

Engineering of red fluorescent proteins and red fluorescent protein-based pH  
biosensors

by

Yi Shen

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

Department of Chemistry

University of Alberta

© Yi Shen, 2015

## Abstract

Fluorescent proteins (FP) and FP-based biosensors have become essential tools for cell biology and neuroscience research. This thesis describes efforts to engineer new FPs with red emission and novel genetically encoded biosensors based on red FPs. Directed evolution and semi-rational design are the main techniques used to develop novel and improved red FP and red FP-based biosensors.

First, A new pH-sensitive red fluorescent protein based on mApple, termed as “pHuji”, has been developed with high pH sensitivity, exhibiting over 20-fold fluorescence intensity change between pH 5.5 and pH 7.5. In live cell imaging, cell surface-displayed pHuji demonstrated high pH sensitivity when exposed to buffers with defined pH values. Collaborators have used pHuji for successful visualization of pH changes during exocytosis and endocytosis.

Following the engineering of intensimetric red pH sensors, we next turned our attention to ratiometric red pH sensors. Through a process of semi-rational design and directed evolution, the red FP mApple was successfully engineered into a series of dual-excitation, ratiometric pH sensors. These new red-shifted ratiometric pH sensors, termed pHlorina, exhibit large ratio change (over 70-fold) for pH changes from 5.0 to 7.5. A series of long Stokes shift variants of mApple ( $\lambda_{\text{ex}} = 450 \text{ nm}$  and  $\lambda_{\text{em}} = 610 \text{ nm}$ ) were also developed and characterized.

A photochromic and thermochromic red FP was serendipitously discovered during the process of engineering the long Stokes shift red FP. This protein, which we designated as switchable hypersensitive red FP (shyRFP), was characterized in terms of light, temperature, and pH dependence. A colour switching mechanism that involves protonation coupled *E-Z* isomerization of the protein chromophore was proposed.

The monomeric RFP mCherry is widely used for live cell fluorescence imaging experiments. Using semi-rational design and random mutagenesis, two new mCherry variants were developed: Long Stokes Shift mCherry (LSSmCherry;  $\lambda_{\text{ex}} = 460$  nm and  $\lambda_{\text{em}} = 610$  nm) and Red-Shifted mCherry (RDSmCherry;  $\lambda_{\text{ex}} = 600$  nm and  $\lambda_{\text{em}} = 630$  nm). These two proteins have distinctively different fluorescence excitation and emission profiles from their predecessor mCherry2. These new additions to the FP toolbox provide templates for the engineering of new colour variants and fluorescent sensors with red emission.

Finally, a filter paper-based screening method has been developed to screen for calcium ion ( $\text{Ca}^{2+}$ ) sensitive RFP variants expressed in *Escherichia coli* colonies. These low-affinity  $\text{Ca}^{2+}$  sensors were semi-rationally designed by altering protein barrel residues near the chromophore to form a  $\text{Ca}^{2+}$  binding pocket. By combining high-throughput screening and rational design, a number of  $\text{Ca}^{2+}$  sensitive variants were successfully identified.

## Preface

A version of Chapter 2 has been submitted for publication as Yi Shen, Morgane Rosendale, Robert E. Campbell and David Perrais, “pHuji, a pH-sensitive red fluorescent protein for imaging of exo- and endocytosis”. In this Chapter, I was responsible for the theoretical modeling of pH sensitivity, design and engineering of the new red pH-sensitive variants, *in vitro* characterization of the new pH sensors as well as manuscript preparation. M. Rosendale performed the live cell imaging and analyzed the imaging data. R. E. Campbell and D. Perrais supervised this project and wrote the manuscript.

A version of Chapter 4 has been submitted for publication as Yi Shen, Matthew D. Wiens, and Robert E. Campbell, “A photochromic and thermochromic fluorescent protein”. I was responsible for the discovery of the protein, design of experiments, mutational study of the protein, data analysis and manuscript writing. M. D. Wiens characterized the light, temperature, and pH dependence of the protein. R.E. Campbell supervised the project and edited the manuscript.

I would like to dedicate this thesis to my parents

Yongjian Shen and Aizhong Chen

## **Acknowledgments**

First of all I would like to thank my research advisor, Dr. Robert E. Campbell, for his continuous support and insightful guidance on my research throughout the past five years. Next, I would like to thank my research colleagues in the Campbell group and all the research collaborators. The group has been a source of useful advices and great collaborations. I would also like to express my appreciation to all the professors my thesis examining committee: Dr. John Klassen, Dr. James Harynuk, Dr. Michael Serpe, and Dr. Roberto Chica from University of Ottawa for agreeing to review my thesis. Last but not least, I would like to thank my family for all their love and encouragement.

## Table of contents

<b>Chapter 1 Red fluorescent protein (RFP) and RFP-based biosensors .....</b>	<b>1</b>
1. 1 Red fluorescent proteins (RFPs) .....	2
1.1.1 Monomerization of RFPs .....	4
1.1.2 Structure and chromophore formation in RFPs .....	7
1.1.3 Classification of RFPs .....	10
1.1.4 Standard RFPs .....	12
1.1.5 Far-red RFPs .....	14
1.1.6 Long Stokes Shift RFPs .....	16
1.2 RFP-based biosensors .....	17
1.2.1 Classification of RFP-based biosensors .....	18
1.2.2 RFP Biosensors based on BiFC .....	19
1.2.3 RFP Biosensors based on FRET .....	21
1.2.4 RFP Biosensors based on Single FP .....	26
1.3 Engineering of RFP and RFP-based biosensors .....	31
1.3.1 Rational, semi-rational, and computer-aided design .....	32
1.3.2 Directed evolution .....	34
1.4 Scope of this thesis .....	39
<b>Chapter 2 Red fluorescent protein-based intensimetric pH sensors .....</b>	<b>41</b>
2.1 Introduction .....	42
2.2 Results .....	43
2.2.1 Characterization of SEP and previously reported red pH sensors .....	43

2.2.2	pHorans: pH-sensitive mOrange variants with fine-tuned $pK_a$ s .....	50
2.2.3	pHuji: a red pH-sensitive FP with near-optimal $pK_a$ and a high apparent Hill coefficient .....	54
2.2.4	Characterization and comparison of pH-sensitive FPs <i>in vitro</i> and <i>in situ</i> .....	56
2.2.5	Detection of single exocytic and endocytic events with TfR-red FPs ..	59
2.3	Discussion .....	65
2.4	Materials and methods .....	68
2.4.1	Mutagenesis, library construction and screening of pH-sensitive FPs ..	68
2.4.2	Protein purification and <i>in vitro</i> characterization .....	69
2.4.3	Cell Culture and transfections .....	70
2.4.4	Live cell fluorescence imaging .....	72
<b>Chapter 3 Red fluorescent protein-based excitation ratiometric pH sensors</b>		<b>74</b>
3.1	Introduction .....	75
3.2	Results .....	78
3.2.1	Engineering ESPT pathway in a monomeric RFP mApple .....	78
3.2.2	pHlorina1, an excitation ratiometric pH sensor with red emission .....	80
3.2.3	pHlorina2, a red ratiometric pH sensor with large inversed intensiometric change .....	83
3.2.4	Engineering of a long Stokes shift variant of mApple .....	86
3.3	Discussion .....	87
3.4	Materials and methods .....	89
3.4.1	General methods and materials .....	90

3.4.2 Mutagenesis, library construction and screening of ratiometric pH indicators.....	90
3.4.3 protein purification and <i>in vitro</i> characterization.....	91
<b>Chapter 4 Photochromic and thermochromic red fluorescent protein .....</b>	<b>93</b>
4.1 Introduction.....	94
4.2 Results.....	95
4.2.1 Discovery of the photochromic and thermochromic RFP .....	95
4.2.2 Photochromism of ShyRFP .....	97
4.2.3 Thermochromism of ShyRFP .....	100
4.2.4 pH dependence of ShyRFP .....	104
4.2.5 Mutational analysis and proposed chromism mechanism .....	107
4.3 Discussion.....	110
4.4 Material and Methods .....	111
4.4.1 Molecular cloning and mutagenesis.....	112
4.4.2 Protein purification and characterization.....	112
<b>Chapter 5 Engineering of long Stokes shift and red-shifted variants of mCherry.....</b>	<b>115</b>
5.1 Introduction.....	116
5.2 Results.....	119
5.2.1 mCherry2 as template for directed evolution.....	119
5.2.2 Directed evolution of a long Stokes shift (LSS) mCherry variant.....	123
5.2.3 Directed evolution of a red-shifted (RDS) mCherry variant .....	125
5.2.4 Spectroscopic characterization of LSSmCherry and RDSmCherry ..	129

5.2.5 Identification of mCherry variant with low cytotoxicity .....	132
5.3 Discussion .....	133
5.4 Materials and Methods .....	137
5.4.1 General methods and materials .....	137
5.4.2 Site directed saturation mutagenesis and random mutagenesis .....	138
5.4.3 Library screening .....	139
5.4.4 Protein purification and <i>in vitro</i> characterization .....	140
5.4.5 Bacterial growth rate comparison .....	141
5.4.6 Live cell imaging .....	141
<b>Chapter 6 Filter Paper Blot Screening for RFP-based Sensor Development</b>	
.....	<b>143</b>
6.1 Introduction .....	144
6.2 Results .....	146
6.2.1 Development of a filter paper blot based assay for screening .....	146
6.2.2 Library construction and Ca <sup>2+</sup> responsive variant screening .....	149
6.3 Discussion .....	153
6.4 Materials and methods .....	154
6.4.1 Mutagenesis library construction and molecular cloning .....	154
6.4.2 Filter paper based screen assay .....	155
6.4.3 Protein purification and <i>in vitro</i> characterization .....	156
<b>Chapter 7 Conclusion and future work .....</b>	<b>158</b>
7.1 Summary of the thesis .....	159
7.2 Future directions .....	160

7.2.1 Optimization of red fluorescent protein based pH sensors .....	160
7.2.1 Optimization of the new red fluorescent proteins.....	160
7.2.3 Application of filter paper based screening method .....	161
7.3 Closing remark.....	162

## List of Tables

<b>Table 1.1</b> Properties of selected RFPs.....	11
<b>Table 2.1</b> <i>In vitro</i> characterization of pH-sensitive FPs.....	45
<b>Table 2.2</b> Emission peak (nm) at different pH values.....	49
<b>Table 2.3</b> Mutations of the pH-sensitive red FPs developed in this study.....	55
<b>Table 5.1</b> Spectral properties of mCherry, mCherry2, LSSmCherry and other LSS RFPs.....	130
<b>Table 5.2</b> Spectral properties of mCherry2, RDSmCherry and other far-red RFPs .....	131
<b>Table 6.1</b> Characterization of Ca <sup>2+</sup> responsive fluorescent proteins.....	152

## List of Figures

Figure 1.1 Conversion of the wild-type tetrameric RFP DsRed to an engineered monomeric RFP. ....	5
Figure 1.2 Schematic representation of RFP as a fusion protein fluorescent tag. ..	6
Figure 1.3 Structure of a representative RFP, mCherry.....	8
Figure 1.4 Mechanism for red chromophore formation in DsRed.....	9
Figure 1.5 Biosensor design based on RFP BiFC.....	20
Figure 1.6 Representative FRET-based biosensors with RFPs. ....	24
Figure 1.7 DdRFP- and ddRFP-based biosensors. A, Fluorescence intensity increase upon the formation of heterodimeric ddRFP pair. B, DdRFP-based caspase-3 sensor. C, DdRFP-based Ca <sup>2+</sup> sensor. ....	26
Figure 1.8 Single FP-based sensors. ....	28
Figure 1.9 Schematic presentation of FP circularly permutation at both the DNA and protein levels. ....	30
Figure 1.10 Schematic representation of the process of directed evolution. ....	35
Figure 2.1 Theoretical optimization of pH sensor properties. ....	47
Figure 2.2 Spectra and pH titration curves of pH-sensitive FPs.....	50
Figure 2.3 Additional characterization of pH-sensitive RFPs. ....	53
Figure 2.4 Sensitivity of SEP, pHoran4 and pHuji in HeLa cells with collapsed pH gradients. ....	58
Figure 2.5 Detection of exocytosis events with TfR-pHuji. ....	60
Figure 2.6 Detection of endocytic vesicles containing TfR-red FPs. ....	64
Figure 3.1 Cellular organelle pH values in a typical mammalian cell. ....	76

Figure 3.2 Engineering an ESPT pathway into the RFP mApple.....	80
Figure 3.3 Red excitation ratiometric pH-sensitive protein pHlorina1.....	82
Figure 3.4 Inverse-response red intensiometric pH-sensitive protein pHlorina0.4 and pHlorina2.....	85
Figure 3.5 Long Stokes shift red FP LSSmApple. ....	87
Figure 4.1 Sequence alignment between shyRFP and mApple. ....	97
Figure 4.2 Photochromic properties of shyRFP.....	99
Figure 4.3 SDS-PAGE analysis of shyRFP under blue light illumination and dark conditions. ....	100
Figure 4.4 Light intensity dependence of photoactivation rate and relative absorbance change. ....	102
Figure 4.5 Thermochromic properties of shyRFP. ....	103
Figure 4.6 pH dependence of shyRFP. ....	105
Figure 4.7 Fluorescence excitation change at different pH. ....	106
Figure 4.8 Absorbance and fluorescence excitation at low pH (pH = 4.6).....	107
Figure 4.9 Proposed mechanism of shyRFP photochromism and thermochromism. .....	110
Figure 5.1 Spectral characteristics of mCherry2, LSSmCherry, and RDSmCherry. .....	120
Figure 5.2 Schematic procedure of bacterial growth comparison method for cytotoxicity evaluation.....	121
Figure 5.3 Percentage of total colony numbers comparison between RFP and EGFP expressing <i>E. coli</i> . ....	122

Figure 5.4 Directed evolution of LSSmCherry.....	125
Figure 5.5 Directed evolution of RDSmCherry.....	128
Figure 5.6 Representative fluorescence images of cells expressing mCherry fusions.....	133
Figure 6.1 Schematic presentation of the filter paper blot screening method procedure.....	147
Figure 6.2 Validation of the on filter fluorescence change.....	149
Figure 6.3 Ca <sup>2+</sup> binding site design and library construction. ....	151

## List of Abbreviations

ATP	adenosine triphosphate
Ala	alanine
Amp	ampicillin
Arg	arginine
Asn	asparagine
Asp	aspartate
avGFP	<i>Aequorea victoria</i> green fluorescent protein
β	beta
BFP	blue fluorescent protein
BiFC	bimolecular fluorescence complementation
BPER	bacterial protein extraction reagent
CaM	calmodulin
CCD	charge coupled device
CFP	cyan fluorescent protein
cm	centimetre
cp	circular permutation or circular permuted
Cys	cysteine
d	dimeric
ddFP	dimerization dependent fluorescent protein
ddRFP	dimerization dependent red fluorescent protein
DIC	differential interference contrast
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleic acid triphosphate
DsRed	<i>Discosoma</i> species red fluorescent protein
EGTA	ethylene glycol tetraacetic acid
EP-PCR	error-prone polymerase chain reaction
eqFP	<i>Entacmaea quadricolor</i> fluorescent protein
ER	endoplasmic reticulum

ESPT	excited-state proton transfer
FBS	fetal bovine serum
FP	fluorescent protein
FRET	Förster resonance energy transfer
<i>g</i>	gravitational force
GECCO	genetically encoded Ca <sup>2+</sup> indicator for optical imaging
GFP	green fluorescent protein
Gln	glutamine
Glu	glutamate
Gly	glycine
h	hour
HeLa	cervical cancer cell originated from Henrietta Lacks
HEK	human embryonic kidney
HEPES	hydroxyethyl piperazineethanesulfonic acid
HHBSS	HEPES buffered hanks balanced saline solution
His	histidine
Ile	isoleucine
K <sub>d</sub>	dissociation constant
kDa	kilodalton
LB	Lennox broth
LED	light emitting diode
Leu	leucine
LSS	long Stokes-shift
Lys	lysine
m	monomeric
M	molar (mol/L)
Met	methionine
min	minute
μL	microlitre
mL	millilitre

$\mu\text{M}$	micromolar
mM	milimolar
MOPS	3-morpholino propane-1-sulfonic acid
mRFP1	monomeric red fluorescent protein 1
MWCO	molecular weight cut-off
ND	neutral density
$n_H$	Hill coefficient
nm	nanometre
nM	nanomolar
NTA	nitrilotriacetic acid
OD	optical density
PA	photoactivation
PALM	photoactivated localization microscopy
PBS	phosphate buffered saline
PDB	Protein Database
<i>Pfu</i>	<i>Pyrococcus furiosus</i>
Phe	phenylalanine
$pK_a$	logarithmic measure of the acid dissociation constant or pH at which a fluorescent protein has 50% fluorescence signal
POI	protein of interest
PPI	protein-protein interaction
Pro	proline
PS	photoswitchable
QY	quantum yield
RDS	red-shifted
RESOLFT	reversible saturable optical fluorescence transitions
RFP	red fluorescent protein
rpm	revolutions per minute
rs	reversible switching

SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEM	standard error mean
SEP	superecliptic pHluorin
s	second
Ser	serine
StEP	staggered extension process
TBS	Tris buffered saline
Thr	threonine
Trp	tryptophan
Tyr	tyrosine
UV	ultraviolet
Val	valine
W	watt
YFP	yellow fluorescent protein

**Chapter 1 Introduction to red fluorescent protein (RFP) and  
RFP-based biosensors**

## 1. 1 Red fluorescent proteins (RFPs)

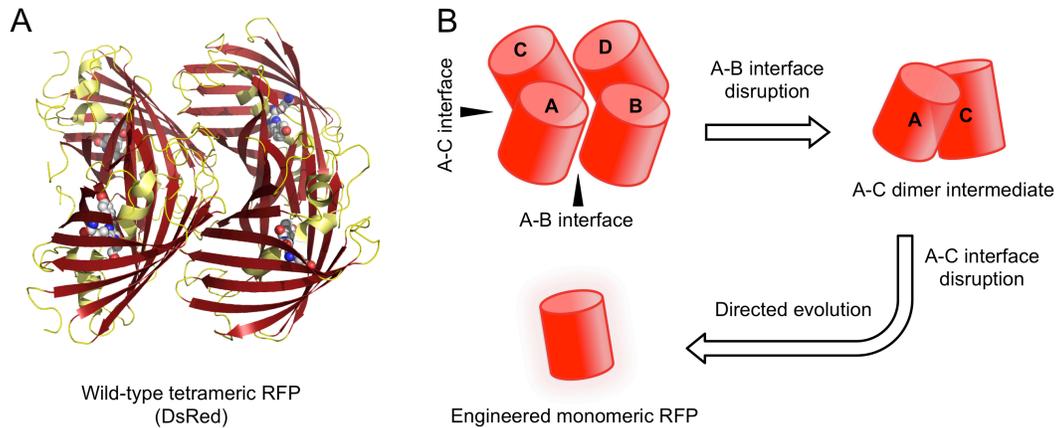
Fluorescent protein (FP)-based techniques have become indispensable tools for life science research and have revolutionized the ability of researchers to undertake real-time live cell imaging of biomolecules. In the 1960s, a protein, subsequently named avGFP (*Aequorea victoria* green FP), was extracted from *Aequorea victoria* jellyfish and purified to demonstrate its green fluorescence<sup>1</sup>. In the following decades, advances in the technology of molecular biology facilitated the elucidation of the DNA sequence as well as the amino acid sequence of this green FP<sup>2</sup>. It was not until 1994 when the recombinant avGFP transgene was first used as a tool for molecular imaging. Specifically, it was used for visualization of gene expression in the worm *Caenorhabditis elegans*, demonstrating the avGFP is self-sufficient to fluoresce in cells from species other than jellyfish<sup>3</sup>. During the next several years, avGFP gained widespread acceptance as a revolutionary tool to visualize and track molecules and biochemical events in living cells and organisms using fluorescence. During this same time period, efforts were being made to engineer improved variants of avGFP by manipulating and modifying the gene sequence. One of the outcomes of these efforts was a significantly brighter and stable variant of the original avGFP, named EGFP<sup>4</sup>. In addition, a number of variants were engineered with fluorescence maxima in the blue, cyan, green, and yellow regions of the visible spectrum<sup>5</sup>.

One of the most significant advances following the initial cloning and early

mutagenesis efforts on the avGFP was the discovery of homologues in non-bioluminescent reef corals and sea anemones. The most important property of these new FPs was their colour diversity, which included variants fluorescent emission hues in the red region of the visible spectrum<sup>6</sup>. The first reef coral-derived red FP to be extensively studied and engineered was isolated from a sea anemone called *Discosoma* sp. This red FP was originally designated drFP583, but is now commonly referred to as DsRed<sup>7</sup>. After expression and purification, mature DsRed has an excitation wavelength maximum ( $\lambda_{\text{ex}}$ ) at 558 nm, and an emission wavelength maximum ( $\lambda_{\text{em}}$ ) at 583 nm. However, DsRed suffers from critical drawbacks when applied as a genetically encoded fluorescent imaging tag. DsRed has a very slow maturation rate ( $t_{0.5} > 10$  hours), and a substantial fraction of the protein molecules form a dead-end green fluorescent chromophore. The green product causes signal bleed-through between different fluorescent channels, which limits the utility of DsRed for multi-colour imaging experiments with GFP variants<sup>8</sup>. More importantly, DsRed is an obligate tetramer in its native form, with a tendency to form higher order oligomers. Any target protein fused to DsRed would become tetrameric itself, which can interfere with the native localization of the target protein and lead to aggregation<sup>9, 10</sup>. For these reasons, the applications of DsRed in cell biology are limited. The obvious solution for addressing the most pressing issues with DsRed was to engineer a monomeric version (see Section 1.1,1).

### 1.1.1 Monomerization of RFPs

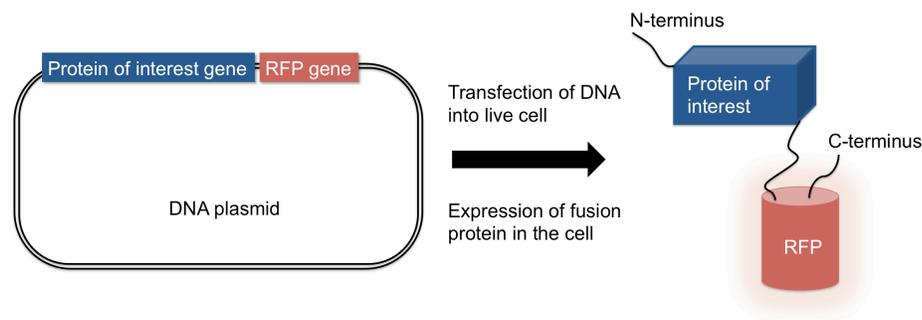
Following a tremendous protein engineering effort, monomerization of DsRed ultimately yielded monomeric RFP 1 (mRFP1;  $\lambda_{\text{ex}} = 584 \text{ nm}$ ;  $\lambda_{\text{em}} = 607 \text{ nm}$ )<sup>11</sup>. In the tetrameric DsRed, each subunit is in association with two neighbouring ones thus forming two types of contacting interface, namely the A-B and A-C interfaces (Figure 1.1). In order to make mRFP1, the interface residues were mutated in order to destabilize the protein-protein contacts at the interface. Disrupting one interface yielded a dimeric intermediate and subsequent disruption of the other interface produced the monomeric FP (Figure 1.1). Following the monomerization of DsRed, an extensive process of directed evolution was used to recover the fluorescence brightness lost during the interface disruption by selecting for variants with improved red fluorescence. A total of 33 mutations were introduced during the course of engineering mRFP1, with 13 interface-disrupting mutations and 20 fluorescence-rescuing mutations<sup>11</sup>.



**Figure 1.1** Conversion of the wild-type tetrameric RFP DsRed to an engineered monomeric RFP. A, cartoon representation of the structure of wild-type tetrameric RFP DsRed. B, disruption of the first A-B interface produces a A-C dimer intermediate and subsequent disruption of the A-C interface produces a monomeric RFP. Interface-disrupting mutations are typically detrimental to the proper folding and chromophore maturation of the intermediate dimer or monomer, therefore these variants must be rescued by directed evolution. Cartoon structures are based on PDB IDs 1G7K<sup>10</sup>.

The monomeric property of mRFP1 overcomes the most critical shortcoming associated with its tetrameric template DsRed. Other favourable properties of mRFP1 include a much shorter maturation time ( $t_{0.5} < 1$  hour) and a 25 nm red-shifted fluorescence emission at 607 nm. These advantages make mRFP1 a suitable RFP in the construction of fusion proteins for fluorescence cell imaging, as well as in the multi-colour fluorescence imaging with EGFP (Figure 1.2)<sup>11</sup>.

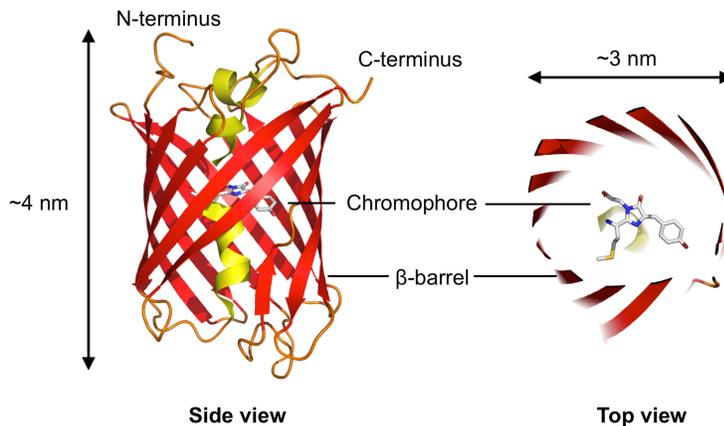
Unfortunately, mRFP1 also exhibits disadvantages including reduced fluorescence brightness compared to DsRed, and relatively poor photostability. Following the development of mRFP1, directed evolution efforts continued, with a focus on higher brightness, colour diversification and better photostability. These efforts eventually produced a number of useful RFP variants which are now known as the mFruit series<sup>12, 13</sup>. The prototypical RFP in the mFruit series is mCherry, which is generally considered to be the successor of mRFP1.



**Figure 1.2** Schematic representation of RFP as a fusion protein fluorescent tag. The gene of an RFP is fused to the gene of the protein of interest. After transfection of the DNA into cell, the fusion protein is expressed. The red fluorescence of the RFP allows the subcellular location of the protein of interest to be visualized using fluorescence microscopy.

### **1.1.2 Structure and chromophore formation in RFPs**

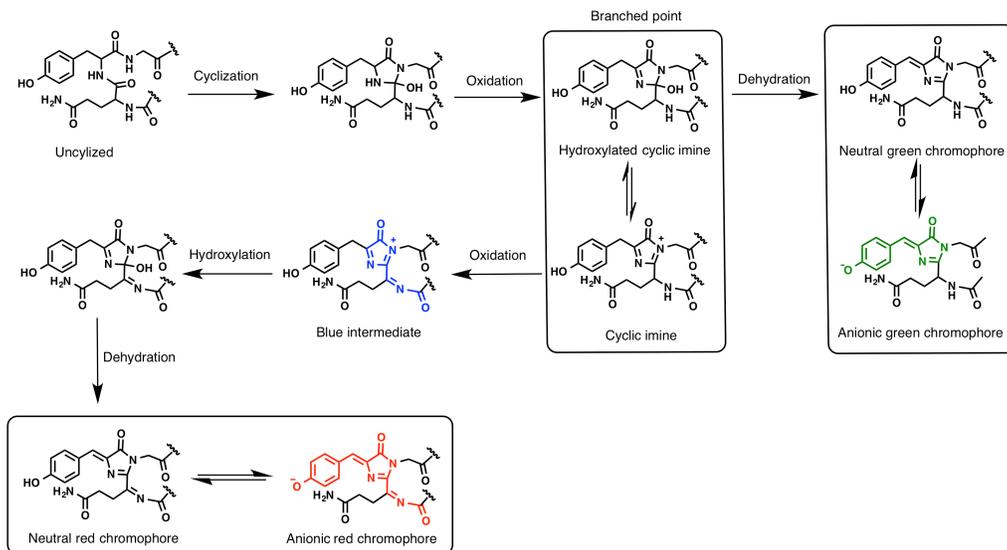
RFPs, either from reef coral or sea anemone, are composed of an eleven-stranded  $\beta$ -barrel that is very similar to the structure of avGFP<sup>9, 10, 14, 15</sup>. The x-ray crystal structures of DsRed and mCherry (as examples of prototypical RFPs) reveal that these proteins have a cylindrical shape created by eleven  $\beta$ -strands wrapped around a central helix<sup>16</sup>. Owing to its unique structure, this eleven-stranded cylindrical protein fold is often referred to as  $\beta$ -can or  $\beta$ -barrel. This  $\beta$ -barrel is ~4 nm in height and ~3 nm in diameter (Figure 1.3). The chromophore is in the middle portion of the central helix and is positioned near the central axis of the  $\beta$ -barrel, protected in the hydrophobic environment created by the surrounding eleven  $\beta$ -strands.



**Figure 1.3** Structure of a representative RFP, mCherry. The secondary structure is in a cartoon representation with the helix coloured in yellow,  $\beta$ -strands coloured in red, and loops coloured in orange. The chromophore is shown in a stick representation with carbon atoms coloured in grey, nitrogen atoms coloured in blue, and oxygen atoms coloured in red (PDB ID# 2H5Q).

