DESIGN, IMPROVEMENT, AND BENCHMARKING OF A LIGHT-INDUCED HETERODIMER PAIR: THE AWAKENING OF iLID

Ryan Alexander Ellis Hallett

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry and Biophysics.

Chapel Hill
2015

Approved by:

Brian Kuhlman

Henrik Dohlman

Klaus Hahn

Kevin Slep

Ashutosh Tripathy

ABSTRACT

Ryan Alexander Ellis Hallett: Design, Improvement, and Benchmarking of a Light-Induced Heterodimer: The Awakening of iLID
(Under the direction of Brian Kuhlman)

Cellular optogenetic tools are engineered protein photoreceptors that allow researchers to probe intricate protein-protein interaction networks with the flip of a light switch. These tools can be turned on with high spatial and temporal resolution to change the activity or localization of a protein inside a cell. In order for a tool to be widely useful it should be generalizable for multiple applications, orthogonal to the system it is used in, and have low levels of activity in the inactive state. Light-inducible heterodimerization is one of the most general optogenetic approaches. Each half of the pair can be fused with any intracellular protein or localization sequence, imparting light induced control over a wide variety of signaling pathways. However, utility of existing light inducible dimers is still limited due to poor dynamic range between active and inactive states or unknown mechanism of action, which can impede analysis.

It is the aim of this thesis to design a general cellular optogenetic tool with large dynamic range, usable in eukaryotic systems, and clear mechanism of activity. In order to create a tool fitting these criteria, we have engineered a light inducible heterodimer pair from the SsrA peptide – SspB protein interaction, using the blue light sensitive photoreceptor, LOV2 from *Avena sativa*. Irradiation of AsLOV2 with blue light induces a conformational change in its C-terminal Jα helix. Our initial incorporation of the SsrA peptide into the Jα helix resulted in a modest change in affinity for SspB with light. Using a protocol of computational library design, phage display screening, and high-throughput binding assays

we were able to engineer an improved light-inducible dimer system, iLID, that exhibits over 50-fold increase in affinity for its partner upon irradiation with blue light. The iLID system comes with two partners, SspB nano and SspB micro, which enables researchers to induce interactions in the nanomolar or micromolar ranges. We have further showed that both iLID pairs can be used to reversibly co-localize proteins of interest in mammalian cells and control small GTPase signaling.

Despite their proposed modularity, successful control of in cell activity depends on compatibility between characteristics of the chosen heterodimer pair and its application. To examine the *in vivo* functional significance of *in vitro* characteristics for light-inducible dimer pairs, we measured *in vitro* affinities and kinetics, light induced gene transcription in *S. cerevisiae*, and lamellipodia protrusion in mammalian cell culture. The results demonstrate a correlation between affinity, kinetics, and dynamic range with cellular activity and highlight the need for thorough benchmarking. This work has yielded valuable insight on how to select optogenetic tools appropriate for specific applications and generated two powerful optogenetic heterodimer pairs, iLID nano and micro, available for use in cell biology.

Seeing as all of us are just pieces of all that was before our time, thank you to all people, places, and things found here on Earth, now and in all times past. In your own little tiny way, you've contributed to the success of this work.

ACKNOWLEDGEMENTS

I would like to sincerely thank the huge team of people that went into making this work a success. Oana Lungu brought the LOV domain to the Kuhlman lab and was an excellent mentor in my days as a lowly roton. Seth Zimmerman is an incredible cell biologist and has been able to make a large number of excellent *in vivo* experiments work that I would not have been able to tackle alone. Gurkan Guntas painstakingly screened a library that would have taken me many lifetimes to explicitly test and because of that, we now have iLID. Hayretin Yumerefendi has been an incredible resource for knowledge both in the photoreceptor field as well as every other molecular biology technique I could think to ask about. Tishan Williams taught me most of what I know about crystallography and has opened my eyes to the wonderful world of boardgames; for that I couldn't be more appreciative. Andrew Leaver-Fay has patiently answered every one of my computer questions over the past 5 years and I owe my ability to operate Rosetta to him and Tim Jacobs. Although it was a long time ago, I wouldn't be where I'm at today without the incredible guidance of Carrie Purbeck Trunzo, she taught me to be bright where ever I am and I don't plan on stopping anytime soon. And finally, Brian Kuhlman has set the bar for what a leader and mentor should be. Working with Brian has taught me the incredible impact a leader can have on the people he works with. I will strive to make whatever space I find myself in as open and encouraging as his lab here at UNC. Also, thanks to him, I will always look at a nearby mountain and think to myself, "eh, I bet we can make it up there".

Outside of lab I have been blessed with the most incredible community of people that a person could ever wish for. Adam Roberts, Jonny Leano, Will Monteith, Raquel Burger-

Calderon, and Tasha Jurgensen are some of the best people you could hope to live life with and made my time here at UNC about more than just work and studying. Thanks to the 2013 Champion Southern Rail Softball team for giving me my all-time top athletic achievement. Thank you to every single last Greenleafer (member of my church, Greenleaf Vineyard) for the years of support, incredible discussions and growth, countless prayers, and showing me what the hands and feet of Jesus look like here around town. I would also like to thank my family, Mom, Dad, Nancy, and TC. My parents have always supported my inquisitive nature (even when it drove them nuts) and because of that I've found my calling in a field where asking the next question is the only constant. And most importantly, a huge thank you to Amy Hallett, my unbelievably loving wife who has not only made me many meals to-go while writing this, but supports and encourages me every day in a way I never thought even possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ChR2	Channelrhodopsin-2
SPARK	Synthetic photoisomerizable azobenzene-regulated K+ channel
LOV	Light Oxygen Voltage
PAS	Per-Arnt-Sim
FMN	Flavin mononucleotide
AsLOV2	*Avena sativa* phot1 LOV2 domain
UV	Ultraviolet
FAD	Flavin adenine mononucleotide
MTHF	5,10-methyltetrahydrofolate
PHR	Photolyase homology region
CCT	Cryptochrome carboxyl-terminus
RMS	Root mean square
ATP	Adenosine triphosphate
FADH•	Neutral semiquinone state of FAD
GAF	Cyclic GMP phosphodiesterase, Adenylate cyclase, FhlA
PHY	Phytochrome specific
PCB	Phycocyanobilin
PΦB	Phytochromobilin
BV	Biliverdin IX alpha
Pfr	Far-red absorbing state of PCB
Pr	Red absorbing state of PCB
BLUF	Blue-Light Sensors Using Flavin Adenine Dinucleotide
AtLOV2	*Arabidopsis thaliana* LOV2 domain
GTPase	Guanosine triphosphate phosphotase
GEF	Guanine Nucleotide Exchange Factor

PA	Photoactivatable
DHFR	Dihydrofolate reductase
PYP	Photoactive yellow protein
GTP	Guanosine triphosphate
GDP	Guanosine diphosphate
TAMRA	Carboxytetramethylrhodamine
iLID	Improved light inducible dimer
oLID	Original light inducible dimer
TULIPs	Tunable, light-controlled interacting protein tags for cell biology
GI	Gigantea
PMSF	Phenylmethanesulfonylfluoride
CPRG	Red-β-D-galactopyranoside
ONPG	o-nitrophenyl-β-D-galactopyranoside
LOVpep+	LOVpep with T406-7A, I532A mutations

Chapter 1

Introduction

This dissertation is primarily focused on advancing the optogenetic field through creation of new light activated tools, the benchmarking of existing tools, and application of such tools to novel cellular uses. This introduction will cover the origins of optogenetics, the underpinnings of blue-light sensitive protein domains, an overview of uses, and evaluation of current limitations.

1.1 The Origin of Optogenetics

1.1.1 Initial Neurobiology Applications

Optogenetics broadly describes the fusion of engineering and biology that uses the input of light to control a particular cellular response. Francis Crick described the need for such tools initially in 1979 and again in 1999 as a way to dissect the function of individual components of the brain[1,2]. It was not until 2002 when this became a reality for the neuroscience field when neurons expressing *Drosophilia* arrestin-2, rhodopsin, and the α subunit of the corresponding G-protein (together, the chARGe system) were shown to produce a light induced action potential[3]. This work was based on importing the invertebrate rhodopsin signal cascade, which is naturally light sensitive, to mammalian neurons, which are typically light insensitive. This advancement, gave researchers the power to control a subset of neurons with light, simply by introducing a few exogenous genes. This technique was expanded upon a few years later with the development of synthetic photoisomerizable azobenzene-regulated K^+ (SPARK) channels[4]. SPARK channels were engineered through

incorporation of a photoactive moiety into the *Drosophilia* K^+ Shaker channel. It was known that Shaker voltage gated ion channels are blocked by the binding of quaternary ammonium ions. Using available structural information, an azobenzene moiety with a cysteine reactive group on one end and an ammonium ion on the other, was incorporated into an extracellular loop of the Shaker channel approximately 15-18 Å away. Azobenzene is a photoreactive molecule, two phenyl groups connected by an N=N double bond. Irradiation with light ~360 nm causes the double bond to isomerize from the *trans* to *cis* conformation[5]. The *trans* state can be reverted to *cis* through exposure to ~500 nm light or will decay naturally through thermal conversion[5]. Once incorporated, the azobenzene acts as a photoactive ion gatekeeper, with *trans*-azobenzene positioning the ammonium ion to block the ion channel. Photoisomerization of the moiety releases inhibition and allows for ion transfer. Using this strategy, researchers demonstrated the ability to induce action potentials with short-wavelength light and silence them with long-wavelength light in rat hippocampal neurons expressing the chemically modified SHAKER constructs[4]. The downside to this method was the need for chemical modification post-expression in addition to incorporation of the SHAKER gene. While these pioneering techniques gave a new level of spatial control to neurobiologists, there was still the issue of timescale. Both techniques operated on the second to minute timescale, about a thousand fold slower than necessary to induce single action potentials. A third innovation to the field less than a year later reduced the timescale of control to the millisecond level[6], and solidified optogenetics as more than just a passing trend. This advancement was made possible through use of an algal photoreceptor, Channelrhodopsin-2 (ChR2). ChR2 had recently been discovered as a naturally occurring light-regulated cation channel from *Chlamydomonas reinhardtii*[7], and using lentiviral incorporation into mammalian neurons, it was shown that ChR2 could be induced in as little as 50 μs after a flash with blue light[6]. This enabled the control of single depolarization events with commonly available microscopy equipment, as the same lasers used for GFP

excitation could be used to excite ChR2. These three initial optogenetic milestones demonstrated simple and viable methods for control of mammalian neurons with light and paved the way for the decade of advancement that would follow.

1.1.2 Promise of Tools in Cell Biology

While initially limited to neurobiology, the promise of optogenetic cellular control extended into almost all cell biology fields. The aforementioned initial experiments set the precedent for what characteristics of optogenetic tools are needed to be useful in studying biological questions. In order to be useful, optogenetic tools need to be genetically encoded, respond quickly to external stimuli, respond reversibly, and be designed such that they are simple to use for a variety of applications. While these initial optogenetic tools were only really useful in the neurobiology field, they highlighted the degree of spatial and temporal control that light afforded researchers. As the field expanded to create generalizable tools that could be used by any cell biologist, the design strategies seen in those first optogenetics tools surfaced again; structure based design of novel photoreceptors, direct harnessing of naturally occurring photoreceptors, and incorporation of photoactive moieties into functional protein domains. A researcher's view of a given cellular pathway is limited by the resolution and capabilities of the technology they use to probe their system. To meet these needs, novel tools with a variety of spectral properties, kinetic rates, and molecular characteristics have been engineered from a wide assortment of natural photoactive macromolecules.

1.2 Light Reactive Protein Domains

The vast majority of currently used photoactivatable protein tools are powered by a naturally occurring protein photoreceptor. Existing photoreceptors are chosen for their ease of genetic incorporation into a host organism and wide array of domains with multiple

functionalities to choose from. Protein photoreceptors vary in the wavelength of activation, timescale of reversion, response upon activation, chromophore usage, and compatibility within host organism. This diversity has allowed for many different cellular uses and highlights the importance of photoreceptor selection for a given application. Surprisingly, use and reengineering of a photoreceptor to become a molecular tool is not contingent upon detailed understanding of its mechanism of action; as such newly discovered and poorly understood protein domains have been added to the optogenetic toolkit. However, for many applications this is a drawback and should contribute to the selection criteria for a given photoreceptor. It is noteworthy to mention that there are a wide variety of photoactive chemical moieties[8] (like azobenzene for one), which have been used to control cellular processes like protein expression[9,10] and cell cycle progression[11], however they are outside the scope of this work.

1.2.1 LOV Domains

Light-oxygen-voltage (LOV) domains are a broad family of proteins involved in relaying extracellular signals of light to intracellular effector domains[12]. LOV domains belong to the Per-ARNT-Sim (PAS) superfamily of proteins that are involved in a wide variety of signaling pathways in archaea, bacteria, and eukaryotic organisms[13]. The first LOV domains were discovered as the blue light photoreceptor domains of phototropin 1 and phototropin 2, two large multi-domain proteins responsible for chloroplast relocation and phototropism in higher plants[14]. Common, but not exclusive to all LOV domains is the binding of a single cofactor, flavin mononucleotide (FMN)[12]. The absorption spectrum of FMN gives LOV domains their ability to perceive blue-light, characterized by maxima at 360 nm and 450 nm[15]. Upon excitation with blue light, the FMN cofactor is excited to a transient triplet state within 30 ns, which decays to a metastable intermediate with a half-life of 4 μs[15]. For some LOV domains, the C(4a) carbon atom in the FMN cofactor forms a covalent bond

with a nearby cysteine residue in the LOV domain while in this metastable intermediate state[15,16]. After formation, the covalent adduct and FMN cofactor thermally reverts back to their ground states. The kinetics of this process vary widely among LOV domains, from seconds to minutes to hours and even days[15,17,18]. In their natural context, the formation of this covalent adduct propagates a conformational change throughout the protein, and modulates the activity of a response domain at the carboxyl-terminus of the LOV domain. Known response elements regulated by LOV domains are histidine kinases, transcription factors, phosphodiesterases, and proteins that regulate circadian rhythm[12]. In the optogenetics engineering field, a handful of LOV domains have been plucked out of their native habitat to be used as the basis for creating novel molecular switches; of most important to this work is the LOV2 domain from *Avena Sativa* (AsLOV2). AsLOV2 regulates its carboxyl-terminal kinase domain through light-dependent displacement of the Jα helix, a helix connecting the FMN-containing PAS fold with the kinase (Figure 1.1)[19,20]. It is currently unknown exactly how the signal is transduced from the FMN-cysteine adduct to the Jα helix, however key residues involved in this process have been identified[21,22]. As its role in the photoactivation process is paramount, many groups have studied the dynamics of the Jα helix extensively. In one study, NMR was used to measure the free energy change of the transition from dark to lit state. They measured the equilibrium of AsLOV2 to be 1.6% Jα-unbound / 98.4% Jα-unbound in the dark, and 8% Jα-unbound / 92% Jα-bound under blue light; an approximately 3.8 kcal mol^{-1} change in free energy between lit and dark states[23]. It was also found that the amino terminal helix of AsLOV2, the A'α helix, unfolds upon irradiation with blue light[22,24]. This was first proposed based on small structural rearrangements of the A'α helix between AsLOV2 structures solved from crystals bathed in blue light and those kept in darkness[22]. Later, the mechanistic relationship between the two helices was more thoroughly probed through extensive mutational analysis of the AsLOV2, including deletions of both the Jα and A'α helices[24]. Based on helicity change as measured

by circular dichroism, change in tryptophan fluorescence, and correlations of residue specific CPMG values all measured in both the light and dark, it was concluded that the A'α helix undocked first, which then allowed the Jα helix to undock[24]. This work was followed up shortly thereafter with an even more extensive list of mutations and various biophysical characterizations, indicating a wide variety of photocycle half-lives, from ~23 minutes to 2 seconds, as well as the promise of some mutations that increase the signal change between the lit and dark states[25]. This breadth of structural information and biophysical characterization concerning AsLOV2 and other LOV domains, have rendered them some of the most popular photoreceptor domains in the optogenetic repertoire.

1.2.2 Cryptochromes

Cryptochromes are another class of blue light sensing photoreceptors and the first to be studied at the protein level[26]. Best studied in *Arabidopsis thaliana*, three cryptochromes have been discovered: CRY1, CRY2, and CRY3[27–29]. In recent years, CRY2 has emerged as one of the most useful photoreceptors in optogenetics. Similar to the LOV domains of the phototropins, cryptochromes absorb blue light, with a maximum absorption at 450 nm and smaller absorption peaks in the near ultraviolet (UV) range[30]. Their absorption spectrum is largely imparted by their cofactor, flavin adenine dinucleotide (FAD), which similar to FMN, absorbs maximally at 450 nm[27]. Distinct from LOV domains, cryptochromes have an additional cofactor, 5,10-methyltetrahydrofolate (MTHF) used in harvesting additional light in the UV range[30]. Cryptochromes generally consist of two domains, an amino terminal photolyase homology region (PHR) and a specialized cryptochrome carboxyl-terminus (CCT), which varies between cryptochrome variants (and is missing from CRY3)[31]. Despite their structural similarity to photolyase domains, the cryptochromes in *Arabidopsis* do not have DNA repair activity. The cryptochrome PHR domain of CRY1 and CRY3 has been resolved by x-ray crystallography (Figure 1.2)[32,33], however there is no

structural information for CRY2 PHR or either CRY1 or CRY2 CCT domain currently. The overall topology of CRY1 PHR and CRY3 PHR are very similar (RMS of 1.9 Å, PDB codes: 1u3d & 2j4d), both consist of an amino terminal α/β domain, a short linker, and an α domain. However, one interesting difference is the presence of a MTHF binding site in CRY3, but not in CRY1 (Figure 1.2B). Additionally, an adenosine triphosphate (ATP) binding site is seen in the CRY1 structure (Figure 1.2A), which corroborates previous *in vitro* data[32]. The cryptochrome PHR domain binds all cofactors noncovalently, which is the most frequent explanation for why either ATP or MTHF are not found in purified protein samples[34]. The FAD binding site is made up of three conserved tryptophan residues within the α domain, but their specific role in photoactivation is somewhat unclear[35]. Mutation of each of these residues in CRY1 results in decreased ability to respond to blue light, however, in CRY2 similar mutations result in constitutive activation[36]. Recently, an electron transfer pathway not utilizing the canonical tryptophan triad has been proposed[37]. In any case, when irradiated with blue light, the FAD cofactor is reduced to form a neutral semiquinone radical (FADH•)[38–40]. This species is considered by some to be the active signaling state of AtCRY1 and AtCRY2. Post excitation and activation, the cryptochromes undergo a conformational change, leading to the displacement of the CCT domain[41]. It has also been shown that CRY2 PHR domain binds CIB1 after exposure to blue light[42]. CRY1 and CRY2 both become phosphorylated after exposure to blue light[43,44], adding another layer of complexity to this mechanism. More recently, CRY2 has been shown to oligomerize in a light dependent fashion, but through currently unknown means[45]. Despite a general uncertainty on the overarching mechanism, site of conformational change, and stoichiometry, both light induced CRY2/CIB1 binding and CRY2 oligomerization have been widely used as optogenetic modifiers[45,46]. Additionally, a variant of CRY2 optimized for oligomerization applications, CRY2olig, has been engineered[47].

1.2.3 Phytochrome Domains

Phytochromes have been found to absorb light across almost the entire spectra of visible light, but most commonly sense red and far-red light[48–51]. They are found in plants, fungi, bacteria, and cyanobacteria. The phytochrome core photoreceptor is made up of three domains: a PAS domain, a GAF (cyclic GMP phosphodiesterase/adenylate cyclase/FhlA) domain, and a PHY (phytochrome specific) domain (Figure 1.3)[49]. Phytochromes form a covalent bond with their chromophore between a conserved cysteine residue in either the PAS or GAF domain and the C3 carbon of the A-ring in the chromophore[49]. The bound phytochrome chromophore varies between organisms, but it is usually a linear tetrapyrrole (bilin) derivative[49]. In plants and cyanobacteria, that derivative is phycocyanobilin (PCB) or phytochromobilin (PΦB); in bacteria or fungi the chromophore is biliverdin Ixα (BV)[49,52]. Irradiation with red light (660 nm - 680 nm) causes an isomerization of the chromophore, most frequently about a double bond connecting the C-ring and D-ring (Figure 1.3C)[49]. This isomerization induces a conformational change in the bound protein and is passed along to an effector domain. The isomerized chromophore is also light sensitive, however the spectrum is red-shifted, now responding to far-red light. Irradiation of the far-red absorbing (Pfr) state with far-red light (740 nm - 760 nm) causes another isomerization of the D-ring back to the red light absorbing (Pr) state[49]. This Pfr to Pr reversion will occur thermally if not activated with far-red light, albeit much slower, on the timescale of hours[53]. Phytochrome use in optogenetics comes with strong advantages and disadvantages. The commonly used photoreceptor, PhyB, from *Arabidopsis thaliana* uses a cofactor, PCB, which is not naturally expressed in other organisms. As a result, cofactor must be exogenously added to cell culture media or made through genetically encoding a set of enzymes needed to synthesize PCB directly into the organism[54]. This method has been demonstrated in *E. coli*[55] and more recently mammalian cells[56], however is not optimal. An advantage of phytochromes is that they fold and remain stable in the absence of their

chromophore, rendering them light insensitive until chromophore addition. This has been sold as an advantage since steps are normally taken to prevent early exposure of a photoreceptor to light before an experiment. The other major advantage of the PhyB photoreceptor is independent control over both the Pr and Pfr states. This gives users direct control of reversion timescale and allows them to be synced in either state with a flash of light. It has been shown that the Pfr state of PhyB interacts with PIF3, but not the Pr state[57]. In the field of optogenetics, the red light dependent PhyB interaction with PIF3 and PIF6 has been utilized for light-dependent heterodimerization[53,57,58]. However, many specific applications involving phytochromes from bacteria (bacteriophytochromes) have been reported utilizing either their native effector domains or engineering domains from other organisms leading to red light induced activity, including transcription[59] and adenylate cyclase activity[60]. An in-depth look at light-induced dimerizing applications will be reviewed in section 1.3.

1.2.4 BLUF proteins and Xanthopsins

While LOV domains, cryptochromes, and phytochromes are the most prevalent photoreceptors used in cell biology, it is important to mention the remaining classes of photoreceptors: blue-light sensors using flavin adenine dinucleotide (BLUF) and xanthopsins. BLUF domains are small protein domains (~100 - 250 amino acids) similar in size to LOV domains (Figure 1.4). The BLUF chromophore, FAD, becomes excited upon absorbing a photon of blue light, which rearranges a hydrogen bond network involving a conserved tyrosine and glutamine[61,62]. This rearrangement is propagated through a nearby β-scaffold, changing interactions with a C-terminal α-helix, Ccap, which changes interactions with effector domains or interaction partners[63]. Interestingly, it has been shown that AppA, PixD, and YcgF all form protein-protein interactions in the dark, but these complexes dissociate in blue light[63]. Xanthopsins are also small protein domains absorbing light in the

blue and near-uv region of visible light[64]. Xanthopsins use 4-hydroxy-cinnamic acid (p-Coumaric acid) as their chromophore[64]. Absorption of a photon causes a cis- to trans-isomerization of the chromophore, which after a number of intermediate states, becomes deprotonated[65]. It is the negative charge formed by this deprotonation that is thought to cause a partial unfolding event in the amino terminal cap that induces signaling[65]. The most widely known xanthopsin, photoactive yellow protein (PYP) from *Halorhodospira halophila*, has been used to control light dependent DNA binding[66,67].

1.3 Successful Optogenetic Dimerization Applications

While traditional optogenetics aims to modulate the potential across a cell membrane, cellular optogenetics aims to modulate protein-protein interaction networks in order to better understand the cellular signaling environment. While many proteins have been engineered to be light sensitive through incorporation of a photoreceptor[68–71], this section will focus on more generic applications. Specifically, the use of dimerization to control a variety of processes like transcription, degradation, and motility will be covered. Additionally, the CRY2/CIB1 and TULIP heterodimerization pairs will be reviewed in detail due to their characterization in chapter 4.

1.3.1 Dimerization Strategies

Cellular homeostasis depends on many protein – protein interactions occurring in specific locations all around the cell simultaneously. In order to better study these interactions, a few strategies using light induced dimerization have been devised to control protein activity including: direct co-localization of two proteins, inducing changes in protein localization, sequestration or activation through oligomerization, and reconstitution of a split protein (Figure 1.5).

Direct co-localization of two proteins is the broadest category, useful for inducing protein – protein interactions that have a direct effect. One great example of direct co-localization was seen in the development of photoactivatable protein degradation tags[72]. In this instance a target protein was fused to a LOV2 domain from *Arabidopsis thaliana* with an ornithine decarboxylase degron incorporated into its carboxyl-terminus. Light induced changes from AtLOV2 allowed the degron to interact with the proteasome and caused degradation of the target protein[72]. Using this technique they showed almost complete ubiquitin-independent degradation of red fluorescent protein (RFP) in yeast after 4 hours of blue light exposure. Follow up studies to this work later showed this mechanism was tunable, allowing for both shorter and longer lived degron variants[73]. Gene transcription has also been controlled with this strategy through light induced localization of a DNA binding protein (zinc finger protein) with an activation domain (VP16)[74–76].

Light induced localization change has most frequently been used as a way to regulate cell motility, whether directly through guanosine triphosphatase (GTPase) localization or indirectly through guanine nucleotide exchange factor (GEF) localization. In the direct application, one half of a heterodimer pair is plasma membrane localized and the other is fused to a constitutively active variant of the GTPase without a membrane localization sequence[53,77]. As GTPases are natively targeted to the plasma membrane through attachment of a prenyl group where they interact with effectors, removal of this makes its activity conditional on induced localization (for the constitutively active GTPase variant)[78]. In the indirect application, a GEF is localized to the plasma membrane with light to activate endogenous GTPase. This strategy has been utilized in yeast to control polarized growth[79]. In this example, the GEF Cdc24 was fused to one half of a heterodimer pair and the other half fused to the transmembrane protein Mid2[79]. Light induced recruitment of the GEF to the membrane led to the formation of a protrusion within 45° of the site of irradiation in the majority of the cells tested[79].

Sequestration through hetero- and homo-oligomerization has only been achieved with a single photoreceptor, CRY2[45,47,80]. In this approach, a single protein is fused to the protein that oligomerizes in response to light. Oligomerization occurs in such a way to sequester or activate the fused proteins' function, creating a generalizable way to control protein function. This set-up is advantageous since it only requires the addition of a single component, instead of the two required from a heterodimer pair. Homo-oligomerization has been used to activate the Wnt/β-catenin pathway, Rac1 activity and RhoA activity through direct fusion to CRY2[45]. Light-induced hetero-oligomerization has also been demonstrated by combining the CRY2 oligomerization with CRY2 light-depended CIB1 binding. Hetero-oligomerization, with CRY2, CIB1, and an additional multimeric protein, has been used to inactivate Tiam1, Vav2, Rac1, RhoG, and Cdc42[80]. While clearly powerful, this method is less precise as the mechanism of CRY2 oligomerization is unknown. A more precise and still generalizable method of light-induced protein activity uses the fluorescent protein Dronpa and its native light inducible monomerization. In this method, fusion of Dronpa to the N- and C- terminus of a protein causes the dimer or tetramer oligomer to occlude access to the protein. Irradiation with light at 490 nm induces monomerization of Dronpa and allows for the protein of interest to interact natively. In this manner, the GEF intersectin and a protease were caged, only active after irradiation with light[81].

Split protein reconstitution has been used as a way to detect protein – protein interactions for many years. Light-induced heterodimerization uses the opposite principle, using light to create a protein - protein interface and reconstitute the split protein activity. By far the most common example of this in optogenetics is the reconstitution of the Gal4 DNA binding domain and Gal4 activation domains to induce transcription. This has almost become the litmus test for light-induced heterodimer pairs, as almost every widely used pair has demonstrated this[82–86]. Light induced reconstitution of the Cre-Lox system using a split Cre protein has been demonstrated using the CRY2/CIB1 pair[86].

1.3.2 CRY2-CIB1

The blue light dependent interaction between CRY2 from *Arabidopsis thaliana* and CIB1 was initially described in 2008[42] and first utilized as an cellular optogenetics tool in 2010[86]. In this initial proof-of-principle and description, the minimal domain elements required for functionality were identified using a yeast two-hybrid approach. The PHR domain from CRY2 (residues 1-498) and the N-terminal fragment (residues 1-170) were enough to confer light induced transcription[86]. This assay showed that CRY2 binding to CIB1 and CIB1N resulted in substantially less background activity in the dark (and less activity in the light) whereas CRY2 PHR binding to CIB1 and CIB1N resulted in much higher dark state activity (and more activity in light). In all experiments, the CRY2 half of the switch was fused to the Gal DNA binding domain and CIB1 fused to the activation domain. Light induced transcription was also shown to be dose-dependent by delivering a series of 10 second pulses of blue light, 8 minutes apart, and showing a proportional increase of protein product to number of blue light pulses[86]. In mammalian cells, the CRY2-CIB1 interaction was demonstrated in two ways. First, plasma membrane bound CIBN1-GFP (green fluorescent protein) and mCherry-CRY2 were transfected and irradiated with blue light. Within 10 seconds of illumination, mCherry-CRY2 had fully accumulated at the membrane and after ~12 minutes of darkness it had returned fully to the cytoplasm[86]. This interaction was repeated in the same cell to determine if the interaction was reversible and repeated full activation was observed. In a final display of protein control, a Cre recombinase split into two halves was reconstituted through CRY2 and CIB1N binding, leading to loxP recombination in 16.4% of cells after 24 hours of blue light[86]. Since its initial debut, this pair has been used by a number of labs in a variety of applications. The pair is functional in zebrafish, where one group showed both light dependent activation and suppression of transcription[83]. Activation was achieved as previously shown, through colocalization of DNA binding and transcription-activating domains while suppression was achieved through light

dependent CRY2 binding and blocking of CIB1N's native transcription activation[42,83]. In a step toward being more biologically adaptable, CRY2/CIB1 was used to modulate transcription of endogenous genes through colocalization of transcription activator-like effectors (TALEs) with VP64 (a 4 copy fusion of the transcription activator, VP16)[75]. TALEs can be engineered to bind any DNA sequence of choice due to a one to one relationship between a repeat in the protein and interaction with a single base pair of DNA[87]. As a result, the CRY2-TALE fusion can be localized to any locus in the genome and CIB1-VP64 recruitment induces transcription there. This system, called LITEs (light inducible transcriptional effectors), was shown to produce more than 15-fold more mRNA after 8 hours of exposure to blue light than when left in the dark[75]. LITEs were also used to site-specifically modify histones and modulate expression through localization of methyltransferases and deacetylases[75]. A similar method of light induced DNA localization is now possible by using a deactivated Cas9 (dCas9) to target a specific stretch of DNA[88]. The CRY2-CIB1 system has been used to control cellular cytoskeletal function as well. CIB1-Lifeact fusions allowed for light induced CRY2 localization to F-actin. Light induced CRY2-(constitutively active)cofilin localized to F-actin and was shown to increase filopodia formation as well as lamellipodia formation and movement[89].

1.3.3 TULIPs: Tunable Light-Controlled Interacting Protein Tags

TULIPs were developed by engineering the AsLOV2 domain to selectively bind an engineered PDZ domain (ePDZb) in response to light[79]. This was achieved by incorporating a PDZ binding peptide into the Jα helix of AsLOV2, called LOVpep. Unfolding of the helix in response to blue light frees the peptide and allows binding to ePDZb. They showed that many mutations that had previously tuned the Jα helix lit/dark equilibrium also changed dynamics of the LOVpep equilibrium, and thus were able to tune their extent to which the peptide was caged[79]. These mutations were: T406-7A, which previously shown to increase

caging (lower dark state binding), I532A, previously shown to increase Jα helicity and increase caging, and V529N, previously shown to decrease caging[90]. Therefore, in order from expected weakest dark state binding to tightest, the mutation pairs tested were: T406-7A+I532A, T406-7A, LOVpep alone, and V529N. This system also came with two other protein-binding partners, the starting PDZ domain that binds weaker than ePDZb, and the affinity-matured ePDZb1, which has a tighter affinity[91]. This modularity gave LOVpep three regimes of affinity it could switch over, each with a number of mutations in AsLOV2 to "fine tune" for a specific application. This wide range of switching was demonstrated by monitoring colocalization of fluorescent protein fusions with each LOVpep and ePDZ variant in the presence and absence of light in yeast. The dynamic range varied from low levels of localization in both the lit and dark states (LOVpepT406-7A,I532A:PDZ) to almost full localization in both states (LOVpep V529N:ePDZb1)[79]. These findings directly correlated relative affinity and caging (albeit absolute affinities values unknown) to *in vivo* function. Applying these tools to functional activity, they were able to light-dependently control MAPK activation and cell polarity[79]. Since their public debut, TULIPs have only been used in a handful of applications. In a truly novel application, researchers fused LOVpep to PEX3, a peroxisome localization sequence and ePDZb1 to either kinesin or dynein. They found that blue light induced localization of kinesin led to peroxisome movement toward the outside of the cell and localization of dynein caused peroxisome movement toward the center of the cell[92]. This was as expected as kinesin directs transport toward the plus end of microtubules and dynein directs towards the minus end[93]. Light-induced motor protein localization to endosome and mitochondria also changed dynamics of each substructure[92].

1.4 Drawbacks and Limitations of the Field

Cellular optogenetics affords a level of spatial and temporal control over cell signaling that has not been possible previously. However, there are many limitations in the

field that next-generation tools and approaches should be mindful of. As described, most existing tools respond to blue light, which does not penetrate more than 1 mm into tissue[94]. Red light penetrates farther, ~3-4 mm[94], but this keeps cellular optogenetic tools relegated to cell culture and epidermal animal applications, unless light sources are implanted. In order to overcome this, tools that respond to longer wavelengths of light need to be developed.

Each tool discussed has a set of limitations that preclude it from one application and ideal for another. This highlights the needs for a wide variety of tools and methods. So far one of the largest drawbacks is a residual activity in the target off state. As most heterodimer and oligomer tools rely on changes in affinity between states, overexpression of one component or both can drastically reduce dynamic range. This has been noticed specifically in a few cases[76,82], but undoubtedly would occur in any case where expression exceeded dark state affinity. Therefore, tools with increased dynamic ranges and methods to increase dynamic ranges are needed. Additionally, cellular optogenetic tools need to have precise characterization to better aid those who use the tools in choosing the correct tools for a given application. Explicit declaration of limitations and methods of use should be provided upon release of any tool. In order to fully compare tools as they are developed, a set of benchmark experiments could be developed to highlight the advantages and disadvantages upon release. Finally, most uses of heterodimerization thus far have been proof of principle and validation using known mechanisms of protein signaling. If the field is to survive and leave an impact, it must use the newfound spatial and temporal control push the current boundary of cellular knowledge.

In the work that follows we describe the design (chapter 2), improvement (chapter 3), benchmarking (chapter 4), and novel application (chapter 5) of a new light-inducible heterodimer tool.

1.5 Figures

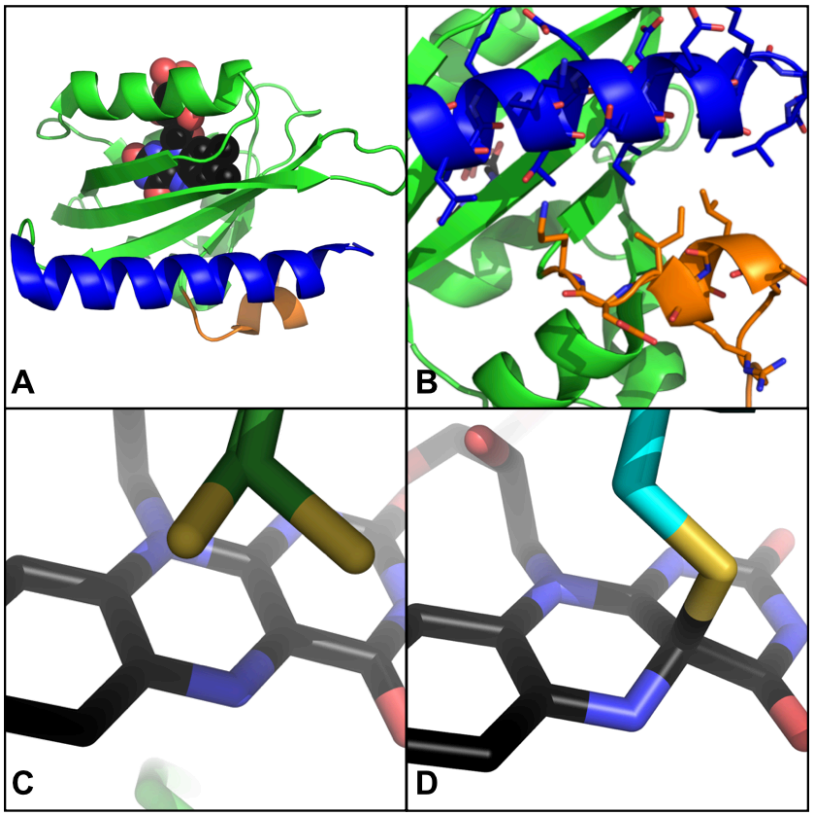


Figure 1.1 Structure of AsLOV2. A) AsLOV2 structure, PAS domain in green, Jα helix in blue, A'α helix in orange, FMN cofactor shown in black spheres B) Interactions between A'α and Jα helix C) C450 in dark state D) C450 in lit state, bound to FMN (PDB IDs A-C: 2v0u; D: 2v0w)

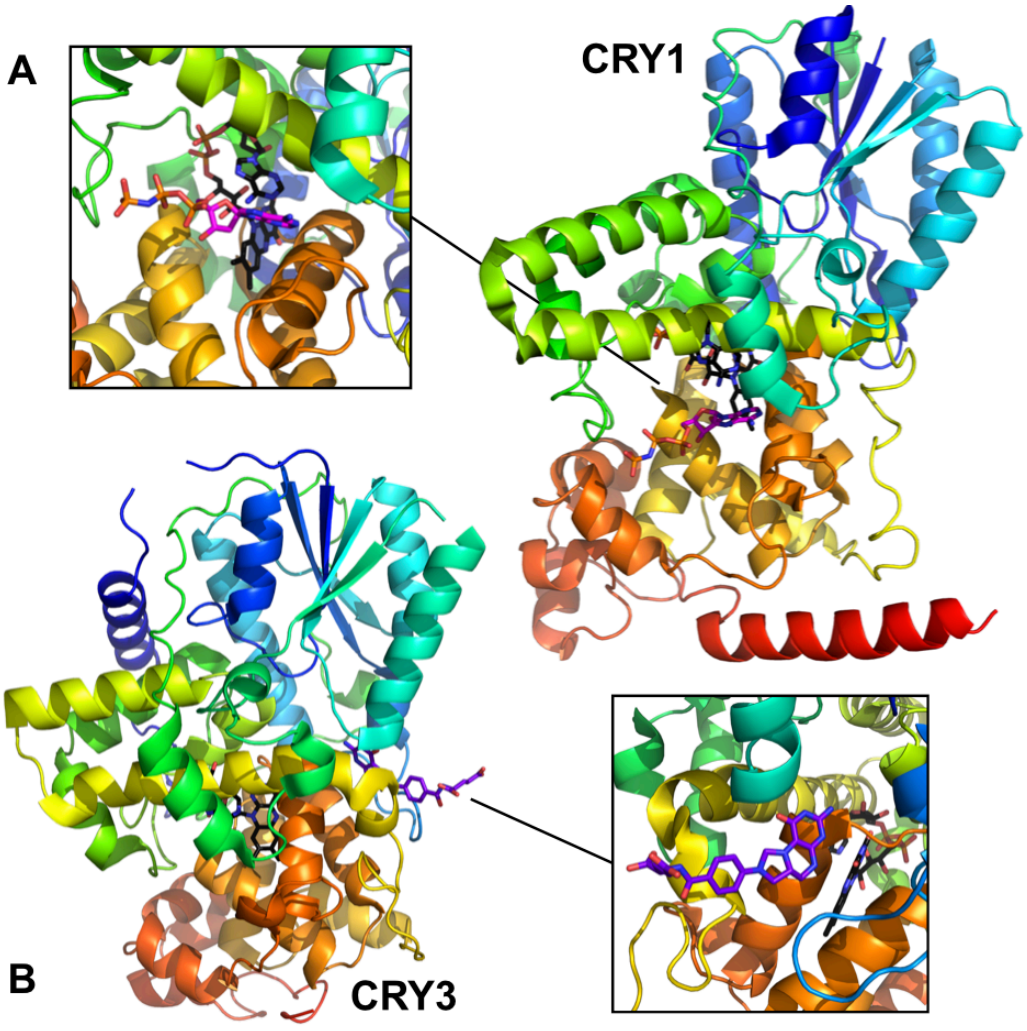


Figure 1.2 Structure of CRY1 and CRY3 PHR domains from Arabidopsis thaliana. A) CRY1 PHR domain with FAD (black) and ATP (magenta) binding site in inset. B) CRY3 PHR domain with MTHF (purple) cofactor binding site in inset. (PDB IDs A: 1u3d; B: 2ijg)

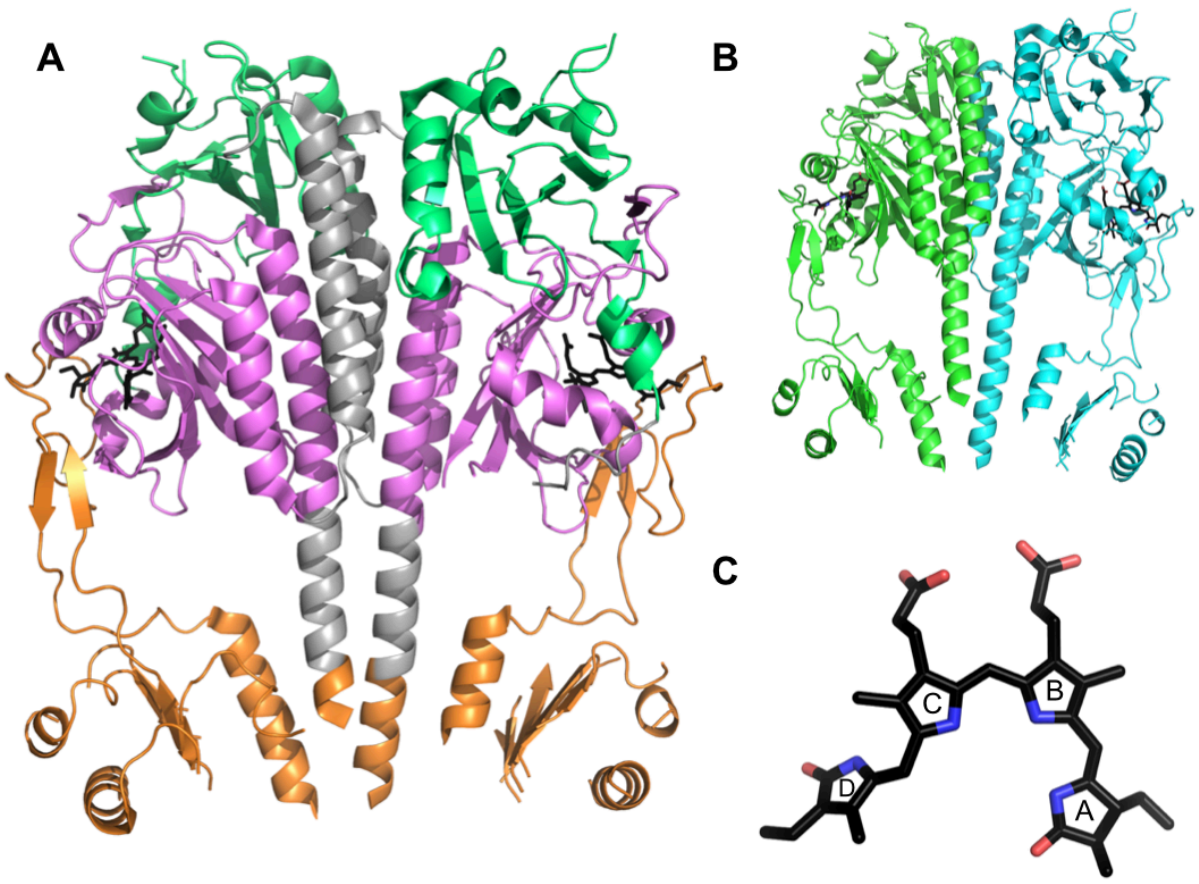


Figure 1.3 Structure of Arabidopsis thaliana PhyB. A) Dimer PhyB structure showing PAS (green), GAF (pink), and PHY (orange) domains. PCB cofactor is shown in black. B) PhyB showing individual monomer subunits. C) PCB cofactor with pyrrole ring naming shown. (PDB: 4our)

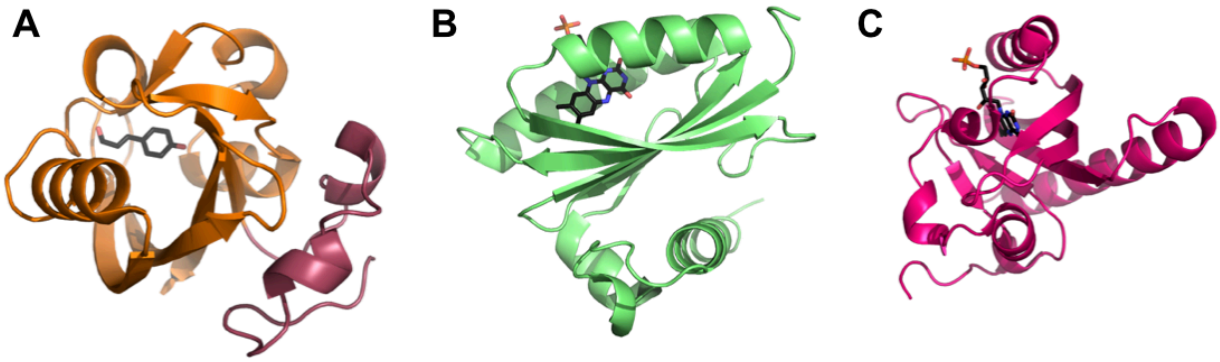


Figure 1.4 Comparison of xanopsin, BLUF, and LOV domain. PAS family members contain similar binding pocket for light sensitive chromophores. A) The xanopsin PYP from *Halorhodospira halophila* (PDB: 2phy), B) the BLUF BlrB from *Rhodobacter sphaeroide*s (PDB: 2byc), and C) AsLOV2 (PDB: 2v0u)

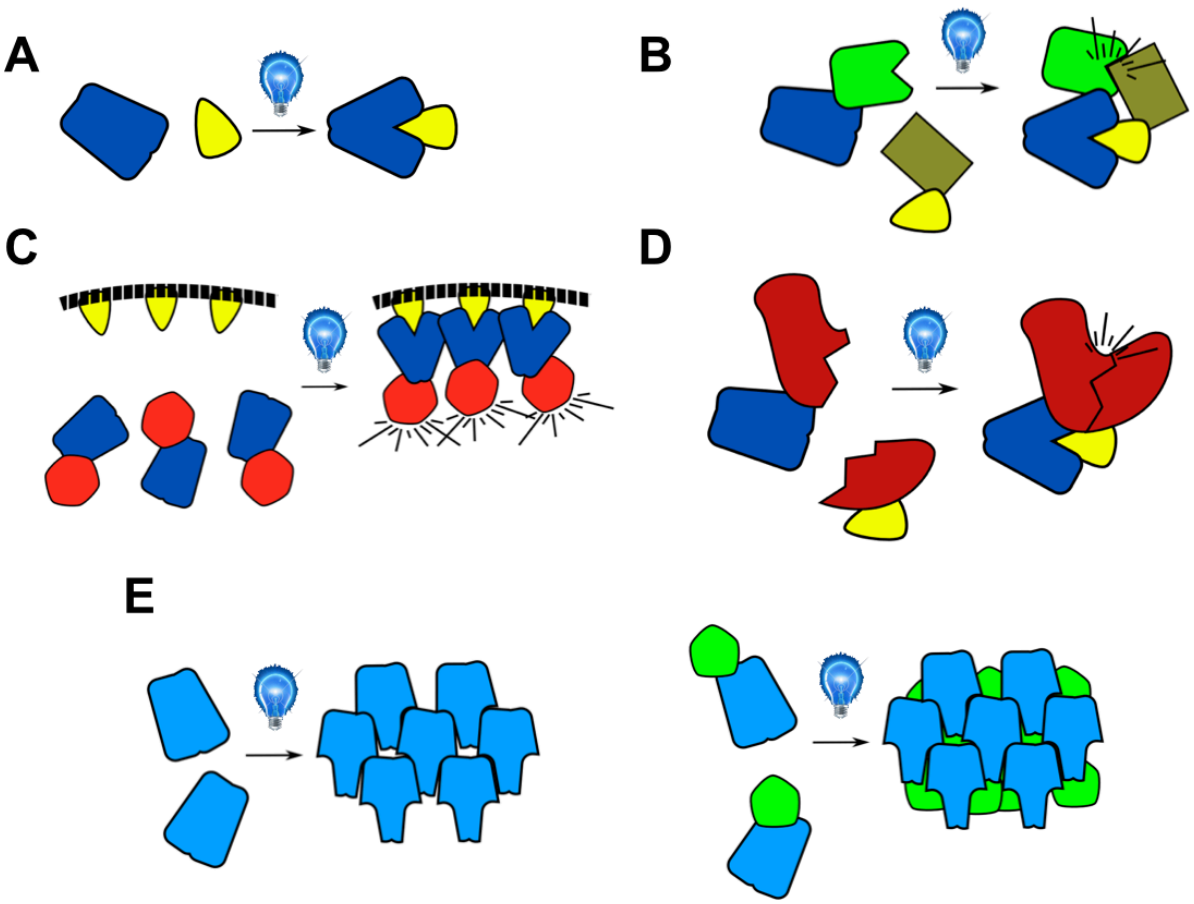


Figure 1.5 Schematic of light induced heterodimerization and oligomerization. A) Light inducible heterodimer pair (Used in B-D) B) Light induced co-localization of green and brown proteins. C) Light induced localization of red protein to membrane D) Light induced reconstitution of split red protein. E) Light induced oligomerization (left) used to sequester the green protein (right).

