

# **Spectroscopic Evaluation of Excited State Depopulation in Red Fluorescent Proteins Developed Using Fluorescence Lifetime Selections**

by

**Srijit Mukherjee**

B.S. and M.S., Indian Institute of Science Education and Research Mohali, 2016

A thesis submitted to the Faculty of the Graduate school of the University of  
Colorado in partial fulfillment of the requirements for the degree of Doctor of  
Philosophy

Department of Chemistry and JILA

2022

## **Committee Members**

Ralph Jimenez (Professor, Department of Chemistry and JILA)

Amy E. Palmer (Professor, Department of Biochemistry and BioFrontiers Inst.)

J. Mathias Weber (Professor, Department of Chemistry and JILA)

Gordana Dukovic (Professor, Department of Chemistry and RASEI)

Joel Kralj (Asst. Professor, Department of MCDB and BioFrontiers Inst.)

Mukherjee, Srijit (Ph.D., Chemistry)

Spectroscopic Evaluation of Excited State Depopulation in Red Fluorescent Proteins Developed Using Fluorescence Lifetime Selections

Thesis directed by Professor (Adjoint) of Chemistry and JILA Fellow Ralph Jimenez

Abstract:

Fluorescent Proteins are intriguing molecules that are ubiquitous tools for biological research. This thesis investigates the relationship between engineering improved fluorescent proteins and fundamental physical chemistry principles that drive the molecular design of brighter and more photostable variants of existing fluorescent proteins. This work primarily focuses on designing brighter variants of the red-fluorescent proteins, FusionRed and mCherry, and the consequent physical investigations that lead to increments in brightness. To achieve this, droplet-based and single stream microfluidic sorters that discriminate members of a cell-based mutagenesis library of variants on their excited state lifetime and fluorescence intensity were developed and utilized. These lifetime-based selections led to variants with greater than 3-fold higher values of fluorescence quantum yield in comparison to their progenitors. Finally, the mechanisms and pathways that lead to excited state depopulation after photoexcitation, which govern observables like fluorescence quantum yield, fluorescence brightness, reversible and permanent photobleaching, and blinking at the single molecule level were investigated. This work highlights the progress and the many knowledge gaps in understanding the interwoven relationships of observed photophysics in fluorescent proteins, and suggests ways in which one might further examine the physical basis of brightness and photostability in these molecules.

## **Dedication**

To my teachers, from my parents to professors – thank you for all I know.

## Acknowledgements

This thesis is dedicated to all of my teachers who have influenced my life, which makes my parents the first and foremost to acknowledge. Their role in supporting me and making me the person I am is beyond instrumental. I have to acknowledge the key role of my uncle, Professor Partha Sarathi Sanyal, who is an icon I have always looked up to. Professor K. S. Viswanathan deserves special thanks for allowing a curious undergrad floating through the hallways of Academic Block-1 in IISER-Mohali to actively contribute in picturing molecules dancing to the melody of infrared and visible light.

I would like to express my sincerest gratitude to my advisor, Professor Ralph Jimenez for imparting his scientific expertise and wisdom. His perfect balance of independence and accessibility has been critical to my progress. Outside mentoring me in methods to perform cutting-edge experiments, his role in making me an effective science communicator has been significant. I have to thank Professor Amy Palmer for being the most attentive co-mentor one can ever have in graduate school. Her guidance has made me a better scientist. I thank the members of the Jimenez Lab across the years. Special mention to Professor ShengTing Hung, who as a postdoctoral mentor and a friend was the best colleague an incoming graduate student could ever ask for. I have to thank Dr. Premashis Manna, on whose foundational work I could build on and explore interesting directions. Thanks to Dr. Nancy Douglas for being the mutagenesis magician of our lab and imparting me each and every bit of biological insight that I possess. I need to thank fellow graduate students at the lab - Dr. Samantha Allen (and Rob Wells) and Kristen Parzuchowski for becoming the best set of work friends I have ever had. Thank you for being there and listening to cries about failed experiments, corrupted data and other shenanigans of being a graduate student. I sincerely thank the other lab members who constantly provided scientific and emotional support across the years – Ryan Wilson, Dr. Richard Erickson, Pia Friis, Dr. Brett Fiedler, Dr. Joshua Slocum and my fellow soccer fan, Dr. Alex Mikhaylov. I thank the wonderful and scientifically motivated undergrads who I have interacted with and mentored. Thank you, Emma Simmerman, Connor Thomas and Annika Ekrem, you have always kept the lab buzzing and my basic pedagogic skills sharp. I thank the members of the Palmer Lab for their help during the countless times I have walked into the cell-culture room with a missing reagent or a pipette (and obviously the best group meeting snacks). I thank each and every one at JILA – with special mentions to Hans Green, Felix



Vietmeyer, Curtis Biemborn and J.R. Raith, whom I have countlessly bugged for every little problem I have ever faced with instrumentation, electronics or computers!

I have to thank the best set of friends who packed their bags from IISER Mohali and showed up in different countries to embrace the love of science. Thank you Saumya Gupta, Vivek Sagar, Nakul Teke and Atul Arora for keeping me sane in the U.S. I thank Yosman Bapat-Dhar, Shwetha Srinivasan and Evelyn Abraham who kept the messages rolling and making me smile every now and then. I need to thank the trio I have grown up with – Tanmoy, Biswadip and Haymanti, it is amazing that we’ve stuck for twenty years. I thank my friends in India, Abhinay and Ashish who have kept in touch and constantly supported me. My wonderful cohort – thanks Nate, Jenny (and Frank), Nicholas, Jona, Marium, Keaten, Tri, Ashutosh, Wyatt (and Livi), Zach (and Katie) and Callum. I thank people from Kolkata in Boulder, who kept the Bengali in me alive – thanks Shambojit (and Nikita), Souradeep, Sannidhya, Sobhan, Sandip, Aritra, and Subhayan! I thank my extended family who have constantly provided me encouragement and feedback, with special mention of my cousin Avishek Chandra who made moving to a new country halfway across the globe an easy transition.

And finally, this thesis would have been incomplete without the constant motivation, love and encouragement from Swarupa. The trust and confidence she instills in me keeps me chugging forward!

# Table of contents

<b>Chapter 1 – Introduction</b>	<b>1– 32</b>
• 1.1 Publication Note	1
• 1.2 Introduction	1– 3
• 1.3 Brightness	3– 22
• 1.4 Photostability	23– 29
• 1.5 Conclusions	29– 30
• 1.6 Thesis Aims and Objectives	30– 32
 <b>Chapter 2 – Instrumentation: Development of a high-throughput microfluidic sorter</b>	 <b>33– 49</b>
• 2.1 Publication Note	33
• 2.2 Introduction	33– 35
• 2.3 Experimental	36– 39
• 2.4 Results and Discussions	39– 48
• 2.5 Conclusions	49
 <b>Chapter 3 – Engineering a brighter version of FusionRed using lifetime flow cytometry and structure guided mutations</b>	 <b>50– 79</b>
• 3.1 Publication Note	50
• 3.2 Introduction	50– 52
• 3.3 Experimental	52– 63
• 3.4 Results	64– 74
• 3.5 Discussion	75– 78

•	3.6	Conclusions	79
<b>Chapter 4 – Characterizing dark state kinetics from ensemble and single molecule fluorescence measurements in FusionRed variants</b>			<b>80– 100</b>
•	4.1	Publication Note	80
•	4.2	Introduction	80– 84
•	4.3	Methods	84– 90
•	4.4	Results	90– 95
•	4.5	Discussion	96– 99
•	4.6	Conclusions	100
<b>Chapter 5 – Characterizing the pathways of non-radiative decay in the lifetime evolution of a brighter variant of mCherry</b>			<b>101– 119</b>
•	5.1	Publication Note	101
•	5.2	Introduction	101– 104
•	5.3	Methods and materials	104– 105
•	5.4	Results	105– 112
•	5.5	Discussion	112– 118
•	5.6	Conclusions	118– 119

## Chapter 6 – Brightening FusionRed-MQV using lifetime selections 120–133

- 6.1 Introduction 120
- 6.2 Results and discussions 121– 131
- 6.3 Conclusions 132– 133

## Chapter 7 – Future directions 134– 141

•	7.1	Development of a sorting platform on dark state conversion	134– 138
•	7.2	Improving maturation of FPs	139– 141

## Bibliography 142–182

## Appendices 183– 312

- Appendix 1 183– 206
- Appendix 2 207– 231
- Appendix 3 232– 249
- Appendix 4 250– 272
- Appendix 5 273– 307
- References 307– 312

<b>List of Figures</b>		
<b>Figure</b>	<b>Title</b>	<b>Page Number</b>
1.1	Structure of a Fluorescent Protein and its chromophore.	3
1.2	A schematic representation of the process for engineering genetically encodable fluorescent biomarkers employed in our group.	6
1.3	Photophysical data from 89 published FPs was obtained from FPBase.	10–11
1.4	Expected trends from the weak-coupling low temperature limit of Englman-Jortner theory.	15
1.5	Excited-state spectral and structural dynamics in RFPs studied using ultrafast spectroscopy.	20
1.6	A graphical summary of results from a TD-DFT study.	22
2.1	Schematic layout of the electronics used in this sorter.	37
2.2	Droplet generation and size.	37
2.3	Lifetime screens of FPs.	40
2.4	Fluorescence lifetime versus brightness scatter plots of mixed cells before and after sorting.	43
2.5	The efficiency of enrichment with various initial fraction of target analyte.	44
2.6	Fluorescence lifetime versus brightness scatter plots of rare mScarlet enrichment.	47
3.1	Lifetime screens for the site directed libraries of FusionRed.	64– 65
3.2	Site-directed mutagenesis based on structural data.	67
3.3	Photophysical properties of the FR family of FPs.	69– 70
3.4	Cellular brightness assays.	71
3.5	Localization Assays.	74
3.6	The overall molecular schematic describing the postulated effects of the substitutions in FR-MQV.	78
4.1	A three-state model of the photophysics of a fluorophore central to this study.	84
4.2	Analysis workflow.	89

4.3	Results from single molecule blinking experiments.	92
4.4	Variation of rate constants.	93
4.5	Photobleaching fits from a three-state model.	95
4.6	Relating blinking to a cis-trans isomerization of the FR chromophore.	99
5.1	Directed evolution of mCherry resulted in variants with blue-shifts and increased lifetime.	106
5.2	The non-radiative rate obtained from the excited state lifetime and fluorescence quantum yield plotted as a function of fluorescence lifetime.	111
5.3	Insights from the crystal structure of mCherry.	114
6.1	Lifetime-brightness screens of the progenitor FusionRed-MQV (WT) with before and after FACS enrichment populations.	121– 122
6.2	The crystal structures indicating the potential interactions of positions 224 and 71.	125
6.3	Photophysical characterization of the brightest clones from site-directed mutagenesis of FR-MQV and FR-HQC.	129
6.4	The brightness of FusionRed variants generated from FR-MQV and FR-HQC in mammalian cells using FACS.	131
7.1	The simulated phase signatures from RFPs.	135– 136
7.2	Off chip phase measurements	136– 138
7.3	Tentative design of microfluidic sorting device to be used for a dark state screening platform.	138
7.4	Elucidating the role of the C-terminus tail in RFPs.	141

## List of Tables

Table	Title	Page Number
3.1	Summary of in vitro photophysical properties of FR mutants developed in this study.	68
3.2	Summary of cellular and molecular brightness for FR variants.	73
4.1	Values of $\tau_{ON}$ and $\tau_{OFF}$ .	91
4.2	Estimation of $k_{DSC}$ and $k_{GSR}$ .	92
4.3	Estimated $k_{DSC}$ values from fitting ensemble traces.	95
5.1	Spectral and photophysical properties of the mCherry variants.	107
5.2	Fit parameters for fluorescence lifetime measurements.	108
5.3	Calculated rate constants of excited-state population decay.	112
6.1	Mean brightness and lifetime from the screen plots provided in Figure 6.3	122
6.2	Lifetime measurements for the FusionRed-L175X clones	123
6.3	Spectral details for the FusionRed-L175X clones	123
6.4	Lifetime measurements for the FusionRed-M42Q L175X <sub>1</sub> C159X <sub>2</sub> clones	124
6.5	Photophysical characterization for the FusionRed-M42Q L175X <sub>1</sub> C159X <sub>2</sub> clones	125
6.6	Lifetime based screens for the FusionRed-MQV/HQC A224X <sub>1</sub> I71X <sub>2</sub> clones	126
6.7	Spectral properties for the high lifetime FusionRed-MQV/HQC A224X <sub>1</sub> I71X <sub>2</sub> clones.	126– 127
6.8	Spectral properties for the high lifetime FusionRed-MQV/HQC A224X <sub>1</sub> I71X <sub>2</sub> clones.	127
6.9	Photobleaching behavior for the high lifetime FusionRed-MQV/HQC A224X <sub>1</sub> I71X <sub>2</sub> clones.	130

## Table of contents for the Appendix

- **Appendix 1:** Addendum to Chapter 2

Section	Title	Page Numbers
1	Optical layout and microfluidic device	183-185
2	Instrument Operation Parameters	186
3	Cell culture and sample preparation	186– 187
4	Data acquisition and signal processing	187– 196
5	Performance of the sorter	196– 198
6	Enriching rare events	198– 201
7	Directed evolution of an RFP library	202– 206

- **Appendix 2:** Addendum to Chapter 3

Section	Title	Page Numbers
1	Amino acid sequence information for FusionRed variants	207
2	Amino acid sequence results from the site-directed library	208
3	Extinction coefficient and quantum yield estimation	209– 212
4	Additional photophysical and biochemical characterization	213– 225
5	Cellular Assays: brightness, maturation and cytotoxicity	226– 230
6	Additional structural information	230– 231
7	FR Evolution Table	231

- **Appendix 3:** Addendum to Chapter 4

Section	Title	Page Number
1	Experimental workflow	232
2	Excitation rate calculations	233
3	Single molecule data analysis addendum	234
4	Ensemble photobleaching	235– 236



5	Simulation results	237– 245
6	Additional fitting results	246– 248
7	Theoretical estimation of the lowest and highest number of photons/frames	248– 249
8	Surface charge on FusionRed	249

- **Appendix 4:** Addendum to Chapter 5

Section	Title	Page Number
1	Mutagenesis Protocols	250– 252
2	Mutagenesis and Directed Evolution of mCherry	252– 257
3	Sequence Alignment of the RFP variants	257
4	Additional in cellulo assessments	258– 261
5	Additional details for in vitro photophysical assessments	262– 266
6	Variation of the radiative rate constant	267
7	Modelling non-radiative rate using the Englman-Jortner Low Temperature/Weak Coupling Formalism	268– 272

- **Appendix 5:** Instrument operation, simulation and data analysis codes

Section	Title	Page Number
1	Code for microfluidic screen dot-plots	273– 275
2	Droplet-sorter operation controls	276– 279
3	Simulation code representing the effects of scattering on phase shift	279– 280
4	Simulation code representing the signal of a droplet in flow	280
5	Codes to analyze in vitro photophysical data	280– 284
6	Code to simulate role of electrostatics on oscillator strength	284
7	Codes for analyzing photobleaching data	284– 287
8	Codes to simulate dark state kinetics in time and frequency domains using numerical solutions to the three-state model	287– 288

9	Codes to simulate dark state kinetics in time and frequency domains using analytical solutions to the three-state model	289– 291
10	Simulation of single molecule kinetics using Monte-Carlo methods	291– 294
11	Detection code for single molecule trajectories from video (TIFF series) file	295– 299
12	Code for binarizing single molecule trajectories	300– 307

<b>List of Appendix Figures</b>		
<b>Figure</b>	<b>Title</b>	<b>Page Number</b>
A1.1	Optical diagram of the microfluidic sorter	183
A1.2	Microfluidic chip design	185
A1.3	Simulated and observed signals of a filled droplet	188
A1.4	Pair matching events to estimate the transit time and the flow speed of the system.	189
A1.5	Comparing signals from filled and empty droplets.	191
A1.6	Screening results of a mix of mCherry, FusionRed-M and Kriek (K2C).	194
A1.7	Simulation of phase shifts obtained from a convolution of scattered and fluorescence signals.	195
A1.8	Poisson distribution for $\lambda = 1$ to $\lambda = 10$ .	199
A1.9	Enrichment efficiency converges as the number of rare events increases.	199
A1.10	mScarlet library sorts with the instrument.	204
A2.1	Gradient (4–20%) SDS-PAGE of FR and FR-MQV.	210
A2.2	Quantum yield measurements.	211
A2.3	pKa measurements.	213
A2.4	Photobleaching measurements.	216
A2.5	Bacterial brightness assay.	217
A2.6	Absorption and excitation spectra.	218
A2.7	Trp280 nm peak for FusionRed variants.	220
A2.8	pH titrations.	221
A2.9	Emission spectra for the FusionRed and mKate variants.	223
A2.10	Excitation dependent emission spectra.	224-225
A2.11	Cytotoxicity assay	228
A2.12	Chromophore maturation kinetics.	229
A2.13	Structural representation of the FR crystal structure.	230

A2.14	Evolution table.	231
A3.1	Schematic of the experimental workflow from sample preparation to data analysis.	232
A3.2	Spatial corrections.	234
A3.3	Intensity variation.	234
A3.4	Ensemble photobleaching kinetics.	235
A3.5	Bi-exponential nature of fluorescence decay from bacteria expressing FusionRed and FusionRed-MQ.	236
A3.6	Schematic representing the algorithm used to simulate single molecule behavior.	237
A3.7	Simulation heatmaps for on and off time with varying rate constants.	238
A3.8	FusionRed blinking simulation (100 ms, 6000 frames)	239
A3.9	FusionRed blinking simulation (50 ms, 12000 frames)	240
A3.10	FusionRed blinking simulation (50 ms, 6000 frames)	241
A3.11	FusionRed blinking simulation (300 ms, 1000 frames)	242
A3.12	Ensemble trends from Monte-Carlo simulations	243
A3.13	Variation in ensemble fluorescence decay traces with respect to $k_{DSC}$ , $k_{ex}$ and $k_{GSR}$ .	244
A3.14	Population based numerical simulation of a 3-state model.	245
A3.15	Quality of fit for varying $k_{DSC}$ and $k_{GSR}$ unbound.	247
A3.16	Keeping $k_{GSR}$ fixed while allowing the fitting algorithm to fit the $k_{DSC}$ .	248
A3.17	Location of positively charged residues on the barrel of FusionRed.	249
A4.1	EP-PCR library of C9 and C12 variants.	253
A4.2	Summary of mutagenesis efforts on the mCherry family to achieve lifetime evolution.	255
A4.3	Role of spatially distant substitutions on the photophysical properties.	256
A4.4	Lifetime & yeast-cell co-evolution trajectory.	258
A4.5	Cellular brightness in mammalian cells.	259
A4.6	Photobleaching trends in E. coli.	260

A4.7	Comparing photostability under excitation rate normalized conditions in <i>E. coli</i> .	261
A4.8	Normalized absorption and emission spectra.	262
A4.9	Fluorescence anisotropy decay and rotational time-constants ( $\tau_r$ ) of the variants.	263
A4.10	Fluorescence decay and the average lifetime ( $\tau$ ) of the variants.	264
A4.11	Fluorescence quantum yield ( $\phi$ ) of the variants.	265
A4.12	Analysis of the radiative rate constants for mCherry variants.	267
A4.13	Fit of the total non-radiative rate constant with the energy gap.	268
A4.14	Schematic representations of relevant observables.	269
A4.15	Absorbance and fluorescence spectra represented on the wavenumber scale.	270
A4.16	The numerator and denominator functions for in the Eq S2.	271
A5.1	The front panel of the NI LabView interface for the droplet sorter operation.	276
A5.2	Additional panels of the NI LabView interface for the droplet sorter operation.	277
A5.3	The block diagram of the sorter GUI displayed in Tables A5.1 and A5.2	278-279

## List of Appendix Tables

<b>Table</b>	<b>Title</b>	<b>Page Number</b>
A1.1	Screening results of sorting efficiency.	196
A1.2	Fluorescence lifetime measured in TCSPC and the microfluidic droplet sorter.	197
A1.3	Screening results of sorting efficiency.	198
A1.4	The enrichment of rare mScarlet from a mixture with EGFP screened at a rate of 2 kHz.	201
A1.5	The enrichment of rare mScarlet from a mixture with mCherry screened at a rate of 2.5 kHz.	201
A1.6	The sequence of mScarlet-I mutants in the mutated positions.	205
A1.7	Photophysical properties of mScarlet-I mutants.	206
A2.2	Sequence information for the FR family with respect to the parental clones from the eqFP578 family	207
A2.2	Sequence and screening results of the FR site-directed (FSD) library clones.	208
A2.3	Measured mean quantum yield and standard deviation errors	212
A2.4	The pKa values measured in this study versus the values reported for known RFPs in the literature.	214
A2.5	TCSPC-based lifetime measurements using pure protein samples.	215
A2.6	Spectral characteristics of the FR family of proteins investigated in this study.	219
A2.7	Mean brightness from FACS measurements with standard deviation error from multiple biological replicates.	226
A2.8	Mean brightness from confocal microscopy measurements with standard deviation error from the number of cells indicated in each dish.	227
A2.9	Number of cells analyzed on Day 2 and Day 6 in the GFP and the RFP channel using FACS.	228

A2.10	Measured versus the reported maturation times ( $t_{50}$ ) for the RFPs investigated in this study.	230
A3.1	Power measurements	233
A3.2	Fitting details for the biexponential decay traces of FusionRed and FusionRed-MQ.	236
A4.1	Sequence alignment of the mCherry variants analyzed in this study.	257
A4.2	Spectral properties of the RFP variants.	262
A4.3	Difference in energies of the peaks of the 0-0 and 0-1 vibronic bands and the relative intensities.	272

<b>List of common Abbreviations and Symbols</b>	
FP	Fluorescent Protein
RFP	Red Fluorescent Protein
GFP	Green Fluorescent Protein
PDB	Protein Data Bank
CFP	Cyan Fluorescent Protein
EP-PCR	Error-Prone Polymerase Chain Reaction
OSER	Organized Smooth Endoplasmic Reticular assay
B-PER	Bacterial Protein Extraction Reagent
TCSPC	Time Correlated Single Photon Counting
FACS	Fluorescence Activated Cell Sorting
FLIM	Fluorescence Lifetime Imaging Microscopy
FRET	Förster Resonance Energy Transfer
QM/MM	Quantum Mechanics/ Molecular Mechanics
MD	Molecular Dynamics
FAPA	Fluorescence Anomalous Phase Advance
FPGA	Field Programmable Gate Array
PMT	Photomultiplier Tube
ADC	Analog to Digital Convertor
DEP	Dielectrophoretic (in context of electrodes)
TIRF	Total Internal Fluorescence Microscopy
SOFI	Super-resolution Optical Fluctuation Imaging
SMLM	Single Molecule Localization Microscopy
CMOS	Complementary Metal Oxide Semiconductor camera detector
EMCCD	Electron Multiplying Charge Coupled Device camera detector
FWHM	Full Width at Half Maximum
SNR	Signal to Noise Ratio
OD	Optical Density
RPM	Rotations Per Minute
FR	FusionRed



mCh	mCherry
EGFP	Enhanced Green Fluorescent Protein
K2C	Kreik
Ex	Excitation
IC	Internal Conversion
ISC	Intersystem Crossing
DSC	Dark State Conversion
GSR	Ground State Recovery
NR/ Non-Rad	Non-radiative pathways of excited state depopulation
R/ Rad	Radiative pathways of excited state depopulation
SS	Stokes shift
SM	Spectral moment
EJ	Englman-Jortner
$\lambda$	Wavelength (unless specified otherwise, e.g., in the context of droplet loading)
$\phi$ / $\varphi$ / QY	Fluorescence quantum yield (unless specified otherwise)
$\varepsilon$ / EC	Extinction coefficient
$\tau$ / LT	Fluorescence lifetime (unless specified otherwise) Subscripts FD and TD denote time or frequency domain measurements.
k	Rate constants. Subscripts denote the nature of the process under consideration. e.g, excitation rate constant is denoted as $k_{ex}$ .
$\eta$	Refractive index of the media
$\sigma$	Absorption cross-section
$\nu$	Frequency
$N_A$	Avogadro's Constant
c	Speed of light
h	Planck's constant

# Chapter 1

## Introduction

### 1.1. Publication Note

Parts of this chapter have been adapted from the article “*Photophysical Engineering of Fluorescent Proteins: Accomplishments and Challenges of Physical Chemistry Strategies.*” Mukherjee, S.; Jimenez, R. J. Phys. Chem. B 2022. <https://doi.org/10.1021/acs.jpcb.1c05629>.

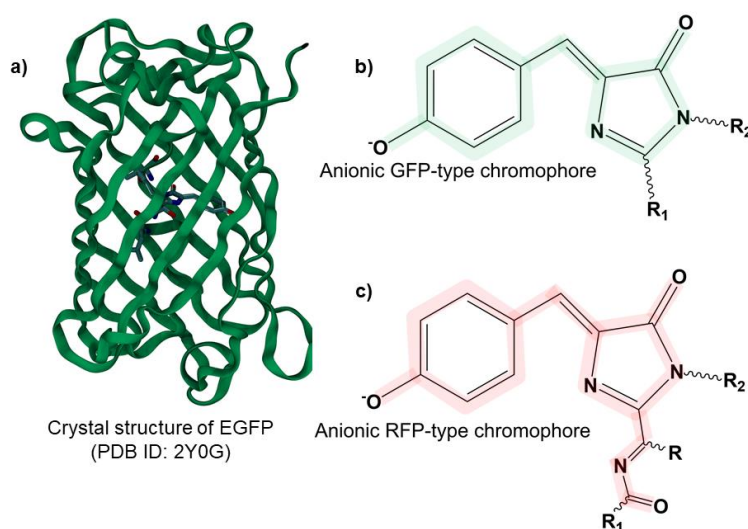
### 1.2. Introduction

Since the first isolation of naturally occurring fluorescent proteins (FPs) from marine organisms several decades ago, subsequent developments in protein engineering have produced a large family of fluorophores spanning the entire visible wavelength spectrum. [1, 2] Applications of FPs in bioimaging and sensing include multi-color microscopy, [3] Förster resonance energy transfer (FRET)-based tools, [4] fluorescence lifetime imaging microscopy (FLIM), [5] voltage sensing, [6] biosensing, [7] catalytic activity monitoring, [8] aggregation studies, [9] and non-biophysical applications such as bio-phosphors for LED lighting. [10] Additionally, the engineering of FPs exhibiting reversible and photo-activated fluorescence and photo-switching between fluorescence bands has enabled imaging with spatial resolution beyond Abbe’s diffraction limit. [11] Despite and because of this widespread use, development of new FPs continues. For example, with the increasing interest in deep-tissue imaging of live animals, significant efforts have focused on the discovery and development of brighter red FPs (RFPs), as longer excitation and emission wavelengths generally provide lower scattering and increased penetration-depth.

FP engineering is inextricably tied to consideration of protein structure. Fortunately, hundreds of X-ray crystallographic structures of FPs at atomic resolution are available (e.g., Figure 1.1). In the canonical green fluorescent protein superfamily, the chromophore—which is comparable in size to synthetic small-molecule fluorophores (~1 nm) – is contained inside a  $\beta$ -barrel with an internal  $\alpha$ -helix. [12] Several chromophore variants can be autocatalytically formed from the reaction of O<sub>2</sub> with a tripeptide in this helix. One common structure typically comprises p-hydroxyphenyl and imidazolinone moieties connected by a methylidyne bridge (Figure 1.1). The electronic

conjugation across this hydrolyzed tripeptide results in a chromophore with a  $\pi \rightarrow \pi^*$  electronic transition excitable at visible wavelengths. The chromophore participates in numerous interactions with amino-acids and the solvent. These short-range interactions (e.g., hydrogen bonding) or long-range effects (e.g., electric fields) tune the energetics and control the electronic structure of the chromophore. Accordingly, amino-acid substitutions, sometimes surprisingly distant from the chromophore, perturb its electronic structure, influence the energetics of its electronic transition, and cause changes in its conformation and the nuclear degrees of freedom coupled to it. Additionally, the overall protein structure governs the movement of diffusing species such as water or O<sub>2</sub>, which can also lead to alterations in chromophore properties. In general, it is difficult to separate the impacts of numerous variables that non-additively contribute to the photophysics. All these factors should be taken into consideration for explaining and tuning properties such as brightness and photostability. [13-16]

In addition to being useful for imaging, FPs provide an incredible molecular framework for investigating diverse photophysical, spectroscopic and dynamical phenomena such as solvation dynamics, light harvesting, photo-induced excited state intramolecular proton and electron transfer, photo-transformations such as photoisomerization, and other radiationless transitions. [17-21] One can examine them with biochemical or physical methods by combining spectroscopic measurement tools including steady-state and time-resolved optical and vibrational spectroscopy, with structural approaches such as nuclear magnetic resonance (NMR) and both static and time-resolved X-ray crystallography. Here, we will discuss the development of FPs, focusing on topics that are of interest to physical chemistry. Though we will focus on the GFP super-family, the topics discussed here apply to FPs of other lineages, such as those containing tetrapyrrole chromophores. [22-24] We further center discussion on brightness and photostability, which are arguably the most rudimentary fluorophore properties. As described below, it has been a difficult task to optimize them in tandem. [16] Studies of the principles behind brightness and photostability have been pursued in our lab for over a decade, and they provide a platform for discussing far more general concepts. For discussion of more complex functionalities, such as photoactivation and photoswitching, we refer readers to other reviews. [25, 26] In describing some of the biggest successes, notable failures, and remaining challenges, we also consider the role of physical principles in guiding design strategies.



**Figure 1.1. Structure of a Fluorescent Protein and its chromophore.** (a.) The crystal structure for enhanced GFP (EGFP; PDB ID: 2Y0G) showing the  $\beta$ -barrel encompassing the internal helix and the chromophore, for which oxygen atoms are indicated in red, nitrogen in blue and carbon in gray. (b.) A ground state structure of the anionic GFP-type p-hydroxyphenyl-imidazolinone chromophore (c.) A ground state structure of the anionic RFP p-hydroxyphenyl-imidazolinone chromophore showing the extension of the electronic conjugation through an acylimine moiety. Green and red highlighting indicate the bonds participating in electronic conjugation (resonance) across the chromophore. Structures were generated using VMD and ChemDraw software suites. [27, 28]

## 1.3. Brightness

### 1.3a. Approaches for improving brightness

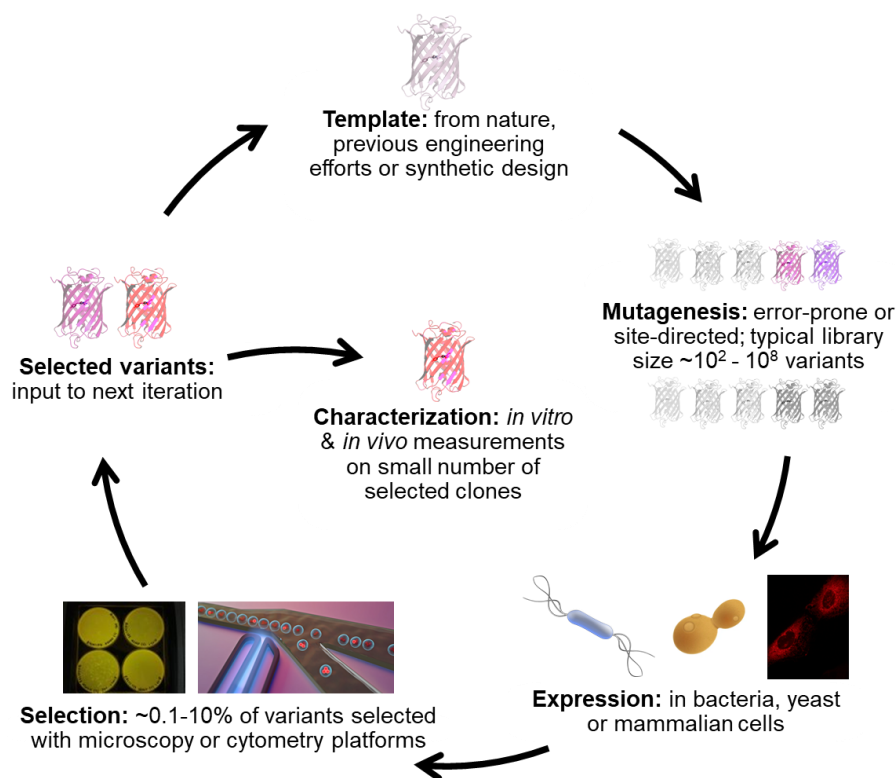
Fluorescence brightness is a molecular property defined as the product of the molar extinction coefficient at maximum absorption ( $\epsilon_{\max}$ ) and the fluorescence quantum yield ( $\phi$ ). However, when FPs are imaged in cells, the observed cellular brightness is a consequence of additional factors such as protein folding, kinetics of chromophore maturation, translational efficiency, expression level (i.e., concentration), environmental sensitivity to pH, nature of the fusion-protein construct and chemical environment of the cellular compartment. [29] The typical development pipeline for

improving brightness or other FP properties involves generating “libraries” (which may range in size from dozens to  $>10^7$  variants) using error-prone PCR or site-directed mutagenesis based on structural, spectroscopic and bioinformatics guidance (Figure 1.2). These libraries are then typically expressed in a host, such as bacteria, yeast, or mammalian cells that allows for screening and selection of clones with desired characteristics – such as higher brightness. For example, Fluorescence Activated Cell Sorting (FACS) and fluorescence imaging are popular tools for selecting on cellular brightness. The mutation and selection cycles are repeated to achieve “directed evolution” in terms of the selection pressure. [30] After several rounds, a small number of selected clones are expressed, and the purified FPs are individually characterized to determine values of molecular properties such as the fluorescence quantum yield and peak molar extinction coefficient. This approach has yielded substantially improved FPs such as EGFP, [31] (1.7-fold brighter than progenitor avGFP) and the mRubys, [32] (mRuby3; 1.7-fold brighter than the naturally occurring progenitor eqFP511), of which the best have peak molar extinction coefficients and fluorescence quantum yields comparable to those of small-molecule dyes. For example, mNeonGreen [33] has fluorescence quantum yield  $\sim 80\%$  and a peak molar extinction coefficient at  $\sim 116,000 \text{ M}^{-1}\text{cm}^{-1}$  (brightness of  $\sim 93$ ) with a peak absorption wavelength comparable to the recently developed, xanthene-based Janelia Fluor (JF) dye JF-503 which has a fluorescence quantum yield  $\sim 87\%$  and an  $\epsilon_{\text{max}} \sim 95,000 \text{ M}^{-1}\text{cm}^{-1}$  (brightness  $\sim 83$ ). [34]

Despite the historical success of this approach, alternative strategies for obtaining brighter FPs have been explored, particularly in recent years as the pace of improvements has slowed. For example, the choice of template has garnered increased attention. In most cases the template is either a naturally occurring FP or a commonly used FP that has undergone previous rounds of development. The notable exception is the brightest RFP as of early 2021, mScarlet, which was developed from a synthetic gene template. [35] Given this achievement, one might be tempted to think natural templates have reached their limits. However, Lambert et al. (2020) reported the discovery of AausFP1, a naturally occurring GFP from the jellyfish *Aequorea victoria*, from which the original GFP was obtained. [36] This FP has the highest ever reported value of fluorescence quantum yield (97%) and a peak molar extinction coefficient of  $\sim 170,000 \text{ M}^{-1} \text{ cm}^{-1}$ . Although the dimeric structure and the very small Stokes shift of AausFP1 may limit its use for imaging, these defects could be addressed, and this discovery suggests that nature will continue to be a source for brighter fluorophores.

Many FPs have been optimized for specific biochemical applications (for example, cellular localization or pH sensitivity) and engineering photophysically-improved variants mandates keeping close attention to their *in vivo* attributes. For instance, Campbell and co-workers proposed that GFP-based fluorophores with the Glycine-Tyrosine-Glycine (G-Y-G) tripeptide chromophore might have attained a local maximum of molecular brightness. Consequently, they focused on improving properties that can increase cellular brightness, such as solubility, translation efficiency, protein folding and chromophore maturation. With this approach, they developed mGreenLantern, which is 6-fold brighter in mammalian cells than EGFP. This improvement in cellular brightness came without an appreciable change in the molecular brightness from its precursor FP - Clover. [37] This study demonstrates that increasing cellular brightness by generating a fast-maturing FP with high copy number can be beneficial in some applications, such as imaging of neurons. However, we found that imaging of small subcellular structures, such as the Golgi apparatus, sometimes benefits more from controlled expression of FPs with high molecular brightness. [38] This suggests that improvements of both molecular and cellular brightness should be pursued in tandem.

The selection step (Figure 1.2) is of particular interest to physical chemists because spectroscopic measurements, such as fluorescence lifetime ( $\tau$ ), can be incorporated on platforms that allow high-throughput screening of cells, such as microfluidic or microscopy-based systems. [30, 35, 38, 39] Relationships between sequence, structure and any spectroscopically-accessible property can be investigated by this approach on  $10^7$  or more variants per day. Fluorescence lifetime-based screening in particular has proven useful because lifetime is independent of concentration, and it is correlated with higher molecular brightness (with caveats, see below). Lifetime-based selection on a microscopy platform led to the development of mScarlet from the dim synthetic template mRed7. [35] Our lab has pioneered the integration of microfluidic flow cytometry with photobleaching, photoswitching, and fluorescence lifetime selection. [40-44] For example, we employed a cell sorter based on fluorescence lifetime to variants of FusionRed with 3-fold higher molecular brightness (discussed later in the text). [38]



**Figure 1.2. A schematic representation of the process for engineering genetically encodable fluorescent biomarkers employed in our group.**

Few efforts have directly aimed to alter the absorption properties by tuning the oscillator strength of electronic transitions in FPs, but initial results in this direction are promising. [45, 46] In a recent study, Myšková and co-workers determined transition dipole moments of absorption (xTDM) and emission (mTDM) of several well-studied, bright FPs. [47] Their approach combined cryogenic x-ray crystallography with room-temperature polarized optical transmission and fluorescence measurements to determine the orientation of these dipole moments. By characterizing the propensity of FPs to crystallize into specific space groups with certain molecular orientations, the authors were able to interpret the optical measurements with two well-defined assumptions. They assumed the TDM of the chromophore was in plane of the two aromatic rings and furthermore that there is a cosine squared relationship of the TDM with the polarized absorption or fluorescence intensity. Information from such studies can inform development of FPs for applications that are sensitive to the directions of the transition dipoles, such as FRET. Another study based on

considering the transition dipole moments was reported by Molina et al. (2020) who developed a high-throughput, fluorescence microscope-based screening device (GIZMO) for screening bacterial cells by two-photon excited fluorescence. [48] The two-photon absorption cross-section can be used to examine the radiative rate and electric fields around the chromophore (the effects of internal electric fields on FP photophysics are discussed below). GIZMO is capable of screening  $10^4$  bacterial cells in ~7 hours, demonstrating great potential for this type of selection in the directed evolution of FPs.

Unconventional selection strategies that do not directly measure brightness have also been explored to develop brighter FPs. For example, many FPs exhibit two-state (on/off) “blinking” of their fluorescence intensities at the single molecule level. These transitions occur when the chromophore has access to non-fluorescent excited electronic states with microsecond to millisecond lifetimes. This “dark state conversion” (also sometimes referred to as “reversible photobleaching”) leads to rapid blinking, and the depopulation of the ground state can result in an apparent reduction in brightness. This connection between blinking, dark-state conversion, and brightness has been exploited. In particular, FLINC (Fluorescence fLuctuation INcrease by Contact) – an imaging technique with threefold higher resolution than the diffraction limit, is based on the observation that the blinking frequency of TagRFP-T varies ~25% as a function of the length of a linker peptide altering its spatial proximity to the non-fluorescent Dronpa FP. [49] Charged residues on the  $\beta$ -barrel participate in electrostatic interactions that control the dark state conversion rates. Consequently, libraries were developed by targeting positions with externally facing side chains (acidic: D159, D196 and basic: R157, R198) and selections on reduced rate of dark-state conversion through FLINC were successful for developing SuperTag-RFP, which is 2-fold brighter than TagRFP-T. [50]

The properties of new variants must be characterized carefully to avoid potential misapplications, e.g., in imaging modalities for which the photophysical properties are unsuitable. For example, population transfer to dark states can be a bottleneck for continuous excitation, especially at high irradiances. Dark states typically are related to chromophore conformational changes or inter-system crossing to the triplet state, and lead to trapping on timescales several orders of magnitude longer than an excitation-emission cycle. [51,52] For detailed discussions of this topic, we refer readers to reviews on dark state conversion and related photo-transformations such as



photoswitching. [16, 53] On a practical note, dark state conversion, which is distinct from saturation of the optical transition in a 2-level system, can lead to inaccuracies in measurements of fluorescence quantum yield, as shown by Ruhlandt and coworkers for photo-switching FPs. [54] Partially in response to this complication, Prangma and co-workers developed methods to accurately determine the fluorescence quantum yield by tuning the local photon density of states near a metal surface. [13] Looking beyond the influences on brightness, investigations of dark-state conversion and engineering of fluorescence blinking are valuable for probe development in super-resolution microscopy. It remains challenging to measure the rates of dark-state conversion, ground-state recovery, and photobleaching which, depending on the FP, may vary by many orders of magnitude (and may overlap in timescales) over the wide range of irradiances relevant to imaging, even for variants with closely related sequences. [55, 56]

### 1.3b. The photophysical basis of increased brightness

#### i. General principles

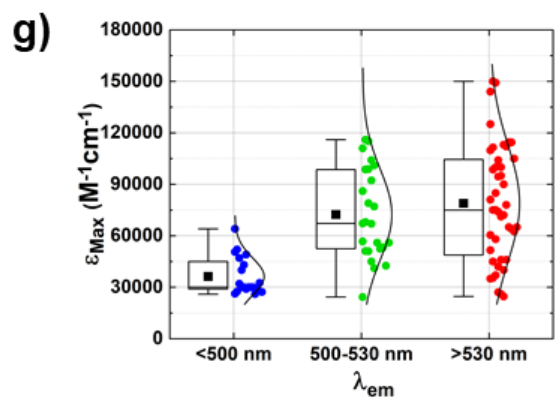
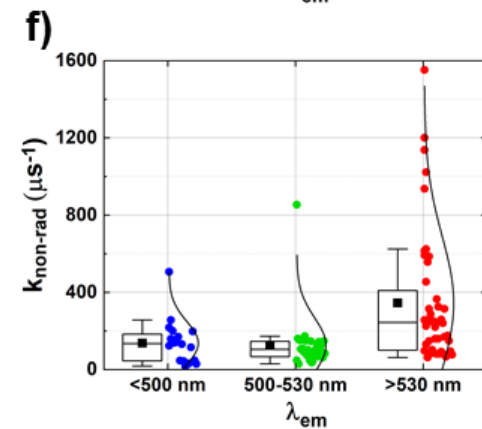
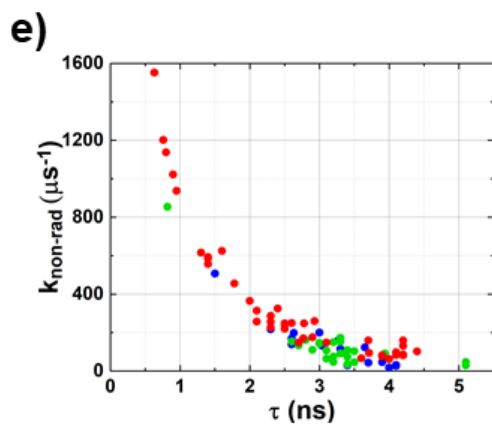
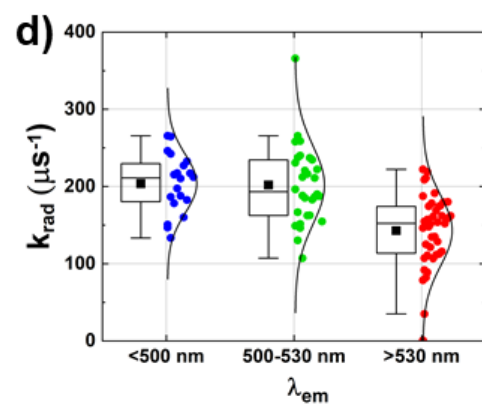
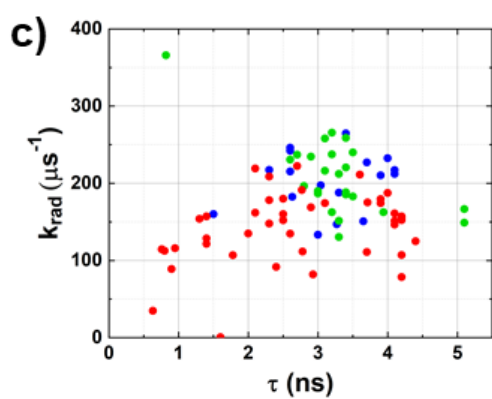
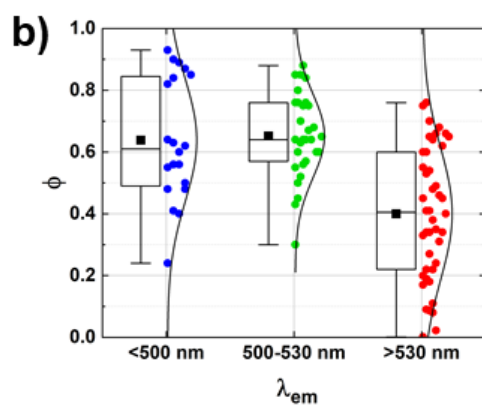
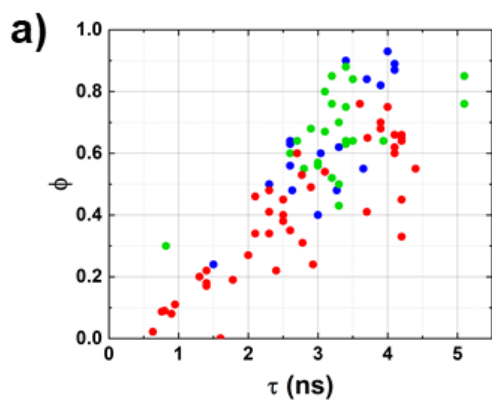
The improved molecular brightness of many newer FPs is attributable to increases in fluorescence quantum yield ( $\phi$ ) and fluorescence lifetime ( $\tau$ ). (Figure 1.3a) Consider the relationship between the fluorescence quantum yield and the radiative ( $k_{rad}$ ) and non-radiative ( $k_{non-rad}$ ) rate constants:

$$\phi = k_{rad} * \tau \quad (1)$$

$$\phi = \frac{k_{rad}}{(k_{rad} + k_{non-rad})} \quad (2)$$

In principle, a particular value of fluorescence lifetime can arise from different combinations of values of  $k_{rad}$  and  $k_{non-rad}$ . We extracted lifetime and quantum yield data from the online repository FPBase [57] to estimate radiative and non-radiative rates. We assumed the reported lifetime values from FPBase can be used to estimate the radiative and non-radiative rates, despite the use of multiple measurement techniques, with FPs in different environments (*in vitro* vs. *in vivo*), and the use of average values to represent what is typically a multiexponential decay. The analysis reveals only a 5-fold variation in the values of the radiative rate constant but a 50-fold variation in the values of the non-radiative rate constant (Figure 1.3c vs. Figure 1.3e). A similar observation for a small set of RFPs was made by Drobizhev et al. [58] Furthermore, a correlation of higher

fluorescence quantum yields with longer fluorescence lifetime is clear, though the relationship is not perfectly linear. While a higher fluorescence quantum yield most often seems to be due to a lower non-radiative rate constant, increasing the radiative rate constant also increases fluorescence quantum yield, as seen for BrUSLEE (evident as out-lying points in Figure 1.3a and 1.3c). [59] This EGFP variant exhibits a short fluorescence lifetime of 0.8 ns but a relatively high fluorescence quantum yield (30%), with a 1.6-fold increase in the radiative rate constant. Mamontova et al., achieved this by reducing the fluorescence lifetime of EGFP while maintaining its brightness. They isolated a triple mutant EGFP T65G-Y145M-F165Y, with spectral properties similar to EGFP, but with a 20% lower brightness and a 70% shorter fluorescence lifetime (cf. 2.6 ns for EGFP). Although this result shows that radiative rate engineering is possible, molecular strategies to guide the design have not been reported. In this context, we briefly consider well-known models for radiative and non-radiative transitions.



**Figure 1.3. Photophysical data from 89 published FPs was obtained from FPBase.** [57] (along with values for FusionRed-Q, -MQ and -MV variants from [38]) Points in blue indicate FPs with peak emission wavelength  $\lambda < 500$  nm, green with peak emission wavelength in the range  $500 \text{ nm} < \lambda < 530$  nm, and red with peak emission wavelength  $\lambda > 530$  nm. The boxplots indicate the distributions of values: the mean of the distribution is indicated by black solid squares; the box indicates the 25th and 75th percentiles and the whiskers the 5th and 95th percentiles of the distribution. The solid line in the box indicates the median value of the distribution. **(a-b.) Trends for the fluorescence quantum yield ( $\phi$ ):** Increases in fluorescence lifetime ( $\tau$ ) generally correlate with a higher fluorescence quantum yield. **(c-d.) Trends for radiative rate constant ( $k_{\text{rad}}$ ):** There is a  $\sim 5$ -fold variation in the radiative rate constant – where higher wavelength emission can potentially lead to lower radiative rate values. The GFP BruSLEE stands out as an exception ( $k_{\text{rad}} > 350 \text{ us}^{-1}$ ). **(e-f.) Trends for the non-radiative rate constant ( $k_{\text{non-rad}}$ ):** The non-radiative rates decrease up to  $\sim 50$ -fold with increasing lifetime. The observed values indicate the means and the standard deviations of  $k_{\text{non-rad Blue}} \sim 137 \pm 112 \text{ } \mu\text{s}^{-1}$ ;  $k_{\text{non-rad Green}} \sim 130 \pm 158 \text{ } \mu\text{s}^{-1}$ ;  $k_{\text{non-rad Red}} \sim 346 \pm 345 \text{ } \mu\text{s}^{-1}$ . **(g.) Trends for the peak extinction coefficient:** Red emitting FPs tend to exhibit larger peak extinction coefficients than green and blue counterparts based on chromophores extended by an acylimine moiety.

## ii. The radiative rate constant ( $k_{\text{rad}}$ )

The experimentally-observable parameters underlying the magnitude of radiative rate can be seen in the Strickler–Berg relationship,

$$k_{\text{rad}} = \frac{1}{\tau_{\text{rad}}} = 8\pi^2 n^2 \langle \nu_f^{-3} \rangle^{-1} \int \epsilon d \ln \nu \quad (3)$$

which shows radiative rate is linearly proportional to the squared value of the refractive index of the medium containing the chromophore ( $n$ ), the strength of the optical transition to the excited electronic state ( $\int \epsilon d \ln \nu$ ; where  $\epsilon$  is extinction coefficient integrated over the absorption band), the inverse of the mean inverse-cubed emission frequency ( $\langle \nu_f^{-3} \rangle^{-1}$ ) and  $c$  denotes the speed of light. [60] For chromophores with relatively narrow fluorescence spectra, the cube of the peak fluorescence frequency, ( $\nu_{\text{peak}}^3$ ) is approximately equal to the value  $\langle \nu_f^{-3} \rangle^{-1}$ . [58] The Strickler-

Berg equation shows that strong absorbers have higher radiative rates than weakly absorbing chromophores. As a secondary effect, there is a linear dependence of radiative rate on the cube of the fluorescence frequency. [61] A long wavelength emitter will radiate slower than a shorter-wavelength emitter with the same integrated absorption cross-section embedded in a medium of the same refractive index. Although this dependence does not represent a direct relationship to the Stokes shift, it does indicate that blue shifting the peak emission can increase the radiative rate. Equation 3 is qualitatively in agreement with Figure 1.3d. The extension of the electronic conjugation through the acylimine moiety is likely to be the largest factor in making red FPs stronger absorbers. However, Figure 1.3g shows that many green FPs absorb as strongly as red FPs. The increased extinction coefficient of the red chromophore can be partially offset by the lower frequency emission to decrease the radiative rates.

Lin *et al.* found good agreement with the Strickler-Berg equation for a set of GFP variants involving residues directly interacting with or forming the chromophore. [62] These FPs showed a nearly constant radiative rate and an inverse relationship of the peak molar extinction coefficient with the width of the absorption band. The first observation is consistent with the trend seen in Figure 1.3c and Figure 1.3e, where the variation in radiative rates is much smaller compared to the variation in non-radiative rates. Accordingly, they suggest that large changes in the fluorescence quantum yields of FPs are unlikely to be the result of variations in  $k_{\text{rad}}$ . However, the example of BrUSLEE shows that tuning the radiative rate by protein engineering is possible. [59] EGFP and BrUSLEE have nearly identical absorption and emission spectra and peak wavelengths, and the increased molar extinction coefficient ( $\epsilon_{\text{max}}$ ) and radiative rate constant of BrUSLEE follow what is expected from the Strickler-Berg relationship: a  $\sim 1.5$ -fold higher value of  $\epsilon_{\text{max}}$  results in a  $\sim 1.6$ -fold increase of the radiative rate.

A final point that should be made about the Strickler-Berg equation is that it might be inaccurate to assume the refractive indices of FPs hardly vary. We used equation 3 and the spectral data from FPbase to calculate values of  $\eta$  for blue (Cerulean), green (EGFP) and red (mScarlet) FPs, all three of which show nearly mono-exponential fluorescence lifetime decays in aqueous environments. [57] The refractive indices show the expected trend of increasing value with decreasing wavelength, but the increase is considerably larger than expected ( $\eta_{\text{mScarlet}} \sim 1.21$ ;  $\eta_{\text{EGFP}} \sim 1.25$ ;  $\eta_{\text{Cerulean}} \sim 1.49$ ). The refractive index of water only varies by 2% over this wavelength range. [63]

The sensitivity of FP fluorescence lifetime to refractive index has been explored as a sensing technique in cellular environments. [64] These observations may provide a route for design of FPs with increased radiative rate by exploring a connection with the polarizability of the amino acids comprising the  $\beta$ -barrel.

### iii. The non-radiative rate constant ( $k_{\text{non-rad}}$ )

In the absence of photochemistry, excited-state non-radiative population loss follows two major routes, internal conversion (IC), and inter-system crossing (ISC). The factors influencing the probability of the transition  $W_{12}$  can be described by Fermi's Golden Rule (Equation 4).

$$W_{12} = \frac{2\pi|M_{12}|^2}{\hbar} \rho_2 \quad (4)$$

For IC between the initial (1) and the final (2) vibronic states of the two electronic manifolds,  $M_{12}$  is the matrix-element for the electronic interaction that couples them, and  $\rho_2$  is the density of vibrational states of the final configuration. This expression also applies to the rate of ISC to a triplet state. In highly fluorescent FP chromophores, the ISC rate is much smaller than IC due to a change in the spin multiplicity. [16] Investigations of ISC and triplet states in FPs merit further attention as such states can play a pivotal role in oxidative photochemistry and photobleaching. [65]

While equation 4 accounts for the physics of the non-radiative transition rate, it does not provide useful insight for FP engineering. Reports on new fluorophores mention a dependence of the non-radiative transition rate on the  $S_1 - S_0$  transition energy (i.e. the "energy-gap law") which results from further theoretical development. [66-68] The classic theory of Englman and Jortner provides a framework for understanding non-radiative transition rates in terms of experimental observables such as the electronic energy gap ( $\Delta E$ ), the Stokes-shift (from vibrational reorganization) and vibrational frequencies. [69] Their theory makes assumptions very similar to those of Marcus theory of electron transfer, [70] namely that population transfer between weakly coupled initial and final states occurs by means of environmental fluctuations. This approach considers the high and low temperature limits of strong and weak coupling between the chromophore vibrations and electronic transitions. The low temperature limit,  $\hbar\langle\omega\rangle \gtrsim k_B T$  is defined in terms of the mean

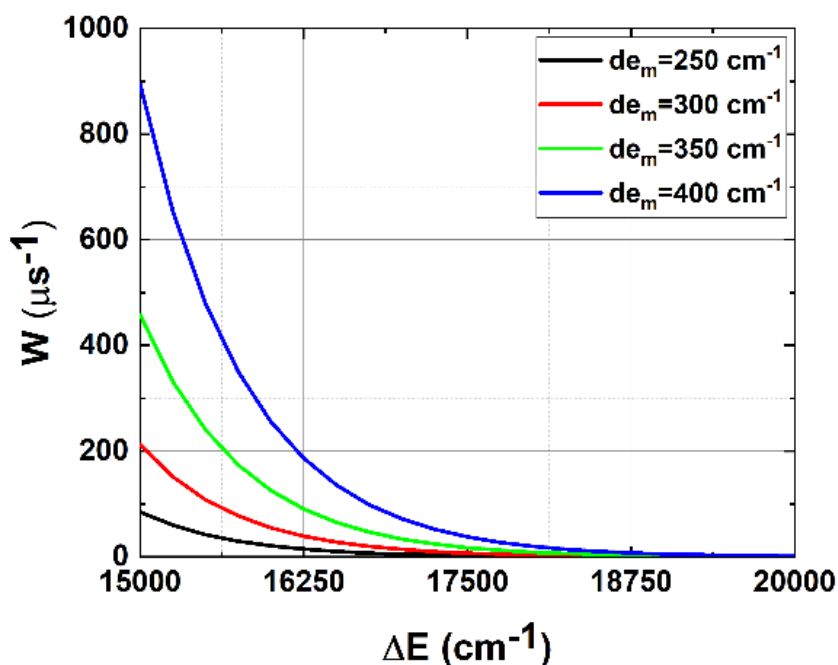
vibrational frequency  $\langle\omega\rangle = N^{-1} \sum_j \omega_j$ , where N is the number of modes and  $\omega_j$  is the frequency of each mode. In the low temperature - strong coupling regime, the coupling strength  $G \sim \frac{E_M}{\hbar\langle\omega\rangle} > 1$ . Where,  $E_M \sim$  Stokes shift/2 in the absence of excited state processes such as excited state proton transfer (ESPT). In the low-temperature/strong coupling limit, the expression for the non-radiative rate resembles that derived by Marcus for the electron transfer rate:

$$W = \frac{1}{\hbar} \frac{C^2 \sqrt{2\pi}}{\sqrt{E_M k_B T^*}} \exp - \left( \frac{(\Delta E - E_M)^2}{4 E_M k_B T^*} \right) \quad (5)$$

Where  $T^*$  is the effective vibrational temperature  $T^* \sim \hbar\langle\omega\rangle/k_B$ ,  $C$  ( $\sim 10^2$ - $10^4$  cm<sup>-1</sup>) represents the Herzberg-Teller coupling of the vibronic transition, and the other variables were previously defined. With experimental Stokes shifts from below 300 cm<sup>-1</sup> to above 3000 cm<sup>-1</sup> and energy gaps from below 15000 cm<sup>-1</sup> to 22500 cm<sup>-1</sup>, FPs can theoretically fall into either weak or strong coupling cases or in between them. In the weak coupling limit ( $G < 1$ ) suited to most FPs, the observed Stokes shift is severalfold smaller than electronic transition energy ( $\Delta E$ ). In this case one arrives at the following expression for the non-radiative rate,

$$W = \frac{1}{\hbar} \frac{C^2 \sqrt{2\pi}}{\sqrt{\hbar \omega_M \Delta E}} \exp - \left( \frac{\gamma \Delta E}{\hbar \omega_M} \right) \quad (6)$$

Here,  $\omega_M$  is the frequency of the normal mode vibration with the maximum frequency in the chromophore, the parameter  $\gamma \sim \log\left(\frac{\Delta E}{d e_m}\right) - 1$ ; where  $d$  is the degeneracy and  $e_m$  is a measure of the reorganization energy of this vibrational mode in the excited state. While more applicable to FPs the weak-coupling form is less intuitive in its relation to experimental observables. Further theoretical work is necessary to address the possible intermediate regimes of this formalism in the context of FP engineering. Both strong and weak-coupling regimes indicate that the non-radiative transfer probability (W) is dependent on the Stokes shift (from vibrational reorganization), vibrational frequencies that couple to the electronic transition and the transition energy gap. To illustrate the effects of these parameters on the non-radiative rate, we plot the trends expected from the weak coupling expression in Figure 1.4, which shows that decreasing energy gap and increased reorganization energy lead to an increased non-radiative decay rate (W).



**Figure 1.4. Expected trends from the weak-coupling low temperature limit of Englman-Jortner theory (equation 6):** The super-exponential increase of the non-radiative rate ( $W$ ) for smaller energy gaps ( $\Delta E$ ) and larger reorganization energies ( $de_m$ ). The values  $\omega_M \sim 3000 \text{ cm}^{-1}$  and  $C^2 \sim 10^7 \text{ cm}^{-2}$  were fixed, as discussed in reference [69].

The Englman-Jortner approach can qualitatively explain the difference in average non-radiative rate for GFPs relative to RFPs. The lower energy emission of the latter is due to their larger chromophores. RFPs contain an acylimine moiety over which the electronic conjugation is extended compared to GFPs (Figure 1.1b and 1.1c). The larger number of vibrational degrees of freedom and smaller electronic energy gap of RFPs could possibly lead to a higher density of states. One can thus predict RFPs to have lower fluorescence quantum yields than GFPs, in-part due to the expected higher values of the non-radiative rate constant. However, an outstanding exception to this trend is mScarlet, which has a smaller non-radiative rate constant than many bright blue-shifted counterparts such as EGFP ( $k_{\text{non-rad mScarlet}} \sim 77 \mu\text{s}^{-1}$  v.  $k_{\text{non-rad EGFP}} \sim 113 \mu\text{s}^{-1}$ ), [57] despite a significantly lower  $\Delta E$  (by  $\sim 3000 \text{ cm}^{-1}$ ) and near identical values of  $E_M$  ( $\sim 375 \text{ cm}^{-1}$ ). This RFP was designed with a focus on conformationally restricting the chromophore, which could



have resulted in a lower density of accessible vibrational states in the ground electronic state. [35] The causes of this exceptional behavior merit further investigation. Further theoretical and experimental studies are warranted to examine the extent to which energy-gap effects are the main contributor to nonradiative relaxation especially in red and far-red emitting chromophores. [58]

#### **iv. Spectral features**

In the descriptions of Strickler-Berg and Englman-Jortner models for the rates of radiative and non-radiative transitions, we pointed out where correlations with steady-state spectral properties are to be expected. For example, small Stokes shifts and blue shifted emission peaks are correlated with high radiative rates and low non-radiative rates. The Stokes shift reflects the reorganization energy for solvation of the excited electronic state (half of the Stokes shift= $E_M$ , as in equation 5 and 6) but this is correct only in the absence of excited state photochemistry such as excited state proton transfer (ESPT) which occurs in the “large Stokes shift (LSS)” FPs. In addition to absorption and emission peak values, photophysical properties are correlated with absorption lineshapes and linewidths. Chromophore spectral lineshapes are sensitive to the underlying femtosecond to nanosecond timescale protein and solvent dynamics and protein conformational heterogeneity. Ultrafast spectroscopy experiments such as time-resolved fluorescence Stokes shift measurements and photon echo techniques resolve the timescales of nuclear motions. [71, 72] An understanding of the interplay between protein structure and dynamics would provide useful insight for FP engineering efforts. [73] To date, most ultrafast spectroscopy of FPs has focused on investigating excited-state photoreactions such as chromophore isomerization, ESPT, and hydrogen-bond dynamics (see below). We now turn to a consideration of structural features and mechanisms for excited state depopulation, such as hydrogen bonding and electrostatic effects. [15, 58]

### 1.3c. Discussion of specific cases:

#### i. Structural arguments

Often, the rationale for mutations to improve brightness is centered on trying to make the chromophore more “planar” and the FP more “rigid.” It is expected from molecular orbital arguments that a planar or flatter chromophore would result in stronger  $\pi \rightarrow \pi^*$  electronic transitions and thus a larger value of extinction coefficient due to improved electronic delocalization through the methine bridge. [74] The second rationale is difficult to examine critically due to the complexity of quantifying “rigidity.” Furthermore, the existence of a conical intersection in the excited-state potential energy surface of FPs, linked to the twisting of the methine bonds between the two rings complicates the straightforward picture of planarity and rigidity. [75] Although they are not certain to capture the average chromophore conformation or excited state distortions, x-ray crystal structures show a correlation between chromophore planarity and brightness. [16] The structures of mCherry and mStrawberry revealed that the decreased fluorescence quantum yield of these FPs are correlated with non-planarity of these chromophores compared to the parent DsRed and sibling mOrange. [76] The x-ray crystal structure of mScarlet (PDB ID:5LK4) shows a dihedral angle of  $\sim 1.9^\circ$  between the methine bridge and phenol ring (or the P-bond rotation), which is significantly smaller than the  $\sim 13.1^\circ$  for the dimmer RFP mCherry (PDB ID:2H5Q). [35] The flatter chromophore has a  $\sim 1.5$ -fold higher  $\epsilon_{\max}$ ,  $\sim 3$ -fold higher fluorescence quantum yield,  $\sim 2.5$ -fold higher radiative rate constant and  $\sim 7$ -fold lower non-radiative rate constant. Along these lines, mutagenesis to introduce residues that improve the packing around the chromophore have often been used to generate brighter variants. This strategy was employed in the development of mTurquoise2, where a fluorescence quantum yield of 93% (the highest of all engineered FPs) was achieved by introducing an I146F substitution in mTurquoise (fluorescence quantum yield  $\sim 84\%$ ). [77] This residue was identified as a target for mutagenesis by a combination of fluorescence lifetime screening, x-ray crystallography, and classical MD simulations, which suggested that a bulkier, non-polar residue might restrict the conformational freedom of the chromophore, resulting in a higher fluorescence quantum yield.

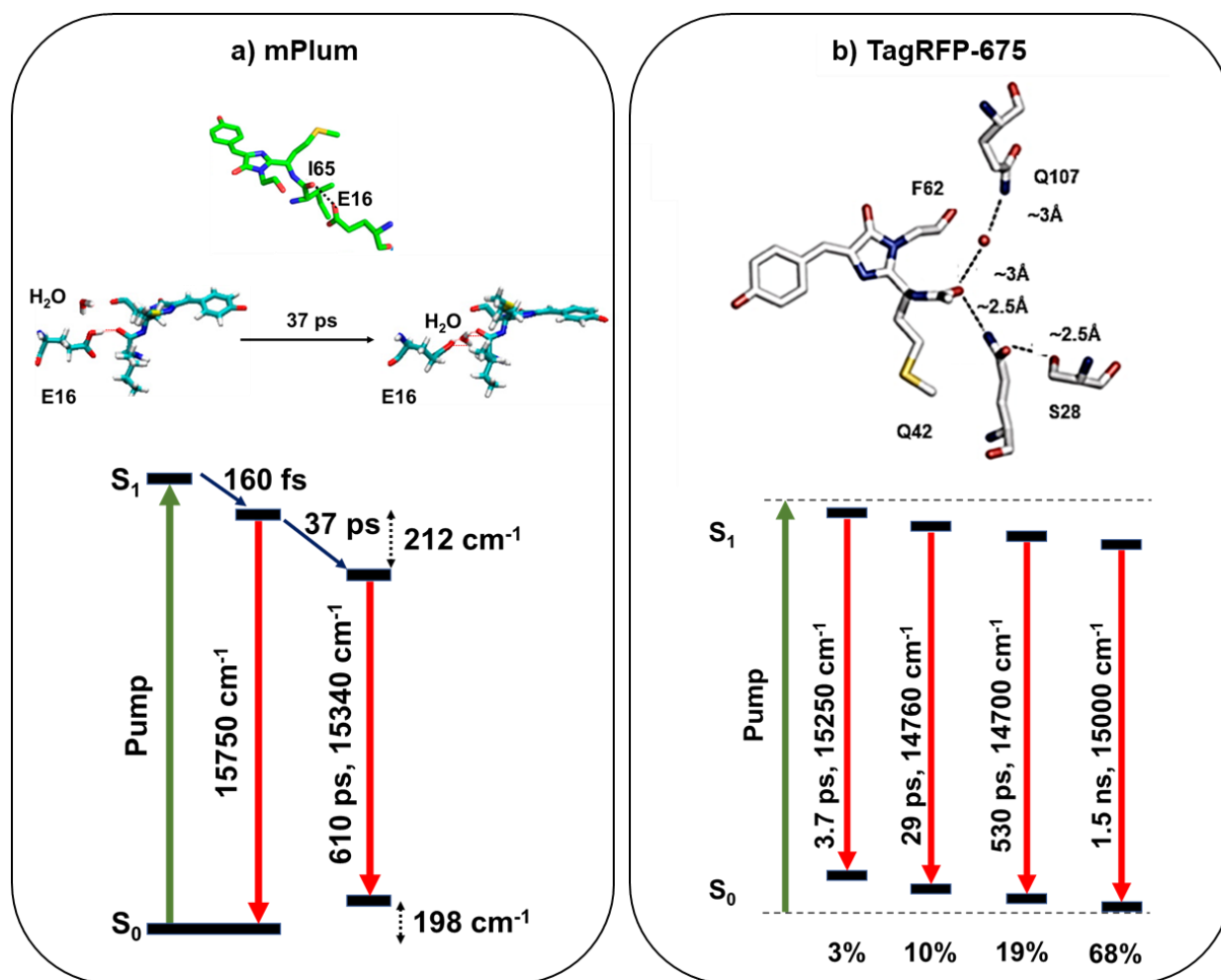
Inspiration for molecular design has been provided by spectroscopic and theoretical studies of specific interactions and excited-state processes in relation to hydrogen bonding and proton transfer - particularly in GFP model chromophores. [78-80] While a majority of these studies have

been based on a combination of vibronic spectroscopy, computational methods, X-ray crystal structure analysis and bioinformatics, [81-84] NMR and electron paramagnetic resonance spectroscopy have also proven useful (and perhaps under-utilized) for providing further insight into FP structure and conformational dynamics. [85-87] For example, high-resolution 2D  $^1\text{H}$ - $^{15}\text{N}$  solution-based NMR backbone relaxation study revealed a light-induced conformational change in the photo-switchable protein rsFolder. [88] The use of kinetic crystallography for revealing the structural transitions associated with long-lived dark states has also been fruitful. [89] This promising approach combines real-time crystallization of FP molecules with optical measurements such as Raman spectroscopy. The work of Bourgeois and co-workers has provided insights into the mechanisms behind photo-transformations in photo-switchable FPs like mEos4B. [90] A structural perspective can be helpful for identifying specific interactions that might govern brightness. [16-19]

## **ii. Influence of local hydrogen-bonding networks**

Though hydrogen bonding has a variety of impacts on the photophysics of various types of FP chromophores, we will limit our considerations to RFPs, and in particular the role of hydrogen bonding with regard to the acylimine moiety. In RFPs, the introduction of a single hydrogen-bonding interaction with the acylimine moiety was proposed for red-shifting the emission. However, the major outcome of this new motif is the introduction of picosecond timescale conformational changes which influence the brightness and the Stokes shift of emission. These processes have been investigated both with experiments and QM/MM simulations. [15, 79, 91-93] In mPlum, which has a large  $1540\text{ cm}^{-1}$  (60 nm) Stokes shift, Boxer and coworkers observed a picosecond timescale red shift of the emission spectrum, which they explained as a dynamic Stokes shift associated with reorganization of the E16 hydrogen bond. [94] We used time-resolved fluorescence experiments to reveal the two-state interconversion between the direct and water-mediated hydrogen-bonding interactions of the acylimine with the E16 residue (Figure 1.5a). [79] QM/MM simulations confirmed the connection between the structural dynamics and the Stokes shift obtained from the experiments. [93] What remains uncertain is exactly the connection between the large Stokes shift and the small fluorescence quantum yield of mPlum (10%).

The availability of multiple hydrogen bonding possibilities in the acylimine region appears to be a detriment to brightness. For example, TagRFP-675 (Figure 1.5b) was engineered from mKate to red shift the emission by introducing hydrogen bonding to the acylimine carbonyl, but these modifications resulted in a decrease in the fluorescence quantum yield (33% to 8%). [95] Our time-resolved emission measurements revealed the presence of four emitting species of varying red shifts, which independently decay to the ground-state. Classical MD simulations also revealed multiple interconverting structures of this hydrogen-bonding network, but it was not possible to assign the spectral forms to individual structures. [15] Interestingly, acylimine hydrogen bonding does not always lead to FPs with low fluorescence quantum yield. In developing the bright FusionRed mutant FR-MQV, [38] we hypothesized that the M42Q mutation would be beneficial by occupying a cavity near the acylimine and thus locking it into a favorable geometry. This interaction increased the molecular brightness of the FP by 2-fold in comparison to FusionRed. It may be productive to focus attention on these issues using quantum and classical molecular simulations to examine models in which specific interactions with the chromophore is predicted to stabilize a single long-lived structure.



**Figure 1.5. Excited-state spectral and structural dynamics in RFPs studied using ultrafast spectroscopy. (a.) mPlum:** Two-state interconversion of direct and water-mediated hydrogen-bonding states. [79, 91, 93] **(b.) TagRFP-675:** The sidechains of residues F62-Q42-Q107-S28-R41 constitute a hydrogen-bonding network leading to four spectral forms (with populations given as percentages) and widely varying excited-state  $\tau$ . [15]

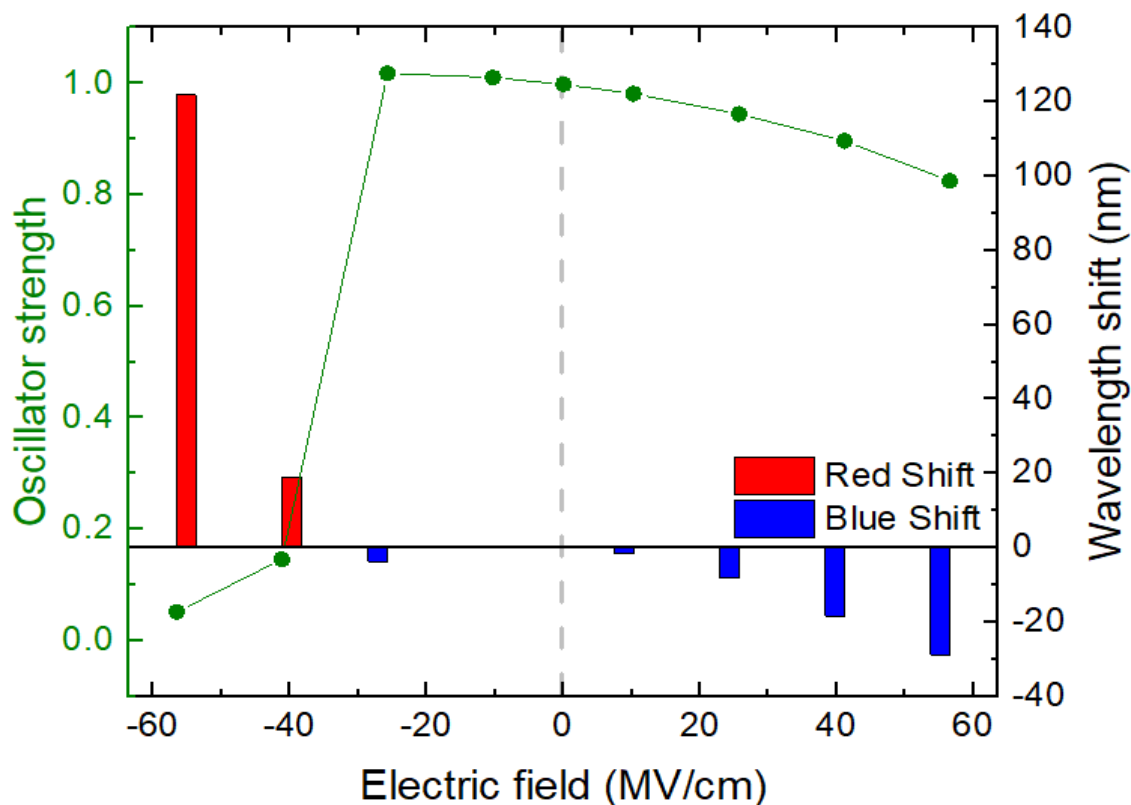
### iii. Electric fields and electrostatic effects

Theoretical and experimental studies have examined the influence of electric fields and electrostatics on radiative and non-radiative rates. [46-48, 62, 82, 96-98] For example, TD-DFT calculations on the isolated GFP chromophore (Figure 1.6) predict that electric fields can change the oscillator strength of an electronic transition, and therefore the radiative rate constant. [97, 98]

Park and Rhee performed non-adiabatic molecular dynamics simulations on the GFP chromophore and found that electrostatic effects can outweigh steric factors impeding the twisting of the methine bridge between the phenoxy and imidazolinone rings, suppressing a major pathway for non-radiative relaxation. [74] Drobizhev and coworkers investigated these issues by measuring two-photon absorption cross-sections of RFPs. [58] This cross-section is sensitive to the change in permanent dipole moment upon electronic excitation, which in part is controlled by the electric field on the chromophore. [58] They proposed that brighter fluorophores could be produced by fine tuning the strength and directionality of the field, especially along the axis from the center of the imidazolinone ring to the phenolate ring. Electric fields along this axis are expected to change the amount of single bond vs. double-bond character in the methine bridge, and therefore tune the rate of non-radiative decay associated with chromophore twisting followed by passage through a conical intersection to the ground state. Since this motion is accompanied by charge transfer across the methine bridge, the rate can be described by Marcus electron transfer theory. [70]

It is clearly a drastic oversimplification to characterize the electrostatic environment around the chromophore by specifying the electric field along different directions. In their *tour de force* 2019 publication, [62] Lin *et al.* circumvented this issue by treating the GFP absorption band as an intervalence charge transfer band between two resonance structures corresponding to the negative charge residing either on the phenolate oxygen or the carbonyl oxygen of the imidazolinone ring. The energetics of this charge transfer is described by Marcus-Hush theory. [99] The ground and excited state potential energy surfaces were described by coupling the diabatic electronic states for the two resonance forms through a bond-length alternation (BLA) coordinate. They demonstrate the electron-donating or electron withdrawing nature of the sidechains at positions 203 and 96 (located at either end of the chromophore) controls the driving force for the charge transfer. This model quantitatively explains trends in absorption maxima, Stokes shifts, molar extinction coefficients, lineshapes, and other properties. For example, eliminating the Thr203 hydrogen bond with the phenolate oxygen reduces the driving force, leading to a decreased transition dipole moment, red-shifted absorption, a decreased Stokes shift, and a smaller vibronic sideband. The effect on the spectral lineshape is due to the relative intensity of 0-0 and 0-1 vibronic peaks of the  $1340\text{ cm}^{-1}$  BLA mode, and occurs in addition to inhomogeneous broadening due to sensitivity of the electronic transition energy to an electric field. The theory can be applied to other chromoproteins such as photoactive yellow protein but it assumes the chromophore undergoes

small geometry changes upon excitation and does not consider mutations that modify the electronic conjugation, so it does not explain the behavior of FPs such as BrUSLEE, where radiative rates are significantly modified by mutations. The model however, provides a significant physical insight for FP design, and it would be very useful to develop a similar approach which reduces the dimensionality of protein electrostatic effects to an energetic coordinate, for modeling other classes of FPs. With this, we move on to a discussion of FP photostability.



**Figure 1.6. A graphical summary of results from a TD-DFT study.** A study performed by Kang et al., [98] demonstrates that the oscillator strength and the absorption maxima for the GFP chromophore can be modulated with changes in the electric field.

## 1.4. Photostability

### 1.4a. Photobleaching measurements and mechanisms

Irreversible photodamage or photobleaching is a hurdle to all applications. Despite the protection offered to the chromophore by encapsulation within the  $\beta$ -barrel, FPs generally photobleach faster than many small molecule fluorophores. This photoreactivity is possibly due to the diversity of chromophore interactions with the surrounding amino acid side chains and with freely diffusing species such as  $O_2$ , which provide opportunities for photo-oxidation and electron-transfer reactions. [100] Identifying the mechanisms and/or structures of the photobleached products has been a major challenge in FP development, most likely because several mechanisms are simultaneously at play and multiple products are formed. As a result, simple physical models have been of limited utility in guiding molecular design for higher photostability.

Photostability is influenced by the photophysics of the fluorophore and by the properties of the excitation source. A wide range of continuous wave (cw) or pulsed lasers, LEDs and arc-lamps have been used for imaging and spectroscopic studies. Furthermore, in an attempt to investigate photobleaching under conditions corresponding to various imaging techniques, measurements have been performed with irradiances spanning several orders of magnitude, from  $\mu W/cm^2$ – $kW/cm^2$ . [101] Often, the decay of fluorescence intensity vs. time is found to be non-exponential. As a result, the bleaching half-life is typically reported as a function of excitation irradiance. However, this half-life, or the corresponding average photobleaching rate, is highly dependent on the particular irradiation conditions such as the photon flux, excitation wavelength, peak intensity and pulse duty cycle. [55, 56] Furthermore, biochemical factors such as the buffer conditions for *in vitro* measurements, the host organism and cellular localization (*e.g.*, cytoplasmic, nuclear or membrane-associated construct) are also important considerations. FPBase provides a note of caution for researchers interested in using posted values for comparing the photostability of different fluorophores. [57] The lack of standards for measuring and reporting photostability complicates decisions on which FP to select for a particular application and thus choices are often made on anecdotal evidence.

To address inconsistencies in reporting of photobleaching and other parameters, Cranfill *et al.* systematically measured the photobleaching and other properties of more than 40 FPs. [101] A



major revelation from their measurements is that the relationship between excitation intensity and photobleaching rates is typically supra-linear, *i.e.*, the average photobleaching rate  $k_{bleach} \propto I^\alpha$  where  $\alpha > 1$  and is nearly quadratic for some FPs. For example, they found that mCherry emission decays with a half-life of 318 seconds when bleached with a 590 nm scanning-laser illumination at 80  $\mu$ W excitation power *vs.* 88 seconds when bleached with a 594 nm widefield LED source at 200  $\mu$ W excitation power. Although this study represents a major step forward in characterizing FPs so that data-driven choices in experimental design can be made, the authors did not aim for a mechanistic understanding of photostability.

The lack of consistency in measuring and reporting photobleaching has posed challenges to creating models that connect photostability with other photophysical properties. To advance this field, it is recommended that quantum yield of photobleaching ( $\phi_{PB}$ ) be adopted as the gold-standard for reporting and comparing FP photostability. Since, as discussed above, the power dependence of FP photobleaching may be nonlinear, and the value of  $\alpha$  might vary in different illumination regimes,  $\phi_{PB}$  should be reported as a function of irradiance, like what is done for the 2-photon upconversion yields of lanthanide nanomaterials. [102] This measure relies on normalizing the bleaching rates with excitation rates, which is determined by multiplying the irradiance on the sample with the absorption cross section at the excitation wavelength (Equation 7).

$$\phi_{PB}(I) = \frac{\text{Number of molecules photolyzed}}{\text{Number of photons absorbed}} = \frac{k_{bleaching}}{k_{excitation}} = \frac{(\ln 2) N_A h c}{(2303) I \times \lambda \times \epsilon_\lambda \times t_{\frac{1}{2}bleaching}} \quad (7)$$

Where,  $N_A$  is the Avogadro's constant,  $h$  is the Planck's constant,  $c$  is the speed of light,  $I$  is the irradiance (power per unit area),  $\lambda$  is the wavelength of excitation,  $\epsilon_\lambda$  is the extinction coefficient at  $\lambda$ , and  $t_{\frac{1}{2}bleaching}$  is the half-life obtained from the bleaching decay profile. Zak et al. followed this approach in considering the potential use of molecular fluorophores for luminescent convertors in LEDs. [103] According to their calculation, one of the most photostable dyes for laser spectroscopy, tetramethylrhodamine, has a  $\phi_{PB}$  of  $3.3 \times 10^{-7}$ . Photostability data from FPBase suggests that the most stable FPs have  $\phi_{PB}$  of  $\sim 10^{-4}$  when analyzed with this approach. [57] Given that photon output before photobleaching limits the imaging duration and quality in many experiments, this 1000-fold difference in photostability between dyes and FPs indicates there is likely far more to be gained by improvements in photostability than in brightness.

The fluorescence profile is often found to decay on multiple timescales. The fastest time constant is typically associated with “reversible photobleaching,” the same process which we already referred to as “dark state conversion,” whereas the slower decay usually is due to permanent photobleaching. The relationship between the photobleaching rate constant and the half-life value is exact in the absence of dark-state conversion, and the photobleaching decay can be fit with a single-exponential, or in cases where the photobleaching is exponential and much slower than dark state conversion (e.g. EGFP and mScarlet). However, many FPs do exhibit long lived dark states, and photo-activation or photo-switching processes compete with the first order kinetics of permanent photobleaching. Moreover, chromophores can permanently bleach from dark states. [104] Analogous dynamics are observed in azobenzene dyes, which like many FPs participate in a trans-cis isomerization under irradiation. [105] For azobenzenes, the dark cis conformer also bleaches without emitting, leading to non-linearity of the photobleaching. We and others have reported methods for decoupling these processes from permanent photobleaching and extracting the corresponding rate constants using other experiments such as fluorescence recovery measurements. [55, 56, 106]

The brightest FPs often suffer from the highest photobleaching rates, in part because the photobleaching quantum yield is inversely proportional to the excited-state lifetime. For example, in the development of FR-MQV, as we selected for progressively longer fluorescence lifetime clones, we observed a trend towards shorter photobleaching half-life under excitation normalized widefield conditions ( $\sim 5 \text{ W/cm}^2$ ). FR-MQV has  $\sim 4$ -fold lower photostability than the parent FusionRed despite a 1.6-fold longer fluorescence lifetime. [38] Another example is mScarlet, which is 5-fold brighter than its variant mScarlet-H but has a 3-fold longer lifetime and a  $\sim 2$ -fold lower photostability (under  $6.9 \text{ W/cm}^2$  widefield excitation and  $1.35 \text{ W/cm}^2$  spinning-disk confocal excitation). [35] However it is interesting to note that the drop in photostability does not scale exactly linearly with the increase in fluorescence lifetime for these two cases: it can be larger or smaller than expected. This observation suggests that different mechanisms of photobleaching may be at play within the two RFP families. More generally, examining the cases which deviate from the expected linear scaling within a closely related series of FPs might provide some insight into the structural features associated with photobleaching.

Turning to molecular mechanisms, computational and experimental studies suggest that photodamage is associated with the availability of molecular oxygen and diffusion of water molecules inside the  $\beta$ -barrel. [100, 107, 108] Excited electronic states have higher reduction potentials than the ground state and are therefore more susceptible to reacting with species such as  $^1\text{O}_2$ , which has been shown to be a major participant in photobleaching in certain cases. [100, 107] Mechanistic investigations of phototoxic FPs that have found applications in chromophore-assisted light inactivation (CALI), have provided important insights into photobleaching. [109] In these systems, photobleaching seems to be directed into one overall reaction. To understand this process, Grigorenko and coworkers employed QM/MM simulations to investigate the reactivity of the FP KillerRed. Their analysis revealed that a model excited-state FP chromophore can react with  $\text{O}_2$  through several possible charge transfer intermediates. By sampling many reaction pathways, their analysis identified a low activation barrier pathway ( $\sim 13$  kcal/mol) in which the hydroxyphenyl moiety is oxidized to a benzoquinone species that can diffuse out of the  $\beta$ -barrel. This result is consistent with electron density maps from an X-ray crystal structure of photobleached KillerRed. [109] Sen et al. recently performed QM/MM simulations to explore the role of T65/S65, the first chromophore-forming residue in EGFP and EYFP, respectively. For EGFP, a T65G substitution reduced the fluorescence lifetime but increased photostability, with a reduction of the characteristic green-to-red oxidative photoconversion, or “redding” behavior commonly observed for EGFP. [110] This phenomenon is observed when the GFP chromophore is oxidized to red-shifted absorption and emission species by high irradiance illumination. The transfer of fluorescing populations to such species has been linked to faster photobleaching. [111, 112] We refer the reader to the work of Acharya *et. al* for a detailed explanation of this phenomenon. [16]

Dark states can also contribute to photobleaching. The lifetimes of triplet states, for example, are orders of magnitudes longer than the  $\text{S}_1$  state and can, in fact, catalyze the formation of  $^1\text{O}_2$ , which greatly speeds up photochemistry promoting chromophore destruction. [55, 65] Therefore, reagents like oxygen scavengers and triplet-state quenchers are promising tools for slowing the rate of photobleaching, as they create apoxia in the system or depopulate non-fluorescent triplet states. [113] Another complication is that cases have been observed where the effect of dark states on photobleaching depends on the illumination conditions. For example, we found that the RFP mCherry was more photostable under pulsed illumination than under cw illumination, particularly

at high irradiances ( $25 \text{ kW/cm}^2$ ), suggesting that mCherry's dark state is photoreactive, whereas for the variant TagRFP R67K S158T, we found that the dark state is more photoprotective than photoreactive. [55]

### **1.4b. Development of photostable FPs**

Despite these challenges, several studies have advanced fluorophore photostability. For example, Single-cell Phenotypic Observation and Tagging with light (SPOTlight), a cell sorting technique, led to the most stable yellow FP (YFP) observed to date, mGold. [114] In this case, eight libraries, each of 8000 variants were produced by targeted mutagenesis of 21 positions in the YFP mVenus (six near the chromophore) and expressed in yeast. SPOTlight then photobleached cells at an irradiance of  $\sim 2 \text{ W/cm}^2$  and selected photostable variants from up to 700,000 cells in multiple rounds of selection. Similarly, high-throughput screening for photostable FPs using a Lego-based robot led the development of another YFP, Citrine2, which is  $\sim 2$ -fold more stable (in widefield and laser bleaching assays) than its precursor, mCitrine. [115] In this case, directed evolution of error-prone mutagenesis libraries was used with photostability selection with a 300 W white light Xenon lamp source that had an effective white light irradiance of  $\sim 3.5 \text{ W/cm}^2$ . This study identified 8 mutations that improved the photostability of Citrine2 in comparison to mCitrine but also led to unfavorable consequences such as dimerization and a drop in fluorescence lifetime from 3.6 ns to 3.3 ns. Surprisingly, these changes also led to a 5% increase in the radiative rate and 25% increase of the non-radiative rate. [115] These examples provide proof-of-principle demonstrations that screening large numbers of variants with new technologies can generate more photostable variants, but much more work remains to be done beyond the demonstration phase.

Our efforts to improve the photostability of mCherry using microfluidic technology led to the development of Kriek, which is 4-fold more photostable in confocal fluorescence microscopy. [108] This variant was selected from a 144,000-member library designed with guidance from MD simulations. The simulations identified a region of the  $\beta$ -7/10 strand interface that showed ns-timescale fluctuations inconsistent with a rigid barrel structure, and we identified additional positions adjacent to this region which appeared tolerant to mutations. Selection was performed to minimize the amplitude of fluorescence signal lost after excitation from four sequential excitation

beams (to mimic pulsed excitation in flow) under irradiance of  $\sim 2 \text{ kW/cm}^2$ . We later found that high irradiance-pulsed excitation tends to populate dark states in mCherry and TagRFP variants. Since the microfluidic-based selection of Kreik was carried out with a similar excitation scheme, we also found that Kreik had lower rate constants for dark state conversion. [55] Previously, we demonstrated that the progenitor mCherry has a photoreactive dark state. Therefore, a reduction in the rate of dark state conversion may have resulted in an additional boost to the photostability of Kreik. Unfortunately, we observe a significant drop in the fluorescence quantum yield of Kreik (8%) with respect to the precursor mCherry (22%) with this strategy. The lesson here is that MD simulations can provide useful insight for library design, but that screening only on photobleaching is likely to produce variants with compromised brightness.

Other site-directed mutagenesis studies have been fruitful, as evident in the development of SiriusGFP - which is 2-fold more stable than its precursor EGFP and was found to be well suited for imaging at high irradiance ( $\sim \text{kW/cm}^2$ ) and long time-lapse imaging assays. [116] In this case, site-directed mutagenesis was guided by considering photochemical reactions of the model GFP chromophore. They found that S147R and S205V mutations in EGFP greatly enhanced the photostability under high irradiance laser-based imaging. [117] As with other efforts to improve photostability, SiriusGFP suffers from reduced brightness in comparison to EGFP due to a  $\sim 3$ -fold reduction in its fluorescence quantum yield. When Ren et al. (2016) introduced a cysteine group near the chromophore of the RFP mKate2, it resulted in the 12-fold improved variant, “mStable.” The photostability of mStable was attributed to the sulfoxidation of a cysteine residue facing the p-hydroxyphenyl moiety of the chromophore. Similar substitutions in mPlum resulted in a 23-fold more stable variant. [118] The development of mGold from mVenus and Citrine2 from mCitrine, demonstrates that photostability can be improved without significant compromises in molecular brightness. This outcome suggests that it is possible to evolve existing bright FPs into more photostable variants, and *vice-versa*, perhaps by screening libraries simultaneously on photostability and fluorescence lifetime.

Inspiration for FP design and applications may draw from effective techniques that minimize photobleaching in molecular fluorophores, such as utilizing triplet state quenchers and using systems that induce anoxia. [119] In addition, a promising strategy was recently reported where deuteration of small molecule dyes such as tetramethyl-rhodamine and JaneliaFluor (JF) dyes

substantially improved photostability without causing detrimental changes to the electronic or photophysical properties. [120] For these types of dye molecules, alkylamine oxidation produces a distinctive secondary isotope effect, where deuteration of hydrogen atoms may reduce a twisted-intermolecular charge transfer process, as is commonly observed for many FPs and in xanthene based JF dyes. Deuteration additionally increases the strength of the C-D bond and also affects the  $^1\text{O}_2$  oxidation rate, which manifests in higher (~1.2 fold) photostability. In addition to these molecular strategies, methods that involve the light source, such as control of photon statistics, pulse shaping, or excitation with squeezed light may also be promising tools for imaging FPs for longer periods of time by increasing the measurement precision attainable for a particular photon dose, even if these techniques do not directly reveal the chemistry behind photobleaching. [121]

## 1.5. Conclusions

A few themes emerge when we consider the evolution of this field over the past 30 years. Initially, FP development relied on relatively straightforward structure-guided design and chemical intuition with screening on cellular brightness. As the productivity of this approach has started to decline, new variants have been introduced by relying on more advanced screening methods, with narrowly-focused photophysical and biochemical goals. Our knowledge of FP structure and the diversity of their photophysics has dramatically expanded, however, at the same time, the pace of new FP development has slowed, e.g. of the more than 1500 FPs collected in FPbase, 270 were reported in the peak years of 2009-2010, decreasing to about half that number in subsequent two-year periods since then. Is this slow-down due to the limits of our technology or of our insight? We think mostly the latter. How will the field continue to advance? Technology for faster, more precise screening of large libraries can partially make up for our knowledge gaps. Screening of RNA aptamer or protein libraries several orders of magnitude larger than those in FP development have frequently been used to successfully evolve high-affinity, specific ligand binding when physical insight is not applied to the library designs. [122, 123] Nevertheless, screening on complex functions such as photophysical properties will require more physical insight applied to the molecular design, and we feel that a renaissance of FP development might follow from investments in new experimental and theoretical approaches.

Experimental studies combining selections that go beyond “skimming the crème off the top” of libraries and simply keeping and studying the best clones could lead to unexpected insights. If the screening is performed such that the evolution of the properties and the structures within the libraries are monitored on a large scale, then bioinformatics and machine-learning techniques could be used to obtain more information about why some selections work and why some selections do not produce improvements. [124] Computational methods such as classical molecular dynamics and hybrid quantum-molecular mechanics would be valuable for creating models that advance our understanding of FP conformational dynamics and photoreactions such as photobleaching. These efforts should be accompanied by refinement of classical physical chemistry theories of, e.g. internal conversion so that it can be understood when we are facing issues particular to a small class of chromophores vs. hitting fundamental physical limits. These investigations could help us develop protein design strategies that bridge structural measures such as chromophore planarity and rigidity, with photophysical parameters such as fluorescence lifetime or dark-state conversion rates. Clearly, physical chemists could make fundamental contributions to the development of new fluorescent proteins or utilize the progress in this field to explore new avenues in chemistry. [125] Progress on these topics may have impacts on areas beyond bio-imaging, such as solar-energy conversion materials, bio-hybrid LED lighting, and bio-photonics such as display technologies.

## **1.6. Thesis aims and objectives**

The central theme of this thesis is to advance fluorescent protein design driven by the fundamentals of physical chemistry, particularly with respect to understanding the basis of brightness and photostable red FPs. This work first involved the development of a multi-parameter, high-throughput, droplet-based microfluidic sorting system, integrating high-throughput fluidics with laser spectroscopy (Chapter 2). The instrument advanced our existing capability by 3-orders of magnitude with respect to multi-parametric cell sorting ( $\sim 10^8$  RFP expressing cells in 3 hours). This device and the pre-existing microfluidic sorting technology enabled selection of FPs with increased fluorescence lifetime (and corresponding increase in fluorescence quantum yield or the brightness) from libraries of rationally or randomly mutated FPs.



Subsequent chapters describe protein design based on crystal structure, bioinformatics, and experimental guidance from ultrafast spectroscopy to identify target residues and generate mutant libraries of FusionRed and mCherry. Consequently, lifetime-based selections lead to the development of bright RFPs - FusionRed-MQV (Chapter 3) and mCherry-XL (Chapter 5). The FusionRed-MQV variant was then characterized in cells to explore the potential for cell biology applications.

Chapter 4 describes investigations of FusionRed-MQV focused on quantitative analysis of photobleaching and dark state conversion *in vitro*. The parental proteins of FusionRed-MQV – FusionRed and FusionRed-MQ displayed efficient dark state conversion tendencies compared to their progeny, and thus were selected for detailed study, with a focus on potential applications in super-resolution imaging. The dark state kinetics were investigated with a new approach combining single molecule imaging, ensemble dark state conversion experiments, Monte-Carlo simulations, and data fitting protocols. It was found that measured single molecule blinking timescales could be accurately predicted from their ensemble dark state conversion timescales.

Chapter 5 describes a study of the evolution pathway from mCherry to mCherry-XL. The increase in quantum yield was largely found to be due to the suppression of non-radiative pathways of excited state depopulation. Modeling the reduction of non-radiative pathways with the Englman-Jortner formalism for excited state non-radiative decay in this lifetime evolution trajectory reveals likely involvement of excited-state ultrafast dynamics rather than internal conversion directly to the ground electronic state.

Chapter 6 describes mutagenesis strategies undertaken to further increase the quantum yield of FusionRed-MQV using increased lifetime selections. Error-prone PCR and saturation mutagenesis of sites 71, 159, 175 and 224 identified variations such as FR-MQV-VA and FR-HQC-TT. These FPs have 1.2-1.4-fold longer lifetimes than FusionRed-MQV, with scaling increments in quantum yield due to a lower non-radiative rate. Furthermore, photobleaching studies on these mutants indicate that the dark state in FusionRed variants with a cysteine residue at position 159 is probably photoprotective.

Finally, Chapter 7 provides a perspective for future work based on preliminary experiments. Exploratory measurements indicate the feasibility of developing a microfluidic sorting system for selecting FPs on dark state conversion using the technique Fluorescence Anomalous Phase



Advance (or FAPA). This is followed by a discussion of chromophore maturation and protein expression in limiting the *in-cellulo* fluorescence of bright FP molecules. Initial mutagenesis efforts and all-atom explicit solvent classical molecular dynamics (MD) simulations elucidate the plausible mechanistic role of the C-terminus of the FP in protein expression and chromophore maturation. Further investigations along these mechanistic lines may present an opportunity for the application of bright FPs like FR-HQC-TT in biological imaging.

## Chapter 2

# Instrumentation: Development of a high-throughput microfluidic sorter

### 2.1. Publication Note

Parts of this chapter have been adapted from the article “*Enrichment of Rare Events Using a Multi-Parameter High Throughput Microfluidic Droplet Sorter.*” Hung, S. T.; **Mukherjee, S.**; Jimenez, R. *Lab Chip* **2020**, 20 (4), 834–843. <https://doi.org/10.1039/c9lc00790c>.

### 2.2. Introduction

Fluorescence lifetime is an intrinsic molecular property that is independent of excitation and emission intensity, local fluorophore concentration, and can be detected even with spectral overlaps among fluorophores and in the presence of cellular auto-fluorescence. Fluorophore lifetime is often sensitive to the solvent and biochemical environment, so it has been used as a detection parameter in imaging and sensing techniques. [126-129] Fluorescence lifetime imaging microscopy (FLIM) is a powerful tool complementing fluorescence brightness-based imaging methods. It has been applied to subcellular pH measurements, [130, 131] intracellular refractive index sensing, [132, 133] molecular interactions in cells, [134-136] drug evaluation and discovery [137-139], drug delivery and cancer studies. [140-143] Nonetheless, FLIM applications are hampered by its throughput. Flow cytometry incorporating fluorescence lifetime measurements could significantly improve the throughput, advancing applications to biological and biomedical research such as directed evolution of FPs, [30] protein subcellular localization, [144] protein-protein interaction, [145] drug discovery, [146] and cellular physiology. [147, 148]

Lifetime-based flow cytometry has been demonstrated at a sorting throughput of hundreds of cells per second. [149] However, there are limitations associated with fluorescence detection in a continuous flow stream. For cellular applications, it restricts the fluorescent markers and reactions to be inside or on the cellular surface and is limited to applications that are insensitive to inter-cellular interactions. One approach for overcoming these limitations is to encapsulate cells or other

analytes into isolated droplets that retain their integrity throughout the analysis, and sorting. The ease with which stable droplets can be formed with pL-scale, tunable volumes make droplet microfluidics particularly useful for analyzing individual molecules, cells or other discrete analytes such as beads. These capabilities have been utilized for studying enzymatic activity *in cellulo* [150, 151] and *in vitro*, [152] single-cell analysis and sorting, [153] screening for antibiotic resistance, [154, 155] directed evolution of enzymes, [156] genetically-encoded biosensors, [157, 158] and quantifying heterogeneity at the single cell level. [159, 160] Moreover, microfluidic droplet platforms can be designed for novel flow cytometry applications such as those simultaneously requiring temporally well-defined mixing of cells with reagents followed by time-resolved detection. Fiedler and coworkers have demonstrated resolution and sorting of genetically-encoded biosensors based on various Förster Resonance Energy Transfer (FRET) ratios measured with delay time in seconds. [158] The same platform can be readily modified for directed evolution of fluorescent proteins or enzymes.

