UNIVERSITY OF CALIFORNIA

Santa Barbara

Engineering flavin-based fluorescent proteins for oxygen-independent biosensing applications

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Chemical Engineering

by

Nolan Thomas Anderson

Committee in charge:

Professor Arnab Mukherjee, Chair

Professor Michelle O'Malley

Professor M. Scott Shell

Professor Maxwell Wilson

December 2023

The dissertation of Nolan Thomas Anderson is approved.

Michelle O'Malley

M. Scott Shell

Maxwell Wilson

Arnab Mukherjee, Committee Chair

December 2023

Engineering flavin-based fluorescent proteins for oxygen-independent biosensing applications

ACKNOWLEDGEMENTS

I have received support from numerous individuals, without which completion of this dissertation research would not have been possible. I would like to thank my family for their unwavering emotional the support and inspiration. I thank my friends, especially for maintaining connections through isolation created by an unprecedented-in-our-lifetimes pandemic. I also thank the Santa Barbara community, especially the particularly welcoming ice hockey and trombone communities, who have provided much-needed camaraderie throughout my PhD work. Members of the Mukherjee group, both past and present, have been indispensable allies throughout my work; providing knowledge, assistance, feedback, ideas, and friendship.

Nolan T. Anderson

PhD Candidate, University of California, Santa Barbara
7323 Bassano Drive, Goleta, California, USA
(502) 681 – 7087 – nolananderson@engineering.ucsb.edu

Education

2017—present	Ph.D., Chemical Engineering, University of California, Santa Barbara; Expected defense date December 4, 2023
2013—2017	B.S., Chemical Engineering, *summa cum laude,* University of Kentucky
2013—2017	B.S., Mathematics, *summa cum laude,* Departmental Honors, University of Kentucky

Professional & Research Experience

2018—present	Graduate Student Researcher, University of California, Santa Barbara (with Arnab Mukherjee), Santa Barbara, California, USA
2013—2017	Undergraduate Student Researcher, University of Kentucky (with Christina M. Payne), Lexington, Kentucky, USA
2016	Visiting Student Researcher, Swedish University of Agricultural Science, Uppsala, Sweden
2015	Startup and Process Engineering Intern, Bechtel Oil, Gas, & Chemicals, Cameron, Louisiana, USA

Honors & Awards

2017—2022	Chancellor's Fellow, University of California, Santa Barbara
2017	NSF Graduate Research Fellowship Program Honorable Mention

Refereed Publications

1. T. Haataja, J. E. Gado, A. Nutt, **N. T. Anderson**, M. Nilsson, M. H. Momeni, R. Isaksson, P. Väljamäe, G. Johansson, C. M. Payne, J. Ståhlberg, "Enzyme kinetics by GH7 cellobiohydrolases on chromogenic substrates is dictated by non-productive binding: insights from crystal structures and MD simulation," *FEBS Journal*, **290(2),** 379-399 (2022).
2. **N. T. Anderson***, K. B. Weyant*, A. Mukherjee, "Characterization of flavin binding in oxygen-independent fluorescent reporters," *AIChE Journal*, **66(12)**, e17083, (2020).
3. H. F. Ozbakir, **N. T. Anderson***, K. C. Fan*, A. Mukherjee, "Beyond the green fluorescent protein: biomolecular reporters for anaerobic and deep-tissue imaging," *Bioconjugate Chemistry*, **31(2)**, 293-302 (2020).

4. C. Wilkens, K. Auger, **N. T. Anderson**, D. A. Meekins, M. Raththagala, M. A. Hachem, C. M. Payne, M. S. Gentry, and B. Svensson, "Plant α-glucan phosphatases SEX4 and LSF2 display different affinity for amylopectin and amylose," *FEBS Letters*, **590(1)**, 118-128 (2016).
   * = *Equally contributing*

Preprint Publications

1. E. S. Lau, J. A. Goodheart, **N. T. Anderson**, V. L. Liu, A. Mukherjee, T. H. Oakley, "Predictable genetic recruitment for luciferin sulfation in the convergent evolution of bioluminescence," *BioRxiv*, **2023.04.12.536614** (2023). †
   † = *Currently under peer review*

Manuscripts in Preparation

1. **N. T. Anderson**, J. Xie*, A. N. Chacko*, V. L. Liu, K.C. Fan, R. Taylor, Z. Imansjah, A. Mukherjee. "Rational design of a circularly permuted flavin-based fluorescent protein." Expected submission November 2023.

Oral Presentations

1. **N. T. Anderson**, J. Xie, A. Mukherjee, "Circular permutation of LOV proteins toward protease activity sensing," *International Conference on Biomolecular Engineering*, Santa Barbara, California, USA; January 2023.
2. **N. T. Anderson**, A. Mukherjee, "Circular permutation of LOV proteins toward protease activity sensing," *American Institute of Chemical Engineers Annual Meeting*, Phoenix, Arizona, USA; November 2022.
3. **N. T. Anderson**, A. Mukherjee, "Engineering genetically-encodable oxygen-independent fluorescent reporters based on LOV proteins," *American Chemical Society Spring Meeting*, San Diego, California, USA; March 2022.
4. **N. T. Anderson**, A. Fenerty, A. Mukherjee, "Oxygen-independent ATP sensing with circularly permuted LOV reporters," **Biomedical Engineering Society Annual Meeting**, Orlando, Florida, USA (virtual); October 2021.
5. **N. T. Anderson**, "Oxygen-independent ATP sensing with circularly permuted LOV reporters," **Clorox-Amgen Graduate Student Symposium**, University of California, Santa Barbara, Santa Barbara, California, USA; September 2021.
6. **N. T. Anderson**, "Engineering Oxygen-Independent Fluorescent Reporters," **Chemical Engineering First-Year Symposium,** University of California, Santa Barbara, Santa Barbara, California, USA; September 2018.
   Underline indicates presenting author

Poster Presentations

1. **N. T. Anderson**, J. Xie, A. N. Chacko, R. Taylor, Z. Imansjah, A. Mukherjee, "Adaptable protease activity sensing using circularly permuted LOV proteins," **Mammalian Synthetic Biology Workshop**, San Jose, California, USA; June 2023.

2. **N. T. Anderson**, Z. Imansjah, R. Taylor, A. Mukherjee, "Directed evolution of a bright oxygen-independent yellow fluorescent protein," **Chlorox-Amgen Graduate Student Symposium,** University of California, Santa Barbara, California, USA; October 2019.
3. **N. T. Anderson**, A. Mukherjee, "Flavin Biosensing: Toward a Biocompatible Oxygen-Independent Fluorescent Reporter," **Chlorox-Amgen Graduate Student Symposium,** University of California, Santa Barbara, Santa Barbara, California, USA; October 2018.
4. **N. T. Anderson**, A. Mukherjee, "Oxygen-Independent Fluorescent Protein Reporters," **Advances in Protein Science,** Thousand Oaks, California, USA; August 2018.
5. **N. T. Anderson**, A. Nutt, M. H. Momeni, P. Väljamäe, G. Johansson, J. Ståhlberg, C. M. Payne, "o-Nitrophenyl Cellobioside as an Active Site Probe for Family 7 Cellobiohydrolases," **AIChE Annual Meeting**, San Francisco, California, USA; November 2016.
6. **N. T. Anderson**, A. Nutt, M. H. Momeni, P. Väljamäe, G. Johansson, J. Ståhlberg, C. M. Payne, "o-Nitrophenyl Cellobioside as an Active Site Probe for Family 7 Cellobiohydrolases," **Swedish Structural Biology Network Annual Conference**, Tällberg, Sweden; June 2016.

Underline indicates presenting author

Teaching Experience

Spring 2021	Teaching Assistant (Instructor: Songi Han), ChE 110B Chemical Engineering Thermodynamics II (UCSB) – 3 units
Winter 2020	Teaching Assistant (Instructor: Siddharth Dey), ChE 154 Engineering Approaches to Systems Biology (UCSB) – 3 units
Fall 2018	Teaching Assistant (Instructor: Matthew Helgeson), ChE 10 Introduction to Chemical Engineering (UCSB) – 3 units
Spring 2017	Undergraduate Assistant (Instructor: Amy Green), MA 111 Contemporary Mathematics (UK) – 3 credit hours
Spring 2017	Section Leader (Instructor: Lisa Blue), CHE 197 Chemistry Excel (UK) – 1 credit hour
Fall 2016	Undergraduate Assistant (Instructor: Nicholas Nguyen), MA 111 Contemporary Mathematics (UK) – 3 credit hours
Spring 2016	Section Leader (Instructor: Lisa Blue), CHE 197 Chemistry Excel (UK) – 1 credit hour
Fall 2015	Physics Tutor, University of Kentucky Department of Academic Excellence
Spring 2015	Section Leader (Instructor: Lisa Blue), CHE 197 Chemistry Excel (UK) – 1 credit hour
Spring 2015	Learning Center Assistant, University of Kentucky Department of Chemistry

Students Mentored

July 2018—June 2020	Rachel Taylor – BS Chemical Engineering 2020 (UCSB); currently PhD candidate at Pennsylvania State University, Chemical Engineering (with Bruce Logan)
January 2019—June 2020	Zoe Imansjah – BS Chemical Engineering 2022 (UCSB); Edison Scholars Program, Summer 2019; currently Production Engineer at Dow Chemical
March—July 2021	Anna Fenerty – BS Biology 2023 (UCSB)
March 2022—June 2023	Jason Xie – BS Biology 2023 (UCSB); currently laboratory technician at University of California, Irvine; Department of Physiology & Biophysics (with Barbara Jusiak)
January 2023—September 2023	Vannie Liu – BS Biology 2023 (UCSB)

Service

Sept 2021	Organizing Co-Chair, Chlorox-Amgen Graduate Student Symposium
March 2021	Category Judge – Middle and High School Bioinformatics and Computational Biology, Kentucky Science and Engineering Fair
March 2021	Category Judge – High School Engineering Mechanics, duPont Manual Regional Science and Engineering Fair
Sept 2020	Session Chair, Chlorox-Amgen Graduate Student Symposium, "Bioengineering"
March 2020	Category Judge – Middle School Biochemistry, Kentucky Science and Engineering Fair
October 2019	Session Chair, Chlorox-Amgen Graduate Student Symposium, "Functional Bio-Inspired Materials"
Sept 2019 – Sept 2020	President, UCSB Chemical Engineering Graduate Student Association
February 2019	Prospective Student Recruitment Coordinator, UCSB Department of Chemical Engineering
Sept 2018 – Sept 2019	Secretary, Executive Board member; UCSB Chemical Engineering Graduate Student Association
Sept 2018 – February 2020	Science as a Career Outreach Project Experiment (SCOPE): educating high school students about state-of-the-art research and career paths in science and engineering
Oct 2017—Sept 2018	First-Year Representative, Executive Board member; UCSB Chemical Engineering Graduate Student Association

ABSTRACT

Engineering flavin-based fluorescent proteins for oxygen-independent biosensing applications

by

Nolan Thomas Anderson

Flavin-based fluorescent proteins (FbFPs) are small, oxygen independent fluorescent proteins derived from naturally occurring phototropins, which fluoresce in the green range. FbFPs have the potential to fill a critical gap in bioimaging of anaerobic and hypoxic biological systems. This gap exists because the green fluorescent protein (GFP) and its derivatives require an oxygen-dependent post-translational modification to cyclize three residues to form the active chromophore. As FbFPs bind flavin, a universally available vitamin in all organisms, as their chromophore, they do not require oxygen to become fluorescent. However, FbFPs do suffer from several weaknesses, including significantly lower brightness than GFP, lack of a varied color palette, and a relative lack of biosensors based on FbFPs. Within, we present engineering efforts toward making FbFPs highly useful bioimaging tools, particularly for use in anaerobic contexts.

One obstacle to engineering FbFPs has been a lack of understanding of flavin-binding dynamics. As flavins are universally available in cells, and flavin-auxotrophic cell lines

require flavin supplementation for survival, obtaining an FbFP not bound to flavin has proven difficult. Here we focus on the FbFPs iLOV and EcFbFP. Unsupported denaturation and buffer exchange of FbFPs has resulted in insoluble, misfolded protein which cannot be reconstituted with flavin. We have devised a protocol by which His-tagged FbFPs are bound to a nickel support, then denatured, buffer exchanged, and refolded, thus yielding properly folded flavin-free protein. We demonstrate that fluorescence can be recovered by adding flavin back to the protein. Through flavin titration, we have determined dissociation constants between both iLOV and EcFbFP, and three naturally occurring flavin cofactors: riboflavin, flavin mononucleotide (FMN), and flavin adenine dinucleotide (FAD). This has shown that FbFPs bind flavins with affinities in the hundreds of nanomolar range, with greatest affinity for FAD and FMN, binding less tightly to riboflavin. Further, we show that deflavinated FbFPs can be used as an effective in vitro sensor for flavin. Understanding of the critical flavin binding dynamics of FbFPs should enable further efforts in engineering the protein-cofactor interaction and help enable incorporation of non-native cofactors into FbFPs.

In order to enable more extensive biosensor construction from FbFPs, we sought to create the first circular permutation of an FbFP. Circular permutation has enabled construction of numerous biosensors from GFP by placing new termini in close proximity, so that fluorescence can be modulated by fused metabolite binding domains. Leveraging structural data and systematic domain insertion, we have determined that iLOV can be circularly permuted about residues G95 and E96, while fusing the original termini with a native phototropin helix and an engineered linker; yielding a dim but fluorescent protein. We then demonstrate that fluorescence can be recovered to approximately the level of linear

iLOV by fusing dimerizing coiled-coil leucine zipper domains to the new termini. Using this "zipped" circularly permuted iLOV, we created fluorescent sensors for the activity of two proteases; tobacco etch virus protease and SARS-CoV-2 main protease. These sensors show a 73 and 61% decrease in fluorescence, respectively, upon expression of the appropriate protease. We expect that our creation of a circularly permuted iLOV will enable building a wide variety of FbFP-based biosensors, much in the same way that circular permutation of GFP has enabled building of a wide variety of GFP-based biosensors.

Finally, we sought to minimize the size of the FbFP iLOV in order to optimize a protein for use when size constraints dominate, such as in a restrictive packaging vector or when fusing to a sensitive protein. Domain insertion and structural analysis suggested that the relatively unstructured termini of iLOV were not critical for fluorescence. Systematic trimming of these termini revealed that five amino acids could be removed from the N-terminus, while four amino acids could be removed from the C-terminus without extensive loss of fluorescence. We searched for internal removable residues and determined that Gly95 could similarly be removed without a great fluorescence penalty. By combining the terminal truncations and deletion of Gly95, we created a protein referred to as "nanoLOV," which retains 86.7 ± 2.2% of iLOV's brightness.

TABLE OF CONTENTS

Chapter 1. Introduction and Background Information

Parts of this chapter are adapted, with permission, from a published paper. I am co-second author on this paper. My contributions include reviewing background and current work on LOV proteins.

Paper: *Bioconjugate Chemistry*, **2020**, *31* (2), 293-302

Authors: Harun F. Ozbakir, **Nolan T. Anderson***, Kang-Ching Fan*, Arnab Mukherjee

* *Each of these authors contributed equally*

DOI: 10.1021/acs.bioconjchem.9b00688

1.1. State of the art in fluorescent biosensing

Genetically encoded fluorescent reporters such as the green fluorescent protein (GFP) provide one of the most powerful techniques for molecular imaging in living systems. GFP has enabled revolutionary progress in cellular biology by allowing labeling of structures, tracking of gene expression, sensing of phenomena, and more in living cells without the need for externally supplied stains. However, GFP and related proteins must undergo an oxidative cyclization reaction to develop their active chromophore, which renders these reporters non-fluorescent in anaerobic conditions.[1–3] As a result, GFP-based reporters are inadequate for studying biological function in anaerobic cell cultures. This is a serious limitation as anaerobic organisms have immense value in both basic research and emerging biotechnologies such as sustainable biomanufacturing, cellular therapies, and cell-based diagnostics.[4–16] Importantly, the study of complex cellular communities that thrive in low-oxygen milieu (e.g., the gut microbiota, which is comprised largely of anaerobic microbes) would benefit tremendously from the availability of nontraditional fluorescent reporters that can be used to interrogate biological function regardless of oxygen availability.[17] Here we summarize recent developments in the class of oxygen-independent fluorescent proteins that respectively enable biological function to be visualized in anaerobic cells. In addition, we identify key avenues where innovative engineering solutions are needed to empower these nontraditional reporters to parallel or even rival the capabilities of GFP.

1.2. Flavin-based fluorescent proteins

Reporters derived from GFP have limited utility for studying gut microbes, hypoxic tumors, and other systems that thrive in low oxygen conditions.[18–22] To this end, efforts to build oxygen-independent fluorescent proteins have led to the discovery of biomolecular reporters

derived from flavin-binding photoreceptors known as light, oxygen, and voltage (LOV) sensing proteins.[23,24] LOV proteins are photochemists by cellular profession. Using flavin as their photoactive cofactor, wild-type LOV proteins convert blue light excitation into conformational rearrangements to modulate effector functions such as kinase activity and DNA binding.[25] Wild-type LOV is dimly fluorescent because the energy from photoexcitation is expended to coordinate electronic and structural changes in the protein. However, by mutating a key photoactive cysteine residue into alanine, it is possible to impair this photochemistry, thereby turning the resultant Cys → Ala mutant fluorescent.[26,27] These non-photoactive mutants are known as flavin-based fluorescent proteins, or FbFPs. A hallmark of this fluorescence mechanism is that light emission is independent of oxygen (**Figure 1.1c**), spectral properties being largely determined by interactions between the FbFP and the noncovalently bound flavin cofactor. Building on this foundation, nearly 20 FbFPs have been derived from bacteria, plants, and environmental metagenomic libraries.[26–31] To potentiate their utility as reporter proteins, engineering techniques such as DNA shuffling and site saturation mutagenesis have been applied to increase photostability, thermal tolerance, and quantum yield.[32–34] Furthermore, using genome mining, new algal FbFPs have been introduced that are characterized by improved molecular brightness, photostability, broad pH range, and thermal stability.[35] At the current time, some of most widely used FbFPs include iLOV, CreiLOV, EcFbFP, and CagFbFP. For purposes of this dissertation, iLOV receive the primary engineering focus. However, all known FbFPs share conserved spectral traits characterized by broad excitation and emission spectra with peaks at 450 and 495 nm **(Figure 1.1d)**, extinction coefficient of ~14 mM^{-1} cm^{-1}, and quantum yields in the 0.17−0.51 range.[23,24,29,31]

LOV proteins and FbFPs exhibit a conserved fold in which a five-strand beta sheet wraps around the flavin cofactor, with the isoalloxazine moiety of the flavin nestled into the bottom of the binding pocket. The sugar of the flavin cofactor extends out of the binding pocket, where it meets alpha helices flanking the beta sheet **(Figure 1.1a-b)**.[32] In LOV proteins, there typically is an additional helix, denoted Jα, which folds along the rear side of the beta sheet, which denatures in response to blue light excitation.[36] Jα is typically deleted in FbFPs in order to direct energy towards fluorescence rather than conformational changes.

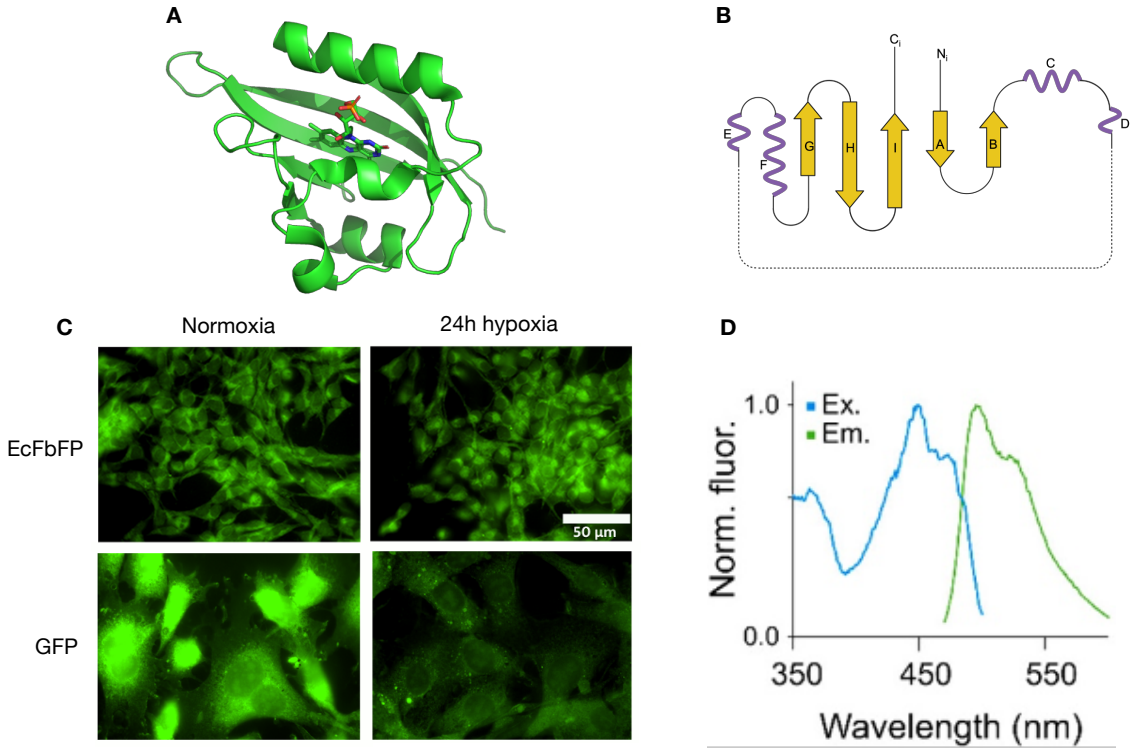

Figure 1.1. Essential characteristics of flavin-based fluorescent proteins. a) The fold of iLOV, typical of FbFPs, is characterized by a beta sheet wrapped around a flavin cofactor, flanked by alpha helices. Adapted from PDB ID 4EES.[32] **b)** iLOV and other FbFPs may be represented with a two-dimensional topology map. Beta strands and alpha helices are designated with letters running from the N-terminus to the C-terminus. **c)** FbFPs show similar fluorescence under normoxic and hypoxic conditions, while GFP fluorescence is nearly completely

eliminated under hypoxic conditions (adapted with permission[37]) **d)** Fluorescence excitation and emission spectra of CreiLOV, typical of an FbFP.

Flavin-based fluorescent proteins have found use for imaging a broad diversity of anaerobic bacteria, fungi, and hypoxically cultured mammalian cells.[37–49] These applications have ranged from tracking horizontal gene transfer and cell division in anaerobic environments to screening promoters for metabolic engineering and genetic circuit designing in anaerobes.[42,43,50–54] Other applications have harnessed the small size of FbFPs (~13 kDa, half that of GFP) to translationally tag proteins in scenarios where the comparatively larger steric footprint of GFP interferes with protein function.[27,46,55,56] In addition, several LOV reporters display remarkable pH stability (pKa ~ 3 for EcFbFP and CreiLOV), which has been exploited to develop pH sensors based on Förster resonance energy transfer (FRET) between LOV and a pH-sensitive GFP variant.[35,57] Aside from FRET, LOV-based biosensors have also been constructed through the incorporation of unnatural amino acids in the flavin binding pocket to selectively modulate fluorescence via photoinduced electron transfer between the amino acid and flavin.[58]

1.3. Key challenges with existing flavin-based fluorescent proteins

As with any technology in its infancy, the existing suite of FbFP reporters comes with limitations as well as exciting opportunities for future research. First and foremost, even the brightest FbFPs are considerably dimmer than GFP, achieving only 10% of GFP fluorescence upon expression in cells.[2,35] Given that cellular brightness of any reporter depends on multiple factors such as protein stability, solubility, and effective fluorescent fraction; it is important for engineering efforts to pursue a holistic approach to development of FbFPs, improving both their spectral properties and their function in cellular contexts. In principle, it should be possible to increase cellular fluorescence at least 4-fold by concurrently maximizing quantum

yield (2- fold increase theoretically possible) and solubility (soluble fraction in E. coli for most FbFPs is less than 50%, unpublished data from our lab) in the brightest available FbFPs. This could be achieved either through rational structure guided mutagenesis or through directed evolution techniques. A second challenge stems from the lack of variously colored FbFPs, particularly ones that are sufficiently red-shifted to avoid overlapping with cellular autofluorescence, which is particularly high at green wavelengths. A recent theoretical study suggested the possibility of red-shifting FbFPs using specific mutations, but these predictions were contradicted by follow-up experiments.[59–61] To our knowledge, efforts to engineer spectral shifts in the existing FbFP repertoire using directed evolution have also proved unsuccessful. Going by the tremendous impact of multicolored GFP variants,[62,63] it is clear that innovative approaches to diversify the FbFP color palette should be a major focus of future research in this area. A final unexplored avenue relates to the possible effects exerted on a cell's endogenous flavin pool due to overexpression of FbFPs. FbFPs bind flavins with moderately strong affinity ($K_d \approx 180-230$ nM),[64] which may be sufficient to deplete free reserves of cellular flavin, estimated to be ≤ 4 μM in E. coli and mammalian cells.[65–68] It has therefore been suggested (albeit not experimentally studied) that flavin sequestration by FbFPs could lead to an increase in flavin biosynthesis via feedback mechanisms.[29] Could FbFPs perturb flavin homeostasis sufficiently to affect cell physiology? Alternatively, could cellular fluorescence of FbFPs be enhanced by augmenting flavin biosynthesis? Rigorous benchmarking of FbFPs against complementary readouts of gene expression (e.g., lacZ based colorimetric assays) in different environments, across multiple cell types, and varying intracellular flavin levels should help address these questions.

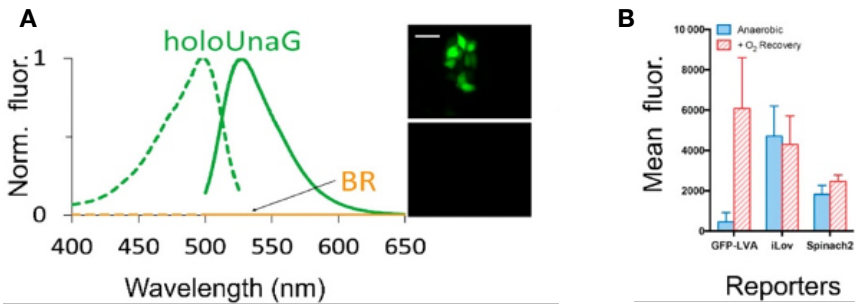

Figure 1.2. Alternative oxygen-independent biomolecular reporters. a) UnaG emits green fluorescence by associating with a heme end-product (biliverdin), which enables imaging of HeLa cells in 0.1% hypoxia (top panel). Under similar conditions, HeLa cells that express mCherry (bottom panel) are nonfluorescent.[69] **b)** Spinach2 is a small molecule dye-binding aptamer that exhibits oxygen-independent fluorescence similar to iLOV, a LOV-based fluorescent reporter. In contrast, GFP is nonfluorescent in anaerobic conditions and needs oxygen to activate fluorescence.[70]

1.4. Alternative biomolecular reporters for anaerobic imaging

In addition to flavin-based fluorescent proteins, several distinct classes of biomolecular reporters hold promise for further expanding the low-oxygen imaging toolbox. The first class consists of heme-based reporters, specifically phytochromes and UnaG, which generate fluorescence upon binding end-products of heme degradation such as biliverdin and bilirubin (**Figure 1.2a**).[69,71] Notably, UnaG and phytochromes have been used to engineer several fluorescent protein biosensors (albeit, yet to be demonstrated in anaerobic systems), which could be useful for probing kinase activity, calcium signaling, redox, protein-protein interactions, and other aspects of cell function in varying oxygen settings.[72–74] However, a key challenge with using heme-based reporters is that bilirubin and biliverdin are not native to all cell types -- for instance, bacteria do not typically synthesize either heme metabolite.[3] In such cases, the chromophores must be supplied exogenously, which is hindered by the poor permeability and limited aqueous solubility of heme compounds. As a partial workaround, phytochromes have been used in conjunction with heme oxygenase to directly synthesize

biliverdin in cells via enzymatic breakdown of heme.[75–77] However, the enzymatic conversion of heme to biliverdin requires oxygen, which makes this approach impractical in anaerobes. Further, heme is also absent in many organisms which lack biliverdin. This multi-gene expression system causes implementation of biliverdin-based fluorescent proteins highly unwieldly in many contexts.

The second class of low-oxygen compatible reporters consists of fluorogenic proteins (e.g., derivatives of photoactive yellow protein or PYP such as FAST) and RNA aptamers engineered to bind synthetic small molecule dyes (e.g., hydroxybenzylidene imidazolinone, coumarin) and activate fluorescence by suppressing fluorescence quenching intramolecular movements in the unbound dye.[70,78–84] Although PYP-based fluorogenic reporter proteins have not yet been demonstrated in anaerobic systems, an RNA-based reporter was recently engineered to develop an anaerobically compatible biosensor for detecting the cellular second messenger, cyclic-di-GMP (**Figure 1.1d**).[70] While promising, PYP and aptamer-based reporters require exogeneously-supplied fluorophores, which poses a potential problem in cell types where membrane permeability is low (e.g., Gram-negative bacteria) as well as in studies where intrinsic membrane permeability is affected by the experimental conditions (e.g., antibiotic use).[85–89] This requirement of exogeneous cofactors compounds the cofactor delivery issue seen with biliverdin-based reporters, as these cofactors are synthetic and thus do not have known biosynthetic pathways. Further, aptamer-based reporters do not readily support fusion to proteins, as RNA and peptides do not become covalently linked during their synthesis.

Finally, a new class of fluorescent reporters has been recently developed based on microbial opsins that bind retinal or synthetic analogs as their chromophore. While this fluorescence mechanism is likely to be oxygen-independent (although yet to be demonstrated

experimentally), the poor solubility of opsin-based fluorescent proteins restricts expression to the plasma membrane, which potentially limits their broad utility as biomolecular reporters.[90,91]

1.5. Outline of dissertation

This dissertation details work I have done to understand, engineer, and advance flavin-based fluorescent proteins as tools in fluorescent bioimaging. The chapters that follow are divided into independent narratives with appropriate background, methods, data, conclusions, references, and supplementary information. Brief summaries of each chapter are provided below.

Chapter 2. Characterization of flavin binding in oxygen-independent fluorescent proteins. While flavin-based fluorescent proteins offer a distinct advantage over other fluorescent proteins in that they require no external factors (neither oxygen nor an exogeneous cofactor), their binding behavior with their flavin chromophore is poorly understood. To fill this knowledge gap, we developed a method to prepare deflavinated FbFPs, and then leveraged these preparations to determine the dissociation coefficients of FbFPs iLOV and EcFbFP with three relevant flavin cofactors, riboflavin, flavin mononucleotide, and flavin adenine dinucleotide. We then demonstrated that deflavinated FbFPs can be utilized as *in vitro* flavin sensors in consumer products.

