Solitary Tract on Sleep Disordered Breathing. *Int. J. Mol. Sci.* **22**, 6742 (2021).

177. Manvich, D. F. *et al.* The DREADD agonist clozapine N-oxide (CNO) is reverse-metabolized to clozapine and produces clozapine-like interoceptive stimulus effects in rats and mice. *Sci. Rep.* **8**, 3840 (2018).

178. Koehl, A. *et al.* Structure of the μ-opioid receptor-Gi protein complex. *Nature* **558**, 547–552 (2018).

179. Wang, W. *et al.* A light- and calcium-gated transcription factor for imaging and manipulating activated neurons. *Nat. Biotechnol.* **35**, 864–871 (2017).

180. Kim, M. W. *et al.* Time-gated detection of protein-protein interactions with transcriptional readout. *eLife* **6**, e30233 (2017).

181. Huang, W. *et al.* Structural insights into μ-opioid receptor activation. *Nature* **524**, 315–321 (2015).

182. Terskiy, A. *et al.* Search of the human proteome for endomorphin-1 and endomorphin-2 precursor proteins. *Life Sci.* **81**, 1593–1601 (2007).

183. Strickland, D., Moffat, K. & Sosnick, T. R. Light-activated DNA binding in a designed allosteric protein. *Proc. Natl. Acad. Sci.* **105**, 10709–10714 (2008).

184. Pham, E., Mills, E. & Truong, K. A synthetic photoactivated protein to generate local or global Ca(2+) signals. *Chem. Biol.* **18**, 880–890 (2011).

185. Strickland, D. *et al.* TULIPs: Tunable, light-controlled interacting protein tags for cell biology. *Nat. Methods* **9**, 379–384 (2012).

186. Lungu, O. I. *et al.* Designing photoswitchable peptides using the AsLOV2 domain. *Chem. Biol.* **19**, 507–517 (2012).

187. Renicke, C., Schuster, D., Usherenko, S., Essen, L.-O. & Taxis, C. A LOV2 domain-based optogenetic tool to control protein degradation and cellular function. *Chem. Biol.* **20**, 619–626 (2013).

188. Yi, J. J., Wang, H., Vilela, M., Danuser, G. & Hahn, K. M. Manipulation of endogenous kinase activity in living cells using photoswitchable inhibitory peptides. *ACS Synth. Biol.* **3**, 788–795 (2014).

189. Niopek, D. *et al.* Engineering light-inducible nuclear localization signals for precise spatiotemporal control of protein dynamics in living cells. *Nat. Commun.* **5**, 4404 (2014).

190. Yumerefendi, H. *et al.* Control of Protein Activity and Cell Fate Specification via Light-Mediated Nuclear Translocation. *PloS One* **10**, e0128443 (2015).

191. Wong, S., Mosabbir, A. A. & Truong, K. An Engineered Split Intein for Photoactivated Protein Trans-Splicing. *PloS One* **10**, e0135965 (2015).

192. Jones, D. C., Mistry, I. N. & Tavassoli, A. Post-translational control of protein function with light using a LOV-intein fusion protein. *Mol. Biosyst.* **12**, 1388–1393 (2016).

193. Mart, R. J., Meah, D. & Allemann, R. K. Photocontrolled Exposure of Pro-apoptotic Peptide Sequences in LOV Proteins Modulates Bcl-2 Family Interactions. *Chembiochem Eur. J. Chem. Biol.* **17**, 698–701 (2016).

194. Yumerefendi, H. *et al.* Light-induced nuclear export reveals rapid dynamics of epigenetic modifications. *Nat. Chem. Biol.* **12**, 399–401 (2016).

195. Niopek, D., Wehler, P., Roensch, J., Eils, R. & Di Ventura, B. Optogenetic control of nuclear protein export. *Nat. Commun.* **7**, 10624 (2016).

196. Sun, W. *et al.* Light-induced protein degradation in human-derived cells. *Biochem. Biophys. Res. Commun.* **487**, 241–246 (2017).