The throughput of lifetime-based droplet sorters is impacted by several factors. First, the statistics of cell loading into droplets typically follows the Poisson distribution. [161] To ensure single cell loading, the proportion of non-empty droplets is often limited to  $< 10\%$  of the whole droplet population. Unfortunately, this sparse loading limits the throughput and is therefore often regarded as a disadvantage of the droplet platform. Deterministic single cell encapsulation methods overcome the limitation imposed by Poisson statistics, but there are other limitations such as increased device complexity, substantial proportion of unsorted or wrongly selected droplets, and high flow rates limiting the ability of integrating with other systems. [162] Second, the throughput of a conventional droplet sorter is limited to 2~3 kHz due to the use of a hard divider to separate the collection and waste channels, but new geometries have been investigated to surpass this limitation achieving brightness-based sorting at 30 kHz. [163] Finally, fast data processing of fluorescence lifetime signatures and real-time sorting decision and actuation components are crucial for achieving kHz sorting rates. Despite advances in incorporating fluorescence lifetime measurements into droplet selection methods, the throughput is much lower than purely brightness-based droplet sorting. For example, the throughput of a recently reported fluorescence lifetime droplet microfluidic sorter is 50 droplets/s. [164] A FACS enrichment step is often used to enrich a subset of targets from a large pool prior to selection or investigation on other parameters and platforms. [30, 165-167] Performing fluorescence lifetime selection with this combination of

methods is disadvantageous. In addition to the restrictions imposed by a continuous flow stream in the FACS step, the use of two different instruments imposes uncertainties into the overall selection because the fluorescence intensity values are difficult to calibrate between instruments. Within the general realm of sorting applications, the analysis, enrichment, and isolation of rare macromolecules, cells and particles from a large population constitutes an important subset that is of great importance across a broad area of biomedical, biotechnological, and environmental science. Several papers have described approaches to this challenge in which a rare population is individually analyzed without isolation, or first enriched rather than attempting one-step, single-particle isolation. For example, fluorescence brightness-based droplet digital detection has been applied to the detection of single bacteria in unprocessed blood [168] and profiling circulating tumor DNA, [169] and the implementation of fluorescence lifetime detection was demonstrated to increase the specificity of particle counting. [170] An ensemble sorting approach which repeatedly analyzes and sorts batches within a sample was recently proposed for enriching or separating fluorescent particles. [171] Many microfluidic systems have been developed to enrich and isolate circulating tumor cells, as reviewed in reference 172. [172] Thus, we quantitatively consider the advantages of a batch sorting technique for increasing the throughput of rare-clone isolation. Here, we discuss the development of a multiparameter high throughput water in oil droplet microfluidic sorter capable of screening and sorting analytes based on emission spectra, emission brightness, and fluorescence lifetime. We raised the throughput of lifetime sorting to the upper limit for a droplet sorter incorporating a hard divider between collection and waste channels, [163] which constitutes a 50-fold increase comparing with the recently reported lifetime droplet sorter. [164] We also describe and demonstrate a novel selection strategy, similar to an ensemble-based approach, which exploits the Poisson statistics of analytes in droplets overloaded with multiple analytes. This method provides a several-fold enhancement in sorting throughput. The strategy can be used to analyze and enrich rare events from a large population in either a qualitative manner without the prior knowledge for the initial frequency of the rare events, or in a quantitative fashion with controls for the efficiency and precision of enrichment when the initial frequency of the rare events is estimated. The enriched sub-population can be subjected to further multiparameter analysis and selection with single-cell resolution on the same microfluidic platform. We demonstrate the power of this multiparameter droplet sorter and the enrichment strategy in the context of directed evolution of red fluorescent proteins (RFPs) expressed in *E. coli*.

## **2.3. Experimental**

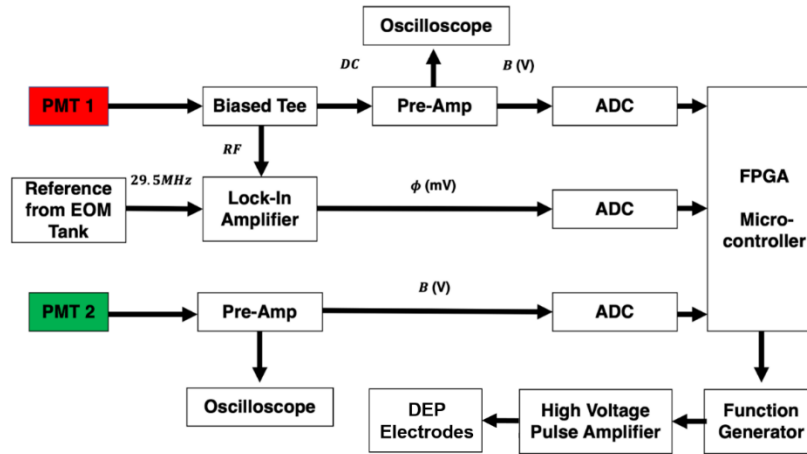
### **2.3a. Optical Layout**

The optical layout of the instrument is depicted in Appendix Figure A1.1. Both 561 nm and 450 nm continuous wave (CW) laser beams excite fluorescence from the cells encapsulated in droplets. The 561 nm beam is focused into an electro-optic modulator that can amplitude modulate the CW beam to a sinusoidal profile. The red and green fluorescence signals are separated by a dichroic mirror and detected by photomultiplier tubes (PMTs).

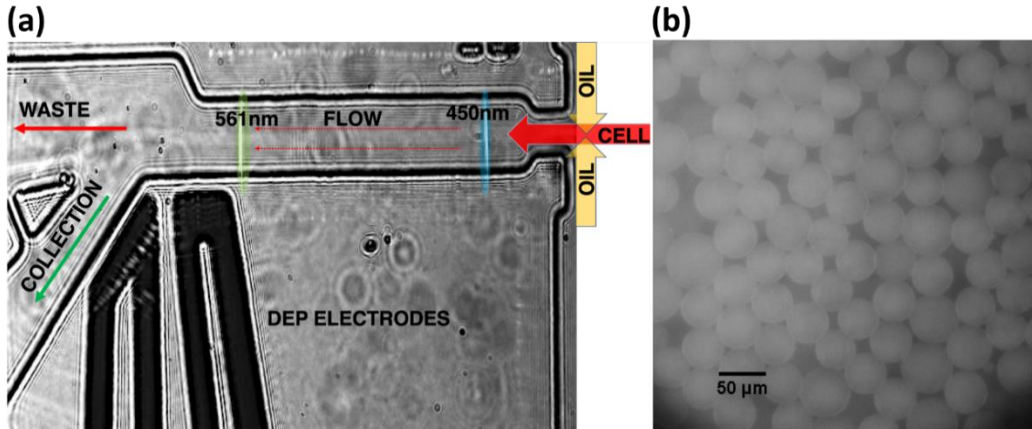
### **2.3b. Electronics and microfluidic device**

The main improvement of sorting throughput in this work is due to the implementation of faster electronics. The layout of the detection electronics is schematically described in Figure 2.1. The electro-optic modulator (EOM; ThorLabs EO-AM-NR-C4) is used to modulate the 561 nm laser light and is driven using a function generator (Agilent 33520B) that provides a 1 V peak to peak sinusoidal signal at 29.5 MHz to a resonator circuit. When screening or sorting based on fluorescence lifetime, the red fluorescence signals from PMT1 are separated into a radio frequency (RF) component (that bears the lifetime information) and the direct current (DC, <83KHz) component using a bias tee. To improve the signal to noise ratio, the DC signals from the biased-tee and from PMT2 (green fluorescence) are amplified using home-built trans-impedance log or linear pre-amplifiers, depending upon the experiment and sample in use, [30] then digitized using Analog to Digital Converter (ADC) boards (Analog Devices, EVAL-AD7986FMCZ, 18 bit). The RF component of the signal is passed onto a commercial high-speed lock-in amplifier (Zurich Instruments UHFLI), which calculates the phase shift of the fluorescence signal relative to the sinusoidal modulation signal to extract information of fluorescence lifetime. The phase shift value from the lock-in amplifier is then digitized using the same type of ADC boards employed for brightness measurements. The digitized signals from the boards are then fed into a customized field programmable gate array (FPGA) board that makes decisions based on user defined parameters interfaced through a LabView program. FPGA has been implemented to enhance the data processing rate for fluorescence lifetime calculation. [173] Brightness and lifetime signals from encapsulated cells in droplets that fulfil the selection criteria are then sorted using dielectrophoresis (DEP) technique. [158] The FPGA sends a sort signal to trigger a function

generator (Keysight 33509B) which provides a square wave pulse which is amplified 1000x in a high voltage amplifier (TREK), before being sent to the electrodes of the microfluidic device. The flow is biased towards the waste channel, so droplets are only directed to the collection channel when the FPGA sends a signal to trigger a high voltage pulse to DEP electrodes. The fluorescence detection and cell selection regions of the device are shown in Figure 2.2a. Further details on the microfluidic device are provided in Appendix 1: Section 1.



**Figure 2.1:** Schematic layout of the electronics used in this sorter.



**Figure 2.2. Droplet generation and size. (a.)** Camera image shows the typical droplet flow with both excitation beams on. The microfluidic chip is designed such that droplets are biased towards the waste channel. **(b.)** Image of droplets containing Rhodamine B generated with the microfluidic chip. The scale bar indicates 50 μm.

### **2.3c. Instrument operation**

The microfluidic sorter is configured with excitation beams at 450 nm and 561 nm, wavelengths which allow for screening based on green and/or red fluorescence signals respectively. The 561 nm excitation beam is modulated at 29.5 MHz, enabling fluorescence lifetime screening in the red channel. To count the number of droplets passing each channel and monitor the flow (number of droplets per second) throughout an experiment, the laser intensities and PMT voltages were set such that a small portion of scattered laser light from each droplet bleeds through the dichroic mirror and the emission filters, and thus can be detected in both PMT channels. We previously reported fluorescence lifetime sorting in a microfluidic flow cytometer, however, the sorting speed was limited to ~30 cells/s because communication among instruments, target and host computers, calculation of fluorescence phase shifts, and sorting decisions relied on software developed on a LabView platform. [30] In the current sorter, the phase shifts are obtained directly from a high-speed lock-in amplifier, and an FPGA coordinates communication among all electronics and performs sorting decisions. A LabView user interface is designed only for setting selection parameters, acquiring data from the FPGA and real time plotting. As a result, the new instrument operates at ~100-fold higher screening and sorting speeds. For both fluorescence-activated droplet sorting (“brightness sorting”) and fluorescence lifetime-activated droplet sorting (“lifetime sorting”), the FPGA and LabView program are designed such that the sorting thresholds can be set to exclude empty and unwanted droplets for sorting purposes, while counting the total number of droplets and monitoring the flow (number of droplets per second). Both brightness and lifetime measurements have been tested at droplet generation rates up to 4 kHz (~0.7 mL/hr volumetric flow rate) for screening and 2.5 kHz (~0.45 mL/hr volumetric flow rate) for sorting. A typical image of droplets generated at ~2.5 kHz (Figure 2.2b), demonstrates their size uniformity and agreement with the estimated droplet volume ~50 pL which is determined from the droplet generation rate and the 0.45 mL/hr volumetric flow rate. More details about instrument operation are available in Appendix 1: Section 2.

### **2.3d. Cell culture and sample preparation**

The droplet microfluidic sorter can be employed to assay diverse cell types, such as bacteria, phytoplankton, yeast, and mammalian cell lines. To test the performance of this sorter, various FPs with distinct fluorescence lifetime, brightness, and spectra were expressed in *E. coli*. Cells

expressing FPs were prepared at desired concentrations according to the measurement of their optical density (OD) and connected to the aqueous inlet of the microfluidic chip. The details of cell culture and sample preparation are described in Appendix 1: Section 3.

## **2.4. Results and Discussion**

This instrument control software is designed such that one can choose the desired combinations of screening and/or sorting based on emission spectrum, brightness, and red fluorescence lifetime. The scattered excitation light from each droplet can be detected by the PMTs, which allows us to monitor the flow, count the number of droplets, and pair-match two events in green and red channels for a particular droplet. Details of data acquisition and signal processing are described in Appendix 1: Section 4. The performance of brightness and lifetime sorting with different screening/sorting criteria is evaluated here. We also present some examples of the strategy for enriching rare events with multiple cell encapsulation.

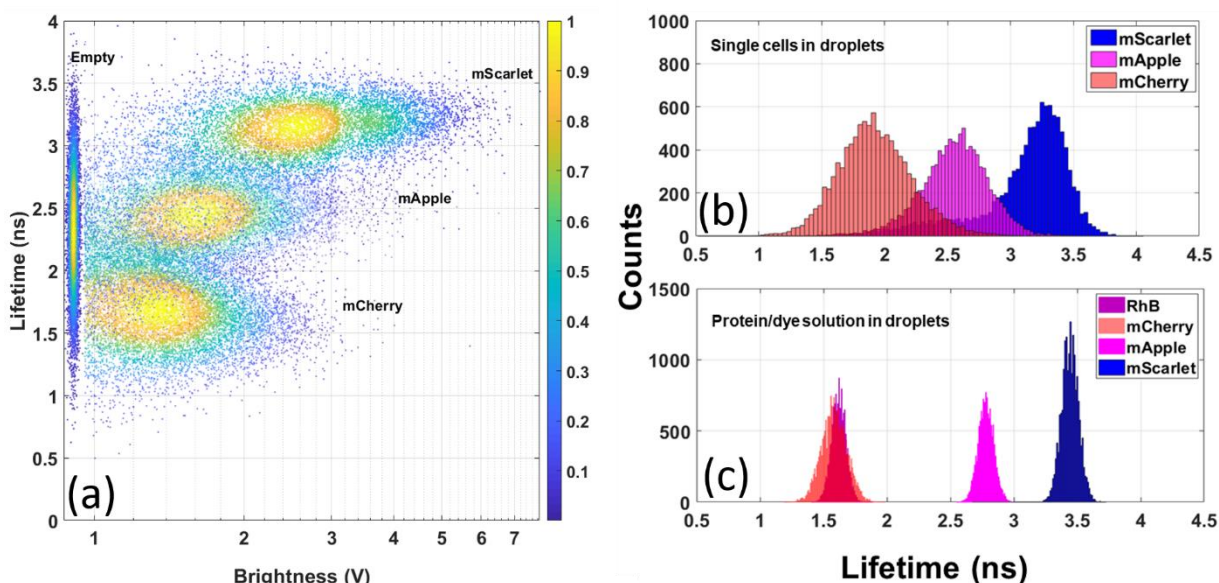
### **2.4a. Performance of two-color brightness-sorting**

To evaluate the performance of brightness detection in the green and red channels, *E. coli* cells expressing EGFP and mScarlet were screened respectively to find the mean brightness in each channel. An approximately 1:1 mixture was prepared and droplets with a brightness threshold greater than mean brightness in red channel was sorted to select mScarlet from  $\sim 10^5$  droplets. The sorted cells were subsequently grown overnight and screened 16 hours after induction of expression to evaluate the sorting efficiency. All screening and sorting experiments were performed at a rate of 2 kHz with an average cell concentration of 0.1 cell/droplet, where 9.5% of the droplets are filled and 95% of filled droplets contain a single cell. The results shown in Appendix 1: Section 5 reveal a sorting efficiency of  $86 \pm 1\%$  averaged from 3 experimental trials, i.e. 86% of re-grown cells have mScarlet and 14% of them have EGFP. The 14% re-grown cells expressing EGFP reflects several factors including the 5% of filled droplets containing multiple cells, varying cytotoxicity for cells expressing different FPs, and the excitation conditions. [174] These issues are discussed in the lifetime sorting section, below.



## 2.4b. Performance of lifetime sorting

The phase shift measured in the frequency domain technique is sensitive to the modulation frequency, [56] transit time of cells passing through the laser beam, and settings of the PMT and lock-in amplifier. Determination of the lifetime and its dependence on these experimental factors is described in Appendix 1: Section 4. *E. coli* cells expressing mCherry, mApple, or mScarlet were screened with brightness and lifetime at a rate of 2.5 kHz. The major population of each RFP is distinguishable by its fluorescence lifetime as shown in Figure 2.3a. The results reveal heterogeneity in both fluorescence brightness and lifetime, as observed in our previous work on other RFPs. [30] The spread of lifetime values is about 0.5-1 ns for these RFPs at full width at half maximum (FWHM) of the histograms in Figure 2.3b.



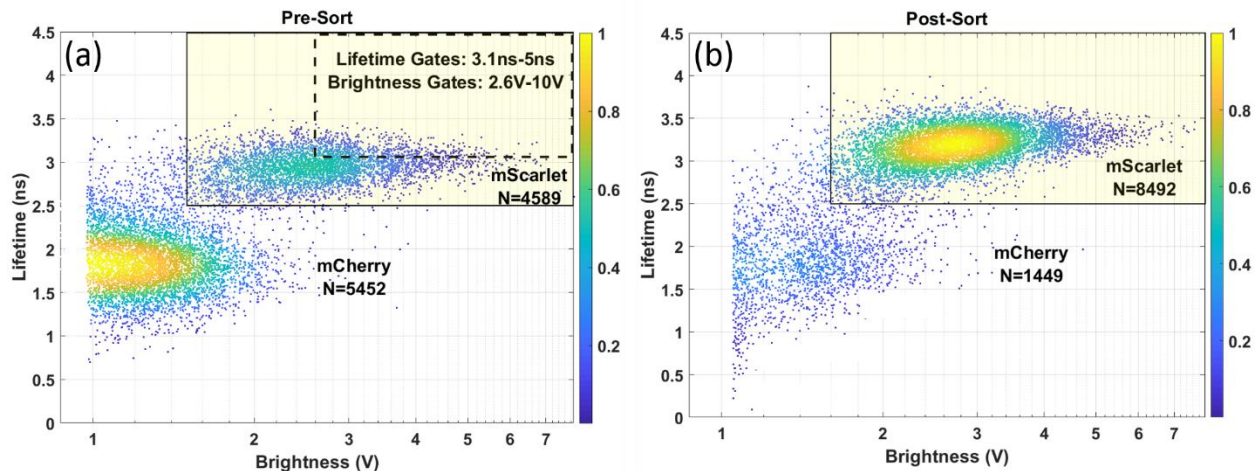
**Figure 2.3. Lifetime screens of FPs.** (a.) Fluorescence lifetime and brightness plots of empty droplets and individual RFPs expressed in *E. coli* screened sequentially ( $10^4$  cells each). Pseudocolor indicates the normalized cell counts with a particular bin of fluorescence lifetime and brightness on the plot, ranging from yellow for the highest to indigo indicating the lowest counts. The mean fluorescence lifetime is 1.7 ns (set as reference), 2.6 ns, and 3.3 ns for mCherry, mApple, and mScarlet respectively. (b.) Corresponding fluorescence lifetime histograms. (c) Fluorescence lifetime histograms of Rhodamine B (RhB) and three purified proteins measured in the microfluidic sorter. The mean fluorescence lifetime is 1.6 ns (set as reference), 1.6 ns, 2.8 ns, and 3.5 ns for RhB, mCherry, mApple, and mScarlet respectively.

The asymmetric histograms of fluorescence lifetime in Figure 2.3b can be understood as an effect resulting from the contribution of scattered excitation light detected along with the fluorescence signal. This effect is modeled with a simulation in Appendix 1: Section 4. Ideally, scattered light has a constant phase shift (which is converted to the fluorescence lifetime) relative to the modulated laser beam due to optical and electronic delays. This is included in the total phase shift by setting the reference phase shift of a bacterial colony expressing mCherry on a plate to 45 degrees. In this particular experiment, the total offset phase shift of empty droplets corresponds to a fluorescence lifetime centered on ~2.35 ns with a wide distribution due to low signal-to-noise ratio (SNR). The scattered light is added to the fluorescence signal and both signals have the same modulation frequency but different phase shift values, so the lock-in amplifier extracts an averaged phase value from the combined signals. The influence of scattered light is more significant at low fluorescence brightness, whereas the average lifetime value approaches the actual fluorescence lifetime value as the fluorescence brightness increases.

The distribution of lifetime measured from a single-FP population can be attributed to cellular heterogeneity, excitation condition and electronics. Cellular heterogeneity is an intrinsic biochemical property that can only be resolved in single cell analysis methods such as this microfluidic droplet sorter. On the other hand, the noise originating from the excitation condition may be further reduced. The diameter of the droplet is estimated to be ~46  $\mu\text{m}$ , but the Rayleigh length of the excitation beam is ~10  $\mu\text{m}$ , hence the location of the cell inside a droplet could lead to variations in fluorescence brightness resulting in uncertainties in lifetime measurement. Theoretically the lifetime is independent of fluorescence signal level, but in practice the scattered excitation light affects weaker fluorescence signals more than stronger ones as discussed above. We further investigated the spread of lifetime due to electronics by performing *in vitro* measurements. In addition to eliminating the cellular heterogeneity, *in vitro* measurements also minimize the fluctuations from excitation condition since a droplet has homogeneous fluorophore concentration and the Rayleigh length is always within the droplet. It is worth noting that various *in vitro* assays can be performed with a droplet platform, but it is difficult to perform them with a continuous stream cytometry. Three purified proteins, mCherry, mApple, and mScarlet, and an organic dye, Rhodamine B, were screened for fluorescence lifetime using the sorter. The histogram of fluorescence lifetime is shown in Figure 2.3c, with FWHM ~0.1 ns for Rhodamine B and ~0.2-0.3 ns for FPs. The widespread in lifetime for mCherry is likely due to its low SNR resulting from

a low quantum yield (hence low molecular brightness). Nonetheless, the FWHM of fluorescence lifetime measured from an *in vitro* experiment is much narrower than that from a cellular measurement. The result indicates that the uncertainty originating from electronics is significantly less than other sources. This also suggests that the lifetime resolution for cellular screening could be improved by reducing the droplet size and/or expanding the beam size to extend the Rayleigh length to ensure that the encapsulated cells are within the Rayleigh length, i.e. an improved uniform excitation condition. This effect will be reduced with larger cell types such as yeast or mammalian cells. Finally, note that the disagreement in the average lifetime among cellular and *in vitro* measurements suggests that the cellular environment differs from the *in vitro* environment. For example, fluorescence lifetime of FPs varying with environmental pH [130, 131] and refractive index [132, 133] has been reported and used for sensing and imaging applications. Details of the *in vitro* experiment including the comparison of fluorescence lifetime measured using the sorter and Time-Correlated Single Photon Counting (TCSPC) are described in Appendix 1: Section 5.

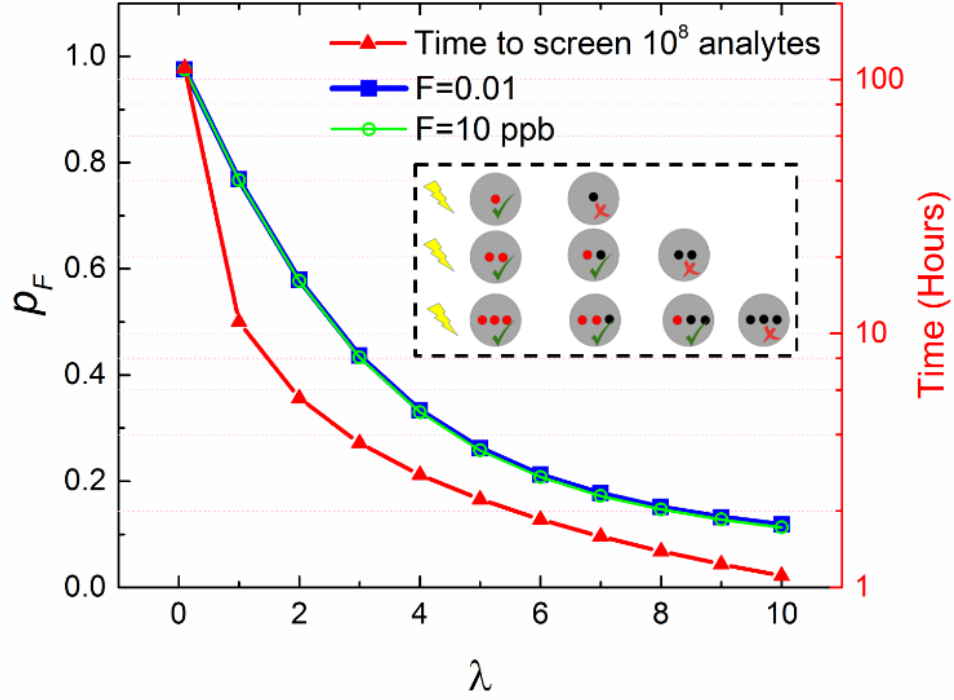
To demonstrate the performance of lifetime-based sorting, *E. coli* cells expressing mScarlet or mCherry were mixed in a ~1:1 proportion and sorted at 2.5 kHz with two parameters, fluorescence lifetime and brightness, at an average concentration of 0.1 cell/droplet. This sort rate represents the fastest fluorescence lifetime droplet sorting reported to date. Approximately  $3 \times 10^3$  droplets were sorted from  $\sim 2.5 \times 10^5$  droplets with the selection gates set to the mean brightness value and mean fluorescence lifetime of mScarlet. The sorted cells were subsequently grown, expressed for 16 hours and re-screened to evaluate the sorting efficiency. The screening results before and after sorting are shown in Figure 2.4, demonstrating an 85% sorting efficiency. The experiment was additionally repeated 3 times with a new mixture, sorting mScarlet or mCherry, and the average efficiencies were  $80 \pm 1\%$  and  $97 \pm 1\%$  respectively, as described in Appendix 1: Section 5. The discrepancy between sorting mScarlet and mCherry can be attributed to the process of re-growing and expressing enriched cells in the experiment with the assumption that bacteria expressing different FPs have the same growth rate, which may not be accurate. Some mCherry mutants, mApple, and EGFP have been reported to show a range of cytotoxicities when expressed in *E. coli*. [174] The difference between two batches of mScarlet enrichment experiments may be due to the flow condition, the biological variation (two biological duplicates in two batches of experiments) and the uncertainty of cell concentration in the sample preparation causing variations in  $\lambda$ , which affects the sorting efficiency that will be further discussed below.



**Figure 2.4. Fluorescence lifetime versus brightness scatter plots of mixed cells before and after sorting.** Solid boxes indicate the thresholds for counting cells expressing mScarlet. N is the number of cells expressing each RFP. **(a.)** Mixture of *E. coli* cells expressing mCherry and mScarlet before sorting. The dashed box indicates the two-parameter sorting gates. **(b.)** Screening results after sorted cells were grown overnight and expressed for 16 hours. The brightness threshold was set slightly higher than pre-sort to exclude the stronger scattered excitation signals from droplets in the post-sort screening, because changing microfluidic chips introduces variations in the focus of the excitation beam and thus the amount of scattered light.

### 2.4c. Strategy for enriching rare fluorescent events

For a large library containing rare events, the overall throughput can be greatly increased by sorting droplets by encapsulating multiple cells in a single droplet as an initial round of enrichment. The efficiency of this strategy can be estimated by considering the Poisson distribution, the combination of cells resulting fluorescent droplets, and the percentages of fluorescent cells in a library. We present a schematic to consider the combination of cells encapsulated in droplets illustrated in the inset of Figure 2.5. A droplet will be detected with fluorescence as long as it contains one or more fluorescent cells. The probability of number of cells (N) encapsulated in a droplet is  $\text{Prob}(N) = (e^{-\lambda} \times \lambda^N) / N!$ , where  $\lambda$  is the average number of cells per droplet.



**Figure 2.5. The efficiency of enrichment with various initial fraction of target analyte.** (cells, molecules, or beads). Presented also, the required enrichment time as a function of average number of cells per droplet. Inset (dashed box): Illustration of cells encapsulated in droplets. The red and black dots indicate fluorescent and non-fluorescent cells, respectively. The green check and red cross marks indicate fluorescent and non-fluorescent droplets.

Assuming a library with initial fraction  $F$  of fluorescent cells, the probability of finding fluorescent cells after sorting,  $p_F$ , is

$$p_F = \sum_{n=1}^{\infty} \frac{\sum_{i=1}^n \binom{n}{i} \cdot i \cdot F^i \cdot (1-F)^{n-i}}{n \cdot \sum_{i=1}^n \binom{n}{i} \cdot F^i \cdot (1-F)^{n-i}},$$

where  $i$  is the number of fluorescent cells and  $n$  is the number of cells per droplet. Since the probability of encapsulated cells per droplet decreases quickly with the increasing number of encapsulated cells, the  $p_F$  can be numerically calculated using  $n \leq 50$  for  $\lambda \leq 10$ . The Poisson distribution for  $\lambda \leq 10$  is plotted in Appendix 1: Section 6. The efficiency of the multiple-cell encapsulation enrichment, which is indicated by the improvement in the fraction of fluorescent cells after sorting (i.e.  $p_F$ ), is estimated with  $F=0.01$  and  $F=10$  ppb for various  $\lambda$  as shown in Figure 2.5. The results indicate that with one round of sorting, the fluorescent cells in the library can be

enriched to about the same fraction regardless of the initial fraction  $F$ , thus this selection strategy is more powerful for enriching rarer events from a large pool (i.e. small  $F$ ). It is not surprising that the enrichment efficiency is significantly affected by the average number of cells per droplet ( $\lambda$ ), but the influence from the fraction of fluorescent cells in the original library is not significant, because the selected droplets all contain fluorescent cells. Assuming a sorting speed of 2.5 kHz, the time required for screening  $10^8$  cells as a function of  $\lambda$  is plotted in Figure 2.5. The result clearly shows that the time can be drastically reduced by including multiple cells in a droplet. The estimation of  $p_F$  only considers the statistical probability, i.e. the number of screened cells is much larger than the inverse of the initial fraction  $F$ . Such enrichment efficiency,  $p_F$ , is estimated to hold for enriching  $\geq 0.5$  ppm targets from  $10^8$  cells, the limit for current throughput to complete enrichment in a few hours without losing cell viability, in Appendix 1: Section 6. However, this does not limit the application of the enrichment strategy from sorting smaller fraction of rare events. With a smaller fraction of rare events, the enrichment efficiency may deviate from the expected value plotted in Figure 2.5, but it still provides approximately the same order of magnitude of enrichment efficiency as illustrated in Appendix 1: Section 6.

To further illustrate the power of this enrichment strategy, we consider two examples of rare events that fluoresce or exhibit a distinct fluorescence lifetime relative to the main fluorescent population. Assume the enrichment is carried out with brightness or lifetime sorting operating at 2.5 kHz with an average 4 cells/droplet encapsulation. In the first example, we assume that the fraction of the rare events is 1 ppm. It would take less than 3 hours to enrich rare events from a  $10^8$  population, resulting in a subset of 100 fluorescent cells mixed with 203 unwanted cells ( $p_F=0.33$ ), i.e.  $3.3 \times 10^5$ -fold enrichment ( $p_F/F$ ) in one round of sorting. The enriched subset can be further cultured, analyzed, or sorted with single cell resolution to isolate the final, purified population. In the second example, we consider a cell-based library containing  $33 \times 10^6$  distinct mutants. To ensure the enrichment covers 95% of this library, at least 3 times of the library size must be screened, [175] which is  $\sim 10^8$  cells. Assuming the desired clones comprise 1% of the original library, this enrichment reduces the library size from  $33 \times 10^6$  down to  $1 \times 10^6$  within 3 hours with  $0.33 \times 10^6$  fluorescent cells, thus a 33-fold enrichment. The enriched library can be further analyzed or sorted at  $\lambda=0.1$  (single-cell resolution) using brightness or lifetime sorting. Using the conventional encapsulation strategy ( $\lambda=0.1$ ) without the enrichment, it would take  $\sim 117$  hours to complete the selection in both examples with brightness or lifetime sorting at the speed of 2.5 kHz developed in

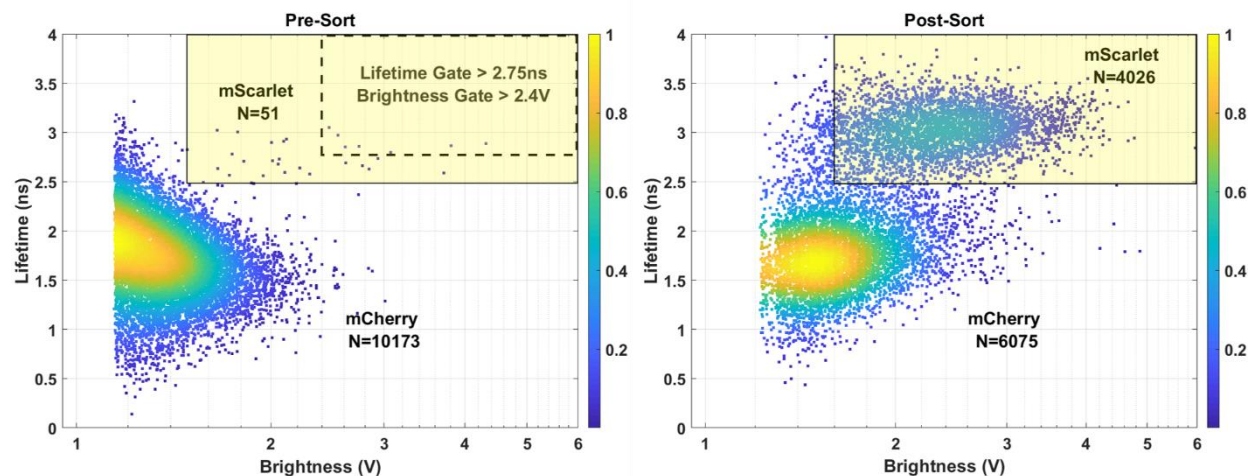


this work. It would take 50 times longer (~5848 hours) for a recently reported lifetime droplet sorting to perform the selection. [164] Using a state-of-the-art droplet sorting at 30 kHz, [163] the selection would require ~10 hours, which is more than 3 times longer than the lifetime enrichment demonstrated here, to complete a brightness-only selection in single cell resolution without fluorescence lifetime information. The combination of this new sorting technology and enrichment strategy enables fast multiparameter analysis and selection of rare events from a  $10^8$ -member population based on fluorescence lifetime, brightness, and spectrum, as a preparation for further investigation and sorting with single cell resolution on a single instrument.

To demonstrate the effectiveness of this strategy, we enriched mScarlet from a mixture of EGFP and mScarlet transformed in *E. coli* using dual color brightness sorting. Since EGFP does not emit red fluorescence, EGFP can be regarded as the non-fluorescent population and mScarlet as the rare fluorescent population observed in the red channel. The number of EGFP cells can be counted in the green channel since only EGFP contributes to the green emission. Thus, the fraction of mScarlet (i.e. the fluorescent events in the red channel) in the mixture was determined to be  $F \sim 0.01$ . After one round of enrichment with  $\lambda=3$  encapsulation, the sorted cells were subsequently grown and screened with  $\lambda \leq 0.1$  encapsulation. The mScarlet population was enriched to an average  $35 \pm 4\%$ , which agrees with the expected value ( $p_F \times 0.86$ )  $\sim 37\%$ , considering the 86% efficiency of single cell two-color sorting described earlier. Detailed experimental protocols can be found in Appendix 1: Section 6.

The enrichment strategy can also be applied in lifetime sorting when the rare events have a distinct fluorescence lifetime from the major population, despite the overlap in emission spectra and brightness. We demonstrate the enrichment of rare cells expressed with mScarlet from a mixture of mCherry and mScarlet, which have large overlap in both emission spectra and cellular brightness. The first test was carried out the same day using the same batch of sample generating results in Figure 2.4. The fraction of mScarlet in the mixture before enrichment was estimated to be  $F \sim 5 \times 10^{-3}$ . The enrichment was performed with  $\lambda=3$  at 2.5 kHz, and the sorted cells were subsequently grown, expressed, and screened with  $\lambda \leq 0.1$ . We attained an enrichment of the mScarlet population to 40% (Figure 2.6), which is consistent with the expected value, including the 85% efficiency of single cell lifetime sorting demonstrated in Figure 2.4, ( $p_F \times 0.85$ )  $\sim 37\%$ . Another enrichment for rare mScarlet was performed using the second batch of sample with  $F \sim 5 \times 10^{-3}$ , resulting in an average enrichment of the mScarlet population  $30 \pm 5\%$ , in agreement

with the expectation ( $p_F \times 0.80$ )  $\sim 35\%$ . Details of experimental protocols for this assay are described in Appendix 1: Section 6.



**Figure 2.6. Fluorescence lifetime versus brightness scatter plots of rare mScarlet enrichment.**

Solid boxes illustrate thresholds for counting cells expressing mScarlet. N is the number of cells expressing each RFP. (Left) Mixture of *E. coli* cells expressing mCherry and mScarlet before enrichment. The dashed box indicates the two-parameter sorting gates. (Right) Screening results after enriched cells were grown overnight and expressed for 16 hours.

## 2.4d. Enrichment of an RFP library

The directed evolution of FPs often involves the screening of cell libraries with rare bright clones. Library size increases exponentially with the number of target residues, and FP libraries are typically found to have a narrow fitness landscape, [124] i.e. the fraction of fluorescent clones dramatically decreases as the mutational space increases due to protein mis-folding, incomplete chromophore maturation, and other photophysical factors. We used this sorter to enrich the population of fluorescent RFP mutants and select the brightest ones for further development. Taking mScarlet-I as the template, we constructed a site-directed library with the size  $\sim 1.7 \times 10^7$ , which requires screening  $> 5.1 \times 10^7$  cells to cover 95% of the library size. In our previous studies of site-directed and/or error-prone PCR libraries of red FPs (RFPs), some non-fluorescent colonies were observed to grow larger than fluorescent ones on plates, likely due to variations in cytotoxicity of various mutations in RFPs. [174] Therefore, we expect reduced sorting efficiency



due to the re-growth and expression processes after enrichment as described above. With this consideration in mind, we decided to load the droplets with  $\lambda=3$ , and a total number of  $\sim 8 \times 10^7$  *E. coli* cells expressing this RFP library was screened in two batches (ensuring the health of cells) to enrich fluorescent cells at  $\sim 2$  kHz. The proportion of fluorescent cells was enriched from initially  $\sim 5\%$  to  $\sim 30\%$ . This is lower than the expected, empirically corrected enrichment efficiency ( $43\% \times 0.86$ ) 37% for  $\lambda=3$ . The enriched population underwent 3 more rounds of enrichments with higher thresholds in fluorescent brightness with  $\lambda=1$  or  $\lambda=0.1$  at 2 kHz, resulting in  $>95\%$  fluorescent population. The final round of sorted cells was re-grown overnight then expressed on agar plates. Three distinct mutants were identified from the agar plates for further development. More information on the library and the detailed enrichment procedure are provided in Appendix 1: Section 7.

This platform is sufficiently flexible to support further enhancements. For example, additional excitation wavelengths with RF modulation can be implemented to expand the information content in both spectral and fluorescence lifetime dimensions. Furthermore, it is possible to increase the sorting speed further by modifying the microfluidic chip design. In particular, brightness sorting at 30 kHz has been demonstrated in a design where the hard divider is replaced with a gapped divider to separate outlets. [163] In addition, increasing the modulation frequency of the excitation beam shortens the phase acquisition time, and therefore increases the fluorescence lifetime detection speed. As such, a modulation frequency of 100 MHz could support a  $\sim 3.4$ -fold increase in sorting speed. However, the modulation frequency may set the limit for the throughput of fluorescence lifetime measurement. When the modulation frequency increases to higher than 100 MHz, the period of the modulation wave becomes less than 10 ns, the same order of magnitude as the fluorescence lifetime of commonly used fluorophores. This may disturb the phase measurement under a strong excitation rate used in frequency domain measurement. On the other hand, to further increase the sorting speed to  $\geq 10$  kHz, the adjoining scattering or fluorescence signals are  $\leq 100$   $\mu\text{s}$  apart. In current setup, the FWHM of the scattering and fluorescence signals is approximately 25  $\mu\text{s}$  at 2 kHz, which is sufficiently small for sorting at 10 kHz. If needed, decreasing the droplet size can not only reduce the noise as previously discussed, but also shorten the transient time of the droplet and cells since they only pass the Rayleigh length region, resulting in narrower FWHM of the scattering and fluorescence signals. Thus, it is feasible to improve the throughput of this multiparameter droplet sorter to  $\geq 10$  kHz.

## 2.5. Conclusion

We developed a multiparameter microfluidic droplet sorter combining the detection of fluorescence lifetime, brightness, and spectra. The throughput of the fluorescence lifetime measurement and sorting, up to 4 kHz for screening and 2.5 kHz for sorting with current chip design, is greatly enhanced by using a FPGA for the communication among all electronics and sorting decisions. This is the fastest fluorescence lifetime droplet screening and sorting speed to date. The high-throughput fluorescence lifetime droplet sorting opens the opportunity of integrating fluorescence lifetime detection with other high throughput detection methods in a microfluidic droplet platform to increase the information content of biological and biomedical assays with single cell resolution. We have also proposed a novel multiple-cell encapsulation strategy enriching the rare events to overcome the obstacle of droplet sorting throughput limited by the nature of Poisson distribution for random cell/molecule encapsulation – by taking the advantage of Poisson statistics. The effect of enrichment increases tremendously as the fraction of rare events decreases. The efficiency and precision of enrichment can be quantitatively controlled if the rare event frequency is estimated before sorting. The enrichment strategy has been demonstrated to be effective in both brightness and lifetime sorting. Combining the enrichment strategy and the multiparameter microfluidic platform allows one to analyze and enrich rare events from a population  $>10^8$  within a few hours. Though the enrichment does not provide single cell/analyte resolution, it greatly reduces the time required to search for rare events, thus is an efficient way to analyze or prepare rare events for further investigation or selection with single cell/analyte resolution. It is also feasible to improve the throughput of the multiparameter sorting to  $\geq 10$  kHz. Together with the new sorting strategy, the speed of droplet-encapsulated rare events analysis and enrichment can potentially exceed FACS, achieving an unprecedented throughput for microfluidics-based cell sorting.

## Chapter 3

# Engineering a brighter version of FusionRed using lifetime flow cytometry and structure guided mutations

### 3.1. Publication Note

Parts of this chapter have been adapted from the article “*Engineering of a Brighter Variant of the FusionRed Fluorescent Protein Using Lifetime Flow Cytometry and Structure-Guided Mutations.*” **Mukherjee, S.**; Hung, S. T.; Douglas, N.; Manna, P.; Thomas, C.; Ekrem, A.; Palmer, A. E.; Jimenez, R. *Biochemistry* **2020**, 59 (39), 3669–3682.  
<https://doi.org/10.1021/acs.biochem.0c00484>.

### 3.2. Introduction

The availability of genetically-encoded fluorophores such as fluorescent proteins (FPs) initiated a revolution in biological imaging. [2] Routine use of FPs in assays involving technologies such as Förster resonance energy transfer (FRET), [176] fluorescence lifetime imaging microscopy (FLIM), [177, 178] molecular sensing, [179] and nanoscopy, [180] make them indispensable tools for biological research. Current efforts focus on developing FPs with excitation and emission in the far-red/near-infrared wavelengths and with photophysical properties such as photoswitching and fluorescence intermittency optimized for super-resolution imaging modalities. [23, 24, 181, 182] Nevertheless, all imaging applications benefit from increased cellular brightness, which is strongly dependent on molecular brightness (defined as the product of the molar extinction coefficient and the fluorescence quantum yield). [57, 101] Lifetime-based selection methods have exploited a correlation between fluorescence lifetime and quantum yield to develop FPs with higher quantum yield such as NowGFP, [183] mTurquoise2, [77] and mScarlet, [35] which is the brightest red FP observed to date.

FPs selected for higher fluorescence lifetime generally show a decreased rate constant of non-radiative decay ( $k_{\text{non-rad}}$ ), ultimately resulting in increased quantum yield and higher brightness. However, quantum yield is also linearly related with  $k_{\text{rad}}$ , therefore there is an interplay between

the absorption probability and radiative emission probability. The Strickler–Berg equation formalizes the relationship between the  $k_{\text{rad}}$  and extinction coefficient, as well as other spectral properties such as the energies and profiles of the absorption and emission bands. [60] For example, blue shifts and decreased spectral width can lead to higher probabilities of radiative decay. [184] The relationship between spectral characteristics and the rate constants of population decay have recently been discussed for FPs. [59, 66] An engineering strategy that maximizes the radiative rate constant while simultaneously minimizing the non-radiative rate constant should be an effective way to brighten a FP.

These issues are particularly acute for red FPs (RFPs), which generally have lower values of fluorescence lifetime and quantum yield compared to shorter wavelength variants, and are therefore not as bright. [185] The RFP chromophores contain an acylimine moiety, which expands their electronic conjugation and leads to a ~50 nm red shift in their absorption and emission spectra with respect to green analogues like EGFP. [53, 186] Many potential non-radiative decay mechanisms that lead to lower quantum efficiencies in such systems have been investigated, including transitions to dark states; [13, 187] charge accumulation and twisting of the acylimine moiety; [92] changes in hydrogen bonding patterns; and electrostatic, steric and conformational effects associated with their increased number of vibrational degrees of freedom. [14, 74]

Cellular brightness depends on both photophysical and non-photophysical factors such as translational efficiency, protein folding and chromophore maturation. [29] FPs like FusionRed-M (FR-M) and mScarlet-I were developed to enhance cellular brightness over their brighter parents FR-1 and mScarlet, respectively. [30] Although there have been many efforts to improve the cellular and molecular brightness of FPs, performance in cell biology applications may still suffer from cellular toxicity, poor or over expression in certain cellular contexts and *in cellulo* oligomerization. [29, 101] As discussed below, in the course of evaluating the performance of bright and commonly used RFPs, we found a trade-off between cellular brightness and localization fidelity of the GalT fusions of mScarlet, which led us to question whether the high cellular brightness associated with very high expression level leads to appearance of mis-localized FPs.

FusionRed (FR) was developed as a non-cytotoxic RFP that shows efficient and correct localization in multiple fusion constructs in different organisms. [188] However, this FP has not been widely adopted because it is relatively dim, in part due to its low quantum yield (0.19). We

recently developed the FR-M variant (FR-L177M, numbered with respect to mCherry and FR-L175M numbered with respect to FR; See Appendix Table A2.1), which is 1.9-fold brighter than FR in HeLa cells. [30] In this study, we used structure-guided engineering strategies to identify substitutions M42Q and C159V (sequence numbering with respect to FR) in the vicinity of the chromophore that individually increase the fluorescence lifetime of FR. These mutations led to the development of the bright triple mutant FR-MQV. We performed detailed *in vitro* photophysical analysis which corroborated the increase in the molecular brightness in FR-MQV based on minimizing  $k_{\text{non-rad}}$  and increasing the  $k_{\text{rad}}$ . Additionally, we investigated the cellular properties of FR-MQV and found it is 5-fold brighter in HeLa cells (using FACS) compared to the parent FR, while preserving its cellular properties such as low cytotoxicity and high-fidelity localization.

### **3.3. Experimental**

#### **3.3a. Mutagenesis, cloning and construct development**

##### **i. Yeast Constructs**

The QuickChange site-directed mutagenesis method was used for making point mutations using PfuTurboDNA polymerase and a thermocycler. Libraries with multiple site-directed targets were created using a splicing overlap extension reaction. Primers were designed to introduce the desired mutations and the initial PCRs generate overlapping gene segments that are used as template DNA for another PCR to create a full-length product. Fresh competent yeast cells (*Saccharomyces cerevisiae* BY4741) were prepared prior to electroporation. Cells, DNA and cut pYestDest52 vector were combined and left on ice for 5 min. Electroporation conditions (Bio-Rad Gene Pulser Xcell) were as follows: C = 25  $\mu$ F, PC = 200 Ohm, V = 1.5 kV (in 0.2 cm cuvettes). Cells were passed twice prior to expression. Mutants were transferred to the pBad-His vector for expression/Ni-NTA protein purification.

##### **ii. Bacterial Expression Constructs**

For bacterial expression, FR-Q and FR-MQ were made using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) with pBad-FR and pBad-FR-M, respectively, as the templates and the

primers tCGCCACACAGGACACAAG and cCACATATGTCTCATCGTCAGC. FR-MV was designed with similar protocols by inserting the C159V substitution into FR-M, with appropriate primers. FR-MQ C159V was also made via Q5 mutagenesis with the above primers and FR-M C159V as the template. The M42Q mutation was also made in pBad-mKate via Q5 mutagenesis and the corresponding A44Q mutation was made in pBad-mScarlet and pBad-mCherry using overlap extension PCR.

### **iii. Mammalian Expression Constructs for FACS**

FR-MQ and FR-MQV were expressed as histone H2B fusion proteins in HeLa cells. Specifically, the FR-MQ and FR-MQV mutants were PCR amplified from the pBad constructs with the upstream primer GGTATGGCTAGCATGACTGGTG and a reverse primer that introduces a NotI site adjacent to the stop codon (acatGCGGCCGCTCATTTCCCTCCATC). Similarly, mScarlet-I (mScarlet T74I) was PCR amplified from a Q5-generated pBad construct (primers AGGGCCTTCATCAAGCACCCC and GCAGCCGTAC-ATGAACTGAGG) with the same upstream primer and a reverse primer that introduces an XcmI site adjacent to the stop codon. The FR-MQ and FR-MQV products were cut with BamHI and NotI and ligated into BamHI/NotI cut piggyBac-H2B and the mScarlet-I product was cut with BamHI and XcmI and ligated into BamHI/XcmI cut piggyBac-H2B.

### **iv. Mammalian Expression Constructs for OSER**

To assess protein aggregation in mammalian cells, the mutants were expressed as fusions with CytERM in U2OS cells. FPs were amplified from the pBad construct with an upstream primer that introduces an AgeI site upstream of the start codon (gcatACCGGTCGCCACCATGGTGTCCG-AGCTGATTAAGG) and the reverse primer described above that introduces a NotI site. The AgeI/NotI cut PCR product was ligated into AgeI/NotI cut CytERM vector and transfected into U2OS cells as described above for HeLa cells.

## **v. Mammalian Expression Constructs for GalT-FP fusions**

To localize our mutants to the Golgi, they were expressed as GalT fusions in U2OS cells. The BamHI/NotI cut FR, FR-M and FR-MQV PCR products described above were ligated into BamHI/NotI cut GalT vector. mScarlet-I was cut from pBad with BamHI and SalI and ligated into BamHI/SalI cut GalT-mScarlet.

## **3.3b. Microfluidic based selection from site directed libraries**

### **i. Library Targets**

We targeted the positions 159, 161, 196 and 198 in FR (using FR numbering; Appendix 2: Section 2), expressed the libraries in yeast (*Saccharomyces cerevisiae*) and screened this library on a lifetime flow cytometer. [30] Screening revealed the presence of brighter clones with longer and shorter lifetime than parent FR (lifetime ~ 2.05 ns). We performed another two rounds of FACS enrichment on a BD FACSAria Fusion Cell Sorter to remove the dim/non-fluorescent clones. Position 159 was mutated to I, L, V, F, M, C, A, G, T, S, W and R; position 161 was mutated to I, L, M, Q, N, H and K; position 196 was mutated to I, V, A and T; and position 198 was mutated to all possible amino acids. The library size was ~ 7,000 clones.

### **ii. Post-Microfluidic Sorting**

After selection, the collected yeast cells were grown in liquid culture and then plated, and 25 clones with unique lifetimes were picked from these sorter-enriched libraries using a lifetime assisted plate-based screen discussed in a previous work. [30] The FR site-directed (FSD) clones with unique DNA sequences were further characterized (Appendix Table A2.2).

### **3.3c. Cell growth, transformation/transfection and protein purification protocols**

#### **i. Bacterial Transformation and Growth**

Bacteria (competent *E. coli*, Top10 strain) were transformed with the DNA encoding a FP of interest in the pBad-His vector. Roughly 2–5  $\mu\text{L}$  of DNA ( $\sim 80\text{ ng/mL}$  concentrations) were slowly pipetted into  $\sim 50\text{ }\mu\text{L}$  of competent cells (Invitrogen). The cells were left on ice for 20 minutes, heat shocked for 45 seconds at  $42^{\circ}\text{C}$  and then grown in antibiotic -free medium for 45 minutes. The cells were then plated ( $\sim 25\text{--}50\text{ }\mu\text{L}$ ) onto ampicillin-containing LB agar plates and grown at  $37^{\circ}\text{C}$  overnight. Colored colonies were picked from these plates and grown in 2XYT medium containing ampicillin overnight at  $37^{\circ}\text{C}$  and 230 RPM. The next morning, 1 mL of this culture was added to 100 mL of fresh 2XYT with ampicillin, grown for 3 hours at  $37^{\circ}\text{C}$  to achieve an OD of  $\sim 0.6$ , and then 1 mL of 20% arabinose was added to the culture to initiate expression. The temperature of incubation was lowered to  $28^{\circ}\text{C}$  to slow down bacterial growth and help protein folding and chromophore maturation. Depending on the maturation rate of the FP, they were grown at this temperature for 20–30 hours.

#### **ii. Bacterial Cell Lysis and Protein Purification**

The induced cell cultures were spun down at 8,000 RPM for 20 minutes at  $4^{\circ}\text{C}$  and the cell pellet frozen at  $-30^{\circ}\text{C}$  to ease lysis. B-PER Bacterial Protein Extraction Reagent (ThermoFisher) was used to lyse the cells in the presence of protease inhibitor. The cells were lysed for 1 hour at room temperature and then spun down at 11,000 RPM and  $4^{\circ}\text{C}$  for 15 and then 20 minutes to remove the cellular debris. The supernatant containing the 6x-His-tagged protein was filtered with a  $0.45\text{ }\mu\text{m}$  polyethersulfone membrane syringe filter and incubated with Ni-NTA agarose for 1 hour on ice. The resin was then loaded into a column, washed with 10- and 20-mM imidazole and then the proteins eluted with 250 mM imidazole. The imidazole was removed using PD-10 desalting columns (GE HealthCare) or 24 hours of dialysis using SnakeSkin dialysis tubing (ThermoFisher) into Tris-HCL buffer (pH 7.4) or saline PBS buffer. These samples were used for *in vitro* photophysical analyses.



### **iii. Mammalian cell Growth and Transfection**

HeLa/U2OS cells were cultured in RPMI medium (Gibco Life Technologies) supplemented with penicillin/streptomycin (Gibco Life Technologies) and 10% heat-inactivated fetal bovine serum (Sigma-Aldrich) at 37°C with 5% CO<sub>2</sub> plus humidity. For imaging experiments, U2OS cells were grown in 35 mm imaging dishes (made in-house from Corning 35 x 10 mm dishes with VWR 18 x 18 mm #1.5 cover slips). All CyTERM constructs were transiently transfected for 18–24 hours using Lipofectamine 3000 (Invitrogen) or TransIT-LT1 transfection reagent (Mirus) according to the manufacturer's instructions. For FACS experiments, we used HeLa cells transiently transfected using the TransIT-LT1 reagent (Mirus, catalog #MIR2304) and prepared for FACS analysis after 48 hours.

### **3.3d. *in vitro* photophysical measurements**

#### **i. Instrumentation**

Absorption spectra were collected on a Cary5000 UV-Vis Near IR Spectrophotometer using a double beam mode with matched cuvettes and blank subtraction. Samples were diluted using 1X -Tris-HCl Buffer (pH ~7.4) and absorbance was measured at optical densities (ODs) between 0.05 and 0.25 to maintain measurements in the linear range of the Beer–Lambert's law. Fluorescence measurements were performed with a HORIBA Jobin Yvon Fluorolog-3 FL3-222.

#### **ii. Extinction Coefficient Measurements**

Alkali denaturation was used to estimate the ratiometric values of maximum extinction coefficient ( $EC_{max}$ ) for the samples. An average of three or more independent measurements was performed for FPs and is reported with a standard deviation error. FPs with fewer measurements are reported without an error bar. To measure the  $EC_{max}$ , the following protocol was used: (a) Blank1: 900  $\mu$ L Tris-HCl buffer (pH 7.4) spectrum was recorded.; (b) Blank2: 900  $\mu$ L Tris-HCl buffer (pH 7.4) + 100  $\mu$ L 10 M NaOH (pH~14); (c) 900  $\mu$ L Tris-HCl buffer (pH 7.4) + a few  $\mu$ L of concentrated pure protein sample was added to adjust the absorbance to a value of OD ~0.1. A spectrum from 300–700 nm was recorded.; (d) 100  $\mu$ L of 10 M NaOH was added to this solution and a spectrum

in the same range was recorded immediately. Kinetic effects start playing a role on delaying the spectral measurement as degradation product peaks are known to drift in amplitude and wavelength over time.

### iii. Calculation of $EC_{max}$

We performed titration-based  $EC_{max}$  calculations for FR-M, FR-MQV and mScarlet-I. The numbers obtained compare well to the values measured by the one-step alkali denaturing method. [101] The values were based on the mathematical relationship stated below, as FPs of the FR family are known to exhibit backbone cleavage. [188] To verify if this method was valid for FR-MQV we also performed SDS-PAGE (Appendix Figure A2.1), where purified proteins (~10  $\mu$ g) were run on TruPAGE precast 4–20% gradient acrylamide gels (Sigma-Aldrich) in TEA–Tricine running buffer with the Spectra Multicolor Broad Range protein ladder (ThermoFisher). For other FPs that do not exhibit backbone cleavage, like mScarlet, the absorption at 380 nm was disregarded.

$$\epsilon_{maxRFP} = \frac{Abs_{maxRFP}}{\left(\frac{Abs_{380\text{ nm}}}{\epsilon_{380\text{ nm}}}\right) + \left(\frac{Abs_{450\text{ nm}}}{\epsilon_{450\text{ nm}}}\right)}$$

### iv. Quantum Yield Measurements

Freshly prepared or flash frozen purified protein was diluted with Tris-HCl buffer (pH 7.4) in a 1 cm path length Quartz cuvette to an OD of ~0.1. A matched cuvette was used for baseline correction to measure absorption spectra. The same cuvette with the solution in it was transferred to the fluorimeter for collecting fluorescence spectra. After each absorption and emission scan, 200–250  $\mu$ L of the sample was removed and replaced with fresh buffer to create a step dilution. This step dilution was repeated 4–5 times for each sample. RFPs were excited at 520 nm such that the entire emission spectrum was recorded for each FP (even for blue-shifted RFPs) with high enough absorption. For each FP, integrated fluorescence was calculated for the area under the RFP emission feature on the emission spectra and was plotted against the corresponding OD at 520 nm from the absorption spectrum. The integrated fluorescence versus OD plot can be fitted with a

straight line (Appendix 2: Figure A2.2) of the form:  $y = slope \times x$ , where  $y$  is the integrated fluorescence and  $x$  is the OD.

$$\phi_{sample} = \frac{\phi_{reference} \times m_{sample} \times \eta_{sample}^2}{m_{reference} \times \eta_{reference}^2}$$

Where the  $m$  is the slope from a linear fit and refractive index of the sample and the reference are  $n_{sample}$  and  $n_{ref}$ , respectively. mCherry ( $\phi = 0.22$  in Tris-HCl, pH 7.4), Cresyl Violet (quantum yield = 0.54 in EtOH) and mScarlet ( $\phi = 0.72$  with respect to mCherry) were used as references. [35, 101]

## v. pKa Measurements

pKa measurements were performed by preparing buffers in the range of pH 2 to 12. The pH was measured for each buffer to confirm the calculated pH values using a pH meter. Fixed amounts of concentrated pure protein samples were added to 1 mL buffer in a quartz cuvette with a 1 cm path length and a fluorescence spectrum was recorded in each case with excitation at 520 nm. The maximum value of fluorescence counts was used to normalize the fluorescence spectra for each protein. The data were plotted and was fit to the sigmoidal curve shown in Appendix Figure A2.3. The calculated pKa values and reported values for published RFPs have been provided in Appendix 2: Table A2.4. Buffers in the pH range 2–3 were prepared with a dilution of glycine and 1 M HCl; in the range 3–6 were prepared using dilutions of 0.1 M citric acid and 0.1 M Na-citrate; in the range 6–8 were prepared with dilutions of 0.2 M Na<sub>2</sub>HPO<sub>4</sub> and 0.2 M KH<sub>2</sub>PO<sub>4</sub>; in the range of 9–12 were prepared with dilutions of 0.2 M glycine and 1 M NaOH and Tris-HCl buffer was used for the pH 7.4 measurement.

## vi. Fluorescence Lifetime Measurements in the Lifetime Flow Cytometer

The lifetime flow cytometer utilizes frequency-domain phase fluorimetry to select analytes based on excited state lifetime. The details of this set-up are discussed in references 30 and 41. [30, 41] The excitation beam is sinusoidally modulated at 29.5 MHz. The observed fluorescence signal is modulated at the same frequency as the excitation beam, but its lower modulation depth and a phase shift ( $\varphi$ ) that corresponds to the average time spent by the analyte in the excited state (average lifetime). A high-speed lock-in amplifier determines the phase shift, which is then converted to a lifetime value in the time domain.

$$\tau_{Analyte\ in\ sorter} = \tau_{Reference}\ (ns) + \frac{\tan(\varphi_{Lock-In\ Amp})}{29.5\ MHz}$$

For the reference, mCherry (~1.6 ns) was used. FR had a mean lifetime  $2.05 \pm 0.15$  ns on this device.

## vii. Steady-State Lifetime Measurements for Pure Proteins and Lysate

All lifetime measurements on purified proteins or filtered cell lysates were performed on a commercial TCSPC system (Fluoro-time 100, PicoQuant) using a 560 nm pulsed laser diode head excitation source with a repetition rate of 5 MHz. Emission was collected either using a red filter set ( $600 \pm 30$  nm) or a far-red filter set ( $670 \pm 30$  nm), to check for multiple species in the excited state or interconverting forms. The instrument response function (IRF) was collected using Ludox (Millipore Sigma) colloidal silica, whereas the samples were diluted to an OD value  $<0.05$  for the measurement. A minimum of 20,000 photon counts were used to generate the fluorescence decays. The fluorescence transient decays were fit to an iterative re-convolution with a bi-exponential function (or-tri exponential depending on the protein). The amplitudes and the components of the fits are provided in Appendix Table A2.5.

### **viii. Photobleaching Measurements**

Photobleaching measurements were performed using an LED excitation source (Lumencor) on an Olympus IX-73 fluorescence microscope with *E. coli* cells expressing the FP of interest. Bacteria on plates were washed and dispersed in aqueous blank buffer containing 0.17% (w/v) yeast nitrogen base (Sigma-Aldrich) and 0.5% (w/v) ammonium sulfate (Sigma-Aldrich) then photobleached with excitation rate-normalized LED light. To bleach the sample, a 560 nm LED was used, whereas for repopulation with blue light, a 438 nm LED light source was used. Details of the measurements are discussed in Appendix Figure A2.4.

### **ix. Brightness in *E. coli***

To assess the cellular performance of the FR clones in this study, we performed a bacterial brightness assay at the single cell level on a droplet microfluidic sorting platform. [41] Two biological replicates with three independent technical triplicates were performed for each measurement. The bacteria were grown, and expression was induced as previously described, cytometry was carried out at ~20–22 hours after starting induction. Each technical replicate involved 10,000 cells. Details of the screening protocol have been discussed in Appendix Figure A2.5 and a reference 41. [41]

## **3.3e. Cellular brightness assays**

### **Brightness in HeLa Cells**

*Flow Cytometry:* The proteins of interest in the FR family, along with some standard RFPs including mScarlet, mScarlet-I and mCherry, were fused to histone H2B and expressed in HeLa cells. Single-cell brightness was assessed by selecting single healthy cells based on forward and side-scattering photon counts on a BD FACSCelesta single cell analyzer after 48 hours of transfection. Untransfected cells were used as a control to background subtract and analyze the fluorescence in the red and green channels for the proteins of interest. In most cases (except mScarlet-I with only one biological replicate) three or more biological replicates with three technical replicates of each were analyzed to determine the mean fluorescence with standard

deviation based on the number of measurements (Appendix Table A2.7). The samples were excited by a 561 nm laser line for collecting red fluorescence through the TRITC filter set (585/30 nm) and a 488 nm laser line for collecting through a GFP filter set (530/30 nm). The residual fluorescence from the green channel was effectively at the background level of the EGFP-H2B control (displaying signal values ~20-fold higher than mScarlet with the highest green fluorescence in the series). Brighter red mutants had higher green fluorescence background, suggesting red fluorescence bleeding through the green channel.

*Confocal Microscopy:* HeLa cells grown in 35 mm imaging dishes (made in-house from Corning 35 x 10 mm dishes with VWR 18 x 18 mm #1.5 cover slips) were imaged 48 hours post transfection to maintain consistency with the FACS measurements. Before imaging, cells were washed three times with 2 mL phosphate-free HEPES-buffered Hanks' balanced salt solution (HHBSS) containing 20 mM HEPES (Sigma), pH 7.4 and resuspended in 1.5 mL of the same buffer. Imaging was performed on a Nikon Ti-E spinning disc confocal microscope system. The imaging dishes were mounted on the microscope in an environmentally controlled chamber (Oko Labs; set to 37°C, 5% CO<sub>2</sub>, 90% humidity) and viewed with a 40x (NA 0.95) air objective. A 560 nm laser was used for illumination and a 590–650 nm band pass filter (TRITC) was used for the detection of fluorescence with 200 ms exposure time and 20% laser power of the instrument. To minimize photobleaching, the focus of the microscope was adjusted to a lower laser power (5%) at only the center spot of the large image. Prolonged exposure of FR and its mutants to laser light can lead to lower brightness, which was a critical factor for this assay. Large images (~1600 x 1000  $\mu m$ ) were taken for each dish for each replicate. Details for this assay are presented in Figure 2.4b and Appendix Table A2.8. The imaged cells were analyzed using the suite CellProfiler. [189] A pipeline was created that would identify objects that are above the noise background. A binarized image was thus created, then a gate to filter objects typical of the size of nuclei was selected (2–25  $\mu m$ ), which selected the fluorescing nuclei in the H2B construct. The filtered objects in the parent cells were then quantified for mean brightness. Untransfected cells with cellular autofluorescence were the first to be analyzed, which gave us the estimation of bleed-through fluorescence.

### **3.3f. Cellular localization assays**

#### **i. Imaging Conditions**

U2OS cells were transfected with the constructs of interest as described previously. The cells were imaged 24 hours post transfection with a Nikon Ti-E spinning disc confocal microscope. Several large images ( $\sim 1,680 \times 1,000 \mu m$ ) were captured while scanning the z-focus (z-stacks) for optimum focusing of all the cells appearing in the field of view. Typical scanned depth was 5.0–7.8  $\mu m$  and was evenly separated into 5–7 layers. Maximum intensity projection of the z-stacks was used for the quantification of OSER score. For GalT scoring, Z-stacks and independent frames were used for quantification of localization of FPs to the Golgi.

#### **ii. Data Analysis**

The data analysis for the quantification of OSER scores was done in accordance with our previous report. [30] In brief, the suite CellProfiler was used to develop a pipeline that can identify cells based on size (20–50  $\mu m$ ), shape and fluorescence intensity. The program identifies oligomerization sub-structures in these cells, filters them and then quantifies them as whorls. The pipeline further continues to relate the number of whorl structures with each cell and a MATLAB code quantifies the number of perfect and imperfect (cells with whorls). A score of 100 indicates 100% of the cells are whorl free and 0% indicates all cells have whorls in them. FR and its mutants display a high OSER score, whereas TagRFP-T, was used as a negative control. The imaging data for the GalT-FP construct was analyzed using a blinding approach where multiple individuals were provided with image sets that were randomized. Individual images were scored as having fluorescence signal localized to the Golgi (characterized by a fist-like structure near the nucleus) or not. The approximate size of a healthy nucleus was estimated using ImageJ. [190] If the cells expressed FP outside this fist like structure in the form of puncta or just smeared across the cytoplasm, they were considered mis-localized.

### **3.3.g. Chromophore maturation kinetics and cytotoxicity**

#### **i. Cytotoxicity Assay**

HeLa cells were transfected with the H2B-FP fusion constructs as discussed previously. Two biological replicates were prepared for each FP. Day 2 was defined as 48 hours post transfection. RFP and EGFP transfected cells were mixed at a 50:50 ratio by volume. A sample of the cell mixture was prepared for a flow cytometry measurement (BD FACSCelesta) to determine the actual RFP:EGFP ratio. The rest of the mixture was re-plated for further growth and proliferation. Independent screens of just EGFP cells and RFP cells were carried out first. Cells with green fluorescence higher than background were classified as “EGFP” and cells with red fluorescence higher than background were classified as “RFP”. On Day 6, the re-plated cells were again subjected to flow cytometry to quantify the change in the RFP and EGFP populations and determine a relative level of cytotoxicity. The cytotoxicity score was then calculated as the change in the RFP:GFP ratio for each sample. EGFP was used for normalization because it has been shown to be minimally cytotoxic. [191] The data for the assay is presented in Appendix Table A2.9.

#### **ii. Maturation Kinetics**

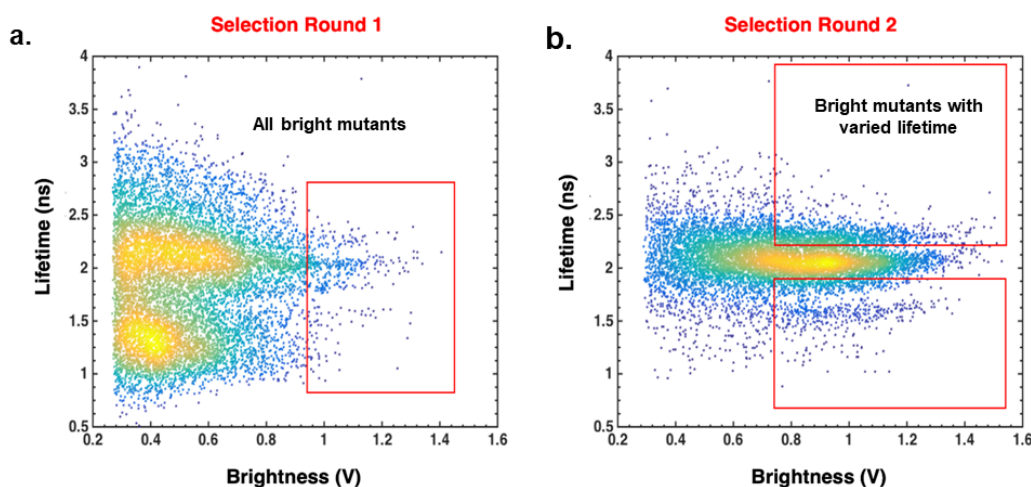
Bacteria (Top 10, *E. coli*) expressing FPs were grown in expression media for 4 hours (28°C, 230 RPM) to facilitate FP production in the exponential growth log phase. Chloramphenicol (250 µg/mL) was added to the cultures to stall bacterial growth and protein production. [192] The cultures were maintained at 37°C and 230 RPM and aliquots were taken every 30-60 minutes to measure the Optical Density (OD) and fluorescence. The  $t_{1/2}$  at 37°C was calculated by calculating the time required for the FP to reach half the maximum fluorescence value. The data for the assay is presented in Appendix Table A2.10.



### 3.4. Results

#### 3.4a. Lifetime-based selection of a site-directed library near the para-hydroxyphenyl moiety of the chromophore identifies the C159V mutation.

Four positions near the para-hydroxyphenyl moiety of the chromophore—C159, M161, V196 and H198—were simultaneously mutated in FR to generate a ~7,000-member library. We employed lifetime-based microfluidic flow cytometry to select variants with increased brightness and values of fluorescence lifetime different from the parent (Figure 3.1). [30] Sequencing of selected clones revealed that H198 and V196 were conserved but positions 159 and 161 showed sequence diversity (Appendix Table A2.2). Furthermore, the selected population contained two clones, FR-C159V and FR-C159L, with higher cellular brightness than the parent FR, but with differing lifetimes. FR-C159V had an increased lifetime of 2.0 ns, whereas FR-C159L had a decreased lifetime of 1.2 ns (c.f. 1.8 ns lifetime for FR: Details in Appendix Section 2). Of the two clones, FR-C159V was selected for further development in combination with the mutations identified below because it had the longer lifetime and because it did not show the complex photoactivation behavior seen in FR-C159L (Appendix Figure A2.4)



**Figure 3.1. Lifetime screens for the site directed libraries of FusionRed.** The fluorescence lifetime (vertical axis) vs. brightness (horizontal axis) screening dot-plots of the FR site-directed library targeting sites near the para-hydroxyphenyl moiety, after two rounds of fluorescence-

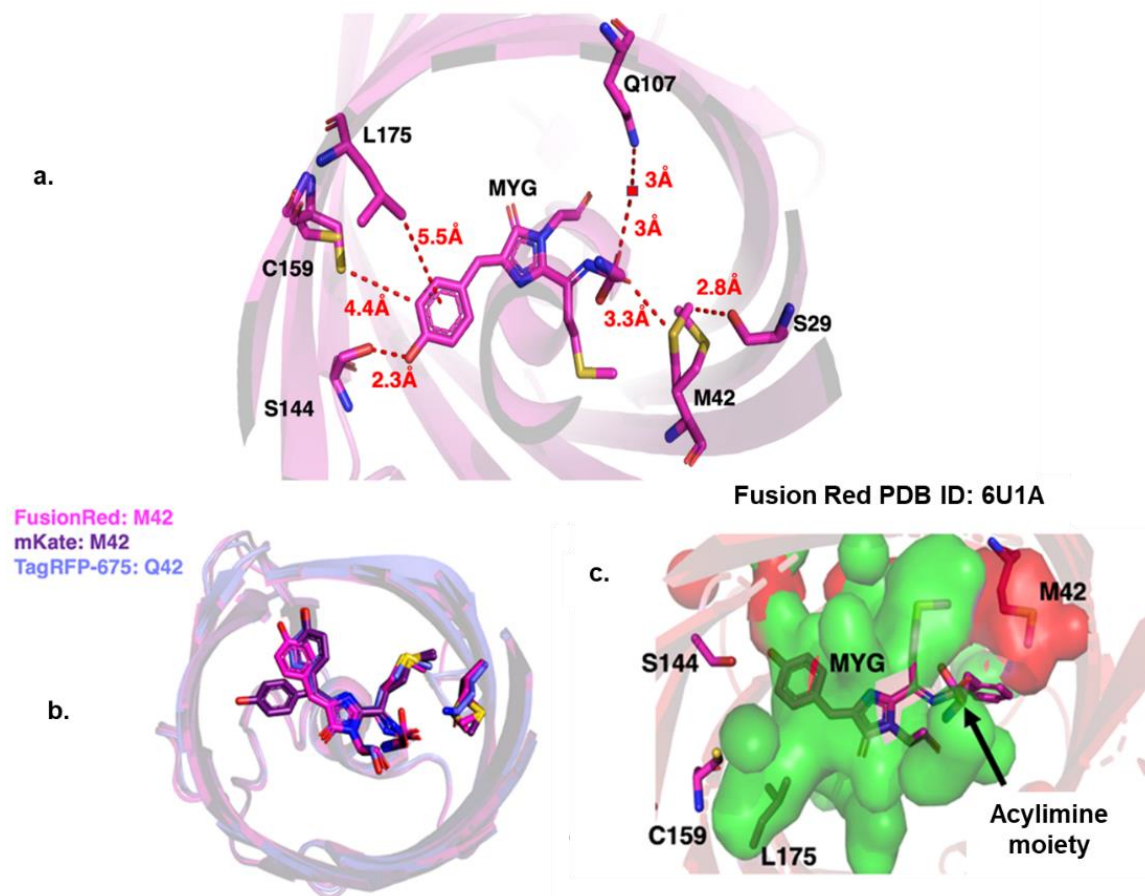
activated cell sorting (FACS) enrichments. Pseudocoloring represents normalized cell counts, where yellow indicates higher cell count and blue indicates a lower cell count. Red boxes indicate the selection gates for multi-parameter microfluidic sorts. **(a.)** The first round of sorting for selecting bright clones. **(b.)** The second round of sorting was performed in two batches using sorting gates for high brightness with different lifetime ranges. The mean fluorescence lifetime of FR measured on this instrument was 2.05 ns.

### **3.4b. The M42Q mutation adjacent to the acylimine moiety of FR increases the extinction coefficient, quantum yield and fluorescence lifetime.**

Mutations that alter the hydrogen-bonding structure surrounding and interacting with the acylimine moiety of the chromophore produce RFPs with extended Stokes shifts. [79, 91, 95] In addition to the site-directed library aimed at mutating the end of the chromophore near the para-hydroxyphenyl moiety, we investigated the end of the chromophore near the acylimine moiety in the crystal structures of mKate (PDB ID: 3BXB), the far-red emitting FP TagRFP-675 (PDB ID: 4KGE) and FR (PDB ID: 6U1A) using PyMol (Figure 3.2 and Appendix Figure A2.13). [193] Based on the structural similarity, conformational freedom and the fact that the sidechain at position 42 is a known hot-spot for altering the hydrogen-bonded chemistry in mKate and TagRFP-675, we incorporated a single point mutation M42Q into FR. [79, 91] The FR-M42Q (or FR-Q) mutant shows a 41% increase in the maximum extinction coefficient, an 18% increase in fluorescence lifetime and a 42% increase in quantum yield (Table 3.1 and Figure 3.3). The mutation also resulted in blue-shifted and narrower absorption and emission spectra and a minor change in Stokes shift (Figure 3.3a and Appendix 2: Sections 4e, f and g). These properties reflect the higher radiative rate constant calculated for this mutant in comparison to FR (Figure 3.3d). Consequently, we performed site-saturation mutagenesis (library size = 20) at this position in the context of FR-M and found that the parent, along with FR-M M42Q (FR-MQ) and FR-M M42I (FR-MI) were the only variants with observable red fluorescence. FR-MQ was the brightest species with a quantum yield of ~43% and a further increase of fluorescence lifetime by 0.3 ns. FR-MI showed decreases in maximum extinction coefficient, quantum yield and fluorescence lifetime, thus we did not pursue engineering of this variant. We investigated the effects of analogous mutations on closely related RFPs and noted rather different outcomes. For example,

mKate M42Q shows a decrease in the molecular brightness due to ~50% decreases in both maximum extinction coefficient and quantum yield, along with a blue-shifted absorption spectrum and a significantly increased Stokes shift, all of which are consistent with previous findings. [79, 91] In addition, analogous mutations at position A44 on mCherry and mScarlet-I RFPs resulted in non-fluorescent clones. It should be noted that mCherry, mScarlet-I and FR show multiple structural differences in the chromophore structure and in the vicinity of the chromophore. Although crystal structures indicate that there seems to be sufficient space to accommodate a 44Q sidechain in mCherry and mScarlet-I, there are multiple structural differences between the three FPs. We cannot explain why this mutation did not produce a fluorescent species in mCherry and mScarlet-I.

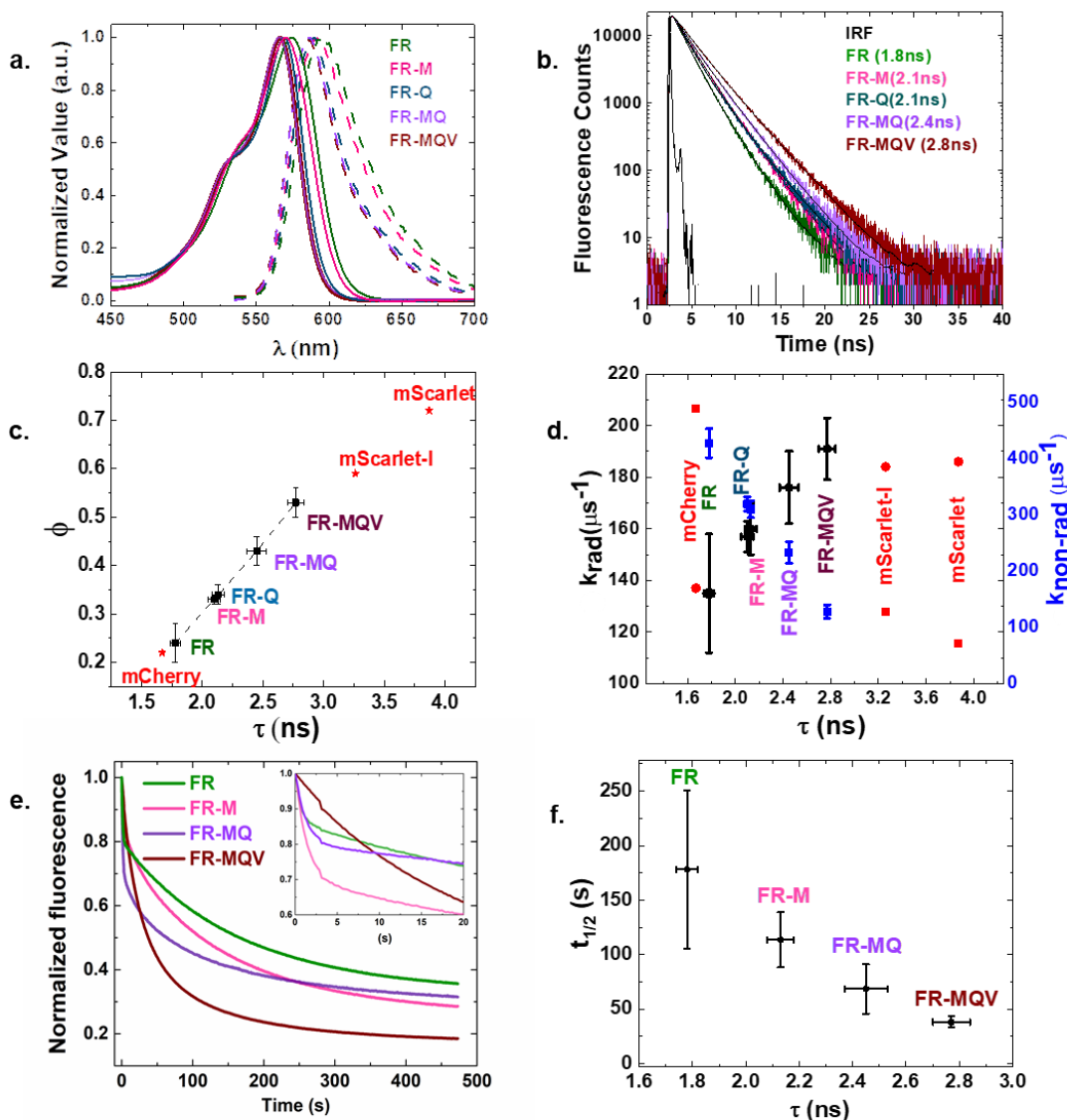
While measuring the maximum extinction coefficient with the alkali-denaturation method, we observed that M42Q containing mutants of FR produce a single hydrolysis product with an absorption band centered at 380 nm (Appendix Figure A2.8). Most RFPs, such as mScarlet, mCherry and mRuby3, display GFP-like degradation with an absorption band at 450 nm. The hydrolysis products of FR variants with the native 42M have both 380 nm and 450 nm absorption peaks. The 380 nm band is assigned to the cleavage of the chromophore from the  $\beta$ -carbon of the Tyr moiety in the chromophore. [194, 195] The absence of the 450nm band in FR-Q and FR-MQ reveals alteration of the chromophore hydrolysis chemistry. This suggests that the 42Q residue interacts strongly with the chromophore in FR.



**Figure 3.2. Site-directed mutagenesis based on structural data. (a.)** Crystal structure of FR with positions around the chromophore relevant to the study generated using PyMol. **(b.)** Comparative locations of the 42 position in TagRFP-675 (PDB ID: 4KGE, blue, 42Q), mKate (PDB ID: 3BXB, purple, 42M) and FR (PDB ID: 6U1A, pink, 42M). The crystal structure of mKate (PDB ID: 3BXB) indicates that the end of the chromophore near the para-hydroxyphenyl moiety may exist in both cis and trans conformations, with the cis conformer overlapping with that of TagRFP-675. For FR, the fluorescent chromophore exists in the cis conformation as well. **(c.)** The occupancy surfaces (based on PyMol) indicate that the chromophore (green surface) and the sidechain of the M42 residue (red surface) come in close contact.

**Table 3.1. Summary of *in vitro* photophysical properties of FR mutants developed in this study** vs. re-measured values for some well-characterized RFPs. The standard deviation error has been reported for measurements where independent triplicates were performed. mCherry and mScarlet were used as references for the quantum yield measurements. [35, 101]

FP	$\lambda_{\text{abs}}$ (nm)	$\lambda_{\text{em}}$ (nm)	$\tau$ (ns)	$\phi$	$\epsilon_{\text{max}}$ (M <sup>-1</sup> cm <sup>-1</sup> )	Molecular brightness	$k_{\text{rad}}$ ( $\mu\text{s}^{-1}$ )	$k_{\text{NR}}$ ( $\mu\text{s}^{-1}$ )
FR	574	596	1.78±0.04	0.24±0.04	94000±7000	1	135±23	427±26
FR-M	571	591	2.13±0.05	0.34±0.02	78000±10000	1.2	160±10	310±15
FR-Q	568	587	2.10±0.05	0.33±0.01	133000±2000	2	157±6	319±13
FR-V	573	594	1.96±0.02	0.31±0.04	84500±8000	1.2	154±4	361±9
FR-MV	569	592	2.39±0.02	0.38±0.09	85000±6500	1.4	166±5	254±9
FR-MI	575	595	1.75±0.02	0.26±0.04	82000	1	149±12	423±13
FR-MQ	567	586	2.43±0.08	0.43±0.08	130000±9000	2.7	176±14	233±19
FR-MQV	566	585	2.77±0.07	0.53±0.03	144000±3000	3.4	191±12	170±16
mScarlet-I	570	591	3.26±0.07	0.59±0.03	102500±4500	2.7	184±10	128±12
mKate	588	634	2.26±0.07	0.33	48000	0.7	130	264
mKate-M42Q	577	654	2.16±0.05	0.17	25000	0.2	79	384
mCherry	586	607	1.67±0.07	0.22 (ref)	75000±5000	0.8	137	488
mScarlet	569	592	3.87±0.07	0.71 (ref)	103000±4500	3.3	186	72



**Photophysical analysis of the FR family of proteins**

**Figure 3.3. Photophysical properties of the FR family of FPs.** (a.) Absorption (solid lines) and emission (dashed lines) spectra. (b.) TCSPC lifetime traces. Longer decay profiles indicate the gradual increase in fluorescence lifetime. The black lines indicate exponential fitting for the decay profiles. (See Appendix 2: Section 4b for fitting details) (c.) Plot of quantum yield versus fluorescence lifetime. The dashed line indicates the linear rise of quantum yield with increments in fluorescence lifetime for the FR family of FPs. The values for mCherry, mScarlet and mScarlet-I are taken from references [35, 101]. (d.) A plot showing the increase in radiative rate constant (black) and decrease in non-radiative rate constant (blue) across the FR family relative to mCherry and mScarlet (red). Vertical error bars (blue and black) in the FR family indicate the uncertainty

in the calculated rate constants, the horizontal error bars (black) are the standard deviation error in the measurement of fluorescence lifetime. (e.) Photobleaching traces for the family of FR mutants, determined by averaging ~10 decay traces in *E. coli* cells for each FP under normalized excitation rates. (f.) The dependence of the fluorescence decay half-life on the excited state lifetime. Error bars indicate standard deviation errors.

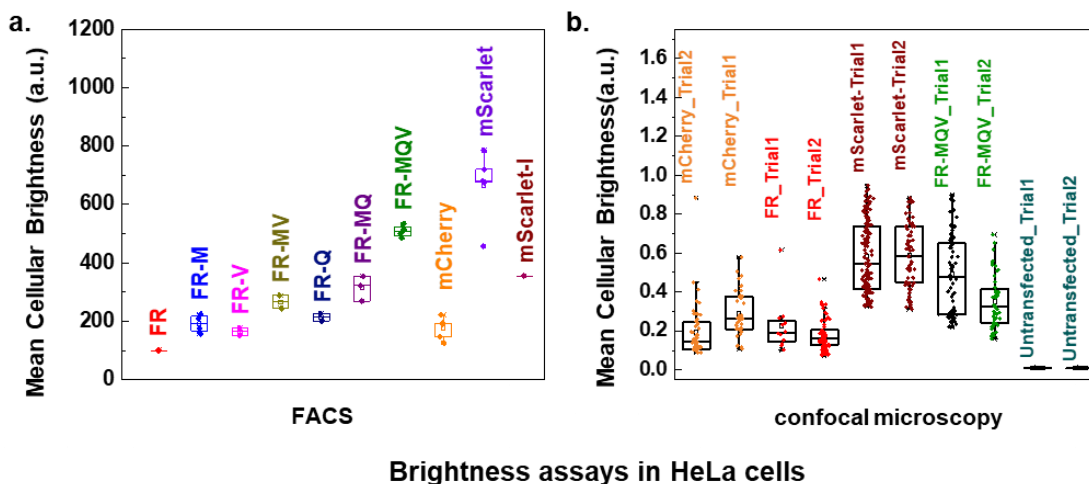
### 3.4c. Combining mutations M42Q, C159V and L175M led to the generation of bright FR-MQV.

Each individual mutation results in an increment of fluorescence lifetime by ~0.3 ns in FR. Therefore, we incorporated all three mutations into FR to yield FR-MQV. The FR-MQV mutant exhibits a fluorescence lifetime of 2.8 ns which results in a high quantum yield of 53% (Figures 3.3b and 3.3c). The absorption and emission spectra are similar, with a slight blue shift and narrowing of the spectra in all variants possessing the M42Q mutation (Figure 3.3a and Appendix 2: Section 4f and g). The molecular brightness of FR-MQV is ~3.4-fold higher than that of FR. Overall, we see increases in fluorescence lifetime and quantum yield, as a result of an increase in the radiative rate constant and a concomitant decrease in the non-radiative rate constant (Figure 3.3d). Steady-state absorbance and fluorescence spectra along with wavelength-dependent lifetime measurements were collected to assess the heterogeneity of chromophore formation. We found that FR-MQV matures predominantly into a red-emitting chromophore which emits as a single red-emissive species (Appendix 2: Section 4b, i and j). FR-MQV has a low *in vitro* pKa of ~4.6, like other members of the FR family (Appendix 2: Section 4a) and breaks down into a single product of alkali hydrolysis, resembling FR-Q and FR-MQ (Appendix 2: Section 4h). We also found that FR-MQV was ~2-fold brighter than FR in *E. coli* (Appendix 2: Section 4d).

The photobleaching traces of FR-MQV and related variants are shown in Figure 3.3e. When photobleaching measurements were performed on multiple FPs illuminated at the same excitation rate, longer fluorescence lifetime is correlated with faster photobleaching (Figure 3.3f). In this case, excitation rates are set by ensuring that the value of excitation power ( $P$ , in mW) from the objective divided by the absorption cross section at 560 nm  $P/\sigma_{560\text{ nm}}$  (defined as  $\sigma_{560\text{ nm}} = 2303 \times \text{EC}_{560} / N_A$ ; where  $\text{EC}_{560}$  is the extinction coefficient at 560 nm and  $N_A$  is the Avogadro constant)



is the same for the FPs being compared. Furthermore, it was observed that under continuous irradiance of  $\sim 5 \text{ W/cm}^2$  at 560 nm, the fluorescence kinetics of FR variants with a C159 residue show two decay timescales. This behavior is typical of most FPs, which show an initial decay due to dark state conversion (also sometimes called reversible photobleaching), along with a slower timescale of permanent photobleaching. [55,56] The C159V and C159L mutations, which incorporate aliphatic groups, significantly reduce the amplitude of the faster decay component (Appendix Figure A2.4). To verify that the faster component corresponds to a reversible process, we used alternating pulses of 560 nm and 438 nm excitation (Appendix 2: Section 4c). The FR-C159V variant showed little recovery and FR-C159L showed slight recovery of the fluorescence, but the recovery was much larger for FR ( $\sim 20\%$  higher). These results show that the reversible photobleaching dynamics of FR are influenced by the chemical nature of the sidechain at position 159.



**Figure 3.4. Cellular brightness assays.** (a.) FACS-based brightness assays of FPs expressed as histone H2B fusions in HeLa cells. Each biological replicate involved three technical replicate screens of  $\sim 10,000$  individual cells. (b.) Mean brightness values of individual HeLa cells obtained from confocal microscopy for histone H2B fusions. Each dish was an independent biological replicate and contained  $\sim 100$  cells. For both a and b, the cells were sampled  $\sim 48$  hours post transfection. Details of the assays are provided in Appendix 2: Section 5a. Error bars indicate standard deviation errors.



### **3.4d. FR-MQV is 5-fold brighter in HeLa cells, retains the cellular properties and high-fidelity localization of FR.**

Flow cytometry and confocal microscopy were employed to quantify the cellular brightness of FR-MQV in HeLa cells (Figure 3.4). FACS screening of HeLa cells expressing FPs fused to histone H2B protein after 48 hours of transfection revealed that FR-MQV is 5-fold brighter than the parent FR, and ~1.4-fold less bright than mScarlet (Figure 3.4a and Appendix 2: Section 5a). Measurements with confocal microscopy on the same construct revealed that FR-MQV is ~2.3-fold brighter than FR and ~1.4-fold dimmer than mScarlet (Figure 3.4b and Appendix 2: Section 5a). We next investigated whether the mutations in FR-MQV would influence cytotoxicity, chromophore maturation and localization of fusion proteins. FR-MQV exhibited low cytotoxicity, like the parent FR and in contrast to mCherry, which showed ~80% relative decrease in cells expressing the FP after 6 days of expression (Details of the assay are discussed in Appendix 2: Section 5b). FR-MQV showed chromophore maturation kinetics identical to the parent FR ( $t_{1/2}$  ~195 min at 37°C; Appendix 2: Section 5c) and similar to mScarlet ( $t_{1/2}$  ~130 min at 37°C), but slower than mScarlet-I ( $t_{1/2}$  ~45 min at 37°C).

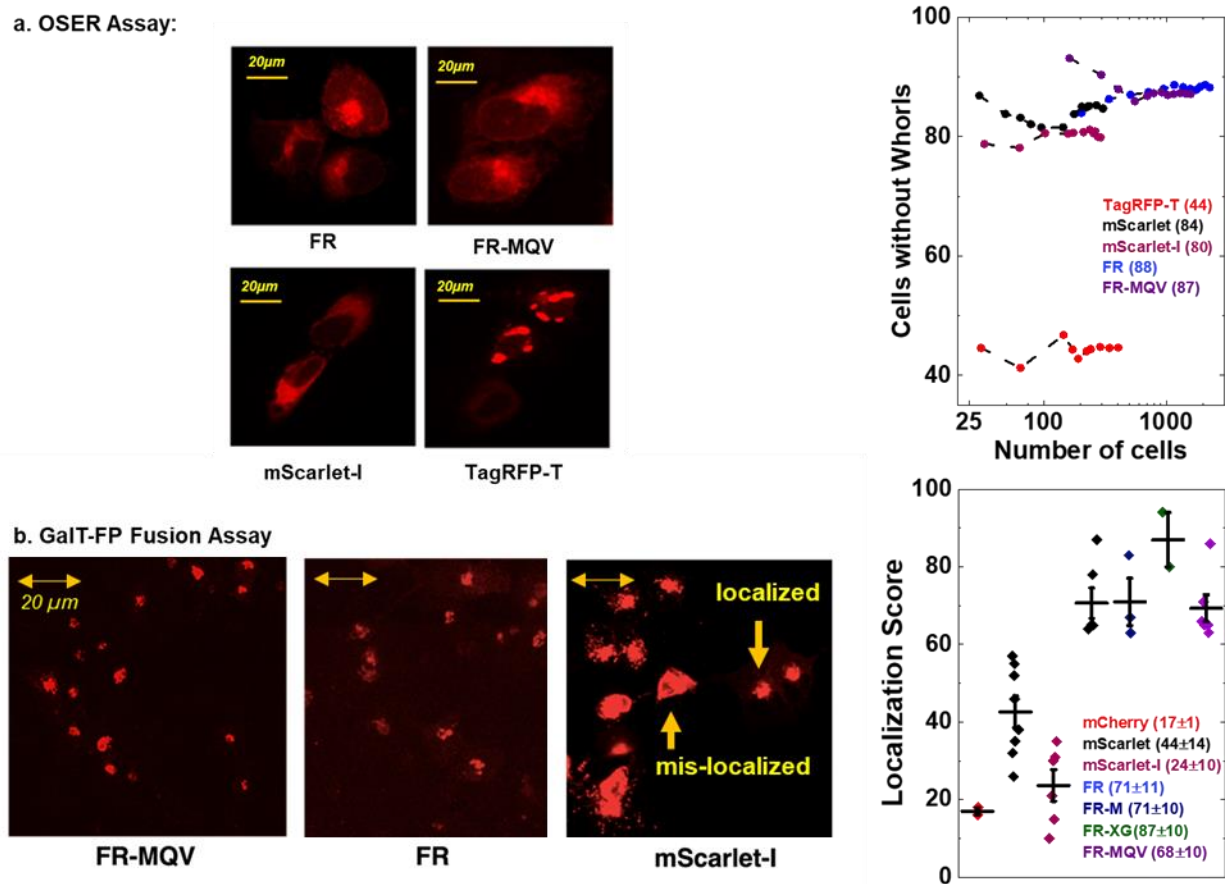
In the OSER assay, FPs are fused to the cytoplasmic end of the endoplasmic reticular signal anchor protein (cytERM). [199] Membrane localization increases the local concentration leading to the formation of dimers or higher order oligomers that distort the ER structure which are observed as cellular sub-structures or “whorls”. Figure 3.5a shows confocal microscopy images of cytERM-FPs expressed in U2OS cells and the corresponding OSER scores. TagRFP-T (positive control) shows characteristic oligomerization in the form of sub-cellular “whorls”, while the FR mutants, mScarlet and mScarlet-I generally lack such structures. In this work, we measured a score of 88 for FR and 87 for FR-MQV which are ~5% higher than measured for mScarlet (84) and ~10% higher than mScarlet-I (80). We also developed GalT-FP fusions to localize the FPs of interest to the Golgi. Proper and mis-localization to the Golgi in U2OS cells imaged using confocal microscopy are shown in Figure 3.5b. Our analysis indicates that FR-MQV and other FR-derived mutants outscore mScarlet and mCherry in terms of proper localization to the Golgi. Our results are qualitatively consistent with previous observations where FR correctly localizes to the Golgi, but GalT-mCherry frequently displays puncta in the cytoplasm. [196] We observed higher occurrences of mis-localization in the bright RFPs mScarlet and mScarlet-I compared to FR

mutants. Thus, cellular imaging studies indicate that FR-MQV retains the high-fidelity localization properties of FR without compromising low cytotoxicity and appropriate maturation and expression.

**Table 3.2. Summary of cellular and molecular brightness for FR variants.** The molecular brightness of FR-MQV is ~3-fold that of FR, while FACS measurements indicate it is ~5-fold brighter in HeLa cells. In comparison, FR-M has only a 20% larger molecular brightness than FR but is ~2-fold brighter in HeLa cells. These observations suggest that the ~5-fold increase in cellular brightness of FR-MQV relative to FR is due to the combination of an increase in molecular brightness and the increased cellular expression possibly from the L175M substitution.

Protein	Molecular Brightness ( $\epsilon_{\max} * \phi$ )	Bacterial Cell Brightness (585±15 nm)	HeLa Cell Brightness (585±15 nm)	Normalized relative to FR-M	
				Molecular Brightness	HeLa Cell Brightness
FR	1	1	1	0.9	0.5
FR-M	1.2	1.3	2.1	1	1
FR-Q	2	1.4	-	1.7	-
FR-MQ	2.6	1.7	3.6	2.3	1.7
FR-MQV	3.4	1.9	5.3	2.9	2.5

Over-expression and oligomerization in cells can lead to unwanted effects such as self-quenching, mis-folding, mis-localization and disruption of organelle structure and function. [197, 198] Hence, we performed the organized smooth endoplasmic reticulum (OSER assay) [199] and examined localization of a GalT-FP fusion to assess the *in cellulo* monomericity and localization properties of FR-MQV relative to FR and the other bright RFPs mScarlet and mScarlet-I.



**Figure 3.5. Localization Assays.** (a.) *The OSER Assay:* Representative maximum intensity projected confocal images for U2OS cells expressing cytERM-FP constructs for FR, FR-MQV, mScarlet-I and TagRFP-T (negative control), and the graphical representation of the OSER scores for these cells. The y-axis on the plot indicates the percent cells that do not express sub-cellular “whorls”. The OSER scores for this assay were: FR, 88 (2,250 cells), FR-MQV, 87 (1,580 cells), mScarlet, 84 (306 cells), mScarlet-I, 80 (291 cells) and TagRFP-T, 45 (406 cells). (b.) *The GalT-FP Fusion Assay:* Confocal microscope images for U2OS cells expressing the GalT-FP fusion constructs for FR-MQV, FR and mScarlet-I, with the localization scores (indicative of the percent cells with the FP localized correctly to the Golgi) plotted for each RFP. The microscopy images show cells expressing GalT-mScarlet localized and mis-localized to the Golgi. The scores for FPs analyzed in the study have been reported with the FR mutants (including previously published FR-XG [30]) outscoring mCherry and mScarlet. Error bars indicate standard deviation errors across biological replicates.