Chapter 3. Rational design of a circularly permuted flavin-based fluorescent protein. In order to enable engineering of a wider variety of FbFP-based biosensors, we created a circular permutation of iLOV. Careful selection of the circular permutation site, engineering of the linker connecting the original termini, and addition of dimerizing domains to the new termini resulted in a protein with similar fluorescence to linear iLOV. We then used this circularly permuted iLOV to create sensors for the activity of two proteases: tobacco etch

virus protease and SARS-CoV-2 main protease. We anticipate that this advance, similar to the previous creation of a circularly permuted green fluorescent protein, will broaden the scope of biosensor engineering using FbFPs.

Chapter 4. Engineering a minimally sized flavin-based fluorescent protein. Here, we seek to enhance a major advantage of FbFPs: their small size. We utilized structural data and domain insertion to identify residues in iLOV which can be deleted without a large loss of fluorescence. We ultimately identified ten residues which can be eliminated, yielding a protein we have termed "nanoLOV," which retains 86.7 ± 2.2% of iLOV's fluorescence, at a length of only 101 amino acids.

Chapter 5. Conclusions. We summarize the advances the preceding chapters have contributed to the advancement of flavin-based fluorescent proteins and the field of molecular bioimaging in general. We also highlight future development that could be enabled by our work as well as broader biophysical questions that are raised by trends observed in engineering FbFPs.

1.6. References

(1) Heim, R.; Prasher, D. C.; Tsien, R. Y. Wavelength Mutations and Posttranslational Autoxidation of Green Fluorescent Protein. *Proceedings of the National Academy of Sciences of the United States of America* **1994**, *91* (26), 12501–12504. https://doi.org/10.1073/pnas.91.26.12501.

(2) Mukherjee, A.; Walker, J.; Weyant, K. B.; Schroeder, C. M. Characterization of Flavin-Based Fluorescent Proteins: An Emerging Class of Fluorescent Reporters. *PLoS ONE* **2013**, *8* (5), e64753. https://doi.org/10.1371/journal.pone.0064753.

(3) Chia, H. E.; Marsh, E. N. G.; Biteen, J. S. Extending Fluorescence Microscopy into Anaerobic Environments. *Current Opinion in Chemical Biology* **2019**, *51*, 98–104. https://doi.org/10.1016/j.cbpa.2019.05.008.

(4) Mimee, M.; Nadeau, P.; Hayward, A.; Carim, S.; Flanagan, S.; Jerger, L.; Collins, J.; McDonnell, S.; Swartwout, R.; Citorik, R. J.; Bulović, V.; Langer, R.; Traverso, G.; Chandrakasan, A. P.; Lu, T. K. An Ingestible Bacterial-Electronic System to Monitor Gastrointestinal Health. *Science* **2018**, *360* (6391), 915–918. https://doi.org/10.1126/science.aas9315.

(5) Hwang, I. Y.; Koh, E.; Wong, A.; March, J. C.; Bentley, W. E.; Lee, Y. S.; Chang, M. W. Engineered Probiotic Escherichia Coli Can Eliminate and Prevent Pseudomonas Aeruginosa Gut Infection in Animal Models. *Nature Communications* **2017**, *8* (1), 15028. https://doi.org/10.1038/ncomms15028.

(6) Ho, C. L.; Tan, H. Q.; Chua, K. J.; Kang, A.; Lim, K. H.; Ling, K. L.; Yew, W. S.; Lee, Y. S.; Thiery, J. P.; Chang, M. W. Engineered Commensal Microbes for Diet-Mediated Colorectal-Cancer Chemoprevention. *Nature Biomedical Engineering* **2018**, *2* (1), 27–37. https://doi.org/10.1038/s41551-017-0181-y.

(7) Hemsley, C. M.; Luo, J. X.; Andreae, C. A.; Butler, C. S.; Soyer, O. S.; Titball, R. W. Bacterial Drug Tolerance under Clinical Conditions Is Governed by Anaerobic Adaptation but Not Anaerobic Respiration. *Antimicrobial Agents and Chemotherapy* **2014**, *58* (10), 5775–5783. https://doi.org/10.1128/aac.02793-14.

(8) Carreau, A.; Hafny-Rahbi, B. E.; Matejuk, A.; Grillon, C.; Kieda, C. Why Is the Partial Oxygen Pressure of Human Tissues a Crucial Parameter? Small Molecules and

Hypoxia. *Journal of Cellular and Molecular Medicine* **2011**, *15* (6), 1239–1253. https://doi.org/10.1111/j.1582-4934.2011.01258.x.

(9) Wilde, A. D.; Snyder, D. J.; Putnam, N. E.; Valentino, M. D.; Hammer, N. D.; Lonergan, Z. R.; Hinger, S. A.; Aysanoa, E. E.; Blanchard, C.; Dunman, P. M.; Wasserman, G. A.; Chen, J.; Shopsin, B.; Gilmore, M. S.; Skaar, E. P.; Cassat, J. E. Bacterial Hypoxic Responses Revealed as Critical Determinants of the Host-Pathogen Outcome by TnSeq Analysis of Staphylococcus Aureus Invasive Infection. *PLOS Pathogens* **2015**, *11* (12), e1005341.

(10) Onderdonk, A. B. Animal Models Simulating Anaerobic Infections. *Anaerobe* **2005**, *11* (4), 189–195. https://doi.org/10.1016/j.anaerobe.2004.12.001.

(11) DeGruttola, A. K.; Low, D.; Mizoguchi, A.; Mizoguchi, E. Current Understanding of Dysbiosis in Disease in Human and Animal Models. *Inflammatory Bowel Diseases* **2016**, *22* (5), 1137–1150. https://doi.org/10.1097/MIB.0000000000000750.

(12) Wei, X.-X.; Shi, Z.-Y.; Yuan, M.-Q.; Chen, G.-Q. Effect of Anaerobic Promoters on the Microaerobic Production of Polyhydroxybutyrate (PHB) in Recombinant Escherichia Coli. *Applied Microbiology and Biotechnology* **2009**, *82* (4), 703–712. https://doi.org/10.1007/s00253-008-1816-4.

(13) Podolsky, I. A.; Seppälä, S.; Lankiewicz, T. S.; Brown, J. L.; Swift, C. L.; O'Malley, M. A. Harnessing Nature's Anaerobes for Biotechnology and Bioprocessing. *Annual Review of Chemical and Biomolecular Engineering* **2019**, *10* (1), 105–128. https://doi.org/10.1146/annurev-chembioeng-060718-030340.

(14) Zeitouni, N. E.; Chotikatum, S.; von Köckritz-Blickwede, M.; Naim, H. Y. The Impact of Hypoxia on Intestinal Epithelial Cell Functions: Consequences for Invasion by

Bacterial Pathogens. *Molecular and Cellular Pediatrics* **2016**, *3* (1), 14. https://doi.org/10.1186/s40348-016-0041-y.

(15) Cummins, E. P.; Taylor, C. T. Hypoxia and Inflammation. *The Biochemist* **2017**, *39* (4), 34–36. https://doi.org/10.1042/BIO03904034.

(16) Hassett, D. J.; Cuppoletti, J.; Trapnell, B.; Lymar, S. V.; Rowe, J. J.; Sun Yoon, S.; Hilliard, G. M.; Parvatiyar, K.; Kamani, M. C.; Wozniak, D. J.; Hwang, S.-H.; McDermott, T. R.; Ochsner, U. A. Anaerobic Metabolism and Quorum Sensing by Pseudomonas Aeruginosa Biofilms in Chronically Infected Cystic Fibrosis Airways: Rethinking Antibiotic Treatment Strategies and Drug Targets. *Advanced Drug Delivery Reviews* **2002**, *54* (11), 1425–1443. https://doi.org/10.1016/S0169-409X(02)00152-7.

(17) Tropini, C.; Earle, K. A.; Huang, K. C.; Sonnenburg, J. L. The Gut Microbiome: Connecting Spatial Organization to Function. *Cell Host & Microbe* **2017**, *21* (4), 433–442. https://doi.org/10.1016/j.chom.2017.03.010.

(18) Guglielmetti, S.; Santala, V.; Mangayil, R.; Ciranna, A.; Karp, M. T. O2-Requiring Molecular Reporters of Gene Expression for Anaerobic Microorganisms. *Biosensors and Bioelectronics* **2019**, *123*, 1–6. https://doi.org/10.1016/j.bios.2018.09.066.

(19) Hansen, M. C.; Palmer, R. J.; Udsen, C.; White, D. C.; Molin, S. Assessment of GFP Fluorescence in Cells of Streptococcus Gordonii under Conditions of Low pH and Low Oxygen Concentration. *Microbiology* **2001**, *147* (5), 1383–1391. https://doi.org/10.1099/00221287-147-5-1383.

(20) Coralli, C.; Cemazar, M.; Kanthou, C.; Tozer, G. M.; Dachs, G. U. Limitations of the Reporter Green Fluorescent Protein under Simulated Tumor Conditions. *Cancer research* **2001**, *61* (12), 4784–4790.

(21) Scott, K. P.; Mercer, D. K.; Glover, L. A.; Flint, H. J. The Green Fluorescent Protein as a Visible Marker for Lactic Acid Bacteria in Complex Ecosystems. *FEMS Microbiology Ecology* **1998**, *26* (3), 219–230. https://doi.org/10.1111/j.1574-6941.1998.tb00507.x.

(22) Zhang, C.; Xing, X.-H.; Lou, K. Rapid Detection of a Gfp-Marked Enterobacter Aerogenes under Anaerobic Conditions by Aerobic Fluorescence Recovery. *FEMS Microbiology Letters* **2005**, *249* (2), 211–218. https://doi.org/10.1016/j.femsle.2005.05.051.

(23) Buckley, A. M.; Petersen, J.; Roe, A. J.; Douce, G. R.; Christie, J. M. LOV-Based Reporters for Fluorescence Imaging. *Current Opinion in Chemical Biology* **2015**, *27*, 39–45. https://doi.org/10.1016/j.cbpa.2015.05.011.

(24) Mukherjee, A.; Schroeder, C. M. Flavin-Based Fluorescent Proteins: Emerging Paradigms in Biological Imaging. *Current Opinion in Biotechnology* **2015**, *31*, 16–23. https://doi.org/10.1016/j.copbio.2014.07.010.

(25) Losi, A.; Gärtner, W. Solving Blue Light Riddles: New Lessons from Flavin-Binding LOV Photoreceptors. *Photochemistry and Photobiology* **2017**, *93* (1), 141–158. https://doi.org/10.1111/php.12674.

(26) Drepper, T.; Eggert, T.; Circolone, F.; Heck, A.; Krauss, U.; Guterl, J. K.; Wendorff, M.; Losi, A.; Gartner, W.; Jaeger, K. E. Reporter Proteins for in Vivo Fluorescence without Oxygen. *Nature Biotechnology* **2007**, *25* (4), 443–445. https://doi.org/10.1038/nbt1293.

(27) Chapman, S.; Faulkner, C.; Kaiserli, E.; Garcia-Mata, C.; Savenkov, E. I.; Roberts, A. G.; Oparka, K. J.; Christie, J. M. The Photoreversible Fluorescent Protein iLOV Outperforms GFP as a Reporter of Plant Virus Infection. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105* (50), 20038–20043. https://doi.org/10.1073/pnas.0807551105.

(28) Nazarenko, V. V.; Remeeva, A.; Yudenko, A.; Kovalev, K.; Dubenko, A.; Goncharov, I. M.; Kuzmichev, P.; Rogachev, A. V.; Buslaev, P.; Borshchevskiy, V.; Mishin, A.; Dhoke, G. V.; Schwaneberg, U.; Davari, M. D.; Jaeger, K.-E.; Krauss, U.; Gordeliy, V.; Gushchin, I. A Thermostable Flavin-Based Fluorescent Protein from Chloroflexus Aggregans: A Framework for Ultra-High Resolution Structural Studies. *Photochemical & Photobiological Sciences* **2019**. https://doi.org/10.1039/C9PP00067D.

(29) Wingen, M.; Jaeger, K.-E.; Gensch, T.; Drepper, T. Novel Thermostable Flavin-Binding Fluorescent Proteins from Thermophilic Organisms. *Photochemistry and Photobiology* **2017**, *93* (3), 849–856. https://doi.org/10.1111/php.12740.

(30) Song, X.; Wang, Y.; Shu, Z.; Hong, J.; Li, T.; Yao, L. Engineering a More Thermostable Blue Light Photo Receptor Bacillus Subtilis YtvA LOV Domain by a Computer Aided Rational Design Method. *PLOS Computational Biology* **2013**, *9* (7), e1003129.

(31) Wingen, M.; Potzkei, J.; Endres, S.; Casini, G.; Rupprecht, C.; Fahlke, C.; Krauss, U.; Jaeger, K.-E.; Drepper, T.; Gensch, T. The Photophysics of LOV-Based Fluorescent Proteins — New Tools for Cell Biology. *Photochemical & Photobiological Sciences* **2014**, *13* (6), 875–883. https://doi.org/10.1039/c3pp50414j.

(32) Christie, J. M.; Hitomi, K.; Arvai, A. S.; Hartfield, K. A.; Mettlen, M.; Pratt, A. J.; Tainer, J. A.; Getzoff, E. D. Structural Tuning of the Fluorescent Protein iLOV for Improved Photostability. *J Biol Chem* **2012**, *287* (26), 22295–22304. https://doi.org/10.1074/jbc.M111.318881.

(33) Mukherjee, A.; Weyant, K. B.; Walker, J.; Schroeder, C. M. Directed Evolution of Bright Mutants of an Oxygen-Independent Flavin-Binding Fluorescent Protein from

Pseudomonas Putida. *Journal of Biological Engineering* **2012**, *6* (1), 20. https://doi.org/10.1186/1754-1611-6-20.

(34) Higgins, S. A.; Ouonkap, S. V. Y.; Savage, D. F. Rapid and Programmable Protein Mutagenesis Using Plasmid Recombineering. *ACS Synthetic Biology* **2017**, *6* (10), 1825–1833. https://doi.org/10.1021/acssynbio.7b00112.

(35) Mukherjee, A.; Weyant, K. B.; Agrawal, U.; Walker, J.; Cann, I. K. O.; Schroeder, C. M. Engineering and Characterization of New LOV-Based Fluorescent Proteins from Chlamydomonas Reinhardtii and Vaucheria Frigida. *ACS Synthetic Biology* **2015**, *4* (4), 371–377. https://doi.org/10.1021/sb500237x.

(36) Koyama, T.; Iwata, T.; Yamamoto, A.; Sato, Y.; Matsuoka, D.; Tokutomi, S.; Kandori, H. Different Role of the Jα Helix in the Light-Induced Activation of the LOV2 Domains in Various Phototropins. *Biochemistry* **2009**, *48* (32), 7621–7628. https://doi.org/10.1021/bi9009192.

(37) Walter, J.; Hausmann, S.; Drepper, T.; Puls, M.; Eggert, T.; Dihné, M. Flavin Mononucleotide-Based Fluorescent Proteins Function in Mammalian Cells without Oxygen Requirement. *PLOS ONE* **2012**, *7* (9), e43921.

(38) Lobo, L. A.; Smith, C. J.; Rocha, E. R. Flavin Mononucleotide (FMN)-Based Fluorescent Protein (FbFP) as Reporter for Gene Expression in the Anaerobe Bacteroides Fragilis. *FEMS Microbiology Letters* **2011**, *317* (1), 67–74. https://doi.org/10.1111/j.1574-6968.2011.02212.x.

(39) Drepper, T.; Huber, R.; Heck, A.; Circolone, F.; Hillmer, A.-K.; Büchs, J.; Jaeger, K.-E. Flavin Mononucleotide-Based Fluorescent Reporter Proteins Outperform Green Fluorescent Protein-Like Proteins as Quantitative In Vivo Real-Time Reporters. *Applied and*

Environmental Microbiology **2010**, *76* (17), 5990–5994. https://doi.org/10.1128/AEM.00701-10.

(40) Landete, J. M.; Langa, S.; Revilla, C.; Margolles, A.; Medina, M.; Arqués, J. L. Use of Anaerobic Green Fluorescent Protein versus Green Fluorescent Protein as Reporter in Lactic Acid Bacteria. *Applied Microbiology and Biotechnology* **2015**, *99* (16), 6865–6877. https://doi.org/10.1007/s00253-015-6770-3.

(41) Landete, J. M.; Peirotén, Á.; Rodríguez, E.; Margolles, A.; Medina, M.; Arqués, J. L. Anaerobic Green Fluorescent Protein as a Marker of Bifidobacterium Strains. *International Journal of Food Microbiology* **2014**, *175*, 6–13. https://doi.org/10.1016/j.ijfoodmicro.2014.01.008.

(42) Clark, I. C.; Youngblut, M.; Jacobsen, G.; Wetmore, K. M.; Deutschbauer, A.; Lucas, L.; Coates, J. D. Genetic Dissection of Chlorate Respiration in Pseudomonas Stutzeri PDA Reveals Syntrophic (per)Chlorate Reduction. *Environmental Microbiology* **2016**, *18* (10), 3342–3354. https://doi.org/10.1111/1462-2920.13068.

(43) Seo, S.-O.; Lu, T.; Jin, Y.-S.; Blaschek, H. P. Development of an Oxygen-Independent Flavin Mononucleotide-Based Fluorescent Reporter System in Clostridium Beijerinckii and Its Potential Applications. *Journal of Biotechnology* **2018**, *265*, 119–126. https://doi.org/10.1016/j.jbiotec.2017.11.003.

(44) Elgamoudi, B. A.; Ketley, J. M. Lighting up My Life: A LOV-Based Fluorescent Reporter for Campylobacter Jejuni. *Research in Microbiology* **2018**, *169* (2), 108–114. https://doi.org/10.1016/j.resmic.2017.10.003.

(45) Tielker, D.; Eichhof, I.; Jaeger, K. E.; Ernst, J. F. Flavin Mononucleotide-Based Fluorescent Protein as an Oxygen-Independent Reporter in Candida Albicans and

Saccharomyces Cerevisiae. *Eukaryotic Cell* **2009**, *8* (6), 913–915. https://doi.org/10.1128/ec.00394-08.

(46) Emmi, S. K.; Benita, K.; Astrid, W.; Martina, P.; Karl-Erich, J.; Ulrich, K. Fusion of a Flavin-Based Fluorescent Protein to Hydroxynitrile Lyase from Arabidopsis Thaliana Improves Enzyme Stability. *Applied and Environmental Microbiology* **2013**, *79* (15), 4727–4733. https://doi.org/10.1128/AEM.00795-13.

(47) Choi, C. H.; DeGuzman, J. V.; Lamont, R. J.; Yilmaz, Ö. Genetic Transformation of an Obligate Anaerobe, P. Gingivalis for FMN-Green Fluorescent Protein Expression in Studying Host-Microbe Interaction. *PLOS ONE* **2011**, *6* (4), e18499.

(48) Gawthorne, J. A.; Reddick, L. E.; Akpunarlieva, S. N.; Beckham, K. S. H.; Christie, J. M.; Alto, N. M.; Gabrielsen, M.; Roe, A. J. Express Your LOV: An Engineered Flavoprotein as a Reporter for Protein Expression and Purification. *PLOS ONE* **2012**, *7* (12), e52962.

(49) Wang, S. E.; Brooks, A. E. S.; Cann, B.; Simoes-Barbosa, A. The Fluorescent Protein iLOV Outperforms eGFP as a Reporter Gene in the Microaerophilic Protozoan Trichomonas Vaginalis. *Molecular and Biochemical Parasitology* **2017**, *216*, 1–4. https://doi.org/10.1016/J.MOLBIOPARA.2017.06.003.

(50) E., K. J.; M., R. L.; M., K. S.; M., T. E. Dual Reporter System for In Situ Detection of Plasmid Transfer under Aerobic and Anaerobic Conditions. *Applied and Environmental Microbiology* **2010**, *76* (13), 4553–4556. https://doi.org/10.1128/AEM.00226-10.

(51) Buckley, A. M.; Jukes, C.; Candlish, D.; Irvine, J. J.; Spencer, J.; Fagan, R. P.; Roe, A. J.; Christie, J. M.; Fairweather, N. F.; Douce, G. R. Lighting Up Clostridium Difficile:

Reporting Gene Expression Using Fluorescent Lov Domains. *Scientific Reports* **2016**, *6* (1), 23463. https://doi.org/10.1038/srep23463.

(52) Lara, A. R.; Jaén, K. E.; Sigala, J.-C.; Mühlmann, M.; Regestein, L.; Büchs, J. Characterization of Endogenous and Reduced Promoters for Oxygen-Limited Processes Using Escherichia Coli. *ACS Synthetic Biology* **2017**, *6* (2), 344–356. https://doi.org/10.1021/acssynbio.6b00233.

(53) Teng, L.; Wang, K.; Xu, J.; Xu, C. Flavin Mononucleotide (FMN)-Based Fluorescent Protein (FbFP) as Reporter for Promoter Screening in Clostridium Cellulolyticum. *Journal of Microbiological Methods* **2015**, *119*, 37–43. https://doi.org/10.1016/j.mimet.2015.09.018.

(54) Immethun, C. M.; Ng, K. M.; DeLorenzo, D. M.; Waldron-Feinstein, B.; Lee, Y.-C.; Moon, T. S. Oxygen-Responsive Genetic Circuits Constructed in Synechocystis Sp. PCC 6803. *Biotechnology and Bioengineering* **2016**, *113* (2), 433–442. https://doi.org/10.1002/bit.25722.

(55) Seago, J.; Juleff, N.; Moffat, K.; Berryman, S.; Christie, J. M.; Charleston, B.; Jackson, T. An Infectious Recombinant Foot-and-Mouth Disease Virus Expressing a Fluorescent Marker Protein. *Journal of General Virology* **2013**, *94* (7), 1517–1527. https://doi.org/10.1099/vir.0.052308-0.

(56) A., G. J.; Laurent, A.; Claire, M.; Paul, D.; M., C. J.; Jost, E.; J., R. A. Visualizing the Translocation and Localization of Bacterial Type III Effector Proteins by Using a Genetically Encoded Reporter System. *Applied and Environmental Microbiology* **2016**, *82* (9), 2700–2708. https://doi.org/10.1128/AEM.03418-15.

(57) Rupprecht, C.; Wingen, M.; Potzkei, J.; Gensch, T.; Jaeger, K.-E.; Drepper, T. A Novel FbFP-Based Biosensor Toolbox for Sensitive in Vivo Determination of Intracellular

pH. *Journal of Biotechnology* **2017**, *258*, 25–32. https://doi.org/10.1016/j.jbiotec.2017.05.006.

(58) Liu, X.; Jiang, L.; Li, J.; Wang, L.; Yu, Y.; Zhou, Q.; Lv, X.; Gong, W.; Lu, Y.; Wang, J. Significant Expansion of Fluorescent Protein Sensing Ability through the Genetic Incorporation of Superior Photo-Induced Electron-Transfer Quenchers. *Journal of the American Chemical Society* **2014**, *136* (38), 13094–13097. https://doi.org/10.1021/ja505219r.

(59) Khrenova, M. G.; Nemukhin, A. V.; Domratcheva, T. Theoretical Characterization of the Flavin-Based Fluorescent Protein iLOV and Its Q489K Mutant. *The Journal of Physical Chemistry B* **2015**, *119* (16), 5176–5183. https://doi.org/10.1021/acs.jpcb.5b01299.

(60) Meteleshko, Y. I.; Nemukhin, A. V.; Khrenova, M. G. Novel Flavin-Based Fluorescent Proteins with Red-Shifted Emission Bands: A Computational Study. *Photochemical & Photobiological Sciences* **2019**, *18* (1), 177–189. https://doi.org/10.1039/c8pp00361k.

(61) Davari, M. D.; Kopka, B.; Wingen, M.; Bocola, M.; Drepper, T.; Jaeger, K.-E.; Schwaneberg, U.; Krauss, U. Photophysics of the LOV-Based Fluorescent Protein Variant iLOV-Q489K Determined by Simulation and Experiment. *The Journal of Physical Chemistry B* **2016**, *120* (13), 3344–3352. https://doi.org/10.1021/acs.jpcb.6b01512.

(62) Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y. A Guide to Choosing Fluorescent Proteins. *Nature Methods* **2005**, *2* (12), 905–909. https://doi.org/10.1038/nmeth819.

(63) Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y. Improved Monomeric Red, Orange and Yellow Fluorescent Proteins Derived from Discosoma Sp. Red Fluorescent Protein. *Nature Biotechnology* **2004**, *22*, 1567.

https://doi.org/10.1038/nbt1037 https://www.nature.com/articles/nbt1037#supplementary-information.

(64) Anderson, N. T.; Weyant, K. B.; Mukherjee, A. Characterization of Flavin Binding in Oxygen-Independent Fluorescent Reporters. *AIChE Journal* **2020**, *66* (12), e17083. https://doi.org/10.1002/aic.17083.

(65) Ulrich, K.; Vera, S.; Astrid, W.; Esther, K.-G.; Karl-Erich, J. Cofactor Trapping, a New Method To Produce Flavin Mononucleotide. *Applied and Environmental Microbiology* **2011**, *77* (3), 1097–1100. https://doi.org/10.1128/AEM.01541-10.

(66) Hühner, J.; Ingles-Prieto, Á.; Neusüß, C.; Lämmerhofer, M.; Janovjak, H. Quantification of Riboflavin, Flavin Mononucleotide, and Flavin Adenine Dinucleotide in Mammalian Model Cells by CE with LED-Induced Fluorescence Detection. *ELECTROPHORESIS* **2015**, *36* (4), 518–525. https://doi.org/10.1002/elps.201400451.

(67) Wilson, A. C.; Pardee, A. B. Regulation of Flavin Synthesis by Escherichia Coli. *Microbiology* **1962**, *28* (2), 283–303.

(68) McAnulty, M. J.; Wood, T. K. YeeO from Escherichia Coli Exports Flavins. *Bioengineered* **2014**, *5* (6), 386–392. https://doi.org/10.4161/21655979.2014.969173.

(69) Kumagai, A.; Ando, R.; Miyatake, H.; Greimel, P.; Kobayashi, T.; Hirabayashi, Y.; Shimogori, T.; Miyawaki, A. A Bilirubin-Inducible Fluorescent Protein from Eel Muscle. *Cell* **2013**, *153* (7), 1602–1611. https://doi.org/10.1016/j.cell.2013.05.038.

(70) Wang, X. C.; Wilson, S. C.; Hammond, M. C. Next-Generation RNA-Based Fluorescent Biosensors Enable Anaerobic Detection of Cyclic Di-GMP. *Nucleic Acids Research* **2016**, *44* (17), e139–e139. https://doi.org/10.1093/nar/gkw580.

(71) Oliinyk, O. S.; Chernov, K. G.; Verkhusha, V. V. Bacterial Phytochromes, Cyanobacteriochromes and Allophycocyanins as a Source of Near-Infrared Fluorescent Probes. *International Journal of Molecular Sciences* **2017**, *18* (8). https://doi.org/10.3390/ijms18081691.

(72) To, T.-L.; Zhang, Q.; Shu, X. Structure-Guided Design of a Reversible Fluorogenic Reporter of Protein-Protein Interactions. *Protein Science* **2016**, *25* (3), 748–753. https://doi.org/10.1002/pro.2866.

(73) Hu, H.; Wang, A.; Huang, L.; Zou, Y.; Gu, Y.; Chen, X.; Zhao, Y.; Yang, Y. Monitoring Cellular Redox State under Hypoxia Using a Fluorescent Sensor Based on Eel Fluorescent Protein. *Free Radical Biology and Medicine* **2018**, *120*, 255–265. https://doi.org/10.1016/j.freeradbiomed.2018.03.041.

(74) Shitashima, Y.; Shimozawa, T.; Asahi, T.; Miyawaki, A. A Dual-Ligand-Modulable Fluorescent Protein Based on UnaG and Calmodulin. *Biochemical and Biophysical Research Communications* **2018**, *496* (3), 872–879. https://doi.org/10.1016/j.bbrc.2018.01.134.

(75) Rumyantsev, K. A.; Shcherbakova, D. M.; Zakharova, N. I.; Verkhusha, V. V.; Turoverov, K. K. Design of Near-Infrared Single-Domain Fluorescent Protein GAF-FP Based on Bacterial Phytochrome. *Cell and Tissue Biology* **2017**, *11* (1), 16–26. https://doi.org/10.1134/S1990519X17010102.

(76) Rodriguez, E. A.; Tran, G. N.; Gross, L. A.; Crisp, J. L.; Shu, X.; Lin, J. Y.; Tsien, R. Y. A Far-Red Fluorescent Protein Evolved from a Cyanobacterial Phycobiliprotein. *Nature Methods* **2016**, *13* (9), 763–769. https://doi.org/10.1038/nmeth.3935.

(77) Shu, X.; Royant, A.; Lin, M. Z.; Aguilera, T. A.; Lev-Ram, V.; Steinbach, P. A.; Tsien, R. Y. Mammalian Expression of Infrared Fluorescent Proteins Engineered from a

Bacterial Phytochrome. *Science* **2009**, *324* (5928), 804–807. https://doi.org/10.1126/science.1168683.

(78) Hori, Y.; Ueno, H.; Mizukami, S.; Kikuchi, K. Photoactive Yellow Protein-Based Protein Labeling System with Turn-On Fluorescence Intensity. *Journal of the American Chemical Society* **2009**, *131* (46), 16610–16611. https://doi.org/10.1021/ja904800k.

(79) Yapici, I.; Lee, K. S. S.; Berbasova, T.; Nosrati, M.; Jia, X.; Vasileiou, C.; Wang, W.; Santos, E. M.; Geiger, J. H.; Borhan, B. "Turn-On" Protein Fluorescence: In Situ Formation of Cyanine Dyes. *Journal of the American Chemical Society* **2015**, *137* (3), 1073–1080. https://doi.org/10.1021/ja506376j.

(80) Plamont, M.-A.; Billon-Denis, E.; Maurin, S.; Gauron, C.; Pimenta, F. M.; Specht, C. G.; Shi, J.; Quérard, J.; Pan, B.; Rossignol, J.; Moncoq, K.; Morellet, N.; Volovitch, M.; Lescop, E.; Chen, Y.; Triller, A.; Vriz, S.; Le Saux, T.; Jullien, L.; Gautier, A. Small Fluorescence-Activating and Absorption-Shifting Tag for Tunable Protein Imaging in Vivo. *Proceedings of the National Academy of Sciences* **2016**, *113* (3), 497–502. https://doi.org/10.1073/pnas.1513094113.