Chapter 2

Designing photoswitchable peptides using the AsLOV2 domain[1]

This work was published in the journal Chemistry & Biology, Volume 19, Issue 4, p507-517 on 20 April 2012. Oana Lungu developed the LOV-ipaA switch and performed experiments relating to LOV-ipaA. Ryan Hallett developed the LOV-SsrA switch and performed all experiments relating to LOV-SsrA. Eun Jung Choi developed the peptide search algorithm. Mary J. Aiken performed LOV-ipaA binding experiments. This work was guided and funded by Klaus Hahn and Brian Kuhlman.

2.1 Overview

Photocontrol of functional peptides is a powerful tool for spatial and temporal control of cell signaling events. We show that the genetically encoded light-sensitive LOV2 domain of Avena Sativa phototropin 1 (AsLOV2) can be used to reversibly photomodulate the affinity of peptides for their binding partners. Sequence analysis and molecular modeling were used to embed two peptides into the Jα helix of the AsLOV2 domain while maintaining AsLOV2 structure in the dark but allowing for binding to effector proteins when the Jα helix unfolds in the light. Caged versions of the ipaA and SsrA peptides, LOV-ipaA and LOV-SsrA, bind their targets with 49- and 8-fold enhanced affinity in the light, respectively. These switches can be used as general tools for light-dependent colocalization, which we demonstrate with photo-activable gene transcription in yeast.

---

[1]This chapter previously appeared as an article in Chemistry & Biology. The original citation is as follows: Lungu, O. et al. Designing Photoswitchable Peptides Using the AsLOV2 Domain. Chem Biol 19, 507–517 (2012)

2.2 Introduction

Peptides regulate a variety of biological processes by acting as competitive inhibitors, allosteric regulators, and localization signals[95–97]. Photocontrol of peptide activity is a powerful tool for precise spatial and temporal control of cellular function[98,99]. Typically, photo-activation of peptides has been achieved by covalently modifying peptides with chemical groups that inhibit function until they are removed by light[10,100]. Because such derivatized peptides must usually be synthesized in vitro, one challenge of this approach is getting the peptides into living cells or animals. Additionally, in most cases the photo-induced reaction is not reversible. Recently, there has been considerable progress in the use of naturally occurring photo-activable proteins to engineer light switches that are genetically encoded and reversible[46,68,101,102]. In the majority of cases, the goal has been to regulate the activity of folded protein domains. Here, we examine if the LOV2 domain from *Avena sativa* phototropin 1 (AsLOV2) can be used to photomodulate the affinity of peptides for binding partners[103].

AsLOV2 is part of the PAS superfamily of domains[16]. It contains a flavin mononucleotide (FMN) cofactor located in the center of the PAS fold, as well as a large α-helical region C-terminal to the fold, termed the Jα helix[19,104]. Upon irradiation with blue light, a covalent adduct is formed between a cysteine side chain in the PAS fold and a carbon atom of the FMN[15,105]. Spectroscopy studies indicate that this leads to a large conformational change in the domain, including the unfolding of the Jα helix[106,107]. When irradiation ceases, reversion of the thiol bond and conformational change back into the dark state occurs spontaneously within seconds to hours, depending on the LOV domain ortholog[108].

The large conformation change that occurs within the Jα helix has been previously harnessed to create a photoswitchable GTPase termed PA-Rac[68]; a photoswitchable variant of the *Escherichia coli* trp repressor that has enhanced affinity for DNA in the light termed

LOV-TAP[69,90]; and a photo-activatable DHFR enzyme[70]. In these studies, entire protein domains were fused to the end of the Jα helix in order to sterically occlude binding with effector molecules or perturb the conformational state of the attached domain. Upon blue light irradiation, unfolding of the Jα helix relieves the steric block or conformational strain. Because peptides are more flexible than are folded domains, it may not be sufficient to place them at the end of the Jα helix to achieve a steric block; tighter caging may be obtained by embedding their functionality within the Jα helix (Figure 2.1). A similar strategy has been successful for caging coiled-coil peptides in the light-sensitive photo-active yellow protein (PYP)[67]. In the case of AsLOV2, the challenge is identifying sequences that incorporate the target binding of the peptide while maintaining the functionality of the Jα helix. One face of the Jα helix is exposed to solvent, while the other face forms hydrophobic interactions with a β sheet in the AsLOV2 domain. Residues on the surface of the helix are expected to be tolerant to mutation, whereas the buried residues should be more conserved. Similarly, most peptides have sets of residues that are required for binding target proteins, whereas other positions can be varied. These observations indicate that it may be possible to identify chimeric sequences for the Jα helix that maintain key interactions with the AsLOV2 domain but incorporate residues critical to peptide function. In this study, we use sequence comparisons along with molecular modeling to create AsLOV2 variants that embed the binding properties of the ipaA[109] and SsrA[110] peptides in the Jα helix.

The usefulness of a photoswitch depends on how much the activity is enhanced by light irradiation (dynamic range) as well as the absolute activity in the dark and in the light. Naturally occurring protein switches vary considerably with regard to absolute activities and dynamic range, indicating that the appropriate switching power for a particular application is likely to be system dependent. For instance, the AsLOV2-derived PA-Rac switch binds to its effector with an affinity of 2 μM in the dark and 200 nM in the light[68], which is appropriate for modulating cellular signaling because a similar change in binding affinity occurs when Rac1

naturally cycles between the GDP and GTP bound state[111]. Studies with the wild-type AsLOV2 domain indicate that it should be possible to create AsLOV2-based switches that show larger changes in activity upon light activation. Nuclear magnetic resonance (NMR) studies have shown that light activation changes the ratio of docked to undocked Jα helix from 98.4:1.6 in the dark to 9:91 in the lit state[23]. This corresponds to a 3.8 kcal mol^{-1}change in free energy, which if efficiently harnessed could be used to create switches that have greater than 100-fold changes in binding affinity for target molecules. In scenarios that involve competitive binding to the Jα helix, that is, the helix is either docked against the LOV domain or bound to an effector molecule, it may be necessary to stabilize the helix docked state in order to take full advantage of the free energy perturbation that light activation provides. Strickland et al. (2010) showed that mutations that stabilize the docked Jα helix could be used to lower the dark state affinity of LOV-TAP for DNA and improve its dynamic range from 5- to 70-fold. Here, we identify further mutations that stabilize a docked Jα helix and show that the peptide switches can also be tuned by varying the location in which the caged sequence is embedded in the Jα helix and by varying the intrinsic affinity of the peptide for its target.

As an application of our tunable peptide photoswitches, we show that the caged ipaA peptide can be used to induce gene expression through light-activated heterodimerization in yeast, demonstrating that the caged peptides can be used as general tools for colocalizing proteins in living cells.

2.3 Materials and Methods

2.3.1 Identifying Peptides Compatible with LOV2 Caging

The protein database was searched for PDB files that have multiple chains and at least one chain with fewer than 30 amino acids. The goal of this search was to create a

comprehensive list of peptides that have been co-crystallized with binding partners. 3137 peptides were identified. To evaluate if a peptide was potentially compatible with caging by the LOV2 domain we used a sliding window of 6 residues to compare the sequence of the Ja-helix with the peptide under consideration. A favorable score was assigned to an aligned residue pair if two residues in the alignment were similar or identical. The following sets of amino acids were considered similar: (K, R), (D, E), (F, Y, W), (N, Q) and (S, T). Additionally, the score for an aligned residue pair was adjusted depending on how buried the LOV2 residue was in the LOV2 domain and how buried the peptide residue was in the peptide/protein co-crystal. The rationale for this scheme was that buried residues are more likely to be important for caging and binding to the target protein. For instance, if the alignment under consideration matched a buried position from the Ja helix with a residue buried in the peptide/protein complex than the match was considered very favorable if they were identical or similar (4 or 3 points respectively), and unfavorable if they were dissimilar (- 4 points). When matching two surface positions there was a weak bonus for identity or similarity (2 or 1 point) and no penalty for dissimilarity. The total score for the alignment was the sum of scores from the six residue pairs. The best scoring alignment had a score of 17 points, and over 200 alignments scored 9 points or better. Despite scoring well in this scheme, a peptide may still be unsuitable for caging if it includes additional critical residues (outside the six residue window) that do not align well with the Ja-helix or do not extend off the end of the Ja-helix as in the case of LOV-ipaA.

2.3.2 Cloning

The LOV-ipaA gene was synthesized with a six histidine N-terminal tag (Genscript, Piscataway, NJ, USA) and cloned into the pET21b vector. The genes for LOV-SsrA and monomeric SspB were synthesized (Genscript) and cloned into pQE-80L and pTriEX4 vectors, respectively. All mutations were performed using site-directed mutagenesis. The

vinculin D1 subdomain (residues 1–258) and full-length vinculin (residues 1–1066) were cloned into a pET15b vector.

2.3.3 Protein Expression and Purification

LOV-ipaA WT and mutants were expressed in E.coli strain BL21(DE3) cells (Genese) at 16°C overnight in the dark. Cells were lysed in buffer containing 50 mM sodium phosphate pH 7.5, 150 mM NaCl and 5 mM 2-mercaptoethanol. Proteins were purified by affinity chromatography over HiTrap HisTrap columns (GE) and eluted with 150 mM Imidazole at pH 7.5. The proteins were further purified through size exclusion chromatography over a Sephadex S75 column (GE) equilibrated with 50 mM sodium phosphate, 150 mM NaCl and 5 mM 2-mercaptoethanol. VinculinD1 subdomain was expressed in E.coli strain BL21(DE3) cells (Stratagene) at 16°C overnight. Cells were lysed in buffer containing 50 mM Tris-HCl pH 8.0, 500 mM NaCl and 5 mM 2-mercaptoethanol. The protein was purified by affinity chromatography over HiTrap HisTrap columns (GE) and eluted with 500 mM Imidazole at pH 8.0. The protein was further purified by ion exchange chromatography using a HiTrap Q column (GE) eluted with a NaCl gradient. VinculinD1 was stored in 20 mM Tris-HCl with 2 mm DTT and 2 mM EDTA.

Vinculin full-length protein was expressed in E.coli strain BL21(DE3) cells (Genese) at 37°C for four hours. Cells were lysed in buffer containing 50 mM Tris-HCl pH 8.0, 500 mM NaCl and 5 mM 2-mercaptoethanol. The protein was purified by affinity chromatography over HiTrap HisTrap columns (GE) and eluted with 500 mM Imidazole at pH 8.0. The protein was then purified by ion exchange chromatography using a HiTrap Q column (GE) eluted with a NaCl gradient. Vinculin was further purifiedthrough size exclusion chromatography over a Sephadex S200 column (GE) equilibrated with 50 mM sodium phosphate, 150 mM NaCl and 5 mM 2-mercaptoethanol.

LOV-SsrA and SspB Constructs were expressed in BL21 *E.coli* cells and induced with 333mM IPTG overnight at 18°C. Proteins were purified through binding Ni-NTA columns and eluted against an imidazole gradient. Size exclusion chromatography using an S75 column was performed as a final clean-up step and to switch buffer to PBS. Protein concentrations for LOV-SsraC, LOV-SsraN, Sspb LOV-ipaA and vinculinD1 were determined using Bradford assays (Thermo). Protein concentrations for full-length vinculin were determined using absorbance at 280 nm measurements with an extinction coefficient of 62,000 $M^{-1}cm^{-1}$.

2.3.4 Peptides

Peptides containing the sequence TANNIIKAAKDATTSLSKVLKNIN, TANNIIKAAKDATTSASKVLNIN, TANNIIKAAKDATTSLSKALKNIN, QIEEAANDENY, LIKKAANDINYAAK, and HVRDAANDEAYMLIK were synthesized at UNC-Chapel Hill and amine labeled using 5-(and-6)-Carboxytetramethylrhodamine (TAMRA) dye (Anaspec, Fremont, CA, USA). Peptide concentration was determined by measuring absorbance of the TAMRA dye at 555 nm using 65,000 $M^{-1}cm^{-1}$ extinction coefficient.

2.3.5 Fluorescence Polarization Experiments

All fluorescence polarization experiments were conducted using a Jobin Yvon Horiba FluoroMax3 fluorescence spectrometer. TAMRA-labeled peptides were excited with polarized light at 555 nm, and the polarization of emitted light was measured at 583 nm. For LOV-ipaA, Competition assays were conducted using 50 nM TAMRA-ipaA peptide and a concentration of vinculin appropriate for binding 85% of the TAMRA- ipaA peptide, as determined from the previously measured dissociation constant for TAMRA-ipaA and vinculin. Varying concentrations of LOV-ipaA were then added to the reaction. The reaction was allowed to proceed for 3 hours at 25°C, with polarization of TAMRA dye emission measured at intervals of 5 min. Competition curves were fit using a first order exponential

decay equation to determine the polarization at time point zero. The fraction of TAMRA-ipaA bound to vinculin was determined by normalizing each curve using the time point zero as the calculated 85% bound point and the polarization of TAMRA-ipaA alone as the zero fraction bound point in the dark. In the light, each curve was normalized with a zero fraction bound point determined by mixing TAMRA-ipaA and LOV-ipaA at the concentration used in the titration under blue light and setting the resulting polarization as the zero fraction bound point. Normalized curves were fit for k_{on} and K_D rates using a numerical integration script in Matlab. Six curves were used per fit, and two fits were averaged for each kinetic measurement.

Binding of SspB to TAMRA-labeled peptides was performed in a 1cm quartz cuvette with a starting peptide concentration of 25 nM in PBS at 25°C. Increasing concentrations of SspB were titrated into the cuvette and polarization was recorded. Binding of LOV-SsrA designs to SspB was measured through competition with the TAMRA-labeled peptide. Competition binding experiments began with a solution containing 25 nM labeled peptide and 40 nM SspB. Titrations of LOV-SsrA designs were exposed to blue light (455nm, 6.0 mW cm^{-2}) for 2 minutes, and polarization readings were taken the instant the light was switched off. After 5 minutes of darkness, a second polarization reading was taken. Competition with the peptide-SspB complex allowed for fitting of an IC50 value for the design and determination of KD according to Nikolovska-Coleska et al. (Nikolovska-Coleska, et al., 2004). Reversibility of LOV-SsrAC binding to SspB was performed by repeatedly irradiating the 1.5 μM LOV-SsrA titration point from the competition assay with blue light (455nm, 6.0 mW cm^{-2}) for 60 seconds and monitoring polarization for 4 minutes.

2.3.6 Illumination

For fluorescence polarization assays, all LOV2-containing proteins were continuously irradiated for at least 1 min using a collimated blue LED with maximum emission wavelength

of 455 nm (ThorLabs, product code LEDC1). The irradiation was directly 10 cm above a sample chamber of 0.9 cm diameter. Illumination power was measured to be 6.0 mW cm^{-2} in the sample holder. For yeast two- hybrid experiments, yeast plates were initially grown using a set of 50 LED wide angle holiday string lights (Target, 120V, 0.04A) placed in a 41x 46 cm array 30 cm above the plates, yielding good colony growth. Optimal growth was observed using an array of 32 blue LED lights (Optek) arranged in series on a 5.5x5.5 cm array placed 30 cm above the plates. Lights had maximal emission wavelength of 470 nm, and were operated at 5 V, 0.5 A.

2.3.7 LOV-SsrA Design

Alignments between the SsrA peptide and LOV2 were modeled with the Rosetta molecular modeling program. The domain assembly protocol was used to assemble the C-terminally aligned designs. The fixed backbone design protocol was used to explore mutations within LOV2-SsrA. C-terminal extensions to LOV2-SsrA were modeled using the floppytail protocol in Rosetta. Designs were sorted by score and the lowest scoring designs were examined. Monomeric SspB was developed using the fixed backbone design and ridged-body docking protocols in Rosetta. SspB was confirmed to be a monomer using SEC-MALS.

2.3.8 LOV-SsrA Thermal Reversion

Photocycle kinetics were determined by monitoring recovery of absorbance at 450nm of 10μM LOV-SsrAC after 60 seconds of irradiation with blue light (455nm, 6.0 mW cm^{-2}) with and without 20 μM SspB. All measurements were taken at 25^{o}C.

2.3.9 Surface Plasmon Resonance

Surface plasmon resonance experiments were conducted using a Biacore 2000 machine (GE Healthcare, Waukesha, WI, USA). Vinculin D1 was immobilized through amine

coupling to the surface of a CM5 chip (GE Healthcare). Different concentrations of LOV-ipaA mutants were flown over the immobilized protein, and the change in response units over time was recorded. Data were fit simultaneously for k_{on} and k_{off} to a pseudo-first-order binding model.

2.3.10 Actin Co-sedimentation Assays

Purified rabbit actin (Invitrogen) was polymerized for 30 min at room temperature in 10 mM Tris-HCl (pH 7.5) containing 100 mM KCl, 2 mM MgCl2, 2 mM DTT, and 1 mM ATP. Vinculin (2 μM) and either ipaA peptide or LOV-ipaA mutants were mixed for vinculin:ipaA rations of 1:0, 1:1, 1:2.5, 1:5, 1:10, 1:20, or 1:50 per sample. Polymerized actin (12 μM) was added to each sample, within a volume of 45 μl per sample. Samples were incubated at room temperature for 1 hr. They were then centrifuged in a TLA-100 rotor in a Beckman Coulter Optimax XP ultracentrifuge at an acceleration of 150,000 g for 30 min cooled to 20°C. Samples were split into supernatant and pellet fractions. Pellets were resuspended into 45 μl 2x tris-glycine SDS buffer. All fractions were denatured and run onto an 8% SDS-PAGE polyacrylamide gel. Gels were Coomassie stained and analyzed using ImageJ software to determine the fraction of vinculin present in the pellet versus total vinculin in each sample[112].

2.3.11 Analysis of Actin Co-sedimentation Assays

Apparent binding affinity of ipaA or LOV-ipaA to vinculin was determined by plotting the fraction of vinculin bound to actin versus concentration of LOV-ipaA and fitting the curve to the equation below using Prism software. A total of 3 gels were quantified and averaged for each binding affinity measurement.

$$F = a + (b - a) \times \left(\frac{(K + X_0 + V_0) - \sqrt{(K + X_0 + V_0)^2 - 4(X_0 V_0)}}{2V_0} \right)$$

Where *F* is the fraction of vinculin bound to actin in the pellet, *a* is the minimum fraction bound, *b* is the plateau for fraction bound (set to 0.39), *K* is the binding affinity, X_0 is the concentration of ipaA or LOV- ipaA added to the reaction, and V_0 is concentration of vinculin in the reaction (set to 2 mM).

2.3.12 Isothermal Calorimetry Binding Experiments

All ITC experiments were performed at UNC Chapel Hill in the Macromolecular Interaction Facility using a MicroCal Auto ITC200. Purified LOV-ipaA L623A lit state (I532E A536E), dark state (C450A) mimetics, and vinculinD1 were dialyzed for four hours in 50 mM sodium phosphate, 150 mM NaCl and 5 mM 2-mercaptoethanol buffer.

2.3.13 Initial LOV-ipaA competition inhibition measurements

The binding affinity of synthesized TAMRA labeled ipaA peptide to vinculinD1 was determined by titrating increasing concentrations of vinculinD1 into 2 nM TAMRA labeled ipaA peptide and measuring fluorescence polarization at the emission wavelength of TAMRA, 583 nm. The curve was fit using a one site binding model in Sigma Plot software. The off rate of TAMRA labeled ipaA peptide binding to vinculinD1 was determined by using 50 nM TAMRA ipaA peptide 85% bound to vinculinD1 in a 0.3 cm path length cuvette. A 30-fold excess of unlabeled ipaA peptide was then added to the reaction. The reaction was allowed to proceed for 3 hours at 25°C, with polarization of TAMRA dye emission measured at intervals of 5 min and was fit to a one phase exponential decay equation to determine the first order rate constant.

2.3.14 Structural Software and Data Presentation Methods

All model and crystal structure figures were made using Pymol. All gel images were taken using a Kodak GelLogic 100 imaging system and slightly adjusted for optimal brightness and contrast equally throughout the gel area using Adobe Photoshop. Yeast

plates were scanned using an HP ScanJet 4850 scanner and resulting images were cropped of whitespace using Adobe Photoshop.

2.3.15 Yeast Two-Hybrid Assays

LOV-ipaA L623A WT, lit-state mutants, dark-state mutants, and ipaA were cloned into a pGADT7 vector and transformed into *S. cerevisiae* Y187 strain, whereas the vinculin D1 subdomain was cloned into a pGBKT7 vector and transformed into *S. cerevisiae* Y2Hgold strain (Clontech, Mountain View, CA, USA). Empty vectors were also transformed into the appropriate strains. Transformed colonies of Y2Hgold and Y187 were mated at 30°C overnight and plated on synthetic dextrose (SD) –Leu –Trp media. For yeast two-hybrid experiments, mated colonies were serially diluted (1:5, from right to left on plates shown) and replica plated onto SD –Leu –Trp media; SD –Leu –Trp media with auerobasidin A and 5-bromo-4-chloro-3-indolyl-α-D-gactopyranoside (x-α-gal); SD –Leu –Trp –His media with auerobasidin A and x-α-gal; and SD –Leu –Trp –His –Ade media with auerobasidin A and x-α-gal. Plates were grown for three days at 30°C. For Miller assays, mated colonies were picked and grown to saturation in SD –Leu –Trp media. Saturated colonies were diluted to low-log phase and grown for 4 hr under dark or blue light conditions. Cells were lysed open and treated with chlorophenol red-β-D-galactopyranoside (CPRG; Roche) substrate to determine β-galactosidase activity in Miller units[46].

2.4 Results

2.4.1 Identifying Peptides Compatible with AsLOV2 Caging

To identify protein binding peptides well suited for caging with the AsLOV2 domain, we searched the protein database (PDB) for peptide sequences similar to portions of the Jα helix. The similarity score for each position in the alignment was weighted based on how likely a match or mismatch was to disrupt caging or peptide binding. For instance, if a

position in the alignment mapped to residues important for both the AsLOV2-Jα helix interaction and the peptide-protein interaction, then a large favorable score was given if the amino acids were identical, and a large unfavorable score was given if the amino acids were dissimilar. The importance of a residue to the AsLOV2-Jα helix interaction or the peptide-protein interaction was assigned based on how buried the residue was; residues that were more buried were considered more important. There was little reward or penalty for conserving amino acids on the surface of the Jα helix. Using a sliding window of six residues, the sequence of the Jα helix was aligned with over 3,000 peptide sequences taken from crystal structures of peptides co-crystallized with their protein binding partners. Peptides from several hundred structures were identified as candidates for caging with the AsLOV2 domain. It is worth noting that this approach can be expanded to also consider known protein binding peptides that have not been crystallized with their binding partners. We picked out two peptide sequences for experimental testing and optimization: a vinculin binding peptide from the invasin protein ipaA[113] and the SsrA peptide from *E. coli,* which binds the protease delivery protein SspB[110]. These sequences were chosen for a variety of reasons. First, at functionally important residues they align well with the Jα helix (Figures 2.2A and B). Indeed, several alternative alignments were identified for the SsrA peptide. Second, they adopt alternative conformations when binding their targets; ipaA adopts a helix, whereas SsrA binds in an extended conformation (Figure 2.3). By testing both peptides, we examine if our approach can be used to cage peptides regardless of the conformation they adopt when bound to their target protein. Third, the peptides have different intrinsic affinities for their target proteins. IpaA binds the D1 domain of vinculin very tightly, Kd < 1 nM, whereas SsrA binds SspB with an affinity of ~30 nM.

Finally, they have complementary advantages for use in controlling cell biology. The SsrA and SspB sequences are specific to bacteria, and therefore it is expected that they will not interact with other proteins and peptides in higher organisms. Fusing proteins of interest to a

photo-activable SsrA and SspB should provide a general approach for light-induced heterodimerization, which can be used to localize proteins in the cell and activate cell-signaling pathways. Vinculin and ipaA-like peptides are found in mammals, and therefore caged ipaA should only be useful as a general tool for colocalization in orthogonal systems, such as yeast. However, caged ipaA may be useful for probing vinculin biology in mammalian cells. Vinculin is a protein that connects integrin binding proteins, such as talin[114], to the actin cytoskeleton at focal adhesions[115] and adherens junctions[116], cellular structures important for determining cell shape and motility. The ipaA peptide is from the IpaA protein of the *Shigella flexneri* bacterium and binds to the talin binding site on vinculin[113]. It has been proposed that IpaA binding prevents vinculin from binding talin, and thus linking integrin signaling to the actin cytoskeleton. Selectively photo-controlling the binding of ipaA to vinculin through the LOV-ipaA photoswitch could render it useful as a dominant negative inhibitor of vinculin in mammalian cells and an integral tool for studying the role of vinculin dynamics in cell motility.

2.4.2 Design of LOV-ipaA

In our search of peptides in the protein database, the first ten residues of the ipaA VBS1 helical peptide were identified as a close match to the last ten residues of the AsLOV2 Jα helix (Figure 2.2A). Five of the ten positions are identical, and the hydrophobic residues on the Jα helix that make critical contacts with the AsLOV2 domain β sheet at residues 539, 542, and 543 are conserved in the alignment with ipaA. Additionally, residues in the ipaA sequence that make extensive contacts with vinculin are conserved in the alignment (Ile 612, Ala 615, Ala 616, and Val 619 in ipaA). This analysis suggests that it should be possible to design a chimeric sequence for the Jα helix that is compatible with the AsLOV2 structure and will bind vinculin when undocked from the AsLOV2 structure.

To design the chimeric sequence, the Rosetta molecular modeling program was used to assess the impact of altering the AsLOV2-Jα and vinculin-ipaA complex sequences[117]. Side-chain optimization simulations were used to thread the first ten residues of ipaA onto the last ten residues of the Jα helix. The Rosetta scores of individual residues were examined to determine if particular residues in the ipaA sequence packed unfavorably against the AsLOV2 domain β sheet. Only the mutation of the residue in position 540 from aspartic acid to tyrosine showed unfavorable scores. To search for an alternative amino acid to place at this position, Rosetta was used to perform a sequence optimization simulation in which position 540 was allowed to adopt alternative identities and neighboring side chains were free to adopt new side-chain conformations. One of the best scoring residues in the 540 position on the Jα helix was isoleucine. A working model for LOV-ipaA was then created by using the threaded AsLOV2-Jα design and adding the remaining eleven residues of ipaA onto the C-terminal end of Jα using the fragment insertion capability of Rosetta's domain assembly protocol (Figure 2.2C).

2.4.3 Dark- and Lit-State Binding between LOV-ipaA and Vinculin D1

We developed a fluorescence polarization competition assay to measure the binding affinity and kinetics of LOV-ipaA to the vinculin D1 subdomain under dark, as well as blue light, irradiation conditions (Figure 2.4A). LOV-ipaA was titrated into a mixture of the vinculin D1 subdomain bound to ipaA peptide labeled with the dye 5-(and-6)-Carboxytetramethylrhodamine (TAMRA). As LOV-ipaA bound to vinculin the dye-labeled peptide was competed off and its fluorescence polarization signal decreased. Initial experiments showed that the slow intrinsic off-rate of the ipaA peptide and LOV-ipaA from vinculin made it possible to monitor both the kinetics and thermodynamics of binding over a time course of several hours. Experiments with varying amounts of LOV-ipaA were fit

simultaneously for the on rate and binding affinity of LOV-ipaA to vinculin D1 in Matlab using a numerical integration protocol.

The same concentrations of LOV-ipaA under either blue light irradiation or dark-state conditions yielded different kinetic curves (Figure 2.4B and C). The binding affinity of LOV-ipaA WT in the light is 3.5 nM, whereas it is 69 nM in the dark, a 19-fold change (Table 2.1). The change in binding affinity upon light activation is primarily due to changes in the on rate for binding; in the dark, the rate constant is 1.3×10^3 M^{-1} s^{-1}, and in the light it is 1.4×10^4 M^{-1} s^{-1}. Similar changes were observed for LOV-ipaA mutants that abolish FMN-thiol bond formation (C450A) and lead to a pseudodark state or mutants that destabilized the Jα helix (I532E A536E), causing a pseudo-lit-state conformation. These results are consistent with the proposed mechanism of caging. In the dark, the Jα helix is primarily docked against the AsLOV2 domain, and the peptide is presented less frequently to the target binding site, slowing the on rate for binding. In general, the on rates for binding are slow both in the lit and dark state. Peptides and proteins often have k_{on} values greater than 1×10^5 M^{-1} s^{-1} [118]. The slow rates observed here are consistent with the helix addition mechanism of ipaA binding to vinculin, wherein the vinculin D1 4-helix bundle is rearranged into a 5-helix bundle through the addition of ipaA[119].

2.4.4 Optimization of LOV-ipaA

We tested two sets of mutations predicted to enhance the dynamic range of the LOV-ipaA photoswitch. The first set of mutations was designed by Strickland et al. (2010) to stabilize the helical structure of Jα. The mutations, G528A and N538E, increased the dynamic range of the LOV-TAP photoswitch for its effector from 5- to 70-fold. When used in the LOV-ipaA system, the mutations did have a large effect on LOV-ipaA dark-state binding to vinculin D1, decreasing affinity more than 7-fold to 475 nM (Table 2.1). However, the mutations also weakened lit-state binding affinity over 10-fold to 110 nM. The net effect of

the G528A and N538E mutations in LOV-ipaA then was a decrease in the photoswitching dynamic range of the protein. The weakened affinity for vinculin in the lit state is likely due to charge-charge repulsion between Glu538 in LOV-ipaA (G528A and N538E) and a glutamate close to the binding site of ipaA in vinculin.

We designed a second set of mutations to increase the dynamic range of the switch, L514K L531E, using the interactive modeling program FoldIt[120]. These mutations replaced two hydrophobic residues, one in the Jα helix, and the other on the β sheet contacting the Jα helix, with a salt bridge. The design was meant to stabilize interactions between the β sheet and Jα helix and lead to a more tightly bound helix in the dark state. Indeed, the binding affinity of LOV-ipaA L514K L531E to vinculin D1 in the dark state weakened to 245 nM, whereas the lit-state affinity increased only slightly to 5 nM. This set of mutations make LOV-ipaA a photoswitch with a 49-fold difference between the lit- and dark-state effector binding affinities.

To independently validate results from the florescence polarization competition assay, surface plasmon resonance experiments were also conducted. The on rate, off rate, and binding affinity of vinculin D1 to LOV-ipaA L514K L531E were measured. LOV-ipaA L514K L531E pseudolit (I532E A536E [20]) and pseudodark (C450A [121]) states were used. Pseudo-dark-state LOV-ipaA L514K L531E (Figure 2.4D) was able to bind vinculin D1 with an on rate of 4.5×10^{2} M^{-1} s^{-1} and an off rate of 2.9×10^{-5} s^{-1}. A binding affinity of 64 nM was calculated from the kinetic data. The on rate was identical to that measured for LOV-ipaA L514K L531E under dark-state conditions using the fluorescence polarization assay. The off rate varied slightly, leading to a 4-fold difference in binding affinity between the two measurements. In part, the discrepancy might be due to the fact that it is difficult to fit an off rate that is so slow. Measuring the LOV-ipaA L514K L531E pseudolit photoswitch binding to vinculin D1, an on rate of 3.8×10^{4} M^{-1} s^{-1} and an off rate of 8.7×10^{-5} s^{-1} were obtained (Figure 2.4E). Fitting indicated that the binding affinity was 2.3 nM. These values are similar

to the kinetic rates measured for the LOV-ipaA lit mimetic using the fluorescence polarization assay.

There is a 49-fold increase in binding affinity for vinculin when LOV-ipaA L514K L531E is activated with light but binding in the dark is still considerable (245 nM). For some applications, it may be useful to have a switch that has weaker binding affinity in the dark. To test if we could switch the range of affinities over which LOV-ipaA functions, we made a mutation to LOV-ipaA that was predicted to reduce affinity for vinculin but should have negligible effect on the interactions between the Jα helix and the LOV domain. This mutant, L623A, weakened the affinity of the LOV-ipaA lit-state mimetic for the vinculin D1 domain from 3 nM to 2.4 μM, whereas the dark-state mimetic had no detectable binding as monitored with isothermal titration calorimetry, suggesting that binding in the dark is weaker than 40 μM (Figure 2.5).

2.4.5 Design and Optimization of LOV-SsrA

We extended the design strategy used to cage ipaA in order to create a second photoswitchable peptide, LOV-SsrA. The SsrA peptide interacts with the protease delivery protein SspB from *E. coli* as a linear epitope using the seven residue sequence AANDENY[110]. Three possible alignments between SsrA and the Jα helix were identified in our PDB-wide search, one toward the N-terminal side of the helix (LOV-SsrAN), one near the C-terminal end (LOV-SsrAC), and one in the middle of the helix (LOV-SsrAM; Figure 2.2B). We first tested the C-terminal alignment LOV-SsrAC.

In LOV-SsrAC, the two alanines from SsrA were aligned with A542 and A543 from the last helical turn of the Jα helix, and the final three residues of the Jα helix were replaced with residues from the SsrA binding sequence (Figure 2.2D). Except for the last leucine, all buried positions in the Jα helix are conserved with this alignment, which only embeds the peptide in a single helical turn of the Jα helix. In comparison, the ipaA peptide is embedded

within two helical turns in LOV-ipaA. To measure the affinity of LOV-SsrA for SspB in the dark and light, a fluorescence polarization competition assay was performed with TAMRA-labeled SsrA peptide[122]. Unlike ipaA, SsrA binds rapidly to SspB and only equilibrium populations could be measured. The initial LOV-SsrA design showed a 2-fold change in affinity of 31 nM to 57 nM when activated with light (Table 2.2). As lit-state binding was near-native in affinity for SspB, approaches to stabilize the Jα helix and therefore decrease dark-state affinity were explored. We tested mutations previously shown to stabilize the Jα helix, G528A and N538E, as well as extensions to the C terminus of LOV-SsrAC. G528A and N538E had the desired effect and lowered dark-state affinity to 570 nM with a 5-fold change upon light activation. Extensions of varying lengths were designed with Rosetta[123] with the goal of binding the surface of the AsLOV2 domain, thereby constraining SsrA to a single, nonbinding conformation. The target surface on AsLOV2 was a hydrophobic patch immediately adjacent to the C terminus of the Jα helix. Two extensions were experimentally tested, the addition of a single phenylalanine and the addition of the sequence GYGNL. The longer extension had no detectable effect, whereas the single phenylalanine increased the dynamic range of the switch to 8-fold (Figure 2.6A).

The LOV-SsrAN designs incorporating the N-terminal alignment of SsrA had similar dynamic ranges to their corresponding C-terminal alignments. However, all had drastically reduced affinity for SspB in both the lit and dark state (>10 μM). The initial N-terminally aligned designs showed an approximate 2-fold increase in affinity after exposure to blue light. The addition of the G528A and N538E mutations slightly weakened the lit-state affinity for SspB to 12.6 μM and the dark-state affinity to 49 μM (Figure 2.6B). Because of constraints imposed by the AsLOV2 domain, the SspB binding sequence used in LOV-SsrAN was AANDEAY instead of AANDENY as used in LOV-SsrAC. To determine if the sequence change was responsible for the lower dark- and lit-state affinities for SspB, a peptide with the same sequence as LOV-SsrAN was synthesized and binding to SspB was

measured. Binding was tight, 38 nM, suggesting that the reduced affinity for SspB was not because of changes to the binding sequence but rather because of the location within the Jα helix that the peptide was embedded. Results with LOV-SsrAM were more ambiguous, as the sequence embedded in the Jα helix, AANDINY, showed reduced affinity for SspB as a peptide (1.9 μM). However, lit- (40 μM) and dark-state (156 μM) state affinities were still significantly lower than that the isolated peptide, suggesting again that embedding sequences deeper in the Jα helix lowers accessibility to the peptide in both dark and light states (Figure 2.7).

The reversibility of LOV-SsrA switches was tested by monitoring their ability to compete with a TAMRA-labeled SsraA peptide for binding to SspB over multiple rounds of irradiation followed by incubation in the dark (Figure 2.6C). As monitored indirectly from following the fluorescence polarization signal from the SsraA peptide, the fraction of LOV-SsrA molecules switching from a bound to unbound state remained unchanged over multiple rounds of illumination, indicating that the switches are reversible. To check that the observed results were not an artifact due to interaction between the blue light and the TAMRA dye, similar reversion experiments were performed with only SspB and labeled peptide. No polarization changes were observed after incubation with blue light. Half-times between 27 and 50 s have been reported for the LOV2 domain from *Avena sativa*[15,121]. Using absorption spectroscopy, we determined the half-life of LOV-SsrAC's photocycle to be 28.4 ± 0.1 s*; Figure 2.7). *Note: this half-life has been updated since a contamination was found. The values here were published in an erratum later in 2012.

2.4.6 LOV-ipaA Binding to Full-Length Vinculin: Actin Cosedimentation Assays

Having established that the AsLOV2 domain could be used to cage peptides, we tested if LOV-ipaA could be used to perturb vinculin in a biologically relevant manner. Actin cosedimentation assays were performed to measure the apparent binding affinity of LOV-

ipaA to full-length vinculin. Vinculin naturally forms an autoinhibited conformation stabilized by interactions between the head (including the D1 domain) and tail domains. Binding of the ipaA peptide to the head domain competes with the autoinhibited state and releases the tail domain to bind with polymerized F-actin[113]. Actin cosedimentation assays were performed to measure the apparent binding affinity of LOV-ipaA in the dark and lit state to full-length vinculin. In the assay, polymerized actin, vinculin, and LOV-ipaA (lit- or dark-state mutants) were incubated together (Figure 2.8A). Only vinculin that bound to LOV-ipaA should be able to interact and bind to F-actin. The resulting mix of bound and unbound vinculin was centrifuged at high speeds, resulting in fractionation of F-actin polymers out of solution, along with the vinculin-LOV-ipaA complexes that had bound to them. A sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel was used to monitor the amount of vinculin that fractionated out of solution along with F-actin as a function of the total concentration of vinculin, thereby determining the fraction of full-length vinculin bound to LOV-ipaA. The assay was repeated at several concentrations of LOV-ipaA to create a binding curve and calculate apparent binding affinities (Figure 2.8B and C). In general, the affinities are expected to be considerably weaker with full-length vinculin than those observed for binding to the isolated D1 domain, because in the full-length protein there is competition with the autoinhibited state of vinculin[112]. For instance, peptides from the protein talin bind to the vinculin D1 domain with an affinity of 15 nM but bind to full-length vinculin with an affinity of ~5 μM[112,124]. LOV-ipaA with the lit-state mutant bound with an apparent affinity of 8 μM to vinculin, whereas the dark-state mimetics bound with an affinity of 56 μM (Table 2.3). The binding of vinculin to ipaA in the presence of polymerized actin also gives us a window into the reactions that occur in vivo, where full-length vinculin must be bound both by talin and by F-actin in order to engage in integrin signaling. The difference in binding affinity between lit and dark states of LOV-ipaA to full-length vinculin in the presence of

polymerized actin suggests that LOV-ipaA may be a relevant photoswitch for probing such signaling in vivo.

2.4.7 Photo-Activatable Yeast Transcription

To demonstrate that the caged peptides can be used to colocalize proteins and activate signaling events in living cells, we used the LOV-ipaA-vinculin D1 interaction as a heterodimerization switch for controlling gene expression. The experiments were performed in *Saccharomyces cerevisiae*, as it is an orthogonal system, lacking proteins that would cross-react with the LOV-ipaA-vinculin D1 interaction. We linked LOV-ipaA L623A to the GAL4 activation domain (AD), while linking vinculin D1 (vinD1) to the GAL4 binding domain (BD), and monitored the GAL4-induced activation of the transcriptional *GAL* promoter through a yeast two-hybrid assay. This is a strategy that has been widely used to identify protein-protein interactions. In our case, LOV-ipaA and vinculin D1 should only interact upon irradiation with blue light, thus bringing the GAL4 AD and BD into proximity (Figure 2.9A). The full GAL4 protein can then activate the expression of reporter genes downstream of the *GAL* promoter. We tested the ability to activate the reporter gene LacZ under both dark- and lit-state conditions through quantification of β-galactosidase activity (Figure 2.9B). We observed an activity of 5 Miller units under blue light conditions, 20 Miller units with the LOV-ipaA L623A lit-state mimetic, and 30 units using the ipaA peptide. Almost no activity (0.4 Miller units) was seen under dark or dark-state conditions, and no activity was seen in empty vector negative controls. We also tested activation of genes *MEL1*, *HIS3*, and *ADE2* under dark- and lit-state conditions by replica plating BD-vindD1 colonies mated with AD-LOV-ipaA L623A lit- and dark-state mimetic mutants or AD-LOV-ipaA L623A WT (Figure 2.9C). We saw strong growth of colonies containing LOV-ipaA L623A lit mimetic or WT grown in blue light on plates lacking His, as well as those lacking His and Ade, indicating expression of both*HIS3* and *ADE2* genes. Furthermore, colonies were blue, indicating expression of

the *MEL1* gene, whose product, α-galactosidase, interacted with the x-α-gal substrate for blue screening. Colonies containing LOV-ipaA L623A dark mimetic or WT grown in the dark grew on control plates but did not grow on plates lacking His or Ade, showing low to no expression of *HIS2*, *ADE2*, or *MEL1* genes. This indicates that LOV-ipaA-vinculin D1 heterodimerization can be used as a tool to photocontrol yeast gene expression.

The LOV-ipaA L623A mutation was critical to achieving a lit- to dark-state change of phenotype in yeast gene expression. When experiments for *MEL1* and *HIS3* were conducted using LOV-ipaA lacking this mutation, no observable growth change was seen between colonies containing lit- and dark-state AD-LOV-ipaA mated with BD-vinD1 (Figure 2.10). The binding affinity of LOV-ipaA to vinculin D1 was hence so tight, even in the dark state, as to allow binding events and subsequent yeast gene expression to occur. The L623A mutation shifted the binding affinity of LOV-ipaA for vinculin D1 to a range in which there was no binding in the dark but significant activation in the light. This result highlights the usefulness of being able to tune the switches for specific applications.

2.5 Discussion

We have shown that protein-binding peptides can be embedded in the AsLOV2-Jα helix so that their affinity for effector proteins is weakened in the dark but is enhanced when light-activation releases the Jα helix from the LOV domain. The approach is applicable to a variety of peptide sequences because only a subset of residues on the Jα helix need to be conserved to maintain favorable interactions with LOV domain. Additionally, most protein-binding peptides contain residues that can be mutated without significantly weakening affinity for binding partners. In the case of LOV-ipaA and LOV-SsrAC, we also took advantage of the fact that not all of the peptide needs to be embedded in the Jα helix to create a steric block against effector binding. In this scenario, only a few residues from the N-terminal portion of the peptide need to be compatible with the folded Jα helix.

For some applications, it may be necessary to tune the designed photoswitch to have a dark-state activity and dynamic range compatible with the desired outcome. In the yeast two-hybrid experiments with LOV-IpaA, we found that we needed to weaken dark-state binding in order to prevent gene expression in the dark. We accomplished this by introducing a mutation that weakens the baseline affinity of IpaA for vinculin. Whereas the initial design switched from 3.5 nM to 69 nM affinity for vinculin, the mutant switched from 2.4 μM to >40 μM. The switching power of the LOV-peptide switches were also manipulated by introducing mutations that stabilize the interaction between the Jα helix and the rest of the LOV domain, as well as by varying how deeply the caged peptide was embedded in the helix. It was somewhat surprising that the SsrA peptide had significantly lower litstate affinity for SspB when embedded in the N-terminal portion of the Jα helix than when placed at the C terminus. NMR experiments with the WT AsLOV2 domain indicate that the helix undocks as a cooperative unit when the protein is activated with light[23]. This suggests that there should be similar levels of access to residues in the N- and C-terminal regions of the Jα helix in the lit state; therefore, we expected that lit-state binding affinity for the peptide would be relatively insensitive to where the peptide was placed in the Jα helix. Our results indicate that the N-terminal portion of the Jα helix may be less accessible in the lit state, perhaps by transient interactions with the hydrophobic face of the LOV domain β sheet. Additionally or alternatively, the mutations we have made to the Jα helix to create LOV-SsrAN may have resulted in stronger interactions between the Jα helix and the rest of the AsLOV2 domain in the light.

Both the LOV-ipaA- vinculin D1 and the LOV-SsrA-SspB photo-activatable binding interactions can be harnessed as tools for photo-activatable heterodimerization. The LOV-ipaA- vinculin switch can be used in bacteria and in yeast cells, as these systems do not contain vinculin or vinculin binders that would affect the LOV-ipaA-vinculin interaction. On the other hand, the LOV-SsrA-SspB photo-activatable heterodimerization binders should be

useful in higher organisms, such as mammalian cells. LOV-ipaA has a very slow off rate for vinculin D1, so it is better suited for long timescale applications, such as yeast mating. In contrast, LOV-SsrA binding to SspB is rapidly reversible so can be used for more transient interactions, such as single cell motility experiments. In these ways, both LOV-ipaA and LOV-SsrA should be useful tools to spatially and temporally bring together proteins for activating signaling cascades.