DsRed forms its chromophore from three sequential amino acids: Gln65, Tyr66, and Gly67<sup>7</sup>. There have been several reports of investigations into the mechanism of the chromophore formation<sup>7, 17, 18</sup>. The latest and preferred proposed mechanism is a branched pathway model<sup>19</sup>. This model suggests that the formation of a DsRed chromophore starts with the cyclization of the main-chain to form a 5-membered ring intermediate. This 5-membered ring intermediate undergoes an initial step of oxidation. The product of oxidation is a hydroxylated cyclic imine, which equilibrates with a cyclic imine. The branched point lies at the equilibrium. Dehydration of the hydroxylated cyclic imine forms the neutral green

chromophore. In another branch, irreversible oxidation of the cyclic imine leads to an intermediate with blue fluorescence. Further dehydroxylation and dehydration leads to the formation of the neutral form of the red chromophore. This phenol chromophore undergoes deprotonation to become the anionic form of the chromophore that gives rise to the characteristic red fluorescence of DsRed (Figure 1.4)<sup>19, 20</sup>. Understanding the RFP chromophore structure and its formation process opens avenues for manipulating the chromophore structure and its surrounding environment to obtain desirable FP phenotypes. A variety of RFP-derived colours have been engineered from the diversification of the chromophores and their immediate environments as will be described below.



**Figure 1.4** Mechanism for red chromophore formation in DsRed.

### 1.1.3 Classification of RFPs

FPs are now available in a wide range of colours spanning the whole visible spectrum, from violet to far-red<sup>21,22</sup>. RFPs have a number of inherent advantages, and nicely complement other spectral FP variants, including the widely used EGFP variant, for multi-colour imaging applications. Red-shifted fluorescence is associated with reduced background auto-fluorescence, lower phototoxicity, and better tissue penetration due to lower absorption<sup>11, 12, 23</sup>. All other factors being the same, these properties should make RFPs superior probes for fluorescence imaging, particularly in deep tissue *in vivo* imaging.

Designing and engineering RFPs to meet the needs of modern biological research, as well as developing unprecedented variants that can drive the development of new or emerging imaging strategies, is a rapidly growing research area. Many RFPs of different spectral phenotypes have been developed since the first monomeric RFP. Currently, there are at 3 main classes of RFPs according to the fluorescence spectral profiles, including standard FPs, far-red FPs, and RFPs with a large Stokes shift (LSS RFPs) (Table 1). It is important to note that the designation of all of these proteins as ‘red’ is misleading, since many emit wavelengths of light that would appear orange to the naked eye.

**Table 1.1** Properties of selected RFPs

<b>Protein</b>	$\lambda_{\text{ex}}$ (nm)	$\lambda_{\text{em}}$ (nm)	<b>EC</b>	<b>QY</b>	<b>Bright ness</b>	<b>p<i>K<sub>a</sub></i></b>	<b>Bleaching (s)</b>	<b>Maturation (min)</b>
<b>Standard Red FPs</b>								
mOrange	548	562	71000	0.69	49	6.5	9	150
mOrange2	549	565	58000	0.6	34.8	6.5	N.D.	N.D.
TagRFP	555	584	100000	0.48	49	3.8	48	100
TagRFP-T	555	584	81000	0.41	33.2	4.6	337	100
mRuby	558	605	112000	0.35	39.2	4.4	N.D.	N.D.
mRuby2	559	600	113000	0.38	43	5.3	123	150
mTangerine	568	585	38000	0.3	11.4	5.7	N.D.	N.D.
mApple	568	592	75000	0.49	36.7	6.5	N.D.	N.D.
mStrawberry	574	596	90000	0.29	26.1	4.5	15	50
FusionRed	580	608	95000	0.19	18.1	4.6	150	130
mCherry	587	610	72000	0.22	15.8	4.5	96	40
<b>Far Red FPs</b>								
mKate	588	635	31500	0.28	8.8	6.2	N.D.	N.D.
mKate2	588	633	62500	0.4	25	5.4	84	20
mPlum	590	649	41000	0.1	4.1	4.5	53	100
mRaspberry	598	625	86000	0.15	12.9	N.D.	N.D.	N.D.
mNeptune	600	650	67000	0.2	13.4	5.4	255	35
TagRFP657	611	657	34000	0.1	3.4	5	N.D.	N.D.
TagRFP675	598	675	46000	0.08	3.7	5.7	N.D.	25
mCardinal	604	659	87000	0.19	16.5	5.3	730	27
<b>LSS RFPs</b>								
mKeima	440	620	14400	0.24	3.5	6.5	N.D.	N.D.
mBeRFP	446	611	65000	0.27	17.6	5.6	N.D.	N.D.

LSSmKate2	460	605	26000	0.17	4.4	2.7	N.D.	N.D.
LSSmKate1	463	624	31200	0.08	2.5	3.2	N.D.	N.D.
LSSmOrange	437	572	52000	0.45	23.4	5.7	N.D.	138

EC: extinction coefficient; QY: quantum yield; Brightness: the product of EC and QY;  $pK_a$ : pH value at which the fluorescence intensity is 50% of maximal; Bleaching: time for fluorescence intensity to be photobleached by 50% under arc-lamp illumination; Maturation: time for fluorescence intensity to reach 50 % maximal with the exposure to oxygen. N.D.: not determined.

### 1.1.4 Standard RFPs

The class of standard RFPs can be further sub-divided into groups of orange FPs with emission maximum at 550–580 nm, and red FPs with emission maximum at 580–620 nm. One of the biggest breakthroughs in generating standard RFPs in the orange and red spectral regions resulted from the further directed evolution of mRFP1. The first amino acid Gln66 in the chromophore, which is a critical determinant of the spectral characteristics, plays an important role in shifting the spectral profiles of mRFP1 derivatives. For example, mRFP1.1 contains a Gln66Met mutation that red-shifts excitation and emission by 5 nm. Continued screening for improved folding and brightness led to the discovery of the Met163Gln mutation that ultimately led to the production of mCherry ( $\lambda_{ex} = 587$  nm;  $\lambda_{em} = 610$  nm)<sup>12</sup>, the most widely used monomeric RFP to date. Introduction of the Met66Cys mutant led to a blue-shifted variant known as mTangerine, with  $\lambda_{em} = 585$  nm. The Met66Thr mutation led to the development of mOrange and the orange-red protein mStrawberry, which emit at 562 and 596 nm, respectively.

The dimeric intermediate created during the process of DsRed monomerization was also improved and genetically fused to a second copy of its own gene to create a “tandem dimer” RFP termed tdTomato ( $\lambda_{\text{ex}} = 554 \text{ nm}$ ;  $\lambda_{\text{em}} = 581 \text{ nm}$ ). Due to the formation of an intramolecular pseudo-dimer, tdTomato behaves like a monomeric RFP, yet the dimer structure offers exceptional brightness<sup>12</sup>.

Later RFP engineering efforts were focused on improving the photostability of the mRFP1 variants, which led to the production of mOrange2 ( $\lambda_{\text{ex}} = 549 \text{ nm}$ ;  $\lambda_{\text{em}} = 565 \text{ nm}$ ) and mApple ( $\lambda_{\text{ex}} = 568 \text{ nm}$ ;  $\lambda_{\text{em}} = 592 \text{ nm}$ ). mOrange2 has a 25-fold increase in photostability from its template mOrange. mApple shares the same chromophore structure as mRFP1.1, but has a slightly blue-shifted  $\lambda_{\text{em}}$  at 592 nm<sup>13</sup>. These new FPs were named after fruits due to the similarity of their respective emission colours and many common fruits.

The mFruit series of RFPs has yielded a tremendous wealth of information about FP structure-function relationships. mOrange and mCherry, owing to their high brightness and photostability, have also become the FPs of choice for most live cell fluorescence imaging experiments where orange or red fluorescence is needed. However, several of the other mFruit FPs, including mTangerine, and mStrawberry, suffer from low intrinsic brightness and poor photostability, and are

not often used for imaging purposes.

In addition to DsRed-derived monomeric variants, a second lineage of standard monomeric RFPs was engineered from the sea anemone *Entacmaea quadricolor* RFPs eqFP578 and eqFP611<sup>24,25</sup>. For example, TagRFP ( $\lambda_{\text{ex}} = 555$  nm;  $\lambda_{\text{em}} = 584$  nm) is a bright monomeric RFP engineered from the dimeric RFP eqFP578<sup>25</sup>. The Ser158Thr variant of TagRFP, designated as TagRFP-T, is approximately 9-fold more photostable than its template TagRFP<sup>13</sup>. The latest addition to the eqFP578 variant family is an RFP named as FusionRed ( $\lambda_{\text{ex}} = 580$  nm;  $\lambda_{\text{em}} = 608$  nm), with decreased cytotoxicity when expressed in mammalian cells<sup>26</sup>. The eqFP611 lineage has also yielded mRuby ( $\lambda_{\text{ex}} = 558$  nm;  $\lambda_{\text{em}} = 605$  nm) along with its brighter second-generation variant mRuby2 ( $\lambda_{\text{ex}} = 559$  nm;  $\lambda_{\text{em}} = 600$  nm). Both variants offer a relatively long Stokes shift ( $\sim 50$  nm) between the excitation and emission maxima<sup>27,28</sup>.

### 1.1.5 Far-red RFPs

Far-red FPs with emission maximum over 620 nm are of particular importance for *in vivo* and deep tissue imaging in small animal models such as mice and rats. The far-red FPs are superior to FPs of other colours for deep tissue and *in vivo* imaging because mammalian tissue is more transparent to far-red light due to the strong absorption by haemoglobin at wavelengths below  $\sim 600$  nm. Between 600

nm and 1200 nm, beyond which infrared absorption by water molecules begins, there exists a relatively transparent “optical window”. This “optical window” has motivated FP engineers to push the excitation and emission wavelengths of RFP into the far-red and even the near infrared region<sup>23</sup>. Thus, developing far-red shifted bright FPs has thus been a widely recognized goal for FP engineering ever since the development of mRFP1.

Early efforts in creating further red shifted RFPs from mRFP1 yielded mPlum ( $\lambda_{\text{ex}} = 590 \text{ nm}$ ;  $\lambda_{\text{em}} = 649 \text{ nm}$ ), and mRaspberry ( $\lambda_{\text{ex}} = 598 \text{ nm}$ ;  $\lambda_{\text{em}} = 625 \text{ nm}$ )<sup>29</sup>. mGrape represented yet another attempt of red shifting mRFP1, but the penultimate variant mGrape3 ( $\lambda_{\text{ex}} = 608 \text{ nm}$ ;  $\lambda_{\text{em}} = 646 \text{ nm}$ ) was relatively dim and displayed a photochromic behaviour with switching between green and far-red fluorescence<sup>23</sup>.

mCherry has also served as a template for engineering of longer wavelength emission with the aid of a computationally designed library<sup>30</sup>. This effort led to the creation of mRouge with maximal excitation at 600 nm and emission at 637 nm. However, the brightness of mRouge was still relatively low, primarily due to a low quantum yield (0.03).

mKate ( $\lambda_{\text{ex}} = 588 \text{ nm}$ ;  $\lambda_{\text{em}} = 635 \text{ nm}$ ) and mKate2 ( $\lambda_{\text{ex}} = 588 \text{ nm}$ ;  $\lambda_{\text{em}} = 633 \text{ nm}$ )

were developed from wild-type eqFP578, and demonstrate bright far-red emission above 630 nm<sup>31, 32</sup>. Further efforts led to the development of two bright far-red mKate derivatives, mCardinal<sup>33</sup> and TagRFP657<sup>34</sup>. These variants have excitation maxima of 604 nm and 611 nm, respectively, and a 659 nm emission maxima for both. To date, the record for the most red-shifted emission maximum for an RFP is 675 nm, attributed to another mKate variant named as TagRFP675<sup>35</sup>.

### **1.1.6 Long Stokes Shift RFPs**

Long Stokes shift (LSS) refers to fluorophores for which the difference between the fluorescence excitation and emission maxima is larger than ~100 nm. The availability of LSS FPs provides researchers with a greater selection of spectrally resolvable colours for multi-colour imaging application. LSS RFPs absorb blue light (usually in the range of ~440 nm to ~460 nm) and fluoresce in the red region of the visible spectrum. Thus, LSS RFPs hold particular promise for two-photon fluorescence imaging due to the fact that they can typically be excited at the same two-photon wavelength as EGFP using widely available pulsed laser systems.

The first LSS RFP, known as mKeima, was developed based on a chromoprotein from the stony coral *Montipora* sp.<sup>36</sup>. Starting from mKate, three new LSS RFPs, LSSmKate1, LSSmKate2<sup>37</sup>, and mBeRFP<sup>38</sup> were engineered by introducing an excited state proton transfer (ESPT) pathway for the protein chromophore. Upon

blue light excitation, the excited state of the chromophore has a lower  $pK_a$  and thus transfers a proton through a hydrogen bond network to generate an excited state anionic chromophore with a reduced energy, which then emits red fluorescence. LSSmKate1 and LSSmKate2 outperform mKeima in terms of pH-stability, photostability, and brightness. Unlike mKeima and mBeRFP, LSSmKates also lack the additional red excitation peak associated with the anionic ground state (i.e., normal Stokes shift) in the yellow region of the spectra. The lack of this peak ensures compatibility with standard RFPs in multi-colour fluorescence imaging. The strategy of engineering an ESPT pathway into a standard RFP has also been applied to some of the mFruit RFPs, including mOrange and mCherry, to generate variants with blue-shifted fluorescence excitation<sup>39</sup>. Further development produced LSSmOrange, which exhibits the highest brightness among all the LSS RFPs<sup>40</sup>.

## **1.2 RFP-based biosensors**

Almost two decades ago, the advent of FPs launched a new and unprecedented era in cell biology by enabling researchers to monitor cellular processes in living systems using fluorescence microscopy and related imaging technologies<sup>41</sup>. For many applications, including monitoring of transcription and/or translation, visualization of organelles and other subcellular structures, and imaging of biomolecule motility and dynamics, the FP has a “passive” role<sup>42</sup>. However, for many other applications the FP is designed to play an “active” role, meaning that

its inherent fluorescence intensity or hue will change in response to a specific cellular process of interest. These “active” FP constructs are often referred to as sensors, biosensors, indicators, or reporters. Examples of such FP-based biosensors include ones for intracellular pH, concentration of various ions, secondary messengers such as ATP<sup>43</sup>, redox potential<sup>44</sup>, membrane voltage<sup>45</sup>, reactive oxygen species<sup>46</sup>, and various enzyme activities<sup>47</sup>. The utility of these biosensors can be further multiplied by combining them with specific promoters and/or targeting signals to target the biosensor to specific organelles, cells, or tissues. In addition, since these biosensors are entirely protein-based then can be expressed in transgenic model organisms to enable true *in vivo* monitoring of a variety of physiological processes. Since red-shifted light penetrates deeper into tissues and is associated with lower levels of autofluorescence, RFPs are attractive components for biosensor design<sup>48, 49</sup>. RFP-based biosensors also complement the existing pool of GFP-based sensors, allowing concurrent multiple parameter imaging.

### **1.2.1 Classification of RFP-based biosensors**

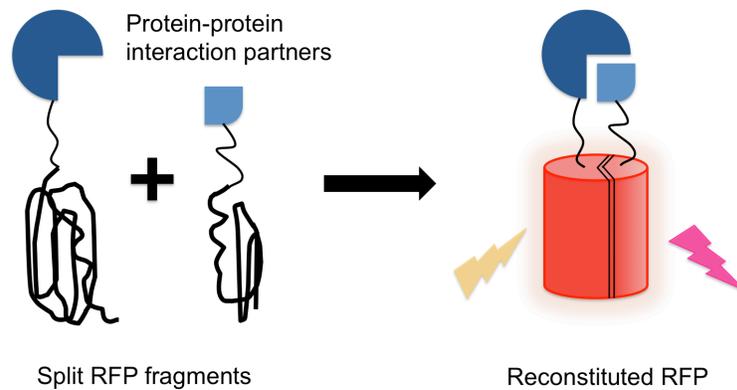
Biosensors generally are composed of two parts: a molecular recognition/binding element that interacts with the sensing target; and a signal transducing element that converts the interaction into a detectable signal, such as fluorescence change. With the highly fluorescent intrinsic chromophore, RFPs are very good choices to serve as signal transducers in biosensors. Several approaches have been employed

to convert RFPs into effective signal output domain for a variety of recognition events. Based on the design strategies, RFP-based biosensors can be categorized into four main classes: bimolecular fluorescence complementation (BiFC)-based biosensors, Förster resonance energy transfer (FRET)-based biosensors, dimerization dependent FP (ddFP)-based biosensors, and single FP-based biosensors.

### **1.2.2 RFP Biosensors based on BiFC**

BiFC-based biosensors are based on complementation, which is the reconstitution of two protein fragments into a functional whole protein. For detection of protein-protein interactions with this type of biosensor, the FP is genetically split into two parts at the DNA level. The split portions are each fused to the gene of the potentially interacting protein pair. Before complementation, the FP fragments are partially unfolded and non-fluorescent. Protein-protein interaction (PPI) between the two protein partners brings the non-fluorescent FP fragments in to close proximity, allowing the formation of the functional FP (Figure 1.5). Following the first successful demonstration of the reconstitution of split avGFP fragments<sup>50,51</sup>, RFPs have been also been shown to be capable of undergoing a similar fragment complementation to produce an alternative fluorescent colour. The first RFP BiFC system was developed using a mRFP1 mutant (Gln66Thr) with excitation and emission wavelengths of 549 and 570 nm, respectively<sup>52</sup>. mCherry, mPlum, and mKate were also used as templates to develop new BiFC systems with longer emission wavelength and brighter reconstituted fluorescence<sup>53-55</sup>. More recently,

an mNeptune-BiFC system was introduced to extend the RFP BiFC system into the far-red region in emission. This system was successfully applied for *in vivo* imaging of RNA–protein and protein–protein interactions<sup>56</sup>.



**Figure 1.5** Biosensor design based on RFP BiFC. Two potentially interacting proteins are fused to the two split RFP fragments respectively. Interaction between the two protein partners bring the RFP fragments in close proximity, leading to reconstitution of an intact RFP to give a red fluorescence signal output.

With the availability of various split RFP systems, construction of BiFC-based biosensors is relatively straightforward and optimization of such biosensors is usually not required. As the complementation is irreversible, it is useful for detection of transient PPI, but not suitable if the reversibility of the protein

complex formation is also of interest. Moreover, formation of a mature chromophore in the reconstituted FP usually takes tens of minutes in the BiFC-based biosensors<sup>55, 57</sup>. Fortunately, the problems of irreversibility and slow kinetics can be overcome using alternative biosensing strategies including FRET-based biosensors and fluorogenic biosensors based on ddFPs.

### **1.2.3 RFP Biosensors based on FRET**

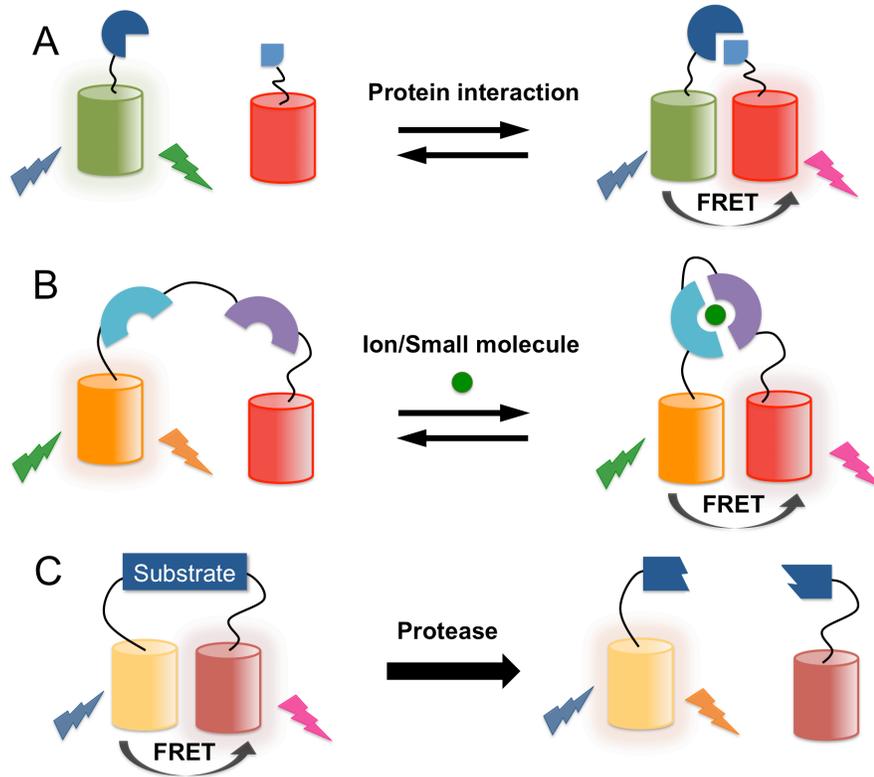
FRET is the radiationless energy transfer through dipole-dipole interaction between two chromophores that have compatible energy levels and are close in distance (<10 nm). The basic design principle of all FRET-based biosensors is to couple a specific binding event or covalent modification of a protein to a change of the energy transfer efficiency between the donor FP and the acceptor FP. A variety of FP FRET-based biosensors for detection of PPI, ion concentrations, small molecule concentrations, and enzyme activities, have been developed (Figure 1.6).

The design of intermolecular biosensors for PPI detection is similar to that of BiFC. Instead of having the interacting proteins of interest each fused to the FP fragments, they are fused to the donor FP and acceptor FP. FRET occurs when the two protein partners form a complex. Unlike BiFC, formation of this complex is completely reversible due to the little or no interaction between the donor and

acceptor FPs. For biosensor designs involving the incorporation of ligand induced PPI or a protein conformational change, an intramolecular FRET-based biosensor can be constructed by attaching both donor and acceptor FPs in a single polypeptide. The intermolecular biosensor design offers a more consistent signal output due to the fixed and constant ratio of donor/acceptor concentration in different cells.

The cyan and yellow FP-based FRET donor and receptor pair is an excellent choice for the construction of genetically encoded FRET biosensors due to the large spectral overlap and their comparable brightness. However, the development of various monomeric RFP variants has provided new and exciting possibilities to construct red-shifted FRET pairs. For example, mRuby2, currently the brightest monomeric RFP, was paired with a bright GFP variant, Clover, to confer a greater dynamic range and photostability in comparison with the existing FRET reporter designs based on CFP and YFP. This makes the Clover-mRuby2 pair a superior alternative for the improvement of existing FRET sensors and the development of new ones<sup>28</sup>. New RFPs also offer the opportunities to construct new FRET pairs with novel spectral properties, especially the ones with spectral compatibility with the existing CFP-YFP pairs<sup>58, 59</sup>. The mOrange-mCherry pair was one of the first to be investigated. Although this pair does provide a red shifted fluorescent spectra, the resulting FRET sensors had only limited responses<sup>60, 61</sup>. Fortunately, the orange-red FRET pair was recently improved by developing self-associating variants of mOrange and mCherry, with reversion of the hydrophobic dimeric

interface breaking mutations<sup>62, 63</sup>. One of the major advantages of a red-shifted FRET pair is the ability to simultaneously image both it and the traditional CFP-YFP FRET pair. For example, the LSSmOrange and mKate2 FRET pair and the CFP-YFP FRET pair have been simultaneously imaged using a single laser excitation wavelength<sup>40</sup>.



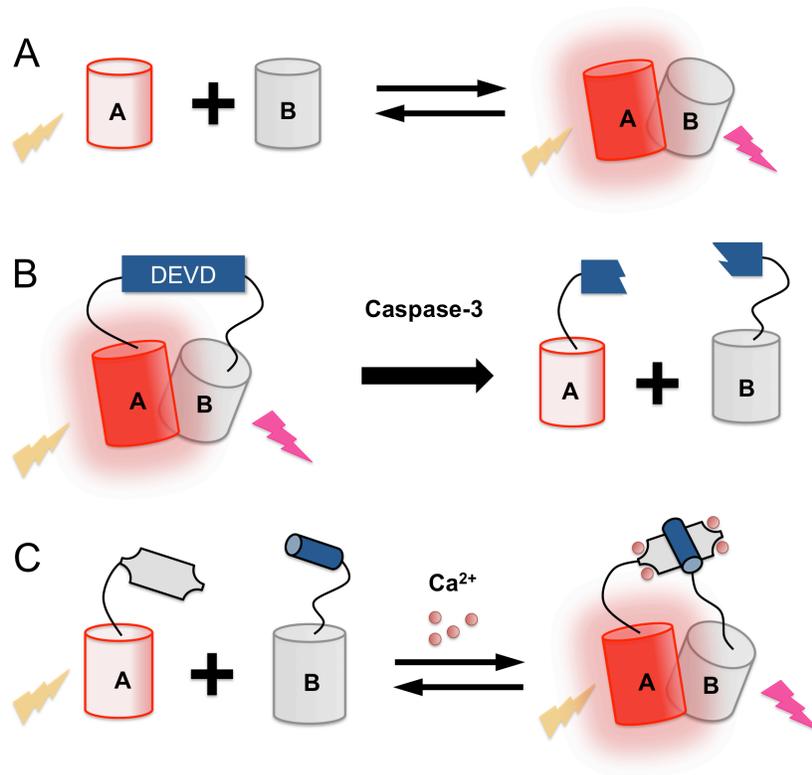
**Figure 1.6** Representative FRET-based biosensors with RFPs. A, Intermolecular biosensors for PPI. Unlike BiFC-based biosensors, the FRET pair does not interrupt the disassembly of the protein complex formed by their fusion partners. B, Ion/small molecule biosensors. An intramolecular protein complex is formed, or a conformation changed, upon the binding of certain ion or small molecule. C, Protease biosensors where the two FPs are linked by a protease substrate sequence.

### 1.2.3 RFP Biosensors based on ddFP

In an approach that is conceptually similar to BiFC, a novel strategy for converting FPs into reporters based on ddFPs has been reported<sup>64, 65</sup>. A ddFP is a

pair of engineered FPs that exhibits minimal fluorescence in their monomeric state, but exhibits bright fluorescence upon dimerization. This strategy is similar to BiFC as a functional FP is being reconstituted from its parts. However, the key difference between ddFPs and BiFC is that, in the case of ddFP, the functional FP is a dimer and the interaction is reversible. ddFP-based biosensor design strategy also share some similarities with the ones based on FRET, since two FPs are required, and the signal output reflects the relative proximity between these FP partners. Again, the key difference between the FRET and ddFP approach is that ddFP retains substantial affinity for each other and will spontaneously dimerize at relatively high concentrations (above 10  $\mu\text{M}$ ). Although some FRET pairs have a weak tendency to dimerize, the dissociation constants are typically much higher ( $<100 \mu\text{M}$ )<sup>63</sup>.

DdFPs are a version of dimeric RFP dTomato that was subjected to an extensive engineering process involving both rational design and directed evolution. This process led to the development of ddRFP, which demonstrates a 10-fold increase in red fluorescence intensity upon heterodimer formation. This dimerization-dependent fluorescence change of ddRFP was used for detection of reversible  $\text{Ca}^{2+}$ -dependent association of calmodulin and M13 in live cells, as well as imaging of caspase-3 activity during apoptosis (Figure 1.7)<sup>65</sup>.



**Figure 1.7** DdRFP- and ddRFP-based biosensors. A, Fluorescence intensity increase upon the formation of heterodimeric ddRFP pair. B, DdRFP-based caspase-3 sensor. C, DdRFP-based  $\text{Ca}^{2+}$  sensor.

#### 1.2.4 RFP Biosensors based on Single FP

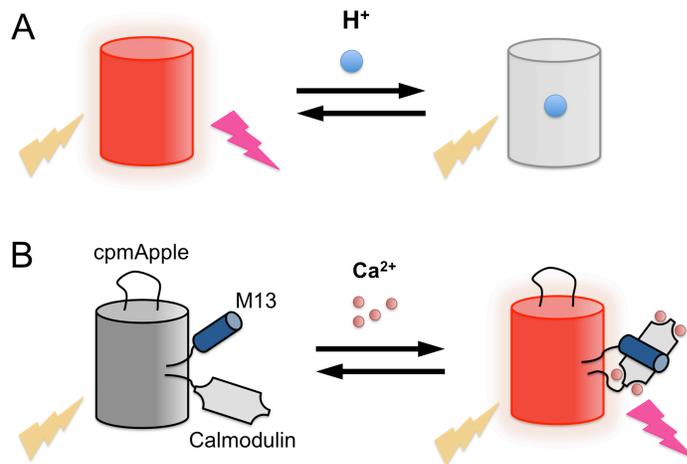
Single FP-based biosensors contain only one engineered FP signal-transducing domain. The goal of the creation of such a biosensor is to engineer this single FP to respond to the biochemical stimulus of interest with a reversible change in fluorescent intensity (intensiometric), or excitation spectral profile (excitation ratiometric), or emission spectral profile (emission ratiometric). The advantage of

single FP-based biosensors with fluorescence intensity change is that there is a typically a much larger intensity change at a single wavelength when compared to a FRET-based biosensor. Furthermore, single FP-based sensors also have the benefit of using up a smaller region of the visible spectrum window, possibly enabling the use of more colours of fluorophore simultaneously. Yet another advantage is a smaller protein size.

One way to create a single FP-based biosensor is to take advantage of the intrinsic sensitivities of certain FP variants (Figure 1.8A). A classic example is to take advantage of the inherent pH sensitivity of FPs to create pH sensors. All FPs exhibit some pH dependence, and some variants happen to have apparent  $pK_a$ s that are close to physiologically relevant pH values. Among the many examples of such FP-based pH sensors, the most widely used are the pHlourin variants of avGFP<sup>66</sup>. The chloride ion sensitivity of YFP is another example of an intrinsic FP sensitivity<sup>67, 68</sup>. It is also proven possible to rationally engineer intrinsic sensitivity into a FP by incorporating an analyte binding site directly on the exterior of the FP barrel. For example, the reduction/oxidation sensitive roGFP and the  $Ca^{2+}$  sensitive CatchER were engineered in this way<sup>69-71</sup>.

