Development of screening systems for enhanced fluorescent protein engineering

by

Matthew D Wiens

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry
University of Alberta

Abstract

Fluorescent proteins (FPs) have dramatically advanced life science research. Since their discovery they have become invaluable tools for imaging living systems and have been developed into precision instruments for measuring normally invisible events such as fluctuations in calcium ion concentration, protein-protein interactions (PPi), and even membrane voltage. More surprisingly, FPs have now been converted into optogenetic actuators capable of manipulating the biochemistry of the cell. There is substantial interest in improving both the diversity and quality of available FPs because of their great utility and potential. In this thesis, I describe my efforts to design better methods to easily improve FPs and I use these methods to create a variety of new FP variants.

First, I explored the potential of developing FPs as tandem dimers. I used directed evolution to create a series of heterodimeric FPs called the vine Tomatoes (vTs). Specifically, I created green-green (GGvT), green-red (GRvT), and red-red (RRvT) heterodimers by genetically fusing two tightly dimerizing FP domains and then evolving as a pair. This allows the two monomers to differentiate, creating tandem heterodimers with advanced characteristics such as exceptionally high FRET efficiency in GRvT of 99%, and the brightest red fluorescent protein to date with RRvT at 120.

Next I developed a robot-assisted screening system for photostability screening. This simplified screening process led to the development of Citrine2, a variant with 9-fold improvement in photostability relative to its precursor, mCitrine. I also observed that concentration plays a significant role in photostability, and I so I attempted to modify this property. With only five mutations, the concentration dependence of photostability switched from being an inverse relationship in mCitrine, to a direct relationship in Citrine2.

From there I switched to developing new screening systems for photocleavable proteins. Several different screening systems based on bimolecular fluorescence complementation and FRET were developed to screen directly for the photoinduced dissociation of a photocleavable protein (PhoCl) developed in our lab.

A second photocleavable protein called SplitOr was created from PSmOrange2. The goal was to develop a spectrally orthogonal photocleavable protein that could be photocleaved with wavelengths of light that do not cause PhoCl photocleavage. PSmOrange2 was circularly permuted and its fluorescence and photoconversion properties rescued, creating three new versions of SplitOr with photophysical characteristics that suggest photocleavage is occurring.

Altogether, these projects have advanced the field of FP development. I created three research-ready FPs, RRvT, GRvT, and Citrine2; I developed and characterized a rapid, low cost, robot-assisted illumination system which will aid our lab and others in the development of photostable FP variants; I advanced our knowledge of photocleavable proteins by developing an evolution system for the photocleavable protein PhoCl; and finally I created an orthogonal photocleavable protein that may find use as an optogenetic actuator.

Preface

Chapter 2 has been published by Matthew D. Wiens, Dr. Yi Shen, Xi Li, Dr. M. Alaraby Salem, Dr. Nick Smisdom, Wei Zhang, Dr. Alex Brown, and Dr. Robert E. Campbell, as "A Tandem Green--Red Heterodimeric Fluorescent Protein with High FRET Efficiency", in Chembiochem, 17, 4, 2361-2367 (2016). I performed the directed evolution of GdT, GGvT, GRvT, and finished RRvT, also called the vine Tomatoes. I also characterized the vine Tomatoes and wrote the manuscript. Dr. Yi Shen started the evolution of RRvT, Xi Li and Dr. Nick Smisdom assisted with data collection for smFRET, Dr. Mohammad Alaraby under the supervision of Dr. Alex Brown performed the FRET dipole orientation calculations, Wei Zhang assisted with mammalian cell imaging, Dr. Robert E. Campbell directed the research and assisted in preparing the manuscript.

Chapter 3 is being prepared for publication and was performed in collaboration with, Friederike Hoffmann, Dr. Yingche Chen and Dr. Robert. E. Campbell. Friederike. Hoffmann assisted in the early evolution of Citrine2, Dr. Yingche Chen performed the HeLa cell imaging, and Dr. Robert. E. Campbell directed the research.

Chapter 4 was performed in collaboration with Wei Zhang and contains data from the submitted paper. "Optogenetic control with a photocleavable protein" Wei Zhang, Alexander W. Lohman, Yevgeniya Zhuravlova, Xiaocen Lu, Matthew D. Wiens, Hiofan Hoi, Sine Yaganoglu, Manuel A. Mohr, Elena N. Kitova, John S. Klassen, Periklis Pantazis, Roger J. Thompson, and Robert E. Campbell, Nature Methods (2017). Wei Zhang helped perform the native PAGE screening of the PhoCl-sfYFP constructs and is the creator of PhoCl.

Acknowledgments

I would like to thanks the various organisations that helped make this possible, University of Alberta Molecular Biology Services Unit (MBSU) for DNA sequencing and technical assistance. This work was supported by grants from the Canadian Institutes for Health Research (CIHR) (MOP-123514) to R.E.C. and the Natural Sciences and Engineering Research Council of Canada (NSERC) (RGPIN 288338-2010) to R.E.C. and (RGPIN-2015-05341) to A.B. M.D.W. was supported in part by a Queen Elizabeth II scholarship from the University of Alberta. Y.S. was supported by an Alberta Innovates scholarship. Y. Chen received support from NIH (BRAIN Initiative).

I am very grateful for the efforts and support given by the faculty and staff at the University of Alberta. Specifically my supervisory committee members Dr. Juli Gibbs-Davis and Dr. Michael Serpe. My candidacy committee members Dr. Jonathan Veinot and Dr. Nils Petersen. I would also like to thank Dr. Klaus Ballanyi, Dr. Alexander Brown, Gareth Lambkin, Wayne Moffat, Xi Li, Dr. M. Alaraby Salem, and Dr. Nick Smisdom for their scientific and technical support.

I would like to thanks all members of the Campbell lab for their scientific, technical, and emotional support in making my time in the lab a wonderful experience. Special thanks go to Dr. Robert Campbell for fostering such a great community and supporting my scientific explorations.

Lastly, I would like to thanks my family and friends which overlap strongly with the above group for keeping me sane, grounded, and happy. Of course, thank you to my wife Nicole for walking beside me throughout this entire journey, and the countless ways you have made both my research and life better.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ϕ	Quantum yield
ε	Extinction coefficient
µL	Microliter
µm	Micrometer
µg	Microgram
3D	3 dimensional
A (Amino acid)	Alanine
A (Nucleobase)	Adenine
AA	Amino acid
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
avGFP	*Aequoria victoria* green fluorescent protein
BFP	Blue fluorescent protein
BiFC	Bimolecular fluorescence complementation
BiLC	Bimolecular luminescence complementation
BPER	Bacterial protein extraction reagent
BRET	Bioluminescence resonance energy transfer
C (Amino acid)	Cysteine
C (Nucleobase)	Cytosine
C. elegans	*Caenorhabditis elegans*
Cα	Alpha carbon
CALI	Chromophore assisted light inactivation

CaMPARI	Calcium modulated photoactivatable ratiometric integrator
CBR	Click beetle red luciferase
CIHR	Canadian Institutes for Health Research
cm	Centimeter
Co	Carbonyl carbon
D (Amino acid)	Aspartic acid
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
ddFPs	Dimerization dependent fluorescent proteins
ddGFP	Dimerization dependent green fluorescent protein
ddRFP	Dimerization dependent red fluorescent protein
ddYFP	Dimerization dependent yellow fluorescent protein
dGTP	Deoxyguanosine triphosphate
DMEM	Dulbecco's modified eagle media
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
DsRed	*Discosoma* red fluorescent protein
dT	Dimer tomato
dTTP	Deoxythymidine triphosphate
E (Amino acid)	Glutamic acid
E. coli	*Escherichia coli*
EB	Empty barrel
eBFP	Enhanced blue fluorescent protein

EDTA	Ethidium diamine tetra-acetic acid
eGFP	Enhanced green fluorescent protein
EMCCD	Electron Multiplying Charge Coupled Device
EP	Error prone
ESI	Electrospray ionization
ESPT	Excited state proton transfer
F (Amino acid)	Phenylalanine
FACS	Fluorescence activated cell sorting
FLIM	Fluorescence lifetime imaging microscopy
fmol	Femtomole
FP	Fluorescent protein
FPX	Fluorescent protein exchange
FRET	Förster resonance energy transfer
G (Amino acid)	Glycine
G (Nucleobase)	Guanine
GdT	Green dimer tomato
GECO	Genetically encoded Ca2+ indicators for optical imaging
GFP	Green fluorescent protein
GGvT	Green-green vine tomato
GRvT	Green-red vine tomato
H (Amino acid)	Histidine
HBSS	Hanks balanced salt solution
HCV	Hepatitis C virus

HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HHBSS	HEPES buffered Hanks balanced salt solution
Hsp	Heat shock protein
HYG	Histidine-tyrosine-glycine
I (Amino acid)	Isoleucine
IDT	Integrated DNA technologies
InPhO	Integrated photon output
IR	Infra-red
IRF	Instrument response function
ITYG	Isoleucine-threonine-tyrosine-glycine
K (Amino acid)	Lysine
k_{cat}	Catalytic rate constant
K_d	Dissociation constant
kDa	Kilo Dalton
L	Liter
L (Amino acid)	Leucine
LB	Lysogeny broth
LED	Light emitting diode
LP	Long pass
LSM	Laser scanning microscope
LSSFP	Long Stokes shift fluorescent protein
m	Meter
M	Molar

M (Amino acid)	Methionine
MBSU	Molecular biology services unit
mg	Milligram
min	Minute
mL	Milliliter
mm	Millimeter
MS	Mass spectroscopy
mTFP1	Monomeric teal fluorescent protein one
mW	Milliwatt
MWCO	Molecular weight cut-off
N (Amino acid)	Asparagine
NA	Numerical aperture
ND	Neutral density
NFT	Secondary dichroic mirror
nH	Hill coefficient
NLS	Nuclear localizing sequence
nm	Nanometer
nM	Nanomolar
ng	Nanogram
NSERC	Natural Sciences and Engineering Research Council of Canada
NTA	Nitrilotriacetic acid
P (Amino acid)	Proline

PACE	Phage-assisted continuous evolution
PAGE	Polyacrylamide gel electrophoresis
PA-GFP	Photoactivatable green fluorescent protein
PAUP	Phylogenetic analysis using parsimony
PCM	Polarizable continuum model
PCR	Polymerase chain reaction
PDB	Protein database
PhoCl	Photocleavable protein
POI	Protein of interest
PPi	Protein-protein interaction
PSmOrange2	Photoswitchable monomeric orange 2
Q (Amino acid)	Glutamine
QM/MM	Quantum mechanical/molecular modelling
R (Amino acid)	Arginine
RCF	Relative centrifugal force
RESOLFT	Reversible saturable optical linear fluorescence transitions
RFP	Red fluorescent protein
RNA	Ribonucleic acid
ROS	Reactive oxygen species
rpm	Rotations per minute
RRvT	Red-red vine tomato
S (Amino acid)	Serine
S. elongates	*Synechococcus elongatus*

SD	Standard deviation
SDS	Sodium dodecyl sulphate
sfGFP	Superfolder green fluorescent protein
sfYFP	Superfolder yellow fluorescent protein
smFRET	Single molecule Förster resonance energy transfer
SR	Steroid receptor
StEP	Staggered extension process
SYG	Serine-tyrosine-glycine
T (Amino acid)	Threonine
T (Nucleobase)	Thymine
Taq	*Thermus aquaticus* polymerase
TBS	Tris buffered saline
TD-DFT	Time-dependent density functional theory
tdTomato	Tandem dimer tomato
UV	Ultraviolet
V (Amino acid)	Valine
VMD	Visual molecular dynamics
W	Watt
W (Amino acid)	Tryptophan
Y (Amino acid)	Tyrosine
YFP	Yellow fluorescent protein

Introduction

1.1 Overview and premise

Understanding the internal functioning of living cells is an incredibly complicated endeavor. Cells function like modern cities. Each biomolecule is a person with a specific job to perform. There are roads of actin and tubulin, specialized communities walled in by membranes. For example, the power plant mitochondria and downtown core nucleus. There are also thousands of enzymatic people. Each enzyme person performs only a small number of jobs, with many using the same resources, creating competition or requiring cooperation, some are saboteurs, while others support their neighbors; together creating a complex economy. Each person is part of a complex supply chain. New people are constantly born in ribosomes, trained by chaperones, work, and are forcibly retired by proteasomes. All play a role in the intricate flow of goods and energy to keep a cell alive, but none can perform in isolation. Altogether, this chaotic dance of creating, destroying, associating, dissociating, activating, and inhibiting is coordinated and controlled without a single person aware.

Trying to understand the apparent chaos in the cell is not something that can be done in one step. First the individual jobs and their required biomolecules must be identified. DNA holding the blueprints, RNA relaying information, membranes creating separated compartments, and proteins sorting and modifying small molecules to extract energy and acquire more raw materials. Without control, supplies are wasted and energy squandered, which could ultimately result in the death of the cell. How does this chaotic living system provide the control that results in persistent, stable, and organized cells? The answer is in their interactions. This is called the interactome, the map of all interactions in a cell. How is this

map made? How are these protein-protein interactions (PPi) experimentally identified? We place GPS tracking tools that report back their precise location.

Many tracking tools exist such as fluorescent dyes and nanoparticles but the most versatile and reliable tool currently is the fluorescent protein (FP). Using genetic engineering techniques we can create a gene which tethers FPs to proteins of interest (POI) and introduce this gene into cells. The FP then becomes a signal that we can detect to report the location of the POI as it dances through the cell. In this thesis I will describe our efforts to improve our ability to visualize and manipulate interactions in living cells, through the genetic engineering of FPs. The rest of this introduction will provide an overview of the relevant background and context for the FP technologies presented here.

1.2 Optical imaging

In order to probe living systems, we need to be able to visualize objects that are on the cells' size scale, nanometers to micrometers. There are several technologies that can create images at this scale including electron microscopy, atomic force microscopy, and optical microscopy.[1] Electron microscopy has the highest resolution but is not practical for imaging living organisms; it is too damaging and there is too little contrast without adding contrast agents that are generally incompatible for use on living cells. Atomic force microscopy is limited to surfaces and requires physically touching the cells. Most optical techniques have resolutions limited to 200-500 nm depending on the wavelengths of light used. Photons in the visible and infrared spectrum penetrate through most living tissue causing little or no damage, allowing imaging of the surface and interior of many cells and tissues.[2,3] Despite the worse

resolution optical imaging is the most compatible with living systems making it the imaging mode of choice for imaging them.

1.2.1 Absorbance-based imaging

Absorbance is the optical imaging techniques that most resembles how human eyes and brains see the world. Differential absorbance is one way our eyes and brains differentiate objects. An external light source generates photons which bounce off of our environment. Some of the photons are absorbed and others reflected or scattered. A small amount of the scattered and reflected photons come to our eyes. In our eyes these photons are absorbed by photoreceptors. There are four types of photoreceptors, three are for bright light conditions and they differentiate colours, the fourth is useful in low light conditions. Each of these four photoreceptors depends on a different pigment or opsin, which upon absorption of a photon, creates a signal cascade that ultimately results in a signal being passed down the optical nerve, to our brains for processing.

So if we look at a black and white image, the white parts absorb fewer photons resulting in more reaching our eyes and the black parts absorb many more photons resulting in fewer photons reaching our eyes. In our eyes the photoreceptors translate the signal into nerve impulses and our brain creates an image based on the differential absorbance.

Absorbance is based on the physical and electronic configuration of the molecule a photon interacts with. A photon can only be absorbed if the energy of the photon matches an electronic transition in the molecule it hits. Upon absorbing a photon the molecule will go from being in the singlet ground state (S_0) into the first singlet electronic state (S_1), taking the energy of the photon into itself.[4] Next, the molecule will relax back to S_0 by transferring this energy

into the surrounding medium as heat or vibrational energy. Now that the molecule is back in S_0 it can once again absorb another photon and repeat the cycle (Figure 1.1A).

To image cells based on absorbance, the cell or structures within the cell must be associated with molecules that have large absorbances. In a normal cell, some organelles such as nuclei, or chloroplasts are visible using absorbance.[5] However, adding stains such as Masson's trichrome (stains keratin and muscle fibers as red, collagen as blue, cytoplasm as pink, and nuclei as purple) allows many tissues and cell components to be visualized and differentiated.[6]

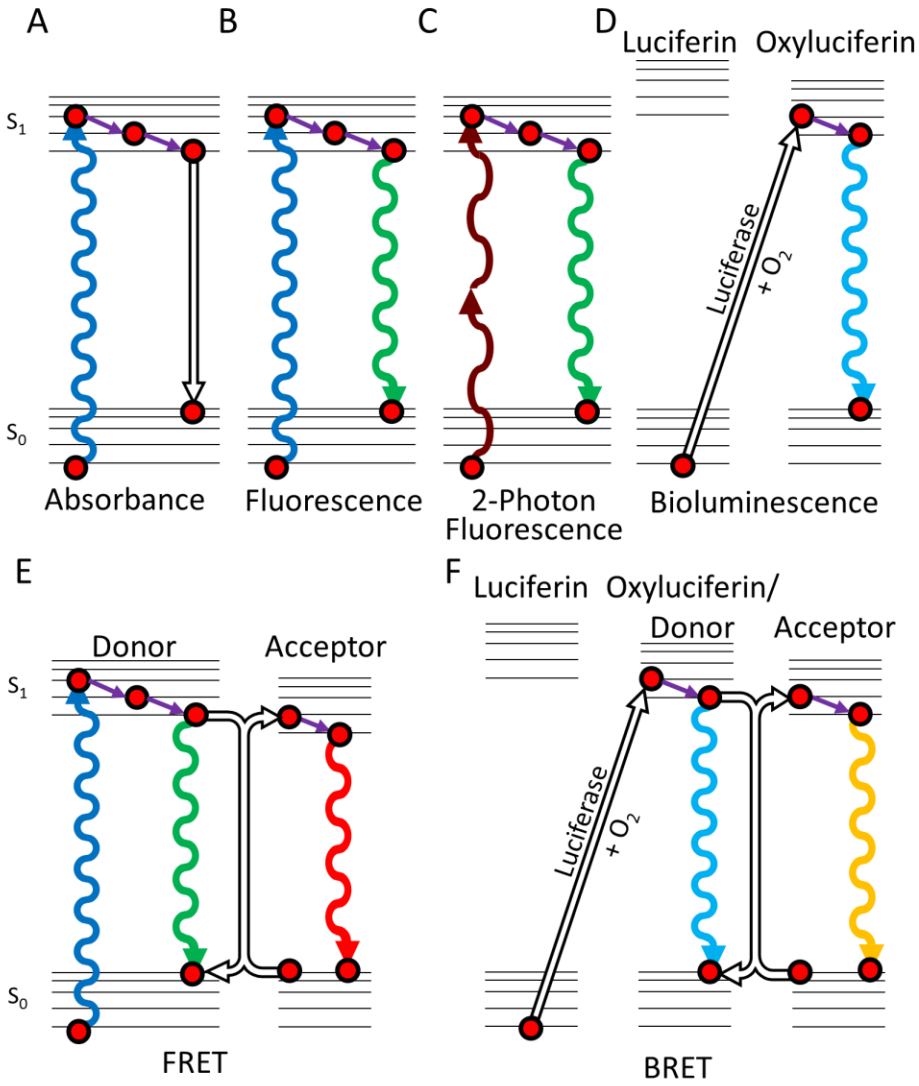

Figure 1.1 Jablonski diagrams of optical imaging modes. A) Absorbance. B) Fluorescence. C) 2-photon fluorescence. D) Bioluminescence. E) Förster resonance energy transfer (FRET). F) Bioluminescence energy transfer (BRET). Undulating lines represent transitions involving photons with up and down arrows representing absorbance or emission, respectively. Hollow arrows represent non-radiative transitions. Purple arrows are internal conversion relaxation events between vibrational states. In paths that bifurcate only one route is taken per event. All emission events have a probability to undergo non-radiative relaxation as seen in A).

1.2.1.1 Fluorescence-based imaging

Fluorescence is a special case where a photon is absorbed as described above, but after absorbance a photon is emitted to relax the molecule from S_1 back to S_0 (Figure 1.1B). When a molecule absorbs a photon, it pushes the molecule into an available excited state with the change in energy matching the energy of the absorbed photon. Each electronic state has several available vibrational states and the molecule often gets excited directly to a higher energy vibrational state of the excited electronic state, broadening the wavelength range that can be absorbed. This rapidly relaxes to the lowest energy vibrational state before fluorescence occurs causing the emitted photon to have less energy than the originally absorbed photon. The ground state has its own set of vibrational states which the molecule can relax into during fluorescence, allowing still lower energy photons to be emitted. The difference in energy/wavelength between the absorbed photon and the emitted photon is called the Stokes shift.

The difference in energy or wavelength between the absorption, also called excitation, and emission is what makes fluorescence microscopy so powerful. As we can create specific wavelengths of light we can illuminate a sample with a wavelength of light that is absorbed by the fluorescent molecule, and filter the light getting to the detector to only allow the lower energy photons resulting from fluorescence. As fluorescent molecules are rare in living organisms, if we add one to a cell the identity of the fluorescing species is generally not in doubt. This creates an image with very high contrast because if there are no fluorescent molecules no photons will be detected, and the background will be black. Stated another way, fluorescence microscopy has a very good signal to noise ratio and a low background signal. In comparison, absorbance techniques have a higher background and a poorer signal to noise.

1.2.1.2 Two-photon excitation

Two-photon excitation is a different way of initiating fluorescence where, instead of using a single photon to excite a fluorescent molecule, two are used. As molecules can only absorb photons that match the transition from one energy level to another, these two photons together must have the same total energy as this transition and be absorbed by the molecule nearly simultaneously.[7] In practice this means that in two-photon fluorescence the excitation wavelength is twice that of a single photon excitation (Figure 1.1C).

Two-photon excitation is beneficial for live cell imaging applications in two ways. First, it allows the use of longer wavelengths which better penetrate living tissue as longer wavelengths avoid the natural absorbance of water and heme. There is also decreased Raleigh scattering at longer wavelengths. Second, two-photon excitation can be limited to a small volume. One-photon systems excite all molecules in the path of the excitation light, with fluorescence scaling linearly with illumination intensity. With two-photon excitation, the fluorescence scales quadratically with illumination intensity.[8] This quadratic scaling results in no fluorescence occurring above and below the focal plane, limiting detectable fluorescence to small volume near the focal point of the illumination (Figure 1.2).

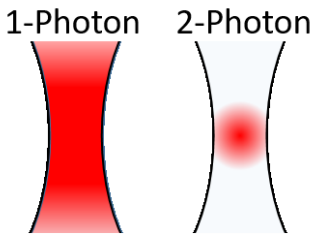

Figure 1.2 Excitation profile of 1- and 2-photon illumination at the illumination's focal point. The black lines show the outer edge of a cross section of the light path at the focal point created by the condenser of a microscope. The red shows the resulting fluorescence intensity of fluorescent molecules excited by the two excitation modes.

1.2.2 Bioluminescence

Bioluminescence is an alternative to fluorescence for imaging of living cells and whole organisms. Bioluminescence is similar to fluorescence in that light is emitted in an electronic relaxation event. However, rather than using the absorbance of light to excite a molecule to S_1, a specific enzyme substrate is enzymatically converted into a product in the excited state which then relaxes and emits (Figure 1.1D). This requires that a supply of substrate be added to a sample in order to create light. The class of enzymes that perform this reaction are called luciferases and oxidize their substrates to create light. As bioluminescence does not require an external light source of any kind it can have even better signal to noise ratio than fluorescence. However as the substrate will be present everywhere it can diffuse to all luciferases which will all be emitting simultaneously. It is estimated that bioluminescent systems have evolved independently more than 50 times,[9] with varying substrates and emission wavelengths allowing multiple luciferases to be used simultaneously (Figure 1.3). The luciferases are classified based on their substrate (*i.e.*, D-luciferin versus coelenterazine) and by their protein structure.

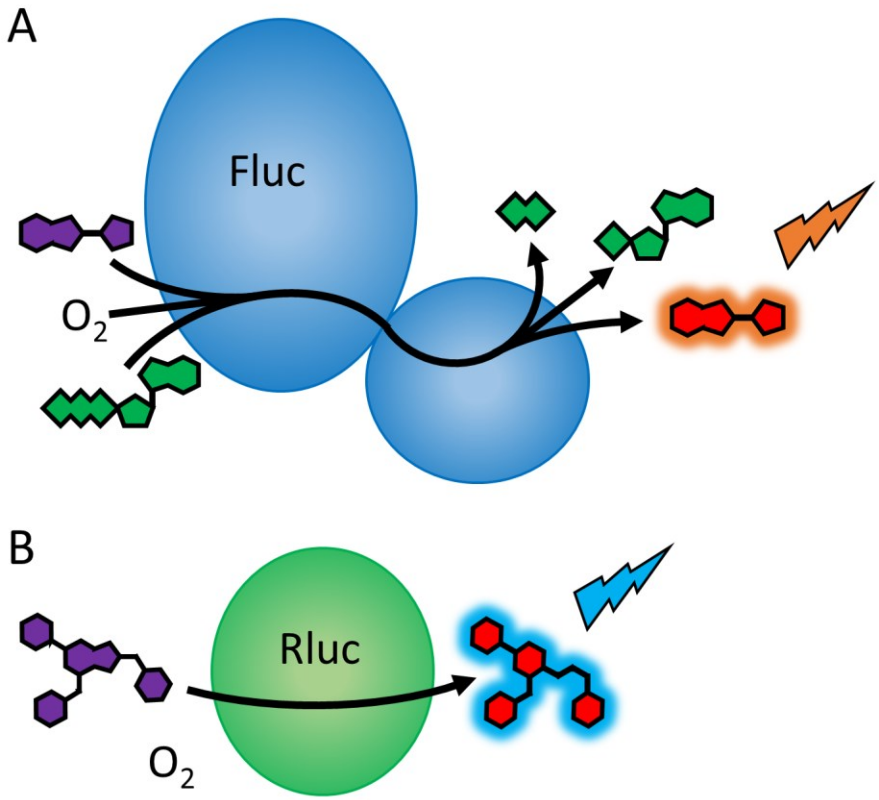

Figure 1.3 Bioluminescent enzymes. A) Firefly Luciferase (Fluc) catalytically converts substrate D-luciferin (purple), molecular oxygen, and adenosine triphosphate (ATP, green), creating pyrophosphate, adenosine monophosphate (AMP) and excited oxy-luciferin (red) which then emits an orange photon. B) *Renilla* luciferase (Rluc) catalytically converts its substrate coelenterazine (purple) and molecular oxygen into excited coelenteramide (red) which then emits a cyan photon.

1.2.3 Förster resonance energy transfer

It is possible to use both bioluminescence and fluorescence to excite another fluorescent molecule through a process called Förster resonance energy transfer (FRET). FRET is the non-radiative transfer of energy between molecules. FRET occurs when an excited molecule called the donor has an emission profile that matches the absorbance profile of another fluorophore, called the acceptor. The donor and acceptor together are called a FRET

pair. FRET depends on the overlap of the donor emission and the acceptor absorbance spectra, the quantum yield of the donor, and the distance and orientation between the donor and acceptor.[4] The donor can be a fluorophore or the bioluminescent product of a luciferase (Figure 1.1E and F). In the case of the bioluminescent product being the donor, the process is also known as bioluminescence resonance energy transfer (BRET) but functions the same.

FRET efficiency is defined by Equation 1-1. With E representing the FRET efficiency, which is the proportion of donor excitation events that result in FRET energy transfer. It depends on r the distance between chromophores and the Förster radius, R_0.[4]

$$E = \frac{1}{1 + (r/R_0)^6}$$

Equation 1-1

The Förster radius is determined by Equation 1-2. It depends several values: the orientation factor, κ; the quantum yield of the donor, ϕ_D; the spectral overlap of the donor and acceptor, J; Avogadro's number, N_A; and the index of refraction of the energy transfer media, n.[4]

$$R_0 = \sqrt[6]{\frac{9 \ln 10}{128 \pi^5 N_A} \frac{\kappa^2 \phi_D}{n^4} J}$$

Equation 1-2

The orientation factor, defined by Equation 1-3, is itself dependant on several angles: the angle between the emission transition dipole of the donor and the absorption transition moment of the acceptor, $\boldsymbol{\theta}_T$; the angle between the line connecting the center of the two chromophores and the emission transition dipole of the donor, $\boldsymbol{\theta}_D$; and the angle between the line connecting the center of the two chromophores and the absorption transition moment of the acceptor, $\boldsymbol{\theta}_A$.[4]

Equation 1-3

$$\kappa^2 = (\cos\theta_T - 3\cos\theta_D \cos\theta_A)^2$$

The spectral overlap is the overlap of the unity normalized fluorescence donor emission spectrum, $\overline{f_D}$ and the acceptor molar extinction coefficient spectrum, ε_A. These are further scaled by the forth power of the wavelength, λ, at each point.[4]

$$J = \int \overline{f_D}(\lambda)\, \varepsilon_A(\lambda)\lambda^4 d\lambda$$

Equation 1-4

In practice the Förster radius describes the distance at which the FRET efficiency is 50%. Measurable FRET changes occur between 1-10 nm, creating an optical ruler that can report the precise distance between the donor and acceptor.

1.3 Fluorescent Proteins

The green fluorescent protein was discovered by Dr. Osamu Shimomura in the jellyfish *Aequorea victoria*, while researching the bioluminescent properties of the jellyfish. It turned out that the jellyfish creates a bioluminescent protein called Aequorin which produces blue photons. Interestingly the jellyfish creates a BRET pair with the Aequorin as donor and *A. victoria* green fluorescent protein (avGFP) as acceptor, which converts the energy via BRET into a green photon.[10] Dr. Martin Chalfie discovered that the gene encoding avGFP was sufficient, without any other cofactors, to create fully mature and fluorescent avGFP and successfully transplanted the gene into *Caenorhabditis elegans* and *Escherichia coli*, proving that avGFP is fully genetically encodable and can be used as a reporter gene.[11] The last of the trio that opened up the field of FP technologies, and won a Nobel prize in 2008 for their efforts, was Dr. Roger Tsien. Dr. Tsien's work elucidated how FPs structure and fluorescence interact,

through mutational analysis, and successfully expanded the palette of FPs,[12–14] creating FPs that light up life across the visible spectrum.

1.3.1 Fluorescent protein structure

Dr. Douglas Prasher determined the primary structure of the 238 amino acid long avGFP.[15] Crystal structures later revealed the avGFP folds into an 11 stranded β-barrel surrounding a core α-helix. The β-barrel is made up of antiparallel β-strands except between strands 1 and 6 which are parallel.[16,17] The protein undergoes a self-sufficient post translational modification to convert three amino acids in the core α-helix, Serine(65)-Tyrosine(66)-Glycine(67) or SYG, into a green fluorescent chromophore, 4-(*p*-hydroxybenzylidene)imidazolidin-5-one (Figure 1.4). The mature chromophore is held in place and protected by the barrel to maintain a planar configuration of the newly created conjugated system, which is required for fluorescence.

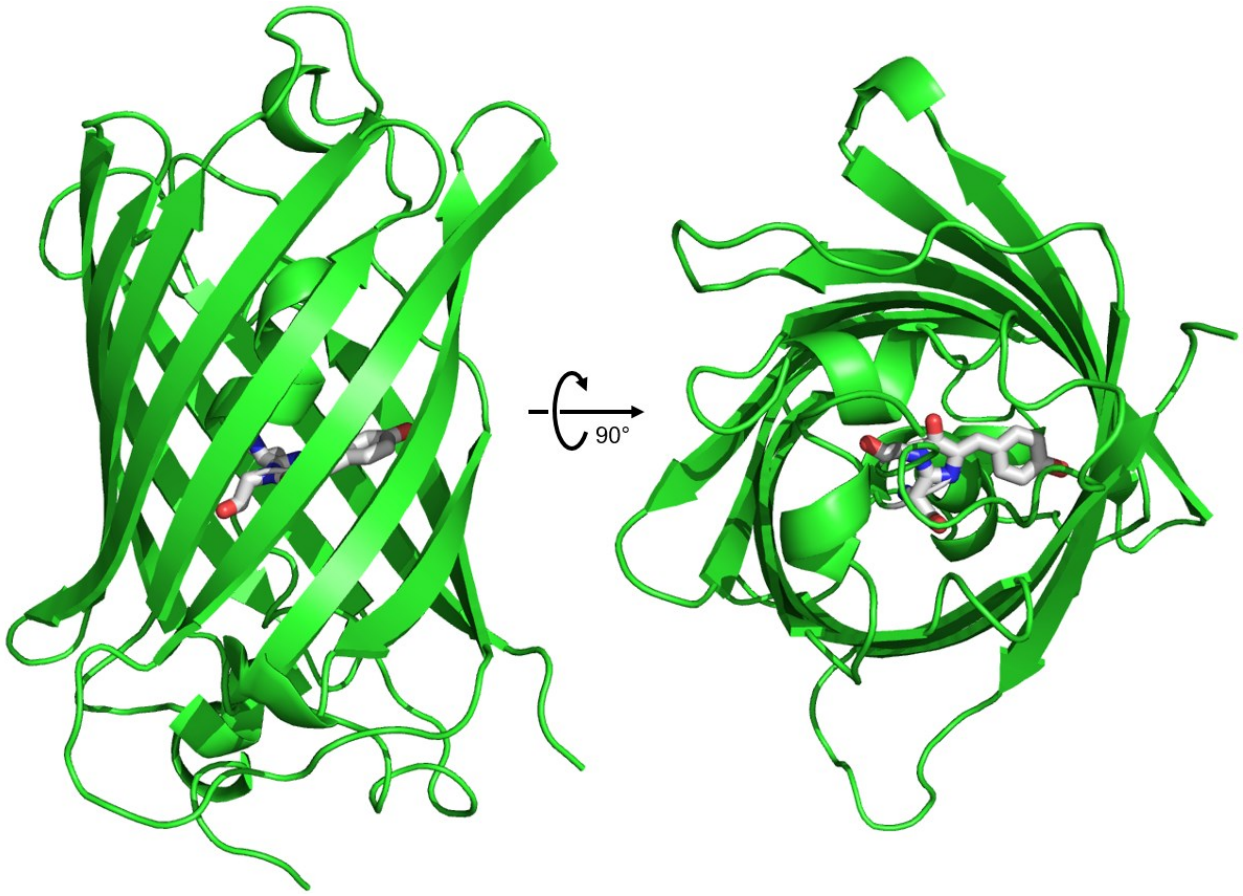

Figure 1.4 Crystal structure of avGFP (PDB: 1GFL).

The β-barrel structure of avGFP has proven tolerant to a large number of mutations. To date more than 32 avGFP variants have been produced as research-ready fluorescent tags, with many more that retained fluorescence but have not proven to be useful. Of the 238 amino acids, 59 have been mutated in at least one research-ready variant. Such diversity now exists that we can analyze the mutation set phylogenetically through the known history from publications but also biologically as if they evolved naturally in a distance tree exclusively based on their mutation set (Figure 1.5).

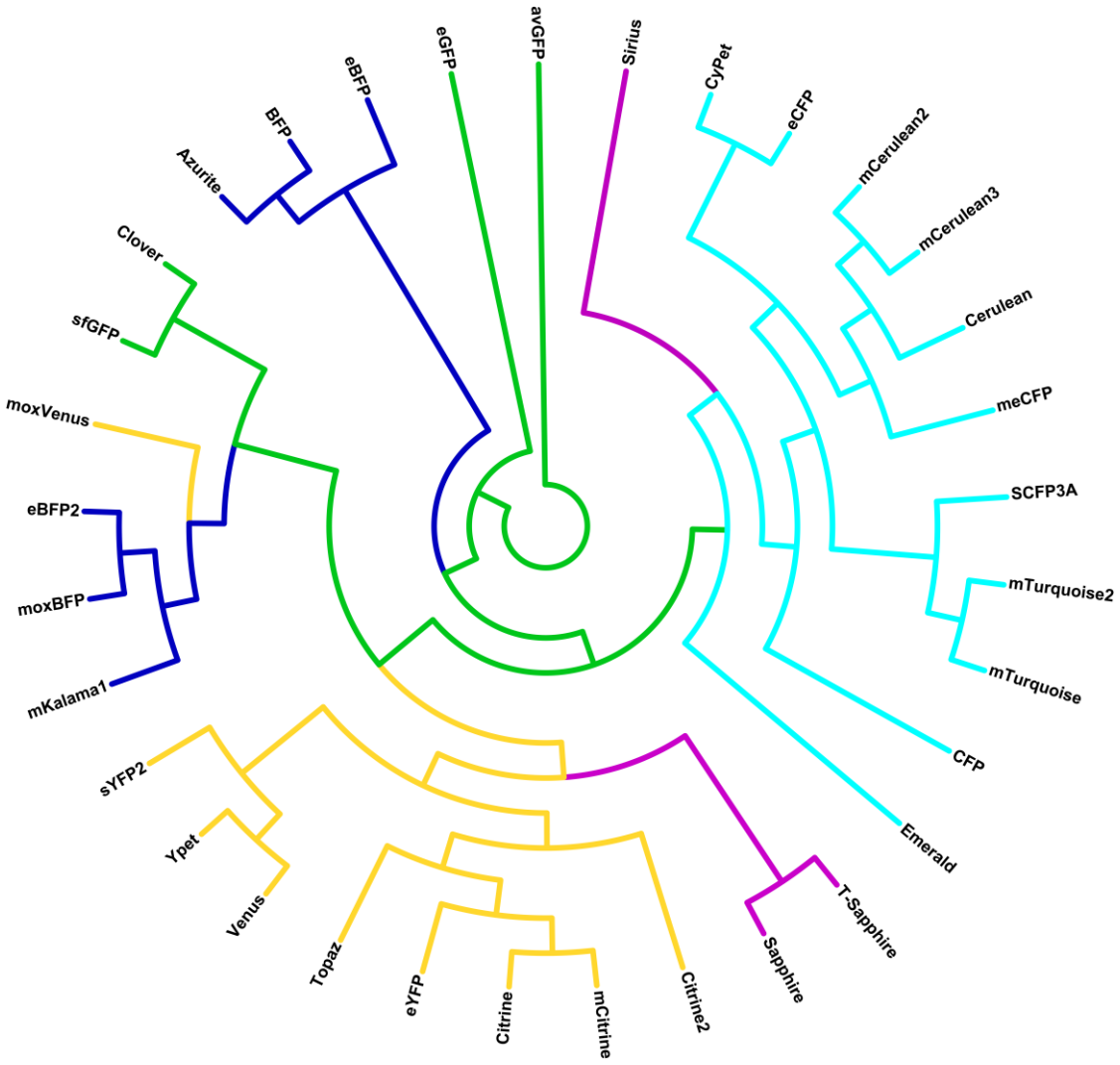

Figure 1.5 Distance tree of the artificially created variants of avGFP. The colours represent the approximate emission colours of each FP. This is the highest scoring tree created in PAUP (V4.0a150) using a neutrally weighted heuristic search with tree bisection and reconnection of the aligned protein sequences.

There is a large natural diversity of FPs and this has branched out into over 400 unique sequences.[18] Today, avGFP variants make up less than a third of all research-ready FPs. The majority of FPs originate from organisms other than *A. victoria*, many come from coral genera such as *Zoanthus*,[19] *Discosoma*,[20] and *Fungia.*[21] However, FPs while rare, exist scattered throughout the tree of life extending even to a chordate *Branchiostoma lanceolatum.*[22]

There is second category of FPs containing members such as UnaG[23] and smURFP[24] which are not homologous to the β-barrel FPs. These have a binding pocket for a substrate that activates the substrates fluorescence. This requires the host organism or the researcher to provide this substrate. This group of FPs are of interest as they include members that are able to fluoresce brightly in the near IR region,[24–26] which is not currently accessible to β-barrel FPs. Care must be taken as their substrates are not present in all hosts nor necessarily at high enough concentration to achieve high brightness. However, as many FPs in this category bind biologically ubiquitous substrates for their chromophore (e.g., biliverdin or bilirubin), engineering improved binding pockets shows promise in enhancing their utility.[24]

1.3.2 Chromophore formation

As stated previously, β-barrel FP chromophores are created through a self-sufficient post-translational modification. This means that within a properly folded FP barrel the chromophore can be created using only its own peptides, water, and oxygen. At least three reactions must occur to form the chromophore of an FP: cyclization, oxidation, and dehydration. After all three reactions occur, the chromophore must then be held in a planar conformation to allow for fluorescence as twisting results in increased non-radiative relaxation. Most enzymes only catalyze one specific reaction andthey do so with an average catalytic rate constant (k_{cat}) of 10 s^{-1}.[27] From this perspective, FPs are quite slow at performing the multiple reactions required for chromophore formation. The fastest FPs are able to mature in ~10 minutes with a k_{cat} of 0.002 s^{-1},[22] with most requiring a few hours with a k_{cat} less than 0.0003 s^{-1}.[14,17,28]

1.3.2.1 Chromophore formation of a green chromophore

The chromophore of avGFP is built from the amino acids SYG in a mechanism first proposed by Heim et al.[29] and refined by Rosenow et al.[30] (Figure 1.6). The backbone of the three residues are cyclized by the nitrogen of the glycine making a nucleophilic attack on the carbonyl carbon of the serine. The oxidation step creates an acylimine connecting the nitrogen and α-carbon of the tyrosine in the newly formed heterocycle. Last is the dehydration reaction which removes the hydroxyl group of what used to be the carbonyl carbon of serine. The dehydration creates a new acylimine double bond between the old serine carbonyl carbon and the tyrosine nitrogen and pushes the double bond off the nitrogen to between the α and β carbons of the tyrosine. This results in the mature form of the avGFP chromophore.

The tyrosine side chain exists in an equilibrium between its neutral phenol form and the anionic phenolate. When protonated, the phenol chromophore excites maximally at 395 nm and emits at 510 nm. When deprotonated, the phenolate chromophore excites maximally at 475 nm and also emits at 510 nm. Both forms emit at the same wavelength due to excited state proton transfer (ESPT). The phenol form, once excited, is rapidly deprotonated into the anionic phenolate from which emits at 510 nm.

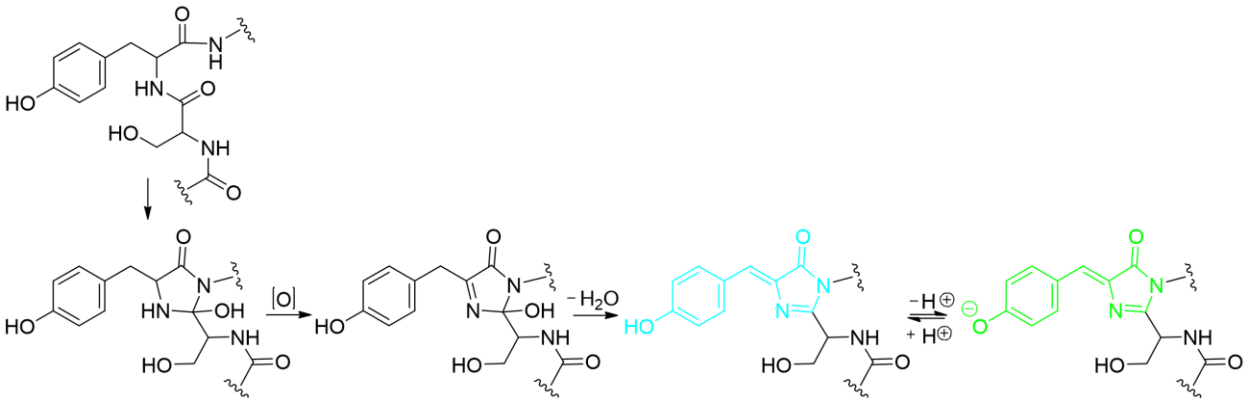

Figure 1.6 Chromophore formation in avGFP. [30]

1.3.2.2 Chromophore formation of a red chromophore

Creating longer wavelength fluorescence requires the creation of a larger conjugated system which is generally achieved in FPs by adding another oxidation step to extend the chromophore with an acylimine moiety. This oxidation extends the chromophore's conjugated system to include the acylimine, (Figure 1.7).[31] In the case of dTomato[14] and its progenitor DsRed[20] it became apparent that the order of the reactions is both important and the reactions occur stochastically rather than progressively. A branched pathway exists such that the chromophore can form a green or a red chromophore and they do not interconvert.[31] This appears to be related to the dehydration reaction which is the irreversible step on the pathway to the green chromophore. A green chromophore is produced if the dehydration reaction happens immediately after the first oxidation (Figure 1.7). A red chromophore is produced if a second oxidation reaction occurs before the dehydration step.

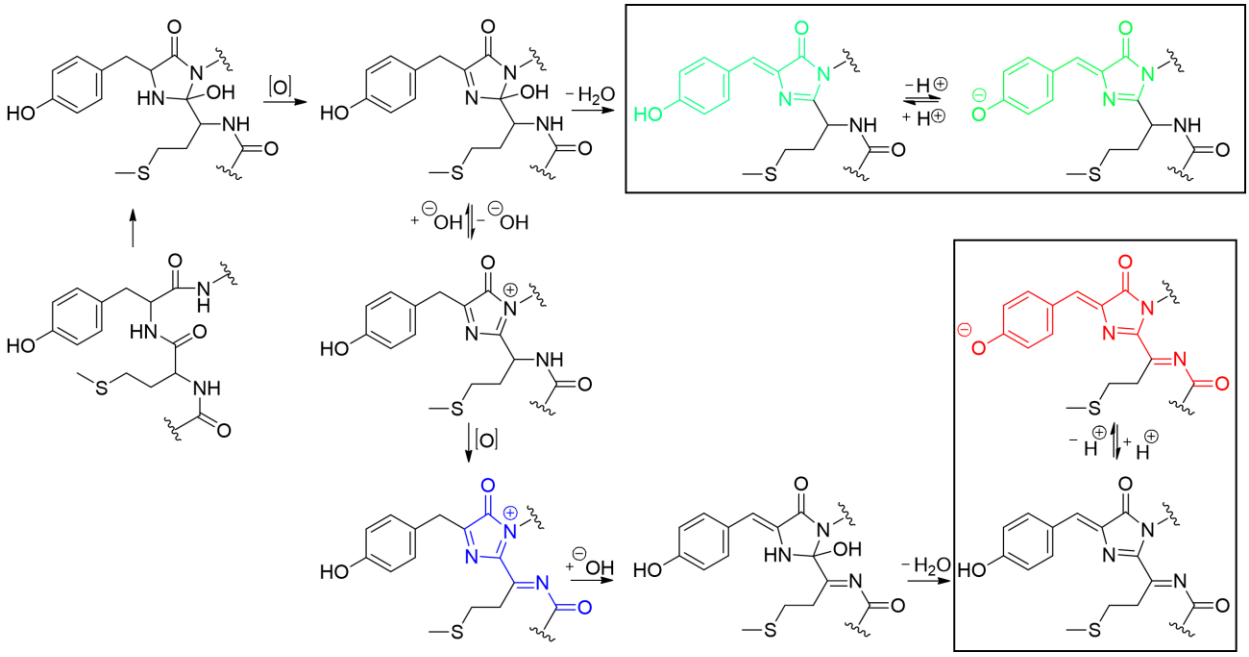

Figure 1.7 Branched chromophore maturation in red fluorescent proteins. Colours represent the emission wavelength of fluorescent conjugated system.