2.6 Supporting Information

2.6.1 FASTA sequences of LOV-ipaA and LOV-SsrA

>LOV-ipaA

MHHHHHHGSLATTLERIEKNFVITDPRLPDNPIIFASDSFLQLTEYSREEILGRNCRFLQGPE
TDRATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNLFHLQPMRDQKGDVQYFIGVQLDGTE
HVRDAAEREGVMLIKKTANNIIKAAKDVTTSLSKVLKNIN

>LOV-ipaA L514K L531E I532E A536E

MHHHHHHGSLATTLERIEKNFVITDPRLPDNPIIFASDSFLQLTEYSREEILGRNCRFLQGPE
TDRATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNLFHLQPMRDQKGDVQYFIGVQKDGTE
HVRDAAEREGVMEEGKTENNIIKAAKDVTTSLSKVLKNIN

>LOV-ipaA L514K L531E I532E A536E L623A

MHHHHHHGSLATTLERIEKNFVITDPRLPDNPIIFASDSFLQLTEYSREEILGRNCRFLQGPE
TDRATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNLFHLQPMRDQKGDVQYFIGVQKDGTE
HVRDAAEREGVMEEGKTENNIIKAAKDVTTSASKVLKNIN

>LOV-SsrAC

MRGSHHHHHHGEFLATTLERIEKNFVITDPRLPDNPIIFASDSFLQLTEYSREEILGRNCRFL
QGPETDRATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNLFHLQPMRDQKGDVQYFIGVQL
DGTEHVRDAAEREAVMLIKKTAEEIDEAANDENYF

>LOV-SsrAM

MRGSHHHHHHGEFLATTLERIEKNFVITDPRLPDNPIIFASDSFLQLTEYSREEILGRNCRFL
QGPETDRATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNLFHLQPMRDQKGDVQYFIGVQL
DGTEHVRDAAEREAVMLIKKAANDINYAAKEL

>LOV-SsrAN

MRGSHHHHHHGEFLATTLERIEKNFVITDPRLPDNPIIFASDSFLQLTEYSREEILGRNCRFL
QGPETDRATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNLFHLQPMRDQKGDVQYFIGVQL
DGTEHVRDAANDEAYMLIKKTAEEIDEAAKEL

2.6.2 Matlab code for fitting LOV-ipaA polarization competition assay

```
clear all
```

```
%  ------------------------------------------
% starting protein concentrations (M)
ipa_tot = 43e-9;
vin_tot = 47e-9;
lov_vin_start = 0.0;
nlov = 6;
lov_array = [80e-9 150e-9 300e-9 1000e-9 3000e-9 5000e-9 ];
% experimental results (t= seconds)
tout_exp = dlmread('time.csv');
ntout_exp = 38;
ipa_vin_exp(:,1) = dlmread('D80nm.csv');
ipa_vin_exp(:,2) = dlmread('D150nm.csv');
ipa_vin_exp(:,3) = dlmread('D300nm.csv');
ipa_vin_exp(:,4) = dlmread('D1um.csv');
ipa_vin_exp(:,5) = dlmread('D3um.csv');
ipa_vin_exp(:,6) = dlmread('D5um.csv');
% time steps (seconds)
dt = 1e-2;    % time step
nits = 12e5;    % number of time steps
nout = 1200;
dt_out = (nits/nout)*dt;
tout = [0:dt_out:nout*dt_out];   % time points for collecting data
nk1r = 1;    % dissociation of ipa_vin
k1r_array = [.0024];
kd_ipa_vin = [ 2e-9 ];  % kd of ipa_vin
nk2f = 1;     %  association of lov_vin
k2f_array = [1e3];%
nkd = 1;
kd_array = [73e-9];  % equilibrium dissociation constant for lov binding to
vin
for i_k1r = 1:nk1r
    for i_k2f = 1:nk2f
        for i_kd = 1:nkd
 % initialize kinetic parameters
        k1r = k1r_array(i_k1r);
        k1f = k1r/kd_ipa_vin;
        kd  = kd_array(i_kd);
for Lov / vin binding
        k2f = k2f_array(i_k2f);
k2r = k2f * kd;

from vin

  % Rate constant for ipa_vin breakup
% Rate constant for ipa_vin association
  % equilibrium dissociation constant
  % rate constant for
  % rate constant for Lov dissociating
% loop over various Lov_tot concentrations
sum_sqr_deviation_pol = 0;
for l = 1:nlov
    %ipa_tot = ipa_tot_start;
    %vin_tot = vin_tot_start;
in _out arrays
ntp = floor(time_point/dt_out)+1;  % how to find time_point
fraction_bound = ipa_vin_out(ntp,l)/ipa_tot;
pol = fraction_bound;
lov_tot = lov_array(l);
lov = lov_tot;
% calculating ipa_vin
a = 1;
b = -(ipa_tot + vin_tot + kd_ipa_vin);
c = vin_tot * ipa_tot;
ipa_vin = (-b - (b*b - 4*a*c)^0.5)/2;
ipa_vin_start = ipa_vin;
lov_vin = lov_vin_start;
vin = vin_tot - ipa_vin - lov_vin;
ipa = ipa_tot - ipa_vin;
% output vectors
ipa_vin_out(:,l) = zeros(1,nout+1);
```

```
pol_out(:,l) = zeros(1,nout+1);
% fill output variable for time = 0
n = 1;
ipa_vin_out(n,l) = ipa_vin;
fraction_bound = ipa_vin/ipa_tot;
%dpol = pol_min - pol_max;
pol_out(n,l) = fraction_bound;
% numerical integration of kinetic reaction
for i = 1:nits
    delta_k1f = dt*k1f*ipa*vin;
    delta_k1r = dt*k1r*ipa_vin;
    delta_k2f = dt*k2f*lov*vin;
    delta_k2r = dt*k2r*lov_vin;
    lov_vin = lov_vin + delta_k2f - delta_k2r;
    ipa_vin = ipa_vin + delta_k1f - delta_k1r;
    vin = vin + delta_k1r + delta_k2r - delta_k1f - delta_k2f;
    ipa = ipa + delta_k1r - delta_k1f;
    lov = lov + delta_k2r - delta_k2f;
    % collect data for making plots
    if mod (i,floor(nits/nout)) == 0
        n = n+1;
        %lov_vin_out(n,l) = lov_vin;
        ipa_vin_out(n,l) = ipa_vin;
        fraction_bound = ipa_vin/ipa_tot;
        %dpol = pol_min - pol_max;
        pol_out(n,l) = fraction_bound;
end end

% compare rmsd of experiment to simulated curves
for nt = 1:ntout_exp;
    time_point = tout_exp(nt);
square_deviation_pol = ((pol) - ipa_vin_exp(nt,l))^2;
                sum_sqr_deviation_pol = sum_sqr_deviation_pol +
square_deviation_pol;
end end


        rmsd = (sum_sqr_deviation_pol/(nlov*ntout_exp))^0.5;
        end % i_kd
    end  % i_k2f
end % i_k1r

plot(tout, pol_out(:,1),'b', tout_exp, ipa_vin_exp(:,1),'ob');
hold on
plot(tout, pol_out(:,2),'r', tout_exp, ipa_vin_exp(:,2),'or');
plot(tout, pol_out(:,3),'g', tout_exp, ipa_vin_exp(:,3),'og');
plot(tout, pol_out(:,4),'c', tout_exp, ipa_vin_exp(:,4),'oc');
plot(tout, pol_out(:,5),'m', tout_exp, ipa_vin_exp(:,5),'om');
plot(tout, pol_out(:,6),'k', tout_exp, ipa_vin_exp(:,6),'ok');
hold off
xlabel('time (seconds)');
ylabel('fraction bound');
```

2.7 Tables

LOV-ipaA Construct	k_{on} ($M^{-1}s^{-1}$)	K_{off} (s^{-1})	K_D (nM)
Dark mimetic (C450A)	$1.3 \pm 0.3 \times 10^{3}$	$8.0 \pm 0.7 \times 10^{-5}$	64 ± 9.5
WT dark	$1.4 \pm 0.1 \times 10^{3}$	$9.6 \pm 0.1 \times 10^{-5}$	69 ± 0.5
Lit mimetic (I532E A536E)	$2.9 \pm 2.0 \times 10^{4}$	$8.5 \pm 0.3 \times 10^{-5}$	3.0 ± 1.0
WT blue light	$1.3 \pm 0.3 \times 10^{4}$	$4.5 \pm 1.2 \times 10^{-5}$	3.5 ± 0.5
L514K L531E dark	$4.5 \pm 1.5 \times 10^{2}$	$1.1 \pm 0.4 \times 10^{-4}$	245 ± 5.0
L514K L531E blue light	$2.5 \pm 0.1 \times 10^{3}$	$1.3 \pm 0.1 \times 10^{-5}$	5.0 ± 0.1
G528A N538E dark	$1.8 \pm 0.3 \times 10^{2}$	$8.2 \pm 0.2 \times 10^{-5}$	475 ± 75
G528A N538E blue light	$8.0 \pm 2.0 \times 10^{2}$	$1.2 \pm 0.1 \times 10^{-4}$	160 ± 40

Table 2.1 Kinetic rates of LOV-ipaA binding vinculin D1

LOV-SsrA construct	Lit K_D (μM)	Dark K_D (μM)
Initial LOV-SsrA Fusion	$3.1 \pm 0.6 \times 10^{-2}$	$5.7 \pm 0.5 \times 10^{-2}$
LOV-SsrA G528A & N538E	$1.2 \pm 0.2 \times 10^{-1}$	$5.7 \pm 0.8 \times 10^{-1}$
LOV-SsrAC	$1.2 \pm 0.1 \times 10^{-1}$	$9.0 \pm 0.7 \times 10^{-1}$
LOV-SsrAM	40 ± 2	156 ± 6
LOV-SsrAN	12.6 ± 0.3	49 ± 3
WT SsrA- AANDENY	NA*	$3.5 \pm 0.2 \times 10^{-2}$
SsrA- AANDEAY	NA*	$3.8 \pm 0.8 \times 10^{-2}$
SsrA- AANDINY	NA*	1.9 ± 0.6

*TAMRA-SsrA binding to SspB only measured in dark. Error shown is standard error to binding curve fit

Table 2.2 Lit and dark binding affinities for LOV-SsrA constructs to SspB

LOV-ipaA Mutant	K_D (mM)	Error (mM)
IpaA Peptide	7.8	1.3
Dark Mimetic (C450A)	54	12
Lit Mimetic (A532E I536E)	7.4	2.3
L514K L531E Dark Mimetic	116	34
L514K L531E Lit Mimetic	8.1	2.1
G528A N538E Dark Mimetic	315	84
G528A N538E Lit Mimetic	12	2.1

Table 2.3 Apparent binding affinity of LOV-ipaA to full-length vinculin

2.8 Figures

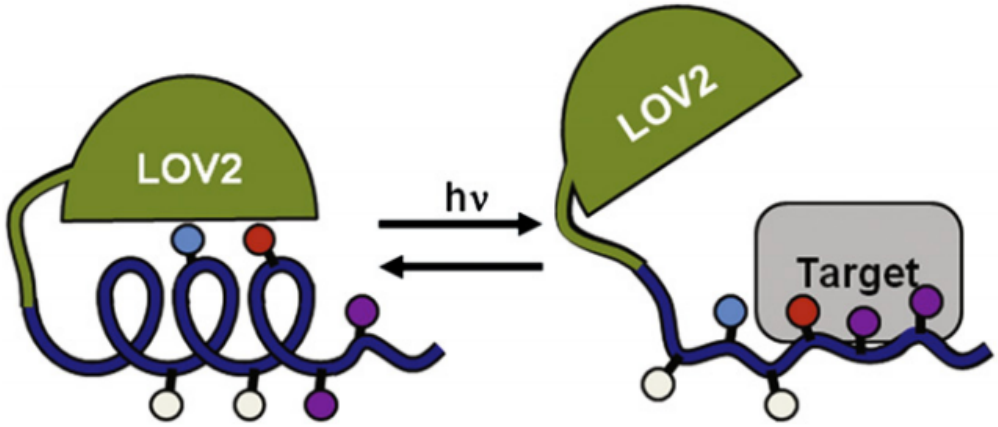


Figure 2.1 Caged peptide schematic. Photoswitches are designed as sequence chimeras between the AsLOV2 Jα helix and the peptide to be caged. Residues that are important to AsLOV2-Jα interactions (cyan), important to peptide-target interaction (purple), important to both interactions (red), and residues that are important to neither interaction (white) are identified and mutated accordingly. Irradiation unfolds the Jα helix, and the peptide can bind its target.

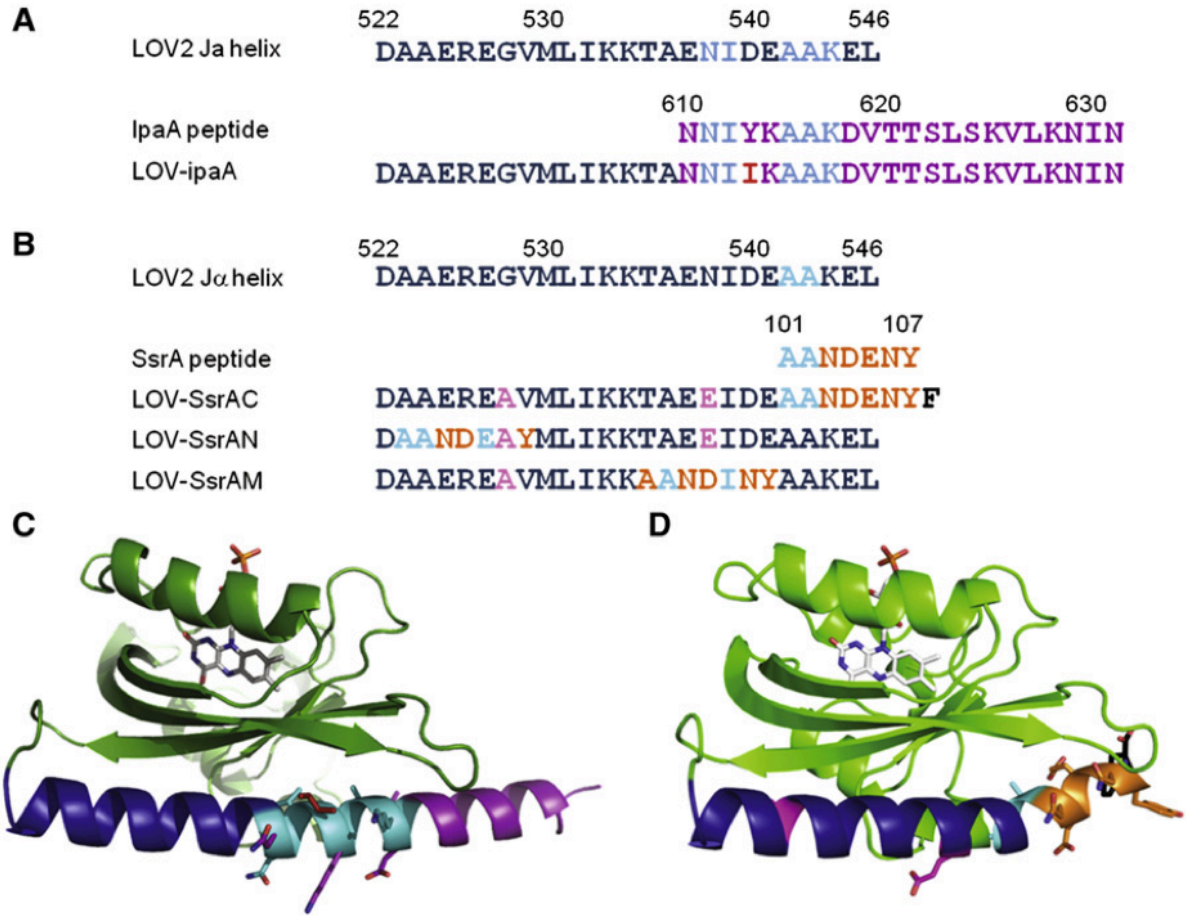


Figure 2.2 Sequence alignment of AsLOV2 and peptides. (A) Sequence alignment of AsLOV2-Ja, ipaA, and LOV-ipaA. Ja sequence (blue), ipaA sequence (purple), chimera sequence (cyan), and designed residues (red) are indicated. (B) Sequence alignment of AsLOV2-Ja helix, SsrA peptide, and three LOV-SsrA designs—LOV- SsrAC, LOV-SsrAN, and LOV-SsrAM. Ja helix sequence (blue), SsrA sequence (orange), chimera sequence (cyan), designed positions (black), and helix-stabilizing mutations (pink) are indicated. (C) Model of LOV-ipaA with residues colored as in (A). Residues N538, I539, A542, and A543 K544 (cyan) as well as residues N537, K541, D545, V546 (purple), and I540 (red) are shown as sticks. (D) Model of LOV-SsrAC with residues colored as in (B). Residues A528 and E538 (pink) as well as residues A542 and A543 (cyan) and N544, D545, E546, N547, and Y548 (orange) are shown as sticks.

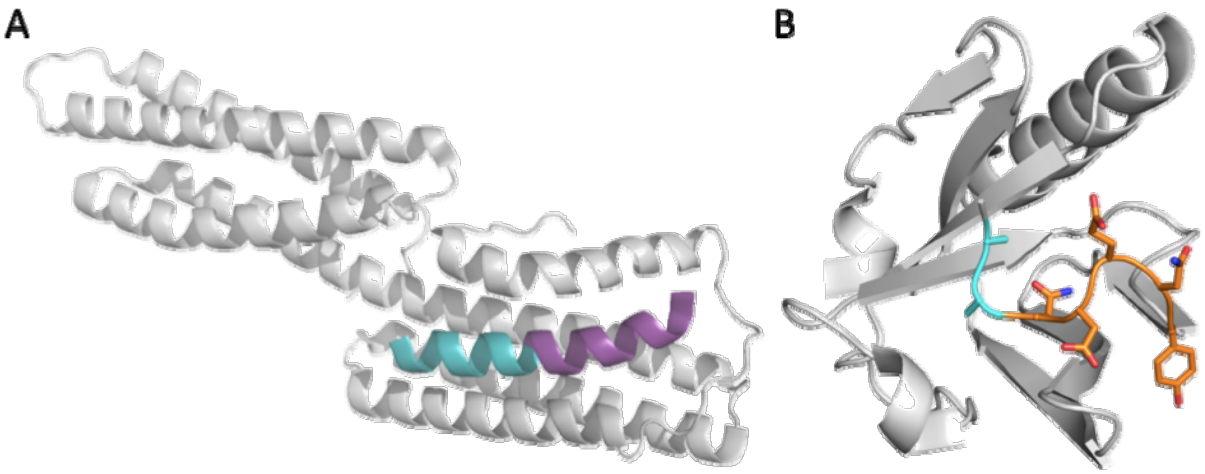


Figure 2.3 Crystal structures of peptides binding native partners. (A) Structure of ipaA (cyan and purple) binding the vinculinD1 subdomain (white). PDB code: 2GWW (B) Crystal structure of SsrA (orange/cyan) bound to SspB (grey). PDB code: 1OU8.

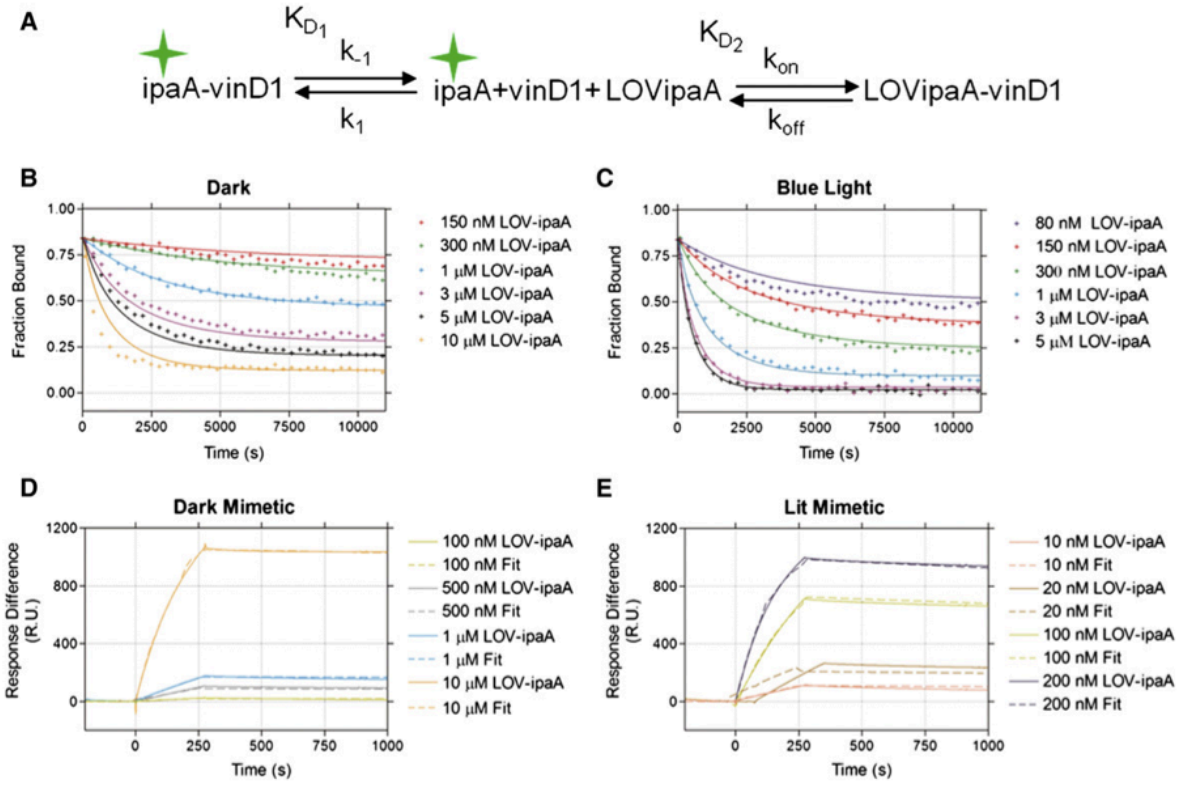


Figure 2.4 LOV-ipaA binding assays. (A) Schematic of fluorescence polarization competition assay is shown. TAMRA-labeled ipaA (ipaA*) is bound to vinculin D1 subdomain (vinD1). Vinculin dissociates from the complex with rates k_{-1} and k_1 and binding affinity K_{D1}. LOV-ipaA (LOVipaA) binds vinculin with rates k_{on} and k_{off} and affinity K_{D2}. Fluorescence polarization decreases as the fraction of TAMRA-ipaA bound to vinculin decreases. (B and C) Fraction of TAMRA-ipaA bound to vinculin over time with varying concentrations of LOV-ipaA titrated in the dark and (C) under blue light. (D and E) Surface plasmon resonance measurements and first-order binding fit of LOV-ipaA L514K L531E C450A pseudodark and (E) LOV-ipaA L514K L531E I532E A536E pseudolit mutants.

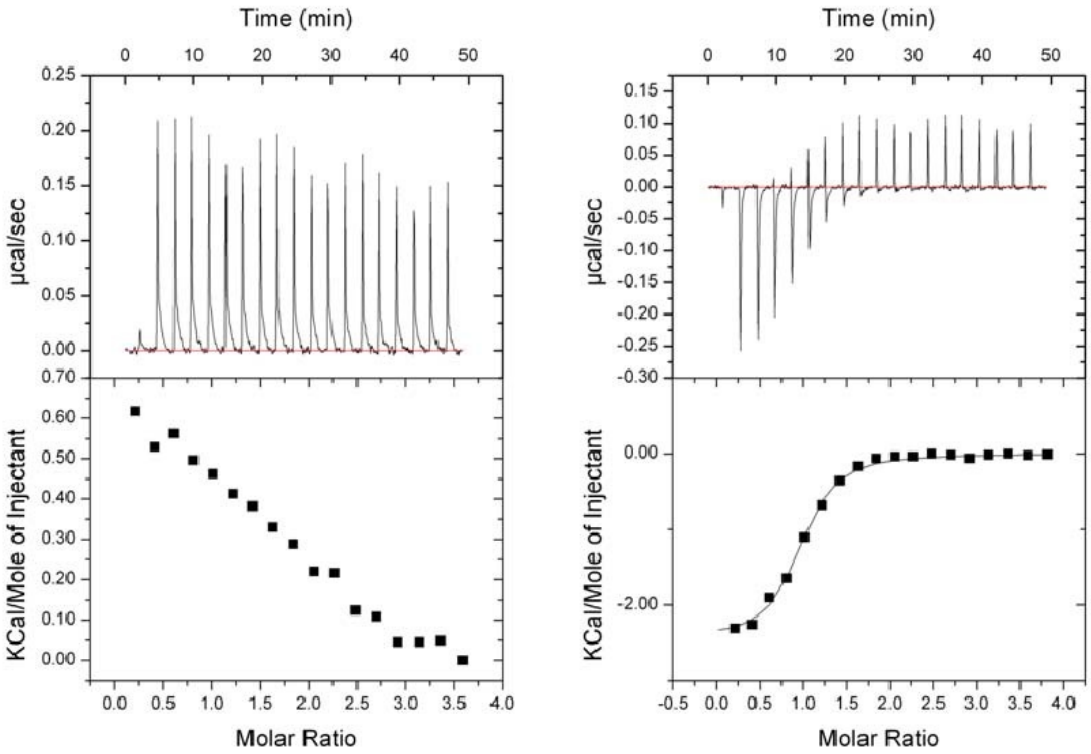


Figure 2.5 LOV-ipaA binding to vinculinD1 measured by ITC. Isothermal Titration Calorimetry experiment of LOV-ipaA L623A mutation binding the vinculinD1 subdomain in the darks state (C450A, left) or lit state (I532E A536E, right). Both experiments were done with identical amounts and concentrations of protein. Binding affinity of dark state was too weak to fit. Binding affinity of lit state was determined to be 2.4 µM.

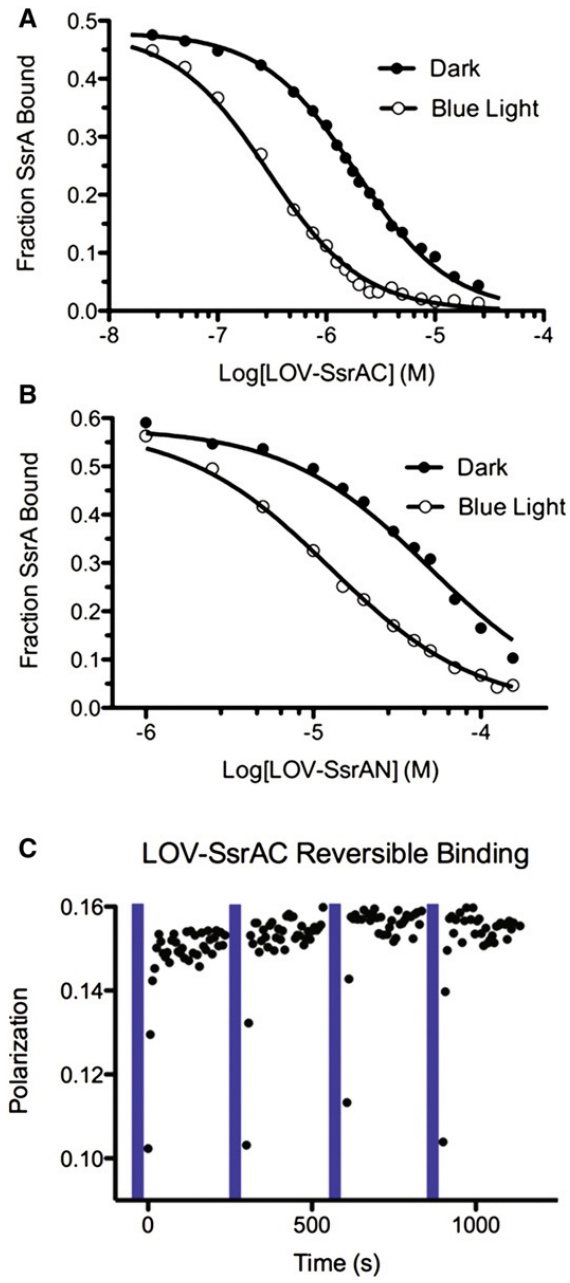


Figure 2.6 LOV-SsrA binding to SspB. (A and B) Competitive binding assay of LOV-SsrAC (A) or LOV-SsrAN (B) to an equilibrium solution of SspB and 5(6)TAMRA-SsrA. 5(6)TAMRA-SsrA becomes unbound as LOV-SsrA competes for SspB binding. Binding to SspB was measured immediately after illumination with blue light (open circles) and after return to dark state (closed circles). (C) Reversible binding of LOV-SsrAC to SspB. A single titration point from the fluorescence polarization competition assay was repeatedly irradiated with blue light (blue bar = 60 s) and reversion to dark-state equilibrium was monitored by polarization.

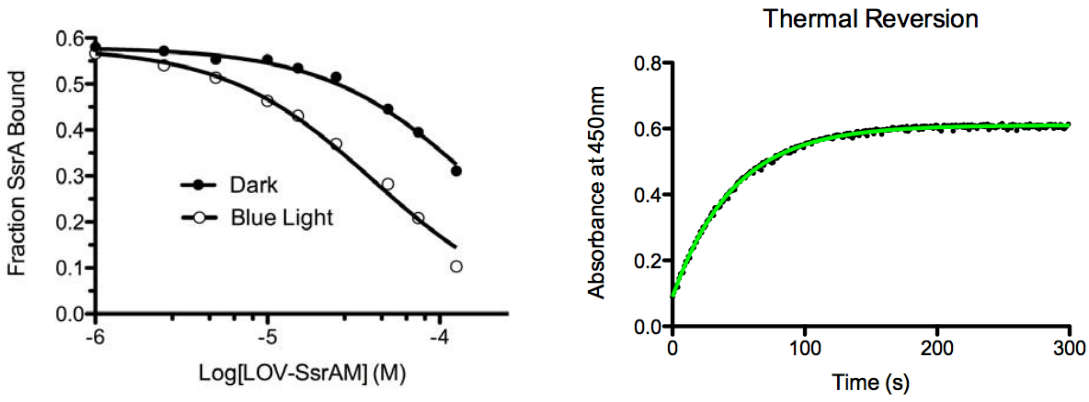


Figure 2.7 LOV-SsrAM binding to SspB and kinetics of LOV-SsrAC. (Left) Binding of LOV-SsrAM to SspB as measured through the fluorescence polarization competition assay described in Experimental Procedures. Binding to SspB was determined to be 40 μM immediately after irradiation with blue light (open circles) and 160 μM after relaxation to the dark state (closed circles). (Right) Thermal reversion of LOV-SsrAC in presence and absence of SspB. The recovery of absorbance at 450 nm by LOV-SsrAC (5μM) after irradiation with blue light was measured (closed circles) and fit to a single exponential (green line). The photocycle half-life was determined to be 28.4 s.

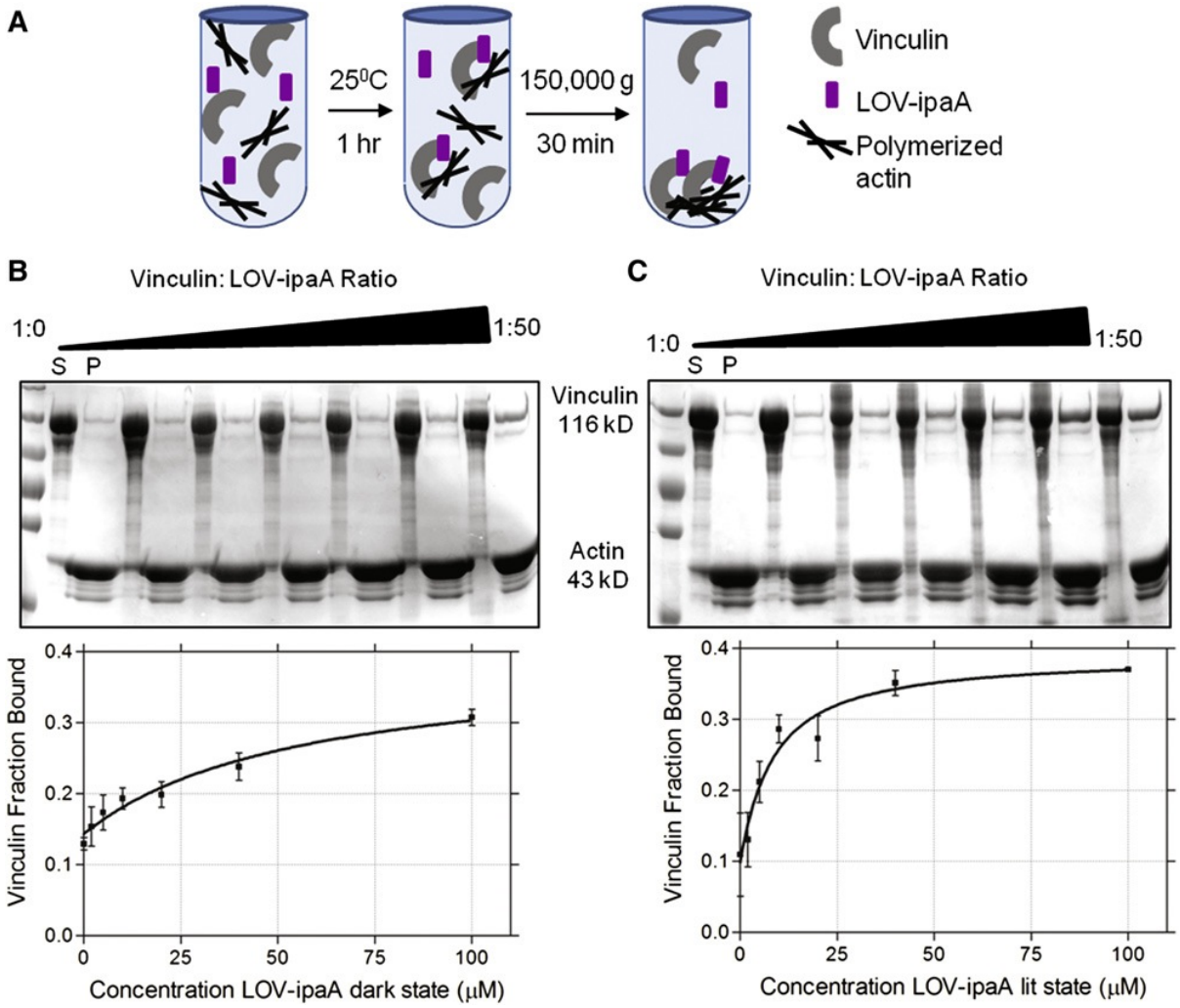


Figure 2.8 LOV-ipaA bound to vinculin preferentially binds actin. (A) Full-length vinculin, LOV-ipaA, and polymerized actin are incubated 1 hr at room temperature. Vinculin that is bound to LOV-ipaA will bind polymerized actin. The mixture is centrifuged at 150,000 g, pelleting polymerized actin and all vinculin bound to it out of solution. (B and C) SDS-PAGE gel of LOV-ipaA C450A and (C) LOV-ipaA I532E A536E actin cosedimentation assay with vinculin. Molar ratios from 1:0 to 1:50 vinculin:LOV- ipaA were used. Supernatant (S) and pellet (P) fractions are shown side by side. Apparent binding affinity curves of fraction of vinculin bound to actin versus concentration of LOV-ipaA are plotted below. Error bars represent standard error of the mean.

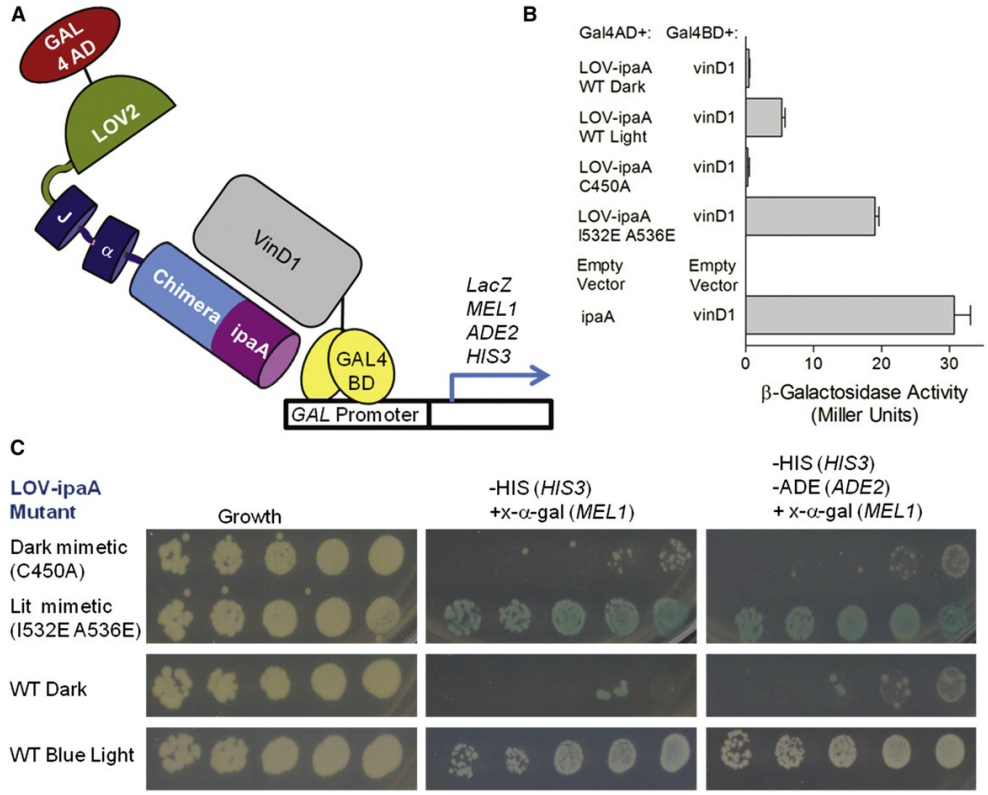


Figure 2.9 LOV-ipaA/VinculinD1 Yeast Two-Hybrid. (A) LOV-ipaA L623A is linked to the GAL4 activation domain (AD), whereas vinculin D1 (vinD1) is linked to GAL4 binding domain (BD). Irradiation with blue light brings AD-LOV-ipaA into proximity to BD-vinD1, allowing for GAL-induced expression of reporter genes LacZ, MEL1, HIS2, and ADE2. (B) LacZ expression is quantified. b-galactosidase activity of S. cerevisiae-mated strains containing BD and AD linked proteins, as specified, is shown. Error bars represent standard error of the mean. (C) S. cerevisiae-mated strains containing BD-vinD1 and AD-LOV-ipaA mutants, as indicated, are grown in dark or blue light conditions on SD plates. Difference in levels of transcription of MEL1, HIS3, and ADE2 in dark- versus lit-state conditions is seen.

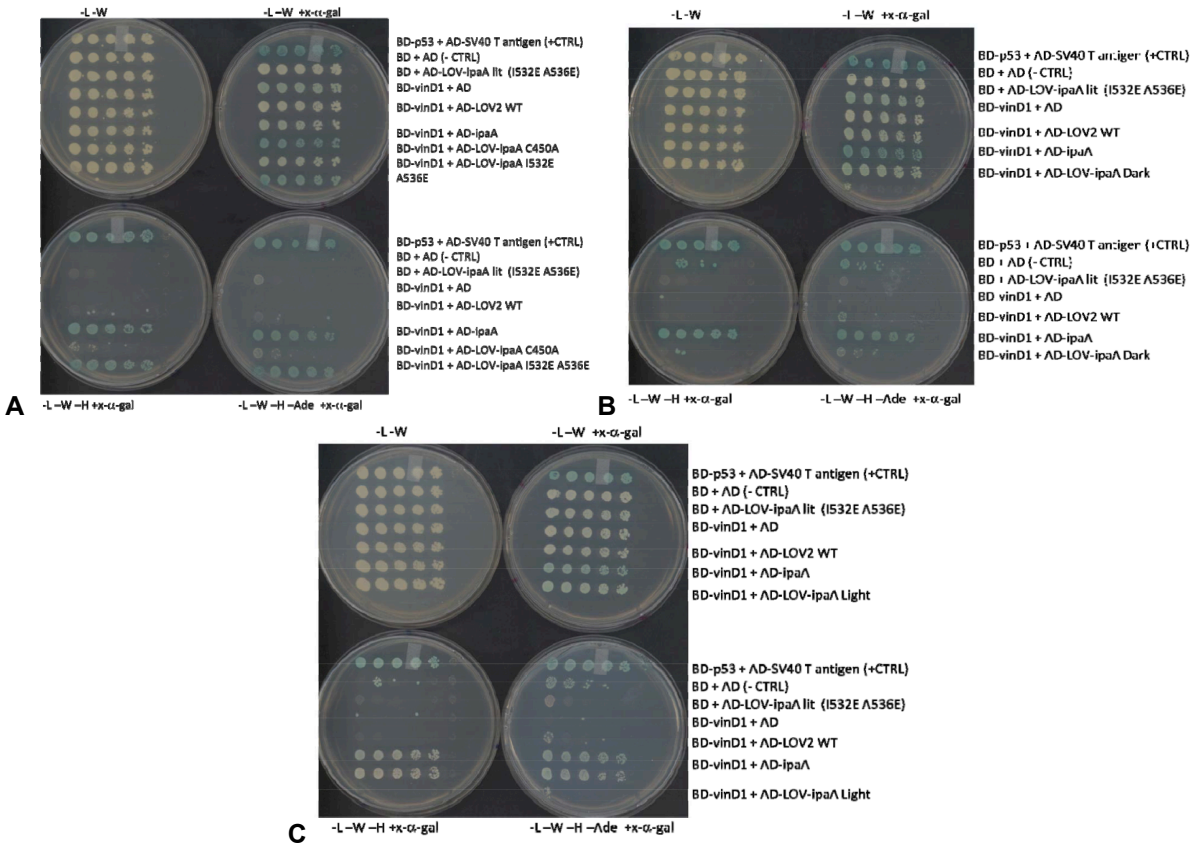


Figure 2.10 LOV-ipaA/VinculinD1 yeast two-hybrid controls and mimetics. (A) LOV-ipaA L623A Lit state and dark state mimetic yeast two-hybrid assay and controls. SD media conditions are indicated. S cerevisiae strain Y2Hgold containing the Gal4 activation domain (AD) linked to the construct indicated was mated with strain Y187 containing the Gal4 binding domain (BD) linked to the construct indicated. Mated colonies were serially diluted (1:5), replica plated, and grown for 3 days at 30° C. (B) LOV-ipaA L623A WT yeast two-hybrid assay and controls (Dark state plates are on top, blue light plates on bottom). SD media conditions are indicated. S cerevisiae strain Y2Hgold containing the Gal4 activation domain (AD) linked to the construct indicated was mated with strain Y187 containing the Gal4 binding domain (BD) linked to the construct indicated. Mated colonies were serially diluted (1:5), replica plated, and grown for 3 days at 30° C in either blue light or in the dark, as indicated. (C) LOV-ipaA WT lit state and dark state mimetic yeast two-hybrid assay and controls. SD media conditions are indicated. S cerevisiae strain Y2Hgold containing the Gal4 activation domain (AD) linked to the construct indicated was mated with strain Y187 containing the Gal4 binding domain (BD) linked to the construct indicated. Mated colonies were serially diluted (1:5), replica plated, and grown for 3 days at 30° C.

Chapter 3

Engineering an improved light-induced dimer (iLID) for controlling the localization and activity of signaling proteins[1]

This work was published in the Proceedings of the National Academy of Sciences, Volume 112, Issue 1, p112-117 on 6 January 2015. Gurkan Guntas constructed the DNA libraries and performed the phage display and ELISA selections. Ryan Hallett designed the computational library and performed all in vitro experiments. Seth Zimmerman performed all of the mammalian cell culture experiments. Tishan Williams solved the structure of iLID and Hayretin Yumerefendi performed proof of principle phage display work with AsLOV2 domains. This work was guided and funded by Brian Kuhlman and James Bear.

3.1 Overview

The discovery of light-inducible protein–protein interactions has allowed for the spatial and temporal control of a variety of biological processes. To be effective, a photodimerizer should have several characteristics: it should show a large change in binding affinity upon light stimulation, it should not cross-react with other molecules in the cell, and it should be easily used in a variety of organisms to recruit proteins of interest to each other. To create a switch that meets these criteria we have embedded the bacterial SsrA peptide in the C-terminal helix of a naturally occurring photoswitch, the light-oxygen-voltage 2 (LOV2) domain from *Avena sativa*. In the dark the SsrA peptide is sterically blocked from binding its

[1]This chapter previously appeared as an article in PNAS. The original citation is as follows: 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).

natural binding partner, SspB. When activated with blue light, the C-terminal helix of the LOV2 domain undocks from the protein, allowing the SsrA peptide to bind SspB. Without optimization, the switch exhibited a twofold change in binding affinity for SspB with light stimulation. Here, we describe the use of computational protein design, phage display, and high-throughput binding assays to create an improved light inducible dimer (iLID) that changes its affinity for SspB by over 50-fold with light stimulation. A crystal structure of iLID shows a critical interaction between the surface of the LOV2 domain and a phenylalanine engineered to more tightly pin the SsrA peptide against the LOV2 domain in the dark. We demonstrate the functional utility of the switch through light-mediated subcellular localization in mammalian cell culture and reversible control of small GTPase signaling.

3.2 Introduction

Inducible protein dimers are complexes that form when a specific stimulus is provided – for instance the protein FRB binds to the protein FKBP12 in the presence of rapamycin [125]. They are powerful research tools because with genetic engineering they can be used to localize and activate proteins in living systems [101,102,126,127]. For example, by fusing one half of an inducible dimer to a DNA binding domain and the other half to a transcription activation domain, transcription of target genes can be initiated by providing the stimulus that induces dimerization. Chemically induced dimers, such as FRB and FKBP12, have been used to control a wide variety of biological processes but are limited by irreversibility and lack of spatial control within a cell. For this reason, there is strong interest in light-inducible dimers that can be activated in specific regions of a cell or an organism using light in a reversible manner.

Several light-inducible dimers are currently available and have been used to control signaling pathways in living cells. In almost all cases, the dimers are derived from naturally occurring photoactivable systems. The most widely used pair thus far is cryptochrome 2

(Cry2) and CIB1 from *Arabidopsis thaliana*. The Cry2/CIB1 pair shows blue light induced dimerization in both yeast and mammalian cell culture. Association occurs on a sub-second time scale and reversion within 10 minutes [126]. The mechanism of light-activation is not fully understood and it has recently been shown that Cry2 oligomerizes into large clusters under blue light in addition to associating with CIB1. This could be a drawback for applications that require precise stoichiometry, but the oligomerization itself has been utilized for control of protein activation [45]. Another dimerization pair is phytochrome B (PhyB) and PIF, also from *Arabidopsis thaliana.* PhyB and PIF interact after irradiation with red light and dissociate with exposure to far-red light [101]. This system requires a chromophore, phycocyanobilin (PCB), that is not naturally present in many organisms, including mammals [55,59]. The tunable light-controlled interacting protein tags (TULIPs) make use of the blue light sensing light-oxygen-voltage (LOV) domain and an engineered PDZ domain [128]. Sub-cellular localization has been shown with TULIPs in both yeast and mammalian cells. However, the presence of a PDZ binding peptide and PDZ domain in the system could lead to cross talk with endogenous signaling pathways. The FKF1/GIGANTEA (GI) heterodimer pair from *Arabidopsis thaliana* also relies on a LOV domain for light activated binding. This pair dimerizes in response to blue light in mammalian cells, associating within a few minutes of activation and dissociating on the order of a few hours [102]. However, GI is a very large protein and the switching power of the pair is sensitive to expression levels.