(81) Bruchez, M. P. Dark Dyes–Bright Complexes: Fluorogenic Protein Labeling. *Current Opinion in Chemical Biology* **2015**, *27*, 18–23. https://doi.org/10.1016/j.cbpa.2015.05.014.

(82) Kellenberger, C. A.; Wilson, S. C.; Sales-Lee, J.; Hammond, M. C. RNA-Based Fluorescent Biosensors for Live Cell Imaging of Second Messengers Cyclic Di-GMP and Cyclic AMP-GMP. *Journal of the American Chemical Society* **2013**, *135* (13), 4906–4909. https://doi.org/10.1021/ja311960g.

(83) Jullien, L.; Gautier, A. Fluorogen-Based Reporters for Fluorescence Imaging: A Review. *Methods and Applications in Fluorescence* **2015**, *3* (4), 12. https://doi.org/10.1088/2050-6120/3/4/042007.

(84) Li, C.; Plamont, M.-A.; Sladitschek, H. L.; Rodrigues, V.; Aujard, I.; Neveu, P.; Le Saux, T.; Jullien, L.; Gautier, A. Dynamic Multicolor Protein Labeling in Living Cells. *Chemical Science* **2017**, *8* (8), 5598–5605. https://doi.org/10.1039/C7SC01364G.

(85) Rowland, B.; Purkayastha, A.; Monserrat, C.; Casart, Y.; Takiff, H.; McDonough, K. A. Fluorescence-Based Detection of lacZ Reporter Gene Expression in Intact and Viable Bacteria Including Mycobacterium Species. *FEMS Microbiology Letters* **1999**, *179* (2), 317–325. https://doi.org/10.1111/j.1574-6968.1999.tb08744.x.

(86) Schneider, J. P.; Basler, M. Shedding Light on Biology of Bacterial Cells. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2016**, *371* (1707), 20150499. https://doi.org/10.1098/rstb.2015.0499.

(87) Na, K.; Dirk, L.; Johan, P.; Mehmet, B. Visualization of Periplasmic and Cytoplasmic Proteins with a Self-Labeling Protein Tag. *Journal of Bacteriology* **2016**, *198* (7), 1035–1043. https://doi.org/10.1128/jb.00864-15.

(88) Banaz, N.; Mäkelä, J.; Uphoff, S. Choosing the Right Label for Single-Molecule Tracking in Live Bacteria: Side-by-Side Comparison of Photoactivatable Fluorescent Protein and Halo Tag Dyes. *Journal of Physics D: Applied Physics* **2019**, *52* (6), 64002. https://doi.org/10.1088/1361-6463/aaf255.

(89) Yang, Z.; Weisshaar, J. C. HaloTag Assay Suggests Common Mechanism of E. Coli Membrane Permeabilization Induced by Cationic Peptides. *ACS Chemical Biology* **2018**, *13* (8), 2161–2169. https://doi.org/10.1021/acschembio.8b00336.

(90) Herwig, L.; Rice, A. J.; Bedbrook, C. N.; Zhang, R. K.; Lignell, A.; Cahn, J. K. B.; Renata, H.; Dodani, S. C.; Cho, I.; Cai, L.; Gradinaru, V.; Arnold, F. H. Directed Evolution of a Bright Near-Infrared Fluorescent Rhodopsin Using a Synthetic Chromophore. *Cell Chemical Biology* **2017**, *24* (3), 415–425. https://doi.org/10.1016/j.chembiol.2017.02.008.

(91) McIsaac, R. S.; Engqvist, M. K. M.; Wannier, T.; Rosenthal, A. Z.; Herwig, L.; Flytzanis, N. C.; Imasheva, E. S.; Lanyi, J. K.; Balashov, S. P.; Gradinaru, V.; Arnold, F. H. Directed Evolution of a Far-Red Fluorescent Rhodopsin. *Proceedings of the National Academy of Sciences* **2014**, *111* (36), 13034–13039. https://doi.org/10.1073/pnas.1413987111.

Chapter 2. Characterization of flavin binding in oxygen-independent fluorescent proteins

This chapter is adapted with permission from a published paper. I am a co-first author on the paper. My contributions include protein expression and purification, *in silico* modeling of cofactor binding, and writing and revision of the manuscript.

Paper: *AIChE Journal* **2020** 66:e17083

Authors: **Nolan T. Anderson***, Kevin B. Weyant*, and Arnab Mukherjee

* *Each of these authors contributed equally*

DOI: 10.1002/aic.17083

2.1. Abstract

Fluorescent proteins based on light, oxygen, and voltage (LOV) sensing photoreceptors are among the few reporter gene technologies available for studying living systems in oxygen-free environments that render reporters based on the green fluorescent protein nonfluorescent. LOV reporters develop fluorescence by binding flavin mononucleotide (FMN), which they endogenously obtain from cells. As FMN is essential to cell physiology as well as for determining fluorescence in LOV proteins, it is important to be able to study and characterize flavin binding in LOV reporters. To this end, we report a method for reversibly separating FMN from two commonly used LOV reporters to prepare stable and soluble apoproteins. Using fluorescence titration, we measured the equilibrium dissociation constant for binding with all three cellular flavins: FMN, flavin adenine dinucleotide, and riboflavin. Finally, we exploit the riboflavin affinity of apo LOV reporters, identified in this work, to develop a fluorescence turn-on biosensor for vitamin B2.

2.2. Introduction

Fluorescence imaging using genetic reporters is one of the best established techniques for monitoring dynamic biological processes in live cells.[1–3] However, flagship reporters based on the green fluorescent protein (GFP) depend on environmental oxygen to emit light,[4] which renders them unusable in gut microbes, archaea, anaerobic communities, marine bacteria, and other biological systems that thrive in oxygen-free environments.[5–9] Unlike GFP, chemogenetic reporters function by pairing proteins with synthetic dyes or biogenic fluorophores such as bilins or flavin mononucleotide (FMN), thereby retaining the ability to fluoresce even in oxygen-free conditions.[10–15] Compared to bilins and synthetic dyes, FMN is an essential metabolite found in all living systems,[16] which makes it possible to use FMN based fluorescent proteins to label cells without requiring external agents to be delivered across the largely impenetrable bacterial and fungal cell walls. FMN-based reporters are derived from light, oxygen, and voltage sensing (LOV) domains that are found in certain photoactive proteins, which use FMN to initiate cell signaling by absorbing light and forming a covalent bond with a cysteine residue located in the FMN binding cavity. Mutating the cysteine to alanine blocks this photochemistry and results in a noncovalent complex that exhibits fluorescence with excitation and emission peaks centered at 450 and 495 nm respectively.[17,18] Previously, we have successfully used saturation mutagenesis of the FMN binding pocket to improve quantum yield of a bacterially sourced LOV reporter.[19] We have also discovered brighter LOV reporters from algae by implementing a genome mining-based approach,[20] which was recently extended to identify a new and highly thermostable LOV reporter from thermophilic bacteria.[21] Taken together with additional benefits such as small size (~12 kDa), acid tolerance,[22] and metal-responsive fluorescence,[23–25] LOV reporters are

making it possible to extend fluorescence imaging to several anaerobes, including gut bacteria, oral flora, parasitic protists, and pathogenic fungi.[26–32]

As the fluorescence properties of LOV reporters are determined by protein-bound FMN, it would be useful to develop a method for separating FMN from the protein and reconstituting the apo protein with FMN (or other flavins) in order to characterize the equilibrium dissociation constant (that is, K_d), specificity, reaction kinetics, binding energetics, and related thermodynamic properties. Such a platform is important for engineering and evolving new fluorescent LOV reporters as well as deciphering mechanisms that lead to improvements in fluorescence properties. However, flavoproteins are generally known to tightly bind flavins, with dissociation constants in the sub-micromolar to nanomolar range, which makes them challenging to deflavinate without affecting stability or solubility of the flavin-free apo protein.[33,34] As a result, deflavination techniques need to be carefully tailored for each unique flavoprotein in order to achieve maximum flavin removal, while producing reconstitutable forms of the apo-protein.[34] Although several methods have been developed for resolving flavins[33] from flavodoxins, photolyases, flavin-based redox sensing proteins, oxidases, reductases, and photoreceptors, to our knowledge, these methods have not been tested in existing fluorescent LOV reporters. In this work, we develop a method to reversibly dissociate FMN from two widely used LOV reporters, iLOV[17] (from *A. thaliana*) and EcFbFP[18] (from *B. subtilis*) and use the distinctive fluorescence properties of flavin bound LOV reporters to measure the K_d for FMN binding. We also demonstrate that LOV reporters are capable of binding riboflavin and flavin adenine dinucleotide (FAD), which represent two other forms of flavins commonly found in biological contexts. Finally, given the demand for easy-to-use riboflavin testing methods in pharmaceutical and nutritional sectors, we make use

of deflavinated LOV reporters to develop a simple turn-on fluorescent biosensor for riboflavin and apply it to quantify riboflavin content in complex mixtures, including a commercial multivitamin dietary supplement.

2.3 Materials and Methods

2.3.1 Cloning of LOV reporter genes

LOV reporter genes were synthesized by GenScript (Piscataway, NJ) or Integrated DNA Technologies (Coralville, IA), based on the originally published sequences of EcFbFP[18] and iLOV.[33] The genes were cloned in the pQE80L expression vector using BamHI and HindIII restriction enzymes, which appended a His_6 tag at the N terminus of each LOV protein. PCR amplification, restriction digestion, and ligation were accomplished using standard protocols. Briefly, PCR was carried out in 50 µl reaction volume using 1–10 ng of template DNA and 0.5 µM primers, 0.2 mM dNTPs, and 2.5 units Taq DNA polymerase. The PCR cycle consisted of an initial denaturation at 94°C for 2 min followed by 25 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 45 s. A final extension step at 72°C for 5–10 min was employed to complete synthesis of full-length templates. In some cases, the PCR products were digested with 10 units of DpnI at 37°C for 1 hour in order to remove the methylated template DNA. Amplicons were digested with 10 units each of BamHI and HindIII restriction endonucleases at 37°C for 1 hour and subsequently ligated into pQE80L expression vector digested with BamHI and HindIII using similar reaction conditions. Ligation reactions were performed using 400 units T4 DNA ligase in a 20 µl reaction volume at room temperature for 1 hour. All plasmid constructs were propagated by transformation in *E. coli* DH5α cells using heat shock at 42°C or electroporation at 1.8 kV using a GenePulser electroporator (Biorad). Cells were plated on LB-agar supplemented with 100 µg/ml ampicillin for selection. Plasmids

were isolated from *E. coli* DH5α transformants and used to transform *E. coli* MG1655 cells for protein expression. All plasmid constructs were verified by sequencing (Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign).

2.3.2 Expression and purification of apo and holo LOV reporters

Single colonies of *E. coli* MG1655 transformants expressing the LOV reporter constructs were inoculated in 5 ml Lennox broth supplemented with ampicillin at 100 μg/ml and grown for 16 hours at 37°C with vigorous (300 r.p.m.) orbital shaking. Cells from the overnight cultures were diluted in 0.5 L medium in a 2 L shake flask and grown under similar conditions as before. Protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM when the culture reached an optical density of 0.4–0.6 at 600 nm (typically, 2 hours after inoculation). Protein expression was continued for another 4–6 hours at 37°C before harvesting cells by centrifugation at 5000 x g for 15 min at 4°C. Pellets were stored at – 80°C until use. For protein purification, frozen pellets were thawed at room temperature and resuspended in 10–15 ml lysis buffer (20 mM Tris hydrochloride, 200 mM sodium chloride, 10 mM imidazole, pH 8.0, 1 mg/ml lysozyme) and incubated for 30 minutes at room temperature, followed by ultrasonication (five cycles of 10 1-second pulses of 17–20 W each). Cell debris was removed by centrifuging the lysate at 10,000 x g for 20 min at 4°C and the supernatant was incubated with 2–4 ml of nickel-nitrilotriacetic acid (Ni-NTA) resin (Qiagen) on a rocker for 1 hour at 4°C. The Ni-NTA resin and supernatant were loaded onto a gravity flow chromatographic column (Fischer Scientific) and washed with 40–50 ml of wash buffer (20 mM Tris hydrochloride, 200 mM sodium chloride, 40 mM imidazole, pH 8.0) to remove nonspecifically bound proteins. Finally, the nickel-bound protein was eluted with 20 ml elution buffer (20 mM Tris hydrochloride, 200 mM

sodium chloride, 500 mM imidazole, pH 8.0) and further purified using anion exchange chromatography. For anion exchange, proteins were loaded onto an anion exchange column (HiTrap, GE Life Sciences) by flowing with the Ni-NTA elution buffer followed by step elution using a high salt buffer (20 mM Tris hydrochloride, 1 M sodium chloride, 500 mM imidazole, pH 8.0). The resulting protein was exchanged into phosphate buffered saline (PBS) by dialyzing (10 kDa MWCO Slide-A-Lyzer cassette) against ~150 volumes of PBS for 2–3 hours following which a fresh volume of PBS was added and dialysis continued overnight. In some cases, we were able to bypass the anion exchange step, proceeding straight from Ni-NTA chromatography to dialysis without noticeable loss of protein quality or yield.

For preparing apo protein, the Ni-NTA loaded column was first washed with 15 ml of denaturing buffer (20 mM Tris hydrochloride, 20 mM sodium chloride, pH 8.0, 6 M guanidine hydrochloride) at 4°C for 2 hours. Following the first wash, the nickel-bound protein was incubated overnight with fresh denaturing buffer (15 ml) at 4°C under mild agitation. At the end of overnight incubation, the bound protein was washed again (1–2 times) using fresh denaturing buffer with 1 hour incubation at each wash step. At this stage, the protein turned visibly nonfluorescent, indicating removal of bound flavin. The resulting apo protein was rinsed with 20 ml imidazole-free wash buffer, to wash out the denaturant, and further incubated in the same buffer for an hour at 4°C before eluting with 20 ml elution buffer. Deflavinated apo protein was dialyzed as before to remove imidazole and remaining traces of denaturant. Protein fractions were run on denaturing PAGE gels and quantified using the Bradford assay. Purified holo and apo proteins were stored at 4°C and typically used within 1–2 days.

2.3.3 Analytical gel filtration

We determined oligomeric states of holo and apo LOV reporters using gel filtration chromatography. To this end, a Superdex 200 (GE Healthcare) size exclusion chromatography column was calibrated with globular proteins standards — bovine thyroglobulin (670 kDa), bovine γ-globulin (158 kDa), chicken ovalbumin (44 kDa), horse myoglobin (17 kDa), and vitamin B_{12} (1.35 kDa). Purified LOV proteins were loaded in the column, washed with 50 ml PBS, and elution volumes ($V_{elution}$) corresponding to peaks in the 280 nm absorption chromatogram were recorded. The void volume (V_{void}) and interstitial pore volume (V_{pore}) were calculated based on the elution volumes of thyroglobulin and vitamin B_{12} respectively. The distribution coefficient (K) of each protein between the mobile phase and stationary phase was calculated via **Equation 2.1**.

$$K = \frac{V_{elution} - V_{void}}{V_{pore}} \qquad \textbf{(Equation 2.1)}$$

Molar mass was estimated by calibrating the logarithm of molecular weight of protein standards against K. Oligomeric state was then determined by dividing the estimated molar mass by the known molecular weight of LOV monomers.

2.3.4 Fluorescence and circular dichroism spectroscopy

Fluorescence intensity measurements were performed using 150 µl quartz cuvettes (Starna Cells) in a Cary Varian Eclipse fluorometer. Excitation and emission slit widths were set at 5 nm and the photomultiplier tube gain was set to medium (600 V). For emission scans, excitation was typically performed at 450 nm, and emission spectra were scanned between 470 nm and 600 nm. Excitation spectra were obtained by monitoring emission at 540 nm and scanning excitation wavelengths between 300 nm and 520 nm. For circular dichroism spectroscopy, 150 µl of iLOV (40 µM) or EcFbFP (20 µM) were loaded in a 1 cm path length

quartz cuvette (Starna Cells). IFar-UV CD spectra were recorded using a Jasco 720 spectrometer between with a resolution of 1 nm and a scan rate of 50 nm/min, averaged over 10 individual spectral scans, and expressed in molar ellipticity. Typically, the HT voltage was maintained at <600 V to minimize noise. Prediction of secondary structure from far-UV CD spectra was made using BeStSel.[35]

2.3.5 Analysis of flavin binding by LOV reporters

For fluorescence titration experiments, apo proteins were diluted to a concentration of 10–20 μM in PBS and stock solutions of FMN, FAD, and riboflavin were prepared at a final concentration of 200 μM (also in PBS). Titration was carried out by adding flavin in 1 μl steps to 200 μl apo protein in quartz microcuvettes. After each addition, the solution was gently mixed using a pipette and incubated for 3 min in the dark to ensure complete equilibration. Fluorescence measurements were performed as described before. Flavin binding was assumed to proceed by simple bimolecular interaction as described in **Equation 2.2**.

$$apo\ LOV + flavin \rightleftharpoons holo\ LOV \qquad \textbf{(Equation 2.2)}$$

A dynamic mass balance model for the above reaction can be constructed as described in **Equations 2.3-2.5**.

$$\frac{d[apo\ LOV]}{dt} = -k_1[apo\ LOV][flavin]^n + k_2[holo\ LOV] \qquad \textbf{(Equation 2.3)}$$

$$\frac{d[flavin]}{dt} = -k_1[apo\ LOV][flavin]^n + k_2[holo\ LOV] \qquad \textbf{(Equation 2.4)}$$

$$\frac{d[holo\ LOV]}{dt} = k_1[apo\ LOV][flavin]^n - k_2[holo\ LOV] \qquad \textbf{(Equation 2.5)}$$

Here, k_1 and k_2 represent on and off rates for complex formation and n is a Hill-like parameter that we initially included to account for possible cooperative effects. As fits derived with n as a free parameter essentially resulted in Hill coefficients ~1, we assumed noncooperative

binding for our final fits. The equilibrium dissociation constant can be calculated with **Equation 2.6**.

$$K_d = \frac{k_2}{k_1} \quad \textbf{(Equation 2.6)}$$

Flavin concentrations were varied between 1–20 μM and **Equations 2.3-2.5** were solved using the ode23s in Matlab (version R2020a) to calculate equilibrium concentrations and determine the fraction of total, *that is*, holo + apo LOV protein, that converts into the fluorescent holo form in the presence of flavin. We fitted these results to experimental data from the fluorescence titration experiments and estimated K_d using nonlinear least squares regression, implemented using lsqcurvefit (Matlab R2020a). We treated the starting concentration of the protein as an unknown, similar to prior studies on apo flavodoxins.[36,37] Concentrations of the apo protein derived in this way were found to be within 50–70% of the actual starting concentration, measured independently using the Bradford assay, indicating that some fraction of the apo protein could not be reconstituted into a fluorescent form.

2.3.6 Fluorescent riboflavin biosensor

To estimate riboflavin concentration in complex mixtures, calibration curves were first constructed by titrating known concentrations of riboflavin in apo LOV and measuring fluorescence using a Cary Varian Eclipse fluorometer. Excitation and emission wavelengths were set to 450 and 495 nm, slit widths were set at 5 nm and the photomultiplier tube gain was kept at medium (600 V). We fixed the concentration of apo proteins at 10 μM, which resulted in linear calibration curves for riboflavin titration from ~0.1–1 μM (concentration range commonly found in vitamin pills, flavin-fortified food and beverages) while avoiding aggregation that we observed at higher concentrations (**Figure 2.S3**). 100X minimal essential medium (MEM) vitamin solution, a mixture of eight water soluble vitamins, was purchased

from Thermo Fisher and used directly in fluorometric assays after dilution. Multivitamin supplements were purchased from a commercial source (Nature Made®), ground to a fine powder (~1 g per pill), mixed in PBS, mildly sonicated for 30–40 min, and clarified using centrifugation at 5000 x g for 15 min, before using the supernatant in the riboflavin assay. In each case, fluorescence recordings were acquired as described above and compared with the standard calibration curve to determine riboflavin content, which was finally validated by comparing with the manufacturer specified concentrations.

2.3.7 Data analysis

Fluorescence titration measurements were conducted using $N \geq 4$ independent replicates from 1 (iLOV + FMN, Ec + FAD titrations) or 2 distinct batches of purified proteins (iLOV + FAD, riboflavin; Ec + FMN, riboflavin titrations). Quality of model fits was judged by inspecting residuals (Figure 2.S3). All K_d values are reported as mean ± standard error. Pairwise comparisons of K_d values were performed using the 2-tailed Student's *t*-test with significance level set at 0.05. A 95% confidence interval for the measured riboflavin content in multivitamin tablets was computed using the Student's t-distribution with 2 (for iLOV) or 3° of freedom (for EcFbFP).

2.4. Results and discussion

2.4.1 Preparation of flavin-free LOV reporters

In our earlier work,[22] we observed that most LOV reporters lost fluorescence following thermal treatment at 70°C for 1 hour, but fluorescence recovered rapidly upon cooling. Based on this observation, we speculated that it should be possible to develop a method for reversibly separating the flavin fluorophore from LOV reporters, while maintaining the ability of the apo protein to recover fluorescence upon supplying flavin. We first attempted to separate bound

flavin from the plant-derived LOV reporter (iLOV) using several deflavination techniques that are well-established in flavin enzymology,[34] including protein precipitation with trichloroacetic acid, ammonium sulfate, treatment with halide salts (up to 2 M KBr), and chaotropic agents (up to 8 M urea). However, these methods either failed to remove flavin or produced apo proteins that aggregated and could not be converted into the fluorescent holo protein by supplying FMN. Attachment of flavoproteins to a chromatographic support has previously been shown to improve chances of generating stable preparations of apo proteins, presumably by preventing the apo proteins from aggregating in solution.[34,38] To implement this technique, we introduced a hexa-histidine motif at the N terminus of iLOV, which enabled binding to a chromatographic nickel column. Using column-immobilized iLOV, we were able to successfully remove flavin after overnight treatment with a strong chaotropic denaturant, guanidine hydrochloride. Following extensive on-column washing, imidazole based elution, and dialysis (see **Table 2.S1** for yields and percent recovery), we obtained purified apo preparations (**Figure 2.S1**) that were nonfluorescent (**Figure 2.1a**) and did not form visible precipitates even after prolonged storage (> 7 days) at 4°C. Using analytical size exclusion chromatography (**Figures 2.1b and 2.S2**), we verified that the purified apo preparation retains the parent holo-protein's monomeric state (although higher concentrations tended to aggregate, **Figure 2.S3**) but lacked the characteristic excitation and emission spectra of LOV reporters (**Figure 2.1a**), indicating removal of protein-bound flavin. Far-UV circular dichroism spectroscopy revealed a shallower trough at ~220 nm for the apo protein compared to holo iLOV, suggesting a partial loss of α helical content in the apo protein (**Figure 2.1c and Table 2.S2**). Importantly, addition of FMN to apo iLOV immediately (within seconds) restored the characteristic fluorescence spectrum of LOV reporters (**Figure 2.1d**), indicating

successful reconstitution. We were also able to extend this approach to separate flavin from the bacterial LOV reporter, EcFbFP (**Figure 2.1e**), which retained its dimeric state (**Figure 2.1f**) but, as before, displayed a partial loss of α helical secondary structure in the far-UV CD spectrum (**Figure 2.1g and Table 2.S2**). As in the case of iLOV, apo EcFbFP readily reverts into the fluorescent holo form upon supplying FMN (**Figure 2.1h**).

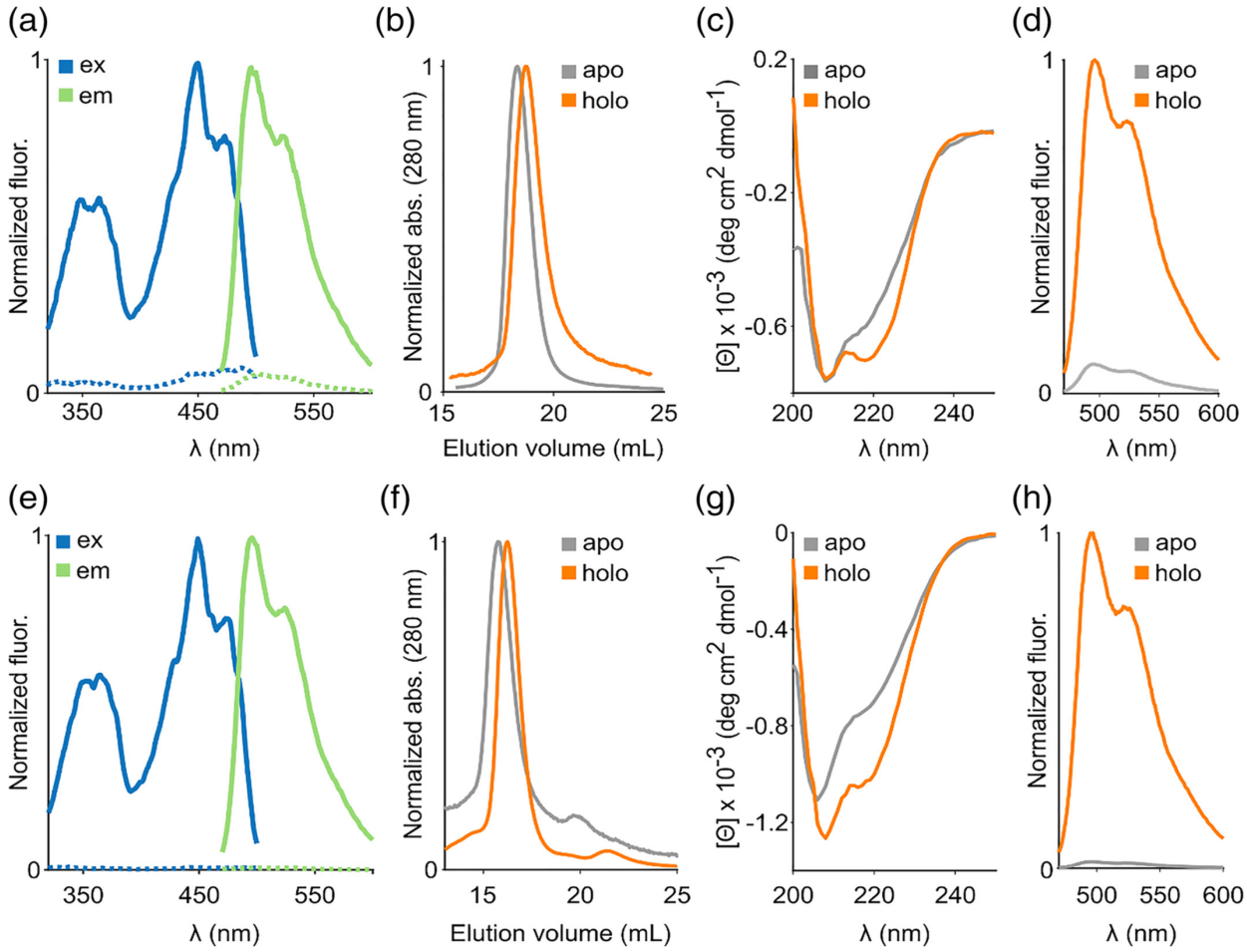

Figure 2.1. Flavin removal in fluorescent LOV reporters. **a)** Excitation and emission spectra of holo (solid lines) and deflavinated (dashed lines) iLOV. **b)** Elution profiles of apo and holo iLOV from gel filtration chromatography. **c)** Far UV CD spectra of apo and holo iLOV. **d)** Recovery of fluorescence emission upon adding FMN to apo iLOV. **e)** Excitation/emission spectra, **f)** elution profiles, **g)** far UV CD spectra, and **h)** fluorescence recovery in the bacterial fluorescent protein EcFbFP.

2.4.2 Flavin binding in LOV reporters

Free FMN demonstrates negligible fluorescence at 495 nm, which corresponds to peak emission in flavin-bound fluorescent LOV proteins. Thus, the distinct fluorescence properties of LOV-bound FMN compared to free FMN, make it possible to use fluorescence spectroscopy to monitor FMN binding in LOV proteins. Similar methods, based on changes in fluorescence intensity between free and protein-bound FMN,[39] have been widely used to study flavin binding in enzymes, including flavodoxin,[37,40] L-amino acid oxidase,[41] and cytochromes,[42,43] albeit in these cases the protein environment quenches FMN fluorescence, unlike for LOV reporters. To measure the equilibrium dissociation constant (K_d), we titrated aliquots of FMN into solutions containing apo iLOV. After each titration, we allowed the solution to equilibrate in the dark before recording emission at 495 nm. We fitted the fluorescence measurements to a simple bimolecular binding model, assuming noncooperative binding behavior (see **Figure 2.S4(a-c)** for fit residuals) and derived a K_d of 0.23 ± 0.02 μM between iLOV and FMN (**Figure 2.2a**). Using a similar approach, we determined a K_d of 0.18 ± 0.03 μM for EcFbFP (**Figure 2.2a**), suggesting that both monomeric (iLOV) and dimeric (EcFbFP) LOV reporters bind FMN with similar affinities (*p-value = 0.2, N = 4 for iLOV, 5 for Ec*) (**Figure 2.2d**). Notably, the calculated K_d values reveal that these LOV reporters bind their fluorophore with a stronger affinity than most biliverdin based fluorescent reporters ($K_d \sim 490$ nM–4.84 μM[44]), but similar to recently evolved fluorogenic reporters such as YFAST, which binds its synthetic fluorophore with a $K_d \sim 0.13$ μM.[45]

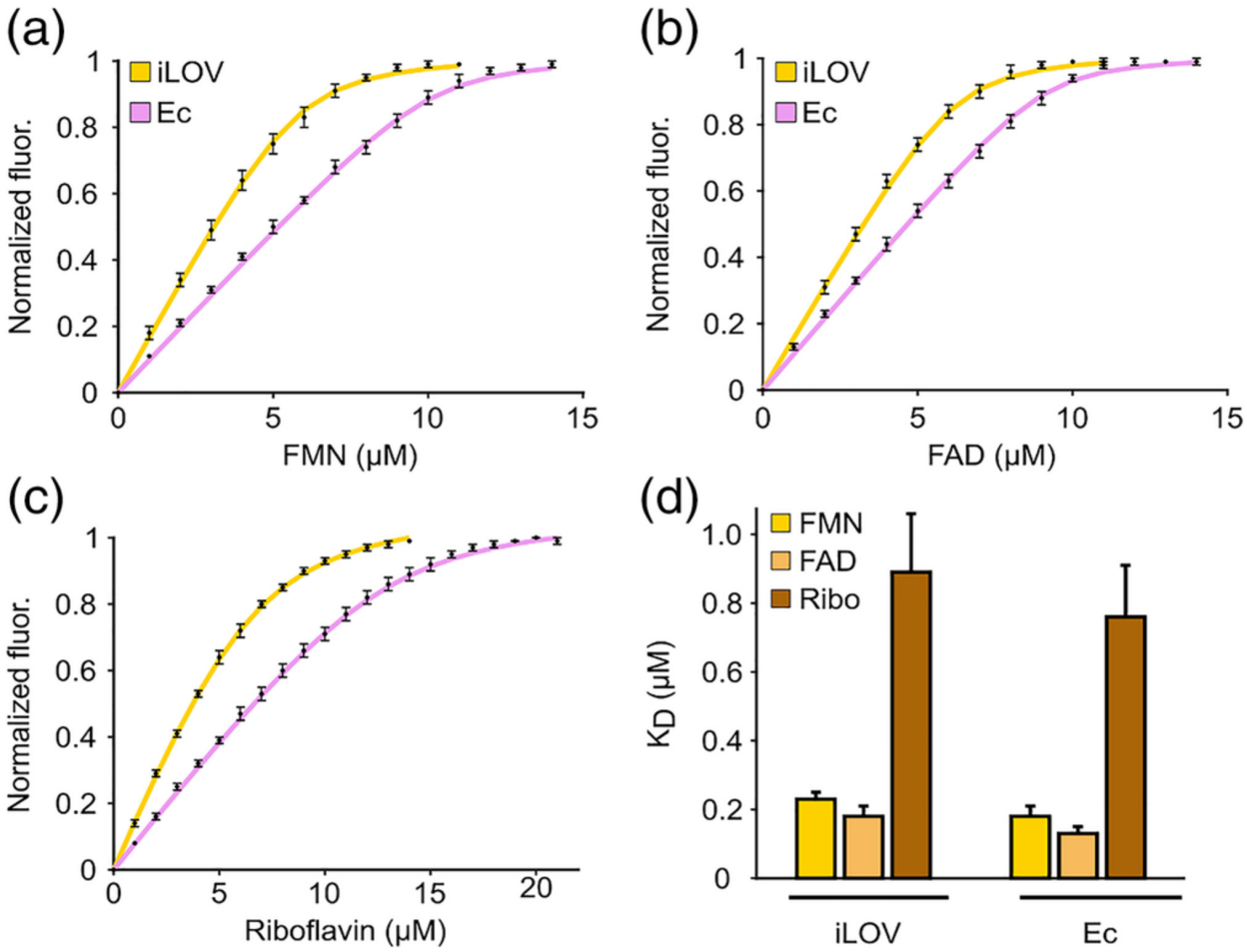

Figure 2.2. Flavin-binding in LOV reporters. Fluorescence titration of, **a)** FMN, **b)** FAD, and **c)** riboflavin in purified preparations of apo iLOV and apo EcFbFP. **d)** Equilibrium dissociation constant (K_D) for binding of LOV reporters to various flavins. All measurements represent means of at least four independent replicates from 1-2 purification batches and error bars correspond to the standard error of mean.