The majority of intrinsic single FP biosensors are green or yellow and only a few intrinsic red single-FP sensors have been described. The first engineered single RFP biosensor with intrinsic sensitivity was the pH-sensitive mNectarine, which

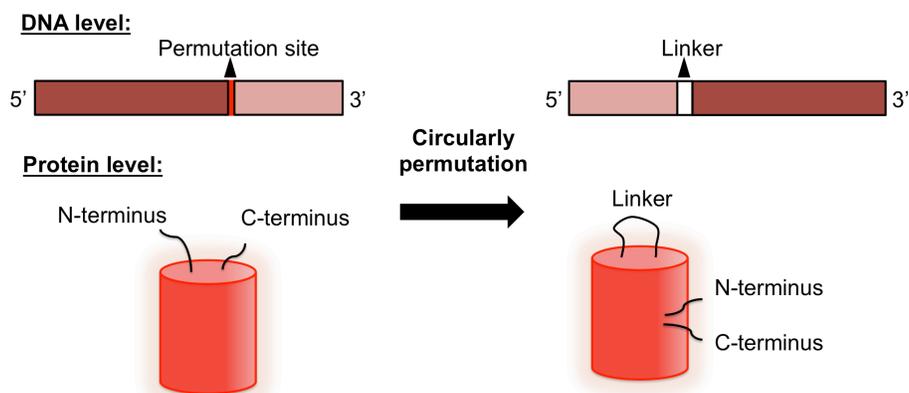
was applied for detection of nucleoside transport<sup>72</sup>. Another pH-sensitive RFP, pHTomato, was developed to report synaptic neurotransmitter release at single nerve terminal<sup>73</sup>. Yet another example is pHRed, which is an excitation ratiometric pH sensor engineered from LSS RFP mKeima to image energy-dependent changes of cytosolic and mitochondrial pH<sup>74</sup>.



**Figure 1.8** Single FP-based sensors. A, Single FP-based pH sensor based on intrinsic sensitivity. B, Single FP-based  $Ca^{2+}$  sensor with an extrinsic  $Ca^{2+}$  binding domain.

Another type of single FP-based biosensor are those that have a genetically incorporated extrinsic analyte recognition domain inserted into the FP. The extrinsic domain is typically fused to one of the termini, or inserted into an exposed loop of the FP, to minimize disruption of the overall FP structure.

However, for the FP to work as an effective signal transducer, the extrinsic recognition domain must be in close proximity to the chromophore in order for it to modulate the chromophore environment and induce a fluorescence change. Unfortunately, the FP chromophore is well protected in the center of the barrel structure, and the termini and loop region are relatively distant from the chromophore. To circumvent this problem researchers rely on a strategy termed circularly permutation. (Figure 1.9) Circularly permuted FPs (cpFP) are generated by genetically linking the original N- and C- termini with a short polypeptide linker and introducing new N- and C- termini at a position elsewhere in the protein<sup>75, 76</sup>. This manipulation is done at the DNA level, thus maintaining the fully genetically encoded property. For the purpose of FP biosensor construction, the new N- and C- termini are introduced close to the chromophore. Any ligand or physiological change induced protein conformational changes in the extrinsic recognition element could now result in alteration of the chromophore environment, thereby changing fluorescence of the cpFP<sup>77-80</sup>.



**Figure 1.9** Schematic presentation of FP circularly permutation at both the DNA and protein levels.

The prototypical biosensor of this design is the  $\text{Ca}^{2+}$  sensor known as GCaMP<sup>81, 82</sup>. In GCaMP, M13 and calmodulin (CaM) were fused to the N- and C- termini of cpGFP. Structural studies reveal that in its  $\text{Ca}^{2+}$  free state, the fluorescence is quenched because the chromophore is exposed to the bulk solvent. In the presence of  $\text{Ca}^{2+}$ , CaM wraps around M13 to form a compact structure, thus blocking the solvent access and restoring the chromophore to its fluorescing state. Based on the same rationale, a red  $\text{Ca}^{2+}$  biosensor known as R-GECO1<sup>83</sup> was created by replacing the cpEGFP in an improved GCaMP variant with a circularly permuted variant of mApple (i.e., cpmApple146-145). R-GECO1 was further optimized and engineered to into spectrally diversified and low-affinity variants, including an improved R-GECO1.2, a blue-shifted O-GECO, a red-shifted CAR-GECO<sup>84</sup>, a long Stokes shift REX-GECO (Jiahu Wu, unpublished results) and low-affinity

red (LAR)-GECO variants<sup>85</sup>. RCaMP, a similar single RFP-based Ca<sup>2+</sup> sensor was engineered from the cpmRuby template<sup>86</sup>.

While single RFP-based Ca<sup>2+</sup> sensors have received the most attention to date, efforts to create a broader selection of red fluorescent sensors are increasing. For example, our group has developed an RFP-based glutamate indicator using the cpmApple from R-GECO1, fused with a glutamate-binding domain (Jiahui Wu, unpublished results). Using the same cpmApple domain, we have developed a red fluorescent voltage sensor, FlicR1 (Ahmed Abdelfattah, unpublished results). FlicR1 exhibits voltage-sensing properties that are comparable to the best of available green fluorescent voltage indicators<sup>79, 87</sup>.

### **1.3 Methods for engineering of RFPs and RFP-based biosensors**

Protein engineering is the creation of proteins with new or improved functions by altering the genetic sequence of the protein using recombinant DNA technology. Changes in the gene sequence lead to corresponding changes in the amino acid sequence. Protein engineering is greatly facilitated by the relatively advanced state of current molecular biology technology, computer aided protein design, protein structure determination methods, and high-throughput screening techniques. Advances in all of these areas have led to protein engineering

becoming increasingly important and widespread for the creation of new molecular tools.

Arguably, engineering of RFPs has been one of the most fruitful chapters in the book of protein engineering. As described in the earlier parts of this chapter, such efforts have led to the creation of a large number of colourful and bright RFPs and RFP-based biosensors. The two strategies that have proven most useful for engineering of RFPs and RFP-based biosensors are rational design and directed evolution. In a typical RFP or RFP-based biosensor engineering process, primary clones with desired phenotype are usually obtained using rational design, and then directed evolution comes into play for further optimization.

### **1.3.1 Rational, semi-rational, and computer-aided design**

Rational design is a knowledge-based, deterministic engineering of proteins with desired properties or novel characteristics. Rational design can be performed at the amino acid level or protein subdomain level. To engineer RFP and RFP-based biosensors, changes at the amino acid level are often made using the information gained from protein structure or computational modeling. This rational design approach requires the use of site-directed mutagenesis, which enables the introduction of specific mutations into the target gene. There are two common methods for site-directed mutagenesis. One is called the “overlap extension”

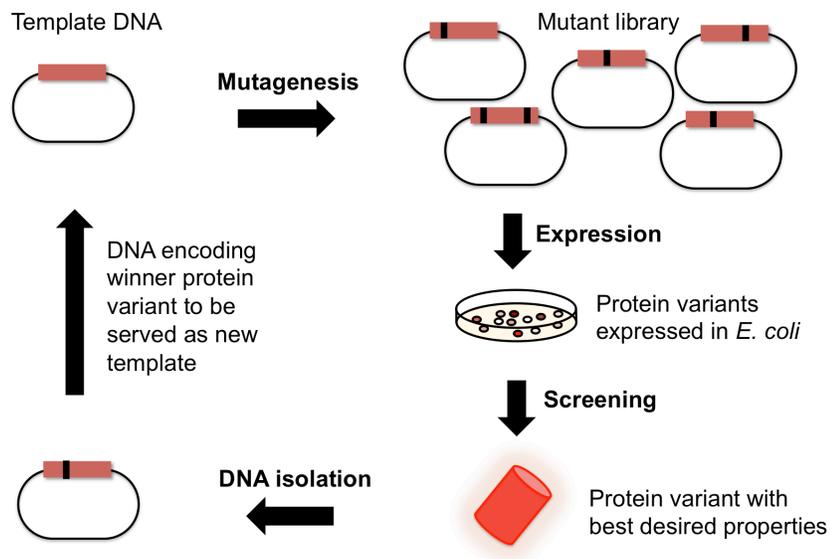
method<sup>88</sup>. This method involves two primer pairs, where one primer of each primer pair contains the desired mutation. These four primers are used in the first polymerase chain reaction (PCR), where two PCRs take place, and two double-stranded DNA products are obtained. Upon denaturation and annealing of them, two heteroduplexes are formed, and each strand of the heteroduplex includes the desired mutant codon. DNA polymerase is then used to fill in the overlapping 3' and 5' ends of each heteroduplex and a second PCR is performed to amplify the new mutated DNA template. The other site-directed mutagenesis method is called “whole plasmid single round PCR”, which is also known commercially as “Quikchange”<sup>89</sup>. It requires two oligonucleotide primers with the desired mutations, which are complementary to the opposite strands of a double-stranded DNA plasmid template. Both strands of the template are replicated using DNA polymerase, without displacing the primers and a mutated plasmid is obtained with breaks that do not overlap. DpnI methylase is then used for selective digestion of the original non-mutated templates, leaving circular and nicked vector with the mutant genes. Upon transformation of the nicked vector into competent cells, the nick in the DNA is repaired, and a circular, mutated plasmid is obtained. A common variation of site-directed mutagenesis is “saturation mutagenesis”. It involves the replacement of a single amino acid within a protein with each of the natural amino acids using the degenerate codon NNK<sup>90</sup>, thus providing all possible amino acid variations at the site.

Circularly permutation is a common strategy in rational design for RFP-based biosensors at the subdomain level (Figure 1.9). Advances in molecular biology have enabled researchers to construct circularly permuted proteins in the laboratory using different methods. When the desired site of permutation is known in advance, fragment PCR followed by overlap-extension PCR can be used. Another method is the use of concatemers, whereby a tandem repeat of the gene is amplified or cloned into plasmid DNA. Using specifically designed primers the desired circularly permuted gene can be amplified<sup>76, 91</sup>.

### **1.3.2 Directed evolution**

Rational design is an effective approach when the structure of the protein of interest is known and the structure-function relationships are well understood. In many cases of protein engineering, however, there is limited amount of information on the structure and mechanisms of the protein of interest. Moreover, the output of rational design often times results in unexpected loss of desired functionality due to the complexity of protein structure and function relationships. Thus, the use of an “irrational” method, which involves random mutagenesis and empirical screening of protein properties, was introduced as a different approach. This approach is known as directed evolution<sup>92</sup>, as it is inspired by the natural process of evolution but with the artificial selection criteria set by the protein engineer. A typical directed protein evolution strategy involves repeated rounds of

the following steps: generation of a gene library with random diversification; expression of the gene library; screening for mutants with an improved property; isolation the DNA of best candidates to be used as templates in the next round(Figure 1.10). These procedures are usually repeated until satisfactory mutants are acquired.



**Figure 1.10** Schematic representation of the process of directed evolution.

The first step in each round of molecular evolution is the generation of a large number ( $>10^5$ ) of mutant genes. DNA diversification for directed evolution can be achieved by *in vitro* random mutagenesis and gene shuffling. Both of these methods are capable of creating libraries of great genetic complexity, which potentially increase the probability to find variants with the desired property

comparing to site-directed mutagenesis. Libraries of random mutagenesis are performed on a gene template (or pool of gene templates) by amplifying genes using PCR under conditions where the polymerase has a high error rate. Such conditions include using low fidelity polymerases in buffer with high  $Mg^{2+}$  concentration,  $Mn^{2+}$ , and unbalanced dNTPs. The error rate of the reaction can typically be adjusted through the change of buffer conditions<sup>93</sup>. However, the size of the library is constrained by the fact that beneficial mutants become ever more rare with increased mutation rate due to the much greater likelihood of acquiring a detrimental mutation than a beneficial one. Thus, directed evolution depends on a large number of evolution cycles with a gradual accumulation of small improvements from each cycle. Gene shuffling is another powerful strategy to gain beneficial mutations by assembling hybrid genes from several different gene templates. Commonly used template sequences can be a group of naturally occurring homologous genes or a mixture of mutagenized genes derived from a single progenitor. Many techniques of gene shuffling have been developed and used in directed evolution, among which, staggered extension PCR (StEP) is the most widely used methods<sup>94</sup>. StEP utilizes full-length sequences as templates, which simplify the procedure without sacrificing the effective accumulation of mutations. Specifically, the hybrid events in StEP are created by the staggered DNA fragments, which are generated in a modified PCR with extremely short annealing and extension steps.

Directed evolution with screening in the context of bacterial colonies, has been effectively used to develop many new red FPs. For bacterial screening, it is relatively convenient to create genetic libraries in vitro, and then introduce them into the bacteria using routine molecular biology methods. For screening in eukaryotic systems, it is substantially more difficult to create large libraries of cells expressing libraries of genes generated in vitro. One strategy for creating libraries in eukaryotic systems is to generate the library in the cells. This can be achieved using the process of somatic hypermutation that occurs naturally in B-lymphocytes. This process is normally used by cells to introduce random point mutations into antibody encoding genes for the purpose of affinity maturation. This approach has been exploited to generate mutations in non-antibody proteins, such as RFP, for the development RFPs variants with red shifted excitation and emission<sup>29</sup>.

The usefulness of directed evolution lies in the fact that detailed structural and mechanistic information is not required for the strategy to be successful. The only requirements for directed evolution are the creation of large libraries of random mutants as well as a suitable and robust screening scheme that selects for the desired protein properties. As the random mutagenesis methods are relatively universal and generalizable, screening method development is now seen as the most critical component of a directed evolution strategy for development of a RFP and RFP-based biosensors.

The predominant method of screening red FPs is to express the protein in *Escherichia coli* and screen the fluorescent colonies on agar plates, using a colony fluorescence imaging system. The characteristic of RFPs that is most commonly optimized in this way is the brightness, which could be increased by improved maturation efficiency, improved folding efficiency, or increased quantum yield or extinction coefficient. Using a custom-built camera system, images of plates are taken with the appropriate excitation and emission filters allowing bright colonies to be differentiated from the dim ones<sup>95</sup>. Bacterial colony screening could be used to examine up to 10s of thousands of colonies in a single day. It also provides the convenience of routine isolation of the candidate protein gene from single colonies through colony picking and culturing. Use of a fluorescence-activated cell sorter is another important screening approach for bacterial libraries of RFP mutants<sup>29, 40, 48, 96</sup>. However, this approach requires a much more expensive instrumentation system, but has the advantage of higher throughput compared to the colony screening method<sup>97</sup>.

For screening RFP-based sensors, the colony-based screening method is still applicable when used in conjunction with a periplasmic export expression system. This system directs the protein sensor candidates to the periplasm of bacterial cells where the target analyte concentration can be controlled<sup>83, 84</sup>. A more common method for screening RFP-based sensors is to test the analyte induced

fluorescence response of the extracted protein sensor solution on a 96-well or 384-well plate using a fluorescence plate reader. This method offers higher accuracy over colony-based screening. However, it is considerably more labour intensive, and the throughput is rather limited<sup>73</sup>. Direct screening of sensor function in a live cell is an emerging trend for developing advanced RFP-based sensors for cell biology and neuroscience research<sup>98, 99</sup>. For example, a neuron-based screening platform was reported with a number of advantages compared with solution measurements<sup>100, 101</sup>. Neuronal screening was able to identify variants with improved response kinetics and large fluorescence signal change under physiologically meaningful conditions.

The power of directed evolution is well established for engineering of RFP and RFP-based biosensors, but still suffers from being highly labour intensive and the requirement of an effective screening protocol. In contrast, rational design requires a relative small workload due to its specificity. However, in practice, rational designs are often not as optimal as expected. Accordingly, protein engineering has significantly benefited from the thoughtful combination of both rational design and directed evolution. This combination has been, and also will continue to be, successful in yielding new and improved RFPs and RFP-based biosensors.

#### **1.4 Scope of this thesis**

In this thesis, we describe our efforts to expand the existing repertoire of RFP and RFP-based biosensors using protein engineering strategies including directed evolution and rational design. Chapter 2 and 3 describe the development of new RFP-based pH sensors. Using monomeric RFP mOrange and mApple as templates, we engineered intensimetric pH sensors pHoran and pHuji, and excitation ratiometric red pH sensors named pHlorina1 and pHlorina2. In Chapter 4, we describe the discovery and characterization of a novel photochromic and thermochromic RFP. Combining with a mutational study, we proposed a molecular mechanism for the chromism of the new RFP. Chapter 5 describe the engineering of a long Stokes shift and a red-shifted RFP variant based on mCherry. We characterized the spectral properties of these new RFPs and compared them with the existing ones. Chapter 6 describes a new method for screening FP-based biosensors through colony blotting on filter paper. We also demonstrate the use of this method in the development of a series of engineered Ca<sup>2+</sup> sensitive RFPs. In the last chapter, we provide a brief summary of this thesis, as well as a brief discussion on the future developments in the field of engineering RFP and RFP-based biosensors.

## **Chapter 2 Red fluorescent protein-based intensimetric pH sensors**

## 2.1 Introduction

Genetically encoded sensors based on fluorescent proteins (FPs) have become essential tools for studying cell physiology. A broad range of FP-based sensors have now been developed to monitor numerous biochemical parameters including  $\text{Ca}^{2+}$  concentration<sup>83</sup>, membrane potential<sup>45</sup>,  $\text{Cl}^-$  concentration<sup>67</sup>, and pH<sup>66, 102</sup>. For studies of exocytosis and endocytosis, the sensor of choice has been superecliptic pHluorin (SEP), a mutant of enhanced green FP (EGFP)<sup>66, 102</sup>. With a  $pK_a$  (i.e., the pH value at which the fluorescence intensity is 50% of maximal) of 7.2, SEP is nearly non-fluorescent at pH 5.5 (the pH of intracellular secretory vesicles and recycling endosomes) but brightly green fluorescent at pH 7.4 (the extracellular pH). The SEP protein, fused to relevant membrane proteins, has been extensively used to detect the exocytosis of synaptic vesicles, secretory vesicles, and recycling endosomes<sup>102-107</sup>. Moreover, it has been used to detect the formation of clathrin coated vesicles (CCVs) using the pulsed pH (ppH) protocol (i.e., alternating the extracellular pH between 7.4 and 5.5), which reveals the location of receptors that have been newly internalized with high temporal accuracy<sup>108</sup>.

A red FP (RFP) with pH dependence similar to SEP would be of great utility because it would enable two colour imaging of membrane trafficking events such as exocytosis and endocytosis. For example, a red FP pH sensor could be imaged in conjunction with a GFP tagged protein of interest in order to analyze the spatial and temporal dynamics of protein recruitment during membrane trafficking<sup>109, 110</sup>.

Alternatively, the trafficking of two membrane cargos could be monitored simultaneously to address their differential sorting. A number of pH-sensitive orange to red *Discosoma* red FP-derived variants have been previously reported, with  $pK_a$ s ranging from 6.5 for some proteins of the mFruit series (i.e, mOrange, mOrange2, and mApple)<sup>12, 13</sup>, to 6.9 for mNectarine<sup>72</sup>, and 7.8 for pHTomato<sup>73</sup>. In addition, an excitation ratiometric pH sensor with an apparent  $pK_a$  of 6.6, designated pHRed, was engineered from the long Stokes shift FP mKeima<sup>74</sup>.

For practical applications, none of the previously reported pH-sensitive red FPs provide *in situ* pH-dependent changes in fluorescence intensity that are comparable to SEP. To the best of our knowledge, no previous report has directly addressed the question of why these red pH sensors have not proven to be as useful as SEP, despite the substantial effort invested in their development. In this work we address this important question and, based on the insights we obtain, engineer a new pH-sensitive red FP. We report pHuji, a new red fluorescent pH sensor that can be used for imaging of exocytic and endocytic events.

## **2.2 Results**

### **2.2.1 Characterization of SEP and previously reported red pH sensors**

*In vitro* characterization of the pH dependence of fluorescence demonstrates that SEP is exceptionally well tuned for large fluorescence changes when transitioning

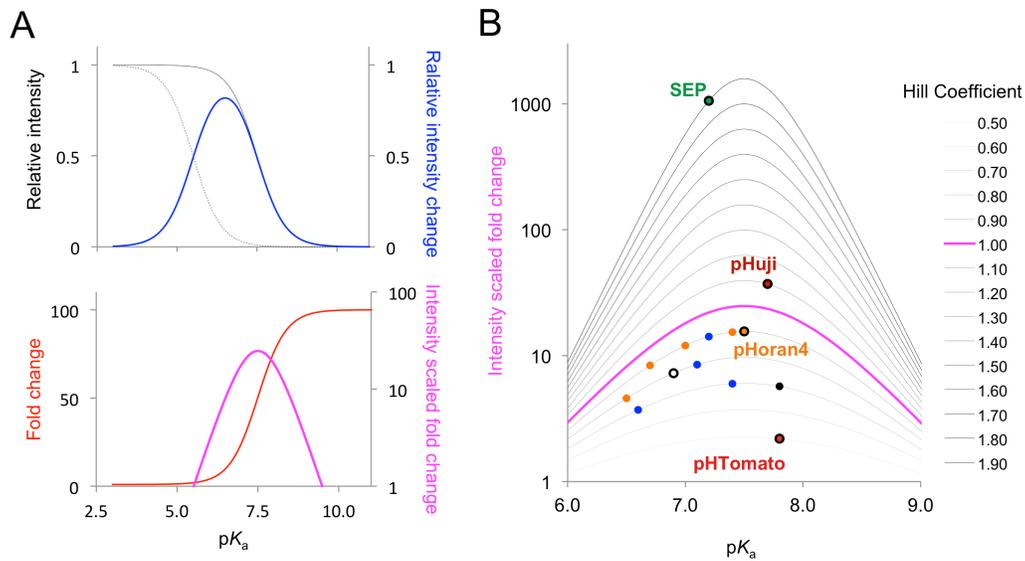
between pH 5.5 and 7.5. SEP is almost entirely quenched at pH 5.5 and exhibits an approximately 50-fold higher fluorescence intensity at pH 7.5 (Table 2.1). In contrast, mNectarine and pHTomato only have approximately 6-fold and 3-fold change, respectively, over the same pH range, despite their relatively high  $pK_a$ s of 7.9 and 6.9, respectively (Table 2.1). This result demonstrates that  $pK_a$  alone is not sufficient to characterize the pH-sensitive FPs.

**Table 2.1** *In vitro* characterization of pH-sensitive FPs

Protein	$pK_a$	$n_H$	Fluorescence		Excitation	Emission
			fold	change	peak at pH	peak at pH
			(pH 5.5-7.5)		7.2 (nm)	7.2 (nm)
SE-pHluorin	7.2	1.90	50		495	512
mNectarine	6.9	0.78	6		558	578
pHTomato	7.8	0.51	3		562	578
mOrange	6.5	0.77	5		548	562
pHoran1	6.7	0.87	10		547	564
pHoran2	7.0	0.89	12		549	563
pHoran3	7.4	0.87	15		551	566
pHoran4	7.5	0.92	17		547	561
mCherry-TYG	7.8	0.73	5		546	568
mApple	6.6	0.68	4		568	592
A-9	7.4	0.73	12		576	596
A-17	7.2	0.87	11		576	596
A-47	7.1	0.83	10		570	592
pHuji	7.7	1.10	22		566	598

\* Biphasic titration curve. To obtain the apparent  $n_H$  the whole titration curve was fit with a monophasic function.

To find the optimal parameters for maximized fluorescence change between pH 5.5 and 7.5, we modeled the theoretical fluorescence changes between these two pH values assuming various  $pK_a$  values. These calculations revealed that the fluorescence fold change ( $F_{7.5}/F_{5.5}$ ) increases with increasing  $pK_a$  (Figure 1A, red curve). However, a higher  $pK_a$  also results in lower brightness of the protein in the physiological pH range (Figure 2.1A, black curve). To account for this limitation, we took the relative fluorescence intensity change  $((F_{7.5}-F_{5.5})/F_{\max})$  into consideration (Figure 2.1A, blue curve). This value is maximal when  $pK_a = 6.5$ . Multiplication of the functions for fold change and relative change leads to a new function, termed as intensity scaled fold change, which represents the optimal compromise of both fluorescence fold change and overall brightness (Figure 2.1A, magenta curve). This calculation revealed that the optimal theoretical  $pK_a$  value of a pH-sensitive FP is 7.5.

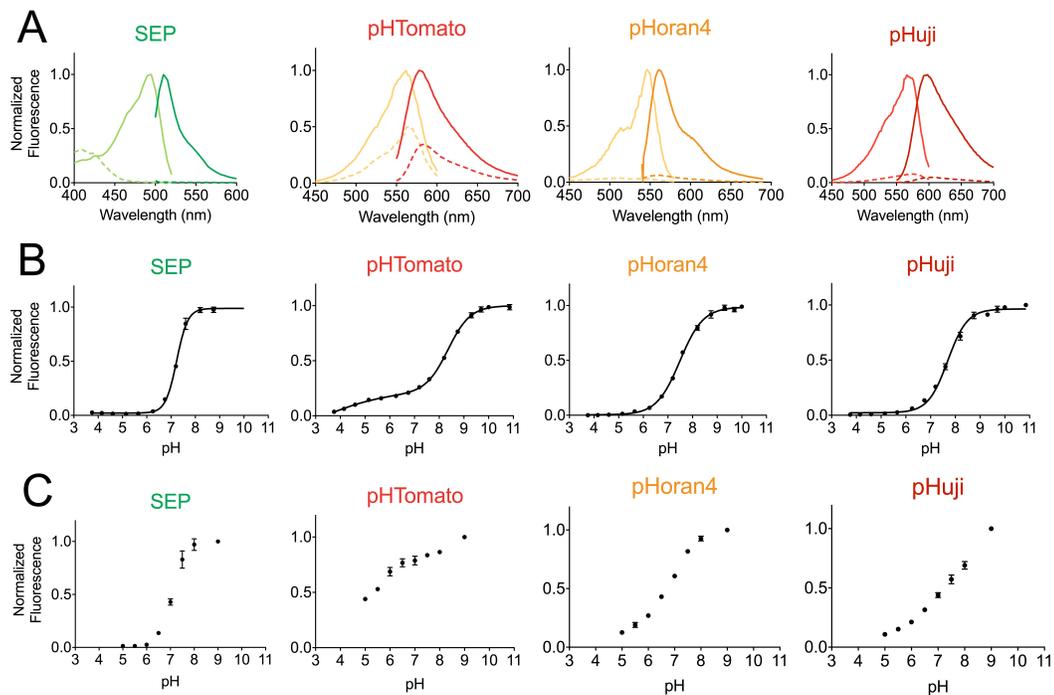


**Figure 2.1** Theoretical optimization of pH sensor properties. A, theoretical calculation of optimal  $pK_a$  for the largest fluorescence change from 5.5 to 7.5, with a  $n_H$  of 1.0. Black solid line, relative fluorescence at pH 7.5 ( $F_{7.5}/F_{max}$ ); black dashed line, relative fluorescence at pH 5.5 ( $F_{5.5}/F_{max}$ ); red line, fluorescence fold change ( $F_{7.5}/F_{5.5}$ ); blue line, relative fluorescence intensity change ( $(F_{7.5} - F_{5.5})/F_{max}$ ); magenta line, intensity scaled fold change, which is the multiplication product of fold change and intensity change. B, effect of  $n_H$  on the intensity scaled fold change at different  $pK_a$  values. The pH-sensitive FPs covered in this work are mapped according to  $pK_a$  and  $n_H$  values. From top to bottom ( $n_H$  descending), right to left ( $pK_a$  descending): SEP; pHuji; pHoran4 (orange, same as mOrange and mOrange variants), pHoran3, A-17 (blue, same as mApple and mApple variants), pHoran2, pHoran1; A-47, mNectarine (white), mOrange; mCherry-TYG (black), A-9, mApple; pHTomato.

While SEP, with a  $pK_a$  of 7.2, is the closest to this ideal value, this factor alone is not sufficient to explain the superior performance of SEP relative to the other variants. It quickly became apparent that an equally important factor in determining the performance of these pH indicators is the slope of the fluorescence versus pH curves, which is determined by the apparent Hill coefficient ( $n_H$ ). Indeed, further modeling revealed that a higher  $n_H$  provides a substantially larger fluorescence change without changing the optimal  $pK_a$  value (Figure 2.1B). Curve fitting revealed that SEP has an exceptionally high  $n_H$  of 1.90. Overall, the excellent performance of SEP is due to the combination of a nearly optimal  $pK_a$  and a high  $n_H$ . In contrast, the  $n_H$ s of mNectarine and pHTomato are 0.78 and 0.51 respectively (Table 2.1), which largely explains the relatively modest changes in fluorescence for these proteins between pH 5.5 and 7.5. Moreover, the emission peaks of these red FPs have more than a 15 nm blue shift over the pH range of 5 to 10, while SEP, in contrast, has a negligible peak shift (Table 2.2, Figure 2.2A). These limitations prompted us to look for red FP pH sensors with optimized  $pK_a$  and  $n_H$  values as well as good optical characteristics such as a red-shifted emission with little pH dependent peak shift.