197. Melero-Fernandez de Mera, R. M. *et al.* A simple optogenetic MAPK inhibitor design reveals resonance between transcription-regulating circuitry and temporally-encoded inputs. *Nat. Commun.* **8**, 15017 (2017).

198. French, A. R., Sosnick, T. R. & Rock, R. S. Investigations of human myosin VI targeting using optogenetically controlled cargo loading. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1607–E1616 (2017).

199. Deisseroth, K. *et al.* Next-generation optical technologies for illuminating genetically targeted brain circuits. *J. Neurosci. Off. J. Soc. Neurosci.* **26**, 10380–10386 (2006).

200. Miesenböck, G. The optogenetic catechism. *Science* **326**, 395–399 (2009).

201. Zemelman, B. V., Lee, G. A., Ng, M. & Miesenböck, G. Selective photostimulation of genetically chARGed neurons. *Neuron* **33**, 15–22 (2002).

202. Zemelman, B. V., Nesnas, N., Lee, G. A. & Miesenbock, G. Photochemical gating of heterologous ion channels: remote control over genetically designated populations of neurons. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 1352–1357 (2003).

203. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nat. Neurosci.* **8**, 1263–1268 (2005).

204. Zhang, F. *et al.* Multimodal fast optical interrogation of neural circuitry. *Nature* **446**, 633–639 (2007).

205. Han, X. & Boyden, E. S. Multiple-color optical activation, silencing, and desynchronization of neural activity, with single-spike temporal resolution. *PloS One* **2**, e299 (2007).

206. Li, X. *et al.* Fast noninvasive activation and inhibition of neural and network activity by vertebrate rhodopsin and green algae channelrhodopsin. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 17816–17821 (2005).

207. Kim, J.-M. *et al.* Light-driven activation of beta 2-adrenergic receptor signaling by a chimeric rhodopsin containing the beta 2-adrenergic receptor cytoplasmic loops. *Biochemistry* **44**, 2284–2292 (2005).

208. Airan, R. D., Thompson, K. R., Fenno, L. E., Bernstein, H. & Deisseroth, K. Temporally precise in vivo control of intracellular signalling. *Nature* **458**, 1025–1029 (2009).

209. Oh, E., Maejima, T., Liu, C., Deneris, E. & Herlitze, S. Substitution of 5-HT1A receptor signaling by a light-activated G protein-coupled receptor. *J. Biol. Chem.* **285**, 30825–30836 (2010).

210. Wang, X., Chen, X. & Yang, Y. Spatiotemporal control of gene expression by a light-switchable transgene system. *Nat. Methods* **9**, 266–269 (2012).

211. Nihongaki, Y., Suzuki, H., Kawano, F. & Sato, M. Genetically engineered photoinducible homodimerization system with improved dimer-forming efficiency. *ACS Chem. Biol.* **9**, 617–621 (2014).

212. Chen, D., Gibson, E. S. & Kennedy, M. J. A light-triggered protein secretion system. *J. Cell Biol.* **201**, 631–640 (2013).

213. Motta-Mena, L. B. *et al.* An optogenetic gene expression system with rapid activation and deactivation kinetics. *Nat. Chem. Biol.* **10**, 196–202 (2014).

214. Jayaraman, P. *et al.* Blue light-mediated transcriptional activation and repression of gene expression in bacteria. *Nucleic Acids Res.* **44**, 6994–7005 (2016).

215. Zhao, E. M. *et al.* Optogenetic regulation of engineered cellular metabolism for microbial chemical production. *Nature* **555**, 683–687 (2018).

216. Grusch, M. *et al.* Spatio-temporally precise activation of engineered receptor tyrosine kinases by light. *EMBO J.* **33**, 1713–1726 (2014).

217. Shimizu-Sato, S., Huq, E., Tepperman, J. M. & Quail, P. H. A light-switchable gene promoter system. *Nat. Biotechnol.* **20**, 1041–1044 (2002).