### 3.5. Discussion

In this work we engineered FR using a structure guided approach, aimed to develop mutants with increased fluorescence lifetime manifesting in higher quantum yield, eventually leading to a FP with higher molecular and cellular brightness. We examined both the para-hydroxyphenyl and the acylimine sites of the chromophore for sites where we could potentially increase steric effects, increase rigidity, interact chemically or electrostatically with the chromophore. This approach led us to mutate multiple residues near the para-hydroxyphenyl moiety, thus motivating the need for lifetime screening of a 7,000-member library. In contrast, near the acylimine moiety we identified the capacity for an interaction to be added at position M42. Crystal structure data suggests conformational freedom of the sidechain at position 42 in FR. [200] The position 42 lies in close proximity to the acylimine moiety of the chromophore and based on our previous work, is critical to the hydrogen-bond network on that end of the chromophore in closely related FPs mKate and TagRFP-675. [79, 91] A smaller site-saturated library of 20 possible mutants at position 42 allowed us to carry out a plate based screen on FR-M, where Gln stood out as the best possible sidechain residue in terms of increases in fluorescence lifetime and quantum yield.

Individually L175M, C159V and M42Q substitutions, increased the fluorescence lifetime of FR by ~0.3 ns (~17% increase), and when incorporated together in FR-MQV exhibited ~1 ns (~55%) increase over FR, suggesting the effect of each mutation in terms of fluorescence lifetime was additive in the triple mutant. FR-MQV closely follows the spectral characteristics of FR-Q and FR-MQ, all of which have reduced spectral widths in absorbance and emission spectra with blue shifts in the absorption spectra with ~20% or higher peak absorbance with respect to FR. Consistent with the Strickler–Berg relationship, we calculate higher rate constants for radiative decay in FR mutants with the M42Q mutation. Based on these values, we successfully engineered the decrease of  $k_{\text{non-rad}}$  with a simultaneous increase of  $k_{\text{rad}}$  in FR, which is an efficient way of engineering the molecular brightness of a FP. FR-MQV has a  $k_{\text{rad}}$  comparable to the brightest RFP to date - mScarlet and its cellularly brighter mutant mScarlet-I. Finally, these mutations translate to high cellular brightness without compromising the cellular properties of the parent FR, as seen from the cytotoxicity, maturation and imaging assays performed in this study. Given this favorable combination of incremental photophysical changes and preserved cellular properties of FR-MQV

with respect to FR, further rounds of engineering could produce a new standard for red fluorescent protein labelling.

Crystal structures indicate that the para-hydroxyphenyl moiety of FP chromophores can adopt either the trans (non-fluorescing) or the cis (fluorescing) conformation [25, 201] and in FR, the latter can potentially form a hydrogen bond with the hydroxy group of the S144 residue (Figure 3.2). [200] The location of C159 suggests that it may play a similar role for the trans conformer of the chromophore. The position 159 was recently discussed to be critical in terms of molecular brightness for FR. [200] The propensity to switch to the trans conformer may be reduced when the residue at this position is substituted for an aliphatic group. The cis to trans isomerization is a reversible process and is manifested as the reversible component of photobleaching under continuous and pulsed (ms to s) illumination. [55, 56] The reduction in dark state conversion in the photobleaching kinetics of FR-C159V and FR-C159L indicate that there is a reduction in the tendency of the chromophore to switch to a dark, trans conformer. Furthermore, FR-C159V did not show photo-activation of fluorescence, [202] which suggests the chromophore might be locked into either the cis- or trans- conformation. Leu has a larger aliphatic sidechain than Val; thus, the shorter lifetime and complicated photo-activation behavior of FR-C159L may be a steric effect, with Val providing a better spatial fit that restricts chromophore movement into the dark, trans conformer.

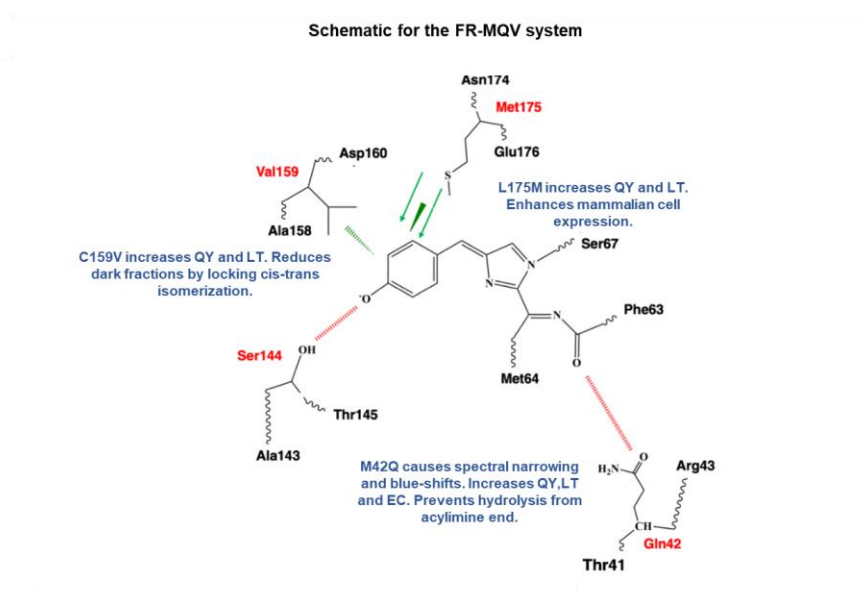
Time-resolved ultrafast spectroscopy previously demonstrated that TagRFP-675 and mKate-Q display large Stokes shifts because there are multiple emissive species in the form of non-interconverting hydrogen bonded conformers. [79, 91] Multiple emissive species tend to broaden spectra and correlate with lower quantum yield and shorter fluorescence lifetime values. [79, 91] Steady state excitation-dependent emission spectra and fluorescence lifetime decay spectra collected in multiple emission windows dismiss the existence of multiple emissive species in M42Q mutants of FR (Details in Appendix 2: Section 4b and j). This is consistent with the minimal change in Stokes shift in the M42Q mutants of FR in comparison to mKate-Q (Appendix 2: Section 4f). The polar sidechain of Gln can exist as one of many possible rotamers. [203] It is possible that this residue at position 42 rearranges itself into a conformation that might restrict the number of emissive species. This is supported by UV-Vis spectral data, which reveal that incorporating M42Q decreases the spectral width, blue shifts the absorption and emission spectra and increases

the radiative rate constant. Other than Gln, we observed red fluorescence from the mutant with Ile at this position. Ile is of similar size to Gln and Met but is aliphatic in nature. Characterization of the FR-MI mutant revealed a relative decrease in maximum extinction coefficient and quantum yield with a decrease in fluorescence lifetime with respect to the parent FR-M. We also observed an undesirable green emitting chromophore in the FR-MI and mKate-Q mutants, which was absent from the M42Q mutants of FR (Appendix 2: Section 4i), suggesting complete maturation to a red chromophore in FR-42Q mutants. Alkali denaturation of FR-Q, FR-MQ and FR-MQV shows cleavage of the chromophore only from the  $\beta$ -carbon of the Tyr sidechain, which is indicative of base access to a single site for hydrolysis (Appendix 2: Section 4h). The  $\beta$ -carbon of the Tyr sidechain in RFPs is spatially distant ( $>10\text{\AA}$  based on crystal structure data) from the acylimine moiety of the chromophore. These molecular and photophysical properties suggest that factors other than steric effects (such as electrostatics, hydrogen bonding, etc.) may be responsible for the changes observed in the M42Q mutants of FR.

The ~5-fold increase in brightness observed in FR-MQV relative to FR (Table 3.2) is likely an amalgamation of the higher molecular brightness of the M42Q and the C159V mutations along with the higher cellular expression seen for FR-M. [30] The trends for brightness measured through cytometry and microscopy are similar, but FACS measurements employ short laser exposure times with fast, sensitive detectors like photomultiplier tubes (PMTs). In microscopy, cells are subject to relatively longer exposure times and higher irradiances (~ms timescales) imposed by acquisition times of cameras, which can lead to photobleaching. The correlation between excited-state lifetime and photobleaching (Figure 3.3f) is evidence that photobleaching occurs due to excited-state absorption, as discussed for other RFPs. [55] A practical consequence is that brightness values measured by FACS are systematically higher than those measured in microscopy. This discrepancy between the absolute brightness values between FACS and confocal microscopy measurements was also observed for mScarlet.

Figure 3.6 is a schematic indicating the postulated effect of each mutation in the development of FR-MQV. Overall, the high extinction coefficient, decrease in the spectral width and blue shifts can be attributed to the M42Q mutation, whereas changes in the dark state behavior with increments in fluorescence lifetime and quantum yield seem to be additive from incorporating the C159V and the L175M mutations into the FR-Q system. All three mutations in FR-MQV face

inward, in the vicinity of the chromophore and therefore seem to minimally perturb the cellular properties of FR. The high scores for the OSER and Golgi-localization assays for all FR mutants indicate cellular properties in FPs are not appreciably affected by mutations of residues with side chains facing towards the chromophore. In the development of FR from mKate and mKate2, most of the mutations that optimized the FP's biological properties were facing out from the  $\beta$ -barrel. We targeted positions to increase brightness, and thus we did not alter these external positions that were involved in the development of FR from mKate2. [188] Consequently, the mutants retained FR's original performance in terms of localization and cytotoxicity. Though chromophore maturation kinetics can be greatly altered with internal mutations, [29] in this case, chromophore maturation was not substantially slowed by these mutations.



**Figure 3.6. The overall molecular schematic describing the postulated effects of the substitutions in FR-MQV.** The L175M mutation increases the quantum yield (QY), fluorescence lifetime (LT) and expression in mammalian cells compared to FR. The C159V mutation reduces the propensity of the cis-trans isomerization of the chromophore and results in subsequent increments in the quantum yield and fluorescence lifetime. The M42Q mutation interacts strongly with the acylimine moiety of the chromophore resulting in blue shifts, spectral narrowing, increased quantum yield, fluorescence lifetime and extinction coefficient (EC) while also changing the hydrolysis products of the chromophore.

### 3.6. Conclusions

In this study we present the development of FR-MQV, a FP with high brightness that maintains favorable cellular properties of FR. The C159V mutation appears to play a critical role in restricting the dark state conversion of the protein and increasing fluorescence lifetime, which leads to an increase in quantum yield. The M42Q mutation plays an important role in enhancing the molecular brightness by increasing both the quantum yield and maximum extinction coefficient. Interestingly, this mutation had different effects in closely related FPs including mKate and TagRFP-675. The previously reported L175M mutation contributes to robust expression levels which improves cellular brightness for FR mutants compared to FR. With its increased brightness and favorable cellular properties, FR-MQV shows promise as a template for further rounds of engineering suited to specific imaging applications like FLIM and FRET. Increasing fluorescence lifetime through further rounds of engineering may increase quantum efficiencies of emission to values near mScarlet with further reduction in the non-radiative rate constant (Chapter 6).



## Chapter 4

# Characterizing dark state kinetics from ensemble and single molecule fluorescence experiments in FusionRed variants

### 4.1. Publication Note

Parts of this chapter have been adapted from the article “*Characterizing Dark State Kinetics and Single Molecule Fluorescence of FusionRed and FusionRed-MQ at Low Irradiances.*” Mukherjee, S.; Thomas, C.; Wilson, R.; Simmerman, E.; Hung, S. T.; Jimenez, *ChemRxiv* 2022. <https://doi.org/10.26434/chemrxiv-2022-2dr03>

### 4.2. Introduction

Fluorescence-based bioimaging and biosensing, such as multi-color microscopy, Förster resonance energy transfer (FRET) and fluorescence lifetime imaging microscopy (FLIM), is the primary non-invasive approach to image biological systems. [3, 4, 204-207] Due to Abbe’s diffraction limit, imaging sub-wavelength biological substructures requires innovations beyond the scope of conventional microscopy. [208] Probing this “super-resolution” scale has high value to modern biology and has garnered the attention of scientists and engineers from a range of specializations - making it a popular interdisciplinary field of research and development. [209-214] Techniques like stimulated emission depletion microscopy (STED), photoactivated localization microscopy (PALM), reversible saturable optical fluorescence transition microscopy (RESOLFT), and other single-molecule localization microscopy (SMLM) overcome Abbe’s limit through biochemical or optical modulation of single emitters. [215-218] Most SMLM methods, including single-molecule active control microscopy (SMACM), bleaching/blinking assisted localization microscopy (BALM) and super-resolution optical fluctuation imaging (SOFI), rely on photo-activatable or photo-switchable emitters. [219] These emitters can be activated and/or converted to a new species upon irradiation, exhibiting either molecular binding/modification or spontaneous stochastic blinking. [219-223] Techniques that exploit stochastic blinking, such as SOFI, rely on post-processed data analysis algorithms where traditionally acquired fluorescence time-lapse images are analyzed to obtain spatio-temporal correlations for stochastic fluctuations

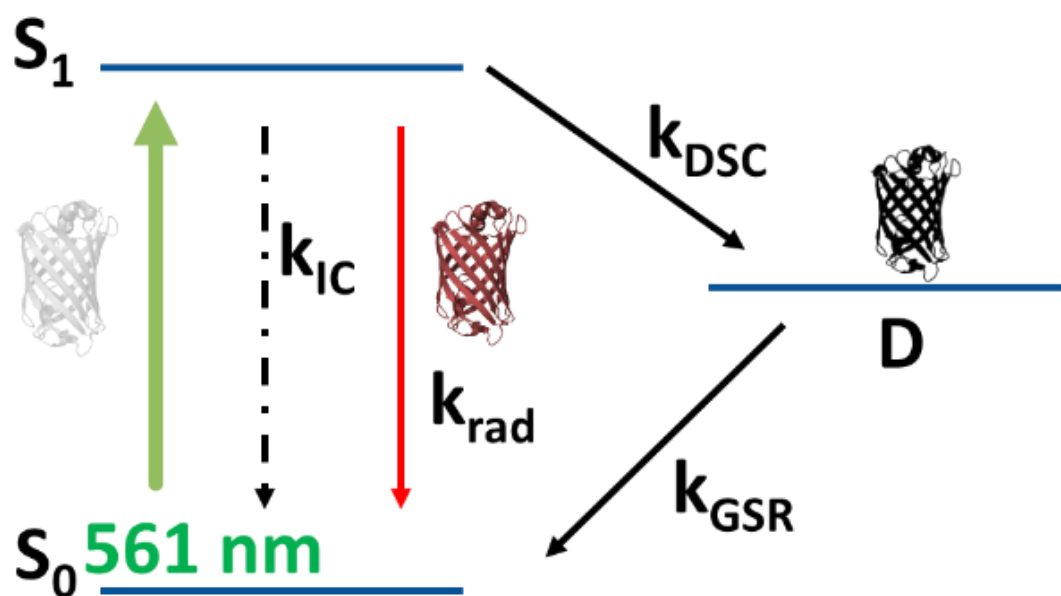
during emission. [219] Current efforts in this domain focus on application-based analysis of such methods and molecular interpretation of the nature of the blinking process for different emitters. A majority of the literature treats emitters simply as tools, with limited analysis of their dark state kinetics and photophysics. [224-226] Unwanted artifacts in the determination of spatial locations for these emitters can thus arise, as a major assumption underlying such stochastic methods is that the signals from single fluorophores are additive as long as the fluorescence dynamics of the molecules are independent. [226]

Genetically encoded fluorophores, such as fluorescent proteins (FPs), are popular and powerful choices to image biological systems and often exhibit dark molecular states which manifest in stochastic blinking behavior at the single molecule level. [227-231] Extensive mechanistic studies have elucidated a diverse range of mechanisms and timescales for stochastic and tuned dark-state conversion, including electron-transfer reactions, excited-state proton transfer, chromophore and sidechain conformational changes. For example, a number of proteins from the avGFP line display a spontaneous, light-driven, pH-dependent dark-state conversion pathway. [232-234] In such cases a pH-dependent dark state conversion is often attributed to a proton transfer between amino acid residues and the chromophore. The red FP (RFP) DsRed exhibits a similar light-driven dark-state conversion through a different, pH-independent process. [233] Photoconvertible FPs (PCFPs) like IrisFP, mEosFP and Dendra derivatives exhibit dark-state conversion via chromophore distortions due to side-chain conformational changes, chromophore twisting motions or by proton transfer from the triplet or a radical ground state. [90, 235-238] Selectively engineered photo-switchable or reversibly-switchable fluorescent proteins (rsFPs) exhibit highly efficient transfer to dark states, allowing the molecules to be switched on and off with either light-driven pH-dependent pathways or reversible chromophore conformational changes such as a *cis-trans* isomerization along the methylidene bridge connecting the two conjugated rings in the chromophore moiety. [202, 239-240] Dark states can be selectively accessed or depopulated using optical or thermal modulation. [241-247] For example, pulsed excitation with resonant frequencies was used to populate and depopulate the dark state in FPs like AcGFP and rsFastLime, providing selective modulation of fluorescence. [244] Varied mechanisms of accessing dark states such as those listed above have made FPs natural choices for several SMLM techniques.

Despite the extensive advances in the study of dark-state dynamics, many of the above-mentioned studies far preceded the development of advanced SMLM techniques such as SOFI. Several previous studies were constrained by experimental challenges to probing dark state dynamics with irradiances on the order of  $\text{kW}/\text{cm}^2$  to attain adequate signal to noise ratios. [227-233] While these studies explore the fundamentals of dark states in FPs, probing dark-state dynamics at high irradiances comes with a three-fold disadvantage. These conditions accelerate permanent photodegradation of the fluorophore, make extrapolation to much lower irradiance as in widefield and SMLM regimes difficult, and cause light-induced photo-toxicity and photodamage to biological systems. [219, 248, 249] It is also worth noting that blue shifting the excitation light increases phototoxicity, thus highlighting the need for development of new bright and red-shifted fluorophores as well as for detailed photophysical characterization of existing red fluorophores – which additionally allow for deeper imaging in comparison to blue shifted analogues. [249, 250] When extrapolating the rate constant of ground-state recovery ( $k_{\text{GSR}}$ ) at low irradiances from high irradiance measurements,  $k_{\text{GSR}}$  is commonly assumed to be proportional to the excitation rate ( $k_{\text{Ex}}$ ). [229, 251] Recovery from a dark to a fluorescent state is often a consequence of conformational switching such as a dark-trans to fluorescent-cis isomerization of the FP chromophore. [38] Since such conformational switches are often energetically controlled, the excitation dependence of  $k_{\text{GSR}}$  may originate from the absorption of the excitation photons by dark state species and/or the rise of local temperature due to high irradiance. [229, 251, 252]

FusionRed and its sibling TagRFP-T exhibit fluorescence intermittency in live-cell imaging using TIRF microscopy with camera acquisition timescales of 50 ms. [253] The study demonstrated the potential to achieve a theoretical spatial resolution beyond the diffraction limit ( $\sim 25\text{--}30\text{ nm}$ ) with FusionRed using SMLM methods like BALM and SOFI. [253] In our previous work, we developed FusionRed-MQV, a FusionRed variant with 3-fold higher molecular brightness developed using a combination of lifetime-based microfluidic selection and site-directed mutagenesis. [38] Additionally, we found that the substitution C159V in FusionRed resulted in a brighter variant which showed a monoexponential photobleaching trace contrary to the biexponential behavior exhibited by the parental RFP FusionRed. Variants lacking the C159V substitution, like the 2.5-fold brighter FusionRed-MQ exhibited biexponential photobleaching traces with a fast fluorescence decay component ( $\sim \text{s}$ ) followed by a significantly slower decay component ( $>100\text{ s}$ ). [38] The faster component was attributed to reversible photobleaching,

where fluorescent molecules are trapped in the dark state. [38] This was verified by employing a high-energy 438 nm pulse (~2 s; 50% duty cycle) with a continuous 560 nm excitation scheme which resulted in distinct reversible photoswitching for FusionRed variants with a Cys residue at position 159. [38] High energy 438 nm light prompts a return to the fluorescent state from a dark state, suggesting that the lower energy 560 nm excitation minimally perturbs the ground state recovery process. [38] Findings from this and other studies, including crystal structure data, indicate a possible interconversion of the FusionRed chromophore from a fluorescent *cis* to a dark *trans* isomer. [182, 200] Moreover, in the low irradiance regime of 1-10 W/cm<sup>2</sup>, the temperature increase in the vicinity of an FP molecule can be considered negligible. [254] Based on these observations, we hypothesized that the rate of recovery to the ground state ( $k_{\text{GSR}}$ ) is independent of the excitation rate ( $k_{\text{EX}}$ ) under low irradiances for FusionRed and FusionRed-MQ. To verify these claims, we explore the relatively uncharted territory of dark state kinetics of these two RFPs under low irradiances, central to widefield and SMLM techniques like SOFI. We do so by extracting on-off statistics with single molecule imaging and exploiting the kinetics of reversible photobleaching using widefield excitation on ensemble RFPs. We combined the strength of each approach to quantitatively extract rate constants of dark-state conversion ( $k_{\text{DSC}}$ ) and ground-state recovery ( $k_{\text{GSR}}$ ) using a three-state model (Figure 4.1). We performed simulations utilizing Monte Carlo methods to bridge the two extremes in imaging modalities and qualitatively validate the three-state model. Finally, we propose a structural model to rationalize the dark-state dynamics observed for these FPs.



**Figure 4.1. A three-state model of the photophysics of a fluorophore central to this study.** The 561 nm excitation source allows access to the excited electronic ( $S_1$ ) state from the ground electronic ( $S_0$ ) state. Following this, the fluorophore can either return to  $S_0$  or access a long lived dark ( $D$ ) state. It was demonstrated in Chapter 3 that the  $D$  state can be depopulated efficiently using a 438 nm light. [38] The arrow labels  $k_{ex}$ ,  $k_{IC}$ ,  $k_{rad}$ ,  $k_{DSC}$  and  $k_{GSR}$  indicate the rate constants for excitation, non-radiative (internal conversion), radiative emission,  $S_1$  to dark state conversion and dark state to  $S_0$  recovery, respectively. Permanent photobleaching from the  $S_1$  and the  $D$  states are ignored in the regime of low irradiances.

## 4.3. Methods

### 4.3a. Experimental methods and data collection

#### i. Cell growth and protein purification

FusionRed and FusionRed-MQ in the pBad-His plasmid were transformed into the *E. coli* Top10 strain via heat shock and grown for 45–60 minutes in LB media in a shaker at 37 °C and 230 rpm. The transformants were plated on agar plates with 100 µg/mL ampicillin and 0.2% arabinose (Sigma Aldrich) overnight at 37 °C. Colored colonies were grown in 200 mL 2XYT (VWR) liquid

cultures with 100 µg/mL ampicillin for 1–3 hours at 37 °C and 230 rpm to an OD of 0.6. Arabinose was then added (0.2%) to induce protein expression for 16–24 hours at 28 °C and 230 rpm. The cells were pelleted, chemically lysed (B-PER, Thermo Fisher Scientific) and the 6-His tagged FPs were isolated on Ni-NTA columns (Thermo Fisher Scientific) by gravity filtration, eluting with 250 mM imidazole (Sigma Aldrich). Excess imidazole was removed with desalting columns (GE Healthcare) with dialysis buffer (150 mM NaCl, 50 mM Tris-HCL, pH 7.4) as an eluent.

## **ii. Single molecule measurements**

*Preparation of glass slides and coverslips:* Minimizing the presence of fluorescent impurities is of particular concern in single-molecule studies. To reduce artifacts from impurities in our measurements, we found that plasma-cleaned glass slide chambers were best suited to single molecule TIRF. [255] Before plasma cleaning, the slides and coverslips (22 x 40 mm, No 1, VWR) were cleaned with dilute HCl then washed with Alconox detergent and rinsed with deionized water, then soaked in methanol overnight to dislodge large contaminants. A custom aluminum slide holder held the slides and coverslips inside a reactive ion etcher, such that both sides were exposed to the plasma. The slides were then exposed to 300 s of O<sub>2</sub> (Bias: 50 W and SCCM: 50 mTorr) plasma to remove organic contaminants and to charge their surfaces, followed by 60 s of Ar plasma (Bias: 50 W and SCCM: 50 mTorr) to minimize presence of remaining reactive oxygen species. The slides were used within 24 hours after plasma cleaning to avoid recontamination and loss of surface charge.

*Sample preparation:* The pure protein samples were diluted with Tris-HCl buffer (pH ~7.4–8.0) and loaded by slow ejection from a 200 µL micropipette. It was determined that FP concentrations >300 pM caused crowding of FPs in the field of view and failure of our spot analysis algorithm to report blinking trajectories, whereas concentrations <100 pM resulted in such sparse distribution that it became difficult to find the correct focus height and provided few data points. Additionally, a washing procedure was developed to minimize the presence of non-adhered FPs in solution and thus minimize free FP diffusion into the imaging plane. The loaded chamber was left in the dark for 10–15 minutes to allow FPs to settle onto the imaging surface, then a volume of imaging buffer (150 mM HEPES, 100 mM NaCl, pH 7.4) equal to the volume of the loaded sample was passed

through the chamber 4–6 times, with 2-minute intervals between washes. The liquid was slowly ejected by a micropipette on one side of the chamber while filter paper was used to absorb the liquid flowing out from the other side. This washing procedure helped to maximize signal-to-background ratios and minimize artifacts from non-specific adhesion to the glass surface for FPs in solution.

*TIRF Imaging:* The samples were imaged with TIRF microscopy on an Olympus IX-73 inverted microscope. The microscope is accessorized with an Olympus cellTIRF-1Line system fiber coupled to a laser (Toptica iChrome MLE). An Olympus 60x-in-oil (NA:1.42) TIRF objective and an EMCCD camera (Andor iXon 897) were used for the single molecule experiments. A schematic of this system has been provided in Appendix Figure A3.1. To measure the excitation rate, the objective focus was first determined by imaging a dye sample under bright-field illumination, then the sample was removed and the laser was focused at the ceiling (approximately 2 m beyond the sample location) for this z-position of the objective. The irradiance measurements were carried in this normal ( $I_{\text{Normal}}$ ) to the imaging plane position using a power meter (X-cite). The excitation intensity of the evanescent field ( $I_{\text{TIRF}}$ ) was calculated from the incident intensity ( $I_{\text{Normal}}$ ), the indices of refraction ( $\eta_2$ ,  $\eta_1$ ), and the incident angle. [256] The calculations of excitation rates for normal and TIRF illumination are presented in Appendix 3: Section 2. To image samples, a cropped area of the imaging plane (~128 x 128 pixels on a 256 x 256-pixel binning) corresponding to the region of highest intensity of the laser profile was selected. Then for the lowest value of irradiance ( $1\text{W}/\text{cm}^2$ ), 100 nm fluorescent beads (TetraSpeck) were used to first determine the approximate z-focus, and the motorized stage (Prior) was moved in the x–y plane to the position of single FP molecules, to determine an accurate focus. The experiment was started after moving to an adjacent spot (~100  $\mu\text{m}$ ) outside this imaging area of the previous step, where drift on the z-axis was minimal. This was done so as to minimize photobleaching of single molecules.

### **iii. Ensemble measurements:**

Bright bacterial colonies on the agar plates described above were chosen for time-lapse photobleaching experiments. Two to three colonies were transferred to microcentrifuge tubes and washed with 500  $\mu\text{L}$  imaging buffer by vortexing for ~20 s. The cells were centrifuged at 3000-

5000 RPM for 60 s, and the washing buffer was removed. The pelleted cells were then resuspended in the same buffer to an OD in the range of 0.1 to 0.5 to get a cell density suitable for imaging. A glass coverslip and slide were cleaned with Alconox detergent, rinsed with deionized water, and blown dry with filtered compressed air. 10–20  $\mu\text{L}$  of the cell mixture was added between the coverslip and slide, which was imaged on an Olympus IX-73 inverted microscope system. Samples were excited by 560 nm continuous wave LED illumination (Lumencor). Fluorescence bleaching measurements were taken with the 20x or 40x-in air objective lens (Olympus). The fluorescence was collected through a 629/56 nm band-pass filter by a sCMOS camera (Andor Zyla). Videos were collected with 10–50 ms exposure times and frame rates of 20–32 FPS for the fast and reversible component of the decay and 10–20 FPS for the slow and irreversible component, and with irradiances ranging from 1–20  $\text{W}/\text{cm}^2$ . We performed three independent trials where each trial for an FP involved a technical replicate with  $\sim 10$ –20 cells to gain consistent bleaching traces.

## 4.3b. Data analysis

### i. Single molecule data analysis

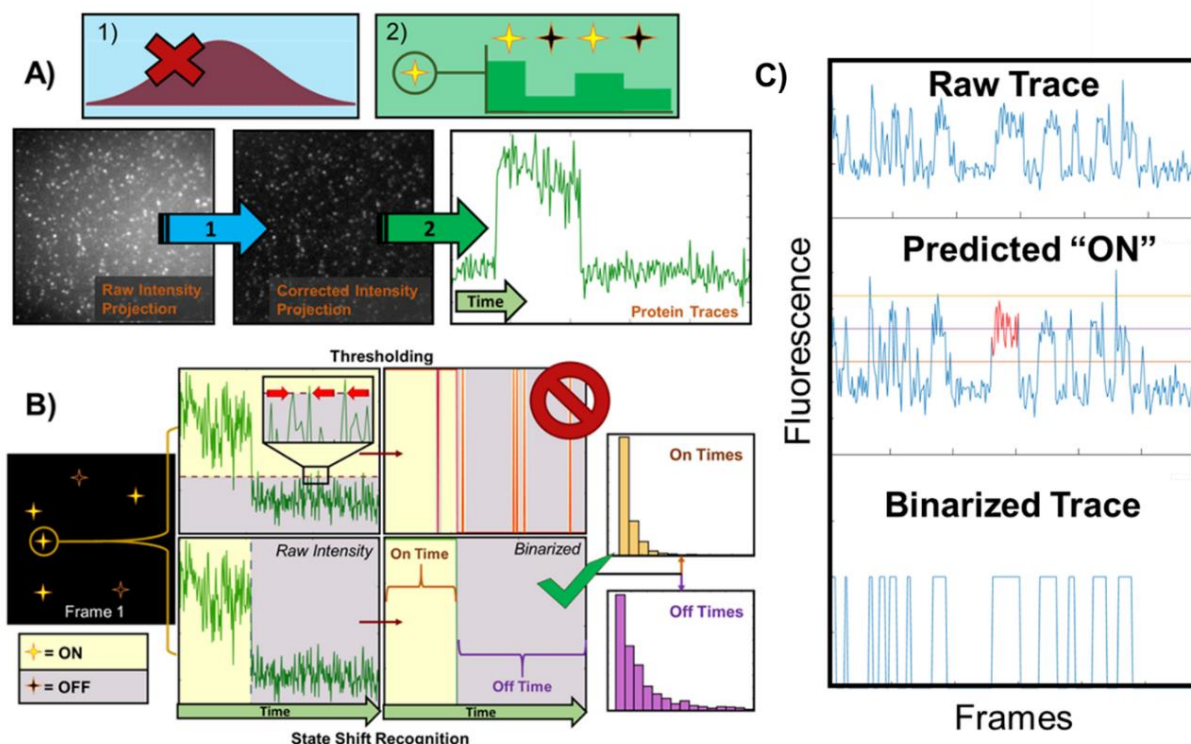
Single molecule data analyses from imaging videos were carried out using two independent scripts: One for spatial identification of bright spots followed by one for temporal and intensity analysis of these bright spots. Figure 4.2 shows a schematic representation of this workflow.

*Spot identification script:* Despite cropping, there is a systematic  $\sim 10\%$  intensity variation across the imaging plane with a Gaussian profile. To account for this, the videos were iteratively fitted to a Gaussian intensity correction function to correct for the laser background, primarily from residual scattering, and normalizing for the variation in intensity due to the spatial mode of the excitation laser. (Appendix Figure A3.1) Following Gaussian correction, our analysis also revealed a biexponential decay of mean intensities across the timeframe of the video. Therefore, the mean intensity of the videos was fitted to a biexponential function, which revealed a fast component of decay ( $\sim 1$  s) along with a slow component of decay ( $> 3$  s). While the timescale of the slow component of decay varied with the incident intensity, the fast component was seen to be fairly consistent (Appendix Figure A3.2). Additional checks with blank solutions also revealed a



consistent timescale for the fast component of decay. Therefore, after the Gaussian correction, a secondary correction was incorporated to account for the quick exponential drop in the overall light intensity. Given that this decay was also found in blank medium and was missing from the laser's temporal profile, we attribute it to diffusion or photobleaching of impurities in our blank medium or the objective oil. Following the Gaussian and exponential corrections, the algorithm identifies a number of bright locations equal to an input of the predicted number of single FPs in the video. This number was set between 50 and 500 FPs depending on the field of view, the efficiency of binding the FPs to the glass surface and the concentration of the protein used. The algorithm extracts the brightest pixels in the maximum intensity projected image of the video from the user defined input value for the number of single FPs. It then iteratively appends the location of maximum value after it passes a check, which involves scanning a pixel grid surrounding the pixel centered at maximum value based on the statistical distribution of the brightness around the grid.

*Temporal and Intensity analysis:* To extract information on real “on” and “off” blinking events we drew inspiration from the work of Watkins and Yang. [257] In order to find single on and off events in a trajectory of a single molecule, we used an intensity change point approach. Our spot analysis script provides us with intensity corrected trajectories with time for the brightest spots arranged in the ascending order of mean intensity. We therefore assessed the first five and the last five trajectories, based on the brightest and the dimmest spots identified through the previous script. In each case, we estimated the average single molecule on intensity to provide as an input for this code. Although many approaches utilize histograms from the intensities of each frame to effectively threshold and binarize a trajectory, it is difficult to use this approach for our data sets at the lower and the upper bounds of irradiances, which are characterized by increasingly longer  $\tau_{\text{ON}}$  or  $\tau_{\text{OFF}}$ , respectively. [258]



**Figure 4.2. Analysis workflow.** (a.) The background correction and molecule selection process. (b.) Our analysis workflow contrasted against standard thresholding methods. Note that the state shift recognition workflow identifies points where possible changes occur and binarizes on the basis of those points. This avoids the noise-based artifacts shown in the thresholding approach, where it can be difficult to set an arbitrary threshold that is not crossed by noise. (c.) The binarization of a raw trace based on the thresholding algorithm.

Following the input of the five potential on events, our algorithm performs two steps. First, it fits the change of intensity between frames for the entire dataset to a Gaussian distribution. It should be noted that a protein turning on or off produces a relatively small change in intensity that falls within the noise distribution. Therefore, it is not possible to separate these events from background noise with equal or higher intensity without additional information, whereas noise below this threshold can be discarded. As a result, all frames with a change value of the threshold or higher are earmarked as possible changes of state. Thus, the primary question is where to set the threshold for optimal recognition. (Appendix Figure A3.3) The theoretical minimum intensity changes for a molecule result from a case where the protein turns on exactly halfway through a frame's acquisition time. This results in a change of intensity of  $\frac{1}{2}$  the protein signal, followed by a second

change of the same magnitude. As such, a good baseline estimate for the noise threshold value is 50% of the expected signal. Empirically, we have found that  $\sim 1.5 \sigma$  (standard deviation) of the overall noise produced consistent results. This boundary excludes 86% of overall noise, but it is worth noting that the on state contributes to higher (shot) noise. As such, this boundary excludes approximately 70% of the larger standard deviation on distribution. The second step of the algorithm considers the intervals between each on and off point individually and binarizes each segment on the basis of a user-defined  $\tau_{\text{ON}}$ . (Appendix Figure A3.2 and Figure 4.2c) This further minimizes the contributions of noise by averaging each segment. Segments that are more than  $2\sigma$  above the model  $\tau_{\text{ON}}$  are considered multi-molecule events and the corresponding trace is discarded. All segments that end at the last frame of the video (on or off) are also discarded to avoid artifacts, notably from permanent photobleaching or denaturation.

### iii. Ensemble data analysis:

Photobleaching data analysis was carried out with a previously reported scheme. [38] In brief, a pipeline in the CellProfiler suite was used to identify bacteria in an imaging plane of uniform irradiance and provide normalized intensity trajectories. [189] The intensity trajectories were analyzed and fit to biexponential traces using a custom fitting program in MATLAB (Appendix Table S4.1).

## 4.4. Results

### 4.4a. Simulations and measurements of single molecule photophysics

We performed simulations of blinking trajectories for single molecules in a three-state ( $S_0$ ,  $S_1$  and D) model with our hypothesis. (Figure 4.1) As expected, the results show that  $\tau_{\text{ON}}$  decreases with both shorter values of  $\tau_{\text{DSC}}$  and higher values of  $k_{\text{EX}}$ , whereas  $\tau_{\text{OFF}}$  depends only on the value of  $\tau_{\text{GSR}}$ . The details of the algorithm and simulated results are described in Appendix 3: Section 5. In brief, the residence of single FPs in each state ( $S_0$ ,  $S_1$  or D) at each time step is simulated using a combination of Monte Carlo methods weighted by probabilities of state change, governed by the  $k_{\text{EX}}$ ,  $k_{\text{rad}}$ ,  $k_{\text{non-rad}}$  (IC),  $k_{\text{DSC}}$  and  $k_{\text{GSR}}$ . An “on” or a fluorescent event is detected every time the

molecule successfully completes an S<sub>0</sub>-S<sub>1</sub>-S<sub>0</sub> cycle, without going through D. (Figure 4.1) The predictions of the simulations accurately represent experiments, where data from three independent trials revealed a hyperbolic dependence of the  $\tau_{ON}$  of FusionRed and FusionRed-MQ on the excitation rates. The  $\tau_{OFF}$  for both FPs showed little or no dependence on excitation rates as summarized in Table 4.1 and Figure 4.3.

Our single molecule measurements were carried out with irradiation on the order of  $\sim 10^3$  photons/s, which is significantly lower than the optical saturation limit for FPs ( $\sim 10^8$  excitation-photons/s). Therefore, the population in the S<sub>1</sub> electronic state can be treated under a steady state approximation. At low irradiances, we assume minimal absorption from the dark states and therefore consider the reverse dark-state conversion to be negligible. [38] The timescales of permanent photobleaching are significantly longer than that of dark-state lifetime (Appendix 3; Section 4), hence it was neglected in this scheme. Incorporating these assumptions, we arrive at equations 1 and 2 to calculate  $k_{DSC}$  and  $k_{GSR}$  for FusionRed and FusionRed-MQ, [259] and the calculated values are presented in Table 4.2 and Figure 4.4.

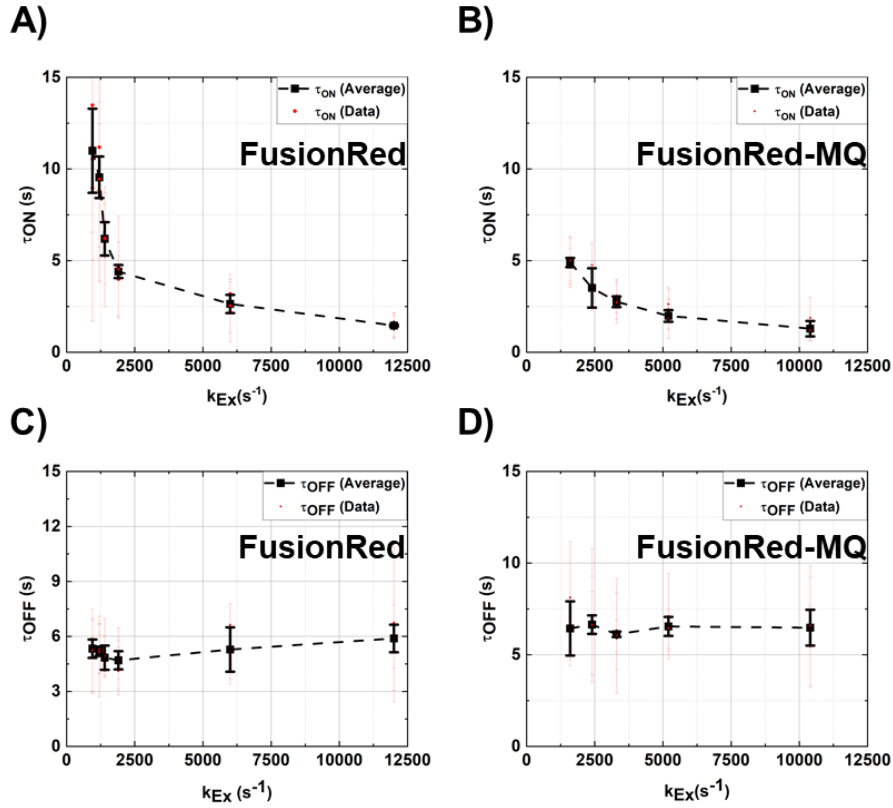
$$\tau_{ON} = \frac{k_{em} + k_{non-rad} + k_{dsc}}{k_{ex} \cdot k_{dsc}} \quad (1)$$

$$\tau_{OFF} = \frac{1}{k_{GSR}} \quad (2)$$

**Table 4.1. Values of  $\tau_{ON}$  and  $\tau_{OFF}$**

FusionRed			FusionRed-MQ		
$k_{Ex}$ (s <sup>-1</sup> )	$\tau_{On}$ (s)	$\tau_{off}$ (s)	$k_{Ex}$ (s <sup>-1</sup> )	$\tau_{On}$ (s)	$\tau_{off}$ (s)
950	$10.9 \pm 2.3$	$5.3 \pm 0.5$	1600	$4.9 \pm 0.3$	$6.4 \pm 1.5$
1200	$9.5 \pm 1.1$	$4.7 \pm 1$	2400	$3.5 \pm 1.1$	$6.6 \pm 0.5$
1400	$6.2 \pm 0.9$	$5.2 \pm 0.2$	3300	$2.8 \pm 0.3$	$6.1 \pm 0.1$
1900	$4.4 \pm 0.4$	$4.8 \pm 0.7$	5200	$2.0 \pm 0.3$	$6.5 \pm 0.5$
6000	$2.6 \pm 0.5$	$5.3 \pm 1.2$	10400	$1.3 \pm 0.4$	$6.5 \pm 1.0$
12000	$1.2 \pm 0.5$	$5.8 \pm 0.8$			

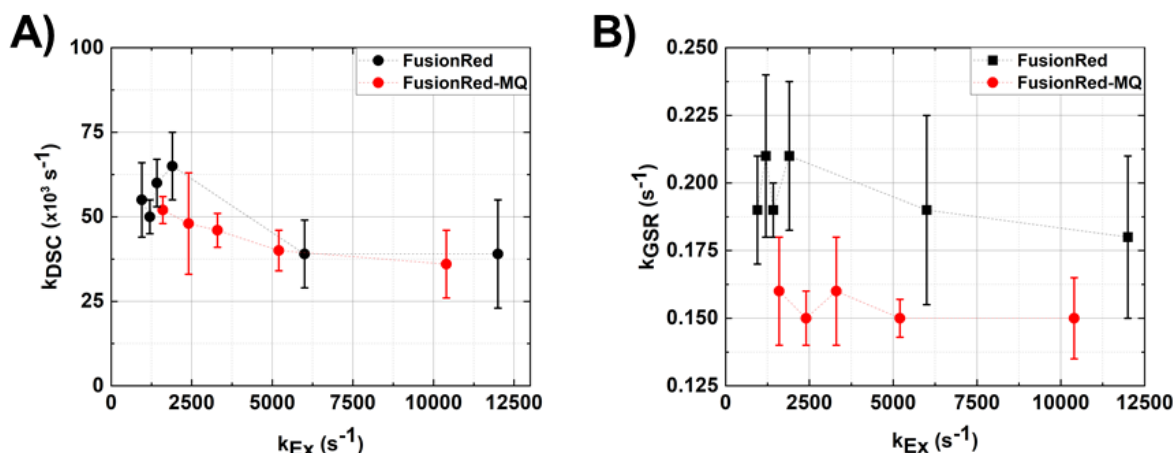
The  $k_{DSC}$  and  $k_{GSR}$  remain consistent under the range of irradiances probed in this study. However, FusionRed clearly shows an  $\sim 30\%$  higher  $k_{GSR}$  than FusionRed-MQ.



**Figure 4.3. Results from single molecule blinking experiments.** Measured  $\tau_{ON}$  with respect to excitation rate for (a.) FusionRed and (b.) FusionRed-MQ; and measured  $\tau_{ON}$  with respect to excitation rate for (c.) FusionRed and (d.) FusionRed-MQ. The black data points are the mean  $\tau_{ON/OFF}$  extracted from three independent experiments (error bars indicate the confidence intervals for the values extracted from fits).

**Table 4.2. Estimation of  $k_{DSC}$  and  $k_{GSR}$**

FusionRed			FusionRed-MQ		
$k_{Ex}$ ( $s^{-1}$ )	$k_{DSC}$ ( $\times 10^3 s^{-1}$ )	$k_{GSR}$ ( $s^{-1}$ )	$k_{Ex}$ ( $s^{-1}$ )	$k_{DSC}$ ( $\times 10^3 s^{-1}$ )	$k_{GSR}$ ( $s^{-1}$ )
950	$55 \pm 11$	$0.19 \pm 0.02$	1600	$52 \pm 4$	$0.16 \pm 0.02$
1200	$50 \pm 5$	$0.21 \pm 0.03$	2400	$48 \pm 15$	$0.15 \pm 0.01$
1400	$60 \pm 7$	$0.19 \pm 0.01$	3300	$46 \pm 5$	$0.16 \pm 0.01$
1900	$67 \pm 10$	$0.21 \pm 0.03$	5200	$40 \pm 6$	$0.15 \pm 0.01$
6000	$39 \pm 10$	$0.19 \pm 0.04$	10400	$36 \pm 10$	$0.15 \pm 0.02$
12000	$39 \pm 16$	$0.18 \pm 0.03$			



**Figure 4.4. Variation of rate constants.** (a.)  $k_{\text{DSC}}$  and (b.)  $k_{\text{GSR}}$  with respect to excitation rate  $k_{\text{Ex}}$  measured from single-molecule blinking experiments. The error bars indicate standard deviation errors calculated from measured  $\tau_{\text{OFF}}$  and  $\tau_{\text{ON}}$ .

#### 4.4b. Simulating ensemble behavior

We first generated normalized sum fluorescence traces for  $\sim 10^3$ – $10^4$  simulated blinking single FP trajectories using the average  $k_{\text{DSC}}$  and  $k_{\text{GSR}}$  from the single molecule experiments. The simulations indicated a larger dark fraction for FusionRed-MQ in comparison to FusionRed. We then performed simulations to obtain normalized sum fluorescence traces for blinking single FP trajectories (3 sets of 150 emitters each) by varying one rate constant at a time, including  $k_{\text{Ex}}$ ,  $k_{\text{DSC}}$  and  $k_{\text{GSR}}$ . The results indicated that the dark fraction increases with increasing  $k_{\text{Ex}}$  and  $k_{\text{DSC}}$  but has no or little dependence on  $k_{\text{GSR}}$  in the vicinity of measured values in the single molecule blinking experiments. This observation encouraged us to pursue an alternate analysis compared to single molecule experiments, to quantify the  $k_{\text{DSC}}$ . The results of these simulations are presented in Appendix 3: Sections 5f, g and h. The normalized fluorescence signals are proportional to the normalized population of FPs on the ground state  $S_0$ , thus the fractions of FPs in  $S_0$  and dark state D as a function of time can be estimated as shown in Appendix Figure A3.13.

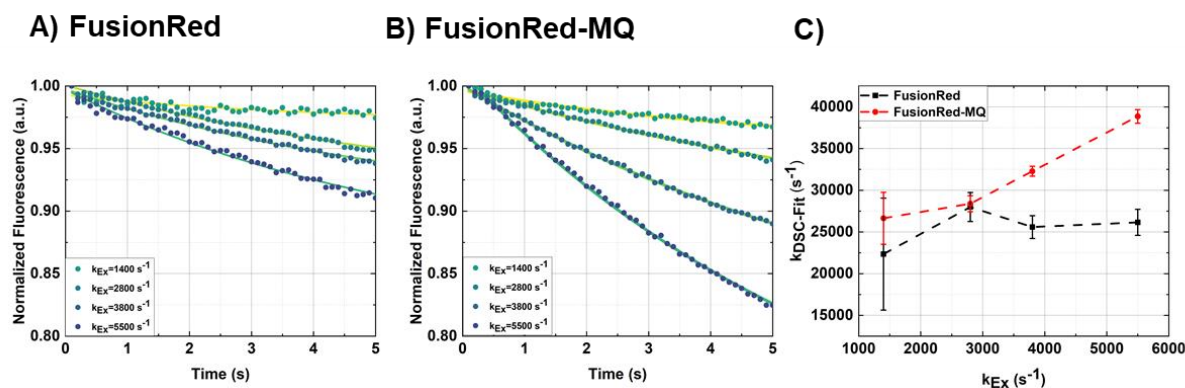
#### 4.4c. Ensemble photobleaching

*E. coli* expressing FusionRed and FusionRed-MQ were exposed to normalized widefield excitation. The decay profiles of both FPs were recorded and fit with a biexponential function as presented in Appendix 3; Section 4. Based on our previously reported pulsed photobleaching measurements, we assigned the slower decay component to permanent photobleaching and the faster component to reversible photobleaching due to dark-state conversion. [38] Experimental results showed a larger dark fraction for FusionRed-MQ compared to FusionRed.

To extract  $k_{DSC}$ , we consider a three-state model to represent the kinetic processes of a fluorescent protein in the first five seconds of a photobleaching trace, where the contribution of the permanent bleaching component is minimal (<10%) and reverse dark-state conversion (D to  $S_1$ ) can be ignored. The rate equations can be written in the matrix form,  $\frac{dn}{dt} = A\mathbf{n}$ , as explicitly shown in Equation 3, where  $n$  is the population for each state.

$$\frac{d}{dt} \begin{bmatrix} n_{S0} \\ n_{S1} \\ n_D \end{bmatrix} = \begin{bmatrix} -k_{Ex} & k_{rad} + k_{non-rad} & k_{GSR} \\ k_{Ex} & -(k_{rad} + k_{non-rad} + k_{DSC}) & 0 \\ 0 & k_{DSC} & -k_{GSR} \end{bmatrix} \begin{bmatrix} n_{S0} \\ n_{S1} \\ n_D \end{bmatrix} \quad (3)$$

The analytical solutions can be obtained by solving the eigenvalues and eigenvectors for rate equations (Appendix 3; Section 6). The observed fluorescence is proportional to the population in  $S_0$  for an ensemble measurement. The resolution of the single molecule measurement in the time domain is limited by the camera acquisition time and is significantly larger than the  $\tau_{DSC}$ , but smaller than the  $\tau_{GSR}$ . Therefore, our single molecule measurements provide an accurate estimation of the  $k_{GSR}$ , but not the  $k_{DSC}$ . However, we can utilize the  $k_{GSR}$  values from the single molecule measurements to fit the reversible bleaching curves of the ensemble measurements to extract an accurate estimation for the  $k_{DSC}$  across an irradiance range. The fluorescence decay data were fit to the analytical expression of  $n_{S0}$  to get  $k_{DSC}$  values (bound  $k_{GSR}$ ). The fitting results and the estimated  $k_{DSC}$  values are provided in Figure 4.5 and Table 4.3. While there seems to be minimal light-driven behavior for  $k_{DSC}$  for FusionRed, fitting revealed an uptick of the  $k_{DSC}$  with increasing excitation rates for FusionRed-MQ.



**Figure 4.5.** Photobleaching fits from a three-state model, for (a.) FusionRed and (b.) FusionRed-MQ, across varying irradiance ranges. (c.)  $k_{DSC}$  obtained from fits. Details of the fit are provided in Appendix 3; Section 6.

**Table 4.3.** Estimated  $k_{DSC}$  values from fitting ensemble traces

Excitation Rate ( $s^{-1}$ )	$k_{DSC}$ ( $\times 10^3 s^{-1}$ )	
	FusionRed	FusionRed-MQ
1400	$22.3 \pm 6.7$	$26.6 \pm 3.1$
2800	$27.9 \pm 1.8$	$28.4 \pm 0.9$
3800	$25.6 \pm 1.4$	$32.3 \pm 0.6$
5500	$26.1 \pm 1.6$	$38.9 \pm 0.8$



## 4.5. Discussion

Conventional models to characterize stochastic blinking usually rely on a two-state “on” and “off” system, [229, 251] when the excitation rate is much higher than the sampling rate of the detector used. [260, 261] When modeling single molecule blinking with such models under these conditions, one can assume emission takes place when the emitter is in the fluorescent ( $S_0$ ) state for two consecutive simulation time steps ( $\ll$  dark state lifetime). Imaging methods such as widefield and TIRF ( $\sim 10^{2-6}$  photons/s) employ low excitation rates for SMLM schemes. In such cases, excitation and fluorescence photons are infrequent and long photon acquisition times are required to quantify blinking of FPs. Therefore, the detector is blind to the molecule entering or exiting the  $S_1$  state, and the process needs to be modeled differently. To address this situation, we devised an algorithm incorporating Monte Carlo methods with acquisition time steps longer than excited-state lifetime but shorter than  $\tau_{\text{GSR}}$  ( $\sim 100$  ms). This algorithm accurately represents our single molecule experiments, where excitation rates and acquisition timescales are comparable. With this simulation algorithm, it is reasonable to assume that one always observes the FP in  $S_0$  or D, and  $S_1$  is only rarely populated in our three-state model. The fluorophore starts from  $S_0$  or D and returns to these states in each time cycle, and the probability of a state change in each cycle is determined by rate constants of radiative emission, internal conversion, DSC and GSR processes. Our simulation of single molecule fluorescence correctly predicted the  $\tau_{\text{ON}}$  dependence on  $k_{\text{EX}}$  and  $k_{\text{DSC}}$ , and the  $\tau_{\text{OFF}}$  dependence on  $k_{\text{GSR}}$ . By repeating multiple cycles of single molecule simulations, the average on and  $\tau_{\text{OFF}}$  can be obtained based on a probability density function of an exponential distribution:  $f(t) = \frac{1}{\mu} e^{-t/\mu}$  (Appendix 3: Sections 5a-e). The model was then extended to mimic ensemble behavior (Appendix 3: Sections 5f-g). When varying one rate constant at a time, the ensemble dark fraction increases with increasing  $k_{\text{EX}}$  and  $k_{\text{DSC}}$ , but the fluorescence decay profile shows minor variation in the dark fraction even with 30-45% variation in  $k_{\text{GSR}}$ . In other words, the dark fraction is controlled by factors affecting processes that populate the dark state, and the rate-determining ground state recovery process dominates the time constant of the faster component of decay in the fluorescence bleaching profile. Thus, simulations in both single molecule and ensemble average of RFPs qualitatively validated this three-state model.

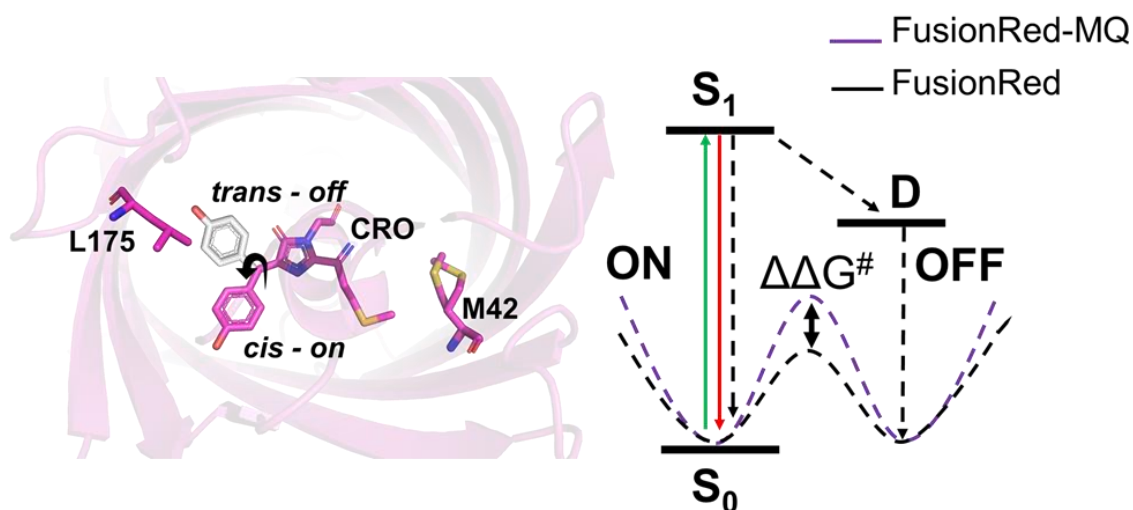
We then developed a new method combining single molecule imaging and ensemble photobleaching measurements of RFPs with low irradiances to quantitatively extract dark state kinetic parameters based on the three-state model. The lowest irradiance in our single molecule measurements corresponds to an excitation rate of  $\sim 1000$  photons/s. At such low irradiances, we expect an emission of less than 100 photons per frame from the dimmer FusionRed (Appendix 3: Section 7). The relatively long  $\tau_{ON}$  and low intensity-change thresholds for on to off (or vice versa) state changes pose a challenge for data analysis under these conditions. For these low irradiances, conventional algorithms to binarize intensity traces inaccurately binarize on and off events due to the low signal to noise ratios. Thresholding a change of state is a challenging step. [257, 262] Instead, we used algorithms based on statistics of intensity variation across frames to predict possible state change points and determine if a “real” on to off (or vice versa) event takes place. [262] The analysis algorithm is also helpful to predict very short  $\tau_{ON}$  for the other extreme case where we use  $\sim 10$ -fold higher excitation rates. Outside sample preparation methods and efficient data analysis algorithms, experimental analysis of such discrete states can also be improved with improved camera technologies like photon-number resolving detections schemes. [263] Additionally, to accurately determine on and off events, our algorithm screened out instances where two molecules might fluoresce at the same time or the molecule undergoes permanent photobleaching or denatures (extracts  $\tau_{OFF}$  only when the event is bounded by an on time and vice-versa). Moreover, it is worth noting that the blinking rates are sensitive to the nature of the electrostatic interaction between the binding surface and the molecule. Glass surfaces under treatments like exposure to plasma are generally negatively charged prompting interactions with the positively charged areas of the protein (Appendix 3: Section 8). [264] For example, the super-resolution imaging modality FLINC (Fluorescence fLuctuation INcrease by Contact) is based on variation of blinking frequency in TagRFP-T (by  $\sim 25\%$ ) through the electrostatic interaction with another non-fluorescent FP (Dronpa). [49] Assuming that FusionRed binds to the glass surface via positively charged surfaces facing outward from the  $\beta$ -barrel, we may expect these effects to be identical for FusionRed-MQ, which has only two internal mutations compared to FusionRed. Therefore, this method of sample preparation might introduce artifacts to blinking dynamics in the context of comparing two FPs with dissimilar residues facing out of the  $\beta$  barrel structure.

Our data analysis algorithms revealed computed rate constants  $k_{DSC}$  and  $k_{GSR}$  obtained from  $\tau_{ON}$  and  $\tau_{OFF}$  histograms, had minimal variation in the irradiance regimes examined for both FPs. The

value of  $k_{\text{GSR}}$  can be accurately determined from the single molecule fluorescence dynamics because the time resolution of the measurement, which is limited by the camera acquisition time, is smaller than the  $\tau_{\text{GSR}}$ . Therefore, we conclusively determined that the  $k_{\text{GSR}}$  for FusionRed is 1.3-fold higher than that of FusionRed-MQ. We also verified our hypothesis that  $k_{\text{GSR}}$  of both FPs is independent of  $k_{\text{Ex}}$  under low irradiances. The  $k_{\text{DSC}}$  values are, however, on the order of kHz, which is  $\sim 2$  orders of magnitude larger than the acquisition rate (Hz) of our single-molecule experiment. While the  $k_{\text{DSC}}$  values are in the range of expected values from previous measurements, [56] and despite high precision, accurate values of  $k_{\text{DSC}}$  under low irradiances had to be determined by a different method. We used an eigenvector-eigenvalue approach to extract analytical expressions for fluorescence decay from the kinetics of our three-state model and fit ensemble biexponential bleaching traces to accurately estimate the  $k_{\text{DSC}}$ . The fitting analysis revealed a light-driven dependency for the  $k_{\text{DSC}}$  ( $\sim 1.5$ -fold increase in the  $k_{\text{DSC}}$  with  $\sim 4$ -fold increase in the  $k_{\text{ex}}$ ) of FusionRed-MQ, whereas  $k_{\text{DSC}}$  of FusionRed remained almost constant with increasing excitation rate. While a degree of heterogeneity can be expected between single-molecule measurements obtained from pure proteins and ensemble measurements in bacterial cytoplasm, it is encouraging to note that FusionRed has low pH sensitivity in the neutral pH range and the pH values of our imaging buffer and bacterial cytoplasm are close ( $\sim 7.4$  vs  $\sim 7.2$ – $7.8$  for *E. Coli*). [265] Thus, we extracted quantitative estimates for the  $k_{\text{DSC}}$  and  $k_{\text{GSR}}$  by combining the strengths of each approach.

Next, we consider the photophysics in the context of protein structure. The crystal structure shows that Met residue at position 42 in FusionRed is located at the imidazolinone end of the chromophore. In FusionRed-MQ position 42 is substituted for a Gln residue, possibly altering the hydrogen bonding patterns at the acylimine end of the chromophore. [38] Meanwhile, the residue at position 175 located above the phenol ring of the chromophore pocket; is a Leu in FusionRed and is substituted to the larger sidechain Met in FusionRed-MQ. [30, 182, 200] The effects of these two mutations on FusionRed's brightness are complementary. Unlike the C159V mutation, Q42M and L175M do not change the profile of the ensemble bleaching, likely preserving the efficient cis to trans dark-state isomerization pathway. [38] It is interesting to note that the amino acid residue analogous to position 175 in FusionRed influences the dark-state behavior of the two photoconvertible FPs IrisFP and mEos4B (S173 vs F173). [237] De-Zitter and co-workers demonstrated that F173 in mEos4B reduced the number of hydrogen bonds maintained by the dark

chromophore in its green form compared to the smaller-sized but hydrogen-bonding capable S173 residue for IrisFP, providing an explanation for the reduced photoswitching contrast. [237] As such, the  $k_{\text{GSR}}$  values can help to identify a possible difference in barrier for dark to fluorescent state interconversion (Figure 4.6) between FusionRed and FusionRed-MQ. The approximate difference in the ground state barrier of bond rotation expected for a trans to cis isomerization can be calculated using a transition state theory approach for interconverting ground state isomers. [125] Under the assumption that the local temperature does not change on irradiation and a similar value of the pre-exponential factor, this difference in barrier ( $\Delta\Delta G^\ddagger$ ) is  $\sim 20$  kJ/mol for FusionRed and FusionRed-MQ. A value of 20 kJ/mol is a reasonable free energy difference that predicts a change of a few possible hydrogen bonds, similar to what was observed for IrisFP, mEos4B and others. [237, 266- 268] An energetic reluctance to switch back to the bright state can therefore be explained by a conformationally restricted chromophore for FusionRed-MQ, suggesting a mechanism for the higher brightness observed for the FusionRed-L175M variant. [30, 38]



**Figure 4.6. Relating blinking to a cis-trans isomerization of the FR chromophore. (a.)** The spatial orientation of the L175 and M42 residues in the FusionRed crystal structure (PDB ID: 6U1A), with the cis and the trans forms of the chromophore. **(b.)** Schematic depiction of a possible barrier for a ground state cis-trans isomerization depicting a switch from an off to an on state.

## 4.6. Conclusions

Bright RFPs with desirable biological properties have a growing role as strong candidates for dual-usage both for SMLM-based imaging and widefield ensemble assays. [250] To address the latter, in this study we quantified the nature and timescales of dark state conversion and ground state recovery for FusionRed and its 2.5-fold brighter progeny FusionRed-MQ. Our approach combined single molecule imaging with ensemble bleaching measurements to extract  $k_{\text{DSC}}$  and  $k_{\text{GSR}}$  of these RFPs. The methods applied in this work are an advancement over our and other previous studies where rate constants were extracted using time and frequency domain measurements at several orders of magnitude higher irradiances, which are often not suitable for widefield fluorescence and SMLM-based imaging assays. [56] To the best of our knowledge, this is the first study to quantify  $k_{\text{DSC}}$  and  $k_{\text{GSR}}$  directly under low irradiances relevant to modern SMLM imaging schemes like SOFI. We verified our hypothesis that  $k_{\text{GSR}}$  of FusionRed and FusionRed-MQ is independent of  $k_{\text{EX}}$  under low irradiances. This is contrary to high irradiance studies where  $k_{\text{GSR}}$  was observed to vary with  $k_{\text{EX}}$ . [229, 251] FusionRed is a plausible candidate for SMLM imaging, [253] and our investigation suggests FusionRed-MQ is a better choice for such applications owing to its 1.8-fold higher quantum yield and higher  $k_{\text{DSC}}$ . [38] The methods and results of this work can be extended to the characterization of other fluorophores with appropriate dark state kinetic models, or incorporated into multi-parametric screening technologies to select FPs with high rates of blinking for methods like SOFI.

## Chapter 5

# Characterizing the pathways of non-radiative decay in the lifetime evolution of a brighter variant of mCherry

### 5.1. Publication Note

Parts of this chapter have been adapted from the article titled “*Directed evolution of a bright variant of mCherry: Suppression of non-radiative decay by fluorescence lifetime selections.*” by Mukherjee S., Manna P., Hung S.T., Vietmeyer F., Friis P., Palmer A. E., and Jimenez R. *ChemRxiv* 2022. <https://doi.org/10.26434/chemrxiv-2022-v3t03>

### 5.2. Introduction

Fluorescence lifetime-based selections on fluorescent protein (FP) libraries, performed on cell-screening platforms such as microscopes or flow cytometers, have been key to the development of the brightest genetically-encoded fluorophores. [39-41] For example, brighter variants of the red FP (RFP) FusionRed have been generated with microfluidic cell-sorting. [30, 38] The brightest RFP and cyan-FP (CFP) to-date, mScarlet ( $\phi = 0.70$ ;  $\tau = 3.9$  ns) and mTurquoise2 ( $\phi = 0.93$ ;  $\tau = 4.0$  ns), respectively were also developed using lifetime selections on imaging platforms. [35, 77] Although these lifetime selections have been unquestionably successful in generating brighter FPs for imaging applications, the intermediates along the evolutionary trajectory generated by these efforts have generally been set aside in favor of focusing on the final product of molecular evolution. Although it is widely acknowledged that lifetime and other photophysical properties such as photobleaching or spectral changes frequently co-evolve, a mechanistic investigation of their interdependence has rarely been pursued. [250]

In the GFP superfamily, the chromophore is formed when a tripeptide of the internal  $\alpha$ -helix embedded in an 11-stranded  $\beta$ -barrel undergoes rearrangement, cyclization, dehydration, and oxidation. [186, 269] This process results in a  $\pi$ -conjugated structure of p-hydroxybenzylidene and imidazolinone rings bridged through a methine carbon. In RFPs, the chromophore conjugation is further extended by a N-acylimine moiety that leads to a red-shift of the absorption and emission spectra. [15, 92] A combination of electrostatic and steric interactions from surrounding amino

acids tune the ground ( $S_0$ ) and the excited ( $S_1$ ) state potential energy surfaces of the chromophore, thus influencing the pathways of excited state depopulation. [58, 62, 74, 82, 125, 270] As a result, the unique chromophore environment within each FP determines its fluorescence lifetime. Each functional (i.e., properly-folded and chromophore-matured) FP generated by an engineering effort provides an opportunity to examine a point within the evolutionary trajectory of fluorescence lifetime and other photophysical properties. These properties may be correlated with a particular pathway across the mutational landscape. [271] The global photophysical assessment of an evolutionary trajectory therefore contains vital information that can inform strategies for FP evolution. This is particularly relevant for developing a new FP tailored to a specific application modality- whether it is fluorescence lifetime imaging (FLIM), Forster Resonance Energy Transfer (FRET), or multicolor widefield imaging. [272]

Many theoretical and spectroscopic studies provide an emerging framework for understanding the factors influencing molecular brightness of FPs. [58, 62, 74, 82, 125, 270] However, understanding the reasons for interdependence of photophysical properties is a major challenge. Observables such as fluorescence lifetime that directly report on excited state depopulation timescales are sometimes found to be correlated with other photophysical properties that are less directly related to these dynamics. For instance, in an attempt to find a bright, far-red emitting FP, Cauty *et al.* generated and sorted mutant libraries in bacteria by gene shuffling mScarlet, which has a long fluorescence lifetime, with mCardinal, which has far-red emission (659 nm). Though the investigators did not meet their original goal, their analysis revealed that fluorescence lifetime and peak emission wavelength are inversely correlated. [66] We employed high-throughput microfluidic screening on mOrange2, TagRFP-T and mCherry and found a similar inverse correlation of fluorescence lifetime and emission wavelength. [40] Moreover, an inverse correlation of fluorescence lifetime and photostability under constant irradiation also suggested that increase in fluorescence lifetime leads to molecules bleaching faster, likely due to the higher reductive potential of the  $S_1$  state. This study also showed that the correlation changed with increasing irradiance, suggesting that mechanisms other than photodegradation from the  $S_1$  state may be at play. An inverse correlation of the fluorescence lifetime and photostability was also reported in our previous study of FusionRed variants (Chapter 3). [38] Further investigations are necessary for gleaning mechanistic insights from these observations.



Spectroscopic, computational, and crystallographic evidence suggest that chromophore planarity and protein rigidity are the primary characteristics of bright FPs. [16, 76, 273] Other investigators have considered more subtle physical factors that influence brightness. [62, 82, 125] Using physical models such as the Marcus-Hush theory, these studies focused on quantifying and understanding the role of the driving force in charge transfer between the two resonance forms of the GFP chromophore in the ground and excited electronic states. [62] This driving force is vulnerable to electrostatic control from the environment, which regulates access to non-radiative pathways of excited state depopulation, and therefore directly influences brightness. Analogous studies of the RFP chromophore have not yet been reported. Park and Rhee utilized hybrid QM-MM approaches to demonstrate the influence of electrostatic effects on controlling the non-radiative rate of GFP. [74] Their analysis revealed that electric fields control access to photoisomerization channels which dictate ultrafast non-radiative decay. [74] For RFPs, Drobizhev and co-workers demonstrated that the electric field pointing from the oxygen atom on the imidazolinone ring to the oxygen atom on the phenol ring in the chromophore controls access to twisted-intermolecular charge-transfer (TICT) states which promote ultrafast non-radiative relaxation. [58] In summary, the control of non-radiative decay pathways is a major bottleneck in “brightening” FPs.

Despite its poor brightness, mCherry is a popular RFP for cellular sensing and imaging because of its widespread availability in a plethora of fusion constructs. [274, 275] This monomeric RFP traces its lineage from the naturally occurring tetramer DsRed. [276] The consequent engineering strategies on DsRed resulted in a FP with high expression and fast chromophore maturation, low phototoxicity, and a favorable red-shift of absorption and emission spectra. However, mCherry is much dimmer than its progenitor, primarily due to a three-fold drop in fluorescence quantum yield. Previous attempts to develop brighter and red-shifted versions of mCherry have been unable to restore its fluorescence quantum yield to that of DsRed. [76, 174, 273, 277] In this chapter, we present a monomeric variant denoted mCherry-XL (**eXtended Lifetime**: W143S, I161V, Q163Y and I197R) which is 3-fold brighter than mCherry, and matches the molecular brightness of DsRed. We provide insight into the evolution from mCherry ( $\phi = 0.22$ ;  $\tau = 1.6$  ns) to mCherry-XL ( $\phi = 0.70$ ;  $\tau = 3.9$  ns). We analyzed functional RFPs with the longest lifetimes at two intermediate steps in the evolution trajectory to assess co-evolution of other photophysical properties such as absorption and emission wavelengths. As expected, we observe a near-linear response of



increasing quantum yield with lifetime, along with blue shifts in the absorption and emission maxima. Although the analysis of radiative rates reveals small but significant changes, it is primarily the 6.5-fold reduction in the non-radiative rate that leads to the major increase in fluorescence lifetime. Furthermore, our analysis reveals the reductions in non-radiative rate are not entirely due to the changes in the excited state reorganization energy or blue shifting of the absorption and emission profiles. Our lifetime evolution trajectory agrees with the mechanistic proposals of Drobizhev *et al.* and Lin *et. al.*, who emphasize electrostatic and steric control of the chromophore, to minimize non-radiative processes beyond the scope of energy-gap-type models. [58, 62]

## **5.3. Methods and materials**

### **5.3a. Mutagenesis, cell growth and sorting**

Appendix 4; Sections 1, 2 and 3 describes our protocols for cell growth, sample preparation and sorting. Yeast cells (*Saccharomyces cerevisiae* BY4741) transformed with the pYestDest52 vector containing FPs were used for lifetime-based screening with a microfluidic sorting system developed in our laboratory. [30, 38, 40, 41]

### **5.3b. Protein purification, *in vitro* measurements**

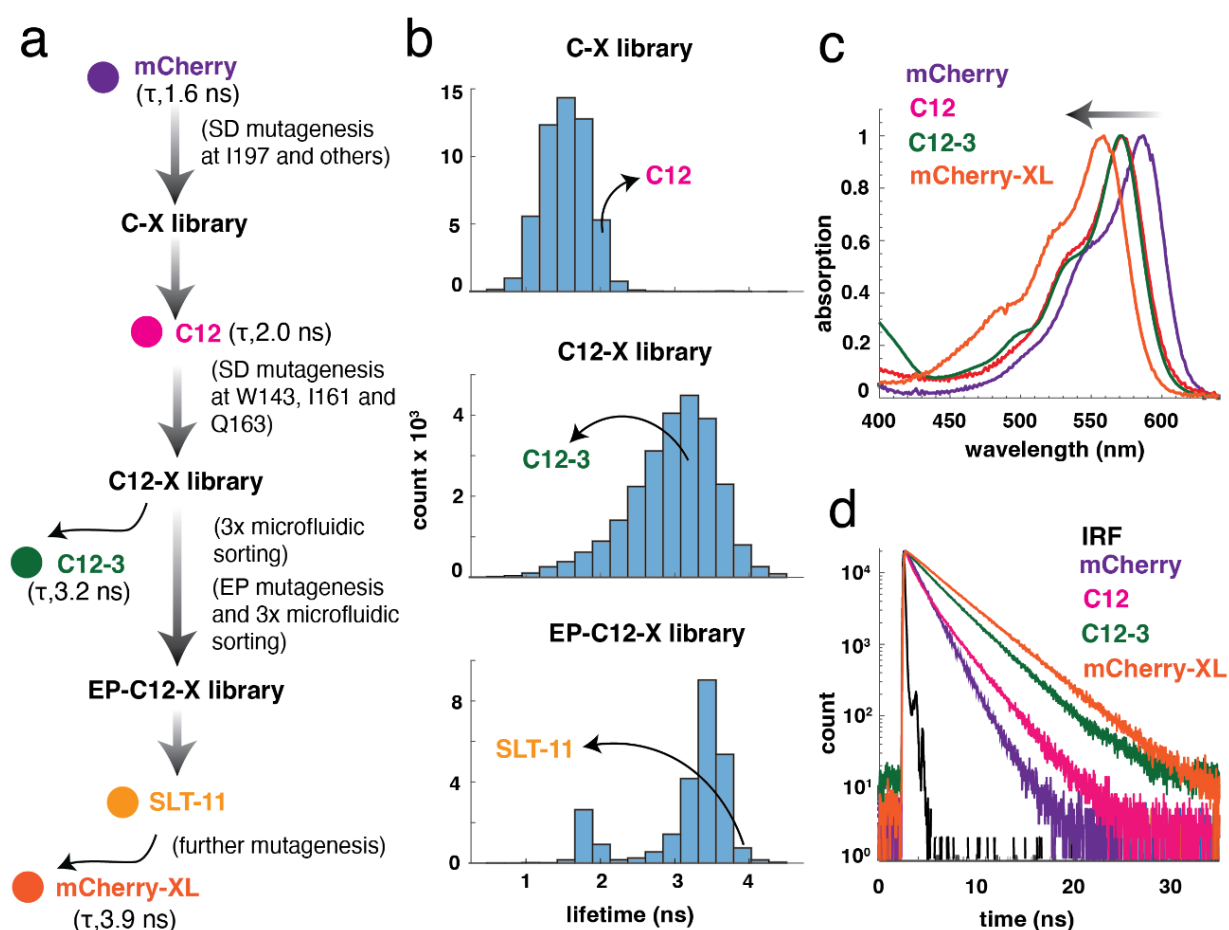
Protocols for *in vitro* measurements have been provided earlier (Chapter 3). [38] In short, DNA of selected FP variants from yeast were cloned into the bacterial p-Bad-His vector and protein was extracted using Ni-NTA-based column chromatography. Steady state data for spectra, quantum yield and extinction coefficient measurements were collected using a Cary 5000 UV–vis near-IR spectrophotometer in the double beam mode for absorption, and a HORIBA Jobin Yvon Fluorolog-3 FL3-222 instrument for fluorescence and excitation spectra. Fluorescence lifetime and anisotropy decay measurements were collected using a commercial time-correlated single photon counting (TCSPC) system (Fluoro-time 100, PicoQuant) with a 560 nm pulsed laser diode head excitation source and a repetition rate of 5 MHz using a spectral filter centered at 600 nm (60 nm FWHM). The methods used in this measurement have been discussed in reference 5 and the details

of the fit and the instrument response function have been provided in the Appendix 4; Section 5. [38] Samples were diluted using 1X-Tris-HCl Buffer (pH ~ 7.4), and measurements were carried out at optical densities (ODs) between 0.05 and 0.25 to ensure measurements fall within the linear regime of the instrument response.

## 5.4. Results

### 5.4a. A 2.5-fold increase in lifetime is achieved using directed evolution strategies on mCherry.

We utilized lifetime-based microfluidic cell sorting to direct the evolution of mCherry in *S. cerevisiae* (Figure 5.1) to longer lifetime. First, we created mutagenesis libraries targeting residue numbers 16, 17, 70, 99, and 197 (numbered with respect to mCherry PDB:2H5Q). The variant C12 (mCherry I197R) was identified from the first round of selection. Next, a site-directed library on variant C12 targeting positions 143, 161 and 163 (C12-X library) resulted in clones with lifetime ranging from 1 ns to 4 ns (Figure 5.1b). A portion of this library was grown on plates and 20 bright variants were selected on lifetime. One of these variants, C12-3 (mCherry W143I, I161C, Q163L and I197R) was selected for further investigation. The C12-X library was then subjected to error-prone mutagenesis followed by another round of lifetime sorting with the sorting gate placed at 3.7 ns (Figure 5.1b). Selection of variants after this round of lifetime enrichment identified the mutant “SLT-11” (mCherry N98S, R125H, F129L, Q137L, W143S, I161V, Q163Y, and I197R) with  $\tau=3.9$  ns. Unfortunately, the SLT-11 variant had poor protein yield in bacteria compared to yeast cells. Therefore, mutations distant from the chromophore ( $>15$  Å) were reverted to those of mCherry (S98N, H125R/K, L129F and L137Q) to boost bacterial expression without compromising the long-lifetime of SLT-11 (Appendix Figure A4.3). This process led to the generation of the quadruple mutant “mCherry-XL” (mCherry W143S, I161V, Q163Y, I197R), which was further characterized. Details of library generation and protocols for mutagenesis have been provided in Appendix 4; Sections 1 and 2. The four FPs (mCherry, C12, C12-3 and mCherry-XL) were characterized *in vitro* and *in cellulo* (Appendix 4; Sections 4 and 5). It was found that mCherry-XL has a 3-fold higher molecular brightness and is 1.5-fold brighter in HeLa cells compared to its progenitor mCherry (Appendix Figure A4.5).



**Figure 5.1. Directed evolution of mCherry resulted in variants with blue-shifts and increased lifetime. (a.)** The lifetime evolution trajectory from mCherry ( $\tau_{av} = 1.6$  ns) to mCherry-XL ( $\tau_{av} = 3.9$  ns), with intermediate clones (shown in filled circles) selected for photophysical analysis (barring the exception of SLT-11). SD and EP mutagenesis refer to site-directed and error-prone mutagenesis respectively. **(b.)** Lifetime histograms of mCherry libraries at the different stages of evolution. The lifetime histograms characterize libraries expressed in yeast cells screened on a microfluidic platform developed in our laboratory. [30, 38, 40, 41] The number of cells screened in each panel were 52382, 24998 and 25617 for C-X, C12-X and EP-C12-X library respectively. **(c & d.)** Absorption and fluorescence lifetime decays of the variants. The arrow in c indicates blue-shift in the absorption of the variants compared to their precursor mCherry.

### 5.4b. Improved brightness is primarily achieved by suppressing the rate of non-radiative relaxation.

We performed *in vitro* photophysical measurements on mCherry, C12, C12-3, and mCherry-XL (Table 5.1). Fluorescence lifetime decays were fit to the appropriate exponential decay functions (Appendix Figure A4.10). Except for C12, a shift from multi-exponential to a mono-exponential decay is observed in this series, where the amplitude and the time constant of the major decay component both increase. The timescale of the shorter component ( $\sim 1$  ns) is approximately the same for all four variants (Table 5.2). Fluorescence anisotropy measurements on these FPs revealed rotational diffusion constants similar to those of monomeric proteins like mCherry, indicating that the mutations did not lead to dimerization or higher order oligomerization (Appendix Figure A4.9).

**Table 5.1.** Spectral and photophysical properties of the mCherry variants. Error bars indicate standard deviations from triplicate measurements. Details of the experimental methods have been presented in Appendix 4; Section 5.

Variant	$\lambda_{\text{abs}}$ (nm)	$E_{\text{abs}}$ ( $\text{cm}^{-1}$ )	$\lambda_{\text{em}}$ (nm)	$E_{\text{em}}$ ( $\text{cm}^{-1}$ )	Stokes shift ( $\text{cm}^{-1}$ )	$\phi$ (%)	Lifetime (ns)	$\epsilon_{\text{max}}$ ( $\text{M}^{-1}\text{cm}^{-1}$ )
mCherry	587	17036	609	16420	615	22 (ref)	$1.67 \pm 0.07$	$76000 \pm 7000$
C12	572	17483	608	16446	1035	$24 \pm 1$	$2.05 \pm 0.05$	$69000 \pm 3400$
C12-3	571	17513	599	16694	819	$40 \pm 1$	$3.15 \pm 0.05$	$59000 \pm 4000$
mCherry-XL	558	17921	589	16978	943	$70 \pm 2$	$3.86 \pm 0.05$	$72000 \pm 4000$

**Table 5.2.** Fit parameters for fluorescence lifetime measurements (decay traces and fits in Appendix 4; Section 5c) along with amino acids at the four positions comprising the sequence variation (full sequences in Appendix 4; Section 3).

Variant	$\tau_1$ (ns)	$a_1$ (%)	$\tau_2$ (ns)	$a_2$ (%)	$\tau_3$ (ns)	$a_3$ (%)	$\tau_{av}$ (ns)	143	161	163	197
mCherry	1.7	83	0.95	17	-	-	1.6	W	I	Q	I
C12	2.6	64	1.1	29	0.3	7	2.0	W	I	Q	R
C12-3	3.4	93	1.1	7	-	-	3.2	I	C	L	R
mCherry-XL	3.9	97	1.1	3	-	-	3.9	S	V	Y	R

The fluorescence quantum yields of these FPs show a clear correlation with the average fluorescence lifetime (Appendix Figure A4.12a), though the trend is not perfectly linear. [38, 250] The two-fold decrease in the half-life of photobleaching with a 2.5-fold increase in fluorescence lifetime also follow expected trends, agreeing with observations of higher photobleaching for FPs with longer fluorescence lifetimes (Appendix Figure A4.6 & A4.7). [38, 250] Additionally, a linear fit ( $R_{adj}^2=0.95$ ) of the lifetime and quantum yield resulted in a slope of  $150 \pm 14 \mu s^{-1}$ . Based on the relation,

$$\phi = k_r * \tau \quad (1)$$

this slope corresponds to the rate constant of radiative decay or  $k_r$ . One can also calculate the individual rate constants for  $k_r$  and non-radiative rate ( $k_{nr}$ ) for these FPs using equation 2. These values are presented in Table 5.3.

$$\phi = \frac{k_r}{(k_r + k_{nr})} \quad (2)$$

Although the 1.3-fold variation in the values of  $k_r$  seems minor, the values lie outside the range predicted by the linear fit (Appendix Figure A4.12b). The value of  $k_r$  can be explained by the Strickler-Berg equation, which relates radiative rate to the peak extinction coefficient and peak fluorescence frequency. [60, 250] The small variation in  $k_r$  is further corroborated by the observation that the peak extinction coefficient for this series only varies 15% from an average value of  $\epsilon_{max \text{ Avg}} \sim 70000 \text{ M}^{-1}\text{cm}^{-1}$ . This variation is modest compared to a 1.5-fold increase of  $\epsilon_{max}$

observed for FusionRed variants reported in our previous study. [38] However, the variation in  $k_r$  can also be attributed to the use of an average lifetime value to represent the multi-exponential fluorescence decay kinetics. The multiple timescales may reflect the presence of multiple chromophore conformations with differing absorption cross-sections. This is bolstered further by the poor linear fit between the cubed value of the emission frequency and the estimated  $k_r$ . This frequency dependence of the radiative rate, predicted by the Strickler-Berg equation, is expected to yield a linear relationship for chromophores with relatively narrow fluorescence spectra (Appendix Figure A4.12c). [60, 61, 250] Most significantly, we found the value of  $k_{nr}$  undergoes the largest change (a 6.5-fold decrease) across this series. A consistent blue-shift in the absorption and emission peak wavelengths with higher brightness, correlated with this decrease in  $k_{nr}$  led us to investigate models that could explain this observation.

#### **5.4c. Lifetime evolution suppresses non-radiative mechanisms beyond the constraints of an energy-gap.**

We first considered an Arrhenius-type dependence of the non-radiative rate constant with the transition energy, as one might expect from an “energy-gap law” (Appendix Figure A4.13). [66, 67, 76] Accordingly, we assumed a single absorbing and emitting species, a lack of excited state photochemistry, and adherence to the mirror image rule of excitation and emission spectra. Under these assumptions, we estimate the 0-0 transition energy gap ( $\Delta E_{00}$ ) for each FP from the intersection point of the absorbance and the fluorescence spectra (Appendix Figure A4.14). However, it is found that the change in non-radiative rate with  $\Delta E_{00}$  is not a perfect fit for the expected exponential dependence. Additionally, we noted the Stokes-shift of mCherry ( $616\text{ cm}^{-1}$ ) is smaller compared to FPs with the I197R substitution ( $\sim 900\text{ cm}^{-1}$ ). This observation encouraged us to examine models that consider the role of the excited state reorganization energy in addition to the energy gap. Englman and Jortner’s treatment of non-radiative rate in terms of the electronic energy gap ( $\Delta E$ ), the excited state reorganization energy and vibrational frequencies provides a broader foundation for understanding nonradiative transition rates than the energy-gap law. [69, 278] Their theory assumes that the electronic transition ( $S_1$  to  $S_0$  relaxation) is coupled to the molecular vibrations and the environmental fluctuations of the bath. The highest-frequency vibrational mode in the excited electronic state serves as the primary path for non-radiative

relaxation based on its coupling strength. The various limiting cases of the theory apply to strong and weak coupling in high and low temperature limits. [69, 278, 279] The strong coupling limit is applicable to systems with a large reorganization energy. In the weak coupling limit, where the role of the energy gap is dominant, the model considers a parameter  $\gamma$  that quantifies the role of the excited state reorganization.

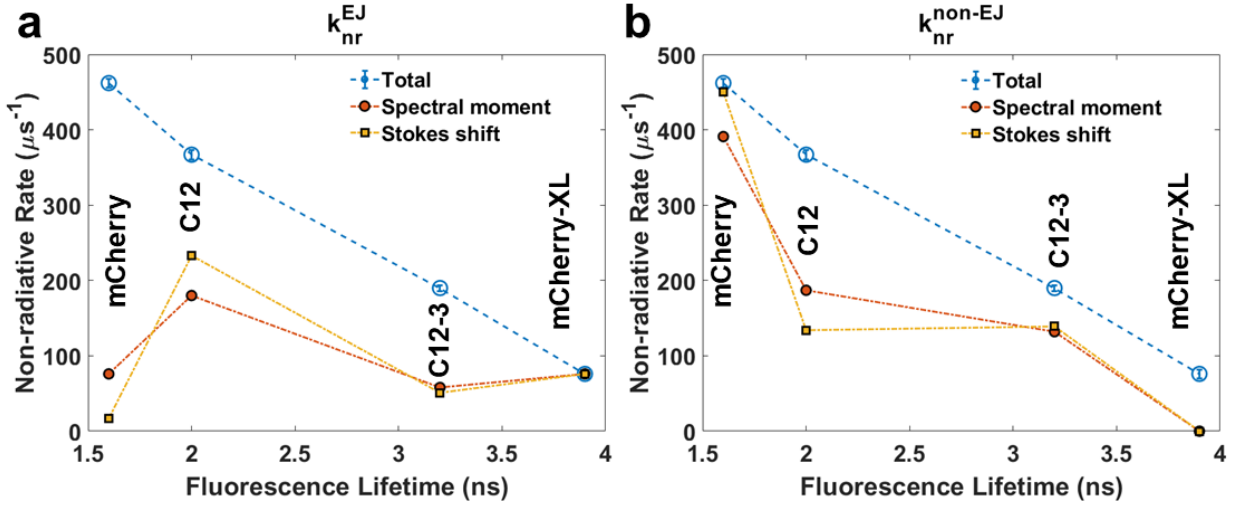
The sum of mode-specific reorganization energies of the chromophore and bath nuclear motions is the reorganization energy ( $\lambda$ ), which is typically assumed to be half of the experimentally-measured Stokes shift ( $\lambda_{SS}$ : Stokes shift). [280] However, this approximation is accurate only for systems with Gaussian absorption and fluorescence spectra. FPs in this study also exhibit a visible 0-1 vibronic sideband along with the primary 0-0 transition, and therefore deviate strongly from Gaussian lineshapes (Figure 5.1c). Thus, we utilized a model provided by Jordanides *et. al* to estimate an upper limit for the reorganization energy from the spectra (Details in Appendix 4; Section 7). [72] The only assumption in this model is that the dielectric medium follows linear response. In this approach, the reorganization energy ( $\lambda_{SM}$ : Spectral Moment) is calculated from the normalized difference of the first moment between the absorption and fluorescence spectra around the transition energy. The values of  $\lambda_{SM}$  for mCherry and mCherry-XL are 812  $\text{cm}^{-1}$  and 924  $\text{cm}^{-1}$ , which are substantially larger than the corresponding values of 308  $\text{cm}^{-1}$  and 472  $\text{cm}^{-1}$  estimated from half the value of Stokes shift.

We employed these values of the reorganization energy in the Englman-Jortner low temperature/weak coupling case, as appropriate for FPs where the 0-0 peak dominates the absorption spectrum. [38, 69, 278, 279]

$$k = \frac{1}{\hbar} \frac{C^2 \sqrt{2\pi}}{\sqrt{\hbar \omega_M \Delta E}} \exp - \left( \frac{\gamma \Delta E}{\hbar \omega_M} \right) \quad (3)$$

Where  $k$  is the rate constant of non-radiative relaxation,  $\omega_M$  is the frequency of the normal mode vibration of the highest frequency in the excited state of the chromophore, the parameter  $\gamma \sim \log\left(\frac{\Delta E}{de_m}\right) - 1$ ;  $d$  is the degeneracy and  $e_m$  is the reorganization energy of the high frequency vibrational mode in the excited state. We then set  $\Delta E = \Delta E_{00}$ ,  $de_m = \lambda$ , and  $\omega_M = \omega_{C-H \text{ Stretch}} = 3000 \text{ cm}^{-1}$ . To estimate the value of Herzberg-Teller coupling ( $C^2$ ), we assume the non-radiative relaxation in mCherry-XL is entirely accounted for by the Englman-Jortner model (Table 5.3;

$k_{nr}^{Total} = k_{nr}^{E-J}$ ). These approximations lead to an estimate for the  $C^2 = 1.45 \times 10^4 \text{ cm}^{-2}$  using  $\lambda_{SM}$  and  $C^2 = 7.19 \times 10^5 \text{ cm}^{-2}$  using  $\lambda_{SS}$ . The minimal ( $\sim 10\%$ ) variation in the peak frequency and intensity of the 0-0 and the 0-1 bands in the absorption spectra for these variants allow us to assume a constant value of  $C^2$  for the other members of the series (Appendix 4; Section 7b). Doing so, we find small variations (of  $\sim 2\%$ ) for the pre-exponent and  $\gamma$  within a range of 2 to 3, which falls within the range predicted by Englman and Jortner for weakly-coupled  $S_1$ - $S_0$  transitions. [38, 69, 278, 279] The results of this analysis are presented in Figure 5.2 and Table 5.3.



**Figure 5.2. The non-radiative rate of the mCherry mutants obtained from the excited state lifetime and fluorescence quantum yield plotted as a function of fluorescence lifetime. (a.)** The non-radiative rate calculated with the low-temperature/weak-coupling limit of Englman-Jortner formalism ( $k_{nr}^{E-J}$ , Eqn. 3) [38, 69, 278, 279] compared to the total non-radiative rate ( $k_{nr}^{Total}$ , Eqn. 1&2) **(b.)** The difference between  $k_{nr}^{Total}$  and  $k_{nr}^{E-J}$  of the variants, denoted as  $k_{nr}^{non-EJ}$  shown in comparison to the  $k_{nr}^{Total}$ . The data in blue circles (with standard deviations) indicate the experimental values of the  $k_{nr}^{Total}$ , whereas the data in gold and red indicate predicted values from the Englman-Jortner relationship using the reorganization energies from the Stokes shift ( $\lambda_{SS}$ ) and the spectral moment methods ( $\lambda_{SM}$ ) respectively. [72]



**Table 5.3.** Calculated rate constants of excited-state population decay.

Variant	$\Delta E_{00}$ (cm <sup>-1</sup> )	$\lambda_{ss}$ (cm <sup>-1</sup> )	$\lambda_{SM}$ (cm <sup>-1</sup> )	$k_r$ ( $\mu s^{-1}$ )	$k_{nr}^{Total}$ ( $\mu s^{-1}$ )	$k_{nr}^{E-J}$ ( $\mu s^{-1}$ )		$k_{nr}^{non E-J}$ ( $\mu s^{-1}$ )	
						Spectral moment	Stokes shift	Spectral moment	Stokes shift
<b>mCherry</b>	16728	308	812	$132 \pm 5$	$467 \pm 6$	76	17	391	450
<b>C12</b>	16965	518	987	$116 \pm 6$	$367 \pm 6$	180	233	187	134
<b>C12-3</b>	17104	409	878	$127 \pm 4$	$190 \pm 4$	58	51	132	139
<b>mCherry-XL</b>	17450	472	924	$179 \pm 6$	$76 \pm 6$	76	76	0	0

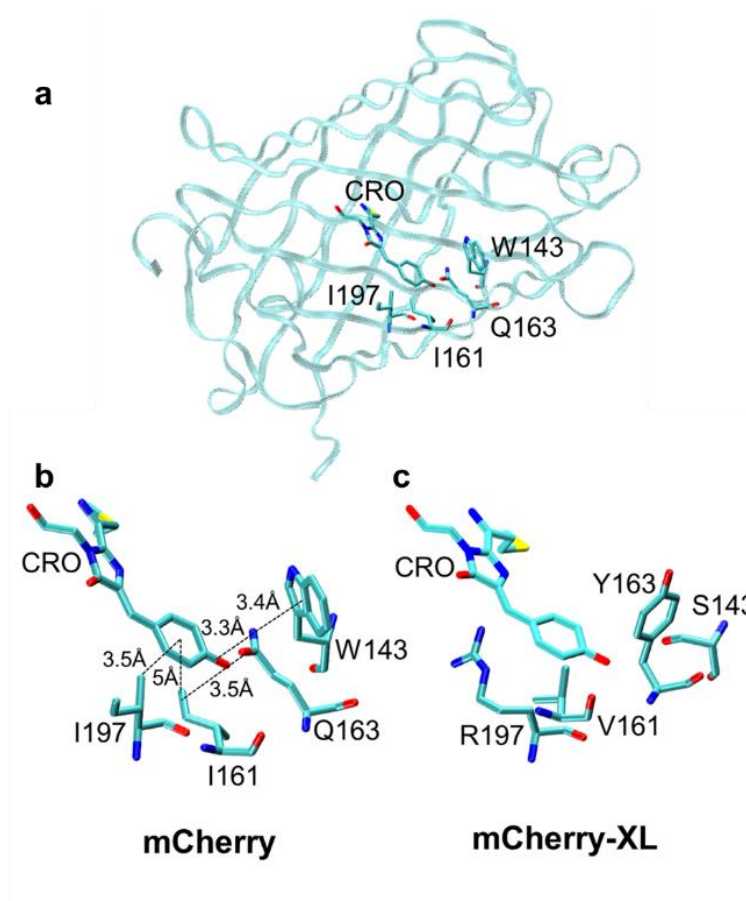
Due to the small variation in the pre-exponential factor, the exponential dependence of the  $k_{nr}$  on the values of  $\Delta E_{00}$  and  $\gamma$  dominate the value of the non-radiative rate in this model. [38, 69, 278, 279] We observe that  $k_{nr}^{E-J}$  differs significantly from the total non-radiative rate ( $k_{nr}^{Total}$ ) for mCherry, C12, and C12-3. As a counter example, if the  $k_{nr}^{Total}$  in mCherry was a consequence of the energy gap and the reorganization energy ( $k_{nr}^{Total} = k_{nr}^{E-J}$ ), this model would have significantly overestimated the  $k_{nr}$  for the brighter variants of this family. Thus, the net positive difference between these two values ( $k_{nr}^{Total} - k_{nr}^{E-J} = k_{nr}^{non E-J}$ ), suggests mechanisms outside the scope of non-radiative relaxation from reorganization and the energy gap might be operational in mCherry and its variants, and this particular trajectory of lifetime evolution successfully suppresses this mechanism.

## 5.5. Discussion

We demonstrated that directed evolution of fluorescence lifetime in mCherry results in a bright fluorescent protein with a high quantum yield. The new “mCherry-XL” RFP is ~1.5-fold brighter than mCherry in mammalian cells, and its lifetime and quantum yield are on par with those of mScarlet, which is the brightest RFP to date. [35] The four mutations identified in this study are located close to the phenol end of the chromophore (Figure 5.3). The rotation of the phenol ring about the methine carbon (P-Bond rotation) is considered to be important for non-radiative decay. [58, 74] The spatial proximity of these mutations to the phenol moiety of the chromophore in

mCherry-XL possibly impacts these motions. For example, substitution I197R was also seen in mScarlet, and the positively charged side chain is an electron withdrawing moiety capable of forming multiple hydrogen bonded contacts at this end of the chromophore. [35] The crystal structure of mScarlet (PDB: 5LK4) shows multiple hydrogen-bonded contacts for the residue R197 which results in a small dihedral angle for the P-bond in comparison to mCherry (2° vs 13°, respectively). Interestingly, all variants of mCherry after incorporating the I197R mutation displayed a higher reorganization energy in comparison to mCherry. The C12 variant (mCherry I197R) also exhibited a tri-exponential fluorescence lifetime decay. The charged nature of the sidechain at position 197 and the possibility of forming multiple hydrogen bonds may result in multiple non-interconverting chromophore conformations as reported for FPs such as TagRFP-675. [15] Additionally, the substitution Q163Y might have a similar role to play in the context of the phenol end of the chromophore. Q163 is a polar residue and the crystal structure of mCherry indicates a weak hydrogen-bond of the amine group of this sidechain (3.3 Å) with the phenolic oxygen on the chromophore (Figure 5.3b). Substitutions such as replacement of Q163 with an electron withdrawing and positively charged K residue have been shown to increase lifetime and quantum yield in DsRed. [76, 277] In a recent report, the substitution W143S was independently and computationally recognized in an effort to develop bright and red-shifted variants of mCherry, which yielded the bright red-shifted RFP, mSandy2 ( $\lambda_{em} = 606$  nm;  $\phi = 0.35$ ). [273] The crystal structure of mCherry reveals an interaction of the indole ring on W143 with the amine group on the Q163 residue (3.4 Å).

The minor but significant variation of the extinction coefficients in these variants is correlated with the small variation of the radiative rates relative to the average value of radiative rate obtained from a linear fit of the quantum yield vs. lifetime (Appendix Figure A4.12). This variation and the observed multi-exponential fits for fluorescence lifetime decays suggests the possibility of multiple emitting chromophore conformations. Structural dynamics slower than the excited state lifetime may lead to conformational diversity of environments that perturb the electronic structure of the chromophore and its radiative rate.