**Table 2.2** Emission peak (nm) at different pH values

<b>Protein</b>	<b>pH = 5.0</b>	<b>pH = 7.2</b>	<b>pH = 10.0</b>	<b>Peak shift from pH 5 to 10</b>
SE-pHluorin	509	512	512	3
pHoran4	558	561	562	4
pHuji	600	598	598	2
mNectarine	584	576	568	<b>16</b>
pHTomato	580	568	562	<b>18</b>
mCherry-TYG	578	578	562	<b>16</b>



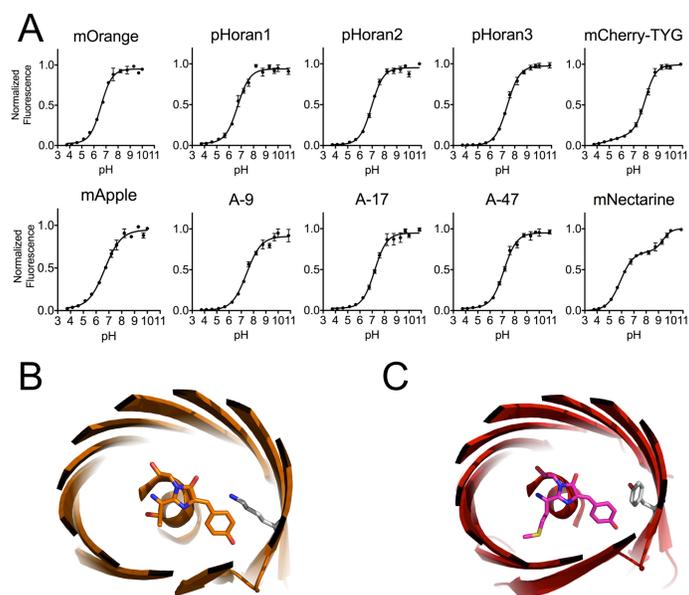
**Figure 2.2** Spectra and pH titration curves of pH-sensitive FPs. A, excitation and emission spectra of pH-sensitive FPs SEP, pHTomato, pHoran4, and pHuji at pH 5.5 (dashed line) and pH 7.5 (solid line). B, pH titration curves of the indicated FP. C, fluorescence of HeLa cells expressing pDisplay proteins fused to the indicated FP as a function of pH. Error bars represent SEM.

### 2.2.2 pHorans: pH-sensitive mOrange variants with fine-tuned $pK_a$ s

As a first step towards developing a second colour of pH sensor with optimized  $pK_a$  and  $n_H$ , we created a series of orange FPs with  $pK_a$ s that approached the ideal value of 7.5. To achieve this goal, we started with mOrange, a bright FP with a  $pK_a$  at 6.5<sup>12</sup>. The previously reported M163K mutation was introduced to raise the

$pK_a$  to 7.5<sup>13</sup>. Using mOrange M163K as a template we performed 4 rounds of random mutagenesis followed by saturation mutagenesis at residue 163 and 161, where beneficial mutations had repeatedly showed up during screening of randomly generated libraries. Both of these residues are in close proximity to the chromophore. Colonies that retained high fluorescent brightness were cultured and the fluorescent brightness of extracted protein was measured at pH 5.5 and 7.5. Two specific mutations, E160K and G196D, which are known to increase the photostability of mOrange<sup>13</sup>, were also rationally introduced during the course of development. Library screening led to the identification of 3 bright variants with  $pK_a$ s of 6.7, 7.0, and 7.4 (Table 2.3). Along with the template mOrange M163K, this series of the pH-sensitive mOrange variants was named as pHorans (pH-sensitive orange FPs) 1 to 4 according to their  $pK_a$  values in ascending order (Table 2.1). Despite the fact that pHoran4 is remarkably bright (extinction coefficient 83,000 M<sup>-1</sup>cm<sup>-1</sup> and quantum yield 0.66) and has a  $pK_a$  value close to SEP, its fluorescence intensity fold change from pH 5.5 to 7.5 is still significantly smaller than that of SEP (Table 2.1 and Figure 2.2B) due to its lower  $n_H$  value of 0.92. Unfortunately, pHoran variants with higher  $n_H$  values were not identified during exhaustive library screening. Yet another drawback of the pHorans is that their orange fluorescent emission spectra (with emission peaks close to 560 nm) result in substantial spectral bleedthrough when combined with SEP for 2-colour imaging.

We also explored the modification the chromophore structure of the red FP mCherry<sup>12</sup> in the effort to identify a pH sensor with an emission peak well separated from that of SEP. We found that a variant with the threonine-tyrosine-glycine (TYG) chromophore, designated as mCherry-TYG, has a significantly higher  $pK_a$  of 7.8 and an apparent  $n_H$  of 0.73 (Table 2.1). Due to its pH-dependent emission peak shift (Table 2.2) and incomplete quenching at low pH (Figure 2.3A), this variant was not further pursued.



**Figure 2.3** Additional characterization of pH-sensitive RFPs. A, pH titration curves of additional pH-sensitive red FPs mentioned in this study. Error bars represent SEM. B, modeled location of the M163K mutation in pHOran4 (based on the crystal structure of mOrange, PDB: 2H50). The FP is viewed from the axis of the  $\beta$ -barrel. Secondary structures are shown in cartoon form. The chromophore and the mutated residue are in stick representation with carbon coloured orange in the chromophore and gray in mutated residue. Nitrogen is shown in blue and oxygen in red. C, modeled location of the K163Y mutation in pHuji (based on the crystal structure of mCherry, PDB: 2H5Q). The chromophore and the mutated residue are in stick representation with carbon coloured magenta in the chromophore and gray in the mutated residue. Nitrogen and oxygen are coloured as in (B).

### **2.2.3 pHuji: a red pH-sensitive FP with near-optimal $pK_a$ and a high apparent Hill coefficient**

For our next attempt at developing an improved red pH sensor with an emission peak well separated from that of SEP, we turned to mApple as a template<sup>13</sup>. Libraries were constructed by randomizing residues in close proximity to the chromophore, including positions 64, 70, 95, 97, 146, 148, 159, 161, 163, 177 and 197 (numbered according to mCherry crystal structure, PDB ID 2H5Q)<sup>12</sup>. After initial screening by comparing fluorescence of protein variants at pH 5.5 and 7.5, it was found that positions 161 and 163 were primarily responsible for modulating the  $pK_a$  of mApple. Accordingly we constructed a targeted library of variants by saturation mutagenesis of both positions 161 and 163 and screened exhaustively for variants with increased  $pK_a$  and higher  $n_H$ .

A number of pH-sensitive variants, including A-9, A-17, and A-47, were identified from the library (Tables 2.1 and 2.2). Among them, one variant with a single mutation of K163Y demonstrated the highest pH sensitivity with a  $pK_a$  of 7.7 and an apparent  $n_H$  of 1.10. Notably, all of the mApple variants as well as pHoran3 and pHoran4 have essential mutations at position 163, predicted from published mFruit structures<sup>111</sup> to be close to the chromophore (Figure 2.3 B-C), clearly demonstrating that this residue plays a key role in modulating the pH sensitivity of variants of the *Discosoma* red FP. Two rounds of random mutagenesis using mApple K163Y as the template did not lead to the

identification of variants with further improvements in pH sensitivity. Accordingly, mApple K163Y was designated as pHuji (pronounced same as Fuji, a common cultivar of apple). It exhibits relatively bright fluorescence with an extinction coefficient of 31,000 M<sup>-1</sup>cm<sup>-1</sup> and a quantum yield of 0.22. More importantly, pHuji demonstrates a more than 20-fold fluorescent intensity change from pH 5.5 to 7.5 which, to the best of our knowledge, is the largest reported intensity change for a red FP in this pH range (Table 2.1 and Figure 2.2B). This is in good accordance with the predictions we could make from its high n<sub>H</sub> and near-optimal pK<sub>a</sub> value (Figure 2.1B). Finally, pHuji has an emission peak (598 nm) that shows negligible pH dependent shift (Table 2.2) and is well separated from SEP emission (512 nm) (Table 2.1, Figure 2.2A). Simultaneous two-colour imaging is therefore possible with these two variants.

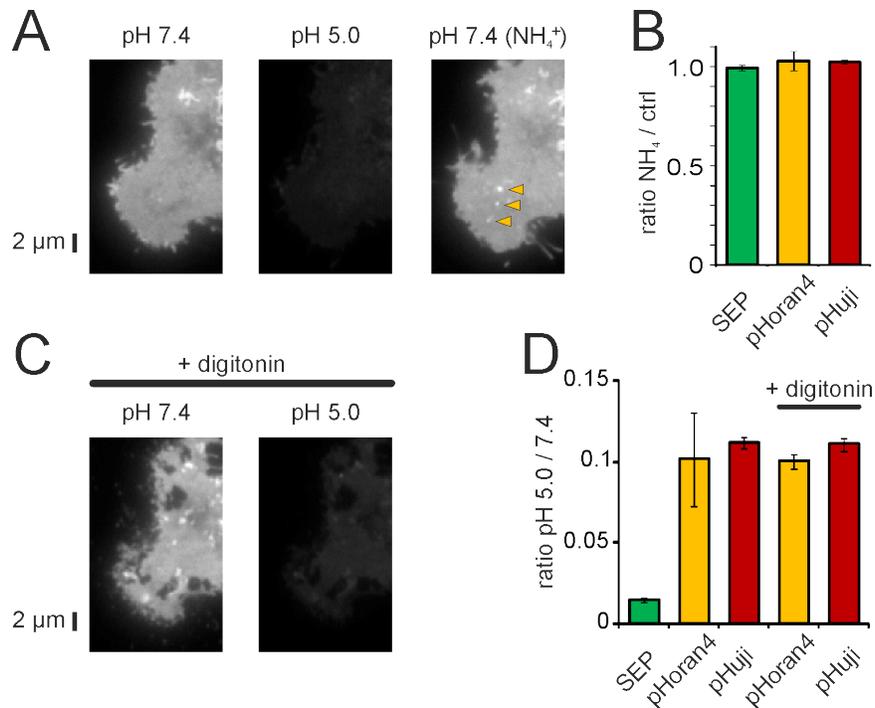
**Table 2.3** Mutations of the pH-sensitive red FPs developed in this study

<b>Protein</b>	<b>Template</b>	<b>Mutations</b>
pHoran1	mOrange	E160K/G196D/K184E/E218V
pHoran2	mOrange	N98S/T127S/E160K/I161V/G196D
pHoran3	mOrange	E160K/I161W/M163G/K184V/G196D/E218V
pHoran4	mOrange	M163K
A-9	mApple	I161G/K163H
A-17	mApple	I161G/K163F
A-47	mApple	I161G/K163V
pHuji	mApple	K163Y
mCherry-TYG	mCherry	M66T

#### **2.2.4 Characterization and comparison of pH-sensitive FPs *in vitro* and *in situ***

Based on the *in vitro* characterization, we decided to focus on comparing only our best candidates, pHuji and pHoran4, with previously described SEP and pHTomato in living cells. Accordingly, we transfected HeLa cells with the pH-sensitive red FP in the pDisplay vector to anchor them to the cell surface as a fusion to the transmembrane domain of platelet-derived growth factor receptor (PDGFR). These cells were imaged with TIRF microscopy in buffered solutions ranging from pH 5.0 to 8.9. Similar to the fluorescence measurements obtained for recombinant proteins, SEP fluorescence was completely quenched at pH 5.0, and maximal at pH 8.9, with a 63-fold change in fluorescence between pH 5.5 and 7.5. The pH sensitivity of red FPs in the same context was qualitatively similar to the *in vitro* data, though their pH sensitivity was somewhat less pronounced than expected. pHTomato was poorly pH-sensitive, with less than 2-fold change in fluorescence between pH 5.5 and 7.5. pHuji and pHoran4 were more strongly pH-sensitive with 5-fold change over the same pH range (Figure 2.2). However, in contrast to the *in vitro* data (Figure 2.2), pHuji and pHoran4 still retained dim fluorescence even at pH 5.0. This could either be due to environmental factors that result in different pH sensitivities *in vitro* compared to in cells, or to the presence of a visible intracellular pool of protein that is inaccessible to extracellular pH changes.

To test for the presence of an intracellular pool of protein that is resistant to extracellular pH changes, we first applied a solution containing ammonium chloride (50 mM) to cancel intracellular pH gradients<sup>66</sup>. This treatment revealed a few intracellular vesicles for all FPs but the overall diffuse fluorescence did not change (Figure 2.4A-B), demonstrating that this diffuse labelling is indeed due to proteins on the plasma membrane and not in mildly acidic intracellular compartments. However, the residual fluorescence at pH 5.0 could also be due to a neutral intracellular compartment. Thus, to collapse all ionic gradients, we permeabilized cells with digitonin (50  $\mu\text{g/ml}$ ). In these conditions, the fluorescence at pH 5.0 was still  $10.0 \pm 0.8\%$  ( $n = 3$ ) and  $11.0 \pm 0.8\%$  ( $n = 4$ ) of that at pH 7.4 for pHoran4 and pHuji, respectively, similar to unpermeabilized cells ( $10.1 \pm 2.9\%$  ( $n = 4$ ) and  $11.1 \pm 0.7\%$  ( $n = 5$ ); Figure 2.4C-D). Therefore, the pDisplay vector correctly anchored the FPs at the plasma membrane, but pHoran4 and pHuji are less sensitive to pH in cells than *in vitro*. Altogether, the strong pH sensitivity of pHuji and pHoran4 led us to further explore their utility as reporters of endocytosis and exocytosis.

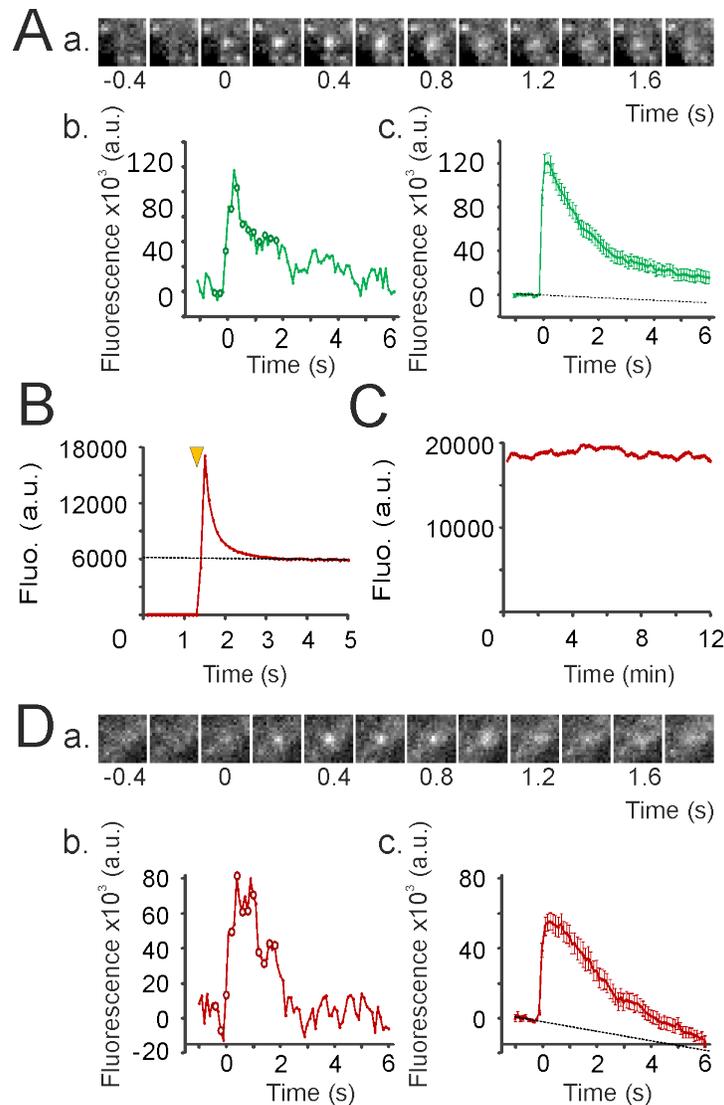


**Figure 2.4** Sensitivity of SEP, pHoran4 and pHuji in HeLa cells with collapsed pH gradients. A, TIRF images of a HeLa cell transfected with pDisplay-pHuji at pH 7.4, pH 5.0 and then pH 7.4 + NH<sub>4</sub>Cl (50 mM), which reveals intracellular vesicles that are normally acidic (yellow arrowheads) while diffuse fluorescence of proteins on the plasma membrane remains unchanged B, fluorescence ratio of cells transfected with the indicated pDisplay-FP, in NH<sub>4</sub><sup>+</sup> containing vs. control solution (n = 3-5 cells). The ratios do not differ from unity, showing that no acidic intracellular organelle contributes to pDisplay-FP fluorescence. C, images at pH 7.4 and then pH 5.0 of the same cell as in A after 2 min incubation with digitonin (50 μg/ml). Note the presence of bright puncta which suggest that intracellular organelles were also permeabilized by digitonin and hence at the same pH as the bath solution. At pH 5.0, similar to untreated cells, pHuji fluorescence is largely, but not completely, decreased. D, Fluorescence ratios of cells at pH 5.0 vs pH 7.4 in control conditions or in the presence of digitonin (n = 3-4 cells). The fluorescence ratios were not

significantly different in control and in digitonin ( $p > 0.85$ ). Error bars represent SEM.

### **2.2.5 Detection of single exocytic and endocytic events with TfR-red FPs**

The transferrin receptor (TfR) has been widely used to detect single endocytosis<sup>108</sup> and exocytosis events of recycling endosomes<sup>112, 113</sup>. Indeed, in cells transfected with TfR-SEP, exocytic events are readily detected in continuous TIRF recordings as sudden bursts of fluorescence as receptors go from an acidic intravesicular to the neutral extracellular compartment, and then rapidly diffuse from the site of exocytosis (Figure 2.5A). However, we noticed that pHuji, like earlier versions of its mApple template<sup>13</sup>, has a pronounced photoswitching behaviour, with fluorescence going down to ~35% of initial fluorescence in less than 2s of continuous illumination in our recording conditions (Figure 2.5B). Nevertheless, this behaviour did not impair the possibility to obtain stable time-lapse recording conditions to study endocytosis (Figure 2.5C and see below), and we could record exocytic events with similar profiles as those recorded with TfR-SEP (Figure 2.5D). Conversely, pHoran4 or pHTomato did not have a measureable photoswitching behaviour, and could also be used to detect exocytosis as shown previously<sup>113, 114</sup>.



**Figure 2.5** Detection of exocytosis events with TfR-pHuji. A, Detection of exocytosis events with TfR-SEP (a) Example of an exocytosis event recorded at 10 Hz. Note the sudden appearance of fluorescence and its diffusion away from the exocytosis site in a few frames. (b) Quantification of the fluorescence of the event shown in (a). A background image (average of 5 frames before the event) was subtracted before quantification. Circles correspond to images in (a). (c) Average of 56 events quantified as in (b) in 3

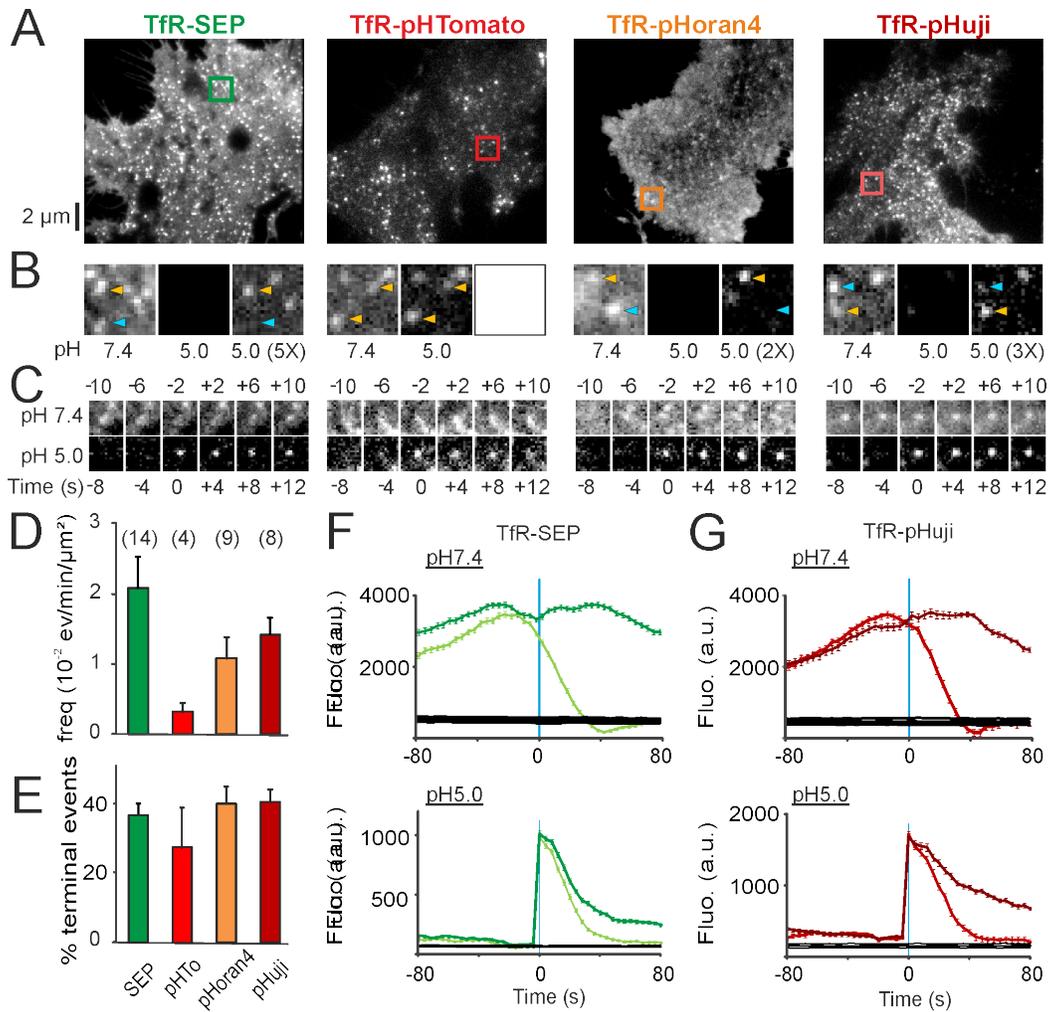
cells. Dotted line is an extrapolation of the linear fit of the 9 data points before event detection to account for photobleaching. Each frame is 2.25  $\mu\text{m}$  wide. B, fluorescence of a cell expressing TfR-pHuji illuminated at the time marked by the yellow arrowhead and imaged continuously at 10 Hz. C, time lapse recording of the same cell as in (B) (100 ms illumination, 0.25 Hz) similar to the ones used for imaging CCV formation. D, Same as in (A) for exocytosis events detected in TfR-pHuji expressing cells. (c) Average fluorescence values for 40 events in 3 cells. Error bars represent SEM.

The ppH protocol<sup>108</sup> was developed to detect the formation of endocytic vesicles loaded with TfR-SEP. This protocol involves alternating the extracellular pH between 7.4 and 5.5 every two seconds. At pH 5.5, receptor-fused SEP on the plasma membrane is not fluorescent, and all observed fluorescence is attributable to receptors sequestered inside endocytic vesicles. Accordingly, the formation of a new endocytic vesicle is revealed when a cluster of proteins visible during a pH 7.4 interval remains visible during the following pH 5.5 interval. Importantly, optimal detection of these vesicles is achieved when the pre-existing cluster is minimally fluorescent at pH 5.5, as is the case with SEP. Cells transfected TfR-red FPs had a punctuate distribution over homogenous fluorescence similar to cells transfected with TfR-SEP (Figure 2.6A). The TfR clusters correspond to clathrin coated pits (CCPs) (see Figure 2.7A-C). Some moving intracellular vesicles were seen in cells transfected with TfR-pHTomato, reflecting the

incomplete quenching of pHTomato in these vesicles that are known to be acidic and usually not seen with SEP, nor with pHoran4 and pHuji.

To visualize the formation of single CCVs using red FPs, we slightly modified the ppH protocol: we switched the low pH solution from pH 5.5 to pH 5.0 so as to maximize the quenching of plasma membrane receptors, as suggested from our data using pDisplay constructs. TfR-SEP clusters were completely quenched during low pH intervals, TfR-pHoran4 and TfR-pHuji clusters were largely, although not completely, quenched, but TfR-pHTomato clusters were only partially quenched (Figure 2.6B). During the ppH protocol, scission events associated with new vesicle formation were clearly seen with all the constructs tested (Figure 2.6C), but these events were very rare in TfR-pHTomato transfected cells. Similar to the original ppH protocol<sup>108, 109</sup>, this modified protocol did not have any apparent adverse effect on cells, as the frequency of detected events was constant for at least 20 minutes of recording (see Figure 7D). We detected and quantified endocytic events using the algorithm developed previously<sup>109</sup> with minor modifications to account for incomplete quenching of clustered surface receptors in pH 5 images. We could thus detect a large number of events in cells transfected with TfR-pHoran4 and TfR-pHuji, with frequencies only slightly lower than those obtained with TfR-SEP. In contrast, very few events were detected in TfR-pHTomato transfected cells (Figure 2.6D). In 3T3 cells, scission events fall into two categories: terminal events, where the cluster visible at pH 7.4 disappears less than 40s after the scission event, and non-

terminal events, where it remains visible<sup>108, 109</sup>. The proportion of terminal events measured in this dataset, about 40%, was very similar for SEP, pHoran4 and pHuji tagged TfR (Figure 2.6E), and comparable to published data<sup>109</sup>. The average fluorescence time course of events detected with TfR-SEP, TfR-pHoran4 and TfR-pHuji were also quite similar (Figure 2.6F-G) although there was some residual fluorescence at pH 5.0 for the pHoran4 and pHuji constructs before scission. Vesicles were also visible a little longer after scission, presumably while they were being acidified (Figure 2.6G). This is consistent with the incomplete quenching of pHoran4 or pHuji fluorescence at acidic pH. We conclude from these results that pHuji and pHoran4 are both useful reporter proteins for the trafficking of TfR and the detection of endocytic vesicles.



**Figure 2.6** Detection of endocytic vesicles containing TfR-red FPs. A, images of NIH-3T3 cells transfected with TfR fused to SEP, pHTomato, pHoran4 or pHuji at pH 7.4. B, details corresponding to the boxed areas in A, at pH 7.4 (left) and 2s later at pH 5.0 (middle, same contrast; right, increased contrast as indicated). Note the complete quenching of some clusters (blue arrowheads) whereas others are still visible (yellow arrowheads) at pH 5.0. C, examples of events detected in cells transfected with the four FPs. Each frame is  $2.25 \mu\text{m}$  wide. D, average frequency of scission events detected with the different markers. The number of cells tested is indicated. E, proportion

of terminal scission events with the different markers. F, average fluorescence of non-terminal (dark green, 984 events) and terminal (light green, 802 events) scission events at pH 7.4 (top) and 5.0 (bottom) aligned to their time of detection in 14 cells transfected with TfR-SEP. The black lines indicate 95% confidence intervals for significant enrichment. G, same as F for 8 cells transfected with TfR-pHuji (598 non-terminal events, dark red and 447 terminal events, light red). Error bars represent SEM.

### 2.3 Discussion

This work was motivated by our realization that the suboptimal performance of red fluorescent pH indicators largely stems from the fact that previous efforts had been largely focused on optimizing  $pK_a$  and had not given appropriate attention to maximizing  $n_H$ . Indeed, our modeling of the response of pH indicators suggests that, for describing pH sensitivity, the value of the  $n_H$  is actually more important than the value of the  $pK_a$ . The majority of previous studies on FP-based pH sensors have assumed an apparent  $n_H$  of 1.0 when fitting the pH titration curve, with two notable exceptions<sup>115, 116</sup>. Two intensity-based pH-sensitive red FPs, mNectarine<sup>72</sup> and pHTomato<sup>114</sup>, have suitably elevated  $pK_a$ s but relatively low  $n_H$  values. SEP, the best available FP-based pH indicator, has a  $pK_a$  that is close to optimal but what really sets it apart from other pH-sensitive FPs is its exceptionally high  $n_H$  of 1.90. This high  $n_H$  value is indicative of the existence of a second residue that can be deprotonated and interacts with the chromophore in a

cooperative manner. Based on the observation that the  $n_H$  value could be significantly different from protein to protein as a result of altered chromophore environments, we set out to engineer variants with larger fluorescence intensity changes to a given pH change.

To engineer a pH-sensitive red FP with a high  $n_H$ , we employed several different strategies including direct mutation of the chromophore, modulation of the chromophore environment by site directed mutagenesis of residues in close proximity, and random mutagenesis. Interestingly, two of the most pH-sensitive red FPs identified in this work, pHuji and pHoran4 both have mutations at residue 163 (K163Y and M163K, respectively). This result indicates the importance of this position for regulating  $pK_a$  and  $n_H$  of red FPs. In the absence of an x-ray crystal structure, we cannot know for certain how these mutations are modulating the chromophore environment. However, the close proximity of the side chain of residue 163 to the chromophore phenolate group leads us to speculate that these mutations are serving to raise the  $pK_a$  by either stabilizing the anionic form or destabilizing the protonated neutral form. Both tyrosine and lysine side chains normally have  $pK_a$ s in the range of 10-11, so deprotonation of the side chain of residue 163 in either pHuji or pHoran is unlikely to be acting cooperatively in the pH 5 to 7.5 range. Indeed, the observed increases in  $n_H$  for both of these variants are relatively modest. They however provide large improvements in their performance for imaging membrane trafficking events.

pHuji has a near-optimal  $pK_a$  of 7.7 and the highest apparent  $n_H$  of any of the so far described orange and pH-sensitive red FPs. These properties make pHuji 6- to 7-fold more pH-sensitive than the previously reported red pH sensor pHTomato<sup>114</sup> in the physiological range. However, pHuji should not be considered a replacement for SEP, since it still falls short of SEP in terms of the magnitude of fluorescence fold change between pH 5.0 and 7.5, especially *in situ*. In addition, pHuji, like earlier versions of its mApple template<sup>13</sup> and mApple-derived R-GECO1<sup>84</sup>, exhibits photoswitching behaviour. This behaviour did not impair our detection sensitivity based on steady state time lapse (for endocytosis) or continuous (for exocytosis) recordings. Nevertheless, pHuji should be used with caution for quantitative imaging of exocytosis such as in synaptic terminals<sup>117</sup> or protocols based on photobleaching (*e.g.*, fluorescence recovery after photobleaching (FRAP)<sup>118</sup>. Notably, pHoran4 does not display photoswitching and should provide a good alternative to pHuji if needed.