An ideal light-inducible dimer would be small and modular, associate quickly in presence of signal and quickly dissociate in its absence, and be fully orthogonal to the organisms they are to be used in. To fulfill these criteria, we chose to use the LOV2 domain of phototropin 1 from *Avena sativa* as the photoactive element of our light-inducible dimer. In its native setting, the AsLOV2 domain senses blue light and activates a C-terminal kinase domain. AsLOV2 can be genetically encoded and its chromophore, flavin mononucleotide, is abundant in most organisms. It is monomeric in dark state and remains so under activating

blue light. The structure of AsLOV2 has been determined, making it amenable to structure guided engineering. AsLOV2 is made up of a core per-arnt-sim (PAS) fold with flanking α-helicies on the N- and C-termini [104]. Upon blue light irradiation, a conserved cysteine residue in the core of AsLOV2 becomes covalently bound to the flavin cofactor, structural rearrangement is passed along through the PAS fold, and ultimately the flanking helices unfold from the PAS core [19,24]. The covalent bond breaks with a half-life of 30-50 seconds, and the helices refold to their dark state conformation [15,129]. To reengineer the AsLOV2 domain to be part of a light-inducible complex, we incorporated a naturally occurring binding element within the protein; 7 residues of the *E. coli* SsrA peptide that bind SspB, a 13kD adaptor protein also from *E. coli*. Both SsrA and SspB have been structurally characterized and the SsrA peptide shares sequence identity with the Jα helix of AsLOV2 [110,130,131]. We previously showed that incorporation of SsrA into the Jα helix of AsLOV2 led to steric occlusion of SspB binding in the dark and uncaging with blue light irradiation, yielding a light-inducible heterodimer pair [132]. Our original light inducible dimer (oLID), while containing some of the characteristics desirable for a versatile tool, did not show large changes in binding affinity with light stimulation. Original fusions of AsLOV2 and SsrA yielded a two-fold change in affinity for SspB. This was improved to an eight-fold switch by incorporating mutations known to stabilize the Jα helix in the dark as well as including a C-terminal phenylalanine predicted by molecular modeling to further hold the Jα helix against the PAS domain in the dark. However, even in its final form, dimerization in the dark at moderate protein concentrations prevented large light induced phenotypic changes *in vivo*.

In order to create a generalizable, versatile, and more powerful light-inducible dimer, we sought to improve oLID with a novel combination of computational library design and phage display screening. Using the Rosetta macromolecular software suite, we ranked point mutations within AsLOV2 and created a library of mutations with the goal of improving the dynamic range of oLID, specifically reducing dark state binding. We then used phage

display and ELISA-based binding assays to screen for mutations that both weakened dark state binding and responded to light activation. Here, we describe two improved light-inducible (iLID) pairs that we identified with this approach, iLID nano, which switches from 130 nM under blue light to 4.7 μM in the dark, and iLID micro, which switches from 800 nM under blue light to 47 μM in the dark. iLID nano and iLID micro both co-localize under blue light within seconds, revert to dark state within minutes, and can be activated sub-cellularly in mammalian cell culture. To demonstrate functional utility in cells, we show light-dependent control of GTPase activity and the cytoskeleton through localization of the DH/PH domains of ITSN and Tiam1.

3.3 Materials and Methods

3.3.1 Design and Construction of Computationally Directed Library

The pmut_scan protocol in the Rosetta molecular modeling suite was used to evaluate all possible point mutations in AsLOV2 (PDB: 2v0u). The set of all possible point mutations was filtered to include only mutations that decreased the Rosetta score or increased it by less than 1 REU (Rosetta Energy Unit). For a few positions on the PAS β sheet mutations with ΔΔG's of +2 REU's were allowed. The list was then filtered to only include mutations that were within 6Å of the Jα helix. The resulting library contained 743 mutations at 49 positions (additional mutations were allowed due to degenerate primer design). The Rosetta-biased point-mutant library was constructed using a comprehensive mutagenesis protocol [133]. Mutagenic oligos with degenerate codons flanked by 15-18 bases were pooled. 200 picomoles of the oligo pool were incubated with 20 units of T4 Polynucleotide kinase (NEB) for 1 hour at 37°C in the presence of 1x T4 PNK buffer. The enzyme was heat-inactivated at 65°C for 20 minutes. Single-stranded template uracil-DNA encoding TorA-LovSsrA-pIII_C-terminal domain was amplified and purified as described [134].

One microgram of template DNA (0.75 picomoles) were combined with 2-fold excess 5'-phosphorylated oligo pool. Template-oligo annealing followed by polymerase extension and ligation of the mutagenic strand was performed as described [133]. Template DNA was degraded with 5 units of Uracil DNA Glycosylase (NEB) and 2 units of Exonuclease III (NEB) before desalting and electroporation into SS320 cells. The experimental library size was 5 x 10^7, which most likely samples all possible point-mutants and a large fraction of double-mutants.

3.3.2 Construction of the shuffled library

Mutagenic oligos encoding substitutions suggested by the primary ELISA screens of the point-mutant library were pooled and assembled using a gene-assembly protocol. Gene-SOE PCR was performed with Q5 DNA Polymerase (NEB), 1 μM total oligo (31 total oligos) and the following cycling conditions: 20 cycles of (98°C 10s, 52°C 30s, 72°C 10s) in 50 μL. One μL of the assembled product was PCR-amplified using outside primers for restriction cloning. Two μg's of linearized and gel purified pFNOM6-tat-pIII plasmid was ligated to 1 μg of shuffled-lovssrA library insert (4:1 insert:vector ratio). Ligated DNA was ethanol precipitated and electroporated into SS320 cells. The number of total transformants was 2 x 10^8, which partially covers the designed sequence space (1.2 x 10^8).

3.3.3 Phage Display Selection Against SspB

The LOV library was expressed as an N-terminal fusion of the phage pIII coat protein using the tat secretion pathway, as export via the DsbA signal peptide pathway was not possible (Figure 3.1). All libraries were subjected to four rounds of panning prior to ELISA screening. Maxisorp 96-well plates were coated overnight at 4°C with 100 μg/mL of 5 micrograms/mL His-MBP-SspB fusion in the presence of 50 mM sodium bicarbonate pH 9.6 buffer. Coated wells were washed with PBS-0.1% (v/v) Tween (PBST) and blocked at room temperature for 2 hours with 200 μL of PBST-BSA (5 mg/mL). Blocked wells are washed

with PBST and incubated with 10^{11} – 10^{12} library phages for 1 hour under a collimated blue LED array (1.2 mW*cm^{-2} at 450nm). Wells were washed 10-times with PBST while keeping the plate under blue light. Plates were moved into dark and incubated in the dark for 10 minutes. Dissociating phages were collected with PBST buffer. Early log SS320 cells (OD_{600}: 0.2-0.5) were infected with eluted phage and infected cells were grown in 25 mL 2xTY media supplemented with 100 µg/mL of ampicillin. When cells reach early log phase, M13KO7 helper phage was added with 20:1 multiplicity of infection. Following a 20-minute incubation at 37°C without shaking, 0.1 mM IPTG and 5 µM flavin mononucleotide (FMN) were added and the culture was moved to a 30°C shaker covered with foil to ensure dark conditions. After 1.5 hour, kanamycin was added to 25 µg/mL and the culture was grown overnight. The following day, the phage was PEG-precipitated and quantified using A_{268} values [134]. Gene pool post-Round-4 selection was PCR amplified and cloned into a previously modified pET21b vector that introduces an N-terminal FLAG epitope. The ligation reaction was directly transformed into BL21(DE3)pLysS cells for protein expression and ELISA screening.

3.3.4 Photoswitch Evaluation by Soluble Protein ELISA

BL21(DE3)pLysS cells carrying pET21b_FLAG_LovssrA clones were plated on LB/Amp plates. Next day, 96-well growth blocks containing 500 µL LB medium supplemented with 25 µg/mL chloramphenicol, 100 µg/mL ampicillin, and 0.6 mM IPTG were inoculated with individual colonies and grown overnight in the dark at 30°C. One well per plate was reserved for the parent for comparison. 200 µl of cells were lysed in the dark with 10 µl Popculture reagent and 1 unit of Benzonase nuclease for 15 minutes. Lysates were centrifuged and supernatant diluted 5-10 fold in PBST with 1mM DTT was used for the assays. Duplicate Maxisorp plates (96-well or 384-well) were coated with His-MBP-SspB as described above. Plates were blocked with PBS-3% (m/v) BSA and incubated with the

lysate supernatant for 1 hour in the dark and under blue-light (1.2 mW*cm^{-2} at 450nm, collimated blue LED array). Plates were washed 5 times with PBST and incubated for 30 minutes with 1:20,000 diluted anti-FLAG antibody HRP conjugate (Sigma A8592). Plates were washed 3 times with PBST and twice with PBS. 100 µl TMB was added and color was developed for 15 minutes. The reaction was quenched with 100 µl 0.5 M sulfuric acid before measuring A_{450}. LovssrA variants for the positive clones were PCR amplified using 4 µl cells as template and sequenced.

3.3.5 Growth and Purification of Recombinant Proteins

All expressed and purified proteins were cloned into the pQE-80L protein expression vector. All LOV clones contained an N-terminal 6x(His) tag and SspB clones contained an N-terminal 6x(His)-MBP-TEV site tag. Sequence-verified clones were transformed into BL21(DE3)pLysS *E.Coli* cells. Cultures were grown to an OD of 0.6 at 37°C, then induced with 333 mM IPTG and moved to 18°C for 16 hours. Cell pellets were resuspended in loading buffer (50 mM sodium phosphate pH 7.5, 500 mM NaCl, 20 mM Imidazole, 100 µM PMSF). Resuspended cells were sonicated and lysate spun down for 30 minutes at 20,000 rpm. Proteins of interest were isolated via Ni^{2+} affinity chromatography using HisTrap HP columns (GE), and eluted with elution buffer (50mM phosphate pH 7.5, 500 mM NaCl, and 500 mM imidazole). Proteins expressed as MBP fusions were then cut with TEV, dialyzed overnight into loading buffer, and re-run over the HisTrap HP column, collecting the tag-less protein. Finally, size exclusion chromatography with a Superdex 75 column (GE) was used for a final cleanup step and buffer exchange to final binding buffer, PBS (10 mM dibasic sodium phosphate, 1.8 mM monobasic potassium phosphate, 137 mM NaCl, 2.7 mM KCl). Protein concentration was determined via BCA protein assay (Pierce). Expression (*E.coli* & mammalian) constructs of iLID and SspB variants used can be found in Table 3.6 and will be available on Addgene.

3.3.5 Fluorescence Polarization Binding Assays

Fluorescence polarization experiments were conducted with a Jobin Yvon Horiba FluoroMax3 spectrofluorometer in a 1 cm quartz cuvette. Polarization of the TAMARA-SsrA peptide was measured through excitation at 555 nm and emission at 584 nm. For direct binding experiments between SspB and TAMARA-SsrA, starting peptide concentration was 25 nM in PBS. For competitive binding assays, 25 nM TAMARA-SsrA and 40 nM SspB (in PBS) were incubated prior to titrating LOV fusions. At each titration, the sample was irradiated with 6.0 $mW*cm^{-2}$ blue light for 1 minute, turned off and a lit state polarization was measured. After 5 minutes in the dark, a second measurement was taken, representing the dark state population.

3.3.6 Thermal Reversion Assay

Purified iLID protein was dialyzed against 4L of PBS buffer twice before measuring reversion kinetics. Protein samples were prepared at 10μM and allowed to equilibrate to room temperature in a 1cm quartz cuvette. Samples were then irradiated with blue light (6.0 $mW*cm^{-2}$ at 450nm, collimated blue LED array) for 30 seconds. Recovery of absorbance at 450nm was then measured every 0.1 seconds for 5 minutes.

3.3.7 Crystallization and Structural Determination of iLID

Initial crystals of iLid were grown using hanging drop vapor diffusion. Drops consisted of 2 μL well solution (800 mM lithium chloride, 100 mM TRIS:HCl pH 8.5, 32% (m/v) PEG 4000) and 1 μL iLid protein (10 mg/mL in 100mM ammonium acetate). Initial conditions were optimized by microseeding. Final crystal drops were 2 μL well solution, 1 μL iLid (10 mg/mL), 0.5 μL microseed solution. Crystals grew to maximum size in 3 days. X-ray diffraction data was collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Data was indexed with XDS and scaled with Scala [135–137]. The structure of iLid was determined

using PAS fold residues N414-D515 of PDB id 2v0u as a molecular replacement model in Phaser [138]. Phenix and Coot were used to iteratively refine the final structure [139,140].

3.3.8 Mammalian Cell Culture and Transfection

IA32 mouse fibroblast cells were cultured at 37° C and 5% CO_2 in DMEM supplemented with 10% FBS (HyClone), 100U/ml penicillin, 100 μg/ml streptomycin and 292 μg/ml L-glutamine. Cells were transiently transfected using NanoJuice (EMD Millipore) as recommended by manufacturer.

3.3.9 Mammalian Cell Localization/GEF Microscopy and Image Analysis

IA32 mouse fibroblasts were transfected in 6 well tissue culture plates with equal amounts of 2 vectors, each containing a component of the switch (ex. 0.5 μg pll7.0 Venus-iLID-CAAX : 0.5 μg pll7.0 tgRFPt –Nano; See Table 3.6 for vector information). 24 hr after transfection and 24 hr before image aquisition, cells were plated on No. 0 glass bottomed dishes (MatTek) coated with a 10 μg/ml solution of Fibronectin. Image acquisition was performed using a FluoView FV1000 scanning confocal inverted microscope equipped with a 1.30 N.A. 40 x oil immersion objective, a Hamamatsu PMT, environmental chamber (Precision Plastics) and controlled by Fluoview software (Olympus, Version 3.1b). A 25 mW argon laser provided the 488 nm and 515 nm laser lines while a 15 mW diode laser provided the 559 nm laser line. The environmental chamber was used to continue culture at 37° C and 5 (v/v) CO_2 throughout image acquisition. The software Time Controller was used to set a timeline of image acquisition and LOV domain excitation within a predefined region of interest. During periods of excitation, the ROIs were continuously scanned with 488 nm light except during the time it took to aquire an image (< 3 sec). 1% power of the 488 nm laser was used for all LOV domain excitation. During all experiments, images were acquired every 10 sec before and after excitation while image acquisition times varied during each

type of excitation, due to the difference in excitation area. For spot activation and whole cell activation LOV excitation and image acquisition required 7.6 sec 14.7 sec respectively. All images were analyzed using FIJI software [141]. For spot activation, the average tgRFPt fluorescence intensity within the activated ROI was measured for each frame and the background subtracted. The values for each image set were normalized to the average of the first three frames before excitation. For half-cell activation, ROIs of similar tgRFPt fluorescence were manually chosen within the activated area and outside the activated area. Mean values were acquired for each frame and background subtracted. The reported values are a ratio of the fluorescence inside to outside the activated area. For whole cell activation, analysis was automated. For each frame, ROIs were produced by applying the default Auto Threshold function to the Venus channel(Mito ROI). To produce a Cyto ROI the Mito ROI was expanded by 10 pixels and then removed from the new ROI. The average tgRFPt intensity was then measured for each ROI and frame. The ratio of mitochondrial to cytoplasmic signal was then calculated for each frame like so: (Mito – Cyto)/Cyto. The cytoplasmic signal was first subtracted from the mitochondrial signal to remove any signal that was contributed to the mitochondrial ROI from the cytoplasm above and below the mitochondria.

3.3.10 Mammalian Cloning and Vector Information

All mammalian constructs were cloned into pLL 7.0 lentiviral vectors. Therefore, a CMV promoter drives expression. Mammalian constructs were assembled through PCR amplification and subsequent restriction enzyme digestion and ligation. The constructs were verified by sequencing. See Table 3.6 for more information and Addgene numbers.

3.4 Results

3.4.1 Computational library screened by phage display and protein ELISA

In our original light-inducible dimer (oLID) there was appreciable binding in the dark, ~800 nM. Previous studies suggest that a major source of dark state "leakiness" comes from transient undocking and unfolding of AsLOV2's Jα-helix, and mutations that increase the intrinsic stability of the Jα helix have proven successful in reducing dark state activity of light-activated AsLOV2 switches [90,128]. We sought to extend this approach and use molecular modeling combined with high-throughput screening to find mutations throughout the domain that stabilize the docked state of the Jα-helix. The pmut_scan protocol in Rosetta [142] was used to model and score all possible point mutations within 6 Å of the Jα helix using the crystal structure of AsLOV2 in its dark state (PDB: 2v0u). All mutations that were predicted to be stabilizing or neutral (within 1 kcal / mole of the wild type residue) were considered potential mutations that might stabilize the dark state and were used as the basis for creating a directed library. The final point mutant library consisted of 743 mutations at 49 positions (Table 3.4), and used the oLID sequence as the starting sequence.

The oLID library was fused with the phage pIII coat protein and an N-terminal Tat (twin-arginine translocation) secretion sequence for display on the surface of phage [143]. The Tat secretion pathway was used because it maintains proteins in the folded state during secretion, which allows them to remain bound to cofactors. In this case this was critical, as FMN binding is needed for LOV domain activity (Figure 3.1, Figure 3.2B). We screened the library by rounds of positive and negative selection; binding under blue light to immobilized MBP-SspB and subsequent elution in the dark (Figure 3.2C). After 4 rounds of lit and dark screening, individual sequences were tested by soluble protein ELISA (Figure 3.2D). We ranked mutations by change in dynamic range, defined as ELISA signal under blue light/ELISA signal in darkness. Single point mutant LOV variants had minimal improvement

to overall switching. The top mutations were pooled and recombined with each other to yield a recombined library. Screening of the recombined library was repeated as before (Figure 3.2E). The top 60 sequences had dynamic ranges ranging from 2.4 to 35.4 (Figure 3.3). Four sequences with the highest dynamic ranges were chosen for further characterization (Figure 3.4).

3.4.2 *In vitro* Characterization

We measured the binding affinity of our four top sequences for SspB in the dark and after irradiation with blue light using a competitive fluorescence polarization binding assay. All four had improved dynamic ranges compared to the parent clone oLID (Figure 3.5A,B) and switched roughly over the same affinity range, ~100nM to <5μM (Figure 3.6). Clones D5 and F2 had the highest dynamic range of the clones tested, 43-fold and 36-fold changes respectively. D5 switched from 77 nM under blue light to 3.4 μM in the dark and F2 switched from 132 nM under blue light to 4.7 μM in the dark. As dark state activity can be the limiting factor in the usefulness of light-inducible tools, we chose to move forward with clone F2, naming it our improved Light Inducible Dimer, iLID (Figure 3.5C,E)

The concentration threshold required for activity varies for different signaling proteins within a cell. Having tools that switch over different ranges of affinities may allow a wider range of targets to be controlled. In order to create an alternate affinity range for iLID as well as retain our improved switching, we made a point mutant, R73Q to SspB, reducing the peptide/protein affinity to 900 nM (Figure 3.7). In the context of SspB R73Q, iLID has an affinity of 800 nM for SspB R73Q under blue light and a dark state affinity of 47 μM, which is a 58-fold change in binding affinity (Figure 3.5D). Thus, two sets of tools are available for light-inducible heterodimerization, iLID nano (makes use of WT SspB) and iLID micro (makes use of SspB R73Q). Comparisons of the affinity ranges of these tools to their predecessors can be seen in Figure 3.5F.

3.4.3 Structural Characterization of iLID

iLID contains 10 mutations when compared to oLID. Four of the mutations are clustered in the hinge loop that connects the PAS domain to the Jα helix, four are located in the Jα helix, and two are in the PAS domain. To investigate how these mutations are improving the dynamic range of the switch, we solved the crystal structure of iLID in its dark state to a resolution of 2.05 Å (Figure 3.8). In general, there are very few structural perturbations when comparing iLID to AsLOV2. Within the core PAS domain (residues 413-516), the Cα RMSD between iLID and AsLOV2 is 0.42 Å, and the hinge loop showed almost no movement in the protein backbone (0.23 Å Cα RMSD)(Figure 3.8B). Several of the side chains in the hinge region are forming interactions that may be important for the stabilizing the dark state of the switch. R519 is packed against W491 from the PAS domain. The isoleucine replacing valine 520 is buried against sidechains from the Jα helix and the PAS domain and packs against a valine at position 493 that is a leucine in AsLOV2. The last residue in the hinge, position 522, is an aspartic acid in AsLOV2 and a glycine in iLID. This is the last residue before the Jα helix and has a positive phi angle, phi = 49.0°, psi = -126°. Aspartic acids are observed in this region of the ramachandran plot, but glycine is by far the most common amino acid in this region of the plot. In a set of high-resolution crystal structure, 47% of the residues with phi near 50° and psi near -130° are glycine, while only 7% are aspartic acids. This suggests that there may be some strain that is relieved when the aspartic acid is mutated to a glycine.

The largest conformational change in the iLID structure was in the N-terminal residues and the A'α helix. The first two resolved residues, T406 and T407, point toward the Jα helix instead of the Dα helix as in AsLOV2 and the A'α helix is shifted by 3 Å (Figure 3.8C). A smaller, ~1 Å movement, is seen in the last 12 residues of the Jα helix toward the A'α helix, allowing for tighter packing between the two. The slight unwinding of the A'α helix allows for the terminal nitrogen of the helix to be capped by E546 of the SsrA sequence.

Additionally, the side chain of E409 forms a hydrogen bond with the tyrosine from the SsrA sequence.

Finally, the structure of iLID revealed the location and conformation of the caged SsrA peptide. The designed C-terminal phenylalanine, F549 fits into a hydrophobic pocket on the surface of the PAS domain made up of I417, I428, F429 and Y508 (Figure 3.8D). SsrA makes three hydrogen bonds with the rest of the LOV domain, all within the A'α helix. The carboxyl group of D545 interacts with the hydroxyl of T406, the carboxyl group of E546 with the backbone amide of L407, and the hydroxyl of Y548 with the carboxyl of E409. In comparison with AsLOV2, F549 occupies the same area as the C-terminal L546. Due to extra residues in iLID, the last turn of the Jα helix buckles out and wraps back in to make this placement of F549 possible. iLID atomic coordinates and structure factors have been deposited to the Protein Data Bank (PDB ID code 4WF0).

3.4.4 Reversion and paired down iLID

Based on the structure of iLID, we chose 6 mutations to revert to determine their role or necessity in the improved switching of iLID (Figure 3.9, 3.10, Table 3.2). Reversion of R519 and D522, led to a lower dynamic range and weaker caging, which is apparent as tighter affinities under blue light and in the dark. Removal of F537 had a similar effect on dynamic range, but weakened blue-light and dark state affinities for SspB to the micromolar range. The remaining three reversions, Y502Q, H521R, and C530M, had little effect on the dynamic range. The double reversion mutant Y502Q/H521R, binds about 3 fold tighter in both states. In addition to keeping the improved dynamic range, C530M has very little effect on the affinity range or reversion kinetics of the switch and can be removed without consequence (Figure 3.11).

3.4.5 Subcellular recruitment

As oLID was previously untested in mammalian cell culture, we were interested in how well it would function and if the improvements of iLID, measured *in vitro,* would be apparent in cells. To test LID function in cells, we designed constructs by fusing the LIDs to the yellow fluorescent protein, Venus, and a well-characterized peptide sequence that would anchor the protein to specific subcellular localizations (CAAX = Cell Membrane, Mito = Mitochondria). Further, SspB (Nano) and SspB R73Q (Micro) were fused to the red fluorescent protein, TagRFP-T [144]. IA32 mouse fibroblasts [145] transiently expressing the LID protein pairs were imaged using a confocal microscope and activated with a 488 nm laser. To assay whole cell activation, cells expressing TagRFP-T-SspB and the Mito anchored LID pairs were imaged during a time course of intermittent whole cell 488 nm light stimulation. Cells expressing each pair of switches showed rapid recruitment of TagRFP-T to the mitochondria during stimulation followed by dissociation into the cytoplasm with the removal of stimulus (Figure 3.12A top row, Video S1). Recruitment for each pair was quantified and compared between switches by measuring the ratio of Mitochondrial/Cytoplasmic TagRFP-T fluorescence intensity throughout time (Figure 3.12B). To prevent bias, mitochondrial and cytoplasmic regions of interest (ROIs) were automatically identified for each frame using the Venus channel and the default Threshold function in FIJI software. The recruitment phase was fit to a single phase exponential and we determined that the iLID/Nano and iLID/Micro pairs showed a 4.4 and 4.3 fold change in comparison to a 2.4 fold change seen for oLID/Nano in this assay; each with a half-life < 30 s. Further, we hypothesized that the ratio of Mitochondrial/Cytoplasmic TagRFP-T fluorescence in the dark state would be representative of the dark state binding affinities observed *in vitro*. As expected, the iLID/Nano dark state ratio is higher than the ratio for iLID/Micro. However, while the *in vitro* evidence would suggest that oLID/Nano would have the tightest binding affinity in the dark, their dark state Mitochondrial/Cytoplasmic TagRFP-T fluorescence ratio is the lowest.

To test if the LID proteins could be used to recruit protein to a subcellular localization based on a region of interest (ROI), the CAAX fused LID pairs were expressed and imaged in fibroblasts (Figure 3.12A bottom row, Video S2). The optical sections obtained using the confocal were thick enough to capture light emitted from both apical and basal membrane bound proteins. This allowed us to stimulate a region of interest on these membranes and image the re-localization of TagRFP-T to these regions of the cell. Again, we imaged cells with intermittent stimulation at a ROI. During stimulation the TagRFP-T signal increased within the ROI while diminished outside the ROI. From this assay it is difficult to determine differences in dark state binding as we do not know what portion of TagRFP-T-SspB is membrane bound versus cytoplasmic. However, by quantifying the TagRFP-T fluorescence intensity in a relatively small ROI (260 um^2) the cytoplasmic concentration of TagRFP-T-SspB should remain constant before and after 488 nm light stimulation. Therefore by calculating pre:post stimulation ratio of TagRFP-T fluorescence intensity within the ROI we determined the relative fold change of binding in the light and dark. These values correlate with *in vitro* measured dynamic ranges (Figure 3.13). These experiments led us to realize that by anchoring the switching half of the dimer pair we obtained a relatively tight spatial resolution of activation. Combined with the minute time scales of association and dissociation we were able to exemplify these attributes of the switch by sequentially writing the letters i,L,I, and D on the membrane of a cell (Figure 3.12C, Video S3).

3.4.6 Light induced GTPase signaling through GEF localization

One area of interest for the use of light-inducible dimers is to control and study cell signaling. The Rho family of small GTPases (consisting of RhoA, Rac1, and Cdc42) are signaling proteins involved in many aspects of cell physiology including actin cytoskeleton remodeling. In general GTPase signaling is considered active when the GTPase is bound to GTP and inactive after the GTPase hydrolyzes GTP to GDP. Canonically, active Rac1 leads

to branched actin formation and lamellipodial morphology; active Cdc42 leads to bundled actin and filopodial morphology; and active RhoA leads to bundled actin and actin stress fiber formation. Other groups have found that small GTPases can be controlled through the localization of guanine nucleotide exchange factor (GEF) Dbl homology (DH) / pleckstrin-homology(PH) domains to the membrane [101,146]. At the membrane the DH/PH domains are colocalized with GTPases where they can aid in exchange of GDP for GTP, activating the GTPases. We hypothesized that the LID switches could be used to locally activate GTPase signaling, leading to a change in actin dynamics and a visible change in cell morphology (Figure 3.14A). Therefore we fused the DH/PH domains of Intersectin (ITSN, Cdc42 GEF) and Tiam1 (Rac GEF) to the N-terminus of the TagRFP-T-SspB constructs (Figure 3.14B). Fibroblasts were transiently transfected with each construct in combination with Venus-iLID-CAAX using a 1:1 ratio of DNA. As each construct is under the control of a CMV promoter we expect that on average each construct is expressed at equal levels. Two days later the cells were imaged and stimulated in a similar manner to the spot stimulation previously described. Here, stimulated areas were at the edge of the cell. As expected, the cells expressing the Tiam1 fusions, within minutes of stimulation, began to ruffle and formed a lamellipodia at the site of recruitment (Figure 3.15A, Video S4). Surprisingly, the cells expressing the ITSN fusions displayed a similar phenotype in the area of stimulation (Figure 3.15B, Video S5). While we would expect a filopodial morphology to form, our results may be explained in two ways; crosstalk between GTPase pathways that leads to an increase in Rac GTPase signaling or a direct catalytic effect on a Rac GTPase such as RhoG from the ITSN DH/PH [147,148]. In addition to forming lamellipodia, most cells displayed increased appearance of vesicles in the stimulated area consistent with the role of Cdc42 in endocytosis (Figure 3.15B arrowheads, Video S6).

3.5 Discussion

A novel aspect of our study was the use of high throughput screening and selection to improve the dynamic range of a photoactivatable protein switch. A specific challenge of engineering a switch is that the selection protocol should be sensitive to protein activity in both the on and off state. In our case, it was critical that the LOV domain be presented on the surface of the phage in a functional form. In phage display it is common to use signal sequences that direct a protein to the Sec-mediated translocation pathway, which transports proteins through membranes in an unfolded state. This was not appropriate for our application because we needed the LOV domain to remain folded during translocation so that it would remain bound to the flavin cofactor. Therefore, we made use of the Tat-secretion pathway that maintains proteins in a folded and cofactor-bound state during secretion. Once the LOV domain was on the surface of phage, we were able to find functional switches by selecting for binding under blue light, and then collecting phage that unbound when the light was turned off. The phage protocol was useful as a first filter in the search for better switches, but explicit screening in the light and dark with the ELISA-based binding assay allowed us to identify the top performing variants. The ELISA assay was particularly informative because the binding properties of single variants could be characterized side-by-side in the light and dark. This allowed us to detect small improvements in switching, which allowed us to select variants to include in the recombination library.

In creating the directed library we used molecular modeling to find mutations predicted to be compatible with the dark state of the switch. We used this approach because we were interested in stabilizing the closed state of the switch. Ideally modeling would be used to find mutations that are compatible with the dark and lit state. This was not possible because there is not detailed structural information about what happens to the PAS domain when it is photoactivated. It is informative to see where the final mutations were located in

iLID. The cluster of mutations in the hinge loop indicates that it has an important role in the docking and undocking of the Jα helix. In a multiple sequence alignment of the LOV domain, there is variability in the hinge loop suggesting that LOV-domain homologs may have different switching properties.

Previous studies have highlighted the role of the A'α helix in stabilizing the docked state of the Jα helix in the dark and releasing it in the light [24]. In the iLID structure the A'α helix rearranges to form several contacts with the Jα helix and the embedded SsrA sequence. Interestingly, the alternate conformation of the A'α helix is almost identical to that of the A'α helix of PA-Rac1, another photoactivatable switch based on the AsLOV2 domain (Figure 3.8C). A shared feature of the iLID and PA-Rac1 structures is the capping of the A'α helix. In the AsLOV2 structure, the A'α helix is capped by a water molecule, however in iLID, the carboxyl group of E546 in SsrA caps the helix and in PA-Rac1 it is capped by the carbonyl group of N595 in Rac1. These capping interactions in Pa-Rac1 and iLID may enhance cooperativity between folding of the A'α helix and caging, and therefore lead to effective switching.

Our goal in creating iLIDs was to create a genetic tool for studying biochemical and cellular processes. We showed that both iLID pairs can be used in mammalian cell culture to form dimers and recruit fusion proteins to sub-cellular structures. Furthermore, the differing affinity ranges of iLID nano and iLID micro measured *in vitro,* translate to different in cell characteristics, and can be used accordingly. We postulate that this inherent tuneability of the switch will lead to use in a broader spectrum of applications. Here, we were able to control local GTPase activity by driving GEF DH/PH domains to a localized spot on the membrane and we expect that the small, modular nature of iLID nano/micro will open up control of a much wider variety of cellular targets.

3.6 Supporting Information

3.6.1 Fasta Sequences of Phage Display Constructs

>Fasta of Tat-oLID-p3

MNNNDLFQASRRRFLAQLGGLTVAGMLGPSLLTPRRATAGSGEFLATTLERIEKNFVITDP
RLPDNPIIFASDSFLQLTEYSREEILGRNCRFLQGPETDRATVRKIRDAIDNQTEVTVQ
LINYTKSGKKFWNLFHLQPMRDQKGDVQYFIGVQLDGTEHVRDAAEREAVMLIKKTAEEI
DEAANDENYFSGLESRGPFSGDFDYEKMANANKGAMTENADENALQSDAKGKLDSVATD
YGAAIDGFIGDVSGLANGNGATGDFAGSNSQMAQVGDGDNSPLMNNFRQYLPSLPQSVE
CRPFVFSAGKPYEFSIDCDKINLFRGVFAFLLYVATFMYVFSTFANILRNKES

>Fasta of DsbA-oLID-p3

MKKIWLALAGLVLAFSASAAELAAAGSGEFLATTLERIEKNFVITDPRLPDNPIIFASDSFLQL
TEYSREEILGRNCRFLQGPETDRATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNLFHLQPM
RDQKGDVQYFIGVQLDGTEHVRDAAEREAVMLIKKTAEEIDEAANDENYFLESRGPFEGKP
IPNPLLGLDSTRPFVCEYQGQSSDLPQPPVNAGGGSGGGSGGGSEGGGSEGGGSEGGG
SEGGGSGGGSGSGDFDYEKMANANKGAMTENADENALQSDAKGKLDSVATDYGAAIDGF
IGDVSGLANGNGATGDFAGSNSQMAQVGDGDNSPLMNNFRQYLPSLPQSVECRPFVFSA
GKPYEFSIDCDKINLFRGVFAFLLYVATFMYVFSTFANILRNKES

3.7 Tables

Construct	Blue Light Affinity	Dark Affinity
Starting LOV-SsrA	0.031 ± 0.006 μM*	0.057 ± 0.005 μM*
LOV-SsrAC	0.120 ± 0.01 μM*	0.9 ± 0.07 μM*
iLID nano	0.132 ± 0.005 μM	4.7 ± 0.7 μM
iLID micro	0.8 ± 0.1 μM	47 ± 13 μM
SspB R73Q	-	0.9 ± 0.2 μM

*Previously published values

Table 3.1. Beginning and Improved iLID affinities

Construct	Blue Light Affinity	Dark Affinity	Fold Change
iLID nano	0.132 ± 0.005 μM	4.7 ± 0.7 μM	36
502/521	0.043 ± 0.005 μM	1.60 ± 0.05 μM	37
519	0.024 ± 0.004 μM	0.49 ± 0.05 μM	20
522	0.032 ± 0.01 μM	0.88 ± 0.04 μM	28
530	0.113 ± 0.004 μM	4.6 ± 0.6 μM	41
537	1.4 ± 0.1 μM	37.3 ± 0.7 μM	27

Table 3.2. Reversion Mutation Affinities

Construct	Blue Light Affinity	Dark Affinity
D5	77 ± 0.001 μM*	3.4 ± 0.02 μM*
C11	97 ± 0.001 μM*	1.7 ± 0.03μM*
H10	123 ± 0.001 μM*	2.0 ± 0.02 μM*
F2 (iLID nano)	0.132 ± 0.005 μM	4.7 ± 0.7 μM

*Error shown is standard error to binding curve fit

Table 3.3. Top ELISA Sequence Affinities

Residue #	Mutations	WT AA	Rosetta Predicted Library
404	13	L	ACDEGKMNQRSTY
406	1	T	S
407	17	T	ACDEFGIKLMNPQRSVY
408	3	L	MTV
411	5	I	KLQRV
413	11	K	ACDEHMNQRSW
414	3	N	AHS
415	1	F	H
417	4	I	VTLA
428	14	I	ACEGHKLMNQRSTV
429	4	F	HKRY
431	1	S	A
475	9	E	ACKMQRSTV
477	3	T	CAV
479	2	Q	EI
493	13	L	ACDEGHIMNQSTV
497	6	Q	AEHKMR
499	14	M	ACDEFHKNQRSTWY
500	9	R	ACEIKMQTV
502	15	Q	ACDEGHIKLMNRSTV
508	7	Y	ACFHKMR
510	4	I	VTLC
512	1	V	T
514	3	L	MQY
519	16	H	ACDEFIKMNPQRSVWY
520	13	V	ACDFHKLMNRSTY
521	13	R	ACEGHKMNQSTVY
522	15	D	ACEFGHKLMNQRSWY
523	15	A	CDEGIKLMNQRSTVY
524	11	A	CDEGKLMNQRS
525	10	E	AFHKLMQRWY
526	17	R	ACDEFGIKLMNQSTVWY
527	16	E	ACDGIKLMNQRSTVWY
529	15	V	ACDEFHIKMNQRTWY
530	18	M	ACDEFGHIKLNQRSTVWY
531	18	L	ACDEFGHIKMNQRSTVWY
532	5	I	AMSTV
533	10	K	AEHLMQRTVY
534	17	K	ACDEFGILMNQRSTVWY
535	18	T	ACDEFGHIKLMNQRSVWY
537	12	E	ACDIKLMNQRTV

539	3	I	LTV
540	16	D	ACEFGHKLMNQRSTWY
541	14	E	ACDIKLMNQRSTVY

Table 3.4. Rosetta predicted point mutations included in library

4WF0 Crystallography Statistics	
Wavelength (Å)	1.0
Resolution range (Å)	31.08 - 1.95 (2.02 - 1.95)
Space group	P 21 21 21
Unit cell	62.156 70.339 80.378 90 90 90
Total reflections	158984
Unique reflections	26276 (2574)
Multiplicity	6.0
Completeness (%)	99.8 (100.00)
Mean I/sigma(I)	9.63 (2.41)
Wilson B-factor	25.4
R-merge	0.088
R-meas	0.038
R-work	0.231 (0.356)
R-free	0.244 (0.398)
Number of non-hydrogen atoms	2711
Macromolecules	2392
Ligands	63
Water	256
Protein residues	288
RMS(bonds)	0.013
RMS(angles)	1.26
Ramachandran favored (%)	99
Ramachandran outliers (%)	0
Clashscore	6.36
Average B-factor	29.5
Macromolecules	28.4
Ligands	30.6
Solvent	39.8

Statistics for the highest-resolution shell are shown in parentheses.

Table 3.5. Data collection and refinement statistics

Construct	Description	Addgene ID
***E. coli* Expression**		
iLID	pQE-80L: iLID (C530M)	60408
SspB Nano	pQE-80L: MBP-SspB WT	60409
SspB Micro	pQE-80L: MBP-SspB R73Q	60410
Mammalian Cell		
Venus-iLID-CAAX	pLL7.0: Venus-iLID-CAAX (from KRas4B)	60411
Venus-oLID-CAAX	pLL7.0: Venus-oLID-CAAX (from KRas4B)	60412
Venus-iLID-Mito	pLL7.0: Venus-iLID-Mito (From ActA)	60413
Venus-oLID-Mito	pLL7.0: Venus-iLID-Mito (From ActA)	60414
tgRFPt-Nano	pLL7.0: tgRFPt-SSPB WT	60415
tgRFPt-Micro	pLL7.0: tgRFPt-SSPB R73Q	60416
Tiam DH/PH-tgRFPt-Nano	pLL7.0: mTiam1(64-437)-tgRFPt-SSPB WT	60417
Tiam DH/PH-tgRFPt-Micro	pLL7.0: mTiam1(64-437)-tgRFPt-SSPB R73Q	60418
ITSN DH/PH-tgRFPt-Nano	pLL7.0: hITSN1(1159-1509)-tgRFPt-SSPB WT	60419
ITSN DH/PH-tgRFPt-Micro	pLL7.0: hITSN1(1159-1509)-tgRFPt-SSPB R73Q	60420

Table 3.6. Construct Information

3.8 Figures

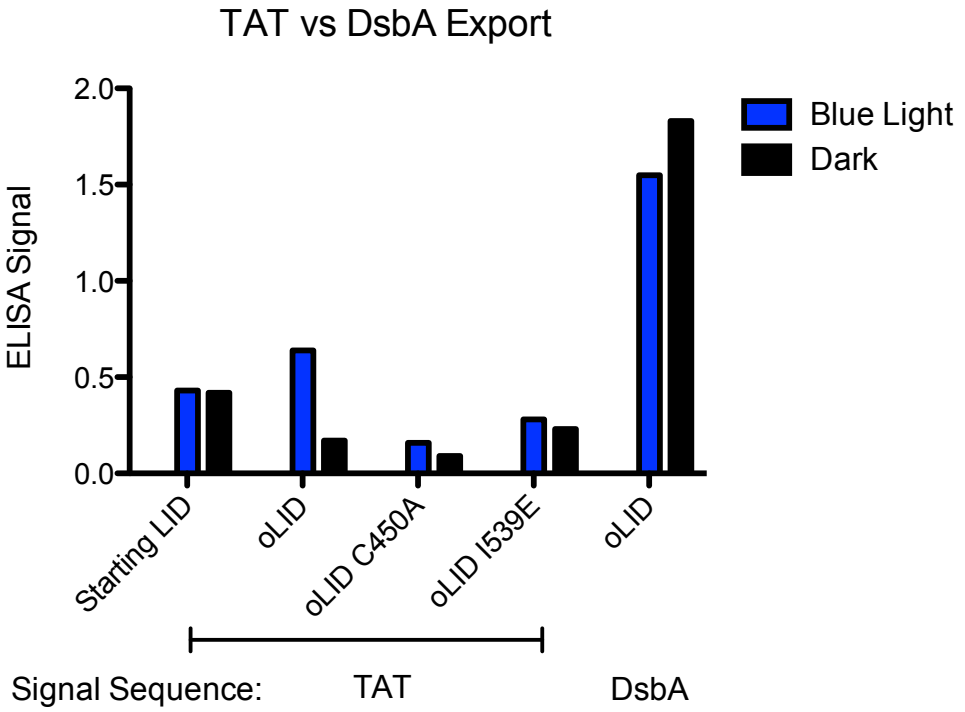


Figure 3.1 Binding of LIDs expressed via Tat and SRP pathways to SspB. Secretion of LID constructs via the TAT pathway yields functional protein that binds immobilized SspB in a light-dependent fashion. Using the DsbA signal sequence, the secreted protein does not bind SspB tighter under blue light than in darkness.

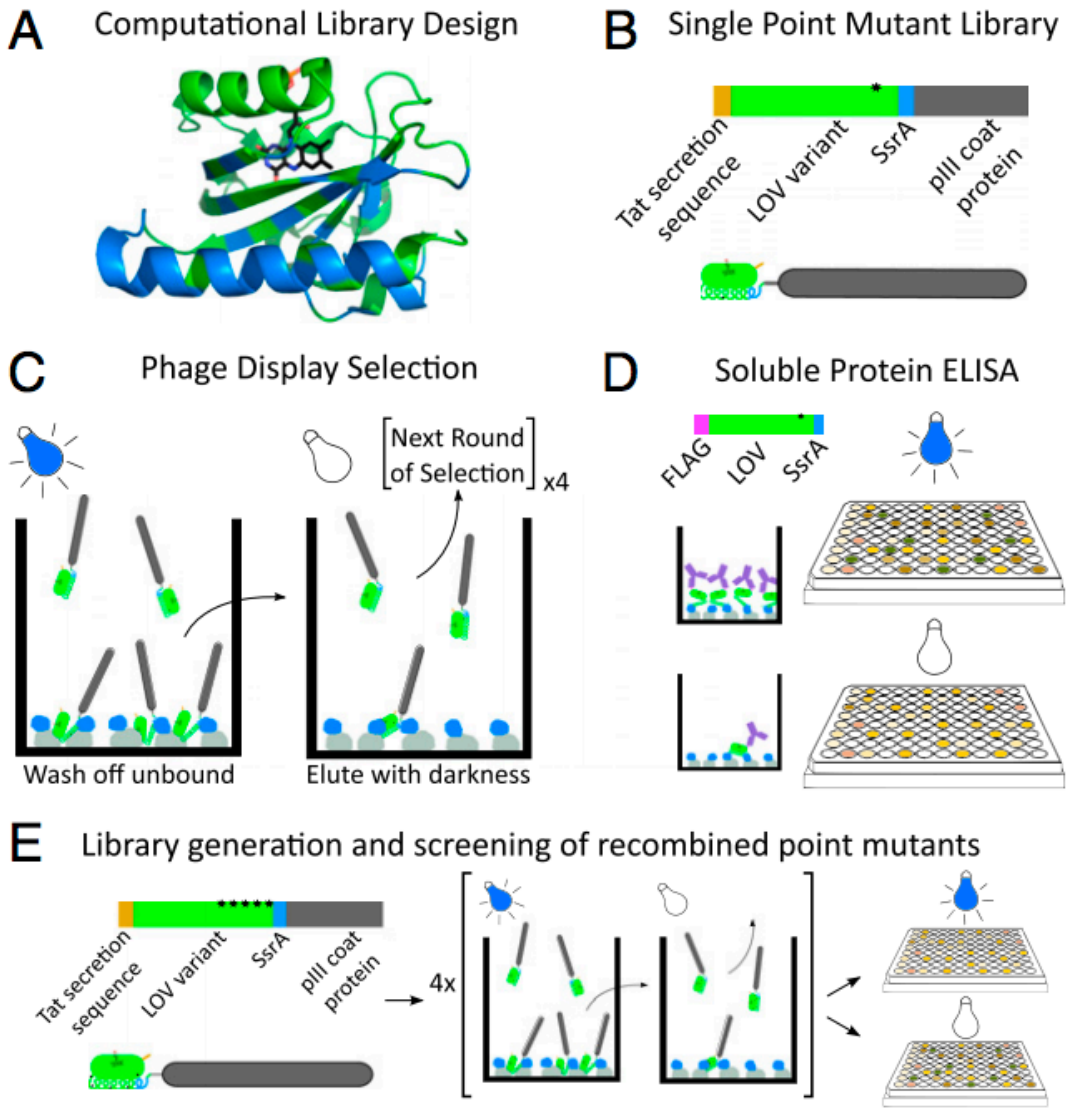


Figure 3.2 Schematic overview of selection and screening protocol for improved LOV variants. (*A*) Using Rosetta's pmut_scan, we generated a library of point mutations at 49 positions within the AsLOV2 domain (blue residues). (*B*) Phage-display construct illustration. (*C*) The phage library was added to plates coated with MBP-SspB (gray-blue) in the presence of blue light and washed. Plates were moved to the dark and eluted phages were collected. (*D*) Top single-mutation sequences were recloned as flag tag fusions and individually expressed. Binding of soluble protein to MBP-SspB coated plates was measured after exposure to blue light and sequestration in the dark by ELISA. (*E*) Mutations with the highest dynamic range were recombined to generate a new library of LOV variants, which was screened using the procedure shown in *C* and *D*.