**Figure 5.3. Insights from the crystal structure of mCherry (a.)** The structure of mCherry (PDB ID: 2H5Q). A detailed view of the chromophore (CRO) and the four positions identified in this study for (b.) mCherry and (c.) and in-silico generated structure of mCherry-XL. The dashed lines indicate the distances measured from the mCherry crystal structure. The in-silico structure was generated using the mutator plugin of the VMD molecular modelling software. [28]

The 3.2-fold increase in quantum yield across this evolution route is dominated by the reduction of the non-radiative rate which exhibits a 6.5-fold decrease across the series. The non-radiative rates of the variants are correlated with the blue-shifts of their absorption peaks. A similar correlation of blue shift with increased lifetime was observed by Canty *et. al*, who were motivated to develop a bright red-shifted FP by gene-shuffling the sequences of the far-red emitting mCardinal with the bright red mScarlet. [66] They performed selections on the resulting library based on the brightness of bacterial cultures in the 620 to 750 nm emission window. The selected variants displayed a near-linear correlation of the emission wavelength with the fluorescence lifetime. They proposed an upper limit of 625 nm for the peak emission maximum of variants

created by this method. Furthermore, they concluded that far-red emitting variants are susceptible to a larger variation in spectral profiles than brighter, blue-shifted variants. Although these spectral correlations are intriguing, they did not quantify or investigate the mechanism of non-radiative decay.

The Englman-Jortner model has recently been invoked in discussions of fluorophore design. For example, modifications to this model were used to explain the low brightness of shortwave infrared fluorophores. [38, 69, 278, 279] Their low quantum yields are partially explained by the high non-radiative rates resulting from high-frequency vibrations coupled to small electronic energy gaps. In contrast, when Drobizhev and co-workers plotted the values of the non-radiative rate on the logarithmic scale against the fluorescence emission frequency for a few RFPs (namely mCherry, XRFP, mPlum, DsRed2, eqFP670, and mScarlet), they found a poor agreement with the linear fit expected by Englman and Jortner's model (Appendix Figure A4.13b). [58] This disagreement prompted them to consider population loss through a conical intersection seam from a twisted intermolecular charge transfer (TICT) state. Furthermore, they analyzed the strength and direction of the local electric field from the imidazolinone to the p-hydroxyphenyl ring using classical two-photon excitation spectroscopy and recognized that the local electric field around the FP chromophore controls access to the TICT states and hence can lead to ultrafast relaxation. Their analysis revealed a smaller electric field from the imidazolinone to the phenol ring potentially leads to a higher quantum yield. This can be achieved by reducing electron density at the phenolic hydroxy group through hydrogen bonds and electron-withdrawing contacts or by providing a larger electron density at the imidazolinone. These design principles corroborate findings from a hybrid QM-MM study that revealed similar ultrafast pathways for non-radiative decay in the GFP chromophore. [74] They also agree with the findings of Lin *et. al*, where stabilization of the negative charge on the phenolic oxygen atom is achieved through hydrogen-bonds between T203 and the phenolic oxygen, and/or through R96 on the imidazolinone oxygen for the GFP chromophore. Substitutions at these positions result in spectral shifts and changes in brightness. [62] These observations agree with our results where substitution of a positively charged arginine residue at the phenolic end of the chromophore in C12 variants suppressed non-radiative pathways, resulted in blue-shifts and an increased Stokes shift. Incorporation of other amino acid residues with hydrogen-bond forming side chains, e.g., serine and tyrosine in mCherry-XL resulted in further decrease in the non-radiative rate. Our efforts possibly led to the suppression of relaxation

pathways that contribute to the value of  $k_{nr}^{non\ E-J}$ . While studies by Drobizhev *et. al.* and others provide foundations of these mechanisms for non-radiative pathways in FPs, [58, 62, 74] they did not address the impact of lifetime based directed evolution on the pathways of excited state depopulation.

In this study, we employed the Englman-Jortner theory to model the variation in non-radiative rate for a closely-related RFP series selected through lifetime based directed evolution. [38, 69, 278, 279] We examined whether the reduction in non-radiative rate can directly be related to the energy gap and the excited state reorganization energy. We approximated the 0-0 transition energy gap using the intersection point of the absorption and the emission spectra, and the upper-limit of excited state reorganization energy using a model provided by Jordanides *et. al.* [72] Furthermore, assuming the entire component of the observed non-radiative rate in mCherry-XL was from this formalism, we estimate a  $k_{nr}^{E-J}$  value for the other members of this series. The intriguing outcome of this analysis revealed almost a 6-fold lower  $k_{nr}^{E-J}$  than calculated  $k_{nr}^{Total}$  for mCherry. Though less pronounced, this observation held true for C12 and C12-3, where this model could only estimate half or less for the values of the calculated non-radiative rate. One also arrives at similar qualitative outcomes using the traditional method to estimate reorganization energy using the Stokes shift and repeating this analysis. In such a case the  $k_{nr}^{E-J}$  is >20-fold lower than the  $k_{nr}^{Total}$  for mCherry. This analysis reveals a large component of non-radiative relaxation is outside the realm of the energy-gap and is suppressed using lifetime selections. This result is consistent with the ultrafast non-radiative decay mechanisms predicted by Drobizhev *et. al* and Park *et. al.* [58, 74]

The development of highly emissive fluorophores frequently involves engineering the environment around the chromophore to make it more rigid or more viscous. [281] An issue that arises in connection with this topic is how to probe or quantify molecular rigidity with spectroscopy (i.e. with parameters beyond the ultimately desired high brightness or emission yield). One approach which has been suggested is to regard the Stokes shift, as a measure of flexibility. In principle, the Stokes shift can be calculated from the first moment of the spectral density of modes,  $\rho(\omega)$  coupled to the optical transition. This  $\rho(\omega)$  function, which can be quantified with time-resolved electronic spectroscopies such as photon echo techniques, represents the vibrational density of states of the system weighted by the coupling to the electronic transition.

To the extent that frequency changes in the vibrational density of states reflect changes in flexibility, e.g. a shift towards high frequencies corresponds to higher rigidity, this change will be reflected in the Stokes shift. This approach to quantifying rigidity of a ligand-binding site was demonstrated using three-pulse photon echo peak shift measurements on antibody-antigen complexes, where decreases in flexibility were correlated with higher affinity binding due to specific protein-ligand interactions. [73]

In this view, a decreased Stokes shift can be interpreted to reflect a reduced flexibility of the environment, if the mutations do not strongly perturb the chromophore electronic structure. [282] This idea is central to the weak-coupling case of the Englman-Jortner theory which predicts a decreased non-radiative rate with the decrease in reorganization energy. [38, 69, 278, 279] Unfortunately this straightforward picture of rigidifying the protein environment around the chromophore may be complicated by the strong interactions of the environment that perturb the chromophore electronic structure. For example, the mutation I197R in mCherry increases the Stokes shift despite a decrease in non-radiative rate. In contrast, an evaluation of the electro-optical properties of the GFP chromophore revealed that blue-shifted absorbers have larger Stokes-shift values. This is in-part due to a larger driving force between the two resonance forms of the chromophore. [62] Another striking counter-example is provided by AsRed2, which has only a ~5% fluorescence quantum yield, despite having a very small Stokes shift ( $\sim 470\text{ cm}^{-1}$ ) and absorption and emission peak wavelengths similar to those of several bright RFPs ( $\lambda_{\text{abs}} = 572\text{ nm}$  and  $\lambda_{\text{em}} = 592\text{ nm}$ ). [283] Although the low apparent brightness of this FP might be attributed to an inaccurate measurement of quantum yield owing to the kindling qualities of its parent asFP595, closer examination of other FPs with small quantum yield and modest Stokes-shifts is warranted.

It is also interesting that other examples of fluorescence lifetime evolution do not show the same spectral trends observed for mCherry-XL. For example, consider the evolution trajectory in the development of FusionRed-MQV (FR-MQV). First, a small but significant ( $70\text{ cm}^{-1}$ ) decrease in Stokes shift resulted from the M42Q mutation in FusionRed. [38] Subsequently, the FR-Q, FR-MQ and FR-MQV variants had similar values of the Stokes shift ( $\sim 570\text{ cm}^{-1}$ ) although consistent decreases in the non-radiative rate (ultimately resulting in a 60% overall decrease) were observed with increased lifetime. Another example is provided by mScarlet and mScarlet-I, which have nearly identical 0-0 transition energies and nearly the same Stokes shift, yet the latter has a two-

fold higher non-radiative rate. [35, 38] Consequently, strategies beyond rigidifying the chromophore are necessary for the design of brighter FPs. [273]

These considerations suggest that large variations in absorption and emission wavelengths are possible for anionic RFP chromophores with similar values of lifetime and quantum yield, because the non-radiative rate does not depend exclusively on the transition energy gap. The RFP mScarlet is red-shifted by 10 nm (or  $315\text{ cm}^{-1}$ ) in maximum absorption and by 6 nm (or  $272\text{ cm}^{-1}$ ) in maximum emission compared to mCherry-XL, yet it matches its quantum yield and fluorescence lifetime values. [35] The two FPs also exhibit nearly identical single exponential fluorescence lifetime decays measured using TCSPC. While the 0-0 transition energy is red-shifted by  $200\text{ cm}^{-1}$ , the Stokes shift for mScarlet is  $280\text{ cm}^{-1}$  smaller than that of mCherry-XL. This suggests the possibility of either red shifting mCherry-XL by reducing the Stokes-shift (while maintaining its quantum yield) or increasing the quantum yield of mScarlet by reducing its red shift. These properties are observed in AusFP1, which has a 2.7-fold smaller Stokes shift and a 1.6-fold higher quantum yield in comparison to EGFP, with the same tripeptide responsible for chromophore formation. [36] Such observations suggest the possibility of designing FPs with higher brightness without compromising on the red-shifted absorption and emission as theoretically predicted by the work of Moron *et al.* [92]

## 5.6. Conclusions

The race to develop bright and red-shifted fluorophores continues to drive FP engineering to newer technologies and selection schemes. Lifetime based evolution is now an established approach for delivering bright FPs such as mScarlet, mTurquoise2 and FusionRed-MQV. In this study, lifetime evolution provides a bright FP in the form of mCherry-XL and reveals the consequent co-evolution of spectral shifts, increased quantum yield, subtle changes in the radiative rate constants and very different mechanisms of suppressing non-radiative pathways. Our results strongly support recent advances by the community to understand the pathways of non-radiative depopulation of FP chromophores that lie outside the description of the energy gap. Moreover mCherry, C12, C12-3 and mCherry-XL have average fluorescence lifetimes of 1.6, 2.0, 3.2 and 3.9 ns respectively with minimal green-absorbing or emitting species, (Appendix Figure A4.8) which makes them suitable

for multi-color FLIM based imaging. Additionally, we do see promise in using mCherry-XL as a template for future engineering and a probe for imaging. Four amino acid substitutions on the mCherry sequence can offer solutions in the form of mCherry-XL for constructs and applications that are limited by the brightness of mCherry. Moreover, a blue-shifted absorption (~28 nm) for mCherry-XL might improve performance relative to mCherry for many FRET related applications where poor spectral overlap is observed between a green donor and a red acceptor pair. This study also shows the strength of microfluidics-based lifetime selections to enrich populations with longer lifetimes, as illustrated by the discovery of the SLT-11 variant in the EP-C12-X library. To summarize, the end-product of our selections drove us to a blue shifted variant mCherry-XL with a specific evolution trajectory based on lifetime. In a broad fitness landscape of mutations, our analysis reveals that there can be other evolution trajectories (such as the discarded clones with poor brightness or maturation) with similar end products in terms of lifetime but with different absorption and emission profiles.



## Chapter 6

### Brightening FusionRed-MQV using lifetime selections

#### 6.1. Introduction

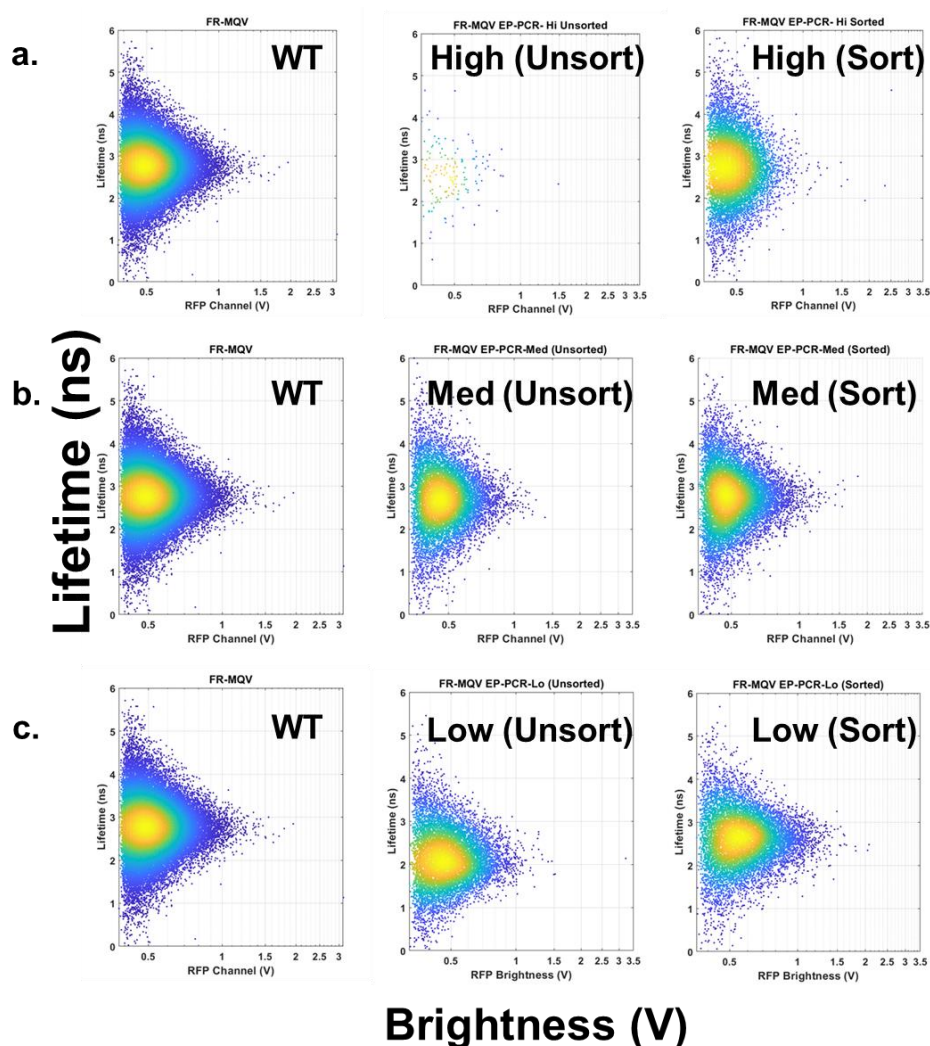
Evolution of brighter FusionRed mutants through engineering of excited state lifetime was discussed in Chapter 3. Several rounds of random, site directed and structure-based mutagenesis strategies on FusionRed generated variants that displayed higher *in vitro* brightness, low oligomerization tendency and high fusion efficiencies. Interesting outcomes such as the 1.5-fold increase of extinction coefficient from the M42Q substitution, drastic changes to the propensity of dark state conversion from the mutation C159V and a 2-fold increase in cellular brightness from the L175M mutation were observed for the FR variants generated in this effort. However, the brightness in mammalian cell lines of many such mutants remained sub-optimal, which we hypothesize is due to their slower chromophore maturation or lower levels of protein expression. Libraries aimed at tackling the protein expression, folding and chromophore maturation in mammalian cells would therefore expand the utility of such FR variants in cellular imaging. Additionally, mutations at positions close to C159 can serve as templates for future library generation efforts focused on improving the photostability or photoswitching properties of FusionRed variants like FR-MQV.

Besides these attractive motivations to generate mutant libraries of FR, the presence of FPs such as mScarlet and mCherry-XL motivate further investigations into the possibility of improving the brightness of FusionRed. This goal can potentially be achieved by further suppression of non-radiative pathways in FusionRed-MQV through mutagenesis and lifetime selections. Therefore, two independent approaches were taken to explore the possibility of further increasing the lifetime and quantum yield in FusionRed-MQV. In this chapter, we discuss results from error-prone PCR on FusionRed-MQV performed at varying error rates and further site directed saturation mutagenesis libraries on positions 71, 159, 175 and 224 to generate variants of FusionRed that can match the quantum yield and lifetime of mScarlet and mCherry-XL.

## 6.2. Results and discussion

### 6.2a. Error-Prone PCR on FR-MQV

In our first approach, EP-PCR libraries of FusionRed-MQV were generated using the GeneMorph II system (as described in Chapter 3 and Appendix 4) and FACS enriched. EP-PCR mutagenesis was performed at low (~2 mutations/kb; at the nucleotide level), medium (~4.5 mutations/kb; at the nucleotide level) and high error rates (~16 mutations/kb; at the nucleotide level) to see a possible diversity of brightness and lifetime. (Figure 6.1) The FACS enriched libraries were screened on the droplet microfluidic device [41] at single-cell resolution.



**Figure 6.1. Lifetime-brightness screens of the FusionRed-MQV (WT) before and after FACS enrichment populations, for (a.) the high-error rate EP-PCR library, (b.) the medium error-rate EP-PCR library and (c.) the low error-rate EP-PCR library. The FACS enrichment greatly enriched**

fluorescent population in the high-error rate EP-PCR library, which shows a 1.2-fold greater variation (See Table 6.1) in lifetime compared to the wild-type encouraging future investigations and lifetime directed sorting. Approximately 100,000 cells were screened for each library.

**Table 6.1. Mean brightness and lifetime for the FusionRed-MQV random mutagenesis libraries screened in Figure 6.3.** The fluorescence lifetime measured on the sorter is in frequency domain represented by  $\tau_{FD}$ .

Library (FACS Enriched)	Mean Brightness (V)	$\tau_{FD}$ (ns)
Lo	0.61 $\pm$ 0.11	2.60 $\pm$ 0.6
Med	0.53 $\pm$ 0.18	2.72 $\pm$ 0.7
Hi	0.51 $\pm$ 0.11	2.73 $\pm$ 0.9
FR-MQV	0.55 $\pm$ 0.11	2.80 $\pm$ 0.4

Despite the low variation in the mean brightness and lifetime of these populations, the consistently higher spread of fluorescence lifetime is an encouraging indication of variants that outperform FR-MQV with longer lifetimes and potentially higher quantum yields. In particular, a 1.25-fold larger variation of lifetime is observed for the high-error rate EP-PCR library (Table 6.1).

## 6.2b. Site-directed saturation mutagenesis

Positions identified with random mutagenesis in the directed evolution efforts detailed in Chapters 3 and 5, such as L175M, can be further saturated to identify better steric or electronic fits (electrostatics or hydrogen bonding), thus further increase lifetime and brightness. [30] Therefore, we employed saturation mutagenesis on some of these positions, to generate small libraries that can be screened with respect to lifetime on plates. Future evolution strategies can be formulated on the basis of these preliminary investigations.

### FusionRed-L175X Library

FusionRed L175M was previously recognized in [30], where we provided a proof-of concept for the directed evolution of fluorescence lifetime. FusionRed-L175M was generated from an error-prone mutagenesis library on FusionRed, which yielded the clone FusionRed-1 (FusionRed; V19I/L52P/L175M numbered based on FusionRed crystal structure; PDB ID: 6U1A). The

mutation L175M in particular was beneficial to the *in vitro* (1.2-fold higher) and *in-cellulo* (2-fold higher) brightness for FusionRed. However, the position L175 was not saturated with all 20 amino acid residues in that work. When this substitution was mapped on FusionRed-M42Q, which was designed with guidance from crystal structure data and ultrafast spectroscopy experiments – it carried over its favorable properties of higher brightness to the clone FusionRed-L175M M42Q. [38] Despite saturating M42, we did not attempt to saturate L175M in that study, in an attempt to retain the favorable biological properties carried over by the L175 substitution. Consequently, this position was targeted with all 20 amino acids to generated variants with, for example, the positively charged histidine, which led to a higher value of lifetime (2.35 ns) and quantum yield (40%). The variants identified in this library, their fluorescence lifetime and spectral details are provided below (Tables 6.2 and 6.3).

**Table 6.2. Lifetime measurements for the FusionRed-L175X clones.** The fluorescence lifetime measured in bacteria on the sorter in frequency domain are represented by  $\tau_{FD}$  and *in-vitro* lifetime measurements using TCSPC in the time domain are represented by  $\tau_{TD}$ .

Protein	$\tau_{FD}$ (ns)	$\tau_{TD}$ (ns)	$\tau_1$ (%)	$\tau_2$ (%)	$\tau_3$ (%)
FR	2.03	1.75	2.55, 30	1.47, 70	-
FR-L175R	2.06	1.83	-	1.99, 84	0.99, 16
FR-L175S	1.43	1.43	2.79, 15	1.29, 75	0.42, 10
FR-L175H	2.35	2.41	3.15, 50	1.83, 50	-
FR-L175M	2.15	2.13	2.54, 47	1.70, 47	-

**Table 6.3. Spectral details for the FusionRed-L175X clones**

Protein	$\lambda_{abs}$ (nm)	$\lambda_{em}$ (nm)	Stokes Shift (cm <sup>-1</sup> )
FR (WT)	575	596	613
FR- L175R	576	601	722
FR-L175S	572	596	704
FR-L175H	568	588	599
FR-L175M	571	591	593

### FusionRed-Q L175X<sub>1</sub> C159X<sub>2</sub> Library

Site-directed mutagenesis on C159, M161, V196, and H198 in FusionRed led to the identification of the bright FusionRed-C159V ( $\tau=1.96$  ns and  $\phi=31\%$ ) variant. [38] As discussed in Chapter 3, this mutation was then mapped onto the bright FusionRed-MQ ( $\tau=2.43$  ns and  $\phi=43\%$ ), resulting in the bright variant FusionRed-MQV ( $\tau=2.77$  ns and  $\phi=53\%$ ). [38] Unlike the position M42, the C159 position also had not been saturated in our previous effort. Therefore, to build on that study and our library on FusionRed-L175X, which yielded the brighter FusionRed-H (L175H), we targeted FusionRed-Q with a library targeting the positions L175 (with H/M) and position C159 (with all 20 amino acids). The clones were plate screened for lifetime and we identified that the variant FusionRed L175H, M42Q and C159 (or FusionRed-HQC), which had a higher lifetime and quantum yield (Table 6.5) in comparison to FusionRed-MQV. The variants identified in this library, their fluorescence lifetime and photophysical details are provided below.

**Table 6.4. Lifetime measurements for the FusionRed-M42Q L175X<sub>1</sub> C159X<sub>2</sub> clones.** The fluorescence lifetime measured in bacteria on the sorter in frequency domain are represented by  $\tau_{FD}$  and *in-vitro* lifetime measurements using TCSPC in the time domain are represented by  $\tau_{TD}$ .

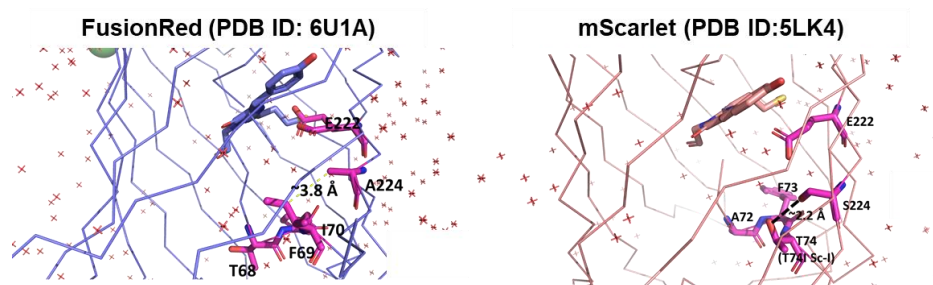
Protein	$\tau_{FD}$ (ns)	$\tau_{TD}$ (ns)	$\tau_1$ (%)	$\tau_2$ (%)	$\tau_3$ (%)
FR-MQV	2.7	2.79	3.19 (65)	1.69 (35)	
FR-MQG	2.3	2.43	3.41 (27)	2.19 (69)	0.59 (4)
FR-MQC (FR-MQ)	2.4	2.36	3.30 (86)	1.72 (14)	
FR-HQV	2.2	2.23	3.39 (28)	1.1 (64)	0.72 (8)
FR-HQC (or FR-HQ)	2.9	3.09	3.45 (88)	1.66 (12)	
FR-HQG	-	1.97	3.08 (35)	1.49 (57)	0.57 (8)
FR-HQH	2.6	2.71	3.13 (72)	1.66 (28)	
FR-HQD	1.9	1.64	3.47 (40)	1.77 (44)	0.65 (16)

**Table 6.5. Photophysical characterization for the FusionRed-M42Q L175X<sub>1</sub> C159X<sub>2</sub> clones**

Protein	$\lambda_{\text{abs}}$ (nm)	$\lambda_{\text{em}}$ (nm)	Stokes Shift (cm <sup>-1</sup> )	$\phi$ (%)	$\tau_{\text{av}}$ (ns)	$\epsilon_{\text{max}}$ (M <sup>-1</sup> cm <sup>-1</sup> )
FR-MQV	566	585	574	53	2.8	140000
FR-HQC	565	583	546	59	3.1	127000
FR-HQH	568	586	540	52	2.7	114000

### FusionRed-MQV/HQ A224X<sub>1</sub> and I71X<sub>2</sub> Library

During the development of FusionRed-M, an alternate evolution pathway for extended lifetime was discovered through the variant FR-13 (H25Y, V49I, F83Y and A224T). [30] Analysis of this variant revealed that the mutation A224T on FusionRed led to a 1.8-fold increase in fluorescence lifetime and a 2-fold increase in quantum yield. This route was not pursued as the incorporation of the A224T mutation resulted in poor maturation for FPs with this substitution. Additionally, the role of this position was also discussed in the effort to develop the bright mScarlet RFP (A218 - numbered with respect to the crystal structure of mScarlet, PDB ID:5LK4). [39] In that study, Bindels and Haarbosch mutated the position T74I on the  $\alpha$ -Helix in mScarlet to generate the fast maturing variant mScarlet-I. Exploring the crystal structure of these proteins indicated a potential interaction between the two residues A224 and I71 in FusionRed and mScarlet (Figure 6.2). Therefore, we saturated the positions A224 and I71 on FusionRed-MQV and FusionRed-HQC with all 20 amino acids to generate a library of ~400 variants to be selected on lifetime. (Table 6.6) Lifetime based screening led to the identification of variants such as FusionRed-HQ A224T and FusionRed-HQ A224T I71T, which exhibited high values of quantum yield (~75%) and long fluorescence lifetimes (~4 ns). The variants identified in this library and their fluorescence lifetimes are provided below (Table 6.6).



**Figure 6.2. The crystal structures indicating the potential interactions of positions 224 and 71 - (numbered with respect to FusionRed) in FusionRed (left) and mScarlet (right).**

**Table 6.6. Lifetime screens for the FusionRed-MQV/HQC A224X<sub>1</sub> I71X<sub>2</sub> clones.** The lifetime measured in bacteria on the sorter in frequency domain are represented by  $\tau_{FD}$ .

Clone Name	71	224	$\tau_{FD}$ (ns)
FR	I	A	1.8
FR-MQV (WT)	I	A	2.8
FR-MQV-IQ	I	<b>Q</b>	3.2
FR-MQV-IG	I	<b>G</b>	2.7
FR-MQV-IV	I	<b>V</b>	3.6
FR-MQV-IL	I	<b>L</b>	2.5
FR-MQV-VA	<b>V</b>	A	3.3
FR-MQV-TA	<b>T</b>	A	3.2
FR-MQV-AA	<b>A</b>	A	3.5
<b>FR-MQV-MA</b>	<b>M</b>	A	2.9
FR-HQC (WT)	I	A	3.2
FR-HQC -IT	I	<b>T</b>	3.5
<b>FR-HQC-TT</b>	<b>T</b>	<b>T</b>	<b>3.8</b>
<b>FR-HQC-VA</b>	<b>V</b>	A	3.4
FR-HQC-LA	<b>L</b>	A	3.5
FR-HQC-CA	<b>C</b>	A	3.5
FR-HQC-AA	<b>A</b>	A	3.5

**Table 6.7. Spectral properties for the high lifetime FusionRed-MQV/HQC A224X<sub>1</sub> I71X<sub>2</sub> clones.**

Protein	$\tau_{TD}$ (ns)	$\phi$ (%)	$\epsilon_{max}$ (M <sup>-1</sup> cm <sup>-1</sup> )	Brightness	$k_{rad}$ ( $\mu$ s <sup>-1</sup> )	$k_{non-rad}$ ( $\mu$ s <sup>-1</sup> )
FR	1.78	24	94000	100	135	427
FR-MQV	2.78±0.06	53 (ref)	140800±9700	330	191	169
FR-MQV-VA	3.27±0.04	64±1	96100±6500	273	196	110
FR-HQC	3.17±0.10	60±1	118500±10500	316	189	126

FR-HQC-T	3.50±0.01	68±2	65000	196	194	91
FR-HQC-VA	3.36±0.05	66±3	123200±1500	360	196	101
FR-HQC-TT	3.74±0.04	75±4	65600±14000	218	200	67
mCherry	1.67±0.07	22 (Ref)	75600±5000	75	137	488
mCherryXL	3.86±0.05	70±2	72000±4000	223	179	76
mScarlet	3.87±0.07	71 (Ref)	104000±4000	327	186	72

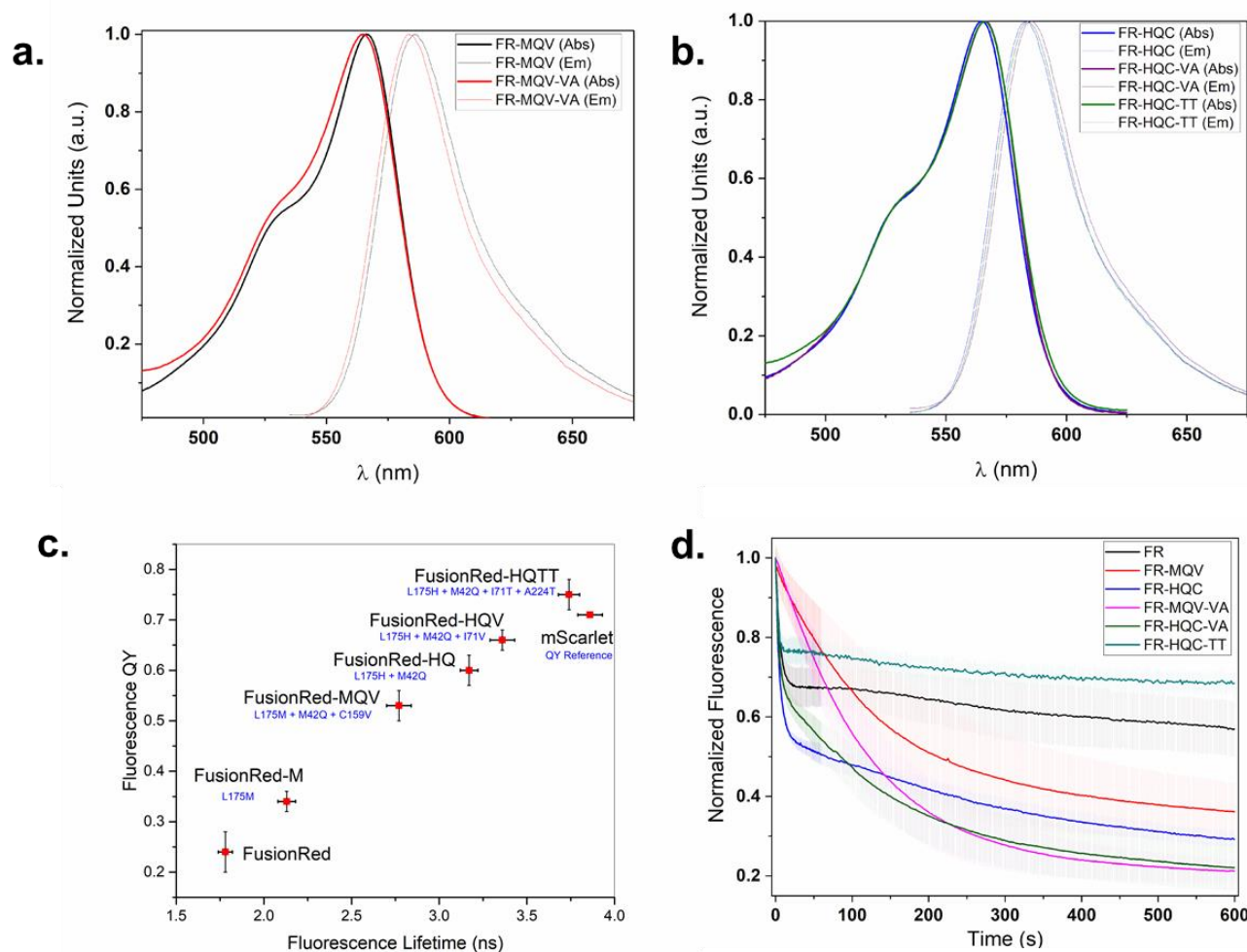
These clones also revealed an interesting observation about chromophore maturation in bacteria. Like the FR-13 variant, FPs that incorporated an A224T substitution suffered from poor maturation. Whenever I71 and A224 were aliphatic residues (G, A, I, V and L) – higher fluorescence was observed in cells. FPs such as FR-MQV-VA and FR-HQC-VA displayed higher lifetime than the parental FPs (FR-MQV/HQC) and exhibited fast maturation, unlike clones such as FR-HQC-TT. A summary of photophysical characterization of relevant FPs with highest lifetimes selected from this library are summarized in Figure 6.3, Tables 6.7 and 6.8. As expected, variants with high quantum yield and lifetime display a consistent drop in the non-radiative rate, as seen for the mCherry variants presented in Chapter 5.

**Table 6.8. Spectral properties for the FusionRed-MQV/HQC A224X<sub>1</sub> I71X<sub>2</sub> clones.**

RFP	$\lambda_{\text{abs}}$ (nm)	FWHM (Abs)	$\lambda_{\text{em}}$ (nm)	FWHM (Em)	$\lambda_{\text{Stokes}}$ (nm)	Stokes shift (cm <sup>-1</sup> )
FR	574	63	596	54	22	645
FR-MQV	566	55	585	41	19	575
FR-MQV-VA	565	57	582	40	17	515
FR-HQC	565	55	583	40	18	550
FR-HQC-T	566	54	583	39	17	515
FR-HQC-VA	566	55	585	39	19	575
FR-HQC-TT	566	57	583	38	17	515



We also performed excitation-normalized photobleaching measurements (with protocols described in Chapter 3) in bacteria and observed that FPs derived from the FR-MQV lineage clearly show a mono-exponential bleaching behavior. These variants retain the mono-exponential bleaching trend by restricting the chromophore to switch from a fluorescent cis to a dark trans isomer. [38] However, FPs from the FR-HQC family show a bi-exponential behavior, characteristic of the wild-type FusionRed. [38] The time constants for the mono and bi-exponential fits are collected in Table 6.9, where one can observe very different values for the longer component of decay, which we have previously assigned as the time constant of permanent photobleaching ( $t_{BL}$ ). Within FR-MQV and FR-MQV-VA, the  $t_{BL}$  drops by 1.2-fold with a 1.2-fold increase in lifetime, which is characteristic of a reduction in photostability with an increase in fluorescence lifetime. These data suggest that the dark state might be photo-protective for FPs with bi-exponential bleaching kinetics because clones of the FR-HQC family with similar lifetimes as those of the FR-MQV family show higher photostability. Additionally, it appears that FR-HQC-TT doesn't follow the trend of increasing lifetime correlating with reduced photostability. This discrepancy is likely caused by an inaccurate photostability measurement possibly due to the high background noise seen due to the poor expression/maturation in bacteria.

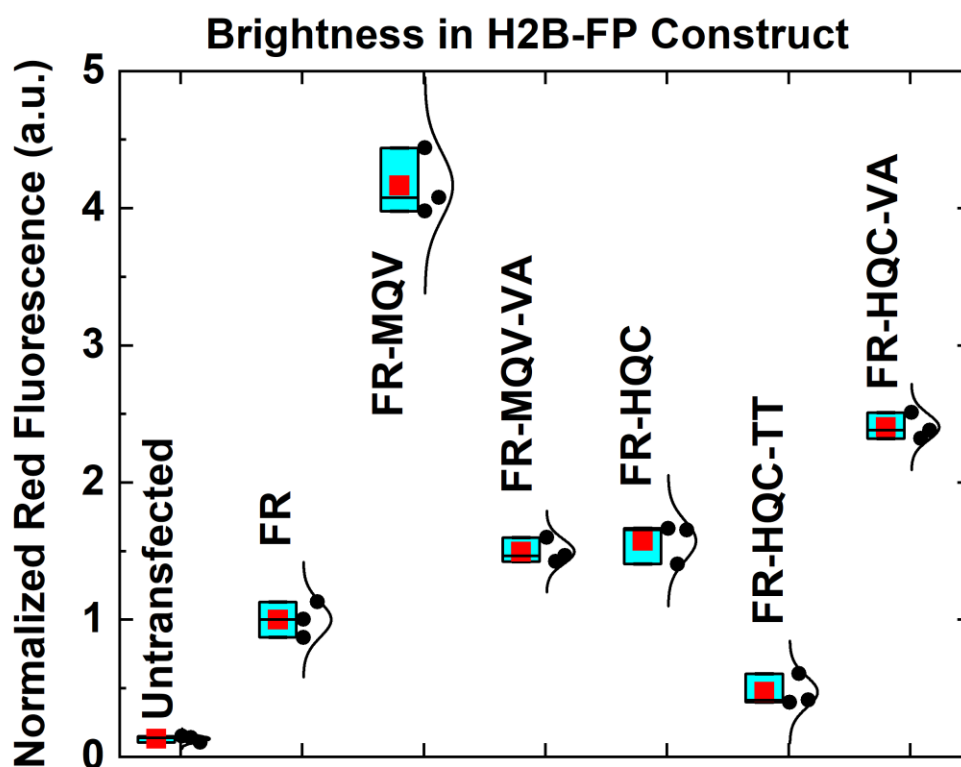


**Figure 6.3. Photophysical characterization of the brightest clones from site-directed mutagenesis of FR-MQV and FR-HQC.** (a.) The absorption and emission profiles of FR-MQV variants (b.) The absorption and emission profiles of FR-HQC variants. (c.) The linearity of the quantum yield and lifetime of the FusionRed family in comparison to mScarlet. (d.) The photobleaching traces of the long lifetime variants from FusionRed-MQV and FusionRed-HQC under excitation normalized conditions (Irradiance  $\sim 5 \text{ W/cm}^2$ ) with a 560 nm widefield excitation source.

**Table 6.9. Photobleaching behavior for the high lifetime FusionRed-MQV/HQC A224X<sub>1</sub> I71X<sub>2</sub> clones.**  $t_D$  and  $t_{BL}$  denote the time constants of reversible and irreversible photobleaching obtained from fitting a photobleaching curve to exponential decay functions.

RFP	Dark State	$t_D$ (s)	$t_{BL}$ (s)
FR	Yes	5.01±0.05	1445±169
FR-MQV	No	-	142±5
FR-HQC	Yes	5.12±0.03	347±3
FR-MQV-VA	No	-	121±1
FR-HQC-VA	Yes	3.84±0.04	163±1
FR-HQC-TT	Yes	2.41±0.04	299±8

We cloned these FPs into the piggyBac-H2B construct for expression in HeLa cells to test mammalian cell expression using FACS. Despite an “apparent” increase in maturation/expression for variants like FR-MQV-VA and FR-HQC-VA in bacterial cultures, these variants were ~2-fold dimmer than the previously generated FR-MQV in mammalian cells. However, most of these variants were brighter than the wild type FusionRed. As expected, the FR-HQC-TT clone underperformed significantly, quite possibly due to the A224T mutation. The results of this assay are presented in Figure 6.4.



**Figure 6.4. The brightness of FusionRed variants generated from FR-MQV and FR-HQC in mammalian cells using FACS.** FR-MQV still outperforms brighter FP molecules like FR-HQC variants and FR-MQV-VA in mammalian cells despite lower values of quantum yield. The data was collected on a BD FACSCelesta single cell analyzer with a TRITC filter set (585/30). Single-cell brightness was assessed by selecting single healthy cells based on forward and side-scattering photon counts after 48 h of transfection. Untransfected cells were used as a control to background subtract and analyze the fluorescence in the red channel for the proteins of interest. The raw counts were normalized with respect to the average number of counts for FusionRed. Each technical replicate (represented as black dots) consisted of a screen of ~20000 cells. The boxplots indicate the distributions of values: the mean of the distribution is indicated by red solid squares; the blue box indicates the 25th and 75th percentiles of the distribution. The solid line in the box indicates the median value of the distribution.

## 6.3. Conclusions

Here, we attempted to utilize random mutagenesis on FusionRed-MQV at varying error rates in bacteria to explore the mutational landscape of this bright RFP. The larger diversity of lifetime identified from the EP-PCR library screens and sorts revealed that further rounds of error-prone or targeted mutagenesis can perhaps lead to brighter versions of FusionRed-MQV. As a consequence, we attempted a targeted mutagenesis strategy that saturates positions identified in previous random EP-PCR libraries on FusionRed. For example, we learned, reverting V159C in FR-HQV unlike FR-MQV, results in a 20% increase in fluorescence lifetime. Moreover, libraries generated with the saturated mutagenesis strategies here were small (20-400 unique variants), in comparison to bigger libraries generated in Chapters 3 and 5 (>1000). These libraries were screened and selected from bacterial colonies on plates. This approach is tedious in comparison to single cell, high-throughput microfluidic screening. However, the fluorescence signal from individual colonies is larger, which results in better resolution of fluorescence lifetime, thus allowing for the identification of dim clones with long lifetimes such as FusionRed-HQC-TT. At single cell resolution, the low signal to noise from clones that show poor expression or maturation in cells can lead to unnecessary rejections even for FPs that display high molecular brightness. [62] These clones are important resources to understand the physical processes that dictate fluorescence. Detailed characterization of the physico-chemical environments of such variants provide excellent opportunities to foresee principles or design strategies that can lead to the development of brighter fluorescent tools. Such strategies can be applied outside the realms of fluorescent proteins from the GFP superfamily. [23] Biochemical issues pertaining to the expression or maturation can be tackled independently, through efforts targeted to improve solubility, expression or maturation. [36]

These efforts confirm that lifetime selections consistently reduce the non-radiative rate and remain the primary pathway for brightening an FP by increasing the fluorescence quantum yield. However, unlike the mCherry family, the role of reorganization energy and energy gap in these variants seems minimal. The absorption and emission wavelengths for the FR-MQV and FR-HQC variants generated in this study vary by <5 nm, and the Stokes shift by <10%. As described in Chapter 5, the Stokes shift (or the reorganization energy) can be treated as a measure of

chromophore flexibility or the environmental rigidity, provided that the electronic structure of the chromophore is not dramatically perturbed. A ~2-fold drop in the non-radiative rate is seen for bright variants like FR-HQC-VA in comparison to FR-MQV, yet the absence of a consequent response on the spectral parameter like the Stokes shift inherently becomes a problem of interest. Additionally, we again see the saturation of the fluorescence quantum yield at the ~75% mark for the FusionRed family. The brightest RFP in literature mScarlet and the brightest variant from the mCherry family, mCherry-XL also saturate around the same values of lifetime and quantum yield. Such observations can suggest the potential saturation of fluorescence lifetime and quantum yield around a theoretical limit, which warrants independent and necessary investigations.

# Chapter 7

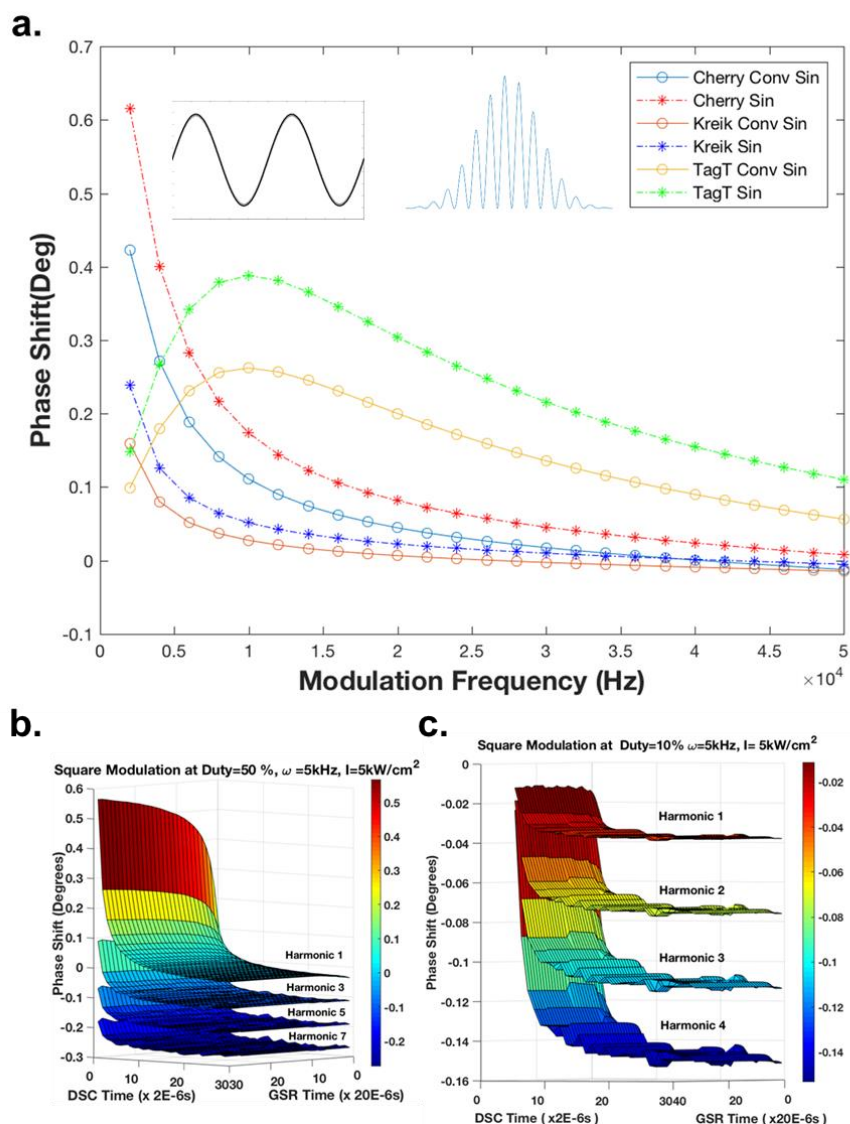
## Future Directions

### 7.1. Development of a sorting platform on dark state conversion

As demonstrated in Chapter 2, our multiparametric high-throughput sorter system is capable of screening and sorting analytes on emission spectra, brightness and fluorescence lifetime. Chapter 3 and 4 discussed the role of dark state conversion on the ensemble and single molecule photophysics of FusionRed variants at low irradiances. These dark state conversion and ground state recovery processes are critical for applications in super-resolution microscopy such as SOFI. [219] The effect of lifetime selections on the dark state conversion and ground state recovery kinetics has not been studied.

Manna *et. al* demonstrated the phenomenon of Fluorescence Anomalous Phase Advance (FAPA) for RFPs mCherry, Kreik and TagRFP-T. [56] Analogous to phase-fluorimetry, this spectroscopic technique incorporates the use of intensity modulated excitation light. However, this method uses lower modulation frequencies (kHz in comparison to MHz for phase-fluorimetry) which is commensurate with the dark state lifetime instead of the  $S_1$  state lifetime. Instead of lagging behind the modulated light, fluorescence signal appears to precede the excitation source in time, leading to an “anomalous phase advance.” FAPA is a promising technique for probing dark state lifetime as a unique frequency spectrum characteristic of the dark time can be generated by changing the intensity or modulation frequency (Figure 7.1). [56] This spectroscopic tool can potentially be used to screen analytes based on dark state lifetime in flow. Addition of an excitation beam, intensity modulated in the kHz regime can thus be used to provide a unique phase signature for the dark time in FPs. Figure 7.1 displays simulation results that predict the signature of FAPA for an analyte flowing through a kHz modulated beam mimicking a possible in-flow experiment. These simulations were performed using numerical solutions to a three state (see Appendix 5) model used previously in Chapter 4. To mimic the excitation in flow, the sinusoidal excitation profile in time was enveloped on a Gaussian function. Correspondingly, the in-phase and out of phase components of the fluorescence signal was used to extract a phase shift – simulating the role of a lock-in amplifier for a frequency domain phase measurement.

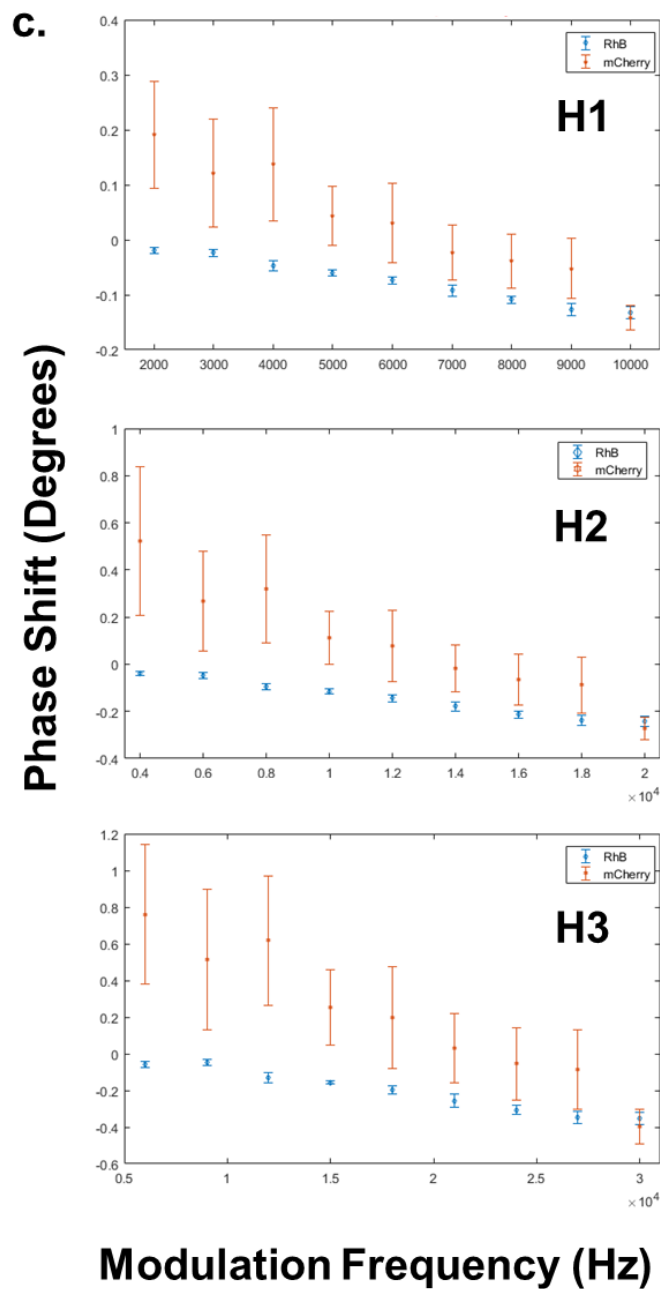
Additionally, using square wave modulation instead of sine waves can provide an opportunity to investigate several harmonics of a sine function with a single excitation beam in flow, leading to the formation of a phase spectra (Figure 7.1; Panels b and c). To validate this method, we used an optical chopper to create 50 percent and 10% duty cycle square wave illumination, and then extracted multi-harmonic phase information from pure proteins. Figure 7.2 depicts the experimental setup schematic and Figure 7.3 depicts pure protein experiments, bolstering the validity of the approach.



**Figure 7.1. The simulated phase signatures from RFPs. (a.)** Phase advances observed for mCherry, Kreik and TagRFP-T with dark times of  $\sim 0.2$  ms, 1 ms and 2.8 ms for continuous sinusoidal excitation at 5 kHz (labeled as “Cherry Sin”, “Kreik Sin” and “TagT Sin”) against the



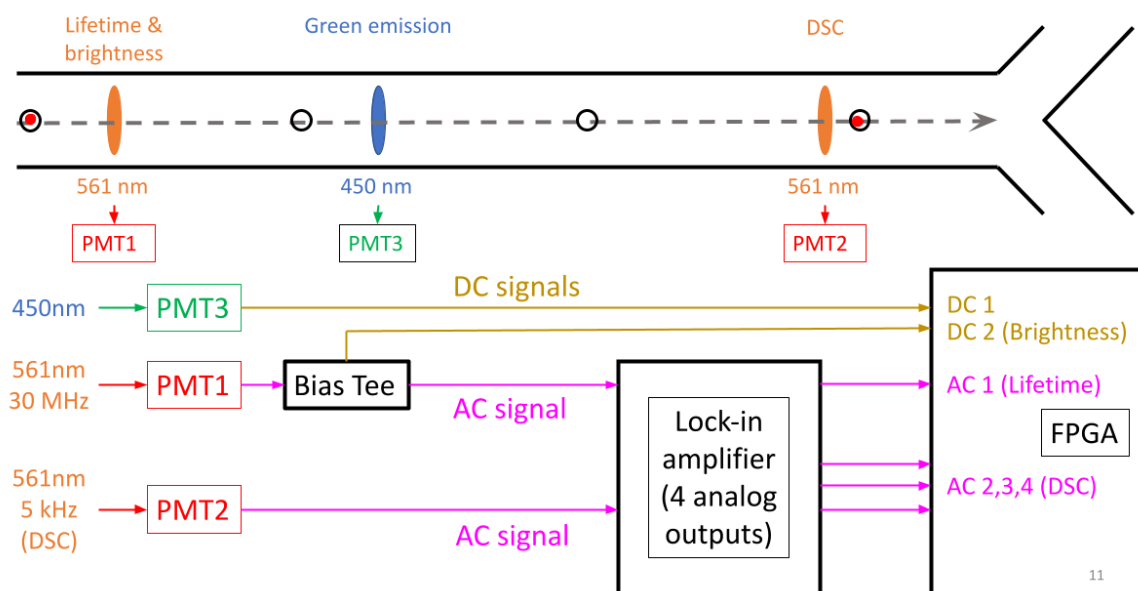
[illegible]



**Figure 7.2. Off chip phase measurements.** (a.) The experimental setup used for performing preliminary phase advance experiments with square waves on pure FP samples in 1:10 FP - Octanol droplets. (b.) Comparative measurements of RFPs excited with sinusoidal modulation at multiple frequencies between 2-10 kHz versus one 10% duty square wave at 2kHz. FPs with long dark times and high dark fractions such as C4PB12 (mCherry variant) and TagRFP-T show higher phase advances with both excitation schemes. Additionally, the amplitudes of excitation change

for each harmonic in a square wave at a certain duty cycle. This results in subtle differences in the phase spectra acquired using the two excitation schemes presented above. (c.) Comparison of phase advance observed in the dye Rhodamine-B (blue) versus mCherry (orange) using a 10% duty cycle square wave. Rhodamine-B shows negligible dark state conversion in comparison to mCherry, which is reflected in the phase advance obtained from the frequency domain multi-harmonic approach proposed here.

The experimental validity of this approach based on preliminary measurements provides exciting opportunities to select FPs on dark times and dark fractions. A tentative design for a future dark state conversion sorting device is provided in Figure 7.3.



**Figure 7.3. Tentative design of microfluidic sorting device to be used for a dark state screening platform.**

## 7.2. Optimizing bright variants of FusionRed and mCherry for imaging applications

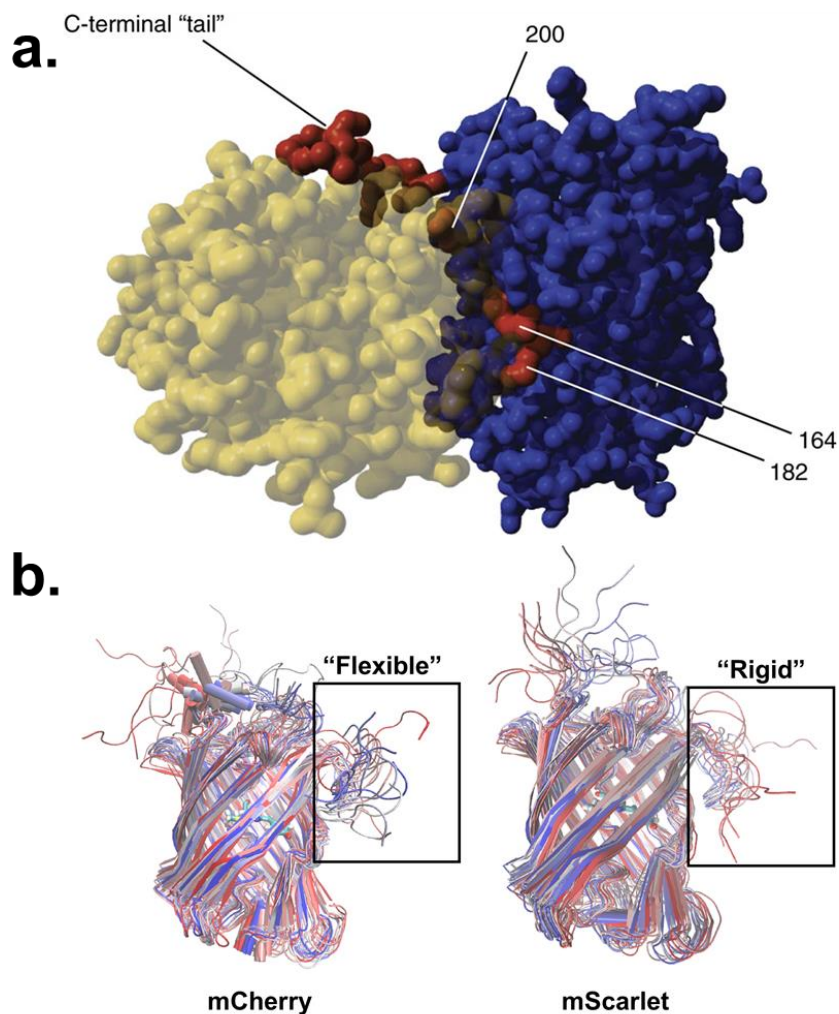
Over the course of this work, we developed FusionRed and mCherry variants with lifetimes ranging from 1 ns to 4 ns and quantum yields from 20% to 75%. While the molecular brightness of a FP is critical, the brightness in cells determines the utility of such molecules as imaging tools. Bright FPs generated in this study such as mCherry-XL and FR-HQC-TT exhibit very high fluorescence quantum yields (~70%), but do not match the performance of mScarlet in cells. [35, 38] This can be attributed to the expression and maturation kinetics of these FPs. Additionally, the chromophore maturation and expression problem in FPs is routinely addressed by targeting amino-acid residues on the FP molecule based on anecdotal or bioinformatics evidence. [35,37] Rarely mechanistic investigations decoupling the kinetics of protein expression and chromophore maturation are pursued in such efforts. [35,37] On the other hand, considerable strides have led protein engineers to concentrate on enhancing biochemical properties that govern cellular brightness such as solubility, translation efficiency, protein folding, and chromophore maturation. These efforts led to mGreenLantern, which is 6-fold brighter in mammalian cells than EGFP. This increase in cellular brightness occurred without a noticeable change in molecular brightness when compared to its predecessor, FP - Clover. [37] These approaches will be beneficial for converting bright molecules such as mCherry-XL and FusionRed-HQC-TT to brighter tools applicable in cell biology.

Balleza and co-workers were successful in quantifying the role of maturation kinetics in the cellular brightness of many FPs, yet a mechanistic investigation of the same was not pursued. [29] Moreover, the work did not address the problem of protein expression, which remains independent of chromophore maturation, and varies within contexts such as cell types (e.g, HeLa vs. yeast cells) and constructs (e.g, fusions to other proteins). Most proteins exhibit a peak at 280 nm in their absorbance spectra - a signature of the tryptophan residues in the molecule. For the same number of tryptophan residues, most FusionRed A224T variants displayed similar peak intensities at 280 nm. This suggests that expression levels of these FPs in bacteria (grown for the same period of time) were similar to the wild-type, yet many variants displayed sufficiently lower peak values for absorbance of the red-absorbing chromophore (considering the peak width and peak wavelength

are similar). This suggests that many FusionRed A224T variants suffered a maturation problem over the expression problem. On the other hand, FusionRed-MQV-VA and FusionRed-HQC-VA appeared to be brighter in bacteria yet did not match the cellular brightness of FusionRed-MQV in HeLa cells. This brightness variability could be the result of an expression problem.

Focusing on the mutational landscape of the FP outside the barrel, for example, the N-terminus and C-terminus tail, can reveal interesting insights. Shemiakina and co-workers swapped the C-terminus tail (a disordered chain of ~10-15 amino acids) of mKate2, to the tail of mCherry, to reduce the oligomerization tendency of mKate2. [188] This modification manifested in reduced maturation rates, fluorescence brightness and pH stability for the generated variant mKate2.5. To rescue the protein's spectral and biochemical properties, several rounds of combined site-directed and random mutagenesis were performed, which led to the identification of substitutions such as T75P and Q76P between the chromophore-carrying  $\alpha$ -helix and the fourth  $\beta$ -strand which come near the C-terminal tail of the FP. This process subsequently led to the identification of FusionRed which shows greatly improved maturation rates compared to mKate2.5. [188]

Motivated by these observations, we performed preliminary all-atom explicit solvent classical molecular dynamics simulations with the NAnoscale Molecular Dynamics (NAMD) package on mScarlet and mCherry. [284] Interestingly, we found the C-terminus tail was interacting with the outer side of the protein barrel in mScarlet across long timescales (~100 ns) and was fairly inflexible in comparison to the mCherry's C-terminus tail. Preliminary experimental data suggests that mutating residues on the C-terminus tail of mScarlet, significantly changes the chromophore maturation rate. These observations, if pursued, can form the mechanistic basis of understanding chromophore maturation in FPs, and aide the development of versions of mCherry-XL and FusionRed-HQC-TT for imaging applications.



**Figure 7.4. Elucidating the role of the C-terminus tail in RFPs. (a.)** The oligomerization tendencies of mKate2 is driven by the C-terminus tail. The image has been adapted from the work of Shemiakina *et. al* [188] **(b.)** Snapshots from MD [284] simulations on mCherry and mScarlet indicating a dynamic “flexible” C-terminus for mCherry and a “rigid” C-terminus for mScarlet. Each simulation is a 1  $\mu$ s run, with a structure snapshot captured every 50 ns. Red indicates snapshots from the beginning of the simulation and blue represents snapshots taken in later timesteps on the simulation trajectory.

## Bibliography

1. Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y. A Guide to Choosing Fluorescent Proteins. *Nat. Methods* **2005**, 2 (12), 905–909. <https://doi.org/10.1038/nmeth819>.
2. Tsien, R. Y. The Green Fluorescent Protein. *Annu. Rev. Biochem.* **1998**, 67, 509–544. <https://doi.org/10.1146/annurev.biochem.67.1.509>.
3. Duwé, S.; Dedecker, P. Optimizing the Fluorescent Protein Toolbox and Its Use. *Current Opinion in Biotechnology.* 2019, pp 183–191. <https://doi.org/10.1016/j.copbio.2019.04.006>.
4. Algar, W. R.; Hildebrandt, N.; Vogel, S. S.; Medintz, I. L. FRET as a Biomolecular Research Tool — Understanding Its Potential While Avoiding Pitfalls. *Nat. Methods* **2019**, 16 (9), 815–829. <https://doi.org/10.1038/s41592-019-0530-8>.
5. Coelho, S.; Poland, S. P.; Devaughes, V.; Ameer-Beg, S. M. Adaptive Optics for a Time-Resolved Förster Resonance Energy Transfer (FRET) and Fluorescence Lifetime Imaging Microscopy (FLIM) in Vivo. *Opt. Lett.* 2020, 45 (10), 2732. <https://doi.org/10.1364/ol.385950>.
6. Bando, Y.; Sakamoto, M.; Kim, S.; Ayzenshtat, I.; Yuste, R. Comparative Evaluation of Genetically Encoded Voltage Indicators. *Cell Rep.* **2019**, 26 (3), 802-813.e4. <https://doi.org/10.1016/j.celrep.2018.12.088>.
7. Tachibana, S. R.; Tang, L.; Zhu, L.; Liu, W.; Wang, Y.; Fang, C. Watching an Engineered Calcium Biosensor Glow: Altered Reaction Pathways before Emission. *J. Phys. Chem. B* **2018**, 122 (50), 11986–11995. <https://doi.org/10.1021/acs.jpcc.8b10587>.
8. Mo, H. M.; Xu, Y.; Yu, X. W. Improved Soluble Expression and Catalytic Activity of a Thermostable Esterase Using a High-Throughput Screening System Based on a Split-GFP Assembly. *J. Agric. Food Chem.* **2018**, 66 (48), 12756–12764. <https://doi.org/10.1021/acs.jafc.8b04646>.

- 
9. Liu, Y.; Wolstenholme, C. H.; Carter, G. C.; Liu, H.; Hu, H.; Grainger, L. S.; Miao, K.; Fares, M.; Hoelzel, C. A.; Yennawar, H. P.; Ning, G.; Du, M.; Bai, L.; Li, X.; Zhang, X. Modulation of Fluorescent Protein Chromophores to Detect Protein Aggregation with Turn-On Fluorescence. *J. Am. Chem. Soc.* **2018**, *140* (24), 7381–7384. <https://doi.org/10.1021/jacs.8b02176>.
  10. Fernández-Luna, V.; Coto, P. B.; Costa, R. D. When Fluorescent Proteins Meet White Light-Emitting Diodes. *Angewandte Chemie - International Edition*. 2018, pp 8826–8836. <https://doi.org/10.1002/anie.201711433>.
  11. Sigal, Y. M.; Zhou, R.; Zhuang, X. Visualizing and Discovering Cellular Structures with Super-Resolution Microscopy. *Science*. 2018, pp 880–887. <https://doi.org/10.1126/science.aau1044>.
  12. Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. Primary Structure of the Aequorea Victoria Green-Fluorescent Protein. *Gene* 1992, *111* (2), 229–233. [https://doi.org/10.1016/0378-1119\(92\)90691-H](https://doi.org/10.1016/0378-1119(92)90691-H).
  13. Prangma, J. C.; Molenaar, R.; Van Weeren, L.; Bindels, D. S.; Haarbosch, L.; Stouthamer, J.; Gadella, T. W. J.; Subramaniam, V.; Vos, W. L.; Blum, C. Quantitative Determination of Dark Chromophore Population Explains the Apparent Low Quantum Yield of Red Fluorescent Proteins. *J. Phys. Chem. B* **2020**, *124* (8), 1383–1391. <https://doi.org/10.1021/acs.jpcc.9b10396>.
  14. Simine, L.; Lammert, H.; Sun, L.; Onuchic, J. N.; Rossky, P. J. Fluorescent Proteins Detect Host Structural Rearrangements via Electrostatic Mechanism. *J. Am. Chem. Soc.* **2018**, *140* (4), 1203–1206. <https://doi.org/10.1021/jacs.7b10851>.
  15. Konold, P. E.; Yoon, E.; Lee, J.; Allen, S. L.; Chapagain, P. P.; Gerstman, B. S.; Regmi, C. K.; Piatkevich, K. D.; Verkhusha, V. V.; Joo, T.; Jimenez, R. Fluorescence from Multiple Chromophore Hydrogen-Bonding States in the Far-Red Protein TagRFP675. *J. Phys. Chem. Lett.* **2016**, *7* (15), 3046–3051. <https://doi.org/10.1021/acs.jpclett.6b01172>.



- 
16. Acharya, A.; Bogdanov, A. M.; Grigorenko, B. L.; Bravaya, K. B.; Nemukhin, A. V.; Lukyanov, K. A.; Krylov, A. I. Photoinduced Chemistry in Fluorescent Proteins: Curse or Blessing? *Chemical Reviews*. 2017, pp 758–795. <https://doi.org/10.1021/acs.chemrev.6b00238>.
17. Sen, T.; Ma, Y.; Polyakov, I. V.; Grigorenko, B. L.; Nemukhin, A. V.; Krylov, A. I. Interplay between Locally Excited and Charge Transfer States Governs the Photoswitching Mechanism in the Fluorescent Protein Dreiklang. *J. Phys. Chem. B* 2021, 125 (3), 757–770. <https://doi.org/10.1021/acs.jpcc.0c09221>.
18. Bogdanov, A. M.; Acharya, A.; Titelmayer, A. V.; Mamontova, A. V.; Bravaya, K. B.; Kolomeisky, A. B.; Lukyanov, K. A.; Krylov, A. I. Turning on and off Photoinduced Electron Transfer in Fluorescent Proteins by  $\pi$ -Stacking, Halide Binding, and Tyr145 Mutations. *J. Am. Chem. Soc.* 2016, 138 (14), 4807–4817. <https://doi.org/10.1021/jacs.6b00092>.
19. Grigorenko, B. L.; Nemukhin, A. V.; Morozov, D. I.; Polyakov, I. V.; Bravaya, K. B.; Krylov, A. I. Toward Molecular-Level Characterization of Photoinduced Decarboxylation of the Green Fluorescent Protein: Accessibility of the Charge-Transfer States. *J. Chem. Theory Comput.* 2012, 8 (6), 1912–1920. <https://doi.org/10.1021/ct300043e>.
20. Vegh, R. B.; Bravaya, K. B.; Bloch, D. A.; Bommarius, A. S.; Tolbert, L. M.; Verkhovsky, M.; Krylov, A. I.; Solntsev, K. M. Chromophore Photoreduction in Red Fluorescent Proteins Is Responsible for Bleaching and Phototoxicity. *J. Phys. Chem. B* **2014**, 118 (17), 4527–4534. <https://doi.org/10.1021/jp500919a>.
21. Klehs, K.; Spahn, C.; Endesfelder, U.; Lee, S. F.; Fürstenberg, A.; Heilemann, M. Increasing the Brightness of Cyanine Fluorophores for Single-Molecule and Superresolution Imaging. *ChemPhysChem* 2014, 15 (4), 637–641. <https://doi.org/10.1002/cphc.201300874>.

- 
22. Kumagai, A.; Ando, R.; Miyatake, H.; Greimel, P.; Kobayashi, T.; Hirabayashi, Y.; Shimogori, T.; Miyawaki, A. XA Bilirubin-Inducible Fluorescent Protein from Eel Muscle. *Cell* 2013, 153 (7). <https://doi.org/10.1016/j.cell.2013.05.038>.
  23. Rodriguez, E. A.; Tran, G. N.; Gross, L. A.; Crisp, J. L.; Shu, X.; Lin, J. Y.; Tsien, R. Y. A Far-Red Fluorescent Protein Evolved from a Cyanobacterial Phycobiliprotein. *Nat. Methods* 2016, 13 (9), 763–769. <https://doi.org/10.1038/nmeth.3935>.
  24. Shcherbakova, D. M.; Baloban, M.; Emelyanov, A. V.; Brenowitz, M.; Guo, P.; Verkhusha, V. V. Bright Monomeric Near-Infrared Fluorescent Proteins as Tags and Biosensors for Multiscale Imaging. *Nat. Commun.* 2016, 7. <https://doi.org/10.1038/ncomms12405>.
  25. Zhou, X. X.; Lin, M. Z. Photoswitchable Fluorescent Proteins: Ten Years of Colorful Chemistry and Exciting Applications. *Current Opinion in Chemical Biology.* 2013, pp 682–690. <https://doi.org/10.1016/j.cbpa.2013.05.031>.
  26. Zhou, X. X.; Lin, M. Z. Photoswitchable Fluorescent Proteins: Ten Years of Colorful Chemistry and Exciting Applications. *Current Opinion in Chemical Biology.* 2013, pp 682–690. <https://doi.org/10.1016/j.cbpa.2013.05.031>.
  27. ChemDraw V 20.1. PerkinElmer Informatics.
  28. Humphrey, W., Dalke, A. and Schulten, K., "VMD - Visual Molecular Dynamics", *J. Molec. Graphics*, **1996**, vol. 14, pp. 33-38.
  29. Balleza, E.; Kim, J. M.; Cluzel, P. Systematic Characterization of Maturation Time of Fluorescent Proteins in Living Cells. *Nature Methods* **2018**, 15 (1), 47–51. <https://doi.org/10.1038/nmeth.4509>.
  30. Manna, P.; Hung, S. T.; Mukherjee, S.; Friis, P.; Simpson, D. M.; Lo, M. N.; Palmer, A. E.; Jimenez, R. Directed Evolution of Excited State Lifetime and Brightness in FusionRed

---

Using a Microfluidic Sorter. *Integr. Biol. (United Kingdom)* **2018**, *10* (9), 516–526. <https://doi.org/10.1039/c8ib00103k>.

31. Cormack, B. P.; Valdivia, R. H.; Falkow, S. FACS-Optimized Mutants of the Green Fluorescent Protein (GFP). In *Gene*; 1996; Vol. 173, pp 33–38. [https://doi.org/10.1016/0378-1119\(95\)00685-0](https://doi.org/10.1016/0378-1119(95)00685-0).
32. Kredel, S.; Oswald, F.; Nienhaus, K.; Deuschle, K.; Röcker, C.; Wolff, M.; Heilker, R.; Nienhaus, G. U.; Wiedenmann, J. mRuby, a Bright Monomeric Red Fluorescent Protein for Labeling of Subcellular Structures. *PLoS ONE* **2009**, *4* (2). <https://doi.org/10.1371/journal.pone.0004391>.
33. Shaner, N. C.; Lambert, G. G.; Chamma, A.; Ni, Y.; Cranfill, P. J.; Baird, M. A.; Sell, B. R.; Allen, J. R.; Day, R. N.; Israelsson, M.; Davidson, M. W.; Wang, J. A Bright Monomeric Green Fluorescent Protein Derived from Branchiostoma Lanceolatum. *Nature Methods* **2013**, *10* (5), 407–409. <https://doi.org/10.1038/nmeth.2413>.
34. Grimm, J. B.; Tkachuk, A. N.; Xie, L.; Choi, H.; Mohar, B.; Falco, N.; Schaefer, K.; Patel, R.; Zheng, Q.; Liu, Z.; Lippincott-Schwartz, J.; Brown, T. A.; Lavis, L. D. A General Method to Optimize and Functionalize Red-Shifted Rhodamine Dyes. *Nature Methods* **2020**, *17* (8), 815–821. <https://doi.org/10.1038/s41592-020-0909-6>.
35. Bindels, D. S.; Haarbosch, L.; Van Weeren, L.; Postma, M.; Wiese, K. E.; Mastop, M.; Aumonier, S.; Gotthard, G.; Royant, A.; Hink, M. A.; Gadella, T. W. J. mScarlet: A Bright Monomeric Red Fluorescent Protein for Cellular Imaging. *Nature Methods* **2016**, *14* (1), 53–56. <https://doi.org/10.1038/nmeth.4074>.
36. Lambert, G. G.; Depernet, H.; Gotthard, G.; Schultz, D. T.; Navizet, I.; Lambert, T.; Adams, S. R.; Torreblanca-Zanca, A.; Chu, M.; Bindels, D. S.; Levesque, V.; Moffatt, J. N.; Salih, A.; Royant, A.; Shaner, N. C. Aequorea's Secrets Revealed: New Fluorescent Proteins with Unique Properties for Bioimaging and Biosensing. *PLoS Biology* **2020**, *18* (11). <https://doi.org/10.1371/journal.pbio.3000936>.

- 
37. Campbell, B. C.; Nabel, E. M.; Murdock, M. H.; Lao-Peregrin, C.; Tsoulfas, P.; Blackmore, M. G.; Lee, F. S.; Liston, C.; Morishita, H.; Petsko, G. A. mGreenLantern: A Bright Monomeric Fluorescent Protein with Rapid Expression and Cell Filling Properties for Neuronal Imaging. *Proceedings of the National Academy of Sciences of the United States of America* **2020**, *117* (48), 30710–30721. <https://doi.org/10.1073/pnas.2000942117>.
38. Mukherjee, S.; Hung, S. T.; Douglas, N.; Manna, P.; Thomas, C.; Ekrem, A.; Palmer, A. E.; Jimenez, R.; Engineering of a Brighter Variant of the Fusionred Fluorescent Protein Using Lifetime Flow Cytometry and Structure-Guided Mutations. *Biochemistry* **2020**, *59* (39), 3669–3682. <https://doi.org/10.1021/acs.biochem.0c00484>.
39. Bindels, D. S.; Postma, M.; Haarbosch, L.; van Weeren, L.; Gadella, T. W. J. Multiparameter Screening Method for Developing Optimized Red-Fluorescent Proteins. *Nature Protocols* **2020**, *15* (2), 450–478. <https://doi.org/10.1038/s41596-019-0250-7>.
40. Dean, K. M.; Davis, L. M.; Lubbeck, J. L.; Manna, P.; Friis, P.; Palmer, A. E.; Jimenez, R. High-Speed Multiparameter Photophysical Analyses of Fluorophore Libraries. *Analytical Chemistry* **2015**, *87* (10), 5026–5030. <https://doi.org/10.1021/acs.analchem.5b00607>.
41. Hung, S. T.; Mukherjee, S.; Jimenez, R. Enrichment of Rare Events Using a Multi-Parameter High Throughput Microfluidic Droplet Sorter. *Lab Chip* **2020**, *20* (4), 834–843. <https://doi.org/10.1039/c9lc00790c>.
42. Lubbeck, J. L.; Dean, K. M.; Ma, H.; Palmer, A. E.; Jimenez, R. Microfluidic Flow Cytometer for Quantifying Photobleaching of Fluorescent Proteins in Cells. *Analytical Chemistry* **2012**, *84* (9), 3929–3937. <https://doi.org/10.1021/ac202825z>.
43. Davis, L. M.; Lubbeck, J. L.; Dean, K. M.; Palmer, A. E.; Jimenez, R. Microfluidic Cell Sorter for Use in Developing Red Fluorescent Proteins with Improved Photostability. *Lab Chip* **2013**, *13* (12), 2320–2327. <https://doi.org/10.1039/c3lc50191d>.

- 
44. Lychagov, V. V.; Shemetov, A. A.; Jimenez, R.; Verkhusha, V. V. Microfluidic System for In-Flow Reversible Photoswitching of near-Infrared Fluorescent Proteins. *Anal. Chem.* **2016**, *88* (23), 11821–11829. <https://doi.org/10.1021/acs.analchem.6b03499>.
45. Zheng, L.; Migliore, A.; Beratan, D. N. Electrostatic Field-Induced Oscillator Strength Focusing in Molecules. *Journal of Physical Chemistry B* **2020**, *124* (29), 6376–6388. <https://doi.org/10.1021/acs.jpcc.0c04783>.
46. Drobizhev, M.; Callis, P. R.; Nifosì, R.; Wicks, G.; Stoltzfus, C.; Barnett, L.; Hughes, T. E.; Sullivan, P.; Rebane, A. Long- and Short-Range Electrostatic Fields in GFP Mutants: Implications for Spectral Tuning. *Scientific Reports* **2015**, *5*. <https://doi.org/10.1038/srep13223>.
47. Myšková, J.; Rybakova, O.; Brynda, J.; Khoroshyy, P.; Bondar, A.; Lazar, J. Directionality of Light Absorption and Emission in Representative Fluorescent Proteins. *Proceedings of the National Academy of Sciences of the United States of America* **2020**, *117* (51), 32395–32401. <https://doi.org/10.1073/pnas.2017379117>.
48. Molina, R. S.; King, J.; Franklin, J.; Clack, N.; McRaven, C.; Goncharov, V.; Flickinger, D.; Svoboda, K.; Drobizhev, M.; Hughes, T. E. High Throughput Instrument to Screen Fluorescent Proteins under Two-Photon Excitation. *Biomedical Optics Express* **2020**, *11* (12), 7192. <https://doi.org/10.1364/boe.409353>.
49. Mo, G. C. H.; Ross, B.; Hertel, F.; Manna, P.; Yang, X.; Greenwald, E.; Booth, C.; Plummer, A. M.; Tenner, B.; Chen, Z.; Wang, Y.; Kennedy, E. J.; Cole, P. A.; Fleming, K. G.; Palmer, A.; Jimenez, R.; Xiao, J.; Dedeker, P.; Zhang, J. Genetically Encoded Biosensors for Visualizing Live-Cell Biochemical Activity at Super-Resolution. *Nature Methods* **2017**, *14* (4), 427–434. <https://doi.org/10.1038/nMeth.4221>.
50. Mo, G. C. H.; Posner, C.; Rodriguez, E. A.; Sun, T.; Zhang, J. A Rationally Enhanced Red Fluorescent Protein Expands the Utility of FRET Biosensors. *Nat. Commun.* **2020**, *11* (1). <https://doi.org/10.1038/s41467-020-15687-x>.

- 
51. Mohr, M. A.; Kobitski, A. Y.; Sabater, L. R.; Nienhaus, K.; Obara, C. J.; Lippincott-Schwartz, J.; Nienhaus, G. U.; Pantazis, P. Rational Engineering of Photoconvertible Fluorescent Proteins for Dual-Color Fluorescence Nanoscopy Enabled by a Triplet-State Mechanism of Primed Conversion. *Angew. Chemie - Int. Ed.* **2017**, *56* (38), 11628–11633. <https://doi.org/10.1002/anie.201706121>.
52. Li, M. J.; Lin, Y. H.; Sung, R.; Sung, K. E-Z Isomerization Mechanism of the Green Fluorescent Protein Chromophore: Remote Regulation by Proton Dissociation of the Phenol Group. *Journal of Physical Chemistry A* **2021**, *125* (17), 3614–3621. <https://doi.org/10.1021/acs.jpca.1c01371>.
53. Subach, F. V.; Verkhusha, V. V. Chromophore Transformations in Red Fluorescent Proteins *Chemical Reviews*. **2012**, pp 4308–4327. <https://doi.org/10.1021/cr2001965>.
54. Ruhlandt, D.; Andresen, M.; Jensen, N.; Gregor, I.; Jakobs, S.; Enderlein, J.; Chizhik, A. I. Absolute Quantum Yield Measurements of Fluorescent Proteins Using a Plasmonic Nanocavity. *Communications Biology* **2020**, *3* (1), 1–7. <https://doi.org/10.1038/s42003-020-01316-2>.
55. Dean, K. M.; Lubbeck, J. L.; Binder, J. K.; Schwall, L. R.; Jimenez, R.; Palmer, A. E. Analysis of Red-Fluorescent Proteins Provides Insight into Dark-State Conversion and Photodegradation. *Biophysical Journal* **2011**, *101* (4), 961–969. <https://doi.org/10.1016/j.bpj.2011.06.055>.
56. Manna, P.; Jimenez, R. Time and Frequency-Domain Measurement of Ground-State Recovery Times in Red Fluorescent Proteins. *Journal of Physical Chemistry B* **2015**, *119* (15), 4944–4954. <https://doi.org/10.1021/acs.jpcb.5b00950>.
57. Lambert, T. J. FPbase: A Community-Editable Fluorescent Protein Database. *Nature Methods*. **2019**, pp 277–278. <https://doi.org/10.1038/s41592-019-0352-8>.
58. Drobizhev, M.; Molina, R. S.; Callis, P. R.; Scott, J. N.; Lambert, G. G.; Salih, A.; Shaner, N. C.; Hughes, T. E. Local Electric Field Controls Fluorescence Quantum Yield of Red

- 
- and Far-Red Fluorescent Proteins. *Frontiers in Molecular Biosciences* **2021**, 8. <https://doi.org/10.3389/fmolb.2021.633217>.
59. Mamontova, A. V.; Solovyev, I. D.; Savitsky, A. P.; Shakhov, A.; Lukyanov, K. A.; Bogdanov, A. M. Bright GFP with Subnanosecond Fluorescence Lifetime. *Scientific Reports* **2018**, 8 (1). <https://doi.org/10.1038/s41598-018-31687-w>.
60. Strickler, S. J.; Berg, R. A. Relationship between Absorption Intensity and Fluorescence Lifetime of Molecules. *The Journal of Chemical Physics* **1962**, 37 (4), 814–822. <https://doi.org/10.1063/1.1733166>.
61. Jung, G., Brockhinke, A., Gensch, T., Hötzer, B., Schwedler, S., & Veettil, S. K. Fluorescence Lifetime of Fluorescent Proteins. In G. Jung (Ed.), *Fluorescent Proteins I: From Understanding to Design* **2011**, (pp. 69–97). Springer Berlin Heidelberg. [https://doi.org/10.1007/4243\\_2011\\_14](https://doi.org/10.1007/4243_2011_14)
62. Lin, C. Y.; Romei, M. G.; Oltrogge, L. M.; Mathews, I. I.; Boxer, S. G. Unified Model for Photophysical and Electro-Optical Properties of Green Fluorescent Proteins. *Journal of the American Chemical Society* **2019**, 141 (38), 15250–15265. <https://doi.org/10.1021/jacs.9b07152>.
63. Terazono, H., Anzai, Y., Soloviev, M., & Yasuda, K. Labelling of live cells using fluorescent aptamers: Binding reversal with DNA nucleases. *Journal of Nanobiotechnology* **2010**, 8, 8. <https://doi.org/10.1186/1477-3155-8-8>
64. Tregidgo, C., Levitt, J. A., & Suhling, K. Effect of refractive index on the fluorescence lifetime of green fluorescent protein. *Journal of Biomedical Optics* **2008**, 13(3), 031218. <https://doi.org/10.1117/1.2937212>
65. Byrdin, B., Duan, C., Bourgeois, D., and Brettel, K., A Long-Lived Triplet State Is the Entrance Gateway to Oxidative Photochemistry in Green Fluorescent Proteins. *Journal of the American Chemical Society* **2018** 140 (8), 2897-2905. <https://doi.org/10.1021/jacs.7b12755>

- 
66. Canty, L., Hariharan, S., Liu, Q., Haney, S.A., Andrews, D. W. Peak emission wavelength and fluorescence lifetime are coupled in far-red, GFP-like fluorescent proteins. *PLOS ONE* **2018**, 13(11): e0208075. <https://doi.org/10.1371/journal.pone.0208075>
67. Mairing, K., Krasnenko, V., & Miller, S. Photophysics of the blue fluorescent protein. *Journal of Luminescence* **2007**, 122–123(1–2), 291–293. <https://doi.org/10.1016/j.jlumin.2006.01.144>
68. Bhalekar, S. B., Bhagwat, A. A., & Sekar, N. Orange-Red Fluorescent (Partially Rigidified) Donor- $\pi$ -(rigidified)-Acceptor System – Computational Studies. *Journal of Fluorescence* **2020**, 30(3), 565–579. <https://doi.org/10.1007/s10895-020-02506-1>
69. Englman, R.; Jortner, J. The Energy Gap Law for Radiationless Transitions in Large Molecules. *Molecular Physics* **1970**, 18 (2), 285–287. <https://doi.org/10.1080/00268977000100171.1>
70. Marcus, R. A. On the theory of oxidation-reduction reactions involving electron transfer. I. *The Journal of Chemical Physics* **1956**, 24(5), 966–978. <https://doi.org/10.1063/1.1742723>
71. Fleming, G. R., & Cho, M. Chromophore-solvent dynamics. *Annual Review of Physical Chemistry* **1996**, 47(1), 109–134. <https://doi.org/10.1146/annurev.physchem.47.1.109>
72. Jordanides, X. J., Lang, M. J., Song, X., & Fleming, G. R. Solvation dynamics in protein environments studied by photon echo spectroscopy. *Journal of Physical Chemistry B* **1999**, 103(37), 7995–8005. <https://doi.org/10.1021/jp9910993>
73. Jimenez, R., Salazar, G., Yin, J., Joo, T., & Romesberg, F. E. Protein dynamics and the immunological evolution of molecular recognition. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, 101(11), 3803–3808. <https://doi.org/10.1073/pnas.0305745101>



- 
74. Park, J. W.; Rhee, Y. M. Electric Field Keeps Chromophore Planar and Produces High Yield Fluorescence in Green Fluorescent Protein. *Journal of the American Chemical Society* **2016**, *138* (41), 13619–13629. <https://doi.org/10.1021/jacs.6b06833>.
75. Bravaya, K. B.; Grigorenko, B. L.; Nemukhin, A. V.; Krylov, A. I. Quantum Chemistry behind Bioimaging: Insights from Ab Initio Studies of Fluorescent Proteins and Their Chromophores. *Accounts of chemical research* **2012**, *45* (2), 265–275. <https://doi.org/10.1021/ar2001556>.
76. Shu, X.; Shaner, N. C.; Yarbrough, C. A.; Tsien, R. Y.; Remington, S. J. Novel Chromophores and Buried Charges Control Color in mFruits. *Biochemistry* **2006**, *45* (32), 9639–9647. <https://doi.org/10.1021/bi060773l>.
77. Goedhart, J.; Von Stetten, D.; Noirclerc-Savoye, M.; Lelimosin, M.; Joosen, L.; Hink, M. A.; Van Weeren, L.; Gadella, T. W. J.; Royant, A. Structure-Guided Evolution of Cyan Fluorescent Proteins towards a Quantum Yield of 93%. *Nature Communications* **2012**, *3*. <https://doi.org/10.1038/ncomms1738>.
78. Conyard, J.; Heisler, I. A.; Chan, Y.; Bulman Page, P. C.; Meech, S. R.; Blancafort, L. A New Twist in the Photophysics of the GFP Chromophore: A Volume-Conserving Molecular Torsion Couple. *Chemical Science* **2018**, *9* (7), 1803–1812. <https://doi.org/10.1039/c7sc04091a>.
79. Konold, P.; Regmi, C. K.; Chapagain, P. P.; Gerstman, B. S.; Jimenez, R. Hydrogen Bond Flexibility Correlates with Stokes Shift in mPlum Variants. *Journal of Physical Chemistry B* **2014**, *118* (11), 2940–2948. <https://doi.org/10.1021/jp412371y>.
80. Steiert, F.; Petrov, E. P.; Schultz, P.; Schwille, P.; Weidemann, T. Photophysical Behavior of mNeonGreen, an Evolutionarily Distant Green Fluorescent Protein. *Biophysical Journal* **2018**, *114* (10), 2419–2431. <https://doi.org/10.1016/j.bpj.2018.04.013>.
81. Tao, A.; Zhang, R.; Yuan, J. Characterization of Photophysical Properties of Photoactivatable Fluorescent Proteins for Super-Resolution Microscopy. *Journal of*

- 
- Physical Chemistry B* **2020**, *124* (10), 1892–1897.  
<https://doi.org/10.1021/acs.jpcc.9b11028>.
82. Romei, M. G.; Lin, C. Y.; Mathews, I. I.; Boxer, S. G. Electrostatic Control of Photoisomerization Pathways in Proteins. *Science* **2020**, *367* (6473), 76–79.  
<https://doi.org/10.1126/science.aax1898>.
  83. Coquelle, N.; Sliwa, M.; Woodhouse, J.; Schirò, G.; Adam, V.; Aquila, A.; Barends, T. R. M.; Boutet, S.; Byrdin, M.; Carbajo, S.; Mora, E. D. La; Doak, R. B.; Feliks, M.; Fieschi, F.; Foucar, L.; Guillon, V.; Hilpert, M.; Hunter, M. S.; Jakobs, S.; Koglin, J. E.; Kovacsova, G.; Lane, T. J.; Lévy, B.; Liang, M.; Nass, K.; Ridard, J.; Robinson, J. S.; Roome, C. M.; Ruckebusch, C.; Seaberg, M.; Thepaut, M.; Cammarata, M.; Demachy, I.; Field, M.; Shoeman, R. L.; Bourgeois, D.; Colletier, J. P.; Schlichting, I.; Weik, M. Chromophore Twisting in the Excited State of a Photoswitchable Fluorescent Protein Captured by Time-Resolved Serial Femtosecond Crystallography. *Nature Chemistry* **2018**, *10* (1), 31–37.  
<https://doi.org/10.1038/NCHEM.2853>.
  84. Liptonok, S. P.; Gil, A. A.; Hall, C. R.; Lukacs, A.; Iuliano, J. N.; Jones, G. A.; Greetham, G. M.; Donaldson, P.; Miyawaki, A.; Tonge, P. J.; Meech, S. R. Infrared Spectroscopy Reveals Multi-Step Multi-Timescale Photoactivation in the Photoconvertible Protein Archetype Dronpa. *Nature Chemistry* **2018**, *10* (8), 845–852.  
<https://doi.org/10.1038/s41557-018-0073-0>.
  85. Mizuno, H.; Mal, T. K.; Wälchli, M.; Kikuchi, A.; Fukano, T.; Ando, R.; Jeyakanthan, J.; Taka, J.; Shiro, Y.; Ikura, M.; Miyawaki, A. Light-Dependent Regulation of Structural Flexibility in a Photochromic Fluorescent Protein. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105* (27), 9227–9232.  
<https://doi.org/10.1073/pnas.0709599105>.
  86. Oscar, B. G.; Zhu, L.; Wolfenden, H.; Rozanov, N. D.; Chang, A.; Stout, K. T.; Sandwich, J. W.; Porter, J. J.; Mehl, R. A.; Fang, C. Dissecting Optical Response and Molecular Structure of Fluorescent Proteins with Non-Canonical Chromophores. *Frontiers in Molecular Biosciences* **2020**, *7* (July), 1–10. <https://doi.org/10.3389/fmolb.2020.00131>.

- 
87. Widder, P.; Schuck, J.; Summerer, D.; Drescher, M. Combining Site-Directed Spin Labeling: In Vivo and in-Cell EPR Distance Determination. *Physical Chemistry Chemical Physics* **2020**, 22 (9), 4875–4879. <https://doi.org/10.1039/c9cp05584c>.
88. Christou, N. E.; Ayala, I.; Giandoreggio-Barranco, K.; Byrdin, M.; Adam, V.; Bourgeois, D.; Brutscher, B. NMR Reveals Light-Induced Changes in the Dynamics of a Photoswitchable Fluorescent Protein. *Biophysical Journal* **2019**, 117 (11), 2087–2100. <https://doi.org/10.1016/j.bpj.2019.10.035>.
89. Bourgeois, D. Deciphering Structural Photophysics of Fluorescent Proteins by Kinetic Crystallography. *International Journal of Molecular Sciences*. **2017**. <https://doi.org/10.3390/ijms18061187>.
90. De Zitter, E.; Thédié, D.; Mönkemöller, V.; Hugelier, S.; Beaudouin, J.; Adam, V.; Byrdin, M.; Van Meervelt, L.; Dedecker, P.; Bourgeois, D. Mechanistic Investigation of mEos4b Reveals a Strategy to Reduce Track Interruptions in SptPALM. *Nature Methods* **2019**, 16 (8), 707–710. <https://doi.org/10.1038/s41592-019-0462-3>.
91. Yoon, E.; Konold, P. E.; Lee, J.; Joo, T.; Jimenez, R. Far-Red Emission of mPlum Fluorescent Protein Results from Excited-State Interconversion between Chromophore Hydrogen-Bonding States. *Journal of Physical Chemistry Letters* **2016**, 7 (12), 2170–2174. <https://doi.org/10.1021/acs.jpcllett.6b00823>.
92. Moron, V.; Marazzi, M.; Wanko, M. Far Red Fluorescent Proteins: Where Is the Limit of the Acylimine Chromophore? *Journal of Chemical Theory and Computation* **2019**, 15 (7), 4228–4240. <https://doi.org/10.1021/acs.jctc.9b00070>.
93. Faraji, S.; Krylov, A. I. On the Nature of an Extended Stokes Shift in the mPlum Fluorescent Protein. *Journal of Physical Chemistry B* **2015**, 119 (41), 13052–13062. <https://doi.org/10.1021/acs.jpcb.5b07724>.

- 
94. Abbyad, P.; Childs, W.; Shi, X.; Boxer, S. G. Dynamic Stokes Shift in Green Fluorescent Protein Variants. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104* (51), 20189–20194. <https://doi.org/10.1073/pnas.0706185104>.
95. Piatkevich, K. D.; Malashkevich, V. N.; Morozova, K. S.; Nemkovich, N. A.; Almo, S. C.; Verkhusha, V. V. Extended Stokes Shift in Fluorescent Proteins: Chromophore-Protein Interactions in a near Infrared TagRFP675 Variant. *Scientific Reports* **2013**, *3*. <https://doi.org/10.1038/srep01847>.
96. Khrenova, M. G.; Mulashkin, F. D.; Bulavko, E. S.; Zakharova, T. M.; Nemukhin, A. V. Dipole Moment Variation Clears up Electronic Excitations in the  $\pi$ -Stacked Complexes of Fluorescent Protein Chromophores. *Journal of Chemical Information and Modeling* **2020**, *60* (12), 6288–6297. <https://doi.org/10.1021/acs.jcim.0c01028>.
97. Kang, B.; Liu, H.; Jang, D. J.; Lee, J. Y. Electric Field Effect on the Ground State Proton Transfer in the H-Bonded HBDI Complex: An Implication of the Green Fluorescent Protein. *RSC Advances* **2014**, *4* (51), 26543–26551. <https://doi.org/10.1039/c4ra00974f>.
98. Kang, B.; Baek, K. Y.; Lee, J. Y. Electric Field Effect on Trans-p-Hydroxybenzylideneimidazolidinone: A DFT Study and Implication to Green Fluorescent Protein. *Bulletin of the Korean Chemical Society* **2015**, *36* (1), 276–281. <https://doi.org/10.1002/bkcs.10063>.
99. Treynor, T. P., Andrews, S. S., & Boxer, S. G. Intervalence band stark effect of the special pair radical cation in bacterial photosynthetic reaction centers. *Journal of Physical Chemistry B* **2003**, *107*(40), 11230–11239. <https://doi.org/10.1021/jp035039f>
100. Grigorenko, B. L.; Nemukhin, A. V.; Polyakov, I. V.; Khrenova, M. G.; Krylov, A. I. A Light-Induced Reaction with Oxygen Leads to Chromophore Decomposition and Irreversible Photobleaching in GFP-Type Proteins. *Journal of Physical Chemistry B* **2015**. <https://doi.org/10.1021/acs.jpcb.5b02271>.

- 
101. Cranfill, P. J.; Sell, B. R.; Baird, M. A.; Allen, J. R.; Lavagnino, Z.; De Gruiter, H. M.; Kremers, G. J. Davidson, M. W.; Ustione, A.; Piston, D. W. Quantitative Assessment of Fluorescent Proteins. *Nat. Methods* **2016**, *13* (7), 557–562. <https://doi.org/10.1038/nmeth.3891>.
102. Meijer, M. S., Rojas-Gutierrez, P. A., Busko, D., Howard, I. A., Frenzel, F., Würth, C., Resch-Genger, U., Richards, B. S., Turshatov, A., Capobianco, J. A., & Bonnet, S. Absolute upconversion quantum yields of blue-emitting LiYF<sub>4</sub>:Yb<sup>3+</sup>, Tm<sup>3+</sup> upconverting nanoparticles. *Physical Chemistry Chemical Physics* **2018**, *20*(35), 22556–22562. <https://doi.org/10.1039/c8cp03935f>
103. Zak, P. P.; Lapina, V. A.; Pavich, T. A.; Trofimov, A. V; Trofimova, N. N.; Tsaplev, Y. B. Luminescent Materials for Modern Light Sources. *Russian Chemical Reviews* **2017**, *86* (9), 831–844. <https://doi.org/10.1070/rcr4735>.
104. Widengren, J., & Rigler, R. Mechanisms of photobleaching investigated by fluorescence correlation spectroscopy. *Bioimaging* **1996**, *4*(3), 149–157. [https://doi.org/10.1002/1361-6374\(199609\)4:3<149::aid-bio5>3.0.co;2-d](https://doi.org/10.1002/1361-6374(199609)4:3<149::aid-bio5>3.0.co;2-d)
105. Serra, F., & Terentjev, E. M. Nonlinear dynamics of absorption and photobleaching of dyes. *Journal of Chemical Physics* **2008**, *128*(22), 224510. <https://doi.org/10.1063/1.2937455>
106. Lew, M. D., Lee, S. F., Ptacin, J. L., Lee, M. K., Twieg, R. J., Shapiro, L., & Moerner, W. E. Three-dimensional superresolution colocalization of intracellular protein superstructures and the cell surface in live *Caulobacter crescentus*. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108*(46), E1102 LP-E1110. <https://doi.org/10.1073/pnas.1114444108>
107. Demchenko, A. P. Photobleaching of Organic Fluorophores: Quantitative Characterization, Mechanisms, Protection. *Methods and Applications in Fluorescence* **2020**. <https://doi.org/10.1088/2050-6120/ab7365>.

- 
108. Dean, K. M.; Lubbeck, J. L.; Davis, L. M.; Regmi, C. K.; Chapagain, P. P.; Gerstman, B. S.; Jimenez, R.; Palmer, A. E. Microfluidics-Based Selection of Red-Fluorescent Proteins with Decreased Rates of Photobleaching. *Integrative Biology (United Kingdom)* **2015**, *7* (2), 263–273. <https://doi.org/10.1039/c4ib00251b>.
109. Bulina, M. E.; Lukyanov, K. A.; Britanova, O. V.; Onichtchouk, D.; Lukyanov, S.; Chudakov, D. M. Chromophore-Assisted Light Inactivation (CALI) Using the Phototoxic Fluorescent Protein KillerRed. *Nature Protocols* **2006**, *1* (2), 947–953. <https://doi.org/10.1038/nprot.2006.89>.
110. Sen, T.; Mamontova, A. V.; Titelmayer, A. V.; Shakhov, A. M.; Astafiev, A. A.; Acharya, A.; Lukyanov, K. A.; Krylov, A. I.; Bogdanov, A. M. Influence of the First Chromophore-Forming Residue on Photobleaching and Oxidative Photoconversion of EGFP and EYFP. *International Journal of Molecular Sciences* **2019**, *20* (20). <https://doi.org/10.3390/ijms20205229>.
111. Bogdanov, A. M.; Bogdanova, E. A.; Chudakov, D. M.; Gorodnicheva, T. V.; Lukyanov, S.; Lukyanov, K. A. Cell Culture Medium Affects GFP Photostability: A Solution. *Nature Methods*. **2009**, pp 859-860. <https://doi.org/10.1038/nmeth1209-859>.
112. Gorbachev, D. A.; Petrusevich, E. F.; Kabylda, A. M.; Maksimov, E. G.; Lukyanov, K. A.; Bogdanov, A. M.; Baranov, M. S.; Bochenkova, A. V.; Mishin, A. S. A General Mechanism of Green-to-Red Photoconversions of GFP. *Frontiers in Molecular Biosciences* **2020**, *7*. <https://doi.org/10.3389/fmolb.2020.00176>.
113. Kwon, J.; Park, J. S.; Kang, M.; Choi, S.; Park, J.; Kim, G. T.; Lee, C.; Cha, S.; Rhee, H. W.; Shim, S. H. Bright Ligand-Activatable Fluorescent Protein for High-Quality Multicolor Live-Cell Super-Resolution Microscopy. *Nature Communications* **2020**. <https://doi.org/10.1038/s41467-019-14067-4>.