Clearly, the most appropriate application of pHuji is as the preferred second colour for simultaneous dual colour imaging of pH-sensitive processes with GFP or SEP-based probes. Importantly, the emission peak of pHuji is well separated from the emission peak of GFP and there is almost no spectral shift associated with the transition from the low to high pH state. Our collaborators demonstrate this possibility by imaging single endocytic events detected with TfR-pHuji together with GFP-tagged dynamin1 and clathrin light chain A. Therefore two-colour experiments with pHuji would be compatible with any of the high number

of existing GFP-fusion proteins and GFP-based sensors, such as ion sensors<sup>45, 67, 83</sup>, or Förster resonance energy transfer (FRET)-based sensors<sup>119</sup>. In conclusion, we expect that pHuji will prove widely useful as the preferred second colour of pH-sensitive probe for simultaneous imaging of multiple receptors and associated proteins during vesicle formation and release at the plasma membrane.

## **2.4 Materials and methods**

### **2.4.1 Mutagenesis, library construction and screening of pH-sensitive FPs**

Directed evolution of pHorans and pHuji was carried out by site directed mutagenesis and multiple rounds of error prone PCR (EP-PCR) using plasmids encoding mOrange and mApple as templates. All site-directed mutagenesis was performed using the Quikchange lightning mutagenesis kit (Agilent) and primers designed according to the manufacturers guidelines. EP-PCR products were digested with XhoI and HindIII and ligated into pBAD/His B vector digested with the same two enzymes, and used to transform electrocompetent *E. coli* strain DH10B (Invitrogen), which were then plated on agar plates containing LB medium supplemented with 0.4 mg/ml ampicillin and 0.02% w/v L-arabinose.

Single colonies were picked, inoculated into 4 ml of LB medium with 0.1 mg/ml ampicillin and 0.02% w/v L-arabinose and then cultured overnight. Protein was extracted using B-PER bacterial extraction reagent (Thermo) as per manufacturer

guidelines. Screening for pH sensitivity of extracted protein was performed with a Safire2 fluorescence plate reader (Tecan) by measuring protein fluorescence intensity in buffers ranging from pH 4.0 to 9.0 by steps of 1.0.  $pK_a$ s and apparent  $n_{HS}$  were estimated according to the screening results. Plasmids corresponding to the variants with relatively higher  $pK_a$ s and larger apparent  $n_{HS}$  were purified with the GeneJET miniprep kit (Thermo) and then sequenced using the BigDye Terminator Cycle Sequencing kit (Applied Biosystems).

#### **2.4.2 Protein purification and *in vitro* characterization**

To purify the pH-sensitive FPs, electrocompetent *E. coli* strain DH10B (Invitrogen) was transformed with the plasmid of interest using Micropulser electroporator (Bio-Rad). Transformed bacteria were cultured overnight on agar plates containing LB and ampicillin. Single colonies were picked and grown overnight in 4 mL LB supplemented with ampicillin at 37 °C. For each colony, the 4 mL culture was then used to inoculate 250 mL of LB medium with ampicillin and grown to an optical density of 0.6. Protein expression was induced with addition of 0.02% L-arabinose and the culture was grown overnight at 37 °C. Bacteria were harvested at 10,000 rpm, 4 °C for 10 min, lysed using a cell disruptor (Constant Systems) and then clarified at 14,000 rpm for 30 min. The protein was purified from the supernatant by Ni-NTA affinity chromatography (ABT) according to the manufacturer's instructions. The buffer of the purified protein was exchanged with 10 mM Tris-Cl, 150 mM NaCl, pH 7.3 with Amicon

ultra centrifugal filter (MWCO 10,000) for a final protein concentration of approximately 10  $\mu\text{M}$ . Molar extinction coefficients (EC) were measured by the alkali denaturation method<sup>9</sup>. Briefly, the protein was diluted into Tris buffer or 1 M NaOH and the absorbance spectra recorded under both conditions. The EC was calculated assuming the denatured RFP chromophore has a EC of 44,000  $\text{M}^{-1}\text{cm}^{-1}$  at 452 nm. Fluorescence quantum yields (QY) were determined using mOrange (for pHOran4) or mApple (for pHuji) as standards.

Fluorescence intensity as a function of pH was determined by dispensing 2  $\mu\text{L}$  of the protein solution into 50  $\mu\text{L}$  of the desired pH buffer in triplicate into a 384-well clear-bottomed plate (Nunc) and measured in a Safire2 plate reader. pH Buffer solutions from pH = 3 to pH = 11 were prepared according to the Carmody buffer system<sup>120</sup>.  $pK_a$  and  $n_H$  were determined by fitting the normalized data to the equation:

$$F = \frac{1}{1 + 10^{n_H(pK_a - pH)}}$$

### **2.4.3 Cell Culture and transfections**

NIH3T3 (ECACC) and HeLa (kind gift from A. Echard, pasteur Institute, paris, France) cells were grown in DMEM supplemented with 10% Fetal Calf Serum, 1% sodium pyruvate and 1% Glutamax with (HeLa) or without (NIH-3T3) 1%

penicillin/streptomycin and maintained at 37 °C in 5% CO<sub>2</sub>. Cells were subcultured every 2-4 days for maintenance. 24 h before imaging, cells were transfected with 1-3 µg total DNA, complexed with either 10 µL Lipofectamine2000 (Invitrogen) (NIH-3T3) or 3-6 µL X-TREMEGene-HP (Roche) (HeLa) in serum free medium. 4 h later, cells were plated on 18 mm glass coverslips (coated with 0.1 mM polylysine for 3 minutes then rinsed with 1× PBS, for NIH3T3) at a density of ~70,000 cells/ml in supplemented DMEM.

TfR-SEP and dyn1-GFP plasmids were kindly provided by C. Merrifield (LEBS, Gif-sur-Yvette, France), Clc-GFP was a kind gift of A. Echard (Pasteur Institute, Paris, France). pDisplay constructs of SEP, pHTomato, pHoran4 and pHuji were made by subcloning the genes of the pH-sensitive proteins into the pDisplay vector (Life Technologies) between the BglIII and SalI restriction sites. The pDisplay vector is translated into a protein consisting of the N-terminal signal peptide, and a C-terminal transmembrane domain of platelet-derived growth factor receptor (PDGFR) which enables anchoring of the fused protein to the cell surface. TfR-pHoran4, TfR-pHTomato and TfR-pHuji were produced by PCR amplification of the relevant red FPs and inserted into the TfR-SEP vector using NheI and AgeI sites.

#### **2.4.4 Live cell fluorescence imaging**

Live cell imaging was done at 37°C. Cells were perfused with HEPES buffered saline solution (HBS) with, in mM: 135 NaCl, 5 KCl, 0.4 MgCl<sub>2</sub>, 1.8 CaCl<sub>2</sub>, 20 HEPES and 1 D-glucose, adjusted to pH 7.4 and 300-315 mOsmol/L and supplemented with 2% fetal calf serum for HeLa cells. For pH titration with pDisplay constructs, solutions were prepared as the HBS solution described above using as pH buffers HEPES for solutions at pH 8.9, 8.0, 7.52 and 6.96, PIPES for solutions at pH 6.5 and 6.0 and MES for solutions at pH 5.5 and 5.09.

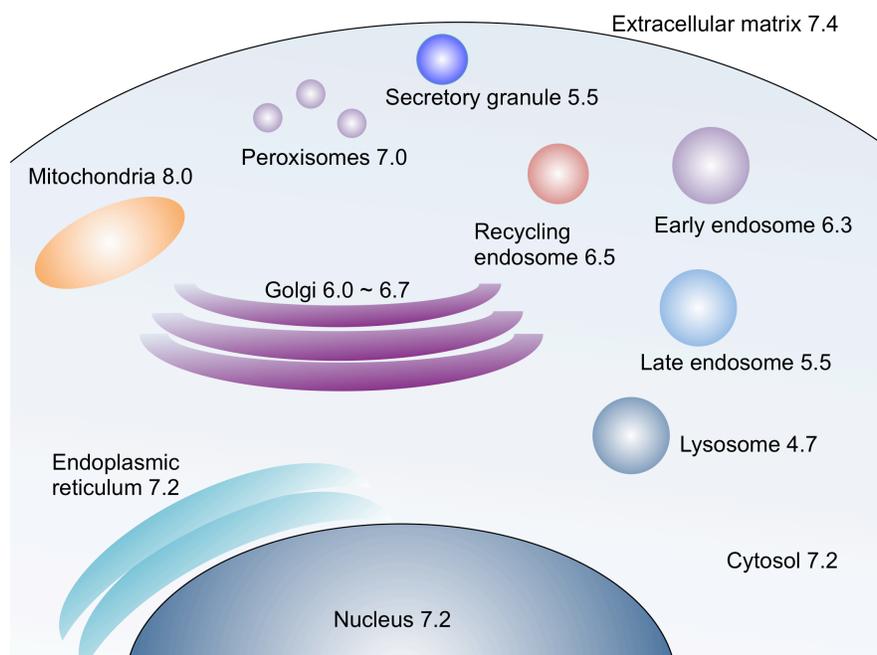
Total Internal Reflection Fluorescence (TIRF) imaging was performed on an Olympus IX71 inverted microscope equipped with an apoN oil 60× objective (NA 1.49), a 1.6× magnifying lens and a EMCCD camera (QuantEM:512SC, Roper Scientific). Samples were illuminated by a 473 nm laser (Cobolt, Sweden) for SEP imaging, as well as by a co-aligned 561 nm laser for red FP imaging. Emitted fluorescence was filtered using filters (Chroma Technology): 595/50m for pHoran4 imaging, 620/60m for pHuji and pHTomato imaging and 525/50m for SEP/GFP imaging. Simultaneous dual colour imaging was achieved using a DualView beam splitter (Roper Scientific, France). In order to correct for x/y distortions between the two channels, images of fluorescently labeled beads (Tetraspeck, 0.2 μm, Invitrogen) were taken before each experiment and used to align the two channels (see 'Image Analysis'). The camera was controlled by MetaVue7.1 (Roper Scientific, France).



# **Chapter 3 Red fluorescent protein-based excitation ratiometric pH sensors**

### 3.1 Introduction

The regulation and homeostasis of pH in living cells and its subcellular organelles (Figure 3.1) is crucial for the viability of cells<sup>121</sup>. An important toolset for probing pH regulation and homeostasis are genetically encoded fluorescent pH-sensitive sensors. Through fluorescence imaging, these sensors can measure the pH dynamics in the cell with high spatial and temporal resolution<sup>122</sup>. The readout of fluorescence intensity-based pH indicators is intuitive. Simply, a change in fluorescence intensity correlates to a change in the pH<sup>115, 123</sup>. However, it can be difficult to quantify pH changes as fluorescence intensity depends on the total concentration of fluorophore<sup>66</sup>. In order to circumvent this disadvantage of intensity-based pH indicators, ratiometric pH indicators are often used instead. FP-based dual-excitation ratiometric pH sensors have bimodal excitation spectra and the excitation peak ratio change as a function of pH<sup>124</sup>. As such, a ratiometric sensor readout is independent of protein concentration and generally less sensitive to photobleaching. This feature offers an important advantage over the single-wavelength FPs such as superecliptic pHluorin<sup>102, 125</sup> and pHuji.



**Figure 3.1** Cellular organelle pH values in a typical mammalian cell. Values in the figure were from various sources.

Wild-type avGFP has traditionally provided the foundation for construction of pH indicators suitable for excitation ratiometric measurements. The excitation spectrum of avGFP is characterized by two major excitation bands, attributed to the internal ground-state equilibrium between the neutral and anionic forms of the chromophore<sup>126</sup>. Illumination with wavelengths of light that are absorbed by the anionic form leads to the emission of fluorescence in the green region of the visible spectra. Intriguingly, illumination with wavelengths of light that are absorbed by the neutral form of the chromophore also leads to green fluorescence that is very similar to that observed for the anionic form. This observation is explained by the process of excited state proton transfer (ESPT) occurring for the

neutral form of the chromophore<sup>127</sup>. In the excited state of the neutral form, the phenol proton is relayed through a hydrogen bond network, hence generating the excited state anionic chromophore. Studies have found that the residue, glutamic acid 222, in avGFP acts as the penultimate proton acceptor in this hydrogen bond network<sup>128, 129</sup>. The existence of ESPT in wild-type avGFP affords unique opportunities for the development of biosensors with ratiometric output<sup>130</sup>.

The ratio of the two absorption bands in wild type avGFP does not depend strongly on pH or ionic strength in the physiological range. Therefore, avGFP variants with increased pH-sensitivity have been engineered for pH monitoring purposes<sup>66, 124, 125</sup>. These variants retain the two absorption bands characteristic of their template, but the ratio of the neutral and anionic populations now depends strongly on the pH change between the physiological range of 5.0 to 7.5. The first reported example of an excitation ratiometric pH-sensitive GFP was a green ratiometric pH indicator named pHluorin. pHluorin displays an increase in green fluorescence for 475 nm excitation, and a decrease in green fluorescence for 395 nm excitation, upon a pH shift from 7.5 to 5.0, with an apparent  $pK_a$  of 6.9. The functionality of pHluorin was initially tested in the *trans*-Golgi network, endosomes, and synapses<sup>66</sup>. Following pHluorin, a number of additional ratiometric pH sensors have been engineered, including: pHluorin2<sup>131</sup>, E2GFP<sup>125</sup>. So far, only one ratiometric pH sensor with red emission, a protein designated pHRed, has been reported<sup>74</sup>. pHRed was developed by mutagenesis of the long Stokes-shifted (LSS) FP mKeima which also has a ESPT pathway and a bimodal

excitation profile<sup>36, 132</sup>. The peak fluorescence emission of pHRed occurs at 610 nm. Upon pH acidification from 7.5 to 5.0, pHRed has a 4-fold increase in the ratio of 440 nm peak intensity over the 585 nm peak intensity ( $F_{440}/F_{585}$ ), with a  $pK_a$  of 6.6<sup>74</sup>.

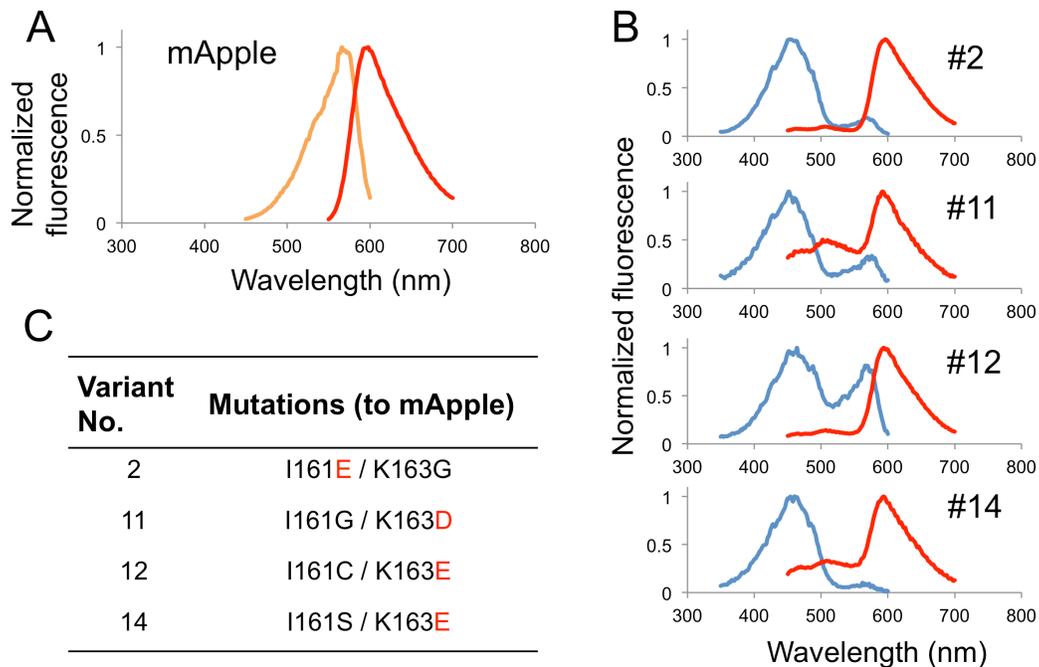
Despite the usefulness of these ratiometric pH sensors, the magnitude of the ratio changes for both pHluorin and pHRed are rather limited<sup>66, 74</sup>. A maximized ratio change would provide better sensitivity for detection and enable new detection applications. Toward this end, by means of semi-rational design and directed evolution, we have successfully engineered the single-excitation red FP mApple<sup>13</sup> into a series of dual-excitation, ratiometric pH sensors. These new red-shifted ratiometric pH sensors termed pHlorina (pH-sensitive fluorescent ratiometric indicator derived from mApple) exhibit a large ratio change of up to 70-fold when the pH changes from 5.0 to 7.5. In addition, a new variant of LSS RFP was developed during the engineering pHlorina variants.

## **3.2 Results**

### **3.2.1 Engineering ESPT pathway in the monomeric RFP mApple**

Unlike mKeima, most common RFPs do not have an ESPT pathway. Thus, the first step towards engineering a ratiometric pH sensor based on an mFruit RFP<sup>12, 13</sup> was to build a potential ESPT pathway in close proximity to the phenol moiety of the protein chromophore. It was previously reported that an acidic residue at

position 161 or 163 in RFP could potentially serve as a proton acceptor<sup>39, 133</sup> to enable an ESPT pathway from the chromophore's phenol group. To achieve this, we chose mApple (Figure 3.2) as the initial FP template and performed saturation mutagenesis on residues 161 and 163. In the visual screening of the mutant library expressed in *E. coli*, colonies with yellow coloured appearance were picked with the rationale that strong blue absorbance of the protonated chromophore would lead to yellow-coloured colonies. After picking up and culturing these yellow colonies, protein variants were extracted and the fluorescence spectra were measured. Four particular variants were found to have distinctive blue excitation and red emission in the fluorescence spectral profile. This indicates the effective introduction of an ESPT pathway in the RFP template (Figure 3.2B). Sequencing results of these variants reconciled well with the hypothesis, revealing that in each case, one of the mutated residues (i.e., 161 or 163) was mutated to either a glutamic acid (E) or an aspartic acid (D) (Figure 3.2C).



**Figure 3.2** Engineering an ESPT pathway into the RFP mApple. A, excitation (orange line) and emission (red line) fluorescence spectra of template protein mApple. B, excitation (blue line) and emission (red line) spectra of mApple variants with blue-shifted and dual-peak excitation. C, mutations of the variants compared to the mApple template, with acidic residues labeled in red.

### 3.2.2 pHlorina1, an excitation ratiometric pH sensor with red emission

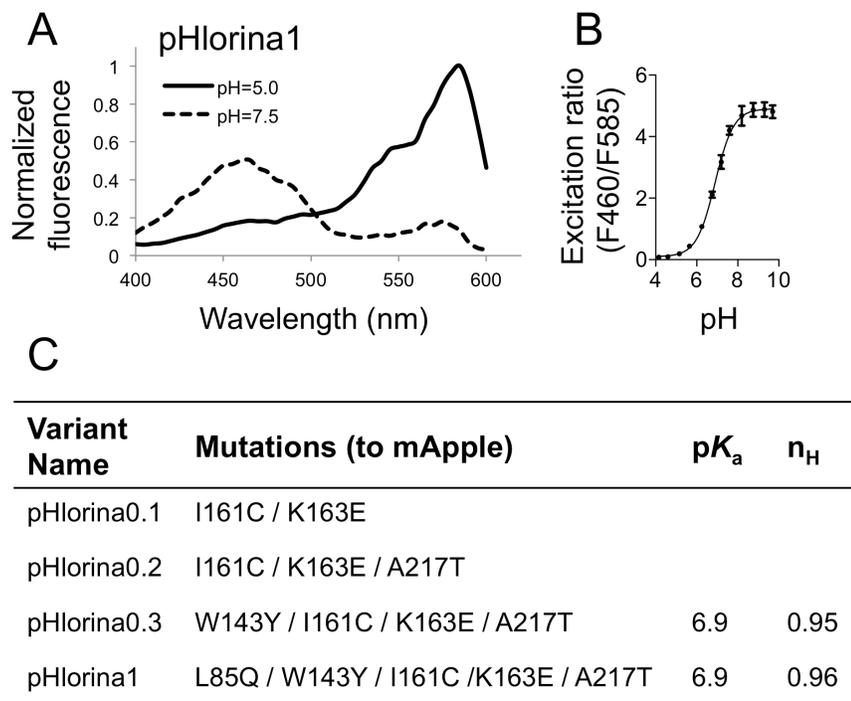
With the ESPT pathway successfully introduced, we then measured the pH sensitivity of the 4 variants with blue excitation. This was done in order to choose the best starting template for the engineering of an excitation ratiometric pH sensor with red emission. Due to a relatively large ratiometric pH response and

high brightness of variant #12, it was chosen as the template for further evolution of the ratiometric pH indicator and named pHlorina 0.1.

To further improve the pH induced fluorescence ratio change of pHlorina0.1, we first randomized residue Ala217. Mutations at this position were reported to be critical for the engineering of pHRed<sup>74</sup>. Specifically, the Ala217Ser mutation in pHRed rendered a larger ratio change compared to its template mKeima. The pHlorina0.1 Ala217X (where X = all 20 common amino acids) saturation mutagenesis library was expressed in *E. coli* followed by screening of the extracted proteins through measuring the excitation spectra at pH 5.0 and 7.5. The variant with the largest excitation ratio change was identified and designated as pHlorina0.2. Sequencing revealed pHlorina0.2 had a mutation of Ala217Thr. This mutation is distinct but similar to the Ala217Ser mutation in pHRed, as both Ser and Thr contain a polar hydroxyl group.

In an effort to increase the brightness of the pHlorina0.2, another saturation mutagenesis was performed at position 143, due to its close proximity to both the chromophore phenol group and glutamic acid residue 163. The brightest variant in the library, pHlorina0.3, possessed a Trp143Tyr mutation. Based on pHlorina0.3, further random mutagenesis and pH sensitivity screening lead to the discovery of pHlorina1, an excitation ratiometric red pH sensor with more than 20-fold excitation (460 nm / 585 nm) ratio change upon a change in pH from 5.0 to 7.5. pHlorina1 had a total of 5 mutations: L85Q/W143Y/I161C/K163E/A217T

compared to its template, mApple. The pH sensitivity of the purified pHlorina1 protein was characterized, demonstrating a high  $pK_a$  of 6.9 and a  $n_H$  of 0.96 (Figure 3.3).



**Figure 3.3** Red excitation ratiometric pH-sensitive protein pHlorina1. A, excitation spectra (emission collected at 620 nm) of pHlorina1 at pH 5.0 (dashed line) and pH 7.5 (solid line). B, pH titration curve of pHlorina1 in terms of excitation ratio (460 nm / 585 nm). C, engineering pathway of pHlorina1, including mutations of variants and pH sensitivity parameters for selected variants.

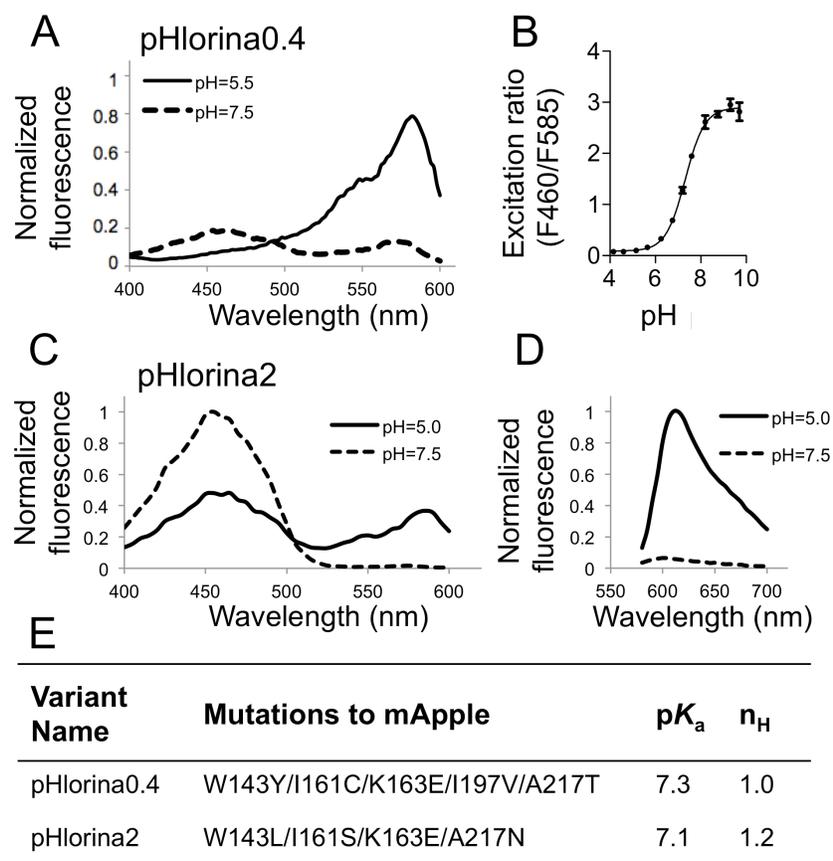
### **3.2.3 pHlorina2, a red ratiometric pH sensor with a large inverse-response intensimetric change**

pHlorina1 and its predecessors have an interesting and counterintuitive pH response. As pH decreases from neutral to acidic conditions, the red fluorescence intensity associated with excitation of the anionic form (excitation at ~585 nm) actually increases. This is in stark contrast to the usually observed fluorescence decrease in fluorescence for the anionic form with decreased pH. Indeed both superecliptic pHluorin (SEP) and pHuji have decreased excitation of the anionic form as pH is decreased. This unique phenomenon, which was previously described as “reversed protonation” in the mKeima ESPT pathway study<sup>134</sup>, led us to attempt to further optimize this aspect of the indicator. Specifically, we aimed to engineer an inverse-response intensimetric pH sensor with minimal red fluorescence at neutral pH and increased red fluorescence intensity upon acidification.

In an attempt to achieve this goal, we first explored pHlorina0.4 (Figure 3.4A-B) as a template for further engineering. pHlorina0.4 had been produced as an intermediate along the evolutionary pathway to pHlorina1. pHlorina0.4 had a  $pK_a$  of 7.3 and, more importantly, a more than 5-fold increase in red fluorescence when pH changes from 7.5 to 5.0, which made it a good candidate as the initial template for engineering the inverse-response intensimetric pH sensor. Unfortunately, after multiple rounds of random and site directed mutagenesis,

little improvement was achieved in terms of low pH induced intensity fold change.

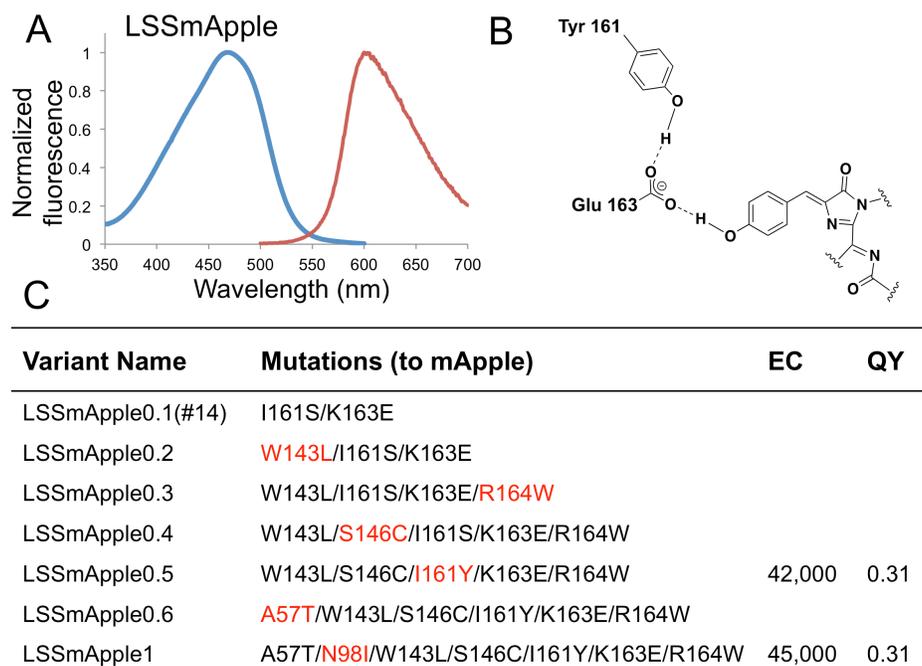
We next turned our focus to the original 4 variants (#2, #11, #12, and #14) with built-in ESPT pathways. Among them, #14 showed minimal red fluorescence from the anionic state at neutral pH, and when pH decreases, this red fluorescence slightly increases as expected. Accordingly, #14 was selected as a new starting point for the engineering. Again, we performed saturation mutagenesis at position 217 and 143 separately. After screening for red fluorescence intensity increase over pH decrease, we were able to obtain a variant, termed as pHlorina2. pHlorina2 exhibits more than 20-fold intensity increase in red fluorescence when the pH is lowered from 7.5 to 5.0 (Figure 3.4C-D). Detailed characterization of the purified protein also showed that it had over 70-fold excitation ratio change over the same pH range, which, to our best knowledge, is the largest ratio change for existing FP-based ratiometric pH indicators. The  $pK_a$  and  $n_H$  were determined as 7.1 and 1.2, respectively (Figure 3.4E).