Some flavoproteins are known to exhibit nonspecific binding to physiological flavins other than their specific prosthetic group.[46–49] To this end, we tested iLOV for its ability to bind riboflavin, the precursor molecule to FMN, as well as FAD, which is a bulkier, adenylated form of FMN. Interestingly, we found that iLOV can be reconstituted into a fluorescent form using either FAD or riboflavin and the resulting spectrum is indistinguishable from the holo iLOV protein. As free FAD is nearly nonfluorescent due to intramolecular stacking between the adenine and isoalloxazine rings,[50] the recovery of FAD fluorescence upon LOV binding

suggests that the adenine and isoalloxazine rings are unstacked as FAD is inserted in the binding pocket of iLOV. Using fluorescence titration we determined a K_d of 0.18 ± 0.03 μM for FAD (**Figure 2.2b**) and 0.89 ± 0.17 μM for riboflavin (**Figure 2.2c**), suggesting that iLOV can bind FAD with a similar affinity as FMN, but its interaction with riboflavin is almost four-fold weaker (*p = .01, N = 4 for FMN, 5 for riboflavin*) (**Figure 2.2d**). We observed a similar trend in the binding properties of bacterial EcFbFP, which was found to bind FAD with K_d = 0.13 ± 0.02 μM and riboflavin with K_d = 0.76 ± 0.15 μM (**Figure 2.2b-d**). The lower affinity for riboflavin has been previously observed in other flavoproteins (e.g., dodecins)[46] and may be caused by the lack of a phosphate group that anchors FMN (and presumably, FAD) in the flavin binding pocket. Consistent with this notion, increasing ionic strength was found to weaken the affinity for FMN (**Figure 2.S5**), which may point to electrostatic screening[51] of interactions involving the phosphate group of FMN and positively charged amino acids (e.g., arginine) in the flavin binding pocket.

2.4.3 Fluorescence "turn-on" biosensors for riboflavin

Riboflavin, (vitamin B2) is an essential nutrient as it serves as the precursor for synthesizing FMN and FAD within cells, which then serve as prosthetic groups for several flavoproteins involved in metabolism. Consequently, riboflavin deficiency in humans is associated with conditions such as fatigue, cheilosis, growth retardation, and night blindness.[52] Animals lack riboflavin biosynthesis pathways and thus typically obtain riboflavin through external means such as dairy products, fortified cereals, bread, energy drinks, and multivitamin pills. For these reasons, the development of easy-to-use assays to estimate riboflavin content in food, infant formulations, dairy products, and bodily fluids (e.g., urine) is of considerable importance to nutritional and pharmaceutical sectors. Existing approaches for

detecting riboflavin make use of liquid chromatography,[37,53] electrochemical techniques (voltammetry,[54,55] immunoassays,[56] surface plasmon resonance[57]), and culture based tests involving riboflavin-dependent microbes grown using the test sample as the only source of riboflavin.[58] These methods, while widely used, suffer from limitations such as slow response, interference from other components in complex mixtures, and limited scalability to a medium or high throughput format. Fluorescent sensors can overcome many of these challenges, which has prompted the development of riboflavin sensors based on Förster Resonance Energy Transfer (FRET) between riboflavin and doped carbon dots[59] or graphene[60] as well as quenching of fluorescence using PEG-dispersed graphene.[61] However, to our knowledge, a fluorescence "turn-on" biosensor for riboflavin has not been reported. As apo LOV reporters are capable of binding riboflavin with low micromolar affinity (**Figure 2.2d**) and exhibit both visually detectable (**Figure 2.3a**) and quantifiable gain of fluorescence, we reasoned that purified apo iLOV or EcFbFP should enable simple, rapid, and quantitative determination of riboflavin content in complex mixtures without a need for extensive sample processing steps such as liquid chromatography. To test this idea, we first verified that LOV fluorescence is specific to flavins and cannot be reconstituted using other biological cofactors such as NAD, deoxynucleotides, and amino acids (**Figure 2.S6**). Next, we used apo iLOV to quantify riboflavin content in a mixture consisting of eight water-soluble vitamins (riboflavin, choline chloride, calcium pantothenate, folic acid, nicotinamide, inositol, thiamine hydrochloride, and pyridoxal hydrochloride) that is commonly used as a nutrient supplement in animal cell culture. We treated various dilutions of the 8-vitamin mixture with a fixed concentration of apo iLOV and estimated the riboflavin content by measuring fluorescence and comparing against a calibration of known concentration standards. We found that apo iLOV could be used to

estimate riboflavin content in the vitamin mixture over a concentration span of 0.04–0.4 μg/ml (**Figure 2.3b**) with reasonable accuracy and precision. Building on these results, we proceeded to use apo iLOV to assay riboflavin content in a more complex milieu involving a commercial multivitamin supplement comprising all water-soluble B-vitamins, vitamins C, E, trace metals, several organic, and inorganic salts. We extracted the pill contents using mild ultrasonication and dissolved it in PBS to achieve a final riboflavin concentration of 41.8 μg/ml (calculated based on manufacture specification). Following treatment with apo iLOV and comparison with calibration standards, we obtained a riboflavin concentration of 47.9 μg/ml (95% CI: 38.7–57.1 ug/ml). Similar results were obtained using apo EcFbFP as a sensor, which determined the riboflavin concentration as 35.5 μg/ml (95% CI: 32.1–38.9 ug/ml) for a multivitamin supplement preparation comprising 32.2 μg/ml riboflavin. Taken together, these results establish the general utility of apo LOV reporters for simple, rapid, and quantitative riboflavin testing in real-world samples.

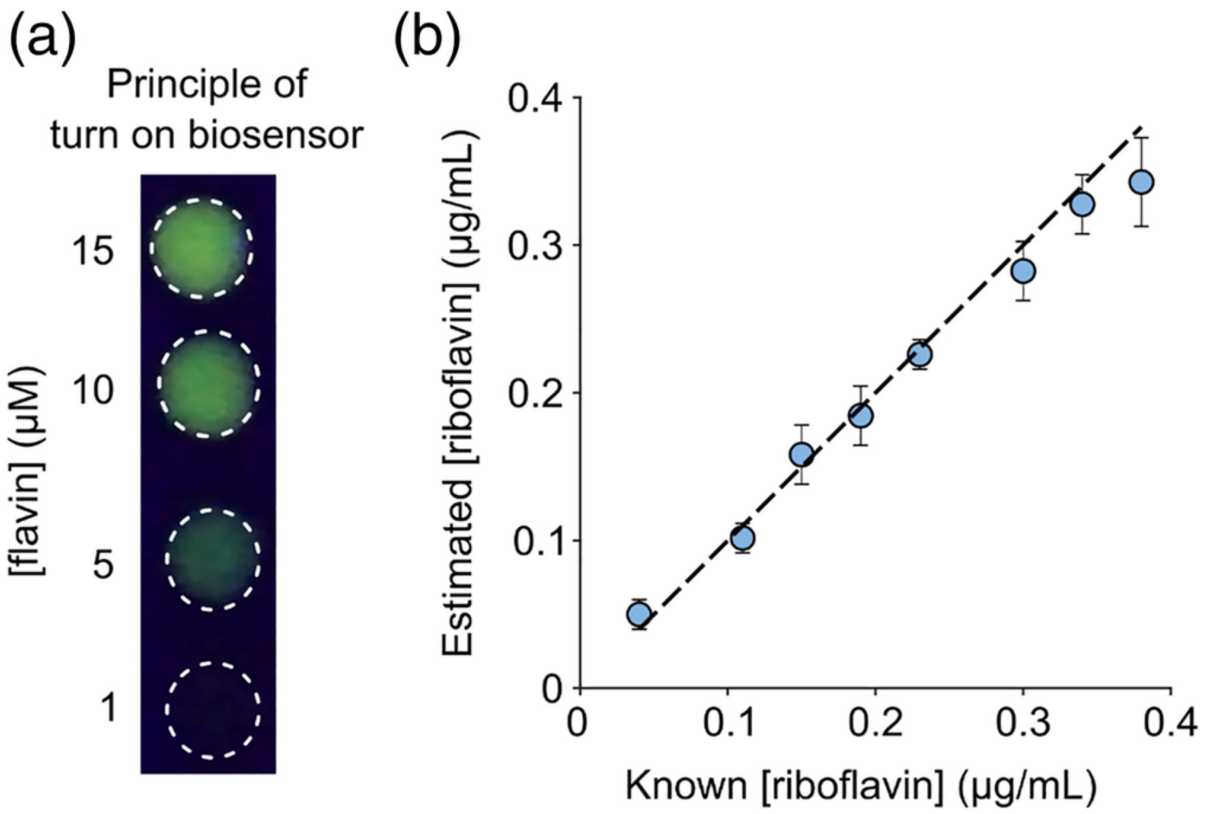

Figure 2.3. Turn-on fluorescent biosensor for riboflavin detection. A, Apo iLOV exhibits a visually detectable gain of fluorescence upon titrating flavin. Images were acquired by spotting ~ 10 µM apo iLOV on a clear UV-transparent plastic surface, exciting with a 320 nm light source in trans-illumination mode and recording with a digital camera. B, Agreement between manufacture-specified and estimated riboflavin concentrations (determined using apo iLOV assay) present in various dilutions of an 8-vitamin cell culture supplement.

2.5 Conclusions

In this work, we develop an approach for the reversible deflavination of fluorescent LOV proteins and apply it to establish the first (to our knowledge) measurements of equilibrium binding between engineered LOV reporters and the three physiological forms of flavins: riboflavin, FMN, and FAD. Although we have developed and validated this method in the context of two commonly used LOV reporters, it should be possible, in principle, to adapt this technique for binding studies in other fluorescent LOV proteins, including brighter proteins such as CreiLOV,[20] photostable variants such as phiLOV,[62] and recently discovered

thermostable LOV proteins such as CagFbFP.[21] Our work has important implications for engineering and applying LOV proteins as self-contained (i.e., no external agent needed) reporter genes in anaerobic microbes. For example, the relatively strong flavin-binding affinities that we measured in this work suggest that over-expression of LOV reporters may cause intracellular flavin (typically present in cells at low micromolar levels[63,64]) to drop, thereby metabolically burdening cells. To this end, we have observed varying levels of growth retardation[22] associated with the expression of different LOV reporters in *E. coli*. Future studies may reveal whether and to what extent growth retardation correlates with flavin affinity of various LOV reporters and if certain organisms can offset this resource constraint by activating feedback regulated flavin biosynthesis[65] in response to LOV expression. From the standpoint of reporter protein engineering, our binding studies suggest that a potential avenue for increasing cellular fluorescence obtained with LOV reporters could be to engineer variants with even tighter K_d (ideally, low nM), which could increase their flavin occupancy by allowing LOV reporters to effectively compete with the native cellular flavoproteome that comprises several tight binding flavoproteins. Notably, equilibrium dissociation constants in the sub-nM range have proved beneficial for developing reporters based on efficient occupancy with fluorescent cofactors, such as the bilirubin binding fluorescent protein, UnaG.[13] As before, potential effects on cell physiology resulting from increased flavin affinity will need to be carefully characterized and mitigated.

In addition to measuring the equilibrium dissociation constants, we also demonstrated that apo iLOV (and apo EcFbFP) could be used to optically assay for riboflavin (i.e., vitamin B2) based on gain of fluorescence. Apo preparations of some flavoproteins such as the riboflavin-binding protein (RfBP) from egg white and apo flavodoxin have been previously

used to quantify vitamin B2 content in cell lysates, dairy products, and beverages – however, these methods either relied on liquid chromatography[47] or quenching of fluorescence for riboflavin detection.[66–68] To our knowledge, the methodology introduced here represents the only known fluorescence turn on biosensor for riboflavin. One potential limitation of this technique is that it cannot distinguish riboflavin from FMN or FAD, which are enzymatically synthesized from riboflavin within cells. While this can be problematic for monitoring specific flavins intracellularly or in lysed cells, we do not believe this will become a major limitation for using the sensors in extracellular or noncellular contexts such as food materials and vitamin supplements where FMN and FAD are typically present in negligible amounts. To this end, we envision that purified preparations of apo LOV proteins could be used to evaluate riboflavin content in common dietary and pharmaceutical sources of vitamin B2 with minimal sample processing, while avoiding time and equipment-intensive methods such as chromatography, mass spectrometry, and microbial culture based assays. Another potential application of our sensor would be for monitoring bioproduction of vitamin B2 in commercial flavin-overproducing microbes such as *Lactobacillus plantarum*, which secrete riboflavin at concentrations (> 1 μg/ml) that should be readily detectable with our sensor. As fluorimetry is readily scalable to medium or high throughput screening applications, one important benefit of these sensors could be in optimizing natural or bioengineered strains for riboflavin overproduction. Although further studies will be required to rigorously characterize apo LOV-based sensors to determine limit of detection, limit of quantification, reusability, response kinetics, and long-term stability as well as validate in a broader set of real-world samples, our results provide a proof-of-concept for a simple, rapid, and scalable fluorescence turn on assay for riboflavin.

2.6 References

1. Terai, T. & Nagano, T. Fluorescent probes for bioimaging applications. *Curr. Opin. Chem. Biol.* **12**, 515–521 (2008).
2. Dean, K. M. & Palmer, A. E. Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nat. Chem. Biol.* **10**, 512–523 (2014).
3. Thorn, K. Genetically encoded fluorescent tags. *Mol. Biol. Cell* **28**, 848–857 (2017).
4. Heim, R., Prasher, D. C. & Tsien, R. Y. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 12501–12504 (1994).
5. Ozbakir, H. F., Anderson, N. T., Fan, K.-C. & Mukherjee, A. Beyond the Green Fluorescent Protein: Biomolecular Reporters for Anaerobic and Deep-Tissue Imaging. *Bioconjug. Chem.* **31**, 293–302 (2020).
6. Guglielmetti, S., Santala, V., Mangayil, R., Ciranna, A. & Karp, M. T. O2-requiring molecular reporters of gene expression for anaerobic microorganisms. *Biosens. Bioelectron.* **123**, 1–6 (2019).
7. Scott, K. P., Mercer, D. K., Glover, L. A. & Flint, H. J. The green fluorescent protein as a visible marker for lactic acid bacteria in complex ecosystems. *FEMS Microbiol. Ecol.* **26**, 219–230 (1998).
8. Coralli, C., Cemazar, M., Kanthou, C., Tozer, G. M. & Dachs, G. U. Limitations of the reporter green fluorescent protein under simulated tumor conditions. *Cancer Res.* **61**, 4784–4790 (2001).
9. Zhang, C., Xing, X.-H. & Lou, K. Rapid detection of a gfp-marked Enterobacter aerogenes under anaerobic conditions by aerobic fluorescence recovery. *FEMS*

Microbiol. Lett. **249**, 211–218 (2005).

10. Buckley, A. M., Petersen, J., Roe, A. J., Douce, G. R. & Christie, J. M. LOV-based reporters for fluorescence imaging. *Curr. Opin. Chem. Biol.* **27**, 39–45 (2015).
11. Chia, H. E., Zuo, T., Koropatkin, N. M., Marsh, E. N. G. & Biteen, J. S. Imaging living obligate anaerobic bacteria with bilin-binding fluorescent proteins. *Curr. Res. Microb. Sci.* **1**, 1–6 (2020).
12. Jullien, L. & Gautier, A. Fluorogen-based reporters for fluorescence imaging: a review. *Methods Appl. Fluoresc.* **3**, 12 (2015).
13. Kumagai, A. *et al.* A Bilirubin-Inducible Fluorescent Protein from Eel Muscle. *Cell* **153**, 1602–1611 (2013).
14. Mukherjee, A. & Schroeder, C. M. Flavin-based fluorescent proteins: emerging paradigms in biological imaging. *Curr. Opin. Biotechnol.* **31**, 16–23 (2015).
15. Shcherbakova, D. M., Shemetov, A. A., Kaberniuk, A. A. & Verkhusha, V. V. Natural Photoreceptors as a Source of Fluorescent Proteins, Biosensors, and Optogenetic Tools. *Annu. Rev. Biochem.* **84**, 519–550 (2015).
16. Abbas, C. A. & Sibirny, A. A. Genetic Control of Biosynthesis and Transport of Riboflavin and Flavin Nucleotides and Construction of Robust Biotechnological Producers. *Microbiol. Mol. Biol. Rev.* **75**, 321–360 (2011).
17. Chapman, S. *et al.* The photoreversible fluorescent protein iLOV outperforms GFP as a reporter of plant virus infection. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 20038–20043 (2008).
18. Drepper, T. *et al.* Reporter proteins for in vivo fluorescence without oxygen. *Nat. Biotechnol.* **25**, 443–445 (2007).

19. Mukherjee, A., Weyant, K. B., Walker, J. & Schroeder, C. M. Directed evolution of bright mutants of an oxygen-independent flavin-binding fluorescent protein from Pseudomonas putida. *J. Biol. Eng.* **6**, 20 (2012).

20. Mukherjee, A. *et al.* Engineering and Characterization of New LOV-Based Fluorescent Proteins from Chlamydomonas reinhardtii and Vaucheria frigida. *ACS Synth. Biol.* **4**, 371–377 (2015).

21. Nazarenko, V. V *et al.* A thermostable flavin-based fluorescent protein from Chloroflexus aggregans: a framework for ultra-high resolution structural studies. *Photochem. Photobiol. Sci.* (2019) doi:10.1039/C9PP00067D.

22. Mukherjee, A., Walker, J., Weyant, K. B. & Schroeder, C. M. Characterization of Flavin-Based Fluorescent Proteins: An Emerging Class of Fluorescent Reporters. *PLoS One* **8**, e64753 (2013).

23. Zou, W., Le, K. & Zastrow, M. L. Live-Cell Copper-Induced Fluorescence Quenching of the Flavin-Binding Fluorescent Protein CreiLOV. *ChemBioChem* **21**, 1356–1363 (2020).

24. Ravikumar, Y. *et al.* FMN-Based Fluorescent Proteins as Heavy Metal Sensors Against Mercury Ions. *J. Microbiol. Biotechnol.* **26**, 530–539 (2016).

25. Ravikumar, Y., Nadarajan, S. P., Lee, C.-S., Rhee, J.-K. & Yun, H. A New-Generation Fluorescent-Based Metal Sensor - iLOV Protein. *J. Microbiol. Biotechnol.* **25**, 503–510 (2015).

26. Buckley, A. M. *et al.* Lighting Up Clostridium Difficile: Reporting Gene Expression Using Fluorescent Lov Domains. *Sci. Rep.* **6**, 23463 (2016).

27. Lobo, L. A., Smith, C. J. & Rocha, E. R. Flavin mononucleotide (FMN)-based

fluorescent protein (FbFP) as reporter for gene expression in the anaerobe Bacteroides fragilis. *FEMS Microbiol. Lett.* **317**, 67–74 (2011).

28. Wang, S. E., Brooks, A. E. S., Cann, B. & Simoes-Barbosa, A. The fluorescent protein iLOV outperforms eGFP as a reporter gene in the microaerophilic protozoan Trichomonas vaginalis. *Mol. Biochem. Parasitol.* **216**, 1–4 (2017).

29. Choi, C. H., DeGuzman, J. V, Lamont, R. J. & Yilmaz, Ö. Genetic Transformation of an Obligate Anaerobe, P. gingivalis for FMN-Green Fluorescent Protein Expression in Studying Host-Microbe Interaction. *PLoS One* **6**, e18499 (2011).

30. Seo, S.-O., Lu, T., Jin, Y.-S. & Blaschek, H. P. Development of an oxygen-independent flavin mononucleotide-based fluorescent reporter system in Clostridium beijerinckii and its potential applications. *J. Biotechnol.* **265**, 119–126 (2018).

31. Teng, L., Wang, K., Xu, J. & Xu, C. Flavin mononucleotide (FMN)-based fluorescent protein (FbFP) as reporter for promoter screening in Clostridium cellulolyticum. *J. Microbiol. Methods* **119**, 37–43 (2015).

32. Tielker, D., Eichhof, I., Jaeger, K. E. & Ernst, J. F. Flavin Mononucleotide-Based Fluorescent Protein as an Oxygen-Independent Reporter in Candida albicans and Saccharomyces cerevisiae. *Eukaryot. Cell* **8**, 913–915 (2009).

33. Chapman, S. K. & Reid, G. A. *Flavoprotein protocols*. vol. 131 (Springer Science & Business Media, 2008).

34. Hefti, M. H., Vervoort, J. & van Berkel, W. J. H. Deflavination and reconstitution of flavoproteins. *Eur. J. Biochem.* **270**, 4227–4242 (2003).

35. Micsonai, A. *et al.* Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc. Natl. Acad. Sci.* **112**, E3095–E3103 (2015).

36. Bollen, Y. J. M., Westphal, A. H., Lindhoud, S., van Berkel, W. J. H. & van Mierlo, C. P. M. Distant residues mediate picomolar binding affinity of a protein cofactor. *Nat. Commun.* **3**, 1010 (2012).

37. Lostao, A. *et al.* Dissecting the Energetics of the Apoflavodoxin-FMN Complex*. *J. Biol. Chem.* **275**, 9518–9526 (2000).

38. Hefti, M. H., Milder, F. J., Boeren, S., Vervoort, J. & van Berkel, W. J. H. A His-tag based immobilization method for the preparation and reconstitution of apoflavoproteins. *Biochim. Biophys. Acta - Gen. Subj.* **1619**, 139–143 (2003).

39. Visser, A. J. W. G., Ghisla, S., Massey, V., Müller, F. & Veeger, C. Fluorescence Properties of Reduced Flavins and Flavoproteins. *Eur. J. Biochem.* **101**, 13–21 (1979).

40. Mayhew, S. G. Studies on flavin binding in flavodoxins. *Biochim. Biophys. Acta - Enzymol.* **235**, 289–302 (1971).

41. Casalin, P., Pollegioni, L., Curti, B. & Simonetta, M. P. A study on apoenzyme from Rhodotorula gracilis D-amino acid oxidase. *Eur. J. Biochem.* **197**, 513–517 (1991).

42. Shumyantseva, V. V *et al.* Fluorescent assay for riboflavin binding to cytochrome P450 2B4. *J. Inorg. Biochem.* **98**, 365–370 (2004).

43. Strittmatter, P. The Nature of the Flavin Binding in Microsomal Cytochrome b5 Reductase. *J. Biol. Chem.* **236**, 2329–2335 (1961).

44. Yu, D. *et al.* A naturally monomeric infrared fluorescent protein for protein labeling in vivo. *Nat. Methods* **12**, 763–765 (2015).

45. Plamont, M.-A. *et al.* Small fluorescence-activating and absorption-shifting tag for tunable protein imaging in vivo. *Proc. Natl. Acad. Sci.* **113**, 497–502 (2016).

46. Bourdeaux, F. *et al.* Flavin Storage and Sequestration by Mycobacterium tuberculosis

Dodecin. *ACS Infect. Dis.* **4**, 1082–1092 (2018).

47. Dorn, M., Jurk, M., Wartenberg, A., Hahn, A. & Schmieder, P. LOV Takes a Pick: Thermodynamic and Structural Aspects of the Flavin-LOV-Interaction of the Blue-Light Sensitive Photoreceptor YtvA from Bacillus subtilis. *PLoS One* **8**, e81268 (2013).

48. Fitzpatrick, P. F., Ghisla, S. & Massey, V. 8-Azidoflavins as photoaffinity labels for flavoproteins. *J. Biol. Chem.* **260**, 8483–8491 (1985).

49. Mathes, T., Vogl, C., Stolz, J. & Hegemann, P. In Vivo Generation of Flavoproteins with Modified Cofactors. *J. Mol. Biol.* **385**, 1511–1518 (2009).

50. Stanley, R. J. & MacFarlane, A. W. Ultrafast Excited State Dynamics of Oxidized Flavins: Direct Observations of Quenching by Purines. *J. Phys. Chem. A* **104**, 6899–6906 (2000).

51. Papaneophytou, C. P., Grigoroudis, A. I., McInnes, C. & Kontopidis, G. Quantification of the Effects of Ionic Strength, Viscosity, and Hydrophobicity on Protein–Ligand Binding Affinity. *ACS Med. Chem. Lett.* **5**, 931–936 (2014).

52. Thakur, K., Tomar, S. K., Singh, A. K., Mandal, S. & Arora, S. Riboflavin and health: A review of recent human research. *Crit. Rev. Food Sci. Nutr.* **57**, 3650–3660 (2017).

53. Zempleni, J. Determination of Riboflavin and Flavocoenzymes in Human Blood Plasma by High-Performance Liquid Chromatography. *Ann. Nutr. Metab.* **39**, 224–226 (2008).

54. Hou, W. & Wang, E. Liquid chromatography with series dual-electrode electrochemical detection for riboflavin. *Analyst* **115**, 139–141 (1990).

55. Ly, S. Y., Yoo, H. S., Ahn, J. Y. & Nam, K. H. Pico molar assay of riboflavin in

human urine using voltammetry. *Food Chem.* **127**, 270–274 (2011).

56. Ravi, G. & Venkatesh, Y. P. Immunoassays for riboflavin and flavin mononucleotide using antibodies specific to d-ribitol and d-ribitol-5-phosphate. *J. Immunol. Methods* **445**, 59–66 (2017).

57. Caelen, I., Kalman, A. & Wahlström, L. Biosensor-Based Determination of Riboflavin in Milk Samples. *Anal. Chem.* **76**, 137–143 (2004).

58. Salvetti, S., Celandroni, F., Ghelardi, E., Baggiani, A. & Senesi, S. Rapid determination of vitamin B2 secretion by bacteria growing on solid media. *J. Appl. Microbiol.* **95**, 1255–1260 (2003).

59. Lin, L., Wang, Y., Xiao, Y. & Chen, X. Ratiometric fluorescence detection of riboflavin based on fluorescence resonance energy transfer from nitrogen and phosphorus co-doped carbon dots to riboflavin. *Anal. Bioanal. Chem.* **411**, 2803–2808 (2019).

60. Kundu, A., Nandi, S., Layek, R. K. & Nandi, A. K. Fluorescence Resonance Energy Transfer from Sulfonated Graphene to Riboflavin: A Simple Way to Detect Vitamin B2. *ACS Appl. Mater. Interfaces* **5**, 7392–7399 (2013).

61. Díez-Pascual, A. M., García-García, D., San Andrés, M. P. & Vera, S. Determination of riboflavin based on fluorescence quenching by graphene dispersions in polyethylene glycol. *RSC Adv.* **6**, 19686–19699 (2016).

62. Christie, J. M. *et al.* Structural tuning of the fluorescent protein iLOV for improved photostability. *J Biol Chem* **287**, 22295–22304 (2012).

63. Pedrolli, D. *et al.* The ribB FMN riboswitch from Escherichia coli operates at the transcriptional and translational level and regulates riboflavin biosynthesis. *FEBS J.*

282, 3230–3242 (2015).

64. Wilson, A. C. & Pardee, A. B. Regulation of flavin synthesis by Escherichia coli. *Microbiology* **28**, 283–303 (1962).

65. Ulrich, K., Vera, S., Astrid, W., Esther, K.-G. & Karl-Erich, J. Cofactor Trapping, a New Method To Produce Flavin Mononucleotide. *Appl. Environ. Microbiol.* **77**, 1097–1100 (2011).

66. Duyvis, M. G., Hilhorst, R., Laane, C., Evans, D. J. & Schmedding, D. J. M. Role of Riboflavin in Beer Flavor Instability: Determination of Levels of Riboflavin and Its Origin in Beer by Fluorometric Apoprotein Titration. *J. Agric. Food Chem.* **50**, 1548–1552 (2002).

67. Wassink, J. H. & Mayhew, S. G. Fluorescence titration with apoflavodoxin: A sensitive assay for riboflavin 5′-phosphate and flavin adenine dinucleotide in mixtures. *Anal. Biochem.* **68**, 609–616 (1975).

68. Zandomeneghi, M., Carbonaro, L. & Zandomeneghi, G. Biochemical Fluorometric Method for the Determination of Riboflavin in Milk. *J. Agric. Food Chem.* **55**, 5990–5994 (2007).