1.4 Photophysical parameters of fluorescent proteins

In order to describe a fluorescent system there are several commonly used sets of parameters: 1) the spectral profile (determined by absorbance, excitation, and emission spectra); 2) the fluorescent brightness (determined by quantum yield and extinction coefficient); 3) fluorescence lifetime; and 4) photochemistry. Each of these parameters gives important information about how a fluorescent molecule performs.

1.4.1 Spectral profile

There are three types of spectra used to describe chromophores: absorbance, excitation, and emission. An absorbance spectrum is created by measuring the amount of light that is

absorbed at each wavelength. An absorbance spectrum shows all wavelengths that are absorbed including those that do not lead to fluorescence (Figure 1.8A).

Excitation and emission spectra deal solely with fluorescence. An excitation spectrum is acquired by illuminating a sample with of a series of wavelengths and measure the intensity of emission at a single chosen wavelength (Figure 1.8B). An emission spectrum is acquired using a single excitation wavelength to illuminate the sample and the detector scans all wavelengths that are emitted (Figure 1.8C).

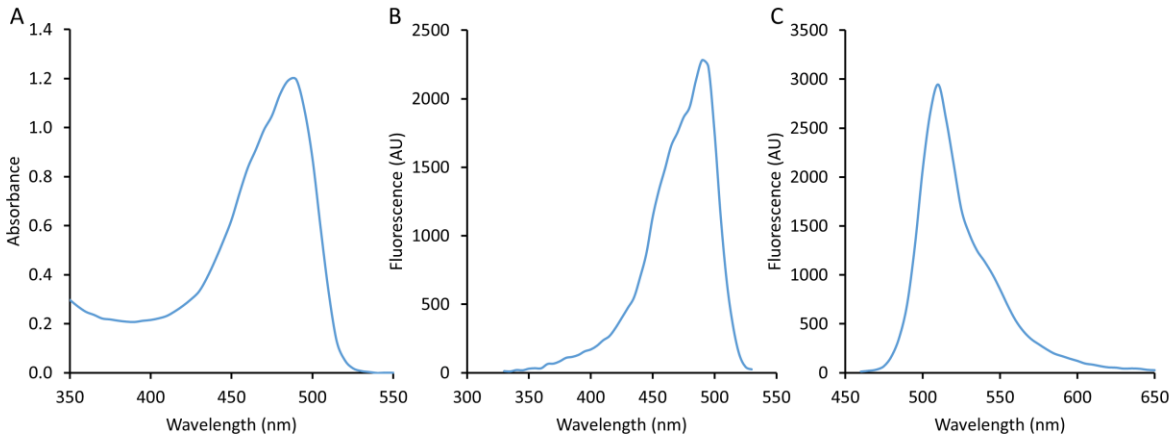

Figure 1.8 Absorbance, excitation, and emission spectra of eGFP. A) Absorbance spectrum of eGFP. B) Excitation spectrum for eGFP (emission at 540 nm). C) Emission spectrum for eGFP (exciting at 450 nm).

Often FPs are described only by the excitation and emission maxima, which for eGFP is 488 nm and 507 nm, respectively. However, the spectral profile better informs a user as to what other wavelengths can be used for excitation and emission. Though rarely used, even more detailed information can be obtained from a three dimensional (3D) scan of excitation and emission. A 3D scan shows every possible combination of excitation and emission and therefore give a complete picture of the full fluorescence landscape (Figure 1.9).

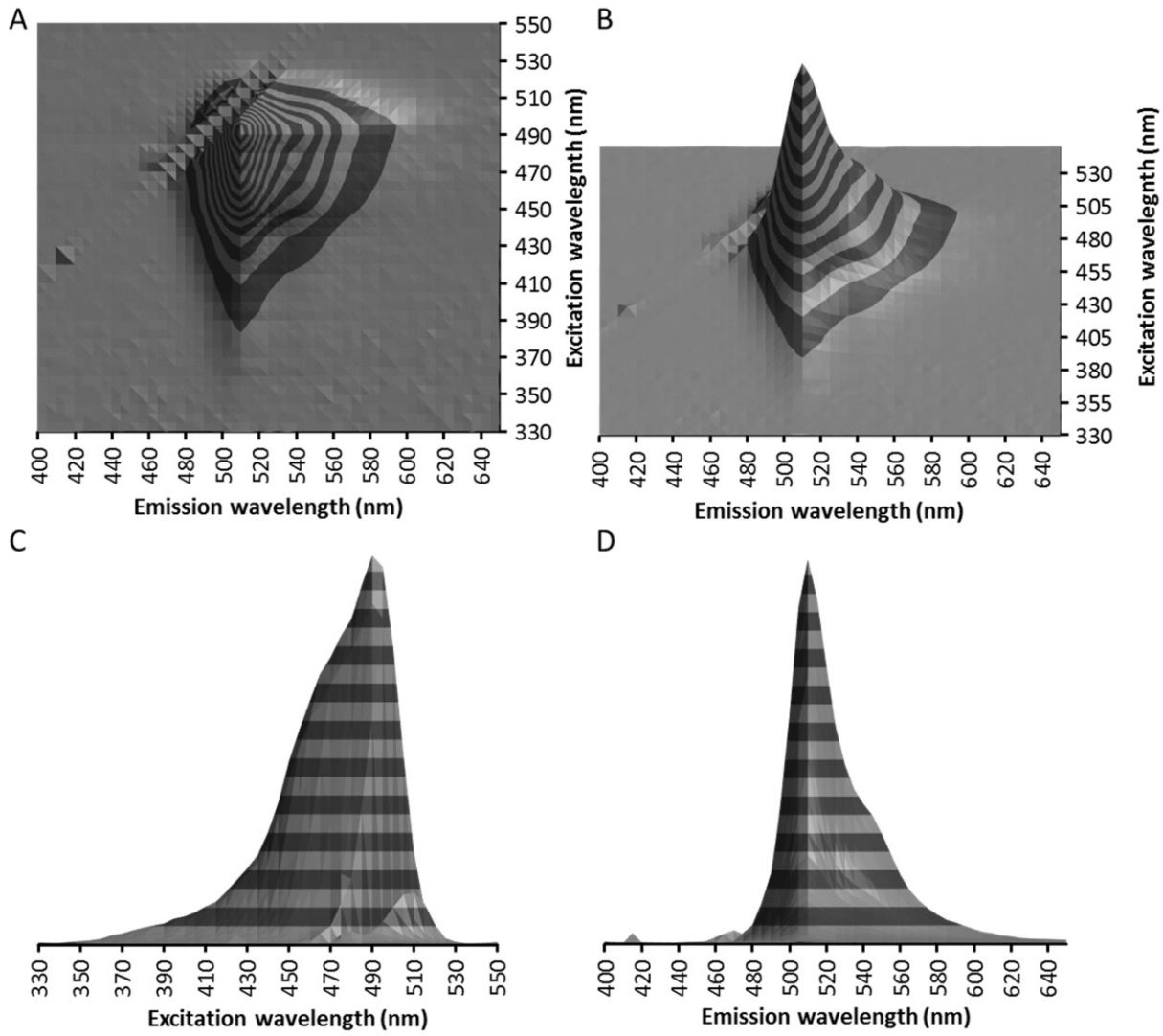

Figure 1.9 3D fluorescence scan of eGFP. A) Top down view of eGFP fluorescence landscape. B) 45 degree rotation on the Y axis of the eGFP landscape. C) Side view of eGFP landscape showing the excitation profile. D) Side view of eGFP landscape showing the emission profile. Each gray band is 300 arbitrary units of fluorescence. When the excitation wavelength is within 5 nm of the emission wavelength the fluorescence was set to zero to remove the bleed though to the detector.

1.4.2 Fluorescent brightness

In order to quantitatively compare chromophores we often use a term called brightness which corresponds to how efficiently the illumination light is converted into fluorescence

emission. Brightness is the product of the extinction coefficient, which describes how well the chromophore is able to absorb light, and the quantum yield, which describes the efficiency which the chromophore is able to convert an absorbed photon into fluorescence. Brightness is reported several ways but the most common two are reporting the product of the extinction coefficient and quantum yield, divided by 1000. Sometimes the resulting brightness is normalized to that of a standard fluorescent molecule such as eGFP or fluorescein. Both quantum yields and extinction coefficients can depend on environmental conditions, and so the conditions under which these numbers are determined are relevant.

Extinction coefficients are used in the Beer-lambert law to relate absorbance and concentration for light passing through a known length of a sample solution.[4] In research-ready FPs the extinction coefficient varies between a low of about 10 000 $M^{-1}cm^{-1}$ and a high of about 128 000 $M^{-1}cm^{-1}$ (for mRuby3).[32] FPs with extinction coefficients of less than 10 000 $M^{-1}cm^{-1}$ would generally be considered too dim to be practically useful. Extinction coefficients are wavelength specific and are reported at their maximal value which is the peak of the absorbance spectra. The extinction coefficient scales directly with absorbance such that a wavelength that has 50% of the absorbance at the peak value will have 50% of the extinction coefficient.

Quantum yield is the answer to the question: How many photons are emitted for every 100 photons absorbed? FPs' quantum yields vary from less than 0.01% (this is approximately where they start to be referred to as chromoproteins rather than FPs) to 93% in mTurquoise2.[33] Quantum yields are not wavelength dependent as long as the illuminating photons are only being absorbed by the chromophore.[4] If other absorbers are present, this value becomes

weighted by the ratio of absorbing entities and their individual quantum yields giving an apparent wavelength dependence but not a real dependence.

1.4.3 Fluorescence lifetime

Fluorescence lifetime determines how fast a chromophore is able to complete an absorbance/emission cycle. Fluorescence lifetimes in FPs are generally between 1.2 and 4 ns for research ready FPs with mCherry at the low end and mTurquoise at the high end.[34] Shorter values can be readily obtained but, as fluorescence lifetimes are directly linked with quantum yield, with shorter lifetime FPs also having low quantum yields and as such they are quite dim and therefore not commonly used.[4] As the absorbance and internal conversion steps are in the attosecond and picosecond timescales, fluorescence lifetime depends mostly on the time spent in the lowest energy vibrational state of S_1 before emission.[4] The initiation of emission is a stochastic process creating a decay curve from which we can calculate a lifetime. With appropriate instrumentation, such as a pulsed illumination source with time correlated single photon counting detectors, can measure two fluorescent molecules with identical excitation and emission spectra and differentiate them by their fluorescence lifetimes. When applied for imaging, this technique is called fluorescence lifetime imaging microscopy (FLIM).[4] Just as with quantum yields, fluorescence lifetime is wavelength independent.

A smaller fluorescence lifetime would imply that a given chromophore could be activated more often making more fluorescence. However, as fluorescent lifetimes are only low when the quantum yield is correspondingly low this generally results in a loss in brightness due to a low quantum yield rather than an increase.

1.4.4 Photochemistry

Photochemistry is the changing of a substance from one form to another using light. In FPs photochemistry comes in four variations: photoactivation, the turning on of fluorescence; photoconversion, the irreversible changing of the colour of fluorescence; photoswitching, a reversible change in the photophysical state; and photobleaching, the irreversible loss of fluorescence.

1.4.4.1 Photoactivation

One example of photoactivation occurs in photoactivatable-GFP (PA-GFP), it acquires fluorescence with 400 nm illumination which induces decarboxylation of a glutamic acid. Before decarboxylation the glutamic acid, E222, supports a hydrogen bond network that holds the chromophore in the neutral state which was rendered non-fluorescent through the T203H mutation. After decarboxylation the hydrogen bonding networking rearranges resulting in a stabilized anionic chromophore which is fluorescent.[35]

1.4.4.2 Photoconversion

Most FP photoconversions are those that undergo green-to-red conversion. All such FPs have a histidine-tyrosine-glycine (HYG) chromophore which initially forms from a green chromophore as represented in Figure 1.6. Illumination using ~400 nm light results in a β-elimination creating a double bond between Cα and Cβ of histidine. This connects to the chromophores conjugated system resulting in red fluorescence and the scission of the histidine N-Cα bond, breaking the backbone of the FP.

Photoconversion was combined with a single FP sensor for calcium, creating CaMPARI.[36] CaMPARI is a green-to-red photoconvertable FP that only converts to red if both Ca^{2+} is bound and 405 nm light is applied, creating an optogenetic AND logic gate. This functions by the conformational change induced by calcium binding controlling the residues responsible for photoconversion.

The other relevant photoconversion is the unique orange to far red conversion in PSmOrange2.[37–39] PSmOrange2 contains an ITYG chromophore which self-sufficiently forms a second heterocycle, shown in Figure 5.1. Illumination with 480 nm light results in a β-elimination and the scission of the C_O-Cα bond in the isoleucine backbone adding a carbonyl to the conjugated system converting in to far-red fluorescence and adding a double bond between the Cα and Cβ of the isoleucine.

1.4.4.3 Photoswitching

Photoswitching is generally controlled by chromophore isomerization. In FP chromophores the aromatic residue has a Cα-Cβ double bond that can be in the Z or E configuration. The vast majority of FPs are in the Z configuration but when in the excited state it is possible to switch between these two confirmations. In order for switching to occur it must be sterically allowed by the surrounding residues which is not generally the case. If the conformational switch is allowed, the isomerization changes the chromophores environment and can activate or deactivate fluorescence[40] and even change the colour.[41] A second excitation event can allow the chromophore to switch back regaining its original fluorescent state.

This photoswitching behavior has been harnessed several ways. For example, it enables super-resolution imaging based on reversible saturable optical linear fluorescence transitions

(RESOLFT).[40,42] Photoswitching as well as photoconversion can also allow for optical highlighting of a specific cell to facilitate tracking it over time.[43]

Photobleaching, is the irreversible loss of fluorescence. Resistance to photobleaching or photostability describes the amount of time a fluorescent molecule can remain functional while being illuminated. Every absorbance event could potentially result in a permanent loss of fluorescence. This can be due to a variety of mechanisms but all of them either change the structure of the chromophore[44,45] or, as in the case of FPs, its surrounding environment.[46] At a given illumination intensity, photostability is often described by the amount of time required to drop to 50% of the original fluorescence. In general, FP photostability half-lives vary from tens to hundreds of seconds at illumination conditions similar to those used to image cells. This is the one characteristic where FPs are generally considered to be inferior to fluorescence dyes with many fluorescent dyes being nearly 10-fold more photostable than FPs.[47] Measuring photostability of FPs can be problematic as it depends on numerous factors including, but not limited to: illumination wavelength, illumination intensity, the buffer composition, and temperature. As the exact mechanisms by which FPs lose their fluorescence remain poorly understood, there are variables that probably play important roles that we cannot account for. This will be discussed in greater detail in Chapter 3.

1.5 Protein complementation

When a protein is divided into two or more polypeptide chains it is sometimes possible to reconstitute the complete folded protein structure by bringing the fragment polypeptides into close proximity. If the split protein is one that normally creates a detectable signal, the proximity-dependent reconstitution can be used as a PPi sensor. To use a split protein to detect

a putative PPi between proteins A and B, protein A must be genetically fused to one half of the split protein and protein B to the second half of the split protein. If A and B interact they will come into close proximity, this drags the two halves of the split protein together allowing them to regain their function, which is the creation of a signal we can detect. By fusing the split protein halves to the proteins of a PPi they become a sensor for the PPi by watching for the function of the split protein. Creating such protein fragment complementation assays requires genetically splitting a detectable protein, such as an FP, into two polypeptides. Each fragment should be soluble and partially fold in such a way that it remains possible to regain function when brought into the proximity of its other half.

1.5.1 Split fluorescent proteins

As fluorescent proteins are only fluorescent after correct folding and the subsequent maturation of the chromophore, splitting an FP into two polypeptides will prevent fluorescence until the two fragments come together. Once brought into close proximity, the complete beta-barrel can form and the chromophore can mature to become fluorescent. This technique for detecting PPi is called bimolecular florescence complementation (BiFC). The first FP to be divided into two fragments was eGFP. In the original implementation, the smaller fragment contained the last 4 β-strands with the larger half containing 7 β-strands and the central helix, these and were brought together using an antiparallel leucine zipper PPi to regain its fluorescence.[48] Since this first version was reported, many more split FPs have been engineered. A few split FPs share the 4 β-strand split location, with others splitting off the last 3 β-strands.[49] In 2008 a third split site that only splits off β-strand 11 was developed using an improved eGFP called superfolderGFP (sfGFP).[50] There has also been a tripartite split where

sfGFP was split into three parts: β-strands 1-9, β-strand 10, and β-strand 11. Reconstitution of the FP requires that all three parts be brought into proximity.[51,52]

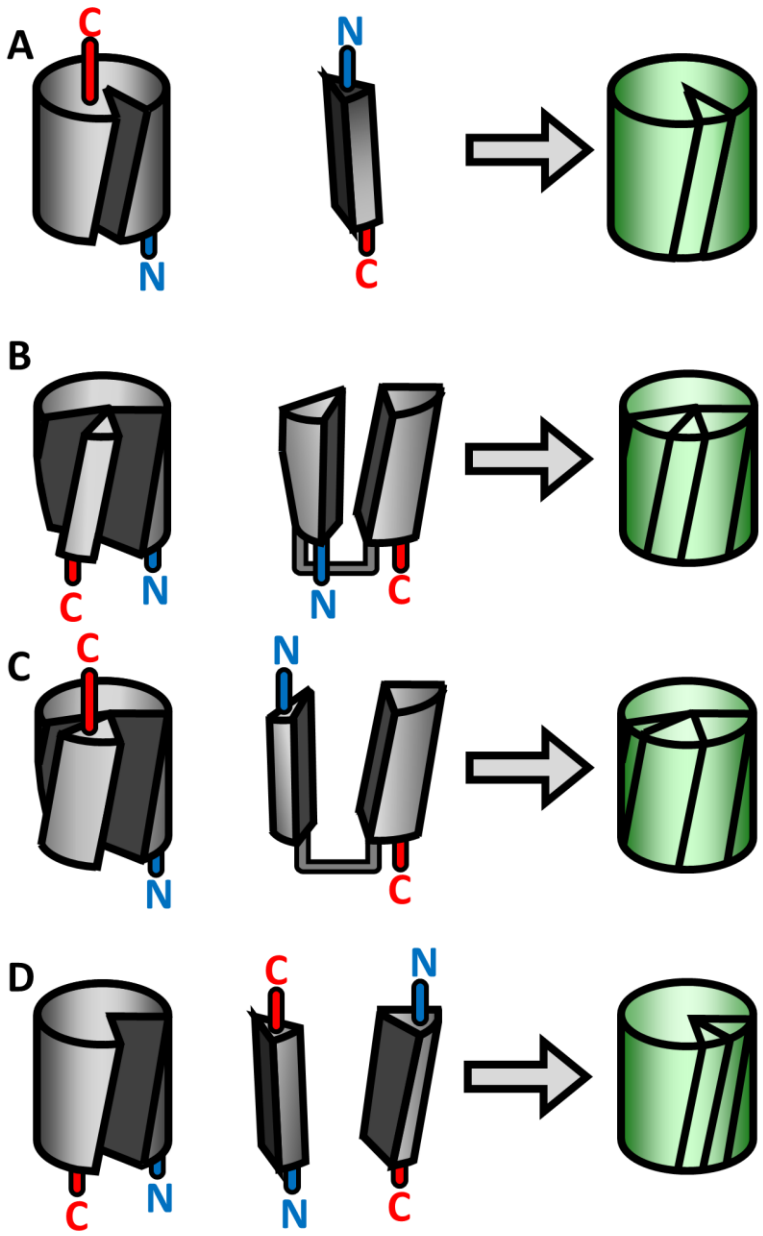

Figure 1.10 Split FP designs. A) Split between, β-strand 10 and 11. B) Split between, β-strand 7 and 8. C) Split between, β-strands 8 and 9. D) Tripartite split between, β-strands 9, 10, and 11. Blue and red lines represent the N- and C- termini, respectively.

Split FPs have limitations that limit how they can be used. First is the tendency to self-assemble creating a false positive signal that reduces their ability to identify weak PPi. Second is the irreversible nature of the reconstitution. When the PPi of interest occurs the split FP reconstitutes, this now whole split FP will not dissociate resulting in the proteins involved in the PPi to be permanently held in close proximity.[53] Third is the lag time between the binding event and the protein folding and maturation of the chromophore such that time resolution for a PPi is in the 10s of minutes.[49] Finally, some fusion proteins only tolerate N or C terminal fusions restricting which split locations are viable. Despite these limitations, split FPs have allowed for the identification of many PPi and even have been used in high throughput screening for inhibitors of specific PPi.[54–57]

There have been many different colours of split FPs developed allowing multicolour PPi assays.[49] Combining split FPs with FRET or BRET partners can be used to visualize three or more interacting molecules. By using a split FP, which when reconstituted can act as part of a FRET pair a three way interaction can be measured by the split FPs fluorescence and the FRET ratio.[49]

1.5.1.1 Dimerization dependent fluorescence

Dimerization dependent fluorescent proteins (ddFPs) are a class of fluorescent proteins modified to detect PPi.[58] The goal in developing ddFPs was to create a reversible version of split FPs by creating an FP pair that change fluorescence upon dimerization. The ddRFP pair increases in brightness by 10-fold upon dimerization and decreases back to basal fluorescence once dissociated (Figure 1.11A). This creates a rapid change in fluorescence and allows the

tracking of dynamic PPi. It does trade the molecular ruler aspect of FRET pairs for a larger change in brightness.

The two parts of a ddFP are generally referred to as A and B. A, is a very dim red FP containing a mature chromophore that is quenched. B was mutated such that it no longer forms a chromophore but still folds into the β-barrel structure. When A and B dimerize their interface interaction causes a subtle change in the conformation of the A protein that reduces the quenching and the fluorescence of A increases by 10 fold.

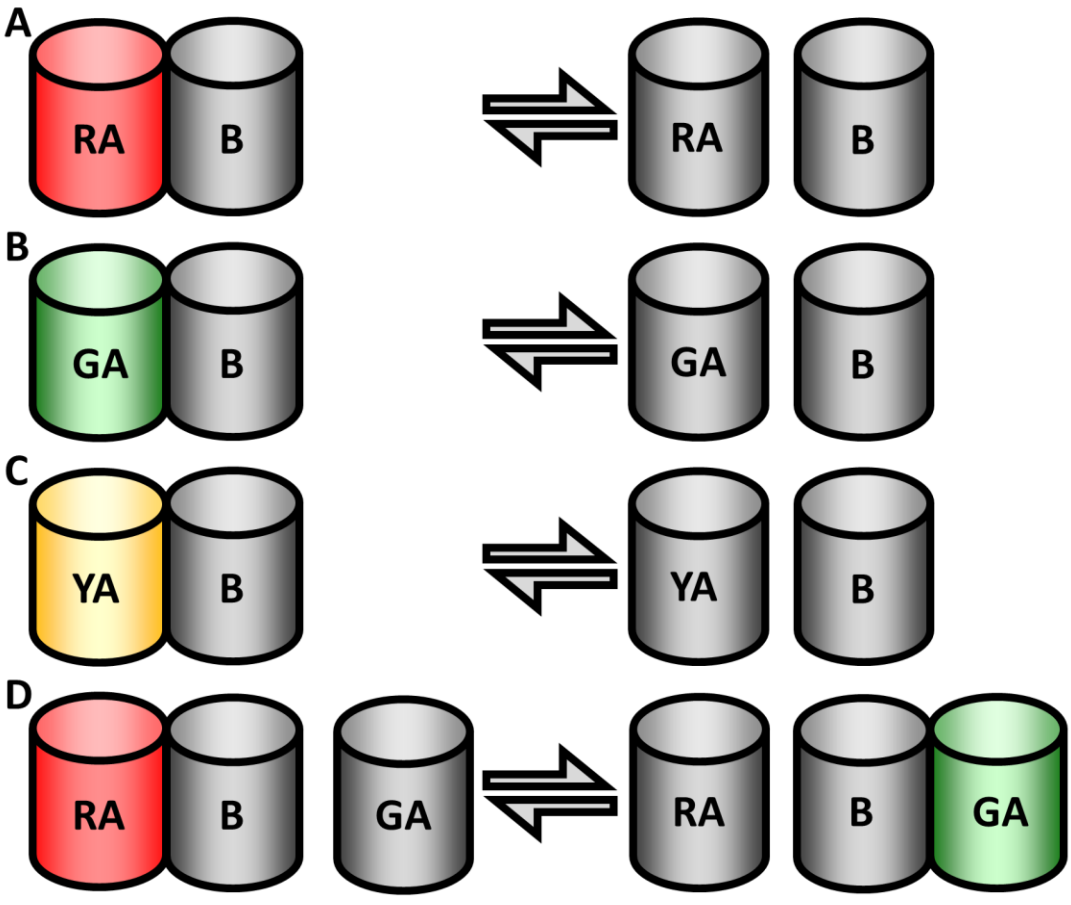

Figure 1.11 Dimerization dependent FPs. A) ddRFP. B) ddGFP. C) ddYFP. D) FPX using ddRFP and ddGFP.

The ddFP design was later extended to create green and yellow pairs (Figure 1.11B and C).[59] It was found that the B copies could activate the A copies associated with all three colours. Use of the green and red A copies with a single B copy allows the exchange of the B copy to switch between bright red or green fluorescence creating a ratiometric assay for a dynamic PPi such as calmodulin and M13 binding Ca^{2+}, in a technique termed fluorescent protein exchange or FPX (Figure 1.11D).[60]

1.5.2 Split luciferases

Similar to fluorescent proteins, luciferases have been converted to make protein fragment complementation assays. For luminescent proteins this is called bimolecular luminescence complementation or BiLC.[61] There are currently four types of split luciferases that have been categorized based on their protein structure and substrate.

Table 1.1 Split luciferases

Organism	Species	Luciferase name	Length	Size (kDa)	Split name	N-terminal split	Split name	C-terminal split	Substrate
Firefly	*Photinus pyralis*	Firefly Luciferase	550	61	NLuc	2-416	CLuc	398-550	D-luciferin
Click beetle	*Pyrophorus plagiophthalamus*	Click Beetle Red	542	64	CBR (1-414)	1-414	CBR (395-542)	395-542	D-luciferin
Click beetle	*Pyrearinus termitilluminans*	Emerald Luciferase	542	64	ELucN	1-415	ELucC	394-542	D-luciferin
Sea pansy	*Renilla reniformis*	*Renilla* Luciferase	311	36	N-Rluc	1-229	C-Rluc	230-311	Coelenterazine
Copepod	*Gaussia princeps*	*Gaussia* Luciferase	185	20	Gluc(1)	1-93	Gluc(2)	94-185	Coelenterazine
Shrimp	*Oplophorus gracilirostris*	Nanoluc	171	19	N65	1-65	66C	66-171	Furimazine

The first type of split luciferases are the D-luciferin-based beetle luciferases which are 542-550 AA in length and include the firefly, *Photinus pyralis*,[62] and click beetle *Pyrophorus plagiophthalamus*[63] and *Pyrearinus termitilluminans*[64] luciferases. These luciferases are homologous and each contains a large N-terminal domain and a smaller C-terminal domain. They have been split such that the N-terminal and C-terminal fragments overlap and both contain the last α-helix of the N-terminal domain.[65] The homology is strong enough that an attempt to improve the CBR(395-542) fragment resulted in a C-terminal split, McLuc1 with three mutations, F420I G421A and E453S, which became capable of activating all three N-terminal fragments.[63] This could in future allow for assays like FPX to be performed using these split luciferases.

The second type is the *Renilla* luciferase which is a coelenterazine-based luciferase that is 311 AA in length and has a globular protein structure.[66] It has been split to create two non-overlapping fragments of 1-229 and 230-311 AA.

The third type is *Gaussia* luciferase which is another coelenterazine-based luciferase, but is much shorter at 185 AA and is the brightest luciferase that uses the coelenterazine substrate. Its structure is yet to be determined but it has low sequence homology to *Renilla* luciferase. However, unlike *Renilla* it is excreted from the cell implying it is optimized for an extracellular environment.[67] *Gaussia* luciferase has been split into two non-overlapping fragments composed of the residues 1-93 and 94-185.

Finally, the fourth type is Nanoluc[68] which is an enhanced version of the large 109 kDa *Oplophorus* luciferase complex which is composed of two 35 kDa proteins and two 19 kDa proteins.[69] Nanoluc is an optimized variant of the 19 kDa protein that no longer forms a complex. This optimization also identified an improved substrate furimazine which is more

stable and has high signal intensity than coelenterazine. Nanoluc has an 10 stranded β-barrel with a 3 α-helix lid (PDB ID: 5IBO).[70] It has been split to create a 1-65 N-terminal fragment and 66-171 C-terminal fragment with both containing part of the barrel and lid. Another split NanoLuc (called NanoBiT) was created by splitting between residue 156 and 157 which removes only the 10th β-strand. Both polypeptides of NanoBiT were subsequently evolved for improved solubility.[71] The NanoBiT's larger N-terminal fragment is named 11S and the C-terminal fragment is called 114.

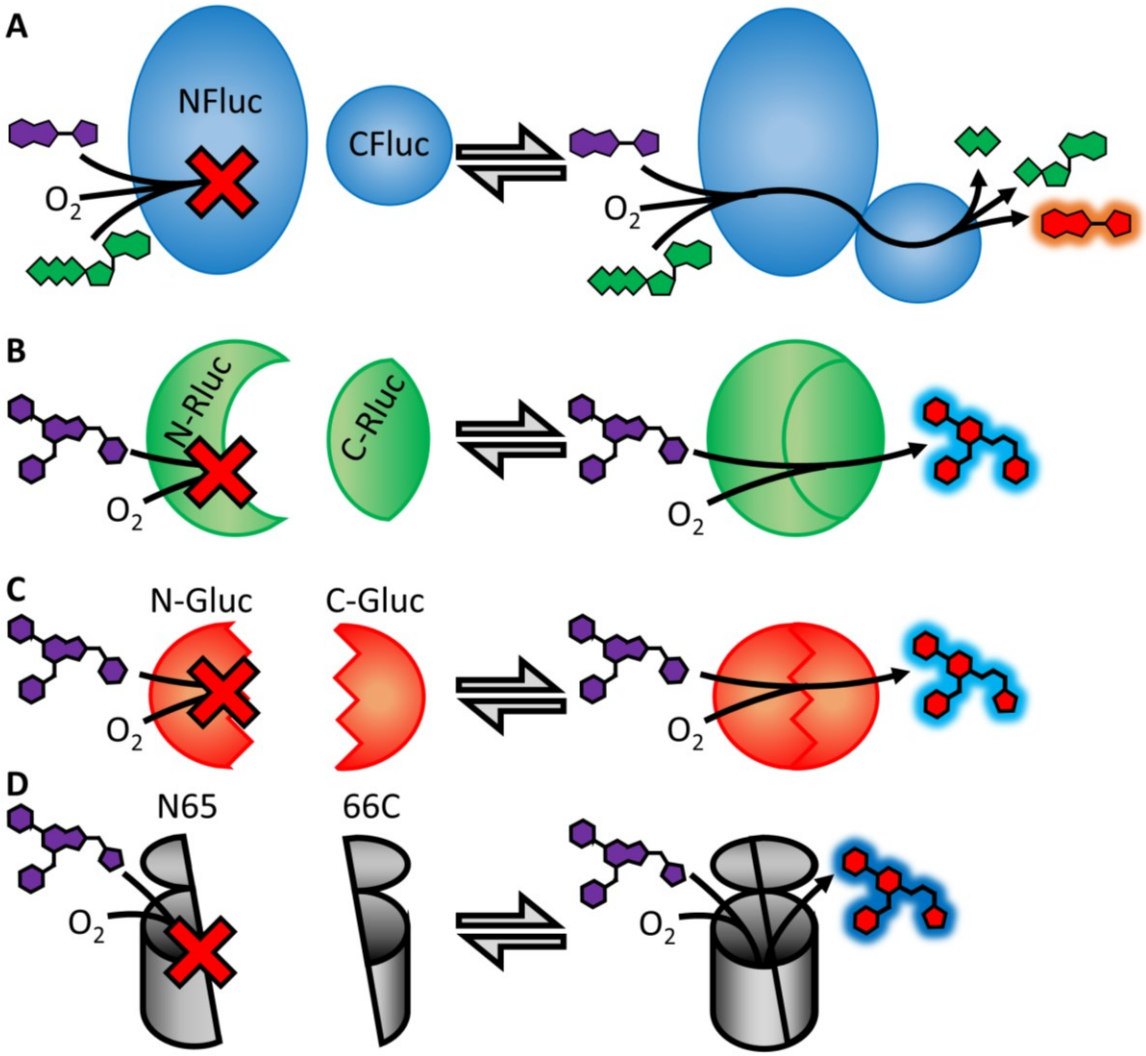

Figure 1.12 The four types of research-ready split luciferases. A) Split D-Luciferin luciferases containing Firefly (shown), click beetle red, and Emerald luciferases. B) *Renilla* luciferase using the substrate coelenterazine. C) *Gaussia* luciferase using the substrate coelenterazine. D) NanoLuc using the substrate furimazine.

There are many other luciferases[72] as well as some minor variations of the split luciferases mentioned above.[73–75] The above list shows the most commonly used and potentially useful variants.

1.6 Optogenetics

Optogenetics refers to the use of genetically-encoded proteins that interact with light, as tools for visualization (sensors) or control (actuators) of biology. Optogenetic sensors are genetic elements that change their emission of light based on an external stimulus such as pH or Ca^{2+} concentration. Optogenetic actuators are proteins that change their structure or function upon exposure to light, such that they cause a meaningful change in a biological process.

1.6.1 Optogenetic sensors based on FPs

There are three main classes of FP-based sensors: single FP sensors, FRET sensors, and BiFC sensors. Single FP sensors are FP barrels that are modified to become sensitive to an external stimulus. The first developed FP sensors were pH sensors such as super-ecliptic pHlorin.[76] FP pH sensors are generally based on the inherent pH sensitivity of the chromophore. These FPs can be mutated to increase their fluorescence dynamic range by optimizing the Hill coefficient and pK_a of fluorescence. A typical application of FP-based pH sensors is monitoring of endo- or exocytosis.[77]

As FPs are generally insensitive to most external stimuli (with pH being an obvious exception), creating a sensor requires that a new sensitivity be engineered into the FP. This can be achieved by genetic fusion of an FP to one or more proteins that are sensitive to the external stimuli in a precise way to influence the fluorescence. In the case of Ca^{2+}, calmodulin and an M13 peptide are fused onto an FP barrel. When calmodulin binds Ca^{2+} this induces a conformational change. This change allows M13 to bind to calmodulin and the resulting conformational shift translates into the attached FP changing its fluorescence. This system has been used to create Ca^{2+} sensors across the visible spectrum.[78,79] While FP-based Ca^{2+} sensors

and pH sensors are the most successful examples of single FP-based sensors, many different FP sensors have been built to detect other ions such as Zn^{2+},[80] or even neurotransmitters.[81]

While single FP biosensors are very useful due to their large dynamic range, their construction is not trivial. If a new sensor is needed, the fastest and easiest way to create it is to construct a FRET sensor. If a PPi or binding event changes the distance between the proteins N- and C-termini, then a pair of FPs can be fused to these termini and the movement or interaction is very likely to result in a measurable change in FRET efficiency. Many FPs have been evolved expressly to be used as FRET pairs in this way.[32,82] With the relative ease with which these sensors can be created, many FP FRET sensors now exist.[83]

In a similar fashion to FRET sensors, many protein fragment complementation assays such as BiLC and ddFPs can be used to make sensors.[58] These sensors often have the advantage of having larger dynamic ranges. However, they have the disadvantage of losing the "molecular ruler" characteristic of FRET. That is, they have binary signal generation and so they are either on or off with nothing in between. Generally speaking, BiFC is not useful for dynamic sensing due to its irreversible nature, but is generally useful for identifying PPi.

1.6.2 FP Optogenetic actuators based on FPs

Optogenetic actuators represent a diverse class of tools[84] and so I will focus only on the types that are specifically relevant to this thesis.

1.6.2.1 Photosensitizers

Photosensitizers are FPs that upon illumination induce a photochemical change in another molecule. The FP KillerRed is a photosensitizer, when illuminated its chromophores

excited state produces singlet oxygen and other reactive oxygen species (ROS) from water and oxygen.[85] These ROS can be used to inactivate or damage nearby proteins and other molecules in a process called chromophore assisted light inactivation (CALI).[86] Inactivation occurs by the ROS oxidizing aromatic residues, cleaving peptide bonds, and crosslinking side chains.[87] The damage caused is controlled by diffusion and is therefore indiscriminate, meaning that strong illumination typically results in much damage and can result in cell death. This has been harnessed in photodynamic therapy to use light to kill cancer cells. This is achieved by adding a photosensitizer and illuminating only the tumor restricting the damage to the tumor.[88]

1.6.2.2 Photooligomerization

Photooligomerization is using light to control the oligomeric state of proteins. The only FP example of this is Dronpa[89] which is a photoswitchable FP from a *Pectiniidae* coral. Dronpa's photoswitching to the green fluorescent state occurs concurrently with oligomerization using 390 nm illumination and switches to the dark state simultaneously with monomerization using 490 nm light.[90] This controllable oligomerization has been harnessed to activate or disable fused proteins such as the hepatitis C virus NS3-4A protease and Cdc42 GEF intersectin.[90]

1.7 Protein engineering

In order to create the sort of genetically encoded tools discussed in this thesis, we need to be able to create, modify and test many different variations of various FPs. By far the most practical method is to work at the DNA level in plasmids then grow and test the variants in the *E. coli* bacteria. *E. coli* grows rapidly, with a doubling time of 30 minutes, accepts transformation easily, and can be controlled by antibiotics. The trade-off is that *E. coli*

functions sufficiently different from cultured human cell lines and model organisms and so proteins developed in *E. coli* often perform differently when transferred to these other systems. Creating new DNA sequences and testing the resulting proteins is the basis of protein engineering.

1.7.1 Library generation

While protein engineering is no longer in its infancy, it also cannot yet be considered to have reached the rationality of adulthood. Accordingly, the trial and error phase of adolescence is still the best way to describe the current state of the field. As such, the best methods are currently based around creation of a large pool or variants and screening for the property we desire. Creating this pool of variants is described as library generation and the success of the engineering process depends on the quality and scope of the library we produce, and our ability to thoroughly screen it for the desired property. There are two general types of libraries: irrational ones which make no assumptions and randomly mutate a whole gene, and semi-rational ones where we use some experimental insight to target specific positions in the protein for randomization.

1.7.1.1 Irrational library generation

Our most used library type is based on error prone polymerase chain reaction (EP-PCR). We use the power of PCR to create many copies of a gene using a polymerase that tends to introduce errors in the copied DNA. Specifically, we use the DNA polymerase from *Thermus aquaticus* (Taq) and increase its error rate by supplementing it with Mn^{2+} which can replace its usual cofactor Mg^{2+}. Incorporation of Mn^{2+} increases the error rate of Taq to 1 every

few thousand bases. After an error prone PCR reaction, each DNA copy contains 0-3 base changes creating a library with several thousand different mutation combinations.

Once several beneficial mutations have been identified, staggered extension process (StEP) PCR[91] is a technique that can be used to randomly combine these mutations to look for synergy between them. If the number of different mutations across the chosen templates is N, StEP PCR allows for N! possible combinations which is 40 320 for 8 mutations. StEP PCR is performed by using a short elongation time in the PCR program resulting in incomplete copies of the template strand. In the next round the incomplete copy can be extended using a different template strand combining the mutations from the two templates. This process repeats many times allowing the mutations to combine in all possible combinations.

1.7.1.2 Semi rational library generation

Site directed saturation mutagenesis is a technique where an amino acid that has been determined to be important is randomized to all possible amino acids to identify the best one, or help elucidate its role. This can be achieved several ways but all of them require a synthetic DNA primer that containing a degenerate codon sequence such as NNK. N refers to any DNA nucleotide and K refers to thymine (T) or guanine (G). An NNK codon codes for all 20 common amino acids and only one of the three stop codons in the included 32 possible base combinations. Such synthetic DNA fabrication is now cheap and fast. Incorporation of the NNK codon into a gene can be achieved through overlap PCR or a process called QuikChange. Both processes use a synthetic DNA primer and extend it using polymerases incorporating the NNK site into the now replicated gene. The gene is then ligated into a plasmid which is then used to transform cells for testing. The same techniques can also be used for insertions,

deletions, and specific amino acid changes by designing primers that encode those changes. Modern molecular biology allows a protein engineer to change a protein in any conceivable way.

1.7.2 Circular permutation

Circular permutation is the process of linking the original N- and C- termini together and introducing new termini at a new location in the protein sequence. In practice, this involves cutting the gene encoding the protein into two parts and reversing their order. This forcefully places the old N- and C- termini adjacent to each other which would generally prevent the protein from folding correctly. To remedy this, a linker is added that connects the old termini to allow the protein to maintain its functional fold.

Moving the N- and C- termini is useful as it can change how a protein interacts with its environment. In the case of single FP calcium sensors, the N- and C- termini are placed in the middle of the seventh β-strand immediately beside the chromophore. This change exposes the chromophore to the environment such that the conformational change induced by the calcium binding to directly affect the chromophore and therefore fluorescence.

1.7.3 Screening

Once a suitable library has been created, examining its contents and selecting novel variants is a process called screening. As the libraries can contain many thousands of members they must first be divided to allow them to be individually observed. If we place a library into plasmids carrying an antibiotic resistance gene we can transform *E. coli* such that each transformed bacteria carries one plasmid and thus one mutated gene from the library. These *E.*

coli are then diluted and spread on Petri dishes that allow each individually transformed cell to grow into a single colony of clones. In the case of FPs, these colonies can be used for screening.

1.7.3.1 Brightness

Brightness is important in nearly all FP uses. Our lab screens for brightness by imaging of colonies of bacteria expressing the library of FPs variants in Petri dishes. The entire plate is illuminated using light from a 300 W xenon arc lamp which is filtered to excite each specific FP. Fluorescence is detected using long pass filter goggles or a camera with an appropriate emission filter. The colonies that appear brightest are picked and grown in liquid culture. This liquid culture is then used to extract purified protein and DNA for further characterization and sequencing and then the best variants are used as the starting material for the next library creation step.

1.7.3.2 Monomerization

As many FPs are natively weak dimers such as eGFP, or obligate tetramers such as DsRed,[92] using such an FP as a fluorescent tag can alter the function of the fused protein by forcing them into close proximity. As such, the process of monomerization has been performed several times to various FPs to make them research-ready.[19,93] As the oligomeric state of a protein is not visible at the colony level, screening has to be performed using polyacrylamide gel electrophoresis (PAGE) or gel chromatography to separate the different oligomeric states. This restricts the size of the library that can be reasonably screened. As such, site-directed mutagenesis must be targeted to locations which come into contact across the protein-protein interfaces. By adding the same charged residues to both sides of the interface, the side chains

can electrostatically repel each other and thereby break the interface. Native PAGE can be used to monitor the success of each particular mutation for decreasing the amount of protein in the oligomeric state. This process invariably diminishes the fluorescence of the FP requiring further brightness screening as described above.

1.7.3.3 Photostability

Screening for photostability involves photobleaching entire Petri dishes of colonies and picking the brightest colonies after photobleaching. While superficially simple, this becomes difficult in that achieving the high intensity and even illumination over an entire Petri dish without killing the cells due to the heat or phototoxicity is non-trivial. White light illumination of 1 W/cm^2 results the death of 50% of colony forming units in 1 hour.[85] Chapter 3 discusses this in greater detail.

1.7.3.4 Sensing

Screening for sensing capabilities involves measuring the fluorescence at high and low levels of the target analyte of interest. As cells exist by maintaining internal conditions different from their environment, simply adding an analyte typically does not alter the internal conditions in a cell. In *E. coli*, Ca^{2+} indicators have been targeted to the periplasm which is more sensitive to the external environment allowing on-plate screening by added solutions of Ca^{2+} or EDTA to achieve high and low levels of Ca^{2+}, respectively.[78] However this process is not very robust, as the periplasm, like the cytoplasm, actively pumps analytes in and out limiting our ability to externally modify it. As such, most sensor screening is performed by individually testing a smaller number of variants in cell lysate in which the analyte

concentration can be controlled. The small number of variants chosen for individual screening can be isolated from larger libraries by using a prescreen for brightness. Accordingly, sensor screening is often a multi-stage process: a prescreen for an easy to measure trait such as brightness, followed by more accurate secondary screens that test for a harder to measure traits such as analyte sensitivity or photostability.

1.8 The scope of the thesis

FPs have become a mainstay of biological and biomedical research and researchers continue to push the limits of what can be achieved using FPs. In this thesis we describe our efforts to create a new generation of FPs with improved properties and functionality.

In Chapter 2, we describe the creation of a new set of FPs called the vine Tomatoes. These include the first green tandem dimer FP GGvT, the current brightest red tandem dimer FP RRvT, and the unique green and red tandem dimer FP GRvT. We analyze the photophysical properties of this new series of FPs and demonstrate their potential in mammalian cell imaging.

In Chapter 3, we describe the development of a robotic photostability screening system which we used to evolve an improved yellow FP designated Citrine2. We demonstrate its photostability in mammalian cells and in vitro. In the process we determined a variant dependent link between FP concentration and photostability and the independence of photostability and fluorescence lifetime.

In Chapter 4, we attempted to develop an improved version of the optogenetic actuator PhoCl, a photocleaving FP. We describe the dynamic fluorescent properties of PhoCl during photocleavage and dissociation and subsequently develop several screening systems designed to measure the photocleavage in colonies and analyze their utility.

In Chapter 5, we describe the creation of a potential second orthogonal photocleavable FP called SplitOr based on the photoconverting FP PSmOrange2. We describe the process of circularly permutating PSmOrange2 and its fluorescence rescue. In addition, we tested various photoconversion methods to confirm its orthogonality to PhoCl and identify the optimal conversion conditions.

In Chapter 6, we summarize the thesis and look at the future prospects of these new variants and screening methods.

A tandem green-red heterodimeric fluorescent protein with high FRET efficiency

2.1 Introduction

The development of improved fluorescent proteins (FP) for live cell fluorescence imaging has primarily focused on engineering variants with brighter intrinsic fluorescence, more efficient folding, and minimal oligomeric interactions. As the majority of wild-type FPs are tetrameric, a key step toward development of an improved FP is disruption of the oligomeric interaction interfaces between β-barrel subunits to convert the protein first to a dimer and ultimately to a monomer.[93] This process invariably decreases brightness and folding efficiency, and further engineering is required to rescue these desirable properties.[19,93–95]

One approach for engineering an effectively monomeric FP from a dimeric FP is to create a genetically fused tandem homodimer that retains the stabilizing interactions of the dimer interface (Figure 2.1). At the protein level, a tandem homodimer is composed of a single polypeptide that encodes two copies of the same protein fused together with a linker long enough to permit the dimerization to occur. Such a protein is effectively monomeric, but doubled in both size and brightness compared to a truly monomeric FP. This process has been used to convert the dimeric progeny of the *Discosoma* sp. red FP (DsRed),[20] dimer2[93] and dTomato,[14] into the tandem dimers tdimer2[93] and tdTomato,[14] respectively. Likewise, a dimeric version of the *Heteractis crispa* red FP (HcRed)[96] and an *Entacmaea quadricolor* derived red FP dimer tdKatushka2,[97] were made into tandem dimers. Notably, tdTomato is one of the brightest and most widely used red FPs for *in vivo* imaging applications.[98–101]

Figure 2.1 Diagram representing the nomenclature used to identify vine Tomato variants.