**Figure 3.4** Inverse-response red intensimetric pH-sensitive protein pHlorina0.4 and pHlorina2. A, excitation spectra (emission collected at 620 nm) of pHlorina0.4 at pH 5.0 (dashed line) and pH 7.5 (solid line). B, pH titration curve of pHlorina0.4 in terms of excitation ratio (460 nm / 585 nm). C, excitation spectra (emission collected at 620 nm) of pHlorina2 at pH 5.0 (dashed line) and pH 7.5 (solid line). D, emission spectra (excited at 560 nm) of pHlorina2 at pH 5.0 (dashed line) and pH 7.5 (solid line). E, mutations and pH sensitivity parameters of pHlorina0.4 and pHlorina2.

### 3.2.4 Engineering of a long Stokes shift variant of mApple

During the course of the engineering of pHlorina2, the bright LSS red fluorescence of the intermediate mApple W143L/I161S/K163E motivated us to develop a bright LSS RFP (Figure 3.5A) by combining semi-rational design and random mutagenesis. In the semi-rational design, saturation mutagenesis was performed at positions 146 and 161, which are both in close contact with the protein chromophore. We also performed saturation mutagenesis at E163, which is the key residue for ESPT (Figure 3.5B). Following several rounds of random mutagenesis and brightness screening, we arrived at a final variant that was designated LSSmApple. LSSmApple had 7 mutations from the starting material mApple: A57T/N98I/W143L/S146C/I161Y/K163E/R164W. With an excitation peak at 465 nm and emission peak at 600 nm, LSSmApple has a long Stokes shift of 135 nm. Brightness measurements revealed that LSSmApple has an extinction coefficient of  $45,000 \text{ M}^{-1}\text{cm}^{-1}$ , and a quantum yield of 0.31 (Figure 3.5C).



**Figure 3.5** Long Stokes shift red FP LSSmApple. A, excitation (blue line) and emission (red line) fluorescence spectra of LSSmApple. B, modeled ESPT pathway in LSSmApple. C, engineering pathway of LSS, including mutations and brightness parameters for selected variants.

### 3.3 Discussion

The motivation of this work was to improve upon the performance of the existing ratiometric red fluorescent pH indicators<sup>74, 122</sup>. Similar to the engineering concept of the intensimetric red pH sensor pHuji, the focus of this study was largely focussed on maximizing the excitation ratio change over the physiological pH

range of 5.0 to 7.5. In our efforts to engineer such a pH-sensitive red fluorescent indicator, we first engineered an ESPT pathway into the mApple RFP. The presence of this pathway resulted in variants with the phenotype of dual excitation peaks at both 460 nm and 585 nm. Interestingly, these 4 variants identified in the initial library all have an acidic amino acid mutation at either position 161 or 163<sup>39, 133</sup>. Although the x-ray crystal structure is unavailable, this result did support our hypothesis that it would be possible to build an ESPT pathway into mApple by engineering an acidic residue as proton acceptor for the phenolate group of the chromophore.

With the prototypes of ratiometric sensors built, the engineering effort shifted to increasing the pH sensitivity as well as the brightness of the fluorescent indicator. Through multiple rounds of site directed mutagenesis and random mutagenesis, the series of pHlorina ratiometric sensors were engineered. pHlorina1 and pHlorina2 have  $pK_a$ s close to 7.0 and relatively high apparent  $n_H$  of 0.96 and 1.2 respectively. These properties make pHlorina more pH-sensitive than the previously reported ratiometric pH sensors, including pHluorin and pHRed, over the physiological pH range.

Live cell imaging of endocytosis and vesicular trafficking, ligand-induced internalization of G protein coupled receptors (GPCRs) in particular, could greatly benefit from the use of pH-sensitive FPs to monitor transit from the neutral extracellular plasma membrane to acidic vesicles. For example,

superecliptic pHluorin and pHuji display a substantial decrease in fluorescence intensity upon acidification. The loss of fluorescence could signify receptor localization to acidic vesicles. However, a potential drawback is that one cannot rule out receptor degradation as a possible explanation for the decreased fluorescence. In such cases, it might be desirable to use the ratiometric pHluorin or pHRed, since ligand-induced receptor transit from the cell surface to the endocytic vesicles would be marked by an increase in fluorescence. However, both ratiometric pHluorin and pHRed have only a moderate fluorescence intensity increase upon acidification with less than 2-fold change. pHrodo dye is a sensitive rhodamine-based fluorogenic dye, that is minimally fluorescent in the neutral pH and dramatically increased in fluorescence as the pH of its surroundings becomes more acidic<sup>135</sup>. As a result of the fluorogenic and pH-sensitive property, the pHrodo dye overcomes the drawback of pHluorin and pHRed and is used as specific sensors for endocytic events<sup>136</sup>. The development of pHlorina2 offers even larger intensity increase than pHrodo dye, and it also retains the advantages of being a genetically encoded biosensor. Such advantages include specific subcellular targeting in live cell, and no need for dye loading or washing. In summary, pHlorina2 is expected to a promising alternative for specific imaging of endocytic events in live cells.

### **3.4 Materials and methods**

### **3.4.1 General methods and materials**

All synthetic DNA oligonucleotides for cloning and library construction were purchased from Integrated DNA Technologies (IDT). *Taq* DNA polymerase (New England Biolabs) was used for error-prone PCR (EP-PCR). PCR products and products of restriction digests were purified using gel extraction kit (Bio Basic) according to the manufacturer's protocols. Restriction enzymes and ligases were purchased from New England Biolabs or Thermo Scientific. The cDNA sequences were confirmed by dye terminator cycle sequencing using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Sequencing reactions were analyzed at the University of Alberta Molecular Biology Service Unit.

### **3.4.2 Mutagenesis, library construction and screening of ratiometric pH indicators**

Engineering of pHlorina1, pHlorina2, and LSSmApple was carried out by site directed mutagenesis and multiple rounds of error prone PCR (EP-PCR) using plasmids encoding mApple as a template. All site-directed mutagenesis was performed using the Quikchange lightning mutagenesis kit (Agilent) and primers designed according to the manufacturers guidelines. EP-PCR products were digested with XhoI and HindIII and ligated into pBAD/His B vector digested with the same two enzymes, and used to transform electrocompetent *E. coli* strain DH10B (Invitrogen), which were then plated on agar plates containing LB

medium supplemented with 0.4 mg/ml ampicillin and 0.02% w/v L-arabinose.

Single colonies were picked, and inoculated into 4 ml of LB medium with 0.1 mg/ml ampicillin and 0.02% w/v L-arabinose and then cultured overnight. Protein was extracted using B-PER bacterial extraction reagent (Thermo) as per manufacturer guidelines. Screening for pH sensitivity of extracted proteins was performed with a Safire2 fluorescence plate reader (Tecan) by measuring protein fluorescence excitation spectra in buffers of pH 5.0 and 7.5. Excitation ratio changes were estimated according to the screening results. Plasmids were purified with the DNA miniprep kit (Bio Basic) and then sequenced using the BigDye Terminator Cycle Sequencing kit (Applied Biosystems). In order to screen libraries for variants that exhibited brighter LSS red fluorescence, a screening method was developed for LSS RFPs when expressed in colonies of *E. coli*. The custom built fluorescent colony imaging system equipped with filter sets (Chroma) for 470/40 nm excitation with 630/60 nm emission (i.e. for LSS fluorescence). Colonies exhibiting the highest intensity of LSS fluorescence were picked and cultured for further spectral confirmation.

### **3.4.3 Protein purification and *in vitro* characterization**

To purify the FPs, electrocompetent *E. coli* strain DH10B (Invitrogen) was transformed with the plasmid of interest using Micropulser electroporator (Bio-

Rad). Transformed bacteria were cultured overnight on agar plates containing LB and ampicillin. Single colonies were picked and grown overnight in 4 mL LB supplemented with ampicillin at 37 °C. For each colony, the 4 mL culture was then used to inoculate 250 mL of LB medium with ampicillin and grown to an optical density of 0.6. Protein expression was induced with the addition of 0.02% L-arabinose and the culture was grown overnight at 37 °C. Bacteria were harvested at 10,000 rpm, 4 °C for 10 min, lysed using a cell disruptor (Constant Systems) and then clarified at 14,000 rpm for 30 min. The protein was purified from the supernatant by Ni-NTA affinity chromatography (ABT) according to the manufacturer's instructions. The buffer of the purified protein was exchanged with 10 mM Tris-Cl, 150 mM NaCl, pH 7.3 with Amicon ultra centrifugal filter (MWCO 10,000) for a final protein concentration of approximately 10  $\mu$ M. Molar extinction coefficients (EC) were measured by the alkali denaturation method<sup>9</sup>. Briefly, the protein was diluted into Tris buffer or 1 M NaOH and the absorbance spectra recorded under both conditions. The EC was calculated assuming the denatured RFP chromophore has an EC of 44,000  $\text{M}^{-1}\text{cm}^{-1}$  at 452 nm. Fluorescence quantum yields (QY) were determined using LSSmKate2 as standard. Fluorescence intensity as a function of pH was determined by dispensing 2  $\mu$ L of the protein solution into 50  $\mu$ L of the desired pH buffer in triplicate into a 384-well clear-bottomed plate (Nunc) and measured in a Safire2 plate reader. pH Buffer solutions from pH = 3 to pH = 11 were prepared according to the Carmody buffer system<sup>120</sup>.

## **Chapter 4 Photochromic and thermochromic red fluorescent protein**

## 4.1 Introduction

Photochromism is a change of colour induced by irradiation with light, while thermochromism is a change of colour due to a change of temperature<sup>137, 138</sup>. Both photochromism and thermochromism have been observed in inorganic and organic compounds, however thermochromism is much less common in biological systems<sup>138, 139</sup>. One mechanism by which molecular photochromism can occur is a photo-induced isomerization to a second stable state with a different absorption wavelength. For such molecules, the rate of conversion is accelerated with increased temperature, and therefore the photochromic and thermochromic properties are intimately coupled<sup>138, 140, 141</sup>.

Only a handful of classes of naturally occurring photochromic proteins have been identified. Representative classes of photochromic proteins include bacteriophytochromes<sup>142</sup>, rhodopsins<sup>143</sup>, and FPs (FPs)<sup>144</sup>. The first photochromic FP to be described was the reversible ‘kindling’ protein, asFP595, from the sea anemone *Anemonia sulcata*<sup>144</sup>. Soon after, the reversibly photochromic FP Dronpa, engineered from a pectiniid coral FP, was reported to be useful as an optical highlighter for biological imaging<sup>145</sup>. Reports of a number of other reversibly photochromic FPs have since followed, including many with faster switching properties or red-shifted emission colour, including: rsFastlime<sup>146</sup>, mTFP0.7<sup>147</sup>, rsCherry<sup>148</sup>, and rsTagRFP<sup>149</sup>. Mechanistic investigations of photochromic proteins have revealed that the photoswitching mechanism is

typically based on coupled protonation changes and *E-Z* isomerizations of the protein chromophores<sup>146, 148, 150-153</sup>. Although the photochromism of FPs has been studied extensively, far fewer examples of FP thermochromism have been reported<sup>154-156</sup>. Here we report an engineered FP that exhibits reversible visible photochromism and thermochromism under ambient light and temperature conditions.

## 4.2 Results

### 4.2.1 Discovery of the photochromic and thermochromic RFP

We serendipitously discovered this photochromic and thermochromic FP during the engineering of a long Stokes shift (LSS) variant of mApple<sup>13</sup>. mApple is a monomeric red FP engineered from *Discosoma* sp. red FP (DsRed)<sup>6</sup>. LSS fluorescence is enabled by the introduction of an excited state proton transfer (ESPT) pathway into the protein. Accordingly, in LSSmApple, blue light excitation of the neutral (protonated) chromophore leads to deprotonation and the formation of the excited state anionic chromophore that emits red light<sup>39</sup>. Screening of a library of randomly mutated variants of an intermediate template (mApple-W143L/I161S/K163E, numbered according to DsRed's sequence), led to the identification of a bright variant that was picked for further characterization. The purified protein appeared yellow due to the blue light absorbing neutral chromophore. However, when this protein solution was left on ice for a few minutes under ambient light, it was observed to turn a bright

magenta colour. Removing the protein from the ice caused a conversion back to the yellow state, revealing its intriguing reversible thermochromic property. DNA sequencing revealed that this protein contained the two additional mutations S146T and R164W (Figure 4.1). The hydroxyl moiety of 146 directly interacts with the phenol/phenolate group of the chromophore, while residue 164 is on the surface of the FP with its side chain directed towards the solvent. The T146S reversion mutation abolished the photochromic and thermochromic character of the protein. Reversion of the mutations at 164 produced a protein that still exhibited chromism. This result clearly demonstrated the importance of T146 for the photo- and thermochromism. This new variant was designated as switchable hypersensitive red FP (shyRFP).

```

                20
                |
SHY RFP  M V S K G E E N N M   A I I K E F M R F K   V H M E G S V N G H   30
mApple  M V S K G E E N N M   A I I K E F M R F K   V H M E G S V N G H   30
                40                60
                |                |
SHY RFP  E F E I E G E G E G   R P Y E A F Q T A K   L K V T K G G P L P   60
mApple  E F E I E G E G E G   R P Y E A F Q T A K   L K V T K G G P L P   60
                80
                |
SHY RFP  F A W D I L S P Q F   M Y G S K V Y I K H   P A D I P D Y F K L   90
mApple  F A W D I L S P Q F   M Y G S K V Y I K H   P A D I P D Y F K L   90
                100                120
                |                |
SHY RFP  S F P E G F R W E R   V M N F E D G G I I   H V N Q D S S L Q D   120
mApple  S F P E G F R W E R   V M N F E D G G I I   H V N Q D S S L Q D   120
                140
                |
SHY RFP  G V F I Y K V K L R   G T N F P S D G P V   M Q K K T M G L E A   150
mApple  G V F I Y K V K L R   G T N F P S D G P V   M Q K K T M G W E A   150
                160                180
                |                |
SHY RFP  T E E R M Y P E D G   A L K S E S K E W L   K L K D G G H Y A A   180
mApple  S E E R M Y P E D G   A L K S E I K K R L   K L K D G G H Y A A   180
                200
                |
SHY RFP  E V K T T Y K A K K   P V Q L P G A Y I V   D I K L D I V S H N   210
mApple  E V K T T Y K A K K   P V Q L P G A Y I V   D I K L D I V S H N   210
                220
                |
SHY RFP  E D Y T I V E Q Y E   R A E G R H S T G G   M D E L Y K   236
mApple  E D Y T I V E Q Y E   R A E G R H S T G G   M D E L Y K   236

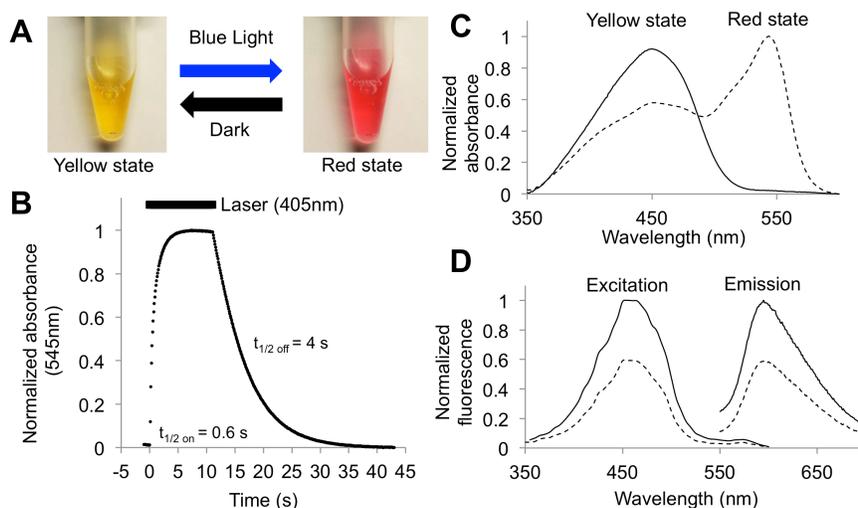
```

**Figure 4.1** Sequence alignment between shyRFP and mApple. Mutated residues are highlighted with a pink background colour.

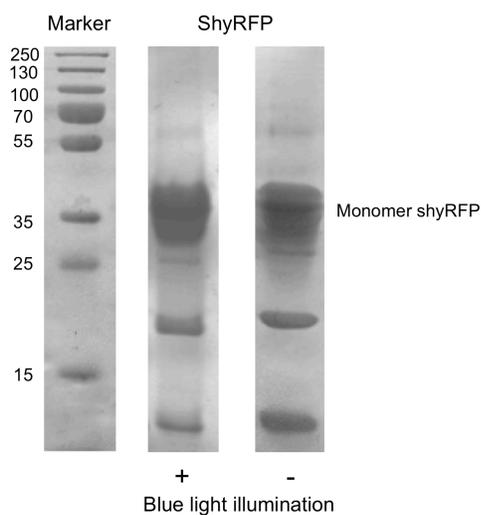
#### 4.2.2 Photochromism of ShyRFP

ShyRFP, similar to earlier generations of its mApple template<sup>13</sup> and the mApple derived genetically encoded biosensor R-GECO1<sup>84</sup>, shows a photoinduced change in spectrum. However, unlike these previous examples where the change is of little visible consequence, shyRFP exhibits a dramatic visible colour shift from a

yellow state (lower energy state) to a red state (higher energy state) upon irradiation with violet or blue light at room temperature (Figure 4.2A). A similar colour change also occurs when the temperature of the protein solution is decreased to 0 °C under normal indoor illumination. When maintained in the dark, the protein thermally reverts from the red state back to the yellow state (Figure 4.2B). Absorbance (Figure 4.2C) and fluorescence spectral measurements (Figure 4.2D) revealed that the yellow state is associated with a single absorbance peak at 450 nm. Excitation at this peak leads to LSS red fluorescence with a peak emission at 596 nm. In contrast, the red state, produced by illumination of the protein solution with a violet laser (405 nm) or blue light (470/40 nm), was associated with a second species with maximal absorbance at 545 nm (Figure 4.2C). No significant fluorescence was observed when this longer wavelength species was excited (Figure 4.2D). After converting to the red state, both the absorbance and fluorescence of the 450 nm absorbing species decreased to ~60% of the intensity in the original yellow state. The protein remained monomeric in both the yellow and red states (Figure 4.3), indicating that the photochromism of this protein, does not involve protein oligomerization through  $\beta$ -barrel structure rearrangement, as observed for Dronpa<sup>157</sup>.



**Figure 4.2** Photochromic properties of shyRFP. A, visible colour appearance of purified protein in the yellow and red states. B, absorbance (545 nm) change upon violet light (405 nm) activation and dark conversion (22 °C, pH = 7.5). Bar above the curve indicates the time when laser illumination is on. C, photochromic absorbance spectrum change. D, photochromic fluorescence spectrum change (yellow state, solid line; red state, dashed line). Note that the 545 nm absorbing species is not apparent in the red state excitation spectrum.

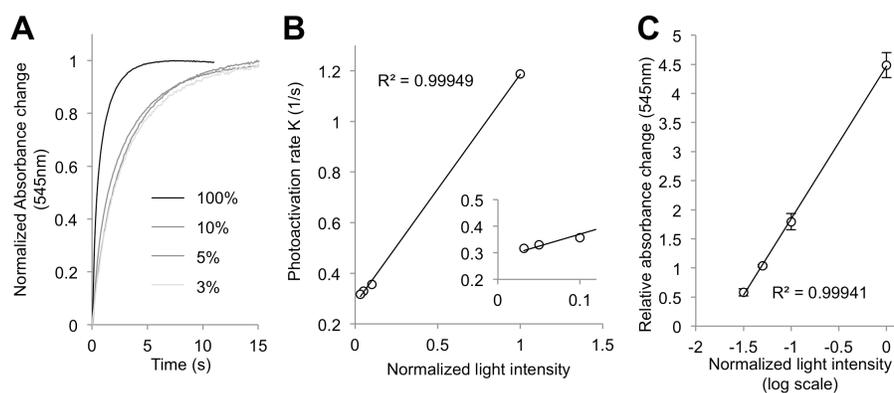


**Figure 4.3** SDS-PAGE (modified, without heat denaturation) analysis of shyRFP under blue light illumination and dark conditions. The lower bands are consistent with the expected ~19 kDa and ~7 kDa fragments produced by hydrolysis of the chromophore imine linkage in RFPs<sup>11</sup>.

#### 4.2.3 Thermochromism of ShyRFP

To further characterize the photochromic properties of shyRFP, the kinetics of the photoswitching and thermal recovery were measured and fit as first order reactions. When illuminated by a violet laser (405 nm, 150 mW) at room temperature and a pH of 7.5, the protein quickly converts to the red state with an activation half-life of 0.6 s. Thermal reversion is slower than the activation with a half-life of 4s (Figure 4.2B). The photoactivation process showed strong light intensity dependence, with both the reaction rate and relative change of

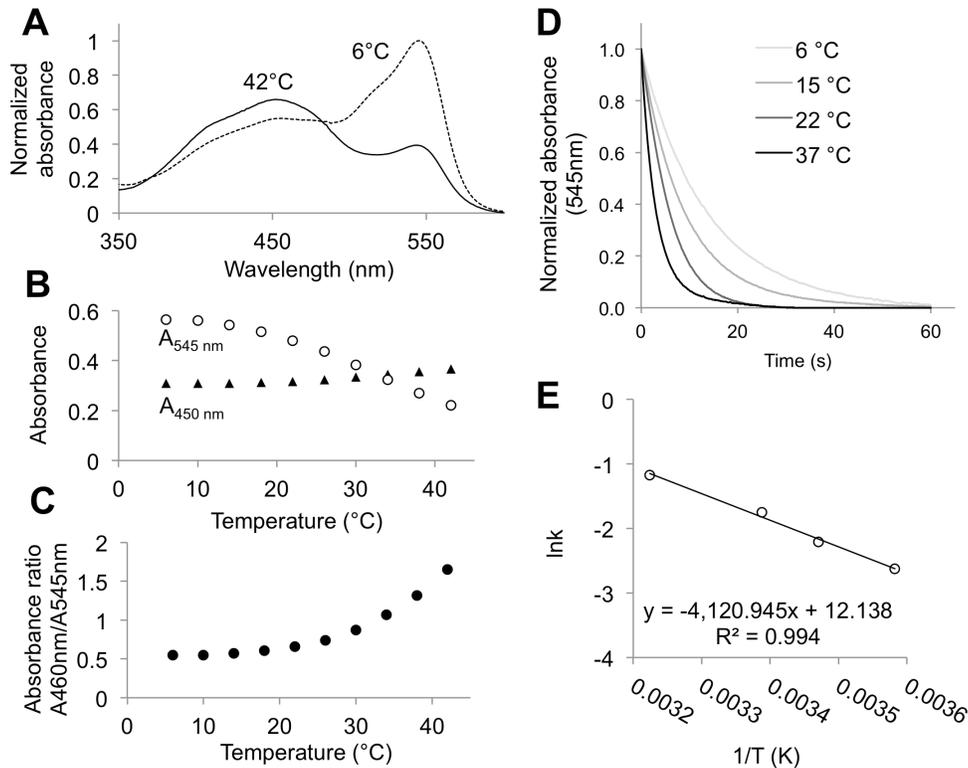
absorbance at 545 nm increasing as light intensity increased (Figure 4.4). The light-dependent thermochromism was characterized by measuring the absorbance spectra of the red state (produced using the 405 nm laser) at different temperature values ranging from 6 to 42 °C (Figure 4.5A). As the temperature decreases from 42 to 6 °C, the absorbance of 450 nm decreased by 20%, while the 545 nm absorbance dramatically increased 3-fold resulting in a red solution (Figure 2B). This result corresponded well to the observation of red colour protein solution after ambient room light illumination on ice. This dramatic shift in absorbance and visible colour can be explained by the difference in extinction coefficients for the yellow ( $30,000 \text{ M}^{-1}\text{cm}^{-1}$ ) and red states ( $142,000 \text{ M}^{-1}\text{cm}^{-1}$ ). This means that, even with a modest degree of conversion of the yellow state to the red state, the solution will appear magenta. The absorbance peak ratio exhibited the largest changes in the 30 to 42 °C range (Figure 4.5C), thus making this protein a possible candidate for a molecular thermometer in mammalian cells or tissues.



**Figure 4.4** Light intensity dependence of photoactivation rate and relative absorbance change. A, photoactivation traces (545 nm absorbance) under different intensities of laser illumination. B, photoactivation rates are linearly correlated with the light intensity. B inset, zoom-in for the first 3 data points in B. C, relative 545 nm absorbance change after illumination under different light intensities.

The dark recovery process was strongly dependent on temperature. Specifically, as the temperature increased, the half-life of reversion decreased (Figure 4.5D). Using the Arrhenius plot (Figure 4.5E), the activation energy of the thermal relaxation was calculated as 34 kJ/mol, which is comparable with other organic photochromic and thermochromic systems such as spiropyrans (9 to 27 kJ/mol)<sup>158</sup>. This relatively low energy barrier explains the short half-life of thermal relaxation for shyRFP (4 s), relative to almost all other photochromic FPs

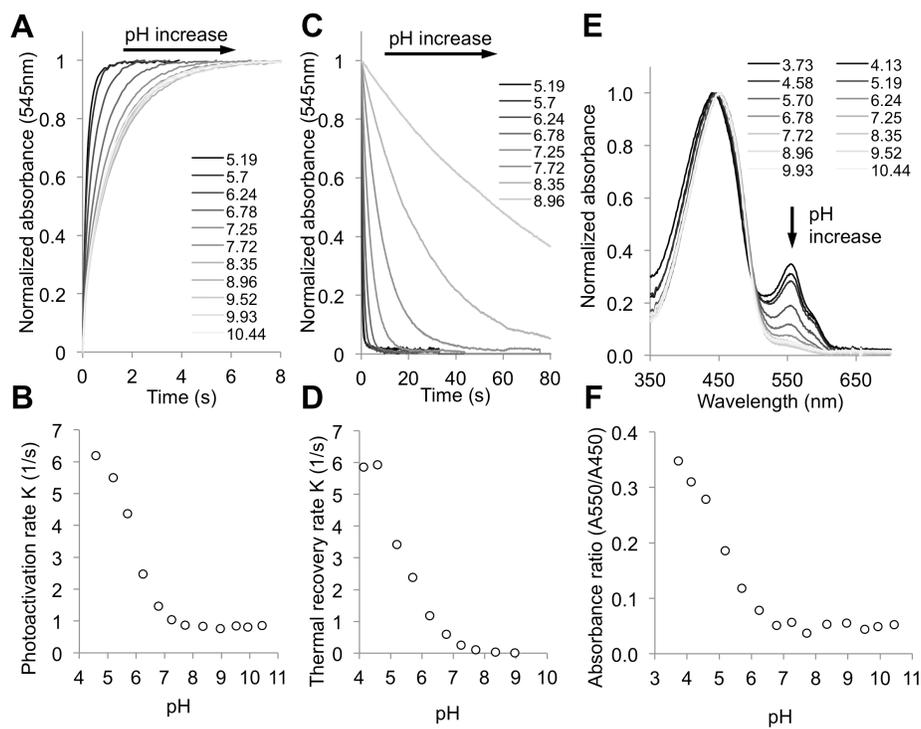
including Dronpa (840 min)<sup>146</sup>, mTFP0.7 (4 min)<sup>147</sup>, rsTagRFP (65 min)<sup>149</sup>, and rsCherry (40 s)<sup>148</sup>.



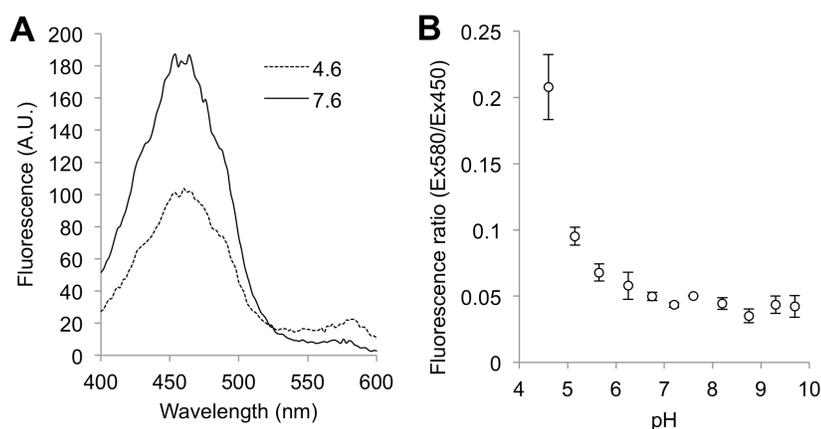
**Figure 4.5** Thermochromic properties of shyRFP. A, absorbance spectra of laser photoswitched protein at 6 and 42°C at pH 7.5. B, change of peak absorbance (545 nm, hollow circle; 450 nm, triangle). C, change of peak absorbance ratio (545 nm / 450 nm) from 6 °C to 42 °C under violet light (405 nm) illumination (pH = 7.5). D, decay kinetics of absorbance at 545 nm under various temperatures at pH 7.5. E, Arrhenius plot of thermal decay kinetics.

#### **4.2.4 pH dependence of ShyRFP**

To investigate the influence of the protein chromophore protonation state on the photo- and thermochromism, kinetic measurements of photoactivation (Figure 4.6A and 4.6B) and thermal recovery (Figure 4.6C and 4.6D) were performed at various pH values between 3 and 11. These measurements revealed that both processes were accelerated at lower pH values, suggesting that the mechanism of photochromism was associated with the chromophore protonation state. Absorbance (Figure 4.6A and 4.6B) and fluorescence spectra (Figure 4.7) of the yellow state protein at various pH values were measured.