2.7. Supplemental information

Table 2.S1. Yield and fractional recovery from purification of apo & holo iLOV

Protein	Protein yield from Ni-NTA purification (*2* batches each using 125 mL culture volume)	% recovery after dialysis of holo or deflavinated elute from Ni-NTA purification
Holo iLOV	283 ± 8 μg	75 ± 16 %

Apo iLOV	184 ± 3 µg	67 ± 24 %

Table 2.S2. Secondary structure content from far-UV CD spectra

	Helical	Beta sheets
Apo iLOV	11.3 %	33.1 %
Holo iLOV	17.4 %	34.8 %
Apo EcFbFP	7.9 %	27.9 %
Holo EcFbFP	19.3 %	29 %

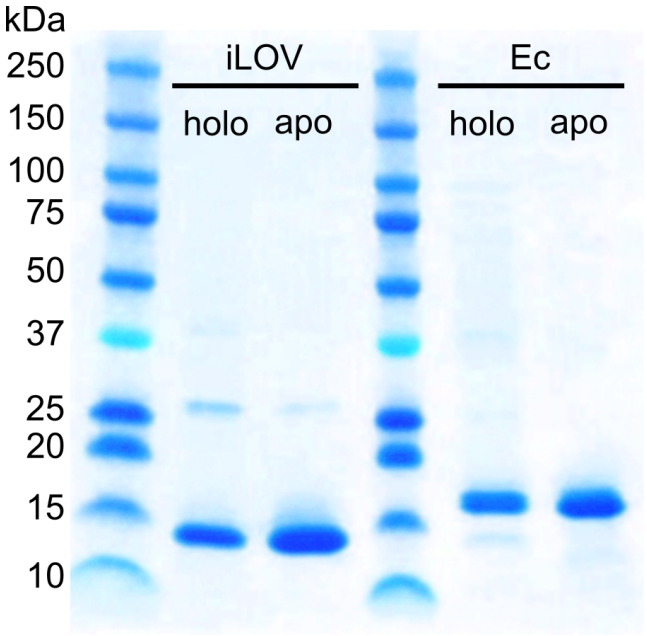

Figure 2.S1. SDS PAGE of purified apo and holo preparations of iLOV and EcFbFP. Molecular weights (kDa) of the ladder in kDa are included for reference.

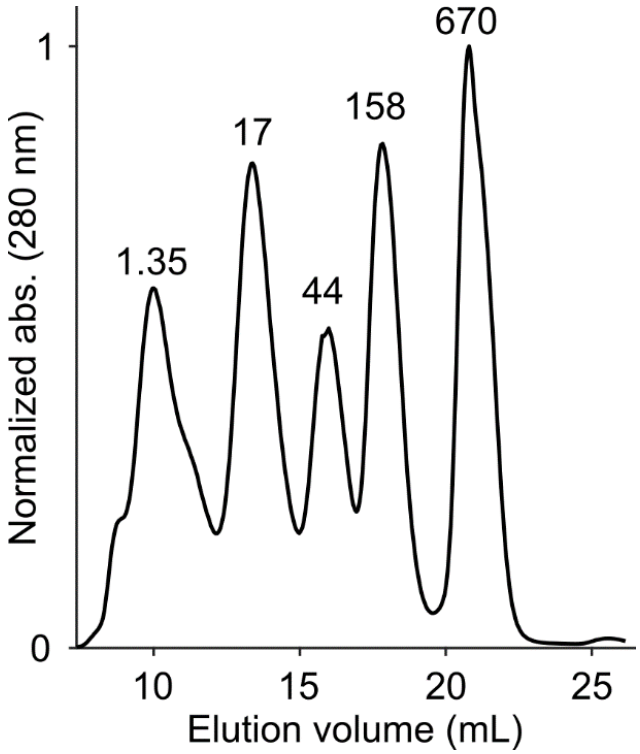

Figure 2.S2. Size exclusion chromatography of analytical standards for determination of molecular weight and oligomeric state of purified LOV proteins. Molecular weights of standards are indicated in kDa for reference.

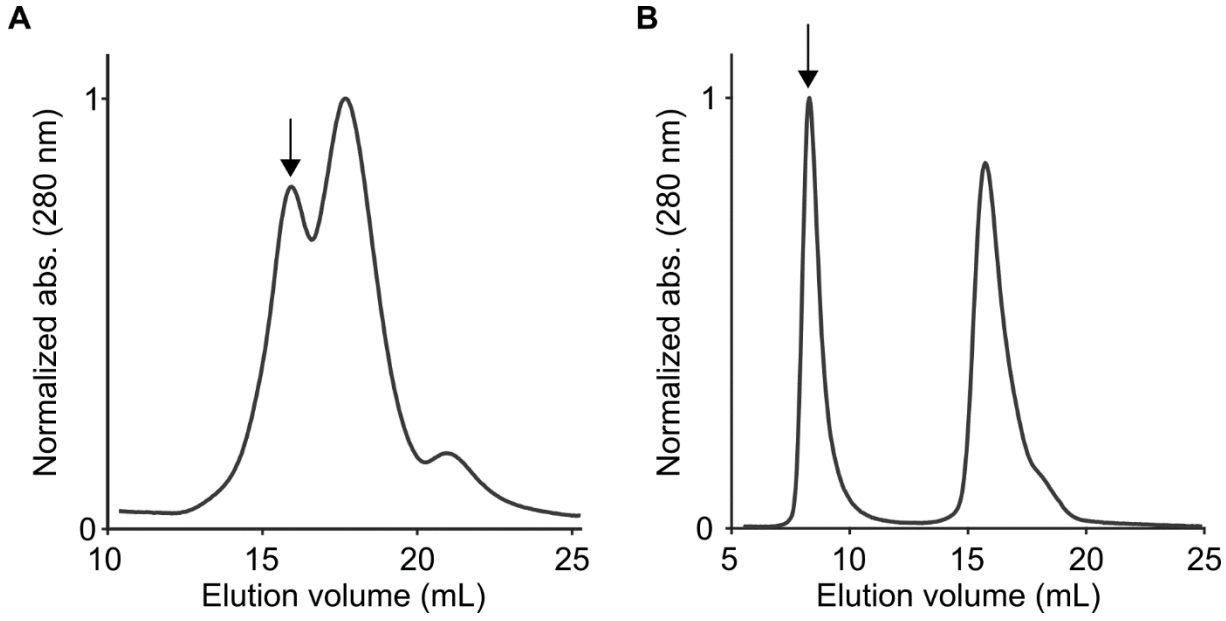

Figure 2.S3. Apo LOV proteins show tendency to aggregate at high concentrations. Size exclusion chromatograms of (A) apo iLOV and (B) apo EcFbFP at concentrations exceeding 50 μM. Arrows indicate peaks corresponding to protein fractions eluting at earlier volumes than expected for monomeric iLOV or dimeric EcFbFP, indicating the formation of higher order oligomers and/or aggregates.

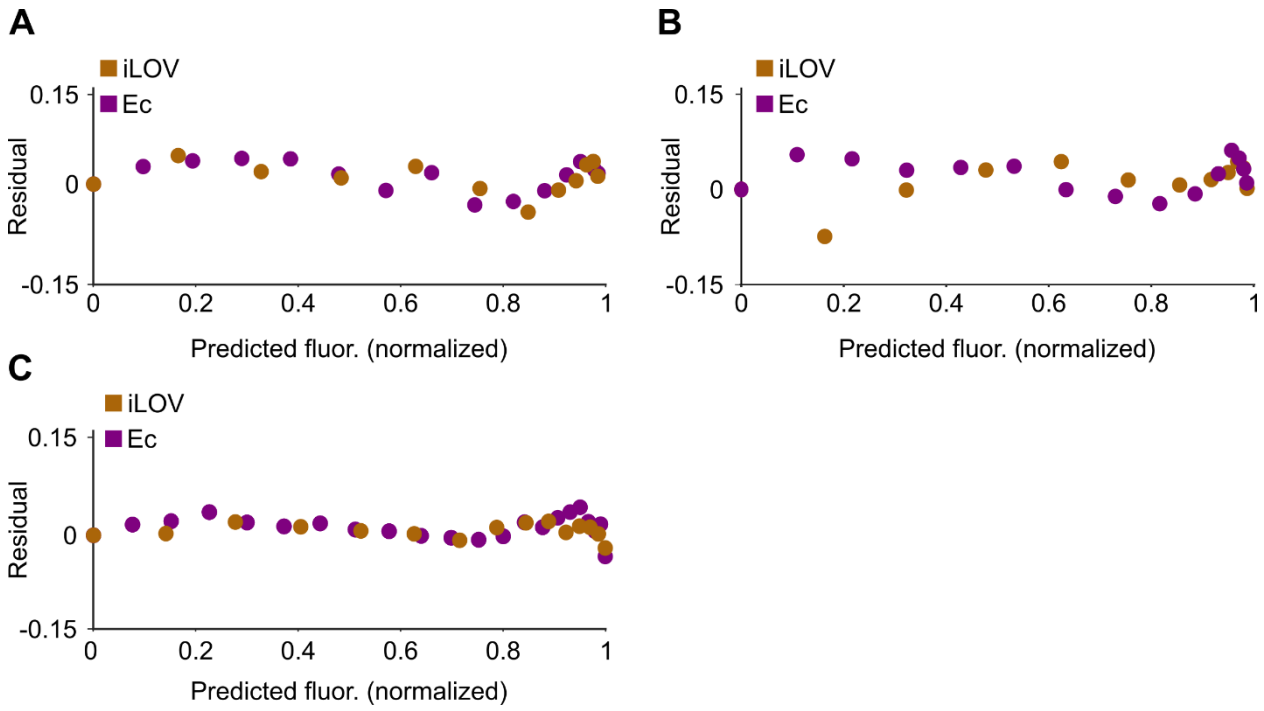

Figure 2.S4. Residuals from fitting experimental titration data (average of N = 4 replicates) to binding model for apo LOV reporters with (A) FMN (B) FAD and (C) riboflavin.

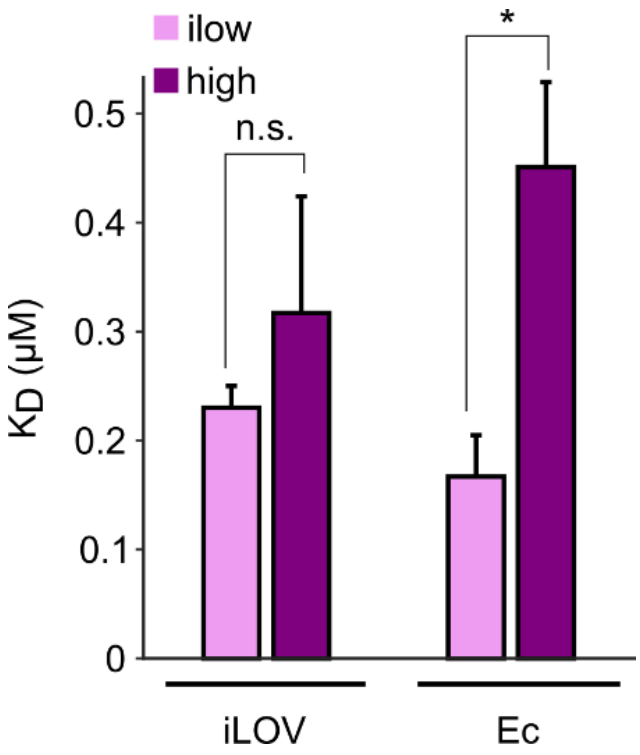

Figure 2.S5. Equilibrium dissociation constant for FMN binding in buffers of low (~ 150 mM) and high (~ 1 M) ionic strength (n.s. indicates non-significant and * indicates $p < 0.05$).

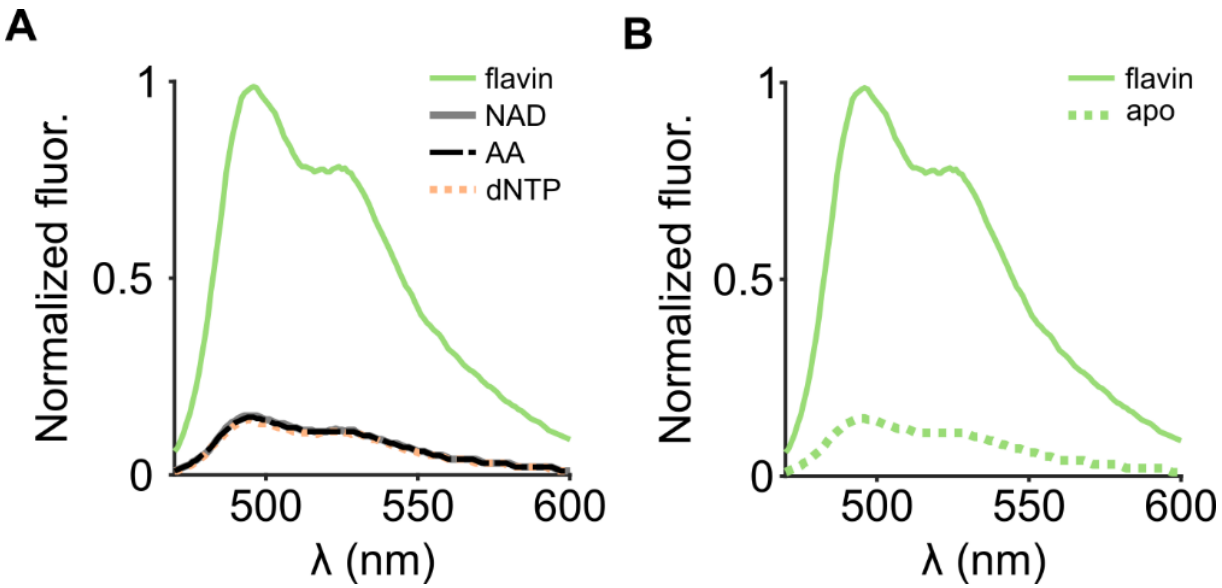

Figure 2.S6. Specificity of apo iLOV fluorescence. (A) Fluorescence of apo iLOV is specific to biological flavins: FAD, FMN, and riboflavin, and cannot be restored by titrating equimolar concentrations (10 μM) of common biological cofactors such as nicotinamide adenine dinucleotide (NAD), a mixture of amino acids (AA), and deoxynucleotides (dNTP). The nonspecific spectrum is found to be nearly identical to the (B) fluorescence of deflavinated apo iLOV, which is provided as a separate figure for reference.

Chapter 3. Rational design of a circularly permuted flavin-based fluorescent protein

This chapter is adapted from a paper pending submission to *Chem Biochem*. I am the first author on this paper. My contributions include project conceptualization, experimental design, performing experiments, data analysis, and writing and revision.

Journal: *Chem Biochem*

Authors: Nolan T. Anderson, Jason S. Xie, Asish N. Chacko, Vannie L. Liu, Kang-Ching Fan, and Arnab Mukherjee

3.1. Abstract

Flavin-based fluorescent proteins (FbFPs) are small oxygen-independent genetically encodable fluorophores which hold great promise for imaging of anaerobic and hypoxic biological systems. However, present unavailability of a circularly permuted FbFP has hindered construction of oxygen-independent fluorescent biosensors, as circular permutation of the Green Fluorescent Protein has helped enable the construction of a broad variety of genetically encoded biosensors. Within, we report the first successful circular permutation of the flavin-based fluorescent protein iLOV. Careful engineering of the linker connecting the protein's original termini yielded a dimly fluorescent circularly permuted iLOV; however, fluorescence can be recovered to approximately the original intensity by fusing dimerizing coiled-coil peptides, most notably leucine zippers, to the newly created termini. We have demonstrated the utility of a circularly permuted iLOV by engineering turn-off biosensors for activity of two proteases, tobacco etch virus protease and SARS-CoV-2 main protease M^{pro}. Insertion of a protease cleavage sequence within the protein's engineered linker yields a protein which is bright in the absence of protease activity, but significantly loses fluorescence upon expression of the respective protease, providing a fluorescent readout which can be observed by flow cytometry or fluorescence microscopy. We expect circular permutation of iLOV will enable creation of a broad library of FbFP-based fluorescent biosensors, thereby extending the usefulness of fluorescent biosensing to anaerobic and hypoxic contexts.

3.2. Introduction

Circular permutation is a protein design method that is utilized both in nature and laboratory engineering which involves the connection of the original N- and C-termini of a parent protein through an amino acid linker, while introducing new termini elsewhere in the protein sequence.[1] This approach has been successfully used to design proteins with improved catalytic efficiency, altered ligand affinity, and increased stability. In addition, circular permutation of the green fluorescent protein (GFP)[2] and its derivatives has enabled the creation of a diverse class of fluorescent biosensors.[3] However, GFP and related reporters require oxygen to mature into a light-emitting state, which limits their usefulness in anaerobic and low-oxygen environments.[4,5] Therefore, alternative reporter proteins have been developed that generate fluorescence by binding to small-molecule chemicals.[4] These oxygen-independent reporter proteins can be broadly classified into two categories: those that bind to and activate the fluorescence of externally added dyes,[6,7] and those that bind to internal cellular chromophores,[7,8] such as flavins. Unlike dye-binding reporters, flavin-based fluorescent proteins (FbFPs), do not require an external supply of chemical fluorogens because flavins are available endogenously in all cells. Consequently, FbFPs have been widely used as fully autonomous genetic reporters in a variety of mammalian,[9] bacterial,[10,11] fungal,[12] and viral systems.[13] Most recently, FbFPs have also been utilized to create split-protein based biosensors by dividing the FbFP backbone into two fragments, each of which is non-fluorescent individually, but becomes fluorescent when brought together by interactions between proteins that are genetically fused to the ends of each fragment.[14] However, in contrast to nearly every major class of fluorescent protein, FbFPs have not been circularly permuted. As a result, it

remains unclear whether FbFPs can tolerate circular permutation and whether this approach could provide a potential avenue for creating fluorescent biosensors based on FbFPs.

In this work, we utilized structural data and site-directed mutagenesis to undertake a systematic investigation of an FbFP, with the objective of identifying potential sites for circular permutation. Through this process, we successfully identified a potential permutation site, and subsequent experiments demonstrated that relocating the N- and C-termini to this position produced a circularly permuted FbFP, although with a significant reduction in fluorescence. However, we were able to show that the fluorescence loss in the circularly permuted construct could be completely restored by fusing dimerizing protein pairs at the newly assigned N- and C-termini. Finally, to demonstrate the utility of circularly permuted FbFP as a scaffold for sensor engineering, we developed and characterized two FbFP-based protease sensors and demonstrated their application in mammalian cells.

3.3. Methods

3.3.1. Molecular biology

For bacterial expression and testing, iLOV and circularly permuted variants were cloned in pJUMP-27-1A (Addgene #126974), an expression plasmid that allows for constitutive expression from the synthetic J23100 promoter. For mammalian expression, iLOV, circularly permuted variants, and protease-sensing constructs were cloned in a lentiviral transfer plasmid used in our previous works (pLenti). Genes encoding TEV protease, low-affinity nerve growth factor receptor (LNGFR), and Mpro were synthesized by Integrated DNA Technologies (Coralville, IA, USA). Gene fragments were amplified using Q5 High-Fidelity 2X Master Mix (New England Biolabs, Ipswich, MA, USA) and assembled using NEBuilder HiFi Gibson Assembly Master Mix (New England Biolabs). The assembled plasmids were propagated by electro-transformation in E. coli NEB10β cells or heat shock

transformation of chemically competent E. coli NEB Stable cells. Plasmids were purified using the PureYield Plasmid Miniprep system (Promega, Madison, WI, USA) and sequence verified using whole plasmid nanopore sequencing (Plasmidsaurus, Eugene, OR, USA). All plasmids used and constructed in this work are described in **Table 3.S1**.

3.3.2. Bacterial expression and whole-cell fluorescence measurements

pJUMP-27-1A vectors harboring circularly permuted iLOV variants were expressed in E. coli, BL21(DE3) cells using electroporation. Transformed cells were grown overnight (~16 hours) in minimal M9 media with 50 μg/mL kanamycin at 37 °C on an orbital shaker (200 r.p.m.). The overnight culture was diluted at 1:100 into fresh M9 media and grown for an additional 3-4 hours to reach an optical density (OD_{600}) of ~0.5. 200 μL of this culture was transferred to a black, clear bottom 96-well plate (Costar, Corning, Kennebunk, ME, USA) and whole-cell fluorescence measurements were acquired using a plate reader (Tecan Spark, Mannedörf, Switzerland). The excitation wavelength was set at 420 nm, while the emission spectrum was scanned from 465 nm to 600 nm in steps of 2 nm. Excitation and emission monochromator bandwidths were 20nm. Total fluorescence was computed by numerical integration of the emission spectrum from 465 nm to 531 nm.

3.3.3. Mammalian cell culture and engineering

Chinese hamster ovary (CHO) cells and HEK 293T cells were cultured in high-glucose Dulbecco's Modified Eagle's Medium (DMEM) (Thermo-Fisher, Waltham, MA, USA) supplemented with 10% fetal bovine serum (R&D Systems, Minneapolis, MN, USA), 100U/mL penicillin-streptomycin (Thermo-Fisher), and 1 mM sodium pyruvate (Thermo-Fisher). Cells were routinely grown at 37°C in a humidified chamber containing 5% CO_2. For

passaging, sorting, and fluorescence analysis, adherent cells were detached from plates using a commercial trypsin-based formulation, TrypLE Express (Thermo-Fisher).

Stable mammalian cell lines were generated by lentiviral transduction, using a transfer plasmid (pLenti) designed to heterologously express proteins from the constitutive human EF1α promoter. To enable enrichment of successfully transduced cells by flow sorting, an internal ribosome entry site (IRES) was used to co-express mCherry together with iLOV, circularly permuted variants, and the protease-sensing constructs. For protease expression, the relevant protease-encoding gene (TEVp or Mpro) was placed under the control of a doxycycline-inducible minimal CMV promoter in the same lentiviral backbone (pLenti). To enable flow sorting of protease-expressing cells, a self-cleaving T2A sequence was used to co-express the low-affinity nerve growth factor receptor (LNGFR).

Lentivirus was synthesized via polyethyleneimine transfection of a 10 cm plate of 50-70% confluent HEK293T cells, using a plasmid mixture comprising a lentiviral packaging plasmid (pPackaging), a broad-tropism conferring envelope plasmid (pVSV-G), and the transfer plasmid (pLenti) carrying the gene of interest. Prior to transfection, all plasmids were purified using the PureYield Plasmid Midiprep System (Promega). 72 hours post transfection, lentivirus particles were concentrated from the supernatant using a commercial PEG-based formulation, LentiX (Takara Bio USA, San Jose, CA, USA). CHO Tet-On 3G cells (Takara Bio USA) were then transduced with the resultant lentivirus using spinfection (1050 x g, 1.5 h, 30°C) in the presence of 8 μg/mL polybrene (Santa Cruz Biotechnology, Santa Cruz, CA, USA), and cultured for 3 days before enriching for stably transduced cells using fluorescence-activated cell sorting.

3.3.4. Fluorescence-activated cell sorting and cytometry

Fluorescence-activated cell sorting (FACS) was used to enrich for stably transduced cells expressing iLOV, circularly permuted iLOV variants, protease-sensing constructs, TEVp, and Mpro. To sort for cells expressing iLOV, circularly permuted constructs, and protease sensor prototypes, we relied on mCherry as a co-expressed fluorescent transduction marker. To sort for cells expressing TEVp or Mpro, we labeled cells with VioBlue REAffinity anti-LNGFR antibody (Miltenyi Biotec, Auburn, CA, USA). Cell sorting was performed using either a Sony MA-900 cell sorter or a Sony SH-800 cell sorter (Sony Biotechnology, San Jose, CA, USA). In preparation for FACS, cells were dissociated at 50-90 % confluence using TrypLE express, pelleted by gentle centrifugation (300 x g, 5 min), resuspended in 0.5 - 1 mL sterile phosphate buffered saline (PBS), and maintained on ice. Cell populations were first gated for live cells (based on forward scatter-area vs. side scatter plot), then singlet cells (based on forward scatter-height vs. width plot), and then for fluorescence in the red (mCherry) or violet channels (anti-LNGFR) based on the appropriate transduction marker. For all constructs, at least 10,000 cells were sorted into 2 mL of DMEM, and then plated on 6 well plates. Once the sorted populations achieved 50-90 % confluence, they were further expanded on 10 cm tissue culture plates and cryo-stocked in BamBanker medium (Nippon Genetics Europe, Dueran, Germany) at – 80 °C or liquid nitrogen for future use.

Fluorescence distributions of engineered cells were assessed by flow cytometry, following a similar protocol as described above, but omitting the cell collection and cryo-storage steps. In order to normalize for variations in viral integration events and reporter gene expression levels across transduced cell lines, the fluorescence in the green channel (iLOV) was divided by fluorescence in the red channel (mCherry) for analysis.

For experiments involving sensing of protease activity, cells were treated with 10 μg/mL doxycycline (Sigma-Aldrich, St. Louis, MO, USA) for 24 hours to induce protease expression. Because doxycycline fluoresces at a wavelength that overlaps with the iLOV spectrum, doxycycline was washed out by changing media to doxycycline-free media approximately 1 hour prior to harvesting cells for flow cytometry as previously described.[26] Protease-driven changes in the fluorescence distribution of cells were assessed by flow cytometry, as described above.

3.3.5. Fluorescence microscopy

Cells were grown in chambered glass-bottom slides (Thermo-Fisher) to approximately 50 % confluence. Cells were washed twice with PBS, then imaged on an inverted fluorescence microscope (Olympus, Shinjuku, Tokyo, Japan) using a 10X 0.25 numerical aperture air-immersion objective (Olympus). Excitation light was provided by a SOLA Light Engine (Lumencor, Beaverton, OR, USA) and filtered through a neutral density filter (Edmund Optics, Barrington, NJ, USA) to an illumination intensity of ~5 mW at the sample plane. EGFP or mCherry filter cubes (Chroma, Bellows Falls, VT, USA) comprising the appropriate excitation/emission filters and dichroic mirrors were used to image iLOV (and related variants) and mCherry-expressing cells, respectively. To minimize photobleaching, exposure was limited to ~ 5 s per field of view.

3.3.6. Statistical analysis

Experimental data are summarized by their mean and standard deviation obtained from multiple ($n \geq 3$) biological replicates defined as measurements performed with distinct cell samples. All tests are 2-sided and a P value of less than 0.05 taken to indicate statistical significance.

3.4. Results and Discussion

3.4.1. iLOV can be circularly permuted between Gly95 and Glu96 in the Hβ-Iβ loop

To engineer a circularly permuted FbFP, we focused our efforts on iLOV, a 111-amino acid FbFP derived by engineering the light, oxygen, and voltage (LOV) sensing domain of *A. thaliana* phototropin.[13] Although the LOV fold is highly conserved among diverse species, there is significant variation in the lengths and conformations of loops connecting the secondary structure elements.[8] Accordingly, to identify a permissive location for circular permutation, we started by testing the ability of various loop regions in iLOV to tolerate the insertion of a protein domain. We chose the human estrogen receptor ligand-binding domain for insertion because this motif has been previously used to identify allosteric insertion 'hotspots' in Cas9.[15] We used Gibson assembly to insert the estrogen-binding domain into 23 loop locations in iLOV (**Fig. 3.1a,b**) and expressed the resulting constructs in *E. coli*. Whole-cell fluorescence measurements revealed a substantial reduction in fluorescence for most of the constructs. However, insertions in the loop connecting the Hβ and Iβ strands of iLOV produced significantly greater fluorescence compared to other insertions (**Fig. 3.1b**). Interestingly, the Hβ-Iβ loop was also recently identified as an optimal location to bisect the backbone of two related LOV proteins, including an FbFP from *Chloroflexus aggregans* (CagFbFP)[14] and a singlet oxygen generator, miniSOG,[16] to create split-protein sensors. Based on these findings, we decided to engineer a circular permutation in the Hβ-Iβ loop between Gly95 and Glu96 (**Fig. 3.1a**). Guided by the crystal structure of iLOV[17] (**Fig. 3.1c**), we initially decided to join the original N- and C-termini (separated by ~ 2.4 nm) using a 34-amino acid Jα helix that is found at the C-terminus of native phototropin's LOV2 domain but is deleted in iLOV.[18] However, the ensuing construct was found to be almost non-fluorescent in *E. coli*

(**Fig. 3.1d**). Therefore, we tested longer linkers comprising 5 or 10 amino acids appended to the Jα helix, both of which were found to significantly increase whole-cell fluorescence in *E. coli* (**Fig. 3.1d**). Accordingly, we proceeded to utilize circularly permuted iLOV with new N- and C-termini at Glu96 and Gly95 respectively, and a 39 amino acid linker connecting the original termini (**Fig. 3.1e**), for expression and further testing in mammalian cell lines.

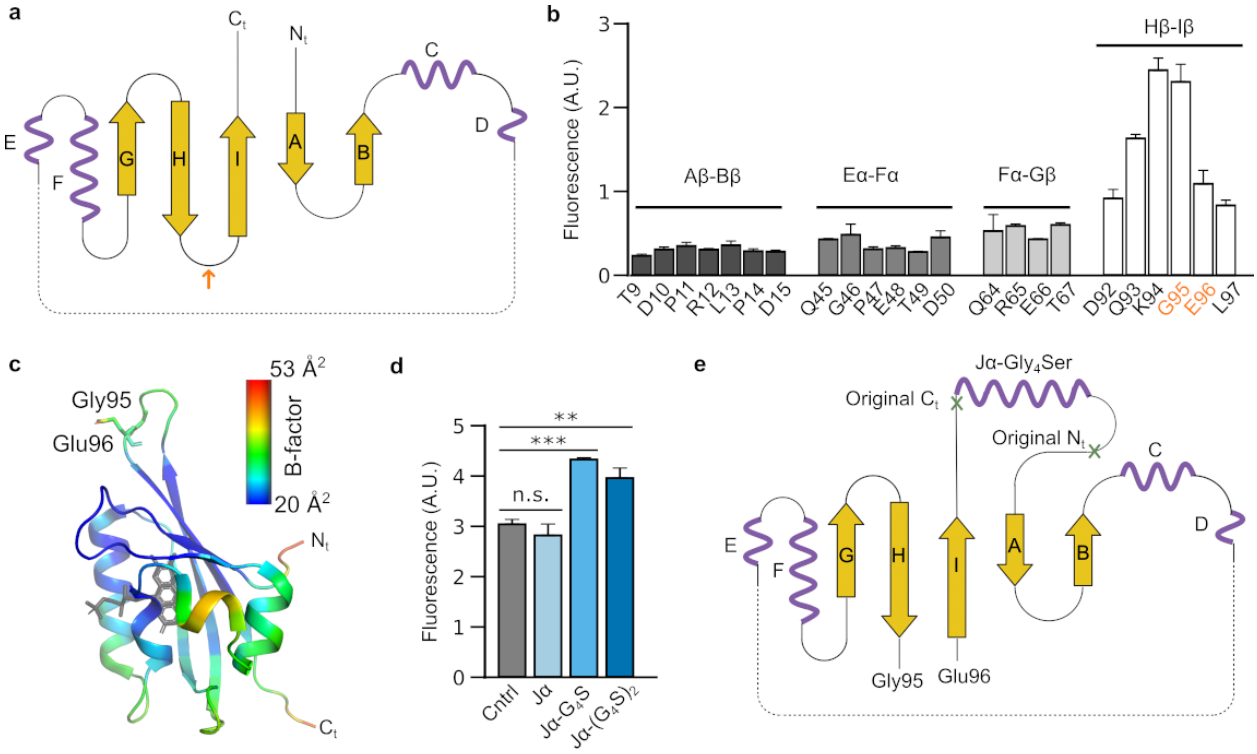

Figure 3.1: Circular permutation of iLOV. a) Topology diagram showing secondary structure elements of iLOV with thick arrows and wavy lines representing β-sheets and α-helices respectively, and the intervening lines indicating loops. The β-sheets and α-helices are numbered alphabetically. The orange arrow indicates the site of circular permutation in the loop connecting the H and I β-sheets. **b)** Whole-cell fluorescence of *E. coli* cells expressing iLOV constructs that harbor insertions of the estrogen receptor ligand-binding domain in various loops. **c)** Crystal structure of iLOV (PDB: 4EES),[17] showing location of circular permutation in the Hβ-Iβ loop. The crystal structure is colored as per B-factors. **d)** Whole-cell fluorescence measurements of *E. coli* cells expressing iLOV constructs that are circularly permuted at the Gly95-Glu96 junction and have the original ends connected by a Jα-based linker. **e)** Topology diagram showing secondary structure elements of circularly permuted iLOV with new termini and Glu96 and Gly95 and the Jα-linker connecting the original termini. Error bars represent the s.e.m. (n = 3 biological replicates). ** is P-value < 0.01; *** is P-value < 0.001.