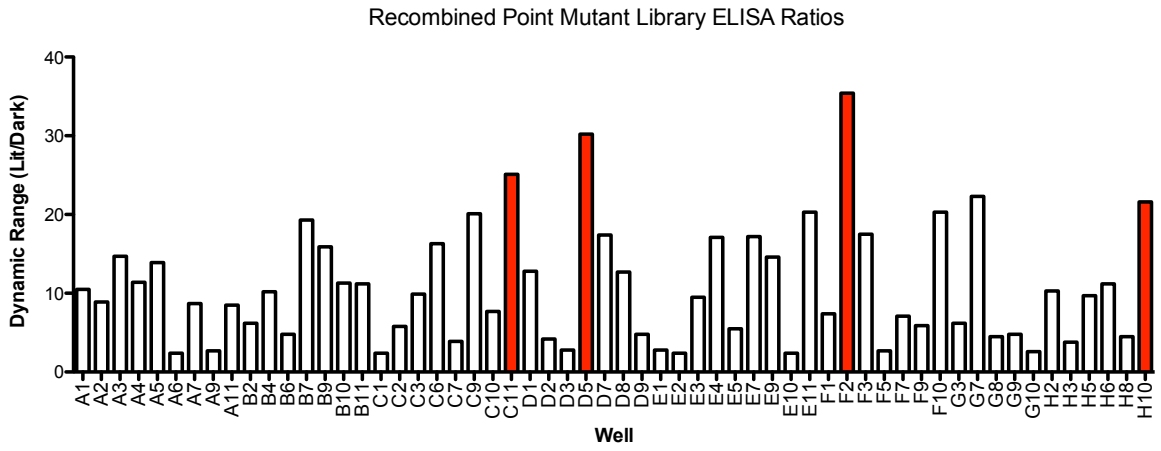


Figure 3.3 ELISA recombined point mutant ratios. Lit/Dark signal ratios for top sequences out of recombined library. C11, D5, H10, and F2 were selected to characterize (red columns).

```
oLID        458 TDRATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNLFHLQPMRDQK
C11         458 TDPATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNMFHLQPMRDYK
D5          458 TDPATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNVFHLQPMRDYK
H10         458 TDRATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNMFHLQPMRDYK
F2 (iLID)   458 TDRATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNVFHLQPMRDYK
                **  ********************************.********▲*

oLID        504 GDVQYFIGVQLDGTEHVRDAAEREAVMLIKKTAEEIDEAANDENYF
C11         504 GDVQYFIGVQLDGTERRRGAGEREAVILIKKTAFEIAEAANDENYF
D5          504 GDVQYFIGVQLDGTERLRGASEREAVILIKKTAFEIAEAANDENYF
H10         504 GDVQYFIGVQLDGTERRRGAGEREAVILIKKTAFEIAEAANDENYF
F2 (iLID)   504 GDVQYFIGVQLDGTERLHGAAEREAVCLIKKTAFQIAEAANDENYF
                ****************▲ .▲* ***** ******▲  *▲*********
```

Figure 3.4 Sequence alignment of top four sequences. Residues left unmutated are shown in black (*), mutations having similar character to original amino acid are shown in grey (□), all other mutations are shown in white. Positions that converged to a single mutation in all 4 sequences have been marked with a red triangle (▲).

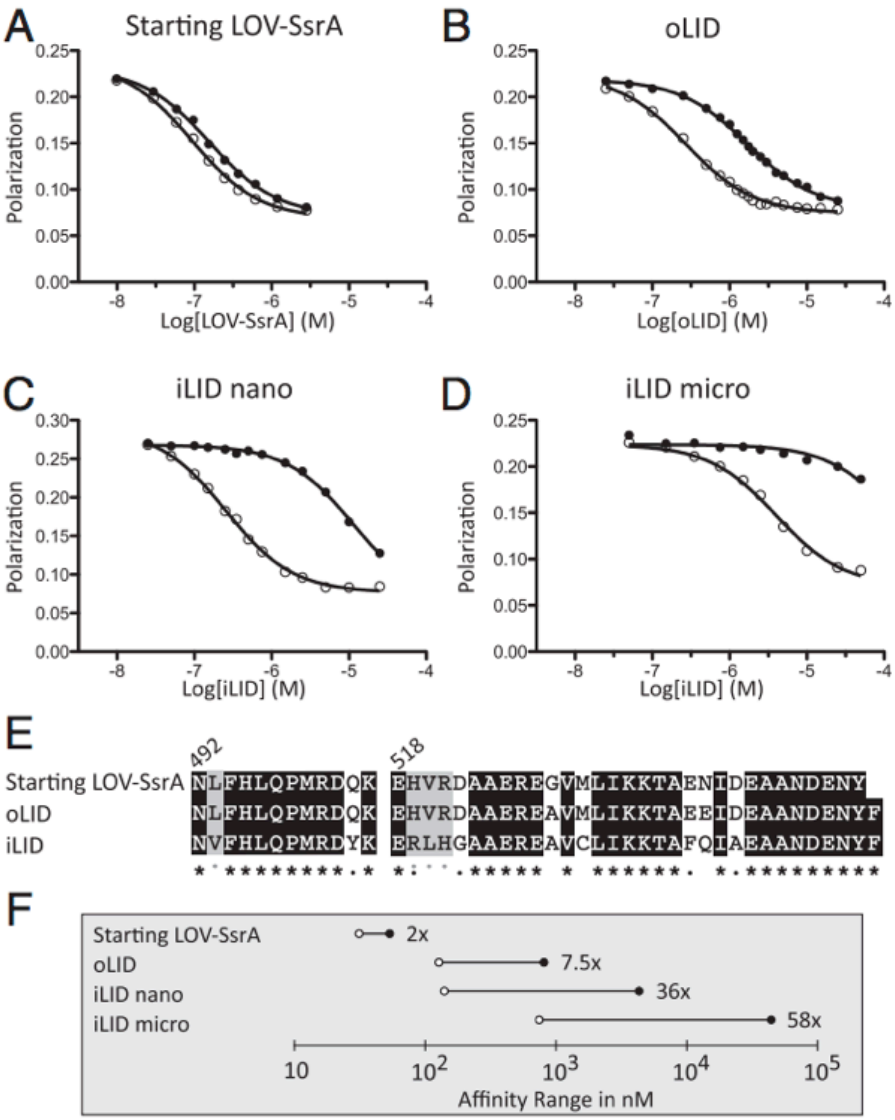


Figure 3.5 Characterization of iLID nano and iLID micro. Competitive fluorescence polarization binding assays measure affinities of heterodimerization under blue light (○) and in the dark (●). (A) Incorporation of SsrA into final turn of the AsLOV2 Jα helix. (B) Addition of helix-strengthening mutations (G528A, N538E) and C-terminal phenylalanine. (C) Top sequence from phage display and screening, iLID, has a 36-fold change in affinity for SspB because of light. (D) A mutation to SspB (R73Q) yields a heterodimer pair that switches over the micromolar range of affinities with a 58-fold change in affinity. (E) Sequence alignment of starting, oLID, and iLID. Black, unmutated; gray, similar amino acid mutation. Gray dots indicate iLID mutations that converged in all top four sequences. (F) Comparison of lit (○) and dark (●) affinities between the original heterodimer pairs and our two new pairs, iLID nano and iLID micro.

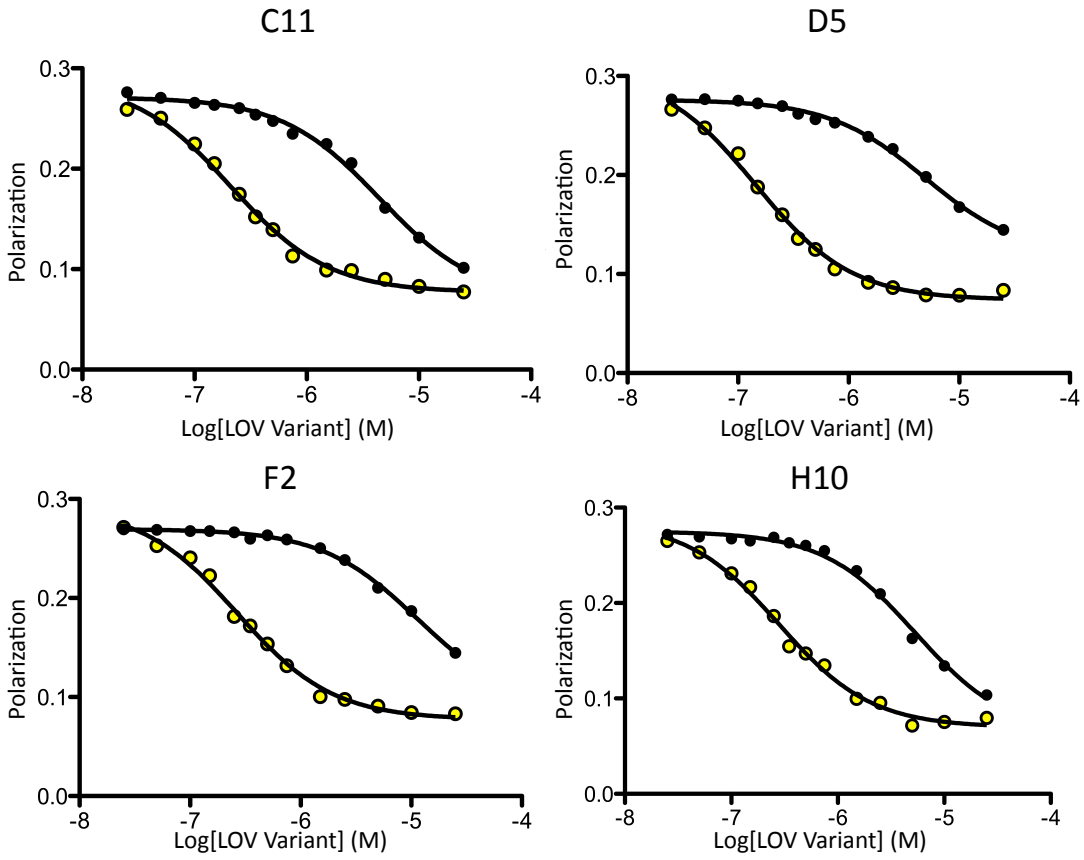


Figure 3.6 In vitro binding data for top four sequences. Florescence polarization competition binding assays for the top four sequences (shown in Fig. S2). All four have a greater fold change in affinity than their parent oLID (Fig. 2B). Binding affinities can be found in Table S3.

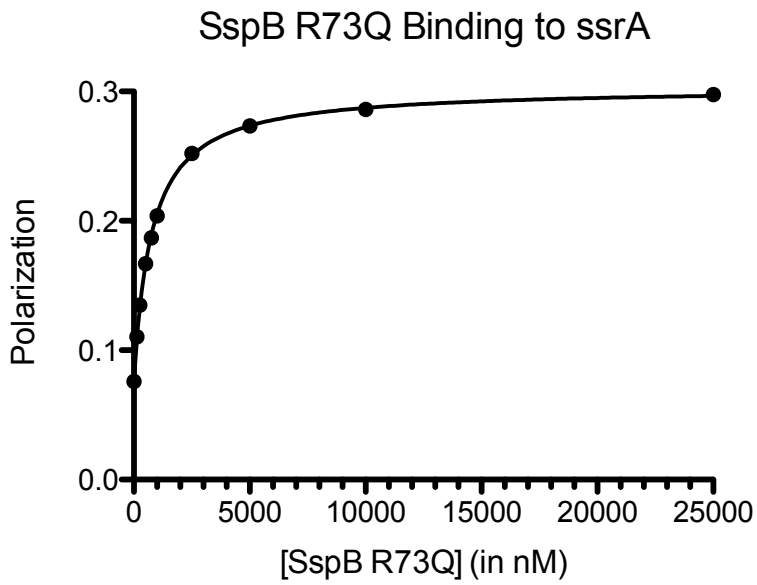


Figure 3.7 Binding of SspB R73Q to ssrA peptide. Fluorescence polarization binding of ssrA labeled peptide to SspB R73Q (Micro). Affinity was measured to be 900 nM ± 200 nM.

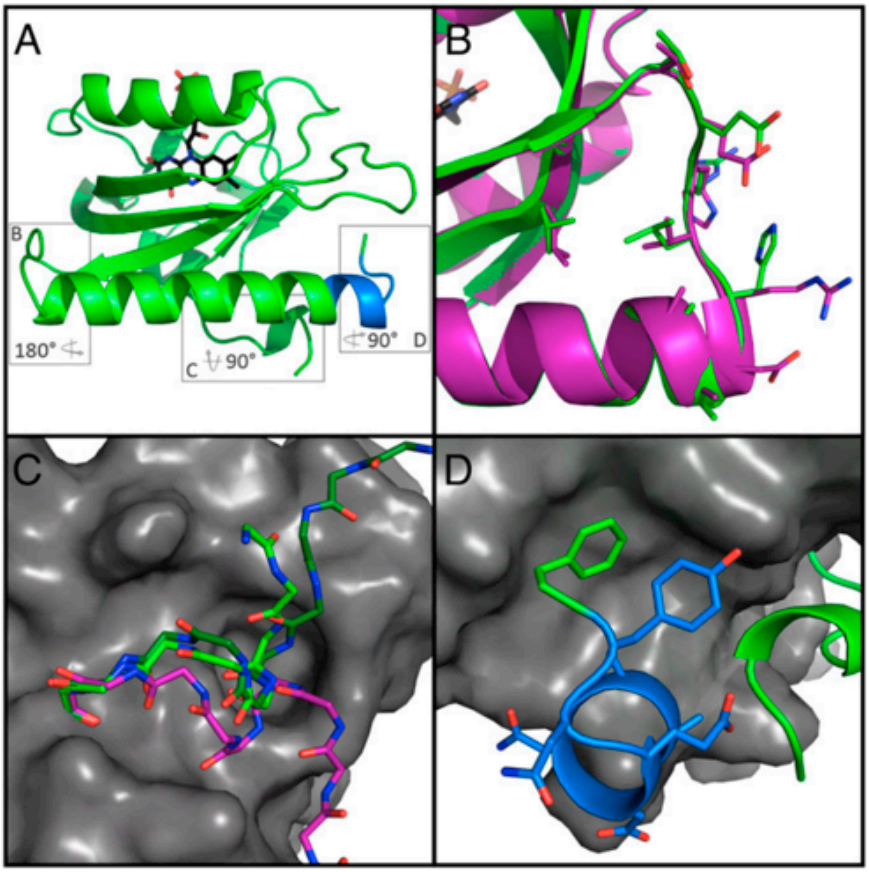


Figure 3.8 Structure of iLID yields insight to role of mutated residues. (A) Overall topology iLID (green) remains unchanged from AsLOV2. Boxes surround three areas of interest to be shown in B, C, and D. (B) Alignment with AsLOV2 (PDB ID code 2V0U, purple) reveals no major backbone rearrangement of the hinge region due to mutations. (C) Close up view of A'α helix from iLID (green), AsLOV2 (purple), and PA-Rac1 (PDB ID code 2WKP, dark green), show different orientation in caged LOV variants from uncaged. (D) SsrA epitope (blue) contains an extra helical turn before wrapping back and interacting with the PAS fold. Designed C-terminal phenylalanine (green) packs nicely into a hydrophobic pocket of the LOV domain.

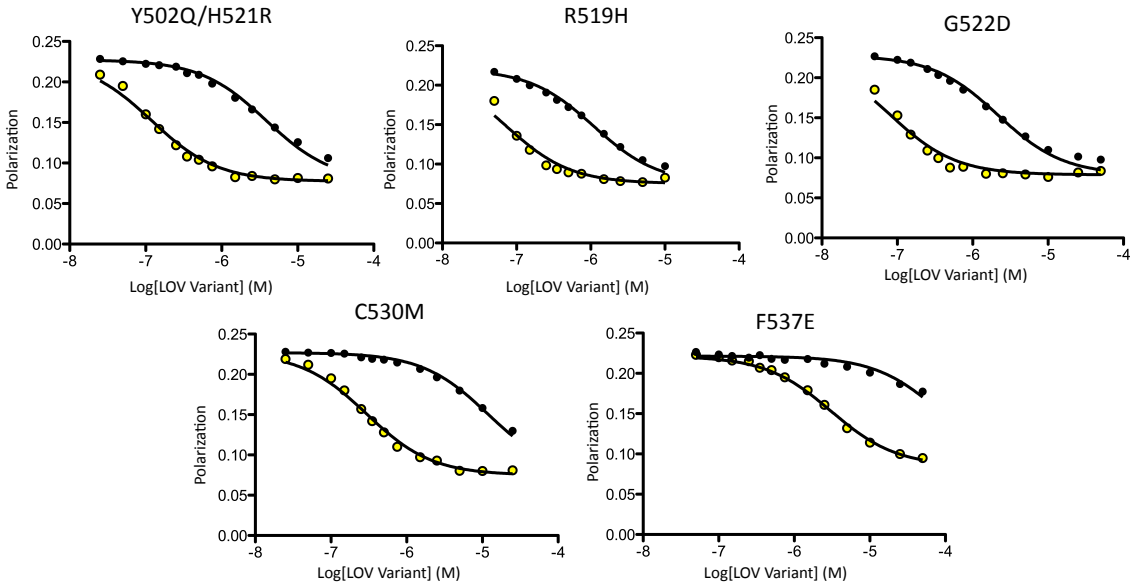
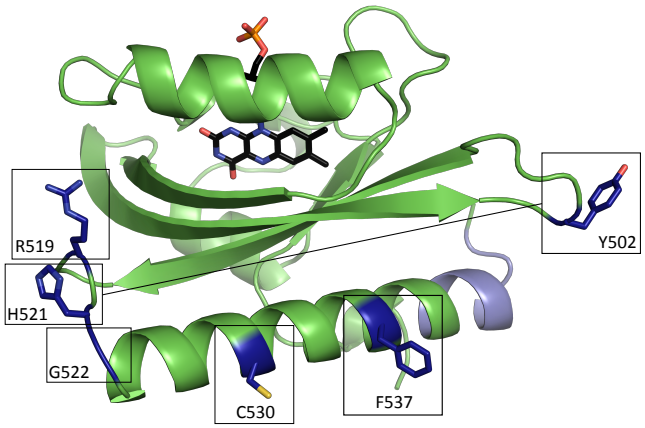


Figure 3.9 In Vitro Binding of Reverted iLID Mutants. Florescence polarization competition binding assays for point mutation reversions show that only C530 can be removed without substantially affecting dynamic range or affinity range. Binding affinities can be found in Table 3.2.



Mutation	Lit Affinity	Dark Affinity	Fold Change
iLID	132 ± 5 nM	4.7 ± 0.7 µM	36
Y502Q/ H521R	43 ± 1 nM	1.6 ± 0.1 µM	37
R519H	23 ± 1 nM	474 ± 1 nM	21
G522D	32 ± 1 nM	885 ± 1 nM	28
C530M	112 ± 1 nM	4.6 ± 0.1 µM	41
F537E	1.4 ± 0.1 µM	37.9 ± 0.3 µM	27

Figure 3.10 Affinities of Reverted iLID Mutations. Reversion of iLID mutations elucidate role each plays in switch improvement. Left: structure of iLID with six mutations shown (boxes, dark blue). Right: Blue light and dark binding affinities due to the reversion of each of the mutations shown on left. C530 appears to have little to no effect on overall switching or range of affinity.

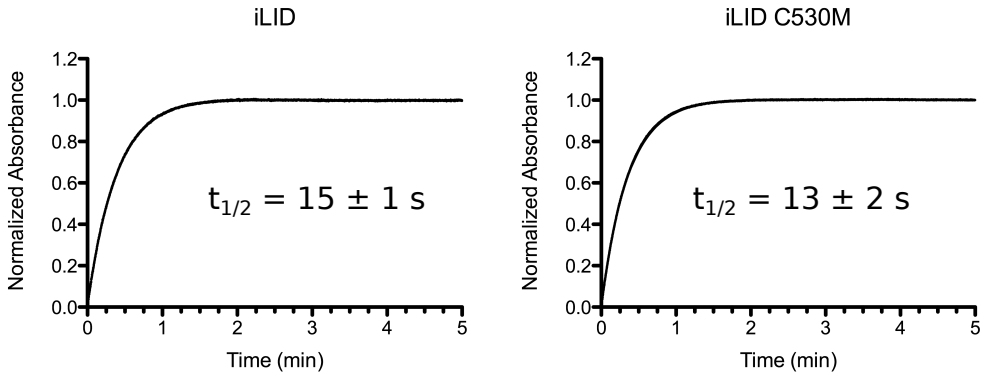


Figure 3.11 Thermal Reversion of iLID and iLID C530M. Reversion of iLID C530M (right) yields similar reversion kinetics as iLID (left). Experiments were performed at room temperature.

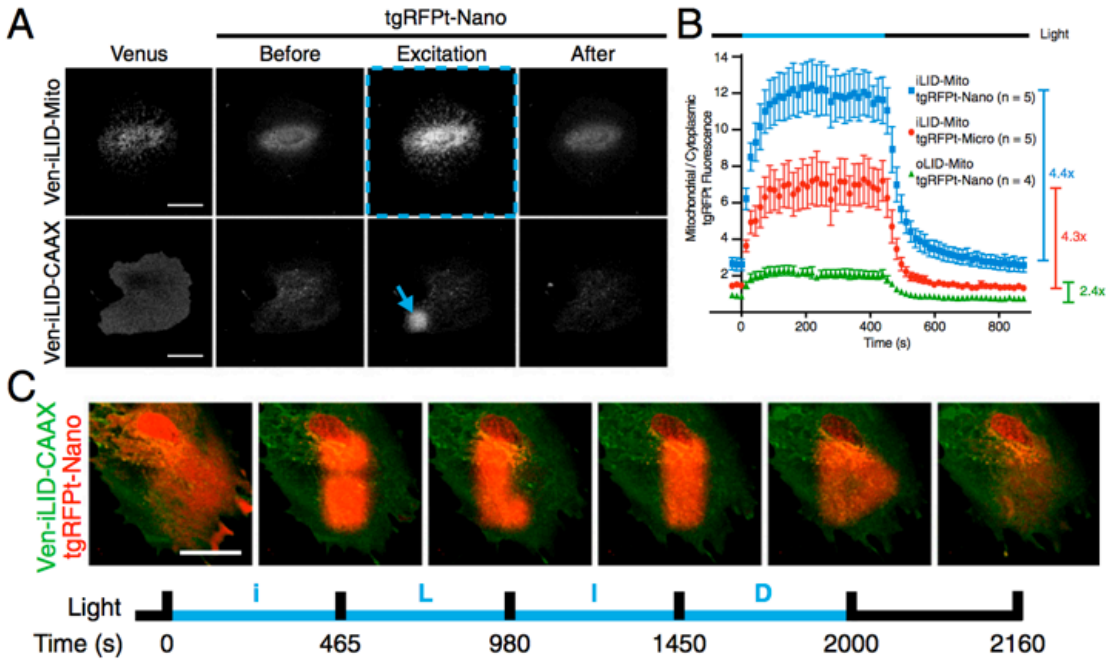


Figure 3.12 iLID provides improved local recruitment in cell culture. (A) IA32 fibroblasts expressing membrane (CAAX) and mitochondrial anchored Venus-iLID and cytoplasmic TagRFP-T-Nano. Localized activation (denoted by blue markings) caused relocalization of TagRFP-T-Nano. (Scale bar, 50 μm.) (B) A ratio of mitochondrial to cytoplasmic TagRFP-T signal intensity during a time course of whole-cell activation as shown in row 1 of A for each pair of mitochondria anchored switches. (C) Patterned activation of Venus-iLID-CAAX shows tight spatial and temporal control of TagRFP-T-Nano localization within a cell. (Scale bar, 50 μm.)

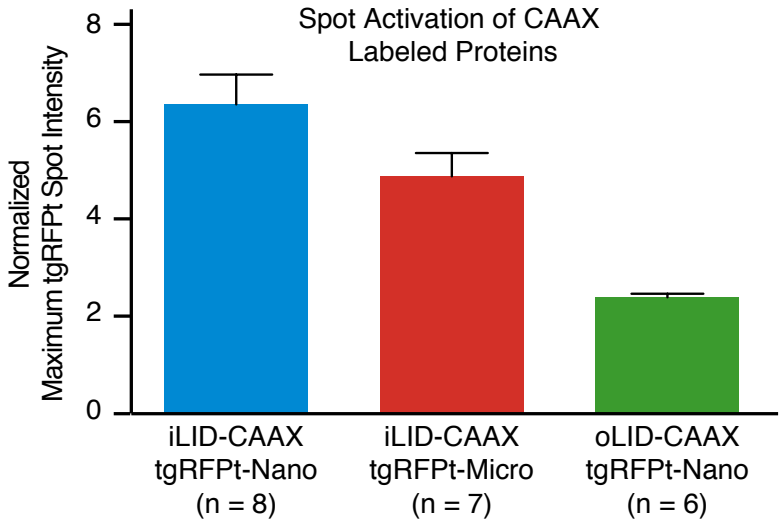


Figure 3.13 The maximum recruitment of sspB to activated LIDs in cells. Measurements were made from a time course of localized activation as shown in Fig. 4A, Row 2. The maximum tgRFPt signal intensity in an activated region of a cell, normalized to the intensity before activation was measured.

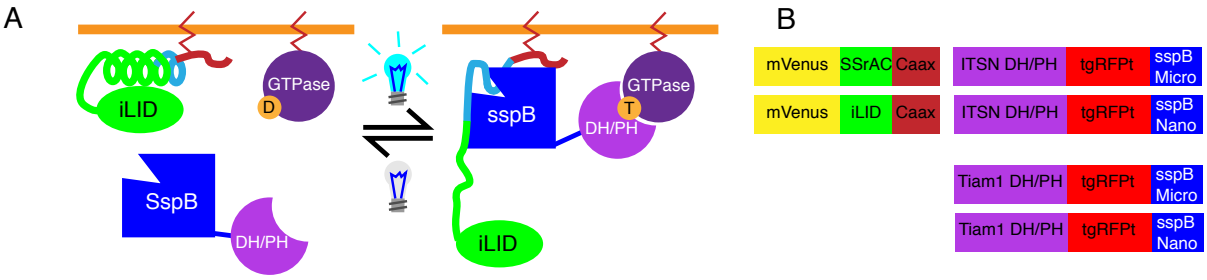


Figure 3.14 Activating GTPase pathways through GEF recruitment. (A) A schematic, demonstrating that in the absence of blue light, Micro fused to a DH/PH domain remains cytoplasmic and inactive. In the presence of blue light, Micro fused to the DH/PH domain binds to iLID, localizing the DH/PH to the membrane, where it increases the rate of GTPase nucleotide exchange to induce signaling. (B) Schematic of constructs made and tested.

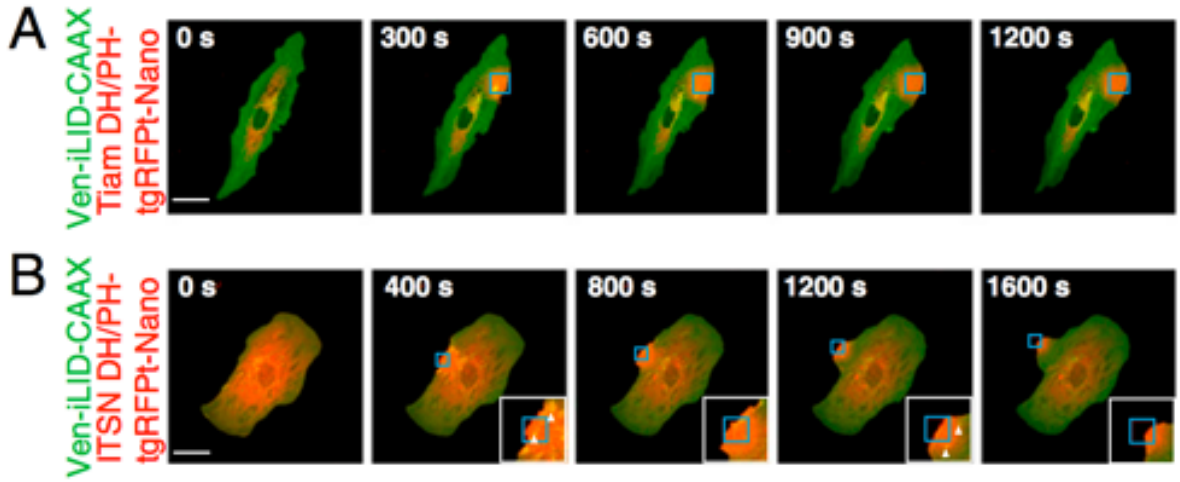


Figure 3.15 Spatial control of GEF DH/PH domains by iLID produces localized control of Rac and Cdc42 activity. IA32 cells expressing Venus-iLID-CAAX and (A) Tiam DH/PH-TagRFP-T-Nano or (B) ITSN DH/PH-TagRFP-T-Nano. iLID was activated in the regions highlighted in blue. White arrowheads mark vesicles. Images are representative (n > 5). (Scale bars, 50 μm.)

Chapter 4

Correlating the in vitro and in vivo activities of light induced dimers: a guide for applications in cellular optogenetics

This chapter covers our efforts to compare in vitro affinities and kinetics to in vivo activity of three light inducible heterodimer systems. Hayretin Yumerefendi performed the yeast gene transcription. Seth Zimmerman performed the mammalian cell colocalization experiments. This chapter is currently being revised for publication. The work was overseen and funded by the guidance of Brian Kuhlman and James Bear.

4.1 Overview

Light inducible dimers are powerful tools for cellular optogenetics as they can be used to control the localization and activity of proteins with high spatial and temporal resolution. Despite the generality of the approach, application of light inducible dimers is not always straightforward as it is frequently necessary to test alternative dimer systems and fusion strategies before the desired biological activity is achieved. To better inform this engineering process we establish correlations between the biophysical properties and the *in vivo* activities of variants of three blue-light inducible dimers: cryptochrome2 (CRY2)/CIB1, iLID/SspB, and LOVpep/ePDZb. We find that the switches vary dramatically in their dark-state and lit-state binding affinities, and that these affinities correlate with activity changes in a variety of *in vivo* assays including transcription control, intra-cellular localization studies and control of GTPase signaling. Additionally, for CRY2 we observe that light induced

changes in homooligomerization can have large effects on activity that are sensitive to alternative fusion strategies.

4.2 Introduction

Optogenetics originally described the use of the light sensitive cation channel, Channelrhodopsin-2, to manipulate the action potential of neurons [6,7]. More recently, optogenetics has extended into the realm of cell biology with the development of cellular optogenetic tools. These tools are not limited to the manipulation of action potentials, but encompass any genetically encoded and light dependent system that can be used to manipulate many cellular processes. Particularly successful has been the use of light induced dimerization to control a variety of processes such as gene transcription, GTPase signaling, protein degradation, and organelle transport [72,76,85,92,101,149–151]. For example, by fusing one half of an inducible dimer to a protein anchored in the plasma membrane and the other half to a guanine nucleotide exchange factor (GEF) it is possible to localize the GEFs to the membrane with light and activate GTPase signaling.

While light inducible dimerization has proven to be a general approach for regulating biological processes, it is frequently necessary to test alternative dimer systems and fusion strategies to determine which approach will be most robust and appropriate for a given application [75]. Part of the challenge is that there are a variety of light inducible dimers that have been described in the literature, but few studies have compared switches side-by-side or characterized their intrinsic biophysical properties. Here, we establish correlations between the *in vitro* and *in vivo* activities of three blue-light inducible dimers: cryptochrome2 (CRY2)/CIB1, iLID/SspB, and LovPep/ePDZb [46,128,152]. These results should provide valuable input for future efforts to control biological pathways with light inducible dimerization.

As a family, blue light inducible dimers provide a powerful experimental platform. Their photosensitive cofactor is abundant in nature making them broadly applicable to many organisms, and the single wavelength of light necessary to manipulate their dimerization makes for a simple experimental setup. CRY2/Cib is a naturally occurring light-dependent heterodimer from *Arabidopsis thaliana*. Additionally, it has been shown that CRY2 forms homooligomers when activated with light [153]. Both CRY2/CIB1 dimerization and CRY2 homooligomerization have been used to control a variety of cellular processes [45,47,83,88,127,154]. However, neither the dark state nor lit state binding affinities between CRY2 and CIB1 have been measured, nor the stoichiometry of oligomerization been determined.

The TULIP (LOVpep/ePDZb) [128] and iLID (iLID/SspB) [152] systems are engineered heterodimer pairs built upon the light-induced conformational change of the *Avena sativa* (*As*) phototropin LOV2 domain [19]. In the TULIP system a PDZ binding motif was encoded in the Jα helix of *As*LOV2, sterically caged from binding an engineered PDZ (ePDZ) domain in the dark [128]. Blue light induces a conformational change within *As*LOV2, relieving this occlusion and increasing affinity to ePDZ. The iLID system works in a similar fashion, caging the *E. coli* ssrA peptide from its binding partner, SspB [132,152]. Despite the mechanistic similarities between TULIPs and iLID, the lack of molecular characterization prevents direct comparison and empirical switch selection when developing a new application.

Recently, the Tucker group began the process of benchmarking light inducible dimers by comparing CRY2/Cib, TULIPs, and Phy/Pif in a set of standardized yeast functional assays [84]. The Phy/Pif pair is a light induced dimer that rapidly forms under red light and rapidly dissociates when illuminated with far red light. The Phy/Pif system requires a cofactor, phycocyanobilin (PCB), which is not readily available in some organisms. These studies demonstrated a wide range of activities when using the switches to co-localize DNA binding and activation domains for control of reporter gene transcription in yeast. To better understand these variations and extend the results to mammalian systems, we continue the

benchmarking process by measuring binding constants for the dimers in the lit and dark states, and performing a variety of activity assays including: co-localization experiments in mammalian cell culture, transcription-control assays in yeast, and the activation of small GTPases via the sub-cellular recruitment of guanine nucleotide exchange factors. In general, we find that the measurements made *in vitro* correspond to what we observe in cells. The switches with the largest changes in *in vitro* binding affinities upon light stimulation make the most effective switches for the in-cell benchmarks.

4.3 Methods

4.3.1 Cloning, Expression, and Purification

All clones are available via Addgene. iLID, LOVpep, and WT AsLOV2 were all cloned into pQE-80L BamH1 and HindIII sites for *E. coli* expression with an N-terminal 6x His tag. The respective binding partners, SspB Nano & Micro, and ePDZb were cloned into a modified pQE-80L vector (BamH1/HindIII sites) with an N-terminal 6xHis-MBP-TEV tag. Both full length *At*CRY2 as well as the PHR domain alone was cloned into the SalI site of the pFastBac HT A vector for insect cell expression. Recombinant bacmid DNA was made in DH10Bac *E. coli* cells and virus amplified in Sf9 insect cells. CIB1N was cloned into the BamH1 and HindIII sites of pQE-80L for expression in *E. coli* with a 6xHis-tag. All mammalian constructs were cloned into the pLL7.0 lentiviral vectors. Expression is therefore driven by a CMV promoter. The constructs were assembled by PCR based overlap extension, enzyme restriction, and ligation or through Gibson assembly.

Bacterial expression was performed as follows: BL21(DE3) cells were transformed through heat shock with each of the expression vectors. For each construct, 1.5L of LB media was inoculated and grown at 37°C to OD 0.6 and induced with 333mM IPTG. iLID, LOVpep, *As*LOV2, SspB nano & micro, and ePDZb were expressed at 18°C for 16 hours.

CIB1N was expressed at 25°C for 6 hours. After expression, cells were spun down at 3500 rpm for 10 minutes and pellets were frozen until purification. Insect cell expression was performed as follows: SF9 cells were inoculated with baculovirus at an MOI of 10 and expressed at 27°C for 48 hours according to [155]. After 48 hours, cells were spun down at 2000 rpm, washed with cold PBS buffer and frozen at -80°C until purification. Bacterial cell pellets of LOV based switches and their binding partners were resuspended in phosphate lysis buffer (50 mM sodium phosphate pH 7.5, 500 mM NaCl, 20 mM Imidazole, 100 μM PMSF) and sonicated. Cell lysates were spun down for 30 minutes at 20,000 rpm. Cell supernatants were filtered with a 5 μm filter, run over HisTrap HP columns (GE) and eluted with elution buffer (50 mM sodium phosphate pH 7.5, 500 mM NaCl, 500 mM Imidazole, 100 μM PMSF). Proteins expressed as 6xHis-MBP fusions were dialyzed overnight in PBS with TEV protease and re-run over HisTrap columns to separate the protein of interest from His-MBP. Finally, all proteins were passed over at Superdex 75 column (GE) as a final clean up and buffer exchange to PBS (10 mM dibasic sodium phosphate, 1.8 mM monobasic potassium phosphate, 137 mM NaCl, 2.7 mM KCl, pH 7.4) for characterization. *At*CRY2 and CIB1N purification was similar to the above protocol except Tris buffers were used instead of phosphates buffers as previously published [155]. Insect cells were lysed by sonication without detergents to prevent contamination. The final size exclusion buffer for both *At*CRY2 and CIB1N was 50 mM Tris-HCl pH 7.4, 250 nM NaCl, 5 mM BME. We also expressed the PHR domain (domain necessary for dimerization) of *At*CRY2 alone, but poor yields precluded *in vitro* experiments with this variant. Most of the dimer systems expressed highly and we had little handling or solubility issues with them. The notable exception was *At*CRY2, which precipitated at concentrations above 20 μM.

4.3.2 Fluorescent Probe Generation

To measure direct binding between *At*CRY2-CIB1N, CIB1N was labeled with 5(6)-TAMRA (Anaspec) at the single cysteine residue. Purified proteins were buffer exchanged on PD-10 desalting columns (GE) into 20mM Tris-HCl pH 7.5, 150mM NaCl, 1mM TCEP. Ten-fold excess dye was added to the prep and was left on a rotator at 4°C overnight. Labeled proteins were then passed through another PD-10 column to remove free dye. Absorbance at 555 nm (ε=65,000 M^{-1} cm^{-1}) was used to quantify dye concentration and BCA assay (Thermo Scientific) was used to quantify protein concentration. Competitive binding assays were used to measure binding for the iLID and LovPep systems. The sequence for the LovPep competitor peptide was 5(6)TAMRA-EEIDKAVDTWV and the sequence for the iLID competitor peptide was 5(6)TAMRA-QIEEAANDENY.

4.3.3 Fluorescent Polarization Binding Assay

Fluorescence polarization measurements were recorded using a Jobin Yvon Horiba FluoroMax3 fluorescence spectrometer. All binding assays except CRY2/CIB1N were performed in PBS buffer in either a 1 cm or 1 mm quartz cuvette at 25°C. CRY2/CIB1N binding was performed in a Tris (20mM Tris-HCl, pH 7.4, 200 mM NaCl, 5 mM BME) buffer due to solubility issues. Polarization of TAMRA was measured with excitation at 555 nm and emission at 584 nm. For *At*CRY2-CIB1N binding, the concentration of TAMRA-CIB1N started at 200 nM and *At*CRY2 was titrated in. At each titration point, the sample chamber was illuminated with 6.0 mW cm^{-2} blue light using a collimated blue led array. A lit state time point was taken immediately after removal of the blue light and another 5 minutes later for *As*LOV2 binding and 10 minutes later for *At*CRY2 binding. Initial affinities of the iLID and LovPep competitor peptides were measured through direct binding titrations. Starting peptide concentrations were 25 nM for the iLID peptide and 250 nM for the LovPep peptide. For iLID nano competitive binding assays, 25 nM peptide and 40 nM SspB nano were

incubated with enough competitor to bind approximately 60% of peptide prior to titration. Competitive binding titrations were illuminated with blue light as in the direct binding assays and dark state measurements were taken after 5 minutes of darkness.

4.3.4 Multi-Angle Light Scattering

SEC-MALS experiments were performed on a Wyatt DAWN HELEOS II light scattering instrument interfaced to an Agilent FPLC System with a Superdex S200 column, Wyatt T-rEX refractometer and Wyatt dynamic light scattering module. CRY2 samples were prepared at 15 μM (in 50 mM Tris-HCl pH 7.4, 250 nM NaCl, 5 mM BME) and run through the S200 either in the presence of blue light (~3 mW cm^{-2}, blue led array) or in darkness.

4.3.5 Dynamic Light Scattering

CRY2 oligomerization was measured in a DynaPro Dynamic Light Scattering Plate Reader at room temperature. CRY2 at 15 μM was illuminated with blue light (6.0 mW cm^{-2} blue light, collimated blue led array) for 1 minute and placed in the instrument. Measurements were taken every 5 seconds for 20 minutes.

4.3.6 Absorption Recovery after Activation

Excited state recovery times were measured using a Cary 50 UV-Visible Spectrophotometer. Samples were irradiated with blue light (6.0 mW cm^{-2} blue light, collimated blue led array) and absorbance at 450 nm was recorded until recovery.

4.3.7 Yeast Plasmids Generation

Clontech pGBKT7 vector was modified to substitute the 2μ origin with CEN4 origin of replication by restriction digest with SacI and XmaI introduced from primers of a vector PCR and CEN4 from pNIA-CEN-MBP[156] yielding pGBKT7-CEN. CRY2PHR (1-498) was cloned with NdeI and NotI, ePDZb with NdeI and BamHI and finally, SspB Nano and Micro were cloned with EcoRI and BamHI into the newly generated pGBKT7-CEN plasmid. Additionally,

CIB1N (1-170) was cloned in the original pGBKT7 vector using NdeI and BamHI. CIB1N was cloned in pGADT7 with NdeI and BamHI, LOV-pep and variant were cloned as well as oLID and iLID were cloned with EcoRI and BamHI. Finally, CRY2PHR was cloned in pGADT7 via NdeI and NotI restriction digest as well. All plasmids were sequence verified using Eurofins DNA sequencing service.

4.3.8 Yeast Transformation and Mating

The resultant plasmids were transformed via high efficiency lithium acetate transformation[157] in Y187 for pGADT7-derived plasmids and Y2HGold for pGBKT7-derived plasmids. After about 72 hours, single colonies for each were isolated and inoculate 0.5 mL YPD culture overnight in order to mate them and generate the respective diploids. The next day, the mated yeast were pelleted at 3000 rpm for 5 minutes and plated on douple dropout plates (SC-Leucine/-Tryptophane).

4.3.9 β -Galactose Assay

β-Galactose assay were performed as follows: Freshly mated yeast colonies were grown for about 36h at 30°C in 5 ml SC-Leu/-Trp. Cell density was measured at OD_{600} and 2.5 mL cultures were diluted to OD_{600} = 0.2 in duplicates – one for a light and another for a dark condition (falcon tubes were wrapped in aluminium foil). Cultures were grown at 30°C in a shaking incubator (250 rpm) for 3 hours in the dark and then for another 4 hours under blue light (465 nm) at 500 μW/cm^2 via LED strip light wrapped around the tube rack. The resulting cultures were pelleted in triplicates and β-Galactose assay using CPRG for a substrate was performed according Clontech yeast handling protocols.

4.3.10 Yeast Growth Assays

Survival assays were performed as follows: Fresh colonies were grown for about 36 h at 30°C in 5 ml SC-Leu/-Trp. Cell density was measured at OD_{600} and cultures diluted in

200 µl of OD_{600} = 1, followed by 8 5-fold serial dilutions. Then, 2 µl of each of the dilutions were pipetted and spotted using a multichannel pipette (Gilson) onto respective dropout plates. The dark condition plates were wrapped in aluminum foil and placed in the same incubator as the lit condition at 30°C. Continuous blue light (465 nm) at 500 µW/cm^2 was provided with LED strip lights attached at the incubator. Yeast plates were imaged after 70 hours incubation, the resulting images were cropped and arranged using Adobe Photoshop.

4.3.11 Mammalian Cell Culture and Transfection

Mouse IA32 fibroblasts were cultured in DMEM supplemented with 10% (vol/vol) FBS (HyClone), 100 U/ml penicillin, 100 µg/mL streptomycin, and 292 µg/mL L-glutamine. Cells were cultured at a constant 37 °C and 5% (vol/vol) CO_2. Cells were transiently transfected in 6 well cell culture dishes using 1 µg total DNA at 1:1 ratio and NanoJuice (EMD Millipore) transfection reagent as recommended by manufacturer.

4.3.12 Mammalian Cell localization/GEF Microscopy

Experiments were performed according to the methods found in Guntas *et al.* Briefly, cells were co-transfected with two vectors containing the sequences encoding each component of the switch in equal parts. 24 hr later tranfected cells were trypsonized and transferred to 3.5 cm MatTek glass bottom dishes coated with a 10 ug/ml solution of fibronectin. 24 – 48 hr later cells were imaged and photo-activated with an Olympus FV1000 confocal microscope equipped with a 1.30 N.A. 40x oil immersion objective. The Fluoview software Time Controller was used to produce a timeline of image acquisition and photo-activation using the same parameters found in Guntas *et al.*

4.3.13 Image analysis and quantification

All images were analyzed using FIJI software. Spot localization was quantified according to Guntas *et al*. Briefly the tgRFPt fluorescence intensity was measured within the

activated ROI and an initial intensity and size matched area outside the activated ROI. A ratio of In : Out was analyzed throughout time. The values that correspond to the period of activation were fit to the equation Y= 1 + Ymax*(1-exp(-K*X)). The values that correspond to the period of reversion were normalized to the maximum values and fit to the equation Y=(Y0 - Plateau)*exp(-K*X) + Plateau. Whole cell activation was quantified with an improved version of the method described in Guntas *et al.* The cytoplasmic ROIs are more accurately determined by removing a small subset of pixels that are representative of the background but previously included. The values described in the paper are the average tgRFPt fluorescence intensities from the algorithmically determined ROIs expressed as (Mito-Cyto)/Cyto. Cytoplasmic values were first subtracted from mitochondrial values to remove any fluosecence signal contributed by the cytoplasm above and below the mitochondria. Curves were fit to the values during the activation and reversion periods using the equations Y=S+Ymax*(1-exp(-K*X)) & Y=(Y0 - Plateau)*exp(-K*X) + Plateau respectively. The fold change was determined by (S+Ymax)/S. All curve fittings were performed using Prism (GraphPad) software. The protrusion distance reported in the Tiam DH/PH localization experiments was measured by kymography. A line one pixel thick was drawn through each of the activated ROIs. The image values through time along that line were concatenated to form a new image. This image was then used to determine the initial and maximal position of the membrane within the time of activation to determine the maximum protrusion distance.

4.4 Results

4.4.1 Dark and Lit-state Binding Affinities

We used fluorescence polarization binding assays to measure lit and dark state affinities of each pair. For the TULIPs and iLID this was performed using a competitive

binding experiment. The photoactivable domains were used to compete off fluorescently labeled peptides from the binding partners. The interaction between *At*CRY2 and CIB1N (the N-Terminus of CIB1 necessary for dimerization) was not amenable to this experimental format, so CIB1N was covalently labeled with a fluorescent dye and direct binding was measured. As we previously reported, the iLID binding partner SspB comes in two variations, Nano and Micro; each with a different affinity range. The iLID Nano system has an affinity of 0.13 μM under blue light and 4.7 μM in the dark. The iLID Micro pair has an affinity of 0.8 μM under blue light and 47 μM in the dark (Figure 4.1B, Table 4.1) [152]. The TULIP switches we examined function over a weaker range of affinities. The LOVpep construct binds to ePDZb with an affinity of 12 μM under blue light and 72 μM in the dark, for a 6-fold change. The presence of additional "caging" mutations, T406-7A + I532A (LOVpep+), weaken the lit state affinity to 18 μM and the dark state affinity to 150 μM, for an 8 fold change (Figure 4.1B, Table 4.1). The different affinity ranges sampled by the TULIP and iLID switches reflects the affinities of the peptides that are being caged in each case. The SsrA peptide used to create iLID binds to its partner, SspB (Nano), with an affinity of 35 nM [82]. The PDZ binding peptide used in TULIP binds to ePDZb with an affinity of 14 μM.

We were not able to observe a light-dependent change in binding affinity for *in vitro* purified CRY2 and CIB1N. In our direct binding assay, we observed low micromolar binding (~4μM) with and without blue-light stimulation (Figure 4.1B, Table 4.1). This result is consistent with the previous observation that CRY2 purified from insect cells did not show differential affinity for CIB1 in pull-down assays performed in the light and the dark [42]. It has been hypothesized that insect cell purified CRY2 is missing an important chromophore, 5,10-methylenetetrahydrofolate (MTHF). However, even in the presence of saturating MTHF concentrations, we did not observe a significant change in binding affinity due to light (Figure 4.2). We ran an electrophoretic mobility shift assay with CRY2 and CIB1N and

found the same result. CRY2 binds CIB1N similarly under blue light as in the dark, both in the low micromolar range (Figure 4.3).