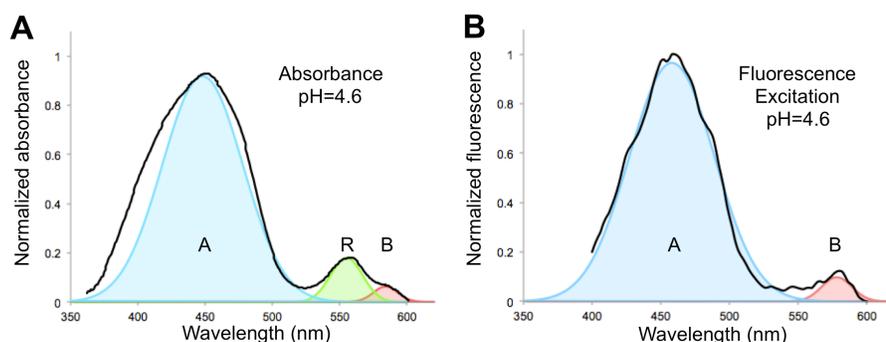


**Figure 4.6** pH dependence of shyRFP. A and B, photoactivation traces and rates at different pH values. C and D, thermal recovery traces and rates at different pH values. E and F, normalized absorbance spectra and absorbance ratio (550 nm / 450 nm) in the yellow state under various pH values.



**Figure 4.7** Fluorescence excitation change at different pH. A, fluorescence excitation spectrum of shyRFP at low pH (4.6) and at neutral pH (7.6). B, excitation ratio (580 nm / 450 nm) as a function of pH.

At low pH (3 to 6), the absorbance and fluorescence spectra revealed the coexistence of three distinct states of chromophore: a blue-absorbing/red-emitting LSS species, a yellow-absorbing/non-fluorescent species, and an orange-absorbing/red-emitting species (Figure 4.8). At neutral to high pH (7 to 10), the chromophore was primarily the blue-absorbing red-emitting LSS species. Higher pH values resulted in denaturation of the protein. The fact that the absorbance spectrum, at any pH, is not identical to the photoactivated form suggested to us that a simple equilibrium between the protonated and deprotonated forms of the chromophore is insufficient to explain the chromic properties of shyRFP.



**Figure 4.8** Absorbance and fluorescence excitation at low pH (pH = 4.6). The three distinct species are indicated using solid coloured curves.

#### 4.2.5 Mutational analysis and proposed chromism mechanism

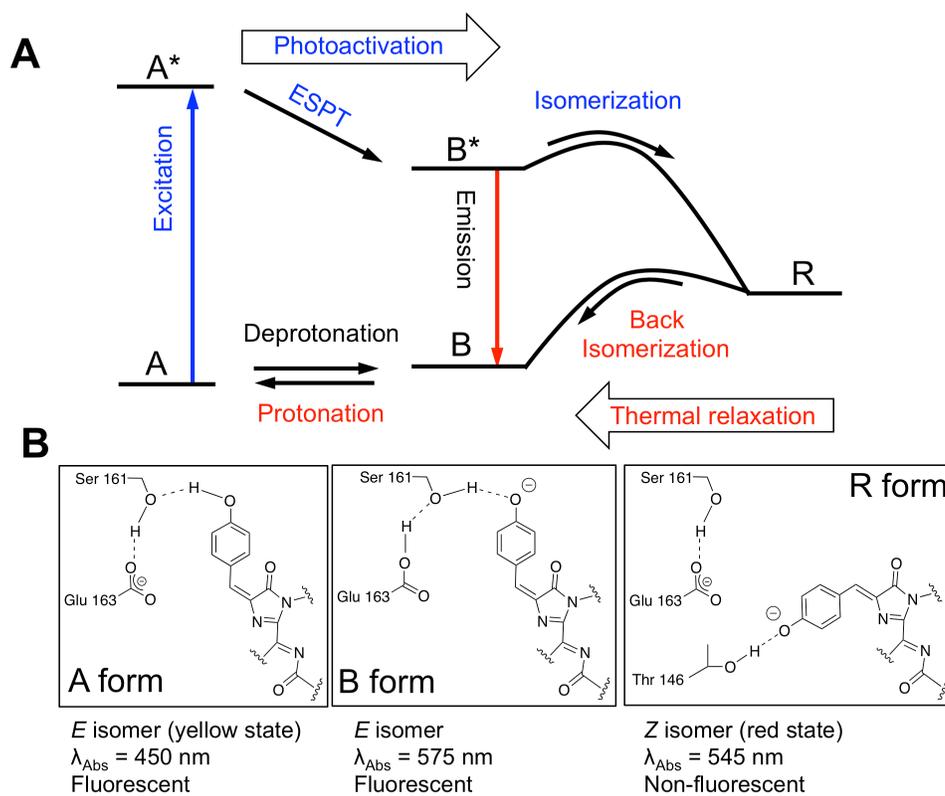
To further explore the mechanism of chromism in shyRFP, saturation mutagenesis was conducted at residues in close proximity to the chromophore (*i.e.*, residues 146, 161 and 163). Mutations at residue 146 resulted in significantly reduced (T146C and T146G) or complete loss (T146S) of photochromism. However, mutations at 146 did not alter the LSS spectral character, which indicates that residue 146 is not involved in the ESPT pathway. At residue 161 and 163, several mutations (S161A, S161F, E163I and E163V) resulted in the disappearance of the blue absorbing/LSS fluorescent species and yielded regular Stokes shift red FP type variants with no chromic changes. This result indicates that residues 161 and 163 are essential for the excited state proton transfer (ESPT) pathway<sup>39</sup> that gives rise to the LSS fluorescence. In addition, considering the

relative positions of residues 161 and 163 to the protein chromophore (Figure 4.9), it appears that the LSS species, associated with the yellow state of shyRFP, is most likely in the *E* conformation<sup>39,40</sup>.

The photophysical characterization and site-directed mutagenesis results with shyRFP led to the proposal of a mechanism that accounts for the key features of this photochromic and thermochromic phenomenon (Figure 4.9A). For the photochromism, violet/blue light illumination changes the chromophore from a 450 nm absorbing LSS fluorescent form (A/A\* form) to a 545 nm absorbing non-fluorescent form (R form), which we conclude is anionic due to its long wavelength absorbance at 545 nm. In the dark, the 545 nm absorbing species R thermally reverts back to its neutral ground state, form A. The ground state equilibrium between A and B shifts towards the red fluorescent B form at low pH due to a “reverse protonation” effect that has been reported for LSS RFP mKeima<sup>36,134</sup>. At low pH the side chain of E163 is protonated and cannot stabilize the hydrogen bond network with the neutral chromophore. Accordingly, the anionic B form of the chromophore is favoured (Figure 4.9B).

The non-fluorescent R form has a blue-shifted absorbance maximum (545 nm) relative to the red fluorescent B form (580 nm). The fact that these two states are not identical at any pH rules out the possibility that violet/blue light illumination simply reversibly shifts the equilibrium between the protonated A state and

deprotonated B state of the protein chromophore. Yet the reaction rate is highly dependent on the pH, indicating that protonation/deprotonation is coupled with chromism. Accordingly, in addition to a change in chromophore protonation state, the mechanism of the photochromism must also involve an alternate structure, conformation, or microenvironment for the illuminated form. Our currently preferred explanation is that violet/blue light illumination induces an *E-Z* isomerization of the chromophore. Our mutational study suggested that the LSS fluorescent A form (yellow state) is in the *E* conformation, and that the photoswitching process changes the chromophore conformation to the *Z* isomer, which is stable in the anionic state (Figure 4.9B). In the dark, the *Z* isomer (red state) is able to thermally relax back to the *E* conformation (yellow state) at room temperature due to the low activation energy. In short, the photochromism and thermochromism of shyRFP are closely coupled, with illumination pushing the chromophore towards the red R state (anionic *Z* isomer) and the thermal relaxation pushing the chromophore towards the yellow A state (protonated *E* isomer).



**Figure 4.9** Proposed mechanism of shyRFP photochromism and thermochromism. A, proposed schematic of the chromophore states involved in conversion. B, proposed isomers of the chromophore and the important amino acids in the chromophore environment.

### 4.3 Discussion

Due to its complex photophysics and rapid photoswitching behaviour, shyRFP is an unlikely candidate for applications in conventional fluorescence microscopy. However, shyRFP could potentially find application in some specialized imaging

experiments. For example, temperature and pH measurements based on the absorbance peak ratios or the kinetics of dark reversion for photoactivated shyRFP might be accessible to absorption microscopy. As both of these variables are inherently coupled, care must be taken to extract the desired information. With its large and distinct reversible absorbance spectra shift, shyRFP could also serve as a model for photoswitching studies and provide insight into the photoactivation mechanism of mApple variants. ShyRFP is the first example of an LSS RFP with a reversible fluorescence change; therefore, like Dronpa, rsCherry, and rsTagRFP, it could in principle be applied for super-resolution microscopy techniques, such as PALM or RESOLFT<sup>159</sup>. Furthermore, the shyRFP photoconversion is accompanied by significant absorbance changes that potentially allows its application in photochromic FRET<sup>149, 160</sup>. ShyRFP also has considerable potential as a teaching tool as it visually changes colour without expensive equipment and reacts strongly to light, temperature, and pH. ShyRFP could also be used in more frivolous pursuits such as transgenic plants that change colour in response to temperature changes. We are certain that further development and evolution of shyRFP will lead to more useful additions to the FP toolbox for use in live cell imaging.

#### **4.4 Material and Methods**

#### **4.4.1 Molecular cloning and mutagenesis**

Site directed mutagenesis and error prone PCR (EP-PCR) was performed using plasmids encoding mApple as templates. All site-directed mutagenesis was performed using the Quikchange lightning mutagenesis kit (Agilent), with primers designed according to the manufacturers guidelines. EP-PCR products were digested with XhoI and HindIII restriction enzymes and ligated into pBAD/His B vector (Life Technologies) digested with the same two enzymes. The ligation product was used to transform electrocompetent *Escherichia coli* strain DH10B (Life Technologies), which were then plated on agar plates containing LB medium supplemented with 0.4 mg/ml ampicillin and 0.02% w/v L-arabinose. Plasmids were purified with the GeneJET miniprep kit (Thermo Scientific) as per manufacturer's instruction. All DNA sequencing was done using the BigDye Terminator Cycle Sequencing kit (Applied Biosystems).

#### **4.4.2 Protein purification and characterization**

A single colony of *E. coli*, transformed with the pBAD/His B plasmid containing the shyRFP gene, was used to inoculate 4 mL LB supplemented with ampicillin. This culture was incubated at 37° in a shaker incubator (220 rpm) for 12 hours. The liquid culture was then added into 500 mL of LB with ampicillin and incubated a further 4 hours. The inducer L-arabinose was added to a concentration of 0.02% and the culture allowed to grow overnight at 30 °C. The cells were then centrifuged at 8,000 rpm for 10 minutes at 4 °C and resuspended in 1× Tris

buffered saline (TBS). The cells were then lysed using a cell disruptor (Constant system). The cell debris was removed by centrifugation at 10,000 rpm for 35 minutes at 4 °C. The protein was then purified using Ni-NTA agarose beads (MC Labs) and buffer exchanged with 1× TBS using a centrifugal device with 10,000 MWCO (Amicon). The shyRFP absorbance profile was obtained by adding 10 µL of purified shyRFP to various pH solutions using the Carmody buffers<sup>120</sup>. Spectra were measured using a UV-Vis spectrometer (Agilent) with the illumination of a 405 nm laser (150 mW, 1200 mW/cm<sup>2</sup>, Changchun New Industries Optoelectronics Tech. Co., Ltd.). The photoswitched profile was obtained by unblocking of the laser illumination 30 seconds before each measurement to allow for full activation, and then reblocking of the laser illumination after each measurement to minimize photobleaching of the FP. Activation and decay plots at various temperatures and pHs were obtained by performing a continuous absorbance reading with an integration time of 0.1 s. Light intensity was modified using neutral density filters (Chroma) with the transmittance values 100% (no filter), 10% (ND 1), 5% (ND 1.3), and 3% (ND 1.5). Photoactivation rates were fitted using one phase association equation:  $Y = 1 - \exp(-K \cdot X)$ , where Y represents the normalized absorbance change, X represents the time (in seconds), and K is the fitted activation rate constant. Decay rates were fitted using one phase decay equation:  $Y = \exp(-K \cdot X)$ , where Y represents the normalized absorbance, X represents the time (in seconds), and K is the fitted decay rate constant.

Extinction coefficients were determined by measuring the absorption spectrum of shyRFP using an UV/Vis spectrometer (Beckman Coulter DU 800) at pH 7.25 in 1× TBS and the same concentration of shyRFP in 1 M NaOH. In 1 M NaOH solution, the FP chromophore is assumed to have an extinction coefficient of  $44,000 \text{ M}^{-1}\text{cm}^{-1}$  at  $446 \text{ nm}^9$ . Through the Beer-Lambert law this gives a concentration of the alkaline denatured chromophore that can be applied to the pH 7.5 FP, allowing the calculation of the yellow state extinction coefficient. The red state extinction coefficient was measured from the absorbance of the shyRFP solution under active 405 nm laser illumination. We assumed that the decrease in yellow state concentration corresponded to the increase in red state concentration. Accordingly, the ratio of yellow state to red state extinction coefficients could be calculated from the relative changes in yellow state and red state absorbance.

**Chapter 5 Engineering of long Stokes shift and red-shifted  
variants of mCherry**

## 5.1 Introduction

The discovery and development of red fluorescent proteins (RFPs)<sup>6</sup> as tools for live cell imaging, represents one of the most important advances in fluorescent protein (FP)-based technologies since the advent of *Aequorea victoria* green FP (GFP)<sup>3</sup>. Protein engineering efforts have produced two predominant lineages of monomeric RFPs derived from different natural sources of RFP. One lineage is derived from the tetrameric RFP DsRed from reef coral *Discosoma* sp. and includes the first monomeric RFP, mRFP1<sup>11</sup>, and the mRFP1 derived mFruit variants including mCherry, mOrange<sup>12</sup>, and mApple<sup>13</sup>. The second lineage is engineered from sea anemone *Entacmaea quadricolor* RFP eqFP611<sup>161</sup> and includes TagRFP<sup>25</sup>, mKate<sup>31</sup>, mKate2<sup>32</sup>, and FusionRed<sup>26</sup>. All of the above mentioned RFPs have been widely used in the cell biology research community for imaging of protein dynamics and localization in live cells.

One of the important reasons behind the popularity of FPs is the availability of a large FP colour palette ranging from ultraviolet to far red, as a result of the researchers' continuing efforts to design and engineer FPs with spectrally shifted excitation and emission. Spectral property modulation of FPs was first demonstrated in the development of yellow FP (YFP)<sup>15</sup>. By introducing the T203Y substitution, located above the phenol ring of chromophore of GFP, a new  $\pi$ - $\pi$  stacking interaction was created. This new interaction resulted in a red-shifted excitation and emission spectra relative to GFP. In addition to adjusting the environment of chromophore, alteration of the excitation and emission of

fluorescence can be achieved by changing the chromophore structure by substitutions of the chromophore constituting residues. This approach was applied in the development of BFP, CFP<sup>5</sup>, mWasabi<sup>162</sup> and mOrange<sup>12</sup>.

DsRed, eqFP587, eqFP611, and many of the variants engineered from these progenitor proteins, have maximal excitation in the range of 570-590 nm and maximal emission in the range of 590-611 nm. Through the use of protein engineering, researchers have created RFP variants with substantially altered excitation and emission maxima. These spectral alterations can include a blue shift<sup>163</sup>, a red shift<sup>23</sup>, or a change in the Stokes shift (the wavelength difference between the excitation and emission maxima)<sup>37</sup>.

Based on past efforts to expand the colour choices of RFPs, two types of FPs with red emission were found to be particularly useful additions to the FP toolbox: long Stokes shift (LSS) RFP (blue excitation, red emission, Stokes shift >100 nm) and red-shifted (RDS) RFP (red excitation, far-red emission). LSS RFPs hold promise for two-photon fluorescence imaging due to their compatibility with EGFP using single Ti-Sapphire laser excitation<sup>37, 132</sup>. The first LSS RFP mKeima was developed based on a chromoprotein from stony coral *Montipora sp.*<sup>36</sup>. Following that, three new LSS RFPs, LSSmKate1, LSSmKate2<sup>37</sup>, and mBeRFP<sup>38</sup>, were engineered from mKate by introducing an excited state proton transfer (ESPT) pathway in proximity to the protein chromophore<sup>39</sup>. Following blue light excitation, the excited state of the chromophore has a lower  $pK_a$  and thus transfers

a proton through a hydrogen bond network to generate an excited state anionic chromophore that emits red light. The strategy of engineering ESPT pathway into FPs was applied to some of the mFruit RFPs, including mOrange and mCherry, to generate variants with blue shifted fluorescence excitation<sup>39</sup>. However, the resulting mCherry LSS variants suffered from the presence of both green and red emitting species and thus no further development was pursued.

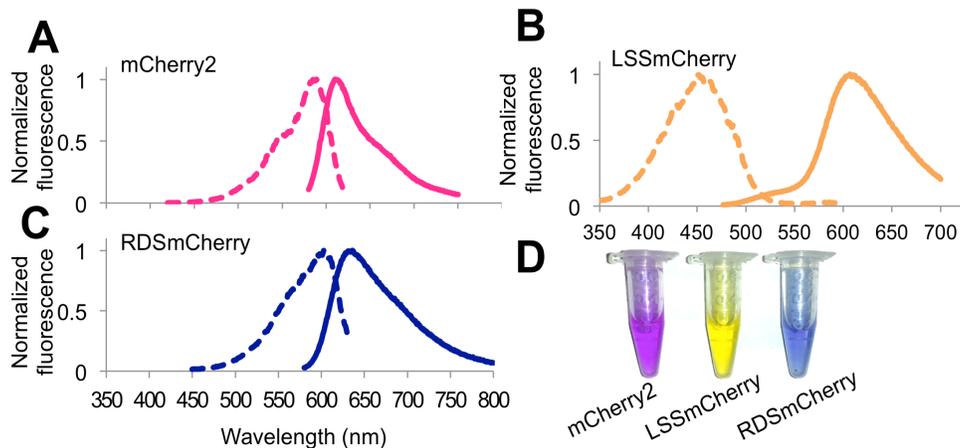
Red-shifted RFPs are of particular importance for deep tissue imaging of whole small mammals, since far-red light penetrates deeper into tissue than more blue shifted wavelengths of light. The great potential utility of such FPs has motivated FP engineers to push the excitation and emission wavelengths of RFP into the near infrared region<sup>23, 31</sup>. Based on mRFP1, the initial effort of creating red shifted RFPs yielded mPlum ( $\lambda_{\text{ex}} = 590 \text{ nm}/\lambda_{\text{em}} = 649 \text{ nm}$ ) and mRaspberry ( $\lambda_{\text{ex}} = 598 \text{ nm}/\lambda_{\text{em}} = 625 \text{ nm}$ )<sup>29</sup>. mGrape<sup>23</sup> represented yet another attempt at red shifting mRFP1, however the final variant mGrape3 ( $\lambda_{\text{ex}} = 608 \text{ nm}/\lambda_{\text{em}} = 646 \text{ nm}$ ) was dim in fluorescence and displayed a photoactivation behaviour. mCherry was also engineered for longer wavelength emission with the aid of a computationally designed library, the resulting variant mRouge<sup>30</sup> shifted the excitation to 600 nm and emission to 637 nm, but the brightness was still low due to a low quantum yield (0.02). In addition, two bright far-red mKate derivatives mCardinal<sup>33</sup> and TagRFP657<sup>34</sup> have been developed. These variants have excitation maximum at 604 nm and 611 nm, respectively, and 659 nm emission maxima.

Here we describe our effort to engineer new spectrally shifted variants based on mCherry, one of the most useful and widely used RFPs currently available. mCherry is widely used in cell biology research due to its monomeric structure, high brightness, faster maturation, and high photostability<sup>12</sup>. Through site-specific saturation mutagenesis and random mutagenesis, combined with spectral property screening, we have now identified two new mCherry variants with long Stokes shift and red-shifted spectral properties respectively (Figure 5.1). These two variants represent the latest additions to the rapidly growing class of LSS FPs and far-red FPs. This work also provided important insight into the amino acid determinants of mCherry fluorescence colour. The information gained in this study should inform and facilitate future efforts to expand the palette of mFruit variants and mFruit-based genetically encoded biosensors<sup>83, 84, 164</sup>.

## **5.2 Results**

### **5.2.1 mCherry2 as template for directed evolution**

mCherry2 (mCherry-K92N/K138C/K139R/S147T/N196D/T202L, courtesy of Nathan Shaner, unpublished), is derived from the RFP mCherry and has a similar excitation maximum at 587 nm and emission maximum at 610 nm (Figure 5.1A). mCherry2 has a slightly higher brightness (Table 5.1) when compared with mCherry.

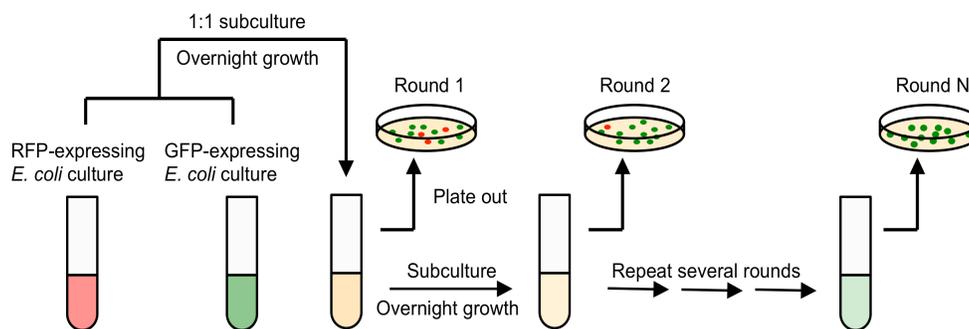


**Figure 5.1** Spectral characteristics of mCherry2, LSSmCherry, and RDSmCherry.

A, Normalized fluorescence excitation (dashed line) and emission spectra (solid line) of mCherry2; B, LSSmCherry and C, RDSmCherry. D, Visible colours of mCherry2, LSSmCherry and RDSmCherry purified protein. The yellow visible colour of LSSmCherry is due to the absorbance of blue light. The blue visible colour of RDSmCherry is due to the absorbance of green and red light.

We next sought to determine if mCherry and mCherry2 exhibited any differences in their cytotoxicity. Accordingly, a bacterial growth rate comparison test was developed to evaluate the cytotoxicity of mCherry and mCherry2, relative to EGFP. Bacterial growth is slowed by the expression of proteins that exhibit toxicity to the bacteria<sup>165, 166</sup>. When cells expressing proteins with various levels of cytotoxicity are grown in the same culture, the faster-growing non-toxic protein expressing cells eventually become more abundant. Due to the fast turnover rate and convenient handling of *E. coli* as host cells, a growth comparison assay was

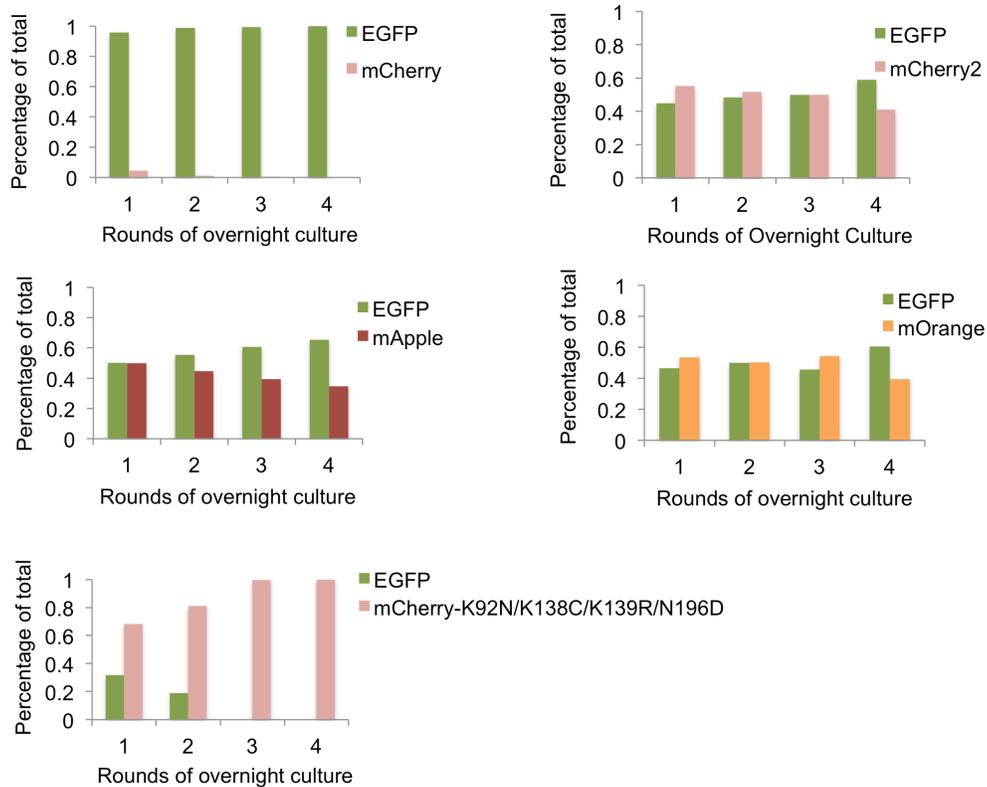
developed to evaluate the cytotoxicity of RFPs, using EGFP as a standard. *E. coli* bearing RFP or EGFP encoding plasmids with the same promoter were first cultured overnight separately without inducing the protein expression. On the following day, the *E. coli* cultures were mixed according to a 1:1 (v/v) ratio and cultured in fresh medium with inducer at the same tube overnight. Subculture was repeated every day for 4 days to ensure a clear comparison result over time. Each overnight culture was plated on agar plates supplemented with inducer for FP expression (Figure 5.2). Fluorescence images of the plates were taken under green and red channel respectively. Percentages of EGFP and RFP-expressing bacteria were then calculated based on the fluorescent colony numbers obtained from the images.



**Figure 5.2** Schematic procedure of bacterial growth comparison method for cytotoxicity evaluation.

Using this assay, we evaluated the relative cytotoxicity of mCherry, mCherry2, mOrange, and mApple (Figure 5.3). These experiments revealed that a bacterial population expressing mCherry was quickly overtaken by the EGFP expressing

population, indicating a higher cytotoxicity of mCherry. *E. coli* expressing other RFPs showed similar growth rates with the ones with EGFP in the time frame of this test, suggesting they had relative lower cytotoxicity compared to mCherry. Due to the high brightness and low cytotoxicity, mCherry2 was chosen as the template for subsequent development.

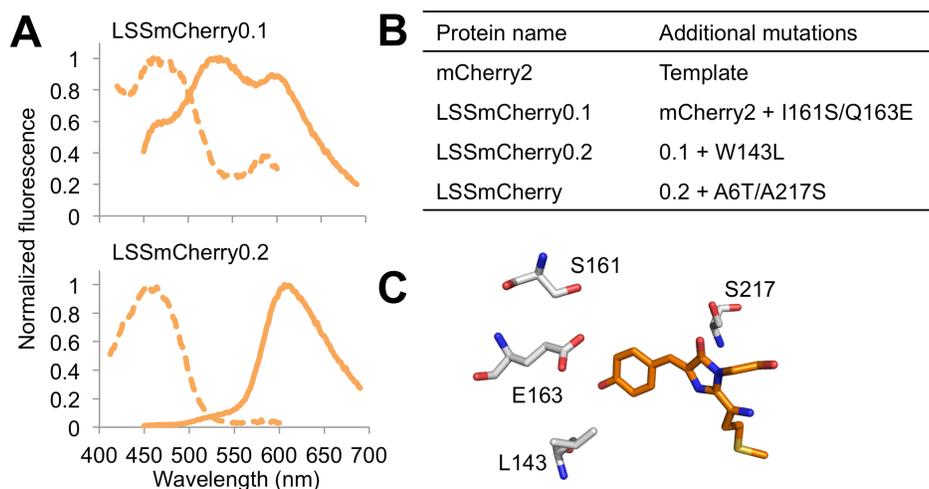


**Figure 5.3** Percentage of total colony numbers comparison between RFP and EGFP expressing *E. coli*.

### 5.2.2 Directed evolution of a long Stokes shift (LSS) mCherry variant

For the development of an RFP with a long Stokes shift (LSS), semi-rational design in combination with random mutagenesis were performed on the template gene encoding mCherry2. In the first step, site-directed saturation mutagenesis was performed on residues 161 and 163 of mCherry2 (numbering according to DsRed) to introduce the excited state proton transfer (ESPT) pathway<sup>39</sup>. The rationale for introducing an ESPT pathway was that this could lead to an LSS phenotype with the excitation maxima shifted to the cyan region (450 nm) of the visible spectrum. The libraries of bacterial colonies expressing mutants were screened using a custom-built fluorescent colony screening system<sup>95</sup>, with LSS red fluorescence ( $\lambda_{\text{ex}} = 470/40 \text{ nm}$ ,  $\lambda_{\text{em}} = 630/60 \text{ nm}$ ) as the positive criterion and red fluorescence of regular Stokes shift ( $\lambda_{\text{ex}} = 560/40 \text{ nm}$ ;  $\lambda_{\text{em}} = 630/60 \text{ nm}$ ) as the negative criterion. The best variant identified in the 161X/163X library, named as LSSmCherry0.1 (Figure 5.4A), possessed mutations of I161S and Q163E, and its fluorescence exhibited a blue shifted excitation, probably due to the establishment of an ESPT pathway from the chromophore to the glutamic acid at 163. This result also corresponded well with previous studies where the LSS RFP variants were discovered with position 163 mutated to acidic residues<sup>37, 39, 40</sup>. However, undesirable green fluorescence and regular Stokes shift red fluorescence still co-existed with the LSS fluorescence in the emission spectrum of LSSmCherry0.1. In the attempt to eliminate the unwanted fluorescence species and further improve mCherry2- I161S/I163E, saturation mutagenesis was performed at position 143, which is in close proximity to the protein chromophore. The resulting variant,

mCherry2-W143L/I161S/Q163E, was designated as LSSmCherry0.2. LSSmCherry0.2 showed exclusive LSS red fluorescence with no regular red fluorescence (Figure 5.4A). This mutant was then subjected to several rounds of random mutagenesis using error-prone polymerase chain reaction (EP-PCR) and screened for better LSS fluorescence brightness. The final variant, LSSmCherry1, had maximal excitation at 450 nm and maximal emission at 610 nm, corresponding to 160 nm Stokes shift. LSSmCherry1 has 5 substitutions (A6T, W143L, I161S, Q163E and A217S) relative to its template mCherry2 (Figure 5.4B). Residues 143,161,163 and 217 are all located inside the  $\beta$ -barrel structure of the FP and pointing towards the protein chromophore (Figure 5.4C); while residue 6 located at the N-terminus of the protein and relatively distant from the chromophore. It is likely that this mutation is facilitating protein folding.