3.4.2. Fluorescence of circularly permuted iLOV is increased by fusing coiled coils

We used lentiviral transduction to stably express circularly permuted iLOV with the above 39-residue linker in Chinese hamster ovary (CHO) cells. We also co-expressed mCherry using an internal ribosome entry site (IRES) to normalize for variations in reporter expression levels across cells. To measure baseline autofluorescence in the green (iLOV) channel, we engineered matched controls by transducing CHO cells with a similar lentiviral vector that expresses mCherry from the same promoter and via an IRES sequence but lacks the iLOV gene. We then measured the normalized fluorescence distribution (i.e., iLOV/mCherry) of test and control cells using flow cytometry. Although cells expressing circularly permuted iLOV showed a statistically significant increase in their mean fluorescence compared to control cells, the fold-change was fairly modest ($F/F_{control} = 1.95 \pm 0.04$, $n = 4$, P-val = 10^{-9}) (**Fig. 3.2a**). We hypothesized that the low fluorescence of circularly permuted iLOV could be caused by the separation of the newly formed N and C-termini, which are otherwise joined in iLOV (**Fig. 3.1e**). To address this issue, we attempted to bring the new termini in proximity through the insertion of a pair of heterodimerizing coiled coils[19,20] at each end (**Fig. 3.2b**). In addition, we also engineered a construct where the N- and C-termini are fused with segments of a split intein (derived from gp41-1),[21] which enables adjacent residues to be joined via a peptide bond. Flow cytometry measurements revealed a substantial increase, ranging from 5-20-fold, in the mean fluorescence of cells expressing the aforementioned constructs (**Fig. 3.2c,d, Fig. 3.S1**), with the highest fluorescence observed in the circularly permuted construct harboring a pair of leucine zippers at the termini (**Fig. 3.2d,e**). In agreement with these results, fluorescence microscopy revealed distinctly fluorescent cells for cultures transduced with circularly permuted iLOV containing leucine zippers at the ends, whereas the cells expressing permuted

iLOV devoid of any coiled coils were substantially dimmer (**Fig. 3.2f**). Furthermore, removal of the Jα-helix connecting the original termini led to a sharp decrease in mean fluorescence of cells expressing the circularly permuted iLOV construct harboring leucine zippers (**Fig. 3.S2**). Based on these findings, we proceeded to use circularly permuted iLOV harboring both the Jα-linker and leucine zippers to build a sensor prototype for detecting protease activity in mammalian cells.

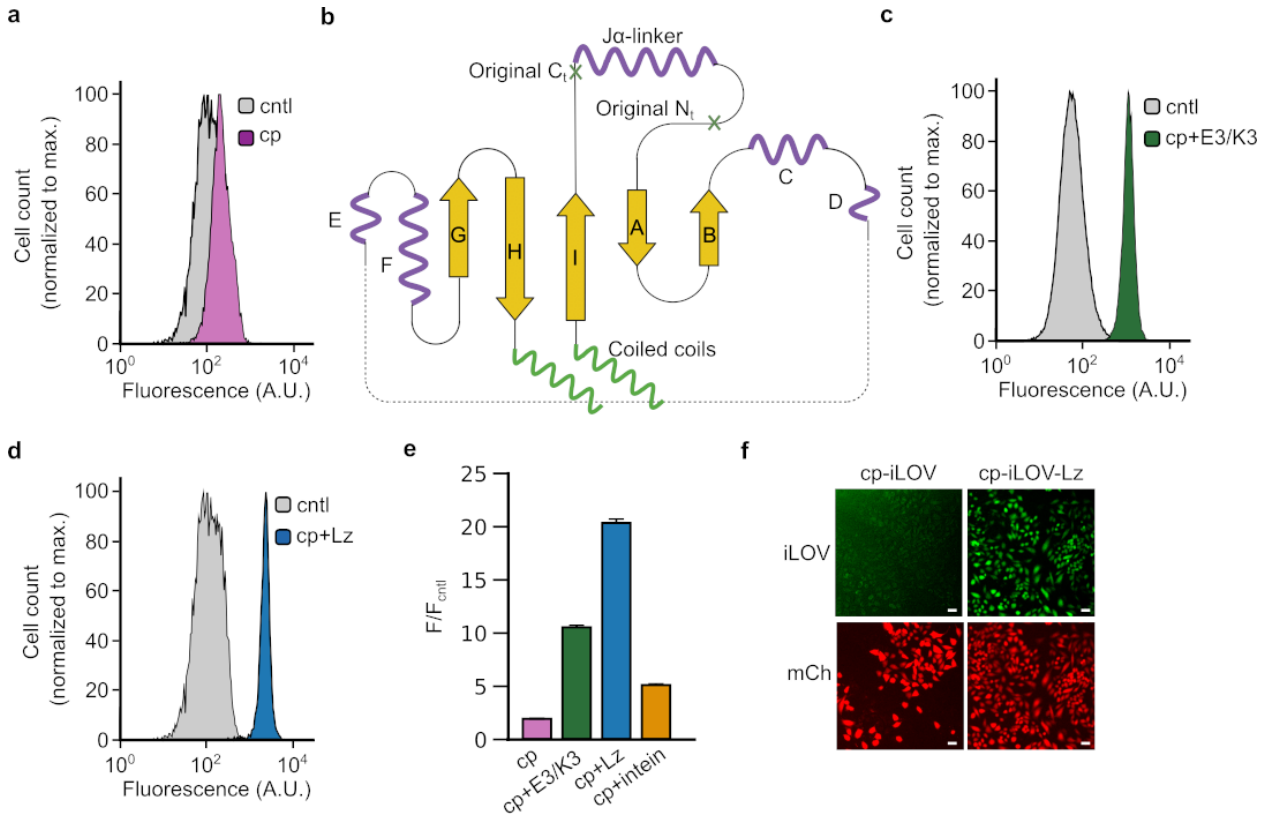

Figure 3.2: Optimization of circularly permuted iLOV. a) Topology diagram showing secondary structure elements of circularly permuted iLOV with thick arrows and wavy lines representing β-sheets and α-helices respectively, and the intervening lines indicating loops. The coiled coil domains fused at the N- and C-termini are depicted as green wavy lines. **b)** Fluorescence distribution of CHO cells transduced to express circularly permuted iLOV (cp). Fluorescence distribution of cells transduced with a similar control vector (cntl) that lacks iLOV expression is also shown for comparison. **c)** Fluorescence distribution of CHO cells transduced to express circularly permuted iLOV harboring a pair of heterodimerizing E3-K3 coiled coils at the N and C termini. **d)** Fluorescence distribution of CHO cells transduced to express circularly permuted iLOV harboring a pair of heterodimerizing leucine zippers at the N and C termini. **e)** Ratio of mean fluorescence levels of CHO cells

expressing various circularly permuted iLOV constructs to the baseline (auto)fluorescence of control cells that lack iLOV expression. **f)** Representative images of CHO cells expressing circularly permuted iLOV (without coiled coils) or circularly permuted iLOV harboring leucine zippers, imaged in the green and red channels. Scale bar is 10 μm. Error bars represent the standard deviation (n = 4 biological replicates).

3.4.3. Circularly permuted iLOV can be engineered to detect protease activity

We sought to engineer protease-sensing function by inserting a cleavage sequence for the tobacco etch virus protease (TEVp) at various locations in circularly permuted iLOV, including at the junctions between the leucine zippers and the newly formed N- and C-termini (**Fig. 3.S3a**) and following the Jα peptide that was introduced to connect the original N and C-termini in (non-permuted) iLOV (**Fig. 3.3a**). Our expectation was that TEV cleavage would lead to a decrease in cellular fluorescence by perturbing the intact circularly permuted structure. We doubly transduced CHO cells to express the candidate TEVp-sensing constructs, as well as TEVp, the latter under the control of a doxycycline-inducible minimal CMV promoter. Using flow cytometry, we measured the fluorescence distribution of cells in the presence and absence of doxycycline. TEVp expression led to a reduction in mean fluorescence of cells expressing circularly permuted iLOV with a TEVp cleavage site at the junctions connecting one or both leucine zippers (**Fig. 3.S3b-d**). However, the largest decrease in fluorescence ($-\Delta F/F_o$ = 73 ± 10 %, n = 3, P-value = 5 x 10^{-4}) was obtained in cells harboring circularly permuted iLOV containing a TEVp cleavage site after the Jα linker (**Fig. 3.3a,b**), which was confirmed by fluorescence spectroscopy on lysates prepared from cells with and without TEVp expression (**Fig. 3.S4**). Control measurements showed that TEVp expression did not affect the fluorescence of cells expressing non-permuted iLOV (**Fig. 3.S5**), indicating that the change in fluorescence was specific to the engineered constructs.

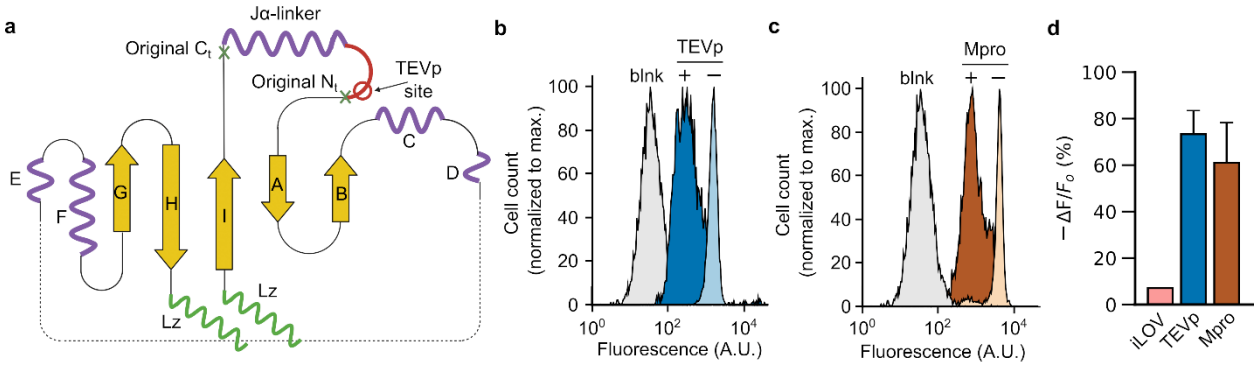

Figure 3.3: Engineering protease-sensing function in circularly permuted iLOV. a) Topology diagram showing secondary structure elements of circularly permuted iLOV harboring leucine zippers (Lz) at the N- and C-termini and a TEVp cleavage site inserted after the Jα linker. **b)** Fluorescence distribution of CHO cells transduced to express doxycycline-inducible TEVp and circularly permuted iLOV with a TEVp cleavage site, in the presence and absence of TEVp expression. Autofluorescence distribution of cells transduced with a control vector (blnk) that lacks iLOV expression is also shown for comparison. **c)** Fluorescence distribution of CHO cells transduced to express doxycycline-inducible Mpro and circularly permuted iLOV with an Mpro cleavage site, in the presence and absence of Mpro expression. Autofluorescence distribution of cells transduced with a control vector (blnk) that lacks iLOV expression is also shown for comparison. **d)** Percent change in mean fluorescence of cells ($-\Delta F/F_o$) expressing iLOV (non-permuted), TEVp and Mpro-sensing constructs, induced by expression of the respective protease. Error bars represent the standard deviation (n = 3 biological replicates).

Next, we sought to determine whether our approach could be adapted for detecting the activity of a different protease. Given that the largest TEVp-dependent decrease in fluorescence was observed in the construct consisting of a TEVp cleavage site following the Jα-linker, we replaced this site with a sequence that is recognized and cleaved by the SARS-CoV-2 main protease, Mpro, a coronavirus protease that is essential for viral replication.[22] We introduced this construct into CHO cell lines that were also transduced for Mpro expression from a doxycycline-inducible promoter. We acquired on- and off-state fluorescence measurements using flow cytometry. Consistent with our previous observations with the TEVp system, expression of Mpro also led to a significant decrease in mean cellular fluorescence ($-\Delta F/F_o$ = 61 ± 17 %, n = 3, P-value = 8 x 10^{-4}) (**Fig. 3.3c,d**). Taken together, these results

confirm the ability of circularly permuted iLOV to serve as a scaffold for engineering the detection of protease activity in mammalian cells.

3.5. Conclusions

In this study, we engineered circularly permuted iLOV by making three rational modifications, including the introduction of new N- and C-termini at Glu96 and Gly95 respectively, connecting the original termini using a Jα-based linker, and attaching heterodimerizing leucine zippers at the new termini. The ensuing construct exhibits significantly improved cellular brightness compared to circularly permuted iLOV that lacks dimerizing coiled coils at the ends. Furthermore, by inserting protease cleavage sequences after the Jα linker, circularly permuted iLOV can be used as a turn-off sensor to detect protease activity in mammalian cells. Given that the LOV fold is conserved among different FbFPs, we anticipate that a similar design approach could be used to create circularly permuted versions of other FbFPs; including thermostable variants, such as CagFbFP;[23] brighter variants, such as CreiLOV;[24] and photostable variants, such as phiLOV.[25] Furthermore, the current work provides a proof-of-concept for engineering FbFP-based sensors, which could be expanded in future studies to detect a range of biological targets - for example, it may be possible to detect protein-protein interactions by replacing the leucine zippers at each terminus of circularly permuted iLOV with protein targets of interest.

The current study has some limitations that we intend to address in future research. First, our experiments were conducted in ambient oxygen conditions. While iLOV fluorescence is independent of oxygen availability, future work should evaluate the performance of circularly permuted iLOV in low-oxygen environments to enable broad application in hypoxic and anaerobic biological systems. In addition, a comprehensive

mechanistic investigation of protease-dependent changes in the fluorescence of circularly permuted iLOV is a crucial avenue for future research. In particular, the protease-driven loss in fluorescence could be caused by a number of factors, including a reduction in the stability of circularly permuted iLOV due to removal of the leucine zippers or the Jα-linker, a decrease in cofactor (i.e., flavin) binding, and change in protein-fluorophore interactions. Detailed biochemical and structural studies are required to test these mechanisms. Lastly, our circular permutation design expands the size of iLOV from 111 to 212 amino acids, diminishing one of the important advantages of iLOV, which is its reduced footprint compared to larger GFP-based reporters (~ 238 amino acids). To address this issue, we have initiated a process of truncating iLOV by systematically removing amino acids from both the N and C-termini, as well as internal residues. However, further research is necessary to determine whether truncated iLOV can tolerate circular permutation and to examine the effects of incorporating shorter dimerizing pairs at the termini.

The discovery of circularly permuted GFP provided a molecular framework for creating a diverse array of sensors which continues to evolve to this present day. Along similar lines, we envision that the discovery of circularly permuted iLOV will synergize with recent advances in developing split FbFPs to create new opportunities for monitoring cellular processes, particularly in scenarios where traditional GFP-based reporters are less effective.

3.6. References

(1) Yu, Y.; Lutz, S. Circular Permutation: A Different Way to Engineer Enzyme Structure and Function. *Trends in Biotechnology* **2011**, *29* (1), 18–25. https://doi.org/10.1016/j.tibtech.2010.10.004.

(2) Baird, G. S.; Zacharias, D. A.; Tsien, R. Y. Circular Permutation and Receptor Insertion within Green Fluorescent Proteins. *Proceedings of the National Academy of Sciences* **1999**, *96* (20), 11241–11246. https://doi.org/10.1073/pnas.96.20.11241.

(3) Kostyuk, A. I.; Demidovich, A. D.; Kotova, D. A.; Belousov, V. V.; Bilan, D. S. Circularly Permuted Fluorescent Protein-Based Indicators: History, Principles, and Classification. *International Journal of Molecular Sciences* **2019**, *20* (17), 4200. https://doi.org/10.3390/ijms20174200.

(4) Ozbakir, H. F.; Anderson, N. T.; Fan, K.-C.; Mukherjee, A. Beyond the Green Fluorescent Protein: Biomolecular Reporters for Anaerobic and Deep-Tissue Imaging. *Bioconjugate Chemistry* **2020**, *31* (2), 293–302. https://doi.org/10.1021/acs.bioconjchem.9b00688.

(5) Craggs, T. D. Green Fluorescent Protein: Structure, Folding and Chromophore Maturation. *Chem. Soc. Rev.* **2009**, *38* (10), 2865–2875. https://doi.org/10.1039/B903641P.

(6) Plamont, M.-A.; Billon-Denis, E.; Maurin, S.; Gauron, C.; Pimenta, F. M.; Specht, C. G.; Shi, J.; Quérard, J.; Pan, B.; Rossignol, J.; Moncoq, K.; Morellet, N.; Volovitch, M.; Lescop, E.; Chen, Y.; Triller, A.; Vriz, S.; Le Saux, T.; Jullien, L.; Gautier, A. Small Fluorescence-Activating and Absorption-Shifting Tag for Tunable Protein Imaging in Vivo. *Proceedings of the National Academy of Sciences* **2016**, *113* (3), 497–502. https://doi.org/10.1073/pnas.1513094113.

(7) Kumagai, A.; Ando, R.; Miyatake, H.; Greimel, P.; Kobayashi, T.; Hirabayashi, Y.; Shimogori, T.; Miyawaki, A. A Bilirubin-Inducible Fluorescent Protein from Eel Muscle. *Cell* **2013**, *153* (7), 1602–1611. https://doi.org/10.1016/j.cell.2013.05.038.

(8) Mukherjee, A.; Schroeder, C. M. Flavin-Based Fluorescent Proteins: Emerging Paradigms in Biological Imaging. *Current Opinion in Biotechnology* **2015**, *31*, 16–23. https://doi.org/10.1016/j.copbio.2014.07.010.

(9) Walter, J.; Hausmann, S.; Drepper, T.; Puls, M.; Eggert, T.; Dihné, M. Flavin Mononucleotide-Based Fluorescent Proteins Function in Mammalian Cells without Oxygen Requirement. *PLOS ONE* **2012**, *7* (9), e43921.

(10) Gawthorne, J. A.; Reddick, L. E.; Akpunarlieva, S. N.; Beckham, K. S. H.; Christie, J. M.; Alto, N. M.; Gabrielsen, M.; Roe, A. J. Express Your LOV: An Engineered Flavoprotein as a Reporter for Protein Expression and Purification. *PLOS ONE* **2012**, *7* (12), e52962.

(11) Lobo, L. A.; Smith, C. J.; Rocha, E. R. Flavin Mononucleotide (FMN)-Based Fluorescent Protein (FbFP) as Reporter for Gene Expression in the Anaerobe Bacteroides Fragilis. *FEMS Microbiology Letters* **2011**, *317* (1), 67–74. https://doi.org/10.1111/j.1574-6968.2011.02212.x.

(12) Tielker, D.; Eichhof, I.; Jaeger, K. E.; Ernst, J. F. Flavin Mononucleotide-Based Fluorescent Protein as an Oxygen-Independent Reporter in Candida Albicans and Saccharomyces Cerevisiae. *Eukaryotic Cell* **2009**, *8* (6), 913–915. https://doi.org/10.1128/ec.00394-08.

(13) Chapman, S.; Faulkner, C.; Kaiserli, E.; Garcia-Mata, C.; Savenkov, E. I.; Roberts, A. G.; Oparka, K. J.; Christie, J. M. The Photoreversible Fluorescent Protein iLOV Outperforms GFP as a Reporter of Plant Virus Infection. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105* (50), 20038–20043. https://doi.org/10.1073/pnas.0807551105.

(14) Yudenko, A.; Smolentseva, A.; Maslov, I.; Semenov, O.; Goncharov, I. M.; Nazarenko, V. V.; Maliar, N. L.; Borshchevskiy, V.; Gordeliy, V.; Remeeva, A.; Gushchin, I. Rational Design of a Split Flavin-Based Fluorescent Reporter. *ACS Synth. Biol.* **2021**, *10* (1), 72–83. https://doi.org/10.1021/acssynbio.0c00454.

(15) Oakes, B. L.; Nadler, D. C.; Flamholz, A.; Fellmann, C.; Staahl, B. T.; Doudna, J. A.; Savage, D. F. Profiling of Engineering Hotspots Identifies an Allosteric CRISPR-Cas9 Switch. *Nat Biotechnol* **2016**, *34* (6), 646–651. https://doi.org/10.1038/nbt.3528.

(16) Boassa, D.; Lemieux, S. P.; Lev-Ram, V.; Hu, J.; Xiong, Q.; Phan, S.; Mackey, M.; Ramachandra, R.; Peace, R. E.; Adams, S. R.; Ellisman, M. H.; Ngo, J. T. Split-miniSOG for Spatially Detecting Intracellular Protein-Protein Interactions by Correlated Light and Electron Microscopy. *Cell Chemical Biology* **2019**, *26* (10), 1407-1416.e5. https://doi.org/10.1016/j.chembiol.2019.07.007.

(17) Christie, J. M.; Hitomi, K.; Arvai, A. S.; Hartfield, K. A.; Mettlen, M.; Pratt, A. J.; Tainer, J. A.; Getzoff, E. D. Structural Tuning of the Fluorescent Protein iLOV for Improved Photostability. *J Biol Chem* **2012**, *287* (26), 22295–22304. https://doi.org/10.1074/jbc.M111.318881.

(18) Harper, S. M.; Christie, J. M.; Gardner, K. H. Disruption of the LOV−Jα Helix Interaction Activates Phototropin Kinase Activity. *Biochemistry* **2004**, *43* (51), 16184–16192. https://doi.org/10.1021/bi048092i.

(19) Ghosh, I.; Hamilton, A. D.; Regan, L. Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein. *J. Am. Chem. Soc.* **2000**, *122* (23), 5658–5659. https://doi.org/10.1021/ja994421w.

(20) Apostolovic, B.; Klok, H.-A. pH-Sensitivity of the E3/K3 Heterodimeric Coiled Coil. *Biomacromolecules* **2008**, *9* (11), 3173–3180. https://doi.org/10.1021/bm800746e.

(21) Beyer, H. M.; Mikula, K. M.; Li, M.; Wlodawer, A.; Iwaï, H. The Crystal Structure of the Naturally Split Gp41-1 Intein Guides the Engineering of Orthogonal Split Inteins from Cis-Splicing Inteins. *The FEBS Journal* **2020**, *287* (9), 1886–1898. https://doi.org/10.1111/febs.15113.

(22) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of Mpro from SARS-CoV-2 and Discovery of Its Inhibitors. *Nature* **2020**, *582* (7811), 289–293. https://doi.org/10.1038/s41586-020-2223-y.

(23) Nazarenko, V. V.; Remeeva, A.; Yudenko, A.; Kovalev, K.; Dubenko, A.; Goncharov, I. M.; Kuzmichev, P.; Rogachev, A. V.; Buslaev, P.; Borshchevskiy, V.; Mishin, A.; Dhoke, G. V.; Schwaneberg, U.; Davari, M. D.; Jaeger, K.-E.; Krauss, U.; Gordeliy, V.; Gushchin, I. A Thermostable Flavin-Based Fluorescent Protein from Chloroflexus Aggregans: A Framework for Ultra-High Resolution Structural Studies. *Photochemical & Photobiological Sciences* **2019**. https://doi.org/10.1039/C9PP00067D.

(24) Mukherjee, A.; Weyant, K. B.; Agrawal, U.; Walker, J.; Cann, I. K. O.; Schroeder, C. M. Engineering and Characterization of New LOV-Based Fluorescent Proteins from Chlamydomonas Reinhardtii and Vaucheria Frigida. *ACS Synthetic Biology* **2015**, *4* (4), 371–377. https://doi.org/10.1021/sb500237x.

(25) Drepper, T.; Eggert, T.; Circolone, F.; Heck, A.; Krauss, U.; Guterl, J. K.; Wendorff, M.; Losi, A.; Gartner, W.; Jaeger, K. E. Reporter Proteins for in Vivo Fluorescence without Oxygen. *Nature Biotechnology* **2007**, *25* (4), 443–445. https://doi.org/10.1038/nbt1293.

(26) Khader, H.; Solodushko, V.; Al-Mehdi, A. B.; Audia, J.; Fouty, B. Overlap of Doxycycline Fluorescence with That of the Redox-Sensitive Intracellular Reporter roGFP. *Journal of Fluorescence* **2014**, *24* (2), 305–311. https://doi.org/10.1007/s10895-013-1331-6.

3.7. Supplementary Information

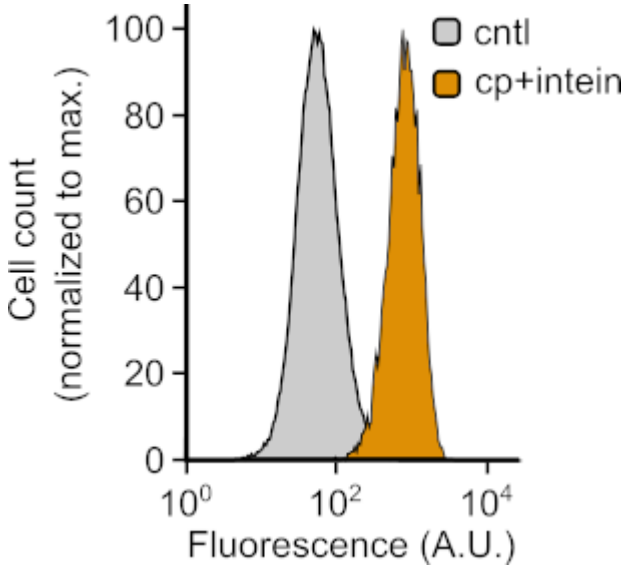

Figure 3.S1: Fluorescence distribution of cells expressing intein-spliced, circularly permuted iLOV. Fluorescence distribution of CHO cells transduced to express circularly permuted iLOV harboring fragments of a split gp41-1 intein at the N and C termini. Fluorescence in the green channel (iLOV) is normalized by fluorescence in the red channel (mCherry) to control for variations in viral integration events and gene expression level across cell lines.

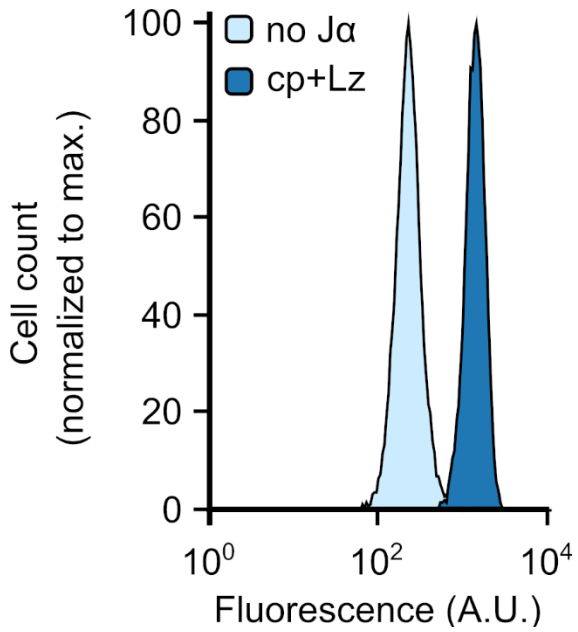

Figure 3.S2: Fluorescence distribution of cells expressing circularly permuted iLOV with leucine zippers, with or without the Jα-helix joining the original termini.

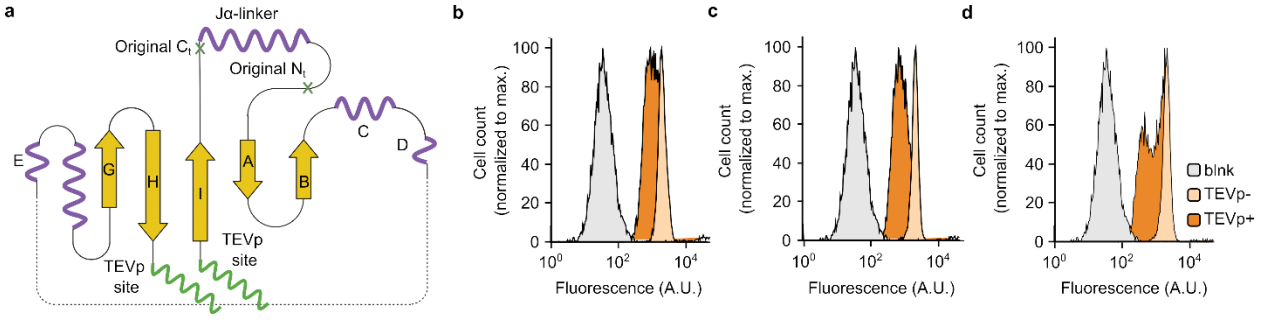

Figure 3.S3: Design of TEVp sensors based on circularly permuted iLOV. a) Topology diagram showing secondary structure elements of circularly permuted iLOV with thick arrows and wavy lines representing β-sheets and α-helices respectively, and the intervening lines indicating loops. The leucine zipper coiled coil domains that are fused at the N- and C-termini are depicted as green wavy lines. TEVp cleavage sites are inserted at one or both junctions formed between the leucine zippers and the termini. **b)** TEVp-driven changes in the fluorescence distribution of CHO cells transduced to express circularly permuted iLOV (cp) harboring TEVp cleavage sites at both junctions, **c)** at the N-terminal junction and **d)** the C-terminal junction. Fluorescence distribution of cells transduced with a similar control vector (cntl) that lacks iLOV expression is shown for comparison. Fluorescence in the green channel (iLOV) is normalized by fluorescence in the red channel (mCherry) to control for variations in viral integration events and gene expression level across cell lines.