4.4.2 Light-dependent CRY2 Homo-oligomerization

Although *in vitro* purified CRY2 did not show light-dependent changes in CIB1N binding, we did observe robust homo-oligomerization of CRY2 with light stimulation as probed by multi-angle light scattering (SEC-MALS). A single symmetric peak was observed both in the light and the dark, with the retention time being delayed in the dark. The light scattering indicated a species with a molecular weight of 75 kD in the dark and 301 kD in the light (Figure 4.4A). The expected molecular weight of monomeric CRY2 is 71 kD, so these results are consistent with the formation of a monomer in the dark and a tetramer in the light. Saturating amounts of MTHF did not change the elution times or molecular weight fits (Figure 4.5). Using dynamic light scattering (DLS), we were able to measure the kinetics of the lit state oligomer to dark state monomer transition. In this assay, there is a 60 second delay between removal of blue light and the first DLS reading, as shown in grey (Figure 4.4B). Factoring in this dead time, the reversion to dark state has a half-life of 90 ± 20 seconds. We also ran co-elution experiments of CRY2 and CIB1N in gel filtration experiments (Figure 4.6). Samples of CRY2 and CIB1N (2:1 molar ratio) were run in the light and dark, however in both states CIB1N did not co-elute with CRY2. This is consistent with the micromolar binding affinities that we observed in the fluorescence experiments.

4.4.3 Reversion Kinetics

We used our purified protein samples to also measure the reversion kinetics of the photoactivated states. All proteins were extensively dialyzed into 50 mM Tris-HCl pH 7.5, 250 mM NaCl buffer for these experiments. Using an absorbance recovery after photoactivation assay, we determined the lit state half-life for each of our photosensitive

domains. In order of fastest to shortest, the half-lives were measured to be 18 ± 2 s for iLID, 23 ± 1 s for LOVpep, 51 ± 2 s LOVpep+, and 92 ± 10 s for CRY2 (Figure 4.7, Table 4.1).

4.4.3 Controlling Sub-cellular Localization

Next, we examined how effective the photoswitches were at recruiting proteins to a specified region of the cell. In particular, we were curious if the *in vitro* binding properties of the dimers would correlate with in-cell behavior. Each half of the switches were fused to a fluorescent protein (Venus or tgRFPt) with spectral properties distinct from the excitation wavelength of the photoactive domain. The Venus labeled half of the switch was also fused to a membrane-anchoring domain (N-Myristoylation (Myr) or C-Farnesylation (CAAX)). The two proteins were then co-expressed in mouse fibroblasts and continuously imaged with a confocal laser-scanning microscope. During imaging a region of interest (ROI) was activated with a 488 nm laser, and changes in protein localization were quantified as a function of time by measuring the ratio of tgRFPt fluorescence intensity inside the activated ROI to the intensity in a ROI of the same size outside the area of activation. The analysis produces a maximum intensity ratio as well as the half-life of activation and reversion (Table 4.2).

The iLID Nano and Micro switches were previously analyzed by this method but our findings are reiterated here. For each of these assays the iLID half of the switch was anchored to the membrane with a Caax motif while Nano and Micro were diffuse in the cytoplasm. Upon activation the tgRFPt-Nano and micro fluorescence intensity increased to produce an average maximum ratio value of 6.36 and 4.88 respectively (Figure 4.8A, B).

Our first experiments with the TULIP system used a similar approach, LOVpep+ was fused to Venus and a Caax motif while ePDZb was fused to tgRFPt. However, with LOVpep+ fused to a Caax motif the C-terminal PDZ binding motif was prevented from binding to ePDZb and no change in fluorescence intensity at the activated ROI was observed (Figure 4.9A). Therefore, we fused the LOVpep+ to an N-Terminal myristoylation

sequence, freeing the PDZ binding motif. Upon expression, we found that the myristoylated sequence localized to the plasma membrane but also localized to other membrane bound organelles. For these experiments we chose ROIs that predominantly consisted of only plasma membrane bound LOVpep+. Upon light stimulation, we observed a small increase in protein localization (average maximum ratio value = 1.34) (Figure 4.8A, B).

Others have reported difficulties in maintaining a functional switch upon anchoring CRY2 to the plasma. However, Pathak *et. al.* recently maintained functionality in yeast by fusing CRY2PHR to the C-terminus of Mid2, a membrane anchored protein [84]. We therefore tested 3 additional experimental approaches varying the switch positions as follows: Venus-CRY2PHR-Caax, tgRFPt-CIB1N; Myr-Venus-CRY2PHR, tgRFPt-CIB1N; and Venus-CIB1N-Caax, tgRFPt-CRY2PHR. While the Caax fused CRY2PHR localized to the plasma membrane, the tgRFPt fluorescence intensity did not increase upon activation within the ROI (Figure 4.9A). Myr-Venus-CRY2PHR had a similar localization pattern as Myr-Venus-LovPep and upon activation induced an increase in tgRFPt intensity within the ROI (Figure 4.8A, B). We measured the average maximum ratio value to be 1.52. Additionally, upon activation Myr-Venus-CRY2 formed large clusters at the membrane (Figure 4.8A inset). To better understand CRY2 cluster formation and dissociation we performed the same ROI analysis on the Venus channel. During activation, we observed an increase in Venus fluorescence intensity within the ROI, which represents cluster formation. Interestingly, we observed a persistent increase in intensity for ~1 min after the light was turned off, suggesting that the clusters continue to form after the blue light is turned off (Figure 4.10B).

In the experiments with Venus-CIB1N-Caax anchored to the membrane, tgRFPt-CRY2PHR is diffuse throughout the cytoplasm before activation. Upon activation with blue light the tgRFPt intensity increases within the ROI and in some cases small tgRFPt-CRY2PHR clusters begin to form (Figure 4.8A inset). Surprisingly, we measured the average maximum ratio value to be 4.98; significantly higher than when CRY2PHR is

anchored to the membrane. However, we hypothesized that CRY2PHR oligomerization alone may be responsible for a portion of the increase in fluorescence measured at the ROI. The idea being that once oligomerized, diffusion of tgRFPt-CRY2PHR would slow while recruiting more monomers, increasing the signal. We tested this hypothesis by expressing and activating tgRFPt-CRY2PHR alone in cells and found that this was indeed the case. CRY2PHR alone had a maximum ratio value of 2.67 (Figure 4.10C and D) Furthermore, we hypothesized that expression of CRY2PHR (the half of the switch that undergoes conformational change upon blue light stimulation) in the cytoplasm would provide less spatial control as compared to membrane anchored CRY2PHR. Our reasoning was that once activated the CRY2PHR could more easily diffuse through the cytoplasm and bind to CIB1N outside of the ROI. However, the gradient of tgRFPt intensity outside the periphery of the ROI was similar to what we measured for the iLID switches where the photoactive domain was anchored to the membrane. (Figure 4.10A).

For all of the switches, the reversion half-lives in cells were longer than the *in vitro* measured half-lives. However, the measurements parallel the *in vitro* patterns and what has been previously observed. The LOV2-based switches are all similar at about 60 sec while the CRY2PHR/CIB1N switches are slower. Interestingly the CRY2PHR/CIB1N reversion half-lives are dependent on the orientation of the switch. The Myr-Venus-CRY2 has a faster half-life at 132 s while the tgRFPt-CIB1N-Caax is significantly slower with a half-life of 242 s (Figure 4.8C).

Except in the case of Venus-CIB1N-Caax with tgRFPt-CRY2PHR the activation half-life seems to correlate with the dynamic range of the switch (larger dynamic range takes longer to reach equilibrium after activation) and not on the kinetics of the protein conformational change (Figure 4.8B). This suggests that the rate-limiting step is diffusion. While we are unsure what causes the slower rate of Venus-CIB1N-Caax with tgRFPt-CRY2PHR, a hypothesis will be presented in the discussion.

4.4.4 Mitochondrial re-localization assay

A limitation of the membrane localization assay is that is difficult to accurately determine the portion of tgRFPt labeled protein that is at the plasma membrane prior to activation due to the axial spatial resolution of the microscope (~600nm). In the relatively flat cultured fibroblasts we used for these experiments, the apical and basal membrane fluorescence values are captured but cannot be distinguished from the cytoplasmic fluorescence. By anchoring the Venus labeled half of the switch to the mitochondrial membrane we were able to more accurately determine the initial amount of tgRFPt labeled protein at the mitochondria relative to the cytoplasm and monitor its change during and after activation. Proteins were anchored to the mitochondrial membrane by fusion to TOM20 at the N-terminus or Mito anchor sequence from *Listeria monocytogenes* ActA protein[158] at the C-terminus. After co-expression of each half of the switch, whole cells were activated with 488 nm light and imaged. Using an automated ImageJ macro, we measured the ratio of mitochondrial to cytoplasmic fluorescence intensity throughout activation and reversion. For each switch this assay produces parameters describing the half-life of activation and reversion, a starting mito/cyto intensity (representative of dark state binding), a maximal mito/cyto intensity and the fold change in intensity (Table 4.3).

Again, the data for the iLID switch has been previously reported. However, the data was reanalyzed using an improved ImageJ macro, which was able to better differentiate the cytoplasm from background. In this assay iLID was fused to the Mito anchoring domain of ActA while Micro and Nano were cytoplasmic. We measured the average fold change for iLID-Nano and Micro to be 5.4 and 5.2 respectively (Figure 4.11A, B). As expected, the initial relative mitochondrial fluorescence intensity for iLID-nano is higher than iLID-micro, paralleling the *in vitro* measured dark state affinity being tighter (Figure 4.11B).

To test the TUPLIP switch we fused Venus labeled LovPep T406,7A I532A to TOM20 at the N-terminus to preserve an accessible C-Terminal PDZ binding motif. Upon co-expression with tgRFPt-ePDZb and activation, the switch produces an average 2.4 fold change in relative mitochondrial tgRFPt fluorescence intensity (Figure 4.11A, B). This fold change parallels the smaller *in vitro* measured dynamic range of binding in comparison to the iLID switches. TULIP also showed a lower starting mitochondrial tgRFPt intensity (0.59); again paralleling TULIP's lower *in* vitro dark state affinity (Figure 4.11B).

We tested the CRY2PHR/CIB1N switch in both orientations. We first tested TOM20-Venus-CRY2PHR with tgRFPt-CIB1N. Unfortunately, TOM20-Venus-CRY2PHR appeared to be toxic to the cells and therefore expression levels in the surviving cells were significantly lower than all other constructs to the extent that the laser power of the microscope had to be substantially increased to obtain a clear image. Additionally, the distribution of mitochondria within the surviving cells was abnormal. Upon activation the cells did not produce a measureable increase in mitochondrial tgRFPt intensity (Figure 4.9B). We therefore reversed the orientation of the switch. By co-expressing and activating Venus-CIB1N-Mito and tgRFPt-CRY2PHR we measured an average 3.1 fold change in relative mitochondrial tgRFPt fluorescence intensity (Figure 4.11A, B). The average initial value of mito/cyto tgRFPt was also high (2.88) compared to the other switches (Figure 4.11B). This suggests a relatively tight dark state binding affinity.

4.4.5 Light Controlled Transcription in Yeast

To examine if our findings from the *in vitro* binding assays and the localization studies correlate with outcomes in a functional assay, we used the light dimerization pairs to control transcription in yeast. The yeast two-hybrid approach has previously been used to demonstrate light dependent transcription for CRY2PHR with CIB1N as well as ePDZb1 and LOVpep[84,86]. We used diploids generated from mating Y187 and Y2HGold strains to test for

the activation of the *lacZ*, *his3* and *ade2* reporter genes (Figure 4.12A) transformed with the split Gal4 transcription factor constructs (Figure 4.12B). We observe an assortment of induced transcription levels dependent on the protein pair used and the reporter gene observed. We identified strong light dependent transcription using β-galactose expression as readout for iLID with Nano (19.5 fold) and iLID with Micro (9.4 fold) (Figure 4.12C and D, Table 4.4). We previously showed that iLID had an improved dynamic range when compared to its parental construct, oLID, by a multitude of measurements. However, we were curious how the two switches compared in their ability to control yeast transcription. As expected light dependent transcription was not detected for the oLID paired with either Nano or Micro (Figure 4.13). In our hands there was also no detectable β-galactose expression for ePDZb paired with LOVpep or LOVpep+.

In addition to monitoring β-galactose expression, we also tested for light-dependent survival on histidine and histidine/adenine dropout plates. Interestingly, iLID when paired with Nano or Micro conveyed growth in the light and dark for single and double dropout plates. In contrast, yeast expressing the LOVpep did not survive in the dark, but there was growth on histidine dropout plates in the light (Figure 4.14). These results are consistent with the survival assays being more sensitive to low levels of expression. iLID Nano and Micro have stronger binding affinities in the dark than LOVpep, and in this context this "leakiness" is sufficient to allow growth even when the switch is in the inactive/dark state.

In previous studies CRY2-DBD paired with CIB1 or CIB1N has been shown to activate transcription in yeast [84,86], although the overall levels of transcription with these constructs were low when compared to results with full-length CRY2. To date the inversed orientation has not been reported to our knowledge. When we paired CIB1N-DBD with CRY2PHR-AD we observed strong light dependent expression of *lacZ* achieving about 9 fold difference, similar to when combining Micro with iLID, but overall lower levels for both

light and dark levels. On the contrary, when testing CRY2PHR-DBD with CIB1N-AD we saw no significant transcriptional activation of *lacZ* but only of *his3* reporter genes, which as expected was in light-dependent manner. Our assay appears less sensitive than the work by Kennedy and co-workers, perhaps because different ß-galactose substrates were used (ONPG vs. CPRG).

4.4.6 Manipulation of lamellipodial protrusion

To test each pair's ability to functionally manipulate a mammalian cell we targeted the Rho GTPase family. The Rho family of small GTPases is known to regulate the dynamics of the actin cytoskeleton and therefore the cells shape. Canonically, activation of the membrane bound Rac family member produces highly branched actin, leading to dynamic lamellipodial protrusions. The inactive/active state of GTPases is determined by the state of the bound nucleotide (GDP/GTP respectively). Guanine nucleotide exchange factors (GEFs) activate GTPase by aiding in the exchange of GDP for GTP [147,159]. Using iLID we have previously shown that by localizing the catalytic DH/PH domain of a Rac GEF (Tiam) to a portion of the plasma membrane we can induce lamellipodial protrusions in that region [77]. We therefore used this approach as a functional test of the CRY2PHR/CIB1N or TULIP switches. To this end, we fused the Tiam DH/PH domain to each of the tgRFPt labeled halves of the switch. The Tiam constructs were then co-expressed with the appropriate membrane bound switch half. Cells were imaged and activated in a similar manner to the previous membrane localization experiments, though here the activation ROIs were located at the edge of the cell. For each cell, the maximal protrusion distance at the ROI was then quantified by kymography. iLID-Nano and Micro produced on average a maximal protrusion distance of 12.0 and 14.5 µm respectively (Figure 4.15A, B). The TULIP switch caused an average protrusion distance of 1.5 µm, significantly less than the iLID switches (Figure 4.15A, B). The CRY2PHR/CIB1N switches were again tested in both

orientations (CRY2PHR or CIB1N anchored at the membrane) and produced an average protrusion distance of 4.2 and 2.2 μm respectively (Figure 4.15A, B).

4.5 Discussion

The protein switches tested here cover a wide range of dark and lit state affinities; each with different dynamic ranges. For the LOV2 based switches, we found a reasonable correlation between *in vitro* affinities and behavior in living cells. Consistent with previous results, the iLID switches had the largest fold-change in binding affinity upon light stimulation and were the most effective at localizing protein to the plasma membrane and inducing cellular protrusions via localization of Tiam. However, the iLID switches also had tighter dark state affinities than the TULIP switch, LOVpep+/ePDZb, and this was evident in the mitochondrial localization assay where more dark-state localization was observed for both iLID pairs than for the TULIP switch. Also, the iLID pairs exhibited more dark-state activity in the yeast two-hybrid survival assays. These results suggest that use of the TULIP switch may be more appropriate when it is critical to avoid background activity.

The results for CRY2 and CIB1 present a more complicated story. First, we did not observe a light dependent change in CIB1N binding with full-length CRY2 purified from insect cells. However, robust homooligomerization of CRY2 with blue-light stimulation was observed. As previously reported, we found that the CRY2/CIB1 system could be used to localize proteins within the cell, but this functionality was dependent on how the protein fusions were constructed and which partner was anchored to the membrane. When CIB1N was fused to the plasma membrane, there was robust recruitment of CRY2PHR to the membrane with light stimulation. When CRY2PHR was fused to the membrane, only small small amounts of CIB1N were recruited with light activation. There may be multiple factors at play here. When CIB1N is fused to the membrane, CRY2PHR oligomerization may allow for multivalent interactions between the CRY2 oligomers and CIB1N, which is already co-

localized via membrane anchoring (Figure 4.16). This multivalency would lead to increased increased affinity between CRY2 and CIB1. Conversely, when CIB1 is cytoplasmic it is monomeric and there is no way to generate a multivalent interaction with CRY2PHR oligomers. This line of thinking would also explain why the light-dependent mitochondrial localization was only observed with CIB1 anchored in the mitochondria. Another possible explanation for the observed behavior is that the large clusters that form when membrane anchored CRY2PHR is light-activated may preclude robust CIB1 binding.

Avidity effects may also help explain the yeast two-hybrid results with CRY2/CIB1. We observed more robust light-dependent changes in transcription when CIB1N was fused to the DNA binding domain. The Gal4 DNA binding domain forms a dimer when bound to DNA, and therefore CIB1N fused to the DNA binding domain is presented as a dimer to CRY2PHR. This may allow for a multivalent interaction when CRY2PHR oligomerizes and therefore enhance the affinity between CRY2PHR and CIB1N. The same multivalent interaction would not be created when CIB1N is fused to the DNA activation domain. When taken together, our results and results from previous studies indicate that CRY2 homooligomerization is likely to play a significant role in the activity of the switch, and this can be used to enhance light-dependent signaling if multivalent interactions can be created.

4.5.1 Reversion Kinetics

The activation and reversion kinetics of the switches become important when planning experiments as they determine how often you must expose the proteins to blue light in order to maintain dimerization. In the context of a cell this may be important in avoiding phototoxicity. In turn, this needs to be balanced with the rate at which the switch needs to be fully off in the context of the experiment. The CRY2 switch reverted with the slowest kinetics. The quicker kinetics of iLID and the LOVpep switches give more precise temporal resolution, allowing for less lag time between light removal and dissociation. One

important point to note is the photocycle of AsLOV2 switches can be tuned with some previously discovered mutations [25,121], so these switches can be altered to fit a variety of contexts. These mutations haven't been tested in the heterodimerization context, but it is likely they would have little impact on dynamic range as they don't directly interact with the Jα helix. The in cell rate of binding, except in one CRY2PHR/CIB1N orientation, seems to be limited by diffusion. The rates of dissociation maintain the same rank order as the *in vitro* measurements but are longer in both assays. This is most likely due to rates of diffusion out of the measured ROIs. Interestingly, in the membrane localization assay, the CRY2PHR/CIB1N reversion rates are orientation dependent.

We hypothesize that the CRY2PHR/CIB1N kinetic inconsistencies mentioned above are due to CRY2PHR's light dependent oligomerization. The Venus-CIB1N-Caax / tgRFPt-CRY2PHR produced a maximal plasma membrane recruitment level similar to the iLID switches. However the rate of activation was significantly slower. This is surprising, as the other switches recruitment half-life were proportional to their dynamic range. This suggests that in addition to CRY2PHR – CIB1N binding a second slower step is driving the increase in tgRFPt fluorescence intensity, such as CRY2 oligomerization. The orientation dependent rate of membrane dissociation may also be explained by oligomerization. In the case where tgRFPt-CIB1N is free in the cytoplasm, we monitored the dissociation of the monomeric CIB1N from the CRY2 clusters. On the other hand, the tgRFPt-CRY2 may have formed oligomers in the ROI, gaining an avidity interaction with the membrane bound CIB1N, slowing the rate of diffusion and increasing the reversion half-life relative to tgRFPt-CIB1N. These results suggest that while the CRY2/Cib switches can function to robustly manipulate in cell protein interactions, many factors need to be considered and significant optimization may be necessary

4.5.2 Functional Considerations

In the functional yeast transcription and Tiam experiments each of the switches was used to manipulate cell biology to varying degrees. These sets of experiments highlight the advantages and limitations of each switch.

TULIPs have been shown to be effective in yeast gene transcription and MAPK activation [79,84]. While low levels of transcription were reported using the higher affinity TULIP (i.e. ePDZb1 paired with LOVpep) [84] for both *lacZ* and *his3*, our results indicate that weakening further the interaction affinity negatively affects the level of transcriptional activation in yeast. In our hands, ePDZb-DBD paired only with LOVpep-AD exhibited light dependent growth in the absence of histidine in the growth medium. The weak interaction appeared insufficient to induce the *ade2* and *lacZ* reporter genes. When ePDZb-DBD was paired with the LOVpep+ – AD no transcription was seen under any condition.

Similarly, with the TULIP switch we were unable to produce significant lamellipodial protrusions by recruiting TIAM DH/PH to the membrane. We hypothesize that this could be due to three factors. First, TULIPs work over a weaker range of affinities and we did not observe robust localization of Tiam to the membrane. Second, Rac is localized to the plasma membrane by a Caax motif. We have tagged LOVpep with a myristoylation motif due to the limitation of not being able to tag it on the C-Terminus. Therefore, in the case of TULIPS, Tiam DH/PH may be recruited to sub-compartment of the membrane that does not contain Rac. Lastly, the linker length may not optimally position the Tiam DH/PH for activation of Rac.

As opposed to the oLID, which exhibited no light dependence but strong transcriptional activation, iLID demonstrated high levels of blue light dependent transcription achieving a ~20 fold increase of transcription in the presence of light for the iLID-Nano pair. In correspondence with the yeast transcriptional assays, both iLID-Nano and iLID-Micro produced robust lamellipodia in the Tiam DH/PH localization assay.

In both orientations CRY2PHR/CIB1N produced smaller lamellipodial protrusions than the iLID switches. We propose that this could be caused by four factors. Like TULIPS CRY2PHR is anchored to the membrane by myristoylatation. Second, the CRY2PHR oligomerization may inhibit Tiam function and has in fact been previously used to inhibit GEFs and GTPases [80]. Third, as we have shown in the membrane recruitment assay, a portion of the increase in tgRFPt-CRY2PHR signal that we see upon stimulation may only be due to oligomer formation and not necessarily recruitment to the membrane. Therefore we obtain a less robust activation of Rac than expected for the amount of signal increase observed. Fourth, the Tiam DH/PH may not be optimally positioned by our constructs and could be further optimized by linker length adjustment.

Each switch contains other characteristics that we found to influence experimental design. While both components of iLID can be tagged on either the N-or C-terminal, we have found that both CRY2 and LOVpep C-terminal fusions inhibit binding to their partners. These stipulations have hindered particular applications in the past [92]. Although untested, a further consideration for the TULIP switch in cells is the use of a PDZ domain and peptide, which could theoretically interfere with endogenous PDZ interactions. In conclusion, through rigorous benchmarking we have determined *in vitro, in vivo* and functional characteristics of three sets of blue light inducible dimers. This information can be used to guide future efforts aimed at using light inducible dimerization to control biological pathways.

4.6 Tables

		In Vitro Binding			
Switch	Binder	Method	Lit Affinity	Dark Affinity	Reversion Halflife (s)
iLID	Nano	FP	0.132 ± 0.005 μM	4.7 ± 0.7 μM	18 ± 2
iLID	Micro	FP	0.8 ± 0.1 μM	47 ± 13 μM	18 ± 2
LovPep	ePDZb	FP	12 ± 3 μM	72 ± 3 μM	23 ± 1
LovPep +	ePDZb	FP	18 ± 1 μM	150 ± 60 μM**	51 ± 2
Cry2	CibN	FP,SEC	~4 μM		92 ± 10
Cry2 PHR	CibN				

Table 4.1 *In vitro* characterization

		Membrane Localization			
Switch	**Binder**	**N**	**Max In/Out**	**Activation Halflife (s)**	**Reversion Halflife (s)**
iLID	Nano	7	6.36 ± 0.21	95.4 ± 18.1	63.8 ± 6.7
iLID	Micro	7	4.88 ± 0.18	72.0 ± 10.5	52.5 ± 2.5
LovPep	ePDZb	13	1.34 ± 0.03	39.4 ± 15.7	53.8 ± 12.6
LovPep, T406-7A, I532A	ePDZb				
Cry2	CibN				
Cry2 PHR	CibN-CAAX	12	4.98 ± 0.47	186.4 ± 30.9	380.7 ± 73.3
Myr-Cry2 PHR	CibN	10	1.52 ± 0.02	132.7 ± 53.4	164.3 ± 30.3

Table 4.2 *In vivo* membrane localization

		Mitochondrial Localization					
Switch	**Binder**	**N**	**Starting Value**	**Max Value**	**Fold Change**	**Activation Halflife (s)**	**Reversion Halflife (s)**
iLID	Nano	5	1.72 ± 0.5	9.23 ± 2.18	5.4	21.2 ± 1.4	32.0 ± 3.7
iLID	Micro	5	1.45 ± 0.07	7.58 ±1.49	5.2	26.7 ± 6.0	23.2 ± 1.5
LovPep	ePDZb	7	0.59 ± 0.09	1.4 ± 0.20	2.4	25.0 ± 6.7	19.5 ± 0.9
LovPep, T406-7A, I532A	ePDZb						
Cry2	CibN						
Cry2 PHR	CibN	6	2.88 ± 0.67	8.84 ± 1.39	3.1	58.8 ± 9.4	317.6 ± 85.25

Table 4.3 *In vivo* mitochondrial localization

Yeast Transcription								
Gal4-DBD	Gal4-AD	β-Gal Light	β-Gal Dark	Fold Change	His Dropout Light	His Dropout Dark	His/Ade Dropout Light	His/Ade Dropout Dark
Nano	iLID	154.7	7.95	19.5	+++	+++	+++	++
Micro	iLID	94.8	10.1	9.4	++	++	++	++
ePDZb	LOVpep	ND	ND	ND	++	-	-	-
ePDZb	LOVPep, T406-7A, I532A	ND	ND	ND	-	-	-	-
Cry2PHR	CIB1N	ND	ND	ND	+++	-	-	-
CIB1N	Cry2PHR	47.1	5.4	8.7	+++	+++	+++	++

+++ = Very Robust Growth ++ = Significant Growth + = Little Growth - = No Growth

Table 4.4 Yeast two-hybrid β-Gal and survival assay

4.7 Figures

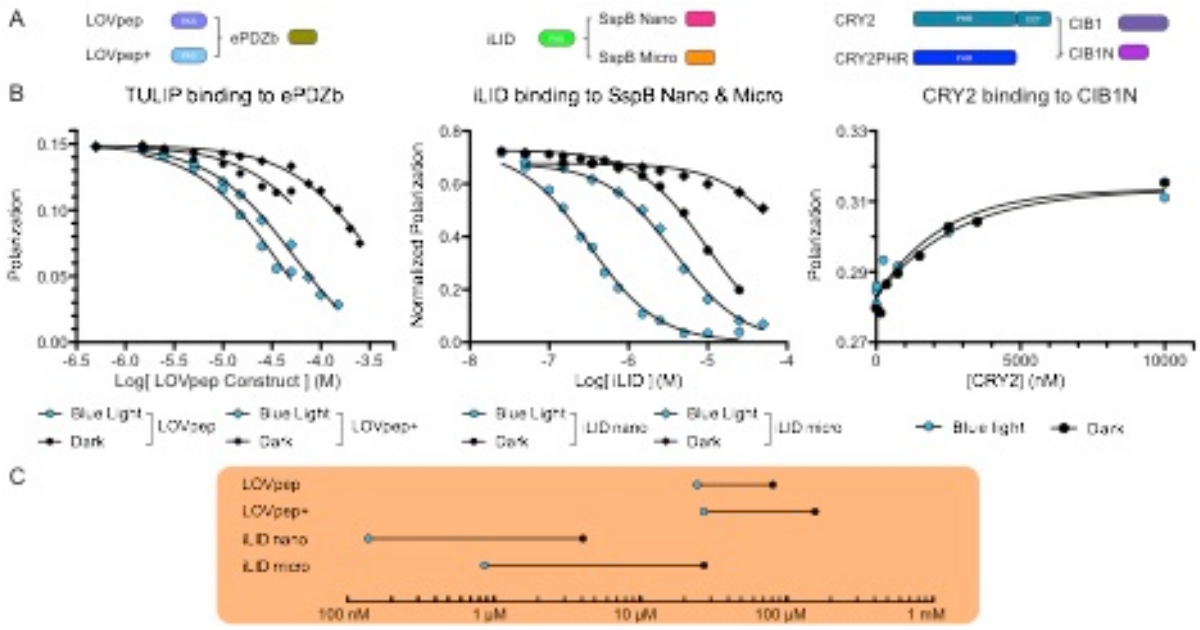


Figure 4.1 Construct schematic and binding affinities of heterodimer pairs. Binding affinities of lit and dark states highlight difference in photoswitch dynamic range. A) Constructs used in this chapter. B) Fluorescence polarization binding plots for LOVpep constructs and ePDZb (left) iLID nano and micro (middle) and CRY2 and CIB1N (right). C) Affinity values from binding data plotted on a Dynagram to highlight dynamic range of each tool.

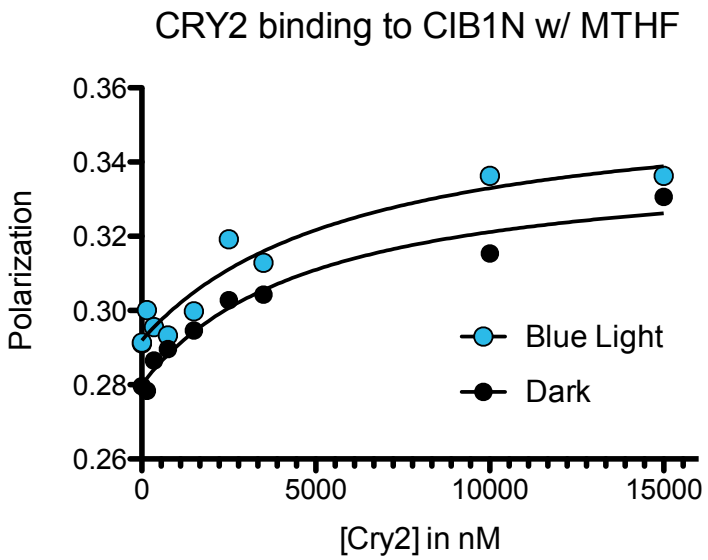


Figure 4.2 CRY binding to CIB1 with MTHF. Addition of MTHF does not significantly change CRY2-CIB1N binding.

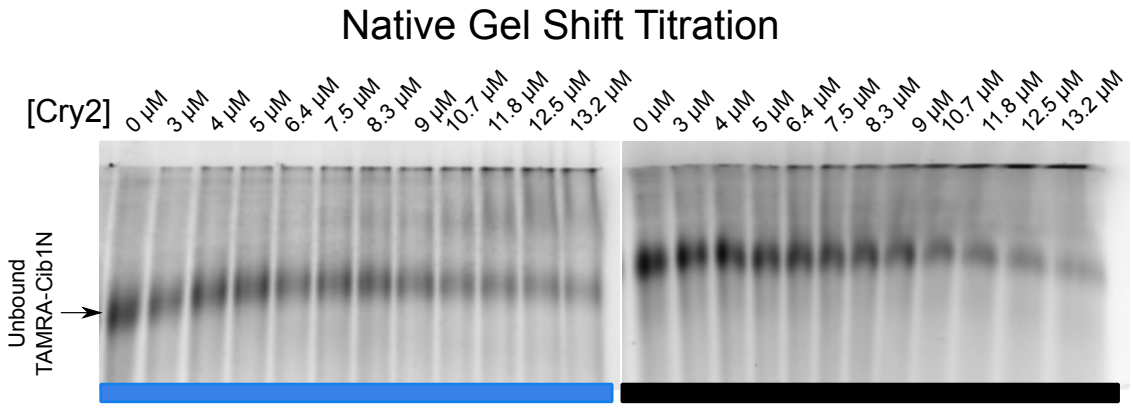


Figure 4.3 CRY2/CIB1 native gel shift binding assay. Binding measured by native gel shift indicates weak binding.

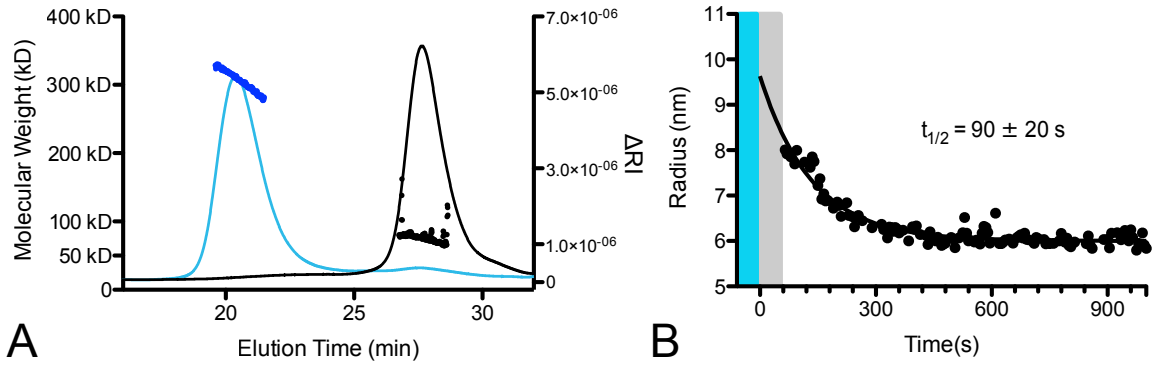


Figure 4.4 CRY2 multi-angle light scattering and dynamic light scattering. Light induces CRY2 oligomerization. Size exclusion chromatography multi-angle light scattering traces for full length CRY2 run under blue light (blue line) or darkness (black line). Fit molecular weight from MALS data for each peak is shown for lit (blue dots) and dark (black dots) peaks.

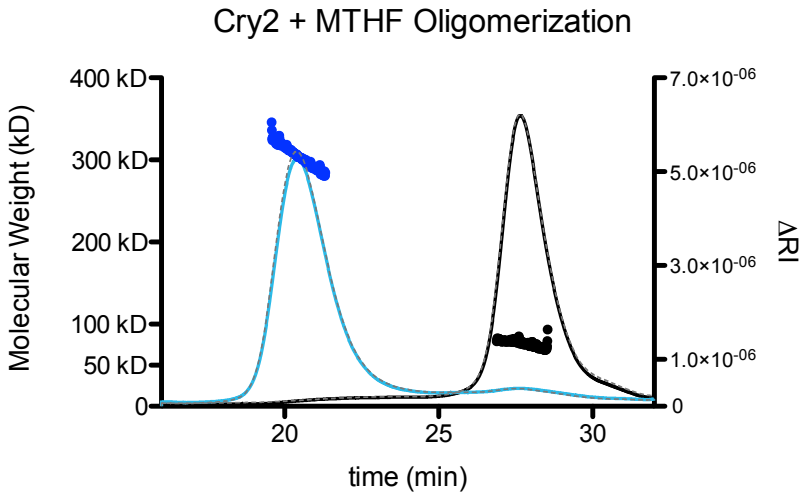


Figure 4.5 CRY2 multi-angle light scattering with MTHF. Addition of MTHF does not change CRY2 oligomerization. Size exclusion chromatography multi-angle light scattering traces for full length CRY2 with saturating MTHF cofactor added run under blue light (blue line) or darkness (black line). Fit molecular weight from MALS data for each peak is shown for lit (blue dots) and dark (black dots) peaks. Elution traces from Figure 4.4 are shown in grey dashed lines for comparison.

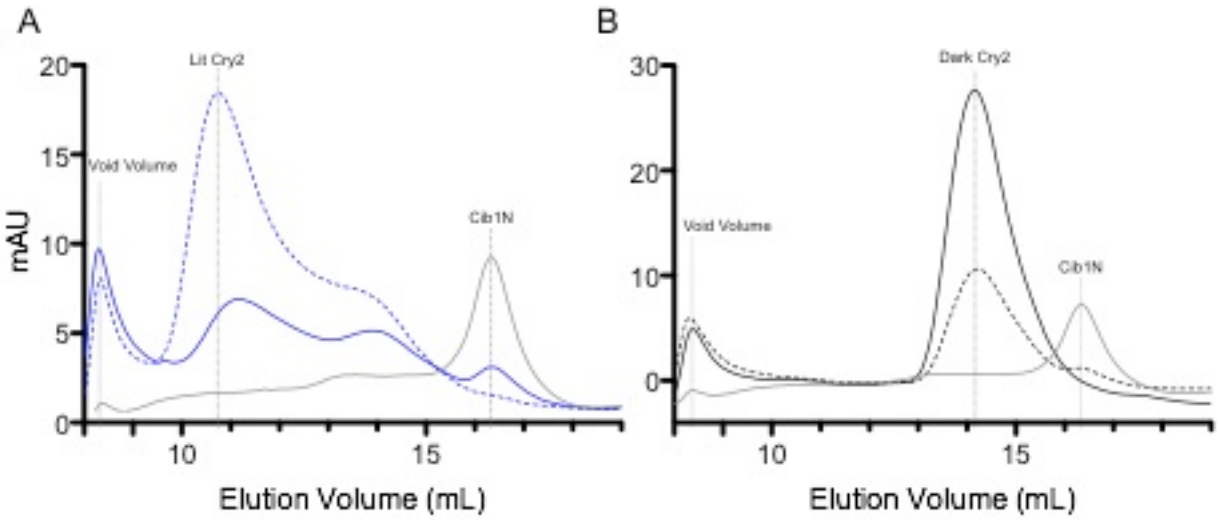


Figure 4.6 CRY2 and CIB1N size exclusion chromatography. CRY2 and CIB1N do not fully co-elute in size exclusion chromatography.

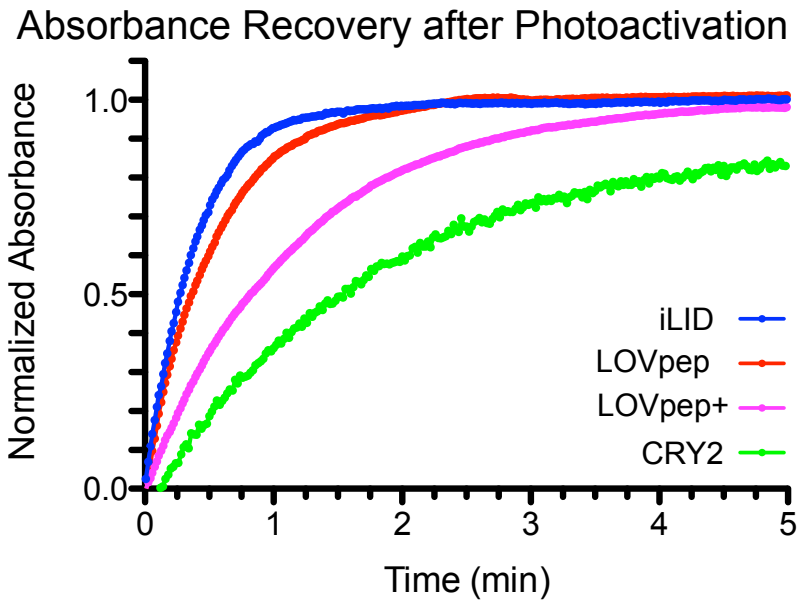


Figure 4.7 Photoreceptor reversion kinetics. Thermal reversion kinetics of the excited state for each photoreceptor show differences in timescale of deactivation. Reversions were measured at room temperature in Tris-HCl buffer.

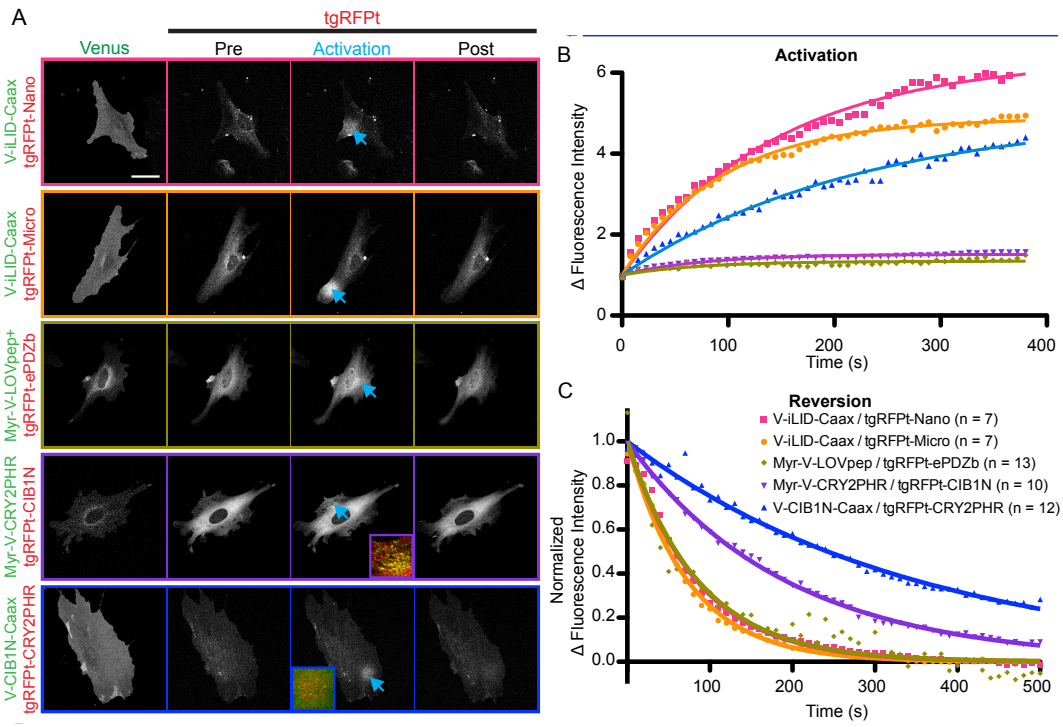


Figure 4.8 Photoswitch membrane localization kinetics. Targeted localization to the plasma membrane shows differences in switch dynamic range and kinetics. A) Representative images of the data analyzed in B and C. Cells transfected with each membrane bound switch pair were visualized and activated by confocal microscopy. Venus labeled constructs are bound to the plasma membrane while tgRFPt labeled constructs are cytoplasmic. The activated ROI is identified by the blue arrow. The activation and post activation images represent the final image of the specified time frame. (Bar = 50 μm) B) A ratio of tgRFPt fluorescence intensity inside the activated ROI to outside the activated ROI during the period of activation as shown in A. C) A normalized ratio of tgRFPt fluorescence intensity inside the activated ROI to outside the activated ROI during the period of activation as shown in A.

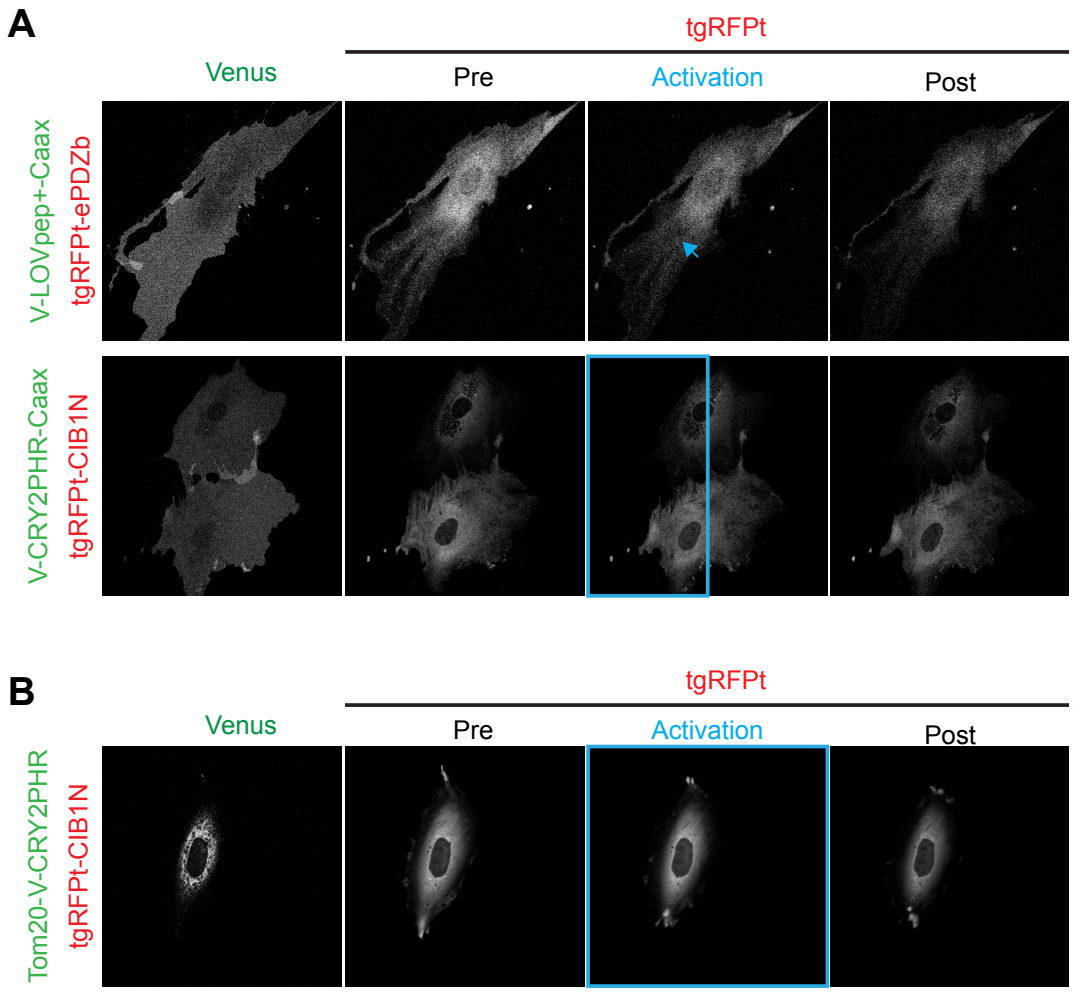


Figure 4.9 Orientation specific localization for CRY2 and TULIPs. For CRY2PHR/CIB1N and TULIPs, orientation matters for function. a) CRY2 and TULIP switches do not function in particular orientations in respect to the plasma membrane localization. b) CRY2PHR anchored to the mitochondria is toxic to cells and does not function in binding CIB1N upon activation. (Bar = 50 μm)

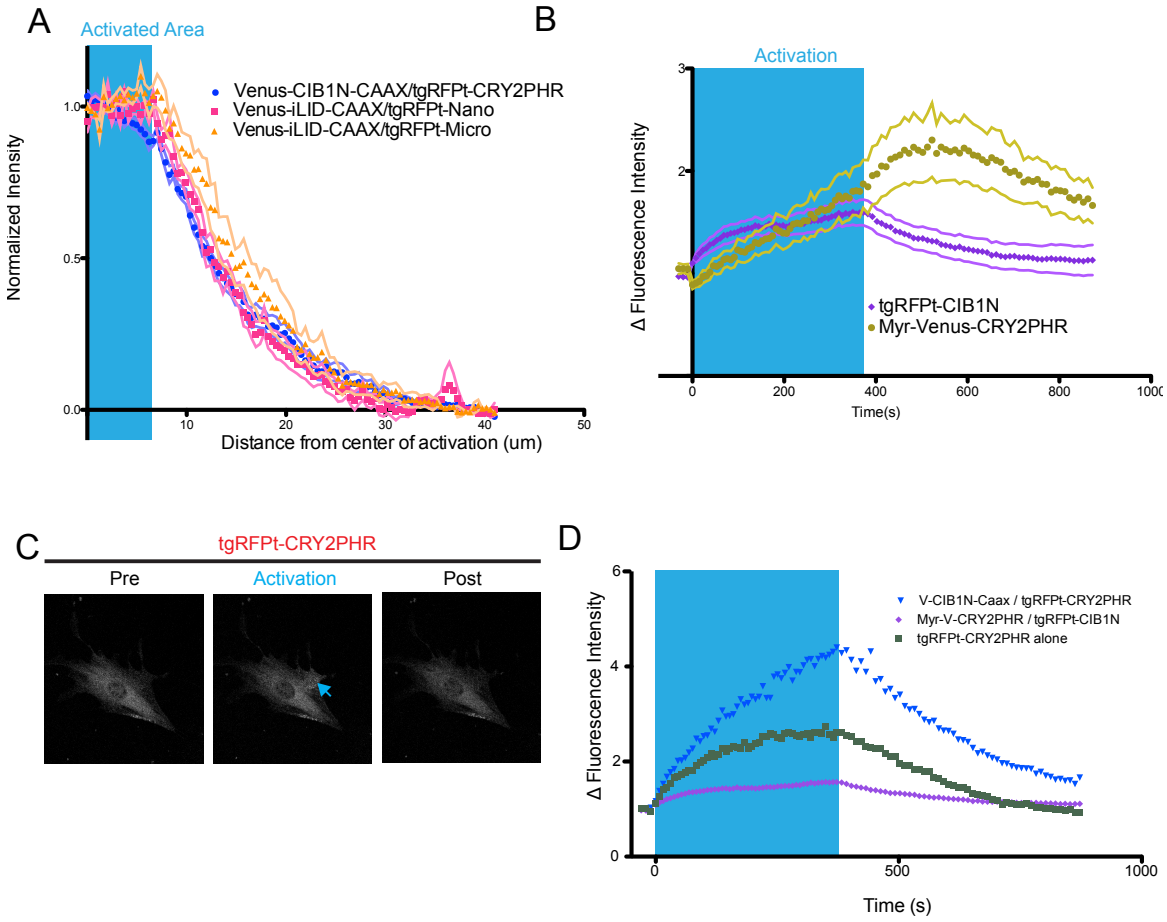


Figure 4.10 CRY2PHR oligomerization effects switching in unexpected ways A) CRY2PHR in the cytoplasm has similar spatial resolution of activation to the iLID switches anchored to the membrane. B) Myr-Venus-CRY2PHR clusters form slower than CIB1N binding occurs and continues to form after stimulation. C) Representative images of tgRFPt-CRY2PHR alone signal increase upon activation. (Bar = 50 μm) D) Quantification of tgRFPt-CRY2PHR alone signal increase upon activation as compared to CRY2PHR in combination with CIB1N.