218. Leung, D. W., Otomo, C., Chory, J. & Rosen, M. K. Genetically encoded photoswitching of actin assembly through the Cdc42-WASP-Arp2/3 complex pathway. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 12797–12802 (2008).

219. Tyszkiewicz, A. B. & Muir, T. W. Activation of protein splicing with light in yeast. *Nat. Methods* **5**, 303–305 (2008).

220. Levskaya, A., Weiner, O. D., Lim, W. A. & Voigt, C. A. Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* **461**, 997–1001 (2009).

221. Kawano, F., Suzuki, H., Furuya, A. & Sato, M. Engineered pairs of distinct photoswitches for optogenetic control of cellular proteins. *Nat. Commun.* **6**, 6256 (2015).

222. Yazawa, M., Sadaghiani, A. M., Hsueh, B. & Dolmetsch, R. E. Induction of protein-protein interactions in live cells using light. *Nat. Biotechnol.* **27**, 941–945 (2009).

223. Crefcoeur, R. P., Yin, R., Ulm, R. & Halazonetis, T. D. Ultraviolet-B-mediated induction of protein–protein interactions in mammalian cells. *Nat. Commun.* **4**, 1779 (2013).

224. Kennedy, M. J. *et al.* Rapid blue-light-mediated induction of protein interactions in living cells. *Nat. Methods* **7**, 973–975 (2010).

225. Wang, H. *et al.* LOVTRAP: an optogenetic system for photoinduced protein dissociation. *Nat. Methods* **13**, 755–758 (2016).

226. Stone, O. J. *et al.* Optogenetic control of cofilin and αTAT in living cells using Z-lock. *Nat. Chem. Biol.* **15**, 1183–1190 (2019).

227. Zhou, X. X., Chung, H. K., Lam, A. J. & Lin, M. Z. Optical control of protein activity by fluorescent protein domains. *Science* **338**, 810–814 (2012).

228. Richter, F. *et al.* Engineering of temperature- and light-switchable Cas9 variants. *Nucleic Acids Res.* **44**, 10003–10014 (2016).

229. Crosson, S. & Moffat, K. Structure of a flavin-binding plant photoreceptor domain: Insights into light-mediated signal transduction. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 2995–3000 (2001).

230. Peter, E., Dick, B. & Baeurle, S. A. Mechanism of signal transduction of the LOV2-Jα photosensor from Avena sativa. *Nat. Commun.* **1**, 122 (2010).

231. Konold, P. E. *et al.* Unfolding of the C-Terminal Jα Helix in the LOV2 Photoreceptor Domain Observed by Time-Resolved Vibrational Spectroscopy. *J. Phys. Chem. Lett.* **7**, 3472–3476 (2016).

232. Schierling, B. & Pingoud, A. Controlling the DNA cleavage activity of light-inducible chimeric endonucleases by bidirectional photoactivation. *Bioconjug. Chem.* **23**, 1105–1109 (2012).

233. Mills, E., Chen, X., Pham, E., Wong, S. & Truong, K. Engineering a photoactivated caspase-7 for rapid induction of apoptosis. *ACS Synth. Biol.* **1**, 75–82 (2012).

234. Baarlink, C., Wang, H. & Grosse, R. Nuclear actin network assembly by formins regulates the SRF coactivator MAL. *Science* **340**, 864–867 (2013).

235. Wu, Y. I. *et al.* A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature* **461**, 104–108 (2009).

236. Cosentino, C. *et al.* Engineering of a light-gated potassium channel. *Science* **348**, 707–710 (2015).

237. Bonger, K. M., Rakhit, R., Payumo, A. Y., Chen, J. K. & Wandless, T. J. General method for regulating protein stability with light. *ACS Chem. Biol.* **9**, 111–115 (2014).