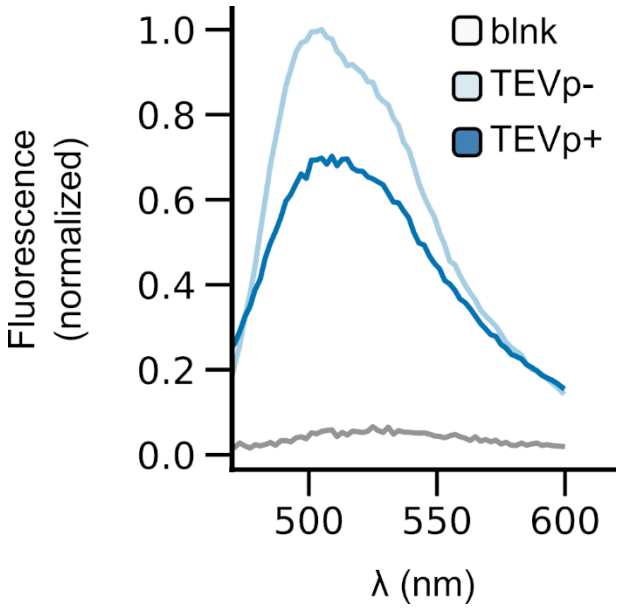

Figure 3.S4. Fluorescence spectra of TEVp sensor. Fluorescence emission spectra of lysates prepared from CHO cells transduced to express circularly permuted iLOV harboring a TEVp cleavage site after the Jα-helix, acquired with and without induction of TEVp expression. The blank spectrum (blnk) is obtained from the lysate of cells transduced with a similar lentivirus that lacks iLOV expression.

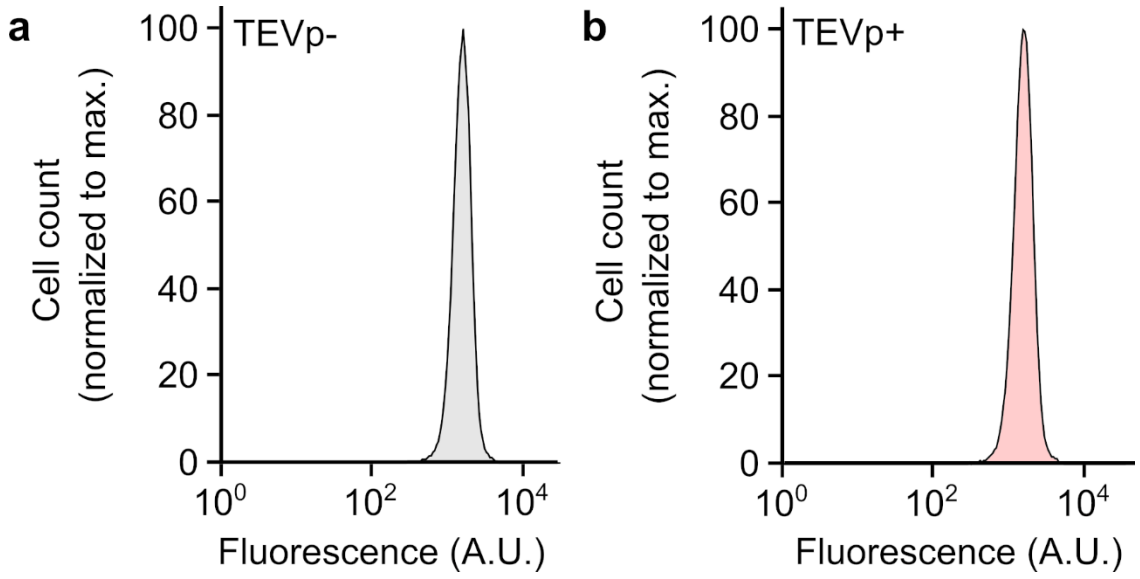

Figure 3.S5: Fluorescence distribution of cells expressing iLOV with and without TEVp expression. The fluorescence profile if iLOV-expressing cells lacking any TEVp cleavage site is identical **a)** with and **b)** without induction of TEVp expression. Fluorescence in the green channel (iLOV) is normalized by fluorescence in the red channel (mCherry) to control for variations in viral integration events and gene expression level across cell lines.

Table 3.S1. List of plasmids

Plasmid	**Description**
pJUMP-MT	Empty bacterial expression vector
pJUMP-iLOV	Constitutive bacterial expression of iLOV
pJUMP-cpiLOV-Jα	Constitutive bacterial expression of cpiLOV with Jα helix, no additional linker
pJUMP-cpiLOV-Jα-Gly_4Ser	Constitutive bacterial expression of cpiLOV with Jα helix and GGGGS linker
pJUMP-cpiLOV-Jα-$(Gly_4Ser)_2$	Constitutive bacterial expression of cpiLOV with Jα helix and GGGGSGGGGS linker
pPkg	Helper plasmid for lentivirus production, supplies viral packaging genes
pVSV-G	Helper plasmid for lentivirus production, expresses VSV-G to confer broad viral tropism

pLenti-MT-IRES-mCh	Empty mammalian (lentiviral) expression vector, using mCherry as a transduction marker
pLenti-iLOV-IRES-mCh	Lentiviral vector for mammalian expression of iLOV
pLenti-cpiLOV-IRES-mCh	Lentiviral vector for mammalian expression of cpiLOV with a 39-amino acid Jα-linker
pLenti-cpiLOV-Jα-Lz	Lentiviral vector for mammalian expression of cpiLOV with a 39-amino acid Jα-linker and leucine zippers on termini
pLenti-cpiLOV-Jα-E3K3	Lentiviral vector for mammalian expression of cpiLOV with a 39-amino acid Jα-linker and E3/K3 electrostatically-dimerizing peptides on termini
pLenti-cpiLOV-Jα-gp41-1	Lentiviral vector for mammalian expression of cpiLOV with a 39-amino acid Jα-linker and fragments of a split gp41-1 intein on termini
pLenti-cpiLOV-Lz	Same as pLenti-cpiLOV-Jα-Lz, but without the Jα helix
pLenti-cpiLOV-E3K3	Same as pLenti-cpiLOV-Jα-E3K3, but without the Jα helix
pLenti-cpiLOV-gp41-1	Same as pLenti-cpiLOV-Jα-gp41-1, but without the Jα helix
pLenti-cpiLOV-Jα-TEVcs	Lentiviral vector for mammalian expression of cpiLOV-Jα with a TEVp cleavage sequence replacing G4S linker downstream of Jα-helix
pLenti-cpiLOV-TEVcs	Same as pLenti-cpiLOV-Jα-TEVcs, but without the Jα helix
pLenti-cpiLOV-Jα-TEVcsUS-Lz	Lentiviral vector for mammalian expression of cpiLOV-Jα-G4S with leucine zippers on termini, and a TEVp cleavage sequence between the N-terminal zipper and cpiLOV
pLenti-cpiLOV-Jα-TEVcsDS-Lz	Lentiviral vector for mammalian expression of cpiLOV-Jα-G4S with leucine zippers on termini, and a TEVp cleavage sequence between the C-terminal zipper and cpiLOV
pLenti-cpiLOV-Jα-TEVcs2-Lz	Lentiviral vector for mammalian expression of cpiLOV-Jα-G4S with leucine zippers on termini, and two TEVp cleavage sequences between the zippers and each terminus
pLenti-cpiLOV-Jα-Mprocs-Lz	Lentiviral vector for mammalian expression of cpiLOV-Jα with an Mpro cleavage sequence replacing G4S linker downstream of Jα-helix, and with leucine zippers on termini
pLenti-TEV-2A-LNGFR	Lentiviral vector for doxycycline-inducible expression of TEV protease

pLenti-Mpro-2A-LNGFR	Lentiviral vector for doxycycline-inducible expression of Mpro protease

Table 3.S2. Sequences of genetic constructs used in this work

Construct	**Amino acid sequence**
N-terminal leucine zipper	ALKKELQANKKELAQLKWELQALKKELAQ
C-terminal leucine zipper	EQLEKKLQALEKKLAQLEWKNQALEKKLAQ
IAAL E3	EIAALEKEIAALEKEIAALEK
IAAL K3	KIAALKEKIAALKEKIAALKE
gp41-1 N terminal fragment	MMLKKILKIEELDERELIDIEVSGNHLFYANDILTHNRDP IRKPPLPARWPIH
gp41-1 C terminal fragment	CLDLKTQVQTPQGMKEISNIQVGDLVLSNTGYNEVLNV FPKSKKKSYKITLEDGKEIICSEEHLFPTQTGEMNISGGL KEGMCLYVKE

Chapter 4. Engineering a minimally sized flavin-based fluorescent protein

This chapter consists of presently unpublished work, though preparation and submission of a manuscript based on this work is anticipated in the coming months. I will be the first author on such a manuscript, and have been responsible for conception of the project, design and execution of experiments, data analysis, and writing. Arnab Mukherjee will appear as the final author on this publication and has also participated in conception and design. This work has also benefited from significant contributions from Kang-Ching Fan, Rachel F. Taylor, Zoe A. Imansjah, and Jason S. Xie in execution of experiments.

4.1. Abstract

Flavin-based fluorescent proteins (FbFPs) are, in general, smaller than more commonly used fluorescent proteins, such as the green fluorescent protein. In particular, the relatively bright and frequently utilized FbFP iLOV, derived from *Arabidopsis thaliana*, consists of only 111 amino acids, making it among one of the smallest existing fluorescent proteins. Size constraints are common in experiments where fluorescent proteins are applied, either due to a genetic packaging limit, such as exists in adeno-associated virus, or due to the fluorescent tag's perturbation of the tagged protein. Therefore, we sought to minimize the size of iLOV. Through rational deletion of amino acids based on structural data, domain insertion, and previous work leading to a circular permutation of iLOV, we discovered ten amino acids that could be deleted from iLOV without major loss of fluorescence: five from the N-terminus, four from the C-terminus, and one internal deletion. Combination of these deletions yields a 101 amino acid protein we have termed "nanoLOV." NanoLOV retains 86.7 ± 2.2% of iLOV's fluorescence, while reducing its length by 9%. We seek to further characterize nanoLOV and apply it in mammalian contexts.

4.2. Introduction

While flavin-based fluorescent proteins (FbFPs) are well-noted for their independence from oxygen or any exogeneous cofactor, they also confer a major advantage over other fluorescent proteins due to their small size. FbFPs are typically in the range of the low hundreds of amino acids,[1] including iLOV at 111 amino acids[2] (**Table 4.1**). This compares favorably to the green fluorescent protein (GFP) and its relatives—GFP itself is 238 amino acids in length.[3] Further, other oxygen independent fluorescent proteins are larger than the smallest FbFPs. UnaG is 139 amino acids[4] and FAST1 is 125 amino acids.[5] Notably, FAST has been reduced in size to 98 amino acids (known as nanoFAST) by truncation of its N-terminal domain.[6] In the similarly light-based but mechanistically distinct imaging methodology of bioluminescence, NanoLuc is one of the smallest available bioluminescent proteins, at 175 amino acids,[7] distinctively larger than the smallest fluorescent proteins.

Reporter	Length (amino acids)	Mass (kDa)
EGFP[3]	238	26.9
iLOV[2]	111	12.9
CreiLOV[8]	119	12.9
CagFbFP[9]	107	11.6
EcFbFP[10]	137	15.7
UnaG[4]	139	15.6
FAST[5]	125	13.6
NanoFAST[6]	98	10.8
NanoLuc[7]	175	19.4

Table 4.1. Sizes of selected genetically encodable reporters.

One of the most widespread applications of fluorescent proteins is the fusion tagging of protein of interest. In this approach, a gene of interest is genetically fused to a fluorescent protein in a translational fusion, either directly or through a flexible linker. In these translational fusions, there is a preference for the smallest possible fluorescent tag in order to minimize perturbation of the tagged protein.[11] Large tags present several obstacles to studying

proteins in their natural contexts: greater genetic length decreases expression of the fusion construct and requires more cell resources for expression,[12] the steric obstacle of the fused protein may restrict access of substrates or binding partners to a protein's binding site,[13] or may disrupt protein folding.[14]

The greater genetic size of a larger fluorescent tag also complicates cell line engineering. Certain genetic engineering tools, such as lentivirus and adeno-associated virus, have limitations on the size of the genetic constructs they can confer. Lentivirus, if tasked with packaging a very large stretch of DNA between its long terminal repeats, loses efficiency in packaging as proviral length increases.[15,16] Contrastingly, adeno-associated virus appears to have a strong upper limit on packaging capacity of about 5kbp.[17] Complex genetic circuits with several co-expressed genes may push the size limits of these genetic engineering tools. Frequently, a fluorescent reporter is necessary in these circuits, either for visualization of a process or confirmation of editing. In such cases, minimization of the genetic size of fluorescent reporters saves valuable space.

As iLOV is an FbFP with relatively high brightness and relatively small size, we view iLOV as a strong candidate for developing a minimally sized fluorescent reporter. Particularly, we targeted for deletion unstructured domains located at termini and in loops between secondary structural features, as well as structural features with minimal flavin contacts and high fluctuation. Further, we utilized a randomized transposon-based codon deletion approach to scan for removable residues. Here we report a version of iLOV which we term nanoLOV, which at 101 amino acids is 10 amino acids smaller than iLOV at 111. Importantly, nanoLOV retains 86.5 ± 2.2% of the brightness of iLOV.

4.3. Materials and methods

4.3.1. Cloning and molecular biology

Aside from constructs used in transposon-based random codon deletion (for these constructs, see section 4.3.3), plasmids for bacterial protein expression were constructed as follows. Genes were cloned into the pJUMP27-1A(sfGFP) backbone (Addgene #126974). iLOV was sourced from a plasmid constructed in our previous work. DNA fragments were amplified via PCR using Q5 High-Fidelity 2X Master Mix (New England Biolabs, Ipswich, MA, USA). Fragments were assembled via Gibson assembly using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs). Assembled constructs were transformed into electrocompetent *Escherichia coli* NEB10β cells (New England Biolabs). Plasmids were purified from single clones using the PureYield Plasmid Miniprep system (Promega, Madison, WI, USA) and sequenced via either Sanger sequencing (Genewiz, South Plainfield, NJ, USA) or whole plasmid nanopore sequencing (Plasmidsaurus, Eugene, OR, USA).

4.3.2. Bacterial protein expression

Cloned plasmids were transformed into electrocompetent *E. coli* BL21(DE3) cells (New England Biolabs). To screen constructs, a culture was started in either 5mL of M9 medium with 50ug/mL kanamycin in a 14mL culture tube, or in 1mL of the same medium in a well of a 96-well deep well plate (2.2mL volume) sealed with a gas permeable membrane and grown overnight (~16-20h) at 37°C with shaking at 200 rpm. Cultures were then diluted 1:100 into 1mL of fresh M9/kanamycin media in deep well plates. These cultures were grown for a further 3-4 hours at 37°C and 200 rpm to an optical density (OD_{600}) of approximately 0.5. For analysis, 200µL of each culture was transferred to a black, clear bottom 96-well plate (Costar, Corning, Kennebunk, ME, USA) and whole-cell fluorescence measurements were

acquired using a plate reader (Tecan Spark, Mannedörf, Switzerland). A fluorescence emission spectrum was acquired using an excitation wavelength of 420nm and emission wavelengths ranging from 465nm to 600nm in step sizes of 2nm. Monochromator slit widths were 20nm. Total fluorescence emission was numerically integrated from 465nm to 531nm.

4.4. Results and Discussion

4.4.1. Truncation of termini

In order to identify regions of iLOV that are conducive to construction of allosteric sensors, we inserted the human estrogen receptor ligand binding domain (ERLBD) at every position in iLOV. Fluorescence screening of these constructs in BL21(DE3) *E. coli* revealed that insertions of ERLBD near the termini of iLOV retain similar fluorescence to wild type iLOV. ERLBD insertions at sites upstream of Val7 and sites downstream of Leu105 have fluorescence at least 95% of wild type iLOV **(Figure 4.1a).** This apparent insensitivity to domain insertion near iLOV's termini led us to infer that these terminal regions are not critical to iLOV fluorescence.

Accordingly, we deleted terminal regions of iLOV of varying lengths. From the N-terminus, we retained M1 (necessary as it is the start codon), and deleted stretches of 5-8 residues, making the first non-methionine residue Val7 to Asp10. Fluorescence screening indicated that a 5 amino acid deletion retains high fluorescence, but truncations of any greater number of residues greatly diminishes fluorescence **(Figure 4.1b).** Similarly, we deleted stretches of 4-9 residues from the C-terminus, making the terminal residue Gly102 to Gly107. Deletion of 4 amino acids from the C-terminus retains high fluorescence, though greater deletions result in gradually decreasing fluorescence **(Figure 4.1b).** Considering iLOV's pre-existing diminished brightness relative to GFP, we preferred not to use deletions that resulted

in any substantial loss of fluorescence in order to maintain a similar magnitude of fluorescence to wild type iLOV; therefore, we concluded that truncation of the first 5 and last 4 residues of iLOV was permissible, resulting in a protein running from Val7 to Gly107.

Viewed in the context of iLOV's known crystal structure, these truncations avoid disruption of secondary structure elements. Near the N-terminus, the Aβ sheet begins at Val7, the first residue which does not tolerate deletion. **(Figure 4.1c)** This implies that an intact Aβ sheet is necessary for iLOV fluorescence. Near the C-terminus, the final residue of the Iβ sheet is Leu105. **(Figure 4.1c)** Though we found that C-terminal truncation beyond Ser108 leads to diminished fluorescence, fluorescence decreases gradually until Leu107 is truncated, at which point fluorescence decreases severely. Similar to the Aβ sheet at the N-terminus, disruption of the Iβ sheet also extinguishes fluorescence.

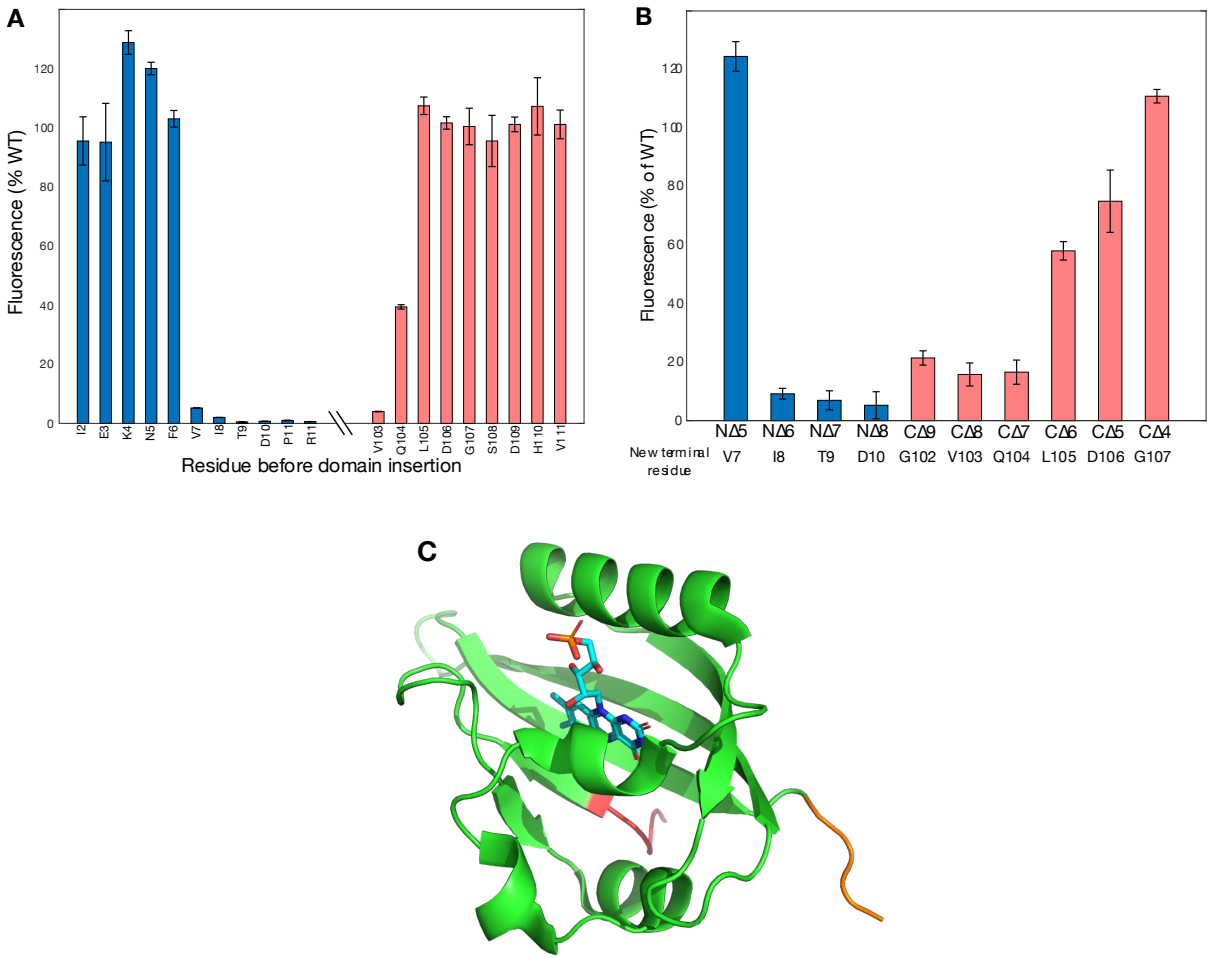

Figure 4.1. Trimming iLOV's termini. **a)** Fluorescence of iLOV with human estrogen receptor ligand binding domain (ERLBD) inserted at positions near its termini. For insertions near each terminus, near-wild type fluorescence is retained. Insertions beyond Val7 lead to a precipitous loss of fluorescence. Insertions upstream of Leu105 similarly dramatically lose fluorescence. **b)** Fluorescence of iLOV with truncated termini. Truncation of five residues from the N-terminus, making Val7 the new N-terminal residue, is tolerated; similarly, truncation of four residues from the C-terminus, making Gly107 the new terminal residue, is tolerated. More severe truncation at the N-terminus result in a sudden loss of fluorescence, while more severe truncations at the C-terminus lead to a gradual decrease in fluorescence. **c)** Successful terminal truncations in the context of iLOV's structure. Successful N-terminal truncations are shown in red and successful C-terminal truncations are shown in orange. Each truncation avoids disturbing or minimally disturbs secondary structure elements.

4.4.2. Structure-guided deletion of Dα helix

Building on the finding that the termini of iLOV may be shortened without major loss of fluorescence, we sought out additional structural elements that may be removed or minimized. The crystal structure of iLOV indicates that the Dα helix has a greater B-value than other secondary structure components **(Figure 4.2a).** The greater fluctuation of this helix suggests it may not be essential to function of iLOV as a fluorescent protein, but rather could be a vestige of LOV's native role as a phototropin. Further, Dα's closest approach to the flavin cofactor is 10Å, suggesting it is unlikely to participate in flavin binding.

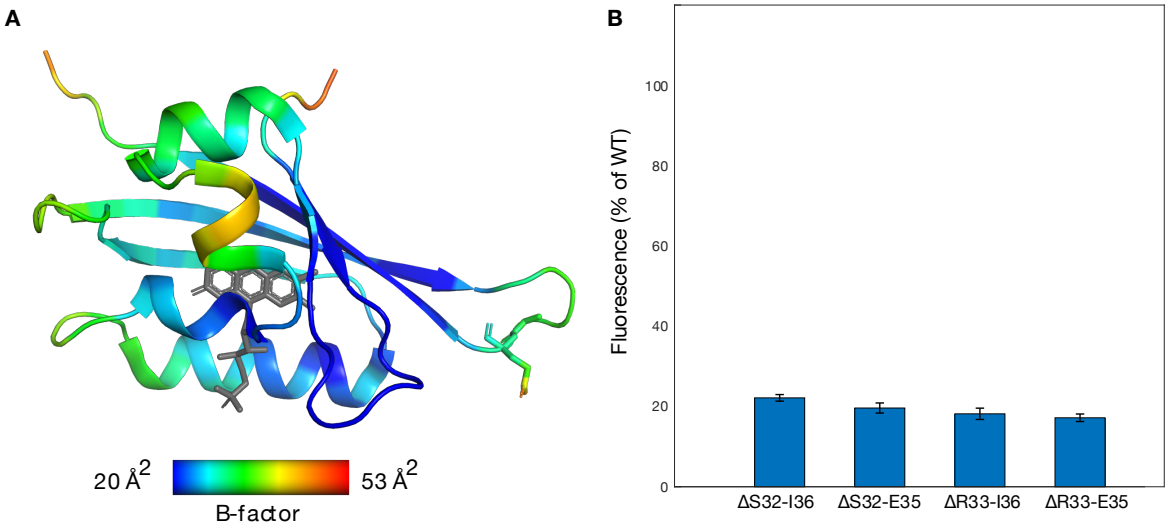

Figure 4.2. Deletion of Dα helix. a) iLOV structure colored by the mean B-factor of each residue. The Dα helix, shown in the foreground, has a higher B-factor than surrounding residues. **b)** Deletion of stretches of the Dα helix all reduce fluorescence to about 20% of wild type.

Dα consists of six amino acids, Ser32-Arg33-Glu34-Glu35-Ile36-Leu37. We constructed deletions of this helix of varying lengths: deleting Ser32 through Ile36, Ser32 through Glu35, Arg33 through Ile36, and Arg33 through Glu35. However, none of these

deletions resulted in fluorescence similar to wild type iLOV **(Figure 4.2b),** so we concluded that Dα is required for iLOV fluorescence.

4.4.3. Deletion of single small hydrophobic residues

As deletion of the Dα helix was not tolerated, we then sought to identify individual residues outside of secondary structural components (i.e. loops) that could be deleted. We focused our search on small, hydrophobic amino acids, namely glycine and proline, as deletion of hydrophilic amino acids from loops may confer a solubility penalty to the protein, and larger, more complex sidechains are more likely to be necessary for folding or function. We identified eight such amino acids as candidates for deletion: Pro11, Pro14, Pro17, Gly38, Gly46, Pro47, Gly78, and Gly95 **(Figure 4.3a).** Deletion of a majority of these candidates led to near total loss of fluorescence. Deletion of Pro14 severely diminishes fluorescent but does not completely eliminate it. However, deletion of Gly95 retains near wild type fluorescence intensity. Interestingly, this deletion parallels our finding that iLOV can be circularly permuted about Gly95 and Glu96 **(Figure 4.3b).** This further highlights findings from our circular permutation work that the Hβ-Iβ loop is particularly pliable for engineering.

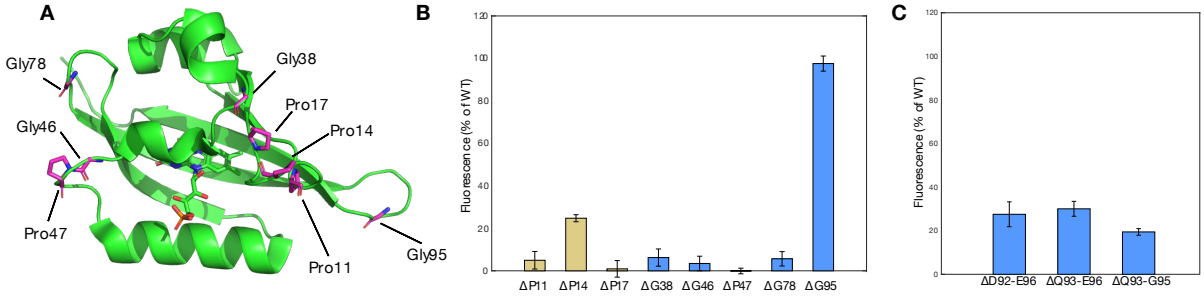

Figure 4.3. Deletion of single amino acids. a) Locations of candidate amino acids for deletion. Prolines and glycines in loops are targeted. **b)** Deletion of Pro14 retains 24.7 ± 1.7% of wild type fluorescence, deletion of Gly95 retains 97.5 ± 3.6% of wild type fluorescence, and each other deletion eliminates nearly all fluorescence. **c)** Deletion of larger stretches of the Hβ-Iβ loop result in diminished fluorescence relative to wild type.

4.4.4. Shortening the Hβ-Iβ loop

Having found that Gly95 within the Hβ-Iβ loop can be deleted while retaining high fluorescence, we sought to further shorten this loop. Similar to our strategy for deleting the Dα helix, we removed stretches of amino acids of varying lengths and positions: Asp92 through Glu96, Gln93 through Glu96, and Gln93 through Gly95. Each of these deletions resulted in a substantial loss of fluorescence **(Figure 4.3c).** Additional investigation via individual deletion of residues in this loop is warranted. We speculate that though deletion of extended stretches of amino acids failed to maintain fluorescence, less severe deletions in this region may be successful, as multiple forms of evidence have indicated its permissibility to manipulation.

4.4.5. Combination of terminal and internal deletions

We have found that, individually, iLOV tolerates deletion of five residues from the N-terminus, deletion of four residues from the C-terminus, and deletion of Gly95. To create the smallest possible protein, we combined each of these three individual deletions in a single protein, removing a total of ten amino acids from iLOV. The combined deletions resulted in high fluorescence, 86.7 ± 2.2% of wildtype iLOV fluorescence **(Figure 4.4).** We have termed this protein "NanoLOV," to date the smallest flavin-based fluorescent protein.

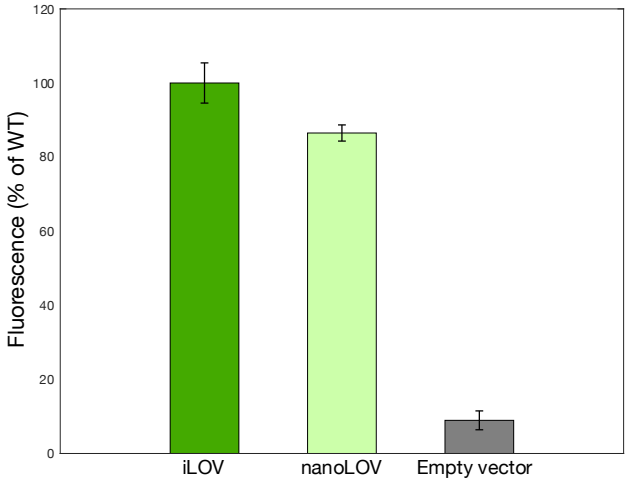

Figure 4.4. NanoLOV retains 86.7 ± 2.2% the fluorescence of iLOV while being ten amino acids shorter.