Directed evolution of pairs of FPs fused together has resulted in enhanced photophysical characteristics. FRET pairs have previously been evolved as tandem heterodimers.[82,102] Several FRET pairs have been enhanced by modifying the dimerization strength to improve the dynamic range of the FRET signal.[103–105] Similarly, bioluminescence resonance energy transfer (BRET) pairs, where FPs are fused to luciferases for enhanced luminescence, such as Nano-Lantern,[106] aequorin-GFP,[107] and BAF-Y,[108] have been evolved as fusion proteins.

The chromophore formation mechanism of DsRed is thought to be a branched pathway, where one branch irreversibly forms a green chromophore and the other branch irreversibly forms a red chromophore at a red-to-green ratio of about 1.0 to 1.5.[13,31,109,110] Mutations of DsRed that bias it towards the green pathway have previously been reported.[111,112] When both green and red chromophores are present in a given DsRed tetramer, highly efficient Förster resonance energy transfer (FRET) from the green donor to red acceptor occurs.[113] Yet another outcome for a DsRed subunit, or any other FP, is that it can fail to form either a green or a red chromophore and end up in a dark state.[114,115]

In this work, I describe new green-green and red-red DsRed-derived tandem dimers, as well as a green-red heterodimer composed of a DsRed variant that predominately forms the green chromophore fused to a DsRed variant that predominately forms the red chromophore.

To reflect the fact that these are variants of dTomato, and that two different copies are included in a single polypeptide, I refer to this series of fused homodimeric and heterodimeric proteins as “vine Tomatoes” (vT). Accordingly, I abbreviate the green-green tandem homodimer as GGvT, the red-red tandem heterodimer as RRvT, and the green-red tandem heterodimer as GRvT.

2.2 Results and discussion

To engineer a green variant of dTomato (GdT), I sought to bias the chromophore maturation to the green pathway and maximize the brightness by introducing mutations that increased the extinction coefficient (Ɛ) and quantum yield (Φ). I initially introduced the A72M and V106A substitutions that have been reported to increase the amount of green fluorescence.[59,112] To further enhance the green fluorescence and increase the green- to-red ratio, I used 10 iterative rounds of directed protein evolution by a combination of error prone PCR (EP-PCR), site-specific codon randomization, and bacteria colony-based screening.[95] The evolutionary path from dTomato to GdT (Table 2.1), along with the improvements in brightness and green-to-red ratio observed in each round, is provided in Figure 2.2. Mutations Q65L, F66L, A72M and V108A are residues with side chains that are buried within the interior of the β-barrel.

Table 2.1 Amino acid mutations in the DsRed-derived Tomato family

Fluorescent Protein	Linker	First half mutations[a]	Second half mutations[b]
tdTomato	HGTGSTGSGSSGTASSEDNN MA	-	-
RRvT	MDELYKGSTGSGSSG(T→P)[c]	K122N	V8A R18S K93R L114I M125V I211V
GRvT	GSTGSGSSGT	Q65L F66L A72M E95V V108A K122E I211R G228D	L114I
GGvT	MDELYKGSTGSGSSG(T→P)	Q65L F66L A72M E95V V108A K122E I211R G228D	Q65L F66L A72M E95V V108A K122E I211R G228D
GdT	Not Applicable	Q65L F66L A72M E95V V108A K122E I211R G228D	Not Applicable

[a] All mutations are relative to tdTomato. [b] Numbering for second half mutations starts at 1 after the linker listed above. [c] (T→P) indicates a Thr to Pro mutation in the unnumbered linker region.

Table 2.2 Photophysical properties of the DsRed-derived Tomato family

Fluorescent Protein	λ_{ex} [b] (nm)	λ_{em} [c] (nm)	ε [d] ($M^{-1}cm^{-1}$)	Φ	Brightness ($\varepsilon \times \Phi/10^3$)	pK_a [e]
tdTomato [a]	554	581	138 000	0.69	95	4.7
RRvT	556	583	134 000	0.88	120	3.9
GRvT (Red)	557	583	71 000	0.97	69	2.5
(FRET)	477	583	89 000[f]	0.41	37	3.3
GGvT	476	500	91 000	0.14	13	3.8
GdT	476	500	59 000	0.12	7.1	3.8

[a] Previously reported data.[14] [b] Wavelength of maximum excitation. [c] Wavelength of maximum emission. [d] Extinction coefficient at λ_{ex} as determined by the alkali-denaturation method. [e] pH at which fluorescence is at half of its maximum value. [f] Sum of green and red chromophore absorbance at 477 nm, 66 000 $M^{-1}cm^{-1}$ and 23 000 $M^{-1}cm^{-1}$ respectively.

These mutations are likely contributing to improved folding efficiency and changes in side chain packing in the protein interior to better accommodate the green chromophore. Late in the evolutionary process, the initial V106A mutation reverted to valine, indicating that it had switched from being beneficial to detrimental. The E95V, K122E, I211R, and G228D mutations are outward facing and may improve the folding or overall stability of the β-barrel. Residue 211 was mutated twice randomly (I211N then N211Y) and a third time by site-specific codon randomization (Y211R). Notably, the dimerization interface was left unchanged through this evolutionary process. Despite this extensive directed evolution, the final version of GdT has only 21% of the inherent brightness of enhanced green fluorescent protein (eGFP)[92] due to a low quantum yield (GdT $\Phi = 0.12$; eGFP $\Phi = 0.60$) (Table 2.2).

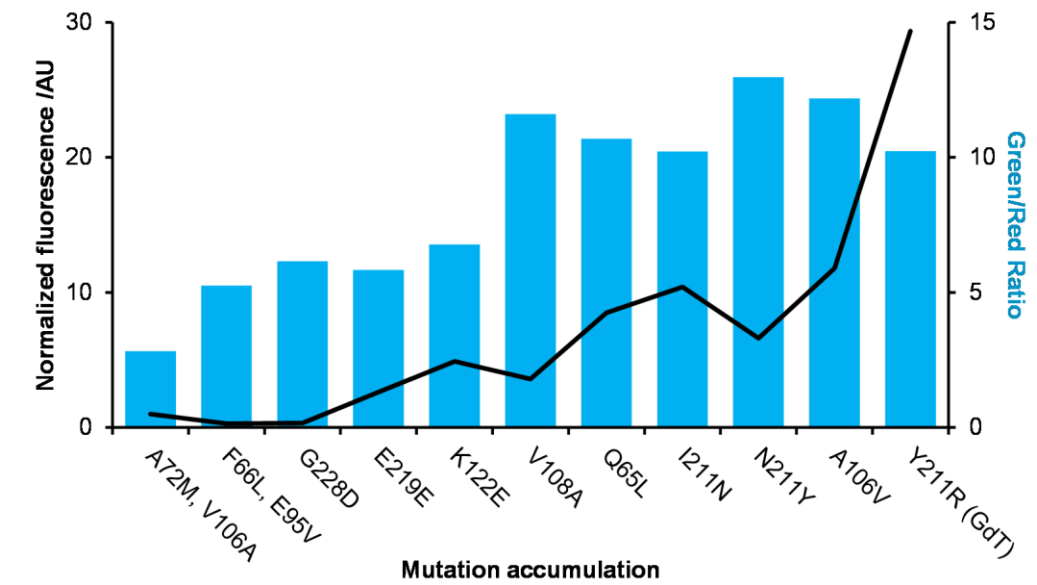

Figure 2.2 Evolutionary path to create a green variant of dTomato. Starting from the initial dTomato A72M/V106A variant (far left column), the evolutionary pathway proceeds through ten rounds from left to right. The black line represents the fluorescence intensity with λ_{ex}=460 nm and λ_{em}=500 nm measured from identically treated liquid cultures of bacteria, normalized to the first variant's fluorescence. The blue histogram represents the green-to-red ratio for the best variant in each round, where green is the fluorescence intensity at 500 nm, and red is the fluorescence intensity at 580 nm, λ_{ex}=460 nm.

Midway through the development of GdT, a first generation of vTs (*i.e.*, tandem dimers) was assembled from synthetic DNA fragments encoding intermediate GdT variants and dTomato (Figure 2.3). To keep this description as simple as possible, I have chosen not to provide unique names or numbers for the intermediate variants. Rather I use the vT names (*i.e.*, GGvT, RRvT, GRvT) to generically refer both to intermediates and the final products. To assemble the vTs, genes for the first (5') and second (3') FP domains of GGvT and RRvT were individually synthesized such that they would be "codon differentiated" with distinct and different DNA sequences. The initial 5' and 3' versions had 78.4% and 69.5% sequence identity, for GGvT and RRvT, respectively. GRvT was assembled using the GdT gene (5') and

the 3’ domain of RRvT (3’). This codon differentiation was done to facilitate site directed mutagenesis on either the 5’ or 3’ FP domains, to minimize the occurrence of homologous recombination in bacteria, and to minimize the switching of partially extended primers between the 5’ and 3’ FP templates during PCR. Such switching during PCR tends to produce half-length hybrid genes that must be separated from full-length tandem genes by gel electrophoresis. The 5’ and 3’ domains of all three vT proteins start with the *Aequorea* GFP-derived N-terminal sequence MVSKGEE.[14] Likewise, all domains of the three vT proteins, with the exception of the 5’ domain of GRvT, end with the GFP-derived C-terminal sequence (MDELYK). A 10-residue linker was used to connect the 5’ and 3’ domains in all vTs.

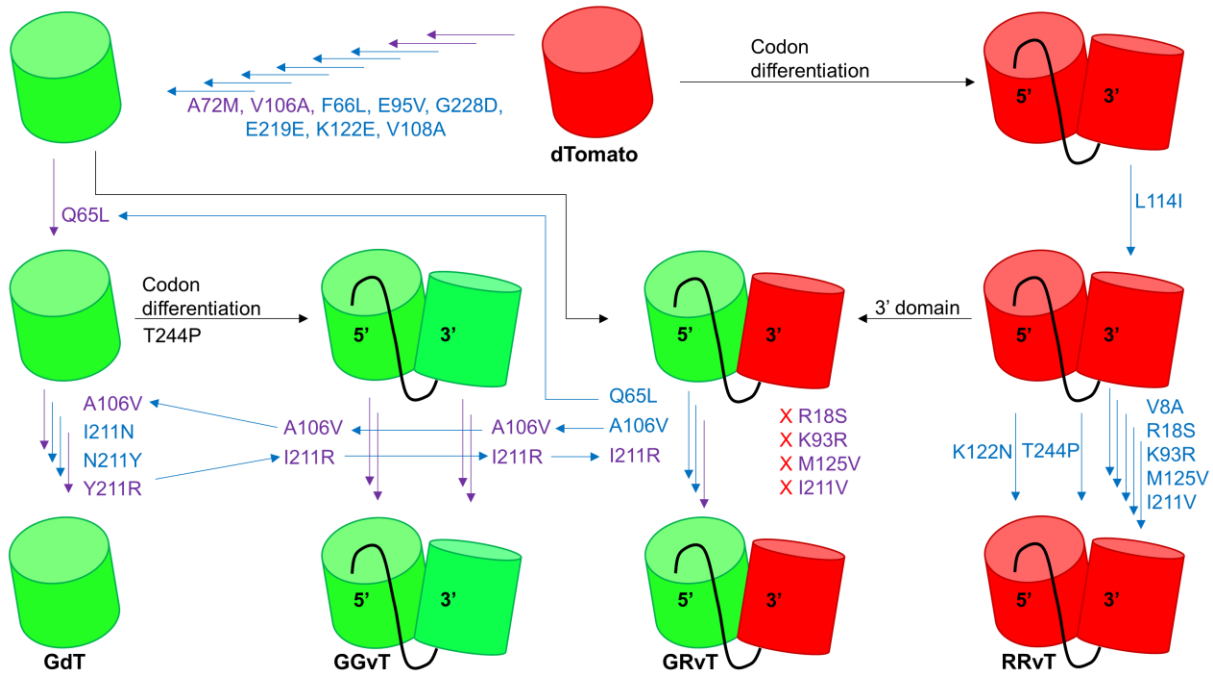

Figure 2.3 Development of vine Tomatoes (vTs). Purple and blue arrows represent mutation events from site-directed mutagenesis and EP-PCR, respectively. Black lines show protein fusion and codon differentiation events. Arrows connecting mutations originate from the source or sources of the mutation's discovery. Red X's indicate mutations that were tested but failed to improve the variant.

The initial GGvT, GRvT, and RRvT templates were subjected to directed protein evolution using EP-PCR. Notably, EP-PCR of the vT genes typically produced both full-length and half-length PCR products, indicating that the codon differentiation was not entirely successful at preventing template switching during PCR. When a beneficial mutation was discovered for one FP domain, it was manually introduced into other domains of the same color, and only retained if the resulting protein was indeed brighter. A graphical summary of the engineering process is provided in Figure 2.3, the full sequences are provided in Figure 2.4, and the structural location of all mutations in the final vTs are provided in Figure 2.5.

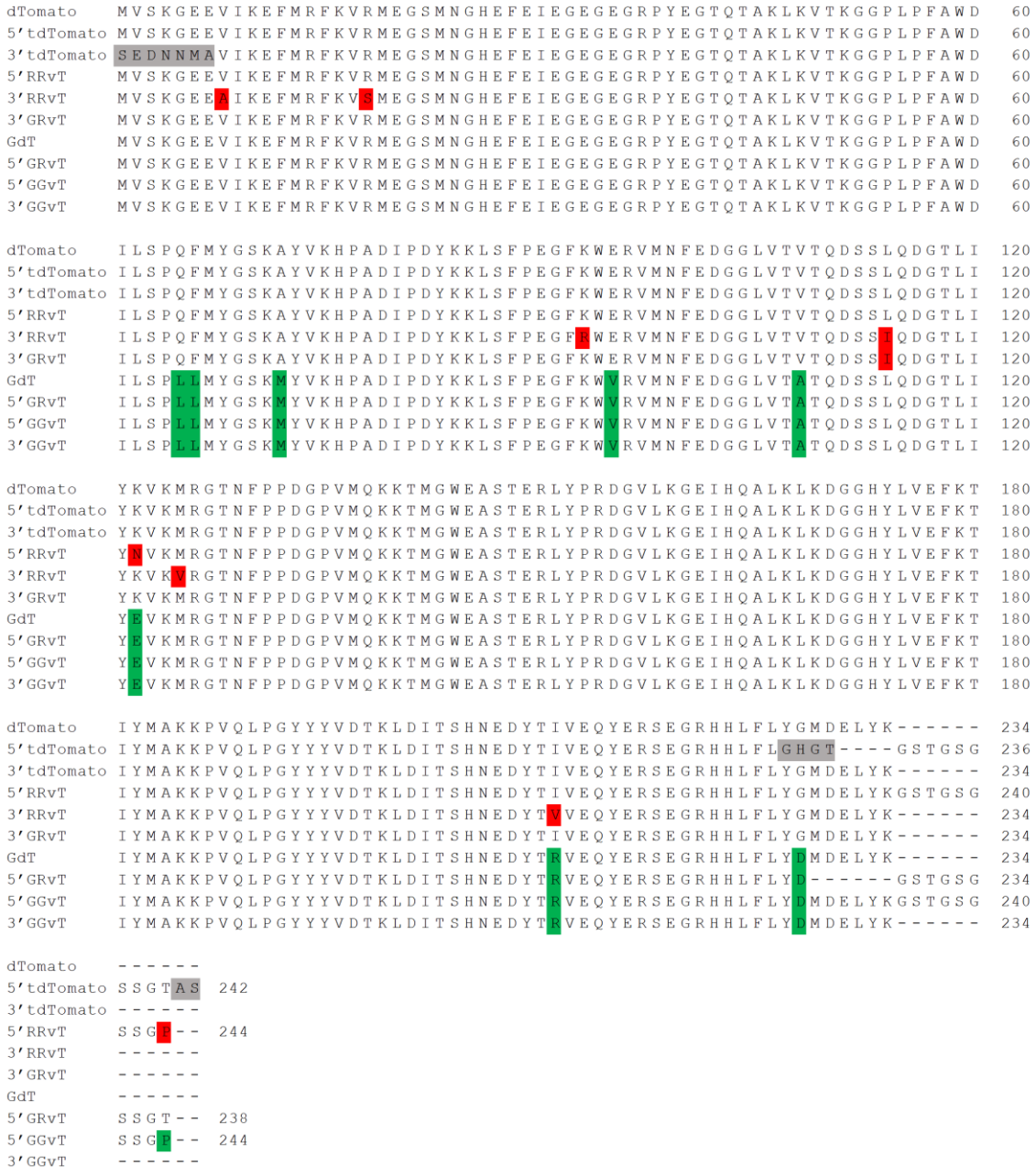

Figure 2.4 Sequence alignment of the vine Tomatoes and their parent sequences. The fused homo- and heterodimers are split into their 5' and 3' domains, with the last amino acid of the 5' domain immediately followed by the first amino acid of its respective 3' domain. Gray highlights represent the unused parts of the tdTomato interdomain linker. Red and green highlights show mutation locations and are colored to match the domain's fluorescence. Domains are sorted by color, with red domains above green domains.

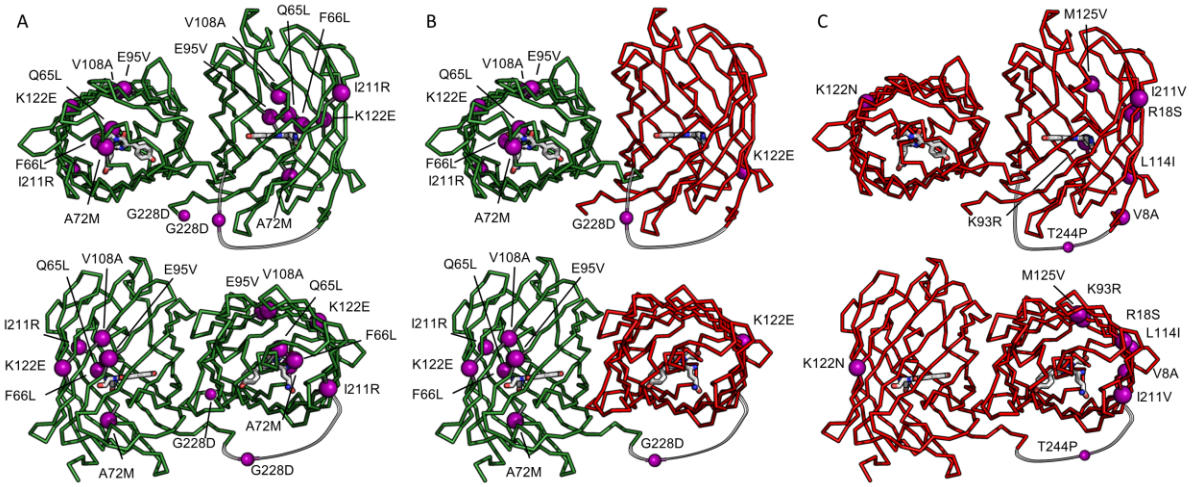

Figure 2.5 Modeled location of mutations in vT variants. A) GGvT, B) GRvT, and C) RRvT. Purple spheres represent the Cα of the mutated residue, and the gray line represents the linker between the tandem FP domains. The structural model is based on chains A and C of tetrameric DsRed (PDB ID: 1G7K).[116]

For the green variants, the Q65L and A106V mutations were initially discovered in the 3' domain of GRvT, and manually introduced into GGvT and GdT. Similarly, the I211R mutation was first discovered in GdT, and then transferred to both domains of GGvT. Screening a library in which position 211 in the 5' domain of GRvT was randomized led to the rediscovery of this mutation. All beneficial mutations in green domains were beneficial to all other green domains. In contrast, mutations discovered in one red domain were not necessarily beneficial to all other red domains. The dimer interface in all vTs remained unchanged.

Relative to tdTomato, the initial codon differentiated RRvT gene had substantially reduced fluorescence in *E. coli* colonies, possibly due to a lower expression level of the FP. Fortunately; this fluorescent brightness was recovered after several rounds of directed evolution (Table 2.1). Recovery required the accumulation of one beneficial mutation in the 5' domain (K122N), and six beneficial mutations in the 3' domain (V8A, R18S, K93R, L114I, M125V, I211V), and one Thr to Pro mutation in the interdomain linker. This Thr to Pro

substitution was incorporated into the GGvT linker. In the 5' domain, the mutation K122N is near the mid-section of the β-barrel and has its side chain directed to the protein exterior. In the 3' domain, M125V is the only mutated residue that has its side chain directed towards the protein interior, where it sits in a hydrophobic pocket near, but not immediately adjacent to, the chromophore (Figure 2.5C). Due to a substantial increase in quantum yield, the inherent brightness of RRvT (Φ = 0.88) is 120% that of tdTomato (Φ = 0.69). GGvT, while still rather dim relative to eGFP, has an inherent brightness that is 180% of GdT (Table 2.2).

While the 5' green FP domain of GRvT is identical to the green domains of GGvT, the 3' red FP domain is distinct from both domains of RRvT. Attempts to transfer beneficial mutations (R18S, K93R, M125V, and I211V) from the 3' domain of RRvT to the red domain of GRvT did not result in the identification of improved variants. While a structural rationale for this apparent incongruity remains unclear, it is likely a consequence of the difference in how RRvT and GRvT libraries were screened and the criteria for identifying improved variants. RRvT libraries were imaged using a 510-560 nm excitation filter and a 660-700 nm emission filter to identify the brightest red FPs. In contrast, GRvT libraries were imaged using a 450-490 nm excitation filter and a 660-700 nm emission filter to find variants that exhibited the brightest red fluorescence due to FRET.

Due to its heterodimeric structure, characterization of GRvT was substantially more complex than characterization of GGvT and RRvT. Much of this additional complexity arose from the occurrence of intramolecular FRET and the fact that incomplete FP chromophore maturation led to substantial fractions of proteins with only a green, or only a red, chromophore. In the remainder of this chapter, I describe our efforts to decouple these complicating factors and provide a complete description of the spectral properties of GRvT.

As GRvT contains two different chromophores, a single excitation spectrum, emission spectrum (Figure 2.6B), quantum yield, and extinction coefficient, are insufficient for describing the most relevant properties of this FP. For GRvT it was necessary to define and measure two distinct quantum yields: the quantum yield for excitation of the green chromophore at 450 nm (Φ_{FRET} = 0.41), and the quantum yield for excitation of the red chromophore at 530 nm (Φ_{red} = 0.97). The value of Φ_{FRET} is the sum of three pathways: green emission resulting from direct excitation of the green chromophore; red emission resulting from direct excitation of the red chromophore; and red emission resulting from green-to-red FRET.

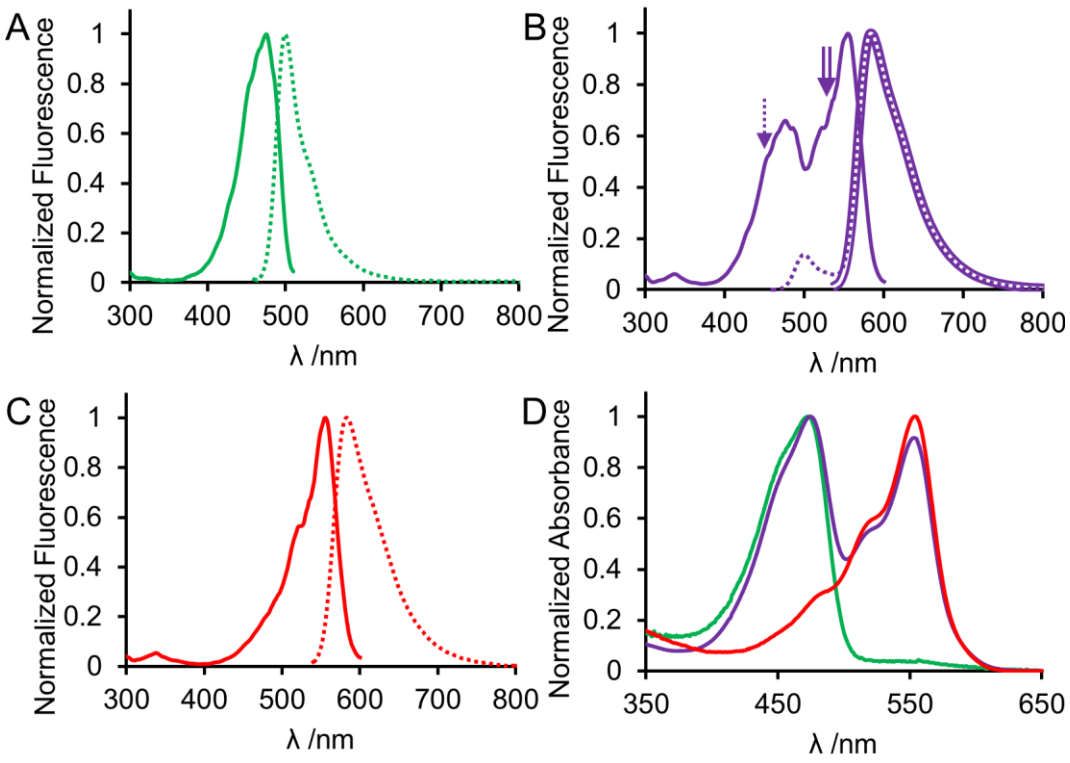

Figure 2.6 Normalized fluorescence spectra of the three vTs. A) Excitation and emission spectra for GGvT: solid line is the excitation spectrum (emission at 520 nm), and dotted line is an emission spectrum (excitation at 450 nm). B) Excitation and two emission spectra for GRvT: solid line is the excitation spectrum (emission at 610 nm), dotted line is an emission spectrum (excitation at 450 nm), double line is an emission spectrum (excitation at 530 nm). Arrows represent the excitation wavelength for the 450 and 530 nm emission spectra. C) Excitation and emission spectra for RRvT: solid line is the excitation spectrum (emission at 610 nm) and dotted line is an emission spectrum (excitation at 530 nm). D) Normalized absorbance spectra: GGvT (green), GRvT (purple), and RRvT (red).

To determine the extinction coefficients for the green and red chromophores of GRvT, the alkali denaturation method was used.[13] Alkali denaturation exposes the chromophores to the solvent allowing the concentration of the chromophores to be determined rather than the protein concentration. This has been previously reported for GFP and DsRed,[13] alkali denaturation of GGvT and RRvT result in the formation of similar species with absorption

maxima close to 450 nm (RRvT at 453 nm; GGvT at 442 nm). Alkali denaturation of GRvT necessarily produces a convolved spectrum (maximum absorbance at 447 nm) that can be fit as a linear combination of the GGvT and RRvT spectra (Figure 2.7). To achieve this the GRvT, GGvT, and RRvT spectra are normalized to their max absorbance. The GGvT and RRvT spectra are then each scaled with one variable each which represents the proportion of the each colour of chromophore. The two scaled spectra are summed and subtracted from the GRvT spectra. The two scaling variables are then modified to minimize the residual difference between summed spectra and the GRvT spectra. Applying this procedure, I determined the extinction coefficients for the red (ε = 71 000 $M^{-1}cm^{-1}$) and green (ε = 66 000 $M^{-1}cm^{-1}$) domains of GRvT.

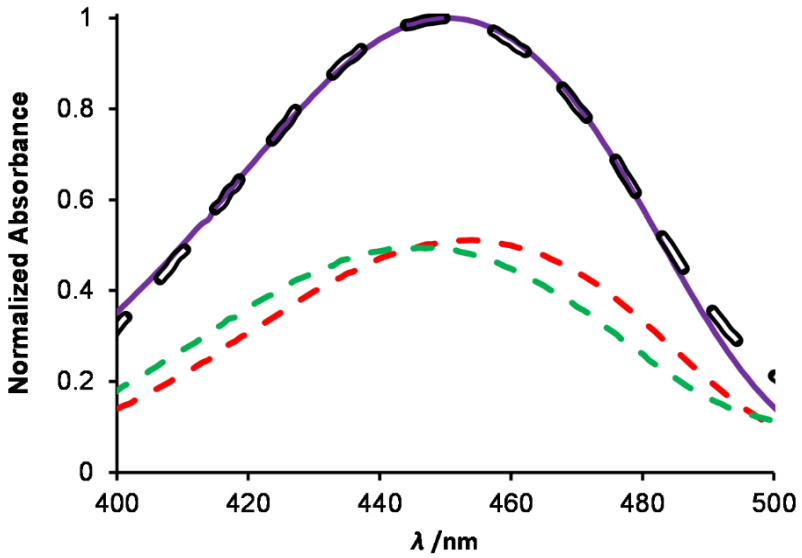

Figure 2.7 Fitting of alkali denaturation of GRvT. The purple line is the denatured absorbance of GRvT. The dashed green and red lines represent the scaled denatured absorbance of GGvT and RRvT, respectively. They are scaled such that, when summed, they fit the denatured absorbance of GRvT. The summed spectrum is represented by the black rectangle line.

The green chromophore of GRvT has a higher quantum yield than GGvT ($\Phi_{FRET} = 0.41$ versus $\Phi_{green} = 0.14$) – a difference that is most likely due to the presence of a higher quantum yield FRET acceptor in GRvT. Similarly, the red chromophore of GRvT has a higher quantum yield than RRvT ($\Phi_{red} = 0.97$ versus $\Phi_{red} = 0.88$), but this cannot be as easily explained. To further probe these differences, variants with one chromophore knocked out were prepared by mutating each of the GRvT's chromophore tyrosines to serines. The resulting variants were designated GrvT and gRvT, with the lower case representing the disrupted chromophore. GrvT had the same quantum yield as GGvT (0.14) and gRvT had the same quantum yield as RRvT (0.88). This result indicates that the presence of the mature green chromophore increases the quantum yield of the red chromophore of GRvT, possibly due to subtle conformational changes induced by the difference in size and shape between the red and green chromophores. If this difference results in a change in the orientations of the residues of the dimer interface the red chromophore would be influenced.

Assuming that the GRvT dimer is identical to the corresponding dimer of the DsRed tetramer, it provides a system to fully describe the observed FRET efficiency in terms of distance and Förster radius. Using the DsRed crystal structure (PDB ID: 1G7K)[116] it is possible to determine the orientation of the donor and acceptor chromophores and calculate a theoretical value for the orientation factor (κ^2).[117] The coordinates of optimized green and red chromophore models[118] were fitted against the original DsRed chromophores.[119] The fit was based on the carbon atoms of the phenyl rings and overlaid with visual molecular dynamics (VMD) software.[120] We then used time-dependent density functional theory (TD-DFT)[121] to compute the absorption transition dipole moments for both chromophores (Figure 2.8).

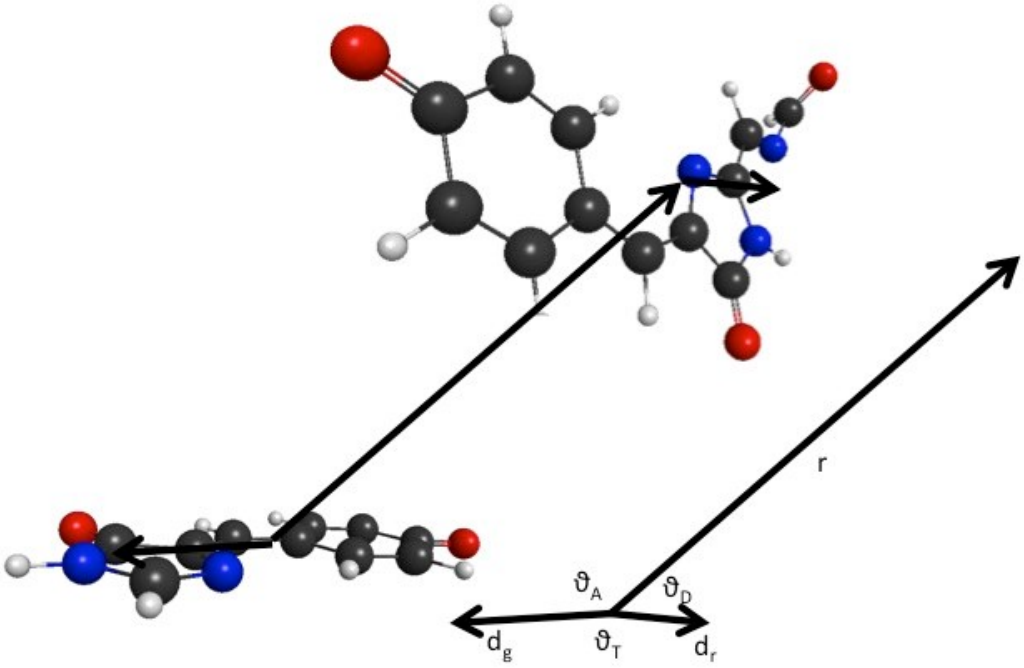

Figure 2.8 Modelling of the transition dipole orientation factor between the green and red chromophores. The optimized chromophore models are fit via the phenyl rings over the original chromophores as oriented in the crystal structure (PDB ID: 1G7K).[116] The phenyl rings should remain in a consistent location even after the mutations to favor the formation of the green chromophore. The vectors $\overrightarrow{d_g}$ and $\overrightarrow{d_r}$ lie along the transition dipole moments of the green and red chromophore models, respectively, while $\vec{r}$ joins the centers of both models. Optimization was done at the PBE0[122,123]/6-31G+(d,p)[124–128] level of theory in the gas phase. Transition dipole moments were determined at the B3LYP[129]/6-31G+(d,p) level of theory applying the polarizable continuum model (PCM)[130–133] with parameters for water.

Based on the approximation that the emission transition dipole of the donor lies along its absorption dipole, a model was made of the relevant dipoles overlaid on the chromophores (Figure 2.8). VMD was used to extract the geometry, which enabled calculation of the

theoretical κ^2. Based on the calculated κ^2 (1.41), and assuming a refractive index of 1.5 for protein,[134] I calculated the Förster radius for intramolecular FRET in GRvT to be 5.8 nm.[135] Using this value and the interchromophore distance of 2.72 nm, I calculate an expected FRET efficiency of 99%.

I fit GRvT's emission spectrum (Figure 2.6B) as a linear combination of GGvT and RRvT (Figure 2.6A and C). The area of the green emission peak is 7% of the total emission. However, with a FRET efficiency of 99% one would expect to observe a green emission peak area that is less than 1% of the total emission. Two factors could be causing the observed green emission to be substantially greater than expected. The first factor is that the red domain will occasionally form a green rather than red chromophore, as can be seen in RRvT (Figure 2.9). The second factor is the failed formation of the red chromophore of GRvT, which results in proteins with only a green chromophore, and thus no FRET. Based on the GRvT emission spectrum and the brightness of the green, red, and FRET components, I calculate that ~20% of GRvT proteins contain only a green chromophore. This ~20% is primarily composed of GRvT proteins where the 5' domain successfully formed a green chromophore and the 3' domain either formed a green chromophore or is dark.

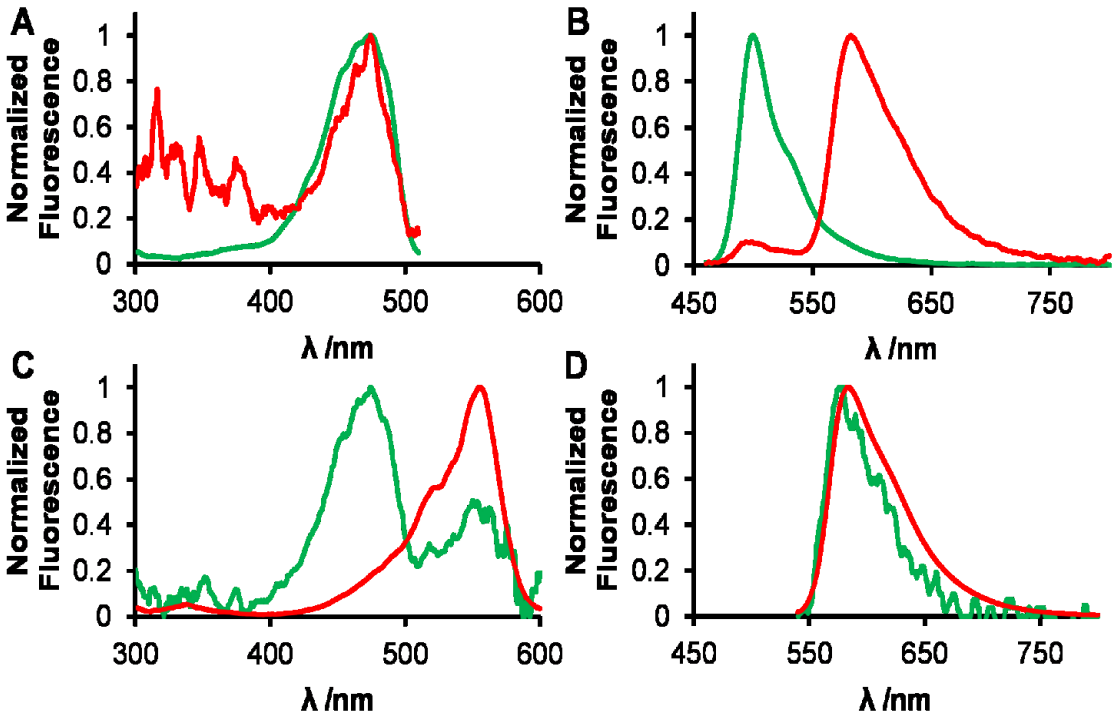

Figure 2.9 Additional vine Tomatoes fluorescence spectra. GGvT (green lines) and RRvT (red lines). A) Excitation spectra (emission at 530 nm). B) Emission spectra (excitation at 450 nm). C) Excitation spectra (emission at 610 nm). D) Emission spectra (excitation at 530 nm). Due to the branched and imperfect chromophore maturation pathway, RRvT and GGvT form small quantities of dim green and dim red chromophores, respectively. All spectra are normalized by each spectra's λ_{max}.

We performed single molecule FRET (smFRET) to independently quantify these populations in GRvT (Figure 2.10). Two peaks were observed: one containing ~80% of the molecules and producing ~100% red photons (relative to total red plus green photons), and one containing ~20% of the molecules and producing ~20% red photons. The larger fraction of the population presumably contains fully mature GRvT and proteins that only formed a red chromophore. The smaller fraction of the population presumably contains proteins that only produced a green chromophore. Proteins with only a green chromophore produce 20% red

photons due to bleed through into the red channel. This 80/20 population distribution is in good agreement with the GRvT emission spectrum (Figure 2.6B).

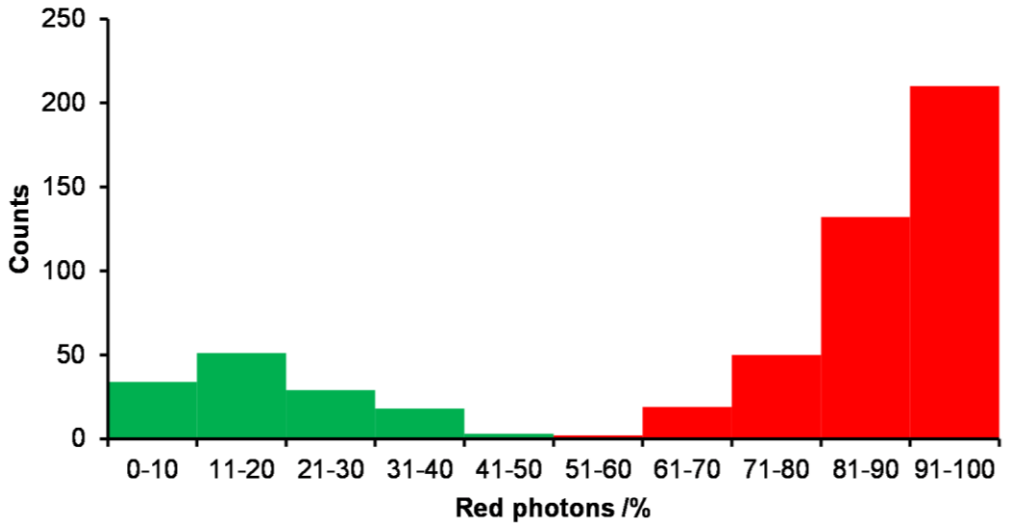

Figure 2.10 Results from single molecule FRET analysis of GRvT. A dilute solution of GRvT was analyzed for a total of 60 s, and the timecourse was divided into 60 000 1 ms bins. For each bin with a number of photons greater than the noise threshold we calculated % red photons = (# red photons) / (# red + # green photons)*100. The excitation lights wavelength was 458 nm. Photons between 470-505 nm were counted as green and all photons of longer wavelength than 585 nm were counted as red.

Two independent tests, smFRET and bulk fluorescence analysis, provided support for the conclusion that ~20% of GRvT proteins contain only a green chromophore and no red chromophore. To gain further support for this conclusion, I turned to a third approach to determine the percentage of GRvT domains that successfully form their chromophore. I used the calculated extinction coefficient at 280 nm (ε = 76 200 $M^{-1}cm^{-1}$) to determine the total protein concentration in a purified sample. It has been previously shown that the chromophore of an FP has a minimal contribution to the absorbance at 280 nm.[136] This concentration includes both the bright and dark protein molecules. For this same sample, I calculated the concentration

of the red and green chromophores using the extinction coefficients determined by alkali denaturation. Based on these concentrations, I calculated that 66% of the total protein molecules contained a red chromophore and 64% contained a green chromophore. Assuming a random distribution of bright and dark proteins between the green and red domains of GRvT, ~45% should contain both chromophores, ~20% only a green chromophore, ~20% only a red chromophore, and ~15% should be completely dark (Figure 2.11). Satisfyingly, this result is consistent with the smFRET and bulk fluorescence analyses that point to approximately 20% of the population of GRvTs containing only a green chromophore. Incorporating our estimates of the dark GRvT population into this analysis slightly decreases the size of the populations for the bright species, but does not change our overall estimates of the relative distribution of the various chromophore combinations. Notably, GRvT proteins, in which the red barrel is not dark but rather forms the green chromophore, would also contribute to this 'green only' population.

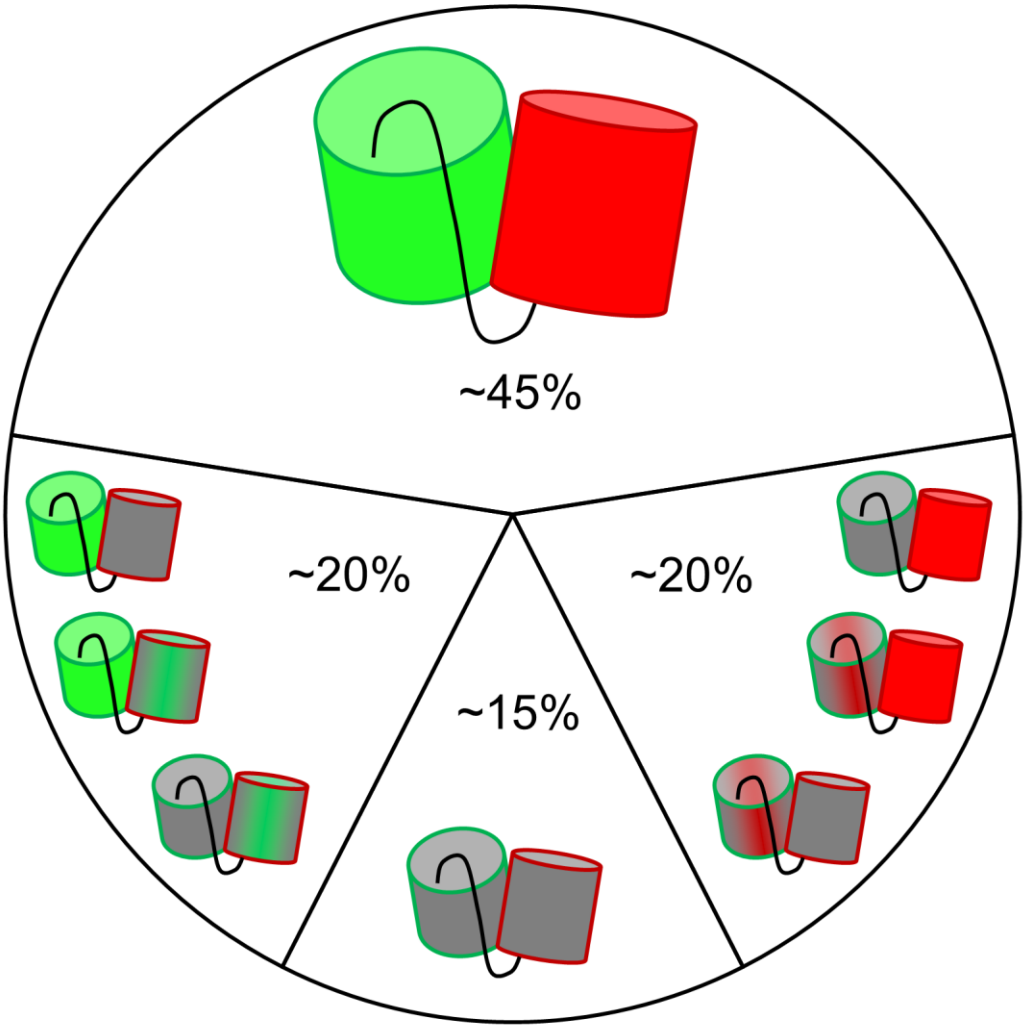

Figure 2.11 Distribution of chromophoric species for GRvT, assuming a random distribution of bright and dark proteins between the green and red domains of GRvT. The barrel outline represents the chromophore color for which the barrel was optimized. Each wedge represents the proportion of all GRvT proteins that share a spectral profile: green-red (45%), red (20%), green (20%), or dark (15%).

Due to its high FRET efficiency, GRvT functions as a long Stokes shift FP (LSSFP). GRvT's Stokes shift is 106 nm (from 477 nm to 583 nm) and has high brightness relative to other LSSFPs. However, the residual green emission and direct excitation of the red chromophore will complicate its use as an LSSFP. The broad excitation profile of GRvT makes

this variant potentially useful for multi-parameter imaging when used in conjunction with other reporters or optogenetic tools.