- 
114. Lee, J.; Liu, Z.; Suzuki, P. H.; Ahrens, J. F.; Lai, S.; Lu, X. Versatile Phenotype-Activated Cell Sorting. *Science Advances* **2020**, No. October, 1–18. <https://doi.org/10.1126/sciadv.abb7438>
115. Wiens, M. D.; Hoffmann, F.; Chen, Y.; Campbell, R. E. Enhancing Fluorescent Protein Photostability through Robot-Assisted Photobleaching. *Integrative Biology (United Kingdom)* **2018**, *10* (7), 419–428. <https://doi.org/10.1039/c8ib00063h>
116. Zhong, S.; Rivera-Molina, F.; Rivetta, A.; Toomre, D.; Santos-Sacchi, J.; Navaratnam, D. Seeing the Long Tail: A Novel Green Fluorescent Protein, SiriusGFP, for Ultra Long Timelapse Imaging. *Journal of Neuroscience Methods* **2019**, *313*, 68–76. <https://doi.org/10.1016/j.jneumeth.2018.12.008>.
117. Chatterjee, T.; Mandal, M.; Gude, V.; Bag, P. P.; Mandal, P. K. Strong Electron Donation Induced Differential Nonradiative Decay Pathways for Para and Meta GFP Chromophore Analogues. *Physical Chemistry Chemical Physics* **2015**, *17* (32), 20515–20521. <https://doi.org/10.1039/c5cp03086b>.
118. Ren, H.; Yang, B.; Ma, C.; Hu, Y. S.; Wang, P. G.; Wang, L. Cysteine Sulfoxidation Increases Photostability of Red Fluorescent Proteins. *ACS Chemical Biology* **2016**, *11* (10), 2679–2684. <https://doi.org/10.1021/acscchembio.6b00579>.
119. Nahidiazar, L.; Agronskaia, A. V.; Broertjes, J.; Van Broek, B. Den; Jalink, K. Optimizing Imaging Conditions for Demanding Multi-Color Super Resolution Localization Microscopy. *PLoS ONE* **2016**, *11* (7). <https://doi.org/10.1371/journal.pone.0158884>.
120. Grimm, J. B.; Xie, L.; Casler, J. C.; Patel, R.; Tkachuk, A. N.; Choi, H.; Lippincott Schwartz, J.; Brown, T. A.; Glick, B. S.; Liu, Z.; Lavis, L. D. Deuteration Improves Small-Molecule Fluorophores. *bioRxiv*. **2020**. <https://doi.org/10.1101/2020.08.17.250027>.
121. Parzuchowski, K. M.; Mikhaylov, A.; Mazurek, M. D.; Wilson, R. N.; Lum, D. J.; Gerrits, T.; Camp, C.H.; Stevens, M. J.; Jimenez, R. Setting Bounds on Two-Photon Absorption

- 
- Cross-Sections in Common Fluorophores with Entangled Photon Pair Excitation. *Physical Review Applied* **2021**, 15 (4). <https://doi.org/10.1103/PhysRevApplied.15.044012>.
122. Kay, B. K., Winter, J., & McCafferty, J. B. T.-P. D. of P. and P. (Eds.). Front Matter. In *Phage Display of Peptides and Proteins* **1996** (p. iii). Academic Press. <https://doi.org/10.1016/b978-0-12-402380-2.50023-x>
123. Ellington, A. D., & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **1990**, 346(6287), 818–822. <https://doi.org/10.1038/346818a>
124. Sarkisyan, K. S., Bolotin, D. A., Meer, M. V., Usmanova, D. R., Mishin, A. S., Sharonov, G. V., Ivankov, D. N., Bozhanova, N. G., Baranov, M. S., Soylemez, O., Bogatyreva, N. S., Vlasov, P. K., Egorov, E. S., Logacheva, M. D., Kondrashov, A. S., Chudakov, D. M., Putintseva, E. V., Mamedov, I. Z., Tawfik, D. S., Kondrashov, F. A. Local fitness landscape of the green fluorescent protein. *Nature* **2016**, 533, 397–401. <https://doi.org/10.1038/nature17995>
125. Lin, C.-Y.; Romei, M. G.; Mathews, I. I.; Boxer, S. G. Energetic Basis and Design of Enzyme Function Demonstrated Using GFP, an Excited-State Enzyme. *J. Am. Chem. Soc.* **2022**. <https://doi.org/10.1021/jacs.1c12305>.
126. Inada, N.; Fukuda, N.; Hayashi, T.; Uchiyama, S. Temperature Imaging Using a Cationic Linear Fluorescent Polymeric Thermometer and Fluorescence Lifetime Imaging Microscopy. *Nat. Protoc.* **2019**, 14 (4), 1293–1321. <https://doi.org/10.1038/s41596-019-0145-7>.
127. Lakner, P. H.; Monaghan, M. G.; Möller, Y.; Olayioye, M. A.; Schenke-Layland, K. Applying Phasor Approach Analysis of Multiphoton FLIM Measurements to Probe the Metabolic Activity of Three-Dimensional in Vitro Cell Culture Models. *Sci. Rep.* **2017**, 7, 42730. <https://doi.org/10.1038/srep42730>.
128. Suhling, K.; Hirvonen, L. M.; Levitt, J. A.; Chung, P. H.; Tregidgo, C.; Le Marois, A.;



- 
- Rusakov, D. A.; Zheng, K.; Ameer-Beg, S.; Poland, S.; Coelho, S.; Henderson, R.; Krstajic, N. Fluorescence Lifetime Imaging (FLIM): Basic Concepts and Some Recent Developments. *Med. Photonics* **2015**, *27*, 3–40. <https://doi.org/10.1016/j.medpho.2014.12.001>.
129. Bücherl, C. A.; Bader, A.; Westphal, A. H.; Laptienok, S. P.; Borst, J. W. FRET-FLIM Applications in Plant Systems. *Protoplasma* **2014**, *251* (2), 383–394. <https://doi.org/10.1007/s00709-013-0595-7>.
130. Benčina, M. Illumination of the Spatial Order of Intracellular PH by Genetically Encoded pH-Sensitive Sensors. *Sensors (Switzerland)* **2013**, *13* (12), 16736–16758. <https://doi.org/10.3390/s131216736>.
131. Schmitt, F. J.; Thaa, B.; Junghans, C.; Vitali, M.; Veit, M.; Friedrich, T. EGFP-PHsens as a Highly Sensitive Fluorophore for Cellular PH Determination by Fluorescence Lifetime Imaging Microscopy (FLIM). *Biochim. Biophys. Acta - Bioenerg.* **2014**, *1837* (9), 1581–1593. <https://doi.org/10.1016/j.bbabi.2014.04.003>.
132. Van Manen, H. J.; Verkuijlen, P.; Wittendorp, P.; Subramaniam, V.; Van Den Berg, T. K.; Roos, D.; Otto, C. Refractive Index Sensing of Green Fluorescent Proteins in Living Cells Using Fluorescence Lifetime Imaging Microscopy. *Biophys. J.* **2008**, *94* (8), L67–L69. <https://doi.org/10.1529/biophysj.107.127837>.
133. Pliss, A.; Peng, X.; Liu, L.; Kuzmin, A.; Wang, Y.; Qu, J.; Li, Y.; Prasad, P. N. Single Cell Assay for Molecular Diagnostics and Medicine: Monitoring Intracellular Concentrations of Macromolecules by Two-Photon Fluorescence Lifetime Imaging. *Theranostics* **2015**, *5* (9), 919–930. <https://doi.org/10.7150/thno.11863>.
134. Margineanu, A.; Chan, J. J.; Kelly, D. J.; Warren, S. C.; Flatters, D.; Kumar, S.; Katan, M.; Dunsby, C. W.; French, P. M. W. Screening for Protein-Protein Interactions Using

- 
- Förster Resonance Energy Transfer (FRET) and Fluorescence Lifetime Imaging Microscopy (FLIM). *Sci. Rep.* **2016**, *6*, 28186. <https://doi.org/10.1038/srep28186>.
135. Rebbeck, R. T.; Essawy, M. M.; Nitu, F. R.; Grant, B. D.; Gillispie, G. D.; Thomas, D. D.; Bers, D. M.; Cornea, R. L. High-Throughput Screens to Discover Small-Molecule Modulators of Ryanodine Receptor Calcium Release Channels. *SLAS Discov.* **2017**, *22* (2), 176–186. <https://doi.org/10.1177/1087057116674312>.
136. Long, Y.; Stahl, Y.; Weidtkamp-Peters, S.; Postma, M.; Zhou, W.; Goedhart, J.; Sánchez-Pérez, M. I.; Gadella, T. W. J.; Simon, R.; Scheres, B.; Blilou, I. In Vivo FRET-FLIM Reveals Cell-Type-Specific Protein Interactions in Arabidopsis Roots. *Nature* **2017**, *548* (7665), 97–102. <https://doi.org/10.1038/nature23317>.
137. Talbot, C. B.; McGinty, J.; Grant, D. M.; McGhee, E. J.; Owen, D. M.; Zhang, W.; Bunney, T. D.; Munro, I.; Isherwood, B.; Eagle, R.; Hargreaves, A.; Katan, M.; Dunsby, C.; Neil, M. A. A.; French, P. M. W. High Speed Unsupervised Fluorescence Lifetime Imaging Confocal Multiwell Plate Reader for High Content Analysis. *J. Biophotonics* **2008**, *1* (6), 514–521. <https://doi.org/10.1002/jbio.200810054>.
138. Kawanabe, S.; Araki, Y.; Uchimura, T.; Imasaka, T. Applying Fluorescence Lifetime Imaging Microscopy to Evaluate the Efficacy of Anticancer Drugs. *Methods Appl. Fluoresc.* **2015**, *3* (2), 25006. <https://doi.org/10.1088/2050-6120/3/2/025006>.
139. Humpolíčková, J.; Weber, J.; Starková, J.; Mašíňová, E.; Günterová, J.; Flaisigová, I.; Konvalinka, J.; Majerová, T. Inhibition of the Precursor and Mature Forms of HIV-1 Protease as a Tool for Drug Evaluation. *Sci. Rep.* **2018**, *8* (1), 10438. <https://doi.org/10.1038/s41598-018-28638-w>.
140. Dai, X.; Yue, Z.; Eccleston, M. E.; Swartling, J.; Slater, N. K. H.; Kaminski, C. F. Fluorescence Intensity and Lifetime Imaging of Free and Micellar-Encapsulated

- 
- Doxorubicin in Living Cells. *Nanomedicine Nanotechnology, Biol. Med.* 2008, 4 (1), 49–56. <https://doi.org/10.1016/j.nano.2007.12.002>.
141. Bakker, G. J.; Andresen, V.; Hoffman, R. M.; Friedl, P. Fluorescence Lifetime Microscopy of Tumor Cell Invasion, Drug Delivery, and Cytotoxicity. In *Methods in Enzymology*; Conn, P. M. B. T.-M. in E., Ed.; Academic Press, 2012; Vol. 504, pp 109–125. <https://doi.org/10.1016/B978-0-12-391857-4.00005-7>.
142. Conway, J. R. W.; Carragher, N. O.; Timpson, P. Developments in Preclinical Cancer Imaging: Innovating the Discovery of Therapeutics. *Nat. Rev. Cancer* 2014, 14 (5), 314–328. <https://doi.org/10.1038/nrc3724>.
143. Ardeshirpour, Y.; Chernomordik, V.; Hassan, M.; Zielinski, R.; Capala, J.; Gandjbakhche, A. In Vivo Fluorescence Lifetime Imaging for Monitoring the Efficacy of the Cancer Treatment. *Clin. cancer Res. an Off. J. Am. Assoc. Cancer Res.* **2014**, 20 (13), 3531–3539. <https://doi.org/10.1158/1078-0432.CCR-13-1826>.
144. Gohar, A. V.; Cao, R.; Jenkins, P.; Li, W.; Houston, J. P.; Houston, K. D. Subcellular Localization-Dependent Changes in EGFP Fluorescence Lifetime Measured by Time-Resolved Flow Cytometry. *Biomed. Opt. Express* **2013**, 4 (8), 1390. <https://doi.org/10.1364/boe.4.001390>.
145. Sambrano, J.; Chigaev, A.; Nichani, K. S.; Smagley, Y.; Sklar, L. A.; Houston, J. P. Evaluating Integrin Activation with Time-Resolved Flow Cytometry. *J. Biomed. Opt.* **2018**, 23 (07), 1. <https://doi.org/10.1117/1.jbo.23.7.075004>.
146. Suzuki, M.; Sakata, I.; Sakai, T.; Tomioka, H.; Nishigaki, K.; Tramier, M.; Coppey-Moisan, M. A High-Throughput Direct Fluorescence Resonance Energy Transfer-Based Assay for Analyzing Apoptotic Proteases Using Flow Cytometry and Fluorescence Lifetime Measurements. *Anal. Biochem.* **2015**, 491, 10–17. <https://doi.org/10.1016/j.ab.2015.08.022>.

- 
147. Alturkistany, F.; Nichani, K.; Houston, K. D.; Houston, J. P. Fluorescence Lifetime Shifts of NAD(P)H during Apoptosis Measured by Time-Resolved Flow Cytometry. *Cytometry. A* **2019**, 95 (1), 70–79. <https://doi.org/10.1002/cyto.a.23606>.
148. Li, W.; Houston, K. D.; Houston, J. P. Shifts in the Fluorescence Lifetime of EGFP during Bacterial Phagocytosis Measured by Phase-Sensitive Flow Cytometry. *Sci. Rep.* **2017**, 7, 40341. <https://doi.org/10.1038/srep40341>.
149. Sands, B.; Jenkins, P.; Peria, W. J.; Naivar, M.; Houston, J. P.; Brent, R. Measuring and Sorting Cell Populations Expressing Isospectral Fluorescent Proteins with Different Fluorescence Lifetimes. *PLoS One* **2014**, 9 (10), e109940. <https://doi.org/10.1371/journal.pone.0109940>.
150. Skilitsi, A. I.; Turko, T.; Cianfarani, D.; Barre, S.; Uhring, W.; Hassiepen, U.; Léonard, J. Towards Sensitive, High-Throughput, Biomolecular Assays Based on Fluorescence Lifetime. *Methods Appl. Fluoresc.* **2017**, 5 (3), 34002. <https://doi.org/10.1088/2050-6120/aa7f66>.
151. Baret, J. C.; Miller, O. J.; Taly, V.; Ryckelynck, M.; El-Harrak, A.; Frenz, L.; Rick, C.; Samuels, M. L.; Hutchison, J. B.; Agresti, J. J.; Link, D. R.; Weitz, D. A.; Griffiths, A. D. Fluorescence-Activated Droplet Sorting (FADS): Efficient Microfluidic Cell Sorting Based on Enzymatic Activity. *Lab Chip* **2009**, 9 (13), 1850–1858. <https://doi.org/10.1039/b902504a>.
152. Fallah-Araghi, A.; Baret, J. C.; Ryckelynck, M.; Griffiths, A. D. A Completely in Vitro Ultrahigh-Throughput Droplet-Based Microfluidic Screening System for Protein Engineering and Directed Evolution. *Lab Chip* **2012**, 12 (5), 882–891. <https://doi.org/10.1039/c2lc21035e>.

- 
153. Mazutis, L.; Gilbert, J.; Ung, W. L.; Weitz, D. A.; Griffiths, A. D.; Heyman, J. A. Single-Cell Analysis and Sorting Using Droplet-Based Microfluidics. *Nat. Protoc.* **2013**, *8* (5), 870–891. <https://doi.org/10.1038/nprot.2013.046>.
154. Churski, K.; Kaminski, T. S.; Jakiela, S.; Kamysz, W.; Baranska-Rybak, W.; Weibel, D. B.; Garstecki, P. Rapid Screening of Antibiotic Toxicity in an Automated Microdroplet System. *Lab Chip* **2012**, *12* (9), 1629–1637. <https://doi.org/10.1039/c2lc21284f>.
155. Liu, X.; Painter, R. E.; Enesa, K.; Holmes, D.; Whyte, G.; Garlisi, C. G.; Monsma, F. J.; Rehak, M.; Craig, F. F.; Smith, C. A. High-Throughput Screening of Antibiotic-Resistant Bacteria in Picodroplets. *Lab Chip* **2016**, *16* (9), 1636–1643. <https://doi.org/10.1039/C6LC00180G>.
156. Kintses, B.; Hein, C.; Mohamed, M. F.; Fischlechner, M.; Courtois, F.; Lainé, C.; Hollfelder, F. Picoliter Cell Lysate Assays in Microfluidic Droplet Compartments for Directed Enzyme Evolution. *Chem. Biol.* **2012**, *19* (8), 1001–1009. <https://doi.org/10.1016/j.chembiol.2012.06.009>.
157. Kintses, B.; Hein, C.; Mohamed, M. F.; Fischlechner, M.; Courtois, F.; Lainé, C.; Hollfelder, F. Picoliter Cell Lysate Assays in Microfluidic Droplet Compartments for Directed Enzyme Evolution. *Chem. Biol.* **2012**, *19* (8), 1001–1009. <https://doi.org/10.1016/j.chembiol.2012.06.009>.
158. Fiedler, B. L.; Van Buskirk, S.; Carter, K. P.; Qin, Y.; Carpenter, M. C.; Palmer, A. E.; Jimenez, R. Droplet Microfluidic Flow Cytometer for Sorting on Transient Cellular Responses of Genetically-Encoded Sensors. *Anal. Chem.* **2017**, *89* (1), 711–719. <https://doi.org/10.1021/acs.analchem.6b03235>.
159. Papalexi, E.; Satija, R. Single-Cell RNA Sequencing to Explore Immune Cell Heterogeneity. *Nat. Rev. Immunol.* **2018**, *18* (1), 35–45. <https://doi.org/10.1038/nri.2017.76>.

- 
160. Chokkalingam, V.; Tel, J.; Wimmers, F.; Liu, X.; Semenov, S.; Thiele, J.; Figdor, C. G.; Huck, W. T. S. Probing Cellular Heterogeneity in Cytokine-Secreting Immune Cells Using Droplet-Based Microfluidics. *Lab Chip* **2013**, *13* (24), 4740–4744. <https://doi.org/10.1039/c3lc50945a>.
161. Moon, S. J.; Ceyhan, E.; Gurkan, U. A.; Demirci, U. Statistical Modeling of Single Target Cell Encapsulation. *PLoS One* **2011**, *6* (7), e21580. <https://doi.org/10.1371/journal.pone.0021580>.
162. Collins, D. J.; Neild, A.; deMello, A.; Liu, A. Q.; Ai, Y. The Poisson Distribution and beyond: Methods for Microfluidic Droplet Production and Single Cell Encapsulation. *Lab Chip* **2015**, *15* (17), 3439–3459. <https://doi.org/10.1039/c5lc00614g>.
163. Sciambi, A.; Abate, A. R. Accurate Microfluidic Sorting of Droplets at 30 KHz. *Lab Chip* **2015**, *15* (1), 47–51. <https://doi.org/10.1039/c4lc01194e>.
164. Hasan, S.; Geissler, D.; Wink, K.; Hagen, A.; Heiland, J. J.; Belder, D. Fluorescence Lifetime-Activated Droplet Sorting in Microfluidic Chip Systems. *Lab Chip* **2019**, *19* (3), 403–409. <https://doi.org/10.1039/c8lc01278d>.
165. Wu, T. J.; Tzeng, Y. K.; Chang, W. W.; Cheng, C. A.; Kuo, Y.; Chien, C. H.; Chang, H. C.; Yu, J. Tracking the Engraftment and Regenerative Capabilities of Transplanted Lung Stem Cells Using Fluorescent Nanodiamonds. *Nat. Nanotechnol.* **2013**, *8* (9), 682–689. <https://doi.org/10.1038/nnano.2013.147>.
166. Lando, D.; Basu, S.; Stevens, T. J.; Riddell, A.; Wohlfahrt, K. J.; Cao, Y.; Boucher, W.; Leeb, M.; Atkinson, L. P.; Lee, S. F.; Hendrich, B.; Klenerman, D.; Laue, E. D. Combining Fluorescence Imaging with Hi-C to Study 3D Genome Architecture of the Same Single Cell. *Nat. Protoc.* **2018**, *13* (5), 1034–1061. <https://doi.org/10.1038/nprot.2018.017>.

- 
167. Braselmann, E.; Wierzba, A. J.; Polaski, J. T.; Chromiński, M.; Holmes, Z. E.; Hung, S. T.; Batan, D.; Wheeler, J. R.; Parker, R.; Jimenez, R.; Gryko, D.; Batey, R. T.; Palmer, A. E. A Multicolor Riboswitch-Based Platform for Imaging of RNA in Live Mammalian Cells. *Nat. Chem. Biol.* **2018**, *14* (10), 964–971. <https://doi.org/10.1038/s41589-018-0103-7>.
168. Kang, D. K.; Ali, M. M.; Zhang, K.; Huang, S. S.; Peterson, E.; Digman, M. A.; Gratton, E.; Zhao, W. Rapid Detection of Single Bacteria in Unprocessed Blood Using Integrated Comprehensive Droplet Digital Detection. *Nat. Commun.* **2014**, *5*, 5427. <https://doi.org/10.1038/ncomms6427>.
169. Ou, C. Y.; Vu, T.; Grunwald, J. T.; Toledano, M.; Zimak, J.; Toosky, M.; Shen, B.; Zell, J. A.; Gratton, E.; Abram, T. J.; Zhao, W. An Ultrasensitive Test for Profiling Circulating Tumor DNA Using Integrated Comprehensive Droplet Digital Detection. *Lab Chip* **2019**, *19* (6), 993–1005. <https://doi.org/10.1039/c8lc01399c>.
170. Hedde, P. N.; Abram, T.; Vu, T.; Zhao, W.; Gratton, E. Fluorescence Lifetime Detection with Particle Counting Devices. *Biomed. Opt. Express* **2019**, *10* (3), 1223. <https://doi.org/10.1364/boe.10.001223>.
171. Turk-MacLeod, R.; Henson, A.; Rodriguez-Garcia, M.; Gibson, G. M.; Camarasa, G. A.; Caramelli, D.; Padgett, M. J.; Cronin, L. Approach to Classify, Separate, and Enrich Objects in Groups Using Ensemble Sorting. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (22), 5681–5685. <https://doi.org/10.1073/pnas.1721929115>.
172. Myung, J. H.; Hong, S. Microfluidic Devices to Enrich and Isolate Circulating Tumor Cells. *Lab Chip* **2015**, *15* (24), 4500–4511. <https://doi.org/10.1039/c5lc00947b>.
173. Lieske, T.; Uhring, W.; Dumas, N.; Skilitski, A. I.; Léonard, J.; Fey, D. Embedded Fluorescence Lifetime Determination for High-Throughput, Low-Photon-Number

- 
- Applications. *J. Signal Process. Syst.* **2019**, 91 (7), 819–831.  
<https://doi.org/10.1007/s11265-018-1372-9>.
174. Shen, Y.; Chen, Y.; Wu, J.; Shaner, N. C.; Campbell, R. E. Engineering of mCherry Variants with Long Stokes Shift, Red-Shifted Fluorescence, and Low Cytotoxicity. *PLoS One* **2017**, 12 (2). <https://doi.org/10.1371/journal.pone.0171257>.
175. Nov, Y. When Second Best Is Good Enough: Another Probabilistic Look at Saturation Mutagenesis. *Appl. Environ. Microbiol.* **2012**, 78 (1), 258–262.  
<https://doi.org/10.1128/AEM.06265-11>.
176. Piston, D. W. & Kremers, G. J. Fluorescent protein FRET: the good, the bad and the ugly. *Trends in Biochemical Sciences.* **2007**, 32(9), p407–414.  
<https://doi.org/10.1016/j.tibs.2007.08.003>
177. York, E. M., Weilingner, N. L., LeDue, J. M., & MacVicar, B. A. Green fluorescent protein emission obscures metabolic fluorescent lifetime imaging of NAD(P)H. *Biomedical Optics Express.* 2019,10(9), p4381–4394. <https://doi.org/10.1364/boe.10.004381>
178. Levchenko, S. M., Pliss, A., & Qu, J. Fluorescence lifetime imaging of fluorescent proteins as an effective quantitative tool for noninvasive study of intracellular processes. *Journal of Innovative Optical Health Sciences.* 2018, 11(1), p1730009(1–9).  
<https://doi.org/10.1142/S1793545817300099>
179. Carter, K. P., Young, A. M., & Palmer, A. E. Fluorescent sensors for measuring metal ions in living systems. *Chemical Reviews.* 2014, 114(8), p4564–4601.  
<https://doi.org/10.1021/cr400546e>
180. Hell, S. W. Toward fluorescence nanoscopy. *Nature Biotechnology.* 2003, 21, p1347–1355.  
<https://doi.org/10.1038/nbt895>



- 
181. Rainey, K. H., & Patterson, G. H. Photoswitching FRET to monitor protein–protein interactions. *Proceedings of the National Academy of Sciences of the United States of America*. 2019, 116(3), p864-873. <https://doi.org/10.1073/pnas.1805333116>
182. Pennacchietti, F., Serebrovskaya, E. O., Faro, A. R., Shemyakina, I. I., Bozhanova, N. G., Kotlobay, A. A., Gurskaya, N. G., Bodén, A., Dreier, J., Chudakov, D. M., Lukyanov, K. A., Verkhusha, V. V., Mishin, A. S., & Testa, I. Fast reversibly photoswitching red fluorescent proteins for live-cell RESOLFT nanoscopy. *Nature Methods*. 2018, 15(8), p601-604. <https://doi.org/10.1038/s41592-018-0052-9>
183. Sarkisyan, K. S., Goryashchenko, A. S., Lidsky, P. V., Gorbachev, D. A., Bozhanova, N.G., Gorokhovatsky, A. Y., Pereverzeva, A. R., Ryumina, A. P., Zherdeva, V. V., Savitsky, A. P., Solntsev, K. M., Bommarius, A. S., Sharonov, G. V., Lindquist, J. R., Drobizhev, M., Hughes, T. E., Rebane, A., Lukyanov, K. A., & Mishin, A. S. Green Fluorescent Protein with Anionic Tryptophan-Based Chromophore and Long Fluorescence Lifetime. *Biophysical Journal*. 2015, 109(2), p380-389. <https://doi.org/10.1016/j.bpj.2015.06.018>
184. Manas, E. S., & Spano, F. C. Absorption and spontaneous emission in aggregates of conjugated polymers. *Journal of Chemical Physics*. 1998, 109(18), p8087-8101. <https://doi.org/10.1063/1.477457>
185. Piatkevich, K. D., Efremenko, E. N., Verkhusha, V. V., & Varfolomeev, S. D. Red fluorescent proteins and their properties. *Russian Chemical Reviews*. 2010, 79(3), p243-258. <https://doi.org/10.1070/rc2010v079n03abeh004095>
186. Miyawaki, A., Shcherbakova, D. M., & Verkhusha, V. V. Red fluorescent proteins: Chromophore formation and cellular applications. *Current Opinion in Structural Biology*. 2012, 22(5), p679-688. <https://doi.org/10.1016/j.sbi.2012.09.002>

- 
187. Dunsing, V., Luckner, M., Zühlke, B., Petazzi, R. A., Herrmann, A., & Chiantia, S. Optimal fluorescent protein tags for quantifying protein oligomerization in living cells. *Scientific Reports*. 2018, 8, 10634 (p1-12). <https://doi.org/10.1038/s41598-018-28858-0>
188. Shemiakina, I. I., Ermakova, G. V., Cranfill, P. J., Baird, M. A., Evans, R. A., Souslova, E. A., Staroverov, D. B., Gorokhovatsky, A. Y., Putintseva, E. V., Gorodnicheva, T. V., Chepurnykh, T. V., Strukova, L., Lukyanov, S., Zeraisky, A. G., Davidson, M. W., Chudakov, D. M., & Shcherbo, D. A monomeric red fluorescent protein with low cytotoxicity. *Nature Communications*. 2012, 3, 1204 (p1-7). <https://doi.org/10.1038/ncomms2208>
189. Kametsky, L., Jones, T. R., Fraser, A., Bray, M. A., Logan, D. J., Madden, K. L., Ljosa, V., Rueden, C., Eliceiri, K. W., & Carpenter, A. E. Improved structure, function and compatibility for cellprofiler: Modular high-throughput image analysis software. *Bioinformatics*. 2011, 27(8), p1179-1180. <https://doi.org/10.1093/bioinformatics/btr095s>
190. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. Fiji: An open-source platform for biological-image analysis. *Nature Methods*. 2012, 9, p676-682. <https://doi.org/10.1038/nmeth.2019>
191. Ansari, A. M., Ahmed, A. K., Matsangos, A. E., Lay, F., Born, L. J., Marti, G., Harmon, J. W., & Sun, Z. Cellular GFP Toxicity and Immunogenicity: Potential Confounders in in Vivo Cell Tracking Experiments. *Stem Cell Reviews and Reports*. 2016, 12(5), p553-559. <https://doi.org/10.1007/s12015-016-9670-8>
192. Liu, B., Mavrova, S. N., Van Den Berg, J., Kristensen, S. K., Mantovanelli, L., Veenhoff, L. M., Poolman, B., & Boersma, A. J. Influence of Fluorescent Protein Maturation on FRET Measurements in Living Cells. *ACS Sensors*. 2018, 3(9), p1735-1742. <https://doi.org/10.1021/acssensors.8b00473>

- 
193. DeLano, W. L. The PyMOL Molecular Graphics System, Version 1.1. Schrödinger LLC. <https://doi.org/10.1038/hr.2014.17>
194. Martynov, V. I., Savitsky, A. P., Martynova, N. Y., Savitsky, P. A., Lukyanov, K. A., & Lukyanov, S. A. Alternative cyclization in GFP-like proteins family. The formation and structure of the chromophore of a purple chromoprotein from *Anemonia sulcata*. *Journal of Biological Chemistry*. 2001, 276(24), p21012-21016. <https://doi.org/10.1074/jbc.M100500200>
195. Tretyakova, Y. A., Pakhomov, A. A., & Martynov, V. I. Chromophore structure of the kindling fluorescent protein asFP595 from *Anemonia sulcata*. *Journal of the American Chemical Society*. 2007, 129(25), p7748-7749. <https://doi.org/10.1021/ja071992c>
196. Costantini, L. M., Baloban, M., Markwardt, M. L., Rizzo, M., Guo, F., Verkhusha, V. V., & Snapp, E. L. A palette of fluorescent proteins optimized for diverse cellular environments. *Nature Communications*. 2015, 6, 7670 (p1-13). <https://doi.org/10.1038/ncomms8670>
197. Kruitwagen, T., Denoth-Lippuner, A., Wilkins, B. J., Neumann, H., & Barral, Y. Axial contraction and short-range compaction of chromatin synergistically promote mitotic chromosome condensation. *ELife*. 2015, 4, e10396 (p1-19). <https://doi.org/10.7554/eLife.10396>
198. Kimura, S., Noda, T., & Yoshimori, T. Dissection of the autophagosome maturation process by a novel reporter protein, tandem fluorescent tagged LC3. *Autophagy*. 2007, 3(5), p452-460. <https://doi.org/10.4161/auto.4451>
199. Costantini, L. M., Fossati, M., Francolini, M., & Snapp, E. L. Assessing the Tendency of Fluorescent Proteins to Oligomerize Under Physiologic Conditions. *Traffic*. 2012, 13(5), p643-649. <https://doi.org/10.1111/j.1600-0854.2012.01336.x>

- 
200. Muslinkina L., Pletnev V.Z., Pletneva N.V., Ruchkin D.A., Kolesov D.V., Bogdanov A.M., Kost L., Rakitina T.V., Agapova Yu. K., Shemyakina I.I., Chudakov D.M., and Pletnev S. Two Independent Routes of Post-Translational Chemistry in Fluorescent Protein FusionRed. *International Journal of Biological Macromolecules*. 2020, 155, p551-559. <https://doi.org/10.1016/j.ijbiomac.2020.03.244>
201. Pletneva, N. V., Pletnev, V. Z., Shemiakina, I. I., Chudakov, D. M., Artemyev, I., Wlodawer, A., Dauter, Z., & Pletnev, S. Crystallographic study of red fluorescent protein eqFP578 and its far-red variant Katushka reveals opposite pH-induced isomerization of chromophore. *Protein Science*. 2011, 20(7), p165-1274. <https://doi.org/10.1002/pro.654>
202. Stiel, A. C., Andresen, M., Bock, H., Hilbert, M., Schilde, J., Schönle, A., Eggeling, C., Egner, A., Hell, S. W., & Jakobs, S. Generation of monomeric reversibly switchable red fluorescent proteins for far field fluorescence nanoscopy. *Biophysical Journal*. 2008, 95(6), p2989-2997. <https://doi.org/10.1529/biophysj.108.130146>
203. Lovell, S. C., Word, J. M., Richardson, J. S., & Richardson, D. C. Asparagine and glutamine rotamers: B-factor cutoff and correction of amide flips yield distinct clustering. *Proceedings of the National Academy of Sciences of the United States of America*. 1999, 96(2), p400-405. <https://doi.org/10.1073/pnas.96.2.400>
204. Elliott, A. D. Confocal Microscopy: Principles and Modern Practices. *Curr. Protoc. Cytom.* **2020**, 92 (1), e68. <https://doi.org/10.1002/cpcy.68>.
205. Ahmed, A.; Schoberer, J.; Cooke, E.; Botchway, S. W. Multicolor FRET-FLIM Microscopy to Analyze Multiprotein Interactions in Live Cells. In *Methods in Molecular Biology*; Poterszman, A., Ed.; Springer US: New York, NY, 2021; Vol. 2247, pp 287–301. [https://doi.org/10.1007/978-1-0716-1126-5\\_16](https://doi.org/10.1007/978-1-0716-1126-5_16).
206. Schouw, H. M.; Huisman, L. A.; Janssen, Y. F.; Slart, R. H. J. A.; Borra, R. J. H.; Willemsen, A. T. M.; Brouwers, A. H.; van Dijl, J. M.; Dierckx, R. A.; van Dam, G. M.;

- 
- Szymanski, W.; Boersma, H. H.; Kruijff, S. Targeted Optical Fluorescence Imaging: A Meta-Narrative Review and Future Perspectives. *Eur. J. Nucl. Med. Mol. Imaging* **2021**, *48* (13), 4272–4292. <https://doi.org/10.1007/s00259-021-05504-y>.
207. Datta, R.; Heaster, T. M.; Sharick, J. T.; Gillette, A. A.; Skala, M. C. Fluorescence Lifetime Imaging Microscopy: Fundamentals and Advances in Instrumentation, Analysis, and Applications. *J. Biomed. Opt.* **2020**, *25* (07), 1. <https://doi.org/10.1117/1.jbo.25.7.071203>.
208. Breuer, G. A Formal Representation of Abbe's Theory of Microscopic Image Formation. *Opt. Acta (Lond)*. **1984**, *31* (6), 661–670. <https://doi.org/10.1080/713821564>.
209. Chojnacki, J.; Eggeling, C. Super-Resolution Fluorescence Microscopy Studies of Human Immunodeficiency Virus. *Retrovirology* **2018**, *15* (1), 41. <https://doi.org/10.1186/s12977-018-0424-3>.
210. Vangindertael, J.; Camacho, R.; Sempels, W.; Mizuno, H.; Dedecker, P.; Janssen, K. P. F. An Introduction to Optical Super-Resolution Microscopy for the Adventurous Biologist. *Methods Appl. Fluoresc.* **2018**, *6* (2), 22003. <https://doi.org/10.1088/2050-6120/aaae0c>.
211. Mishin, A. S.; Lukyanov, K. A. Live-Cell Super-Resolution Fluorescence Microscopy. *Biochem.* **2019**, *84* (Suppl 1), 19–31. <https://doi.org/10.1134/S0006297919140025>.
212. Schermelleh, L.; Ferrand, A.; Huser, T.; Eggeling, C.; Sauer, M.; Biehlmaier, O.; Drummen, G. P. C. Super-Resolution Microscopy Demystified. *Nat. Cell Biol.* **2019**, *21* (1), 72–84. <https://doi.org/10.1038/s41556-018-0251-8>.
213. Chen, L.; Chen, X.; Yang, X.; He, C.; Wang, M.; Xi, P.; Gao, J. Advances of Super-Resolution Fluorescence Polarization Microscopy and Its Applications in Life Sciences. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 2209–2216. <https://doi.org/10.1016/j.csbj.2020.06.038>.

- 
214. Schubert, V.; Neumann, P.; Marques, A.; Heckmann, S.; Macas, J.; Pedrosa-Harand, A.; Schubert, I.; Jang, T.-S.; Houben, A. Super-Resolution Microscopy Reveals Diversity of Plant Centromere Architecture. *Int. J. Mol. Sci.* **2020**, *21* (10). <https://doi.org/10.3390/ijms21103488>.
215. Hell, S. W.; Wichmann, J. Breaking the Diffraction Resolution Limit by Stimulated Emission: Stimulated-Emission-Depletion Fluorescence Microscopy. *Opt. Lett.* **1994**, *19* (11), 780. <https://doi.org/10.1364/ol.19.000780>.
216. Betzig, E.; Patterson, G. H.; Sougrat, R.; Lindwasser, O. W.; Olenych, S.; Bonifacino, J. S.; Davidson, M. W.; Lippincott-Schwartz, J.; Hess, H. F. Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science (80-.)*. **2006**, *313* (5793), 1642–1645. <https://doi.org/10.1126/science.1127344>.
217. Moerner, W. E. Microscopy beyond the Diffraction Limit Using Actively Controlled Single Molecules. *J. Microsc.* **2012**, *246* (3), 213–220. <https://doi.org/10.1111/j.1365-2818.2012.03600.x>.
218. Möckl, L.; Moerner, W. E. Super-Resolution Microscopy with Single Molecules in Biology and Beyond-Essentials, Current Trends, and Future Challenges. *J. Am. Chem. Soc.* **2020**, *142* (42), 17828–17844. <https://doi.org/10.1021/jacs.0c08178>.
219. Dertinger, T.; Colyera, R.; Iyer, G.; Weiss, S.; Enderlein, J. Fast, Background-Free, 3D Super-Resolution Optical Fluctuation Imaging (SOFI). *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (52), 22287–22292. <https://doi.org/10.1073/pnas.0907866106>.
220. Simonson, P. D.; Rothenberg, E.; Selvin, P. R. Single-Molecule-Based Super-Resolution Images in the Presence of Multiple Fluorophores. *Nano Lett.* **2011**, *11* (11), 5090–5096. <https://doi.org/10.1021/nl203560r>.
221. Burnette, D. T.; Sengupta, P.; Dai, Y.; Lippincott-Schwartz, J.; Kachar, B. Bleaching/Blinking Assisted Localization Microscopy for Superresolution Imaging Using

- 
- Standard Fluorescent Molecules. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (52), 21081–21086. <https://doi.org/10.1073/pnas.1117430109>.
222. Dahlberg, P. D.; Saurabh, S.; Sartor, A. M.; Wang, J.; Wang, J.; Mitchell, P. G.; Chiu, W.; Chiu, W.; Shapiro, L.; Moerner, W. E. Cryogenic Single-Molecule Fluorescence Annotations for Electron Tomography Reveal in Situ Organization of Key Proteins in *Caulobacter*. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (25), 13937–13944. <https://doi.org/10.1073/pnas.2001849117>.
223. Lelek, M.; Gyparaki, M. T.; Beliu, G.; Schueder, F.; Griffié, J.; Manley, S.; Jungmann, R.; Sauer, M.; Lakadamyali, M.; Zimmer, C. Single-Molecule Localization Microscopy. *Nat. Rev. Methods Prim.* **2021**, *1* (1), 39. <https://doi.org/10.1038/s43586-021-00038-x>.
224. Vandenberg, W.; Leutenegger, M.; Duwé, S.; Dedecker, P. An Extended Quantitative Model for Super-Resolution Optical Fluctuation Imaging (SOFI). *Opt. Express* **2019**, *27* (18), 25749. <https://doi.org/10.1364/oe.27.025749>.
225. van den Eynde, R.; Sandmeyer, A.; Vandenberg, W.; Duwé, S.; Hübner, W.; Huser, T.; Dedecker, P.; Müller, M. Quantitative Comparison of Camera Technologies for Cost-Effective Super-Resolution Optical Fluctuation Imaging (SOFI). *JPhys Photonics* **2019**, *1* (4), 44001. <https://doi.org/10.1088/2515-7647/ab36ae>.
226. Moeyaert, B.; Vandenberg, W.; Dedecker, P. SOFIevaluator: A Strategy for the Quantitative Quality Assessment of SOFI Data. *Biomed. Opt. Express* **2020**, *11* (2), 636. <https://doi.org/10.1364/boe.382278>.
227. Dickson, R. M.; Cubitt, A. B.; Tsien, R. Y.; Moerner, W. E. On/off Blinking and Switching Behaviour of Single Molecules of Green Fluorescent Protein. *Nature* **1997**, *388* (6640), 355–358. <https://doi.org/10.1038/41048>.
228. Garcia-Parajo, M. F.; Segers-Nolten, G. M. J.; Veerman, J. A.; Greve, J.; Van Hulst, N. F. Real-Time Light-Driven Dynamics of the Fluorescence Emission in Single Green

- 
- Fluorescent Protein Molecules. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, 97 (13), 7237–7242. <https://doi.org/10.1073/pnas.97.13.7237>.
229. Schwille, P.; Kummer, S.; Heikal, A. A.; Moerner, W. E.; Webb, W. W. Fluorescence Correlation Spectroscopy Reveals Fast Optical Excitation-Driven Intramolecular Dynamics of Yellow Fluorescent Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, 97 (1), 151–156. <https://doi.org/10.1073/pnas.97.1.151>.
230. Malvezzi-Campeggi, F.; Jahnz, M.; Heinze, K. G.; Dittrich, P.; Schwille, P. Light-Induced Flickering of DsRed Provides Evidence for Distinct and Interconvertible Fluorescent States. *Biophys. J.* **2001**, 81 (3), 1776–1785. [https://doi.org/10.1016/S0006-3495\(01\)75828-6](https://doi.org/10.1016/S0006-3495(01)75828-6).
231. Garcia-Parajo, M. F.; Koopman, M.; Van Dijk, E. M. H. P.; Subramaniam, V.; Van Hulst, N. F. The Nature of Fluorescence Emission in the Red Fluorescent Protein DsRed, Revealed by Single-Molecule Detection. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98 (25), 14392–14397. <https://doi.org/10.1073/pnas.251525598>.
232. Haupts, U.; Maiti, S.; Schwille, P.; Webb, W. W. Dynamics of Fluorescence Fluctuations in Green Fluorescent Protein Observed by Fluorescence Correlation Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, 95 (23), 13573–13578. <https://doi.org/10.1073/pnas.95.23.13573>.
233. Heikal, A. A.; Hess, S. T.; Baird, G. S.; Tsien, R. Y.; Webb, W. W. Molecular Spectroscopy and Dynamics of Intrinsically Fluorescent Proteins: Coral Red (DsRed) and Yellow (Citrine). *Proc. Natl. Acad. Sci. U. S. A.* **2000**, 97 (22), 11996–12001. <https://doi.org/10.1073/pnas.97.22.11996>.
234. Sinnecker, D.; Voigt, P.; Hellwig, N.; Schaefer, M. Reversible Photobleaching of Enhanced Green Fluorescent Proteins. *Biochemistry* **2005**, 44 (18), 7085–7097. <https://doi.org/10.1021/bi047881x>.



- 
235. Berardozi, R.; Adam, V.; Martins, A.; Bourgeois, D. Arginine 66 Controls Dark-State Formation in Green-to-Red Photoconvertible Fluorescent Proteins. *J. Am. Chem. Soc.* **2016**, *138* (2), 558–565. <https://doi.org/10.1021/jacs.5b09923>.
236. Thédié, D.; Berardozi, R.; Adam, V.; Bourgeois, D. Photoswitching of Green MEos2 by Intense 561 Nm Light Perturbs Efficient Green-to-Red Photoconversion in Localization Microscopy. *J. Phys. Chem. Lett.* **2017**, *8* (18), 4424–4430. <https://doi.org/10.1021/acs.jpcllett.7b01701>.
237. De Zitter, E.; Ridard, J.; Thédié, D.; Adam, V.; Lévy, B.; Byrdin, M.; Gotthard, G.; Van Meervelt, L.; Dedeker, P.; Demachy, I.; Bourgeois, D. Mechanistic Investigations of Green MEos4b Reveal a Dynamic Long-Lived Dark State. *J. Am. Chem. Soc.* **2020**, *142* (25), 10978–10988. <https://doi.org/10.1021/jacs.0c01880>
238. Krueger, T. D.; Tang, L.; Zhu, L.; Breen, I. L.; Wachter, R. M.; Fang, C. Dual Illumination Enhances Transformation of an Engineered Green-to-Red Photoconvertible Fluorescent Protein. *Angew. Chemie - Int. Ed.* **2020**, *59* (4), 1644–1652. <https://doi.org/10.1002/anie.201911379>.
239. Ando, R.; Mizuno, H.; Miyawaki, A. Regulated Fast Nucleocytoplasmic Shuttling Observed by Reversible Protein Highlighting. *Science* (80-. ). **2004**, *306* (5700), 1370–1373. <https://doi.org/10.1126/science.1102506>.
240. Grotjohann, T.; Testa, I.; Leutenegger, M.; Bock, H.; Urban, N. T.; Lavoie-Cardinal, F.; Willig, K. I.; Eggeling, C.; Jakobs, S.; Hell, S. W. Diffraction-Unlimited All-Optical Imaging and Writing with a Photochromic GFP. *Nature* **2011**, *478* (7368), 204–208. <https://doi.org/10.1038/nature10497>.
241. Donnert, G.; Eggeling, C.; Hell, S. W. Major Signal Increase in Fluorescence Microscopy through Dark-State Relaxation. *Nat. Methods* **2007**, *4* (1), 81–86. <https://doi.org/10.1038/nmeth986>.

- 
242. Mahoney, D. P.; Owens, E. A.; Fan, C.; Hsiang, J. C.; Henary, M. M.; Dickson, R. M. Tailoring Cyanine Dark States for Improved Optically Modulated Fluorescence Recovery. *J. Phys. Chem. B* **2015**, *119* (13), 4637–4643. <https://doi.org/10.1021/acs.jpccb.5b00777>.
243. Jablonski, A. E.; Hsiang, J. C.; Bagchi, P.; Hull, N.; Richards, C. I.; Fahrni, C. J.; Dickson, R. M. Signal Discrimination between Fluorescent Proteins in Live Cells by Long-Wavelength Optical Modulation. *J. Phys. Chem. Lett.* **2012**, *3* (23), 3585–3591. <https://doi.org/10.1021/jz3016414>.
244. Chen, Y. C.; Jablonski, A. E.; Issaeva, I.; Bourassa, D.; Hsiang, J. C.; Fahrni, C. J.; Dickson, R. M. Optically Modulated Photoswitchable Fluorescent Proteins Yield Improved Biological Imaging Sensitivity. *J. Am. Chem. Soc.* **2015**, *137* (40), 12764–12767. <https://doi.org/10.1021/jacs.5b07871>.
245. Fleischer, B. C.; Petty, J. T.; Hsiang, J. C.; Dickson, R. M. Optically Activated Delayed Fluorescence. *J. Phys. Chem. Lett.* **2017**, *8* (15), 3536–3543. <https://doi.org/10.1021/acs.jpcclett.7b01215>.
246. Mahoney, D. P.; Demissie, A. A.; Dickson, R. M. Optically Activated Delayed Fluorescence through Control of Cyanine Dye Photophysics. *J. Phys. Chem. A* **2019**, *123* (16), 3599–3606. <https://doi.org/10.1021/acs.jpca.9b01333>.
247. Peng, B.; Dikdan, R.; Hill, S. E.; Patterson-Orazem, A. C.; Lieberman, R. L.; Fahrni, C. J.; Dickson, R. M. Optically Modulated and Optically Activated Delayed Fluorescent Proteins through Dark State Engineering. *J. Phys. Chem. B* **2021**, *125* (20), 5200–5209. <https://doi.org/10.1021/acs.jpccb.1c00649>.
248. Waldchen, S.; Lehmann, J.; Klein, T.; Van De Linde, S.; Sauer, M. Light-Induced Cell Damage in Live-Cell Super-Resolution Microscopy. *Sci. Rep.* **2015**, *5* (1), 15348. <https://doi.org/10.1038/srep15348>.

- 
249. Laissue, P. P.; Alghamdi, R. A.; Tomancak, P.; Reynaud, E. G.; Shroff, H. Assessing Phototoxicity in Live Fluorescence Imaging. *Nat. Methods* 2017, 14 (7), 657–661. <https://doi.org/10.1038/nmeth.4344>.
250. Mukherjee, S.; Jimenez, R. Photophysical Engineering of Fluorescent Proteins: Accomplishments and Challenges of Physical Chemistry Strategies. *J. Phys. Chem. B* 2022, 126 (4), 735–750. <https://doi.org/10.1021/acs.jpcb.1c05629>.
251. Schenk, A.; Ivanchenko, S.; Röcker, C.; Wiedenmann, J.; Nienhaus, G. U. Photodynamics of Red Fluorescent Proteins Studied by Fluorescence Correlation Spectroscopy. *Biophys. J.* **2004**, 86 (1 I), 384–394. [https://doi.org/10.1016/S0006-3495\(04\)74114-4](https://doi.org/10.1016/S0006-3495(04)74114-4).
252. Weber, W.; Helm, V.; Mccammon, J. A.; Langhoff, P. W. Shedding Light on the Dark and Weakly Fluorescent States of Green Fluorescent Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, 96 (11), 6177–6182. <https://doi.org/10.1073/pnas.96.11.6177>.
253. Klementieva, N. V.; Pavlikov, A. I.; Moiseev, A. A.; Bozhanova, N. G.; Mishina, N. M.; Lukyanov, S. A.; Zagaynova, E. V.; Lukyanov, K. A.; Mishin, A. S. Intrinsic Blinking of Red Fluorescent Proteins for Super-Resolution Microscopy. *Chem. Commun.* 2017, 53 (5), 949–951. <https://doi.org/10.1039/c6cc09200d>.
254. Honda, M.; Saito, Y.; Smith, N. I.; Fujita, K.; Kawata, S. Nanoscale Heating of Laser Irradiated Single Gold Nanoparticles in Liquid. *Opt. Express* 2011, 19 (13), 12375. <https://doi.org/10.1364/oe.19.012375>.
255. Davis, J. L.; Dong, B.; Sun, C.; Zhang, H. F. Method to Identify and Minimize Artifacts Induced by Fluorescent Impurities in Single-Molecule Localization Microscopy. *J. Biomed. Opt.* 2018, 23 (10), 1. <https://doi.org/10.1117/1.jbo.23.10.106501>.

- 
256. Tregidgo, C.; Levitt, J. A.; Suhling, K. Effect of Refractive Index on the Fluorescence Lifetime of Green Fluorescent Protein. *J. Biomed. Opt.* **2008**, *13* (3), 031218. <https://doi.org/10.1117/1.2937212>.
257. Watkins, L. P.; Yang, H. Detection of Intensity Change Points in Time-Resolved Single-Molecule Measurements. *J. Phys. Chem. B* **2005**, *109* (1), 617–628. <https://doi.org/10.1021/jp0467548>.
258. Ng, H. F. Automatic Thresholding for Defect Detection. *Pattern Recognit. Lett.* **2006**, *27* (14), 1644–1649. <https://doi.org/10.1016/j.patrec.2006.03.009>.
259. Yip, W. T.; Hu, D.; Yu, J.; Vanden Bout, D. A.; Barbara, P. F. Classifying the Photophysical Dynamics of Single- and Multiple-Chromophoric Molecules by Single Molecule Spectroscopy. *J. Phys. Chem. A* **1998**, *102* (39), 7564–7575. <https://doi.org/10.1021/jp981808x>.
260. Peterman, E. J. G.; Brasselet, S.; Moerner, W. E. The Fluorescence Dynamics of Single Molecules of Green Fluorescent Protein. *J. Phys. Chem. A* **1999**, *103* (49), 10553–10560. <https://doi.org/10.1021/jp991968o>.
261. Lounis, B.; Deich, J.; Rosell, F. I.; Boxer, S. G.; Moerner, W. E. Photophysics of DsRed, a Red Fluorescent Protein, from the Ensemble to the Single-Molecule Level. *J. Phys. Chem. B* **2002**, *105* (21), 5048–5054. <https://doi.org/10.1021/jp010116x>.
262. Yang, H.; Li, H.; Liu, T. Photobleaching Statistics in Single-Molecule on-/off-Time Distributions. *J. Chem. Phys.* **2019**, *151* (17), 174101. <https://doi.org/10.1063/1.5126500>.
263. Qi, L.; Just, F.; Leuchs, G.; Chekhova, M. V. Autonomous Absolute Calibration of an ICCD Camera in Single-Photon Detection Regime. *Opt. Express* **2016**, *24* (23), 26444. <https://doi.org/10.1364/oe.24.026444>.

- 
264. Behrens, S. H.; Grier, D. G. The Charge of Glass and Silica Surfaces. *J. Chem. Phys.* **2001**, *115* (14), 6716–6721. <https://doi.org/10.1063/1.1404988>.
265. Wilks, J. C.; Slonczewski, J. L. PH of the Cytoplasm and Periplasm of Escherichia Coli: Rapid Measurement by Green Fluorescent Protein Fluorimetry. *J. Bacteriol.* 2007, *189* (15), 5601–5607. <https://doi.org/10.1128/JB.00615-07>.
266. Rafiq, S.; Rajbongshi, B. K.; Nair, N. N.; Sen, P.; Ramanathan, G. Excited State Relaxation Dynamics of Model Green Fluorescent Protein Chromophore Analogs: Evidence for Cis-Trans Isomerism. *J. Phys. Chem. A* **2011**, *115* (47), 13733–13742. <https://doi.org/10.1021/jp206815t>.
267. Carrascosa, E.; Bull, J. N.; Scholz, M. S.; Coughlan, N. J. A.; Olsen, S.; Wille, U.; Bieske, E. J. Reversible Photoisomerization of the Isolated Green Fluorescent Protein Chromophore. *J. Phys. Chem. Lett.* **2018**, *9* (10), 2647–2651. <https://doi.org/10.1021/acs.jpcllett.8b01201>.
268. Chang, J.; Romei, M. G.; Boxer, S. G. Structural Evidence of Photoisomerization Pathways in Fluorescent Proteins. *J. Am. Chem. Soc.* **2019**, *141* (39), 15504–15508. <https://doi.org/10.1021/jacs.9b08356>.
269. Bravaya, K. B.; Subach, O. M.; Korovina, N.; Verkhusha, V. V.; Krylov, A. I. Insight into the Common Mechanism of the Chromophore Formation in the Red Fluorescent Proteins: The Elusive Blue Intermediate Revealed. *J. Am. Chem. Soc.* 2012, *134* (5), 2807–2814. <https://doi.org/10.1021/ja2114568>.
270. List, N. H.; Jones, C. M.; Martínez, T. J. Internal Conversion of the Anionic GFP Chromophore: In and out of the I-Twisted S1/S0conical Intersection Seam. *Chem. Sci.* **2022**, *13* (2), 373–385. <https://doi.org/10.1039/d1sc05849e>.

- 
271. Romero, P. A.; Arnold, F. H. Exploring Protein Fitness Landscapes by Directed Evolution. *Nat. Rev. Mol. Cell Biol.* **2009**, *10* (12), 866–876. <https://doi.org/10.1038/nrm2805>.
272. Specht, E. A.; Braselmann, E.; Palmer, A. E. A Critical and Comparative Review of Fluorescent Tools for Live-Cell Imaging. *Annu. Rev. Physiol.* **2017**, *79* (1), 93–117. <https://doi.org/10.1146/annurev-physiol-022516-034055>.
273. Legault, S.; Fraser-Halberg, D. P.; McAnelly, R. L.; Eason, M. G.; Thompson, M. C.; Chica, R. A. Generation of Bright Monomeric Red Fluorescent Proteins via Computational Design of Enhanced Chromophore Packing. *Chem. Sci.* **2022**. <https://doi.org/10.1039/d1sc05088e>.
274. Onukwufor, J. O.; Trewin, A. J.; Baran, T. M.; Almast, A.; Foster, T. H.; Wojtovich, A. P. Quantification of Reactive Oxygen Species Production by the Red Fluorescent Proteins KillerRed, SuperNova and mCherry. *Free Radic. Biol. Med.* **2020**, *147*, 1–7. <https://doi.org/https://doi.org/10.1016/j.freeradbiomed.2019.12.008>.
275. Scaramuzzino, C.; Cuoc, E. C.; Pla, P.; Humbert, S.; Saudou, F. Calcineurin and Huntingtin Form a Calcium-Sensing Machinery That Directs Neurotrophic Signals to the Nucleus. *Sci. Adv.* **2022**, *8* (1), eabj8812. <https://doi.org/10.1126/sciadv.abj8812>.
276. Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y. Improved Monomeric Red, Orange and Yellow Fluorescent Proteins Derived from *Discosoma* Sp. Red Fluorescent Protein. *Nat. Biotechnol.* **2004**, *22* (12), 1567–1572. <https://doi.org/10.1038/nbt1037>.
277. Chica, R. A.; Moore, M. M.; Allen, B. D.; Mayo, S. L. Generation of Longer Emission Wavelength Red Fluorescent Proteins Using Computationally Designed Libraries. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (47), 20257–20262. <https://doi.org/10.1073/pnas.1013910107>.

- 
278. Jang, S. J. A Simple Generalization of the Energy Gap Law for Nonradiative Processes. *J. Chem. Phys.* **2021**, *155* (16), 164106. <https://doi.org/10.1063/5.0068868>.
279. Friedman, H. C.; Cosco, E. D.; Atallah, T. L.; Jia, S.; Sletten, E. M.; Caram, J. R. Establishing Design Principles for Emissive Organic SWIR Chromophores from Energy Gap Laws. *Chem* **2021**, *7* (12), 3359–3376. <https://doi.org/10.1016/j.chempr.2021.09.001>.
280. Mertz, E. L.; Tikhomirov, V. A.; Krishtalik, L. I. Stokes Shift as a Tool for Probing the Solvent Reorganization Energy. *J. Phys. Chem. A* **1997**, *101* (19), 3433–3442. <https://doi.org/10.1021/jp963042b>.
281. Lavis, L. D.; Raines, R. T. Bright Ideas for Chemical Biology. *ACS Chem. Biol.* **2008**, *3* (3), 142–155. <https://doi.org/10.1021/cb700248m>.
282. Shaner, N. C.; Patterson, G. H.; Davidson, M. W. Advances in Fluorescent Protein Technology. *J. Cell Sci.* **2007**, *120* (24), 4247–4260. <https://doi.org/10.1242/jcs.005801>.
283. Shkrob, M. A.; Yanushevich, Y. G.; Chudakov, D. M.; Gurskaya, N. G.; Labas, Y. A.; Poponov, S. Y.; Mudrik, N. N.; Lukyanov, S.; Lukyanov, K. A. Far-Red Fluorescent Proteins Evolved from a Blue Chromoprotein from *Actinia Equina*. *Biochem. J.* **2005**, *392* (3), 649–654. <https://doi.org/10.1042/BJ20051314>.
284. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802. <https://doi.org/10.1002/jcc.20289>.

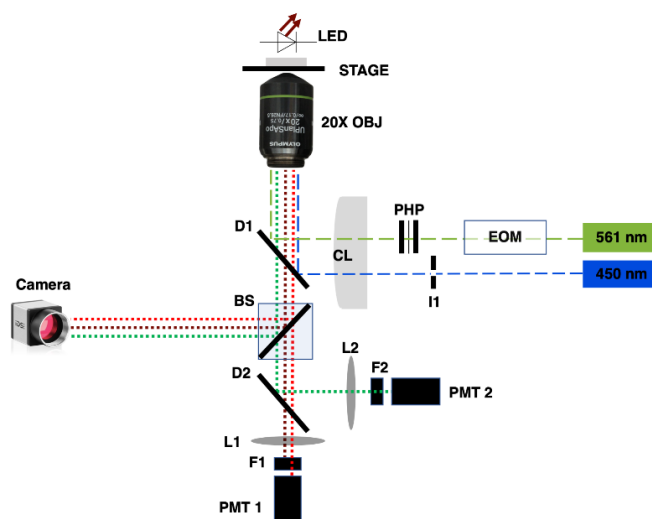
## Appendix 1

This is an addendum to Chapter 2 and parts of this appendix have been adapted from the Supplementary Information section of the article “*Photophysical Engineering of Fluorescent Proteins: Accomplishments and Challenges of Physical Chemistry Strategies.*” **Mukherjee, S.;** Jimenez, R. J. Phys. Chem. B 2022. <https://doi.org/10.1021/acs.jpcb.1c05629>

## Section 1. Optical layout and microfluidic device

### 1a. Optical Layout

The optical layout of the instrument is depicted in Figure A1.1.



**Figure A1.1. Optical diagram of the microfluidic sorter.** The excitation light sources are two CW lasers with wavelengths 561 nm (green dashed line, Coherent Genesis MX) and 450 nm (Blue dashed line, ThorLabs L450P1600MM laser diode on a TCLDM9 temperature control mount shaped by a C230TMD-A f=4.51mm NA=0.55 aspheric lens). EOM is an electro-optic modulator (ThorLabs EO-AM-NR-C4) with a modulation frequency set to 29.5 MHz. PHP are a polarizer, a half-wave plate and another polarizer. I1 is an iris with an adjustable pinhole size. CL is a cylindrical lens with a focal length 150 mm. D1 (Semrock FF493/574-Di01-25x36) and D2 (Semrock FF573-Di01- 25x36) are dichroic mirrors. The laser beams are directed into an inverted objective lens (Olympus UPLSAPO 20X NA=0.75) under a 3-dimensional translation stage where the microfluidic device is located. L1 and L2 are convex lenses, and F1 (Semrock FF01- 629/56)



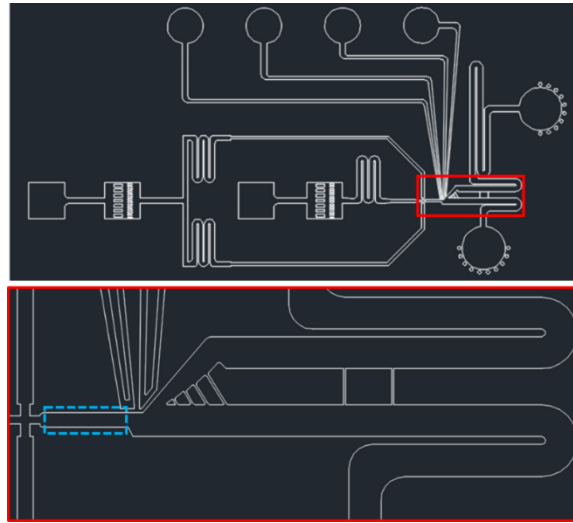
and F2 (Semrock FF01-520/35) are band pass filters. PMT1 and PMT2 are photomultiplier tubes (Hamamatsu R9880-U-20). The far-red LED (dark red dotted line, Thorlabs M730D2) with a long pass filter (Thorlabs FGL715) allows one to visualize the flow on the Camera (Edmund EO-23122M). BS is a 10:90 beam splitter (ThorLabs BSN10R). The green and red fluorescence signals are shown with dotted green and red lines, respectively.

Both 561 nm and 450 nm continuous wave (CW) laser beams are the primary excitation beams to induce fluorescence in the cells encapsulated in droplets. The 561 nm beam is focused into an electro-optic modulator that can amplitude modulate the CW beam to a sinusoidal profile. The half-wave plate and polarizers can be used to control the laser power and the polarization of the sinusoidally modulated beam. The 450 nm beam is spatially filtered using an adjustable iris (I1). Both beams are directed onto a cylindrical lens CL that reshapes the circular beams to elliptical ones. The beams are directed into a 20X inverted objective lens using a dichroic mirror D1. The epifluorescence from the sample staged on the objective passes through D1 into dichroic mirror D2 that reflects the green fluorescence towards a band pass filter F2 and focused onto a photomultiplier tube (PMT) PMT2, which is referred as the green channel. The red fluorescence that passes through D2 and band pass filter F1 is focused onto PMT1 referred to as the red channel. A far-red LED and a long pass filter are mounted above the objective lens, and the LED light is partially reflected off a beam splitter and collected onto a monochrome USB camera to provide a real-time display of the system in flow.

## **1b. Microfluidic device**

The microfluidic devices were fabricated in a clean room environment using standard soft lithography techniques. The photoresist (MicroChem SU-8 3050) was spin-coated evenly on a clean silicon wafer according to the protocol provided with the photoresist SU-8 3050 aiming for depth 50  $\mu\text{m}$ . The coated wafer was illuminated with a uniform UV light source under an acetate mask printed with an inverted image of the device design that was created using CAD software as shown in Figure A1.2. The exposed wafers were baked shortly, rinsed with the developer (MicroChem SU-8 developer), and subsequently hard baked overnight according to the protocol provided with the photoresist SU-8 3050. These wafers were exposed to (1H,1H,2H,2H-

perfluorooctyl) trichlorosilane (TCS) for more than 3 hours under vacuum for surface passivation. Polydimethyl-siloxane (PDMS) pre-polymer and bonding agent (Dow Corning Sylgard 184) were mixed in a 10:1 ratio and around 20-25 g of the mix was poured to the master wafer that was laid flat on a petri dish to achieve devices that were ~4 mm thick. These were baked in an oven at 70 °C overnight to harden the PDMS. The PDMS slabs with the device design were cut out from the master and holes for the tubing to be connected to the devices were punched under a clean room hood. No. 1.5 glass coverslips and these slabs were washed with isopropanol and dried on a hot plate at 75 °C, and subsequently plasma cleaned their surfaces in a Reactive Ion Etcher (AXIC PlasmaSTAR) with O<sub>2</sub> plasma (50 W power, 50 SCCM and 5-10 sec exposure). The PDMS slabs were bonded to the coverslips immediately after plasma cleaning and heated on a hot plate at 75 °C for a couple of hours, then were stored at room temperature for ~24 hours before filling the electrode channels with In-Sn solder (In 52% and Sn 48%, melting point 118 °C). Terminal pins (Male, GT-150) were used to connect the solder filled channels with wire strips that can be clipped to high voltage power sources that are necessary for DEP sorting. The pins are fixed to the devices using epoxy glue to ensure a secure connection. Pneumatic pressure controllers (Parker Electronics) were used to control pressure from cell and oil reservoirs that are connected to the input channel of the devices.



**Figure A1.2. Microfluidic chip design.** The droplet formation and selection junctions are indicated within the solid red box and zoomed in. The laser beams are placed in the dashed blue box to excite the fluorophores in droplets, where the dimension of the channel is  $530 \times 90 \times 50 \mu\text{m}^3$  (length  $\times$  width  $\times$  depth).

## Section 2. Instrument Operation Parameters

Excitation intensities for optical transitions are usually about 1-10 kW/cm<sup>2</sup> for both beams. The illumination is a Gaussian beam in the z- profile (out of imaging plane) and is elliptical in the x-y plane. The full width at half maximum (FWHM) of the elliptical beams have the dimension of ~2.5  $\mu$ m and ~4.5  $\mu$ m in the in-flow direction, and ~45  $\mu$ m and ~35  $\mu$ m in the perpendicular direction for 561 nm and 450 nm beams, respectively. The PMT voltages are usually set to be - 1025 V. Electronically controlled pneumatic pressure controllers (Parker Inc.) are operated in the range of 0.5-5 psi for cell channel and 5-10 psi for oil channel in current chip configuration, which allows uniform droplet generation at the frequency tested up to 4 kHz. All sorting modes are tested for droplet frequencies up to 2.5 kHz in the current microfluidic chip configuration, and the droplet volume for such operation frequencies is measured to be around 50 pL. To generate effective force to sort droplets flowing at 1-2.5 kHz, 1.5 kV peak to peak pulses at 5 kHz frequency are applied on the electrodes of the chip. The high voltage pulses applied on the electrodes are set to 2-5  $\mu$ s with variable delay time (>1  $\mu$ s) allowing the droplet to travel from the last beam to the collection-waste junction depending on the position of the last beam in the dashed blue box in Figure A1.2.

## Section 3. Cell culture and sample preparation

For this study, FPs were expressed in the unicellular bacterial system, *Escherichia coli* (*E. coli*). Bacterial systems offer high expression efficiencies and fast doubling times (~20 minutes) and are robust systems for screening and selecting fluorescent proteins with improved photo-physical properties. Bacterial cells (Top10) are transformed with the plasmid DNA of a fluorescent protein of interest and grown on agar plates with ampicillin resistance and arabinose for expression in an incubator at 37 °C overnight, then a single bacterial colony is selected and passed into a 5 mL LB media with ampicillin resistance (referred as LB-amp) culture tube and grown overnight in a shaker at 37 °C, 230 rpm. 100  $\mu$ l of this culture is then passed into a 10 mL of LB-amp flask and grown in the shaker at 37 °C, 230 rpm for 3 hours till optical density (OD) ~0.6 at 600 nm, post which arabinose is added to the culture for expression for 16 hours in the shaker at 28 °C, 230 rpm. After 16 hours expression, 1~2 ml of cells is pelleted (8000 rpm for 2 minutes) and washed with aqueous blank buffer (composed of 0.17% Nitrogen Base and 0.5% Ammonium Sulfate both in wt%), then

diluted to desired concentrations determined by the absorbance of the re-suspended cell solution in blank buffer at 600 nm. These cell solutions are filtered with sterile cell strainers (40  $\mu\text{m}$ ) before being used for screening or sorting. Sorted cells are usually collected on a suspension in oil and added to a flask with 25 mL LB-amp for overnight growth in a shaker at 37 °C, 230 rpm, and can be expressed for further analysis and sorting depending on the needs.

## Section 4. Data acquisition and signal processing

### 4a. Characterizing scattering from droplets

The laser beam focused on the microfluidic chip using a cylindrical lens has a Gaussian intensity profile  $I(x, y)$  with FWHM intensity spreads  $\sigma_x$  and  $\sigma_y$  as

$$I = I_0 \exp\left(-\frac{2x^2}{\sigma_x^2} - \frac{2y^2}{\sigma_y^2}\right), (1)$$

where  $I_0$  is the peak intensity at the center of the Gaussian beam.

Droplets flowing through this beam result in scattering signals. Despite the presence of emission filters some scattered signals of excitation light leak into the PMT. The scattered light is used to our advantage, for tracking droplet flow speeds and stability, counting number of droplets, and pair-matching two signals in green and red channels from the same droplet. Depending on the concentration of cells used in an experiment – droplets can be filled or empty. Only scattering signals are observed when the droplet is empty, but when the droplet contains a fluorophore – both scattering, and fluorescence signals are recorded on the PMT. These cases may be represented by Gaussian profiles that follow the intensity profile in time: the scattering signal  $Sc(t)$  from an empty droplet, fluorescence signal  $Fl(t)$  and the total signal  $C_0(t)$  from a filled droplet (assuming a 10:1 fluorescence: scattering) as presented in Eqs. (2 – 4). In an instance when a droplet is filled with a single fluorophore, the signals are a convolution of the two functions. The ratio of scattering and fluorescence signals can be varied by adjusting the pump laser intensity, flow speeds, and the PMT gains for an experiment.

$$Sc(t) = Sc_0 \exp\left(-\frac{t^2}{2\sigma_{st}^2}\right), (2)$$

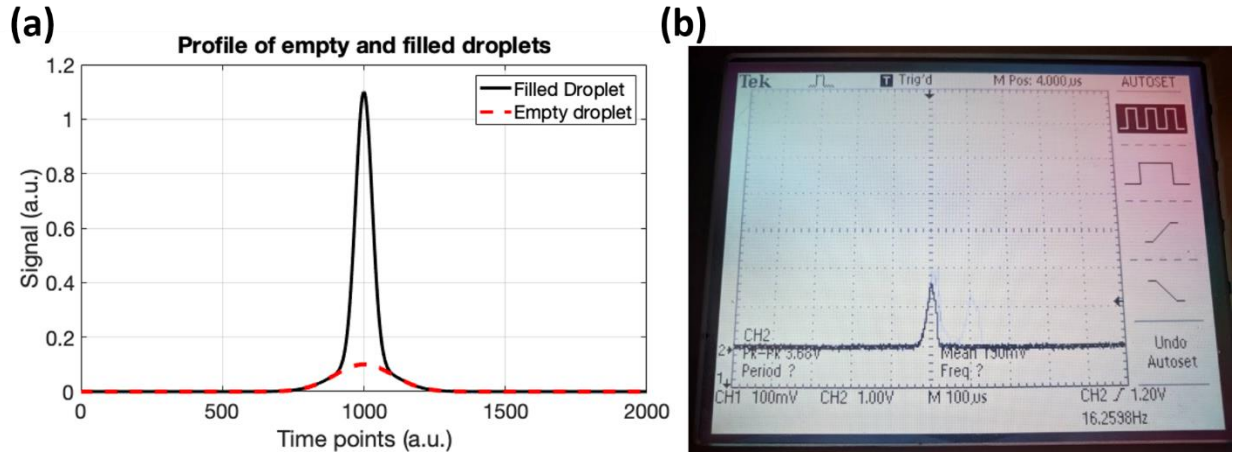
$$Fl(t) = Fl_0 \exp\left(-\frac{t^2}{2\sigma_{ft}^2}\right), (3)$$

$$C_0(t) = 0.1Sc_0 + Fl_0, (4)$$

The distributions in time for the beams  $\sigma_{st}$  and  $\sigma_{ft}$  are dependent on the flow speed  $v_x$ , which in turn is related to the droplet generation frequency. An example of simulated and experimentally observed signals is shown in Figure A1.3.

$$\sigma_{st} = \frac{\sqrt{\sigma_x^2 + radius\_droplet^2}}{2 v_x}, (5)$$

$$\sigma_{ft} = \frac{\sqrt{\sigma_x^2 + radius\_cell^2}}{2 v_x}, (6)$$



**Figure A1.3. Simulated and observed signals of a filled droplet.**

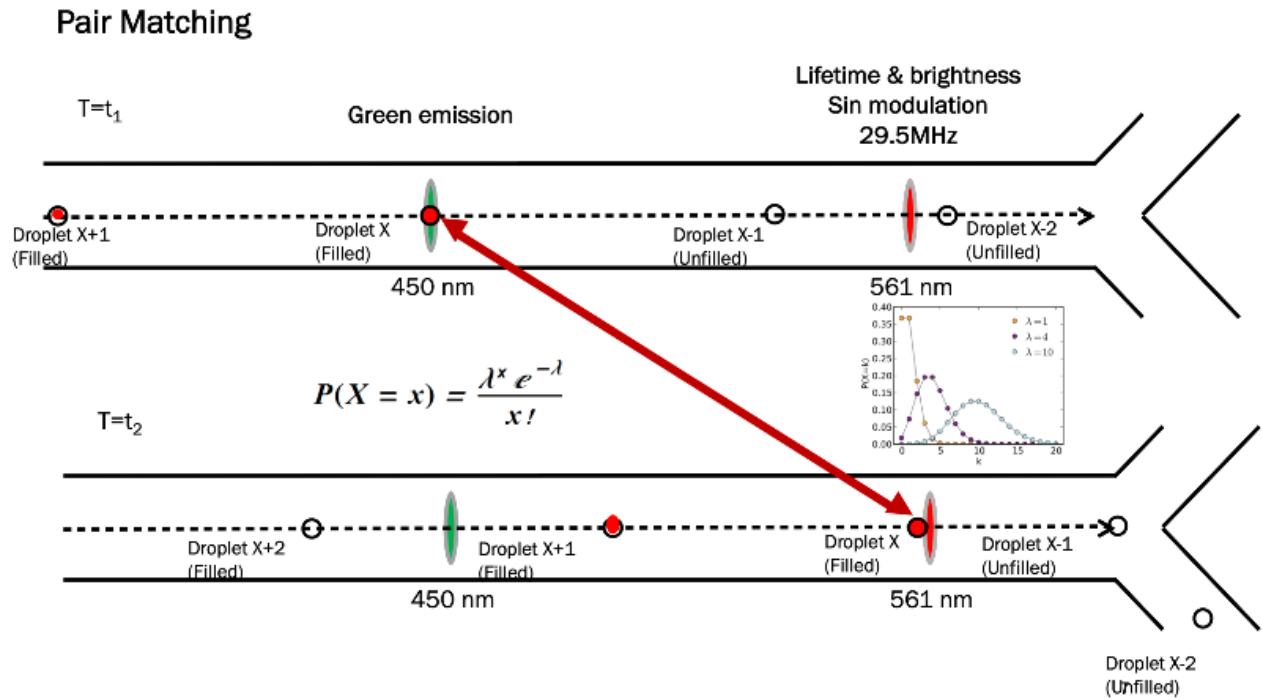
## 4b. Pair-matching

Our instrument has two beams in the current configuration and since we are interested in spectroscopic signatures from the same analyte from both beams, these two events need to be “pair matched” in time. The pair matching event depends upon the flow speed. Either using a distribution feature on the custom-built LabView program or through real time tracking on an oscilloscope we can calculate the flow speed of droplets. A distribution tab on the program allows us to estimate

the pair matching times, and a narrow distribution of pair matching times indicates a steady and stable flow. The distance between the two beams can be calculated through a camera image by a calibrated pixel to length conversion ratio. These allow us to calculate the flow speed of the droplets in the system.

$$v_x = \frac{\text{Distance between the beams}}{\text{Pair matching time}}, (7)$$

Pair-matching events are schematically illustrated in Figure A1.4. For example, when the separation distance of two beams  $d \sim 250 \mu\text{m}$  and the pair matching time  $\sim 500 \mu\text{s}$ , the flow speed of the droplets is  $v_x \sim 0.5 \text{ m/s}$ . The averaged droplet counts per second from a single beam can be used to check whether the pair matching time is reasonable.



**Figure A1.4. Pair matching events to estimate the transit time and the flow speed of the system.**

Repeated measurements indicate that our droplets are around  $\sim 50 \text{ pL}$ , which gives the estimated droplet sizes to be  $\sim 46 \mu\text{m}$  in diameter. Our analyte is usually an *E. coli* cell, which is roughly assumed to be of the size of  $5 \mu\text{m}$  in diameter. The FWHM of our beams on the device are usually

$\sim 3 \mu\text{m}$  on the x-(in-flow) dimensions. Using eqns. (5) and (6),  $\sigma_{st} \sim 34 \mu\text{s}$  and  $\sigma_{ft} \sim 25 \mu\text{s}$ . Thus, a fluorescent signal usually has an FWHM of  $\sim 25 \mu\text{s}$  whenever it passes a beam when we have a steady flow velocity  $\sim 0.5 \text{ m/s}$ .

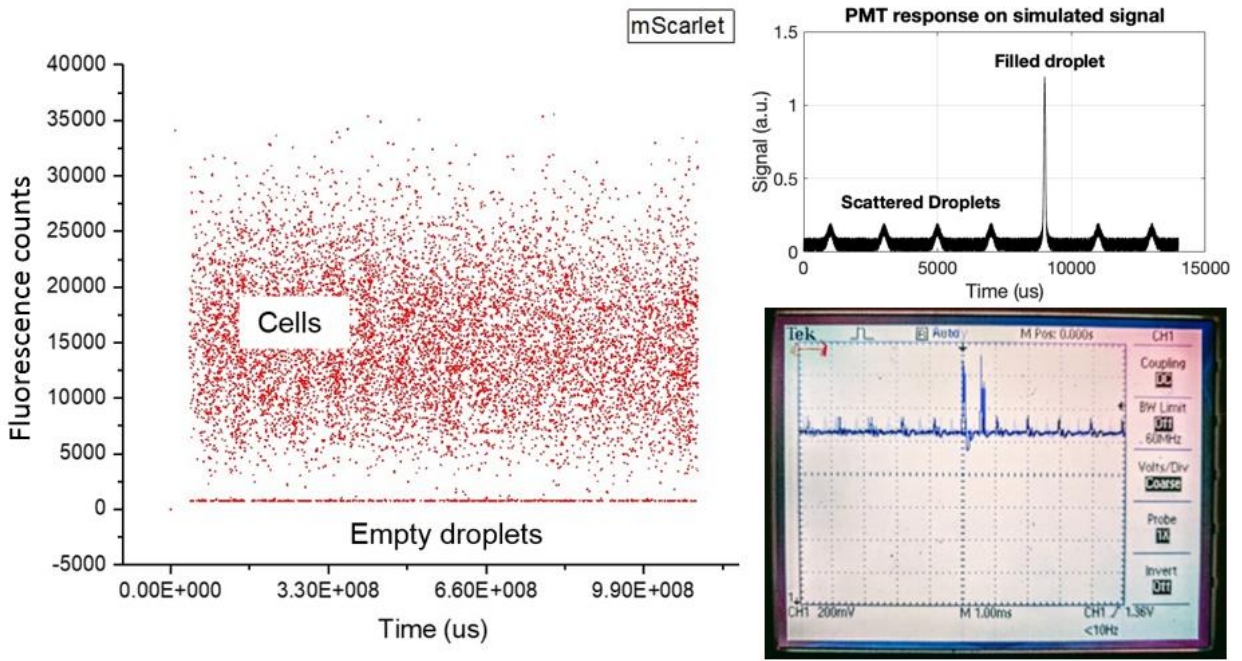
#### 4c. Raw Fluorescence Brightness

To operate the instrument in each mode, one needs to screen the distribution of empty droplets signals to separate from droplets that contain the fluorophore. Then by subtracting the background of empty droplets, one can estimate the distribution of the fluorophores in the respective photophysical parameters that can be analyzed. The operation of the instrument to get the raw peak brightness of an event is carried out in the following fashion.

- The ADCs digitize the analog signals from the PMT.
- The data is then handled by the FPGA board.
- FPGA starts a search algorithm if the data point is higher than threshold.
- FPGA searches for the next 8 data points to see if they are above threshold.
- **YES:** Peak detection – 8 continuous points above threshold. Selects the maximum as the brightness reading from continuous data points above the threshold.
- **NO:** Abandons peak search and looks for next point above the threshold.

Figure A1.5 exhibits an example of screening fluorescence brightness of *E. coli* cells expressed with mScarlet, simulated signals and experimental signals output on an oscilloscope.





**Figure A1.5. Comparing signals from filled and empty droplets.** (Left) Fluorescence counts obtained in Red channel shows the signals from empty droplets and mScarlet. (Upper right) Signals simulated on MATLAB. (Lower right) Similar responses in flow are also noted real time on an oscilloscope during an experiment with average number of cells encapsulated per droplet  $\lambda=0.1$  (9.5% of droplets filled).

#### 4d. Fluorescence Lifetime

For phase fluorimetry our laser is amplitude modulated in sine by the EOM. The modulation frequency of the EOM is set to  $\nu=29.5$  MHz, ( $\omega = 2\pi\nu$ ) which resonates with the fluorescence lifetime of fluorophores in the order of ns. The fluorescence signal is therefore also amplitude modulated in time but is recorded with a phase delay and a lower modulation depth in comparison to the reference signal. The phase delay corresponds to the average time a fluorophore spends in the excited state before it emits a photon back to the ground state (eqn. (8)).

$$Fl(t) \propto A m \sin(\omega t - \phi - \delta) \exp\left(-\frac{t^2}{2\sigma_{ft}^2}\right), (8)$$



The phase shift  $\phi$  corresponds to the phase delay in the signal caused by the intrinsic fluorescence lifetime of the analyte whereas the constant phase shift  $\delta$  can be attributed to the delay of the signal due optics, electronics etc.

$$\tau = \frac{1}{\omega} \tan(\phi), (9)$$

$$m = \frac{1}{\sqrt{1+\omega^2\tau^2}}, (10)$$

In the experimental setup of the system, a biased-Tee filters the high frequency component (>83 kHz) of the signal from the larger gaussian envelope. The gaussian envelope is sent to a pre-amplifier (logarithmic/linear). The signal from the pre-amplifier is then digitized by the ADCs and a peak of the signal is determined as discussed previously.

The high frequency component is fed into one of the input ports of the lock-in amplifier. A reference signal is sent to the lock-in amplifier from a function generator that drives the EOM. The lock-in amplifier calculates the in-phase (cosine component of the signals) and the out-of-phase (sine component of the signal) with respect to the reference signal, and the phase shift is calculated from the inverse tangent of these two parameters.

It is critical to note that the phase value of empty droplets is usually a distribution that is close to the mean of the total phase shift observed from delay due to the optical and electronic lines and the offset value that sets the reference phase shift of a known RFP to be around 45 degrees. Whenever a fluorescence event takes place, i.e., the analyte is in a droplet that passes the beam, an intrinsic phase shift is observed that corresponds to the average fluorescence lifetime of the analyte. Once the FPGA determines the peak fluorescence value, it picks up the corresponding phase value that is reported by the lock-in amplifier through an auxiliary output on the lock-in amplifier. However, the accuracy and stability of phase values are critical to many parameters including

1. The low-bandpass filter for the signal inputs
2. The time constant (TC) of the corresponding filter
3. The spread  $2\sigma_{ft}$  which is critical in determining the order of the low-pass filter and the time constant of the UHF-Lock-in amplifier

4. The PMT gain and the excitation intensity: the scattered signals are not negligible, if fluorescent events are not very prominent, one can observe a skewing of the phase shift towards the mean value of the empty droplets (see Example 2 below).

It is critical to note that the Equivalent Noise Bandwidth (ENBW) is dependent on the time constant and the order of the filter in use. Lower order filters result in greater noise in the lifetime measurements, but the wait times for the phase to stabilize to the 99% of the accurate values are much lower (as an example, see SRS-830 Lock-in-amplifier Manual).

<u>Slope</u>	<u>ENBW</u>	<u>Wait Time</u>
6 dB/oct	1/(4T)	5T
12 dB/oct	1/(8T)	7T
18 dB/oct	3/(32T)	9T
24 dB/oct	5/(64T)	10T

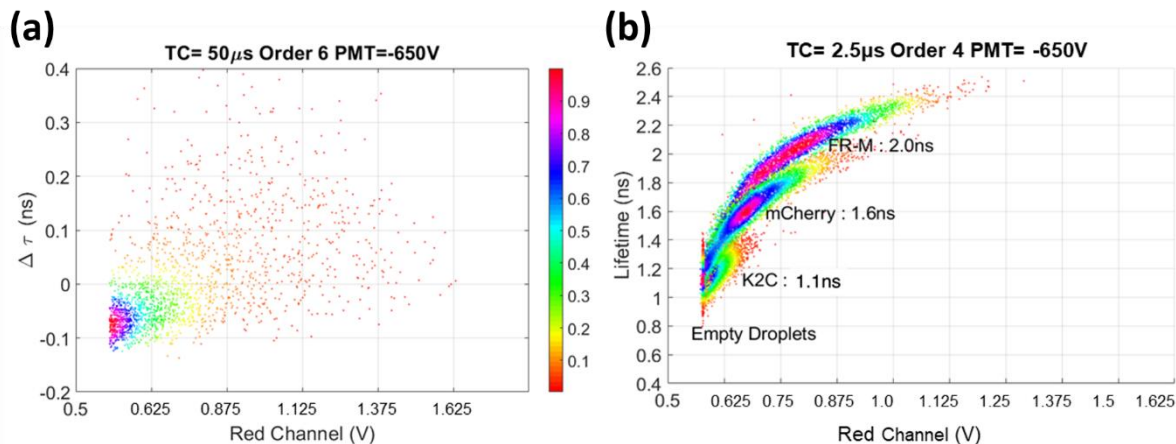
(SRS-830 Lock-in-amplifier Manual)

Hence it is critical that we select a filter and a time constant that provides us enough resolution in lifetime but does not require a long wait time. The noise reduces when the time constant increases, but the wait time also increases. As flow speed is increased the transient time of the fluorescence signal keeps dropping. Thus, it is critical to note that for a certain PMT setting and flow speed the appropriate time constant and the filters are selected for a reasonable phase resolution. Some examples are shown below to illustrate the effects of these settings.

Example 1: A mixture of different fluorescent proteins (mCherry, Kriek and FusionRed-M) are mixed and are screened in a chip that operates at a speed of 2 kHz droplet generation frequency. The PMT was set at -650 V and the laser irradiance was at 75 kW/cm<sup>2</sup> to obtain strong scattering signals. The FWHM of the signal is around 25 μs, and the selection of a TC of 50 μs with an order 6 filter results in a noisy phase signal that is centered around the mean of the empty droplets. This means that the FPGA does not pick up the right phase and the values are centered around the distribution of the empty droplet. The results are shown in Figure A1.6 (a).

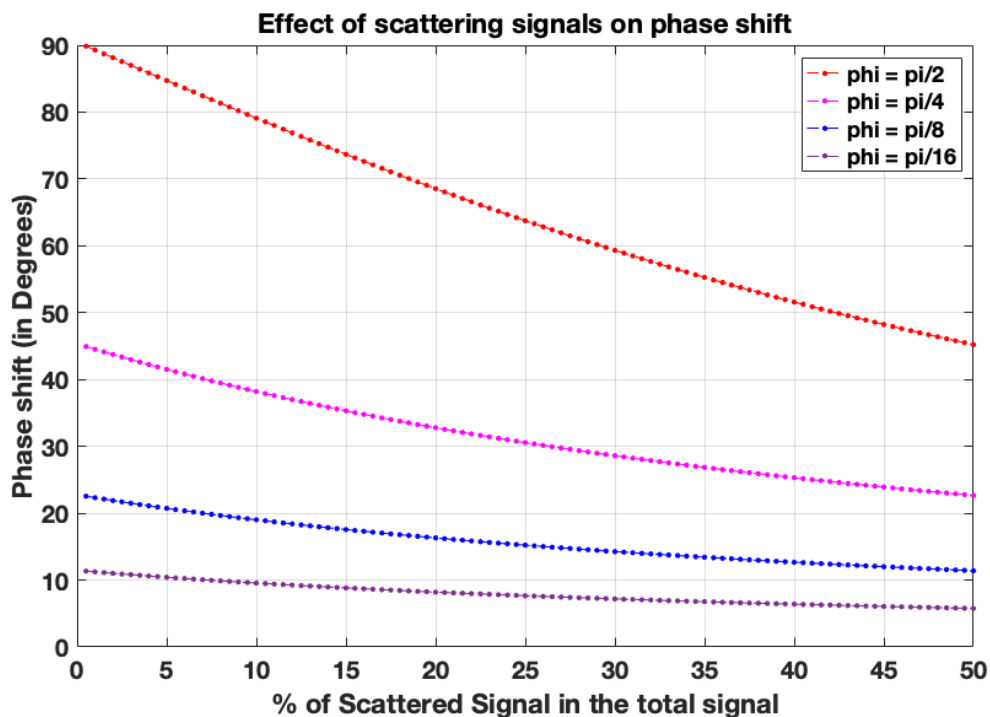
Example 2: The same flow settings were used with a time constant of 2.5 μs and an order 4 filter. While the main populations of each RFP are distinguishable in the mixture, the heterogeneity in lifetime and brightness makes it difficult to distinguish each RFP clearly. Therefore, each RFP was screened individually and plotted on the same scatter plot in Figure A1.6(b). The lifetimes could

be resolved but the scattering signals are also high due to the strong irradiance, resulting in fluorescence signals that are skewed in the lifetime axis. Here mCherry was selected as a reference and its mean lifetime was set at 1.6 ns.



**Figure A1.6. Screening results of a mix of mCherry, FusionRed-M and Kriek (K2C) from (a.) Example 1 and (b.) Example 2.** The fluorescence lifetime is not resolvable in the settings of Example 1, whereas the main population of each RFP is resolvable in Example 2 but the lifetime data is skewed.

The major reason for the data skewing on the lifetime axis in Example 2 can be attributed to the fact that the scattered signals are comparable with the fluorescence signals. The scattered light is added to the fluorescence signal and both signals have the same modulation frequency but different phase shift values, so the lock-in amplifier extracts an averaged phase value from the combined signals. Therefore, the averaged lifetime becomes closer to the actual fluorescence lifetime as the fluorescence brightness increases, but the influence from the scattered light becomes noticeable as the fluorescence brightness decreases. A simulation of the reported phase values was carried out by assuming the phase values being the convolution of the offset phase shift of scattered light (set to zero in this simulation) and different intrinsic phase shifts from fluorescence lifetime as presented in Figure A1.7. As the fraction of scattered signal increases, i.e. the fluorescence is weaker, the measured phase shift of the total signals is skewed towards the phase of scattered light (zero phase shift).



**Figure A1.7. Simulation of phase shifts obtained from a convolution of scattered and fluorescence signals.** This qualitatively indicates the brightness dependent lifetime data observed in Figure A1.6. It clearly demonstrates the effect is more dominant for greater phase shifts and operations at lower fluorescence: scattering.

#### 4e. Calculation of lifetimes from phase shifts

The lifetime mode of the instrument works on the concept of phase fluorimetry based on relative lifetime shifts. Therefore, lifetime characterization in the system needs a reference value for calibration. In most cases either RhodamineB containing fluorescent beads or a fluorescent protein of known lifetime is used as a reference lifetime value. We have used mCherry as the reference in most of our measurements. mCherry is set to 1.7 ns as a reference.

The phase shifts are reported to the FPGA in terms of voltages that are digitized using an ADC board. The least significant bit (LSB) value for the boards is 76  $\mu$ V. The lock-in amplifier reports every degree of a phase shift as +10 mV increments.

Given our resonant frequency is 29.5 MHz. Each degree of phase shift is expressed as

$$\Delta \phi = 1^\circ = \frac{1}{360 \times 29.5} \times 10^{-6} \text{s},$$

and phase shift is expressed in LSB values,

$$\Delta \phi \text{ (V)} = \Delta(\text{LSB}) * 76 * 10^{-6} \text{ V}.$$

Equating the above two equations,

$$\Delta \tau = \Delta(\text{LSB}) \times 7.15 \times 10^{-4} \text{ ns}.$$

Or the fluorescence lifetime of an analyte can be expressed as

$$\tau = \tau(\text{reference}) + \Delta(\text{LSB}) \times 7.15 \times 10^{-4} \text{ ns}.$$

## Section 5. Performance of the sorter

### 5a. Two-color brightness sorting

The samples were prepared as described in Section 3 to obtain *E. coli* expressed with mScarlet and EGFP. *E. coli* cells expressing mScarlet or EGFP were mixed in the proportion of ~1:1 before sorting. The mixture was screened to obtain the mean brightness of mScarlet, then we sorted ~2000 droplets from ~10<sup>5</sup> droplets with the selection gate greater than the mean brightness of mScarlet. The sorting process was repeated 2 times to collect totally 3 sorted samples. The sorted cells were subsequently grown overnight and screened 16 hours after induction of expression. The screening results are tabulated in Table A1.1.

**Table A1.1. Screening results of sorting efficiency.** The sorted cells were re-grown, expressed, and screened for totally ~1000 cells.

# of trials	# of EGFP	# of mScarlet	Sorting efficiency
1	153	1016	87%
2	192	1057	85%
3	143	965	86%

## 5b. Lifetime screening with purified proteins

Three purified proteins, mCherry, mApple, and mScarlet, and an organic dye, Rhodamine B, were prepared with concentrations about 10s  $\mu\text{M}$  to obtain brightness in the range comparable with cellular screening in the sorter. Pure proteins were prepared in Tris-HCl buffer (pH=7.4), and Rhodamine B was dissolved in deionized water (pH~7). Fluorescence lifetime of each sample was measured using Time-correlated Single Photon Counting (TCSPC), then transferred to the microfluidic droplet sorter to screen for the fluorescence lifetime. The results are tabulated in Table A1.2. The deviation between TCSPC and sorter measurements is due to the sorter averaging the phase shifts from the scattered excitation light and the fluorescence signals as described in Section 4. As the fluorescence phase shift (lifetime) being further away from the scattering phase offset, the averaged phase shift deviates from the actual fluorescence lifetime more, in agreement with the simulation shown in Section 4.

**Table A1.2. Fluorescence lifetime measured in TCSPC and the microfluidic droplet sorter.**

The lifetime of Rhodamine B was set as the reference for the sorter measurement, and the uncertainties are the standard deviations from ~25,000 screened droplets.

Sample	TCSPC Intensity weighted lifetime in cuvette (ns)	Lifetime from frequency domain measurement in flow (ns)
Rhodamine B	1.63	1.63 $\pm$ 0.05 (ref)
mCherry	1.54	1.58 $\pm$ 0.09
mApple	3.07	2.78 $\pm$ 0.09
mScarlet	3.78	3.45 $\pm$ 0.08

## 5c. Lifetime-based sorting

The samples were prepared as described in Section 3 to obtain *E. coli* expressed with mScarlet and mCherry. In each batch of experiment, both cell solutions were grown from their single colony, thus there were two biological duplicates for each RFP in two batches of experiments. In the first batch of experiment, *E. coli* cells expressing mScarlet or mCherry were mixed in the proportion of ~1:1 before sorting. With one round of sorting on mScarlet, an 85% sorting efficiency was obtained as described in the Chapter 2. In the second batch of experiment, *E. coli* cells expressing

mScarlet or mCherry were mixed in the proportion of ~40:60 before sorting. The mixture was screened to obtain the mean brightness and fluorescence lifetime of mScarlet and mCherry, then sorted 3,000 droplets from ~250,000 droplets with the selection gate greater than the mean brightness and fluorescence lifetime of mScarlet. The sorting process was repeated 2 times to collect totally 3 sorted samples. Another set of sorting was performed similarly by gating at brightness and fluorescence lifetime lower than the mean of mCherry. The sorted cells were subsequently grown overnight and screened 16 hours after induction of expression. The screening results are tabulated in Table A1.3.

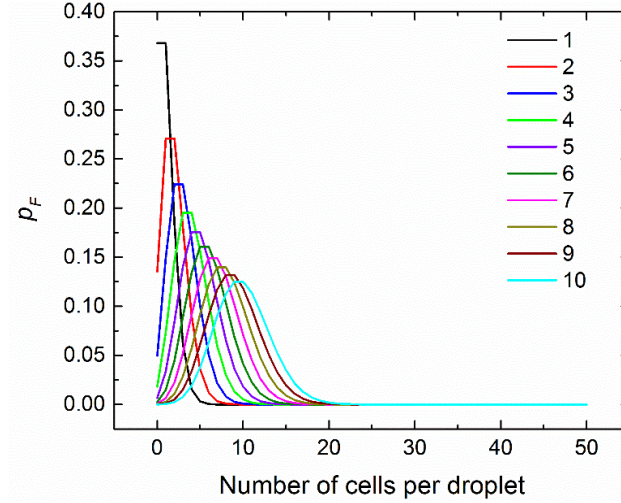
**Table A1.3. Screening results of sorting efficiency.** The sorted cells were re-grown, expressed, and screened for ~50,000 cells

	# of trials	# of mCherry	# of mScarlet	Sorting efficiency
<b>Sort for mScarlet</b>	1	9688	40030	81%
	2	10745	39288	79%
	3	9780	38522	80%
<b>Sort for mCherry</b>	1	46799	1208	97%
	2	50053	1833	96%
	3	47740	1341	97%

## Section 6. Enriching rare events

### 6a. Poisson distribution

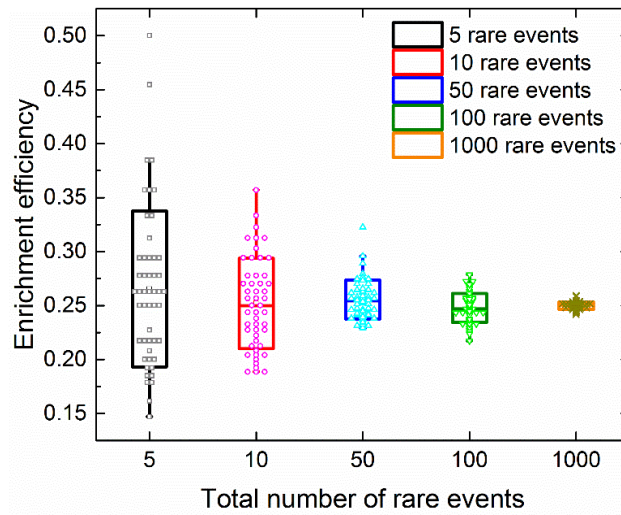
The Poisson distribution for  $\lambda = 1$  to  $\lambda = 10$  are plotted in Figure A1.8. The probability of encapsulated cells per droplet decreases quickly with the increasing number of encapsulated cells, so the  $p_F$  can be numerically calculated using  $n \leq 50$  for  $\lambda \leq 10$ .