4.5. Conclusions

Through rational structure-based deletion of amino acids from iLOV, we have generated an FbFP that is ten amino acids smaller than iLOV. At 101 amino acids and 11.8 kDa, nanoLOV is the smallest FbFP by length to date. Notably, due to varying masses of amino acids, CagFbFP, though longer at 107 amino acids, has a slightly smaller mass at 11.6 kDa.[9] Despite this, nanoLOV's smaller genetic footprint serves as an inherent advantage over CagFbFP.

Though there is conceptual similarity between minimization of iLOV and minimization of FAST (which resulted in NanoFAST), the challenges are distinct. NanoFAST was produced from FAST simply by truncating a 26 amino acid domain from the N-terminus. In contrast, critical structural elements exist near the termini of iLOV, allowing elimination of only four to five amino acids from each terminus. Therefore, searching for internal amino acids that may be eliminated is necessary to further reduce its size. In this study, we were only able to identify one internal deletion, Gly95.

We believe there may be additional amino acids which could be deleted from nanoLOV to further decrease its size. We have preliminarily tested TRIAD,[18] a transposon-based method for randomly deleting single codons from a gene; however, in our attempt most clones contained either additional DNA inserted, larger than expected deletions, or frameshift mutations. Further refinement of our implementation of TRIAD is necessary for successful application of this technique. Expansion of throughput by coupling TRIAD with fluorescence-activated cell sorting and next-generation sequencing could better survey the sequence space and enable directed evolution.

As an alternative to TRIAD, individual deletion of each amino acid remaining in nanoLOV would conclusively determine which amino acids could be deleted. This approach could be supported by a high-throughput robot-enabled cloning system, known as a biofoundry. With such a resource, straightforward cloning and screening of order 100 constructs is possible. If either of these approaches identifies three or more additional amino acids that may be deleted, NanoFAST's mark of 98 amino acids could be matched or surpassed, resulting in, to our knowledge, the smallest fluorescent protein.

We plan to soon demonstrate the broad utility of nanoLOV by demonstrating its expression in mammalian cells. Upon expression in mammalian cells, we will measure nanoLOV's photobleaching rate relative to iLOV. Further, we plan to fuse nanoLOV to various subcellular localization sequences and observe their proper localization via confocal microscopy. We anticipate that this will enable investigation of the trafficking of arbitrary proteins with a minimally small fluorescent tag, while avoiding need for externally introduced cofactors.

We also plan to further characterize nanoLOV through *in vitro* methods. We will purify nanoLOV through nickel-affinity chromatography; and determine fluorescence spectra, approximate quantum yield, oligomeric state, and melting temperature. This investigation will help determine if drawbacks are introduced by the deletion of amino acids from iLOV. We anticipate that nanoLOV's melting temperature will decrease from that of iLOV, resulting in somewhat less efficient folding and expression, leading to the slightly diminished brightness of iLOV. However, for applications where fluorophore size is severely limited, nanoLOV's small size may outweigh any challenges.

We are interested in pursuing a circular permutation of nanoLOV. As described in Chapter 3, cpiLOV is 44 amino acids longer than iLOV itself. This lengthening thus diminishes iLOV's inherent small size advantages. Could nanoLOV be circularly permuted through the same strategy used to circularly permute iLOV? Circular permutation of iLOV required addition of a flexible linker, thus it is unclear if a longer linker would be required with the terminal regions of iLOV truncated in nanoLOV. It could perhaps be more effective to minimize cpiLOV itself, such as by finding a shorter substitute for the Jα helix.

4.6. References

(1) Mukherjee, A.; Schroeder, C. M. Flavin-Based Fluorescent Proteins: Emerging Paradigms in Biological Imaging. *Current Opinion in Biotechnology* **2015**, *31*, 16–23. https://doi.org/10.1016/j.copbio.2014.07.010.

(2) Christie, J. M.; Hitomi, K.; Arvai, A. S.; Hartfield, K. A.; Mettlen, M.; Pratt, A. J.; Tainer, J. A.; Getzoff, E. D. Structural Tuning of the Fluorescent Protein iLOV for Improved Photostability. *J Biol Chem* **2012**, *287* (26), 22295–22304. https://doi.org/10.1074/jbc.M111.318881.

(3) Cormack, B. P.; Valdivia, R. H.; Falkow, S. FACS-Optimized Mutants of the Green Fluorescent Protein (GFP). *Gene* **1996**, *173* (1), 33–38. https://doi.org/10.1016/0378-1119(95)00685-0.

(4) Kumagai, A.; Ando, R.; Miyatake, H.; Greimel, P.; Kobayashi, T.; Hirabayashi, Y.; Shimogori, T.; Miyawaki, A. A Bilirubin-Inducible Fluorescent Protein from Eel Muscle. *Cell* **2013**, *153* (7), 1602–1611. https://doi.org/10.1016/j.cell.2013.05.038.

(5) Plamont, M.-A.; Billon-Denis, E.; Maurin, S.; Gauron, C.; Pimenta, F. M.; Specht, C. G.; Shi, J.; Quérard, J.; Pan, B.; Rossignol, J.; Moncoq, K.; Morellet, N.; Volovitch, M.; Lescop, E.; Chen, Y.; Triller, A.; Vriz, S.; Le Saux, T.; Jullien, L.; Gautier, A. Small Fluorescence-Activating and Absorption-Shifting Tag for Tunable Protein Imaging in Vivo. *Proceedings of the National Academy of Sciences* **2016**, *113* (3), 497–502. https://doi.org/10.1073/pnas.1513094113.

(6) S. Mineev, K.; A. Goncharuk, S.; V. Goncharuk, M.; V. Povarova, N.; I. Sokolov, A.; S. Baleeva, N.; Yu. Smirnov, A.; N. Myasnyanko, I.; A. Ruchkin, D.; Bukhdruker, S.; Remeeva, A.; Mishin, A.; Borshchevskiy, V.; Gordeliy, V.; S. Arseniev, A.; A. Gorbachev, D.; S. Gavrikov, A.; S. Mishin, A.; S. Baranov, M. NanoFAST: Structure-Based Design of a Small Fluorogen-Activating Protein with Only 98 Amino Acids. *Chemical Science* **2021**, *12* (19), 6719–6725. https://doi.org/10.1039/D1SC01454D.

(7) England, C. G.; Ehlerding, E. B.; Cai, W. B. NanoLuc: A Small Luciferase Is Brightening Up the Field of Bioluminescence. *Bioconjugate Chemistry* **2016**, *27* (5), 1175–1187. https://doi.org/10.1021/acs.bioconjchem.6b00112.

(8) Mukherjee, A.; Weyant, K. B.; Agrawal, U.; Walker, J.; Cann, I. K. O.; Schroeder, C. M. Engineering and Characterization of New LOV-Based Fluorescent Proteins from

Chlamydomonas Reinhardtii and Vaucheria Frigida. *ACS Synthetic Biology* **2015**, *4* (4), 371–377. https://doi.org/10.1021/sb500237x.

(9) Nazarenko, V. V.; Remeeva, A.; Yudenko, A.; Kovalev, K.; Dubenko, A.; Goncharov, I. M.; Kuzmichev, P.; Rogachev, A. V.; Buslaev, P.; Borshchevskiy, V.; Mishin, A.; Dhoke, G. V.; Schwaneberg, U.; Davari, M. D.; Jaeger, K.-E.; Krauss, U.; Gordeliy, V.; Gushchin, I. A Thermostable Flavin-Based Fluorescent Protein from Chloroflexus Aggregans: A Framework for Ultra-High Resolution Structural Studies. *Photochemical & Photobiological Sciences* **2019**. https://doi.org/10.1039/C9PP00067D.

(10) Drepper, T.; Eggert, T.; Circolone, F.; Heck, A.; Krauss, U.; Guterl, J. K.; Wendorff, M.; Losi, A.; Gartner, W.; Jaeger, K. E. Reporter Proteins for in Vivo Fluorescence without Oxygen. *Nature Biotechnology* **2007**, *25* (4), 443–445. https://doi.org/10.1038/nbt1293.

(11) Gelman, H.; Wirth, A. J.; Gruebele, M. ReAsH as a Quantitative Probe of In-Cell Protein Dynamics. *Biochemistry* **2016**, *55* (13), 1968–1976. https://doi.org/10.1021/acs.biochem.5b01336.

(12) Lemos, B.; Bettencourt, B. R.; Meiklejohn, C. D.; Hartl, D. L. Evolution of Proteins and Gene Expression Levels Are Coupled in Drosophila and Are Independently Associated with mRNA Abundance, Protein Length, and Number of Protein-Protein Interactions. *Molecular Biology and Evolution* **2005**, *22* (5), 1345–1354. https://doi.org/10.1093/molbev/msi122.

(13) Hoffmann, C.; Gaietta, G.; Bünemann, M.; Adams, S. R.; Oberdorff-Maass, S.; Behr, B.; Vilardaga, J.-P.; Tsien, R. Y.; Ellisman, M. H.; Lohse, M. J. A FlAsH-Based FRET Approach to Determine G Protein–Coupled Receptor Activation in Living Cells. *Nat Methods* **2005**, *2* (3), 171–176. https://doi.org/10.1038/nmeth742.

(14) Sokolovski, M.; Bhattacherjee, A.; Kessler, N.; Levy, Y.; Horovitz, A. Thermodynamic Protein Destabilization by GFP Tagging: A Case of Interdomain Allostery. *Biophysical Journal* **2015**, *109* (6), 1157–1162. https://doi.org/10.1016/j.bpj.2015.04.032.

(15) Kumar, M.; Keller, B.; Makalou, N.; Sutton, R. E. Systematic Determination of the Packaging Limit of Lentiviral Vectors. *Human Gene Therapy* **2001**, *12* (15), 1893–1905. https://doi.org/10.1089/104303401753153947.

(16) Sweeney, N. P.; Vink, C. A. The Impact of Lentiviral Vector Genome Size and Producer Cell Genomic to Gag-Pol mRNA Ratios on Packaging Efficiency and Titre. *Molecular Therapy - Methods & Clinical Development* **2021**, *21*, 574–584. https://doi.org/10.1016/j.omtm.2021.04.007.

(17) Wu, Z.; Yang, H.; Colosi, P. Effect of Genome Size on AAV Vector Packaging. *Molecular Therapy* **2010**, *18* (1), 80–86. https://doi.org/10.1038/mt.2009.255.

(18) Emond, S.; Petek, M.; Kay, E. J.; Heames, B.; Devenish, S. R. A.; Tokuriki, N.; Hollfelder, F. Accessing Unexplored Regions of Sequence Space in Directed Enzyme Evolution via Insertion/Deletion Mutagenesis. *Nature Communications* **2020**, *11* (1), 3469. https://doi.org/10.1038/s41467-020-17061-3.

Chapter 5. Conclusions

This chapter summarizes the advances this work has contributed toward the advancement of flavin-based fluorescent proteins, and discusses further development and fundamental scientific questions raised by the developed technologies and the lessons learned in developing them.

5.1. Advancement of flavin-based fluorescent proteins

The ultimate goal of this dissertation research program was to adapt existing flavin-based fluorescent proteins (FbFPs) for more effective and advanced bioimaging applications. Throughout this work, we have sought to preserve the intrinsic strengths of FbFPs – namely small size and independence from oxygen or other exogeneous cofactors,[1] while mitigating or eliminating their weaknesses. We have made significant progress in understanding FbFPs' relationship with their flavin chromophore, enabling advanced biosensor construction via circular permutation, increasing cellular brightness, and further minimizing the size of FbFPs. However, major challenges remain. Particularly, FbFPs lack a varied color palette, allosteric biosensors, a strong biophysical understanding of factors impacting their fluorescence spectra and yield, and widespread usage across varied organisms and contexts.[2]

We have developed a generalized method for preparing FbFPs without their flavin cofactor in such a manner that they spontaneously can reassociate with flavins to restore fluorescence. Unsupported denaturation and buffer exchange of an FbFP results in an insoluble aggregate that cannot be fluorescently reconstituted. However, binding an FbFP to a nickel resin followed by chemical denaturation, washing, refolding, and elution results in a soluble, non-fluorescent apo FbFP. Addition of flavin mononucleotide (FMN) restores FbFP fluorescence, with excitation and emission spectra that match those of FbFPs that did not undergo deflavination. This held for both of the FbFPs we attempted to deflavinate, iLOV and EcFbFP.

Our preparation of apo FbFPs allowed us to characterize the binding interaction between FbFPs and their flavin cofactors. We observed dissociation constants (K_D) for iLOV and EcFbFP and their native FMN cofactor on the order of 100nM. Considering typical

intracellular flavin concentrations are on the order of 1-10μM,[3] nearly all FbFPs expressed in cells should be flavin-bound.

We were also able to investigate binding interactions between FbFPs and other flavins, namely riboflavin (RF) and flavin adenine dinucleotide (FAD). Interestingly, both iLOV and EcFbFP showed slightly greater affinity for FAD than for FMN. As FAD is somewhat more prevalent in cells than FMN,[4] this raises an interesting question: Are a large portion FbFPs in their cellular context bound to FAD rather than FMN? Purification of FbFPs followed by liquid chromatography and mass spectrometry is needed to elucidate the relative abundance of FMN and FAD binding.

Each FbFP showed less affinity for RF than for FMN or FAD, with dissociation constants of 890nM and 760nM. Riboflavin differs from FMN and FAD by the group on the end of their sugar chain: FMN has a phosphate group, FAD has an adenine dinucleotide (which includes two phosphate groups), while RF only has a hydroxyl group. This suggests that the phosphate group interacts with residues on the edge of the binding pocket, thereby increasing affinity. Understanding this interaction can guide selection of non-native cofactors for further engineering of FbFPs.

We have successfully engineered a circular permutation of iLOV, one of the first such permutations of FbFPs and the first extensively engineered for application itself. While a circular permutation of MiniSOG has been created,[5] it was used chiefly as a means to engineer a split MiniSOG, and useful application of the circular permutation has not been demonstrated. We embarked upon a campaign to convert a naïve, dimly fluorescent circularly permuted iLOV into a bright, highly useful fluorescent protein. We elucidated three major principles that enhance circularly permuted iLOV's fluorescence: (1) a helical connection between iLOV's

original termini, such as the Jα helix, is necessary, (2) an additional flexible linker is needed in addition to the Jα helix and optimization of its length influences fluorescence, and (3) addition of dimerizing domains to the new termini of cpiLOV greatly increases fluorescence. We are curious if these design rules will hold to enable circular permutation of additional FbFPs – and optimization of these additional circularly permuted FbFPs may help uncover the biophysical phenomena that underlie these design principles.

Interestingly, addition of dimerizing domains (namely, leucine zippers) to the termini of unpermuted iLOV results in a boost to its fluorescence. The biophysical mechanism of this increase requires additional investigation, but gives rise to interesting hypotheses. The coiled-coil domains are predicted to tightly hold the termini of both iLOV and cpiLOV close to one another. We speculate that the entropic penalty that would be imposed by separating the coiled coils prevents the protein from visiting conformations in which water can access the flavin chromophore. As FbFPs fluoresce by modulating the intrinsic fluorescence of flavin,[6] both blue shifting it and increasing intensity relative to flavin in an aqueous solution. Thus, limiting solvent access to the bound flavin may cause FbFP-bound behavior to dominate over aqueous-like behavior.

We have utilized our zipped, circularly permuted iLOV (cpiLOV-Lz) to develop sensors for protease activity. Simple replacement of the flexible linker of cpiLOV-Lz with a protease cleavage sequence results in protease-induced loss of fluorescence. We have developed two such sensors, one for tobacco etch virus protease and one for SARS-CoV-2 main protease. These designs serve as a template to construct sensors for other proteases. Given the centrality of proteases to viral reproduction,[7] we envision that this design paradigm could enable rapid development of a sensor for activity of proteases implicated in propagation of an

emergent pathogen. Such a sensor could then enable a high-throughput screen of protease inhibitor candidates toward identification of an antiviral drug.

We have also developed a minimized version of iLOV which we have termed "nanoLOV." NanoLOV eliminates ten amino acids from iLOV, bringing its length down to 101 amino acids, while retaining 86.7 ± 2.2% of iLOV's fluorescence. This makes nanoLOV one of the smallest known fluorescent proteins, to our knowledge exceeded only by NanoFAST (at 98 amino acids).[8] Notably, nanoLOV enjoys a distinctive advantage over NanoFAST because it does not require introduction of an outside cofactor, allowing it to act completely autonomously. Meanwhile, nanoLOV enjoys an advantage over other FbFPs and GFP-derived FPs due to its decreased size, enabling less perturbative tagging of fusion proteins and better accommodating complex gene circuits when genetic size is severely limited.

Our development of cpiLOV and nanoLOV highlights an important parallel: Gly95 appears to be a particularly pliable site for engineering iLOV. The site we successfully circularly permuted iLOV about was also the only internal residue we were able to successfully delete from iLOV in construction of nanoLOV. This apparent pliability suggests that future manipulations of iLOV will be most successful if targeted to the Hβ-Iβ loop in which Gly95 resides. These manipulations may include development of a split iLOV, where two independently non-fluorescent fragments of iLOV can be fluorescently reconstituted; or creation of an allosteric sensor by insertion of a metabolite binding domain at this site, either in iLOV or at the termini of cpiLOV. Careful investigation of other loops of iLOV may help identify other such pliable sites.

4.2. Further investigation enabled by this work

A major outstanding shortcoming of FbFPs is their lack of a varied color palette. GFP, dsRed, FAST, and other classes of fluorescent proteins have a diverse selection of colors, allowing them to be multiplexed or adapted to systems with limitations on certain wavelengths. We see replacement of the native flavin cofactor with a modified flavin as a promising avenue toward a varied color palette. Indeed, some progress has been made in this direction, by incorporating roseoflavin into the FbFP Slr1694,[9] resulting in red shifted fluorescence, though the quantum yield is severely diminished. Our development of a method to reversibly remove flavins from FbFPs *in vitro* may enable screening of large libraries of modified flavins for FbFP binding and fluorescence without the obstacles of cellular transport, biosynthesis, toxicity, or competition from native flavins. If such an approach identifies a promising binding partner for an FbFP, these obstacles can then be mitigated as needed toward developing a bright, autonomous, non-green FbFP.

The Jα helix included in cpiLOV raises an interesting biophysical issue: in its native context, while Jα remains a helix in the dark, it unravels when exposed to blue light.[10] Therefore, it is possible that cpiLOV and sensors based on cpiLOV could alter their behavior upon illumination due to changes in the Jα helix. Potentially, Jα could be partly responsible for the diminished fluorescence of cpiLOV relative to iLOV by diverting some energy that would otherwise be emitted as a fluorescent photon into a conformational change in Jα. We suggest further study of illumination-related behavior of Jα in the context of cpiLOV through several avenues: (1) determining if light-based unfolding of Jα occurs in cpiLOV via electron paramagnetic resonance (EPR), (2) detecting changes in the behavior of Jα between cleaved and uncleaved states of our protease sensors, and (3) replacing Jα with an inert, non-light-

reactive helix of similar length. These biophysical studies could greatly expand our understanding of the mechanisms affecting FbFP function.

In the cases of both iLOV and cpiLOV, addition of dimerizing domains to termini dramatically increased fluorescence, more than doubling the fluorescence absent dimerizing domains. While this increased fluorescence provides a major boon to utility of iLOV, the mechanism underlying it remains unknown. At the most basic level, *in vitro* characterization of zipped iLOV and cpiLOV is necessary to distinguish whether the protein is intrinsically more fluorescent or is just more highly expressed with zippers. We can also utilize deflavination and melting curves to determine dimerizing domains' impact on flavin affinity and thermostability, both of which we anticipate will be bolstered. Furthermore, we suggest utilizing EPR and X-ray crystallography to determine if the coiled coils fold back upon iLOV, potentially having an impact on fluorescence beyond just stabilizing the termini.

A significant complication in the use of FbFPs is their propensity for rapid photobleaching,[11] which can occur on the scale of 30 seconds under normal microscopy conditions due to high intensity of excitation lamps or lasers. This can result in false negative results, as the time a researcher needs to focus the microscope and identify a region of interest can result in a total loss of fluorescence prior to image capture. As such, with a rapidly photobleaching fluorophore, researchers may need to reduce excitation power or exposure time, both of which decrease signal to noise ratio. We intend to evaluate if any of the manipulations we have made to FbFPs impacts their photobleaching rate, either positively or negatively.

While there is not great utility in sensing activity of TEV protease itself, we believe our TEV protease sensor can be adapted to a broad range of phenomena through the use of a split

TEV protease. In short, TEV protease has been split into two pieces which individually are catalytically incompetent, but if the pieces are reconstituted by other protein-protein interactions, TEV protease can cleave its substrate.[12] Therefore, any protein-protein interaction that can cause reconstitution of split TEV protease should be able to be detected with our TEV protease sensor. Conditionally dimerizing peptides, such as FKBP12/FRB (which dimerize in the presence of rapamycin)[13] or calmodulin/RS20 (which dimerize in the presence of calcium)[14] could be used as proof of concept for such "indirect" sensors.

Throughout this work, we have maintained a vision of tool-building with FbFPs. The greatest impact of GFP has come from the tools built upon its foundation; as such we seek to begin a library of allosteric FbFP-based biosensors. These sensors would involve fusion of a binding domain to an FbFP in a manner that would modify FbFP fluorescence depending on analyte binding. Allosteric sensors are particularly attractive due to their reversibility and the rapid timescale of their response –signal can change nearly as soon as the analyte binds, avoiding delays involving transcription, translation, diffusion, or degradation. These sensors could take the form of a whole or split binding domain, translationally fused to either an FbFP or circularly permuted FbFP. As such, cpiLOV allows access to a greater design space for these sensors. Calcium is a prime target for a first cpiLOV-based allosteric sensor, just as calcium was the target of cpGFP-based allosteric sensors.[15]

We hope to develop further synergy between cpiLOV and nanoLOV. These two projects have had opposite impacts on the size of iLOV: nanoLOV reduced its size by 10 amino acids, while cpiLOV increased its size by 45 amino acids (or by 104 amino acids for cpiLOV-Lz). This raises the complementary questions (1) Can nanoLOV be circularly permuted? and

(2) Can cpiLOV be minimized. We hope such a pursuit will prevent researchers from needing to trade off size against versatility in selecting an FbFP.

We desire to bring our work on nanoLOV to its natural conclusion through further characterization and demonstration. We are in the process of purifying nanoLOV, which can then be used to determine fluorescence spectra, quantum yield, melting temperature, oligomeric state, flavin affinity, and other critical parameters. Similarly to zipped iLOV and cpiLOV, it is an outstanding question whether varying fluorescence is due to intrinsic fluorescence of the protein or altered expression. We plan to demonstrate that nanoLOV can be expressed in mammalian cells, first targeting CHO cells as we have with cpiLOV. Upon accomplishment of that objective, we also plan to demonstrate that nanoLOV can be used to visualize the localization of proteins by tagging nanoLOV with various localization sequences. We anticipate there may be difficulties caused by rapid photobleaching coupled with the high laser power needed for confocal microscopy, so we may explore use of total internal reflection microscopy to decrease the amount of excitation light incident on the sample.

The experimentation we report here was performed under normoxic conditions, though one particular advantage of FbFPs is their ability to function in the absence of oxygen. Though it has been shown that FbFPs function similarly under normoxic and hypoxic conditions,[16] we hope to complete the narrative of developing advanced oxygen-independent fluorescent proteins by demonstrating their use in hypoxic contexts. Especially enticing is completing development of a sensor for caspase activity, then demonstrating its use in a model of the hypoxic core of a solid tumor,[17] in which caspase disorders often lead to malignancy.[18]

In conclusion, we have developed several technologies that have advanced the usefulness of FbFPs. In developing these technologies, we have gained important biophysical

insights into the function of FbFPs and the "knobs" that can be turned to manipulate the qualities of FbFPs. We hope that these technologies will find widespread use in molecular biology and enable interrogation of previously inaccessible tissues, systems, and phenomena. Further, we hope that our tools will form the basis for a toolbox of FbFP-based sensors developed by others.

5.3. References

(1) Ozbakir, H. F.; Anderson, N. T.; Fan, K.-C.; Mukherjee, A. Beyond the Green Fluorescent Protein: Biomolecular Reporters for Anaerobic and Deep-Tissue Imaging. *Bioconjugate Chemistry* **2020**, *31* (2), 293–302. https://doi.org/10.1021/acs.bioconjchem.9b00688.

(2) Mukherjee, A.; Schroeder, C. M. Flavin-Based Fluorescent Proteins: Emerging Paradigms in Biological Imaging. *Current Opinion in Biotechnology* **2015**, *31*, 16–23. https://doi.org/10.1016/j.copbio.2014.07.010.

(3) Hühner, J.; Ingles-Prieto, Á.; Neusüß, C.; Lämmerhofer, M.; Janovjak, H. Quantification of Riboflavin, Flavin Mononucleotide, and Flavin Adenine Dinucleotide in Mammalian Model Cells by CE with LED-Induced Fluorescence Detection. *ELECTROPHORESIS* **2015**, *36* (4), 518–525. https://doi.org/10.1002/elps.201400451.

(4) Macheroux, P.; Kappes, B.; Ealick, S. E. Flavogenomics – a Genomic and Structural View of Flavin-Dependent Proteins. *The FEBS Journal* **2011**, *278* (15), 2625–2634. https://doi.org/10.1111/j.1742-4658.2011.08202.x.

(5) Boassa, D.; Lemieux, S. P.; Lev-Ram, V.; Hu, J.; Xiong, Q.; Phan, S.; Mackey, M.; Ramachandra, R.; Peace, R. E.; Adams, S. R.; Ellisman, M. H.; Ngo, J. T. Split-miniSOG for Spatially Detecting Intracellular Protein-Protein Interactions by Correlated Light and

Electron Microscopy. *Cell Chemical Biology* **2019**, *26* (10), 1407-1416.e5. https://doi.org/10.1016/j.chembiol.2019.07.007.

(6) Mukherjee, A.; Walker, J.; Weyant, K. B.; Schroeder, C. M. Characterization of Flavin-Based Fluorescent Proteins: An Emerging Class of Fluorescent Reporters. *PLoS ONE* **2013**, *8* (5), e64753. https://doi.org/10.1371/journal.pone.0064753.

(7) STRAUSS, J. H.; STRAUSS, E. G. Overview of Viruses and Virus Infection. *Viruses and Human Disease* **2008**, 1–33. https://doi.org/10.1016/B978-0-12-373741-0.50004-0.

(8) S. Mineev, K.; A. Goncharuk, S.; V. Goncharuk, M.; V. Povarova, N.; I. Sokolov, A.; S. Baleeva, N.; Yu. Smirnov, A.; N. Myasnyanko, I.; A. Ruchkin, D.; Bukhdruker, S.; Remeeva, A.; Mishin, A.; Borshchevskiy, V.; Gordeliy, V.; S. Arseniev, A.; A. Gorbachev, D.; S. Gavrikov, A.; S. Mishin, A.; S. Baranov, M. NanoFAST: Structure-Based Design of a Small Fluorogen-Activating Protein with Only 98 Amino Acids. *Chemical Science* **2021**, *12* (19), 6719–6725. https://doi.org/10.1039/D1SC01454D.

(9) Zirak, P.; Penzkofer, A.; Mathes, T.; Hegemann, P. Absorption and Emission Spectroscopic Characterization of BLUF Protein Slr1694 from Synechocystis Sp. PCC6803 with Roseoflavin Cofactor. *Journal of Photochemistry and Photobiology B: Biology* **2009**, *97* (2), 61–70. https://doi.org/10.1016/j.jphotobiol.2009.08.002.

(10) Harper, S. M.; Neil, L. C.; Gardner, K. H. Structural Basis of a Phototropin Light Switch. *Science* **2003**, *301* (5639), 1541–1544. https://doi.org/10.1126/science.1086810.

(11) Buckley, A. M.; Petersen, J.; Roe, A. J.; Douce, G. R.; Christie, J. M. LOV-Based Reporters for Fluorescence Imaging. *Current Opinion in Chemical Biology* **2015**, *27*, 39–45. https://doi.org/10.1016/j.cbpa.2015.05.011.

(12) Wehr, M. C.; Laage, R.; Bolz, U.; Fischer, T. M.; Grünewald, S.; Scheek, S.; Bach, A.; Nave, K.-A.; Rossner, M. J. Monitoring Regulated Protein-Protein Interactions Using Split TEV. *Nat Methods* **2006**, *3* (12), 985–993. https://doi.org/10.1038/nmeth967.

(13) Banaszynski, L. A.; Liu, C. W.; Wandless, T. J. Characterization of the FKBP·Rapamycin·FRB Ternary Complex. *J. Am. Chem. Soc.* **2005**, *127* (13), 4715–4721. https://doi.org/10.1021/ja043277y.

(14) Hill, T. J.; Lafitte, D.; Wallace, J. I.; Cooper, H. J.; Tsvetkov, P. O.; Derrick, P. J. Calmodulin−Peptide Interactions: Apocalmodulin Binding to the Myosin Light Chain Kinase Target-Site. *Biochemistry* **2000**, *39* (24), 7284–7290. https://doi.org/10.1021/bi000139m.

(15) Nagai, T.; Sawano, A.; Park, E. S.; Miyawaki, A. Circularly Permuted Green Fluorescent Proteins Engineered to Sense Ca2+. *Proceedings of the National Academy of Sciences* **2001**, *98* (6), 3197–3202. https://doi.org/10.1073/pnas.051636098.

(16) Walter, J.; Hausmann, S.; Drepper, T.; Puls, M.; Eggert, T.; Dihné, M. Flavin Mononucleotide-Based Fluorescent Proteins Function in Mammalian Cells without Oxygen Requirement. *PLOS ONE* **2012**, *7* (9), e43921.

(17) Emami Nejad, A.; Najafgholian, S.; Rostami, A.; Sistani, A.; Shojaeifar, S.; Esparvarinha, M.; Nedaeinia, R.; Haghjooy Javanmard, S.; Taherian, M.; Ahmadlou, M.; Salehi, R.; Sadeghi, B.; Manian, M. The Role of Hypoxia in the Tumor Microenvironment and Development of Cancer Stem Cell: A Novel Approach to Developing Treatment. *Cancer Cell International* **2021**, *21* (1), 62. https://doi.org/10.1186/s12935-020-01719-5.

(18) Olsson, M.; Zhivotovsky, B. Caspases and Cancer. *Cell Death Differ* **2011**, *18* (9), 1441–1449. https://doi.org/10.1038/cdd.2011.30.