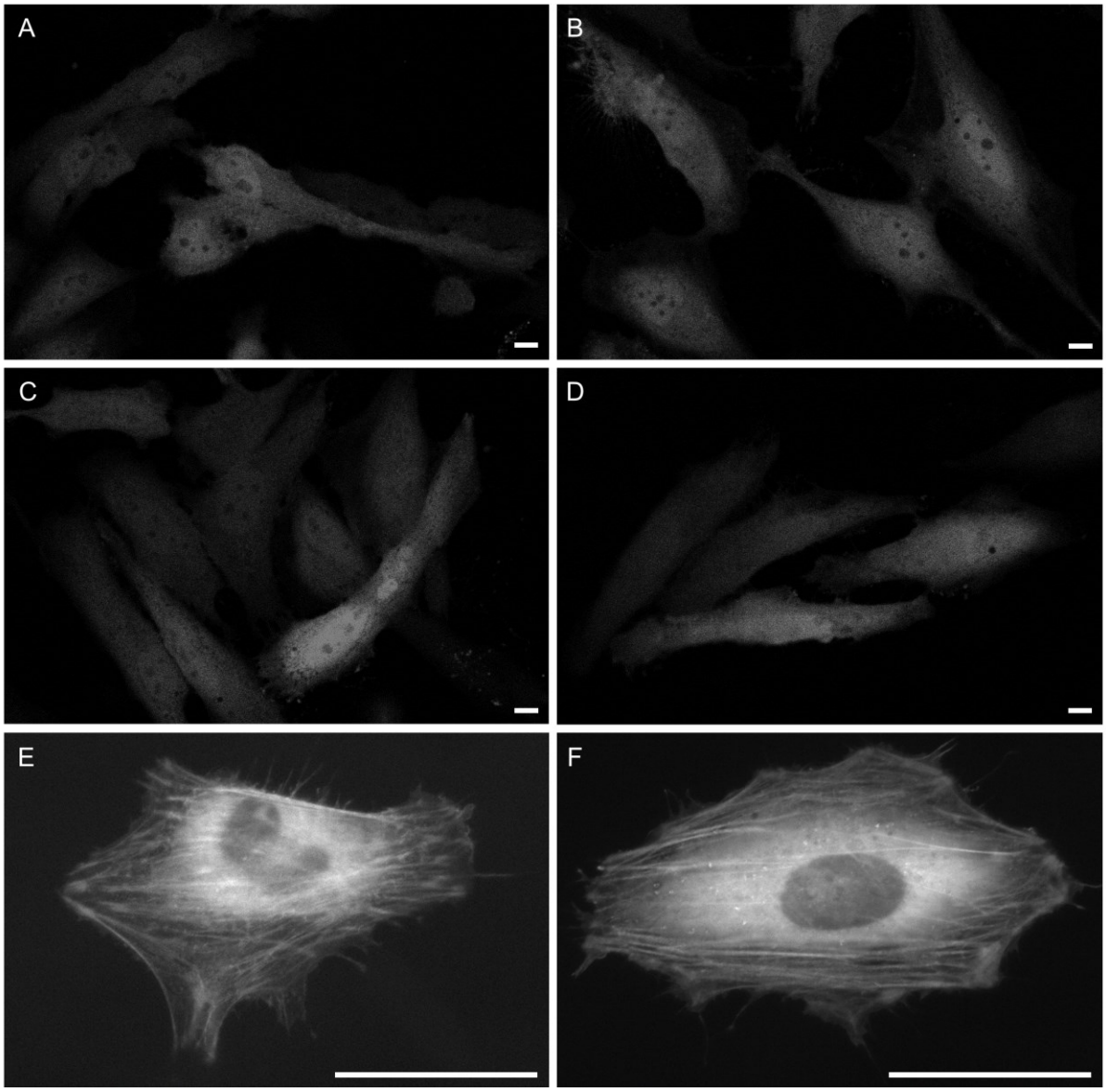

Figure 2.12 Fluorescence imaging of HeLa cells. A-D) Two photon imaging with excitation at 910 nm. A) GRvT, 30 mW. B) RRvT, 60 mW. C) LSSmKate2, 50 mW. D) mKeima, 100 mW. Illumination intensities were varied to provide approximately equivalent fluorescence brightness. E-F) Widefield imaging with 570-600 nm emission filter. E) GRvT, 450-490 nm excitation filter. F) RRvT, 510-560 nm excitation filter. Due to its low brightness, we did not attempt live cell imaging of GGvT. Scale bars are 10 µm.

2.3 Experimental Section

2.3.1 Molecular cloning and mutagenesis

EP-PCR was performed using Taq polymerase, (New England Biolabs). 50 µL reactions were created using: 5 µL of 10x Taq buffer; 2 µL of an NTP mix containing, dATP (5 mM), dGTP (5 mM), dCTP (25 mM), and dTTP (25 mM); 4 µL $MgSO_4$ (25 mM), 1.5 µL of each DNA primer (10 µM), 1 µL of template (~2 fmol), 1 µL of Taq, 0.5-1 µL $MnCl_2$ (10 mM), and deionized H_2O up to 50 µL. The $MnCl_2$ was added last to ensure that the Mn^{2+} will always be at the correct dilution to avoid over incorporation in Mn^{2+} into Taq which would result in an increased error rate. Polymerase chain reaction (PCR) was performed using Pfu polymerase and Q5 high fidelity polymerase from New England Biolabs using their recommended protocols. DNA primers were purchased from IDT. All PCR products were run on 1% agarose gels with 1 µL of 10 mg/mL ethidium bromide, the relevant bands were visualized with UV light and cut out with a razor blade and the DNA extracted using gel extraction kits (Thermo Fisher Scientific or BioBasic) using the manufacturer's recommended protocols.

Gibson Assembly (New England Biolabs) was used to assemble the plasmids for transformation. A reduced scale protocol was used with 1 µL vector, 1.5 µL of insert, and 2.5 µL of 2× Gibson Assembly Master Mix. This mix was then incubated for 4 hours at 50 °C and left at room temperature until needed for transformation, 0-48 hours. It was then diluted with 5 µL of deionized H_2O before transformation. Alternatively, QuikChange lightning, single or multi (Agilent), were used for site directed mutagenesis using the manufacturer's recommended protocols. The resulting plasmids either pBAD/His B (Thermo Fisher Scientific) or pcDNA3.1(+) (Thermo Fisher Scientific) were assembled and then used to

transform *E. coli* strain DH10B (Thermo Fisher Scientific) by electroporation. The codon differentiated templates for RRvT and GGvT were each synthesized (Thermo Fisher Scientific) in two parts that were combined using Gibson Assembly.

2.3.2 Screening

The transformed cells containing the pBAD/His B plasmids were plated onto agar with LB medium and 0.4 mg/ml ampicillin and 0.02% w/v L-arabinose, and grown overnight at 37 °C. On-plate colony screening was performed using illumination from a 300W xenon arc lamp with a 450-490 nm excitation filter for green and 510-560 nm excitation filter for red. Selection was done using a combination of visual inspection with long pass filter goggles and digital fluorescence imaging using a 500-520 nm emission filter for green and 660-700 emission filter for red. Screening was based on the appropriate wavelength fluorescence brightness and the green-to-red ratio. Picked colonies were transferred to liquid cultures and were grown in 2-6 mL Lysogeny Broth (LB) and 0.1 mg/ml ampicillin and 0.02% w/v L-arabinose and incubated in a shaker at 240 rpm and 37 °C. If further maturation time was required for the FPs, further incubation at 225 rpm and 25 °C was used. The liquid cultures were then spun down at 15 000 relative centrifugal force (RCF) for 2 minutes and the supernatant was poured off. Then 100 µL of Bacterial Protein Extraction Reagent (BPER; Thermo Fisher Scientific), was added and the samples were vortexed for 5-60 minutes. The resulting suspension was then spun down at 15 000 RCF for 2 minutes and the supernatant was collected and tested at various dilutions for fluorescence, absorbance and, if relevant, green-to-red ratio, in a fluorescence plate reader (Tecan Safire2). The 2-6 mL liquid cultures described above and the pellet left after the BPER

extract were used as starting points for plasmid purification. Miniprep kits from Thermo Fisher Scientific and BioBasic were used according to their respective standard protocols.

2.3.3 Protein purification and characterization

A single colony of *E. coli*, transformed with the pBAD/His B plasmid containing a vT gene, was used to inoculate 4 mL LB supplemented with 0.4 mg/ml ampicillin. This culture was incubated at 37 °C in a shaker incubator (220 rpm) for 12 hours. The liquid culture was then added into 500 mL of LB with 0.4 mg/ml ampicillin and incubated a further 4 hours. The inducer L-arabinose was added to a concentration of 0.02% and the culture was allowed to grow overnight at 30 °C in a shaker incubator (220 rpm). The cells were then centrifuged at 15 000 RCF for 10 minutes at 4 °C, the supernatant was discarded and the pellet resuspended in 25 mL of 1× Tris buffered saline (TBS). The cells were then lysed using a cell disruptor (Constant System). The cell debris was removed by centrifugation at 15 000 RCF for 35 minutes at 4 °C. The protein was then purified using Ni-NTA agarose beads (MC Labs) according to the Native conditions protocol, and buffer exchanged with 1× TBS using a centrifugal filter unit with 10 000 MWCO (Amicon). Extinction coefficients were determined by measuring the absorption spectrum of vTs using a UV/Vis spectrometer (Beckman Coulter DU 800) at pH 7.25 in 1× TBS and the same concentration of the vT in 1 M NaOH. In 1 M NaOH solution, the both GFP-type and RFP-type chromophores have an extinction coefficient of 44 000 $M^{-1}cm^{-1}$ at their absorption peaks near 450 nm.[13] Application of the Beer-Lambert law gives the concentration of the alkaline denatured chromophore. This concentration is then used as the intact protein concentration at pH 7.5. Inputting this into the Beer-Lambert law

along with the maximal absorbance of the intact protein, yields the extinction coefficient for the intact protein at pH 7.5.

Quantum yields were determined by creating a dilution series of the FPs ranging in absorbance from approximately 0.02 to 0.05, in 1×TBS at pH 7.0. For red an emission scan from 550 nm to 800 nm with an excitation wavelength of 540 nm was used. The sum of the emission fluorescence values over this range were then plotted against their absorbance creating a linearly increasing graph set of data points. The slope of this set directly corresponds to the quantum yield. To convert the slope into a quantum yield, standard samples with known quantum yields must be prepared and measured to same as the unknown samples, resulting in a slope for each of the standards. The slope of the unknown red FP was then divided by the slope of a known FP and multiplied by the quantum yield of the known FP, which gives the quantum yield of the unknown FP. The red FPs mApple and tdTomato were used as standards which have known quantum yields of 0.49 and 0.69 respectively. The same pattern was followed for green except using 460 nm to 800 nm to the emission scan at an excitation of 450 nm and using eGFP as the standard. All fluorescence spectra were obtained using a fluorescence plate reader (Tecan Safire2).

The pK_a was determined in a series of pH buffers (pH 3 to 11) using the Carmody buffer system.[137] For each protein solution 2 µl was added into 50 µl of the desired pH buffer into a 396-well clear-bottomed plate (Thermo Fisher Scientific) and the fluorescence measured in a Safire2 plate reader. pK_a values were extracted by fitting the data with a theoretical curve and assuming a Hill coefficient (n_H) of 1.

2.3.4 Single molecule FRET

Single molecule FRET was performed using a laser-scanning microscope LSM 510/ConfoCor 2 (Carl Zeiss) equipped with a 458 nm laser. A diluted sample of purified GRvT (~100 pM) such that only one protein molecule would diffuse into the detection volume during each 1 ms bin of the experiment. The emission was split to the two detectors using a LP505 beamsplitter. The green and red detectors were equipped with a band pass filter of 470-540 nm and a long pass NFT 565 filter, respectively. Photon arrivals times were sorted into 1 ms bins and data was collected over six 10-second runs. Events were counted if they surpassed the noise threshold of 5 photons. This threshold was determined by first running a blank sample. If either the green or red channel surpassed 5 photons, both the green and red bins were counted.

2.3.5 HeLa cell imaging

HeLa cells were transfected with pcDNA3.1(+) plasmids, encoding the fluorescent protein genes of interest, using Turbofect (Thermo Fisher Scientific) according to the manufacturer's recommended protocol. Imaging of cells expressing actin fusions was done using an Axiovert 200M Zeiss inverted fluorescent microscope equipped with a CoolSNAP HQ2 CCD Camera. For RRvT-actin and GRvT-actin, a 510-560 nm excitation filter and a 450-490 nm excitation filter were used, respectively. For both proteins, a 570-600 nm emission filter was used. Due to its low brightness, I did not attempt live cell imaging of GGvT. 2-photon imaging was performed using an upright FV1000 confocal microscope (Olympus Canada) equipped with FluoView1000 software (Olympus Canada) and a MaiTai DeepSee Ti:sapphire laser tuned to 910 nm (Spectra Physics).

2.4 Conclusions

In this Chapter I described the creation of, at the time of writing, the brightest red FP (RRvT), green (GGvT), and green-red heterodimeric (GRvT), tdTomato variants. GRvT exhibits 99% intramolecular Förster resonance energy transfer (FRET) efficiency, resulting in long Stokes shift red fluorescence. These new variants may prove useful for multicolor live cell imaging applications.

We have developed a set of three new tandem dimer FPs that are interesting both in terms of their fundamental photophysical properties and their potential utility for live cell imaging. As a photophysical model system, the vT series is particularly tractable for probing the FRET phenomenon due to the fact that the donor and acceptor fluorophores are at a fixed distance and orientation from each other. For live cell imaging, preliminary investigations suggest that all three vTs give bright fluorescence, using both 1- and 2-photon excitation, when expressed in mammalian cells and as actin fusions (Figure 2.12). While it is difficult to predict whether these new FPs will prove to be as robust as currently preferred FPs, RRvT could serve as brighter alternative to the popular tdTomato FP, and GRvT could serve as a bright LSSFP.

Enhancing fluorescent protein photostability through robot assisted photobleaching

3.1 Introduction

The use of FPs in imaging cellular functions has revolutionized cellular biology and pushed the boundaries of what it is possible to observe in cells. For example, FPs have been engineered to be sensors for protein-protein interactions,[50] second messengers,[79] neurotransmitters,[81] and membrane voltage changes.[138] Even though FP tools now allow us to probe the inner workings of cells with unprecedented precision and specificity, there remains substantial room for improvement in their fundamental properties. Photostability, the ability for fluorescence to resist photobleaching during illumination, is one property where FPs generally lag behind alternative fluorophore technologies such as synthetic dyes and fluorescent nanoparticles.[139]

Although FP photobleaching is a ubiquitous phenomenon, the emerging message from previous mechanistic studies is that multiple mechanisms are involved and their relative importance may be FP-variant dependent. In avGFP it has been noted that glutamate decarboxylation contributes to photobleaching, but this has not been reported in other FPs.[46] In the enhanced green FP (eGFP), an improved variant of avGFP, photobleaching was linked to the presence of molecular oxygen, radicals, and singlet oxygen,[140] whereas avGFP was found to be insensitive to the same compounds.[141] A QM/MM simulation on GFP-like proteins predicted photo-induced chromophore decomposition in the presence of molecular oxygen.[44] This photobleaching occurs by an excited chromophore undergoing an intersystem crossing into the excited triplet state which can react with oxygen, possibly requiring an additional excitation event, which results in damage to the chromophore.[45] In the orange FPs mOrange

and mOrange2 (derived from *Discosoma* sp. RFP), and the red FPs TagRFP and TagRFP-T (derived from *Entacmaea quadricolor*), oxygen-dependent photobleaching was also observed.[142] This observation led to the creation of Kriek, a more photostable version of mCherry (another red FP derived from *Discosoma* sp.), by reducing oxygen's access to the chromophore.[47] In contrast the protein mStable, another *E. quadricolor* derived red FP, shows increased photostability by the oxidation of a cysteine residue that interacts with the chromophore. It is thought that this oxidation reduces the chance of a cis-trans isomerization of the chromophore, which can result in a bleached FP.[143] Although a role for oxygen appears to be a theme common to most photobleaching mechanisms, removing oxygen prior to illumination tends to cause a reduction in photobleaching but by no means prevents it.[47,142]

Reporting photostability half-lives has been problematic as results from different labs are rarely identical.[92] One method involves suspending cell sized aqueous droplets containing FPs in oil and measuring the time to photobleach from 1 000 to 500 photons/s/molecule, where the light intensity is adjusted to achieve the initial 1 000 photons/s/molecule.[14] This is designed to match how end users image FPs, where the illumination intensity gets scaled to acquire the optimal signal to noise and minimize bleaching. Such experiments have revealed a linear relationship between light intensity and photobleaching, on a log-log graph, such that higher intensities result in increased photobleaching.[144] The slope of this relationship varies between FPs from a low of 0.72 in mTagBFP2 and a high of 1.77 in tdTomato, such that the ranked order of FP photostabilities can change as a function of intensity.[144]

Most of the FPs used in research today are the result of directed evolution efforts to maximize intrinsic fluorescent brightness (a product of quantum yield and extinction coefficient). However, a very bright FP that suffers from poor photostability is not generally

useful. Characterizing the photobleaching of FPs is fraught with challenges, as the photobleaching half-life depends on the wavelength and intensity of the excitation light, as well as a multitude of other factors. Efforts to account for many of these factors have produced rankings of photostability for the commonly used FPs.[144]

Higher intrinsic fluorescence brightness allows an FP to be imaged with better signal to noise ratio and allows a lower required intensity of illumination to be used, which can allow longer imaging times.[144] However, as quantum yields approach unity and extinction coefficients start to plateau, other characteristics must be used to improve FPs. Our approach is to increase the photon budget by improving the photostability of the FP. Increased photostability means reducing the chance that a given absorption event will result in photobleaching. Improving photostability allows studies to be performed longer and with higher temporal resolution.

Even in the absence of a complete mechanistic understanding of photobleaching, empirical screening of large libraries of variants has led to the identification of FP variants with improved photostability.[47,142,145,146] Two general approaches have been used to screen large libraries of variants for improved photostability. The first is microfluidic fluorescence activated cell sorting (FACS)-type systems in which the fluorescence from single cells is measured before and after illumination with high intensity light.[47,147–150] The second is on-plate colony screening where a fluorescence image of an entire Petri dish of bacterial colonies is acquired after photobleaching, and the brightest colonies are manually picked.[94,142]

As an alternative and convenient metric for assessing FP utility, I propose Integrated Photon Output (InPhO), which I define as the product of the integral of the photobleaching decay curve and the FP intrinsic fluorescent brightness. The InPhO value is a measure of the

total number of photons that equal concentrations of FPs will produce until complete bleaching occurs. This differs from other similar measurements, like total photon count, as it is based on a bulk solution measurements rather than single FPs measurements averaged. As with all measures of photostability, InPhO values can only be compared under identical illumination conditions.

A major challenge with this second approach is achieving enough illumination intensity to uniformly photobleach colonies over the large area (42 cm^2) of a standard Petri dish. To address this challenge, I now report a screening system in which colonies on a Petri dish are photobleached using a beam of high intensity white light and a robot that systematically moves the plate to achieve even illumination. I have applied this system to the directed evolution of a yellow FP mCitrine for improved photostability. mCitrine is widely used as both a fusion tag and as a FRET acceptor. Despite mCitrine's popularity, it, like all yellow FPs, is vulnerable to photobleaching – mCitrine photobleaches ten times faster than eGFP.[144] For this reason, a bright and photostable yellow fluorescent protein is desirable.

3.2 Results

3.2.1 Robot assisted plate bleach

The photobleaching robot used is made entirely from one Lego Mindstorms EV3 kit and common laboratory items. The plate bleaching configuration drives a rotating Petri dish in front of a 300 W xenon arc lamp light source. The light source is initially offset from the centre of the Petri dish by a distance equal to the radius of the illumination. The robot then rotates the Petri dish rapidly in front of the light source tracing and retracing a circle around the centre for a specific length of time. The robot then drives forward one step causing the light to trace a

larger circle. Each circle is bleached for a period of time that is directly proportional to the circumference of the circle. Equation 3-1 is the general formula to determine the time required on each successively bigger circle to create even photobleaching.

Equation 3-1

$$t_x = \frac{t_i}{2\pi r} 2\pi(r + sx)$$

Variable t_x is the time spent on circle x, where x is the circle number, with the smallest radius circle being 0 and the next largest being 1, and so on. Variable t_0 is the time spent on the smallest circle, r is the radius of the smallest circle, and s is the step size (the increase in circle radius from circle x to x+1).

To visualize the evenness of the photobleaching pattern, Petri dishes were prepared with fluorescein added to the solid media. In this way, a uniform and photobleachable fluorescent surface was prepared. Fluorescein-loaded Petri dishes were then subjected to photobleaching under conditions that should bleach ~40% of the initial fluorescence. Following photobleaching, the green fluorescence of the Petri dish is imaged to reveal the evenness of the photobleaching pattern. For library screening it is desirable to have as even of photobleaching as possible so that all colonies receive the same dose of light.

Initial testing was performed using a non-overlapping concentric circle protocol (Figure 3.1A). When this protocol was used to bleach a Petri dish (58% fluorescence remaining), the fluorescence was observed to be uneven across the Petri dish and a line profile through the centre revealed a distinct saw-tooth pattern (Figure 3.2E and F). The difference

between the maximum and minimum remaining fluorescence was 43% ± 10% standard deviation (Figure 3.2G).

To address this problem, I designed a second bleaching protocol that used overlapping circles. The overlapping circle protocol uses a smaller step size such that each point on the Petri dish was bleached by five consecutively larger circles, rather than just one in the non-overlapping concentric circle protocol (Figure 3.1B). A fresh fluorescein Petri dish was bleached using the overlapping circle protocol to the same average of 58% fluorescence remaining, resulting in a smooth line profile, and a max-min difference of 17% ± 3% standard deviation (Figure 3.2C). To better understand and visualize the difference in the evenness of photobleaching for these two protocols, I simulated both bleaching profiles (Figure 3.3). The model uses a line as the illumination spot shape, rather the complex donut like spot shape of the light source. Based on these results, the overlapping circle protocol was used in all subsequent screens.

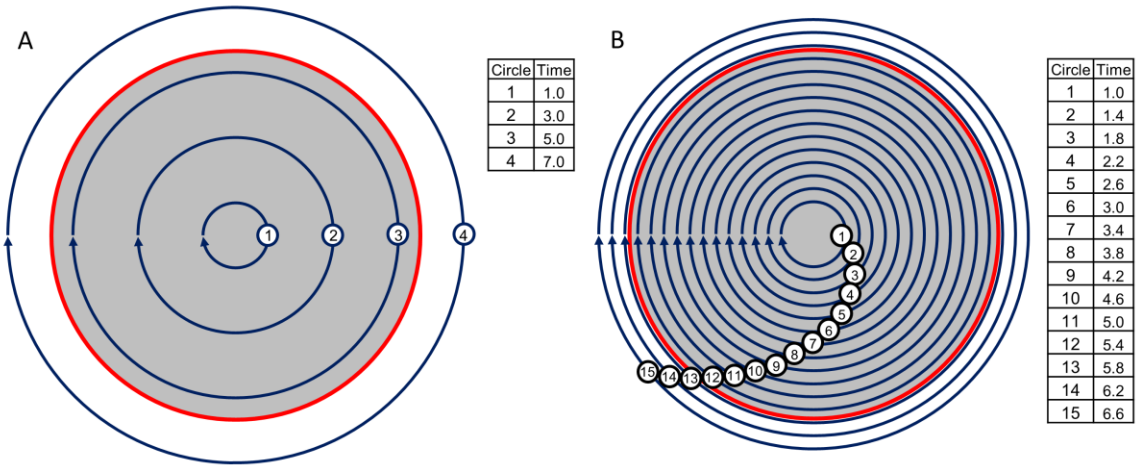

Figure 3.1 Robot photobleaching patterns in the plate bleach configuration. Inset tables show the time required on each circle to achieve even illumination. A) Concentric circle protocol. Red circle represents the edge of a Petri dish. Blue line shows the path of the center of the illumination spot. Each larger circle's radius increases in size by the smallest circle's diameter, which is equal to the illumination spot size. Time is normalized to the time spent on the smallest circle. B) Overlapping circle protocol each larger circle's radius increases in size by 20% of the smallest circle's diameter, which is determined by the illumination spot size.

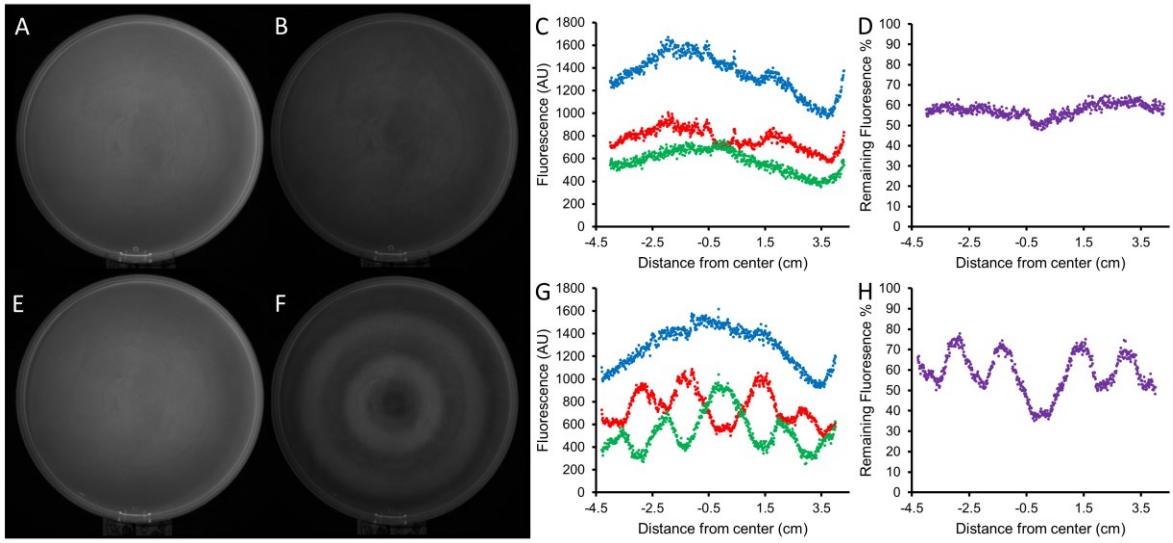

Figure 3.2 Fluorescein plate photobleaching. A-D) Photobleaching using the overlapping circle protocol. E-F) Photobleaching using the concentric circle protocol. A, E) Image of fluorescein Petri dish before photobleaching. B, F) Image of fluorescein Petri dish after photobleaching. C, G) Line profile of fluorescence. Blue is the prebleach profile. Red is the postbleach profile. Green is difference between prebleach and postbleach. D, H) Line profile of the percent fluorescence remaining after photobleaching. Image scale (A, B, E, F) matches the distance axis (C, D, G, H).

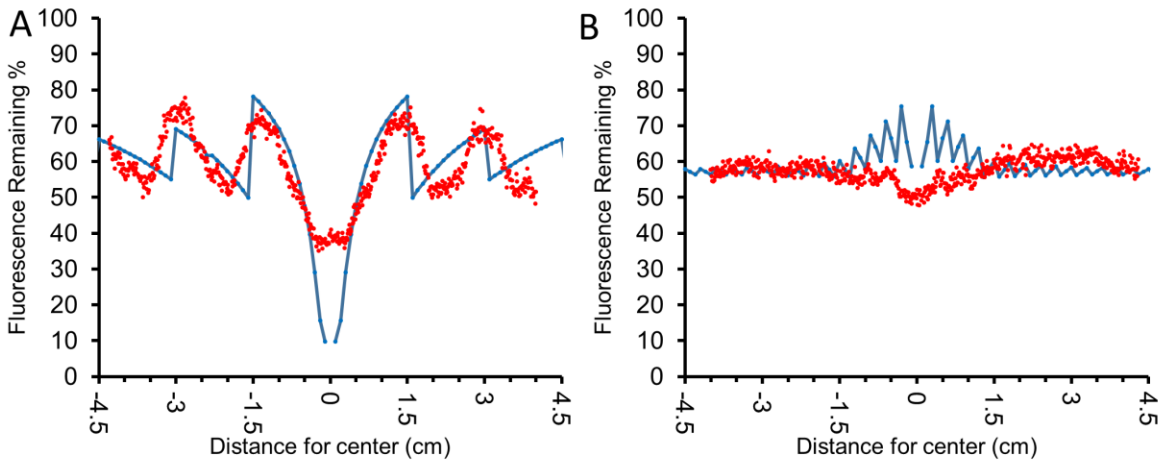

Figure 3.3 Simulated and observed photobleaching profiles. A) Concentric circle protocol. B) Overlapping circle protocol. Blue represents theoretical photobleaching. Red observed photobleaching.

3.2.2 Robot assisted tube bleach configuration

The robotic system was designed to be modular such that the plate spinning motor used in the plate bleach could be used separately from the stepping motor. This was done to allow the system to be easily converted into a format that was suitable for *in vitro* photobleaching of cell lysates or purified FPs in tubes. A PCR tube holder was made of the plastic from tip rack holders and glued to the outer edge of a Petri dish. This custom holder can attach to the plate spinning motor and support 28 tubes. Attaching this holder to plate spinning motor, placing it in front of the light source, and turning on the motor to spin the holder, results in equal light intensity illumination of all 28 samples. Another motor was added to create a shutter to control the illumination time accurately. This setup is referred to as the tube bleach configuration (Figure 3.4).

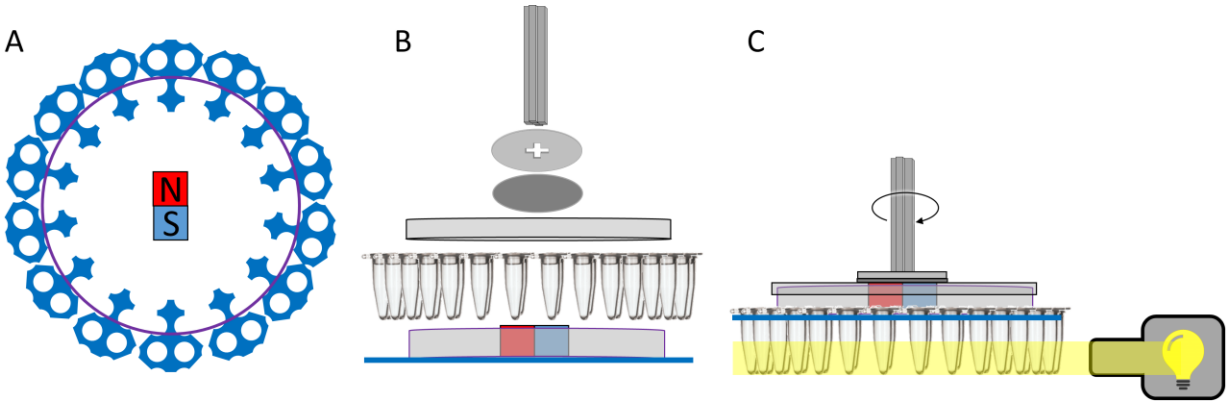

Figure 3.4 Tube bleach configuration. A) Custom PCR tube holder. Constructed using tip rack cut-outs as PCR tube holders, a Petri dish base, and a magnet all glued together. B) Expanded tube bleach apparatus. Top to bottom axle, then three piece glued together: a Lego socket, steel disk, and Petri dish lid. The PCR tubes slide into the tip rack cut-outs and the custom PCR tube holder magnetically attaches the Petri dish base into the Petri dish lid via the steel disk. C) Assembled tube bleach apparatus showing the rotation and light path. The samples are illuminated twice per rotation as they pass through the illumination path.

3.2.2.1 Half-life determination

To generate photobleaching half-lives, BPER extracted cell lysate containing the FP was diluted with H_2O to have a peak absorbance of 0.1 and aliquoted into five 50 µL PCR tubes. The first tube was not bleached, the second tube was placed into the holder and spun in front of the light source for a predetermined length of time, t. The third tube was added and bleached along with the second tube, again for time t, and so on, resulting in a series of tubes bleached for 0t, 1t, 2t, 3t, and 4t. The fluorescence intensity of these samples were then measured in a fluorescence plate reader (Tecan Safire2) and their photobleaching half-lives calculated. FPs at this light intensity have photobleaching half-lives that typically fall between approximately 1 and 60 min. This tube bleach configuration allows seven sets of four tubes,

(and one unbleached) to be run simultaneously so that accurate comparisons can be made between all samples.

The tube bleach configuration was validated by running three sets of mCitrine and Citrine2, three times each. The resulting average error in half-life measurements within a single run of the tube bleach protocol was 0.7% and from three runs was 0.6%. There was no meaningful difference in the percent error between the mCitrine and Citrine2 data sets despite the large difference in half-life.

To maximize the number of variants that could be assessed using this method during library screening, I used an end-point version of this tube bleach configuration to screen 28 samples simultaneously. The protocol is the same as described above except that rather than bleaching several samples to get a time course, a non-bleached sample was compared to a single sample bleached for time *t*. This protocol provides the percent fluorescence remaining value and can be used to calculate an approximate photobleaching half-life. The best variants from this screen were then measured again using the more accurate half-life determination method described above.

3.2.3 Concentration and photostability

During initial experiments with the tube bleach configuration, the importance of diluting samples to an absorbance of 0.1 had not yet been recognized. These early experiments revealed that samples at higher concentrations consistently had longer half-lives than samples with lower concentrations. This was surprising considering that all of the concentrations were below the threshold where inner filter effect should start reducing the effective light intensity. To further explore this phenomenon, a dilution series of the Citrine variants were created and they were all bleached equally (Figure 3.5). All variants showed very strong dependence of

photobleaching on concentration. Even more interesting, the trend switched from increasing photobleaching with increasing concentration in mCitrine, to decreasing photobleaching with increasing concentration in all the newer variants. This dependence was strong enough to change the half-life by approximately four-fold for all variants. The take-away message from these experiments was that concentration must be held constant in order to get accurate photobleaching half-lives that could be compared between variants. As such, absorbance was the simplest way to quickly match concentrations, and an absorbance of 0.1 was chosen because that approximately matches the intracellular concentration of FPs when expressed in mammalian HeLa cells using the strong CMV promoter (~10 μM).[151,152]

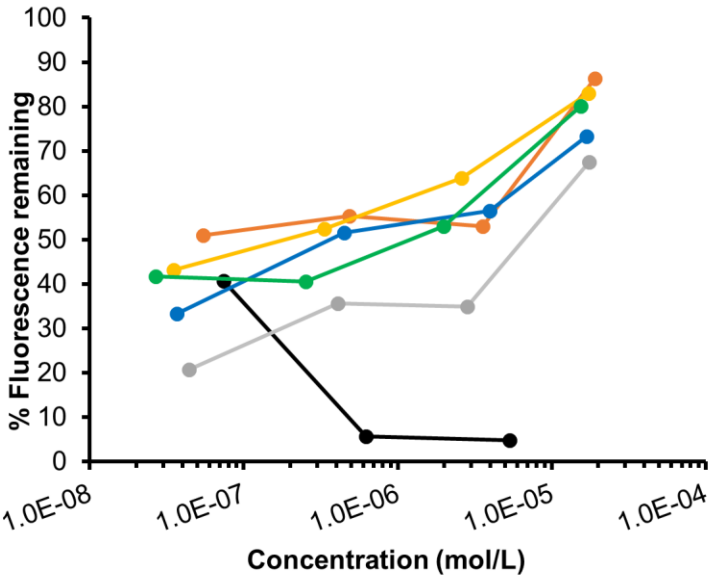

Figure 3.5 Concentration dependence of photostability. All samples were bleached for 20 min in the tube bleach configuration under the white light illumination of a 300 W xenon arc lamp. Black is mCitrine, orange is Citrine 1.7, grey is Citrine 1.9, yellow is Citrine 1.10, blue is Citrine 1.11, and green is Citrine 2.

3.2.4 Mutations and spectroscopic analysis

The FP mCitrine is the result of three distinct directed evolution and engineering efforts. The first was the conversion of avGFP into eYFP, then came the engineering of Citrine from eYFP, and finally the engineering of monomeric mCitrine from Citrine. Through these 3 steps the avGFP gene acquired a total of eight mutations.[10,17,153,154] My photostability screening led to the introduction of nine additional mutation: S30T, M69T, Y145H, N149Y, V163A, K206Q, K214E, M218T, and D234G. The intermediate variants and their specific mutations are listed in Table 3.1. All mutations were found by robot assisted photostability screening of libraries generated through error prone PCR. Four mutations, M69T, Y145H, V163A, and M218T, are inward facing. M69T and Y145H are in the first shell of residues around the chromophore, while V163A and M218T are in the second shell of residues and in contact with first shell residues. N149Y is outward facing and with its backbone adjacent to the chromophore. K206Q is also outward facing and this position 206 is known to be important for dimerization. S30T is outward facing and close at the chromophore but out of direct interaction range. K214E is in the loop region between beta strands S10 and S11. D234G is in the unstructured region near the C- terminus (Figure 3.6). Overall, the chromophore environment of Citrine2 was extensively modified relative to that of mCitrine. The absorbance and emission spectra are subtly different as both are blue-shifted by 5 nm. The most substantial difference between mCitrine and Citrine2 is the large 8.9-fold improvement in photostability (Table 3.2) (Figure 3.6). The brightness of Citrine2 is slightly less than mCitrine due to a decrease in quantum yield from 0.74 in mCitrine to 0.70 in Citrine2. However, this decreased quantum yield is compensated for by an increase in extinction coefficient from 94 000 $M^{-1}cm^{-}$

[1] in mCitrine to 98 000 $M^{-1}cm^{-1}$ in Citrine2, resulting in essentially equivalent fluorescent brightness for the two proteins.

Table 3.1 mCitrine variant mutations

Cumulative Mutations									
Citrine 1.7			Y145H		V163A		K214E	M218T	D234G
Citrine 1.9			Y145H	N149Y	V163A		K214E	M218T	D234G
Citrine 1.10			Y145H	N149Y	V163A	K206Q	K214E	M218T	D234G
Citrine 1.11		M69T	Y145H	N149Y	V163A	K206Q	K214E	M218T	D234G
Citrine 2	S30T	M69T	Y145H	N149Y	V163A	K206Q	K214E	M218T	D234G

Table 3.2 Spectral properties of mCitrine variants

Fluorescent Protein	**λ_{ex} (nm)**	**λ_{em} (nm)**	**ε_{max}[b] (M^{-1} cm^{-1})**	**ϕ_n[c]**	**Brightness[d]**	**Fluorescence lifetime (ns)**	**Bleaching half-life[e] (min)**	**pK_a / Hill[f] coefficient**	**InPhO fold improvement[g]**
mCitrine	513	527	94 000	0.74[a]	66	3.6	3.3	5.3 / 0.6	1
Citrine 1.7	506	524	78 000	0.53	39	2.6	19.1		3.4
Citrine 1.9	508	525	88 000	0.38	32	2.9	15.0		2.2
Citrine 1.10	509	524	106 000	0.7	71	2.8	23.7		7.7
Citrine 1.11	509	524	98 000	0.59	55	3.2	31.3		7.9
Citrine 2	509	522	98 000	0.7	65	3.3	29.8	5.9 / 0.5	8.9

[a] Previously reported data.[144] [b] ε_{max} is the peak extinction coefficient determined by alkali denaturation. [c] ϕ_n is the fluorescence quantum yield. [d] Brightness is the product of ε and ϕ/1000. [e] Bleaching half-life is the time to reduce fluorescence of a 0.1 absorbance FP solution to half its maximum value using the robotic white light tube bleach configuration. [f] pK_a is the pH at which the fluorescence is at half of its maximum value and the Hill coefficient is the cooperativity of the pK_a of fluorescence. [g] InPhO fold improvement is InPhO value normalized to that of mCitrine.

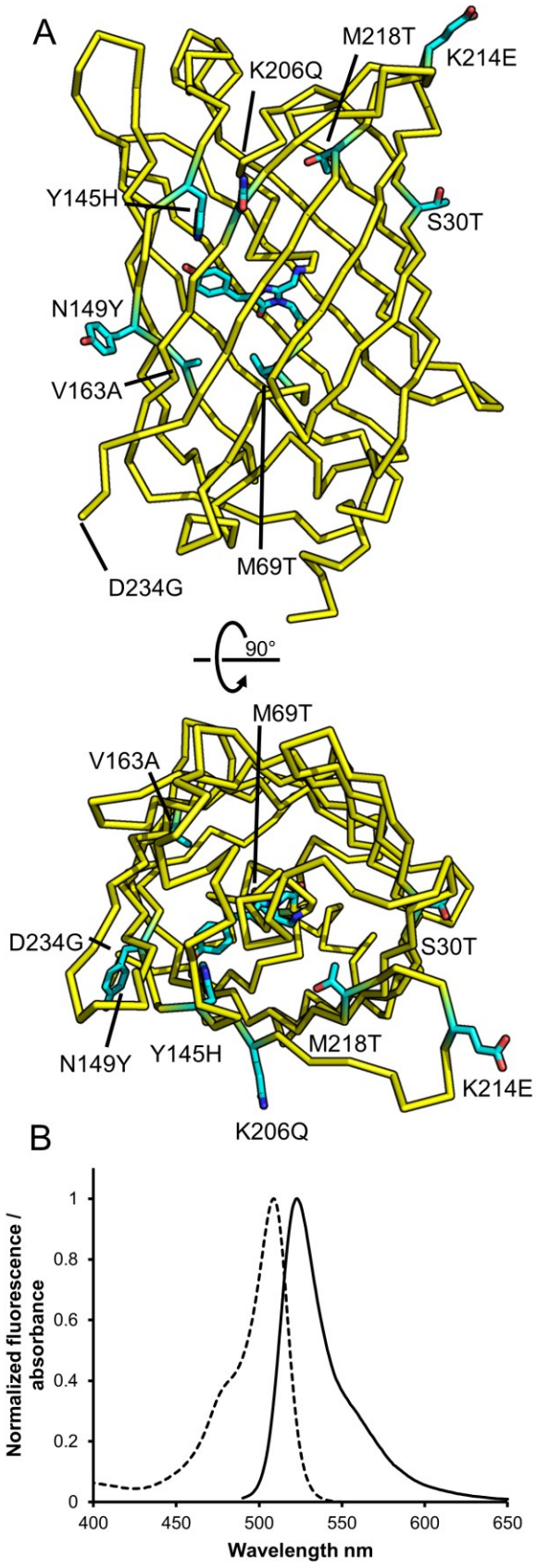
A
M218T
K214E
K206Q
Y145H
S30T
N149Y
V163A
M69T
D234G
90°
M69T
V163A
D234G
S30T
N149Y
Y145H
M218T
K214E
K206Q
B
Normalized fluorescence / absorbance
0
0.2
0.4
0.6
0.8
1
400
450
500
550
600
650
Wavelength nm

Figure 3.6 Modeled location of mutations in Citrine2 and spectra. A) Location of mutations in Citrine2 modeled based on the crystal structure of mCitrine (PDB ID: 1HUY).[153] B) Normalized absorbance and emission spectra of Citrine2. Dashed line is absorbance and solid line is emission. All spectra are normalized to their peak absorbance or emission.

3.2.5 Environmental sensitivity

The pK_a of fluorescence increased slightly in Citrine2 compared to mCitrine, from 5.3 to 5.9. The Hill coefficient decreased from 0.60 for mCitrine at 0.60 to 0.51 for Citrine2. The amino acid Met69 in mCitrine was the result of efforts to reduce the environmental sensitivity of YFPs, as the Q69M mutation greatly decreased the sensitivity towards chloride anion.[153,155] Despite the mutation M69T in Citrine2, it remained chloride insensitive with a small fluorescence increase of 6% when increasing the concentration of Cl^- from 1 mM to 101 mM mCitrine exhibited a similar 14% increase under the same conditions. Iodide was also tested at 1 mM to 101 mM I^-. Both mCitrine and Citrine2 exhibited a small increase in fluorescence of 8% for mCitrine and slighter larger increase of 18% for Citrine2. To examine the effect of oxygen on photobleaching rate, I added EC-Oxyrase to the PCR tubes when measuring the half-life in the tube bleach configuration. The reduction in dissolved oxygen concentration resulted in an increase in mCitrine's photobleaching half-life from 3.3 to 4.0 min (a 21% increase). For Citrine2, the photobleaching half-life increased from 29.8 to 35.2 (an 18% increase). Even after removal of dissolved oxygen the photobleaching half-life mCitrine is 7.45-fold faster than Citrine2 in the presence of dissolved oxygen.

3.2.6 Fluorescence lifetime

I determined the fluorescence lifetime of the best variants from each of the rounds of screening that led to Citrine2 (Table 3.3). The mCitrine to Citrine1.7 lifetime trend matched with the previously observed pattern of decreasing fluorescence lifetime (3.6 ns to 2.6 ns) as photobleaching half-life increases (3.3 min to 19.1 min).[47] However, variants beyond Citrine 1.7 showed simultaneous increases in both photobleaching half-life and the fluorescence lifetime. Citrine2 has a fluorescence lifetime of 3.3 ns and a photobleaching half-life of 29.8 min. If fluorescence lifetime and photobleaching are linked, the mutation set and the order in which the mutations were acquired should be irrelevant to their relationship. Comparing all variants from mCitrine to Citrine2 pairwise for their relative change in photostability and fluorescence lifetime should identify if previous experiments[47] found a correlation between the two or causation. Of all possible pairs of variants, eight have an inverse relationship between photostability and fluorescence lifetime, and seven have a direct relationship (Figure 3.7A and C).

Table 3.3 Fluorescence lifetime

	Monoexponential fitting			Biexponential fitting								
	τ [a]	±[b]	χ^2	τ1	±	A1[c]	±	τ2	±	A2	±	χ^2
mCitrine	3.627	0.020	1.152	-	-	-	-	-	-	-	-	-
Citrine 1.7	2.578	0.023	0.891	3.355	0.562	0.32	2.6E-01	1.921	4.0E-01	0.68	0.202	0.783
Citrine 1.9	2.904	0.030	1.060	3.173	0.261	0.61	2.7E-01	1.859	7.6E-01	0.39	0.187	1.048
Citrine 1.10	2.753	0.020	0.934	-	-	-	-	-	-	-	-	-
Citrine 1.11	3.239	0.031	0.951	-	-	-	-	-	-	-	-	-
Citrine 2	3.314	0.021	0.878	1.437	0.040	0.26	1.2E-02	3.487	6.0E-03	0.74	0.003	0.8625

[a] All lifetimes (τ) are reported in ns. [b] ± are the standard deviations. [c] A1 and A2 are fractions of each lifetime making up the biexponential decay curves.

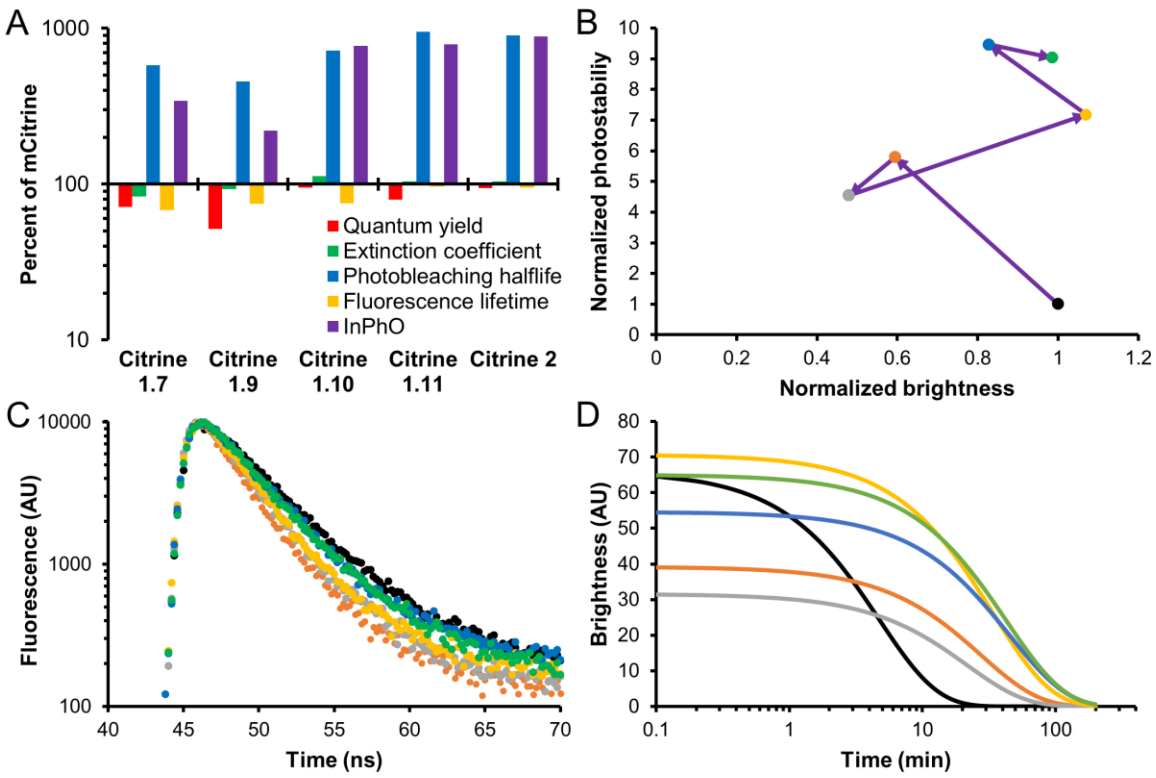

Figure 3.7 Evolution of *in vitro* photophysical properties. A) The percentage difference of each variant's photophysical properties relative to mCitrine. B- D) Black is mCitrine, orange is Citrine 1.7, grey is Citrine 1.9, yellow is Citrine 1.10, blue is Citrine 1.11, and green is Citrine2. B) *in vitro* photostability and brightness changes during directed evolution of Citrine2. Purple arrows represent mutation events. Brightness and photostability are normalized to the values for mCitrine. C) Fluorescence decay lifetime of the Citrine variants. D) Modelled photobleaching curve of equal concentrations of Citrine variants.