**Figure A1.8. Poisson distribution for  $\lambda = 1$  to  $\lambda = 10$ .**

## 6b. Enrichment efficiency

The sorter with current microfluidic chip configuration is capable of enriching rare events from  $10^8$  population within 3 hours without losing cell viability, when choosing  $\lambda \geq 4$ . As the targeted events become rarer, the enrichment efficiency varies depending on how many unwanted cells are encapsulated with the targeted cells in one droplet. We estimate the convergence of the enrichment efficiency below.



**Figure A1.9. Enrichment efficiency converges as the number of rare events increases.**



Assume that there are totally  $n$  targeted cells in  $10^8$  population for enrichment. The droplet encapsulation obeys Poisson distribution, we assume that there is only one targeted cell found in each sorted droplet since they are rare events. This is only an approximation, but it does not lose the generality in the convergence of enrichment efficiency. We simulate sorted  $n$  ( $n = 5, 10, 50, 100$ , and  $1000$ ) rare cells (i.e. droplets) by generating random numbers based on Poisson distribution with  $\lambda = 4$ , as the numbers of cells encapsulated in droplets. Repeating the process 50 times for each case to attain a distribution of enrichment efficiency and the mean and standard deviation, as plotted in Figure A1.9. The average enrichment efficiency is lower than the predicted values shown at Figure 2.6 in the Chapter 2, because this is only an approximation considering only 1 targeted cell per droplet. The standard deviation is  $\sim 7\%$  of the mean for enriching 50 cells from  $10^8$  population, and  $\sim 5\%$  of the mean for enriching 100 cells from  $10^8$  population. The results indicate that the calculated enrichment efficiency can be expected when sorting  $0.5\sim 1$  ppm of rare events from  $10^8$  population. Please note that it does not limit the enrichment efficiency for sorting smaller fraction of rare events. With smaller fraction, the enrichment efficiency may deviate from the expected value, but it still provides approximately same order of magnitude of enrichment. For example, sorting 5 rare cells from  $10^8$  population, the enrichment efficiency varies between approximately 15-50% in Figure A1.9, i.e. enriching rare events from fraction  $10^{-8}$  to  $\sim 10^{-1}$  resulting in  $10^7$  times enrichment.

## **6c. Rare RFP enrichment**

### **i. Enrichment of rare mScarlet from a mixture with EGFP**

The estimated droplet size is approximately 50 pL, so *E. coli* cells expressing EGFP and mScarlet were individually prepared at a concentration less than  $2 \times 10^6$  cells/ml to ensure an average of  $\lambda \leq 0.1$  cell/droplet to screen and determine their mean brightness. To prepare the sample for enrichment experiment, the concentration of the mixed cells was estimated to be  $5.9 \times 10^7$  cells/ml for an average of  $\lambda = 3$  cells/droplet encapsulation. A portion of the mixed cells were diluted to less than  $2 \times 10^6$  cells/ml to ensure an average of  $\lambda \leq 0.1$ , and was used to determine the fraction of mScarlet in the mixture being  $F \sim 0.01$  by screening  $\sim 10,000$  mixed cells. The selection threshold was set to be greater than the mean brightness of mScarlet. After one round of enrichment by

sorting ~2,000 droplets with  $\lambda = 3$  encapsulation, the sorted cells were subsequently grown, expressed and screened with  $\lambda \leq 0.1$  encapsulation. The same process was repeated twice in parallel to obtain totally 3 measurements. The results are tabulated in Table A1.4. The mScarlet population was enriched from  $F \sim 0.01$  to a weighted average  $35 \pm 4\%$ .

**Table A1.4. The enrichment of rare mScarlet from a mixture with EGFP screened at a rate of 2 kHz.**

# of trials	EGFP cells	mScarlet cells	%RFP in mix
1	1785	1218	41
2	5109	2722	35
3	1070	299	22

## ii. Enrichment of rare mScarlet from a mixture with mCherry

The sample preparation is similar to the mScarlet and EGFP mixture described above. The mean of fluorescence brightness and lifetime was determined by screening individual mScarlet and mCherry cells at  $\lambda \leq 0.1$ . The fraction of mScarlet in the mixture before enrichment was estimated to be  $F \sim 5 \times 10^{-3}$  by screening ~10,000 mixed cells at  $\lambda \leq 0.1$  for both batches of experiments. The sorting gates were set to be greater than the mean of fluorescence brightness and lifetime of mScarlet. After one round of enrichment by sorting ~2,000 droplets with  $\lambda = 3$  encapsulation, the sorted cells were subsequently grown, expressed and screened with  $\lambda \leq 0.1$  encapsulation. We attained an enrichment of the mScarlet population to 40% in the first batch of experiment. In the second batch of experiment, the same process was repeated 3 times in parallel to obtain totally 3 measurements. The results of the second batch of experiment are tabulated in Table A1.5. The mScarlet population was enriched from  $F \sim 5 \times 10^{-3}$  to a weighted average  $30 \pm 5\%$ .

**Table A1.5. The enrichment of rare mScarlet from a mixture with mCherry screened at a rate of 2.5 kHz.**

# of trials	mCherry cells	mScarlet cells	%RFP in mix
1	31846	17623	36
2	29775	15519	34
3	32251	7750	19

## Section 7. Directed evolution of an RFP library

### 7a. Site directed libraries and the sorting for bright mutants

We saturated mScarlet-I with the RRN codon that codes for R, D, E, K, G (2 copies), N and S amino acids (see Table A1.6). The library size with equal probability for each available amino acid in this case is  $8^8$  possible mutants ( $\sim 1.7 \times 10^7$ ). To ensure the search covering 95% of a library with equal probability for each available amino acid, at least 3 times of the library size must be screened, which is  $> 5.1 \times 10^7$  cells. [1]

The library DNA was transformed to bacterial cells (Top10 strain), with a standard heat-shock protocol. These transformed cells were added to LB+ampicillin ( $\sim 7$  mL) and grown overnight for the cells to reach higher optical densities. 1.5 mL of this culture was then added to 100 mL 2XYT+ampicillin media, grown for  $\sim 2$  hours to achieve an OD  $\sim 0.6$ . Arabinose was then added to this culture to initiate expression of RFPs in the system. The library selection/sorting was carried out  $\sim 16$ -20 hrs post induction to select mutants with fast RFP maturation.

Before each sorting round, the to-be sorted library was screened with low cellular concentrations  $\lambda=0.1$  to ensure single cell encapsulation in droplets in order to estimate the percent fluorescent cells in the given population. The first two rounds of sorting namely 'S1' and 'S2' were carried out using multiple cells per droplet ( $\lambda=3$  for S1 and  $\lambda=1$  for S2), because each round had significant non-fluorescent mutants. S1 and S2 led to significant enrichment of the library with fluorescent mutants. Then onwards, the library was sorted on two independent pathways at  $\lambda=0.1$  with higher thresholds in fluorescence brightness. The sorts were repeated to ensure library convergence. After every sort, the cells in oil were transferred to a culture with 25 mL LB+ampicillin to grow overnight, consequently, expressed in 100 mL of 2XYT+ampicillin for the next round of sorting. A few ml of the re-grown cells from each sort were made into glycerol stocks stored at  $-80$  °C for future use.

## 7b. Sorting processes and results

### S1

Size  $\sim 1.7 \times 10^7$  mutants ( $\sim 5\%$  fluorescent from screening  $\sim 40,000$  cells)  
 Sort Speed: 2000 droplets/s ( $\sim 2.2 \times 10^7$  cells/hr)  
 Sorted at  $\lambda=3$  cells/droplet in 2 batches  
 Brightness sorting gate : 0.38V (Blank Subtracted just above empty droplet brightness)  
 Screened Cells:  $\sim 8 \times 10^7$  ( $\sim 4.5\times$  coverage)  
 Sorted Cells:  $\sim 5.6 \times 10^6$  ( $\sim 7\%$  of screened cells)

### S2

$\sim 30\%$  of initial population are fluorescent from screening  $\sim 10,000$  cells.  
 Sort Speed: 2000 droplets/s ( $\sim 2.2 \times 10^6$  cells/hr)  
 Sorted at  $\lambda=1$  cell/droplet (in 1 batch)  
 Brightness sorting gate : 0.38V (Blank Subtracted just above empty droplet brightness)  
 Screened Cells:  $1.7 \times 10^7$   
 Sorted Cells:  $2.8 \times 10^6$

### S3A

$\sim 90\%$  of initial population are fluorescent from  $\sim 20,000$  cells.  
 Sort Speed: 2000 droplets/s ( $\sim 7.2 \times 10^5$  cells/hr)  
 Sorted at  $\lambda=0.1$  cells/droplet  
 Mean Brightness =  $0.71 \pm 0.6$  V (Empty droplet subtracted)  
 Brightness sorting gate: 1.9V  
 Screened Cells:  $2.3 \times 10^6$   
 Sorted Cells:  $\sim 1 \times 10^6$

### S3B

$\sim 90\%$  of initial population are fluorescent from  $\sim 20,000$  cells.  
 Sort Speed: 2000 droplets/s ( $\sim 7.2 \times 10^5$  cells/hr)  
 Sorted at  $\lambda=0.1$  cells/droplet  
 Mean Brightness =  $0.76 \pm 0.5$  V (Empty droplet subtracted)  
 Brightness sorting gate : 1.9V  
 Screened Cells:  $1.7 \times 10^6$   
 Sorted Cells:  $0.9 \times 10^5$

### S4A

$>90\%$  of initial population are fluorescent from  $\sim 20,000$  cells.  
 Sort Speed: 2000 droplets/s ( $\sim 7.2 \times 10^5$  cells/hr)  
 Sorted at  $\lambda=0.1$  cells/droplet  
 Mean Brightness =  $1.25 \pm 0.8$  V  
 Brightness sorting gate : 2.3V  
 Screened Cells:  $8 \times 10^5$   
 Sorted Cells:  $4.5 \times 10^4$

### S4B

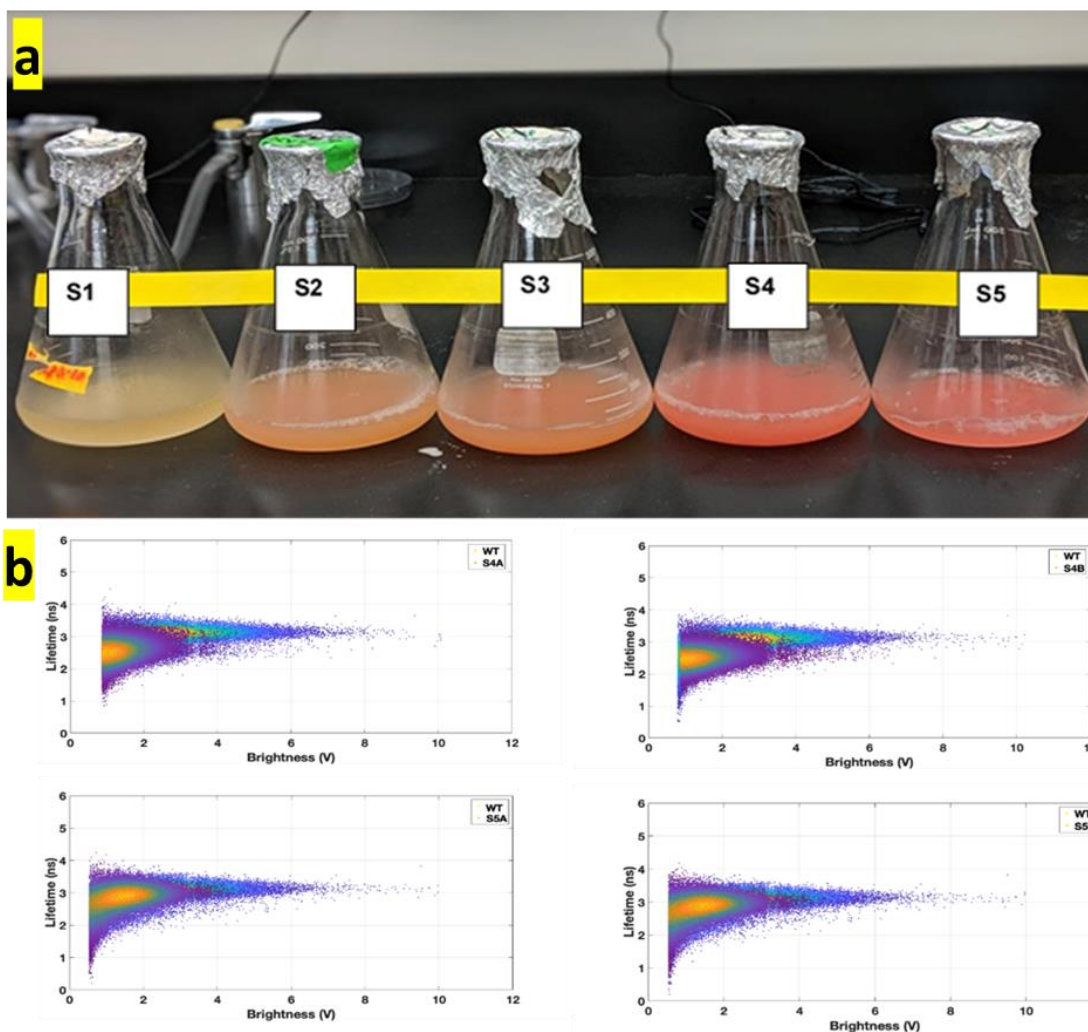
$>90\%$  of initial population are fluorescent from  $\sim 20,000$  cells.  
 Sort Speed: 2000 droplets/s ( $\sim 7.2 \times 10^5$  cells/hr)  
 Sorted at  $\lambda=0.1$  cells/droplet  
 Mean Brightness =  $1.25 \pm 0.8$  V  
 Brightness sorting gate : 2.3V  
 Screened Cells:  $7.8 \times 10^5$   
 Sorted Cells:  $5.1 \times 10^4$

### S5A

$>95\%$  of initial population are fluorescent from  $\sim 20,000$  cells.  
 Sort Speed: 2000 droplets/s ( $\sim 7.2 \times 10^5$  cells/hr)  
 Sorted at  $\lambda=0.1$  cells/droplet  
 Mean Brightness =  $1.8 \pm 1.1$  V  
 Brightness sorting gate : 2.7V  
 Screened Cells:  $7.6 \times 10^5$   
 Sorted Cells:  $7 \times 10^4$

### S5B

$>95\%$  of initial population are fluorescent from  $\sim 20,000$  cells.  
 Sort Speed: 2000 droplets/s ( $\sim 7.2 \times 10^5$  cells/hr)  
 Sorted at  $\lambda=0.1$  cells/droplet  
 Mean Brightness =  $1.9 \pm 1$  V  
 Brightness sorting gate : 2.8V  
 Screened Cells:  $7.6 \times 10^5$   
 Sorted Cells:  $2.5 \times 10^4$



**Figure A1.10. mScarlet library sorts with the instrument.** (a) The progressive enrichment of the library across the rounds of sorting. All samples were re-grown from their pre-sort glycerol stocks for expression. (b) The lifetime and brightness screening plots of the post-sort S4A/S4B and S5A/S5B sorting with respect to the wild type m-Scarlet-I population.

## 7c. Clones and photophysical characterization

We plated 5 plates of S5A and S5B respectively. Both S5A and S5B plates displayed mostly pink colonies. We picked 20 colonies of varying brightness with a check on lifetime using our lifetime beam. All picked colonies had phase shifts that were very close to the shifts observed in mScarlet-I. We had also plated one plate from the S4B and S4A. Interestingly one of the S4 plates showed a quick maturing bright colony. On sequencing it turned out to be unique and different from the wild type sequence, we named that clone S4-1 and purified it for further characterization. From the 20 clones we picked from the S5A and S5B plates, we had converged to 2 clones. Both plates had these clones 7:3 in one case and 8:2 in the other. This was sufficient evidence that our library had converged from both pathways. We named these clones S5-1 and S5-2 and selected them for further photophysical characterization. The mutated amino acids of these clones are listed in Table A1.6.

**Table A1.6. The sequence of mScarlet-I mutants in the mutated positions.**

Protein	114	119	121	147	164	174	196	198
mSc-I	E	I	K	T	A	L	D	K
S4-1	L	I	K	T	S	S	D	E
S5-1	E	I	K	T	K	K	N	N
S5-2	L	N	K	R	N	S	D	K

The protein purification and photophysical characterization were described in our previous work. [2] Fluorescence lifetime measurements were made with the PicoQuant TCSPC with a 560nm pulsed picosecond laser head. The photophysics of each RFP is listed in Table A1.7.

**Table A1.7. Photophysical properties of mScarlet-I mutants.**

Protein	QY				Lifetime (ns)			
	(3x trial / Avg)				(3x trial / Avg)			
<b>mSc-I (ref)</b>	54	54	54	54	3.22	3.29	3.27	3.26±0.03
<b>S4-1</b>	40	42	-	41±1.4	2.88	2.93	-	2.90±0.04
<b>S5-1</b>	49	53	49	50.3±2.3	3.00	3.04	3.01	3.02±0.02
<b>S5-2</b>	48	47	-	47.5±0.7	3.00	3.03	-	3.02±0.01



## Appendix 2

This is an addendum to Chapter 2 and parts of this appendix have been adapted from the Supplementary Information section of the article “*Engineering of a Brighter Variant of the FusionRed Fluorescent Protein Using Lifetime Flow Cytometry and Structure-Guided Mutations.*”

**Mukherjee, S.; Hung, S. T.; Douglas, N.; Manna, P.; Thomas, C.; Ekrem, A.; Palmer, A. E.; Jimenez, R.** *Biochemistry* **2020**, *59* (39), 3669–3682.  
<https://doi.org/10.1021/acs.biochem.0c00484>.

### Section 1. Amino acid sequence information for FusionRed variants

10	20	30	40	50	60	70	80	90	100	110	120	
MSKGEELFTGVVPIELVLDGQVNGHKFSVSEGEEDATYKLTAKFICTTG-KLPVPMPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKEFGDTL												avGFPx0
MRSKSNVKEFMRFKVRMEGTNGHFEIEEGEGRPYEGHNTVKLVTKGGPLPFAMDILSPQFYGSVYVYKPADIP---DYKLSFPEGFKWERVMNFDGGVVTVTQDSSLQDGC												DsRedx1
-MS---ELIKENMHMKLYMEGTNNHHFKCTSEGEKPYEGTQTRIKVVEGGPLPFADILATSFNYGSKTFINHTQGIP---DLFKQSFPEGFTWERITTYEDGGVLTATQDSSLQNGCI												eqFP578
-MS---VLTENMHMKLYMEGTNNHHFKCTSEGEKPYEGTQTRIKVVEGGPLPFADILATSFNYGSKTFINHTQGIP---DFFKQSFPEGFTWERITTYEDGGVLTATQDSSLQNGCL												Katushka
-MS---ELIKENMHMKLYMEGTNNHHFKCTSEGEKPYEGTQTRIKVVEGGPLPFADILATSFNYGSKTFINHTQGIP---DFFKQSFPEGFTWERITTYEDGGVLTATQDSSLQDGL												TagRFP
-MS---ELIKENMHMKLYMEGTNNHHFKCTSEGEKPYEGTQTRIKVVEGGPLPFADILATSFNYGSKTFINHTQGIP---DFFKQSFPEGFTWERITTYEDGGVLTATQDSSLQDGL												mKate5
MVS---ELIKENMHMKLYMEGTNNHHFKCTSEGEKPYEGTQTRIKVVEGGPLPFADILATSFNYGSKTFINHTQGIP---DFFKQSFPEGFTWERITTYEDGGVLTATQDSSLQDGL												mKate2
MVS---ELIKENMHMKLYMEGTNNHHFKCTSEGEKPYEGTQTRIKVVEGGPLPFADILATSFNYGSKTFINHTQGIP---DFFKQSFPEGFTWERITTYEDGGVLTATQDSSLQDGL												mKate25
MVS---ELIKENMHMKLYMEGTNNHHFKCTSEGEKPYEGTQTRIKVVEGGPLPFADILATSFNYGSKTFINHTQGIP---DFFKQSFPEGFTWERITTYEDGGVLTATQDSSLQDGL												FusionRed
MVS---ELIKENMHMKLYMEGTNNHHFKCTSEGEKPYEGTQTRIKVVEGGPLPFADILATSFNYGSKTFINHTQGIP---DFFKQSFPEGFTWERITTYEDGGVLTATQDSSLQDGL												FusionRedM
MVS---ELIKENMHMKLYMEGTNNHHFKCTSEGEKPYEGTQTRIKVVEGGPLPFADILATSFNYGSKTFINHTQGIP---DFFKQSFPEGFTWERITTYEDGGVLTATQDSSLQDGL												FusionRedV
MVS---ELIKENMHMKLYMEGTNNHHFKCTSEGEKPYEGTQTRIKVVEGGPLPFADILATSFNYGSKTFINHTQGIP---DFFKQSFPEGFTWERITTYEDGGVLTATQDSSLQDGL												FusionRedQ
MVS---ELIKENMHMKLYMEGTNNHHFKCTSEGEKPYEGTQTRIKVVEGGPLPFADILATSFNYGSKTFINHTQGIP---DFFKQSFPEGFTWERITTYEDGGVLTATQDSSLQDGL												FusionRedMQ
MVS---ELIKENMHMKLYMEGTNNHHFKCTSEGEKPYEGTQTRIKVVEGGPLPFADILATSFNYGSKTFINHTQGIP---DFFKQSFPEGFTWERITTYEDGGVLTATQDSSLQDGL												FusionRedMV
1	10	20	30	40	50	60	70	80	90	100	110	
VNRTELKGDIGKEDGNILGHK-LEYNYSNHYVIMADQKNGIKVNFKIRVNIEDGGSVQLADHYQNTPIGDGPVLLPDNHLSTQSALSKDPNEKRDHMLVLEFVTAAGITHGHDELTK												avGFPx0
IYVKVFLGVNFPSSGPMQKKTGLWEASTETLYPADGGLEGADMAKLVGGG-HLIONETTYRSKKPATN---LPGYVYVDSKLDITSHNEDYT-IVEQYERTEGRHHLFL-----												DsRedx1
IYVKVINGVNFPSNGPMQKKTGLWEASTETLYPADGGLEGADMAKLVGGG-YLHCSLTKTYRSKKPATN---LKMPGFYFVDRRLRIKEADKET--YVEQHEVAVARYCDLPSKLGH												eqFP578
IYVKVINGVNFPSNGPMQKKTGLWEASTETLYPADGGLEGADMAKLVGGG-YLHCSLTKTYRSKKPATN---LKMPGFYFVDRRLRIKEADKET--YVEQHEVAVARYCDLPSKLGH												Katushka
IYVKVIRGVNFPSSGPMQKKTGLWEASTETLYPADGGLEGADMAKLVGGG-HLIONETTYRSKKPATN---LKMPGVYVDRRLRIKEADKET--YVEQHEVAVARYCDLPSKLGH												TagRFP
IYVKVIRGVNFPSSGPMQKKTGLWEASTETLYPADGGLEGADMAKLVGGG-HLIONETTYRSKKPATN---LKMPGVYVDRRLRIKEADKET--YVEQHEVAVARYCDLPSKLGH												mKate5
IYVKVIRGVNFPSSGPMQKKTGLWEASTETLYPADGGLEGADMAKLVGGG-HLIONETTYRSKKPATN---LKMPGVYVDRRLRIKEADKET--YVEQHEVAVARYCDLPSKLGH												mKate2
IYVKVIRGVNFPSSGPMQKKTGLWEASTETLYPADGGLEGADMAKLVGGG-HLIONETTYRSKKPATN---LKMPGVYVDRRLRIKEADKET--YVEQHEVAVARYCDLPSKLGH												mKate25
IYVKVIRGVNFPSSGPMQKKTGLWEASTETLYPADGGLEGADMAKLVGGG-HLIONETTYRSKKPATN---LKMPGVYVDRRLRIKEADKET--YVEQHEVAVARYSTGGAGDGGK												FusionRed
IYVKVIRGVNFPSSGPMQKKTGLWEASTETLYPADGGLEGADMAKLVGGG-HLIONETTYRSKKPATN---LKMPGVYVDRRLRIKEADKET--YVEQHEVAVARYSTGGAGDGGK												FusionRedM
IYVKVIRGVNFPSSGPMQKKTGLWEASTETLYPADGGLEGADMAKLVGGG-HLIONETTYRSKKPATN---LKMPGVYVDRRLRIKEADKET--YVEQHEVAVARYSTGGAGDGGK												FusionRedV
IYVKVIRGVNFPSSGPMQKKTGLWEASTETLYPADGGLEGADMAKLVGGG-HLIONETTYRSKKPATN---LKMPGVYVDRRLRIKEADKET--YVEQHEVAVARYSTGGAGDGGK												FusionRedQ
IYVKVIRGVNFPSSGPMQKKTGLWEASTETLYPADGGLEGADMAKLVGGG-HLIONETTYRSKKPATN---LKMPGVYVDRRLRIKEADKET--YVEQHEVAVARYSTGGAGDGGK												FusionRedMQ
IYVKVIRGVNFPSSGPMQKKTGLWEASTETLYPADGGLEGADMAKLVGGG-HLIONETTYRSKKPATN---LKMPGVYVDRRLRIKEADKET--YVEQHEVAVARYSTGGAGDGGK												FusionRedMV
120	130	140	150	160	170	180	190	200	210	220	230	233

**Table A2.1. Sequence information for the FR family with respect to the parental clones from the eqFP578 family.** [3] The amino acid sequences of avGFP and DsRed are also listed in the table for comparison. [4, 5] The relevant mutations in this study have been highlighted in gold. The chromophore residues are highlighted in blue all other colors are based on the custom classification for the nature of amino acid sidechains using the suite Clustal-W. [6] The numbering above is with respect to avGFP while below is with respect to FR. All mutation position numbers in the text are with respect to FR numbering in this table.



## Section 2. Amino acid sequence results from the site-directed library

### 2a. Selection of the C159V clone

The eight distinct mutants were screened for brightness in yeast cells in the green and red channel on a BD-FACS Celesta flow cytometer at the BioFrontiers Flow Cytometry core facility at the University of Colorado, Boulder. The green channel is an indicator for the undesirable immature green fluorescence peak seen in RFPs such as SDC-5 (an unpublished mCherry mutant with mutations of W143L, I161T, Q163C, I197R using mCherry sequence numbering). Filtered cell lysates of these mutants were also used for measuring lifetime on a time-correlated single photon counting (TCSPC) system that provided us with lifetimes to select the clones of interest. Sequence, lifetime and brightness data are summarized in Table A2.2.

**Table A2.2. Sequence and screening results of the FR site-directed (FSD) library clones.** The lifetime was measured using cell lysates on a TCSPC system. All fluorescence measurements were normalized to FR.

Clones	159	161	Lifetime (ns)	Green FI (a.u.)	Red FI (a.u.)
FR	C	M	1.78	100	100
mCherry	-	-	1.67	103	191
SDC-5	-	-	-	811	421
FR-1	-	-	-	185	352
FSD-4	T	M	1.70	112	201
FSD-5	?	?	1.94	100	193
FSD-9	V	M	2.03	163	232
FSD-11	L	M	1.21	107	177
FSD-14	A	H	1.01	147	68
FSD-15	T	M	1.73	105	134
FSD-18	T	I	1.70	111	154
FSD-19	G	M	1.94	105	87

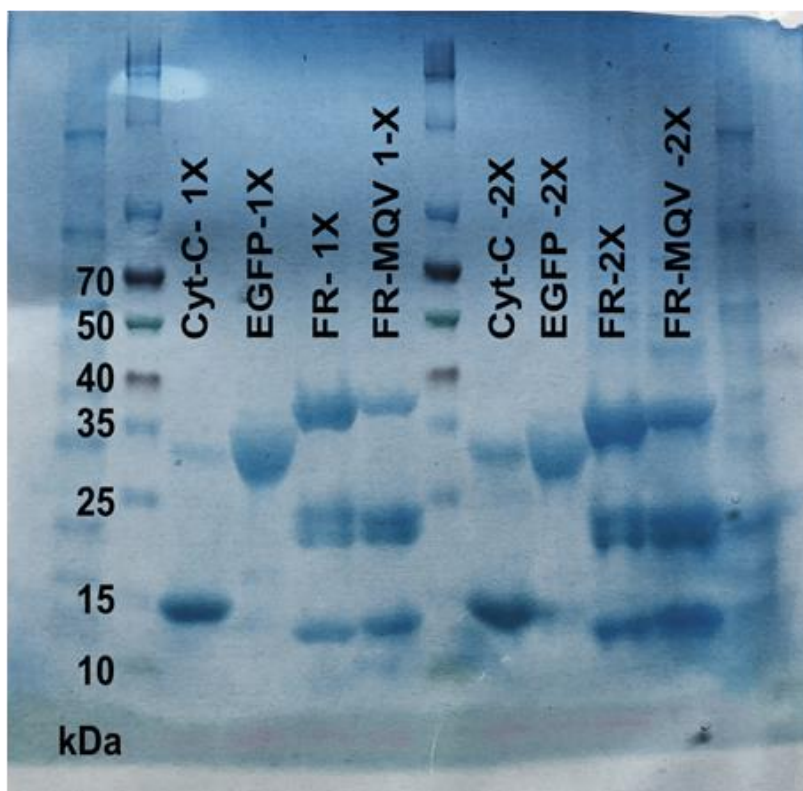
## Section 3. Extinction coefficient and quantum yield estimation

### 3a. Extinction coefficient calculation based on SDS-PAGE results

*Calculation of extinction coefficient based on published observations of FPs exhibiting backbone cleavage, including FusionRed [9]:*

$$\epsilon_{\text{RFP}} = \frac{\text{Abs}_{\text{max RFP}}}{\left(\frac{\text{Abs}_{380\text{nm}}}{\epsilon_{380\text{nm}}}\right) + \left(\frac{\text{Abs}_{450\text{nm}}}{\epsilon_{450\text{nm}}}\right)}$$

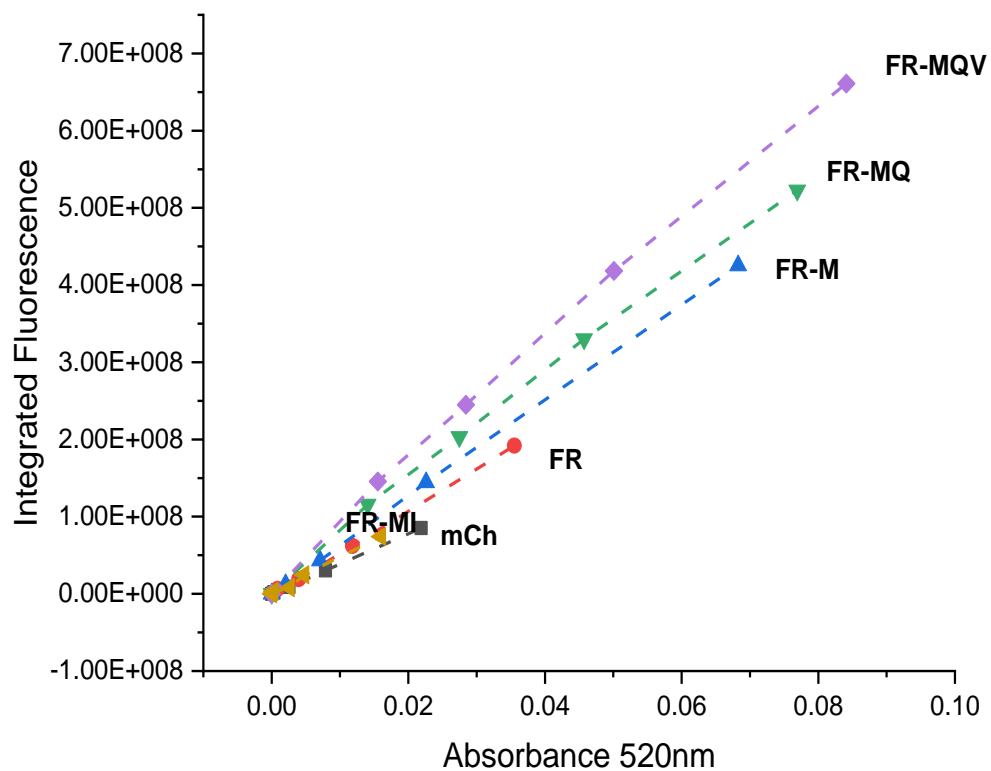
Where  $\text{Abs}_{\text{max RFP}}$  is the maximum absorbance for the RFP in the undenatured absorption spectrum and  $\text{Abs}_{380\text{nm}}$  and  $\text{Abs}_{450\text{nm}}$  were the values of absorbance at 380 nm and 450 nm in the alkali denatured spectrum. We used previously reported values of  $\epsilon_{380\text{nm}} = 70,500 \text{ M}^{-1}\text{cm}^{-1}$  and  $\epsilon_{450\text{nm}} = 44,000 \text{ M}^{-1}\text{cm}^{-1}$ . [2, 9] SDS-PAGE was used to assess the purity of proteins extracted from *E. coli* and to examine whether FR-MQV exhibited backbone cleavage like the parent FR such that the mathematical formula presented above could be used to calculate the extinction coefficient of the RFP. Proteins were boiled at 95°C for 10 minutes in the presence of SDS and DTT before loading onto the gel at 1X and 2X (~10  $\mu\text{M}$ ) concentrations. Also, to confirm that boiling alone was not responsible for the cleavage observed, a control gel was performed without boiling the samples (data not shown). Both gels showed the existence of cleaved fractions for FR mutants and none for EGFP (control). Cytochrome-C (Cyt-C) was included as an additional size marker. The results for FR are consistent with previously reported data, therefore we used the above-mentioned relationship to estimate the extinction coefficient of FR and its family of mutants.



**Figure A2.1. Gradient (4–20%) SDS-PAGE of FR and FR-MQV.** Gels with appropriate controls show backbone cleavage for FR-MQV. Full FP fractions are bands located ~25k Da, while the cleaved fractions are seen as bands at 16 kDa and 9 kDa for FR mutants.

### 3b. Quantum yield measurements

The quantum yield for purified protein samples was measured as described in the Methods and Materials. The data follows expected linear trends as shown in Figure A2.2. Table A2.3 presents the number of independent trials and the standard deviation error observed in measurements for each FP.



**Figure A2.2. Quantum yield measurements.** The linear trend of higher integrated fluorescence against absorption measurements under serial dilution for one of the independent measurements. Higher slope values indicate higher quantum yields.

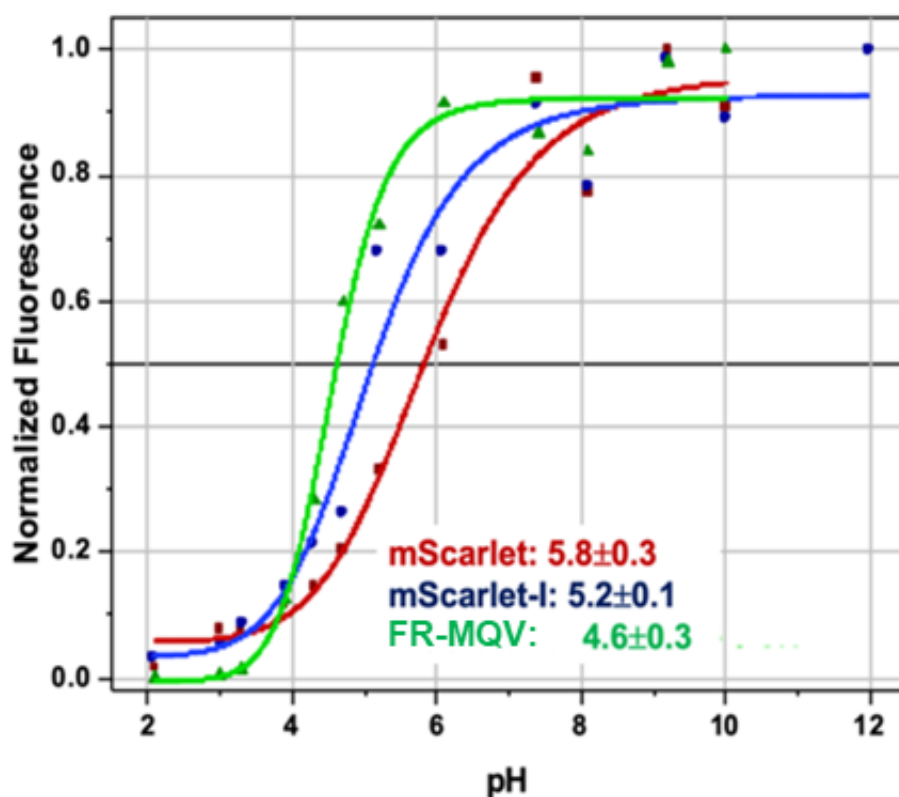
**Table A2.3. Measured mean quantum yield and standard deviation errors.** Quantum yield measurements are susceptible to random and systematic errors. Cranfill *et al.* described protocols to minimize such errors. [7] Steps such as pH control, minimizing temperature fluctuations, gentle thawing of FPs and usage of fresh protein were taken into consideration. Replicate measurements for quantum yield of relevant proteins were taken from batches of FPs prepared through independent transformations, growth and purification protocols. Cresyl Violet, mCherry and mScarlet were usually used as multiple references for the measurements.

<b>Protein</b>	<b>Mean QY</b>	<b>Trials</b>
FR-MI	26±4	3
FR-MQ	43±3	4
FR-MQV	53±3	4
FR-Q	33±1	4
FR	24±4	4
FR-M	34±2	4
mCherry	22 (ref)	4
mKate	33	1
mKate-42Q	17	1
mRuby3	43	1
mScarlet	72±4	3
mScarlet-I	59±2	3
FR-V	31±4	3
FR-MV	38±3	3

## Section 4. Additional photophysical and biochemical characterization

### 4a. pKa measurements

Table A2.4 shows the pKa values of relevant mutants in this study. The pKa was obtained by taking three measurements for each pH data point and fitting the curve to a sigmoidal function (Figure A2.3). The half value of the maximum fluorescence was calculated to be the pKa of the protein. The traces for mScarlet, mScarlet-I and FR-MQV are shown in Figure A2.3. Both mScarlet and mScarlet-I scale well with the reported values by Bindels *et al.* [8]



**Figure A2.3. pKa measurements.** Plot of fluorescence versus pH for FR-MQV, mScarlet and mScarlet-I. The pKa is obtained from the 0.5 fractional fluorescence point. FR-MQV retains the low pKa value of the FR family of proteins.

**Table A2.4. The pKa values measured in this study versus the values reported for known RFPs in the literature.**

Protein	pKa Calculated	pKa Reported
FR	4.5±0.2	4.5 [Ref 9]
FR-M	4.7±0.3	4.8 [Ref 2]
FR-Q	4.3±0.2	-
FR-MQ	4.4±0.1	-
FR-MQV	4.6±0.2	-
mScarlet	5.8±0.3	5.3 [Ref 8]
mScarlet-I	5.2±0.1	5.4 [Ref 8]

#### **4b. Lifetime measurements using time-correlated single photon counting (TCSPC)**

Fluorescence lifetime was measured using TCSPC by exciting with a 560 nm pulsed laser and collecting emission in two spectral windows using appropriate bandpass filters. Differences in the amplitudes and time constants from multi-exponential fits of the decay of mKate-42Q are consistent with the presence of multiple emitting species. Data are presented in Table A2.5.

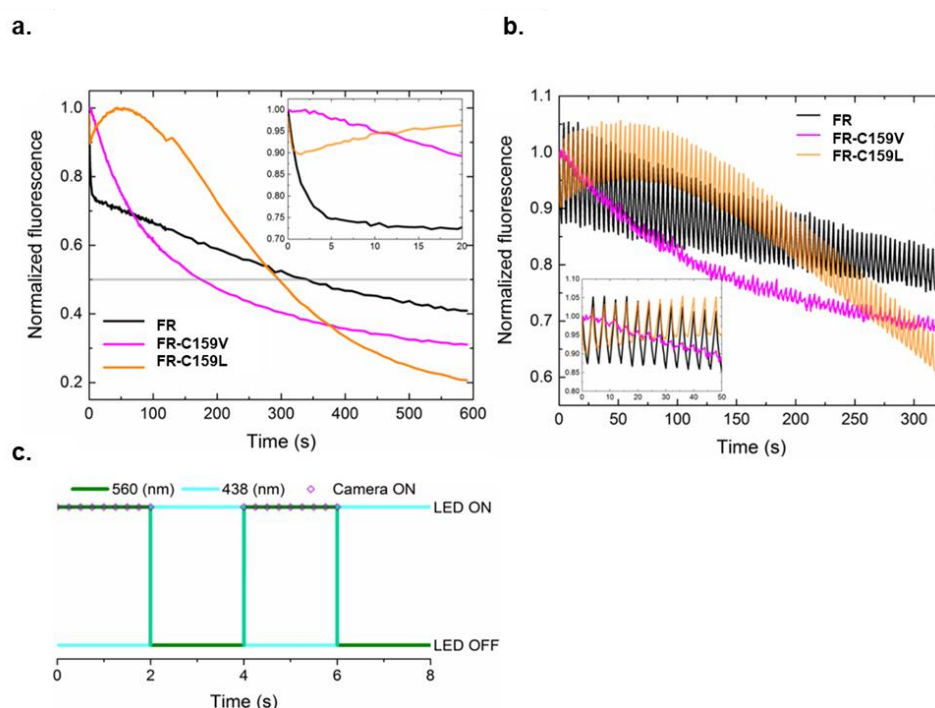
**Table A2.5. TCSPC-based lifetime measurements using pure protein samples.** Independent trials are separated by commas, where  $\tau$  indicates the component of fluorescence lifetime in ns and the % intensity for that component is reported in brackets. As the L175M, C159V and M42Q mutations are added on FR, there is a progressive increment of  $\tau$  in the major component of fluorescence decay. The decays for mKate-42Q are fit to a tri-exponential function and are significantly different in the two emission spectral windows. In contrast, fluorescence decays of the other FPs are fit with bi-exponential functions and do not show significantly different decays in the two emission windows.

Protein	Emission window (600 ± 30 nm)				Emission window (670 ± 30 nm)			
	$\tau_1$ (%Int)	$\tau_2$ (%Int)	$\tau_3$ (%Int)	$\tau_{av}$	$\tau_1$ (%Int)	$\tau_2$ (%Int)	$\tau_3$ (%Int)	$\tau_{av}$
<b>FR</b>	1.40(64), 1.44(63), 1.42(64)	2.49(36), 2.51(37), 2.51(37)	-	1.78±0.04	1.37(39)	2.36(61)	-	1.76
<b>FR-M</b>	1.55(38), 1.56(33), 1.49(33)	2.43(62), 2.47(67), 2.43(67)	-	2.13±0.05	1.70(53)	2.54(47)	-	2.09
<b>FR-Q</b>	1.50(55), 1.52(53), 1.56(52)	2.92(45), 2.93(47), 2.97(48)	-	2.10±0.05	2.93(44)	1.51(56)	-	2.13
<b>FR-MQ</b>	1.74(35), 1.87(40), 2.05(40)	2.82(65), 2.91(60), 3.02(56)	-	2.45±0.08	2.98(56)	1.85(44)	-	2.47
<b>FR-MI</b>	2.22(55), 2.21(54)	1.21(45), 1.29(46)	-	0.26±0.04	2.16(56)	1.16(44)	-	1.72
<b>FR-MQV</b>	3.23(68), 3.15(69), 3.14(69)	2.03(32), 1.94(31), 2.01(31)	-	2.77±0.07	3.13(73)	1.91(27)	-	2.80
<b>mKate</b>	2.40(92), 2.42(90), 2.37(93)	0.87(8), 0.86(10), 0.80(8)	-	2.26±0.07	2.38(92)	8.21(8)	-	2.25
<b>mKate-Q</b>	3.72(52), 3.61(48), 3.65(55)	1.29(40), 1.25(41), 1.32(39)	0.21(8), 0.19(9), 0.46(6)	2.16±0.05	3.62(36)	1.27(61)	0.24(3)	2.05
<b>mCherry</b>	1.63(88), 1.75(70)	0.59(12), 1.06(30)	-	1.67±0.07	-	-	-	-
<b>mScarlet</b>	3.82(98), 3.80(97), 3.93(99)	1.27(2), 1.18(3), 1.31(1)	-	3.87±0.07	-	-	-	-
<b>mScarlet-I</b>	3.41(92), 3.57(86), 3.51(88)	0.92(8), 1.63(14), 1.51(12)	-	3.26±0.07	-	-	-	-
<b>mRuby3</b>	2.79(95)	1.16(5)	-	2.71	-	-	-	-



## 4c. Photobleaching experiments

The photobleaching traces of FR, FR-C159V and FR-C159L are shown in Figure A2.4a. The fast fluorescence decay process may result from transition to a dark state, as has been observed previously for FPs. [10] The incorporation of a small aliphatic group at position 159 in FR-C159V eliminates the tendency of the FP to undergo the dark state conversion process. Thus, the photobleaching profile changed from a bi-exponential-like decay in FR to a mono-exponential-like decay in FR-C159V and FR-C159L. It is worth noting that FR-C159L shows a slow photoactivation process in ~50 s under the irradiance regime used in this study. The tendency for photo-switching, or reversible photobleaching, is the lowest for FR-C159V, followed by FR-C159L, and the highest in FR, suggesting that an aliphatic group at the 159 position may help to reduce the cis–trans isomerization of the chromophore.

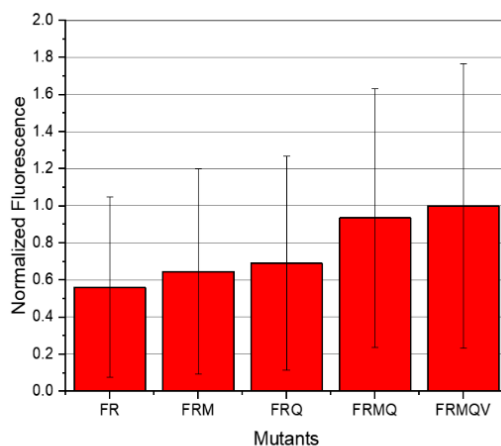


**Figure A2.4. Photobleaching measurements.** (a.) FR, FR-C159V and FR-C159L expressed in bacteria were photobleached with 560 nm LED at a constant irradiance of  $\sim 4.9 \text{ W/cm}^2$ . The inset exhibits the traces of the first 20 seconds. (b.) Fluorescence signals under alternating illumination—each cell was illuminated with 560 nm LED for 2 s, then only 438 nm LED for another 2 s and then the illumination cycle was repeated. The camera only recorded signals when 560 nm LED was turned on. The inset shows the traces of the first 50 minutes. (c.) The alternating

LED light sources are illustrated for the first 8 s. The irradiance was  $\sim 4.9 \text{ W/cm}^2$  for 560 nm and  $\sim 4.5 \text{ W/cm}^2$  for 438 nm throughout the experiment. The camera recorded nine frames in 2 s when the 560 nm LED was turned ON.

#### 4d. Bacterial brightness

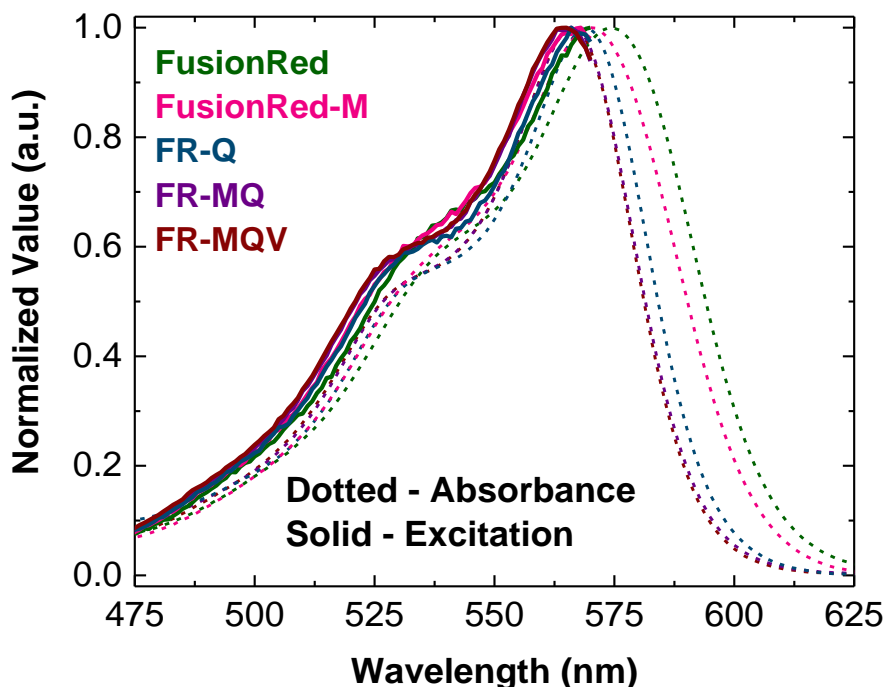
The family of FR mutants relevant to this study was expressed in bacteria to test bacterial brightness in a droplet microfluidic screening platform (summarized in Figure A2.5). [11] Brightness seems to scale well with increasing molecular brightness with FR-MQV being  $\sim 2$ -fold brighter than FR. Details of the experiment are described in the methods and materials section of the Chapter 3 of the article. Note that quantitative comparison of brightness across cellular systems is impacted by many factors. Whereas *in vitro* brightness is dependent solely on molecular parameters whereas in cells the brightness is additionally the result of multiple processes that involve DNA to mRNA transcription, mRNA stability, translational efficiencies, protein folding kinetics, chromophore maturation, cellular pH, and protein stability. This makes a quantitative comparison across systems not possible. Discrepancies in the levels of brightness across systems (molecular, bacterial, yeast and mammalian) have been discussed in our previous publication and by others. [2, 13]



**Figure A2.5. Bacterial brightness assay.** The brightness (normalized with respect to FR-MQV) recorded for the family of FR mutants relevant to this study. We see a progressive increase in bacterial brightness for the family, with FR-MQV being  $\sim 2$ -fold brighter than FR.

#### 4e. Excitation versus absorption spectra

Excitation and absorption spectra overlaid with each other for FR and the FR mutants generated in this study. In Chapter 3, our analysis for absorption wavelengths, calculations of Stokes shift, etc., was done with respect to the absorption spectra for the relevant FPs.



**Figure A2.6. Absorption and excitation spectra.** Overlaying the absorption and the excitation spectra for the FPs relevant to this study. The excitation spectra (solid) and absorption spectra (dotted). The y-axis indicates the normalized value of absorption/excitation in normalized arbitrary units (a.u.).

#### 4f. Spectral characteristics

We observed that both the absorption and emission spectral width decreased as mutations were progressively incorporated into FR in the development of FR-MQV. The spectra progressively narrow with decreases in the absorption and emission FWHM (full width at half maximum). Table A2.6 below summarizes the spectral changes that we observed in this series of mutants.

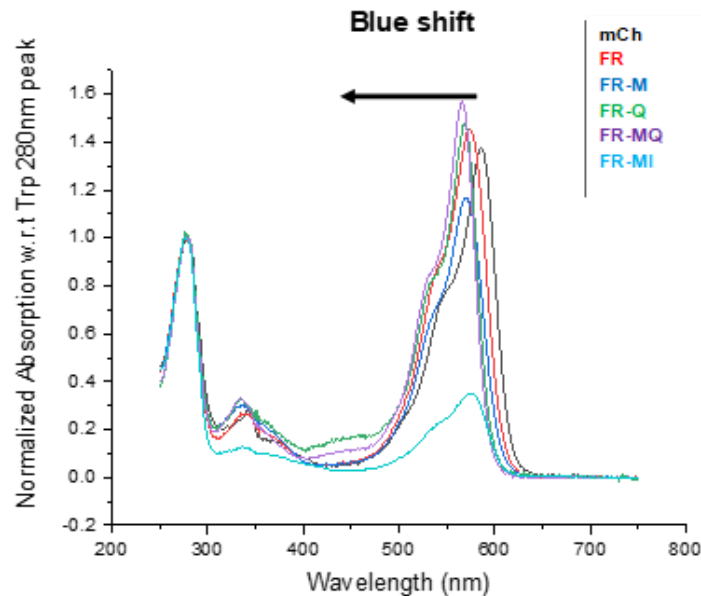
**Table A2.6. Spectral characteristics of the FR family of proteins investigated in this study.**

There is a progressive blue shift and narrowing of the spectra as one moves from the single, to the double and then to the final triple mutant FR-MQV. The change in Stokes shift (both in nm and energy units), unlike mKate-Q, is minimal for the M42Q mutants of FR.

<b>Protein</b>	<b><math>\lambda_{\text{abs\_max}}</math> (nm)</b>	<b>FWHM absorption (nm)</b>	<b><math>\lambda_{\text{em\_max}}</math> (nm)</b>	<b>FWHM emission (nm)</b>	<b>Stokes shift (nm)</b>	<b>Stokes shift (cm<sup>-1</sup>)</b>
FR-MQV	566	53	585	39	19	574
FR-MQ	567	54	586	40	19	572
FR-MV	569	60	591	47	22	654
FR-V	573	63	594	53	21	617
FR-Q	568	55	587	40	19	570
FR-M	571	61	590	52	19	595
FR	574	63	596	54	22	643

#### **4g. Normalization with respect to the 278nm Trp peak**

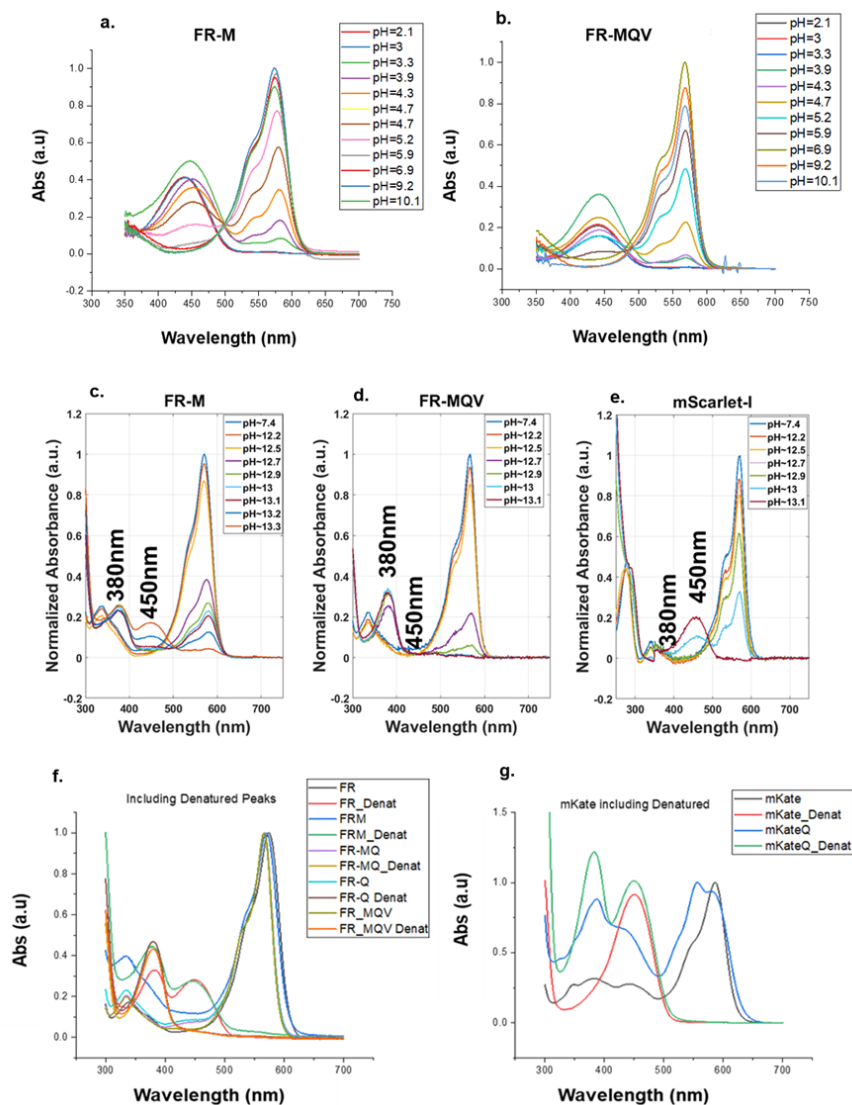
FR and FR mutants with 42Q have the same number of Trp residues. Normalizing with respect to the absorption peak of tryptophan at ~280 nm should reflect the true nature of spectral width for the RFP peak. On doing so we observe the red emission peak is indeed narrower and shifted blue relative to the parent protein.



**Figure A2.7. Trp280 nm peak for FusionRed variants.** Normalizing the absorption spectra with respect to the Trp 280 nm feature reveals that M42Q changes the peak absorption. FR-MQ has a ~20% higher peak absorbance compared to FR.

#### 4h. pH titrations and alkali denaturation

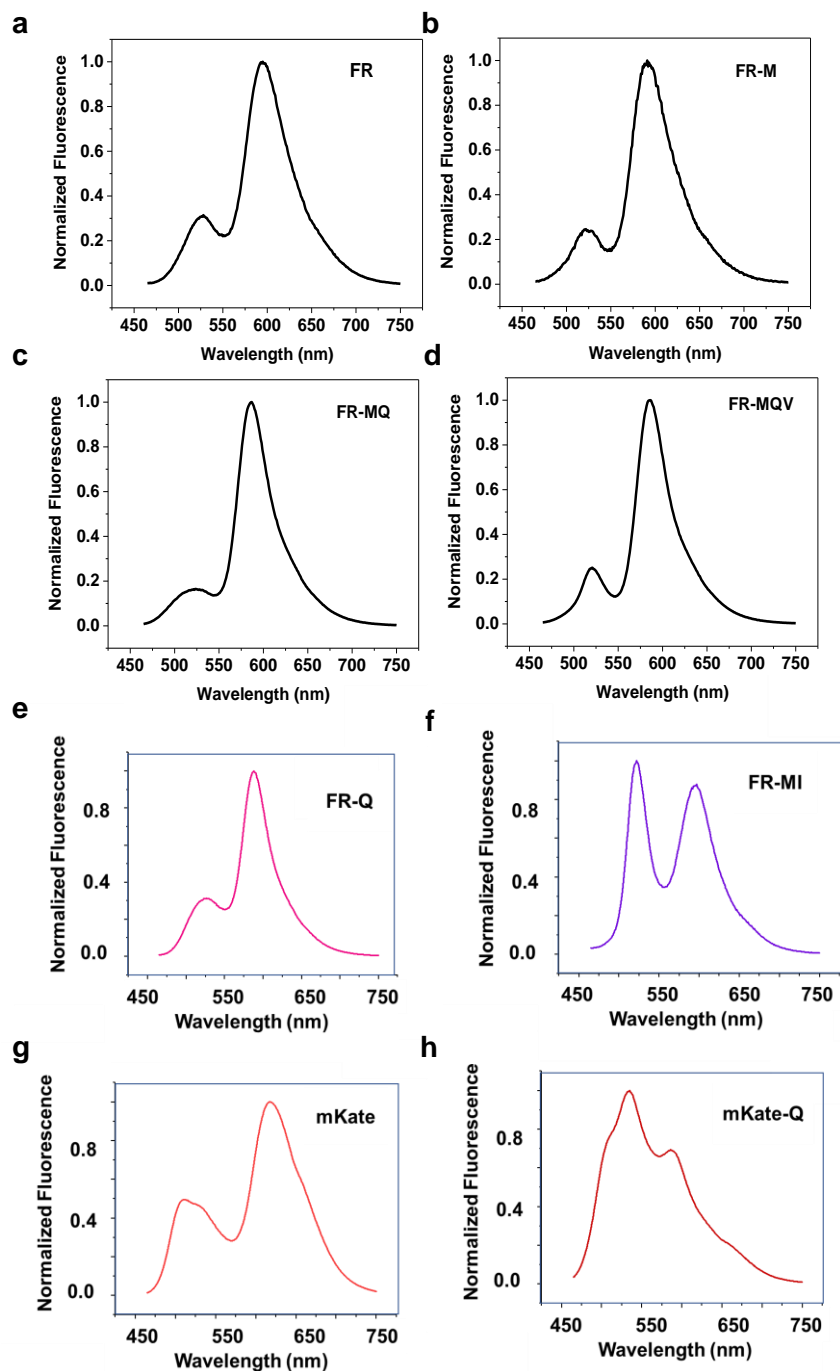
We collected absorption spectra for FPs *in vitro* at pH values 2–14. At acidic pH values, FR-MQV resembles FR and FR-M, but at basic pH values, FR-MQV lacks the 450 nm degradation product that is present in FR-M and many other red FPs, such as mScarlet-I. Instead, there is a single product of alkali denaturation at 380 nm whenever the 42Q mutation is incorporated into the FR protein. However, this effect is not seen in mKate-Q, where we see both the 380 and the 450 nm peaks. mKate and other RFPs, such as mScarlet-I (also mCherry, mScarlet, mRuby3; data not shown), degrade to only the 450 nm hydrolysis product.



**Figure A2.8. pH titrations.** from pH 2 to 10 of (a) FR-M and (b) FR-MQV show similar behavior for the functional form and the acid-degraded form of the FPs. Titration in the basic range (pH >12) shows that (c) FR-M breaks down into two products of alkali hydrolysis, with the 380 nm peak formed at a lower pH than the 450 nm peak; (d) FR-MQV displays only one product of alkali degradation at 380 nm and; (e) mScarlet-I (along with other RFPs) shows only the 450 nm peak for alkali degradation. (f) All constructs with the M42Q mutation behave similarly, but such an effect is not seen in (g) mKate-Q where both the 450 and the 380 nm peaks are observed on denaturation. The y-axis indicates the normalized value of absorption in normalized arbitrary units (a.u.).

#### **4i. Green emission peak**

The formation of some RFP chromophores is accompanied by the formation of a green chromophore that lacks the extension of the acylimine moiety. [12] This is undesirable as it decreases the amount of mature red species present in a protein sample. Protein engineering efforts, particularly those that involve mutations internal to the beta barrel often disrupt the chromophore maturation pathway and hence inadvertently result in an increase in the immature or green species. FR mutants developed in this work employed multiple internal mutations; emission spectra were collected in the 465–750 nm windows, with excitation at 450 nm for Figure A2.9. It is evident that the M42Q mutation does not increase the green emission peak relative to FR (both show a shoulder ~20% of the major red peak) suggesting a properly mature red chromophore. However, the same mutation in mKate changes the spectrum considerably. Even the incorporation of a similar-sized, but aliphatic residue, isoleucine, produces a large green emission peak.

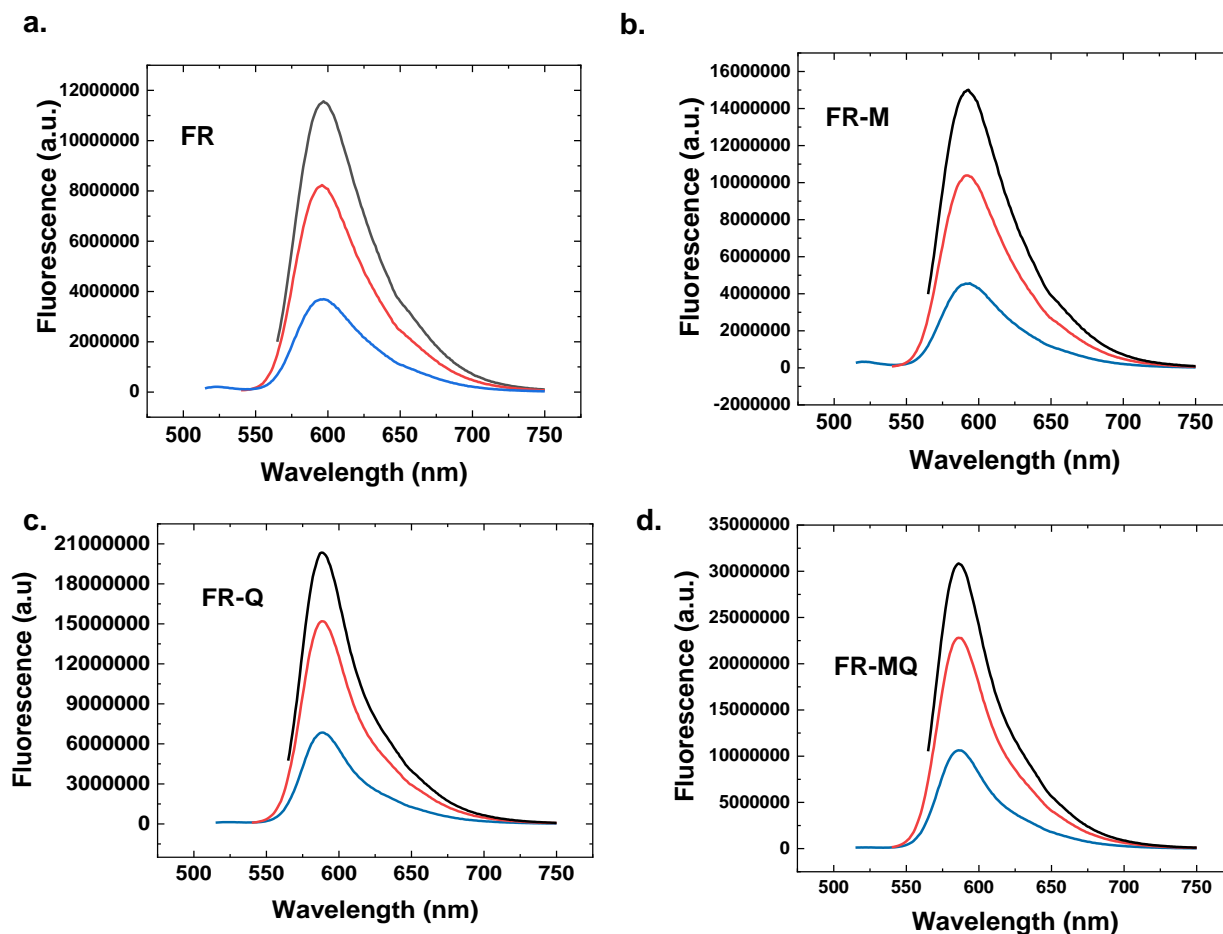


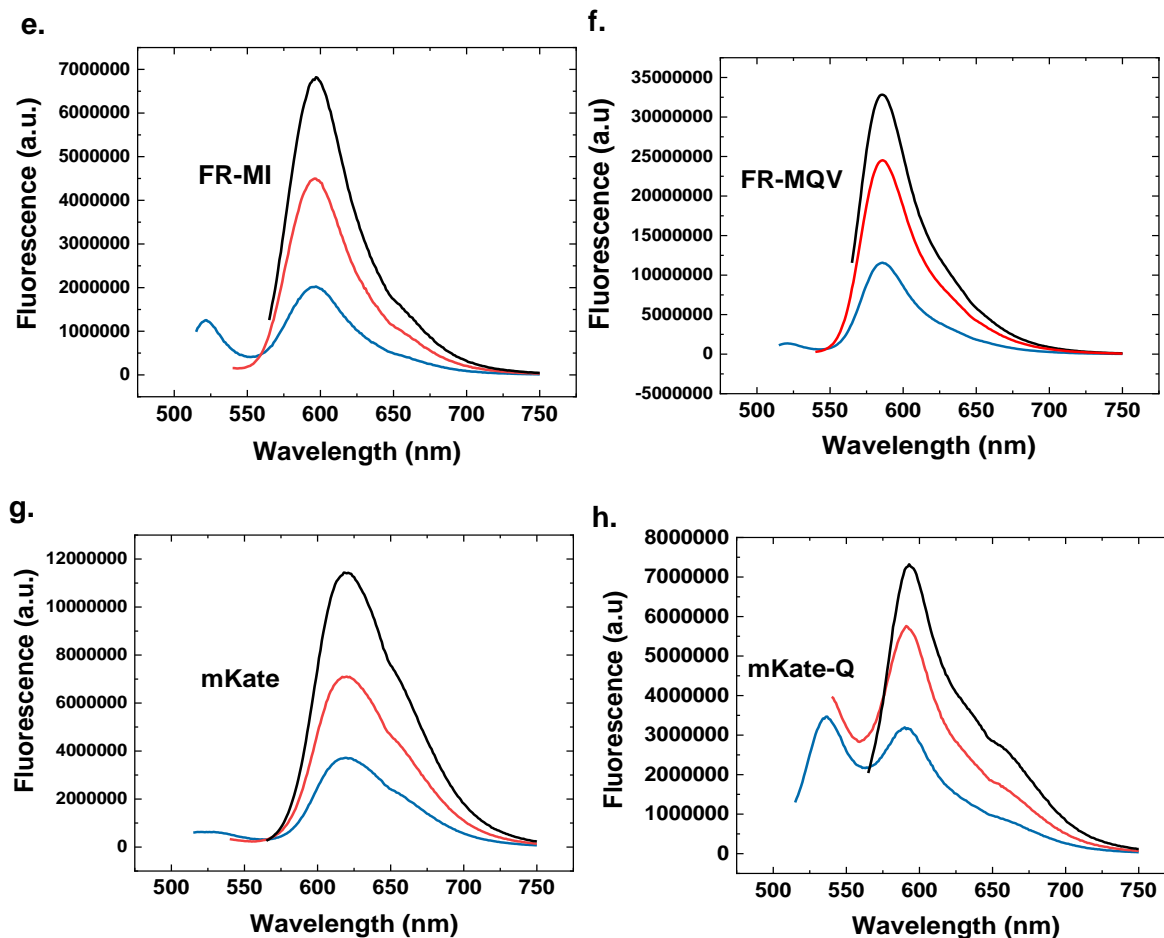
**Figure A2.9. Emission spectra for the FusionRed and mKate variants.** Emission spectra of (a.) FR, (b.) FR-M, (c.) FR-MQ, (d.) FR-MQV, (e.) FR-Q, (f.) FR-MI, (g.) mKate, and (h.) mKate-Q excited with 450 nm light. FR mutants except FR-MI exhibit minimal green fluorescence, suggesting proper red chromophore maturation. mKate-Q also displays a significant green peak.



#### 4j. Excitation-dependent emission

Konold and co-workers previously demonstrated the existence of multiple non-interconverting hydrogen-bonded conformations of TagRFP-675 and mKate-Q. [14] We excited the proteins in this study at 500 nm, 525 nm and 550 nm to find evidence for different emission species in the red window as shown in Figure A2.10. Although mKate-Q shows excitation-dependence of the emission spectral shape (Figure A2.10h), mKate and the FR variants do not show significant excitation-wavelength dependence of their emission spectra.





**Figure A2.10. Excitation dependent emission spectra.** Emission spectra of (a.) FR, (b.) FR-M, (c.) FR-Q, (d.) FR-MQ, (e.) FR-MI, (f.) FR-MQV, (g.) mKate, and (h.) mKate-Q, with excitation at 500 (blue), 525 (red) and 550 (black) nm. Consistent with previous results, mKate-Q has multiple emissive species.

## Section 5. Cellular Assays: brightness, maturation and cytotoxicity

### 5a. Brightness assays

Brightness was measured using FACS and confocal microscopy. The Methods and Materials section of Chapter 3 describes the measurement protocols. Table A2.7 reports the values and number of biological replicates used for measuring the brightness through FACS. Each biological replicate had ~3 technical replicates of ~10000 HeLa cells each. Table A2.8 reports the number of HeLa cells analyzed in each sample dish and the mean intensity of each dish.

**Table A2.7. Mean brightness from FACS measurements with standard deviation error from multiple biological replicates.** Mean brightness measurements were normalized to FR.

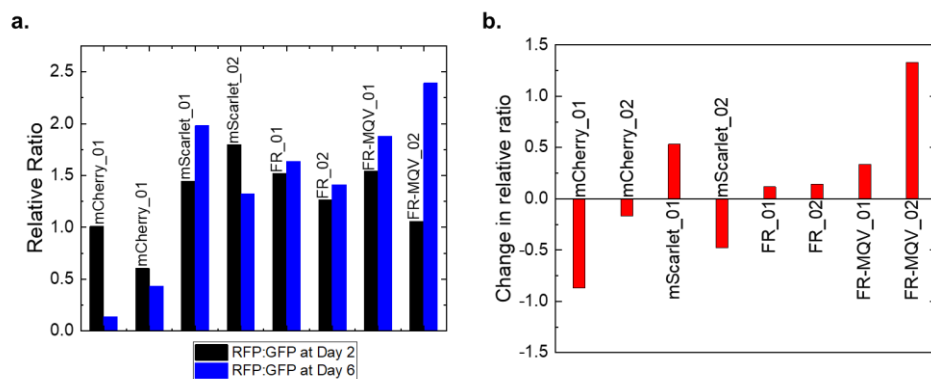
Protein	Biological Replicates	Mean Brightness
FR	5	100±0
FR-M	4	191±31
FR C159V	2	164±17
FR-M C159V	2	266±31
FR- Q	2	214±18
FR-MQ	3	315±43
FR-MQV	4	509±20
mCherry	5	176±39
mScarlet	4	715±51
mScarlet-I	1	355

**Table A2.8. Mean brightness from confocal microscopy measurements with standard deviation error from the number of cells indicated in each dish.** Mean intensity measurements were normalized to FR.

Protein	# Cells	Mean Intensity (x100)
mCherry-1	51	77±55
mCherry-2	38	171±137
FR-1	12	100±68
FR-2	61	100±78
mScarlet-1	131	250±194
mScarlet-2	71	379±273
FR-MQV-1	62	231±154
FR-MQV-2	48	215±187
Untransfected-1	513	4±2
Untransfected-2	469	5±3

## 5b. Cytotoxicity assay

The detailed protocol for this assay is reported in the Materials and Methods section of Chapter 3. Briefly, mammalian cells expressing either EGFP or one of the RFP clones were mixed in a 50:50 ratio by volume. Initially, cells expressing EGFP and each RFP were FACS screened individually (~5000 cells) as controls. Part of mixture was analyzed by FACS to quantify the number of cells carrying EGFP and the RFP at the start (Day 2). The remaining mixture was re-plated. The screens for the re-plated mixtures were repeated after another 4 days of growth (Day 6). Ratios of RFP:EGFP mixtures were calculated for technical replicates for Day 2 and Day 6. The change in the ratio of RFP to EGFP for each replicate is a measure of the relative cytotoxicity of the RFP to EGFP. Figure A2.11 is a graphical representation of the assay. We found that FR and FR-MQV are reproducibly less cytotoxic than EGFP. In contrast, mCherry was consistently more cytotoxic than EGFP, whereas for mScarlet, in one case the RFP was more cytotoxic than EGFP and in another case it was less cytotoxic.



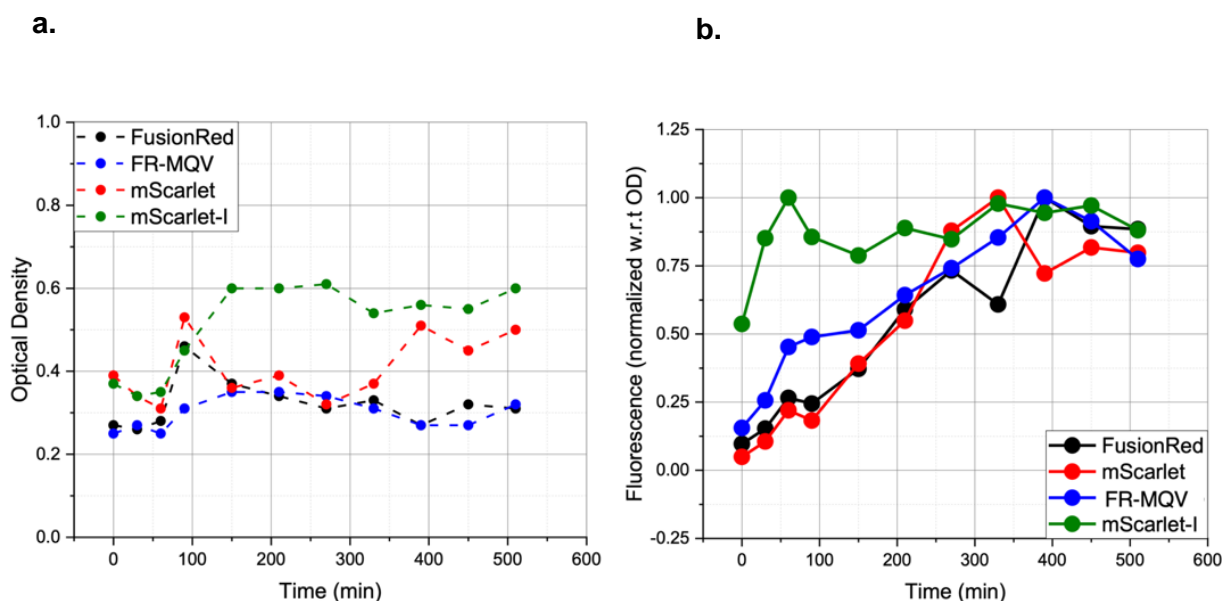
**Figure A2.11. Cytotoxicity assay (a.)** The black bars represent the RFP:EGFP ratio in a mix on Day 2. The blue bars represent the RFP:EGFP ratios measured on Day 6. On each day cells were measured using FACS. **(b.)** The change in the ratio is represented as red bars. FR and FR-MQV are consistently less cytotoxic than EGFP.

**Table A2.9. Number of cells analyzed on Day 2 and Day 6 in the GFP and the RFP channel using FACS.**

FP	Day 2 RFP	Day 2 GFP	Day 6 RFP	Day 6 GFP
mCherry_01	8880	8806	1652	12043
mCherry_02	2584	4277	3848	8852
mScarlet_01	5834	4034	10057	5079
mScarlet_02	2013	1119	10081	7621
FR_01	4703	3096	7353	4493
FR_02	1903	1504	9154	6491
FR-MQV_01	4302	2785	14252	7576
FR-MQV_02	2047	1932	10475	4380

## 5c. Chromophore maturation kinetics

Details of maturation kinetics are provided in the Methods and Materials section of Chapter 3. Briefly, FPs were expressed in *E. coli*. After induction of protein expression, cultures were treated with chloramphenicol to halt new protein production and both the fluorescence and optical density were measured over time. An increase in fluorescence (after normalization to optical density) indicates an increase in chromophore formation. FR and FR-MQV show similar maturation kinetics. The mScarlet and mScarlet-I values are comparable to those reported in the literature. Hence, the mutations in FR-MQV do not appear to perturb maturation and folding of the FP at 37°C (temperatures used for mammalian cell growth/imaging experiments).

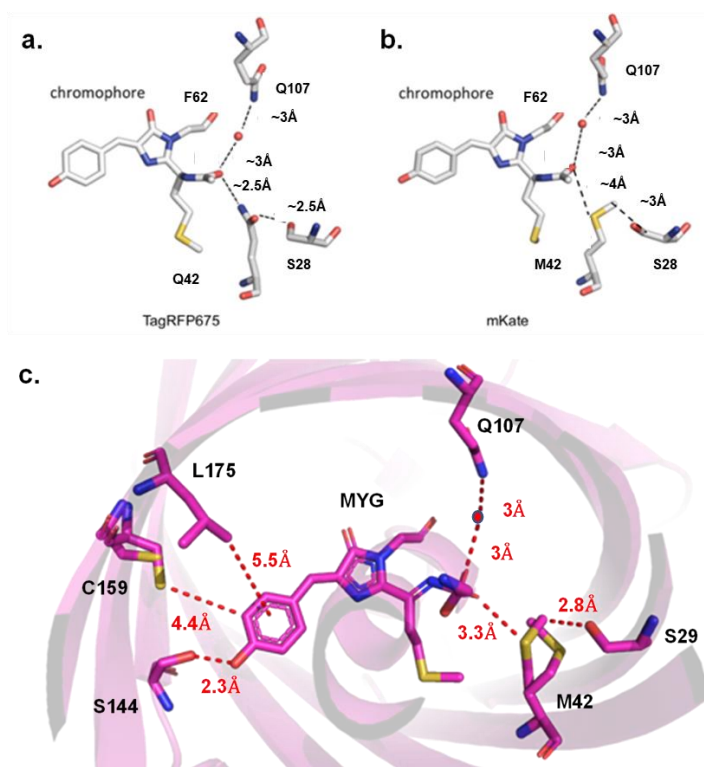


**Figure A2.12. Chromophore maturation kinetics.** (a.) The optical density (OD) of bacterial cultures over time after addition of chloramphenicol at  $t = 0$ . The ODs remained fairly constant, suggesting the action of chloramphenicol stalling the growth of bacterial cells in the log phase. (b.) The measured fluorescence normalized and scaled with respect to the observed ODs at each time point.

**Table A2.10. Measured versus the reported maturation times ( $t_{50}$ ) for the RFPs investigated in this study.**

Fluorescent Protein	$\sim t_{50}$ 37°C	Reported $t_{50}$ (min)
mScarlet-I	45 min	36, 25 [Refs 8,13]
mScarlet	165 min	174, 132 [Refs 8, 13]
FR	195 min	130 [Ref 9]
FR-MQV	195 min	-

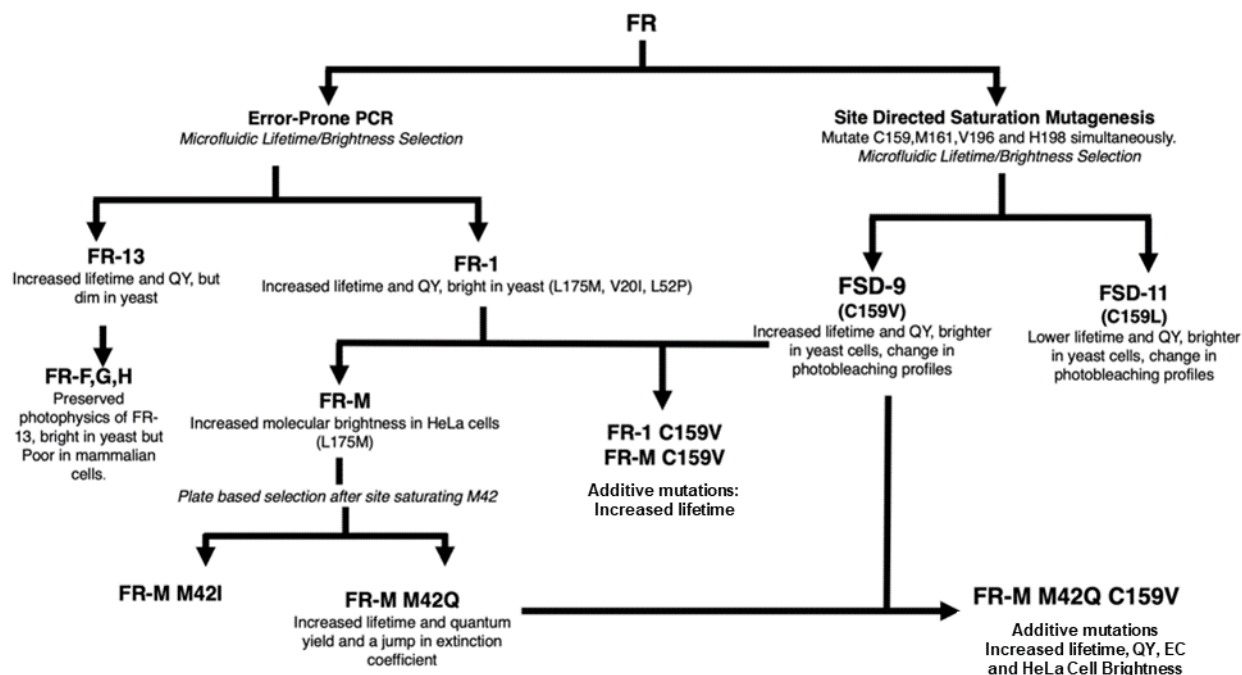
## Section 6. Additional structural information



**Figure A2.13. Structural representation of the FR crystal structure.** The hydrogen bonding network in (a.) TagRFP-675 and (b.) mKate. Images modified from Konold *et al.* [14] Distances were calculated using the crystal structures of TagRFP-675 (PDB ID: 4KGE) and mKate (PDB ID:3BXA). Konold and co-workers describe an extensive network involving the Q106, S28 and M/Q41 residues with a crystallized water molecule at the acylimine end of these RFPs (c.) The

crystal structure of FR (PDB ID: 6U1A) also reveals a similar arrangement. Relevant positions and the distances from the chromophore in the FR structure are shown.

## Section 7. FR Evolution Table



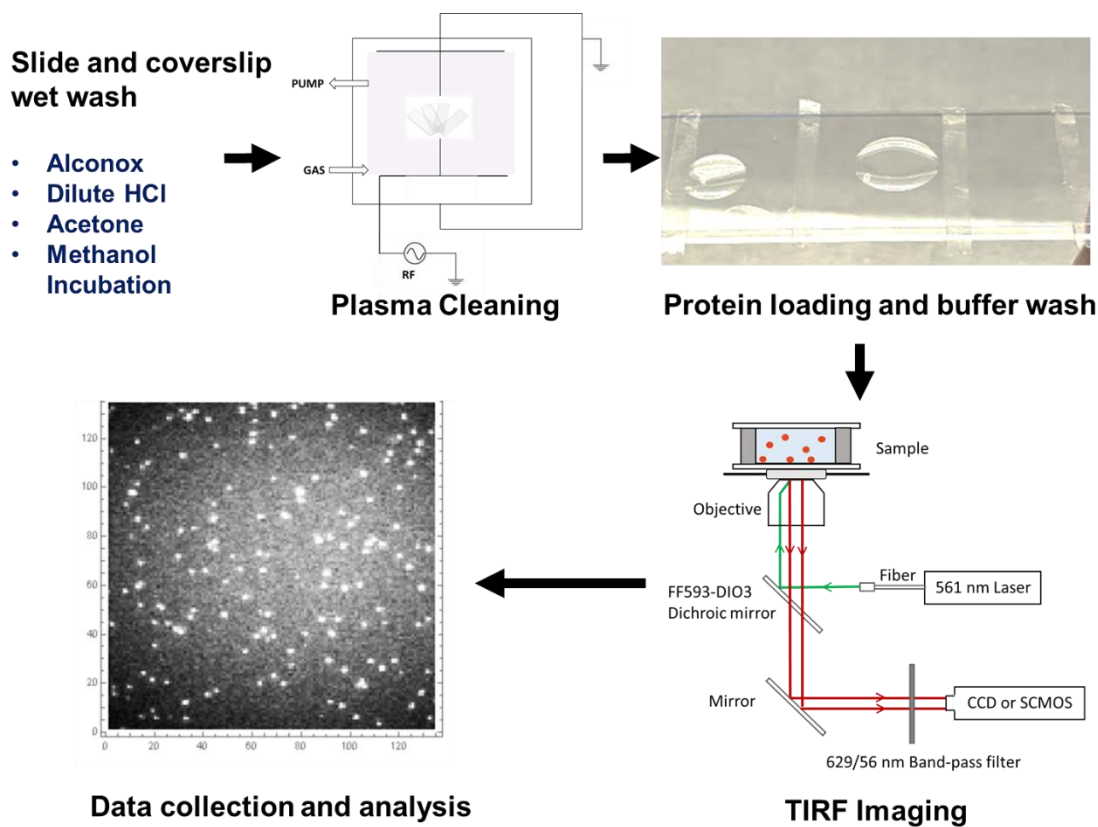
**Figure A2.14. Evolution table.** The evolution of the FR family of proteins. The pathways indicate the various engineering strategies that led to the development of FR-MQV. All amino acid positions have been numbered with respect to the parent FR numbering as per Table A2.1.



## Appendix 3

This is an addendum to Chapter 4 and parts of this appendix have been adapted from the article “Characterizing Dark State Kinetics and Single Molecule Fluorescence of FusionRed and FusionRed-MQ at Low Irradiances.” Mukherjee, S.; Thomas, C.; Wilson, R.; Simmerman, E.; Hung, S. T.; Jimenez, ChemRxiv 2022. <https://doi.org/10.26434/chemrxiv-2022-2dr03>

### Section 1. Experimental workflow



**Figure A3.1.** Schematic of the experimental workflow from sample preparation to data analysis.

## Section 2. Excitation rate calculations

Our TIRF measurements are for proteins (~1–2 nm) bound to the surface of the coverslip, therefore we estimate the irradiance (I) of the evanescent wave at the interface surface using Fresnel's equations, given below. [15]

$$I_e = I \frac{4(\cos\theta)^2(2(\sin\theta)^2 - n^2)}{n^4(\cos\theta)^2 + (\sin\theta)^2 - n^2} \quad (1)$$

$$k_{ex} = \frac{2.303 * I_e * \lambda * \epsilon_\lambda}{\ln(2) * N_A * h * c} \quad (2)$$

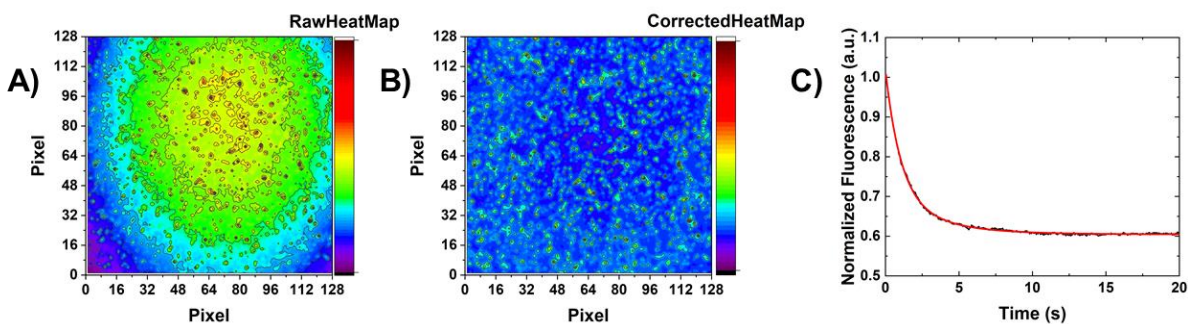
**Table A3.1. Power measurements**

Power (mW)	I (W/cm <sup>2</sup> )	I <sub>e</sub> (W/cm <sup>2</sup> )	Excitation Rate <sub>FR</sub> (Hz)	Excitation Rate <sub>FR-MQ</sub> (Hz)
0.28	0.49	1.23	940	1622
0.42	0.74	1.85	1420	2450
0.49	0.63	1.57	1206	2080
0.56	0.99	2.48	1902	3282
1.24	1.57	3.93	3014	5200
2.48	3.15	7.87	6035	10412
4.96	6.29	15.73	12062	20811

FR = Fusion Red, FR-MQ = FusionRed-MQ

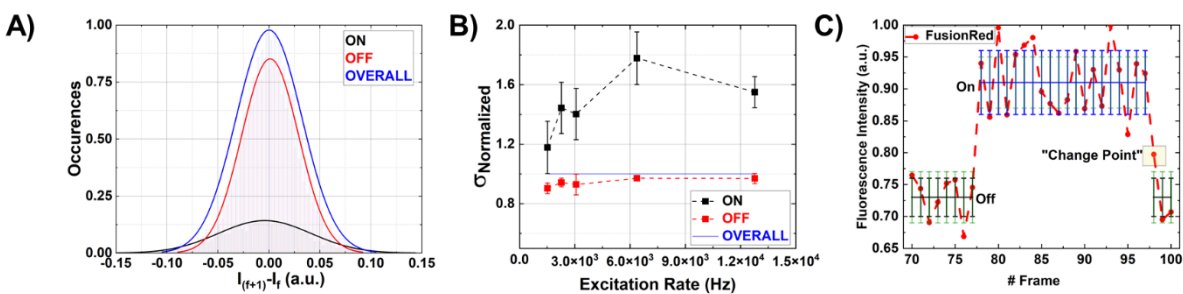
## Section 3. Single molecule data analysis addendum

### 3a. Spot identification script corrections



**Figure A3.2. Spatial corrections.** (a.) Gaussian profile, (b.) after removal of the Gaussian profile, (c.) after removal of the fast component on a blank sample.

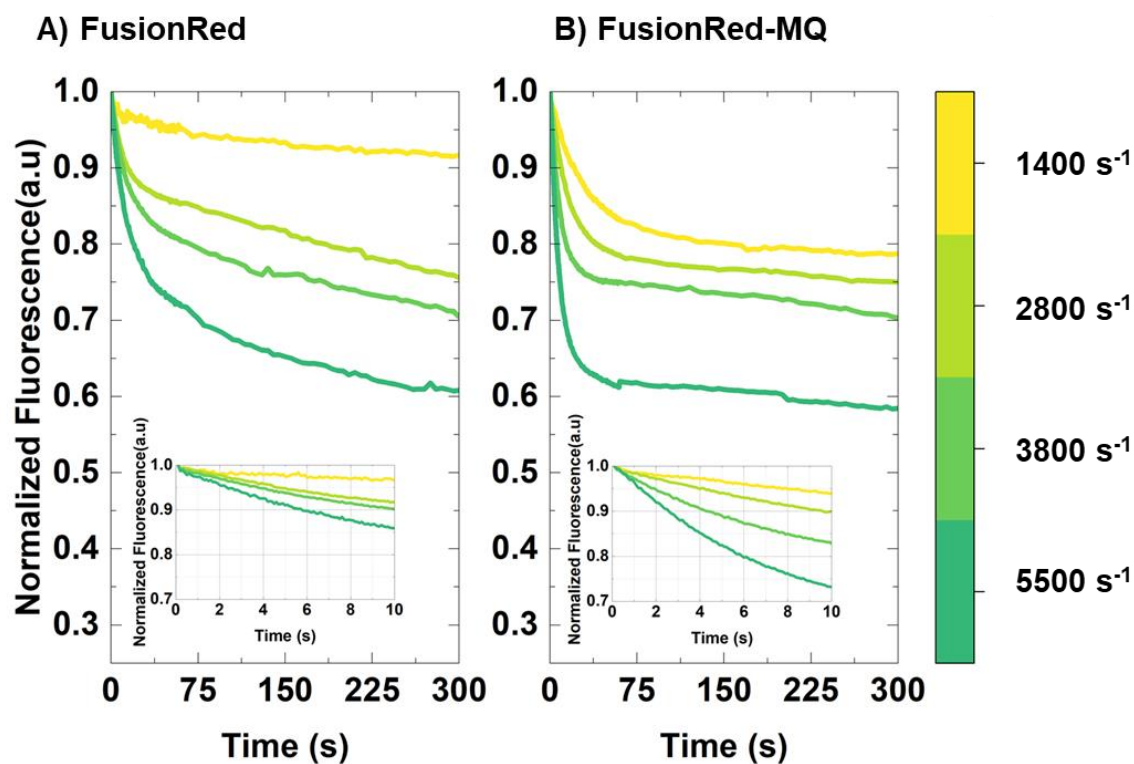
### 3b. Detection of state changes



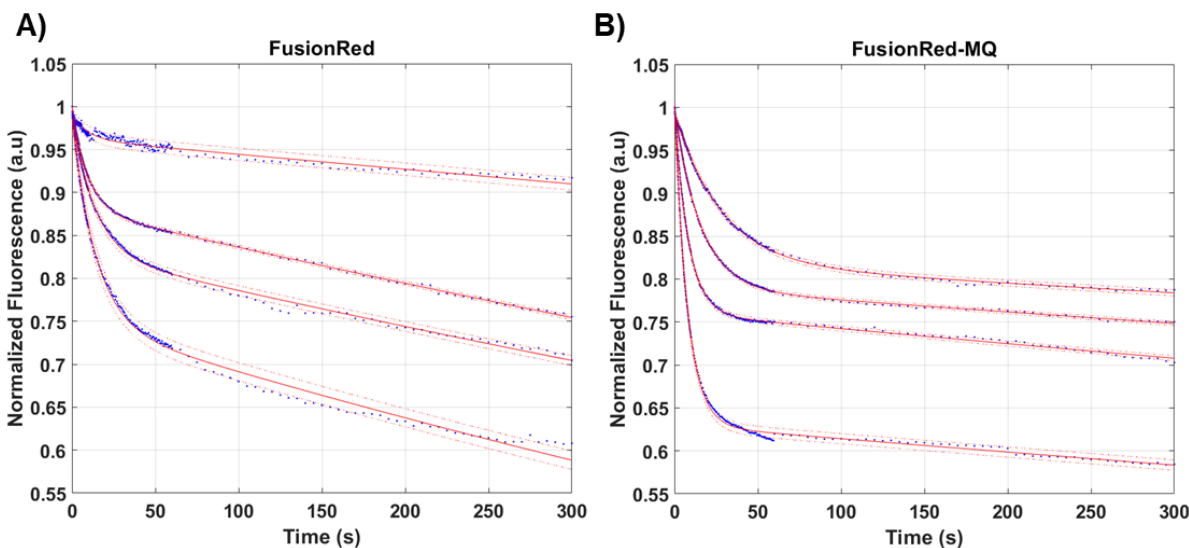
**Figure A3.3. Intensity variation.** (a.) On/ Off and Overall fluctuation histograms, (b.) Levels of fluctuations from the on and off state (read and shot noise), (c.) Change point definition.

## Section 4. Ensemble photobleaching

Figure A3.4 shows the ensemble bleaching of these two proteins in bacteria.



**Figure A3.4. Ensemble photobleaching kinetics.** Photobleaching traces of *E. coli* expressing (a.) FusionRed and (b.) FusionRed-MQ, across varying irradiance ranges.



**Figure A3.5. Bi-exponential nature of fluorescence decay from bacteria expressing FusionRed and FusionRed-MQ.**

**Table A3.2. Fitting details for the biexponential decay traces of FusionRed and FusionRed-MQ**

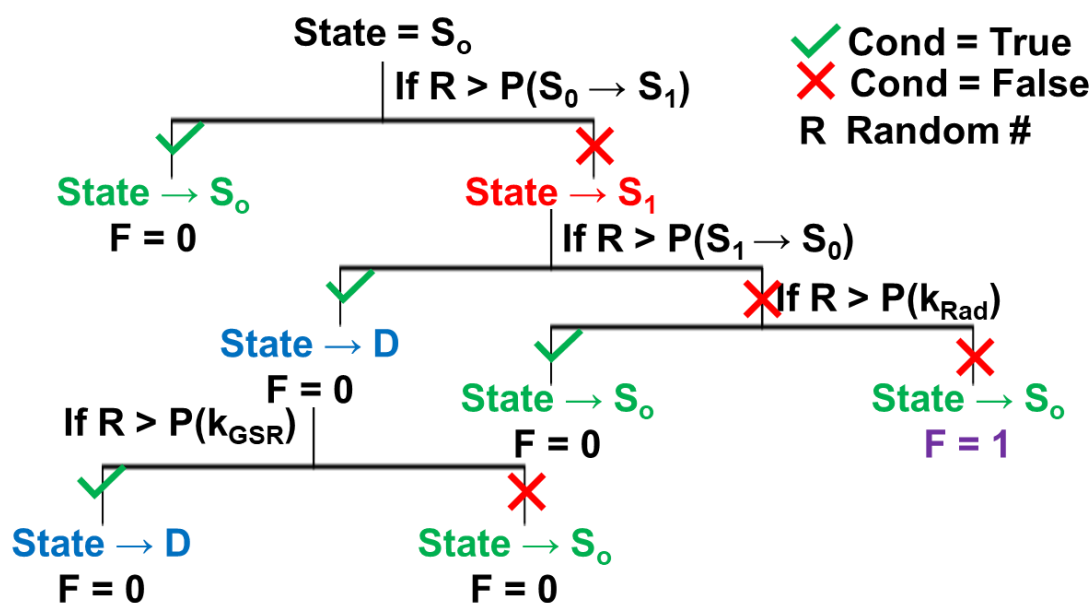
$k_{\text{Ex}}$ ( $\text{s}^{-1}$ )	FusionRed				FusionRed-MQ			
	$\tau_1$ (s)	$A_1$ (%)	$\tau_2$ ( $\times 10^3$ s)	$A_2$ (%)	$\tau_1$ (s)	$A_1$ (%)	$\tau_2$ ( $\times 10^3$ s)	$A_2$ (%)
1400	$14.1 \pm 2.4$	2	$5.4 \pm 0.3$	97	$27.9 \pm 0.5$	16	$7.1 \pm 0.3$	84
2800	$12.6 \pm 2.2$	12	$2.4 \pm 0.6$	88	$15.6 \pm 0.1$	21	$5.6 \pm 0.1$	79
3800	$9.4 \pm 1.8$	16	$1.8 \pm 0.1$	84	$8.2 \pm 0.1$	24	$4.1 \pm 0.1$	76
5500	$9.9 \pm 2.6$	24	$1.6 \pm 0.6$	76	$7.7 \pm 0.1$	34	$3.4 \pm 0.1$	66

In both cases, the larger time constant  $\tau_2$  (most likely for permanent) is  $\sim 500$ -fold larger than the shorter  $\tau_1$  (likely reversible component) of photobleaching. Also, with increasing irradiance we observe accelerated photobleaching for both systems.