238. Spiltoir, J. I., Strickland, D., Glotzer, M. & Tucker, C. L. Optical Control of Peroxisomal Trafficking. *ACS Synth. Biol.* **5**, 554–560 (2016).

239. Stratton, M. M. & Loh, S. N. Converting a protein into a switch for biosensing and functional regulation. *Protein Sci. Publ. Protein Soc.* **20**, 19–29 (2011).

240. Christie, J. M., Gawthorne, J., Young, G., Fraser, N. J. & Roe, A. J. LOV to BLUF: flavoprotein contributions to the optogenetic toolkit. *Mol. Plant* **5**, 533–544 (2012).

241. Cunningham, B. A., Hemperly, J. J., Hopp, T. P. & Edelman, G. M. Favin versus concanavalin A: Circularly permuted amino acid sequences. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 3218–3222 (1979).

242. Bowles, D. J. & Pappin, D. J. Traffic and assembly of concanavalin A. *Trends Biochem. Sci.* **13**, 60–64 (1988).

243. Goldenberg, D. P. & Creighton, T. E. Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **165**, 407–413 (1983).

244. Bliven, S. & Prlić, A. Circular Permutation in Proteins. *PLOS Comput. Biol.* **8**, e1002445 (2012).

245. Yu, Y. & Lutz, S. Circular permutation: a different way to engineer enzyme structure and function. *Trends Biotechnol.* **29**, 18–25 (2011).

246. Halavaty, A. S. & Moffat, K. N- and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from Avena sativa. *Biochemistry* **46**, 14001–14009 (2007).

247. Kawano, F., Aono, Y., Suzuki, H. & Sato, M. Fluorescence imaging-based high-throughput screening of fast- and slow-cycling LOV proteins. *PloS One* **8**, e82693 (2013).

248. Guntas, G. *et al.* Engineering an improved light-induced dimer (iLID) for controlling the localization and activity of signaling proteins. *Proc. Natl. Acad. Sci.* **112**, 112–117 (2015).

249. Sanchez, M. I. & Ting, A. Y. Directed evolution improves the catalytic efficiency of TEV protease. *Nat. Methods* **17**, 167–174 (2020).

250. Kapust, R. B., Tözsér, J., Copeland, T. D. & Waugh, D. S. The P1' specificity of tobacco etch virus protease. *Biochem. Biophys. Res. Commun.* **294**, 949–955 (2002).

251. Lam, S. S. *et al.* Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat. Methods* **12**, 51–54 (2015).

252. Boder, E. T. & Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.* **15**, 553–557 (1997).

253. Chao, G. *et al.* Isolating and engineering human antibodies using yeast surface display. *Nat. Protoc.* **1**, 755–768 (2006).

254. Hoffmann, M. D., Bubeck, F., Eils, R. & Niopek, D. Controlling Cells with Light and LOV. *Adv. Biosyst.* **2**, 1800098 (2018).

255. Strickland, D. *et al.* Rationally improving LOV domain-based photoswitches. *Nat. Methods* **7**, 623–626 (2010).

256. Strader, C. D. *et al.* Allele-specific activation of genetically engineered receptors. *J. Biol. Chem.* **266**, 5–8 (1991).

257. Coward, P. *et al.* Controlling signaling with a specifically designed Gi-coupled receptor. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 352–357 (1998).

258. Redfern, C. H. *et al.* Conditional expression and signaling of a specifically designed Gi-coupled receptor in transgenic mice. *Nat. Biotechnol.* **17**, 165–169 (1999).

259. Zhao, G. Q. *et al.* The receptors for mammalian sweet and umami taste. *Cell* **115**, 255–266 (2003).

260. Mueller, K. L. *et al.* The receptors and coding logic for bitter taste. *Nature* **434**, 225–229 (2005).

261. Magnus, C. J. *et al.* Ultrapotent chemogenetics for research and potential clinical applications. *Science* **364**, eaav5282 (2019).