3.2.7 Mammalian cell bleaching

We tested mCitrine and Citrine2 in HeLa cells to determine how these FPs perform side by side (Figure 3.8). Under widefield illumination, mCitrine had a biexponential decay curve while Citrine2's decay curve was monoexponential. The average half-lives for mCitrine and Citrine2 are 42 and 59 s respectively, resulting in an InPhO increase of 1.38-fold. Under confocal illumination conditions, both mCitrine and Citrine2 are biexponential, with average

half-lives of 2.1 and 2.8 s respectively, and an InPhO increase of 1.33-fold, relative to mCitrine (Table 3.4, Figure 3.9).

Table 3.4 HeLa cell photobleaching

Imaging mode		$t_{1/2}$[a] (s)	±[b]	A1[c]	±	$t_{1/2}$ (s)	±	A2	±	InPhO fold change[d]
widefield	mCitrine	14.0	0.1	0.36	0.003	57.9	0.4	0.64	0.003	
	Citrine 2	7.4	1.2	0.01	0.0006	59.6	0.09	0.99	0.0003	1.38
confocal	mCitrine	1.2	0.009	0.80	0.004	5.8	0.1	0.20	0.004	
	Citrine 2	1.4	0.02	0.56	0.006	4.6	0.04	0.44	0.007	1.33

[a] The $t_{1/2}$ are photobleaching half-lives. [b] ± are the standard deviations. [c] A1 and A2 are fractions of each half-life making up the biexponential decay curves. [d] InPhO fold change was compared to the InPhO of mCitrine in the same imaging mode.

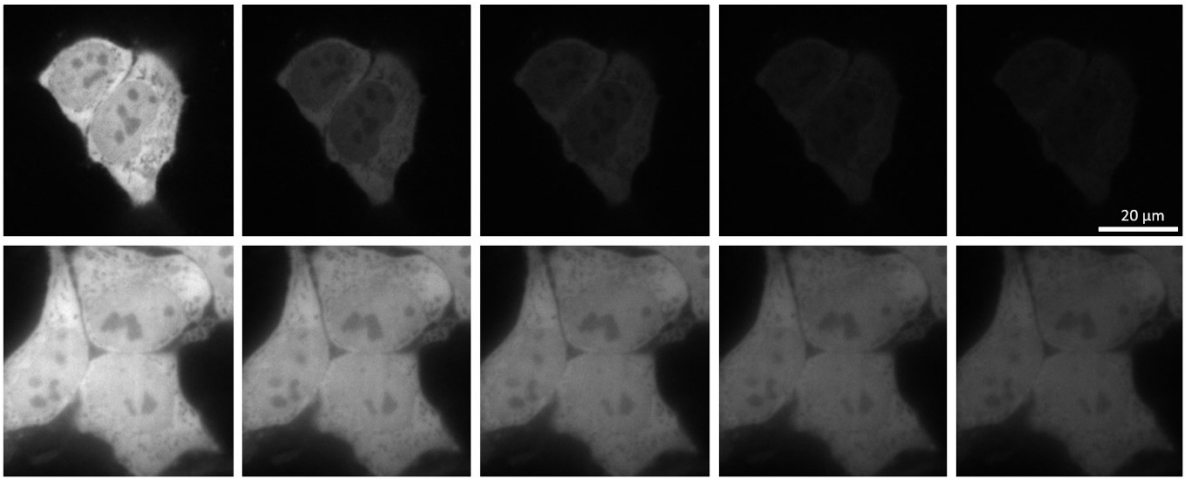

Figure 3.8 HeLa cell photobleaching. Top images are mCitrine, bottom are Citrine2. From left to right the images are at 0s, 2s, 4s, 6s, and 8s. All are confocal images taken at the same settings, 488 nm, 50 mW illumination, and shown at the same scale.

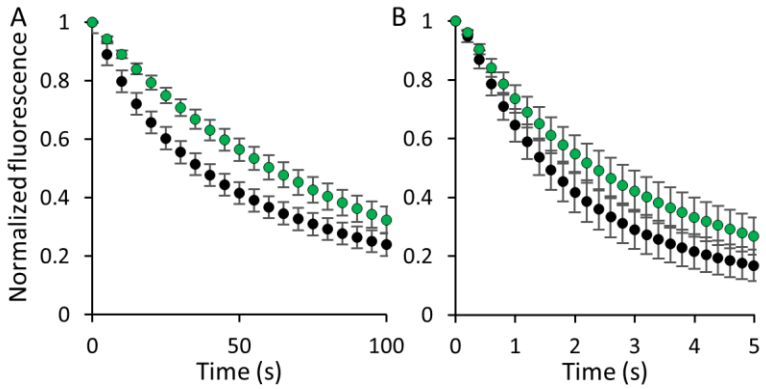

Figure 3.9 Fluorescence decay in HeLa cells. A) Widefield fluorescence with 75 W xenon lamp illumination using 488/10 nm excitation and 525/25 nm emission filters. B) Confocal fluorescence illuminated using a 50 mW, 491 nm pumped diode laser with a 540/30 nm emission filter. Decay curves are the average of at least 16 cells. The error bars are the standard deviation. Black is mCitrine, green is Citrine 2. Fitted half-lives and InPhO are in Table 3.4.

3.3 Discussion

All photobleaching screening systems must balance a number of issues. For example, high intensity illumination causes faster photobleaching, but can also rapidly heats the cells and result in reduced viability. The rotation of the Petri dish allows for additional air convection, and the rotation gives time for radiative cooling while the light is illuminating the other parts of the Petri dish. The use of white light, as compared to a specific wavelength or a small range of wavelengths, matches more closely with the current trends of multicolour imaging, as fluorescent proteins are often exposed to multiple illumination wavelengths, which can change the photobleaching rate.[144,156]

The plate bleach configuration using the overlapping circle protocol is a significant improvement over the initial concentric non-overlapping circle protocol. The overlapping circle protocol results in a much more even bleach and decreases the total variation across the

plate from 43% to 17% and the standard deviation from 10% to 3%. Much of the remaining asymmetry observed in the fluorescence profile of the fluorescein Petri dish is based on three factors: 1) the reflectance of the aluminum foil backing of the tube bleach configuration and the Petri dish holder of the plate imaging system are both uneven and can locally modify the illumination and the observed fluorescence in both the imaging and photobleaching steps; 2) the incident light directly from the illumination source is not even; and 3) the fluorescent gel is not even across the Petri dish and has a large meniscus which, when coupled with the slight angle of the incident light required for imaging, results in the apparent increased fluorescence observed on the right side of the Petri dishes (Figure 3.2A, B, E, F).

As our screening system is aimed at optimizing both brightness and photostability, I had to choose a scoring system for the characteristics that reflected our desired end point. For example, I needed to be able to decide whether losing 5% of the brightness to gain 5% on photostability half-life was worthwhile. InPhO is an unbiased comparison of equal concentrations of FPs that takes into account both brightness and photostability at a given illumination level. Mathematically, the InPhO value for photobleaching of a particular FP at a particular light intensity is described by Equation 3-2:

$$InPhO = \varepsilon\Phi \sum_{i=1}^{\infty} \alpha_i \int_0^{\infty} (0.5)^{\left(x/t_{0.5}^{i}\right)} dx$$

Equation 3-2

where ε is the extinction coefficient, ϕ is the quantum yield, *i* is the identifier of each exponential required to describe the fluorescence decay due to photobleaching, α is the fraction of each exponential component, and $t_{0.5}^{i}$ is the half-life of each exponential component.

Equation 3-2 simplifies to Equation 3-3 for monoexponential decays and to Equation 3-4 for biexponential decays.

$$InPhO = \frac{\varepsilon \Phi t_{0.5}}{\ln 2}$$

Equation 3-3

$$InPhO = \frac{\varepsilon \Phi}{\ln 2}\left(\alpha_1 t_{0.5}^{1} + \alpha_2 t_{0.5}^{2}\right)$$

Equation 3-4

When starting directed evolution to improve an FP, there is typically a trade-off between brightness and photostability. TagRFP-T, mOrange2, and Kriek are three FPs specifically evolved for photostability.[47,142] These FPs acquired 1, 4, and 3 amino acid substitutions from their parental variants, TagRFP, mOrange and mCherry, respectively. All three variants gained photostability at the cost of brightness. Calculating their InPhO values reveals that TagRFP-T went up 2.3-fold, mOrange2 increased 14.7-fold, and Kriek decreased by 1.25-fold. In cases where the starting point is not optimized, such as starting from BFP (p4-3), a dim blue fluorescent protein created from avGFP, both brightness and photostability can be increased.[12,146] The creation of eBFP2 from BFP required 11 mutations and led to an overall increase in InPhO of 2190-fold, with both brightness (4-fold) and photostability (550-fold) improving greatly. In this work, the evolution of Citrine2 follows both of these routes. Starting with an optimized FP mCitrine, there was an initial increase in photostability at the expense of brightness. In later rounds the brightness was recovered with further photostability increases by acquiring several additional mutations. Overall, this process led to the discovery of a variant

with an 8.9-fold InPhO improvement over mCitrine. Interestingly, the quantum yield, extinction coefficient, fluorescence lifetime, and photostability all fluctuate over the directed evolution of Citrine2. Furthermore, Citrine2 is not optimal in any one individual photophysical parameter other than its InPhO value (Figure 3.7).

Photostability has been experimentally linked to the fluorescence lifetime such that, when comparing two similar variants of a particular fluorophore, a decrease in fluorescence lifetime is accompanied by an increase in the photostability.[47] This proposed inverse relationship between fluorescence lifetime and photostability is likely a side effect of starting directed evolution projects with an optimized FP. An optimized FP will most likely be at a local maximum of the fitness landscape for the evolutionarily selected trait, which is most typically the brightness (proportional to the product of quantum yield and the extinction coefficient). As a result, most mutations that favour a different trait, such as photostability, will come at a cost to the brightness. If this loss of brightness if its due to a decreased quantum yield, then the fluorescence lifetime will also decrease according to Equation 3-5:[157]

$$\Phi = \frac{k_r}{\tau}$$

Equation 3-5

where ϕ is the quantum yield, k_r is the rate of radiative decay, and τ is the fluorescence lifetime. However, if evolution is continued for several rounds, the protein may evolve out of its local minimum and in to a new one that is associated with increases in both brightness and photostability. There is a wide range of fluorescence lifetimes, quantum yields, and photostabilities between all FPs derived from avGFP and a relatively small number of

mutations differentiating them. Acquiring more than three or four mutations seems to be able to dramatically change the photophysical properties.

Removing molecular oxygen using EC-Oxyrase results in mCitrine's photostability increasing by 7% and Citrine2's by 37%. As mCitrine's photostability after EC-Oxyrase treatment is 7.5-fold lower than Citrine2 without EC-Oxyrase treatment, a reduction of molecular oxygen access is not the primary cause of the photostability enhancement in Citrine2. Instead, the most likely driver is reducing the creation of reactive oxygen species by repacking of amino acid side chains around the chromophore. Two error prone rounds resulted in the removal of methionine residues, mCitrine to Citrine1.7 and Citrine1.10 to Citrine 1.11, may reduce photobleaching by simply removing favourable targets for oxidation, as methionine is prone to oxidization. Although, as all residues are vulnerable to reactive oxygen species, the mutations may just be creating an internal structure tolerant to damage such that several oxidation events are required, on average, to quench or destroy the chromophore's ability to fluoresce.

The concentration dependence of Citrine 1.7 and later variants created a bias during screening for variants that expressed the FP at a higher concentration. All of the new variants appeared much brighter than mCitrine in colonies and in liquid culture due to higher expression. All of the variants also yielded higher protein concentrations after protein extraction suggesting that screening under this white light bleaching condition also puts selection pressure on protein expression along with photostability and brightness. This expression increase is coupled to the concentration dependence of photostability, as higher concentrations led to an apparent increase in photostability for Citrine 1.7 and later variants. As the on-plate screening cannot normalize the concentration in colonies, this strongly favours

high expressing variants when concentration and photostability have a direct relationship. When an FP has an indirect relationship between concentration and photostability, this screen will not function well as variants with lower expression would be more photostable and dimmer. In order to screen such an FP, the concentration dependence needs to be converted to have a direct relationship, such that higher concentrations result in higher photostability, this was observed in the error prone round that created Citrine1.7 from mCitrine.

The two most likely causes of concentration dependence in FP photobleaching are dimerization and reactive oxygen species generation. As FPs dimerize, there is a subtle rearrangement of the protein structure which can alter FP characteristics and result in either increased or decreased photostability. The dimerization of avGFP-based FPs is very weak, with the strongest dimer having a K_d of 0.11 mM.[154] This is five-fold more concentrated than the most concentrated solution of any Citrine variant tested, implying that very little dimerization is occurring. The creation of reactive oxygen species by FPs can damage neighbouring proteins.[85] As the concentration of FPs increase, more reactive oxygen species are produced, which is expected to cause an apparent decrease in photostability for the FP. This reactive oxygen species hypothesis may explain the reduced photostability seen at high concentrations for mCitrine.

Citrine2 maintains similar environmental sensitivity to mCitrine. In terms of pH, Citrine2 has a smaller Hill coefficient than mCitrine (0.51 versus 0.60). Accordingly, changes in pH should have less influence on Citrine2's fluorescence than on mCitrine's fluorescence. However, this improvement is mitigated by a higher pK_a of 5.9 for Citrine2 relate to 5.3 for mCitrine. Overall, Citrine2 ends up marginally more sensitive to pH than mCitrine but still bright across the physiologically relevant pH range of 5.5 to 7.5. Sensitivity to anions such as

chloride and iodide was noted in precursors (*i.e.*, eYFP) to Citrine2. At high concentrations, Cl^- and I^- caused minor increases of fluorescence for both mCitrine and Citrine2. This is in stark contrast to the 40% and 80% decreases in fluorescence seen in eYFP for both Cl^- and I^-, respectively, over the same concentration range.[158] The Q69M mutation that abolished this sensitivity in Citrine was modified again to M69T in Citrine2. This supports the assertion of Griesbeck et al.[153] that the halide sensitivity is abolished by repacking the protein interior to block a putative halide-binding cavity, rather than a specific interaction with the methionine or now also a threonine side chain in location 69.

In mammalian cell imaging experiments, Citrine2 proved to be more photostable and equally bright relative to mCitrine. The Citrine2 photobleaching half-lives observed both in confocal and widefield imaging are 1.35- and 1.40-fold better than mCitrine, respectively. The deviation of the fold-change in the half-lives these two imaging modalities appears to be due to the different light intensities of illumination,[144] the exponential behaviour of the photobleaching decay also changes. Under widefield illumination conditions, Citrine2 photobleached monoexponentially but under confocal illumination it photobleached biexponentially. mCitrine remains biexponential in both, but the fraction of each exponent changes significantly, with 36% at a 14 s half-life and 64% at a 57.9 s half-life under widefield illumination. For confocal illumination mCitrine photobleaches with 80% with a 1.2 s half-life and 20% at a slower 5.8 s half-life. These results suggest that different photobleaching mechanisms become dominant in different illumination intensity regimes.

3.4 Methods

3.4.1 Molecular cloning and mutagenesis

To create the libraries for photostability screening, EP-PCR was performed using Taq polymerase (New England Biolabs). 50 µL reactions were created using: 5 µL of 10× Taq buffer,2 µL of an NTP mix containing dATP (5 mM), dGTP (5 mM), dCTP (25 mM), and dTTP (25 mM), 4 µL MgSO4 (25 mM), 1.5 µL of each DNA primer (10 µM), 1 µL of template (~2 fmol), 1 µL of Taq, 0.5-1 µL $MnCl_2$ (10 mM), and deionized H_2O up to 50 µL. The $MnCl_2$ was added last. Standard PCR were performed using Pfu polymerase and Q5 high fidelity polymerase (New England Biolabs) using their recommended protocols. DNA primers were purchased from IDT. All PCR products were run on 1% agarose gels with 1 µL of 10 mg/mL ethidium bromide. The relevant bands were visualized with UV light, cut out with a razor blade, and the DNA was extracted using gel extraction kits (Thermo Fisher Scientific or BioBasic) using the manufacturer's recommended protocols.

Gibson Assembly (New England Biolabs) was used to assemble the pBAD His B plasmids (Thermo Fisher Scientific) for transformation. A reduced scale protocol was used with 1 µL vector, 1.5 µL of insert, and 2.5 µL of 2× Gibson Assembly Master Mix. This mix was then incubated for 4 hours at 50 °C and left at room temperature until needed for transformation, 0-48 hours. It was then diluted with 5 µL of deionized H_2O before transformation. *E. coli* strain DH10B (Thermo Fisher Scientific) was then transformed with the resulting plasmids using electroporation.

To insert constructs into pcDNA3.1(+), the pBAD template containing the insert, and pcDNA3.1(+) vector were digested using FastDigest XhoI and HindIII (Thermo Fisher Scientific) in 1× FastDigest buffer for 15 min with no thermal inactivation, then run on an

agarose gel and extracted, as described above. The insert and vector were combined in a 6:1 ratio and ligated using T4 ligase (Thermo Fisher Scientific) in 1× T4 ligase buffer for 15 min at room temperature.

The transformed cells containing the pBAD/His B plasmids were plated onto agar with LB medium and 0.4 mg/ml ampicillin and 0.02% w/v L-arabinose, and grown overnight at 37 °C. On-plate photostability colony screening was performed using illumination from a 300 W xenon arc lamp and a Lego Mindstorms EV3 robot in the plate bleach configuration. The arc lamp is placed 2 cm from the Petri dish gel surface and aimed such that the light just touches the center point around which the Petri dish rotates, and on the center line such as the robot drives forward it will move directly away from the center of rotation. The robot needs to be aligned such that it will drive perpendicularly to the illumination such that the distance between the light source and the Petri dish does not change.

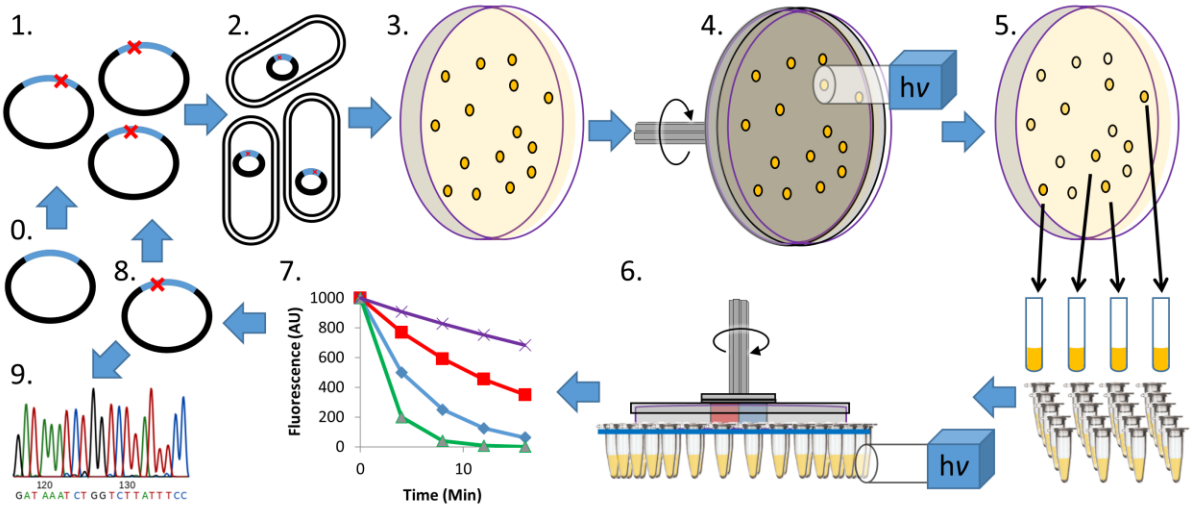

Figure 3.10 Robot assisted photobleaching screening. Clockwise from the middle left shows 0. The initial plasmid. 1. Plasmid library generation. 2. Transformation of *E. coli*. 3. Growth on a Petri dish. 4. Plate bleach configuration bleaching. 5. Selection and growth of brightest variants and followed by BPER protein extraction. 6. Bleaching using the tube bleach configuration. 7. Determination of half-lives. 8. Selection of new templates for the next round. 9. Sequencing.

Using the concentric circle bleach the time spent on the inner most circle was 6 seconds initially but after Citrine 1.10 a longer time of 10 s was used to differentiate between the more photostable variants. The distance that the robot drives should be calibrated periodically to ensure that the robot drives 3 mm, $1/5^{th}$ of the illumination diameter, each step.

The brightest colonies after photobleaching were selected grown in LB and their cell lysate acquired through BPER extraction. The BPER extracted cell lysate was then diluted to an absorbance of 0.1 at peak absorbance, ~510 nm, with H_2O and a 50 µL aliquot was photobleached using the 300 W xenon arc lamp and a Lego Mindstorms EV3 robot for 10-20 min in the tube bleach configuration, described below. Both a photobleached and non-bleached control sample were measured to obtain the percentage of fluorescence remaining. Winners were selected using InPhO, first a rough two point half-life was calculated from the percentage of fluorescence remaining and then a rough brightness value was calculated using the absorbance normalized fluorescence of the non-bleached sample. The variants with the highest InPhO went into the plasmid pool for the next round of EP-PCR as depicted in Figure 3.10.

3.4.2 Protein purification and characterization

A single colony of *E. coli*, transformed with the pBAD/His B plasmid containing an mCitrine variant, was used to inoculate 4 mL LB supplemented with 0.1 mg/mL ampicillin. This culture was incubated at 37 °C in a shaker incubator (220 rpm) for 12 hours. The liquid culture was then added into 500 mL of LB with 0.1 mg/mL ampicillin and incubated for a further 4 hours. The inducer L-arabinose was added to a concentration of 0.02% and the culture was allowed to grow overnight at 30 °C in a shaker incubator (220 rpm). The cells were then

centrifuged at 15 000 RCF for 10 min at 4 °C, the supernatant was discarded and the pellet resuspended in 25 mL of 1×TBS. The cells were then lysed using a cell disruptor (Constant System). The cell debris was removed by centrifugation at 15 000 RCF for 35 min at 4 °C. The protein was then purified using Ni-NTA agarose beads (MC Labs) according to the Native conditions protocol, and buffer exchanged with 1×TBS using a centrifugal filter unit with 10 000 MWCO (Amicon). Extinction coefficients were determined by measuring the absorption spectrum using a UV/Vis spectrometer (Beckman Coulter DU 800) at pH 7.25 in 1×TBS and the same concentration of the FP in 1 M NaOH. In 1 M NaOH, the chromophore has an extinction coefficient of 44 000 $M^{-1}cm^{-1}$ at their absorption peak near 450 nm.[13] Application of the Beer-Lambert law gives the concentration of the alkaline denatured chromophore, which can be used to calculate the extinction coefficient for the intact protein at pH 7.25. Quantum yields were determined as described in chapter 2, using mCitrine (quantum yield = 0.74) as the standard.[144]

3.4.2.1 BPER extraction of cell lysate

The brightest colonies after photobleaching were selected using a combination of visual inspection under 480-500 nm illumination with long pass filter goggles and digital fluorescence imaging using a 515-545 nm emission filter. Picked colonies were transferred to liquid cultures and grown in 2-6 mL LB and 0.1 mg/ml ampicillin and 0.02% w/v L-arabinose in a shaker at 240 rpm and 37 °C, for 16 hours. The liquid cultures were then spun down at 15 000 RCF for 2 min and the supernatant was poured off. Then 100 µL of BPER (Thermo Fisher Scientific) was added and the samples were vortexed for 5-60 min. The resulting suspension was then spun down at 15 000 RCF for 2 min and the supernatant, BPER extracted

cell lysate, was collected and tested for fluorescence and absorbance in a fluorescence plate reader (Tecan Safire2). The pellet left after the BPER extraction was used as the starting point for plasmid purification. Miniprep kits (Thermo Fisher Scientific and BioBasic) were used according to their respective standard protocols.

3.4.2.2 Environmental sensitivity

The pK_a was determined in a series of pH buffers (pH 3 to 11) using the Carmody buffer system.[137] For each protein solution, 2 µl was added to 50 µl of the desired pH buffer in a 396-well clear-bottomed plate (Thermo Fisher Scientific) and the fluorescence measured in a Tecan Safire2 plate reader. pK_a values were extracted by fitting the data against a theoretical curve. Chloride and iodide sensitivity was determined by adding 2 µL of purified protein; the protein starts in 1×TBS, then 5 µl of 1M NaCl or 1M NaI, and then diluted with deionized H_2O up to 50 µL. This was compared against the same solution omitting the NaCl or NaI.

The effect of oxygen on photobleaching was probed using EC-Oxyrase to catalyse the conversion of dissolved oxygen into water. 300 µL BPER cell lysate diluted with H_2O to a peak absorbance of 0.1, then 1.5 units of EC-Oxyrase were added in a closed 2 ml microcentrifuge tube. After 5 min the solution was aliquoted into PCR tubes and half-lives determined via the tube bleach protocol described above.

3.4.2.3 Fluorescence lifetime

Fluorescence lifetimes were measured on a TimeMaster time-resolved spectrofluorimeter (Photon Technology International) that uses the stroboscopic optical boxcar technique.[159] All lifetimes were determined in 1×TBS at protein concentrations of

approximately 0.1 μM. A nitrogen dye laser at using Coumarin 500 (Exciton Inc.) was used to create the 500 nm excitation, and the emission monochromator was set to 530 nm with a slit width of ~1 nm. A time window of 40 ns to 100 ns was measured over 301 channels. Each channel was measured with 10 shots that were averaged. The time window was scanned 3 times and averaged again. Various ND filters were used to match the max photon counts for each sample. The IRF was determined using a dilute solution of LUDOX® (Sigma-Aldritch) with 500 nm excitation and emission.

3.4.3 HeLa cell imaging

HeLa cells were maintained in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum (Gibco), Penicillin-Streptomycin (Gibco), GlutaMAX (Gibco) at 37 °C with 5% CO_2. Transient transfections of pcDNA3.1(+) expression plasmids were performed using Lipofectamine 2000 (ThermoFisher Scientific). HeLa cells were grown to 60-70% confluency on 35 mm glass bottom dishes and were transfected with 1 μg of plasmid DNA and 2 μL lipofectamine 2000 according to the manufacturer's instructions. The cells were imaged 24−48 h after the transfection. Immediately prior to imaging, cells were washed twice with Hanks balanced salt solution and then 1 mL of 20 mM HEPES buffered HBSS was added. All imaging was performed at room temperature.

Epifluorescence/widefield cell imaging was performed with an inverted Eclipse Ti microscope (Nikon) equipped with QuanEMCCD camera (Photometrics) and driven by NIS-Elements AR software package (Nikon). Cells were imaged with a 60× oil objective lens (numerical aperture (NA) = 1.49). The fluorescence was detected by 525/25 nm emission and

488/10 nm excitation filters. The photobleaching was done with 75 W xenon lamp and was recorded for 3 min with 1 s intervals with a 30 ms exposure time.

Confocal imaging was performed with an IX-81 motorised microscope base equipped with a 50 mW 491 nm pumped iodide laser, a CSU10 spinning disk (Yokagawa), a confocal scan head, and C9100-13 EMCCD camera (Hamamatsu). Cells were imaged with 60× oil objective lens (NA = 1.42) and a 540/30 nm emission filter. The photobleaching was done with full laser power (50 mW) and was recorded for 1 min with 0.2 s intervals with an exposure time of 50 ms.

3.4.4 Fluorescein plates

The fluorescein Petri dishes, used for testing the amount of bleaching received across the plate, were made of 3.75 g of Agar placed into a 1 L Erlenmeyer flask with 500 mL of H_2O. The solution was microwaved until all the agar had melted. Fluorescein was added to the hot agar until the solution had obvious green fluorescence (~1 μM). This was then poured to cover the bottom of each Petri dish with approximately equal amounts of the fluorescein agar solution then cooled before use.

3.5 Conclusions

The robotic white light photostability screening system rapid and facile screening of a large number of variants on the basis of their photobleaching half-lives. This system was successfully applied to the creation of a bright and photostable yellow FP designated as Citrine2. This new FP represents a useful new addition to the tool box of photostable fluorescent proteins for live cell imaging.

Progress toward a screening assay for enhanced PhoCl variants

4.1 Introduction

The traditional use of FPs has been as a fluorescent tag for proteins in living cells. This use revolutionized cellular biology by vastly expanding our understanding of living systems. As FPs are such useful tools, significant work has been done to understand them fully and to enhance their utility. It has become possible to modify FPs to the point that they are no longer merely passive imaging tags, but active tools to modify cellular function. These are a type of genetically encoded tool that are controlled by light, which as a group are called optogenetic actuators.

One of the earliest FP-based optogenetic actuators is KillerRed,[85] an FP derived from the hydrozoan chromoprotein anm2CP.[160] It is a genetically encoded photosensitizer, meaning that it catalyzes the production of reactive oxygen species (ROS) when illuminated with 540-580 nm light. Using CALI[161] KillerRed is used as an off switch, disabling fused proteins with the ROS it produces by oxidizing aromatic residues, cleaving peptide bonds, and crosslinking side chains.[87] Although, it would be more accurate to describe it as a grenade. While both a grenade and a switch are effective ways to turn off a light bulb, the side effects are slightly more severe when using a grenade. As its name implies, KillerRed and its progeny KillerOrange[162] and SuperNova[163] kill cells.

Efforts to create less destructive FP optogenetic actuators have resulted in useful constructs. Dronpa[89] is a reversibly switchable FP from a Pectiniidae coral. A reversibly switchable FP is one that can have its fluorescence turned off and on using specific wavelengths of light. It was later determined that Dronpa's photoswitching was coupled with a multimerization (k_d, 10 μM to 100 μM) in the fluorescent on state with 390 nm illumination

and monomerization (k_d > 100 μM) in the non-fluorescent off state with 480 nm illumination.[90] This controllable multimerization was harnessed to control the function of proteins by caging them in the fluorescent on state and freeing them to function in the fluorescent off state (Figure 4.1).

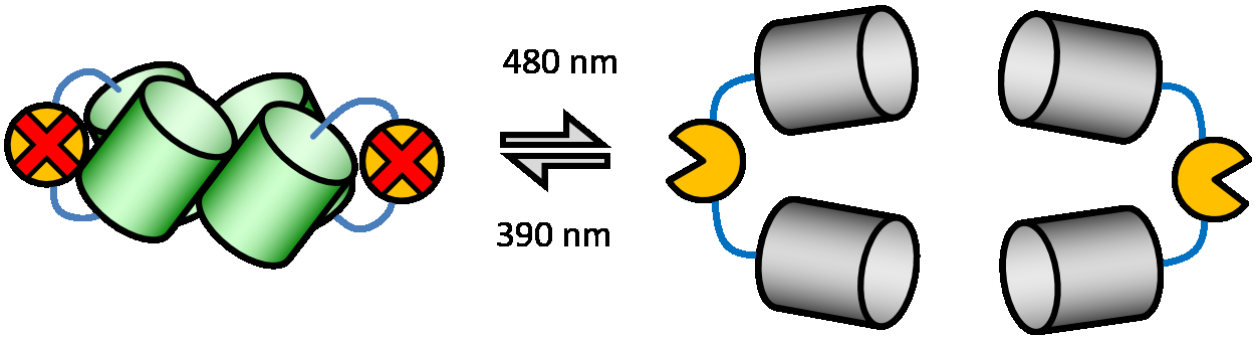

Figure 4.1 Dronpa caging and uncaging. 480 nm illumination results in monomeric Dronpa with an uncaged protease (Right). 390 nm illumination results in tetrameric Dronpa with caged protease (Left).

PhoCl, a PhotoCleavable protein, was developed from a *Clavularia* coral FP (cFP484) by Wei Zhang in the Campbell lab. The natural protein cFP484, was evolved to be monomeric and bright creating mTFP1.[94] This mTFP1 was then engineered to be a green-to-red photoconvertible FP, mMaple.[164] The mMaple template is the starting point from which evolution to create a photocleavable protein was commenced. In the green state, mMaple is excited maximally with 489 nm light and emits maximally at 505 nm. After photoconversion to the red state, this shifts to 566 nm excitation with 583 nm emission. The green-to-red conversion occurs upon illumination with ~400 nm light. This 400 nm light is absorbed and results in a β-elimination reaction that induces a cleavage of the FP backbone. This cleavage also extends the conjugated system of the green chromophore converting it into a red chromophore.[89,165] Both mMaple and mTFP1 are tolerant of circular permutation

demonstrating that this lineage of proteins is highly amenable to manipulation.[43,166] After photo-induced cleavage thePhoCl would separate into two pieces (Figure 4.2).

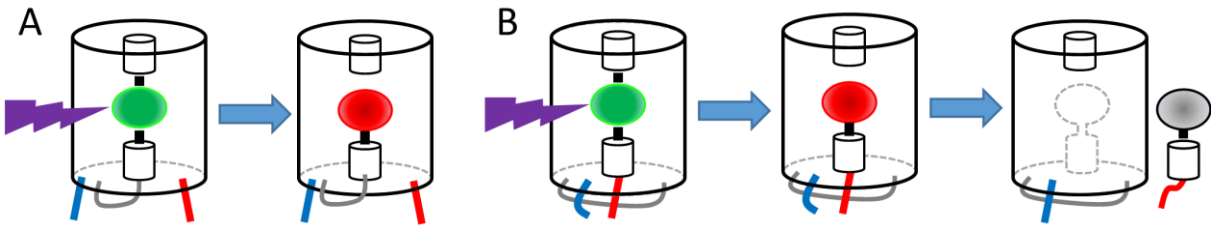

Figure 4.2 The modification of mMaple into PhoCl. A) mMaple. B) PhoCl. Purple lightning is 400 nm light inducing conversion of the chromophore of mMaple and PhoCl. The blue and red lines represent the N and C termini respectively. The grey line is the loop/linker which was modified during circular permutation.

PhoCl's circular permutation puts the N terminus at location 79, and the C terminus at position 78, using mMaple numbering. This location is in the loop connecting the central helix to the 4th β-stand of the β-barrel. A 5 amino acid linker GGSGG was used to connect the old N and C termini. After photocleavage the chromophore and 9 amino acids are no longer covalently attached to the rest of the FP barrel and are able to dissociate. The chromophore and attached residues are called the releasing peptide. Once outside of the FP barrel, the chromophore is protonated and non-fluorescent with an absorbance maximum shifted to 446 nm. This matches the absorbance seen in all FP tyrosine-based chromophores when exposed to solvent.[13]

Dronpa's one major advantage over PhoCl is that it is fully reversible allowing dynamic control of the caged protein of interest (POI). PhoCl functions like Pandora's Box. Once open/photocleaved the contents will freely act on the environment never to be caged again. This does give PhoCl an advantage in generalizability, if the PhoCl construct can completely block the function of the fused POI, after cleavage the POI's function should return as the two

dissociate. In Dronpa the fused POI never fully leaves the cage requiring the attached Dronpa cage to strongly interfere as a multimer and weakly interfere as a monomer, while still remaining tethered together. Dronpa function is dependent on conformational changes for caging, which are difficult to predict and control.

One application of PhoCl is fusing an inhibiting peptide to the N-terminus of PhoCl and the POI to the C-terminus. Before cleavage, the inhibitory peptide is held close to the POI allowing strong inhibition due to the high effective concentration. After cleavage, the POI will dissociate from the PhoCl barrel/inhibitor removing the high effective concentration. The cleavage induced dissociation must shift the effective concentration from well above the k_d of the inhibitor, where all of the POI is constantly bound and inactivated by the inhibitor, to a concentration where the inhibitor will dissociate and the POI become active (Figure 4.3). This works best with low nM k_d binding as it ensures that all the POI is inactivated. This design was successfully used to cage HCV protease with its inhibitor.[167] This approach was also not largely generalizable, as not all POI would have an appropriate k_d, peptide or protein, inhibitor.

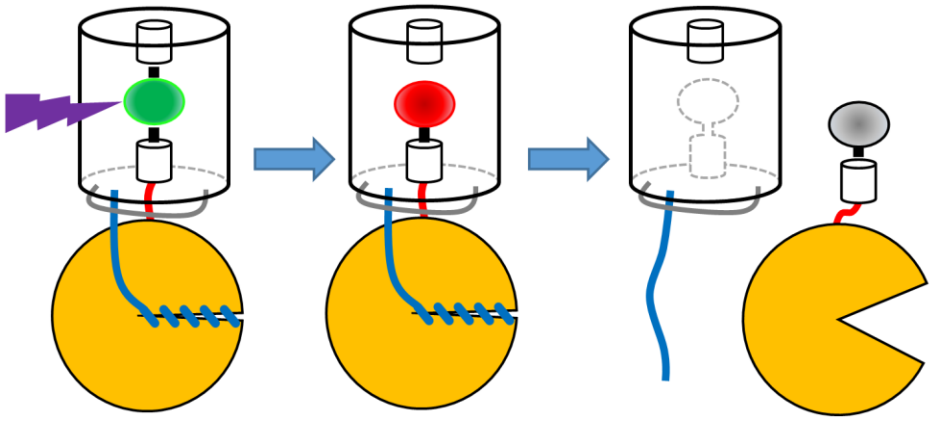

Figure 4.3 Scheme of the PhoCl-HCV protease construct. The yellow circle is HCV protease, the blue line is the inhibitor for the HCV protease. 405 nm illumination converts PhoCl's chromophore to red and breaks the protein backbone. The photocleavage allows the release of HCV and subsequent dissociation from the inhibitor causing HCV to be active.

Steroid receptor (SR) domains remove the requirement for POI specific inhibitors as they act as general protein inactivators.[168] Rather than directly inhibiting a POI they are thought to recruit heat shock protein 90 (Hsp90) which results in inactivation of the fused POI, and is natively present in mammalian cells.[169] Fusing two copies of the SR to a POI resulted in more consistent inactivation,[170] as such a doubled PhoCl system was created. This SR-PhoCl-POI-PhoCl-SR construct recruits two Hsp90 proteins and after illumination with 400 nm light cleaves into three fragments: 1) a SR domain fused to a PhoCl releasing peptide, 2) an SR domain fused to an empty PhoCl barrel, and 3) the now uncaged POI fused to an N-terminal PhoCl releasing peptide with a C-terminal empty PhoCl barrel (Figure 4.4). Fortunately, the Hsp90-SR complex does not interfere with the function of PhoCl itself allowing its use as a general caging system. It has been successfully applied to the Gal4-VP16 transcription factor, Cre recombinase, and should be viable for many other proteins.

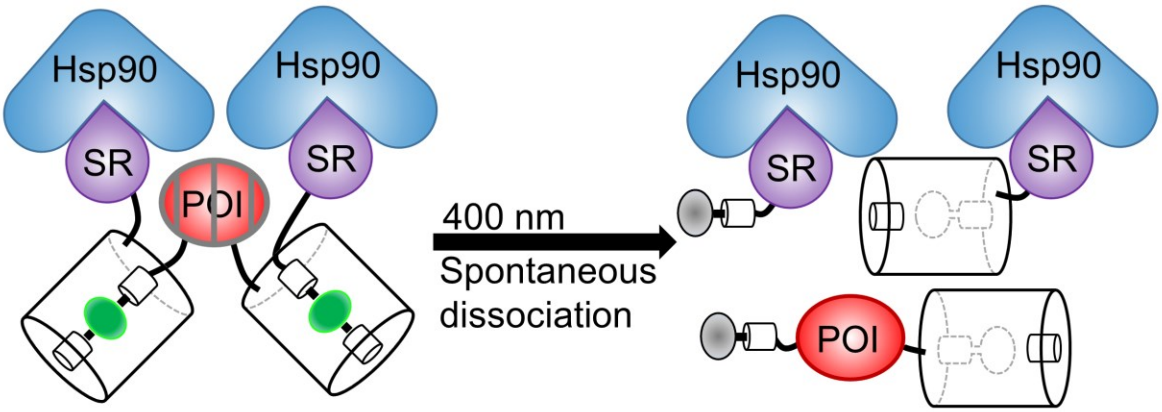

Figure 4.4 Doubled PhoCl Steroid receptor caging of POI.

Despite the success of PhoCl there are a number of key properties that could be further improved. The screen that was used to evolve PhoCl was based on fast maturation of green fluorescence, the conversion efficiency of green-to-red, and the loss of red fluorescence after photoconversion (Figure 4.5). This was done in *E. coli* with cytoplasmic expression of PhoCl with no fused proteins.

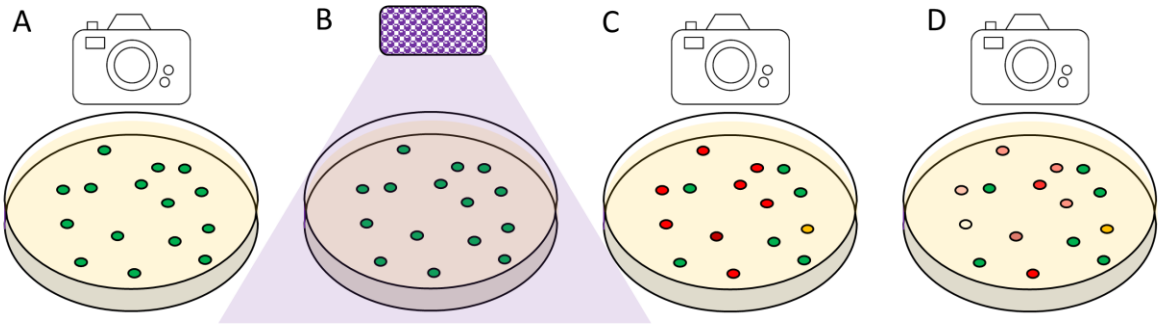

Figure 4.5 Screening procedure use to develop PhoCl. A) A library of PhoCl variants imaged for green fluorescence brightness. B) The Petri Dish of variants was photoconverted for 5 minutes in the 405 nm LED photoconversion chamber. C) The Petri Dish was imaged for red fluorescence. D) After waiting 5-10 minutes the red fluorescence was measured again. Samples with high green fluorescence in A) high red fluorescence in C) and low red fluorescence in D) were selected to make the next PhoCl library.

The reason this screen is flawed is that it does not actually screen for the separation of the releasing peptide and the barrel. To test for separation, PhoCl was fused to the maltose binding protein, expressed in E. coli, purified, and then photoconverted. The resulting protein solution was then analyzed using gel filtration chromatography and fractions were then run on SDS polyacrylamide gel electrophoresis to check the sizes of the fragments. This process take days and many hours per sample so was only done once in the evolution of PhoCl. The following work was aimed at developing a screening method that could, in a high throughput manner, be used to screen PhoCl variants according to their ability to dissociate after photoconversion in order to create a better PhoCl.

4.2 Results and discussion

4.2.1 PhoCl kinetics

Tracking the decay of red fluorescence was the original proxy for dissociation. These tests were done by measuring the red fluorescence immediately after photoconversion and then 5-15 minutes later. In the first 15 minutes PhoCl, with no fusion partner, loses about 50% if its red fluorescence and selection was based on the variation of this value. However, if the red fluorescence is tracked over several hours a very different pattern emerges (Figure 4.6B). The red fluorescence of converted PhoCl initially decays rapidly then gradually recovers ~60% of the lost red fluorescence up to the 5 hour mark where it then slowly loses the red fluorescence. Twenty hours after photocleavage, only 5% of the red fluorescence remained (data point not shown). Photocleavage caused the green fluorescence to decrease to 20% of its initial value and did not fluctuate further (Figure 4.6A).

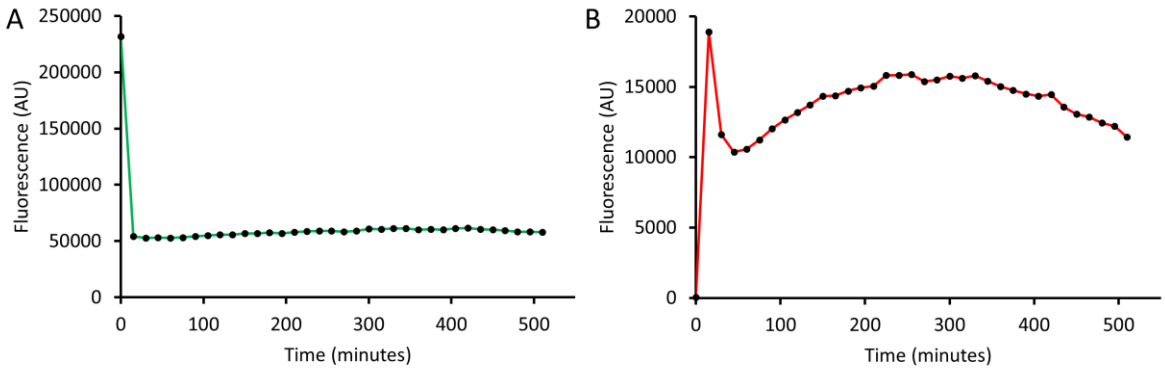

Figure 4.6 PhoCl fluorescence after photoconversion. PhoCl was photocleaved for 5 minutes in the 405 nm LED photoconversion chamber in the gap between time points 0 and 15 minutes. A) Green fluorescence using 470 nm excitation and 505 nm emission. B) Red fluorescence using 540 nm excitation and 575 nm emission.

The recovery of the red fluorescence implies that the release of the cleaved PhoCl chromophore is a multi-staged process when observed using fluorescence. When tested using electrospray ionization mass spectroscopy (ESI MS) the initial drop in red fluorescence is due to some of the releasing peptides successfully dissociating (Figure 4.7).[171] However, as ESI puts the cleaved PhoCl complex into the gas phase changes the dissociation characteristics possibly over representing the dissociated state.[172] The recovery of red fluorescence could potentially be the protein finding a more stable conformation around the red chromophore, if this is that case the red fluorescence could then either slowly decay away by release of the chromophore, or a gradual shifting to a more stable non-fluorescent state. The hours long time scale for the red fluorescence changes is far longer than the average protein conformational change. It is possible that the chromophore, after release, slowly creates a hydrophobic aggregate resulting in excimers capable of fluorescence. In *E. coli* colonies, the red fluorescence of PhoCl remains easily detectable even days after photocleavage. In HeLa cells, the red fluorescence disappears overnight. While it is convenient that PhoCl functions better

in mammalian cell types, its dependence on environmental conditions to fully dissociate or the potential formation of excimers is not desirable. Regardless of which mechanism or combination of mechanisms are occurring PhoCl does not reach a post cleavage equilibrium in terms of fluorescence for several hours.