## Section 5. Simulation results

### 5a. Three state single molecule blinking simulator algorithm

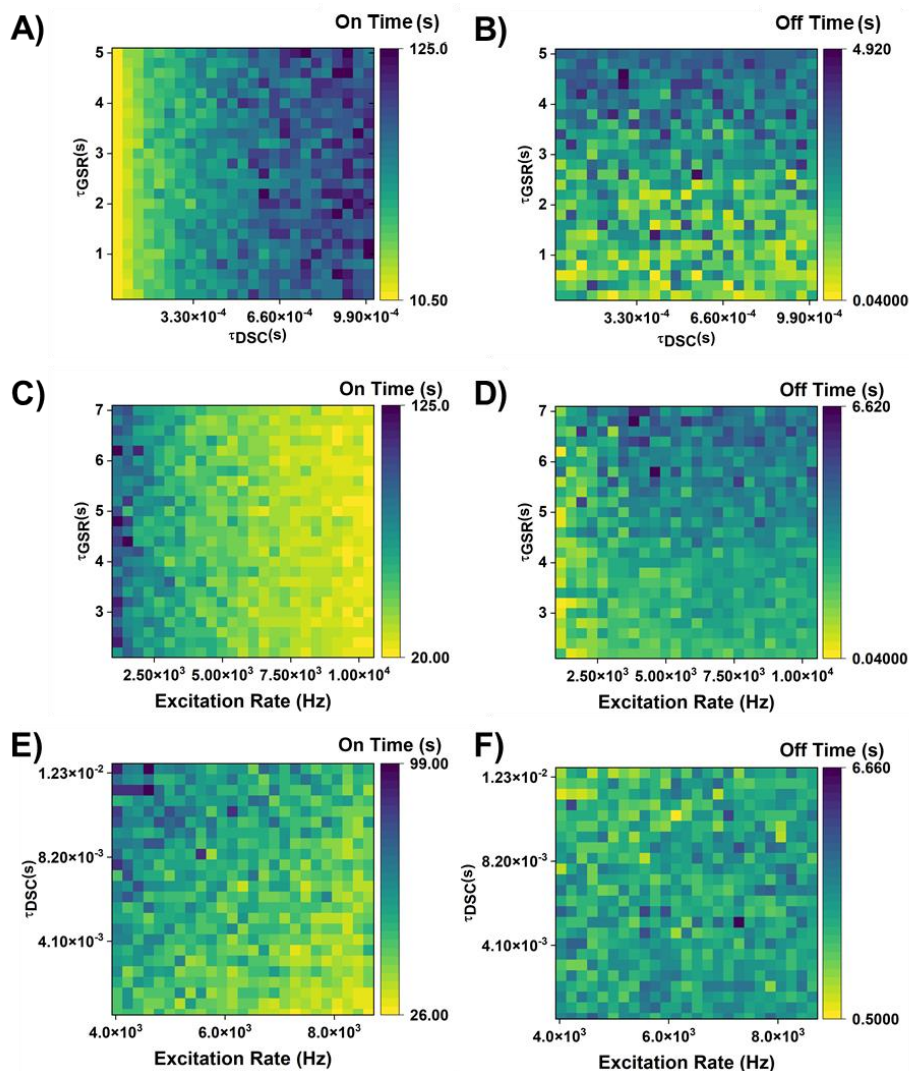
The schematic below indicates the behavior of a single FP molecule at each time step (defined by the user). The code uses random sampling using Monte-Carlo methods to decide the residence of the molecule in each state ( $S_1$ ,  $S_0$  or  $D$ ) based on the probabilities of making a transition from one state to the other. The probabilities of these transitions are in-turn dependent on the user defined rate constants for  $k_{Ex}$ ,  $k_{Em}$ ,  $k_{IC}$ ,  $k_{DSC}$  and  $k_{GSR}$ .



**Figure A3.6. Schematic representing the algorithm used to simulate single molecule behavior.** The schematic for a three ( $S_0$  = Ground electronic,  $S_1$  = First excited and  $D$  = Dark) state model represented above. The value of  $F=1$  indicates a fluorescence photon has been generated in the process.

## 5b. On/Off time dependence on DSC and GSR times and excitation rates

We observe a shorter on-time for both shorter dark-state conversion times and higher excitation rates (Figure A3.7, Panels A, C and E). We also observed that off-times depend only on the ground-state recovery times (Figure A3.7.2, Panels B, D and F).



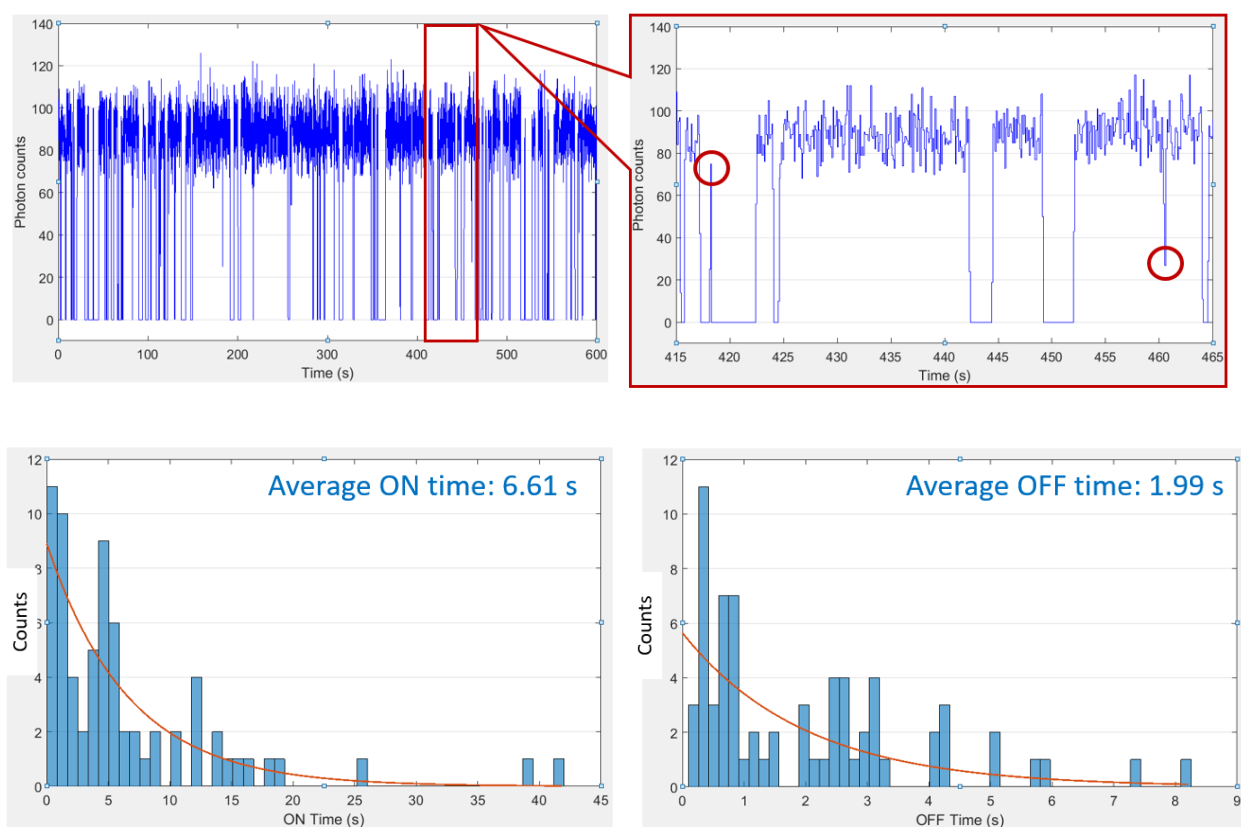
**Figure A3.7. Simulation heatmaps for on and off time with varying rates constants.**

Variations of on and off times with respect to  $\tau_{\text{DSC}} (=1/k_{\text{DSC}})$  and  $\tau_{\text{GSR}} (=1/k_{\text{GSR}})$  expected from Monte-Carlo simulations for FusionRed. **(a & b.)** Variation of on and off times with DSC and GSR times at a fixed excitation rate (1923 Hz). **(c & d.)** Variation of on and off times with excitation rate and GSR times at a fixed DSC time (0.5 ms). **(e & f.)** Variation of on and off times with excitation rate and DSC times at fixed GSR time (3 s).

### 5c. Effects of changing time steps on this algorithm

This model considers the time step to be longer than the excited state lifetime and the dark state conversion time but slower than the ground state recovery time. This is done to mimic millisecond acquisition times for commercial EMCCD and CMOS cameras employed for single molecule imaging. We ignore photobleaching in these simulations. We used  $\tau_{\text{DSC}} \cong 40 \mu\text{s}$ ,  $\tau_{\text{GSR}} \cong 2 \text{ s}$ ,  $\Phi \cong 0.24$ ,  $\tau_{\text{FL}} \cong 1.78 \text{ ns}$  and  $k_{\text{ex}} \cong 3500 \text{ Hz}$  for these simulations.

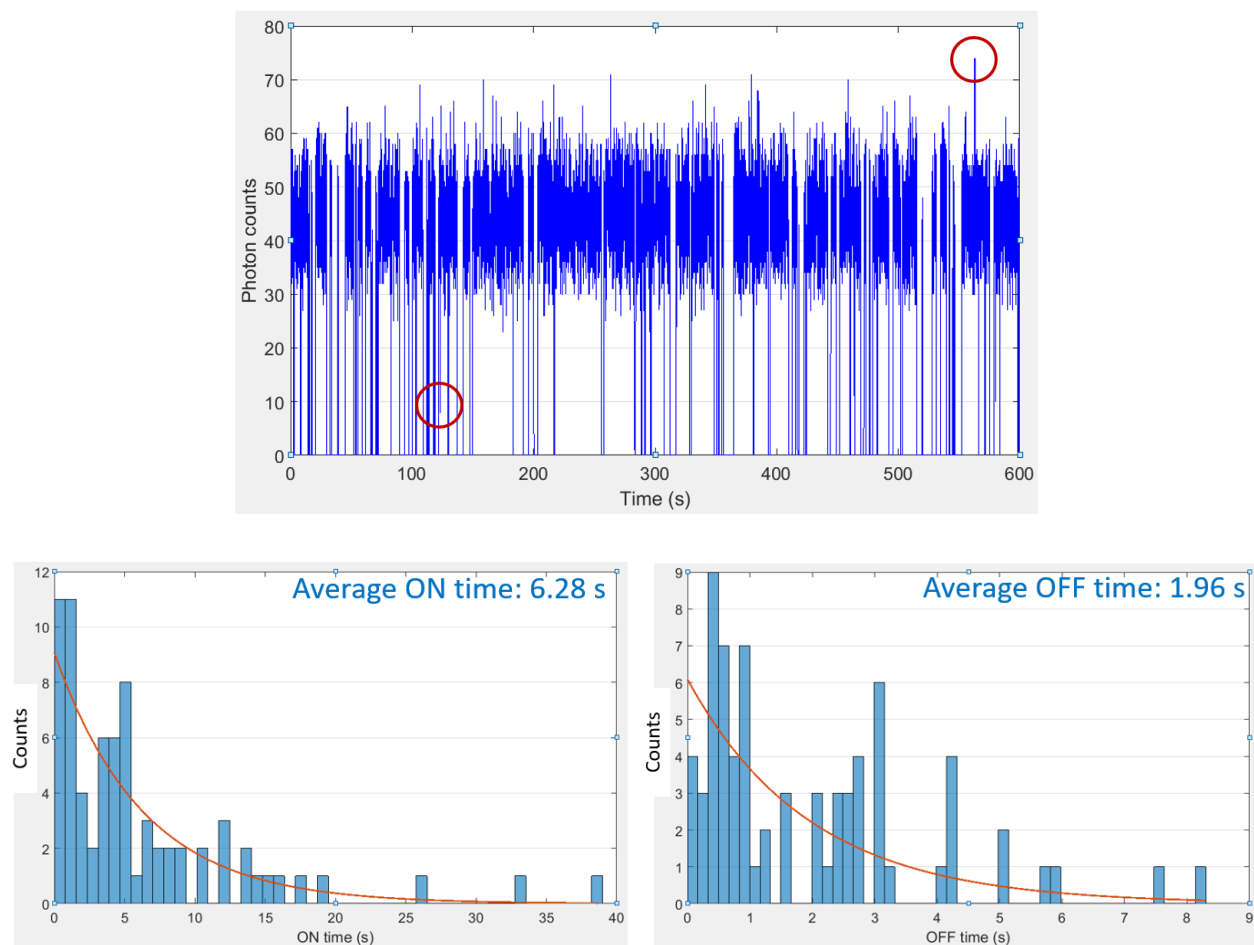
*Case 1:* One FusionRed protein, image acquisition time 100 ms, 6000 frames.



**Figure A3.8. FusionRed blinking simulation (100 ms, 6000 frames)** (Above) Single molecule trajectories. The red circles indicate frames where photon counts are lower than most on events, indicating the switching off of the molecule mid-frame of the acquisition step. Our single molecule binarization change-point algorithm can recognize these real fluctuations and jumps and distinguish these from experimental background noise. (Below) The histograms of binarized on and off traces.



Case 2: One FusionRed protein, image acquisition time 50 ms, 12000 frames.

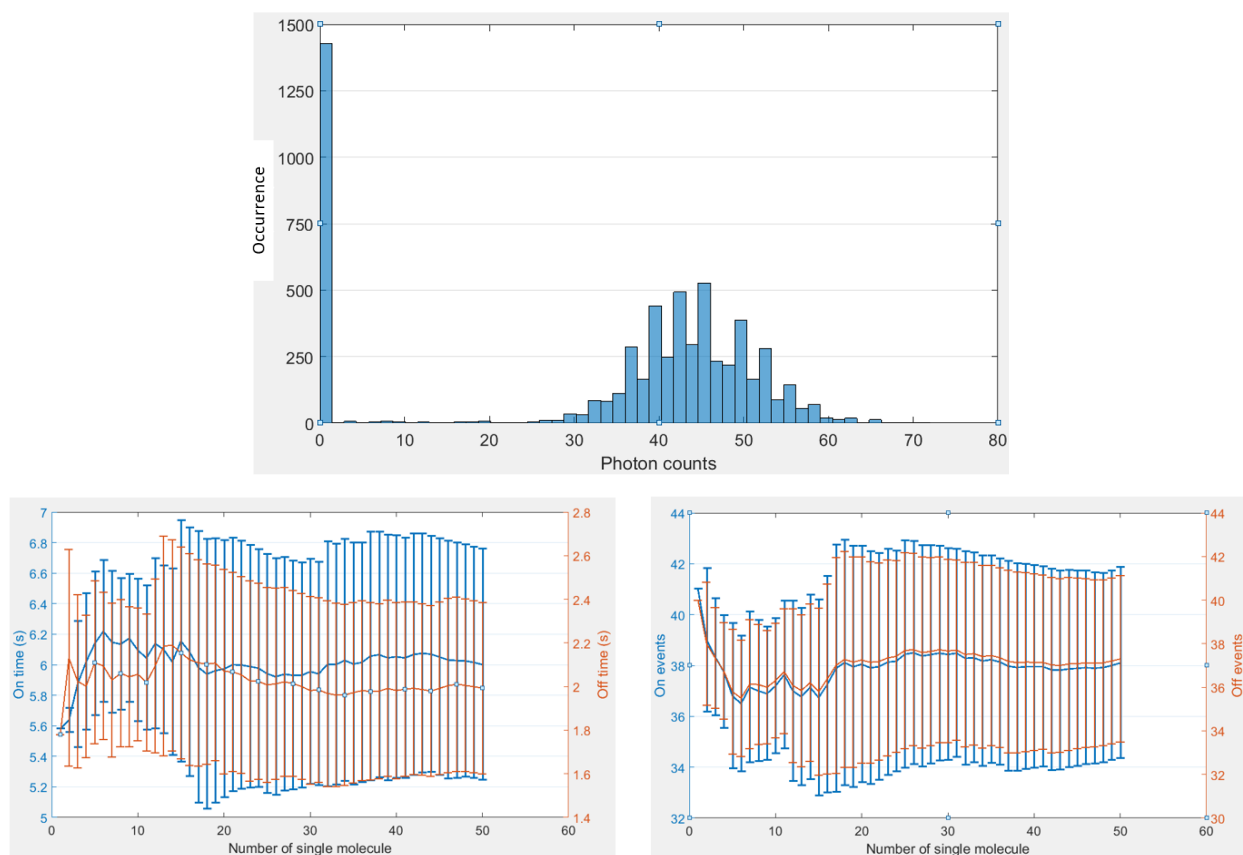


**Figure A3.9. FusionRed blinking simulation (50 ms, 12000 frames)** (Above) The photon count on the camera reduces by ~50% when the acquisition time is halved. (Below) The statistics of on and off times do not change on changing the frame rate.

Similarly, when the image acquisition time is reduced to 20 ms, the photon counts also reduce accordingly, but the on/off times are similar to the above two cases.

## 5d. Statistics of on/off time for a single FP blinking versus multiple FPs under the same time step

*Case 3:* Repeat 50 simulations of single FusionRed protein, image acquisition time 50 ms and taking 6000 frames for each FusionRed protein simulation. In comparison to case 2 presented in the previous simulation.



**Figure A3.10. FusionRed blinking simulation (50 ms, 6000 frames)** (Above) The photon count histograms are centered at ~45 photons for either case. (Below) The fluctuations for on and off times become relatively consistent after ~5 FPs.

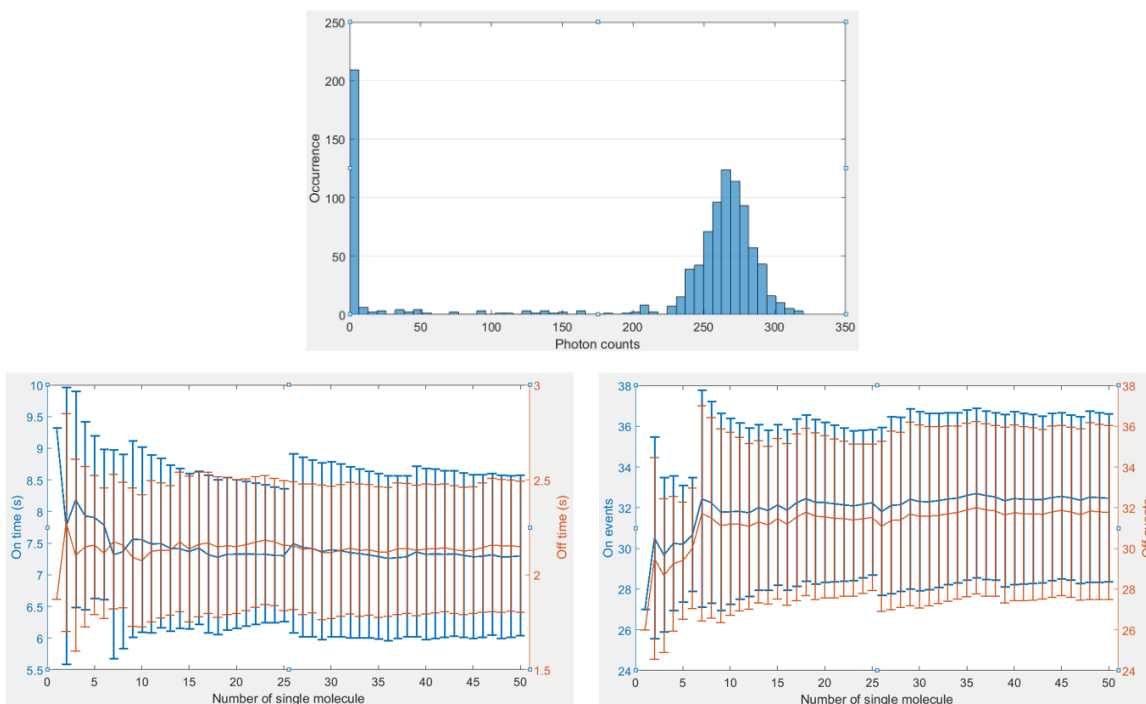
The on/off statistics do not change as the photon counts drop. We still remain below the optical saturation limit and observe near linear photon counts.

## 5e. Statistics of on/off time relevant to an actual single molecule experiment

Our experiments were carried out using camera acquisition times of ~100–300 ms for 50–500 FPs. This case describes the simulated data for these settings.

*Case 4:* In total, 50 simulations of single FR proteins, image acquisition time 300 ms and taking 1000 frames for each FR protein simulation (~5-minute trajectories).

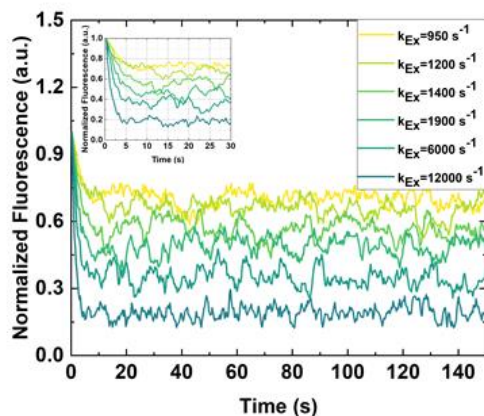
Photon count distribution is shifted ~6 times higher than Case 3, since the acquisition time is 6 times longer. The average ON/OFF time and number of ON/OFF events are also shown below. The average ON/OFF time is slightly longer than Case 3 because the ON/OFF time is added in the unit of the image acquisition time. The average number of ON/OFF events observed in this case are slightly less than Case 3, also because of the longer acquisition time, i.e., the shorter the acquisition time, the more the ON/OFF switching events can be observed. However, the number of events does not change significantly since the average ON/OFF times are much longer than the acquisition time.



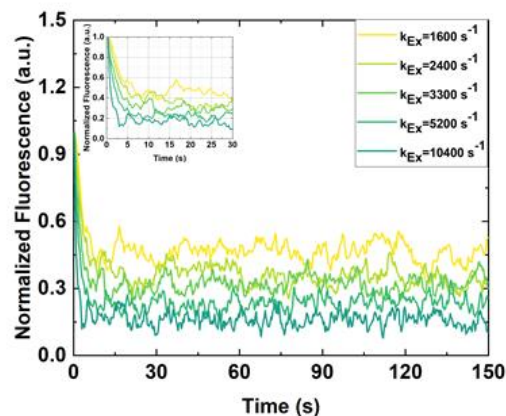
**Figure A3.11. FusionRed blinking simulation (300 ms, 1000 frames)** (Above) At 300 ms, our photon counts are significantly higher than at shorter acquisition time. (Below) The fluctuations for on and off times are almost consistent after ~5 FPs.

5f. Ensemble behavior averaged from simulated single molecule blinking with respect experimental rate constants obtained from experiments.

### A) FusionRed

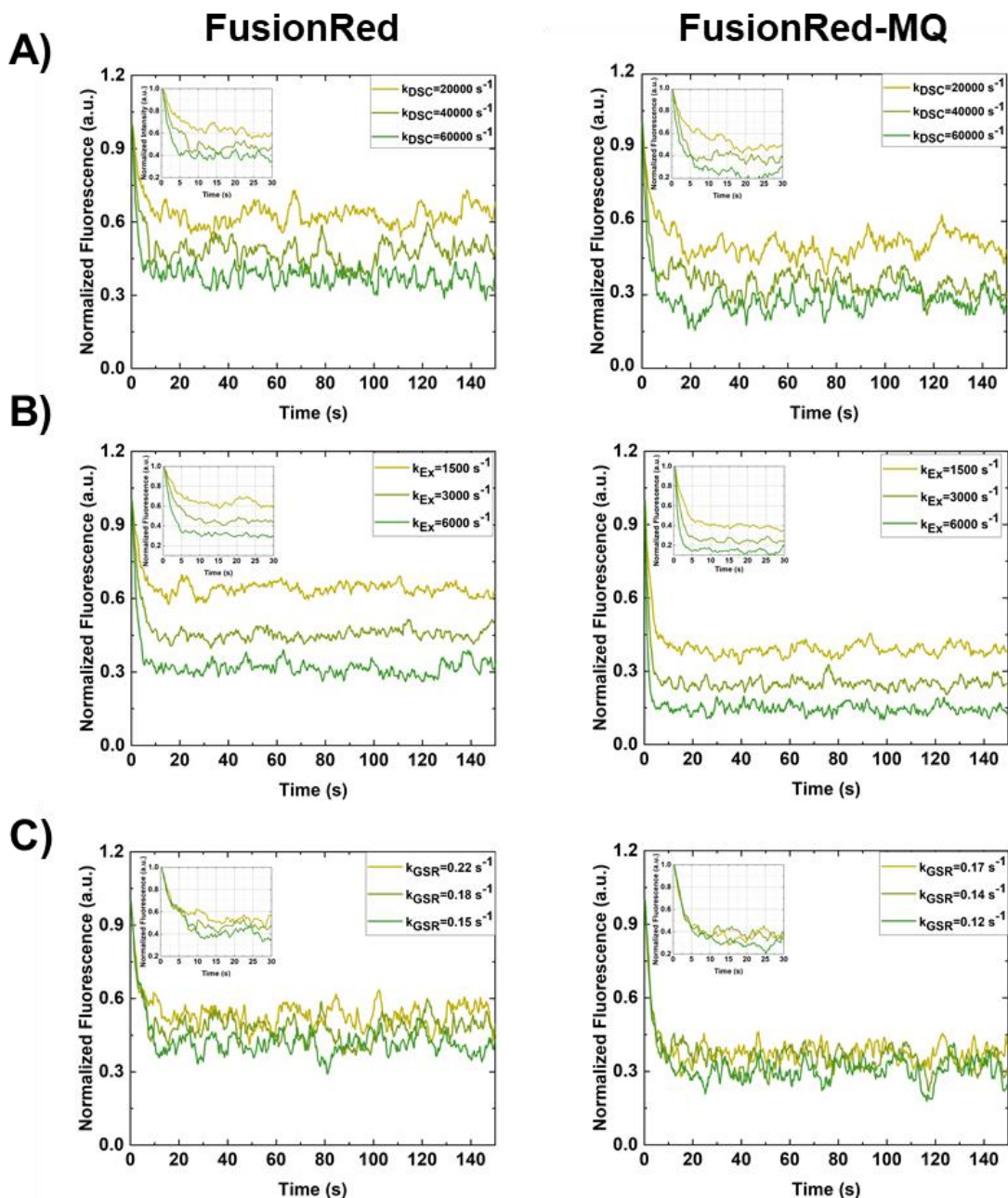


### B) FusionRed-MQ



**Figure A3.12. Ensemble trends from Monte-Carlo simulations.** Simulation results are obtained from rate constants reported in Chapter 3; Table 3.2, without permanent photobleaching. FusionRed-MQ shows a larger dark fraction and faster dark state conversion in comparison to FusionRed. Each trace represents the normalized sum fluorescence from three simulated single molecule blinking video with ~150 emitters.

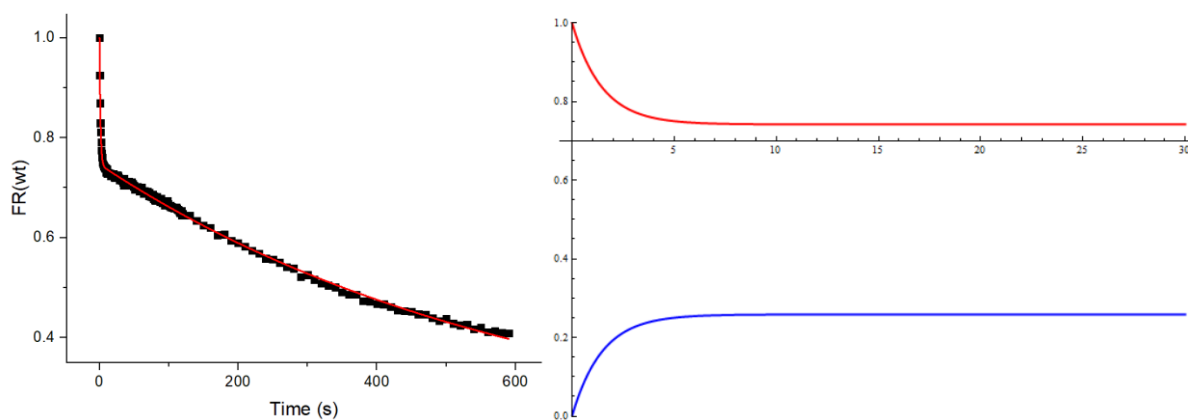
**5g. Simulation of ensemble behavior averaged from simulated single molecule blinking with respect to varying  $k_{\text{EX}}$ ,  $k_{\text{DSC}}$  and  $k_{\text{GSR}}$  while holding other rate constants fixed.**



**Figure A3.13. Variation in ensemble fluorescence decay traces with respect to  $k_{\text{DSC}}$ ,  $k_{\text{EX}}$  and  $k_{\text{GSR}}$ .** Each trace represents the normalized sum fluorescence from three simulated single molecule blinking videos, each with  $\sim 150$  emitters. The panel on the left represents the

simulated traces with respect to photophysical properties of FusionRed and on the right, photophysical properties of FusionRed-MQ. For each panel, the one rate constant was varied at a time while keeping the other two rate constants fixed, for example – for FusionRed panel A,  $k_{\text{DSC}}$  was varied keeping  $k_{\text{ex}}=3000 \text{ s}^{-1}$  and  $k_{\text{GSR}}=0.18 \text{ s}^{-1}$  fixed. These results provide insight into how each rate constant involved in dark state population and depopulation manifest at an ensemble level.

### 5h. Ensemble behavior of the system based on numerical simulations for a three-state system



**Figure A3.14. Population based numerical simulation of a 3-state model.** A population based numerical simulation (non-Monte Carlo methods) indicates that the  $k_{\text{DSC}}$  controls the dark fraction at the ensemble level. (Left) Photobleaching curve of FusionRed fit with two exponential functions. (Right) Normalized populations in  $S_0$  (red) and D (blue) over time (x axis, in seconds) without considering permanent photobleaching obtained from by utilizing rate constants from these fit on the left and numerical simulations for a three-state model.

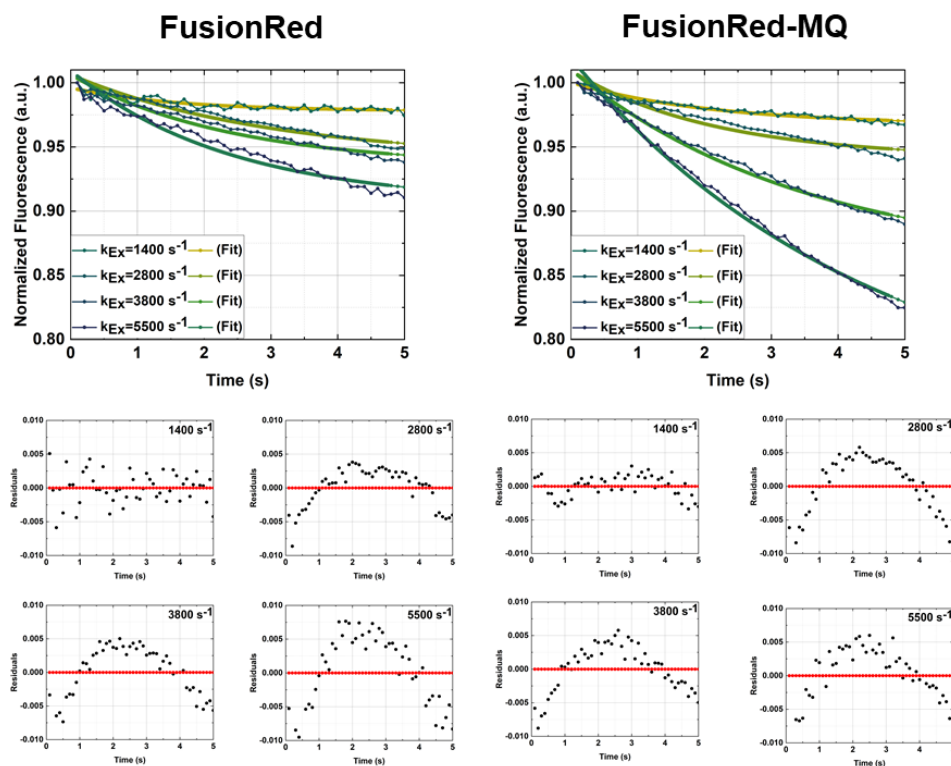
## Section 6. Additional fitting results

### 6a. Analytical expression for fitting the ensemble fluorescent population:

Three-state model equation

$$\begin{aligned}
 S_0(t) = & A * k_{DSC} * ( 4 * k_{GSR} * ( k_{DSC} + k_{Em} + k_{IC} ) * \sqrt{ ( ( k_{DSC} + k_{Em} + k_{Ex} + k_{GSR} + k_{IC} )^2 - \\
 & 4 * ( k_{DSC} * ( k_{Ex} + k_{GSR} ) + k_{GSR} * ( k_{Em} + k_{Ex} + k_{IC} ) ) ) + \exp ( - 0.5 * t * ( k_{DSC} + k_{Em} + k_{Ex} + \\
 & k_{GSR} + k_{IC} + \sqrt{ ( ( k_{DSC} + k_{Em} + k_{Ex} + k_{GSR} + k_{IC} )^2 - 4 * ( k_{DSC} * ( k_{Ex} + k_{GSR} ) + k_{GSR} * ( k_{Em} + \\
 & k_{Ex} + k_{IC} ) ) ) ) * k_{Ex} * ( k_{DSC} + k_{Em} + k_{Ex} + k_{GSR} + k_{IC} - \sqrt{ ( ( k_{DSC} + k_{Em} + k_{Ex} + k_{GSR} + k_{IC} )^2 \\
 & - 4 * ( k_{DSC} * ( k_{Ex} + k_{GSR} ) + k_{GSR} * ( k_{Em} + k_{Ex} + k_{IC} ) ) ) * ( - k_{DSC} + k_{Em} + k_{Ex} - k_{GSR} + k_{IC} + \\
 & \sqrt{ ( ( k_{DSC} + k_{Em} + k_{Ex} + k_{GSR} + k_{IC} )^2 - 4 * ( k_{DSC} * ( k_{Ex} + k_{GSR} ) + k_{GSR} * ( k_{Em} + k_{Ex} + k_{IC} ) ) } \\
 & ) - \exp ( 0.5 * t * ( - k_{DSC} - k_{Em} - k_{Ex} - k_{GSR} - k_{IC} + \sqrt{ ( ( k_{DSC} + k_{Em} + k_{Ex} + k_{GSR} + k_{IC} )^2 - 4 * \\
 & ( k_{DSC} * ( k_{Ex} + k_{GSR} ) + k_{GSR} * ( k_{Em} + k_{Ex} + k_{IC} ) ) ) ) * k_{Ex} * ( - k_{DSC} + k_{Em} + k_{Ex} - k_{GSR} + k_{IC} \\
 & - \sqrt{ ( ( k_{DSC} + k_{Em} + k_{Ex} + k_{GSR} + k_{IC} )^2 - 4 * ( k_{DSC} * ( k_{Ex} + k_{GSR} ) + k_{GSR} * ( k_{Em} + k_{Ex} + k_{IC} ) } \\
 & ) ) ) * ( k_{DSC} + k_{Em} + k_{Ex} + k_{GSR} + k_{IC} + \sqrt{ ( ( k_{DSC} + k_{Em} + k_{Ex} + k_{GSR} + k_{IC} )^2 - 4 * ( k_{DSC} * ( \\
 & k_{Ex} + k_{GSR} ) + k_{GSR} * ( k_{Em} + k_{Ex} + k_{IC} ) ) ) ) ) / ( 4 * ( k_{DSC} * ( k_{Ex} + k_{GSR} ) + k_{GSR} * ( k_{Em} + k_{Ex} \\
 & + k_{IC} ) ) * \sqrt{ ( ( k_{DSC} + k_{Em} + k_{Ex} + k_{GSR} + k_{IC} )^2 - 4 * ( k_{DSC} * ( k_{Ex} + k_{GSR} ) + k_{GSR} * ( k_{Em} + \\
 & k_{Ex} + k_{IC} ) ) ) ) + A_{PB} * k_{PB} + C_{Offset}
 \end{aligned}$$

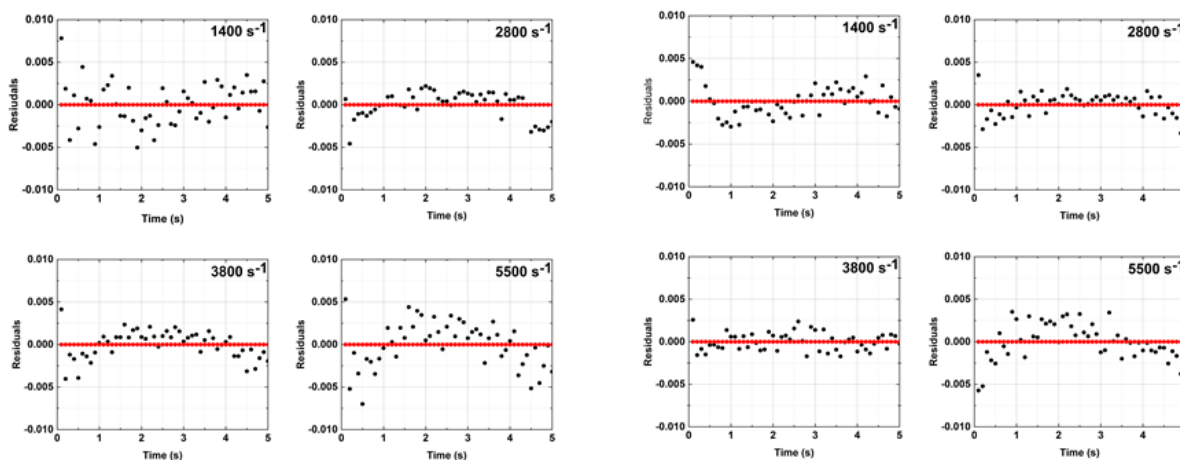
## 6b. Inaccuracies in fitting the ensemble bleaching with a fixed dark-state conversion rate constant



**Figure A3.15. Quality of fit for varying  $k_{DSC}$  and  $k_{GSR}$  unbound.** Residuals indicate a poor quality of fitting. Since single-molecule experiments provide precise measurements for the GSR in comparison to DSC, fixing the  $k_{GSR}$  within the experimental bounds of the single-molecule measurement and then fitting for  $k_{DSC}$  provided better insight and accurate estimation of the kinetics for the dark-state conversion process.



### 6c. Residuals for fitting $k_{DSC}$ with respect to a bound value of $k_{GSR}$



**Figure A3.16. Keeping  $k_{GSR}$  fixed while allowing the fitting algorithm to fit the  $k_{DSC}$ .** This provides a better quality of fit. Residuals indicate better fitting, with  $\text{adj-}R^2 > 0.95$ .

## Section 7. Theoretical estimation of the lowest and highest number of photons/frames

**Lower bound:** Lowest excitation rate for FusionRed

$$k_{ex} = 940 \text{ photons/s}$$

Fluorescence QY of FusionRed = 0.24

Total fluorescence photons = 226 photons/s

Acquisition time = 300 ms, Quantum Efficiency [16] of Andor iXon at 561 nm ~0.8

Maximum number of fluorescence photons per frame ~60 photons/ frame

Numerical aperture of objective ~ 1.42

% light collected by the objective ~ 0.4

Realistic lower bound ~ **25 photons/ frame**

**Upper bound:** Highest excitation rate for FusionRed-MQ

$k_{\text{ex}} = 20800 \text{ photons/s}$

Fluorescence QY of FusionRed = 0.43

Total fluorescence photons = 8950 photons/s

Acquisition time = 300 ms, Quantum Efficiency [16] of Andor iXon at 561 nm ~0.8

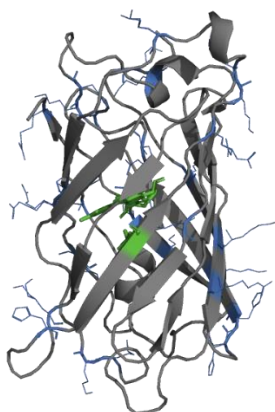
Maximum number of fluorescence photons per frame ~2300 photons/ frame

Numerical aperture of objective ~ 1.42

% light collected by the objective ~ 0.4

Realistic upper bound ~ **900 photons/ frame**

## Section 8. Surface charge on FusionRed



**Figure A3.17. Location of positively charged residues on the barrel of FusionRed.** The crystal structure of FusionRed (PDB ID: 6U1A) indicating the location of positively charged amino acid sidechains pointing out of the  $\beta$ -barrel (blue lines). These provide the opportunity to effectively bind the molecule with a negatively charged glass surface. The mutations acquired in FusionRed-MQ point into the barrel facing the chromophore (green). Therefore, we expect similar perturbations in FusionRed and FusionRed-MQ with respect to electrostatic interactions of the glass surface for both FPs.

## Appendix 4

This is an addendum to Chapter 5. Parts of this addendum will be adapted into the Supplementary Information section of the manuscript titled “*Directed evolution of a bright variant of mCherry: Suppression of non-radiative decay by fluorescence lifetime selections.*” by **Srijit Mukherjee**, Premashis Manna, Sheng-Ting Hung, Felix Vietmeyer, Pia Friis, Amy E. Palmer, and Ralph Jimenez

### Section 1. Mutagenesis Protocols

In order to develop RFP variants with higher brightness, we started with mCherry as a template. mCherry displays high photostability, excellent pH-resistance, fast maturation and a low cytotoxicity. However, mCherry is significantly dimmer than other recently developed RFPs, e.g. mRuby3, [17] mScarlet, [8] and FusionRed-MQV [18] in-part due to its lower quantum yield. To tackle this, we created several site-directed and random mutagenesis libraries and screened for higher excited state lifetime. Lifetime evolution was carried out using our microfluidic sorting platforms as previously reported. [19] We present a short discussion of the mutagenesis and library generation protocol below. Detailed protocols have been reported in our previous works. [2, 11, 18, 19]

#### 1a. Yeast and bacterial constructs for libraries and pure proteins

**Template construction:** The original DNA templates for mCherry and other FPs were amplified with gene-specific primers and cloned into pDonr221 using Gateway recombination system (Life Technologies) and sequenced. The forward primer for each FP included a recombination recognition sequence (attB1), a Shine-Dalgarno sequence for prokaryotic expression, a BamH1 restriction endonuclease site, a Kozak sequence for mammalian expression and ~ 30 nucleotide sequence complementary to the FP. The reverse primer for each FP included ~ 30 nucleotides complementary to the FP, a stop codon, an EcoR1 restriction endonuclease site and an attB2 recombination recognition sequence. Using the Gateway recombination system, the LR reaction

was performed to clone the FP into pYestDest52 vector (Life technologies). After sequences were confirmed the FP/pYestDest52 plasmids were used as template for library construction.

**Error-prone libraries:** GeneMorph II Random Mutagenesis kit (Agilent Cat No. 200550) was used to create the error-prone libraries. The kit protocol was followed with varying amounts of template DNA and cycles depending on the error rate. Like our work with FusionRed a typical error-rate is used that incorporates ~5 mutations (at the nucleotide level) per template. [2, 18] T7 and V5 universal primers (both located on pYesDest52 vector) were used for the amplification. After first round of PCR, the gel-extracted PCR product was used for a second round of PCR to create enough DNA for homologous recombination. After PCR purification the library DNA was isopropanol precipitated and eluted in a few  $\mu$ l of water.

**Site directed libraries:** QuikChange site-directed mutagenesis method was used to make point mutations or switch single amino acids using PfuTurboDNA polymerase and a Thermo cycler. PfuTurboDNA polymerase replicates both plasmid strands with high fidelity and without displacing the mutant oligonucleotide primers. The basic procedure utilizes a supercoiled double-stranded DNA (dsDNA) vector with the FP of interest and two synthetic oligonucleotide primers containing the desired mutation. The oligonucleotide primers, each complementary to opposite strands of the vector, are extended during temperature cycling by PfuTurboDNA polymerase. Incorporation of the oligonucleotide primers generates a mutated plasmid containing staggered nicks. Following temperature cycling, the product is treated with DpnI. The DpnI endonuclease digests the parental DNA template and makes it possible to select for mutation-containing synthesized DNA. The nicked vector DNA containing the desired mutations is then transformed into *E. coli* (Top10). Libraries with multiple site-directed targets were created using SOE reaction (Splicing overlap extension). Primers were designed to introduce the desired mutations. The initial PCRs generate overlapping gene segments that are used as template DNA for another PCR to create a full-length product. The internal primers generate overlapping, complementary 30 ends on the intermediate segments and introduce the desired nucleotide changes for site-directed mutagenesis. Overlapping strands of these intermediate products hybridize at these 30 regions in a subsequent PCR and are extended to generate the full-length product amplified by flanking primers. The full-length product is gel-extracted and isopropanol-precipitated and eluted in a few  $\mu$ l water.

**Electroporation:** Fresh Competent yeast cells (*Saccharomyces cerevisiae* BY4741) were prepared prior to electroporation. Cells, DNA and cut pYestDest52 vector were combined and left on ice for 5 min. Electroporation conditions (Bio-Rad Gene Pulser Xcell): C = 25  $\mu$ F, PC = 200-ohm, V = 1.5 kV (in 0.2 cm cuvettes). Cells were passed twice prior to expression. Mutants of were transferred to pBad-His vector for expression in *E. coli* and consequent Ni-NTA protein purification.

## **1b. Mammalian Expression Constructs for FACS**

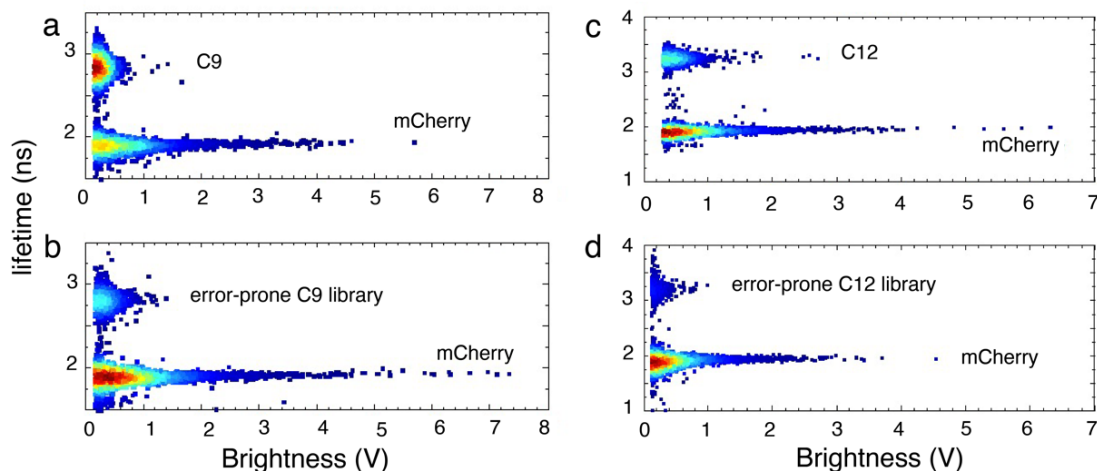
FPS were expressed as histone H2B fusion proteins in HeLa cells. Specifically, the mutants were PCR amplified from the pBad constructs with the upstream primer. Consequently, these were cut and ligated on piggyBac-H2B for expression in HeLa/U2OS cells. Cells were cultured in RPMI medium (Gibco Life Technologies) supplemented with penicillin/streptomycin (Gibco Life Technologies) and 10% heat-inactivated fetal bovine serum (Sigma-Aldrich) at 37°C with 5% CO<sub>2</sub> plus humidity. HeLa cells transiently transfected using the TransIT-LT1 reagent (Mirus, catalog #MIR2304) were prepared for FACS analysis after 48 hours.

## **Section 2. Mutagenesis and Directed Evolution of mCherry to mCherry-XL**

We summarize our evolution efforts below and in Figures A4.1, A4.2, and A4.3.

- Site-directed mutagenesis on 16 (E/D/A/K/N/T/V), 17 (H/K/R) , 70 (K/R) , 99 (All 20), 143 (All 20), 161 (L/M/V/I/F), 163 (All 20 AA), 173 (All 20), 175 (All 20 AA) and 197 (All 20 AA) were triply- FACS enriched and revealed variations at positions 16, 17, 70, 99, 175 and 197.
- The positions 16, 17, 70, 99, and 197 were thus targeted and the triply FACS enriched library was addressed as the C-X (Cherry-X) library. Among these variants C9 and C12 were found to have longer lifetimes, but lower brightness in yeast than that of mCherry. In

order to restore the brightness while maintaining their increased lifetimes, at first, we performed random mutagenesis. However, the random mutagenesis on C9 and C12 did not produce bright variants. (Figure A4.1)

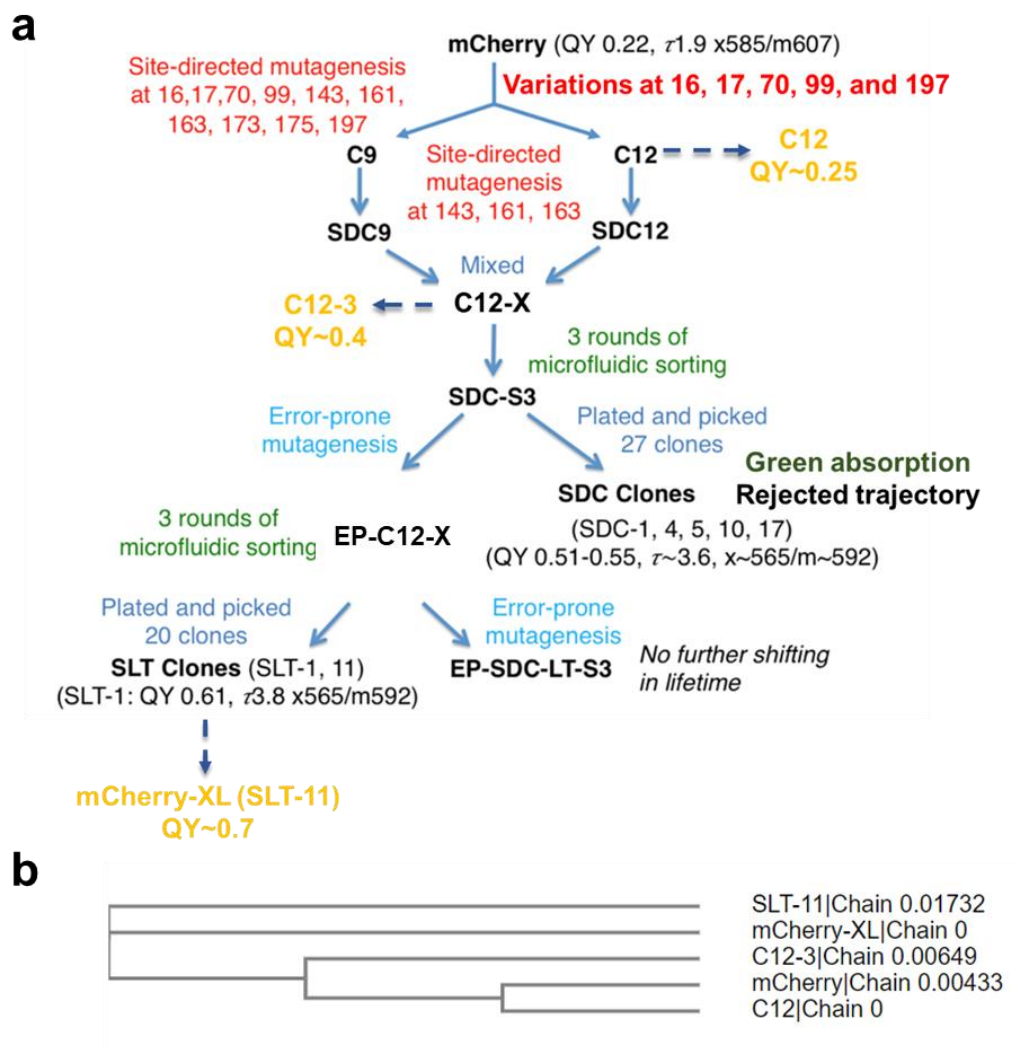


**Figure A4.1. EP-PCR library of C9 and C12 variants.** Microfluidic screening dot plots displaying lifetime and brightness (fluorescence signal in volts) of mCherry, C9, C12 and error-prone PCR libraries of C9 and C12. The pseudocolor indicates normalized cell counts with respect to the total number of screened/sorted cells at a certain value of brightness and lifetime on the plot – black indicating the highest and blue indicating the lowest. Each dot represents an individual yeast cell. It is evident from these plots that error-prone mutagenesis of C9 and C12 did not improve the lifetime and brightness of these variants. To collect these data, mCherry, the variants (C9, C12) and their EP-PCR libraries were expressed in yeast cells and screened on our microfluidic system with ~5000 cell counts. [2, 18]

- We, then performed targeted saturated mutagenesis on C9 and C12 at positions W143, I161 and Q163 based on bioinformatics guidance and sequence diversity of these amino-acid residues in mFruit series of FPs. [20] We mixed the site-directed C9 and C12 libraries, (Figure 1; Chapter 5) which contained variants with high lifetime and brightness. This library was called C12-X. The C12-3 mutant was selected from this library, based on plate-based lifetime screening.
- An alternate evolution trajectory from the C12-X library led us to Site-Directed Clones (SDC), where we sorted the variants to enhance the populations with lifetime longer than

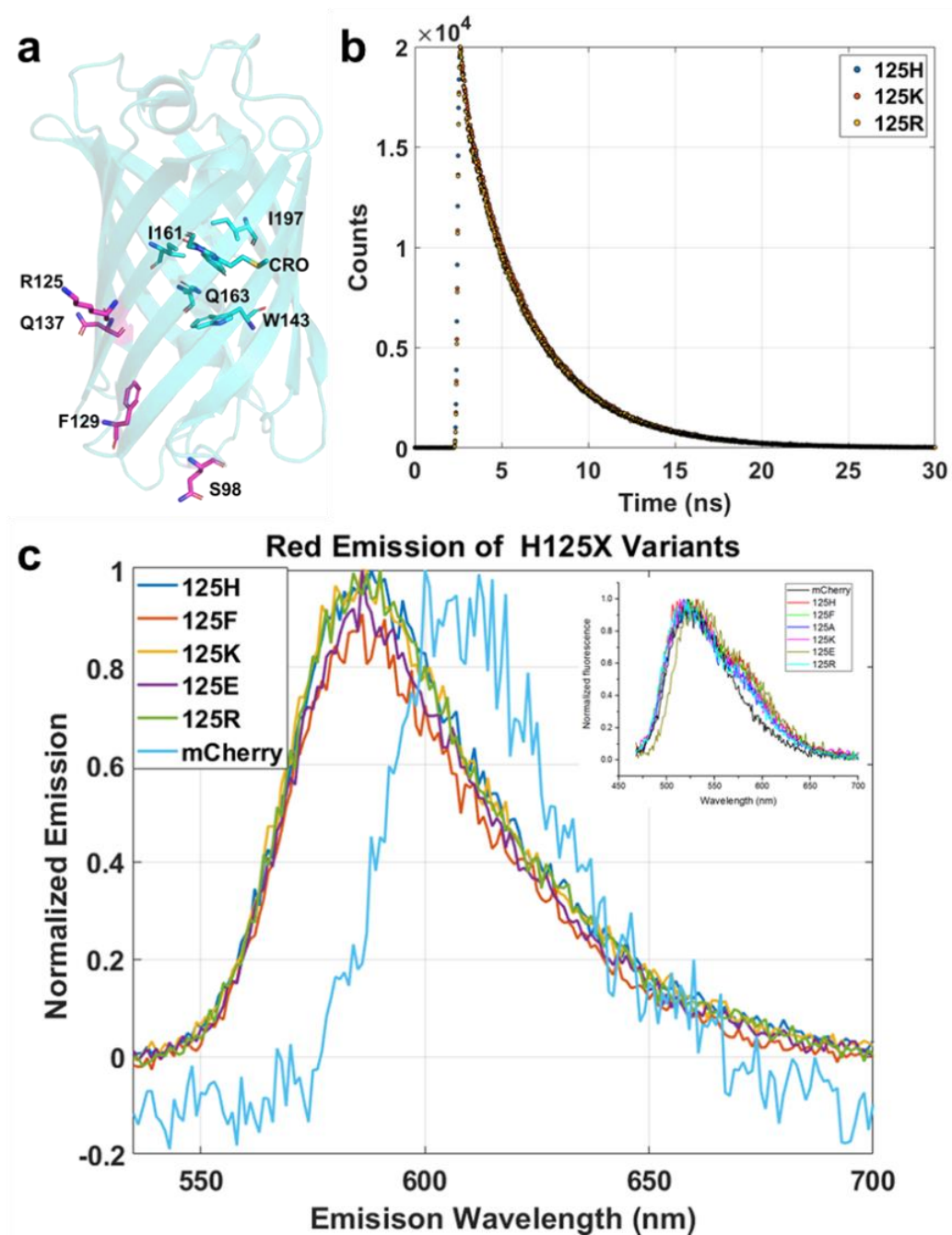
3.5 ns. We selected a few variants from this lifetime-enriched library and called them SDC variants. (e.g., variant SDC-5; QY~0.55, W143L, I161A, Q163L and I197R). The variants of the SDC-clones exhibit a green absorption peak – this evolution trajectory was consequently not pursued in the context of this study. Further analysis of this lineage can be found in Reference 21. [21]

- The C12-X library was subjected to two rounds of error-prone mutagenesis and selections for higher lifetime variants (with selection gate, lifetime > 3.7 ns) to generate the SLT (Selected-LifeTime) library. The SLT-11 mutant (mCherry N98S, R125H, F129L, Q137L, W143S, I161V, Q163Y, and I197R) was selected using plate-based lifetime screening in yeast from the lifetime enriched SLT library.
- A repeated error-prone effort on the SLT library yielded the EP-C12-X library without any improvement in lifetime or brightness. Consequently, we concluded that we arrived at the saturation of lifetime with the trajectory for directing the evolution of lifetime in mCherry.
- The octuple mutant SLT-11 had low protein yield in bacteria, consequently it could not be purified and fully characterized. Mutations at 98, 125, 129 and 137 were found to be external and far from the chromophore. Consequently, these positions were sequentially reverted to that of mCherry. Detailed investigations (single point reversions) revealed neither of these mutations had significant impacts on the excited state lifetime of the SLT-11 variant. We refer the reader to Section 3 of the SI and Reference 8 for further detailed investigations.<sup>21</sup> This led to the quadruple mutant that we address as mCherry-XL (mCherry W143S, I161V, Q163Y and I197R).
- Throughout the lifetime evolution a 561 nm excitation source (Coherent Genesis MX) and a 629/56 emission window (Semrock FF573-Di01- 25x36 / Semrock FF01- 629/56) provided additional selection pressures for selecting on a higher fluorescence brightness in yeast cells along with increased lifetime. [2, 11, 18, 21]



**Figure A4.2. Summary of mutagenesis efforts on the mCherry family to achieve lifetime evolution. (a.)** Evolution tree representative of phenotypic (photophysical) outcomes of the selections. The variants in gold were selected for characterization and study of co-evolution of other photophysical properties. **(b.)** A genetic representation of the evolution tree for the mutants relevant to this study.





**Figure A4.3. Role of spatially distant substitutions on the photophysical properties.** (a.) mCherry's crystal structure (PDB ID: 2H5Q) suggests positions that do not impact the photophysics i.e., 98, 125, 129 and 137 are spatially distant ( $>15 \text{ \AA}$ ) from the chromophore (denoted in purple). The chromophore and sidechains that perturb the photophysics of mCherry are denoted in cyan. Sidechains of 98, 125 and 137 also face out from the  $\beta$ -barrel, barring the exception of F129 which is located in the loop region with the sidechain facing inwards. The sidechain at 125 plays a role in the dimerization tendencies of DsRed derivatives like mCherry.<sup>20</sup>

A positively charged sidechain e.g. R or K at 125 maintains the monomericity and does not change the photophysics of the FP. To elucidate this, we present; **(b.)** The similarity in fluorescence lifetime decay traces for the 125X variants of mCherry-XL measured using TCSPC. To test the monomericity, we present time-resolved fluorescence anisotropy measurements in Figure A4.9. Details for the other reversions are available in Reference 21. [21] & **(c.)** A minimal effect of substitutions at 125 on variants of mCherry-XL for the red emission with excitation at 520 nm. The inset indicates similarity in minimal green emission for these variants along with mCherry for excitation at 480 nm.

### Section 3. Sequence Alignment of the RFP Variants

```

SLT-11|      EEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDI 60
mCherry-XL| EEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDI
C12-3|      EEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDI
C12|         EEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDI
mCherry|     EEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDI
*****

SLT-11|      LSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMSFEDGGVVTVTQDSSLQDGEFIY 120
mCherry-XL| LSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMSFEDGGVVTVTQDSSLQDGEFIY
C12-3|      LSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMSFEDGGVVTVTQDSSLQDGEFIY
C12|         LSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMSFEDGGVVTVTQDSSLQDGEFIY
mCherry|     LSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMSFEDGGVVTVTQDSSLQDGEFIY
*****

SLT-11|      KVKLRGTNFPDSDGPVMQKKTMGSEASSERMPEDGALKGEVKYRLKLDGGHYDAEVKTT 180
mCherry-XL| KVKLRGTNFPDSDGPVMQKKTMGSEASSERMPEDGALKGEVKYRLKLDGGHYDAEVKTT
C12-3|      KVKLRGTNFPDSDGPVMQKKTMGSEASSERMPEDGALKGECKLRKLDGGHYDAEVKTT
C12|         KVKLRGTNFPDSDGPVMQKKTMGSEASSERMPEDGALKGEIKQRLKLDGGHYDAEVKTT
mCherry|     KVKLRGTNFPDSDGPVMQKKTMGSEASSERMPEDGALKGEIKQRLKLDGGHYDAEVKTT
*****

SLT-11|      YKAKKPVQLPGAYNVNRKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYK 231
mCherry-XL| YKAKKPVQLPGAYNVNRKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYK
C12-3|      YKAKKPVQLPGAYNVNRKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYK
C12|         YKAKKPVQLPGAYNVNRKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYK
mCherry|     YKAKKPVQLPGAYNVNRKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYK
*****

```

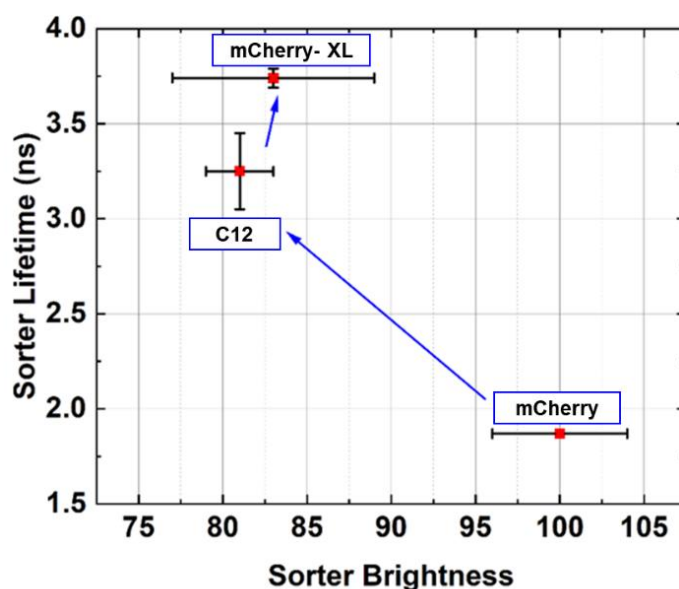
**Table A4.1. Sequence alignment of the mCherry variants analyzed in this study.** Variation in sequences have been highlighted with yellow/red. The positions highlighted in red show amino-acid substitutions that impact the photophysics of mCherry and its progeny, while the ones in yellow do not.

## Section 4. Additional *in cellulo* assessments

*Note: The version of mCherry-XL with the R125K substitution was used for cellular assessments.*

### 4a. Lifetime selection trajectory: Screening yeast cells for lifetime and cellular brightness

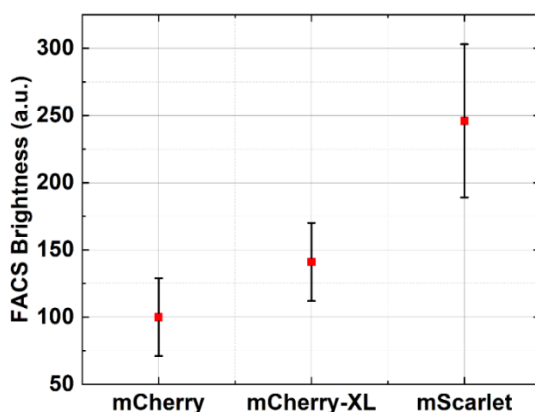
- **Screening on the microfluidic sorting device:** The proteins of interest, once selected out of the respective libraries were subjected to a round of microfluidic screening on the lifetime-sorter. Screenings were carried out at ~18-20 h post induction of FP in yeast cells with protocols used in the sorting experiments. Lifetime and brightness of biological triplicates of ~10,000 cells were recorded to report the screening data in yeast.



**Figure A4.4. Lifetime & yeast-cell co-evolution trajectory.** A screen of lifetime and brightness of single yeast cells on our microfluidic sorting devices [2, 11, 18, 21] indicate the lifetime evolution trajectory from mCherry to mCherry-XL. Means and standard deviations of biological triplicates are reported. Despite a 2.5-fold higher lifetime, mCherry-XL is roughly 90% of the brightness of mCherry presumably due to slower chromophore maturation, low expression efficiency, etc in yeast. [18]

## 4b. Brightness in mammalian cells

- **Flow cytometry:** The proteins of interest were fused to histone H2B and expressed in HeLa cells. Single-cell brightness was assessed by selecting single healthy cells based on forward and side-scattering photon counts on a BD FACSCelesta single cell analyzer after 48 h of transfection. Untransfected cells were used as a control to background subtract and analyze the fluorescence in the red and green channels. The samples were excited by a 561 nm laser line for collecting red fluorescence through the TRITC filter set (585/30 nm) and a 488 nm laser line for collecting through a GFP filter set (530/30 nm), to test for green fluorescence.

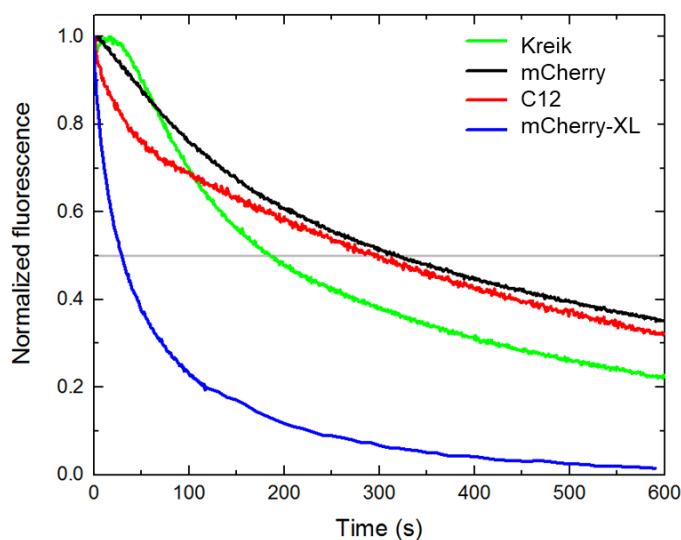


**Figure A4.5. Cellular brightness in mammalian cells.** A fluorescence activated cell sorting (FACS) based brightness assay in the H2B-FP construct in HeLa cells indicates mCherry-XL is 1.5-fold brighter than mCherry. FPs were cloned into a pcDNA3 plasmid C-terminal to H2B. Screenings were carried out at ~48 h post transfection. The error bars indicate the standard deviations from mean values of three biologically triplicate measurements. Each biological replicate measurement was comprised of a technical triplicate of 10000-20000 cells. The detailed methods for this assay have been discussed in a previous work. [2, 18]

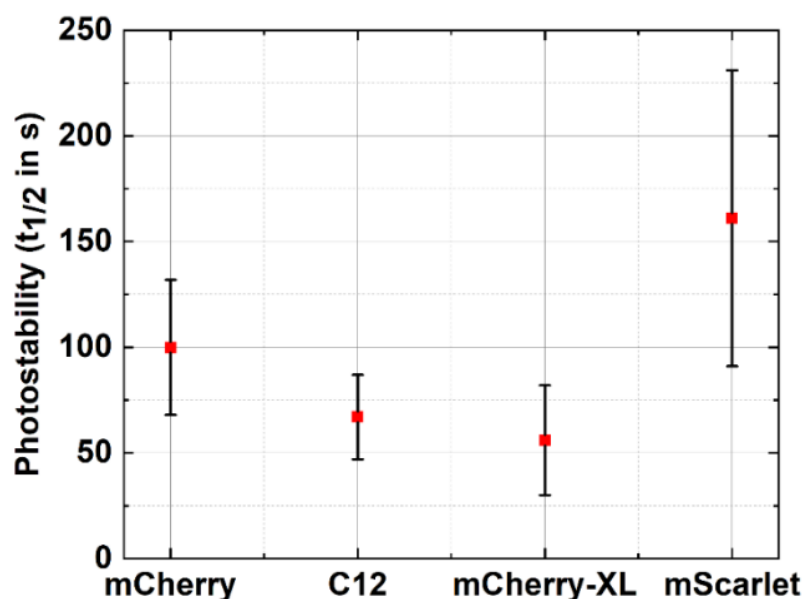
## 4c. Photostability in bacteria

Photobleaching experiments and analysis were carried out in a similar fashion as our previous studies. [2, 18, 23] We provide short descriptions to the sample preparation and data analysis below.

- **Sample Preparation and Experiment:** 2-3 fluorescent bacterial colonies from plates were washed with 500  $\mu\text{L}$  imaging buffer (150 mM HEPES, 100 mM NaCl, pH 7.4) and centrifuged at 3000-5000 RPM for 60 s, and the washing buffer was removed. These cells were then resuspended in the same buffer to an OD in the range of 0.1 to 0.5 to get a cell density suitable for imaging. 10–20  $\mu\text{L}$  of the cell mixture was added between a clean glass coverslip and slide. Imaging was carried on a widefield Olympus IX-73 inverted microscope system. Samples were excited by 560 nm continuous wave LED illumination (Lumencor). Time-lapse fluorescence bleaching measurements were collected through a 629/56 nm band-pass filter by a sCMOS camera (Andor Zyla) using an Olympus 20X objective. Irradiances ranging from 1–20  $\text{W}/\text{cm}^2$  were employed.
- **Data Analysis:** In brief, two pipelines in the CellProfiler suite (V2.2.0) was used to (a) identify bacteria on an imaging plane of uniform irradiance and (b) provide normalized intensity trajectories for the identified objects. The intensity trajectories were analyzed and fit to mono/bi/tri-exponential traces using a custom fitting program in MATLAB.



**Figure A4.6. Photobleaching trends in *E. coli*.** Bleaching profiles of mCherry, C12 and mCherry-XL are largely monoexponential and lack dark state conversion unlike FusionRed mutants investigated in our previous studies. [2, 18, 23] Data was collected at a continuous irradiation of  $\sim 10 \text{ W}/\text{cm}^2$  (not excitation rate normalized). Kreik is a previously published mCherry variant selected out of photostability selections on a microfluidic platform. [22]

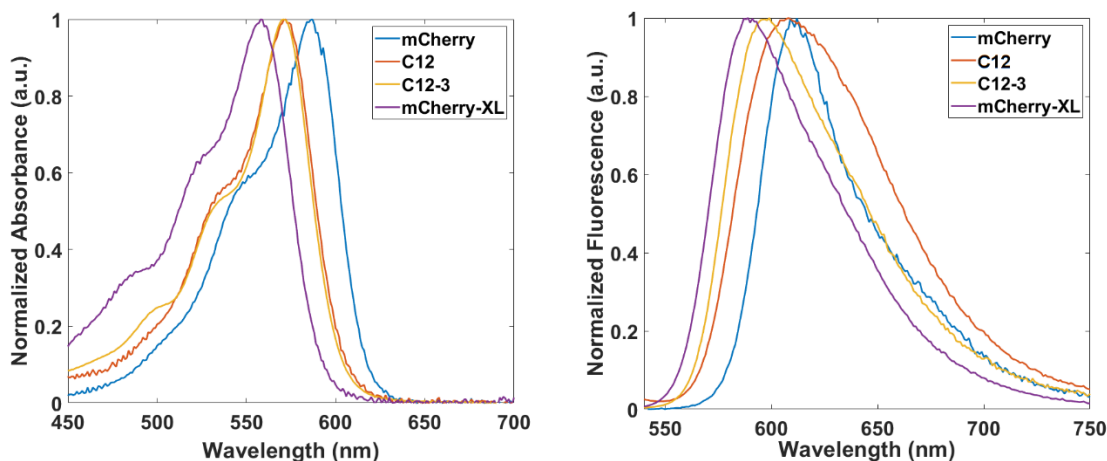


**Figure A4.7. Comparing photostability under excitation rate normalized conditions in *E. coli*.** Bacteria expressing RFPs of interest were photobleached under excitation normalized conditions using a 560 nm LED source at irradiances in the range of  $\sim 10 \text{ W/cm}^2$ . These FPs showed a near mono-exponential photobleaching trace (Figure A4.6). [23] Therefore, the  $t_{1/2}$  (s) is an accurate representation of photostability or the quantum yield of photobleaching under excitation rate normalized irradiation. [24] The expected trend of lower  $t_{1/2}$  with increased excited state lifetime is also seen in this evolution study – in accordance with previous observations from the lifetime evolution efforts directed at FusionRed and mScarlet. [2, 8, 18]



## Section 5. Additional details for *in vitro* photophysical assessments

### 5a. Spectral data of the variants

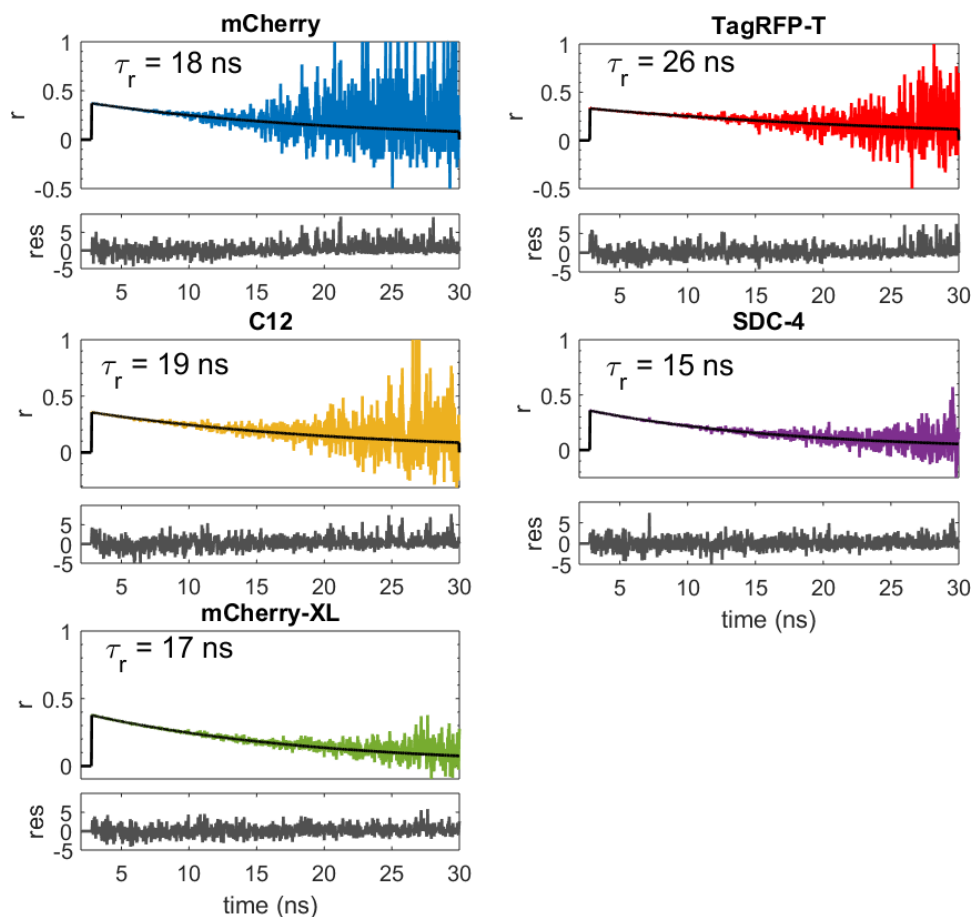


**Figure A4.8. Normalized absorption and emission spectra.** Normalized emission spectra (**right**) of the variants show blue-shifts from mCherry  $\rightarrow$  mCherry-XL as it is observed in the absorption spectra (**left**).

**Table A4.2. Spectral properties of the RFP variants.** Systematic blue shifts in absorption and emission peaks are observed with increments in fluorescence lifetimes. Integrated fluorescence, calculated from the normalized emission spectra of the variants correlates with the broadness of the emission. It is interesting to note that these variants have increased emission broadness and Stokes shift compared the parent mCherry.

RFP	$\lambda_{abs}$ (nm)	$\lambda_{em}$ (nm)	Stokes Shift (nm)	Int. Fluorescence
mCherry	587	609	22	66
C12	572	608	35	88
C12-3	571	599	28	77
mCherry-XL	558	589	31	79

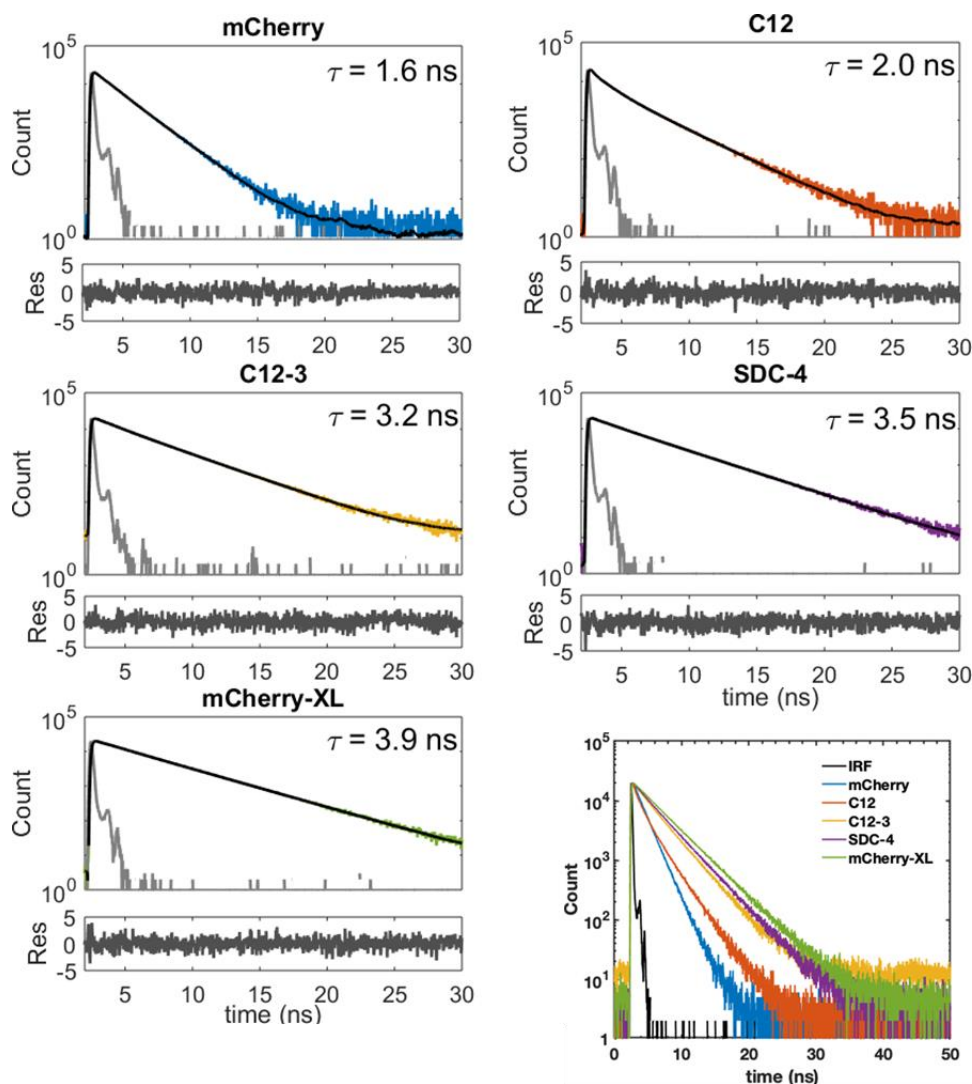
## 5b. Time-resolved fluorescence anisotropy



**Figure A4.9. Fluorescence anisotropy decay and rotational time-constants ( $\tau_r$ ) of the variants.** Anisotropy decay of the purified FPs were measured with our TCSPC instrument (Fluoro Time 100, PicoQuant) as mentioned in the Chapter 5. 560 nm excitation and 5 MHz repetition rate were used. The rotational time-constant ( $\tau_r$ ) of each variant was obtained by fitting the fluorescence anisotropy decay with a single exponential function. All the variants were found to have  $\tau_r$  similar to the parent mCherry, suggesting they likely retained *in vitro* monomeric character. TagRFP-T was characterized with slower  $\tau_r$ , relative to the other variants. This is consistent with its reported *in cellulo* oligomerization tendencies. [2, 18] The alternate evolution trajectory into SDC clones was not pursued due to a substantial green absorption peak. Details of the SDC variants are described elsewhere. [21]

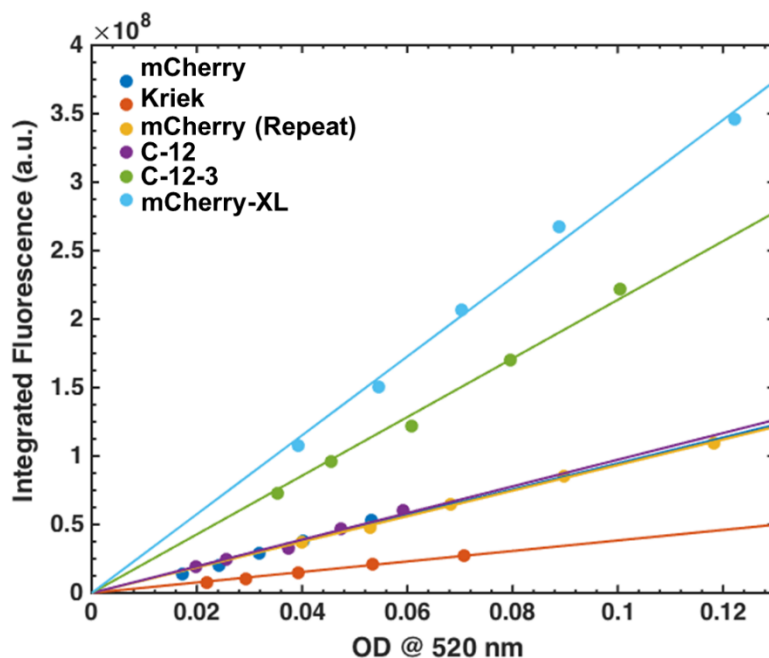


## 5c. Excited State Lifetime of the Variants



**Figure A4.10. Fluorescence decay and the average lifetime ( $\tau$ ) of the variants.** Excited state lifetimes are measured with our TCSPC system (Fluoro Time 100, PicoQuant) using 560-nm laser excitation with a repetition rate of 5 MHz. Lifetime measurements were performed with purified proteins. The fluorescence transients of the FPs were fitted with iterative reconvolution with a bi or tri-exponential function (solid black line) and using the measured instrument response function (IRF) of the system (shown in grey). The values of lifetime were obtained by intensity weighted average of the fitted lifetime constants. [25] The alternate evolution trajectory into SDC clones was not pursued due to a substantial green absorption peak. Details of the SDC variants are described elsewhere. [21]

## 5d. Measurement of fluorescence quantum yield

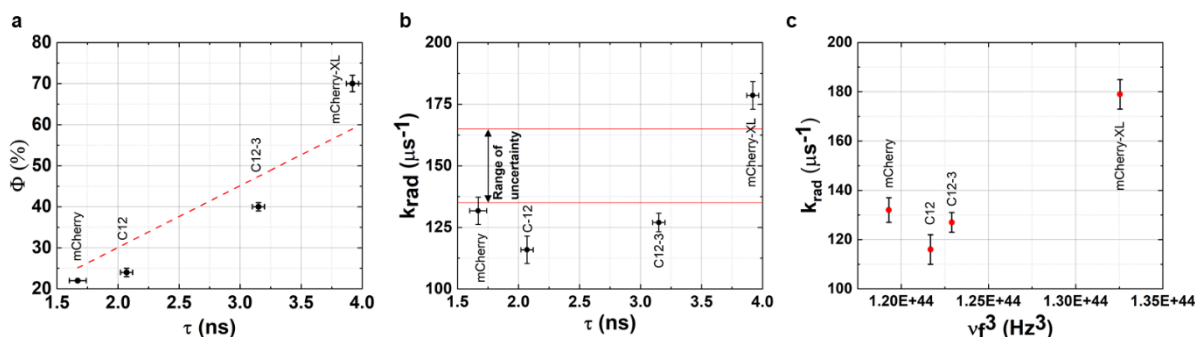


**Figure A4.11. Fluorescence quantum yield ( $\phi$ ) of the variants.** The quantum yield measurements were carried out by measuring the absorbance at 520 nm, and then collecting the integrated fluorescence over the entire red emission band (range 535-800 nm) across a step-dilution series. [2, 18] A higher slope indicates a higher quantum yield, with mCherry ( $\phi = 0.22$ ) and Cresyl Violet in EtOH ( $\phi = 0.54$ ; data not shown) were used as references. [2, 18]

## 5e. Estimation of the extinction coefficient

- The values of peak extinction coefficient ( $\epsilon_{\max}$ ) were determined using a ratiometric method as in our previous reports. [2, 18] To measure the  $\epsilon_{\max}$ , the following protocol was used:
- Sample 1: Blank A - 900  $\mu\text{L}$  of Tris-HCl buffer (pH 7.4). An absorption spectrum was recorded in the 250-750 nm range.
- Sample 2: Blank B - 900  $\mu\text{L}$  of Tris-HCl buffer (pH 7.4) + 100  $\mu\text{L}$  of 10 M NaOH (pH  $\sim$  14). An absorption spectrum was recorded in the 250-750 nm range
- Sample 3: 900  $\mu\text{L}$  of Tris-HCl buffer (pH 7.4) + a few  $\mu\text{L}$  of concentrated purified protein sample was added to adjust the absorbance to a value of OD  $\sim$  0.1. An absorption spectrum was recorded in the 250-750 nm range.
- Sample 4: 100  $\mu\text{L}$  of 10 M NaOH was added to this solution. An absorption spectrum was recorded in the 250-750 nm range.
- In each case a 1 cm path length quartz cuvette was used.
- The ratiometric relationship provided below was used to calculate the value of the peak extinction coefficient. Three independent measurements were performed. Table 1 in the Chapter 5 represents the value of the mean and the standard deviation errors.
- $$\epsilon_{\max\text{RFP}} = \frac{\text{Abs}_{\max\text{RFP}}}{\left(\frac{\text{Abs}_{380\text{ nm}}}{\epsilon_{380\text{ nm}}}\right) + \left(\frac{\text{Abs}_{450\text{ nm}}}{\epsilon_{450\text{ nm}}}\right)}$$
- Unlike members of the FusionRed family,<sup>18</sup> mCherry variants hydrolyze with a single product of denaturation, with an absorbance peak centered at 450 nm.

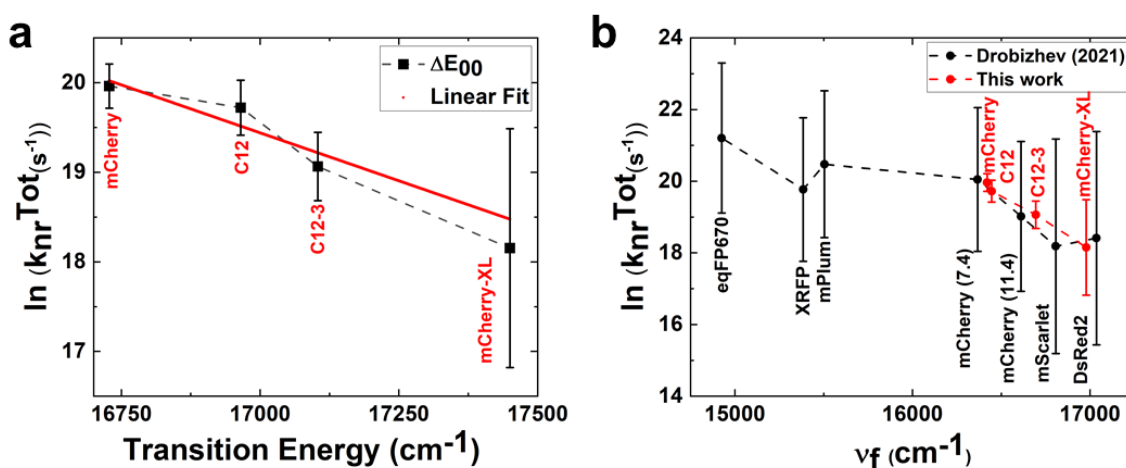
## Section 6. Variation of the radiative rate constant



**Figure A4.12. Analysis of the radiative rate constants for mCherry variants.** (a.) A linear fit of the fluorescence quantum yield with the fluorescence lifetime indicates a good fit ( $R^2_{adj} \sim 0.95$ ). The slope of the linear fit ( $150 \pm 14 \mu s^{-1}$ ) is indicative of the average radiative rate constant. (b.) Values of calculated radiative rate constant with the fluorescence lifetime. The calculated values of the radiative rate constant lie outside the estimates provided by the average radiative rate (provided by the range of uncertainty calculated from the linear fit in panel a. on the figure). (c.) Values of calculated radiative rate constant with the cubed value of the peak fluorescence frequency. The calculated values of the radiative rate constants do not have a linear response with the cubed-value of the fluorescence frequency, thereby suggesting the change in the radiative rate constant is possibly due to multiple inter- or non-interconverting chromophores. This observation is further bolstered by multiexponential fluorescence lifetime decays. [24]

## Section 7. Modelling non-radiative rate in the Englman-Jortner Low Temperature/Weak Coupling limit

A fit of the total non-radiative rate with the energy gap (or the 0-0 energy) shows a poor agreement with an Arrhenius type exponential decay. (Figure A4.13) This observation prompts investigating other models that incorporate the role of excited state reorganization along with the energy gap, such as the Englman-Jortner model for radiationless decay in large organic molecules. Details of this model and its treatment in FP systems can be found in ref # 24. [24]



**Figure A4.13. Fit of the total non-radiative rate constant with the energy gap.** (a.) A linear fit (red line) of the logarithmic values of non-radiative rate (black squares) with the energy gap indicates a low-quality fit ( $R^2_{adj} \sim 0.79$ ). (b.) A comparison of non-radiative rate versus the fluorescence frequency for the variants investigated in this study with the analysis performed by the recent work of Drobizhev et. al. (2021) [26] The variants generated in this study display trends similar to what was investigated by Drobizhev et. al for a subset of unrelated RFPs.

The shapes of absorption and fluorescence spectra for FPs of this family (Figure A4.8), indicate that the Huang-Rhys factor is  $< 1$  because the 0-0 transition is more pronounced than the 0-1 vibronic shoulder. Additionally, the 0-0 transition energy  $\gg$  Stokes-shift, and assuming these FPs lack any excited state photochemistry or excited state proton-transfer encourages the treatment of the non-radiative rate with a model where two electronic potential energy manifolds couple weakly. Therefore, we utilize the treatment of excited state vibrational relaxation from the low

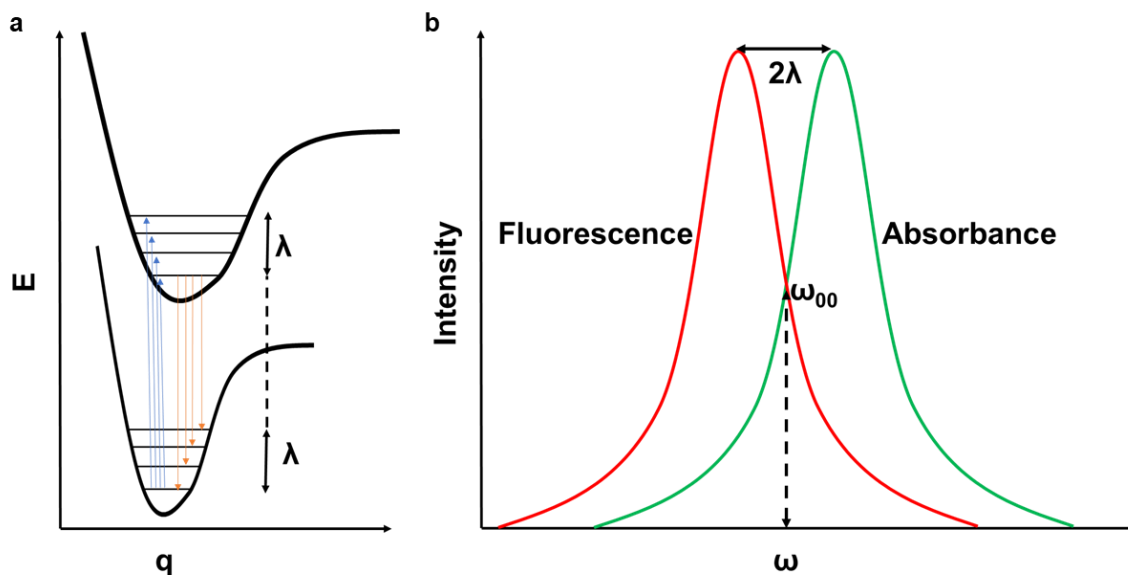
temperature weak/coupling case of the Englman-Jortner formalism. We now proceed to the estimation of the excited state reorganization energy.

### 7a. Estimation of reorganization energy

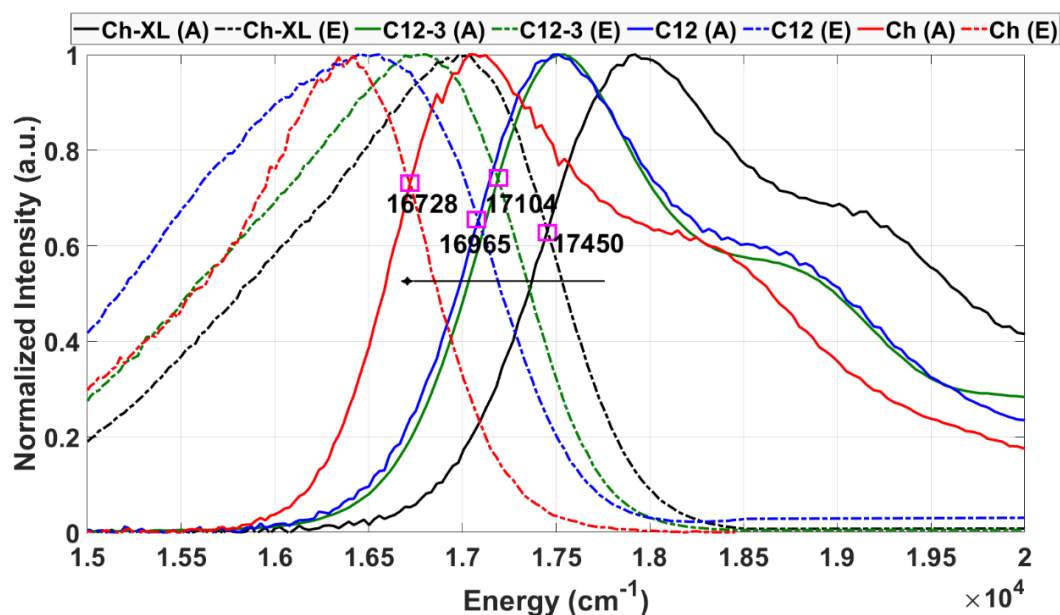
The reorganization energy was estimated using the model provided by Jordanides *et. al.* [27] This method can provide an accurate representation for the reorganization energy for species that do not exhibit Gaussian absorption and fluorescence spectra (Figure A4.14). For species that do exhibit near Gaussian absorption and emission spectra a traditional method of estimating the reorganization energy from the Stokes-shift ( $\lambda_{SS}$ ) can be used, [28]

$$\lambda_{SS} = 0.5 (\max[\sigma_{abs}(\omega)] - \max[\sigma_{fl}(\omega)]), \quad \text{Eq. S1}$$

Where  $\sigma_{abs}(\omega)$  and  $\sigma_{fl}(\omega)$  are the absorbance and fluorescence spectral lineshape functions in the frequency domain.



**Figure A4.14. Schematic representations of relevant observables (a.)** A two state potential energy surface indicating the visual depiction of the excited state reorganization energy. **(b.)** Representative Gaussian absorbance and fluorescence spectra with a 0-0 transition where the reorganization energy is exactly equal to half the Stokes Shift.



**Figure A4.15. Absorbance and fluorescence spectra represented on the wavenumber scale.**

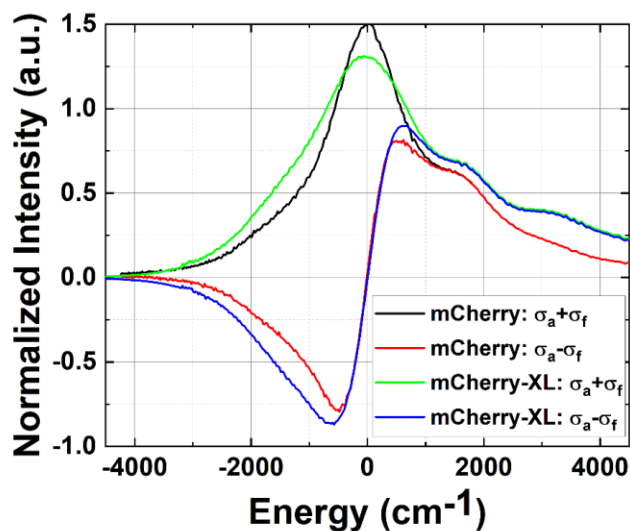
The spectra were scaled, also considering the nonlinear transformation of the wavelength scale to the wavenumber scale. [29] The pink boxes highlight the 00-transition energy and the arrow indicates the decrease in the energy gap of the transition.

However as shown in Figure A4.15, the presence of strong vibronic sidebands in the spectra for these FPs indicates deviation from Gaussian behavior. The model by Jordanides et. *al* provides a means of estimating the reorganization energy when this deviation from Gaussian behavior is observed. The only assumption of this model is that the polarization fluctuations of the dielectric medium follow Gaussian statistics (or the medium follows linear response). [27] The reorganization energy is calculated from the normalized difference of the first moment between the absorption and fluorescence spectra around the transition energy ( $\omega_{00}$ ). With this approach we arrive at the following expression for the reorganization energy:

$$\lambda_{SM} = \frac{\int d\omega \sigma_{abs}(\omega - \omega_{00}) - \sigma_{fl}(\omega - \omega_{00}) (\omega - \omega_{00})}{\int d\omega \sigma_{abs}(\omega - \omega_{00}) + \sigma_{fl}(\omega - \omega_{00})}, \text{ Eq. S2}$$

Using this method to calculate the reorganization energy, we arrive at significantly higher reorganization energy estimates. In reference 27, the reorganization energy of Eosin in water was estimated as  $359\text{ cm}^{-1}$  using half the value of Stokes shift and  $877\text{ cm}^{-1}$  and utilizing the spectral moments. [27] We estimated values for EosinY in ethanol and observe similar trends with the reorganization energy from the Stokes-shift being  $317\text{ cm}^{-1}$  and  $706\text{ cm}^{-1}$  using the spectral moments. It should be noted that the tail ends of the absorption and emission spectrum must be carefully baseline-corrected to ensure accurate integration over the areas of the sum and the difference spectrum. Poor baseline correction can result in erroneous values of the reorganization energy.

We performed this analysis for the RFP variants and observe that the spectral moments consistently provide larger estimates of reorganization energy than the Stokes-shift but the trends are consistent with both approaches (Table 3 and Figure 2, Chapter 5). In both models, mCherry displays the lowest reorganization energy in this series of variants.



**Figure A4.16.** The numerator and denominator functions for in the Eq S2. The plots indicate the expected odd and even function behavior for mCherry and mCherry-XL respectively about  $\omega_0$ .



## 7b. Validity of similar coupling constants for Herzberg-Teller coupling

In our analysis of non-radiative rates, we assume that the entire non-radiative rate in mCherry-XL derives from the values estimated from the Englman-Jortner approach. We then utilize this value, the energy gap, and the reorganization energy to estimate the coupling constant in the pre-exponent. Though the pre-exponent exhibits a very minor dependence on the overall non-radiative rate (as described in the original work of Englman and Jortner), it is important to provide a quantitative assessment for this assumption. These FPs exhibit a strong vibronic progression for the 0-1 vibronic transition in their absorption spectra (Figure A4.8). Therefore, to identify the strength of intensity borrowing of the 0-1 band from the 0-0 band, we fit the bands to two Gaussian functions and analyze their energy difference and the relative intensities. The energy difference can also be calculated using the second derivatives of the absorption spectra that reveal peak positions of these vibronic progressions. Using the spectral derivative method Lin et. *al* arrived at a value of  $1340 \pm 40 \text{ cm}^{-1}$  for the bond-length alteration (BLA) mode of the GFP chromophore. In our analysis (results are presented in Table A4.3), we found this difference to be a value of  $966 \pm 45 \text{ cm}^{-1}$ , which can indicate a weaker BLA mode for the extended RFP chromophore system. [30]

**Table A4.3. Difference in energies of the peaks of the 0-0 and 0-1 vibronic bands and the relative intensities.**

FP	$\Delta E_{01} - \Delta E_{00} (\text{cm}^{-1})$	$I_{00}/I_{01}$
mCherry	1008	2.05
C-12	924	2.31
C-12-3	930	2.17
mCherry-XL	1003	2.42

The minor (~10%) variation in the energy and the relative intensities provide quantitative validity to the assumption of a fixed Herzberg-Teller coupling constant for the RFP family.