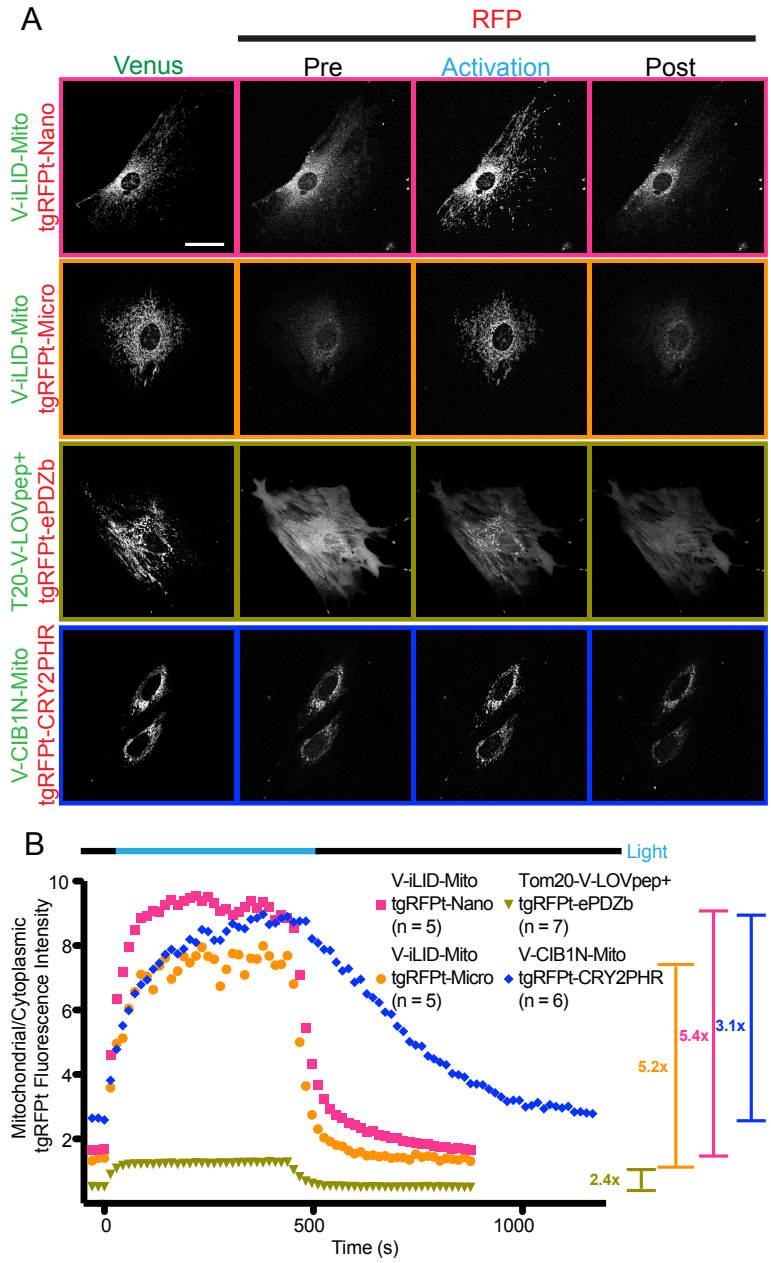


Figure 4.11 Photoswitch mitochondrial localization. Targeted mitochondrial localization identifies differences in dark state binding dynamic range and kinetics. A) Representative images of the data analyzed in B. Cells transfected with each mitochondrial bound switch pair were visualized and activated by confocal microscopy. Venus labeled constructs are bound to the plasma membrane while tgRFPt labeled constructs are cytoplasmic. The entire field of view is activated. The activation and post activation images represent the final image of the specified time frame. (Bar = 50 μm) B) A ratio of mitochondrial to cytoplasmic tgRFPt fluorescence intensity throughout the experiments shown in A.

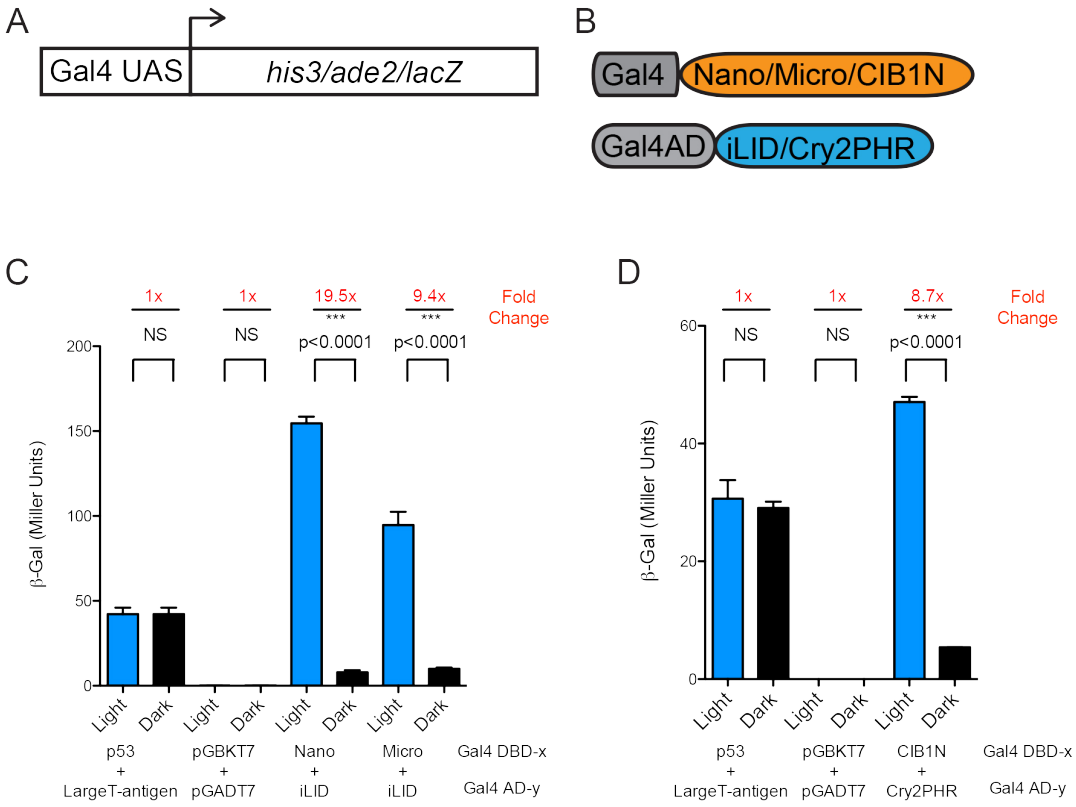


Figure 4.12 Yeast two hybrid transcription comparison A) A schematic of the genome reporters. B) A schematic of the constructs tested. C) ß-galactose transcription induced with the iLID paired with Nano or the Micro (triplicates from 3 independent experiments, p<0.0001) and D) CIB1N with CRY2PHR (triplicates from 3 independent experiments, p<0.0001) (Blue Bars – growth under continuous blue light at 465nm, Black Bars – growth in the dark).

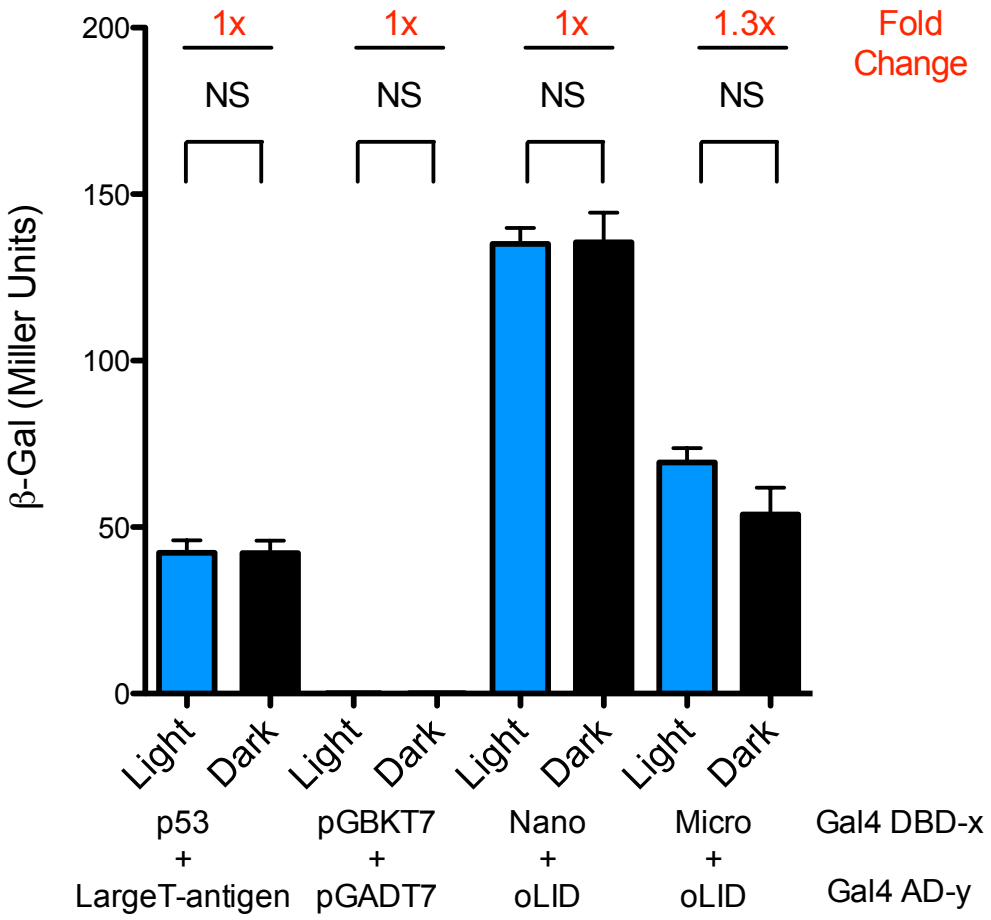


Figure 4.13 oLID yeast two-hybrid. oLID does not allow light dependent transcription in yeast

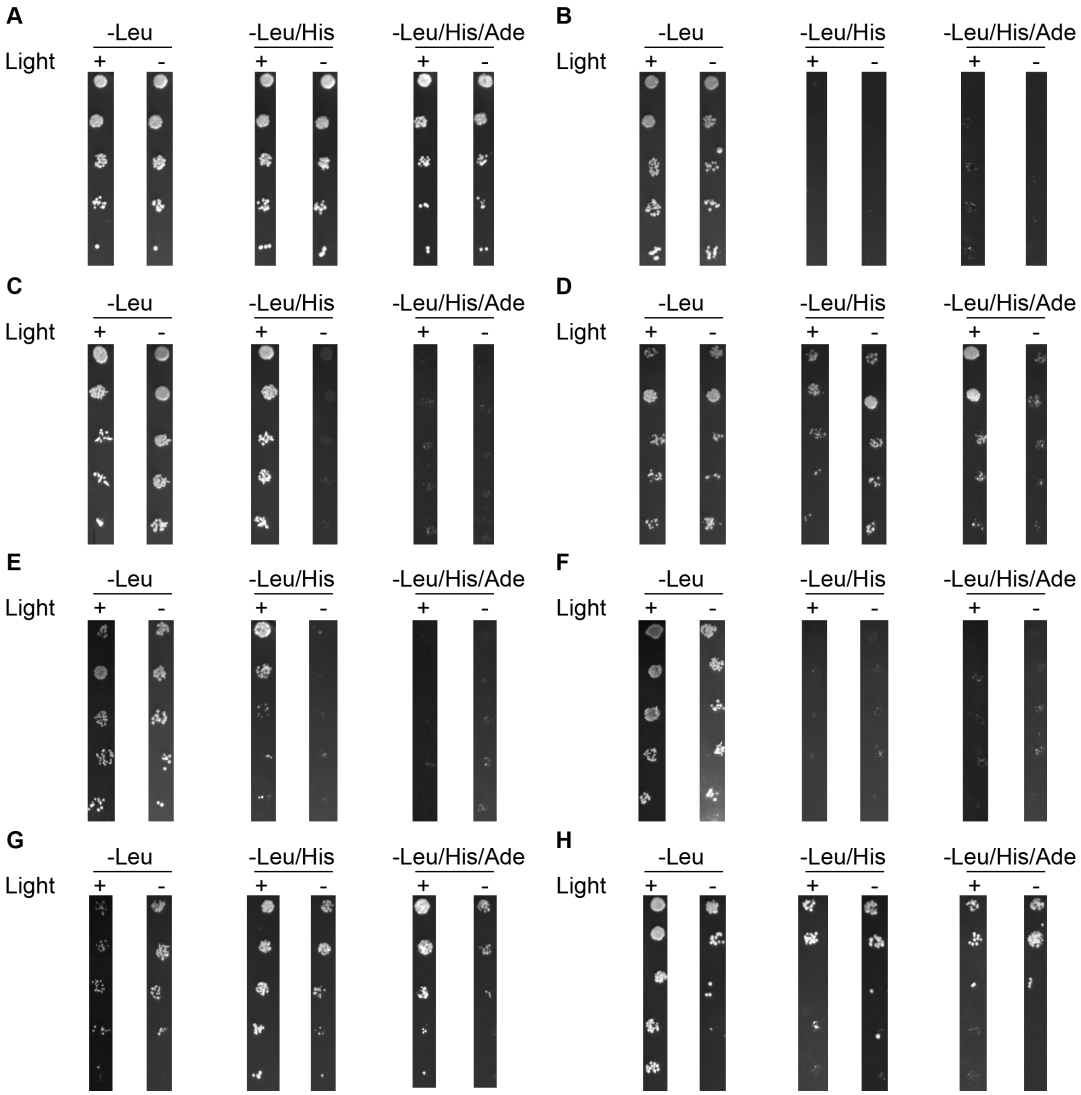


Figure 4.14 Light dependent yeast growth. Light dependent yeast growth on dropout plates demonstrates that low-level transcription is achieved for ePDZb-LOVpep and CRY2PHR-CIB1N A) Y2H positive control p53-DBD paired with Large T-antigen-AD, B) Y2H negative control empty pGBKT7 and pGADT7 vectors, C) CRY2PHR-DBD paired with CIB1N-AD, D) CIB1N-DBD paired with CRY2PHR-AD, E) ePDZb-DBD paired with LOVpep, F) ePDZb-DBD paired with LOVpep+, G) Nano-DBD paired with iLID-AD, H) Micro-DBD paired with iLID-AD.

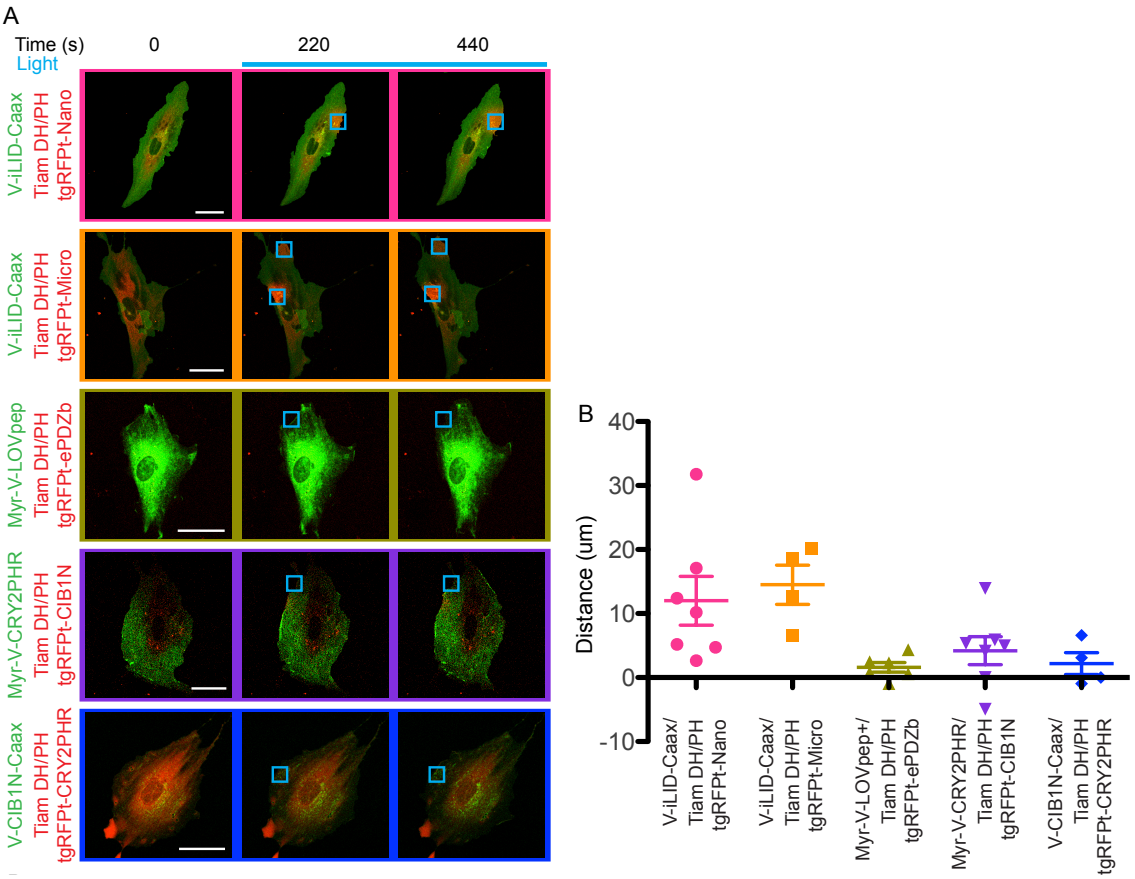


Figure 4.15 DHPH domain targeting to membrane. Targeting Tiam DH/PH domains to the plasma membrane with each switch causes varying degrees of protrusion. A) Representative images of the data analyzed in B. Cells transfected with each membrane bound Tiam DH/PH switch pair were visualized and activated by confocal microscopy. Venus labeled constructs are bound to the plasma membrane while tgRFPt labeled constructs are cytoplasmic. The activated ROI is is represented by the blue square. (Bar = 50 μm) B) Protrusion distances for each cell were measured by kymography.

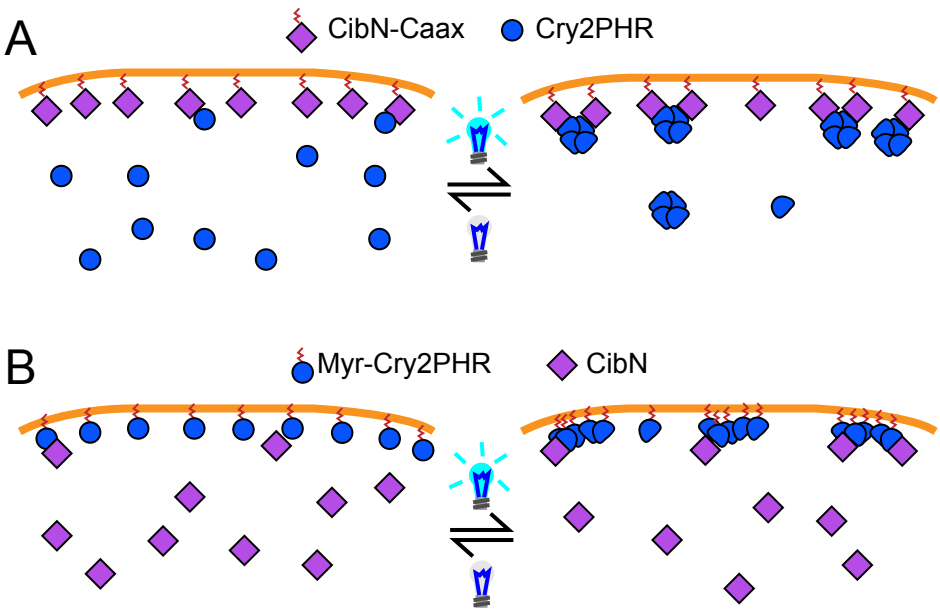


Figure 4.16 Proposed avidity mechanism for CRY2 binding to CIB1N. A) CIB1N imbedded in the membrane has increased affinity for oligomers of CRY2. B) CRY2 clustering on the membrane has minimal increased for affinity for CIB1N and possible occludes CIB1N binding.

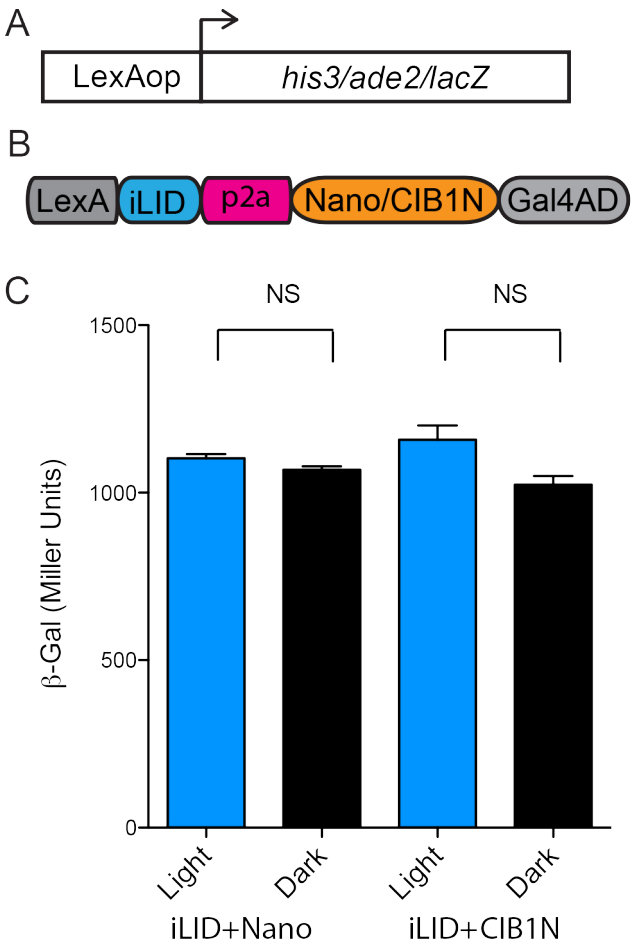


Figure 4.17 Incomplete P2A cleavage activates transcription in yeast two-hybrid. Light-inducible dimers are not functional if expressed with p2a self-cleavage peptide in yeast transcription.

Chapter 5

Light Mediated Ubiquitin Transfer

We aim for iLID to become widely available and useful; a generalizable heterodimerization tool used to modulate and study as many pathways and organisms as possible. However, from onset we have sought to use this pair in a very specific application, the development of a light-controlled ubiquitin pathway. This chapter covers our unpublished and ongoing efforts to engineer this pathway. Joseph Harrison aided in design and choice of E3 ligases. The LOV Trap is currently unpublished and was developed in the laboratory of Klaus Hahn. This work was funded by and performed under the direction of Brian Kuhlman.

5.1 Introduction

One of the most robust methods for determining protein function *in vivo* has simply been to remove expression at the genetic level and look for observable changes in cell function. As making knockout cell lines tends to be costly and time consuming, considerable effort has been directed toward developing simpler methodologies like CRISPR/Cas9 and RNA interference (siRNA or shRNA) that allow for quicker turnaround times[160,161]. However, even with these next generation methods there is a need for a set of tools that allows for the control of protein levels in real time. This is especially apparent when probing the function of embryonic lethal genes where knockouts do not live long enough to conduct experiments. Methods to conditionally control protein degradation afford the ability to probe protein function in normal healthy cells[162]. Conditional control has the

added benefit of working on a short enough time scale to minimize compensation form alternate pathways. An ideal tool to control protein levels inside cells would have a short response time, be reversible, not interfere with native pathways and target endogenous proteins.

Initial attempts at conditional control of protein stability have been achieved through fusions with conditionally stable proteins (temperature sensitive DHFR, N-end rule sequences), localization with ubiquitin E3 ligases, and localization with the proteasome[163–166]. As changes in temperature can alter cellular function and N-end rule instability is irreversible, we have chosen to create a tool that causes localization with an ubiquitin E3 ligase to induce degradation[165,167,168]. This route has the added benefit of being useful to probe the differences in E3 ligase function and activity. Previous methods of induced localization with an E3 ligase have used protein-protein interactions induced by small molecules to control degradation[165,169]. In one example a chimeric small molecule that binds FKBP12 was fused to a peptide that binds the VHL E3 ligase[169]. Upon addition of the small molecule, the group was able to localize a target protein fused to FKB12 with the E3 ligase[169]. In this way, a target protein can be conditionally targeted for ubiquitination and subsequent degradation. The limitation of this approach is that the interaction is irreversible once the small molecule is added. To increase the temporal and spatial resolution of this approach we aim to replace the FKBP/small molecule/E3 ligase interaction with the light inducible heterodimer interaction developed in the previous chapters (LOV-SsrA/oLID and iLID) as well as a second AsLOV2 based heterodimer, LOV Trap[82,152].

Both LOV-SsrA (or iLID) and LOV Trap are light mediated heterodimeric interactions, however they work in opposing ways. LOV-SsrA forms a heterodimer with its partner, SspB, upon irradiation with blue light and dissociates back to monomers in the dark[82,152]. LOV Trap is a heterodimeric interaction between AsLOV2 and an engineered Z-domain, ZA127. In the dark AsLOV2 binds ZA127 and upon irradiation with blue light, dissociates into

monomers (unpublished, Hahn lab). We have chosen these photoreceptors for their small, modular nature, and their known mechanism of action. Additionally, each of the binding partners, ZA127 and SspB, only have a few lysine residues, allowing us to remove them to prevent self-ubiquitination. Finally, both of these switches could be used together since AsLOV2 doesn't bind SspB and ZA127 doesn't bind LOV-SsrA (unpublished data). We have chosen to fuse E3 RING ligases from three proteins IAP2, UHRF1, and RNF8 to the binding partners. The protein of interest will then be fused to the respective AsLOV2 photoreceptor. In this way we hope to achieve light activated ubiquitin transfer with LOV-SsrA (or iLID) and light de-activated ubiquitination with LOV Trap (Figure 5.1).

We have demonstrated with *in vitro* ubiquitination assays that both the LOV-SsrA and LOV Trap switches allow for increased ubiquitination in their heterodimer state over the monomeric state. Thus far we have not seen light mediated ubiquitin transfer or degradation in mammalian cells, however, this work is ongoing. We hope to complete this work in the future and create both light induced and light inactivated degradation tools as well as probe the function of each E3 ligase.

5.2 Materials and Methods

5.2.1 Protein Expression & Purification

All AsLOV2 clones were expressed as MBP fusions in a modified pQE-80L vector. In these constructs, MBP is N-terminal to AsLOV2 separated by a TEV protease cleavage site. SspB or ZA127 / Ring E3 fusions were expressed with N-terminal 6xHis tags in the pQE-80L vectors. In cases where expression was poor, they were re-cloned and expressed as MBP fusions. All proteins were expressed in BL21 *E. coli* cells for approximately 16 hours overnight. Post-expression, cells were spun down at 3000rpm and frozen until purification. Cells were resuspended in phosphate lysis buffer and purified via nickel affinity

chromatography as described in previous chapters. MBP-AsLOV2 and MBP-LOV-SsrA remained uncleaved, while any ring fusion was cleaved from a MBP tag with TEV protease, and re-run over a nickel column to remove MBP, TEV, and uncleaved fusion. As a final step of purification and buffer exchange, size exclusion chromatography on a S75 column equilibrated in HEPES buffer compatible with ubiquination assays (25 mM HEPES, 150 mM NaCl, pH 7.5) was performed.

5.2.2 *In vitro* Ubiquitination Assays

All *in vitro* ubiquitination assays were performed in 25 mM HEPES, 150 mM NaCl, pH 7.5 buffer at 25° C. Reactions consisted of 0.2 μM Uba1 (E1), 0.5 μM Ubc5A (E2), 0.2 μM (IAP2 Fusion), 16 μM flag-ubiquitin, 50 mM ATP, 1 mM DTT, and 0.05 μM target protein (MBP-AsLOV2 or MBP-SsrA). This assay was loosely based on previous ubiquitination experiments[123]. Enough reaction mixture for 2 full experiments, minus ATP, was added to a 0.5 mL tube. The mixture was then split into two clear tubes, one placed under blue light and one in darkness. Once ATP was added, time points were taken at intervals shown and quenched in 6x SDS-loading buffer. Samples were run on 12% SDS-PAGE gels, transferred to nitrocellulose membrane at 30 volts overnight and probed with fluorescent anti-Flag or anti-MBP antibodies. Gels were imaged with a GE lifesciences Typhoon imager and quantified with Image Quant software.

5.2.3 Mammalian Clones and Methods

Mammalian clones were generated according to Figure 5.2. A single plasmid system was devised with the following scheme; [SspB Nano/Micro]-HA-[$IAP2_{ring}$/$UHRF1_{ring}$/ $RNF8_{ring}$]-2A peptide-[mKate2]-[6xHis MBP]-[iLID/SsrA/Empty] or [SspB Nano/Micro]-HA-[$IAP2_{ring}$/$UHRF1_{ring}$/ $RNF8_{ring}$]-2A peptide-[N-terminal Stargazin fragment]-[iLID/SsrA/Empty]-[C-terminal stargazing fragment]-[mCherry] (Figure 5.2). Note, thus far only the positive (SsrA) and negative (empty) controls have been tested with RNF8 and IAP2 rings. HEK

293T cells were transfected with 1 µg plasmid using lipofectamine 2000 (Invitrogen). After 16 hours, cells were checked for fluorescence to indicate successful transfection. In order to prepare for flow cytometry, cells were trypsinized for 5 minutes at 37° C, and then quenched with FBS containing serum. Cells were gently mixed to break up clumping and then transferred to a 15 mL conical to pellet. Cells were pelleted at 1000 rpm for 3 minutes, decanted, and resuspended in phosphate buffered saline supplemented with 10 µg/mL DNAse, 100 units/mL penicillin, 100 µg/mL streptomycin and 5 mM EDTA. Resuspended cells were passed through a 30 µM CellTrics sterile filter. Fluorescence was measured on a BioRad S3 Cell Sorter and data was analyzed with BioRad ProSort software.

5.3 Results

5.3.1 In Vitro Ubiquitination Assays

The IAP2 RING domain SspB and ZA127 fusions were well behaved and functional in *in vitro* ubiquitination assays. In both fusions we noticed the formation of poly-ubiquitin chains in reactions missing the target protein, however this background activity was small in comparison to the total ubiquitination seen when target protein was present. In both cases, we demonstrated preferential light mediated ubiquitin transfer in the heterodimer state, under blue light for LOV-SsrA and in darkness LOV Trap. For MBP-LOV-SsrA/SspB-IAP2 RING, we saw depletion of unmodified MBP-LOV-SsrA in between 120 and 180 seconds when under blue light (Figure 5.3). The same time points had approximately 55% and 35% unmodified target remaining when left in darkness (Figure 5.3, 5.4). Not only did the unmodified target band disappear, we noticed the time dependent appearance of higher molecular weight species indicating that ubiquitin was being covalently attached to the target. We were able to confirm the presence of FLAG-ubiquitin in these higher molecular weight species by western blot (Figure 5.3, bottom). For the LOV Trap heterodimer system,

we saw the opposite change due to light just as expected (Figure 5.5). Probing for MBP, we noticed almost complete depletion of MBP-AsLOV2 in the dark after 120 seconds (Figure 5.5). In comparison, after 120 seconds > 55% of MBP-AsLOV2 remained unmodified when kept under blue light (Figure 5.6). Again, an anti-FLAG western confirmed the presence of ubiquitin in the higher molecular weight species (Figure 5.5, bottom). Unfortunately, RNF8 suffered from poor expression both when fused with SspB and ZA127. Purification of both clones yielded almost no protein at all. We re-cloned both as MBP-fusions (for stability and increased expression), however, upon cleavage of the MBP tag, both crashed out of solution again. The UHRF1 clones are cloned but remain untested in the *in vitro* assays.

5.3.2 Mammalian Ubiquitination

We transfected HEK 293T cells with the mammalian expression clones shown in Figure 5.2. It is expected that the polypeptide chain will be cleaved into two pieces during translation at the 2A peptide leaving an N-terminal fragment containing SspB-IAP2 RING and a C-terminal fragment containing the target protein, mKate2-MBP-iLID/SsrA/Nothing. After confirmation of successful transfection we quantified mKate2 fluorescence with flow cytometry (Figure 5.5). However, we did not see the expected change in fluorescence levels for the mKate2-MBP-SsrA and mKate2-MBP targets (Figure 5.7A). In fact, it appears that the SsrA tagged target protein is found at higher levels than its untagged counterpart. We were also able to test a second fluorescent target, the transmembrane Stargazin-mCherry fusion (+ or - SsrA tag), with a set of SspB-RNF8 E3 ligase fusions (Figure 5.2). In this case, the SsrA tagged Stargazin-mCherry was slightly higher in fluorescence than the untagged version (Figure 5.7B). In this experiment, we also transfected the Stargazin-mCherry clone without an E3, to determine if the SspB-Ring fusions were having any effect. It does appear that the fluorescence is higher without an E3 present; however, this could

simply be due to the expression of a single protein instead of two (Figure 5.7B). None of the LOV Trap clones or UHRF1 clones have been cloned or tested currently.

5.4 Discussion

The *in vitro* ubiquitination assays performed are encouraging in that the switches appear to function in the intended fashion. Our LOV-SsrA system preferentially transfers ubiquitin in the light and the LOV Trap system preferentially transfers ubiquitin in the dark. The two systems switch over different binding affinity ranges, which gives credence to this as a robust process *in vitro* and likely to work for other light-inducible heterodimerization. *In vitro* binding affinities measure the LOV Trap dark state interaction to be 9 ± 2 nM and 600 ± 200 nM under blue light, a sixty fold change in affinity (unpublished, communication with Ryan Hallett). The LOV-SsrA/SspB lit state interaction has been measured at 120 nM, which drops to 900 nM in the dark, an 8-fold change[82]. As such we would expect the LOV Trap to have a higher Lit/Dark fraction of unmodified target at each timepoint than the corresponding LOV-SsrA Dark/Lit fraction of unmodified target. This appears to be the case (Figure 5.6 bottom and Figure 5.4 bottom), as the LOV-SsrA ratios lag behind the LOV Trap ratios. Since these experiments were performed, new versions of LOV-SsrA, iLID micro and iLID nano, have been developed with increased dynamic ranges. Due to increased dark state caging, we would expect iLID switches would outperform its predecessor. One interesting artifact of the different affinity ranges appears to be that the tighter affinity of dark and lit state binding of LOV Trap leads to longer ubiquitin chain formation (Figure 5.5). In the LOV-SsrA assays, MBP-LOV-SsrA's with one, two, and three ubiqutins are visible (Figure 5.3). These bands are not seen at all in the LOV Trap experiments. We hypothesize that for a tight affinity complex with a slow k_{off}, residence time near the E3 will be increased leading to longer poly ubiquitin chains. This subtle, but interesting, difference in the two switches highlights the information that could be revealed with a working set of

light mediated ubiquitin transfer components. The iLID micro pair has the weakest affinity range we have developed to day and would be perfect to test this hypothesis. The instability of purified RNF8 fusions prevented any direct comparisons between the E3 components. Most likely, the fragment of RNF8 used in these experiments is unstable alone and may need more of the native protein to properly fold.

Despite the successful *in vitro* light-mediated ubiquitin transfer, we have not seen a robust degradation phenotype. The cytometry data in Figure 5.7 is preliminary. We have yet to probe for expressed proteins by western blot analysis, so it is unclear even if and to what efficiency the 2A peptide system is working as intended. Additionally, we have not tested any of the LOV Trap constructs *in vivo* yet. However, even without adding the light conditional control, the positive (SsrA tag) and negative (no tag) controls do not behave as expected (Figure 5.7A). This is very discouraging because these constructs should be constitutively localized to the E3 or not localized to it at all. We do see lower fluorescence when the E3 is absent as compared to when it is present (Figure 5.7B), indicating the E3 is doing something, however it is unclear what that might be. These data suggest that the RING E3 ligases chosen may not be suitable for an *in vivo* degradation pathway.

The attempts at developing a light-mediated ubiquitin transfer pathway have been a mix of success and failure. The IAP2 RINGS are excellent mediators of ubiquitin transfer *in vitro*, however do not appear to be fully modular, as the same constructs do not induce degradation *in vivo*. The *in vitro* assays shown could be an interesting way to probe E3 ligase function, especially as a function of affinity to substrate proteins, however it is limited to E3 ligases that can be expressed. Moving forward with the light-induced ubiquitination and degradation pathway, new E3 proteins from a variety of families should be attempted as well as E3 ligase degrons as seen in previous work[165,169]. As we have shown *in vitro* and *in vivo* activities do not necessarily correlate, an *in vivo* selection strategy should be devised to rigorously test E3s as well as linkers and target proteins.

5.5 Figures

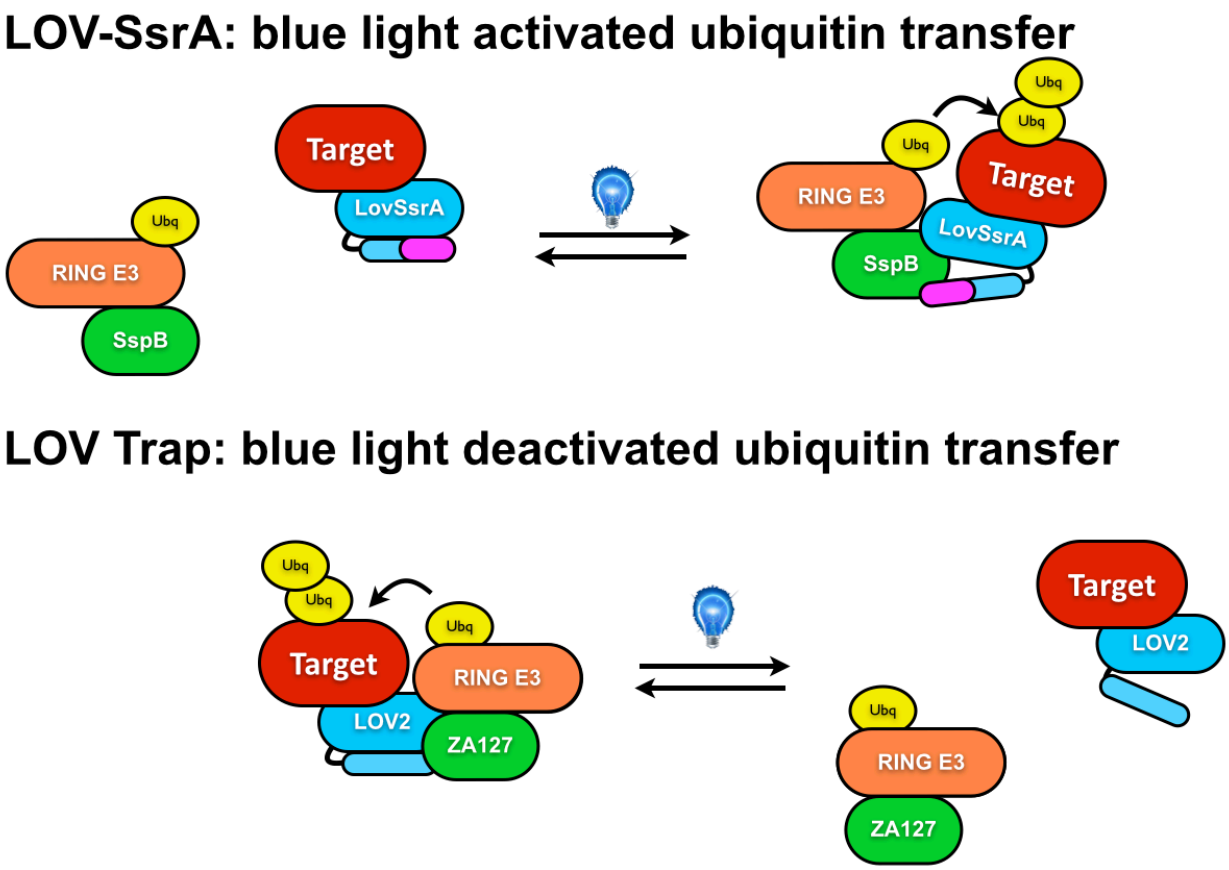


Figure 5.1 Light Mediated Ubiquitination Schematic. Light mediated ubiquitin transfer through co-localization of an RING E3 ligase with a target protein by light induced (top) or light disrupted (bottom) heterodimer.

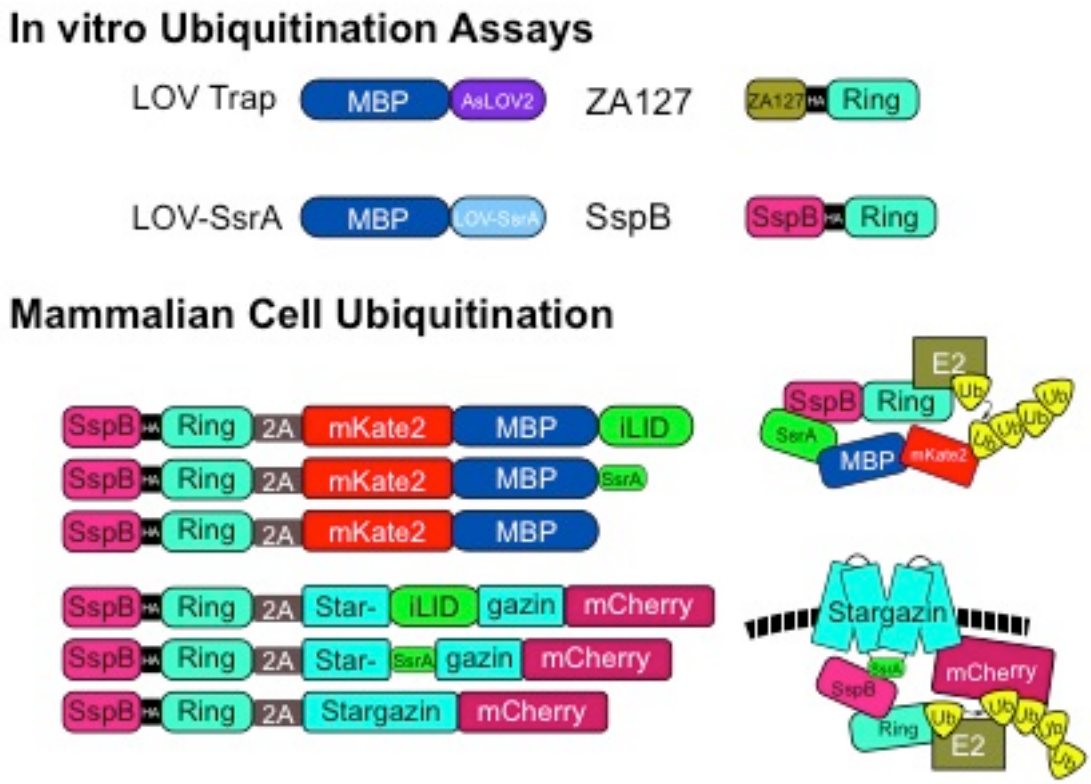


Figure 5.2 Clone schematic. Clone schematic for *in vitro* (Top) and *in vivo* (Bottom) ubiquitination assays.