262. Harding, M. W., Galat, A., Uehling, D. E. & Schreiber, S. L. A receptor for the immunosuppressant FK506 is a cis-trans peptidyl-prolyl isomerase. *Nature* **341**, 758–760 (1989).

263. Siekierka, J. J., Hung, S. H., Poe, M., Lin, C. S. & Sigal, N. H. A cytosolic binding protein for the immunosuppressant FK506 has peptidyl-prolyl isomerase activity but is distinct from cyclophilin. *Nature* **341**, 755–757 (1989).

264. Spencer, D. M., Wandless, T. J., Schreiber, S. L. & Crabtree, G. R. Controlling signal transduction with synthetic ligands. *Science* **262**, 1019–1024 (1993).

265. Zheng, X. F., Florentino, D., Chen, J., Crabtree, G. R. & Schreiber, S. L. TOR kinase domains are required for two distinct functions, only one of which is inhibited by rapamycin. *Cell* **82**, 121–130 (1995).

266. Stanton, B. Z., Chory, E. J. & Crabtree, G. R. Chemically induced proximity in biology and medicine. *Science* **359**, eaao5902 (2018).

267. Cunningham-Bryant, D. *et al.* A Chemically Disrupted Proximity System for Controlling Dynamic Cellular Processes. *J. Am. Chem. Soc.* **141**, 3352–3355 (2019).

268. Karginov, A. V., Ding, F., Kota, P., Dokholyan, N. V. & Hahn, K. M. Engineered allosteric activation of kinases in living cells. *Nat. Biotechnol.* **28**, 743–747 (2010).

269. Dagliyan, O. *et al.* Rational design of a ligand-controlled protein conformational switch. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 6800–6804 (2013).

270. Farrants, H. *et al.* Chemogenetic Control of Nanobodies. *Nat. Methods* **17**, 279–282 (2020).

271. Rose, J. C. *et al.* A computationally engineered RAS rheostat reveals RAS-ERK signaling dynamics. *Nat. Chem. Biol.* **13**, 119–126 (2017).

272. Bonger, K. M., Chen, L., Liu, C. W. & Wandless, T. J. Small-molecule displacement of a cryptic degron causes conditional protein degradation. *Nat. Chem. Biol.* **7**, 531–537 (2011).

273. Negi, S., Pandey, S., Srinivasan, S. M., Mohammed, A. & Guda, C. LocSigDB: a database of protein localization signals. *Database J. Biol. Databases Curation* **2015**, bav003 (2015).

274. Ruggiero, E., Alonso-de Castro, S., Habtemariam, A. & Salassa, L. Upconverting nanoparticles for the near infrared photoactivation of transition metal complexes: new opportunities and challenges in medicinal inorganic photochemistry. *Dalton Trans. Camb. Engl. 2003* **45**, 13012–13020 (2016).

275. Broichhagen, J. & Trauner, D. The in vivo chemistry of photoswitched tethered ligands. *Curr. Opin. Chem. Biol.* **21**, 121–127 (2014).

276. Jendryka, M. *et al.* Pharmacokinetic and pharmacodynamic actions of clozapine-N-oxide, clozapine, and compound 21 in DREADD-based chemogenetics in mice. *Sci. Rep.* **9**, 4522 (2019).

277. Auffenberg, E. *et al.* Remote and reversible inhibition of neurons and circuits by small molecule induced potassium channel stabilization. *Sci. Rep.* **6**, 19293 (2016).

278. Clackson, T. *et al.* Redesigning an FKBP-ligand interface to generate chemical dimerizers with novel specificity. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 10437–10442 (1998).

279. Banaszynski, L. A., Sellmyer, M. A., Contag, C. H., Wandless, T. J. & Thorne, S. H. Chemical control of protein stability and function in living mice. *Nat. Med.* **14**, 1123–1127 (2008).

280. Choi, J., Chen, J., Schreiber, S. L. & Clardy, J. Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science* **273**, 239–242 (1996).