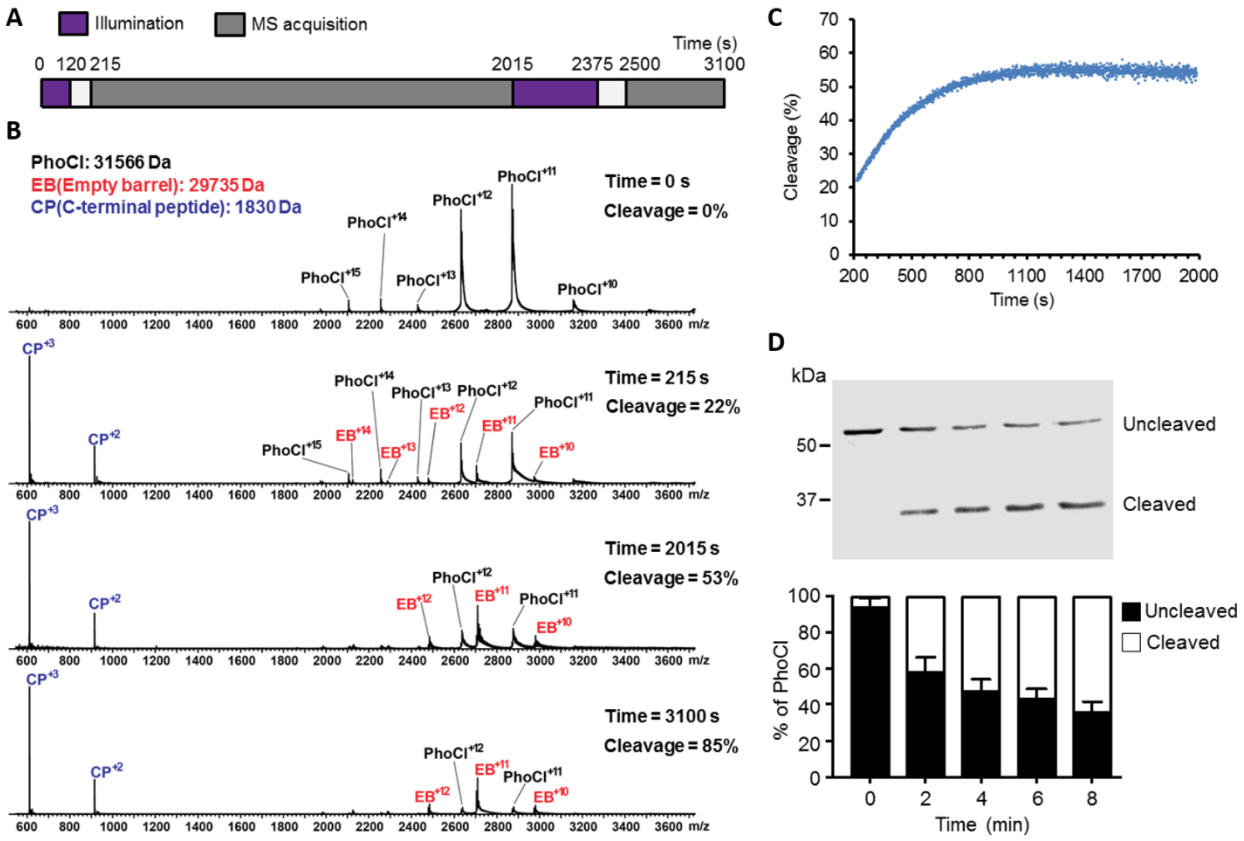

Figure 4.7 Demonstration of PhoCl photocleavage by mass spectrometry and western blot.

Schematic representation of illumination and time-lapse electrospray mass spectrometry (ESI-MS) data acquisition. PhoCl was photoconverted in the 405 nm LED chamber (0.15 mW/mm2) with a 2 min illumination starting at t = 0 s, and a 6 min illumination starting at t = 2015 s. ESI mass spectra were acquired every 1 s for 30 min starting at t = 215 s, and for 10 min starting at t = 2500 s. B) ESI mass spectra of PhoCl acquired at t = 0 s (before photoconversion), 215 s, 2015 s and 3100 s. Percent cleavage was calculated as EB/(EB+PhoCl) × 100, where EB and PhoCl are the sum of the peak heights for various corresponding ionization states. C) Percent cleavage versus time, from t = 215 s to t = 2015 s. D) Quantitative western blot of PhoCl-mCherry-myc photocleavage following 405 nm illumination (2, 4, 6, 8 min at 0.15 mW/mm2). Membranes were probed with an anti-myc primary antibody. Violet light induced a band shift from ~60 k_d (full length PhoCl-mCherry-myc) to ~30 k_d (mCherry-myc) indicating the cleavage of PhoCl. Values are means ± standard deviation (n = 3). Figure and caption reproduced with permission.[171]

The green state of PhoCl has a quantum yield of 0.46 and an extinction coefficient of 28 000 $M^{-1}cm^{-1}$ for a total brightness of 13. As PhoCl's red state fluorescence changes over the period of time required to perform the extinction coefficient and quantum yield tests, the red state values were not determined. mMaple's green state quantum yield of 0.74 and extinction coefficient of 15 000 $M^{-1}cm^{-1}$,[164] are significantly different from PhoCl's and thus we assume mMaple's red state does not match PhoCl's. This means we cannot readily measure the fraction of PhoCl that is converted and mutations that change the green-to-red ratio before and after conversion, may be changing quantum yield and extinction coefficients of the two states rather than the conversion efficiency.

4.2.2 Caging split fluorescent proteins

In an effort to create an improved version of PhoCl, several different constructs were built. The first two designs are based on caging a split FP (Figure 4.8). The idea of these constructs is to make a fluorogenic version of PhoCl for screening purposes. Once photocleaved the caging will cease and the split FP will reconstitute becoming fluorescent. This functions as a screen for PhoCl, as a brighter fluorogenic signal should directly correlate to increased PhoCl photoinduced dissociation.

In the caged-S11 construct the S11 strand is caged by inserting it into the loop connecting the releasing side of the central helix to β-strand 3 (Figure 4.8A). This is like performing a second circular permutation, using S11 as the linker, resulting in a configuration matching mMaple's. The S11-plug construct uses the traditional PhoCl configuration. The S11 fragment is caged by attaching it to the releasing peptide with a minimal linker. The linker

should be so short that the PhoCl barrel will block sfGFP 1-10 from associating with the S11 fragment.

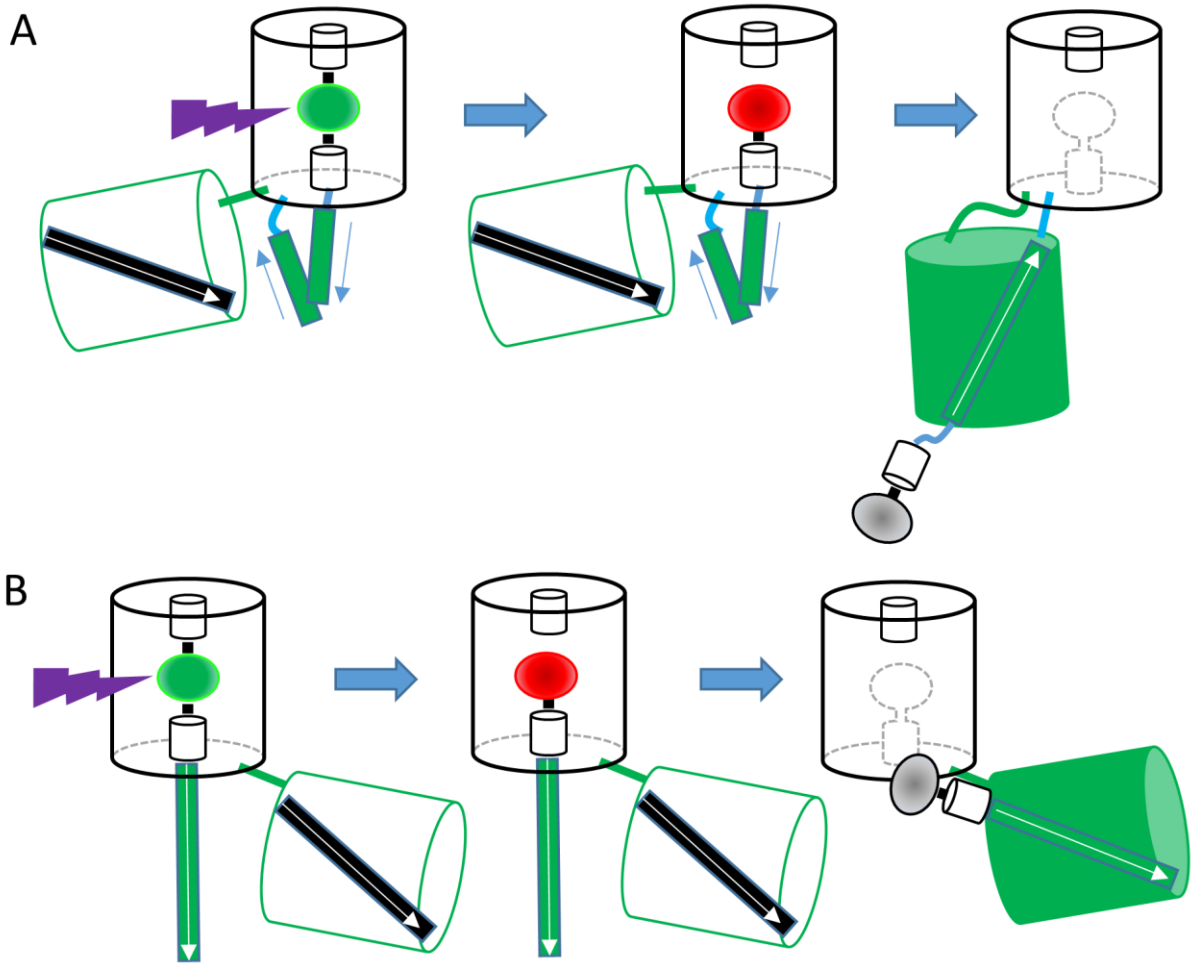

Figure 4.8 Fluorogenic PhoCl constructs. A) The caged-S11 construct uses a cpPhoCl with S11 of sfGFP acting as the linker with S1-10 of sfGFP on the C-terminus of the cpPhoCl. B) The S11-plug construct use PhoCl with S11 of sfGFP on the C-terminus with S1-10 on the N-terminus.

Both constructs initially have PhoCl's green fluorescence, which has a brightness of 13. After photocleavage the split sfGFP can reconstitute and fabricate its chromophore, which has a brightness of 38.[144] Meaning that for every PhoCl chromophore that is cleaved it can be

replaced by a chromophore that is 3-fold brighter. As PhoCl's green fluorescence is destroyed in the photocleavage, these designs were intended to allow screening for the release of the chromophore using green fluorescence after photocleavage. For this to function in *E. coli*, the green fluorescence would have to increase above the precleavage fluorescence. As the *E. coli* colonies continually express more of the PhoCl construct the increase in fluorescence must happen fast enough that it is not masked by the maturation of new PhoCl proteins.

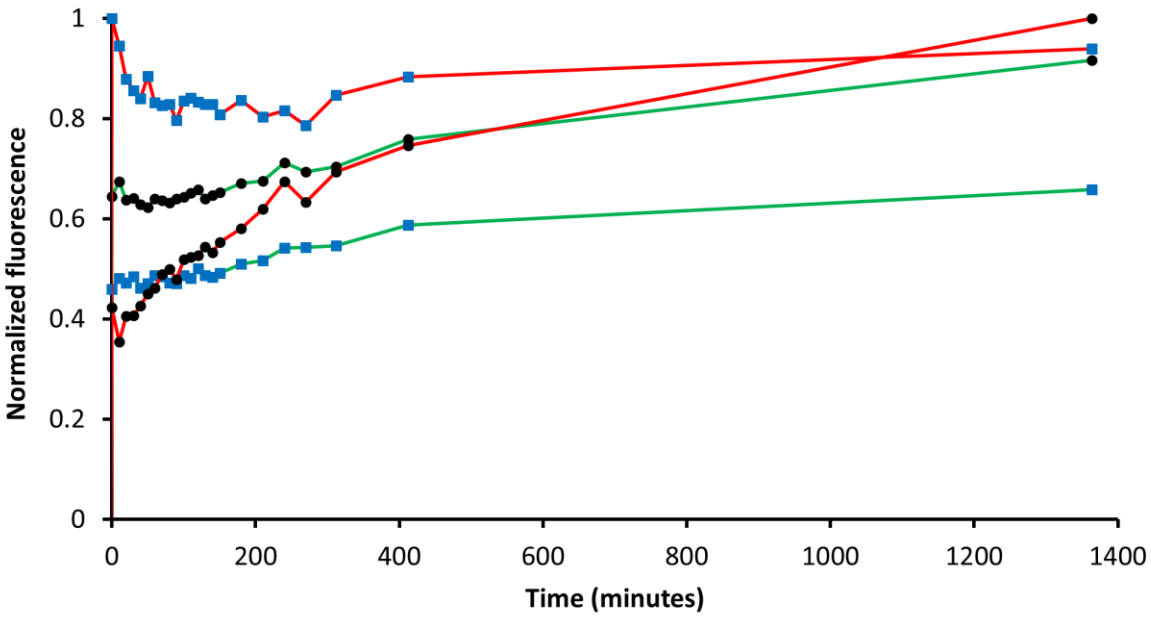

Figure 4.9 Kinetics of Caged-S11 and S11-plug constructs. Red and green lines represent the colour of fluorescence 575 nm and 505 nm respectively. Black circles and blue squares are the S11-plug and Caged-S11 constructs respectively. Time 0 is the first measurement after photocleavage for 5 minutes in the 405 nm LED photoconversion chamber. Fluorescence is normalized to the max fluorescence including the pre-photocleavage fluorescence.

Unfortunately, neither in the colonies nor the *in vitro* test shown in Figure 4.9, did the green fluorescence after conversion surpass the initial green fluorescence. Also the red fluorescence slowly recovers over the entire 24 hour period tested implying that the kinetics of

release have been altered in both constructs. On the positive side the green fluorescence did increase after photocleavage implying that both constructs functioned to the extent that split sfGFP does reconstitute, but not to an extent that allows screening for brightness at the colony level. These split sfGFP designs were abandoned as turnover of cleaved constructs for fresh whole constructs overwhelms the small change in fluorescence from split FP reconstitution.

4.2.2.1 Engineering blue and yellow colour shifted variants of split sfGFP in caged PhoCl constructs

As both PhoCl and sfGFP share nearly identical spectral profiles it was impossible to determine how much each FP contributed to the fluorescence. If the spectral profile of the split FP and PhoCl were sufficiently different it would become possible to track how much of each was fluorescing. We tried two separate single amino acid changes to sfGFP to change its colour to yellow or blue. T203Y creates a π-π stacking interaction that results in a red shift of GFP chromophores to yellow fluorescence. T65S blue shifts the excitation by allowing the protonated GFP chromophore to absorb the blue wavelength, then an excited state proton transfer creates the anionic form of the GFP chromophore emitting green light. Although both mutations resulted in colour shifted variants, the T203Y mutation, which we refer to as sfYFP proved to express and mature better and red-shifts the emission maxima by 15 nm to 520 nm.

Both the caged-S11 and S11-plug designs were mutated to include the T203Y mutation in the sfGFP1-10 fragment, becoming caged-S11-Y and S11-plug-Y respectively. The S11-plug-Y construct does not appear to successfully cage the S11 strand as even before photocleavage the fluorescence emission maxima is at 520 nm. The Caged-S11-Y construct

appears to have caged some of the S11 and has an initial fluorescence emission maxima at 510 nm. These constructs were then photoconverted for 300 s using150 mW 405 nm laser illumination and an optical chopper, to try to convert all of the PhoCl (Figure 4.10). The S11-plug-Y construct bleaches slightly but its fluorescence maxima remains at 520 nm. The Caged-S11-Y construct shifts 10 nm redder to 520 nm after photocleavage. This red shift matches the expected loss of green fluorescence leaving behind sfYFP that was not successfully caged. As this spectra was taken immediately after photocleavage the split sfYFP would not have enough time to form its chromophore so any yellow fluorescence existed before photocleavage. One hypothesis is that there a form of mutually exclusive folding between PhoCl and split sfYFP.[173] If PhoCl folds first it prevents split sfYFP from folding and vice versa. In the construct S11-plug-Y PhoCl sfYFP successful reconstitutes preventing the formation of PhoCl's chromophore. In Caged-S11-Y PhoCl completes it maturation such the PhoCl produces 70% of the fluorescence with sfYFP producing 30%.

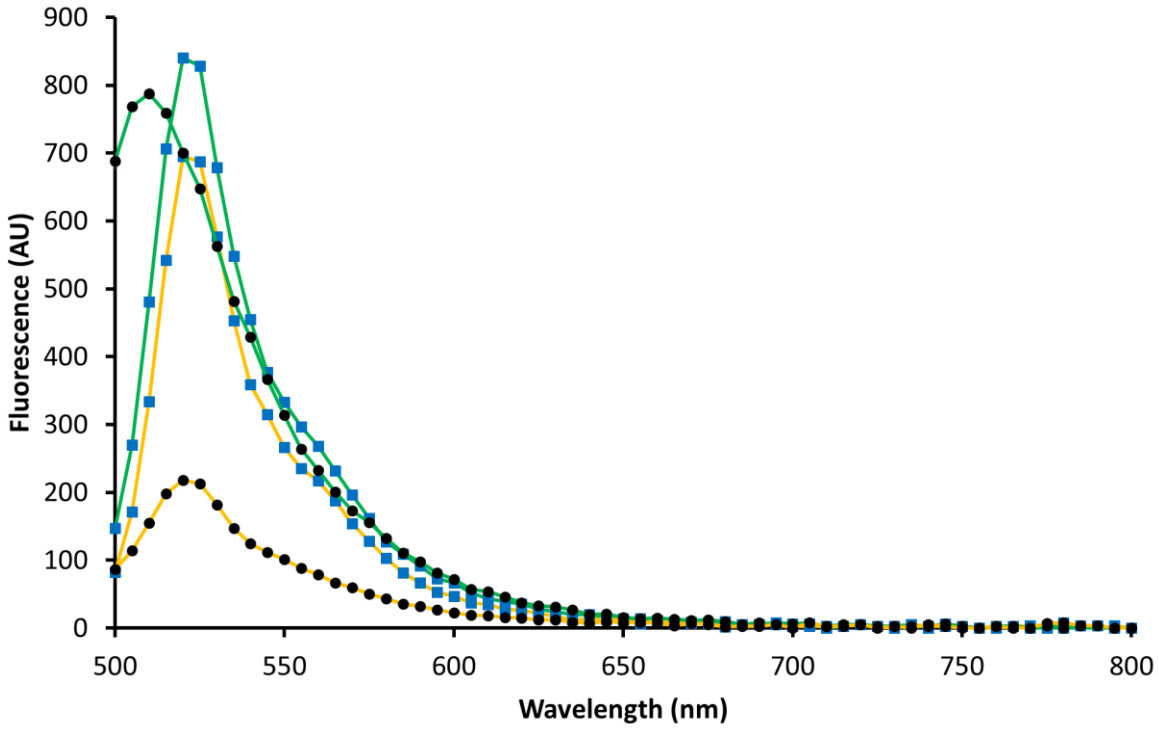

Figure 4.10 Photocleavage of caged-S11-Y and S11-plug-Y. Emission spectra (excitation at 475 nm). The black circles and blue squares are caged-S11-Y and S11-plug-Y respectively. The green lines are fluorescence before photocleavage by 300 s of optically chopped 150 mW 405 nm laser illumination; the yellow lines are after photocleavage.

The yellow fluorescence in the Caged-S11-Y construct never surpassed any part of the tail of the green fluorescence under any illumination or detection conditions making screening with this construct unfeasible. Attempts to improve the caging by shortening the linker between S11 and PhoCl in both constructs did not improve caging or yellow fluorescence recovery.

We later discovered that if samples that were not photoconverted were aliquoted and tested on two subsequent days, the split sfYFP's yellow fluorescence will have increased relative to the green suggesting that split sfYFP is capable of reconstituting at the cost of green fluorescence and properly folded PhoCl. This suggests that split sfYFP can pull the central

helix of PhoCl out of place, destroying its fluorescence, to complete the reconstitution of sfYFP. For this to occur PhoCl must fold first then slowly lose the tug-of-war with sfYFP. This reassembly of split sfYFP before photoconversion, despite the attempt at caging, rendered this an unsuitable screening method.

4.2.3 PhoCl FRET-based screening

The next attempt was to fuse a full FP to the releasing peptide of PhoCl such that the FP will dissociate from the empty PhoCl barrel after photocleavage. I initially chose mPapaya1, a further red shifted yellow FP with maximal excitation at 530 nm and emission at 541 nm.[19] This screening system is FRET-based. Initially, exciting mPapaya1 will result in full yellow fluorescence. After photocleavage the green chromophore of PhoCl is converted to a red chromophore. The absorbance of the red chromophore strongly overlaps with the emission of mPapaya1 making this into a FRET system. Some energy absorbed by mPapaya1 will be transferred to the red chromophore in PhoCl, resulting in decreased mPapaya1 fluorescence. Upon photoinduced dissociation the red chromophore is exposed to solvent. Red chromophores exposed to solvent have a hypsochromic shift from 530 nm to 446 nm. The exposed chromophore cannot be a FRET acceptor for mPapaya1. This breaks the FRET system causing mPapaya1's fluorescence to increase back to full. The result of this is that tracking the increase in yellow fluorescence from mPapaya1 after photocleavage should directly correspond to the red chromophore of PhoCl becoming incapable of absorbing yellow wavelengths (Figure 4.11).

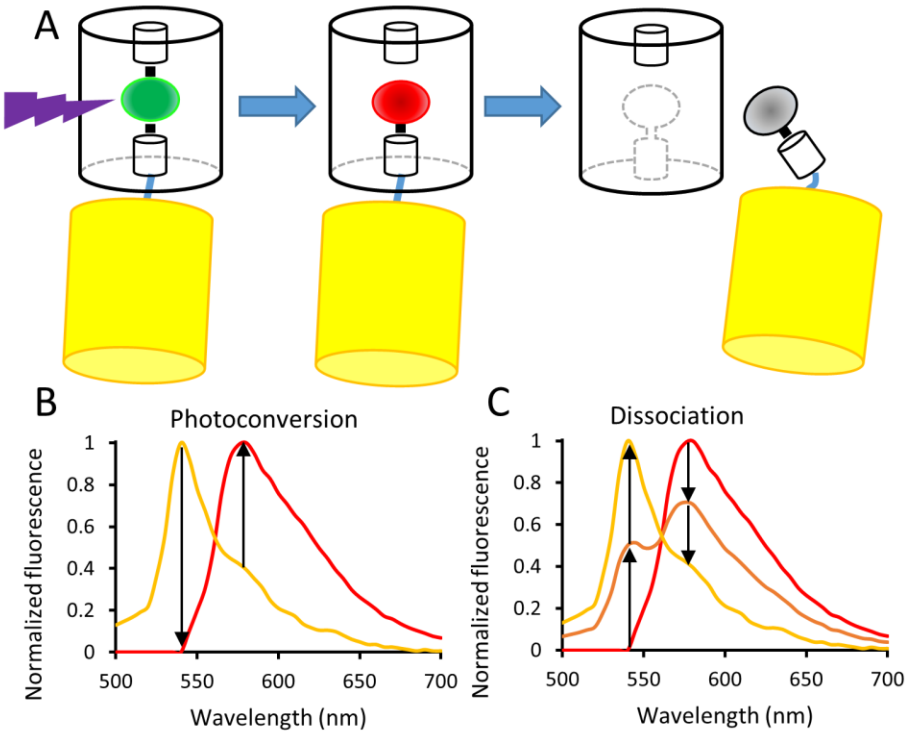

Figure 4.11 PhoCl-mPapaya1 scheme and theory. A) PhoCl-mPapaya1 construct from left to right before conversion, after conversion with yellow to red FRET, and after chromophore release with no yellow to red FRET. B) Theoretical fluorescence emission spectra of PhoCl-mPapaya1 before photoconversion (yellow line) and post photoconversion (red line). C) Theoretical fluorescence emission spectra of PhoCl-mPapaya1 dissociation after 100% photoconversion. Red, orange and yellow lines are 0%, 50%, and 100% dissociation, respectively. Black arrows show the direction of change.

In order to select for improved variants of PhoCl using PhoCl-mPapaya1 the increase in yellow fluorescence was measured. I cannot simply choose the brightest yellow variants after photoconversion because if no photoconversion occurs mPapaya1 fluorescence will remain at full. If only a single measurement of the yellow fluorescence was acquired then variants that disabled the photoconversion of PhoCl would display the highest yellow fluorescence. If yellow fluorescence was measured multiple times, after photoconversion, then the increase in yellow can be used as the selection criteria. This was based on the assumption

that increased yellow fluorescence correlates with the red chromophores increasing exposure to solvent, which should only occur when the PhoCl has cleaved and dissociated

The construct requires a high FRET efficiency as it allows larger changes in post photoconversion fluorescence. High FRET in PhoCl-mPapaya1 is feasible if a very short linker is used, the interchromophore distance can be as short as 4-5 nm given the β-barrel's dimensions, which is much shorter than the Förster radius of 6.5 nm for the mPapaya1-mMaple red FRET pair, this results in a FRET efficiency of 80-95%. This screening system is based on the combination of photoconversion and release by tracking the change in the yellow and red fluorescence when exciting mPapaya1.

As the purpose of PhoCl is photo-induced dissociation the green and red fluorescent properties of PhoCl are not required. If all PhoCl fluorescence was abolished during evolution, while maintaining the dissociation PhoCl have improved utility. This screen is unfortunately limited to variants that maintain a high extinction coefficient for the red fluorescence immediately after photocleavage. To combat this a second time intensive screen was added on top of this one to confirm release independently of the red state.

This second screen uses the liquid subculture of the colonies picked using the FRET prescreen described above. The N-terminal His-tag was used to bind PhoCl-mPapaya1 to Ni-NTA beads, which were then washed to remove any unbound proteins. The beads with PhoCl-mPapaya1 bound were then photoconverted which releases mPapaya1 from the Ni-NTA beads into surrounding buffer solution which was then run on native PAGE. The native PAGE preserves the fluorescence of mPapaya1 and the fluorescence intensity of the mPapaya1 band should correlate to the photo-induced dissociation. Selection of the brightest bands should correlate with improved PhoCl variants and further evolution performed.

The PhoCl-mPapaya1 construct was built, but it failed to express well and no photoconversion was detected. A linker GTGSG between PhoCl and mPapaya1 was added to attempt to remove the potential steric clash between the FPs. This linker comes at the cost of FRET efficiency, but no detectable improvement of protein expression was observed. The sfYFP which previously expressed well with PhoCl was used to replace mPapaya1. Despite the smaller difference in colour with precleaved PhoCl, sfYFP should still have a FRET efficiency between 70-90%. The Förster radius between sfYFP and the converted red chromophore is 6 nm, assuming sfYFPs quantum yield is 0.6. PhoCl-sfYFP was constructed, and it expressed and converted successfully. The resulting change in yellow fluorescence observed was small implying a combination of incomplete photoconversion and dissociation. This implies that PhoCl in this construct and environment is inefficient at photo-induced dissociation. This PhoCl-sfYFP construct provides the potential room for improvement in conversion and release for screening to select.

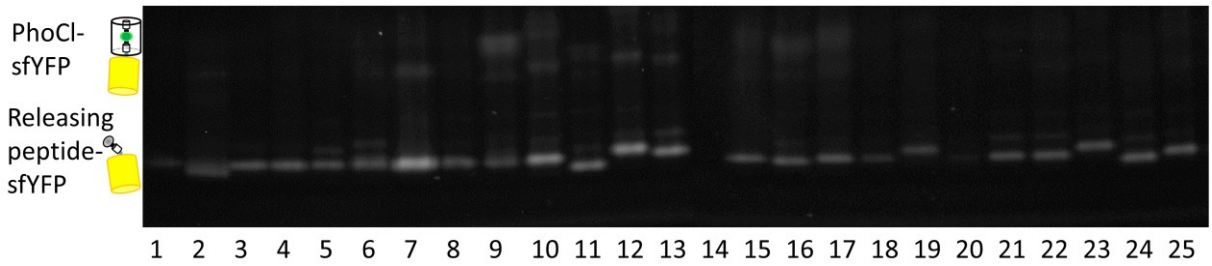

Figure 4.12 Native PAGE gel from EP PCR screening round 2. The native PAGE was imaged using 480-500 nm illumination and a 515-545 nm emission filter. The bright variant in lane 7 contains three mutations: F41I, from round 1; E52V and E95V, from round 2. The faint PhoCl-sfYFP band at the top of the gel contains whole constructs that got stripped off the Ni-NTA beads during photoconversion and collection. The Releasing peptide-sfYFP band derives from successful photocleavage and dissociation of PhoCl. The changes in height in band travel distance is due to changes in the size and charge of the peptides due to mutations.

The PhoCl-sfYFP construct was then used for FRET-based colony prescreening followed by native PAGE screening. Three error-prone rounds were screened and the variants had large differences in the brightness on the native PAGE gel. Unfortunately, all tests against the initial PhoCl showed PhoCl to be superior. With error-prone PCR not giving good results we attempted saturation mutagenesis on histidine 121. The crystal structure of mTFP1 (PDB ID: 2OTB)[174] shows histidine 121 siting adjacent to the chromophore potentially blocking the chromophore from exiting. Randomizing location 121 led to the discovery of the best two variants, which were the two different codons that encode for histidine.

The fact that the FRET screen recapitulated the original histidine variants suggests that the screen does function as intended. However, without any improvements to PhoCl identified screening could be failing two major ways. First, in the case where there were improvements the library, the screening may depend too strongly on conversion efficiency and the red chromophores extinction coefficient rather than dissociation. This is problematic, as variants

need a net improvement in the combination of all three variables. This is supported, as improved red extinction coefficients are often associated with a more rigid chromophore environment, which are likely to hold the red chromophore in place preventing dissociation. Changing the chromophore environment will likely change the photoconversion characteristics, simultaneously with the extinction coefficient. As such, this screen was halted, as it seemed unlikely to find improved variants, after screening four libraries without success. The second way the screening could be failing, is through poor library construction, the error prone and site direction saturation mutagenesis libraries may simply not have contained any improved variants. PhoCl may require mutations in sets of two or more to improve its characteristics, or must acquire one or more negative mutations before starting back towards to an improved PhoCl variant.

4.2.4 Mutational analysis

The mutations discovered during evolution using the PhoCl-sfYFP construct did not result in improvements to PhoCl. These variants demonstrate as much as triple the brightness in the releasing peptide-sfYFP band of the average PhoCl variant (Figure 4.12). Also these locations are tolerant to mutation without complete loss of function (Table 4.1). These mutations may have some as yet undiscovered synergy that may be of value in future effort to improve PhoCl.

Table 4.1 Mutations found in PhoCl evolution.

Mutation	2° structure	Residue direction[a]	Releasing peptide interaction[b]	Interaction change[c]
F41I	β-strand	In	Direct	Remove π-π
M73V	β-strand	In	Direct	Steric
L89P	β-strand	Out	Long range	-[d]
G92S	Loop	Out	Long range	-
F118L	β-strand	Out	Direct	-
I123F	β-strand	In	Direct	Add π-π
N145H	Loop	Out	Direct	-
E152V	Loop	Unclear	Direct	Lose H-bonds
E195V	β-strand	Out	Long range	-

[a] Residue direction is whether the chromophore points in towards the chromophore or out away from the chromophore. [b] Releasing peptide interaction is direct when the sidechain or backbone of residue is in contact with the releasing peptide, it is long range when there is no contact. [c] The interaction changes suggests a putative modification to the interaction with the releasing peptide. [d] The '–' implies that there is no obvious change in the interaction with the releasing peptide.

The three mutations L89P, G92S, and E195V are distant from the releasing peptide and likely only modify the folding and expression of PhoCl-sfYFP. The mutations F118L and N145H are residues that point away from the barrel but their backbone may still interact with the releasing peptide. These two mutation may slightly modify the fluctuations of the barrel itself modifying the release. F41I removes a possible π-π interaction with a phenylalanine on the releasing peptide. I123F does the opposite and introduces a possible π-π interaction with the chromophore, this mutation could be beneficial in stabilizing the empty barrel or affecting photoconversion. M73V is a sterically large change that would cause some reshuffling of the protein interior. E152V removes several possible hydrogen bonding locations, which could encourage release. These mutation locations were mapped onto mTFP1 in Figure 4.13.

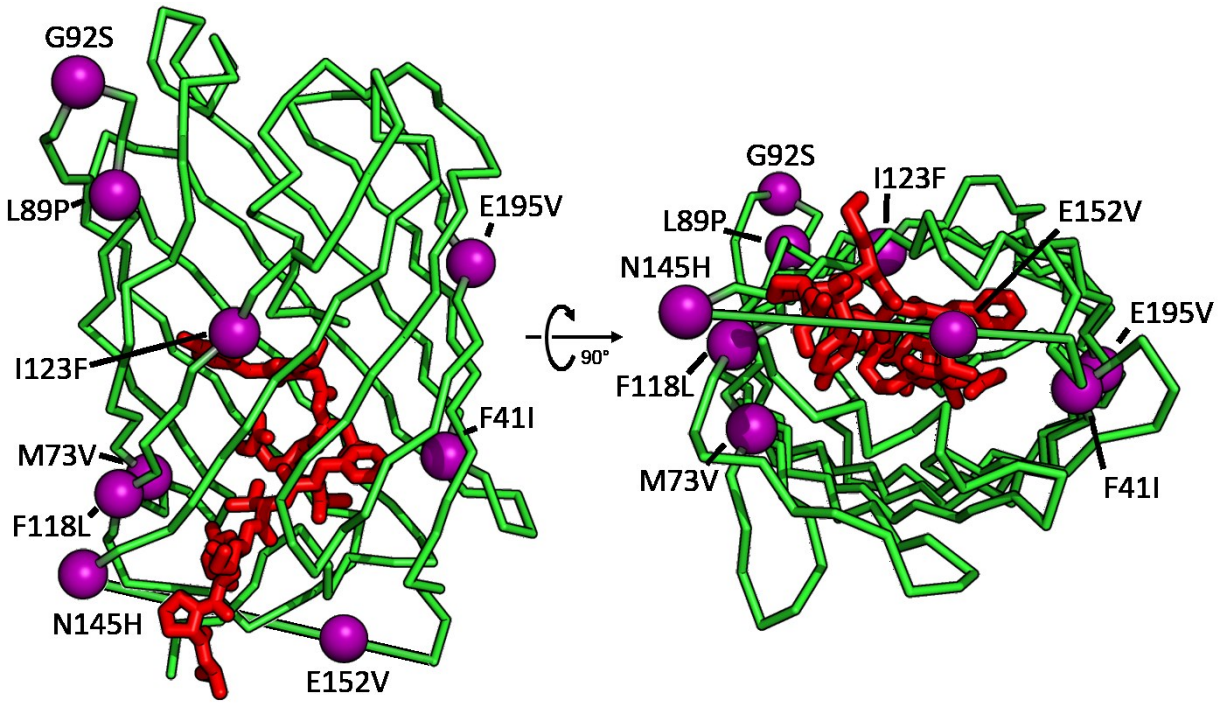

Figure 4.13 Modeled location of mutations in PhoCl. Purple spheres represent the Cα of the mutated residue. The red sticks show the chromophore and the releasing peptide of PhoCl with mTFP1 residues. The structural model is based on mTFP1 (PDB ID 2OTB).[174]

Despite the release process itself being inherently dynamic, all of this analysis is based on the initial position of the chromophore and the releasing peptide from the crystal structure of a distant precursor to PhoCl, mTFP1 (PDB ID 2OTB).[174] As there are significant differences between mTFP1 and PhoCl, the interpretations made must be taken with a grain of salt. On the trajectory that the releasing peptide takes leaving the barrel some intermediate stopping points may exist and be of critical importance to the release, which are currently unidentified. These problems support the continued use of irrational evolution methods such as using error prone PCR libraries as the effect of any given mutation in PhoCl is harder to predict than other static FPs.

4.3 Methods

4.3.1 General molecular engineering

Polymerase chain reactions (PCR) were performed using several different polymerases: Q5, Pfu, and Taq from New England Biolabs. Each polymerase was used according the manufacturer's protocols. EP-PCR were performed using: 5 µL of 10x Taq buffer, 2 µL of an NTP mix containing, dATP (5 mM), dGTP (5 mM), dCTP (25 mM), and dTTP (25 mM), 4 µL $MgSO_4$ (25 mM), 1.5 µL of each DNA primer (10 µM), 1 µL of template (~2 fmol), 1 µL of Taq (New England Biolabs), 1 µL $MnCl_2$ (10 mM), and deionized H_2O up to 50 µL. The $MnCl_2$ was added last. PCR products were purified using 1% agarose gel separation using 1 µL of 10 mg/mL ethidium bromide per 50 mL of gel to mark the bands, followed by gel extraction using Biobasic and Thermo Fisher Scientific gel extraction kits using the recommended protocols.

Gibson assembly was performed using 1 µL of the DNA vector, 1.5 µL of the DNA insert, and 2.5 µL of 2× Gibson Master Mix. The DNA vector used is the plasmid pBAD/His B (Thermo Fisher Scientific) digested with Fast digest enzymes XhoI and HindIII (New England Biolabs), the DNA insert if the PCR or EP-PCR product, both are agarose gel purified and extracted. The Gibson assembly mixture was incubated for 1-4 hours at 50 °C and left at room temperature for up to 48 hours, 5 µL of deionized H_2O was used to dilute the mix then 4 µL was used to transform *E. coli* strain DH10B (Thermo Fisher Scientific) by electroporation.

Transformed *E. coli* were cultured on Petri dishes of agar with Lysogeny Broth (LB) medium and 0.4 mg/ml ampicillin and 0.02% w/v L-arabinose, and grown overnight at 37 °C. Picked colonies were subcultured in liquid culture and were grown in 2-6 mL Lysogeny Broth

(LB) and 0.1 mg/ml ampicillin and 0.02% w/v L-arabinose and incubated in a shaker at 240 rpm and 37 °C overnight.

4.3.2 Protein extraction, and SDS PAGE

Protein was extracted from *E. coli* from the liquid cultures by being spun down at 15 000 RCF for 2 minutes and the supernatant was poured off. Then 100 µL of Bacterial Protein Extraction Reagent (BPER; Thermo Fisher Scientific), was added and the samples were vortexed for 5-60 minutes. The resulting suspension was then spun down at 15 000 RCF for 2 minutes. The supernatant was collected and tested at various dilutions for fluorescence, absorbance, and far red conversion using a fluorescence plate reader (Tecan Safire2). The 2-6 mL liquid cultures described above, and the pellet left after the BPER extraction, were used as starting points for plasmid purification. Miniprep kits from Thermo Fisher Scientific and BioBasic were used according to their respective standard protocols.

4.3.3 Ni-NTA purification

Ni-NTA protein purification was performed using 6 mL liquid cultures grown overnight and BPER extracted as described above. Then 200 µL of Ni-NTA beads were added to the BPER extract and shaken on ice for 1 hour. The beads were then washed using 1×TBS with 20 mM imidazole at pH 8.0 three times, then eluted into 1.5 mL microfuge tubes using 100 µL of 1×TBS with 250 mM imidazole at pH 7.8.

4.3.4 Photophysics

The extinction coefficient of PhoCl was determined by alkali denaturation comparing equal concentrations of PhoCl in 1 M NaOH, and another in 1×Tris buffered saline (TBS).[13] Quantum yield of PhoCl was determined by diluting the 1×TBS + FP solution, from the alkali denaturation test, with 1×TBS at pH 7.4 to create a dilution series with absorbances from 0.001 to 0.002 at 450 nm. The emission spectrum was then measured at 450 nm excitation. All emission wavelengths were summed and plotted and the slope determined. This slope represents to quantum when scaled by a standard, eGFP with a quantum yield of 0.6 was used.

4.3.5 Photoconversion

Photoconversion tests were done in several ways the most common was the use of a custom built photoconversion chamber using six arrays of liquid cooled 405 nm LEDs (0.15 mW/mm^2).[164] This chamber was used to convert Petri dishes and PCR tubes for the varying times, usually of about 5 minutes in length. Photoconversion of liquid samples in PCR tubes was performed with a 405 nm laser (150 mW, 1200 mW/cm^2, Changchun New Industries Optoelectronics Tech. Co., Ltd.). The laser was also used in combination with a Lego Mindstorms EV3 robot, which acted as an optical chopper, blocking the laser 20 times per second with equal periods of illumination and blocking.

4.3.6 PhoCl-FP colony prescreen and native PAGE screening

Libraries of FP-PhoCl variants were cultured on Petri dishes of agar with Lysogeny Broth (LB) medium with 0.4 mg/ml ampicillin and 0.02% w/v L-arabinose, and grown

overnight at 37 °C. The Petri dishes were photoconverted for 5 minutes in the 405 nm LED photoconversion chamber. The Petri dishes were left for one hour and then imaged for yellow and red fluorescence, samples with yellow fluorescence and some red fluorescence were picked into 6 mL liquid subculture.

Ni-NTA purification was performed on the liquid subcultures as described above with the final elution step skipped. The FP-PhoCl variants attached to Ni-NTA beads were resuspended using 1× Tris pH 7 and transferred into transparent 2.0 mL microfuge tubes, and photoconverted using the 405 nm LED photoconversion chamber for 5 minutes. The variants were spun down at 15 000 RCF for 2 minutes, and the supernatant collected, and then diluted with 50 µL of 2× sample loading buffer, 1M Tris pH 7, 20% v/v Glycerol, and 1 µg bromophenol blue. These samples were the run on a native PAGE with a 7-8% separating gel and a 4% stacking gel. The result gels were then imaged for yellow fluorescence and the brightest bands selected for sequencing and further evolution.

4.4 Conclusions

In the process of determining that several screening constructs and protocols were insufficient to find improved variants of PhoCl, some useful information was acquired. PhoCl's dissociation of the releasing peptide does not reach equilibrium for several hours and has two distinct phases of red fluorescence during the release. Caging by distorting a β-strand, as in the construct caged-S11-Y, is more efficient than caging with a partially obscured β-strand, as in S11-plug-Y. Superfolder-based split FPs are more stable than a folded PhoCl protein, and sfGFP either fold first and prevent PhoCl from folding, as in the S11-plug-Y construct, or distort PhoCl enough to render it non-fluorescent, which happened over time in

the caged-S11-Y construct. Finally, when developing a screening system, avoid including variables that can negatively affect the screen that are not related to your end goal. For example the strength of caging or the requirement of a large extinction coefficient in PhoCl's converted chromophore.

Toward SplitOr, an orange PhoCl

5.1 Introduction

Based on the success of the original PhoCl introduced in Chapter 3, a second photocleavable protein, that can be cleaved orthogonally to PhoCl (*i.e.*, by illumination with different wavelengths of light), is under development. As many optogenetic actuators such as PhoCl, Dronpa, and channelrhodopsin[175] are activated with 400 nm illumination, they cannot be used simultaneously.[90] Development of actuators that are activated at wavelengths other than 400 nm would greatly enhance our ability to probe cells optogenetically.

Despite the large variety of FPs discovered, there are is only small class that has a light induced backbone break when illuminated with a wavelength other than 400 nm. PSmOrange[37] and PSmOrange2[38] are both derived from *Discosoma* sp. DsRed.[14] Both start as orange FPs that photoconvert to a far-red state with 480 nm illumination.[176] Compared to PSmOrange, PSmOrange2 has a faster rate of photoconversion, requires lower intensity illumination to induce photocleavage, and has a greater proportion of the protein that converts to far-red.[38]

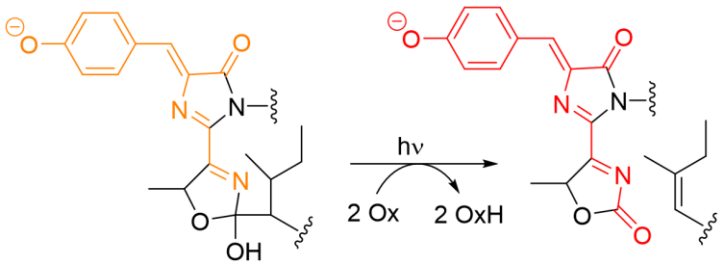

Figure 5.1 Photocleavage of the PSmOrange2 chromophore. Left shows the orange fluorescent state of PSmOrange2. Right is the cleaved far-red fluorescent state of PSmOrange2.

PSmOrange2 is reported to be an orange fluorescent protein with peak excitation at 546 nm, peak emission at 561 nm, a quantum yield of 0.61, and an extinction coefficient of 51 000

$M^{-1}cm^{-1}$. The far-red state has a peak excitation at 619 nm, maximal emission at 651 nm, a quantum yield of 0.38, and an extinction coefficient of 18 900 $M^{-1}cm^{-1}$.[38]

The chromophore of PSmOrange2 contains the standard GFP 4-(p-hydroxybenzylidene)-5-imidazolone heterocyclic core that has been further post-translationally modified to include an N-acylimine extension connecting the conjugated system to a dihydrooxazole ring.[37,39] The far-red state is created by a two-step photo-oxidation in which the hydroxyl group on the dihydrooxazole ring is converted to a carbonyl that further extends the conjugated system. This oxidation causes a main chain break that leaves the chromophore as a new N-terminus for the cleaved protein (Figure 5.1).

To induce a cleaved FP to dissociate into two separate pieces is not a trivial undertaking. The split FPs used for BiFC never dissociate once they have reconstituted, illustrating the high stability of the intact FP structure. However all of the split FPs are split such that the β-barrel itself is disrupted prior to reconstitution.[53] In the case of PhoCl, only the chromophore and half of the central helix must dissociate. This process leaves the β-barrel intact after dissociation. At a molecular level, PhoCl dissociation requires the large chromophore and several amino acids to work their way out of the FP barrel, breaking old intramolecular hydrogen bonds and exposing the hydrophobic chromophore and central helix of PhoCl to the polar solvent. In designing and engineering of PhoCl, the process of circular permutation was used to physically separate the cleaved chromophore and its peptide tail from the rest of the β-barrel. Random mutagenesis was then used to reduce impediments to release.

The following chapter describes my efforts to engineer PSmOrange2 to be a second photocleavable protein that is spectrally orthogonal to PhoCl.