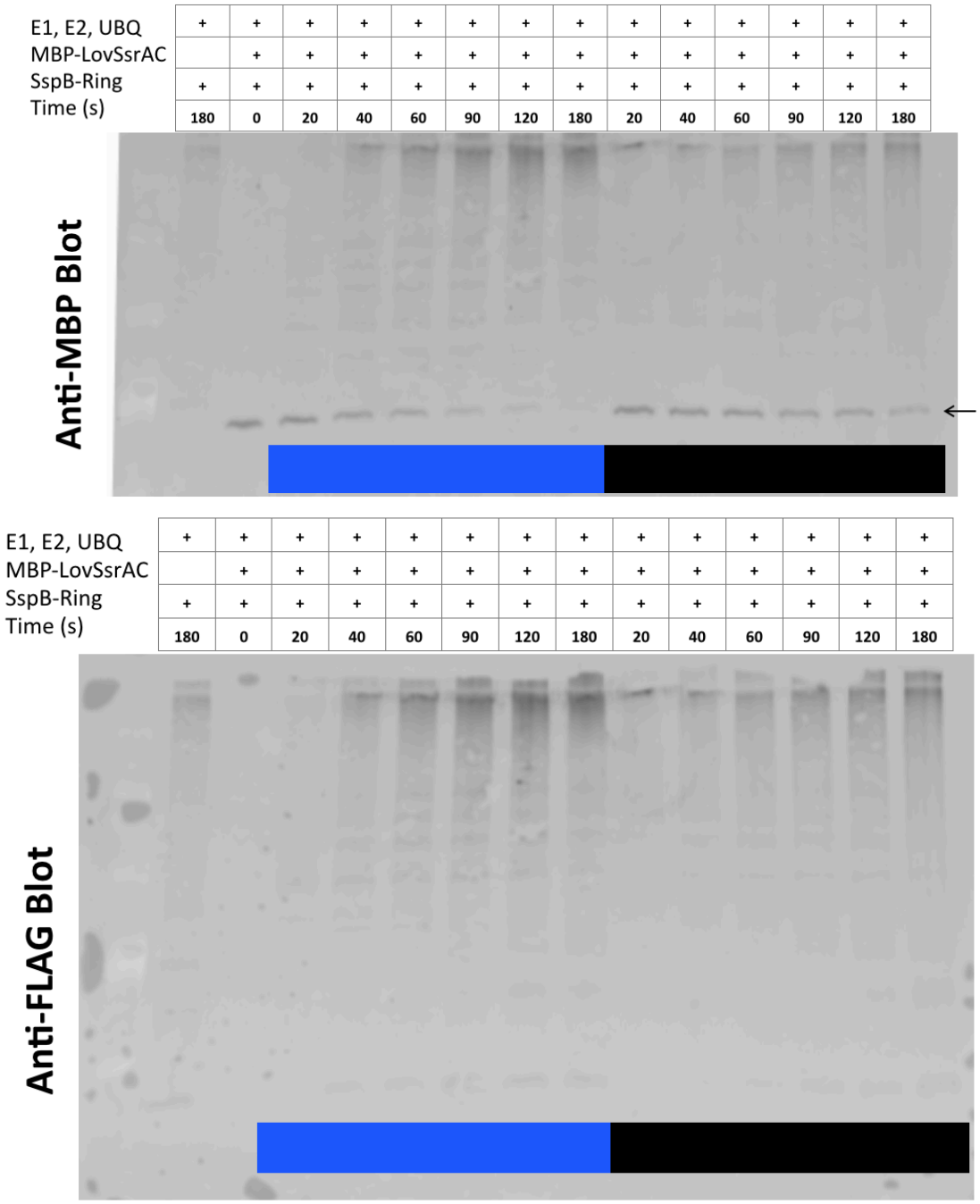


Figure 5.3 LOV-SsrA *in vitro* ubiquitination assay. Blue bar denotes samples irradiated with blue light, black bar denotes those kept in darkness. Arrow marks weight of unmodified target. Top) anti-MBP blot probing the decrease of unmodified of MBP-LOV-SsrA as ubiquitin transfer continues. Bottom) anti-Flag blot probing showing presence of ubiquitin in higher molecular weight species.

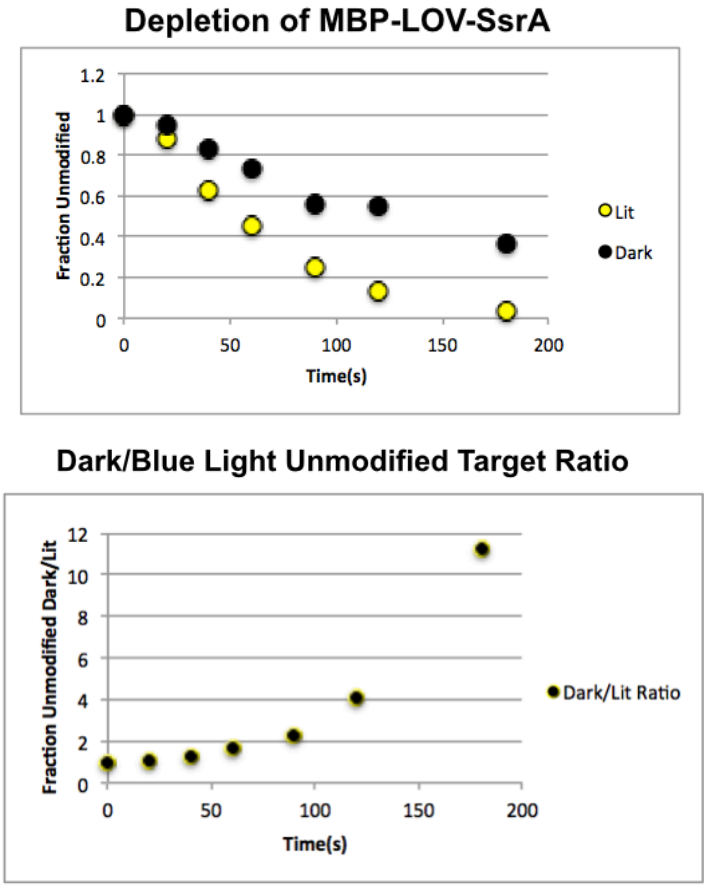


Figure 5.4 Quantification of LOV-SsrA *in vitro* ubiquitination assay. LOV-SsrA preferentially transfers ubiquitin to MBP protein in the light compared to the dark.

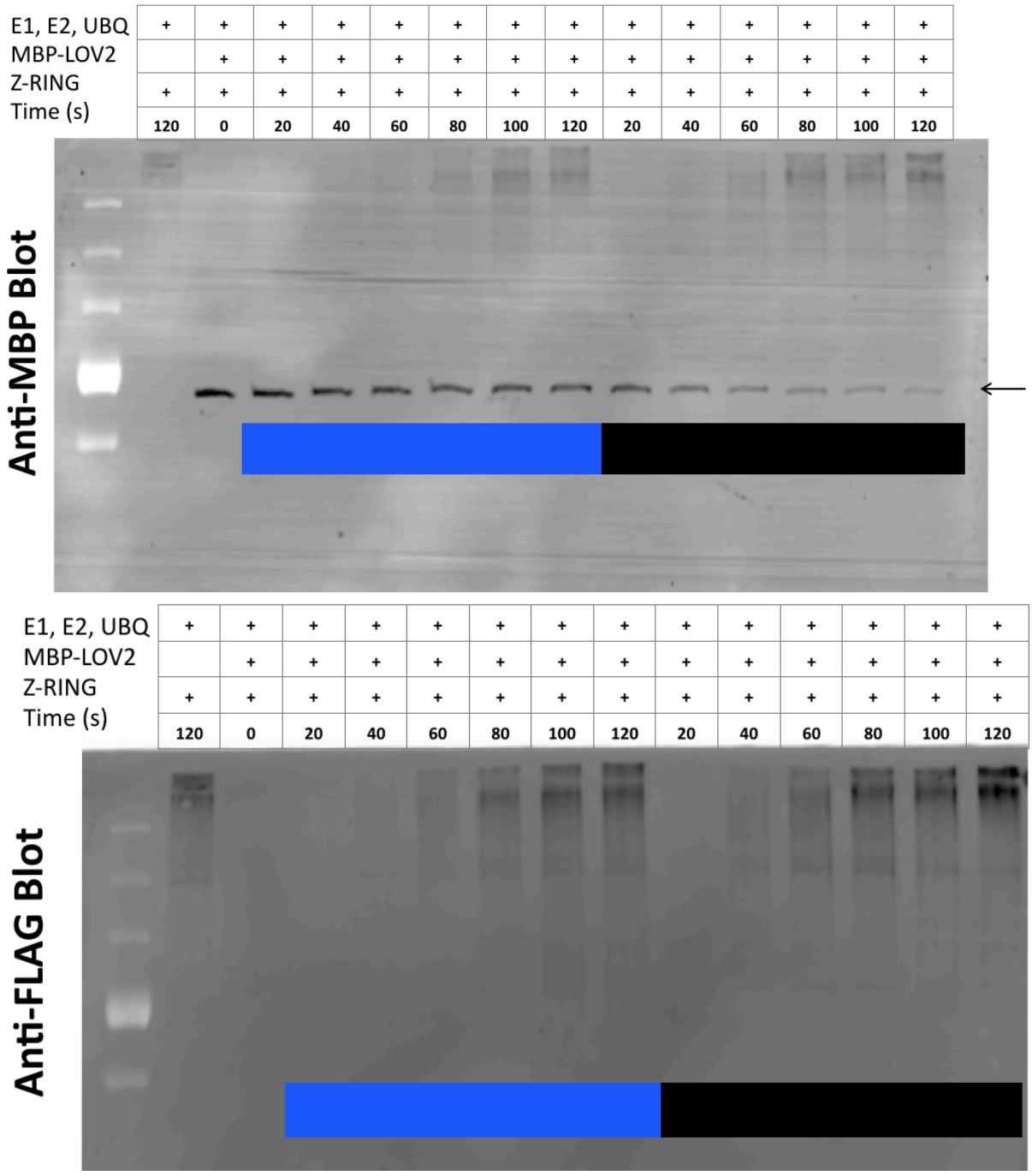


Figure 5.5 LOV Trap *in vitro* ubiquitination assay. Blue bar denotes samples irradiated with blue light, black bar denotes those kept in darkness. Arrow marks weight of unmodified target. Top) anti-MBP blot probing the decrease of unmodified MBP-AsLOV2 as ubiquitin transfer continues. Bottom) anti-Flag blot probing showing presence of ubiquitin in higher molecular weight species.

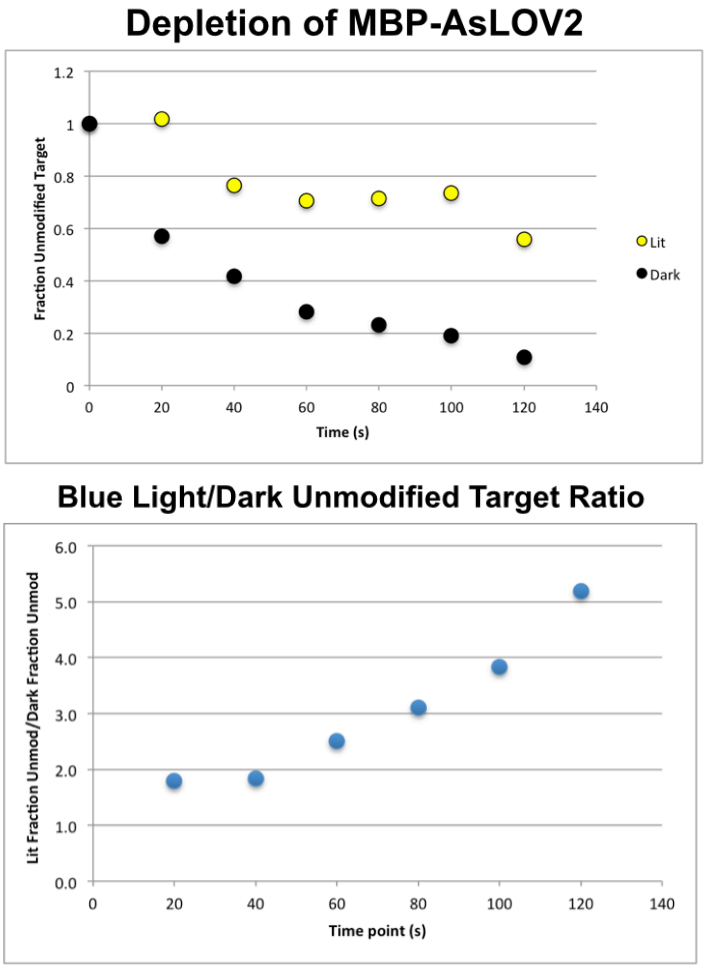


Figure 5.6 Quantification of LOV Trap *in vitro* ubiquitination assays. LOV TRAP preferentially transfers ubiquitin to MBP protein in the dark compared to the light.

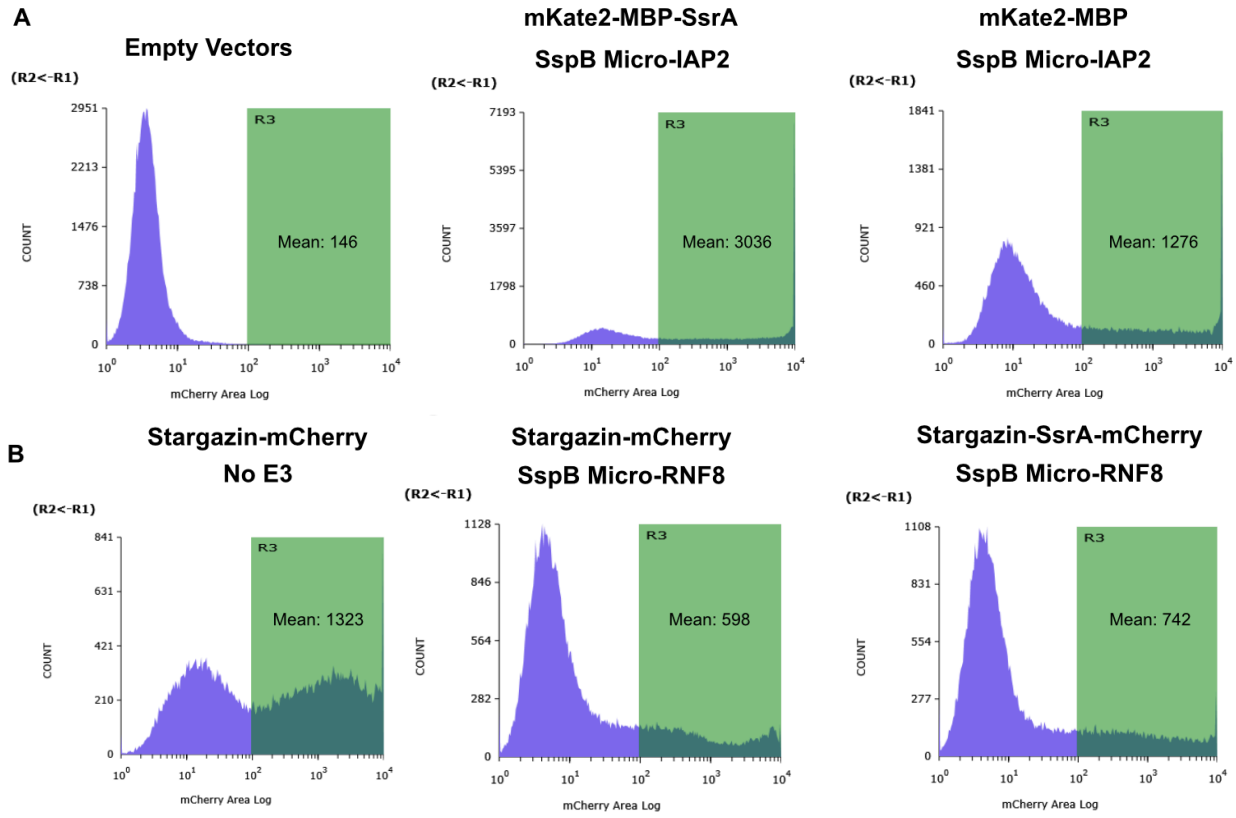


Figure 5.7 Flow cytometry of mammalian cell constructs. As measured by fluorescence, SsrA tagged substrates are not degraded when expressed with SspB-E3 RING fusion

Chapter 6

Conclusions and Final Thoughts

6.1 Light Inducible Peptide Caging

Light inducible caging has been shown to be an effective method of controlling functional protein surfaces, peptide sequences, and entire protein domains. Peptide caging thus far has relied primarily on the undocking of the Jα helix of the *As*LOV2 domain. The work described in chapters 2, 3, and 4 shows functional peptide caging for three peptide sequences. In each case, peptide binding to its target protein is substantially increased in the presence of light and this change in affinity can be used to induce specified phenotype changes in an organism or cell. This method is highly versatile, as the iLID and TULIP switches come in a variety of dynamic ranges and switch over multiple affinity ranges. These factors are critical when choosing the appropriate tool for a specific application. The variety of peptide-based switches available directly translates to a greater number of pathways that can be targeted and studied.

However, this methodology is largely limited by sequence identity to the protein it is to be incorporated within. As such, this technique does not have the power to study the biology of any given peptide simply by attachment to *As*LOV2. Peptides hung off the end of the Jα helix have no change of affinity/activity as a function of light. Peptide caging in other regions of AsLOV2, namely the A'α helix and the hinge connecting the PAS fold to Jα helix, do not appear to be generalizable sites of caging in our hands. One way around this limitation may be to specifically engineer AsLOV2 for a peptide with low sequence identity.

Our work in chapter 3, demonstrates AsLOV2 to be tolerant to a number of mutations, allowing for our modest 2-fold LOV-ssrA switch to be improved to over 50-fold as seen in iLID micro. Using a similar method of high-throughput library screening, it may be possible to engineer a set of compensatory mutations allowing for incorporation of any given peptide, but this most likely be more trouble than it's worth.

We believe the true power of peptide caging lies in the generalizable heterodimer tool generated in the process of caging the SsrA peptide. Our iLID nano and micro systems allow for precise colocalization of two proteins on the second timescale. This system can be used for light induced protein activation by fusing a functional domain of interest to one half of the switch and localizing the other half to the site of activity. The reverse, light induced sequestration is just as easily achieved by localizing the second half to an inactive area of the cell. The proteins have no restrictions N- or C-terminal fusions, resulting in a broad applicability with little to no use-to-use optimization. The iLID system has the added benefit of being orthogonal in higher organisms, meaning they won't interact with endogenous pathways and create experimental artifacts. All of these factors combined with one of the largest dynamic ranges measured (Chapter 3 & 4), indicate iLID nano & iLID micro to be some of the most powerful and versatile light-inducible heterodimer pairs to date.

6.2 Computationally Guided Improvement of Functional Protein Domains

In chapter 3, we used a protocol of computational library design, phage display, and ELISA screening to improve a light activated heterodimer from an 8-fold change in affinity to 36-fold. This method was very successful for improving our photoswitch, as all 4 of the top clones selected showed considerable improvement of dynamic range over the parental sequence. This is actually quite impressive as there were around a hundred rational designs tested with this same goal, but none were able to increase the dynamic range above 12-fold. This is an incredibly difficult protein design problem due to the lack of

structural information known about the lit state of AsLOV2. When analyzing the mutations that ended up in iLID, we get the picture that there are factors we don't fully understand that govern AsLOV2's allosteric mechanism. When looking at the Rosetta score of the final mutations, we noticed that half were predicted to be stabilizing and half as destabilizing. When reverting these mutations one by one, we see that they are all actually contributing to the improvement, however even by solving the structure of iLID, their specific function remains unknown. This explains the failure of rational or discrete computational design; however, highlights the strength for the computationally directed library and selection strategy.

Moving forward, there seems to be room for even greater improvement. We've only touched roughly a third of the residues in AsLOV2, leaving many more to optimize. We focused our initial library on regions within interacting distance of the Jα helix; however, many of the mutations in iLID don't appear to directly contact the Jα helix or SsrA sequence. One feature that may have contributed to success was the tight localization of the library, allowing for mutations to pack together as seen between the mutations in the hinge and Hβ strand. Future libraries could focus on alternate regions of AsLOV2, the A'α helix for one or FMN binding pocket for another.

This protocol could be extended to further improve AsLOV2 based photoswitches or other allosteric protein switches for which there is little known mechanistic information. The dual selection strategy, a single mutation library followed by recombination of top mutations, allows for increased coverage of protein sequence. The high-throughput phage display followed by low-throughput ELISA was useful in narrowing down the library to top clones. In a case where information about mechanism is missing, and most mutations will ablate function, this is especially important. Key requirements are an assay sensitive enough to measure small increases in improvement and a starting construct that with enough signal difference between states to optimize from. In the case with LOV-SsrAC, both of these were

satisfied, and mutations that conferred beneficial properties could be enriched. As many labs are using LOV domains to power optogenetic tools with a variety of functions, this method could be a useful way to improve that desired functionality, whatever it may be.

6.3 Leveling the Optogenetic Playing Field; Is LOV All You Need?

Our look into alternative blue-light mediated heterodimer pairs was enlightening. The CRY2 and TULIP switches, which have been used in a number of cell biology applications, work in a much weaker affinity range than expected and have lower dynamic ranges than expected. The side-by-side comparison from *in vitro* to yeast to mammalian cell gives us an idea of what biophysical characteristics make for the best tool. In applications where you expect a phenotypic change quickly, it appears that tighter affinity switches will have a larger effect. However, in a multi-day experiment such as yeast colony growth, a weak lit state affinity is much better. In both of these cases, the limiting factor is background activity in the off state. Over the course of a few seconds or minutes, micromolar interactions will contribute little to overall effect, however on a longer timescale, even weak micromolar affinities are enough to elicit responses on par with the active state. Finally, this work gives valuable insight to the mechanism of CRY2 heterodimerization and oligomerization. With so many labs using the CRY2 system, both with and without CIB1N, careful understanding of how the tool is inducing activity is paramount. For this reason, in the head-to-head tests the iLID switches demonstrate the utility of a simple, well-characterized mechanism in multiple formats and organisms. Together, this work clearly demonstrates that while LOV isn't all you need, everyone could use a little more of it.

REFERENCES

1. Crick, F. H. Thinking about the brain. *Sci. Am.* **241,** 219–32 (1979).

2. Crick, F. The impact of molecular biology on neuroscience. *Philos. Trans. R. Soc. B Biol. Sci.* **354,** 2021–2025 (1999).

3. Zemelman, B. V., Lee, G. A., Ng, M. & Miesenböck, G. Selective Photostimulation of Genetically ChARGed Neurons. *Neuron* **33,** 15–22 (2002).

4. Banghart, M., Borges, K., Isacoff, E., Trauner, D. & Kramer, R. H. Light-activated ion channels for remote control of neuronal firing. *Nat. Neurosci.* **7,** 1381–6 (2004).

5. Knoll, H. *CRC Handbook of Organic Photochemistry and Photobiology, Volumes 1 & 2, Second Edition*. (CRC Press, 2003). at <https://books.google.com/books?id=ryZtbp6CX34C&pgis=1>

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

7. Nagel, G. *et al.* Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 13940–5 (2003).

8. Brieke, C., Rohrbach, F., Gottschalk, A., Mayer, G. & Heckel, A. Light-controlled tools. *Angew. Chem. Int. Ed. Engl.* **51,** 8446–76 (2012).

9. Chou, C., Young, D. D. & Deiters, A. Photocaged t7 RNA polymerase for the light activation of transcription and gene function in pro- and eukaryotic cells. *Chembiochem* **11,** 972–977 (2010).

10. Young, D. D. & Deiters, A. Photochemical control of biological processes. *Org. Biomol. Chem.* **5,** 999–1005 (2007).

11. Parker, L. L., Kurutz, J. W., Kent, S. B. H. & Kron, S. J. Control of the yeast cell cycle with a photocleavable alpha-factor analogue. *Angew. Chem. Int. Ed. Engl.* **45,** 6322–5 (2006).

12. Crosson, S., Rajagopal, S. & Moffat, K. The LOV domain family: photoresponsive signaling modules coupled to diverse output domains. *Biochemistry* **42,** 2–10 (2003).

13. Taylor, B. L. & Zhulin, I. B. PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol. Mol. Biol. Rev.* **63,** 479–506 (1999).

14. Kinoshita, T. *et al.* Phot1 and phot2 mediate blue light regulation of stomatal opening. *Nature* **414,** 656–60 (2001).

15. Swartz, T. E. *et al.* The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin. *J Biol Chem* **276,** 36493–36500 (2001).

16. 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).

17. Kasahara, M. *et al.* Photochemical properties of the flavin mononucleotide-binding domains of the phototropins from Arabidopsis, rice, and Chlamydomonas reinhardtii. *Plant Physiol* **129,** 762–773 (2002).

18. Circolone, F. *et al.* Structural basis for the slow dark recovery of a full-length LOV protein from Pseudomonas putida. *J. Mol. Biol.* **417,** 362–74 (2012).

19. Harper, S., Neil, L. C. & Gardner, K. Structural basis of a phototropin light switch. *Science* **301,** 1541–1544 (2003).

20. Harper, S. M., Christie, J. M. & Gardner, K. H. Disruption of the LOV-Jalpha helix interaction activates phototropin kinase activity. *Biochemistry* **43,** 16184–92 (2004).

21. Nash, A. I., Ko, W. H., Harper, S. & Gardner, K. A conserved glutamine plays a central role in LOV domain signal transmission and its duration. *Biochemistry* **47,** 13842–13849 (2008).

22. 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,** 14001–9 (2007).

23. Yao, X., Rosen, M. & Gardner, K. Estimation of the available free energy in a LOV2-J alpha photoswitch. *Nat Chem Biol* **4,** 491–497 (2008).

24. Zayner, J. P., Antoniou, C. & Sosnick, T. R. The Amino-Terminal Helix Modulates Light-Activated Conformational Changes in AsLOV2. *J Mol Biol* **419,** 61–74 (2012).

25. Zayner, J. P. & Sosnick, T. R. Factors That Control the Chemistry of the LOV Domain Photocycle. *PLoS One* **9,** e87074 (2014).

26. Ahmad, M. & Cashmore, A. R. HY4 gene of A. thaliana encodes a protein with characteristics of a blue-light photoreceptor. *Nature* **366,** 162–166 (1993).

27. Cashmore, A. R., Jarillo, J. A., Wu, Y. J. & Liu, D. Cryptochromes: blue light receptors for plants and animals. *Science* **284,** 760–5 (1999).

28. Kleiner, O., Kircher, S., Harter, K. & Batschauer, A. Nuclear localization of the Arabidopsis blue light receptor cryptochrome 2. *Plant J.* **19,** 289–96 (1999).

29. Kleine, T., Lockhart, P. & Batschauer, A. An Arabidopsis protein closely related to Synechocystis cryptochrome is targeted to organelles. *Plant J.* **35,** 93–103 (2003).

30. Hoang, N., Bouly, J.-P. & Ahmad, M. Evidence of a light-sensing role for folate in Arabidopsis cryptochrome blue-light receptors. *Mol. Plant* **1,** 68–74 (2008).

31. Yang, H. Q. *et al.* The C termini of Arabidopsis cryptochromes mediate a constitutive light response. *Cell* **103,** 815–27 (2000).

32. Brautigam, C. A. *et al.* Structure of the photolyase-like domain of cryptochrome 1 from Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 12142–7 (2004).

33. Klar, T., Pokorny, R., Moldt, J., Batschauer, A. & Essen, L.-O. Cryptochrome 3 from Arabidopsis thaliana: Structural and Functional Analysis of its Complex with a Folate Light Antenna. *J. Mol. Biol.* **366,** 954–964 (2007).

34. Christie, J. M., Blackwood, L., Petersen, J. & Sullivan, S. Plant Flavoprotein Photoreceptors. *Plant Cell Physiol.* **56,** 401–413 (2015).

35. Liu, B., Liu, H., Zhong, D. & Lin, C. Searching for a photocycle of the cryptochrome photoreceptors. *Curr. Opin. Plant Biol.* **13,** 578–586 (2010).

36. Li, X. *et al.* Arabidopsis cryptochrome 2 (CRY2) functions by the photoactivation mechanism distinct from the tryptophan (trp) triad-dependent photoreduction. *Proc Natl Acad Sci USA* **108,** 20844–20849 (2011).

37. Engelhard, C. *et al.* Cellular Metabolites Enhance the Light Sensitivity of Arabidopsis Cryptochrome through Alternate Electron Transfer Pathways. *Plant Cell Online* **26,** 4519–4531 (2014).

38. Giovani, B., Byrdin, M., Ahmad, M. & Brettel, K. Light-induced electron transfer in a cryptochrome blue-light photoreceptor. *Nat. Struct. Biol.* **10,** 489–490 (2003).

39. Bouly, J.-P. *et al.* Cryptochrome Blue Light Photoreceptors Are Activated through Interconversion of Flavin Redox States. *J. Biol. Chem.* **282,** 9383–9391 (2007).

40. Banerjee, R. *et al.* The signaling state of Arabidopsis cryptochrome 2 contains flavin semiquinone. *J. Biol. Chem.* **282,** 14916–22 (2007).

41. Partch, C. L., Clarkson, M. W., Özgür, S., Lee, A. L. & Sancar, A. Role of Structural Plasticity in Signal Transduction by the Cryptochrome Blue-Light Photoreceptor †. *Biochemistry* **44,** 3795–3805 (2005).

42. Liu, H. *et al.* Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in Arabidopsis. *Science* **322,** 1535–9 (2008).

43. Shalitin, D. *et al.* Regulation of Arabidopsis cryptochrome 2 by blue-light-dependent phosphorylation. *Nature* **417,** 763–7 (2002).

44. Shalitin, D., Yu, X., Maymon, M., Mockler, T. & Lin, C. Blue light-dependent in vivo and in vitro phosphorylation of Arabidopsis cryptochrome 1. *Plant Cell* **15,** 2421–9 (2003).

45. Bugaj, L. J., Choksi, A. T., Mesuda, C. K., Kane, R. S. & Schaffer, D. V. Optogenetic protein clustering and signaling activation in mammalian cells. *Nat. Methods* **10,** 249–52 (2013).

46. Kennedy, M. J. *et al.* Rapid blue-light-mediated induction of protein interactions in living cells. *Nat. Methods* **7,** 973–5 (2010).

47. Taslimi, A. *et al.* An optimized optogenetic clustering tool for probing protein interaction and function. *Nat. Commun.* **5,** 4925 (2014).

48. Davis, S. J., Vener, A. V & Vierstra, R. D. Bacteriophytochromes: phytochrome-like photoreceptors from nonphotosynthetic eubacteria. *Science* **286,** 2517–20 (1999).

49. Rockwell, N. C., Su, Y.-S. & Lagarias, J. C. Phytochrome structure and signaling mechanisms. *Annu. Rev. Plant Biol.* **57,** 837–58 (2006).

50. Ikeuchi, M. & Ishizuka, T. Cyanobacteriochromes: a new superfamily of tetrapyrrole-binding photoreceptors in cyanobacteria. *Photochem. Photobiol. Sci.* **7,** 1159 (2008).

51. Rockwell, N. C., Martin, S. S. & Lagarias, J. C. Mechanistic Insight into the Photosensory Versatility of DXCF Cyanobacteriochromes. *Biochemistry* **51,** 3576–3585 (2012).

52. Rockwell, N. C. & Lagarias, J. C. A brief history of phytochromes. *Chemphyschem* **11,** 1172–80 (2010).

53. 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).

54. Gambetta, G. A. & Lagarias, J. C. Genetic engineering of phytochrome biosynthesis in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 10566–71 (2001).

55. Tabor, J. J., Levskaya, A. & Voigt, C. A. Multichromatic control of gene expression in Escherichia coli. *J Mol Biol* **405,** 315–324 (2011).

56. Müller, K. *et al.* Synthesis of phycocyanobilin in mammalian cells. *Chem. Commun. (Camb).* **49,** 8970–2 (2013).

57. Ni, M., Tepperman, J. M. & Quail, P. H. Binding of phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light. *Nature* **400,** 781–4 (1999).

58. Khanna, R. A Novel Molecular Recognition Motif Necessary for Targeting Photoactivated Phytochrome Signaling to Specific Basic Helix-Loop-Helix Transcription Factors. *PLANT CELL ONLINE* **16,** 3033–3044 (2004).

59. Levskaya, A. *et al.* Synthetic biology: engineering Escherichia coli to see light. *Nature* **438,** 441–2 (2005).

60. Ryu, M.-H. *et al.* Engineering adenylate cyclases regulated by near-infrared window light. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 10167–72 (2014).

61. Kraft, B. J. *et al.* Spectroscopic and mutational analysis of the blue-light photoreceptor AppA: a novel photocycle involving flavin stacking with an aromatic amino acid. *Biochemistry* **42,** 6726–34 (2003).

62. Laan, W., van der Horst, M. A., van Stokkum, I. H. & Hellingwerf, K. J. Initial characterization of the primary photochemistry of AppA, a blue-light-using flavin adenine dinucleotide-domain containing transcriptional antirepressor protein from Rhodobacter sphaeroides: a key role for reversible intramolecular proton transfer f. *Photochem. Photobiol.* **78,** 290–7 (2003).

63. 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).

64. Kort, R. *et al.* The xanthopsins: a new family of eubacterial blue-light photoreceptors. *EMBO J.* **15,** 3209–18 (1996).

65. Van der Horst, M. A. *et al.* From primary photochemistry to biological function in the blue-light photoreceptors PYP and AppA. *Photochem. Photobiol. Sci.* **4,** 688–93 (2005).

66. Morgan, S.-A., Al-Abdul-Wahid, S. & Woolley, G. A. Structure-based design of a photocontrolled DNA binding protein. *J. Mol. Biol.* **399,** 94–112 (2010).

67. Fan, H. Y. *et al.* Improving a designed photo-controlled DNA-binding protein. *Biochemistry* (2011). at <http://pubs.acs.org/doi/abs/10.1021/bi101432p>

68. Wu, Y. I. *et al.* A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature* **461,** 104–8 (2009).

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

70. Lee, J. *et al.* Surface sites for engineering allosteric control in proteins. *Science* **322,** 438–42 (2008).

71. Möglich, A., Ayers, R. a & Moffat, K. Design and signaling mechanism of light-regulated histidine kinases. *J. Mol. Biol.* **385,** 1433–44 (2009).

72. 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,** 619–26 (2013).

73. Usherenko, S. *et al.* Photo-sensitive degron variants for tuning protein stability by light. *BMC Syst. Biol.* **8,** 128 (2014).

74. Polstein, L. R. & Gersbach, C. A. Light-inducible spatiotemporal control of gene activation by customizable zinc finger transcription factors. *J. Am. Chem. Soc.* **134,** 16480–3 (2012).

75. Konermann, S. *et al.* Optical control of mammalian endogenous transcription and epigenetic states. *Nature* **500,** 472–6 (2013).

76. Yazawa, M., Sadaghiani, A. M., Hsueh, B. & Dolmetsch, R. E. Induction of protein-protein interactions in live cells using light. *Nat. Biotechnol.* **27,** 941–5 (2009).

77. 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).

78. Schmidt, A. & Hall, A. Guanine nucleotide exchange factors for Rho GTPases: turning on the switch. *Genes Dev.* **16,** 1587–609 (2002).

79. Strickland, D. *et al.* TULIPs: tunable, light-controlled interacting protein tags for cell biology. *Nat. Methods* **9,** 379–84 (2012).

80. Lee, S. *et al.* Reversible protein inactivation by optogenetic trapping in cells. *Nat. Methods* **11,** 633–636 (2014).

81. Zhou, X. X., Chung, H. K., Lam, A. J. & Lin, M. Z. Optical control of protein activity by fluorescent protein domains. *Science* **338,** 810–4 (2012).

82. Lungu, O. *et al.* Designing Photoswitchable Peptides Using the AsLOV2 Domain. *Chem Biol* **19,** 507–517 (2012).

83. Liu, H., Gomez, G., Lin, S., Lin, S. & Lin, C. Optogenetic control of transcription in zebrafish. *PLoS One* **7,** e50738 (2012).

84. Pathak, G. P., Strickland, D., Vrana, J. D. & Tucker, C. L. Benchmarking of optical dimerizer systems. *ACS Synth. Biol.* **3,** 832–8 (2014).

85. Shimizu-Sato, S., Huq, E., Tepperman, J. M. & Quail, P. H. A light-switchable gene promoter system. *Nat. Biotechnol.* **20,** 1041–4 (2002).

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

87. Moscou, M. J. & Bogdanove, A. J. A simple cipher governs DNA recognition by TAL effectors. *Science* **326,** 1501 (2009).

88. Polstein, L. R. & Gersbach, C. A. A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.* **11,** 198–200 (2015).

89. Hughes, R. M. & Lawrence, D. S. Optogenetic engineering: light-directed cell motility. *Angew. Chem. Int. Ed. Engl.* **53,** 10904–7 (2014).

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

91. Huang, J., Makabe, K., Biancalana, M., Koide, A. & Koide, S. Structural basis for exquisite specificity of affinity clamps, synthetic binding proteins generated through directed domain-interface evolution. *J. Mol. Biol.* **392,** 1221–31 (2009).

92. Van Bergeijk, P., Adrian, M., Hoogenraad, C. C. & Kapitein, L. C. Optogenetic control of organelle transport and positioning. *Nature* **518,** 111–4 (2015).

93. Hirokawa, N. Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science* **279,** 519–26 (1998).

94. Fodor, L., Elman, M. & Ullmann, Y. *Aesthetic Applications of Intense Pulsed Light*. (Springer Science & Business Media, 2010). at <https://books.google.com/books?id=AgelDbucQIkC&pgis=1>

95. Singh, N. *et al.* Detection of native peptides as potent inhibitors of enzymes. Crystal structure of the complex formed between treated bovine alpha-chymotrypsin and an autocatalytically produced fragment, Ile-Val-Asn-Gly-Glu-Glu-Ala-Val-Pro-Gly-Ser-Trp-Pro-Trp, at 2.2 ang. *FEBS J.* **272,** 562–72 (2005).

96. Lockless, S. W. & Ranganathan, R. Evolutionarily conserved pathways of energetic connectivity in protein families. *Science* **286,** 295–9 (1999).

97. Conti, E., Uy, M., Leighton, L., Blobel, G. & Kuriyan, J. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell* **94,** 193–204 (1998).

98. Gautier, A. *et al.* Genetically encoded photocontrol of protein localization in mammalian cells. *J. Am. Chem. Soc.* **132,** 4086–8 (2010).

99. Nguyen, A., Rothman, D. M., Stehn, J., Imperiali, B. & Yaffe, M. B. Caged phosphopeptides reveal a temporal role for 14-3-3 in G1 arrest and S-phase checkpoint function. *Nat. Biotechnol.* **22,** 993–1000 (2004).

100. Möglich, A. & Moffat, K. Engineered photoreceptors as novel optogenetic tools. *Photochem. Photobiol. Sci.* **9,** 1286–300 (2010).

101. 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).

102. 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).

103. Huala, E. *et al.* Arabidopsis NPH1: a protein kinase with a putative redox-sensing domain. *Science* **278,** 2120–3 (1997).

104. Halavaty, A. & 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).

105. Crosson, S. & Moffat, K. Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch. *Plant Cell* **14,** 1067–75 (2002).

106. Harper, S., Neil, L. C., Day, I. J., Hore, P. J. & Gardner, K. Conformational changes in a photosensory LOV domain monitored by time-resolved NMR spectroscopy. *Journal of the American Chemical Society* **126,** 3390–3391 (2004).

107. Swartz, T. E., Wenzel, P. J., Corchnoy, S. B., Briggs, W. R. & Bogomolni, R. A. Vibration spectroscopy reveals light-induced chromophore and protein structural changes in the LOV2 domain of the plant blue-light receptor phototropin 1. *Biochemistry* **41,** 7183–9 (2002).

108. Zoltowski, B. D., Vaccaro, B. & Crane, B. R. Mechanism-based tuning of a LOV domain photoreceptor. *Nat. Chem. Biol.* **5,** 827–34 (2009).

109. Tran Van Nhieu, G., Ben-Ze'ev, A. & Sansonetti, P. J. Modulation of bacterial entry into epithelial cells by association between vinculin and the Shigella IpaA invasin. *EMBO J.* **16,** 2717–29 (1997).

110. Levchenko, I., Grant, R. a, Wah, D. a, Sauer, R. T. & Baker, T. a. Structure of a delivery protein for an AAA+ protease in complex with a peptide degradation tag. *Mol. Cell* **12,** 365–72 (2003).

111. Thompson, G., Owen, D., Chalk, P. A. & Lowe, P. N. Delineation of the Cdc42/Rac-binding domain of p21-activated kinase. *Biochemistry* **37,** 7885–91 (1998).

112. Le Clainche, C., Dwivedi, S. P., Didry, D. & Carlier, M.-F. Vinculin is a dually regulated actin filament barbed end-capping and side-binding protein. *J. Biol. Chem.* **285,** 23420–32 (2010).

113. Izard, T., Tran Van Nhieu, G. & Bois, P. R. J. Shigella applies molecular mimicry to subvert vinculin and invade host cells. *J. Cell Biol.* **175,** 465–75 (2006).

114. Cohen, D. M., Kutscher, B., Chen, H., Murphy, D. B. & Craig, S. W. A conformational switch in vinculin drives formation and dynamics of a talin-vinculin complex at focal adhesions. *J. Biol. Chem.* **281,** 16006–15 (2006).

115. Saunders, R. M. *et al.* Role of vinculin in regulating focal adhesion turnover. *Eur. J. Cell Biol.* **85,** 487–500 (2006).

116. Carisey, A. & Ballestrem, C. Vinculin, an adapter protein in control of cell adhesion signalling. *Eur. J. Cell Biol.* **90,** 157–63

117. Das, R. & Baker, D. Macromolecular modeling with rosetta. *Annu. Rev. Biochem.* **77,** 363–82 (2008).

118. Alsallaq, R. & Zhou, H.-X. Electrostatic rate enhancement and transient complex of protein-protein association. *Proteins* **71,** 320–35 (2008).

119. Nhieu, G. T. Van & Izard, T. Vinculin binding in its closed conformation by a helix addition mechanism. *EMBO J.* **26,** 4588–96 (2007).

120. Cooper, S. *et al.* Predicting protein structures with a multiplayer online game. *Nature* **466,** 756–60 (2010).

121. Salomon, M., Christie, J. M., Knieb, E., Lempert, U. & Briggs, W. R. Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor, phototropin. *Biochemistry* **39,** 9401–10 (2000).

122. Nikolovska-Coleska, Z. *et al.* Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarization. *Anal. Biochem.* **332,** 261–73 (2004).

123. Kleiger, G., Saha, A., Lewis, S., Kuhlman, B. & Deshaies, R. J. Rapid E2-E3 assembly and disassembly enable processive ubiquitylation of cullin-RING ubiquitin ligase substrates. *Cell* **139,** 957–968 (2009).

124. Izard, T. & Vonrhein, C. Structural basis for amplifying vinculin activation by talin. *J. Biol. Chem.* **279,** 27667–78 (2004).

125. Rivera, V. M. *et al.* A humanized system for pharmacologic control of gene expression. *Nat. Med.* **2,** 1028–1032 (1996).

126. Kennedy, M. & Hughes, R. Rapid blue-light-mediated induction of protein interactions in living cells. *Nat. Methods* (2010). at <http://www.nature.com/nmeth/journal/vaop/ncurrent/full/nmeth.1524.html>

127. Konermann, S. *et al.* Optical control of mammalian endogenous transcription and epigenetic states. *Nature* **500,** 472–6 (2013).

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

129. Salomon, M., Christie, J. M., Knieb, E., Lempert, U. & Briggs, W. R. Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor, phototropin. *Biochemistry* **39,** 9401–9410 (2000).

130. Wah, D. A., Levchenko, I., Baker, T. & Sauer, R. Characterization of a specificity factor for an AAA+ ATPase: assembly of SspB dimers with ssrA-tagged proteins and the ClpX hexamer. *Chem Biol* **9,** 1237–1245 (2002).

131. Flynn, J. M. *et al.* Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. *Proc Natl Acad Sci USA* **98,** 10584–10589 (2001).

132. Lungu, O. I. *et al.* Designing photoswitchable peptides using the AsLOV2 domain. *Chem. Biol.* **19,** 507–17 (2012).

133. Firnberg, E. & Ostermeier, M. PFunkel: Efficient, Expansive, User-Defined Mutagenesis. *PLoS One* **7,** (2012).

134. Tonikian, R., Zhang, Y., Boone, C. & Sidhu, S. S. Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries. *Nat. Protoc.* **2,** 1368–86 (2007).

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

136. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography* **67,** 235–242 (2011).

137. Evans, P. Scaling and assessment of data quality. in *Acta Crystallographica Section D: Biological Crystallography* **62,** 72–82 (2006).

138. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).

139. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 213–221 (2010).

140. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 486–501 (2010).

141. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9,** 676–682 (2012).

142. Leaver-Fay, A. *et al.* in *Methods in Enzymology, Volume 487* **487,** 545–574 (2011).

143. Speck, J., Arndt, K. M. & Müller, K. M. Efficient phage display of intracellularly folded proteins mediated by the TAT pathway. *Protein Eng. Des. Sel.* **24,** 473–84 (2011).

144. Shaner, N. C. *et al.* Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nat. Methods* **5,** 545–551 (2008).

145. Cai, L., Makhov, A. M., Schafer, D. A. & Bear, J. E. Coronin 1B antagonizes cortactin and remodels Arp2/3-containing actin branches in lamellipodia. *Cell* **134,** 828–842 (2008).

146. Inoue, T., Heo, W. Do, Grimley, J. S., Wandless, T. J. & Meyer, T. An inducible translocation strategy to rapidly activate and inhibit small GTPase signaling pathways. *Nat. Methods* **2,** 415–8 (2005).

147. Nobes, C. D. & Hall, A. Rho, rac, and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. *Cell* **81,** 53–62 (1995).

148. Jaiswal, M., Dvorsky, R. & Ahmadian, M. R. Deciphering the molecular and functional basis of Dbl family proteins: a novel systematic approach toward classification of selective activation of the Rho family proteins. *J. Biol. Chem.* **288,** 4486–500 (2013).

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

150. Rao, M. V., Chu, P.-H., Hahn, K. M. & Zaidel-Bar, R. An optogenetic tool for the activation of endogenous diaphanous-related formins induces thickening of stress fibers without an increase in contractility. *Cytoskeleton (Hoboken).* **70,** 394–407 (2013).

151. Xu, Y. *et al.* Optogenetic control of chemokine receptor signal and T-cell migration. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 6371–6 (2014).

152. Guntas, G. *et al.* Engineering an improved light-induced dimer (iLID) for controlling the localization and activity of signaling proteins. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 112–117 (2014).

153. Más, P., Devlin, P. F., Panda, S. & Kay, S. A. Functional interaction of phytochrome B and cryptochrome 2. *Nature* **408,** 207–11 (2000).

154. Hughes, R. M., Bolger, S., Tapadia, H. & Tucker, C. L. Light-mediated control of DNA transcription in yeast. *Methods* **58,** 385–91 (2012).

155. Ozgür, S. & Sancar, A. Analysis of autophosphorylating kinase activities of Arabidopsis and human cryptochromes. *Biochemistry* **45,** 13369–74 (2006).

156. Marshall, K. S., Zhang, Z., Curran, J., Derbyshire, S. & Mymryk, J. S. An improved genetic system for detection and analysis of protein nuclear import signals. *BMC Mol. Biol.* **8,** 6 (2007).

157. Gietz, R. D. & Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2,** 31–4 (2007).

158. Bear, J. E. *et al.* Negative regulation of fibroblast motility by Ena/VASP proteins. *Cell* **101,** 717–28 (2000).

159. Rossman, K. L., Der, C. J. & Sondek, J. GEF means go: turning on RHO GTPases with guanine nucleotide-exchange factors. *Nat. Rev. Mol. Cell Biol.* **6,** 167–80 (2005).

160. Elbashir, S. M. *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411,** 494–8 (2001).

161. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337,** 816–21 (2012).

162. Banaszynski, L. A. & Wandless, T. J. Conditional control of protein function. *Chem Biol* **13,** 11–21 (2006).

163. Park, E. C., Finley, D. & Szostak, J. W. A strategy for the generation of conditional mutations by protein destabilization. *Proc. Natl. Acad. Sci. U. S. A.* **89,** 1249–52 (1992).

164. Dohmen, R. J., Wu, P. & Varshavsky, A. Heat-inducible degron: a method for constructing temperature-sensitive mutants. *Science* **263,** 1273–6 (1994).

165. Sakamoto, K. M. *et al.* Protacs: chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 8554–9 (2001).

166. Janse, D. M., Crosas, B., Finley, D. & Church, G. M. Localization to the proteasome is sufficient for degradation. *J. Biol. Chem.* **279,** 21415–20 (2004).

167. Zhou, P., Bogacki, R., McReynolds, L. & Howley, P. M. Harnessing the ubiquitination machinery to target the degradation of specific cellular proteins. *Mol Cell* **6,** 751–756 (2000).

168. Sakamoto, K. M. *et al.* Development of Protacs to target cancer-promoting proteins for ubiquitination and degradation. *Mol. Cell. Proteomics* **2,** 1350–8 (2003).

169. Schneekloth, J. S. *et al.* Chemical Genetic Control of Protein Levels: Selective in Vivo Targeted Degradation. *J. Am. Chem. Soc.* **126,** 3748–3754 (2004).