281. Gomez, J. L. *et al.* Chemogenetics revealed: DREADD occupancy and activation via converted clozapine. *Science* **357**, 503–507 (2017).

282. Castellano, F., Montcourrier, P. & Chavrier, P. Membrane recruitment of Rac1 triggers phagocytosis. *J. Cell Sci.* **113 (Pt 17)**, 2955–2961 (2000).

283. Nguyen, T., Pappireddi, N. & Wühr, M. Proteomics of nucleocytoplasmic partitioning. *Curr. Opin. Chem. Biol.* **48**, 55–63 (2019).

284. Ptashne, M. & Gann, A. Transcriptional activation by recruitment. *Nature* **386**, 569–577 (1997).

285. Song, H. K. & Eck, M. J. Structural basis of degradation signal recognition by SspB, a specificity-enhancing factor for the ClpXP proteolytic machine. *Mol. Cell* **12**, 75–86 (2003).

286. Hall, M. P. *et al.* Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem. Biol.* **7**, 1848–1857 (2012).

287. Loh, J. M. S. & Proft, T. Comparison of firefly luciferase and NanoLuc luciferase for biophotonic labeling of group A Streptococcus. *Biotechnol. Lett.* **36**, 829–834 (2014).

288. Rozbeh, R. & Forchhammer, K. Split NanoLuc technology allows quantitation of interactions between PII protein and its receptors with unprecedented sensitivity and reveals transient interactions. *Sci. Rep.* **11**, 12535 (2021).

289. Stoeber, M. *et al.* A Genetically Encoded Biosensor Reveals Location Bias of Opioid Drug Action. *Neuron* **98**, 963-976.e5 (2018).

290. Fan, F. *et al.* Novel genetically encoded biosensors using firefly luciferase. *ACS Chem. Biol.* **3**, 346–351 (2008).

291. He, L. *et al.* Circularly permuted LOV2 as a modular photoswitch for optogenetic engineering. *Nat. Chem. Biol.* **17**, 915–923 (2021).

292. Harper, S. M., Christie, J. M. & Gardner, K. H. Disruption of the LOV-Jalpha helix interaction activates phototropin kinase activity. *Biochemistry* **43**, 16184–16192 (2004).

293. Valentino, R. J. & Volkow, N. D. Untangling the complexity of opioid receptor function. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **43**, 2514–2520 (2018).

294. Yang, D. *et al.* G protein-coupled receptors: structure- and function-based drug discovery. *Signal Transduct. Target. Ther.* **6**, 7 (2021).

295. Wu, F., Song, G., de Graaf, C. & Stevens, R. C. Structure and Function of Peptide-Binding G Protein-Coupled Receptors. *J. Mol. Biol.* **429**, 2726–2745 (2017).

296. Scott, J. D. & Pawson, T. Cell signaling in space and time: where proteins come together and when they're apart. *Science* **326**, 1220–1224 (2009).

297. Rajakulendran, S. & Hanna, M. G. The Role of Calcium Channels in Epilepsy. *Cold Spring Harb. Perspect. Med.* **6**, a022723 (2016).

298. Kaur, S., Maslov, L. N., Singh, N. & Jaggi, A. S. Dual role of T-type calcium channels in anxiety-related behavior. *J. Basic Clin. Physiol. Pharmacol.* **31**, /j/jbcpp.2020.31.issue-3/jbcpp-2019-0067/jbcpp-2019-0067.xml (2019).

299. U.S. Department of Health and Human Services, Food and Drug Administration. Technical Considerations for Medical Devices with Physiologic Closed-Loop Control Technology: Draft Guidance for Industry and Food and Drug Administration Staff. https://www.fda.gov/media/154994/download (2021).

300. Goutaudier, R., Coizet, V., Carcenac, C. & Carnicella, S. DREADDs: The Power of the Lock, the Weakness of the Key. Favoring the Pursuit of Specific Conditions Rather than Specific Ligands. *eNeuro* **6**, ENEURO.0171-19.2019 (2019).