5.2 Results and discussion

5.2.1 PSmOrange2 photophysics

To track the evolution of PSmOrange2 photocleaving variants, a set of baseline spectra for PSmOrange2 were measured (Figure 5.2). PSmOrange2 exhibits a substantial amount of green fluorescence, with maximal excitation and emission at 490 nm and 510 nm, respectively. This is likely due to either an equilibrium between the protonated and deprotonated states of the chromophore, or incomplete/branched maturation as in DsRed (Figure 1.7).[31]

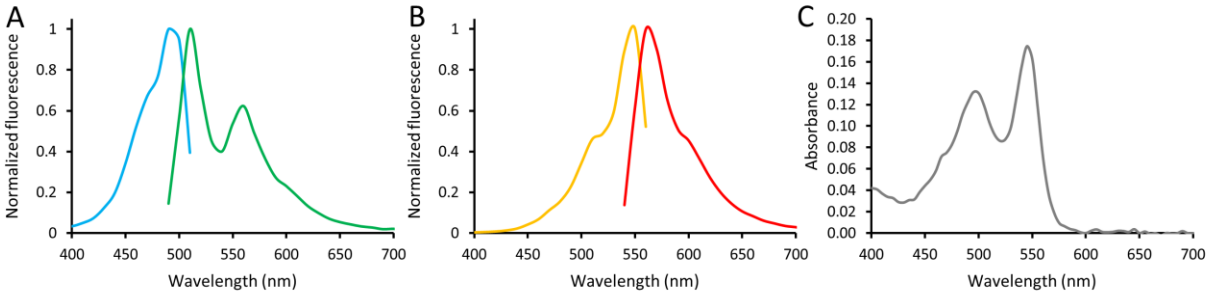

Figure 5.2 PSmOrange2 spectra. A) Spectrum of the green fluorescence of PSmOrange2. The blue line is the excitation spectrum (emission at 530 nm). The green line is the emission spectrum (excitation at 470 nm). B) Spectrum of the orange fluorescence of PSmOrange2. The yellow line is the excitation spectra (emission at 580 nm). The red line is emission spectrum (excitation at 520 nm). C) Absorbance spectrum of PSmOrange2. All spectra were measured on the same sample. A) and B) were normalized to their respective maximal excitation or emission wavelength.

5.2.2 Circular permutation of PSmOrange2

The loop regions at both ends of the central helix of PSmOrange2 were the targets of circular permutation which, based on the precedent of PhoCl, could enable dissociation after photocleavage. Four adjacent locations on each side of the central helix were chosen and the 6-residue linker GGSGGT was used to bridge the gap between the original N- and C- termini. A total of eight circular permutation break points were created: four break points between

residues 83 and 87 (analogous to the permutation site of PhoCl), and another four break points between residues 57 and 61 on the opposite side of the central helix. These eight different circular permutation sites were tested by making one library for each side of the central helix. Each library consisted of 16 members, containing all possible combinations of four possible N-termini and the four possible C-termini. Screening of these two libraries led to the identification of two viable circular permutation combinations from each the two libraries. Specifically, I discovered fluorescent variants with N- and C- termini at 84/84, 84/85, 60/59, and 58/58 (C-terminal position/N-terminal position) (Figure 5.3).

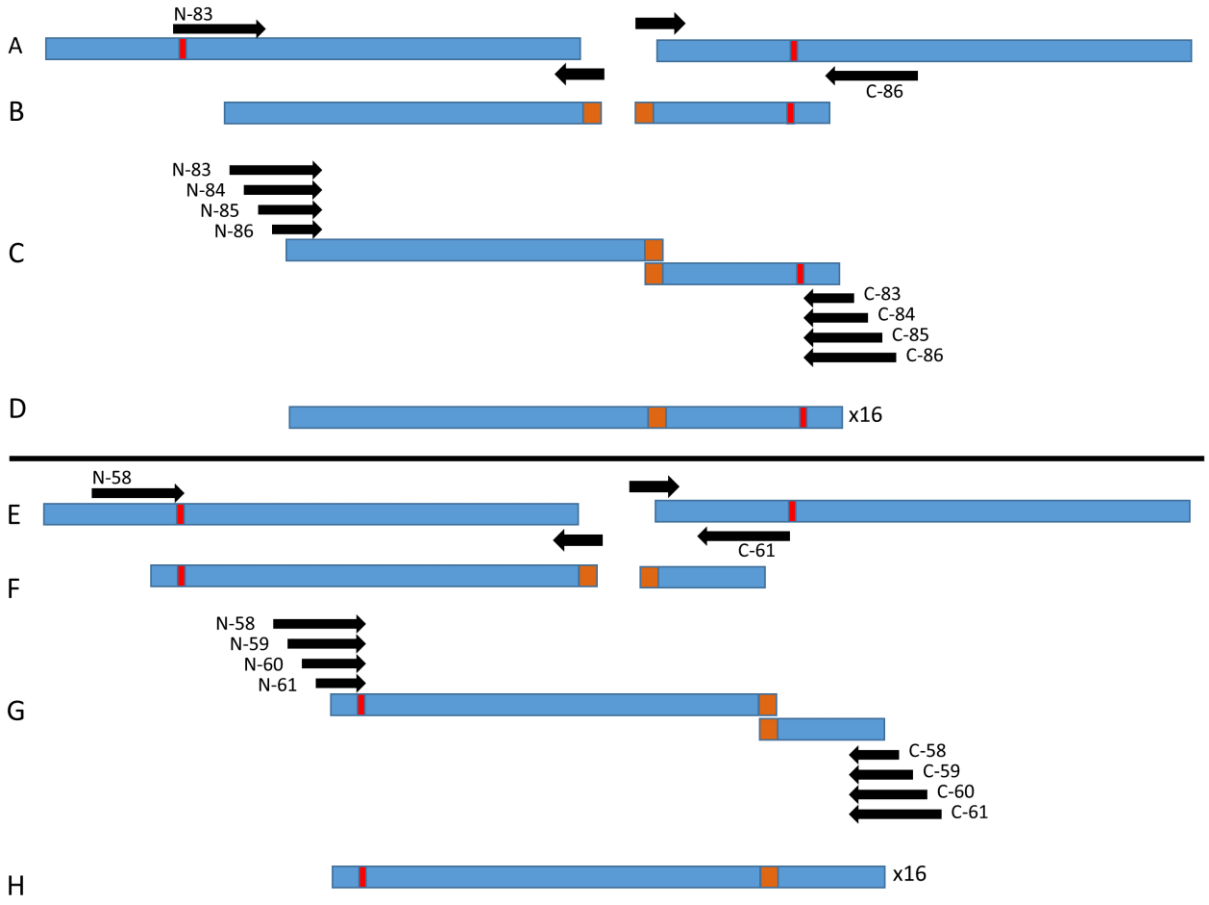

Figure 5.3 Circular permutations of PSmOrange2. Blue is the PSmOrange2 gene, red is the chromophore, and orange is the linker. Black arrows represent the primers and their extension direction. The labels identify which N- or C- terminus location each primer codes for. A-D) The production of the circular permutation library analogous to PhoCl. A) PCR creating the N- and C- terminal fragments and adding the linker. B) Products of the PCR in A. C) Overlap-PCR using four forward primers and four reverse primers. D) Products of the overlap PCR in C this contains 16 different circular permutation variants from the 16 possible pairs of forward and reverse primers. E-H) The production of the circular permutation library on the opposite side of the central helix. E) PCR creating the N- and C- terminal fragments and adding the linker. F) Products of the PCR in E. G) Overlap-PCR using four forward primers and four reverse primers. H) Products of the overlap PCR in G this contains 16 different circular permutation variants from the 16 possible pairs of forward and reverse primers.

As all of these variants were very dim, I performed one round of error prone PCR in an attempt to rescue their fluorescence. Only the evolution of the 58/58 circular permutation led to a substantially improved variants with 3 mutations (D78V, K92T, and Q213L, using PSmOrange2 numbering). Narrowing the pool down to the two most promising circular permutation variants in terms of expression and brightness, the original 84/84 circular permutation and the improved 58/58 circular permutation variant emerged as the two most promising scaffolds for further evolution.

I speculated that the negative charge associated with the new C-terminal position could be disruptive to protein folding and function. The positive charge of the N-terminal position is already distant as all variants are evolved with an N-terminal His-tag, however the amino acid sequence of the His-tag is likely still disruptive. To attempt to rescue the fluorescence of the 84/84 and 58/58 circular permutations, I inserted a randomized amino acid at the N- and C-termini, respectively. No improved variants of the 84/84 circular permutation were discovered in the resulting library. Upon screening this 58/58 circular permutation insertion library, a brighter variant was identified that contained an inserted isoleucine at the N-terminus, the C-terminus inserted the DNA bases GTG but also acquired an unplanned single nucleotide deletion in residue 54, upstream of the inserted bases, this introduced a frameshift mutation that converted the C-terminus from the original …TKGGP* into …TRVAPCKLGCFGG*.

Although the 84/84 circular permutation was fluorescent, it was not pursued any further as no improvements were found after the rounds of error prone mutation and end insertions. Accordingly, the 58/58 circular permutation frameshift variant became the focus for our future efforts. From here on, variants have been renumbered such that the inserted isoleucine at the N-terminus is now location 1. Thus the three mutations already incorporated became D27V,

K41T, and Q162L. This 58/58 circular permutation variant was named SplitOr 0, as the location of the backbone break happens in the residue word SPLIT between the I and the T, and the Or is from the orange fluorescence.

5.2.3 SplitOr evolution

SplitOr 0 was only dimly fluorescent (despite the 3 mutations that had improved the brightness) but the primary goal of this project was to engineer photocleavage and dissociation, as opposed to bright fluorescence. To this end, the ability to break the backbone is more important than fluorescence prior to the photocleavage. Unlike PhoCl, the chromophore of SplitOr would not leave the barrel for dissociation to occur, and the far-red fluorescence created by photocleavage could potentially remain (Figure 5.4). This prevents the possibility of tracking the far-red fluorescence decay as a proxy for dissociation, as was done with red fluorescence in PhoCl screening.

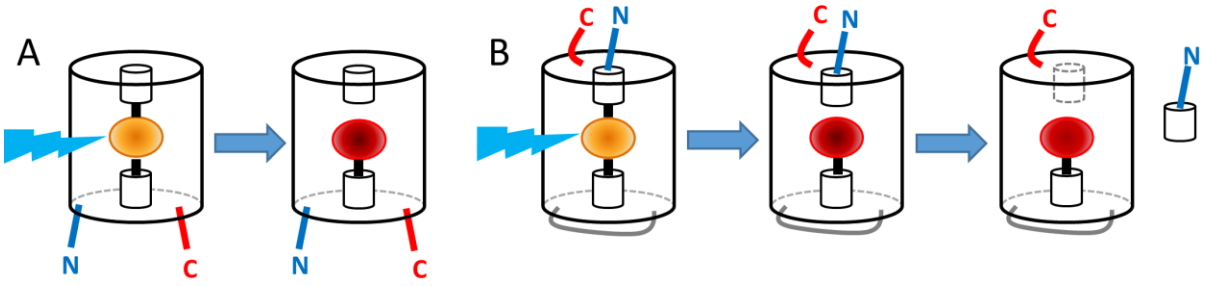

Figure 5.4 The modification of PSmOrange2 into SplitOr. A) PSmOrange2. B) SplitOr. Blue lightning is 488 nm light inducing photoconversion of the chromophore. The grey line is the linker which was added to circularly permutated PSmOrange2.

I decided to evolve SplitOr primarily for far-red fluorescence after photocleavage. The directed evolution workflow involved first screening of error prone PCR libraries for fluorescence brightness of the orange state. The selected variants were then photoconverted to

the far-red fluorescent state using white light illumination in the tube bleach configuration, using the robot described in Chapter 3. In rounds 1 and 2, I found that mutations V27D, a reversion to the PSmOrange2 residue, and Y163C, both improved the brightness from SplitOr 0. However, I was unable to observe any far-red fluorescence following photoconversion. SplitOr 0.3 was came from round 3 with in the discovery of the E221V mutation that successfully rescued the photoconversion to far-red fluorescence. However, the emission maximum of the resulting far-red state had blue shifted about 20 nm to ~630 nm from PSmOrange2's 651 nm (Figure 5.5).

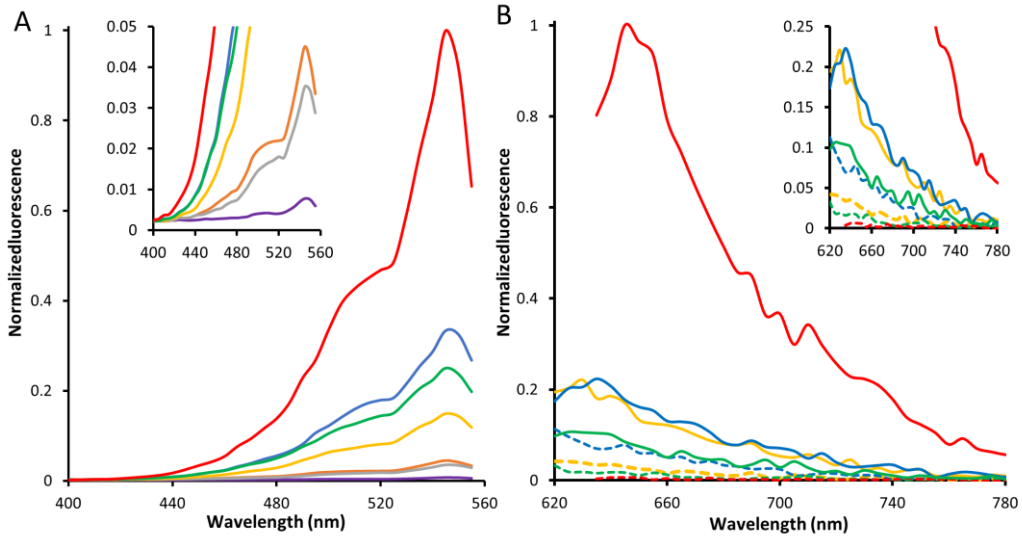

Figure 5.5 Orange excitation and far-red emission spectra. A) Excitation spectra of the PSmOrange2 variants at 570 nm emission. The red line is PSmOrange2, the purple line is 58/58 circular permutation, the orange line is SplitOr 0, the grey line is SplitOr 0.4, the yellow line is SplitOr 1.1, the green line is SplitOr 1.2, and the blue line is SplitOr 1.3. The fluorescence spectra were normalized two ways. First all spectra were divided by their individual autofluorescence at 400 nm, then divided by the maximum fluorescence of the brightest variant. Inset is zoomed in on the dimmer variants. B) The far-red emission spectra of PSmOrange2 variants. The red line is PSmOrange2 far-red emission at 620 nm excitation. The yellow, green and blue lines are SplitOr 1.1, SplitOr 1.2, and SplitOr 1.3 respectively and show the far-red emission spectra from 610 nm excitation. The dashed lines represent the variants before photoconversion and the solid lines are after 5 min of photoconversion using the white light tube bleach configuration. Normalization of the emission spectra occurred by dividing all variants by the maximum fluorescence of the brightest variant, PSmOrange2. Inset is zoomed in on the SplitOr variants.

Rounds four to seven of screening were focused on improving photoconversion by selection for the brightest far-red variants, in addition to selection for brighter orange fluorescence. In these rounds, the colonies picked for brighter orange fluorescence proved to also have the brightest far-red fluorescence. As these bright orange colonies were removed from the plate by picking before the photoconversion selection they could not be picked in the

far-red fluorescence screen. Suggesting that the brightness of the orange and far red fluorescent states are tightly linked. Following the completion of all the seventh round of screening, three different versions of SplitOr were selected for detailed characterization. SplitOr 1.1 was the best variant from round six, with the highest far-red fluorescence and acquired two additional mutations E109K acquired in round 4, with SplitOr 0.4 and Y179D is new from round 6. SplitOr 1.2 had the highest orange fluorescence and two additional mutations, I14F and G182V, relative to SplitOr 1.1. SplitOr 1.3 had the best ratio between orange fluorescence lost and red fluorescence gained and three additional mutations relative to 1.1 (and different from 1.2), N47Y, T55A, and E66V (Figure 5.6 and Table 5.1).

Table 5.1 SplitOr variant mutations

FP variant	Cumulative mutations												
58/58 CP mutant			D27V	K41T				Q162L					
SplitOr 0	^1I[a]		D27V	K41T				Q162L					240TRVAPCKLGCFGG*
SplitOr 0.1	^1I		D27V	K41T				Q162L	Y163C				240TRVAPCKLGCFGG*
SplitOr 0.2	^1I		-	K41T				Q162L	-				240TRVAPCKLGCFGG*
SplitOr 0.3	^1I		-	K41T				Q162L	Y163C			E221V	240TRVAPCKLGCFGG*
SplitOr 0.4	^1I		-	K41T			E109K	Q162L	Y163C			E221V	240TRVAPCKLGCFGG*
SplitOr 1.1	^1I		-	K41T			E109K	Q162L	Y163C	Y179D		E221V	240TRVAPCKLGCFGG*
SplitOr 1.2	^1I	I14F	-	K41T			E109K	Q162L	Y163C	Y179D	G182V	E221V	240TRVAPCKLGCFGG*
SplitOr 1.3	^1I	-[b]	-	K41T	N47Y	E66V	E109K	Q162L	Y163C	Y179D	-	E221V	240TRVAPCKLGCFGG*

[a] ^ indicates an insertion. [b] - indicates the parental amino was reincorporated.

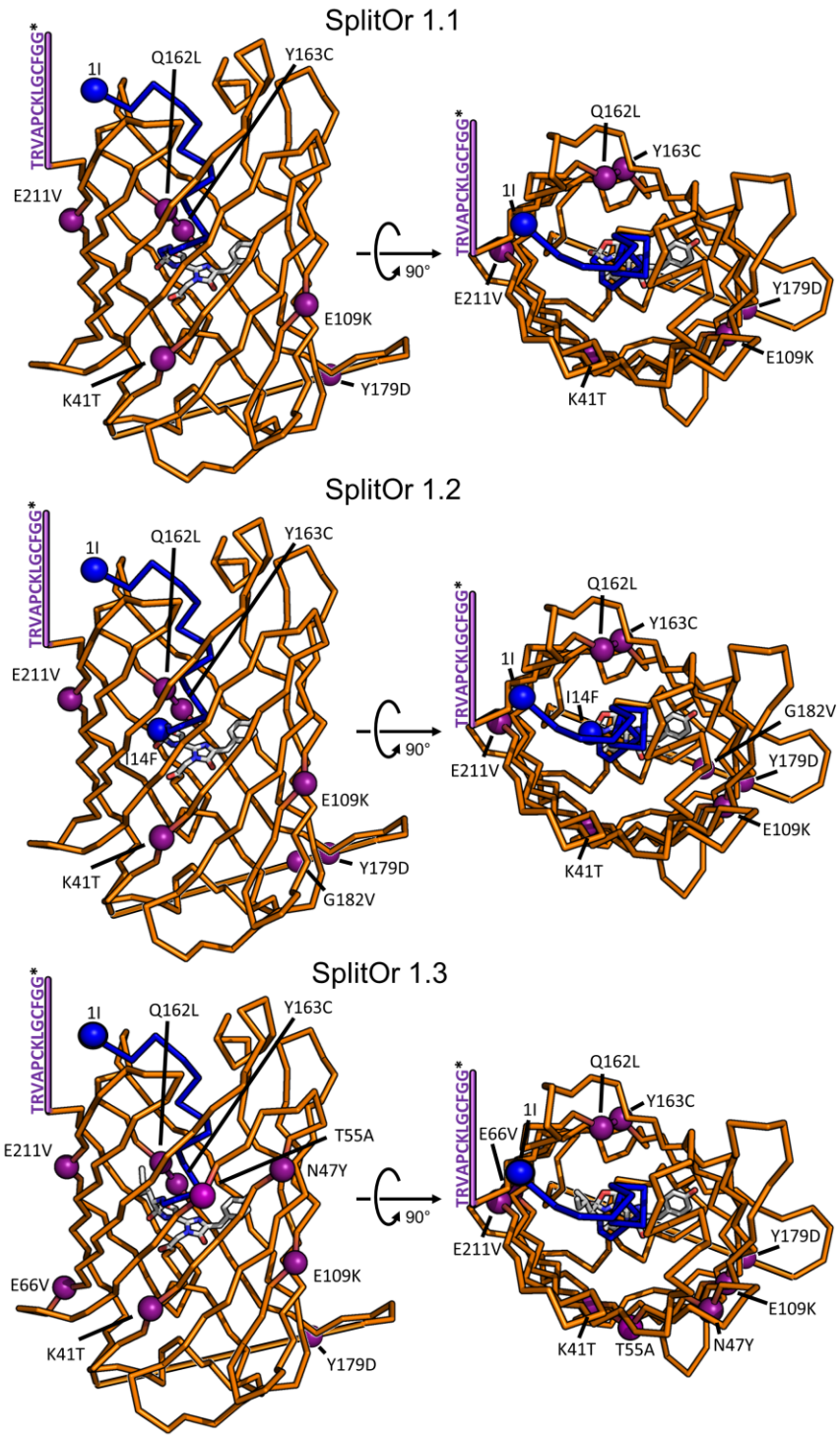

Figure 5.6 Modelled location of mutations in SplitOr variants. Purple and blue spheres represent the Cα of the mutated residues, with the blue additionally representing the peptide that would be released after photocleavage. The structural model is based on PSmOrange2 (PDB ID 4Q7U).[39]

5.2.4 Mutational analysis

SplitOr 1.1 is the base from which SplitOr 1.2 and 1.3 were derived. SplitOr 1.1 contains a single inward facing mutation on the β-barrel, Q162L (Figure 5.6). This mutation helped recover some of orange fluorescence lost in the 58/58 circular permutation and it points towards the chromophore on the side of the releasing peptide. The addition of leucine prevents the possibility of several hydrogen bonds forming due to the loss of the glutamine amide. Isoleucine is a smaller and more hydrophobic side chain than glutamine, which may cause some repacking of the interior structure. The mutations K41T, E109K, Y163C, and E221V are all outward pointing residues on the β-barrel and are potentially involved in folding or more subtle changes to the protein interior. Y179D is a mutation just before the circular permutation linker in the loop created by the circular permutation, and likely helps accommodate the conformational distortion resulting from the introduction of the circular permutation. SplitOr 1.2 has two additional mutations. The first of these is the inward facing I14F, which is a reversion to PSmOrange. This mutation seemed to improve photoconversion in PSmOrange2 at the cost of orange fluorescence brightness. The second SplitOr 1.2 mutation, G182V, is in the circular permutation linker. SplitOr 1.3 contains three mutations, N47Y, T55A, and E66V, relative to SplitOr 1.1. These mutations are all outward pointing residues and seem to influence conversion efficiency through subtle conformational changes in the protein. A sequence alignment of these variants is shown in Figure 5.7.

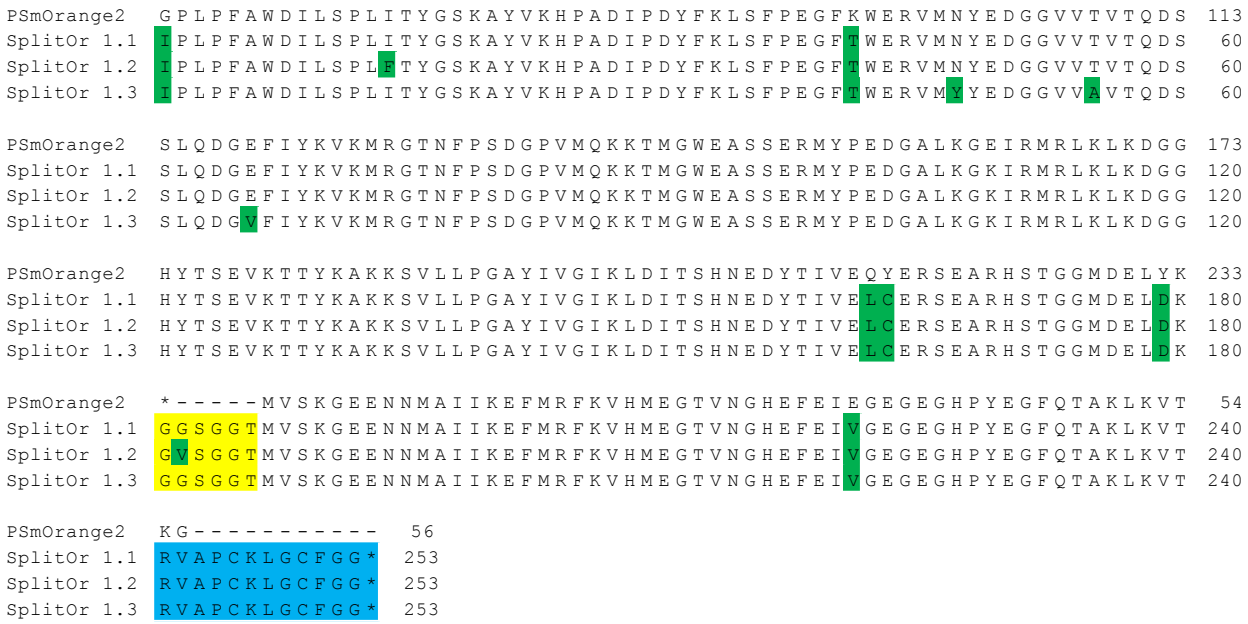

Figure 5.7 Sequence alignment of SplitOr variants. Green highlights mutations found through screening, blue highlights the frameshift created C-terminus, and yellow highlights the inserted linker.

5.2.5 Photocleavage

I speculated that the observed 20 nm blue shift in the far-red fluorescence after photoconversion could potentially be the result of the change in chromophore environment that occurred when the peptide fragment dissociated from photoconverted SplitOr. When the releasing peptide dissociates, the chromophore environment must change, potentially causing the observed blue shift. However, it is also possible that the circular permutation and the subsequent mutations required to rescue the far-red fluorescence led to the shift in far-red fluorescence and dissociation is not occurring. At an intermediate stage of the directed evolution, I tested SplitOr for dissociation and cleavage by fusing the maltose binding protein (MBP) to the N-terminus (the end that is released) of SplitOr 0.4 (Figure 5.8).

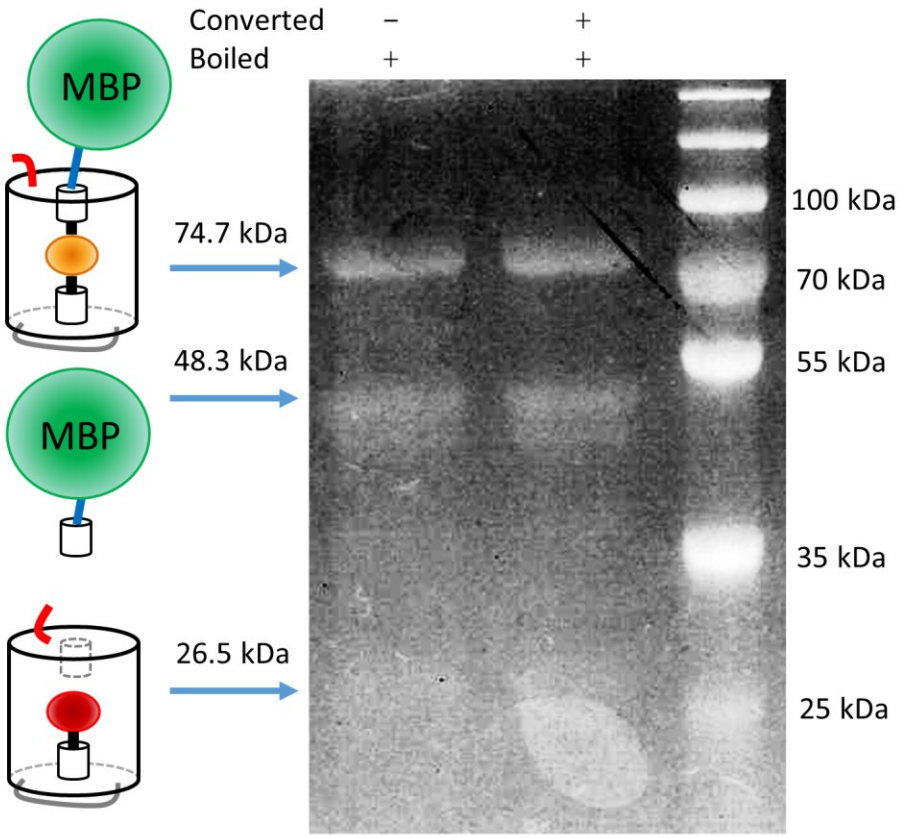

Figure 5.8 Photocleavage of SplitOr 0.4. The leftmost columns contain SplitOr 0.4 with MBP fused to the N-terminus. The middle lane sample was converted using the white light tube bleach configuration for 5 minutes. The left and middle lane samples were boiled in SDS for 10 minutes. The full MBP-SplitOr has an expected molecular weight of 74.7 kDa, after cleavage the MBP-releasing peptide and the empty SplitOr barrel are 48.3 and 26.5 kDa respectively.

The SDS-PAGE of MBP-SplitOr 0.4 with and without photoconversion shows that SplitOr cleaves at the expected location. However, it appears that illumination is not the primary factor leading to dissociation, as both illuminated and non-illuminated samples showed the same banding pattern. Thermally induced cleavage may explain this observation, though the precursors to SplitOr, PSmOrange and mOrange, have not been reported to exhibit the thermally induced backbone cleavage seen in other DsRed-based red FPs.[13,37] The

mutations required to convert PSmOrange into PSmOrange2 may have reintroduced thermally induced cleavage into SplitOr. Further experiments with the later variants were aimed at determining whether the SplitOr protein cleaves only when illuminated and, if so, if the peptide fragment dissociates from the remainder of the barrel.

To determine if the photocleavage of SplitOr is spectrally orthogonal to the photocleavage of PhoCl, photoconversion tests with 405 nm illumination were performed. Conversion was tested using the 405 nm conversion chamber from Chapter 4 and using an optically chopped 150 mW 405 nm laser. Both conversion modes successfully converted PhoCl, but neither illumination mode resulted in far-red fluorescence in SplitOr. Two other photoconversion modes proved to be ineffective at producing far-red fluorescence: the use of a 510-560 nm excitation filter on the 300 W xenon arc lamp, and optically chopped 532 nm laser light (450 mW). Only two illumination modes resulted in far-red fluorescence: a 450-490 nm excitation filter on the 300 W xenon arc lamp very slowly created far-red fluorescence, and an unfiltered 300 W xenon arc lamp white light illumination resulted in a relatively high level of photoconversion (Figure 5.9). A possible explanation for the increased photoconversion using white light is that SplitOr photoconversion is optimal when absorbing two or more different wavelengths in rapid succession, such as in primed conversion.[156] Additionally, the white light can result in significant heating such that it may thermally enhance photoconversion or induce thermal cleavage.

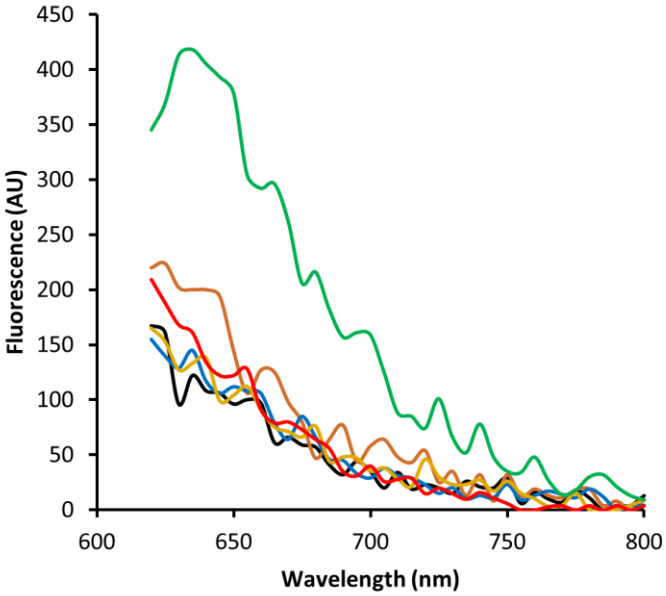

Figure 5.9 Photoconversion of SplitOr 1.2 using various illumination modes. Green, 10 minutes of 300 W xenon arc lamp white light illumination. Orange, 20 minutes of 300 W xenon arc lamp illumination filter to 510-560 nm. Yellow, 450 mW 532 nm optically chopped laser illumination. Blue, 5 minutes of 405 nm illumination in the LED conversion chamber. Red heating to 72 °C for 10 minutes. Black, no illumination.

The thermally enhanced photoconversion hypothesis was tested by using an extended axle to submerge the rotating arm of the tube bleaching system in water cooled to 4°C. The water was contained in a transparent glass cylinder 10 cm in diameter, and the xenon arc lamp illuminated the samples through the water and glass. Two aliquots of each variant, SplitOr 1.1, 1.2, and 1.3, were tested at room temperature in the liquid cooled 4 °C chamber. None of the variants showed a significant difference in far-red fluorescence in response to the temperature difference (Figure 5.10). A second test was done by heating SplitOr to 72 °C for 10 minutes with no illumination, which resulted in no far-red fluorescence (Figure 5.9). Together these results imply that thermally enhanced photoconversion is not occurring or is negligible.

However, direct thermal cleavage without forming a far-red fluorescent species may be occurring as observed in the SDS-PAGE.

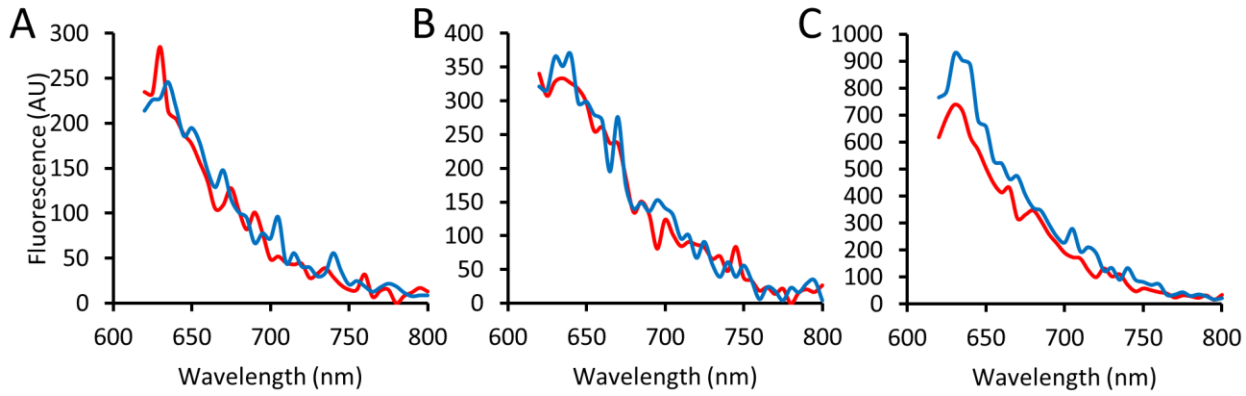

Figure 5.10 Thermal photoconversion of SplitOr variants. A) SplitOr 1.1, B) SplitOr 1.2, C) SplitOr 1.3. Red lines were photoconverted at room temperature and blue lines at 4 °C water cooled. All photoconversion is done with a 300W xenon arc lamp with no filters.

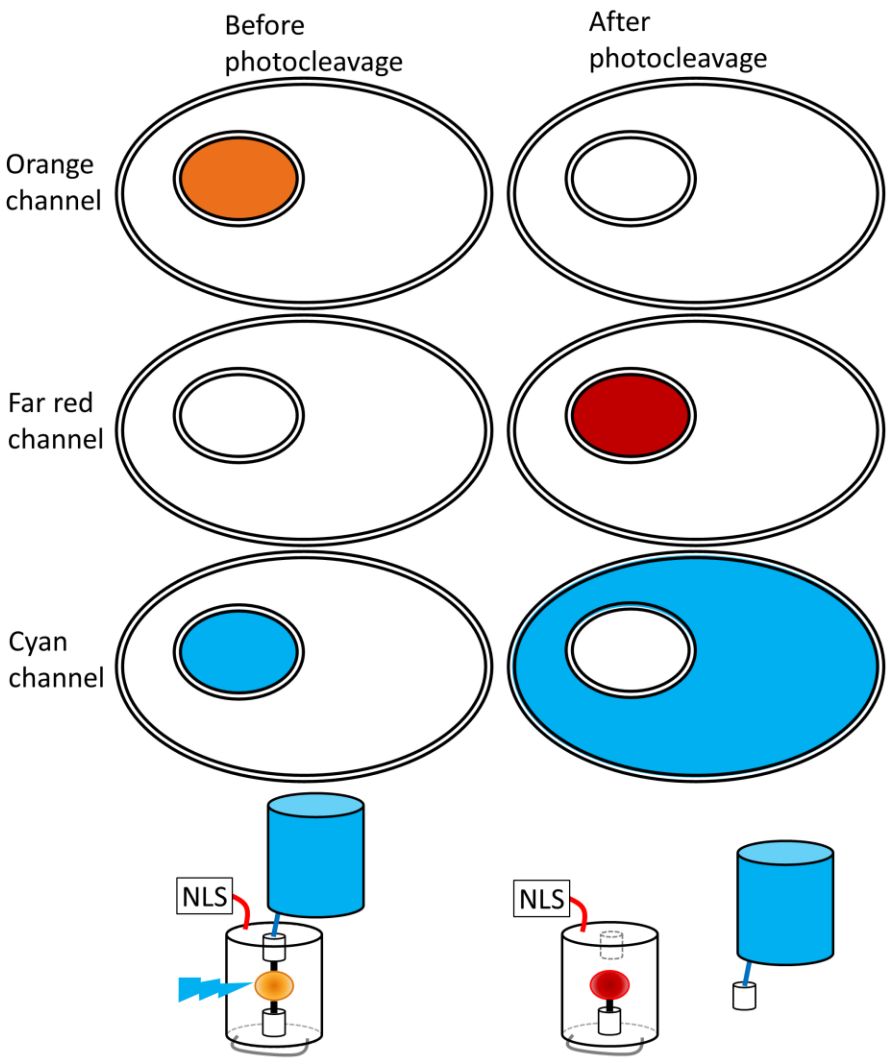

Figure 5.11 Scheme for SplitOr photocleavage confirmation *in cellulo*.

For future work, the next step is to try the SplitOr variants in constructs that were successful used to demonstrate PhoCl, such as fusing a nuclear localizing sequence (NLS) to the SplitOr barrel and a cyan FP to the releasing strand (Figure 5.11). This SplitOr-cyan FP construct would need to be imaged in cells using three different fluorescence channels. Before photocleavage, the NLS will pull the construct into the nucleus, causing the orange and cyan fluorescence to be localized there. After photocleavage, the orange chromophore will be largely converted to far-red. When dissociation occurs, the far-red chromophore stays attached

to the NLS, keeping it in the nucleus, while the releasing strand and the Cyan FP will leave the nucleus and spread into the cytoplasm. This test plays two roles: it could confirm that SplitOr does in fact dissociate after photocleavage, and it could confirm that the trends observed in *E. coli* and *in vitro* hold true in mammalian cells. Together these two traits would make SplitOr into a powerful optogenetic tool.

5.3 Methods

5.3.1 General molecular engineering

PCR were performed using several different polymerases: Q5, Pfu, and Taq (New England Biolabs), each used according the manufacturer's protocols. EP-PCR was performed using Taq polymerase (New England Biolabs). 50 µL reactions were created using: 5 µL of 10×Taq buffer; 2 µL of an NTP mix containing, dATP (5 mM), dGTP (5 mM), dCTP (25 mM), and dTTP (25 mM); 4 µL $MgSO_4$ (25 mM), 1.5 µL of each DNA primer (10 µM), 1 µL of template (~2 fmol), 1 µL of Taq, 1 µL $MnCl_2$ (10 mM), and deionized H_2O up to 50 µL. The $MnCl_2$ was added last. PCR products were purified using 1% agarose gel separation using 1 µL of 10 mg/mL ethidium bromide per 50 mL of gel to mark the bands, followed by gel extraction using gel extraction kits (Biobasic and Thermo Fisher Scientific), using the recommended protocols.

Gibson assembly was performed using 1 µL of the DNA vector, 1.5 µL of the DNA insert, and 2.5 µL of 2×Gibson Master Mix (New England Biolabs). The DNA vector used was the plasmid pBAD/His B (Thermo Fisher Scientific) digested with Fast Digest enzymes XhoI and HindIII (New England Biolabs), the DNA insert was the PCR or EP-PCR product, and both were agarose gel purified and extracted. The Gibson assembly mixture was incubated

for 1-4 hours at 50°C and left at room temperature for up to 48 hours, 5 µL of deionized H_2O was used to dilute the mix then 4 µL was used to transform *E. coli* strain DH10B (Thermo Fisher Scientific) by electroporation.

Transformed *E. coli* were cultured on Petri dishes of agar with LB medium, 0.4 mg/ml ampicillin, and 0.02% w/v L-arabinose, and grown overnight at 37°C. Picked colonies were subcultured in liquid culture and were grown in 2-6 mL LB, 0.1 mg/ml ampicillin, and 0.02% w/v L-arabinose and incubated in a shaker at 240 rpm at 37°C overnight.

5.3.2 Protein extraction, purification, and SDS PAGE

Protein was extracted from *E. coli* in liquid cultures by spinning the cultures down at 15 000 RCF for 2 min and pouring the supernatant off. Then 100 µL of BPER; Thermo Fisher Scientific), was added and the samples were vortexed for 5-60 min. The resulting suspension was then spun down at 15 000 RCF for 2 min. If the pellet retained the orange-red colour of the FP, then the sample was vortexed to resuspend the cells, underwent a freeze-thaw cycle, and spun down again at 15 000 RCF for 2 min. The final supernatant, in either case, was collected and tested at various dilutions for fluorescence, absorbance, and far-red conversion using in a fluorescence plate reader (Tecan Safire2).The 2-6 mL liquid cultures described above and the pellet left after the BPER extraction were used as starting points for plasmid purification. Miniprep kits (Thermo Fisher Scientific and BioBasic) were used according to their respective standard protocols.

Ni-NTA protein purification was performed prior to the SDS-PAGE. Two 6 mL liquid cultures of MBP-PhoCl 0.4 were grown overnight and BPER extracted as described above. Then, 200 µL of Ni-NTA beads were added to the BPER extract and the extracts were shaken

on ice for 1 hour. The beads were then washed using 1×TBS with 20 mM imidazole at pH 8.0 three times, then eluted into 1.5 mL microfuge tubes using 100 µL of 1×TBS with 250 mM imidazole at pH 7.8.

Two 50 mL aliquots of Ni-NTA purified protein were prepared. One was photoconverted using the white light tube bleach configuration for 10 min. Both samples were then diluted with 50 µL of 2×sample loading buffer, 1M Tris pH 7, 5% w/v SDS, 20% v/v Glycerol, and 1 µg bromophenol blue. These samples were then heated in a PCR machine to 98°C for 10 min. These samples were then run on an SDS-PAGE with a 12% separating gel and a 10% stacking gel. The ladder used was PageRuler™ Prestained Protein Ladder, 10 to 180 kDa (Thermo Fisher Scientific).

5.3.3 Circular permutation library creation

The circular permutation libraries were built in stages. First, four forward and reverse circular permutation primers were designed for each of the four possible circular permutation sites with Gibson assembly overhangs for pBAD included. Then the forward and reverse linker primers were designed to allow a future overlap PCR between them. Two PCRs were performed, one contained the maximum length forward circular permutation primer and the reverse linker primer, and the other PCR contained the maximum length reverse circular permutation primer and the forward linker primer, such that the circular permutation region was duplicated. These PCRs were run on an agarose gel and the appropriate sized bands were extracted. The DNA from those two bands were combined and run using overlap PCR containing the four forward and four reverse circular permutation primers. This was again run

on a 1% agarose gel with the appropriate band extracted and fused into plasmids using Gibson assembly. This was repeated for the second circular permutation site.

5.3.4 Screening protocol

On-plate colony screening was performed using illumination from a 300 W xenon arc lamp with a 510-560 nm excitation filter for orange and 660-700 nm excitation filter for far-red. Selection was done using a combination of visual inspection with long pass filter goggles and digital fluorescence imaging using a 600-660 nm emission filter for orange and 660-700 nm emission filter for far-red. Colonies with the highest orange brightness were picked and then the plates were photoconverted using the plate bleach configuration or the robot used in Chapter 3, using the white light illumination and a base time setting of 10 s with a 3 mm step size, after which the brightest far-red fluorescent colonies were picked. The picked colonies were grown in liquid culture as described above to acquire the BPER extract. These extracts were then aliquoted such that each sample was tested fresh and photoconverted using the white light tube bleach configuration for 10 min as described in Chapter 3.

5.3.5 Conversion testing

Photoconversion was attempted using the various illumination conditions listed below. The 405 nm photoconversion chamber described in Chapter 4. A 405 nm laser, (150 mW, 1200 mW/cm^2, Changchun New Industries Optoelectronics Tech. Co., Ltd.), and a 532 nm laser (450 mW, 3600 mW/cm^2, Changchun New Industries Optoelectronics Tech. Co., Ltd.) were used as the light source for the tube bleach configuration. The 300W xenon arc lamp was used with the tube bleach configuration with no filters for white light or with 450-490 nm or 510-

560 nm filters. For the hot thermal conversion tests, a sample was placed in a PCR machine at 72°C for 10 min. For the cold conversion tests, a custom glass container was constructed to submerge protein samples into a controlled temperature water bath while allowing simultaneous cooling and illumination (Figure 5.12). This apparatus is used as an accessory of the tube bleach configuration described in Chapter 3.

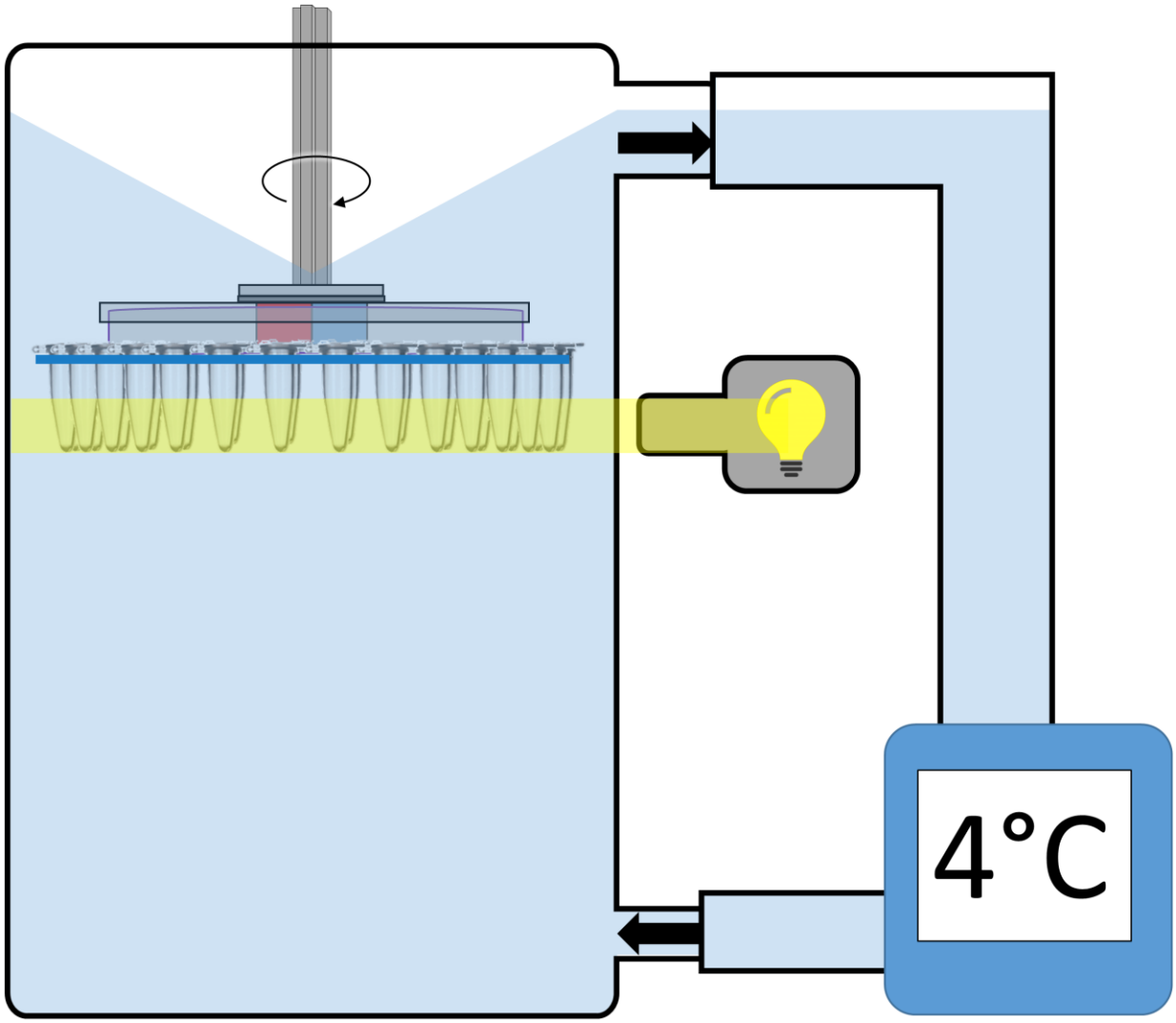

Figure 5.12 Design for water bath for temperature controlled illumination using the tube bleach configuration.