# Appendix 5

## Section 1. Code for microfluidic screen dot-plots

```
function hAxes = dscatter(X,Y, varargin)
%   DSCATTER(X,Y) creates a scatterplot of X and Y at the locations
%   specified by the vectors X and Y (which must be the same size), colored
%   by the density of the points.
%
%
%       [data, params] = fcsread('SampleFACS');
%       dscatter(data(:,1),10.^(data(:,2)/256),'log',1)
%       % Add contours
%       hold on
%       dscatter(data(:,1),10.^(data(:,2)/256),'log',1,'plottype','contour')
%       hold off
%       xlabel(params(1).LongName); ylabel(params(2).LongName);
%
% Reference:
% Paul H. C. Eilers and Jelle J. Goeman
% Enhancing scatterplots with smoothed densities
% Bioinformatics, Mar 2004; 20: 623 - 628.

lambda = [];
nbins = [];
plottype = 'scatter';
contourFlag = false;
msize = 10;
marker = 's';
logy = false;
filled = true;
if nargin > 2
    if rem(nargin,2) == 1
        error('Bioinfo:IncorrectNumberOfArguments',...
            'Incorrect number of arguments to %s.',mfilename);
    end
    okargs = {'smoothing','bins','plottype','logy','marker','msize','filled'};
    for j=1:2:nargin-2
        pname = varargin{j};
        pval = varargin{j+1};
        k = strmatch(lower(pname), okargs); %#ok
        if isempty(k)
            error('Bioinfo:UnknownParameterName',...
                'Unknown parameter name: %s.',pname);
        elseif length(k)>1
            error('Bioinfo:AmbiguousParameterName',...
                'Ambiguous parameter name: %s.',pname);
        else
            switch(k)
                case 1 % smoothing factor
                    if isnumeric(pval)
                        lambda = pval;
                    else
                        error('Bioinfo:InvalidScoringMatrix','Invalid smoothing parameter.');
```

```

        case 6
            marker = pval;
        case 7
            msize = pval;
        case 8
            filled = pval;
    end
end
end
end

minx = min(X,[],1);
maxx = max(X,[],1);
miny = min(Y,[],1);
maxy = max(Y,[],1);

if isempty(nbins)
    nbins = [min(numel(unique(X)),200) ,min(numel(unique(Y)),200) ];
end

if isempty(lambda)
    lambda = 10;
end

edges1 = linspace(minx, maxx, nbins(1)+1);
ctrsl = edges1(1:end-1) + .5*diff(edges1);
edges1 = [-Inf edges1(2:end-1) Inf];
edges2 = linspace(miny, maxy, nbins(2)+1);
ctrsl = edges2(1:end-1) + .5*diff(edges2);
edges2 = [-Inf edges2(2:end-1) Inf];

[n,p] = size(X);
bin = zeros(n,2);
% Reverse the columns to put the first column of X along the horizontal
% axis, the second along the vertical.
[dum,bin(:,2)] = histc(X,edges1);
[dum,bin(:,1)] = histc(Y,edges2);
H = accumarray(bin,1,nbins([2 1])) ./ n;
G = smooth1D(H,nbins(2)/lambda);
F = smooth1D(G',nbins(1)/lambda)';
% = filter2D(H,lambda);

if logy
    ctrsl = 10.^ctrsl;
    Y = 10.^Y;
end
okTypes = {'surf','mesh','contour','image','scatter'};
k = strmatch(lower(plottype), okTypes); %#ok
if isempty(k)
    error('dscatter:UnknownPlotType',...
        'Unknown plot type: %s.',plottype);
elseif length(k)>1
    error('dscatter:AmbiguousPlotType',...
        'Ambiguous plot type: %s.',plottype);
else
    switch(k)
        case 1 %'surf'
            h = surf(ctrsl,ctrsl,F,'edgealpha',0);
        case 2 % 'mesh'
            h = mesh(ctrsl,ctrsl,F);
        case 3 %'contour'
            [dummy, h] =contour(ctrsl,ctrsl,F);
        case 4 %'image'
            nc = 256;
            F = F./max(F(:));
            colormap(repmat(linspace(1,0,nc)',1,3));
            h =image(ctrsl,ctrsl,floor(nc.*F) + 1);
        case 5 %'scatter'
            F = F./max(F(:));
            ind = sub2ind(size(F),bin(:,1),bin(:,2));

```

```

        col = F(ind);
        if filled
            h = scatter(X,Y,msize,col,marker,'filled');
        else
            h = scatter(X,Y,msize,col,marker);
        end
        %colormap(repmat(linspace(1,0,256)',1,3));

    end

end

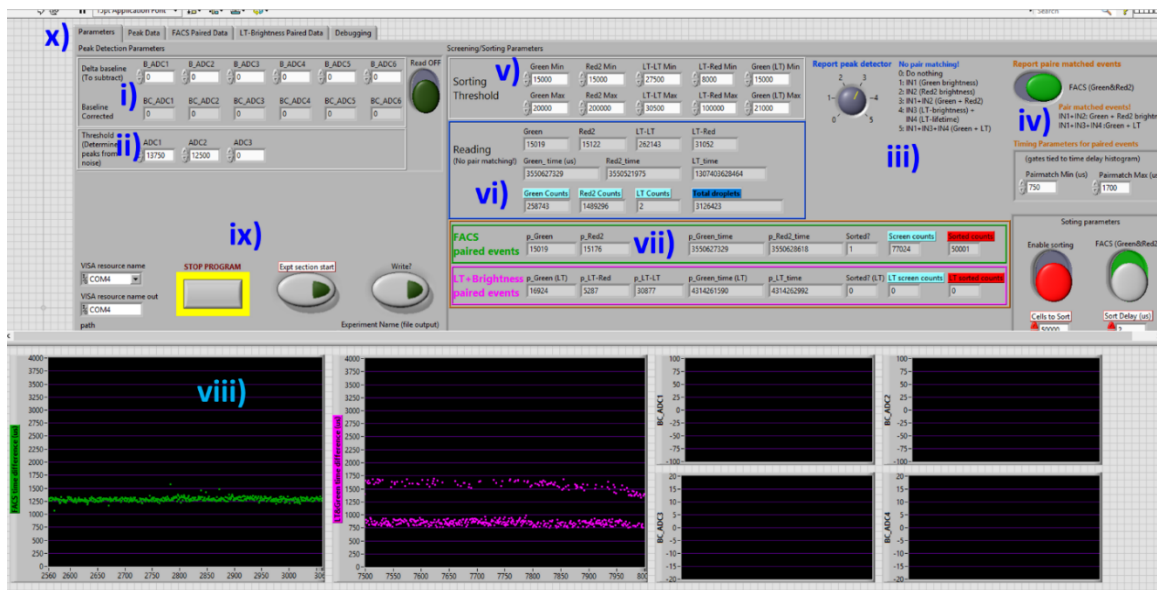
if logy
    set(gca,'yscale','log');
end
if nargout > 0
    hAxes = get(h,'parent');
end
%%% This method is quicker for symmetric data.
% function Z = filter2D(Y,bw)
% z = -1:(1/bw):1;
% k = .75 * (1 - z.^2);
% k = k ./ sum(k);
% Z = filter2(k'*k,Y);

function Z = smooth1D(Y,lambda)
[m,n] = size(Y);
E = eye(m);
D1 = diff(E,1);
D2 = diff(D1,1);
P = lambda.^2 .* D2'*D2 + 2.*lambda .* D1'*D1;
Z = (E + P) \ Y;

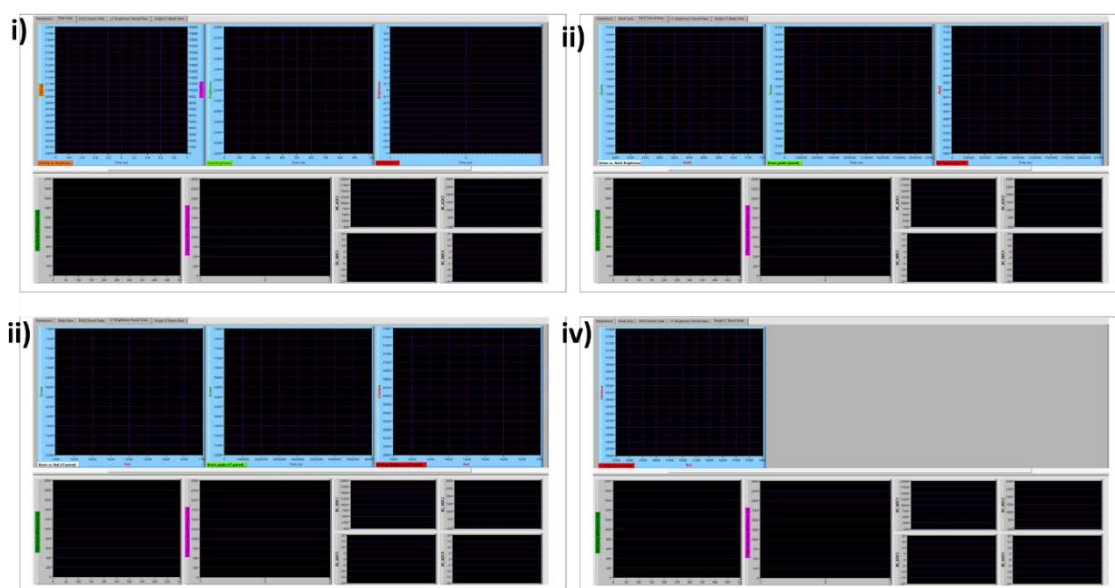
```

## Section 2. Droplet-sorter operation controls

### 2a. Graphical user interface



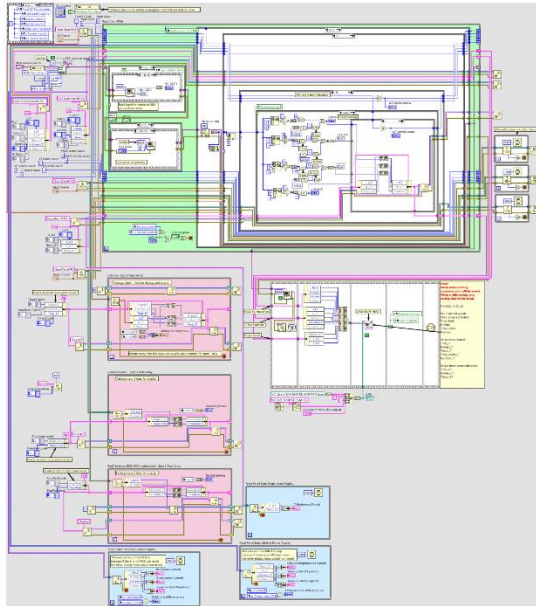
**Figure A5.1. The front panel of the NI LabView interface for the droplet sorter operation.** i) Setting baseline values in least-significant bit (LSB) units for background corrections in each ADC (analog to digital convertor card) channel. Standard mode of operation involves ADC1 for red channel brightness, ADC2 for green channel brightness and ADC3 for lifetime readout from the lock in amplifier. Other ADCs 4,5, and 6 allow for the data acquisition of the FPGA board for other modes of future operation such as dark state screening. ii) Threshold value for peak determination for ADC channels 1,2 and 3. iii and iv) Sorter operation mode selector: red brightness, green brightness, red+green brightness, red brightness+lifetime etc. Various combinations of these multiparameter modalities allow for screening/sorting analytes. v) Sort-threshold gates for ADC channels. vi and vii) Real time read outs for sort/screen events and file writing modalities. viii) Graphical displays of real time screening/sorting events on each ADC channel. Green dots indicate the events above threshold in the green channel (ADC1) and the pink dots indicate the events above the threshold ix) Program on/off and data saving controls x) Tabs that allow to real time data visualization tools for each mode of operation like lifetime, event pairing etc. The panels are presented in Figure A5.2.



**Figure A5.2. The additional panels of the NI LabView interface for the droplet sorter operation.** i) Graphical representation of peak data for green brightness, red brightness and lifetime data. This window is an active real time data display mode irrespective of the mode of operation ii) Graphical representation of dual color pair matching events. This window is active on dual color operation. iii) Graphical representation of pair matched events. This window follows data from window ii only if the events are pair matched iv) Lifetime and brightness data in a single mode of operation (red brightness and lifetime). This mode is active whenever red channel brightness and lifetime data are being collected by the sorter.

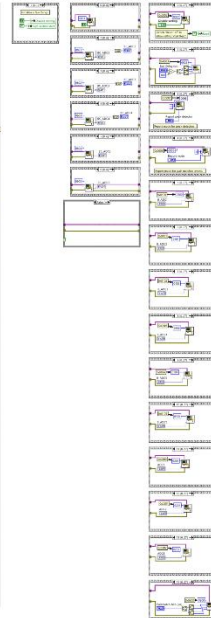
## 2b. Block diagram for the sorter

Sorter\_SP53\_v9.vi  
C:\Users\ZDMONSTER\_SORTER2\Sheng\DSC sorter program\Sorter\_SP53\_v9.vi  
Last modified on 12/23/2018 at 12:11 PM  
Printed on 2/21/2022 at 7:36 PM

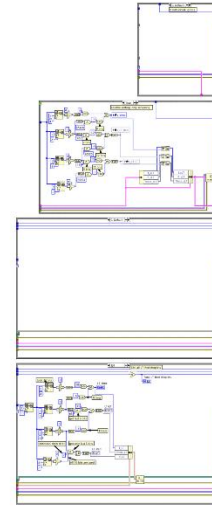


Page 1

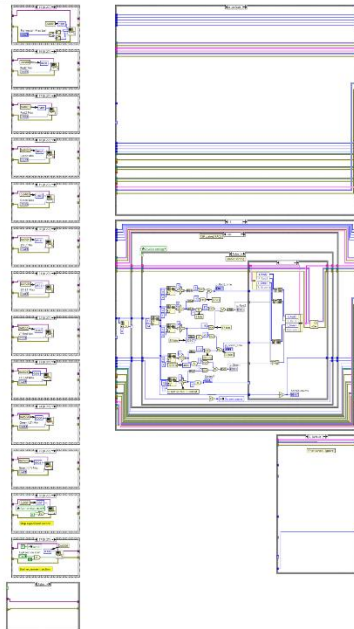
Sorter\_SP53\_v9.vi  
C:\Users\ZDMONSTER\_SORTER2\Sheng\DSC sorter program\Sorter\_SP53\_v9.vi  
Last modified on 12/23/2018 at 12:11 PM  
Printed on 2/21/2022 at 7:36 PM



Page 2

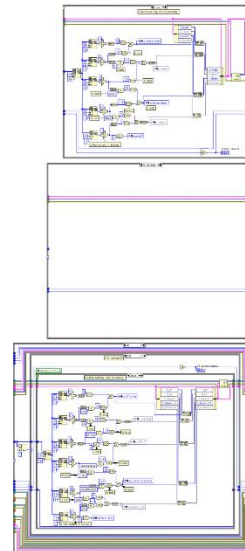


Sorter\_SP53\_v9.vi  
C:\Users\ZDMONSTER\_SORTER2\Sheng\DSC sorter program\Sorter\_SP53\_v9.vi  
Last modified on 12/23/2018 at 12:11 PM  
Printed on 2/21/2022 at 7:37 PM



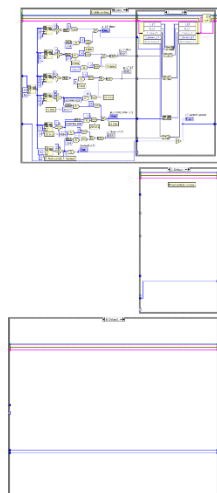
Page 3

Sorter\_SP53\_v9.vi  
C:\Users\ZDMONSTER\_SORTER2\Sheng\DSC sorter program\Sorter\_SP53\_v9.vi  
Last modified on 12/23/2018 at 12:11 PM  
Printed on 2/21/2022 at 7:37 PM



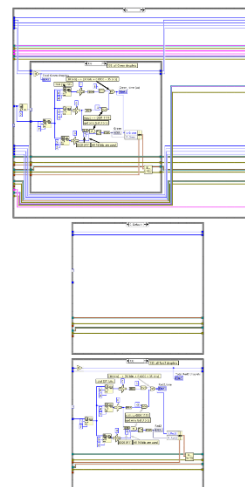
Page 4

Sorter\_SP53\_v9.vi  
 C:\Users\ZDMONSTER\_SORTER2\Sheng\DSC sorter program\Sorter\_SP53\_v9.vi  
 Last modified on 12/23/2018 at 12:11 PM  
 Printed on 2/21/2022 at 7:37 PM



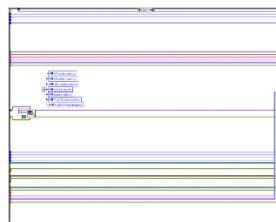
Page 5

Sorter\_SP53\_v9.vi  
 C:\Users\ZDMONSTER\_SORTER2\Sheng\DSC sorter program\Sorter\_SP53\_v9.vi  
 Last modified on 12/23/2018 at 12:11 PM  
 Printed on 2/21/2022 at 7:37 PM



Page 6

Sorter\_SP53\_v9.vi  
 C:\Users\ZDMONSTER\_SORTER2\Sheng\DSC sorter program\Sorter\_SP53\_v9.vi  
 Last modified on 12/23/2018 at 12:11 PM  
 Printed on 2/21/2022 at 7:37 PM



Page 7

**Figure A5.3. The block diagram of the sorter GUI displayed in Figures A5.1 and A5.2**

## Section 3. Simulation code representing the effects of scattering on phase shift

% This program simulates the effect of scattering on phase shift observed by a lock in amplifier

```
t=0.001:0.001:2; %0-200us Time step 1ns
phi=pi/16; % 1.8 degree phase shift
f1=0.5*sin(2*pi*29.65*t+phi);
sc=0.5*sin(2*pi*29.65*t);
%plot(t(1,1:500),f1(1,1:500));
%plot(t(1,1:500),sc(1,1:500));
phase=zeros(1,100)
for st=1:1:100;

sig=((st-1)*sc)/100+f1;

I=zeros(1,2000);
Q=zeros(1,2000);
```



```

I=(sin(2*pi*29.65*t));
Q=(cos(2*pi*29.65*t));
Xem=sig*I.';
Yem=sig*Q.';
phase(1,st)=atand(Yem/Xem)

```

```
end
```

## Section 4. Simulation code representing the signal of a droplet in flow

```

% This program simulates a scattering and a fluorescent signal seen on flow
droplet=0.1*scat(1,499001:501000);
cell=co_sig(1,499001:501000);

```

```

r=randi([0,5],1,3); %generating random integers

sig=zeros(1,5*2000);
for j=0:1:6
    if j == r(1,1)
        disp('filled')
        sig(1,(2000*j)+1:2000*(j+1))=cell(1,1:2000);
    end
    if j == r(1,2)
        disp('filled')
        sig(1,(2000*j)+1:2000*(j+1))=cell(1,1:2000);
    end
    if j == r(1,3)
        disp('filled')
        sig(1,(2000*j)+1:2000*(j+1))=cell(1,1:2000);
    else
        j
        sig(1,(2000*j)+1:2000*(j+1))=droplet(1,1:2000);
    end
end

noise=rand(1,14000)
s=sig+0.1*noise;
time=0.5:0.5:7000;
plot (time(1,:),s(1,:));

```

## Section 5. Codes to analyze *in vitro* photophysical data

### 5a. Quantum yield

%Notes for Use: ASCII data is imported in the form of a wavelength column, the first sample absorbance for four consecutive dilutions. The first block just extracts data from the spreadsheet and plots it. Data is characteristic of the csv format of absorption data from the Agilent scan program of the Cary 5000 UV-Vis spectrophotometer.

```

SamplAbs=mCherry;
SamplAbs=table2array(SamplAbs);
x=SamplAbs(:,1);
y1=SamplAbs(:,2);
y2=SamplAbs(:,4);
y3=SamplAbs(:,6);
y4=SamplAbs(:,8);
plot(x,y1,x,y2,x,y3,x,y4)

```

%This is a way of averaging a specified set of points and subtracting the mean from their original dataset, i.e. setting the tail of the absorption spectrum where no absorbance is expected to zero.

```
Baseline=y1(1:50,1);
```

```

Mean=mean(Baseline);
y1=y1-Mean;
Baseline=y2(1:50,1);
Mean=mean(Baseline);
y2=y2-Mean;
Baseline=y3(1:50,1);
Mean=mean(Baseline);
y3=y3-Mean;
Baseline=y4(1:50,1);
Mean=mean(Baseline);
y4=y4-Mean;
plot(x,y1,x,y2,x,y3,x,y4)

%This is for making and visualizing normalized spectra.

Smax=max(y1);
Y1Norm=y1/Smax;
Smax=max(y2);
Y2Norm=y2/Smax;
Smax=max(y3);
Y3Norm=y3/Smax;
Smax=max(y4);
Y4Norm=y4/Smax;
RCB2NormAbs=table(x,Y1Norm,Y2Norm,Y3Norm,Y4Norm);

% This section is relevant to absorbance values used for exciting fluorescence, for example row
181 in this data set corresponds to the wavelength 520 nm at which fluorescence was excited at.

WL=x(181,1)
A1=y1(181,1);
A2=y2(181,1);
A3=y3(181,1);
A4=y4(181,1);

%This extracts and blanks fluorescence data from the Horiba-Yvon Jobin fluorimeter exported to a
csv file an ASCII format. This presumes that there are 5 spreadsheets per sample which need to be
imported: the blank and measurements 1-4. Each of these spreadsheets should have 4 columns. The
first specifies the wavelength and the fourth states Slc/Rlc; signal CPS corrected for the current.

x=Blank(:,1);
Baseline=Blank(:,4);
y1=mCherryFlu1(:,4);
y1=table2array(y1);
y1=y1-Baseline;
Flumin=nanmin(y1);
y1=y1-Flumin;
y2=mCherryFlu2(:,4);
y2=table2array(y2);
y2=y2-Baseline;
Flumin=nanmin(y2);
y2=y2-Flumin;
y3=mCherryFlu3(:,4);
y3=table2array(y3);
y3=y3-Baseline;
Flumin=nanmin(y3);
y3=y3-Flumin;
y4=mCherryFlu4(:,4);
y4=table2array(y4);
y4=y4-Baseline;
Flumin=nanmin(y4);
y4=y4-Flumin;
plot(x,y1,x,y2,x,y3,x,y4);

%This is largely identical to the block above that makes normalized spectra only this time for
fluorescence.

Smax=max(y1);
Y1Norm=y1/Smax;
Smax=max(y2);
Y2Norm=y2/Smax;
Smax=max(y3);

```

```

Y3Norm=y3/Smax;
Smax=max(y4);
Y4Norm=y4/Smax;
RCB2NormFlu=table(x, Y1Norm,Y2Norm,Y3Norm,Y4Norm);

% Trapezoidal integration of the area under the fluorescence curve.

F1=trapz(y1);
F2=trapz(y2);
F3=trapz(y3);
F4=trapz(y4);

% The variable QY is the linear fit of the integrated fluorescence and the recorded absorbance
that provides the slope value. Comparative values of the slope against a reference value provide
the estimated quantum yield.

Absorbance=table(A1, A2,A3,A4,0);
Absorbance=table2array(Absorbance);
IntegratedFluorescence=table(F1,F2,F3,F4,0);
IntegratedFluorescence=table2array(IntegratedFluorescence);
Absorbance=Absorbance';
IntegratedFluorescence=IntegratedFluorescence';
RCB2Table=table(Absorbance,IntegratedFluorescence)
scatter(Absorbance,IntegratedFluorescence)
QY=fitlm(Absorbance,IntegratedFluorescence)

```

## 5b. Extinction coefficient

%Notes for Use: ASCII data is imported in the form of a wavelength column, the first sample absorbance and consequently the denatured sample. Data is characteristic of the csv format from the Agilent scan program of the Cary 5000 UV-Vis spectrophotometer. The data is collected and graphed out for viewing.

```

SamplAbs=mCherry;
SamplAbs=table2array(SamplAbs);
x=SamplAbs(:,1);
y1=SamplAbs(:,2);
y2=DenatAbs(:,4);
plot(x,y1,x,y2,x,y3,x,y4)

```

```

%This is a way of averaging a specified set of points and subtracting the mean from their original
dataset, i.e. setting the tail of the absorption spectrum where no absorbance is expected to zero.
Baseline=y1(1:50,1);
Mean=mean(Baseline);
y1=y1-Mean;
Baseline=y2(1:50,1);
Mean=mean(Baseline);
y2=y2-Mean;
Baseline=y3(1:50,1);
Mean=mean(Baseline);

```

%This is for making and visualizing normalized spectra.

```

Smax=max(y1);
Y1Norm=y1/Smax;
Smax=max(y2);
Y2Norm=y2/Smax;

```

```

%This is for finding the peak absorbance wavelength of the native FP
result = find(Y1Norm(:,1)==1);

```

%This extracts absorbance (y-axis) values for the peak of the native FP, the value of denatured species at 380 nm and the value of the denatured species at 450 nm.

```

A1=y1(result,1); %Peak Absorption
A2=y2(41,1); %Absorption at the 41st data point or 380 nm
A3=y3(111,1); %Absorption at the 111st data point or 450 nm
%The variable EC is your extinction coefficient
EC= A1/ ((A2/70000) + (A3/44000));

```

## 5c. Stokes shift and reorganization energy

```
% Code for calculating the reorganization energy using Stokes shift and the spectral moment model
proposed by Jordanides et. al, JPC-B (1999)
% Section 1: Loading absorption and emission data from csv files
% Make sure the absorption and emission are on the same x-axis
load abs_cherry.csv
SamplAbs=abs_cherry;
load em_cherry.csv
SamplEm=em_cherry;

% Section 2: Normalize absorption and emission data
norm_abs=SamplAbs(:,2)/max(SamplAbs(:,2));
norm_em=SamplEm(:,2)/max(SamplEm(:,2))

plot(SamplAbs(:,1),norm_abs);
hold on;
plot(SamplEm(:,1),norm_em);
figure;

% Section 3: Find peak absorption and emission to calculate Stokes shift

peak_abs = find(norm_abs==1);
peak_em = find(norm_em==1);
peak_abs_lambda = SamplAbs(peak_abs,1);
peak_em_lambda = SamplEm(peak_em,1);
Stokes_nm= peak_em_lambda-peak_abs_lambda;
Stokes_energy=1E7*((1/peak_abs_lambda)-(1/peak_em_lambda));
reorg_trad=Stokes_energy/2;

%Section 4: Jacobian Transformation for wavelength to wavenumber scale
len_x_abs= length(SamplAbs);
len_x_em= length(SamplEm);
X1=zeros(len_x_abs,1);
X2=zeros(len_x_em,1);
X1(:,1)= 1E7;
X2(:,1)= 1E7;
X1= X1./SamplAbs(:,1);
X2= X2./SamplEm(:,1);
abs_sq_lam=(SamplAbs(:,1)*SamplAbs(:,1).');
em_sq_lam=(SamplEm(:,1)*SamplEm(:,1).');
abs_cm_scale= norm_abs./abs_sq_lam (:,1);
em_cm_scale= norm_em./em_sq_lam (:,1);
abs_cm_norm = abs_cm_scale (:,1) / max (abs_cm_scale (:,1));
em_cm_norm = em_cm_scale (:,1) / max (em_cm_scale (:,1));
plot (X1,abs_cm_norm);
hold on
plot (X2,em_cm_norm);
figure;

%Section 5: Calculate the 0-0 transition energy
cross= (1E7*(1/peak_abs_lambda))-reorg_trad;
X1_shift=X1-cross;
X2_shift= X2-cross;
plot (X1_shift,abs_cm_norm);
hold on
plot (X2_shift,em_cm_norm);
figure

%Section 6: Calculate the 0-0 transition energy

%Flipping the emission matrix to match data structure
em2=flipud(em_cm_norm);
x_em2=flipud(X2_shift);
plot(x_em2,em2);
diff=abs_cm_norm-em2;
sum=abs_cm_norm+em2;
plot (X1_shift,diff);
hold on
plot (X1_shift,sum);
figure
```

```
%Numerical integration

num= trapz(diff*cross);

denom= trapz (sum);
reorg_jordanides= num/denom;
```

## Section 6. Code to simulate role of electrostatics on oscillator strength

```
% Energy and wavefunction of a particle in a box under a constant electric field boundary on both
sides to estimate the contribution of a change in electric field on the extinction coefficient.
clc; clear;

m=9.10938356e-31; % mass of electron in kg
q=1.60217662e-19; % charge of electron in Coulomb
l=12e-10; % length of the box --> length of the chromophore, ~ 12 Angstrom
F=-1e10; % electric field in V/m
h= 6.62607004e-34; % Planck's constant in m^2kg/s
hbar=(h/(2*pi));

El_inf=(hbar^2/(2*m))*(pi/l)^2;
v_L=(q*F*l/El_inf);

syms D(epsilon)
D(epsilon)=(airy(-(pi/v_L)^(2/3))*epsilon)*airy(2,((pi/v_L)^(2/3))*(v_L-epsilon))-(airy(2,-
((pi/v_L)^(2/3))*epsilon)*airy((pi/v_L)^(2/3))*(v_L-epsilon));

epsilon=double(vpasolve(D,epsilon,-21))

%A=[13.984445 24.450580528 33.03615085559 40.9371256]; % solutions for epsilon
%epsilon=A(4);

zeta_0=-((pi/v_L)^(2/3))*epsilon;
%zeta_1= ((pi/v_L)^(2/3))*(v_L-epsilon);

ratio = -airy(zeta_0)/airy(2,zeta_0); % ratio = b/a; psi= a*Ai(zeta)+b*Bi(zeta)

% Normalization

x=0:1e-11:1;

%f(x)=(airy(((pi/v_L)^(2/3))*((v_L*x/l)-epsilon))+(ratio*airy(2,((pi/v_L)^(2/3))*((v_L*x/l)-
epsilon))))).^2;

fun = @(x) (airy(((pi/v_L)^(2/3))*((v_L*x/l)-epsilon))+(ratio*airy(2,((pi/v_L)^(2/3))*((v_L*x/l)-
epsilon))))).^2;

a=sqrt(1/integral(fun,0,1));
b=ratio*a;

psi=a*airy(((pi/v_L)^(2/3))*((v_L*x/l)-epsilon))+(b*airy(2,((pi/v_L)^(2/3))*((v_L*x/l)-
epsilon)));

A= psi*psi';
plot(x,abs(psi.^2))
hold on;
```

## Section 7. Codes for analyzing photobleaching data

### 7a. Using Cellprofiler V2.2.0

- **Pipeline 1:** Select objects to bleach

### Input modules

**Images:** Use only the first image of the experiment for analysis

### Analysis modules

**Crop:** Set the coordinates and radius in pixels

**IdentifyPrimaryObjects:** Set the "Typical diameter of objects"

### Output

**View output settings:** Select output folder

- **Pipeline 2:** Export intensity timelapse for each selected object

### Input modules

**Images:** Use all images including the one obtained from SelectBleachObjects

**NamesAndTypes:** Select "Single image location" that contains the image  
obtained from SelectBleachObjects

### Analysis modules

**ExportToSpreadsheet:**

1. Select Sub-folder
2. Set "Filename prefix"

After obtaining the results, open the "Filenam\_FirstFrame.csv", copy the data from column "Intensity\_MeanIntensity\_BleachingFrames" (Column K on the csv output file) without the title cell and paste in a newly created .mat file in MATLAB. Then analyze with "TimeLapseExp\_PB.m" in MATLAB.

## 7b. Using Matlab code “TimeLapseExp\_PB.m”

```
% This program splits a column into multiple columns and transpose the  
% results.  
  
%%
```

```

% This section creates "Time" column.
% 1. Create the variable column "Time".
% 2. Choose the proper codes for various number of time intervals in the experiment.
% 3. The formula of Time for 2 time intervals = 0.25s*12(frames) + 1s*897(frames).
% 4. The formula of Time for 3 time intervals = 0.25s*12(frames) + 1s*897(frames) +
15s*60(frames).

dt1 = 0.25; % 1st time interval, sec.
dt2 = 1; % 2nd time interval, sec.
nFdt1 = 12; % Number of frames with dt1 (including the first frame, t=0).
nFdt2 = 897; % Number of frames with dt2.

%----- For 2 time intervals -----
nFrames = nFdt1+nFdt2;
Time = zeros(nFrames,1);
for iit1 = 1:nFdt1
    Time(iit1) = Time(iit1)+dt1*(iit1-1);
end
for iit2 = nFdt1+1:nFrames
    Time(iit2) = Time(nFdt1)+dt2*(iit2-nFdt1);
end

%----- For 3 time intervals -----
% dt3 = 15; % 3rd time interval, sec.
% nFdt3 = 60; % Number of frames with dt3.
% nFrames2 = nFdt1+nFdt2;
% nFrames = nFdt1+nFdt2+nFdt3;
% Time = zeros(nFrames,1);
% for iit1 = 1:nFdt1
%     Time(iit1) = Time(iit1)+dt1*(iit1-1);
% end
% for iit2 = nFdt1+1:nFrames2
%     Time(iit2) = Time(nFdt1)+dt2*(iit2-nFdt1);
% end
% for iit3 = nFrames2+1:nFrames
%     Time(iit3) = Time(nFrames2)+dt3*(iit3-nFrames2);
% end

%%
% This section converts MeanIntensity back to 16 bit.
% 1. Set the variable name of measured mean intensity to "MeanIntensity" and save as
"FileName.mat".
% 2. Run the above section to create nFrames or manually input nFrames.

%load('mCh03.mat')
%nFrames = 720;
Background = 240;
% Convert back to 16 bit image intensity.
TimeLapse = transpose(reshape(MeanIntensity,[],nFrames))*2^16-Background;

%%
% This section normalizes the 16 bit MeanIntensity.
% !!! NOTE: need to subtract the background before normalize the signal (NormTlapse)!!!

nObjects = size(TimeLapse,2); % size(object, dimension); (dim1,dim2)=(row,col).
NormTlapse = zeros(nFrames,nObjects);
for ii = 1:nObjects
    NormTlapse(:,ii) = TimeLapse(:,ii)/norm(TimeLapse(:,ii),Inf);
end

%%
% This section finds t1/2 for single measurement on one FP with continuous excitation.
% !!! NOTE: need to create "Time" column!!!

% 1. [C,I] = min(____) finds the indices of the smallest elements (C), and returns them (indices)
in output vector I.
%     If there are several identical smallest values, this syntax returns the index of the first
smallest element that it finds.
% 2. Execute this section with the line containing "NormTlapseSelected" after
removing bad data points.

```

```

%----- Choose the proper [dNormT, idxPBhalf] and nObjectsFinal. -----
[dNormT, idxPBhalf] = min(abs(NormTlapse-0.5));
nObjectsFinal = size(NormTlapse,2);
% === NormTlapseSelected: NormTlapse with bad data removed ===
% [dNormT, idxPBhalf] = min(abs(NormTlapseSelected-0.5));
% nObjectsFinal = size(NormTlapseSelected,2);

% TimeHalf = Time(idxPBhalf); % Just to check...
t_half = mean(Time(idxPBhalf));
t_SD = std(Time(idxPBhalf));

%%
% Plot NormTlapse vs Time.

plot(Time, NormTlapse(:,1:nObjectsFinal))
% plot(Time, NormTlapseSelected(:,1:nObjectsFinal))

%%
% This section finds t1/2 for All measurements (combined final results) on one FP with
continuous excitation.

% [dNormTime, idxPBhalfT] = min(abs(NormTlapseAll-0.5));
% nObjectsAll = size(NormTlapseAll,2);
% Time_half = mean(Time(idxPBhalfT));
% Time_SD = std(Time(idxPBhalfT));
% plot(Time, NormTlapseAll(:,1:nObjectsAll))

```

## Section 8. Codes to simulate dark state kinetics in time and frequency domains using numerical solutions to the three-state model

- **Part 1: Define function body for the rate equations**

```

%Three state: Ground, dark and excited state only
%This function allows the user to run the dark state code
% This version allows user inputs for DSC/GSR times in the main code
function xprime = dsc_gsr_10(t,x,tsld,tgsr)
I=10E3; %laser intensity in W/cm*cm

%Define the constants
h=6.626E-34; %Planck's constant J/s
c=2.998E8; %Speed of light m/s
na=6.022E23; %Avogadro's number
lambda=561E-9; %wavelength of light m
epsilon=100*100; %Molar extinction coeff in 1/(Mcm)
sigma=2.303*epsilon*1000/(na); %Absorption cross section cm*cm

%Define the modulation frequency of light if using modulation
%Comment out if using CW
nu=5E3;

% Defining time constants
% Uncomment the DSC/GSR time if using other modalities
tslg=2.00E-9; % Excited State Lifetime
%tgsr=500E-6; % GSR Time = User defined
%tsld=50E-6; % DSC Time = User defined
tds1=500E-6; % Reverse DSC = 500 microseconds

% Conversion to rate constants
kslg=1/tslg;

```



```

kgsr=1/tgsr;
ksld=1/tsld;
kds1=1/tds1;
kb=1000;           % Bleaching times in order of a few seconds

kge0=I*sigma*lambda/(h*c);

% Using a sin-modulated excitation source
% This can be changed to any waveform of interest
% Modulation not needed if a CW source is used
kgs1=0.5*kge0*(sin(2*pi*nu*t)+1);

% The rate equations involved in the three state model
xprime=[-kgs1*x(1)+kslg*x(2)+kgsr*x(3); kgs1*x(1) - (kslg+ksld)*x(2)+kds1*x(3); ksld*x(2)-
(kds1+kgsr)*x(3)];

```

- Part 2: Main code for extracting state populations in the time-domain and frequency shifts**

```

% Main body of the DSC/GSR code
tic; cla; clc; clear;

% Please define the range of the DSC/GSR interested to simulate
phi=zeros(100,100); % Phase difference matrix
dsc=zeros(100,1);   % Frequency matrix
tsld=0;              % Initial frequency
gsr=zeros(100,1);   % Lifetime matrix
nu=5E3;

for i=1:1:100;
dsc(i,1)=tsld+2E-6;
tsld=dsc(i,1)
tgsr=0;
for j=1:1:100;
    a=1/(nu*1E4);           %time step size
    b=2*pi/(nu);            %total time
    tot_steps= b/a;         %total number of steps
    t=0:a:b;                %time span for 10 full cycles
    gsr(j,1)=tgsr+20E-6;
    tgsr=gsr(j,1)
    tspan=t;

    x0=[1 0 0];
    [t,x]=ode45(@dsc_gsr_10,tspan,x0,[],tsld,tgsr);
    xdash=(sin(2*pi*nu*t));
    ydash=(cos(2*pi*nu*t));
    Xem=xdash.*x(:,2);
    Yem=ydash.*x(:,2);
    Xex=xdash.*(sin(2*pi*nu*t)+1);
    Yex=ydash.*(sin(2*pi*nu*t)+1);

    phi(i,j)=atand(Yem/Xem)-atand(Yex/Xex);
end
end
save('phi')
save('nul')
save('tau')
toc

```

## Section 9. Codes to simulate dark state kinetics in time and frequency domains using analytical solutions to the three-state model

```

%(*This program is to simulate the population based on 3-state model, including ground state (S0),
excited state (S1) and dark state (D), but not considering reverse DSC*). Note this code is a
Mathematica notebook (*.nb)
ClearAll;
Eigensystem[{{
  {-kEx, kEm+kIC, kGSR},
  {kEx, -(kEm+kIC+kDSC), 0},
  {0, kDSC, -kGSR}
}}];
lambda1=0;
lambda2=0.5 (-kDSC-kEm-kEx-kGSR-kIC-\[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm
kGSR+kEx kGSR+kGSR kIC)));
lambda3=0.5 (-kDSC-kEm-kEx-kGSR-kIC+\[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm
kGSR+kEx kGSR+kGSR kIC)));
v1={ (kGSR (kDSC+kEm+kIC))/(kDSC kEx), kGSR/kDSC, 1};
v2={-(1/(2 kDSC)) (kDSC-kEm-kEx+kGSR-kIC-\[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC
kGSR+kEm kGSR+kEx kGSR+kGSR kIC))), -(1/(2 kDSC)) (kDSC+kEm+kEx-
kGSR+kIC+\[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm kGSR+kEx kGSR+kGSR kIC))), 1};
v3={-(1/(2 kDSC)) (kDSC-kEm-kEx+kGSR-kIC+\[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC
kGSR+kEm kGSR+kEx kGSR+kGSR kIC))), -(1/(2 kDSC)) (kDSC+kEm+kEx-kGSR+kIC-
\[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm kGSR+kEx kGSR+kGSR kIC))), 1};
%(* n[t_]={nS0[t], nS1[t], nD[t]}; *)
nC[c1_, c2_, c3_] := c1*v1 + c2*v2 + c3*v3; (* t=0 *)
nBC={1, 0, 0}; (* Initial condition *)
Solve[nC[c1, c2, c3] == nBC, {c1, c2, c3}];
c1=-(\[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm kGSR+kEx kGSR+kGSR kIC)))/(kDSC
(-1/kDSC(\[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm kGSR+kEx kGSR+kGSR kIC))) -
1/kDSC2kGSR \[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm kGSR+kEx kGSR+kGSR kIC))) -
1/(kDSC kEx)kGSR \[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm kGSR+kEx kGSR+kGSR
kIC)))-1/(kDSC2 kEx)kEm kGSR \[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm kGSR+kEx
kGSR+kGSR kIC))-1/(kDSC2 kEx)kGSR kIC \[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm
kGSR+kEx kGSR+kGSR kIC)));
c2=-(kDSC kEx (-kDSC-kEm-kEx-kGSR-kIC+\[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm
kGSR+kEx kGSR+kGSR kIC)))/(2 (kDSC kEx+kDSC kGSR+kEm kGSR+kEx kGSR+kGSR kIC)
\[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm kGSR+kEx kGSR+kGSR kIC))));
c3=-(kDSC kEx (kDSC+kEm+kEx+kGSR+kIC+\[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm
kGSR+kEx kGSR+kGSR kIC)))/(2 (kDSC kEx+kDSC kGSR+kEm kGSR+kEx kGSR+kGSR kIC)
\[Sqrt]((kDSC+kEm+kEx+kGSR+kIC)^2-4 (kDSC kEx+kDSC kGSR+kEm kGSR+kEx kGSR+kGSR kIC))));
nS0[t_]:=c1 Exp[lambda1 t] v1[[1]]+c2 Exp[lambda2 t] v2[[1]]+c3 Exp[lambda3 t] v3[[1]];
nD[t_]:=c1 Exp[lambda1 t] v1[[3]]+c2 Exp[lambda2 t] v2[[3]]+c3 Exp[lambda3 t] v3[[3]];
nS1[t_]:= c1 Exp[lambda1 t] v1[[2]]+c2 Exp[lambda2 t] v2[[2]]+c3 Exp[lambda3 t] v3[[2]];
%(*Constants*)
h = 6.626*10^-34;
cLight=2.998*10^8;
NA=6.022*10^23;
%(*Set parameters*)
lambda=561*10^-9;
%(* Excitation wavelength in m *)
I0=24*10^4;
%(* Laser Intensity in W/m^2 (W/cm^2=1e4 W/m^2) *)
QY=0.41;
%(* Quantum yield *)
tau=2.45*10^-9;
%(* Excited state lifetime in sec *)
epsilon=72.9*10^3;
%(* Molar extinction coefficient in M^-1cm^-1 at lambda *)
sigma=2.303*epsilon*1*10^3*1*10^-4/NA;
%(* Absorption cross-section in m^2 *)
kEx=(I0*sigma*lambda)/(h*cLight);
%(* Excitation rate, in Hz *)
kEm=QY/tau;
%(* Radiative emission rate constant, in Hz *)
kDSC=1/(24*10^-6);
%(* DSC rate constant, in Hz *)
kGSR=1/(20*10^-3);
%(* GSR rate constant, in Hz *)

```

```

kIC=(1/tau)-kEm;
%(* Non radiative decay rate constant in Hz *)
p1=Plot[nS0[t],{t,0,5}, PlotStyle->Directive[RGBColor[0,1,0.],AbsoluteThickness[2.]], PlotRange->
{0.7,1.0}]
pD1=Plot[nD[t],{t,0,5}, PlotStyle->Directive[RGBColor[0,1,0.],AbsoluteThickness[2.]], PlotRange->
{0,0.3}]
p1S1 = Plot[nS1[t],{t,0,5}, PlotStyle->Directive[RGBColor[0,0,1.],AbsoluteThickness[2.]],PlotRange-> {0,0.0001}]
%(*Constants*)
h = 6.626*10^-34;
cLight=2.998*10^8;
NA=6.022*10^23;
%(*Set parameters*)
lambda=561*10^-9;
%(* Excitation wavelength in m *)
I0=24*10^4;
%(* Laser Intensity in W/m^2 (W/cm^2=1e4 W/m^2) *)
QY=0.41;
%(* Quantum yield *)
tau=2.40*10^-9;
%(* Excited state lifetime in sec *)
epsilon=72.9*10^3;
%(* Molar extinction coefficient in M^-1cm^-1 at lambda *)
sigma=2.303*epsilon*1*10^3*1*10^-4/NA;
%(* Absorption cross-section in m^2 *)
kEx=(I0*sigma*lambda)/(h*cLight);
%(* Excitation rate,in Hz *)
kEm=QY/tau;
%(* Radiative emission rate constant,in Hz *)
kDSC=1/(18*10^-6);
%(* DSC rate constant,in Hz *)
kGSR=1/(20*10^-3);
%(* GSR rate constant,in Hz *)
kIC=(1/tau)-kEm;
%(* Non radiative decay rate constant in Hz *)
p2=Plot[nS0[t],{t,0,5}, PlotStyle->Directive[RGBColor[0,0,1.],AbsoluteThickness[2.]], PlotRange->
{0.7,1.0}]
pD2=Plot[nD[t],{t,0,5}, PlotStyle->Directive[RGBColor[0,0,1.],AbsoluteThickness[2.]], PlotRange->
{0,0.3}]
p2S1 = Plot[nS1[t],{t,0,5}, PlotStyle->Directive[RGBColor[0,0,1.],AbsoluteThickness[2.]],PlotRange-> {0,0.0001}]
%(*Constants*)
h = 6.626*10^-34;
cLight=2.998*10^8;
NA=6.022*10^23;
%(*Set parameters*)
lambda=561*10^-9;
%(* Excitation wavelength in m *)
I0=4.4*10^4;
%(* Laser Intensity in W/m^2 (W/cm^2=1e4 W/m^2) *)
QY=0.20;
%(* Quantum yield *)
tau=2.01*10^-9;
%(* Excited state lifetime in sec *)
epsilon=70.7*10^3;
%(* Molar extinction coefficient in M^-1cm^-1 at lambda *)
sigma=2.303*epsilon*1*10^3*1*10^-4/NA;
%(* Absorption cross-section in m^2 *)
kEx=(I0*sigma*lambda)/(h*cLight);
%(* Excitation rate,in Hz *)
kEm=QY/tau;
%(* Radiative emission rate constant,in Hz *)
kDSC=1/(38.12*10^-6);
%(* DSC rate constant,in Hz *)
kGSR=1/(1.9678);
%(* GSR rate constant,in Hz *)
kIC=(1/tau)-kEm;
%(* Non radiative decay rate constant in Hz *)
p3=Plot[nS0[t],{t,0,10}, PlotStyle->Directive[RGBColor[1.,0.,0.],AbsoluteThickness[2.]],
BaseStyle->{FontSize->18},PlotRange-> {0.70,1.0},Frame->True,FrameLabel->{"Time (s)","Normalized
population"}]

```

```

pD3=Plot[nD[t],{t,0,10}, PlotStyle->Directive[RGBColor[0.,0.,1.],AbsoluteThickness[2.]],
BaseStyle->{FontSize->16},PlotRange-> {0,0.3},Frame->True,FrameLabel->{"Time (s)","Normalized
population"}]
p3S1= Plot[nS1[t],{t,0,10}, PlotStyle->Directive[RGBColor[0.,1.,0.],AbsoluteThickness[2.]],
BaseStyle->{FontSize->16},PlotRange-> {0.000004,0.000008},Frame->True,FrameLabel->{"Time
(s)","Normalized population"}]

Show[p1,p2,p3, PlotRange-> {0.70,1.0}]
Show[pD1,pD2,pD3, PlotRange-> {0,0.3}]
Show[p3,pD3, p3S1, PlotRange->{0,1}]

```

## Section 10. Simulation of single molecule kinetics using Monte-Carlo methods

```

%Single Molecule Behavior Simulator based on Monte-Carlo Simulations on low irradiance trends
collected on a CCD/CMOS camera with acquisition times much slower than radiative and non-radiative
decay. The unit time (or time step) in this simulation is longer than fluorescence lifetime, but
shorter than dark-state lifetime. Therefore, it can be assumed that the molecule must stay or
return to S0 or D in each unit time cycle.
clc; clear; cla;
% % Number of trajectories
nSim = 500;
sumton=zeros(nSim);
sumtoff=zeros(nSim);
savetraj=zeros(500,nSim);
for nnn=1:nSim

% Ensemble parameters for the FP of interest: Tip Use SI
h = 6.626*10^-34; c_light = 2.998*10^8; NA = 6.022E23; % constants
lambda = 561*10^-9; % Excitation wavelength in m
I0 = 1.24e4; % Laser Intensity in W/m^2 (W/cm^2 = 1e4 W/m^2)
epsilon = 71E3; % Molar extinction coefficient in M^-1cm^-1 at lambda;
sigma = 2.303*epsilon*1E3*1E-4/NA; % Absorption cross-section in m^2
kEx = (I0*sigma*lambda)/(h*c_light); % Excitation rate, in Hz
QY = 0.22; % Quantum yield
tau = 1.78e-9; % Excited state lifetime in sec
tGSR=3.76; % D to S0 time in seconds
tDSC=1/(171509); % S1 to D time in Seconds
kEm = QY/tau; % Radiative emission rate constant, in Hz
kDSC = 1/tDSC; % DSC rate constant, in Hz
kGSR = 1/tGSR; % GSR rate constant, in Hz
kIC = (1/tau)-kEm; % Non-radiative decay rate constant, in Hz
%Time step data
unitT = 1e-6; % Unit time of the simulation in sec now I am considering each time unit is 10us
NunitT = 1.5e8; % Total number of timesteps you are running the simulation for

% At each time step
Sig = zeros(NunitT,1); % Possible fluorescence at each time point [Sig(x,1)=0 (off) or 1 (on)]
S = zeros(NunitT,1); % Possible electronic state, 0=S0, 1=S1, 2=D, where D is dark state
TimeAxis = zeros(NunitT,1);
Sig(1) = 0; % Initial fluorescence signal
S(1) = 0; % Initial electronic state
OnCount = 0; % ON state counts

% Quick check
kON_En = kEx*kDSC/(kEm+kIC+kDSC);
tauON_En = 1/kON_En;
kON_SM = kEx*kIC*kDSC/(kEm+kIC+kDSC);
tauON_SM = 1/kON_SM;
% Parameters below are for ON/OFF time caculation

%Defining your experimental detection settings on the camera
time = unitT*NunitT; % Total time of the experiment

```

```

nFrame = 500; % Number of frames your camera will detect
tExpo = time/nFrame; % Total time/total frames will be the exposure time per frame
Ndt = round(tExpo/unitT); % Number of unitT in one tExpo [This is the number of events hidden in
each frame]

%Defining variables that will be needed for the MC simulations
tExpoAxis = zeros(nFrame,1); % Time axis with respect to the frame rate
PhotonAcc = zeros(nFrame,1); % Accumulated ON/OFF counts in each frame
Status = zeros(nFrame,1); % ON/OFF status, 0=OFF, 1=ON
Switching = 1; % Just give a number
OnTime = zeros(length(Switching),1); % Continuous ON time
OffTime = zeros(length(Switching),1); % Continuous OFF time
OnAcc = zeros(length(Switching),1); % Accumulated ON counts
OffAcc = zeros(length(Switching),1); % Accumulated OFF counts
SigExpo = 0; % Total Sig in one tExpo
mm = 1; % Initial index of OnTime
nn = 1; % Initial index of OffTime

mCount = 0; % mCount default value = 1, so mCount+1 is to prepare for next consecutive ON event
nCount = 0; % nCount default value = 1, so nCount+1 is to prepare for next consecutive OFF event
Sw = 0; % Number of ON/OFF switching events
NoSw = 0; % Number of No switching events

% Define probability ProbS is the probability of the molecule leaving state S. Prob_StoF is the
probability of the molecule leaving from S to state F, emProb is the probability of radiative
decay from S1 to S0. First order rate kinetics defining the probability to exist in a state at
a time t
ProbS0 = 1-exp(-kEx*unitT);
ProbS1 = 1-exp(-(kEm+kIC+kDSC)*unitT);
ProbD = 1- exp(-kGSR*unitT);
S1_Exit_Rate=kEm+kIC+kDSC; %Exit pathway from the S1 state

% Exit rates from each state define the movement of the system from a
% particular state
Prob_S0toS1 = 1; % Assume every time a photon comes in you'll excite to S1
Prob_S1toS0 = (kEm+kIC)/S1_Exit_Rate; % QY of recovery to GS [Should be ~ 1 - major pathway]
Prob_S1toD = kDSC/S1_Exit_Rate; % QY of DSC - [This number is tiny- minor pathway]
emProb = kEm/S1_Exit_Rate; % QY of Fluorescence [Part of the major pathway that is fluorescence]

% Monte Carlo simulation on single molecule electronic state. You run the iii loop for the number
of steps you initially defined, so your catch details for each timestep (ms,us or ns)
%Remember Sig is the fluorescence signal at time point
%Remember S is the electronic state at the time point
for iii=1:NunitT-1; %Time Axis is a point in time based on your stepsize

    TimeAxis(iii)=(iii-1)*unitT;

    %S is your initial electronic state
    if S(iii)==0; %If in the ground state

        if rand>ProbS0;
            Sig(iii+1)=0;
            S(iii+1)=0;
        else
            if rand>Prob_S1toS0;
                Sig(iii+1)=0;
                S(iii+1)=2;
            else
                S(iii+1)=0;
                if rand>emProb;
                    Sig(iii+1)=0;
                else
                    Sig(iii+1)=1;
                    OnCount=OnCount+1;
                end
            end
        end
    end
end

```

```

        end
    end
end

if S(iii)==2;
    if rand>ProbD;
        Sig(iii+1)=0;
        S(iii+1)=2;
    else
        Sig(iii+1)=0;
        S(iii+1)=0;
    end
end
end

end
TimeAxis(NunitT) = (NunitT-1)*unitT;
% save('FR_SigTest', 'Sig');

% Calculate ON/OFF time
for ii=1:nFrame;
    tExpoAxis(ii) = ii*tExpo;
    SigExpo = 0;
    for jj=1:Ndt; % Signal counts in frame ii
        kk = (ii-1)*Ndt+jj;
        SigExpo = SigExpo+Sig(kk);
    end
    PhotonAcc(ii) = SigExpo; % Accumulated ON/OFF counts in each frame

    if SigExpo>0;
        Status(ii) = 1;
    elseif SigExpo==0;
        Status(ii) = 0;
    end

    if ii==1;
        if Status(ii)==1;
            mCount = mCount+1;
            OnTime(mm) = mCount*tExpo;
            OnAcc(mm) = mCount;
        elseif Status(ii)==0;
            nCount = nCount+1;
            OffTime(nn) = nCount*tExpo;
            OffAcc(nn) = 0;
        end
    else % ii>1;
        if Status(ii)==Status(ii-1); % No ON/OFF switching
            if Status(ii)==1; % Continue ON
                mCount = mCount+1;
                OnTime(mm) = mCount*tExpo;
                OnAcc(mm) = mCount;
            elseif Status(ii)==0; % Continue OFF
                nCount = nCount+1;
                OffTime(nn) = nCount*tExpo;
                OffAcc(nn) = 0;
            end
            NoSw = NoSw+1;
        else % Switching ON/OFF
            if Status(ii)==1; % OFF to ON
                nn = nn+1;
                nCount = 0;
                mCount = mCount+1;
                OnTime(mm) = mCount*tExpo;
                OnAcc(mm) = mCount;
            elseif Status(ii)==0; % ON to OFF
                mm = mm+1;
                mCount = 0;
                nCount = nCount+1;
                OffTime(nn) = nCount*tExpo;
                OffAcc(nn) = 0;
            end
        end
    end
end

```

```

        end
        Sw = Sw+1;
    end
end
end
TotOnTime = sum(OnTime);
TotOffTime = sum(OffTime);
nOnTime = length(OnTime);
nOffTime = length(OffTime);
% This section is to plot & fit ON/OFF time histogram, obtain the statistical (average) ON/OFF
time, & histogram of photon counts

nBinsOn = 50;
nBinsOff = 50;
nBinsPhoton = 50;
% For ON time
figure(1);
HistgOn = histogram(OnTime,nBinsOn); hold on % Plot histogram
% HistgOn = histogram(OnTime,'BinWidth',0.09) % Alternative option
[HcountOn,edgesOn] = histcounts(OnTime,nBinsOn); % Get counts in each bin
FitOn = fitdist(OnTime,'Exponential');
tOn = FitOn.mu;
AreaOn = sum(HcountOn*HistgOn.BinWidth);
% Plot fit curve
XvalsOn = 0:max(OnTime)/nBinsOn:max(OnTime);
FitCurveOn = AreaOn*exp(-XvalsOn/tOn)/tOn; % Compute fit curve. FitOn.mu is to get fit parameter
mu.
plot(XvalsOn,FitCurveOn); hold off
% Just to check...
% hold on
% histfit(OnTime,nBinsOn,'Exponential'); hold off
% For OFF time
figure(2);
HistgOff = histogram(OffTime,nBinsOff); hold on % Plot histogram
[HcountOff,edgesOff] = histcounts(OffTime,nBinsOff); % Get counts in each bin
% HistgOff = histogram(OffTime,'BinWidth',0.09) % Alternative option
FitOff = fitdist(OffTime,'Exponential');
tOff = FitOff.mu;
AreaOff = sum(HcountOff*HistgOff.BinWidth);
% Plot fit curve
XvalsOff = 0:max(OffTime)/nBinsOff:max(OffTime);
FitCurveOff = AreaOff*exp(-XvalsOff/tOff)/tOff; % Compute fit curve. FitOn.mu is to get fit
parameter mu.
plot(XvalsOff,FitCurveOff); hold off
% Just to check...
% hold on
% histfit(OffTime,nBinsOff,'Exponential'); hold off
% For Photon counts
figure(3);
HistgPhoton = histogram(PhotonAcc,nBinsPhoton); hold on
[HcountPhoton,edgesPhoton] = histcounts(PhotonAcc,nBinsPhoton);
hold off
%save('FR_50ms3','OnTime','OffTime','nOnTime','nOffTime','tOn','tOff','HcountPhoton')

% This section is to plot single molecule ON/OFF trace
figure(4);
stairs(tExpoAxis,PhotonAcc)% Actual ON/OFF trace observed in the experiment
savetraj(:,nnn)=PhotonAcc;
sumton(:,nnn)=tOn;
sumtoff(:,nnn)=tOff;
end
save('trajectories.txt','savetraj','-ASCII','-append')

```

## Section 11. Detection code for single molecule trajectories from video (TIFF series) file

```
(* This code uses a TIF/TIFF image series, corrects background, selects
single emitters and saves trajectories as csv files. This is a Mathematica
notebook. *)
Data analysis
In[173]:=
(* Data import *)
pics=Import["C:\\Users\\Single_Mol_061420\\8.0_02\\*.tif"];
Length[pics]
pic0=Reverse[ImageData[pics[[1]]]];
Dimensions[pic0]
ListDensityPlot[pic0,ColorFunction->GrayLevel,ImageSize->600,InterpolationOrder->0]
(* Raw data conversion using import above *)
pixelBorderCutoff=1;
(* If your heat map image has a "dark line" on one of the borders, use this
to cut off pixels around the border*)
rawData=Table[Reverse[ImageData[pics[[i]]]],{i,1,Length[pics]}];
rawData=rawData[[All,1+pixelBorderCutoff;;-1-pixelBorderCutoff,1+pixelBorderCutoff;;-1-pixelBorderCutoff]];
rawData=rawData-Min[rawData];
rawData=rawData+1/10 Mean[Flatten[rawData]];
rawData=rawData/Max[rawData];
additive=Total[rawData];
additiveGaussian=Total[rawData[[Round[Length[rawData]/2];;-1]]];
Dimensions[additive]
ListDensityPlot[additive,ColorFunction->GrayLevel,InterpolationOrder->0,PlotLegends->All]
ListPlot3D[additive,PlotRange->All]
(* Plotting raw data for background Gaussian fit *)
ydata=Table[{n,Mean[additiveGaussian[[n,All]]]}, {n,1,Dimensions[additiveGaussian][[1]]}];
FindFit[ydata,a E^(-(y-[Mu])^2/(2[Sigma]^2))+b y,{a,[Mu],[Sigma],b},y];
yparam={a,[Mu],[Sigma],b}/.%;
ygauss=Table[{y,yparam[[1]] E^(-(y-yparam[[2]])^2/(2yparam[[3]]^2))+yparam[[4]]y},{y,1,Dimensions[additiveGaussian][[1]]}];
xdata=Table[{n,Mean[additiveGaussian[[All,n]]]}, {n,1,Dimensions[additiveGaussian][[2]]}];
FindFit[xdata,a E^(-(x-[Mu])^2/(2[Sigma]^2))+b x,{a,[Mu],[Sigma],b},x];
xparam={a,[Mu],[Sigma],b}/.%;
xgauss=Table[{x,xparam[[1]] E^(-(x-xparam[[2]])^2/(2xparam[[3]]^2))+xparam[[4]]x},{x,1,Dimensions[additiveGaussian][[2]]}];
Print["x-fit: a E^(-(x-[Mu])^2/(2[Sigma]^2))+b x","\na = ",xparam[[1]],"\t b = ", xparam[[4]],"\n\ [Mu] = ",xparam[[2]],"\t \ [Sigma] = ",xparam[[3]]];
ListPlot[{xdata,xgauss}]
Print["y-fit: a E^(-(y-[Mu])^2/(2[Sigma]^2))+b y","\na = ",yparam[[1]],"\t b = ", yparam[[4]],"\n\ [Mu] = ",yparam[[2]],"\t \ [Sigma] = ",yparam[[3]]];
ListPlot[{ydata,ygauss}]
```



```

background=Transpose[Table[1/Sqrt[xparam[[1]]yparam[[1]]] (xparam[[1]]E^(-(x-
xparam[[2]])^2/(2xparam[[3]]^2))+xparam[[4]]x) (yparam[[1]] E^(-(y-
yparam[[2]])^2/(2yparam[[3]]^2))+yparam[[4]]y),{x,1,Dimensions[additiveGaussi
an][[2]]},{y,1,Dimensions[additiveGaussian][[1]]}]];
background=Table[background,{i,1,Length[rawData]}];
ListDensityPlot[background[[1]],ColorFunction->GrayLevel,InterpolationOrder-
>0,PlotLegends->Automatic]
>(* Defining functions and variables *)
calculateDenoisedData[backgroundImage_]:= (
tempDenoised=rawData-backgroundImage;
tempDenoised=tempDenoised-Min[tempDenoised];
tempDenoised=tempDenoised+1/10 Mean[Flatten[tempDenoised]];
tempDenoised=tempDenoised/Max[tempDenoised];
Return[tempDenoised];
);
picGeometricAvg[pic_]:=Times@@(Times@@pic)^(1/Length[Flatten[pic]]);
picArithmeticAvg[pic_]:=Total[Total[pic]]/Length[Flatten[pic]];
spectralFlatness[pic_]:=picGeometricAvg[pic]/picArithmeticAvg[pic];
backgroundSubtractionResolution=0.05;
backgroundSubtractionStartingFactor=1;
>(* Maximizes flatness by optimizing n *)
\[Epsilon]=backgroundSubtractionStartingFactor;
previousSpectralFlatness=spectralFlatness[additive];
flatnessVariation={};
flatnessVariationPics={};
deNoised=calculateDenoisedData[\[Epsilon]/Length[rawData]*background];
deNoisedAdd=Total[deNoised];
deNoisedSpectralFlatness=spectralFlatness[deNoisedAdd];
AppendTo[flatnessVariation,{\[Epsilon],deNoisedSpectralFlatness}];
While[deNoisedSpectralFlatness>previousSpectralFlatness,
previousSpectralFlatness=deNoisedSpectralFlatness;
AppendTo[flatnessVariationPics,deNoisedAdd];
\[Epsilon]=\[Epsilon]+backgroundSubtractionResolution;
deNoised=calculateDenoisedData[\[Epsilon]/Length[rawData]*background];
deNoisedAdd=Total[deNoised];
deNoisedSpectralFlatness=spectralFlatness[deNoisedAdd];
AppendTo[flatnessVariation,{\[Epsilon],deNoisedSpectralFlatness}];
];
BSMF=\[Epsilon]=\[Epsilon]-backgroundSubtractionResolution;
deNoisedAdd=flatnessVariationPics[[-1]];
ListPlot[flatnessVariation]
ListDensityPlot[deNoisedAdd,PlotRange->All,PlotLegends-
>Automatic,ColorFunction->GrayLevel,InterpolationOrder->0]
>(* Time-dependence *)
tempAnalysisArray=deNoised;
timeAvgBrightness=Table[Mean[Flatten[tempAnalysisArray[[ii]]]],{ii,Length[deN
oised]}];
timeStDevBrightness=Table[StandardDeviation[Flatten[tempAnalysisArray[[ii]]]]
,{ii,Length[deNoised]}];
ListLinePlot[{timeAvgBrightness,timeStDevBrightness,Table[Max[tempAnalysisArr
ay[[u]]],{u,Length[tempAnalysisArray]}],Table[Min[tempAnalysisArray[[u]]],{u,
Length[tempAnalysisArray]}]},PlotRange->All,PlotLegends-
>{"Mean","StDev","Max","Min"}]
>(* Bi exponential *)
tempAnalysisArray=deNoised;
timeAvgBrightness=Table[Mean[Flatten[tempAnalysisArray[[ii]]]],{ii,Length[deN
oised]}];

```

```

exponentialFit=FindFit[timeAvgBrightness,{a1 E^(- t/[Tau]1)+a2 E^(-
t/[Tau]2)+b,{a1>0, b>0, \[Tau]2>\[Tau]1>1,a2>0}},{a1, \[Tau]1,a2,
\[Tau]2,b},t]
timeParam={a1, \[Tau]1,a2, \[Tau]2,b}/.exponentialFit;
timeFit1=Table[timeParam[[1]] E^(-
t/timeParam[[2]])+timeParam[[5]],{t,1,Length[timeAvgBrightness]}};
timeFit2=Table[timeParam[[3]] E^(-
t/timeParam[[4]])+timeParam[[5]],{t,1,Length[timeAvgBrightness]}};
timeDecayFit=Table[timeParam[[1]] E^(- t/timeParam[[2]])+timeParam[[3]] E^(-
t/timeParam[[4]])+timeParam[[5]],{t,1,Length[timeAvgBrightness]}};
ListLinePlot[{timeAvgBrightness,timeDecayFit,timeFit1,timeFit2},PlotRange-
>All]
avgBackground=Mean[Mean[background[[1]]]];
backgroundTD=Table[(timeFit1[[i]]/avgBackground)background[[i]],{i,Length[bac
kground]}}; (* Change how many exponentials you want here by changing
"timeFit1" array to "timeFit2" or "timeDecayFit" *)
ListLinePlot[{Table[Mean[Mean[backgroundTD[[j]]]],{j,Length[backgroundTD]}],t
imeAvgBrightness}]
deNoisedTD=calculateDenoisedData[backgroundTD];
tempAnalysisArray=deNoisedTD;
timeAvgBrightness=Table[Mean[Flatten[tempAnalysisArray[[ii]]]],{ii,Length[tem
pAnalysisArray]}};
timeStDevBrightness=Table[StandardDeviation[Flatten[tempAnalysisArray[[ii]]]]
,{ii,Length[tempAnalysisArray]}};
ListLinePlot[{timeAvgBrightness,timeStDevBrightness,Table[Max[tempAnalysisArr
ay[[u]]],{u,Length[tempAnalysisArray]}],Table[Min[tempAnalysisArray[[u]]],{u,
Length[tempAnalysisArray]}]},PlotRange->All,PlotLegends-
>{"Mean","StDev","Max","Min"}]
>(* Variables to make protein detection easier to work with *)
desiredNumberOfProteins=500;
analysisArray=deNoisedTD;
blinkTraceSmoothing=1;
maxProteinRadius= 2;
brightnessThreshold=Mean[Flatten[Total[analysisArray]]]+StandardDeviation[Fla
tten[Total[analysisArray]]];
>(* Molecule detection code *)
tempAnalysisArray=Total[analysisArray];
width=maxProteinRadius-1;
proteins={};
(* Keep going until it finds the proper number of traces *)
For[iterator=0,iterator<desiredNumberOfProteins,iterator++,
i=j=1;
>(* Iterate over x to find location of maximum *)
While[i<=Length[tempAnalysisArray]&&iterator<desiredNumberOfProteins,
j=1;
>(* Iterate over y to find location of maximum *)
While[j<=Length[tempAnalysisArray[[i]]]&&iterator<desiredNumberOfProteins,
>(* Once you find maximum value, do this *)
If[tempAnalysisArray[[i,j]]==Max[tempAnalysisArray],
>(* Append location of maximum value and begin scan of (2*width + 1)^2 pixel
grid centered at maximum value *)
AppendTo[proteins,{}];
For[ip=Max[i-width,1],ip<=Min[i+width,Length[tempAnalysisArray]],ip++,
For[jp=Max[j-width,1],jp<=Min[j+width,Length[tempAnalysisArray[[ip]]]],jp++,
If[tempAnalysisArray[[ip,jp]]>brightnessThreshold,
AppendTo[proteins[[Length[proteins]]],{ip,jp}];
];

```

```

tempAnalysisArray[[ip,jp]]=0;
];
];
i=Length[tempAnalysisArray];
j=Length[tempAnalysisArray[[i]]];
];
j++;
];
i++;
];
];
ClearAll[i,j,ip,jp,iterator];
blink={};
For[prot=1,prot<=desiredNumberOfProteins,prot++,
AppendTo[blink,Table[
{t,Mean[
Table[
analysisArray[[t]][[proteins[[prot,i,1]]]][[
proteins[[prot,i,2]]]],
{i,Length[proteins[[prot]]}]]
]]
],{t,1,Length[analysisArray]}]];
];
ClearAll[prot];
graphicsGridSize=Ceiling[Sqrt[desiredNumberOfProteins]];
plots=Table[{},{ii,graphicsGridSize}];
trace=1;
While[trace<=desiredNumberOfProteins,
index=Mod[trace,graphicsGridSize];
If[index==0,index=graphicsGridSize];
AppendTo[plots[[index]],ListLinePlot[Mean[Transpose[Partition[blink[[trace]],
blinkTraceSmoothing]]],PlotRange->All,PlotLabel->ToString[trace]]];
trace++;
];
GraphicsGrid[plots,ImageSize->1200]
Out[174]= 600
Out[176]= {270,270}
Out[177]=
Out[185]= {268,268}
Out[186]=
Out[187]=
During evaluation of In[173]:= x-fit:  $a E^{-((x-\mu)^2/(2\sigma^2))}+b x$ 
a = 23.2441 b = -0.0149449
 $\mu$  = 223.885  $\sigma$  = 187.199
Out[197]=
During evaluation of In[173]:= y-fit:  $a E^{-((y-\mu)^2/(2\sigma^2))}+b y$ 
a = 21.6005 b = -0.0111732
 $\mu$  = 139.553  $\sigma$  = 146.977
Out[199]=
Out[202]=
Out[220]=
Out[221]=
Out[225]= Mean
StDev
Max
Min

```

```

Out[228]= {a1->0.00300059,\[Tau]1->14.0738,a2->0.00697043,\[Tau]2->19.3128,b-
>0.0398134}
Out[233]=
Out[235]=
Out[240]= Mean
          StDev
          Max
          Min
>(* Exports blink traces*)
fileName="TrFRMQNoOx_25_7_7"
exportData=Table[i,{i,1,Length[blink[[1]]]}, {j,1,Length[blink]-1}];
For[ii=1,ii<=Length[blink]-2,ii++,
For[jj=1,jj<=Length[blink[[ii]]],jj++,
exportData[[jj,ii+1]]=blink[[ii,jj,2]];
];
];
Export[fileName<>".csv",exportData];
Out[270]= NewEGFPNoOx_10_02_7_23
>(* Brightness histogram video *)
brightnessHistVid=Table[
Image[
Histogram[
Round[Flatten[tempAnalysisArray[[u]]],0.01],
100,
PlotRange->{{0,1},{0,2750}},
Epilog->Inset[Framed[Grid[{
{"Mean =",Mean[Mean[tempAnalysisArray[[u]]]}],
{"Max =",Max[tempAnalysisArray[[u]]]},
{"Min =",Min[tempAnalysisArray[[u]]]}],
Alignment->{{Left}}],
Background->White],
{Right,Top},
{Right,Top}]],
ImageSize->600],
{u,Length[tempAnalysisArray]}
];
ListAnimate[brightnessHistVid]
>(* Plots all pixel brightnesses in all frames *)
allPixelTrace=Flatten[Table[{ii,tempAnalysisArray[[ii,jj,kk]]},{ii,Length[tem
pAnalysisArray]}, {jj,Length[tempAnalysisArray[[ii]]]}, {kk,Length[tempAnalysis
Array[[ii,jj]]}],2];
DensityHistogram[allPixelTrace,allPixelTrace[[-1,1]],PlotRange-
>{{0,allPixelTrace[[-1,1]]},{0,1}},ColorFunction->"Rainbow"]

```

## Section 12. Code for binarizing single molecule trajectories

```
% Single Molecule Trajectory Binarizer Readme File
% Matlab Live Script Version Info: Matlab R2021 with fit toolbox
% %-----%
%
% File Name: LoopedBin_5.mlx
% Input Data format: time trace tarjectories *.csv format. (col (1): frame/time col (2,:):
fluorescence intensity info)

% %-----%
% Description: This code accepts noisy single-molecule fluorescence time traces and tries to
binarize them based on an input guess "on" value. The code binarizes the on and off time traces,
provides information on the statistics of on and off events, fits the binned on and off data to a
biexponential decay to give a time constant based on a Poissonian distribution.

% %-----%
% # Section 1: Choose upto 8 fluorescence traces to track the frame/time information with
fluorescence intensity
% * Select upto 5 traces with a "guess value" for a single molecule on event.
% * Note this is a guess "on" only - program will go through re-iterative protocols to polish
this value
% * Run this section only.

% # Section 2: Defines selected "On" time and displays it and associated stats

% Set the trace number (upto 5) for the 1D array "ChosenOne" and "BlinksStart" and "BlinksEnd" to
define the guess "on" time values
% * Check on the plot in section 2 to see if the mean and standard deviations reported for the
guess "on" make sense for the trace.
%Example: is the off-event included in standard deviation window of the mean guessed on
% * If not: re-define the guess "on"
% * If yes: Run the rest of the code.

%
% # Section 3: Extracts noise of dataset and fits it to a random distribution.
%
%
% # Section 4: Identifies state changes and binarizes based on them.
%
%
% # Section 5: Binarized trace storage
%
%
% Section 6:
% * Fits on and off times to mono or biexponential fits and reports relevant stats
% * If fits are not converging, starting values can be changed

%Section 1:

RawTraces=E3Negative(:,:); % Picks a preimported dataset. Dataset should have Frame # as the
first column
H=height(RawTraces);
W=width(RawTraces);
RawTraces=table2array(RawTraces);
x=RawTraces(:,1);
Displaytraces = [10,9,8,7,6,5,4,3] %Fill in the #s of 8 traces you would like to look at

for a = 1:8 % Plots 8 traces for viewing. At least 5 on times must be manually identified
subplot(4,2,a)
```

```

    plot(x,RawTraces(:,Displaytraces(a)))
    title(Displaytraces(a))
    ylabel('Fl(a.u)');
    xlabel('Frame')
end

RawTraces=array2table(RawTraces);
traces = figure();
clf

%Section 2: Defines selected "On" time and displays it and associated stats

ChosenOnes = [ 5, 5, 5, 5, 5] % Chosen Blinks - Maintain same order throughout
BlinksStart = [ 261, 261, 262, 264, 265] %Start of on time -Frame Number (Can't be 0)
BlinksEnd = [ 299, 299, 298, 297, 295] % Ending Number
%Try to pick traces with clear on times that aren't significantly higher
%than other blinks (i.e. potential multi fluorophore events)
Mean= zeros(1,5);
SD= zeros(1,5);
LowB= zeros(1,5);
UpB= zeros(1,5);

for a=1:5 %Run stats on selected on times.
    chosen = table2array(RawTraces(:,ChosenOnes(a)));
    blinkduration = linspace(BlinksStart(a),BlinksEnd(a),BlinksEnd(a)-BlinksStart(a)+1);
    Samblink = (0);
    SamBlink=table2array(RawTraces(BlinksStart(a):BlinksEnd(a),ChosenOnes(a)));
    Mean(a)=mean(SamBlink);
    SD(a)=std(SamBlink);
    LowB(a)=Mean(a)-2*SD(a);%Segments above this value are considered "on" during that time
    UpB(a)=Mean(a)+2*SD(a);

%Segments above this value are considered multiprotein and the offending trace removed from
analysis

    Meana= linspace(Mean(a),Mean(a),numel(x));%Makes lines for plotting bounds
    LowBa= linspace(LowB(a),LowB(a),numel(x));
    UpBa= linspace(UpB(a),UpB(a),numel(x));

    subplot(3,2,a)

%The selected on time is shown in red. The lower bound should be high enough to distinguish
between an on and an off event (i.e above background)

    plot(x,chosen,x,LowBa,x,UpBa,x,Meana,blinkduration,SamBlink,'red')
    ylabel('Fl(a.u)');
    xlabel('Frames')
    title(ChosenOnes(a))
end
subplot(3,2,6) %Plots a key with a legend to make sure all other plots are legible
plot(x,chosen,x,LowBa,x,UpBa,x,Meana,blinkduration,SamBlink,'red')
legend('Trajectory','LowerBound','UpperBound','Mean')

title("Key")

MeanAr =Mean %Shows the means of each selected on time. Should all probably be comparable
SDAr = SD
LowBAR = LowB %Boundary values, by their associated reference trace
UpBAR = UpB
UpB = 0;
LowB = 0;
RawTrM=RawTraces{:,,:}; % Converts the table to a Matrix
TrajAve=sum(sum(RawTrM(:,2:end)))/( (W-1)*H) %Average of entire dataset for comparison

```

```

clf
Section 3
Extracts noise of dataset and fits it
DeltaI=RawTrM(1:end-1,2:end)-RawTrM(2:end,2:end); %Finds difference in intensity between every
consecutive frame in the dataset

B=DeltaI(DeltaI~=0); %Histograms noise, removes any zero values to avoid a large peak @ zero
h=histogram(B,1500,'BinLimits',[min(B),max(B)]); %Always makes 1500 bins, size varies based on
noise spread
xlabel('Delta Intensity')
counts=h.Values.';

x3=[min(B):(max(B)-min(B))/1499:max(B)].';
x3=x3(counts~=0);
counts=counts(counts~=0);
[NoiseFit,gofnoise] = fit(x3,counts,'gauss1') %Displays fit for noise and goodness of fit stats
plot(NoiseFit,x3,counts) %Plots noise fit
xlabel('Delta in A.U.')
coeffNoise=coeffvalues(NoiseFit);

```

```

%Section 4: Identifies state changes and binarizes based on them

```

```

Fmean=coeffNoise(1,2); %Finds values of noise fit
FSD=coeffNoise(1,3)/(2^(1/2));
FLB=Fmean-1.5*(FSD) %Frames that change by more than one of the bounds are considered points of
interest
FUB=Fmean+1.5*(FSD) %Points of interest are possible state changes
Points=zeros(1,W);%Where points will be stored
for k=1:1:W-1 %Checks for frames where change is larger than the bounds
    FrameNum=1;
    for l=1:1:H-1
        if DeltaI(l,k)>=FUB
            Points(FrameNum,k)=1;
            FrameNum=FrameNum+1;
        elseif DeltaI(l,k)<=FLB
            Points(FrameNum,k)=1;
            FrameNum=FrameNum+1;
        else
            end
    end
end
ExperimentBins=zeros(H,W-1);
A=mat2dataset(Points);
A=dataset2table(A);
H2=height(A);

```

```

%Section 5: Binarized trace storage

```

```

savedtraces = cell(1,5); %Where extracted on times, off times and binarized traces will be stored
savedontimes = cell(1,5);
savedofftimes = cell(1,5);
for a=1:5
    UpB = UpBAr(a);
    LowB = LowBAr(a);
for k=1:1:W-1 %Actually binarizes
    for l=1:1:H2 %Checks between every point of interest and finds the average between them
        POI=Points(l,k); % Compares averages to the bounds set by selected trace
        POI2=0; %If it is below the low bound = 0
        if l~=1 % Above the upper bound =2
            POI2=Points(l-1,k); % Otherwise = 1
        end
        if l==1

```

```

        Section=RawTrM(1:POI,k+1);
        SectionMean=mean(Section);
        if SectionMean>UpB
            ExperimentBins(1:POI,k)=2;
        elseif SectionMean<LowB
            ExperimentBins(1:POI,k)=0;
        else
            ExperimentBins(1:POI,k)=1;
        end
    elseif POI==0 && POI2~=0
        Section=RawTrM(POI2:H,k+1);
        SectionMean=mean(Section);
        if SectionMean>UpB
            ExperimentBins(POI2:H,k)=2;
        elseif SectionMean<LowB
            ExperimentBins(POI2:H,k)=0;
        else
            ExperimentBins(POI2:H,k)=1;
        end
    elseif POI==0
    else
        Section=RawTrM(POI2:POI,k+1);
        SectionMean=mean(Section);
        if SectionMean>UpB
            ExperimentBins(POI2:POI,k)=2;
        elseif SectionMean<LowB
            ExperimentBins(POI2:POI,k)=0;
        else
            ExperimentBins(POI2:POI,k)=1;
        end
    end
end
end
for k=1:1:W-1 %Checks for and sets any 1 frame events to neighboring values
    for l=1:1:H % i.e. 1110111 = 1111111
        if l==1
            elseif l==H
            elseif ExperimentBins(l,k)~=ExperimentBins(l-1,k) && ExperimentBins(l-
1,k)==ExperimentBins(l+1,k)
                ExperimentBins(l,k)=ExperimentBins(l-1,k);
            end
        end
    end
end
TS=ExperimentBins(:,ChosenOnes(a)-1); %Shows each chosen trace binarized based on that traces
metrics
subplot(3,2,a)
plot(x,TS);
title(['Trace ', num2str(ChosenOnes(a))])
ylabel('Fl(a.u)');
xlabel('Frame')

% Section 6
% Throws out multi protein events and extracts on and off times

TraceNum=1;
TraceNums = zeros(1,5);
SelectBlinks =[];
for k=1:1:W-1
    if max(ExperimentBins(:,k))==1
        SelectBlinks(:,TraceNum)=ExperimentBins(:,k);
        TraceNum=TraceNum+1;
    else
    end
end
TraceNum=TraceNum-1;
OffNum=1;
OnNum=1;
Ontime=0;
Offtime=0;
for k=1:1:TraceNum

```



```

    for l=1:1:H
    if l==1 && SelectBlinks(l,k)==1
        Ont=1;
    elseif l==1 && SelectBlinks(l,k)==0
        Offt=1;
    elseif SelectBlinks(l,k)==1 && SelectBlinks(l-1,k)==0
        Offtime(OffNum,1)=Offt;
        OffNum=OffNum+1;
        Ont=1;
    elseif SelectBlinks(l,k)==0 && SelectBlinks(l-1,k)==1
        Ontime(OnNum,1)=Ont;
        OnNum=OnNum+1;
        Offt=1;
    elseif SelectBlinks(l,k)==1
        Ont=Ont+1;
    elseif SelectBlinks(l,k)==0
        Offt=Offt+1;
    end
    end
end
SelectBlinks = logical(SelectBlinks);
TraceNums(a)=TraceNum;
savedtraces(a) = SelectBlinks;
savedontimes{a} = Ontime;
savedofftimes{a} = Offtime;
end
TraceNums
Binned =figure();
clf

%Section 7

% Fits on and off times to mono and biexponential fits and reports relevant stats
% If fits are not converging, starting values can be changed

Fitinfo = cell(6,8);
Fitinfo{6,1} = "Onfit";
Fitinfo{6,2} = "OnStats";
Fitinfo{6,3} = "Bionfit";
Fitinfo{6,4} = "BiOnStats";
Fitinfo{6,5} = "Offfit";
Fitinfo{6,6} = "OffStats";
Fitinfo{6,7} = "BioffFit";
Fitinfo{6,8} = "Bioffstats";
OnOffRatio=zeros(5,3);

for a=1:5
    Ontime = cell2mat(savedontimes(a));
    Offtime = cell2mat(savedofftimes(a));
    [~,edges] = histcounts(Ontime,50);

    subplot(5,1,1)
    h=histogram(Ontime,50,'FaceColor','yellow');
    title(['Trace ', num2str(ChosenOnes(a))])
    legend('On times')
    xlabel('Frames')

    xfit = fittestype('a*exp(-u/t)','independent','u');
    fo = fitoptions(xfit);
    fo.Lower = [0 0];
    xfit = fittestype('a*exp(-u/t)','independent','u','options',fo);
    %Start Points for On mono fit
    startPoints = [20, 20];

    counts=h.Values.';

```

```

x3=[edges(1,1):((edges(1,51)-edges(1,1))/49):edges(1,51)].';
[Onfit,OnStats]=fit(x3,counts,xfit,'Start',startPoints);

subplot(5,1,2)
plot(Onfit,x3,counts)
title('On MonoExponential')
xlabel('Frames')
ylabel('On Events')
onfunc= Onfit.a*exp(-x3/Onfit.t);
onresid=counts-onfunc;
subplot(5,1,3)
plot(x3,onresid,'*')
title('On Residuals')
xlabel('Frames')
ylabel('Delta')

bifit = fittype('a*exp(-u/t)+b*exp(-u/t2)','independent','u');
bifo = fitoptions(bifit);
bifo.Lower = [0 0 0 0];
bifo.Upper = [1000,1000,1000,1000];
bifit = fittype('a*exp(-u/t)+b*exp(-u/t2)','independent','u','options',bifo);
%Start Points for "on" biexp fit
bistartPoints = [30, 12, 1, 50];
[biOnfit,biOnStats]=fit(x3,counts,bifit,'Start',bistartPoints);

subplot(5,1,4)
plot(biOnfit,x3,counts)
title('On BiExponential')
xlabel('Frames')
ylabel('On Events')
bionfunc= biOnfit.a*exp(-x3/biOnfit.t)+biOnfit.b*exp(-x3/biOnfit.t2);
bionresid=counts-bionfunc;
subplot(5,1,5)
plot(x3,bionresid,'*')
title('On Bi Residuals')
xlabel('Frames')
ylabel('Delta')
On =figure();
clf
OnLT=(biOnfit.a*biOnfit.t+biOnfit.b*biOnfit.t2)/(biOnfit.a+biOnfit.b);
PercentFast=(biOnfit.a/(biOnfit.a+biOnfit.b))*100;
PercentSlow=(biOnfit.b/(biOnfit.a+biOnfit.b))*100;

[N,edges] = histcounts(Offtime,50);
subplot(5,1,1)
h=histogram(Offtime,50,'FaceColor','#7E2F8E');
title(['Trace ', num2str(ChosenOnes(a))])
legend('Off times')
xlabel('Frames')

counts=h.Values.';
x3=[edges(1,1):((edges(1,51)-edges(1,1))/49):edges(1,51)].';
%Start Points for "off" mono fit
startPoints = [40, 20];
[Offfit,OffStats]=fit(x3,counts,xfit,'Start',startPoints);

subplot(5,1,2)
plot(Offfit,x3,counts)
title('Off MonoExponential')
xlabel('Frames')
ylabel('Off Events')
offfunc= Offfit.a*exp(-x3/Offfit.t);

```

```

offresid=counts-offfunc;

subplot(5,1,3)
plot(x3,offresid,'*')
title('Off Residuals')
xlabel('Frames')
ylabel('Delta')

%Start Points for "off" biexp fit
bistartPoints = [75, 3, 20, 60];
[biOfffit,biOffStats]=fit(x3,counts,bifit,'Start',bistartPoints);

subplot(5,1,4)
plot(biOfffit,x3,counts)
title('Off BiExponential')
xlabel('Frames')
ylabel('Off Events')
biofffunc= biOfffit.a*exp(-x3/biOfffit.t)+biOfffit.b*exp(-x3/biOfffit.t2);
bioffresid=counts-biofffunc;

subplot(5,1,5)
plot(x3,bioffresid,'*')
title('Off Bi Residuals')
xlabel('Frames')
ylabel('Delta')

Off =figure();
clf

OffLT=(biOfffit.a*biOfffit.t+biOfffit.b*biOfffit.t2)/(biOfffit.a+biOfffit.b);
PercentFastoff=(biOfffit.a/(biOfffit.a+biOfffit.b))*100;
PercentSlowoff=(biOfffit.b/(biOfffit.a+biOfffit.b))*100;

Ondata(a,:)= [biOnfit.t,PercentFast,biOnfit.t2,PercentSlow,biOnStats.adjrsquare,OnLT];
Offdata(a,:)= [biOfffit.t,PercentFastoff,biOfffit.t2,PercentSlowoff,biOffStats.adjrsquare,OffLT];

Fitinfo{a,1} = (Onfit);
Fitinfo{a,2} = (OnStats);
Fitinfo{a,3} = (biOnfit);
Fitinfo{a,4} = (biOnStats);
Fitinfo{a,5} = (Offfit);
Fitinfo{a,6} = (OffStats);
Fitinfo{a,7} = (biOnStats);
Fitinfo{a,8} = (biOnfit);

Offtraces = 0;
Ontraces = 0;
for e = 1:TraceNums(a)
    if min(SelectBlinks(1:2,e)) == 0
        Offtraces= Offtraces +1;
    else
        Ontraces = Ontraces +1;
    end
end
OnOffRatio(a,:)= [Ontraces,Offtraces, Ontraces/(Offtraces+Ontraces)];
end
SelectionInfo = [ChosenOnes',BlinksStart',BlinksEnd']

%Shows chosen traces and frames, in case the settings need to be reproduced
Ondata

%Shown as Component 1, % Component 1, Component 2, % Component 2, Adjusted R^2, Lifetime

```

```
OfficialOn = [mean(Ondata(:,6)),std(Ondata(:,6))]  
  
%Average of the lifetimes of the five traces and standard deviation  
  
Offdata  
OfficialOff = [mean(Offdata(:,6)),std(Offdata(:,6))]  
OnOffRatio
```

## References

1. Nov, Y. When Second Best Is Good Enough: Another Probabilistic Look at Saturation Mutagenesis. *Appl. Environ. Microbiol.* **2012**, 78 (1), 258–262.  
<https://doi.org/10.1128/AEM.06265-11>.
2. Manna, P.; Hung, S. T.; Mukherjee, S.; Friis, P.; Simpson, D. M.; Lo, M. N.; Palmer, A. E.; Jimenez, R. Directed Evolution of Excited State Lifetime and Brightness in FusionRed Using a Microfluidic Sorter. *Integr. Biol. (United Kingdom)* **2018**, 10 (9), 516–526. <https://doi.org/10.1039/c8ib00103k>.
3. Merzlyak, E. M.; Goedhart, J.; Shcherbo, D.; Bulina, M. E.; Shcheglov, A. S.; Fradkov, A. F.; Gaintzeva, A.; Lukyanov, K. A.; Lukyanov, S.; Gadella, T. W. J.; Chudakov, D. M. Bright Monomeric Red Fluorescent Protein with an Extended Fluorescence Lifetime. *Nat. Methods* **2007**. <https://doi.org/10.1038/nmeth1062>.
4. Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. Primary Structure of the Aequorea Victoria Green-Fluorescent Protein. *Gene* **1992**, 111 (2), 229–233. [https://doi.org/10.1016/0378-1119\(92\)90691-H](https://doi.org/10.1016/0378-1119(92)90691-H).
5. Baird, G. S.; Zacharias, D. A.; Tsien, R. Y. Biochemistry, Mutagenesis, and Oligomerization of DsRed, a Red Fluorescent Protein from Coral. *Proc. Natl. Acad. Sci. U. S. A.* **2000**. <https://doi.org/10.1073/pnas.97.22.11984>.
6. Thompson, J. D., Higgins, D. G., & Gibson, T. J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research.* **1994**, <https://doi.org/10.1093/nar/22.22.4673>
7. Cranfill, P. J.; Sell, B. R.; Baird, M. A.; Allen, J. R.; Lavagnino, Z.; De Gruiter, H. M.; Kremers, G. J.; Davidson, M. W.; Ustione, A.; Piston, D. W. Quantitative Assessment of Fluorescent Proteins. *Nat. Methods* **2016**, 13 (7), 557–562.  
<https://doi.org/10.1038/nmeth.3891>.

- 
8. Bindels, D. S.; Haarbosch, L.; Van Weeren, L.; Postma, M.; Wiese, K. E.; Mastop, M.; Aumonier, S.; Gotthard, G.; Royant, A.; Hink, M. A.; Gadella, T. W. J. MScarlet: A Bright Monomeric Red Fluorescent Protein for Cellular Imaging. *Nat. Methods* **2016**, *14* (1), 53–56. <https://doi.org/10.1038/nmeth.4074>.
  9. Shemiakina, I. I.; Ermakova, G. V.; Cranfill, P. J.; Baird, M. A.; Evans, R. A.; Souslova, E. A.; Staroverov, D. B.; Gorokhovatsky, A. Y.; Putintseva, E. V.; Gorodnicheva, T. V.; Chepurnykh, T. V.; Strukova, L.; Lukyanov, S.; Zaraisky, A. G.; Davidson, M. W.; Chudakov, D. M.; Shcherbo, D. A Monomeric Red Fluorescent Protein with Low Cytotoxicity. *Nat. Commun.* **2012**, *3*, 1204. <https://doi.org/10.1038/ncomms2208>.
  10. Dean, K. M.; Lubbeck, J. L.; Binder, J. K.; Schwall, L. R.; Jimenez, R.; Palmer, A. E. Analysis of Red-Fluorescent Proteins Provides Insight into Dark-State Conversion and Photodegradation. *Biophys. J.* **2011**, *101* (4), 961–969. <https://doi.org/10.1016/j.bpj.2011.06.055>.
  11. Hung, S. T.; Mukherjee, S.; Jimenez, R. Enrichment of Rare Events Using a Multi-Parameter High Throughput Microfluidic Droplet Sorter. *Lab Chip* **2020**, *20* (4), 834–843. <https://doi.org/10.1039/c9lc00790c>.
  12. Miyawaki, A.; Shcherbakova, D. M.; Verkhusha, V. V. Red Fluorescent Proteins: Chromophore Formation and Cellular Applications. *Curr. Opin. Struct. Biol.* **2012**, *22* (5), 679–688. <https://doi.org/10.1016/j.sbi.2012.09.002>.
  13. Balleza, E.; Kim, J. M.; Cluzel, P. Systematic Characterization of Maturation Time of Fluorescent Proteins in Living Cells. *Nat. Methods* **2018**. <https://doi.org/10.1038/nmeth.4509>.
  14. Konold, P. E.; Yoon, E.; Lee, J.; Allen, S. L.; Chapagain, P. P.; Gerstman, B. S.; Regmi, C. K.; Piatkevich, K. D.; Verkhusha, V. V.; Joo, T.; Jimenez, R. Fluorescence from Multiple Chromophore Hydrogen-Bonding States in the Far-Red Protein TagRFP675. *J. Phys. Chem. Lett.* **2016**. <https://doi.org/10.1021/acs.jpcllett.6b01172>.

- 
15. Martin-Fernandez, M. L.; Tynan, C. J.; Webb, S. E. D. A “pocket Guide” to Total Internal Reflection Fluorescence. *J. Microsc.* **2013**, *252* (1), 16–22. <https://doi.org/10.1111/jmi.12070>.
  16. van den Eynde, R.; Sandmeyer, A.; Vandenberg, W.; Duwé, S.; Hübner, W.; Huser, T.; Dedeker, P.; Müller, M. Quantitative Comparison of Camera Technologies for Cost-Effective Super-Resolution Optical Fluctuation Imaging (SOFI). *JPhys Photonics* **2019**, *1* (4), 44001. <https://doi.org/10.1088/2515-7647/ab36ae>.
  17. Bajar, B. T.; Wang, E. S.; Lam, A. J.; Kim, B. B.; Jacobs, C. L.; Howe, E. S.; Davidson, M. W.; Lin, M. Z.; Chu, J. Improving Brightness and Photostability of Green and Red Fluorescent Proteins for Live Cell Imaging and FRET Reporting. *Sci. Rep.* **2016**, *6* (1), 20889. <https://doi.org/10.1038/srep20889>.
  18. Mukherjee, S.; Hung, S. T.; Douglas, N.; Manna, P.; Thomas, C.; Ekrem, A.; Palmer, A. E.; Jimenez, R. Engineering of a Brighter Variant of the Fusionred Fluorescent Protein Using Lifetime Flow Cytometry and Structure-Guided Mutations. *Biochemistry* **2020**, *59* (39), 3669–3682. <https://doi.org/10.1021/acs.biochem.0c00484>.
  19. Dean, K. M.; Davis, L. M.; Lubbeck, J. L.; Manna, P.; Friis, P.; Palmer, A. E.; Jimenez, R. High-Speed Multiparameter Photophysical Analyses of Fluorophore Libraries. *Anal. Chem.* **2015**, *87* (10), 5026–5030. <https://doi.org/10.1021/acs.analchem.5b00607>.
  20. Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y. Improved Monomeric Red, Orange and Yellow Fluorescent Proteins Derived from *Discosoma* Sp. Red Fluorescent Protein. *Nat. Biotechnol.* **2004**, *22* (12), 1567–1572. <https://doi.org/10.1038/nbt1037>.
  21. Manna, P. Development and Characterization of Improved Red Fluorescent Protein Variants, University of Colorado at Boulder, Boulder, 2018.

22. Dean, K. M.; Lubbeck, J. L.; Davis, L. M.; Regmi, C. K.; Chapagain, P. P.; Gerstman, B. S.; Jimenez, R.; Palmer, A. E. Microfluidics-Based Selection of Red-Fluorescent Proteins with Decreased Rates of Photobleaching. *Integr. Biol. (United Kingdom)* 2015, 7 (2), 263–273. <https://doi.org/10.1039/c4ib00251b>.
23. Mukherjee, S.; Thomas, C.; Wilson, R.; Simmerman, E.; Hung, S. T.; Jimenez, R. Characterizing Dark State Kinetics and Single Molecule Fluorescence of FusionRed and FusionRed-MQ at Low Irradiances. *ChemRxiv* 2022. This content is a preprint and has not been peer-reviewed. <https://doi.org/10.26434/chemrxiv-2022-2dr03>
24. Mukherjee, S.; Jimenez, R. Photophysical Engineering of Fluorescent Proteins: Accomplishments and Challenges of Physical Chemistry Strategies. *J. Phys. Chem. B*, **2022**. <https://doi.org/10.1021/acs.jpcb.1c05629>
25. Sillen, A.; Engelborghs, Y. The Correct Use of “Average” Fluorescence Parameters. *Photochem. Photobiol.* **1998**, 67 (5), 475–486. <https://doi.org/https://doi.org/10.1111/j.1751-1097.1998.tb09082.x>.
26. Drobizhev, M.; Molina, R. S.; Callis, P. R.; Scott, J. N.; Lambert, G. G.; Salih, A.; Shaner, N. C.; Hughes, T. E. Local Electric Field Controls Fluorescence Quantum Yield of Red and Far-Red Fluorescent Proteins. *Front. Mol. Biosci.* 2021, 8 (February), 1–21. <https://doi.org/10.3389/fmolb.2021.633217>.
27. Jordanides, X. J.; Lang, M. J.; Song, X.; Fleming, G. R. Solvation Dynamics in Protein Environments Studied by Photon Echo Spectroscopy. *J. Phys. Chem. B* **1999**, 103 (37), 7995–8005. <https://doi.org/10.1021/jp9910993>.



- 
28. Mertz, E. L.; Tikhomirov, V. A.; Krishtalik, L. I. Stokes Shift as a Tool for Probing the Solvent Reorganization Energy. *J. Phys. Chem. A* **1997**, *101* (19), 3433–3442. <https://doi.org/10.1021/jp963042b>.
29. Mooney, J.; Kambhampati, P. Get the Basics Right: Jacobian Conversion of Wavelength and Energy Scales for Quantitative Analysis of Emission Spectra. *J. Phys. Chem. Lett.* **2013**, *4* (19), 3316–3318. <https://doi.org/10.1021/jz401508t>.
30. Lin, C. Y.; Romei, M. G.; Oltrogge, L. M.; Mathews, I. I.; Boxer, S. G. Unified Model for Photophysical and Electro-Optical Properties of Green Fluorescent Proteins. *J. Am. Chem. Soc.* **2019**, *141* (38), 15250–15265. <https://doi.org/10.1021/jacs.9b07152>.

ProQuest Number: 29067913

INFORMATION TO ALL USERS

The quality and completeness of this reproduction is dependent on the quality and completeness of the copy made available to ProQuest.



Distributed by ProQuest LLC (2022).

Copyright of the Dissertation is held by the Author unless otherwise noted.

This work may be used in accordance with the terms of the Creative Commons license or other rights statement, as indicated in the copyright statement or in the metadata associated with this work. Unless otherwise specified in the copyright statement or the metadata, all rights are reserved by the copyright holder.

This work is protected against unauthorized copying under Title 17,  
United States Code and other applicable copyright laws.

Microform Edition where available © ProQuest LLC. No reproduction or digitization of the Microform Edition is authorized without permission of ProQuest LLC.

ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 - 1346 USA