5.4 Conclusions

The three SplitOr variants 1.1, 1.2, and 1.3 are all functioning circular permutation variants of PSmOrange2. As none of the three candidate SplitOr variants photoconvert under 405 nm illumination, they appear to be spectrally orthogonal to PhoCl cleavage. Also, the 20 nm blue shift in far-red fluorescence that arose during the circular permutation evolution supports the hypothesis that the cleaved strand leaves the SplitOr barrel. However, it is also possible that this change is simply due to the mutations required to rescue SplitOr and no dissociation is occurring. As such, these variants will next be tested by collaborators for photocleavage in mammalian cells. While its future has yet to be determined, it looks hopeful that SplitOr will expand the toolbox of optogenetic actuators by enabling spectrally distinct, longer wavelength, photocleavage systems.

Conclusions

6.1 Summary of the thesis

FPs have revolutionized our ability to image and understand the inner functioning of cells and tissues. Despite the hundreds of FPs that have already been found or engineered, there are always new uses that require a different set of photophysical properties. As nature optimizes the photophysical properties of FPs according to their environmental conditions, they tend to include properties that don't match our research ideals. As such, we have developed methods to redesign and evolve FPs to fit our purposes.

In terms of brightness, red FPs have generally lagged behind their blue-shifted counterparts due to having much lower quantum yields and a tendency is mislocalize. One of the few exceptions to this general trend was tdTomato, a fused homodimeric red FP. This design was created to take advantage of the inherent oligomerization of its progenitor, DsRed, and maintain the critical dimer interface interactions, albeit within a single polypeptide chain. I decided to explore the possibilities of evolving FPs in this tandem dimer structure to allow the two FP domains to acquire different mutations to create tandem heterodimers. In order to allow this evolution, we codon differentiated the two halves of the tandem heterodimer by swapping the codons encoding for each amino acid. This allows the DNA encoding for the two nearly identical protein domains to be modified individually at the DNA level. Without codon differentiation, site specific modifications would target and change both FP domains simultaneously, limiting our ability to perform controlled modifications. I created a series of variants, the vine Tomatoes, which carry the vT suffix. The first was RRvT, which is the brightest red FP at the time of this writing. With the success of RRvT, I decided to explore the possibilities of tandem heterodimers and evolved a colour switched variant called GGvT that

unfortunately had only middling brightness. As GGvT and RRvT contain the same dimerization interface, I was able to mix and match the green and red parts creating GRvT. GRvT had some surprising properties with the red portion of the FP displaying a quantum yield of 97%. By assuming that the dimeric structure is identical to that of DsRed, I was able to characterize and model the FRET efficiency of this system in unprecedented detail. This system has a very high FRET efficiency which results in photophysical properties resembling a bright, long Stokes shift FP by efficiently converting blue excitation to red emission.

In terms of photostability, FPs have generally lagged behind their fluorescent dye and quantum dot counterparts. The low photostability of FPs limits our ability to perform experiments requiring long term and high speed imaging, both of which require large amounts of illumination and therefore are often limited by FP photobleaching. Despite the importance of photostability, we still have a very poor understanding of the mechanisms of photobleaching, which make rational efforts to improve photostability unreliable. To resolve this, I developed a low-cost robotic screening system that may democratize methods to improve photostability in new FP variants. This screening platform gives even illumination over large areas (*i.e.*, the dimensions of a Petri dish), enabling screening of photostability and photoconversion. During evolution, it became apparent that the concentration of the samples tested was strongly influencing the photobleaching half-life of the FPs. This concentration dependence can positively or negatively correlate to the photobleaching half-life. Using this robotic screening system, I created an FP variant called Citrine2 which is the most photostable yellow FP. Additionally, I proposed a new metric for FP utility called InPhO which incorporates both brightness and photostability and therefore enables unbiased comparisons between FPs.

The field of optogenetics is rapidly increasing our ability to gather information from, and take control of, cellular functions using light. In Chapter 4, I attempted to improve the properties of the photocleaving optogenetic actuator PhoCl by developing a screening system that can report the amount of photocleavage and subsequent dissociation. I designed a caged BiFC construct in an effort to create a fluorogenic sensor to report PhoCl dissociation. Unfortunately, both the caging and the fluorogenic signal were only partially successful as the caging was leaky and the fluorescent signal was too weak. I then switched to a FRET based screen that relies on the fluorescence colour change in PhoCl to create an increasing fluorescence signal as dissociation occurs. This was used as a pre-screen to identify variants, with potentially beneficial mutations, that were then subjected to a secondary PAGE-based screen. Unfortunately, this approach was not successful at finding improved variants but did identify several mutations which decreased the amount of dissociation. In Chapter 5, I developed a spectrally orthogonal photocleavable protein that I call SplitOr. To create SplitOr, I circularly permuted PSmOrange2 and used directed evolution to recover its fluorescence and photoconversion properties. Three viable SplitOr variants have been identified. These variants can all be photoconverted with 480 nm illumination, which is spectrally orthogonal to the 405 nm light used to photoconvert PhoCl. All three viable SplitOr variants displayed colour differences in the far red state, relative to PSmOrange2, that are consistent with a change in the chromophore environment that could be attributed to dissociation.

6.2 Future directions

6.2.1 vine Tomatoes

In terms of tool development, the future of vine Tomatoes could proceed in several ways. For example, it may yet be possible to evolve a blue fluorescent version of the vine Tomatoes. There is a transient blue fluorescent state that forms during the maturation of the red vine Tomatoes. Evolving the protein to remain trapped in this blue fluorescent state, similar to the way I evolved a green fluorescent protein from a red fluorescent protein, may create more interesting photophysical properties.

Like the vine Tomatoes, the ddFPs are also based on tdTomato/dTomato. Accordingly, mutations found that make the vine Tomatoes bright could be carried over into a next generation of brighter and more dynamic ddFPs. In addition, a future blue vine Tomato could potentially be converted into a blue ddFP. Sheng Yi in the Campbell lab is currently modifying a teal FP (TFP) such that it with dimerize with the Tomato series to create a TFP-Tomato FRET pair and potentially a ddTFP variant. The primary limitation of the GRvT is the relatively dim fluorescence of the green state. This problem could potentially be circumvented in the by replacing the green domain with a teal one TRvT, as TFP is one of the brightest FPs yet reported. Further evolution of the green chromophore could potentially create a powerful LSS type FP out of GRvT, as GRvT already has equivalent brightness to the best LSSFPs.

6.2.2 Photostability screening

The screening system reported in this thesis could easily be applied to essentially any FP and thereby enable photostability screening to contribute to the next generation of FPs. It has already been used to find an improved version of the red Ca^{2+} sensor K-GECO which

increased K-GECO's photostability by ~3 fold and maintains its fluorescence fold change (data not shown). However, as this work shows, the use of bacterial screening is problematic, as Citrine2 displays significantly less photostability improvement in HeLa cells than in *E. coli*. Developing a screening system around yeast or some other fast growing eukaryotic cell might enable better FPs to be developed.

The other avenue this research opened up is the effect of concentration on photostability. While the effect has been empirically confirmed, the rationale provided in this thesis is speculative and currently lacks experimental support. Delving into how mCitrine and Citrine2 exhibit different trends could greatly increase our understanding of photostability.

If photostability can be dynamically controlled, it may be possible to make a sensor where extrinsic factors, like calcium ions, change the photostability half-life of the FP. Such a sensor could potentially serve as the basis for an integrator of calcium ion concentration in a cell. If paired with a FRET donor, calcium ion-dependent bleaching would result in increasing fluorescence of the FRET donor. The photostability sensor FP must switch between high and low photobleaching half-life states. In the high half-life state, this theoretical sensor would see minimal change in the FRET donor fluorescence and in the low half-life state the fluorescence change would be rapid. The photostability sensor FP would need to be illuminated at a constant intensity to bleach the acceptor. The fluorescence of the FRET donor FP could be measured periodically throughout the experiment to report how much time was spent in the high and low half-life states. To be useful as an integrator, this would require a large magnitude half-life change in the photostability sensor FP. If successful, an FP with calcium-dependent photobleaching could serve as the basis for a novel imaging modality and a new integrator system.

6.2.3 PhoCl enhancement

The improvements to PhoCl that I sought to realize have not yet proven to be readily accessible using fluorescence based assays. The use of BiLC would remove the difficulties of overlapping fluorescence profiles in the BiFC designs and create a detectable signal much more rapidly. One of the difficulties of such a system is supplying the substrate cheaply and equally to all colonies in an expressed library. This challenge would need to be addressed before BiLC could be used for large scale or high throughput screening. Further development of PhoCl screening systems could potentially prove useful for other photocleaving proteins such as SplitOr and other optogenetic actuators.

6.2.4 SplitOr development

Currently, SplitOr is being tested in the lab of Dr. Eric Schreiter (Janelia Research Campus, Ashburn, Virginia) for photocleavage in mammalian cells. The Schreiter lab has a long term goal of creating an integrator system for Ca^{2+} based on its cleavage, and have already made substantial progress towards this goal using PhoCl analogs. To create a version based on SplitOr, they will need to develop a variant that photocleaves in the presence of Ca^{2+} and the dissociation must create a persistent fluorescent or luminescent signal. They have independently confirmed that all three SplitOr variants photoconvert under 50 mW/cm^2 470 nm LED illumination, and preparing to perform further studies to confirm its photocleavage and dissociation properties, this may require further evolution to make these traits as robust as possible.

6.2.5 Continuous evolution

This thesis is focussed on creating high throughput screening systems that allow screening of hard-to-visualize properties such as photostability and photocleavage. A major limitation of all of these screens is that they require human intervention to select the best variants. If I could design a system in which survival is dependent on fluorescence, a population of variants could be left alone and variants would evolve to thrive under the given conditions.

Phage-assisted continuous evolution (PACE) was used successfully to evolve T7 RNA polymerase.[177] This was achieved by linking the infection rate of the phage to the function of T7 RNA polymerase, which is needed to create the essential pIII phage protein. The more pIII is produced, the more efficiently the phage infects cells. So, if the phage carries a mutation that enhances T7 RNA polymerase, it will infect more cells and produce a next generation that can carry further improvements. As such, any enzyme function that can be linked to the production of pIII can be evolved using PACE.[177,178]

As fluorescence can't be easily linked to pIII production, PACE cannot be used directly. Another form of continuous evolution is forcing survival in an increasing toxic or stressful environment. This was done to evolve antibacterial resistance using increasing antibiotic concentration across an two dimensional growth environment.[179] This forces bacteria to evolve antibacterial resistance to continue growth by expanding into higher antibiotic concentration territory and out of high density, low antibiotic regions.

I propose that FPs fluorescent properties could be linked to survival if the fluorescence is linked to photosynthesis. If an FP can be fused to photosystem II of a fast growing photosynthetic organism, the photons that the FP absorbs can be transferred via FRET to

chlorophyll, powering photosynthesis. Photosynthetic organisms do not utilize green wavelengths efficiently for energy production. Accordingly, I could use FPs to convert unused green photons into useful red photons. By providing only green photons as a light source for bacterial growth, the survival of the photosynthetic organism would depend on the FP fluorescence and FRET as the primary source of energy. As a higher FP extinction coefficient and quantum yield would result in more efficient photosynthesis, such a system would necessarily select for bacteria with brighter fluorescence.

To be useful in a continuous evolution set-up, I would need to use photosynthetic bacteria with a fast doubling time. *Synechococcus elongatus* UTEX 2973[180] doubles in less than 2 hours in optimal conditions and it grows optimally in BG11 media at 41°C with 3% CO_2 in the atmosphere.[181] It can be genetically manipulated using triparental mating with *E. coli,*[182] so only trivial genetic engineering of *E. coli* is required.

In order for this system to work, an LSSFP such as LSSmKate2,[183] would need to be used. LSSmKate2 excites at 460 nm and emits at 605 nm and has good spectral overlap for FRET with chlorophyll a. To have good FRET efficiency, it would need to be fused to the N-terminus of photosystem II so that it is held close to the associated chlorophyll a cofactors.[184]

Energetically, the FP must provide the bacteria with enough energy to "pay" for its own creation. This will require the transfer of ~3200 photons to produce the ~1000 ATP required for translation[185] of 250 AA of FP, and a further 200 ATP to transcribe $1/10^{th}$ of an mRNA[186] as they are on average translated 10 times (may be as high as 10 000).[187] If LSSmKate2 photobleaches too quickly to might not provide enough benefit to pay its energy cost. FPs such as eGFP and YFP each have a photon budget of 10^5 photons before they photobleach.[188] If the same holds for LSSmKate2, a FRET efficiency of only 32% is required to create more ATP

than was spent creating it. This is a conservative estimate, as a FRET donor is partially protected from bleaching by the acceptor.[189]

These rough calculations suggest that the FP could provide energy to *S. elongatus*, making this continuous directed evolution feasible. Regulation of the wavelength and intensity of the incident light would allow for further increasing of the selection pressure on the FP to more efficiently convert the light. Thus, any improvements to LSSmKate2's brightness or photostability would provide increased benefit to *S. elongatus* and create a continuous evolution system for FP development.

6.3 Final thoughts

The field of FP engineering has never looked brighter. As FPs are further improved, creating high throughput screening systems for novel properties will become increasingly vital to advancing the field. This will mean finding new ways to screen in tissues and specific cell types. Improved techniques to make libraries that contain more variants and if possible more valuable variants, and ways to speed up the process. In more general terms, we need to see the limits of our current approaches to move beyond them.

Bibliography

1 A. R. Kherlopian, T. Song, Q. Duan, M. a Neimark, M. J. Po, J. K. Gohagan and A. F. Laine, *BMC Syst. Biol.*, 2008, **2**, 74.

2 V. J. Pansare, S. Hejazi, W. J. Faenza and R. K. Prud'homme, *Chem. Mater.*, 2012, **24**, 812–827.

3 S. L. Jacques, *Phys. Med. Biol.*, 2013, **58**, R37–R61.

4 J. R. Lakowicz, Ed., *Principles of Fluorescence Spectroscopy*, Springer US, Boston, MA, 2006.

5 I. Itzkan, L. Qiu, H. Fang, M. M. Zaman, E. Vitkin, I. C. Ghiran, S. Salahuddin, M. Modell, C. Andersson, L. M. Kimerer, P. B. Cipolloni, K.-H. Lim, S. D. Freedman, I. Bigio, B. P. Sachs, E. B. Hanlon and L. T. Perelman, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 17255–17260.

6 G. L. Humason, *Animal tissue techniques:4th.*, W.H. Freeman, San Francisco, 1979.

7 M. Göppert-Mayer, *Ann. Phys.*, 1931, **401**, 273–294.

8 W. Kaiser and C. Garrett, *Phys. Rev. Lett.*, 1961, **7**, 229–231.

9 S. H. D. Haddock, M. A. Moline and J. F. Case, *Ann. Rev. Mar. Sci.*, 2010, **2**, 443–93.

10 O. Shimomura, F. H. Johnson and Y. Saiga, *J. Cell. Comp. Physiol.*, 1962, **59**, 223–239.

11 M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward and D. C. Prasher, *Science*, 1994, **263**, 802–805.

12 R. Heim and R. Y. Tsien, *Curr. Biol.*, 1996, **6**, 178–182.

13 L. A. Gross, G. S. Baird, R. C. Hoffman, K. K. Baldridge and R. Y. Tsien, *Proc. Natl. Acad. Sci.*, 2000, **97**, 11990–11995.

14 N. C. Shaner, R. E. Campbell, P. A. Steinbach, B. N. G. Giepmans, A. E. Palmer and

R. Y. Tsien, *Nat. Biotechnol.*, 2004, **22**, 1567–72.

15 D. C. Prasher, V. K. Eckenrode, W. W. Ward, F. G. Prendergast and M. J. Cormier, *Gene*, 1992, **111**, 229–233.

16 F. Yang, L. G. Moss, G. N. Phillips, G. N. Phillips Jr. and G. N. Phillips, *Nat. Biotechnol.*, 1996, **14**, 1246–1251.

17 M. Ormö, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien and S. J. Remington, *Science*, 1996, **273**, 1392.

18 S. Simeon, W. Shoombuatong, N. Anuwongcharoen, L. Preeyanon, V. Prachayasittikul, J. E. S. Wikberg and C. Nantasenamat, *J. Cheminform.*, 2016, **8**, 72.

19 H. Hoi, E. S. Howe, Y. Ding, W. Zhang, M. A. Baird, B. R. Sell, J. R. Allen, M. W. Davidson and R. E. Campbell, *Chem. Biol.*, 2013, **20**, 1296–1304.

20 M. V Matz, A. F. Fradkov, Y. a Labas, A. P. Savitsky, A. G. Zaraisky, M. L. Markelov and S. A. Lukyanov, *Nat. Biotechnol.*, 1999, **17**, 969–973.

21 S. Karasawa, T. Araki, T. Nagai, H. Mizuno and A. Miyawaki, *Biochem. J.*, 2004, **381**, 307–12.

22 N. C. Shaner, G. G. Lambert, A. Chammas, Y. Ni, P. J. Cranfill, M. A. Baird, B. R. Sell, J. R. Allen, R. N. Day, M. Israelsson, M. W. Davidson and J. Wang, *Nat. Methods*, 2013, **10**, 407–409.

23 A. Kumagai, R. Ando, H. Miyatake, P. Greimel, T. Kobayashi, Y. Hirabayashi, T. Shimogori and A. Miyawaki, *Cell*, 2013, **153**, 1602–1611.

24 E. A. Rodriguez, G. N. Tran, L. A. Gross, J. L. Crisp, X. Shu, J. Y. Lin and R. Y. Tsien, *Nat. Methods*, 2016, **13**, 763–769.

25 D. Shcherbo, I. I. Shemiakina, A. V Ryabova, K. E. Luker, B. T. Schmidt, E. A.

Souslova, T. V Gorodnicheva, L. Strukova, K. M. Shidlovskiy, O. V Britanova, A. G. Zaraisky, K. A. Lukyanov, V. B. Loschenov, G. D. Luker and D. M. Chudakov, *Nat. Methods*, 2010, **7**, 827–829.

26 K. A. Rumyantsev, D. M. Shcherbakova, N. I. Zakharova, A. V Emelyanov, K. K. Turoverov and V. V Verkhusha, *Sci. Rep.*, 2015, **5**, 18348.

27 A. Bar-Even, E. Noor, Y. Savir, W. Liebermeister, D. Davidi, D. S. Tawfik and R. Milo, *Biochemistry*, 2011, **50**, 4402–4410.

28 D. S. Bindels, L. Haarbosch, L. van Weeren, M. Postma, K. E. Wiese, M. Mastop, S. Aumonier, G. Gotthard, A. Royant, M. A. Hink and T. W. J. Gadella Jr, *Nature Methods*, 2017.

29 R. Heim, D. C. Prasher and R. Y. Tsien, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 12501–4.

30 M. A. Rosenow, H. N. Patel and R. M. Wachter, *Biochemistry*, 2005, **44**, 8303–8311.

31 R. L. Strack, D. E. Strongin, L. Mets, B. S. Glick and R. J. Keenan, *J. Am. Chem. Soc.*, 2010, **132**, 8496–8505.

32 B. T. Bajar, E. S. Wang, A. J. Lam, B. B. Kim, C. L. Jacobs, E. S. Howe, M. W. Davidson, M. Z. Lin and J. Chu, *Sci. Rep.*, 2016, **6**, 20889.

33 J. Goedhart, D. von Stetten, M. Noirclerc-Savoye, M. Lelimousin, L. Joosen, M. A. Hink, L. van Weeren, T. W. J. Gadella and A. Royant, *Nat. Commun.*, 2012, **3**, 751.

34 K. a Walther, B. Papke, M. B. Sinn, K. Michel and A. Kinkhabwala, *Mol. Biosyst.*, 2011, **7**, 322–336.

35 G. H. Patterson and J. Lippincott-Schwartz, *Science*, 2002, **297**, 1873–1877.

36 B. F. Fosque, Y. Sun, H. Dana, C. T. Yang, T. Ohyama, M. R. Tadross, R. Patel, M.

Zlatic, D. S. Kim, M. B. Ahrens, V. Jayaraman, L. L. Looger and E. R. Schreiter, *Science*, 2015, **347**, 755–760.

37 O. M. Subach, G. H. Patterson, L.-M. Ting, Y. Wang, J. S. Condeelis and V. V Verkhusha, *Nat. Methods*, 2011, **8**, 771–777.

38 O. M. Subach, D. Entenberg, J. S. Condeelis and V. V. Verkhusha, *J. Am. Chem. Soc.*, 2012, **134**, 14789–14799.

39 S. Pletnev, D. M. Shcherbakova, O. M. Subach, N. V. Pletneva, V. N. Malashkevich, S. C. Almo, Z. Dauter and V. V. Verkhusha, *PLoS One*, 2014, **9**, 1–12.

40 T. Grotjohann, I. Testa, M. Leutenegger, H. Bock, N. T. Urban, F. Lavoie-Cardinal, K. I. Willig, C. Eggeling, S. Jakobs and S. W. Hell, *Nature*, 2011, **478**, 204–208.

41 Y. Shen, M. D. Wiens and R. E. Campbell, *RSC Adv.*, 2014, **4**, 56762–56765.

42 C. Ding and J. Wei, *Sci. Rep.*, 2016, **6**, 18845.

43 H. Hoi, T. Matsuda, T. Nagai and R. E. Campbell, *J. Am. Chem. Soc.*, 2013, **135**, 46–49.

44 B. L. Grigorenko, A. V. Nemukhin, I. V. Polyakov, M. G. Khrenova and A. I. Krylov, *J. Phys. Chem. B*, 2015, **119**, 5444–5452.

45 G. Donnert, C. Eggeling and S. W. Hell, *Nat. Methods*, 2007, **4**, 81–86.

46 J. J. van Thor, T. Gensch, K. J. Hellingwerf and L. N. Johnson, *Nat. Struct. Biol.*, 2002, **9**, 37–41.

47 K. M. Dean, J. L. Lubbeck, L. M. Davis, C. K. Regmi, P. P. Chapagain, B. S. Gerstman, R. Jimenez and A. E. Palmer, *Integr. Biol.*, 2015, **7**, 263–73.

48 I. Ghosh, A. D. Hamilton and L. Regan, *J. Am. Chem. Soc.*, 2000, **122**, 5658–5659.

49 Y. J. Shyu and C. D. Hu, *Trends Biotechnol.*, 2008, **26**, 622–630.

50 S. Cabantous, T. C. Terwilliger and G. S. Waldo, *Nat. Biotechnol.*, 2005, **23**, 102–107.

51 S. J. Kellermann, A. K. Rath and A. Rentmeister, *ChemBioChem*, 2013, **14**, 200–204.

52 G. S. Waldo, S. Cabantous, US 7,666,606 B2, 2010.

53 T. J. Magliery, C. G. M. Wilson, W. Pan, D. Mishler, I. Ghosh, A. D. Hamilton and L. Regan, *J. Am. Chem. Soc.*, 2005, **127**, 146–157.

54 C. Zych, A. Domling and V. Ayyavoo, *Drug Des. Devel. Ther.*, 2013, **7**, 403–412.

55 Z. Ding, J. Liang, Y. Lu, Q. Yu, Z. Songyang, S.-Y. Lin and G. B. Mills, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 15014–15019.

56 T. Kojima, S. Karasawa, A. Miyawaki, T. Tsumuraya and I. Fujii, *J. Biosci. Bioeng.*, 2011, **111**, 397–401.

57 J. N. Offenborn, R. Waadt and J. Kudla, *New Phytol.*, 2015, **208**, 269–279.

58 S. C. Alford, A. S. Abdelfattah, Y. Ding and R. E. Campbell, *Chem. Biol.*, 2012, **19**, 353–360.

59 S. C. Alford, Y. Ding, T. Simmen and R. E. Campbell, *ACS Synth. Biol.*, 2012, **1**, 569–575.

60 Y. Ding, J. Li, J. R. Enterina, Y. Shen, I. Zhang, P. H. Tewson, G. C. H. Mo, J. Zhang, A. M. Quinn, T. E. Hughes, D. Maysinger, S. C. Alford, Y. Zhang and R. E. Campbell, *Nat. Methods*, 2015, **12**, 195–198.

61 P.-A. Vidi and V. J. Watts, *Mol. Pharmacol.*, 2009, **75**, 733–739.

62 K. E. Luker, M. C. P. Smith, G. D. Luker, S. T. Gammon, H. Piwnica-Worms and D. Piwnica-Worms, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 12288–12293.

63 N. Hida, M. Awais, M. Takeuchi, N. Ueno, M. Tashiro, C. Takagi, T. Singh, M. Hayashi, Y. Ohmiya and T. Ozawa, *PLoS One*, 2009, **4**.

64 N. Misawa, A. K. M. Kafi, M. Hattori, K. Miura, K. Masuda and T. Ozawa, *Anal. Chem.*, 2010, **82**, 2552–2560.

65 E. Conti, N. P. Franks and P. Brick, *Structure*, 1996, **4**, 287–298.

66 A. M. Loening, T. D. Fenn and S. S. Gambhir, *J. Mol. Biol.*, 2007, **374**, 1017–1028.

67 I. Remy and S. W. Michnick, *Nat. Methods*, 2006, **3**, 977–979.

68 M. P. Hall, J. Unch, B. F. Binkowski, M. P. Valley, B. L. Butler, M. G. Wood, P. Otto, K. Zimmerman, G. Vidugiris, T. MacHleidt, M. B. Robers, H. A. Benink, C. T. Eggers, M. R. Slater, P. L. Meisenheimer, D. H. Klaubert, F. Fan, L. P. Encell and K. V. Wood, *ACS Chem. Biol.*, 2012, **7**, 1848–1857.

69 S. Inouye, K. Watanabe, H. Nakamura and O. Shimomura, *FEBS Lett.*, 2000, **481**, 19–25.

70 K. V. Lovell, S., Mehzabeen, N., Battaile, K.P., Wood, M.G., Encell, L.P., Wood, 2016.

71 A. S. Dixon, M. K. Schwinn, M. P. Hall, K. Zimmerman, P. Otto, T. H. Lubben, B. L. Butler, B. F. Binkowski, T. MacHleidt, T. A. Kirkland, M. G. Wood, C. T. Eggers, L. P. Encell and K. V. Wood, *ACS Chem. Biol.*, 2016, **11**, 400–408.

72 T. Xu, D. Close, W. Handagama, E. Marr, G. Sayler and S. Ripp, *Front. Oncol.*, 2016, **6**, 1–8.

73 T. Ozawa, a Kaihara, M. Sato, K. Tachihara and Y. Umezawa, *Anal. Chem.*, 2001, **73**, 2516–2521.

74 T. Endoh, M. Mie, H. Funabashi, T. Sawasaki, Y. Endo and E. Kobatake, *Bioconjug. Chem.*, 2007, **18**, 956–962.

75 S. B. Kim, A. Kanno, T. Ozawa, H. Tao and Y. Umezawa, *ACS Chem. Biol.*, 2007, **2**, 484–492.

76 S. Sankaranarayanan, D. De Angelis, J. E. Rothman and T. A. Ryan, *Biophys. J.*, 2000, **79**, 2199–208.

77 Y. Shen, M. Rosendale, R. E. Campbell and D. Perrais, *J. Cell Biol.*, 2014, **207**, 419–432.

78 Y. Zhao, S. Araki, J. Wu, T. Teramoto, Y.-F. Chang, M. Nakano, A. S. Abdelfattah, M. Fujiwara, T. Ishihara, T. Nagai and R. E. Campbell, *Science*, 2011, **333**, 1888–1891.

79 T.-W. Chen, T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman, L. L. Looger, K. Svoboda and D. S. Kim, *Nature*, 2013, **499**, 295–300.

80 T. Mizuno, K. Murao, Y. Tanabe, M. Oda and T. Tanaka, *J. Am. Chem. Soc.*, 2007, **129**, 11378–11383.

81 J. S. Marvin, B. G. Borghuis, L. Tian, J. Cichon, M. T. Harnett, J. Akerboom, A. Gordus, S. L. Renninger, T. Chen, C. I. Bargmann, M. B. Orger, E. R. Schreiter, J. B. Demb, W. Gan, S. A. Hires and L. L. Looger, *Nat. Methods*, 2013, **10**, 162–170.

82 A. W. Nguyen and P. S. Daugherty, *Nat. Biotechnol.*, 2005, **23**, 355–360.

83 L. Lindenburg and M. Merkx, *Sensors (Switzerland)*, 2014, **14**, 11691–11713.

84 D. Tischer and O. D. Weiner, *Nat. Rev. Mol. Cell Biol.*, 2014, **15**, 551–8.

85 M. E. Bulina, D. M. Chudakov, O. V Britanova, Y. G. Yanushevich, D. B. Staroverov, T. V Chepurnykh, E. M. Merzlyak, M. A. Shkrob, S. Lukyanov and K. A. Lukyanov, *Nat Biotechnol*, 2006, **24**, 95–99.

86 A. F. Fradkov, V. V Verkhusha, D. B. Staroverov, M. E. Bulina, Y. G. Yanushevich, V. I. Martynov, S. Lukyanov and K. A. Lukyanov, *Biochem. J.*, 2002, **368**, 17–21.

87 E. R. Stadtman and R. L. Levine, *Amino Acids*, 2003, **25**, 207–218.

88 T. J. Dougherty, J. E. Kaufman, A. Goldfarb, K. R. Weishaupt, D. Boyle and A. Mittleman, *Cancer Res.*, 1978, **38**, 2628–35.

89 R. Ando, H. Mizuno and A. Miyawaki, *Science*, 2004, **306**, 1370–3.

90 X. X. Zhou, H. K. Chung, A. J. Lam and M. Z. Lin, *Science*, 2013, **338**, 810–814.

91 H. Zhao, L. Giver, Z. Shao, J. A. Affholter and F. H. Arnold, *Nat. Biotechnol.*, 1998, **16**, 258–261.

92 N. C. Shaner, P. A. Steinbach and R. Y. Tsien, *Nat. Methods*, 2005, **2**, 905–909.

93 R. E. Campbell, O. Tour, A. E. Palmer, P. A. Steinbach, G. S. Baird, D. A. Zacharias and R. Y. Tsien, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 7877–7882.

94 H. Ai, J. N. Henderson, S. J. Remington and R. E. Campbell, *Biochem. J.*, 2006, **400**, 531–40.

95 H.-W. Ai, M. A. Baird, Y. Shen, M. W. Davidson and R. E. Campbell, *Nat. Protoc.*, 2014, **9**, 910–28.

96 N. G. Gurskaya, A. F. Fradkov, A. Terskikh, M. V. Matz, Y. A. Labas, V. I. Martynov, Y. G. Yanushevich, K. A. Lukyanov and S. A. Lukyanov, *FEBS Lett.*, 2001, **507**, 16–20.

97 D. Shcherbo, C. S. Murphy, G. V Ermakova, E. A. Solovieva, T. V Chepurnykh, A. S. Shcheglov, V. V Verkhusha, V. Z. Pletnev, K. L. Hazelwood, P. M. Roche, S. Lukyanov, A. G. Zaraisky, M. W. Davidson and D. M. Chudakov, *Biochem. J.*, 2009, **418**, 567–574.

98 J. Hasegawa, Y. Sakamoto, S. Nakagami, M. Aida, S. Sawa and S. Matsunaga, *Plant Cell Physiol.*, 2016, **57**, pcw027.

99 L. F. Sobala, Y. Wang and P. N. Adler, *Development*, 2015, **142**, 3974–81.

100 S. Wei, Q. Zou, S. Lai, Q. Zhang, L. Li, Q. Yan, X. Zhou, H. Zhong and L. Lai, *Sci. Rep.*, 2016, **6**, 19648.

101 H. Yagi, Y. Oka, M. Komada, M.-J. Xie, K. Noguchi and M. Sato, *Neurosci. Lett.*, 2015, **612**, 18–24.

102 A. J. Lam, F. St-Pierre, Y. Gong, J. D. Marshall, P. J. Cranfill, M. A. Baird, M. R. McKeown, J. Wiedenmann, M. W. Davidson, M. J. Schnitzer, R. Y. Tsien and M. Z. Lin, *Nat. Methods*, 2012, **9**, 1005–1012.

103 R. Grünberg, J. V Burnier, T. Ferrar, V. Beltran-Sastre, F. Stricher, A. M. van der Sloot, R. Garcia-Olivas, A. Mallabiabarrena, X. Sanjuan, T. Zimmermann and L. Serrano, *Nat. Methods*, 2013, **10**, 1021–7.

104 I. Kotera, T. Iwasaki, H. Imamura, H. Noji and T. Nagai, *ACS Chem. Biol.*, 2010, **5**, 215–222.

105 L. H. Lindenburg, A. M. Hessels, E. H. T. M. Ebberink, R. Arts and M. Merkx, *ACS Chem. Biol.*, 2013, **8**, 2133–2139.

106 K. Saito, Y.-F. Chang, K. Horikawa, N. Hatsugai, Y. Higuchi, M. Hashida, Y. Yoshida, T. Matsuda, Y. Arai and T. Nagai, *Nat. Commun.*, 2012, **3**, 1262.

107 V. Baubet, H. Le Mouellic, A. K. Campbell, E. Lucas-Meunier, P. Fossier, P. Brulet and P. Brúlet, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 7260–5.

108 H. Hoshino, Y. Nakajima and Y. Ohmiya, *Nat. Methods*, 2007, **4**, 637–639.

109 M. F. Garcia-Parajo, M. Koopman, E. van Dijk, V. Subramaniam and N. F. van Hulst, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 14392–14397.

110 V. V. Verkhusha, D. M. Chudakov, N. G. Gurskaya, S. Lukyanov and K. A. Lukyanov, *Chem. Biol.*, 2004, **11**, 845–854.

111 G. S. Baird, D. A. Zacharias and R. Y. Tsien, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 11984–9.

112 A. V. Terskikh, A. F. Fradkov, A. G. Zaraisky, A. V. Kajava and B. Angres, *J. Biol. Chem.*, 2002, **277**, 7633–7636.

113 B. J. Bevis and B. S. Glick, *Nat. Biotechnol.*, 2002, **20**, 83–87.

114 N. Durisic, A. G. Godin, C. M. Wever, C. D. Heyes, M. Lakadamyali and J. A. Dent, *J. Neurosci.*, 2012, **32**, 12915–20.

115 M. H. Ulbrich and E. Y. Isacoff, *Nat. Methods*, 2007, **4**, 319–21.

116 D. Yarbrough, R. M. Wachter, K. Kallio, M. V Matz and S. J. Remington, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 462–467.

117 T. Kato, H. Kashida, H. Kishida, H. Yada, H. Okamoto and H. Asanuma, *J. Am. Chem. Soc.*, 2013, **135**, 741–750.

118 M. A. Salem and A. Brown, *J. Chem. Theory Comput.*, 2014, **10**, 3260–3269.

119 Schrödinger LLC, The PyMOL Molecular Graphics System, Version 1.8.

120 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.

121 E. Runge and E. K. U. Gross, *Phys. Rev. Lett.*, 1984, **52**, 997–1000.

122 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

123 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158.

124 R. Ditchfield, *J. Chem. Phys.*, 1971, **54**, 724.

125 W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.*, 1972, **56**, 2257–2261.

126 M. M. Francl, *J. Chem. Phys.*, 1982, **77**, 3654.

127 P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta*, 1973, **28**, 213–222.

128 T. Clark, J. Chandrasekhar, G. W. Spitznagel and P. V. R. Schleyer, *J. Comput. Chem.*,

1983, **4**, 294–301.

129 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.

130 V. Barone and M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.

131 M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem.*, 2003, **24**, 669–681.

132 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3093.

133 M. Cossi and V. Barone, *J. Chem. Phys.*, 2001, **115**, 4708–4717.

134 J. Benesch, A. Askendal and P. Tengvall, *J. Colloid Interface Sci.*, 2002, **249**, 84–90.

135 B. W. van der Meer, D. M. van der Meer and S. S. Vogel, *FRET - Förster Resonance Energy Transfer*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2013.

136 G. Jung, J. Wiehler and A. Zumbusch, *Biophys. J.*, 2005, **88**, 1932–1947.

137 W. R. Carmody, *J. Chem. Educ.*, 1961, **38**, 559.

138 A. S. Abdelfattah, S. L. Farhi, Y. Zhao, D. Brinks, P. Zou, A. Ruangkittisakul, J. Platisa, V. A. Pieribone, K. Ballanyi, A. E. Cohen and R. E. Campbell, *J. Neurosci.*, 2016, **36**, 2458–72.

139 M. L. Martin-Fernandez and D. T. Clarke, *Int. J. Mol. Sci.*, 2012, **13**, 14742–14765.

140 L. Greenbaum, C. Rothmann and R. Lavie, *Biol. Chem.*, 2000, **381**, 1251–1258.

141 R. Swaminathan, C. P. Hoang and A. S. Verkman, *Biophys. J.*, 1997, **72**, 1900–7.

142 N. C. Shaner, M. Z. Lin, M. R. McKeown, P. A. Steinbach, K. L. Hazelwood, M. W. Davidson and R. Y. Tsien, *Nat. Methods*, 2008, **5**, 545–551.

143 H. Ren, B. Yang, C. Ma, Y. S. Hu, P. G. Wang and L. Wang, *ACS Chem. Biol.*, 2016, **11**, 2679–2684.

144 P. J. Cranfill, B. R. Sell, M. A. Baird, J. R. Allen, Z. Lavagnino, H. M. de Gruiter, G.-J. Kremers, M. W. Davidson, A. Ustione and D. W. Piston, *Nat. Methods*, 2016, 1–7.

145 M. a Mena, T. P. Treynor, S. L. Mayo and P. S. Daugherty, *Nat. Biotechnol.*, 2006, **24**, 1569–1571.

146 H. W. Ai, N. C. Shaner, Z. Cheng, R. Y. Tsien and R. E. Campbell, *Biochemistry*, 2007, **46**, 5904–5910.

147 Y. Zhao, A. S. Abdelfattah, Y. Zhao, A. Ruangkittisakul, K. Ballanyi, R. E. Campbell and D. J. Harrison, *Integr. Biol.*, 2014, **6**, 714–25.

148 L. M. Davis, J. L. Lubbeck, K. M. Dean, A. E. Palmer and R. Jimenez, *Lab Chip*, 2013, **13**, 2320.

149 J. L. Lubbeck, K. M. Dean, H. Ma, A. E. Palmer and R. Jimenez, *Anal. Chem.*, 2012, **84**, 3929–3937.

150 B. P. Cormack, R. H. Valdivia and S. Falkow, *Gene*, 1996, **173**, 33–38.

151 N. Carlson and M. Rechsteiner, *J. Cell Biol.*, 1987, **104**, 537–546.

152 G. J. Pielak, C. Li, A. C. Miklos, A. P. Schlesinger, M. Slade, G. Wang and I. G. Zigoneanu, 2010, **48**, 226–234.

153 O. Griesbeck, G. S. Baird, R. E. Campbell, D. A. Zacharias and R. Y. Tsien, *J. Biol. Chem.*, 2001, **276**, 29188–29194.

154 D. A. Zacharias, J. D. Violin, A. C. Newton and R. Y. Tsien, *Science*, 2002, **296**, 913–6.

155 R. M. Wachter and S. J. Remington, *Curr. Biol.*, 1999, **9**, 628–629.

156 W. P. Dempsey, L. Georgieva, P. M. Helbling, A. Y. Sonay, T. V Truong, M. Haffner and P. Pantazis, *Nat. Methods*, 2015, **12**, 645–8.

157 S. Raut, J. Kimball, R. Fudala, H. Doan, B. Maliwal, N. Sabnis, A. Lacko, I. Gryczynski, S. V Dzyuba and Z. Gryczynski, *Phys. Chem. Chem. Phys.*, 2014, **16**, 27037–42.

158 S. Jayaraman, P. Haggie, R. M. Wachter, S. J. Remington and A. S. Verkman, *J. Biol. Chem.*, 2000, **275**, 6047–6050.

159 D. R. James, A. Siemiarczuk and W. R. Ware, *Rev. Sci. Instrum.*, 1992, **63**, 1710–1716.

160 D. A. Shagin, E. V. Barsova, Y. G. Yanushevich, A. F. Fradkov, K. A. Lukyanov, Y. A. Labas, T. N. Semenova, J. A. Ugalde, A. Meyers, J. M. Nunez, E. A. Widder, S. A. Lukyanov and M. V. Matz, *Mol. Biol. Evol.*, 2004, **21**, 841–850.

161 M. E. Bulina, K. A. Lukyanov, O. V. Britanova, D. Onichtchouk, S. Lukyanov and D. M. Chudakov, *Nat. Protoc.*, 2006, **1**, 947–953.

162 K. S. Sarkisyan, O. A. Zlobovskaya, D. A. Gorbachev, N. G. Bozhanova, G. V. Sharonov, D. B. Staroverov, E. S. Egorov, A. V. Ryabova, K. M. Solntsev, A. S. Mishin and K. A. Lukyanov, *PLoS One*, 2015, **10**, 1–11.

163 K. Takemoto, T. Matsuda, N. Sakai, D. Fu, M. Noda, S. Uchiyama, I. Kotera, Y. Arai, M. Horiuchi, K. Fukui, T. Ayabe, F. Inagaki, H. Suzuki and T. Nagai, *Sci .Rep.*, 2013, **3**, 2629.

164 A. L. McEvoy, H. Hoi, M. Bates, E. Platonova, P. J. Cranfill, M. A. Baird, M. W. Davidson, H. Ewers, J. Liphardt and R. E. Campbell, *PLoS One*, 2012, **7**.

165 H. Mizuno, T. K. Mal, K. I. Tong, R. Ando, T. Furuta, M. Ikura and A. Miyawaki, *Mol. Cell*, 2003, **12**, 1051–1058.

166 R. D. Fritz, M. Letzelter, A. Reimann, K. Martin, L. Fusco, L. Ritsma, B. Ponsioen, E. Fluri, S. Schulte-Merker, J. van Rheenen and O. Pertz, *Sci. Signal.*, 2013, **6**, rs12-rs12.

167 J. Kügler, S. Schmelz, J. Gentzsch, S. Haid, E. Pollmann, J. Van Den Heuvel, R. Franke, T. Pietschmann, D. W. Heinz and J. Collins, *J. Biol. Chem.*, 2012, **287**, 39224–39232.

168 D. Picard, *Methods Enzymol.*, 2000, **327**, 385–401.

169 L. C. Scherrer, D. Picard, E. Massa, J. M. Harmon, S. S. Simons, K. R. Yamamoto and W. B. Pratt, *Biochemistry*, 1993, **32**, 5381–6.

170 T. Matsuda and C. L. Cepko, *Proc. Natl. Acad. Sci.*, 2007, **104**, 1027–1032.

171 J. S. . Zhang, Wei; Lohman, Alexander W.; Zhuravlova, Yevgeniya; Lu, Xiaocen; Wiens, Matthew D.; Hoi, Hiofan; Yaganoglu, Sine; Mohr, Manuel A.; Kitova, Elena N.; Klassen and R. E. Pantazis, Periklis; Thompson, Roger J.; Campbell, *Nat. Methods, Submitted*.

172 A. Tjernberg, S. Carno, F. Oliv, K. Benkestock, P.-O. Edlund, W. J. Griffiths and D. Hallen, *Anal. Chem.*, 2004, **76**, 4325–4331.

173 T. L. Radley, A. I. Markowska, B. T. Bettinger, J.-H. Ha and S. N. Loh, *J. Mol. Biol.*, 2003, **332**, 529–536.

174 J. N. Henderson, H.-W. Ai, R. E. Campbell and S. J. Remington, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 6672–6677.

175 C. Bamann, T. Kirsch, G. Nagel and E. Bamberg, *J. Mol. Biol.*, 2008, **375**, 686–694.

176 G. Kremers, K. L. Hazelwood, C. S. Murphy, M. W. Davidson and D. W. Piston, *Nat. Methods*, 2009, **6**, 355–358.

177 K. M. Esvelt, J. C. Carlson and D. R. Liu, *Nature*, 2011, **472**, 499–503.

178 A. H. Badran, V. M. Guzov, Q. Huai, M. M. Kemp, P. Vishwanath, W. Kain, A. M. Nance, A. Evdokimov, F. Moshiri, K. H. Turner, P. Wang, T. Malvar and D. R. Liu, *Nature*, 2016, **533**, 58–63.

179 M. Baym, T. D. Lieberman, E. D. Kelsic, R. Chait, R. Gross, I. Yelin and R. Kishony, *Science*, 2016, **353**, 1147–1151.

180 J. Yu, M. Liberton, P. F. Cliften, R. D. Head, J. M. Jacobs, R. D. Smith, D. W.

Koppenaal, J. J. Brand and H. B. Pakrasi, *Sci. Rep.*, 2015, **5**, 8132.

181 R. Y. Stanier, R. Kunisawa, M. Mandel and G. Cohen-Bazire, *Bacteriol. Rev.*, 1971, **35**, 171–205.

182 N. F. Tsinoremas, A. K. Kutach, C. A. Strayer and S. S. Golden, *J. Bacteriol.*, 1994, **176**, 6764–6768.

183 K. D. Piatkevich, J. Hulit, O. M. Subach, B. Wu, A. Abdulla, J. E. Segall and V. V Verkhusha, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 5369–74.

184 S. Vasil'ev, P. Orth, a Zouni, T. G. Owens and D. Bruce, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 8602–8607.

185 M. Lynch and G. K. Marinov, *Proc. Natl. Acad. Sci.*, 2015, **112**, 15690–15695.

186 O. Yarchuk, N. Jacques, J. Guillerez and M. Dreyfus, *J. Mol. Biol.*, 1992, **226**, 581–596.

187 J. J. Li, P. J. Bickel and M. D. Biggin, *PeerJ*, 2014, **2**, e270.

188 Q. Zhao, I. T. Young and J. G. S. de Jong, *J. Biomed. Opt.*, 2011, **16**, 86007.

189 R. C. Patel, D. C. Lange and Y. C. Patel, *Methods*, 2002, **27**, 340–348.