**Figure 5.4** Directed evolution of LSSmCherry. A, Normalized fluorescence excitation (dashed line) and emission spectra (solid line) of LSSmCherry0.1 and LSSmCherry0.2. B, Mutation table of LSSmCherry. C, Modeled critical mutation positions of LSSmCherry (based on mCherry crystal structure, PDB: 2H5Q)<sup>16</sup>.

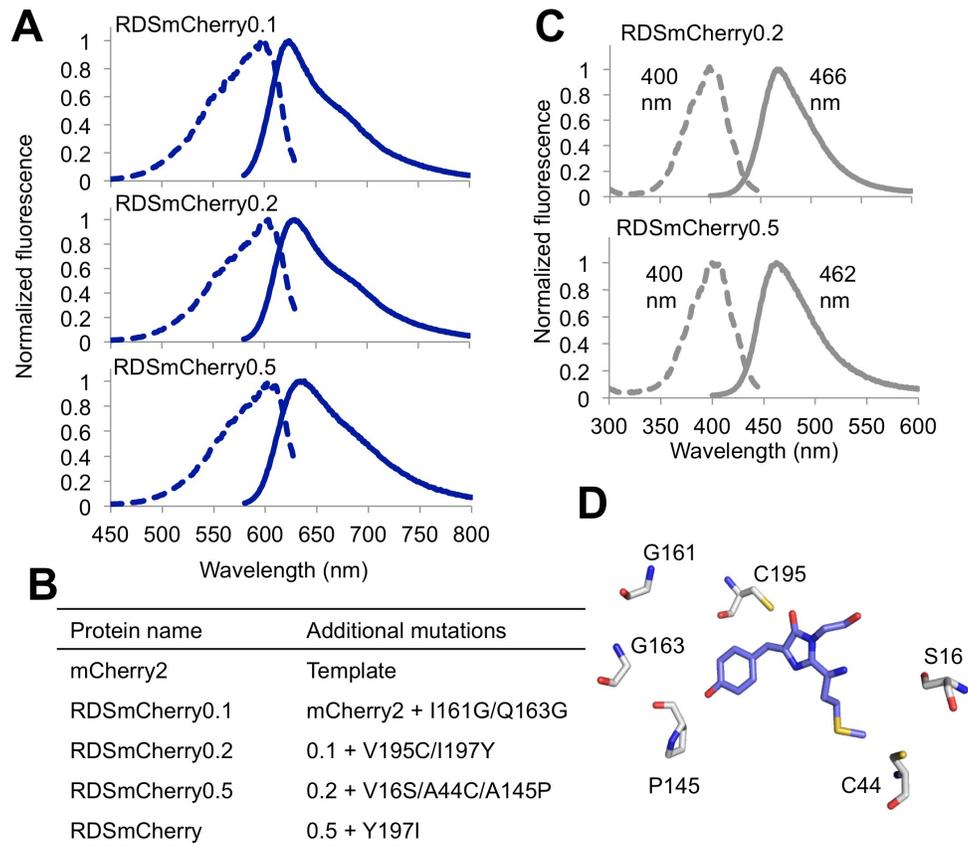
### 5.2.3 Directed evolution of a red-shifted (RDS) mCherry variant

At the initial step of developing LSSmCherry, the mCherry2 161X/163X gene library was created. Through careful visual inspection of the bacterial colonies in this mutant library, a number of colonies were found to have blue/purple coloured appearance, which indicates a red shift in absorbance spectrum due to decreased red reflection and increased blue and green reflection. The far red fluorescence brightness of bacterial colonies expressing these red shifted phenotype variants were compared using the fluorescent colony screening system, and the brightest

variant was selected and sequenced. This mCherry2-I161G/Q163G variant had a significantly red shifted fluorescence excitation maximum at 598 nm, and emission maximum at 625 nm (Figure 5.5C), which was similar to the fluorescence spectral profile of mRaspberry<sup>29</sup>. This mCherry2-I161G/Q163G far-red fluorescent variant was named as RDSmCherry0.1 and used as the template for further development and optimization. In an attempt to mimic the strategy of creating a  $\pi$ - $\pi$  stacking interaction with the chromophore (as used in avGFP to create YFP) for further red shifted fluorescence, saturation mutagenesis was performed at residues 195 and 197. Residue 197 is structurally aligned with the key  $\pi$ - $\pi$  stacking residue (tyrosine 203) of YFP. Screening for the most red-shifted mutant in the 195X/197X library led to the variant mCherry2-I161G/Q163G/V195C/I197Y, termed as FSSmCherry0.2 (Figure 5.5A). Despite the successful introduction of an aromatic amino acid tyrosine in the position 197, only a moderate red shift occurred to the variant, in both excitation (to 600 nm) and emission (to 630 nm). Based on RDSmCherry0.2, multiple rounds of saturation mutagenesis and EP-PCR produced RDSmCherry0.5 with its excitation maximum red-shifted to 604 nm and emission maximum to 636nm (Figure 5.5A). Compared with its template mCherry2, RDSmCherry0.5 has 7 mutations: V16S, A44C, A145P, I161G, Q161G, V195C, and I197Y (Figure 5.4B). All of these mutations are inside the  $\beta$ -barrel pointing towards the chromophore with the exception of A145P. We presume that the proline substitution may help to improve fluorescence by slight adjustment of the positions of adjacent amino

acids, such as residue 146 which does have its side chain directed towards the chromophore.

During the detailed spectral characterization of RDSmCherry0.5, it was found that this variant had a distinctive blue fluorescent component with excitation at 400 nm and emission at 462 nm (Figure 5.5C). Characterizing RDSmCherry0.1 and 0.2 showed that RDSmCherry0.2 had a similar blue fluorescence of excitation at 400 nm and emission at 466 nm (Figure 5.5C), while RDSmCherry0.1 had only a single excitation peak at 598 nm. This led to the speculation that mutations at position 195 and 197 caused the undesirable blue fluorescence. Therefore, another round of saturation mutagenesis on 195 and 197 was performed on the template of RDSmCherry0.5 with the goal of preventing formation of the blue fluorescent species. This effort led to the identification of mCherry2-V16S/A44C/A145P/I161G/Q161G/V195C, which was designated as RDSmCherry. RDSmCherry is slightly blue shifted from RDSmCherry0.5 with red excitation at 600 nm and far-red emission at 630 nm. The key reversion mutation of tyrosine 197 back to the original isoleucine abolished the blue fluorescence while increasing the intensity of the far-red fluorescence.



**Figure 5.5** Directed evolution of RDSmCherry. A, Normalized fluorescence excitation (dashed line) and emission spectra (solid line) of RDSmCherry0.1, RDSmCherry0.2 and RDSmCherry0.5. B, Mutation table of RDSmCherry. C, Blue fluorescence of RDSmCherry0.2 and RDSmCherry0.5. D, Critical mutation positions of RDSmCherry (based on mCherry crystal structure, PDB: 2H5Q).

#### 5.2.4 Spectroscopic characterization of LSSmCherry and RDSmCherry

The spectroscopic properties of LSSmCherry were characterized and compared with those of LSSmKate1 ( $\lambda_{\text{ex}} = 463 \text{ nm}$ ;  $\lambda_{\text{em}} = 624 \text{ nm}$ ), LSSmKate2 ( $\lambda_{\text{ex}} = 460 \text{ nm}$ ;  $\lambda_{\text{em}} = 605 \text{ nm}$ ), mKeima ( $\lambda_{\text{ex}} = 440 \text{ nm}$ ;  $\lambda_{\text{em}} = 620 \text{ nm}$ ), and mBeRFP ( $\lambda_{\text{ex}} = 460 \text{ nm}$ ;  $\lambda_{\text{em}} = 615 \text{ nm}$ ) (Table 1). The excitation spectrum of LSSmCherry has a single excitation peak, similar to those of LSSmKate1 and LSSmKate2<sup>37</sup>. In contrast, mKeima<sup>132</sup> and mBeRFP<sup>38</sup> excitation spectra revealed additional peaks corresponding to regular Stokes shift red fluorescence phenotypes. The molar extinction coefficient (EC) and quantum yield (QY) of LSSmCherry was determined to be  $37\,000 \text{ M}^{-1} \text{ cm}^{-1}$  and 0.17, respectively (Table 1). LSSmCherry has a fluorescence brightness (the product of the extinction coefficient and quantum yield) substantially higher than that of LSSmKate and LSSmKate2<sup>37</sup>.

**Table 5.1** Spectral properties of mCherry, mCherry2, LSSmCherry and other LSS RFPs

FP	Excitation Maxima (nm)	Emission Maxima (nm)	Stokes Shift (nm)	Extinction Coefficient (M <sup>-1</sup> cm <sup>-1</sup> )	Quantum Yield	Brightness* (10 <sup>3</sup> )
mCherry	587	610	23	72,000	0.22	15.8
mCherry2	589	610	21	79,400	0.22	17.5
LSSmCherry	450	610	160	35,000	0.29	10.1
LSSmApple	460	595	135	45,000	0.31	13.9
LSSmKate1	463	624	161	31,200	0.08	2.5
LSSmKate2	460	605	145	26,000	0.17	4.4
mKeima	440	620	180	13,400	0.24	3.2
mBeRFP	446	611	165	65,000	0.27	17.5

\* Brightness is defined as the product of extinction coefficient and quantum yield.

RDSmCherry, along with its predecessors 0.1, 0.2, and 0.5, were characterized and compared with other typical far red FPs derived from mCherry or mKate, including mRouge ( $\lambda_{\text{ex}} = 600 \text{ nm}$ ;  $\lambda_{\text{em}} = 637 \text{ nm}$ )<sup>30</sup>, mCardinal ( $\lambda_{\text{ex}} = 604 \text{ nm}$ ;  $\lambda_{\text{em}} = 659 \text{ nm}$ )<sup>33</sup> and TagRFP659 ( $\lambda_{\text{ex}} = 611 \text{ nm}$ ;  $\lambda_{\text{em}} = 659 \text{ nm}$ )<sup>34</sup> (Table 2). RDSmCherry has an EC of 55,400 M<sup>-1</sup>cm<sup>-1</sup> and a QY of 0.09. Compared with the mKate derivatives mCardinal and TagRFP659, RDSmCherry was less red-shifted in its emission wavelength, and significantly dimmer in term of the fluorescence

brightness. RDSmCherry was similar to mCherry-based far-red FP mRouge in terms of excitation and emission spectra, with a relatively higher overall brightness.

**Table 5.2** Spectral properties of mCherry2, RDSmCherry and other far-red RFPs

<b>FP Name</b>	<b>Excitation Maxima (nm)</b>	<b>Emission Maxima (nm)</b>	<b>Extinction Coefficient (M<sup>-1</sup> cm<sup>-1</sup>)</b>	<b>Quantum Yield</b>	<b>Brightness* (10<sup>3</sup>)</b>
mCherry2	589	610	79,400	0.22	17.5
RDSmCherry0.1	598	625	59,600	0.10	6.0
RDSmCherry0.2	600	630	34,400	0.03	1.0
RDSmCherry0.5	604	636	23,300	0.02	0.5
RDSmCherry	600	630	55,500	0.09	5.0
mRouge	600	637	43,000	0.02	0.9
TagRFP659	611	659	29,000	0.10	2.9
mCardinal	604	659	87,000	0.19	16.5

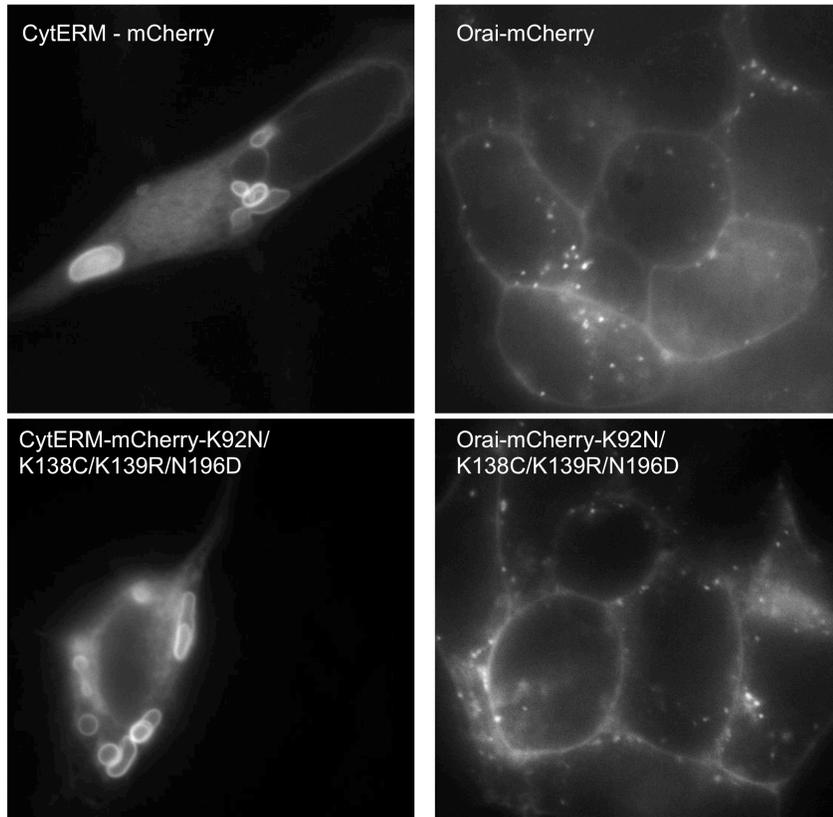
\* Brightness is defined as the product of extinction coefficient and quantum yield.

### 5.2.5 Identification of mCherry variants with low cytotoxicity

In the attempt to investigate the residues in mCherry2 responsible for the reduction in cytotoxicity, a number of variants were generated with different combinations of mCherry2 mutations (comparing with mCherry) using site directed mutagenesis. Unexpectedly, one of the variants, mCherry-K92N/S147T/N196D, allowed faster growth rate when compared with EGFP (Figure 5.3), which suggested that it had an even lower cytotoxicity than EGFP and mCherry2.

It is commonly assumed that cytotoxicity of FPs is due to formation of oligomers or higher order aggregates that can interfere with normal cellular functions. This impression is bolstered by the fact that many FPs, including mCherry, are sometime observed to form bright puncta, which are often assumed to be aggregates, when expressed in mammalian cells. Motivated by such reasoning, we wondered the low-cytotoxicity mCherry variants might be particularly useful for imaging of fusion proteins in mammalian cells. Accordingly, we tested two particular fusion protein constructs: CytERM and Orai1. The CytERM is an endoplasmic reticulum (ER) signal-peptide which anchors its downstream protein to the cytoplasmic side of ER<sup>167</sup> and Orai1 is a Ca<sup>2+</sup> selective ion channel named as Ca<sup>2+</sup> release-activated Ca<sup>2+</sup> channel protein 1<sup>168</sup>. These two fusion partners were chosen because they were previously reported as “challenging” targets for FPs. Both CytERM-mCherry and Orai-mCherry displayed mislocalization (bright circular structure) and aggregation (bright puncta) when expressed in mammalian

cells. Unfortunately, the low-cytotoxic mCherry variant showed similar behaviour to mCherry in these two specific fusions (Figure 5.6).



**Figure 5.6** Representative fluorescence images of cells expressing mCherry fusions.

### 5.3 Discussion

In the engineering of LSSmCherry, similar to LSSmKate1<sup>39</sup>, the introduction of a glutamic acid at position 163 resulted in the introduction of an ESPT pathway from the chromophore tyrosine phenol group. At the ground state of the

chromophore, the  $pK_a$  of the glutamic acid 163 carboxylate is much lower than the  $pK_a$  of the tyrosine 66 side chain hydroxyl. This resulted in a stabilized RFP chromophore in the protonated form, which has a characteristic excitation wavelength in the blue region at  $\sim 450$  nm. When excited, a deprotonated chromophore is formed due to ESPT using the glutamic 163 as a proton acceptor, which is followed by the red emission at around 610 nm. Mutation W143L in LSSmCherry0.2 greatly facilitated the LSS fluorescence through elimination of the green and regular Stokes shift red fluorescent species mixture that co-existed in LSSmCherry0.1. The leucine residue might stabilize the chromophore in a conformation that is preferred for LSS fluorescence.

When developing LSSmCherry, the screening criteria during directed evolution was bright LSS red fluorescence and minimal regular Stokes shift red fluorescence. By enforcing these criteria, we hoped to develop a variant that would be compatible with other RFPs for multi-colour imaging. Although the final variant LSSmCherry is somewhat dimmer than the recently developed mBeRFP<sup>38</sup>, it has the advantage of having no red fluorescence from regular Stokes shift. On the other hand, mBeRFP has a substantial red excitation peak at 576 nm, making it incompatible for concurrent imaging in combination with regular RFPs. Due to the large excitation emission separation in the fluorescence spectra, LSS RFPs could serve as a third colour in fluorescence imaging with GFP and RFP. Among all the single-excitation LSS RFPs, LSSmCherry will be a favourable alternative to LSSmKate1 and LSSmKate2, due to the advantage of higher fluorescence

brightness. LSSmCherry also offers a larger Stokes shift and a similar fluorescence brightness when comparing to LSSmApple,

The RDSmCherry0.1 was first identified by the blue appearance of a colony expressing this red shifted mCherry mutant. The two glycines at position 161 and 163 were responsible for shifting the excitation and emission by 11 nm and 15 nm. This supposedly occurs by destabilizing the ground state of the negatively charged chromophore. The introduction of tyrosine at position 197 in the second round of directed evolution, further red-shifted the fluorescence profile, but at the same time a species with blue fluorescence was introduced. Although the tyrosine at 197 resulted in a red shift by stabilizing the excited state with a  $\pi$ -stacking interaction, very likely it caused a trapped chromophore intermediate with blue fluorescence as well<sup>19</sup>. This phenomenon is as described previously in the development of TagRFP derived blue FPs mTagBFP and mTagBFP2<sup>163, 169</sup>. The following rounds of evolution led to the most red-shifted variant, RDSmCherry0.5, but the blue fluorescent species was still present. Thus in the final round, the tyrosine 197 was mutated back and this change successfully eliminated the blue fluorescence in the final variant RDSmCherry. This result confirmed that Y197 was the cause of the blue fluorescence. Unfortunately, reversion of the Y197 mutation caused the far-red fluorescence of RDSmCherry to be modestly blue-shifted compared with RDSmCherry0.5. Presumably this shift is due to the loss of chromophore  $\pi$ - $\pi$  interaction with the tyrosine residue.

The development of RDSmCherry0.5 parallels the development of mRouge in terms of the choice of starting template, screening criteria, critical mutation positions and the final variant spectroscopic properties<sup>30</sup>. However, the differences between these two studies are also clear. First of all, mRouge was developed using a subset of the full amino acid library at a given residue with the aid of computational library construction. In contrast, positions in RDSmCherry0.5 were experimentally saturated with all of the 20 amino acids using degenerate codon NNK and then subjected to library screening. Despite the fact that these two variants share some of the mutated positions, the actual mutations are drastically different (Table 3). These variations could be attributed to the differences in the library design strategies as well as the initial templates: mCherry2-I161G/Q163G for RDSmCherry and mCherry-V195T/I197T/A217N for mRouge. Unfortunately, due to its low fluorescence brightness, it is unlikely for RDSmCherry to find applications in deep tissue imaging. There are a number of superior variants, such as the mKate far-red derivative mCardinal, that have already been demonstrated to be useful for intravital imaging<sup>33</sup>.

The mutagenesis library of mCherry2-161X/163X contained a number of spectrally varied phenotypes, including LSSmCherry0.1 and RDSmCherry0.1 as described in this thesis. The close proximity between these two residues and the protein chromophore phenolate group provides a unique opportunity for spectral property modulation by mutating these two amino acids<sup>9</sup>. Notably, residues 161 and 163 have been heavily mutated all through the development of DsRed derived

FPs. For example, in the mFruit series, residue 163 was mutated from lysine in DsRed to methionine in mRFP1 and mOrange, and glutamine in mCherry and dTomato<sup>12</sup>. In addition, position 161 and 163 were also substituted during the engineering of photoswitchable mCherry variants rsCherry and rsCherryRev<sup>148</sup>.

The development LSSmCherry and RDSmCherry presents new opportunities for multi-colour fluorescence imaging as well as FRET pair construction. Pairing LSSmCherry with RDSmCherry as a FRET donor and acceptors could enable single wavelength excitation four colour FRET when combined with a CFP/YFP FRET pair. LSSmCherry and RDSmCherry could also be new templates for genetically encoded biosensor engineering. Finally, the mutational insights acquired in this study could be applied in the future RFP design as well as the further colour palette expansion of currently available RFP-based biosensors.

## **5.4 Materials and Methods**

### **5.4.1 General methods and materials**

All synthetic DNA oligonucleotides for cloning and library construction were purchased from Integrated DNA Technologies. *Taq* DNA polymerase (New England Biolabs) was used for error-prone PCR (EP-PCR). PCR products and products of restriction digests were purified using GeneJET gel extraction kit (Thermo Scientific) according to the manufacturer's protocols. Restriction enzymes and ligases were purchased from New England Biolabs or Thermo

Scientific. The cDNA sequences were confirmed by dye terminator cycle sequencing using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Sequencing reactions were analyzed at the University of Alberta Molecular Biology Service Unit.

#### **5.4.2 Site directed saturation mutagenesis and random mutagenesis**

Directed evolution of LSSmCherry and RDSmCherry was carried out by site directed saturation mutagenesis and EP-PCR using plasmids encoding mCherry2 as template. All site-directed mutagenesis was performed using the Quikchange lightning mutagenesis kit (Agilent) and primers designed according to the manufacturers guidelines with the degenerate codon (NNK) for the intended mutation positions. EP-PCR was performed using unbalanced dNTP concentrations and increased concentration of MgCl<sub>2</sub> (50 mM), and the introduction of MnCl<sub>2</sub> (10 mM) to further decrease the fidelity of *Taq* DNA polymerase. Products were digested with XhoI and HindIII and ligated into pBAD/His B vector (Life technologies) digested with the same two enzymes, and used to transform the electrocompetent *E. coli* strain DH10B ElectroMax (Life technologies) using a Micropulser electroporator (Bio-Rad), which were then plated on agar plates containing LB medium supplemented with 0.4 mg/ml ampicillin and 0.02% (w/v) L-arabinose. Plates were incubated for overnight at 37 °C prior to screening.

### 5.4.3 Library screening

*E. coli* colonies expressing the mutation libraries were grown on 10 cm Petri dishes. In order to screen libraries for variants that exhibited brighter LSS red fluorescence, a screening method was developed for LSS RFPs when expressed in colonies of *E. coli*. The custom built fluorescent colony imaging system equipped with filter sets (Chroma) for 470/40 nm excitation with 630/60 nm emission (i.e. for LSS fluorescence) and 560/40 nm excitation with 630/60 nm emission (i.e. for red fluorescence) is used to acquire both LSS and red fluorescence images of the Petri dish. Colonies exhibiting the highest intensity ratio (LSS/Red) between these channels were picked and cultured for further spectral confirmation. For far-red FP screening, colony images were acquired with filter set 622/36 nm excitation and 680/40 nm emission. Colonies exhibiting the highest intensities were picked and cultured for further spectral confirmation. Single colonies were picked, inoculated into 4 ml of LB medium with 0.1 mg/ml ampicillin and 0.02% (w/v) L-arabinose and then cultured overnight. Protein was extracted using B-PER bacterial extraction reagent (Thermo) as per manufacturer guidelines. Screening of the extracted protein was performed with a Safire2 fluorescence plate reader (Tecan) by measuring protein fluorescence. The brightest LSS variant or most red shift variant was selected to be the template for the following round of library creation.

#### 5.4.4 Protein purification and *in vitro* characterization

To purify the FPs, electrocompetent *E. coli* strain DH10B Electromax (Life technologies) was transformed with the plasmid of interest using Micropulser electroporator (Bio-Rad). Transformed bacteria were cultured overnight on agar plates containing LB and 0.4 mg/ml ampicillin. A single colony was picked and grown overnight in 4 mL LB supplemented with 0.1 mg/ml ampicillin at 37 °C. The 4 mL culture was then used to inoculate 250 mL of LB medium with ampicillin and grown to an optical density of 0.6. Protein expression was induced with the addition of 0.02% L-arabinose and the culture was grown overnight at 37 °C. Bacteria were harvested at 10,000 rpm, 4 °C for 10 min, lysed using a cell disruptor (Constant Systems) and then clarified at 14,000 rpm for 30 min. Protein was purified from the supernatant by Ni-NTA affinity chromatography (ABT) according to the manufacturer's instructions. The buffer of the purified protein was exchanged with 10 mM Tris-Cl, 150 mM NaCl, pH 7.3 with Amicon ultracentrifugal filter (MWCO 10,000). Molar extinction coefficients (EC) were measured by the alkali denaturation method. Briefly, the protein was diluted into Tris buffer or 1 M NaOH and the absorbance spectra recorded under both conditions. The EC was calculated assuming the denatured RFP chromophore has an EC of 44,000 M<sup>-1</sup>cm<sup>-1</sup> at 452 nm. Fluorescence quantum yields (QY) were determined using LSSmKate2 or mCherry as standards. Fluorescence intensity as a function of pH was determined by dispensing 2 µL of the protein solution into 50 µL of the desired pH buffer in triplicate into a 384-well clear-bottomed plate (Nunc) and measured in a Safire2 plate reader. pH buffer solutions from pH = 3 to

pH = 11 were prepared according to the Carmody buffer.

#### **5.4.5 Bacterial growth rate comparison**

Electrocompetent *E. coli* was transformed with RFP or EGFP encoding plasmid with the same pBAD/His B vector. The transformed bacteria were then cultured separately in two tubes of LB medium with 0.1 mg/ml ampicillin overnight. 50  $\mu$ L of *E. coli* cultures from the each tubes was mixed and cultured in a tube with 4 mL fresh LB medium supplemented with 0.1 mg/ml ampicillin and 0.02% (w/v) L-arabinose for 14 to 16 hours. Inoculation from the mixed culture to fresh medium was repeated every day for 4 days. Each overnight culture was plated on agar plates supplemented with 0.4 mg/ml ampicillin and 0.02% (w/v) L-arabinose. All the plates were imaged using a custom built colony imaging system under green (470/40 nm excitation and 510/20 nm emission) and red (560/40 nm excitation and 630/60 nm emission) channel respectively. Percentages of EGFP and RFP-expressing bacteria were calculated based on the fluorescent colony counts obtained from the images.

#### **5.4.6 Live cell imaging**

To construct Orai1-mCherry and CytERM-mCherry, the gene of mCherry was amplified with a 5' primer with an *AgeI* site and a 3' primer with a *NotI* site. The purified PCR products were digested and ligated into a similarly digested pOrai1-YFP or pCytERM-mEGFP. The plasmid encoding Orai-YFP was acquired from

Addgene, plasmid number 19756. The plasmid encoding CytERM-mEGFP was gifted from professor Erik Snapp at Albert Einstein College of Medicine. HeLa cells and HEK cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS, Sigma) and Glutamax (Life Technologies) and incubated at 37°C with 5% CO<sub>2</sub>. Transient transfections were performed using Turbofect (Thermo Scientific) according to the manufacturer's guidelines. Transfected cells were imaged using either an Axiovert 200M (Zeiss) or a Nikon Eclipse Ti. The Axiovert 200M (Zeiss) was equipped with a 75 W xenon-arc lamp, a 40× objective lens (NA = 1.3, oil), a 14-bit CoolSnap HQ2 cooled CCD camera (Photometrics), and driven by open source Micro-Manager software. The Nikon Eclipse Ti microscope was equipped with a 150 W Lumen 200 metal halide lamp (Prior Scientific), a 16-bit 512SC QuantEM CCD (Photometrics), a 25% neutral density filter, a 40× objective (NA = 0.95, air), and driven by a NIS-Elements AR 3.0 software package (Nikon).

## **Chapter 6 Filter Paper Blot Screening for RFP-based Sensor Development**

## 6.1 Introduction

Fluorescent proteins (FPs) are members of a structurally homologous class of proteins that share the unique property of being self-sufficient to form a visible wavelength chromophore from a sequence of three amino acids within their own polypeptide sequence<sup>9, 15</sup>. The inherent properties of the chromophore, coupled with the influence of the surrounding protein structure, make it possible to engineer mutant variants sensitive to particular analytes<sup>170</sup>. Various genetically encoded biosensors based on FPs have been created, including  $\text{Ca}^{2+}$ <sup>83</sup>, pH<sup>66</sup>, ATP<sup>43</sup>, membrane potential<sup>45</sup> and enzyme activities<sup>171</sup>. These FP-based sensors have become indispensable imaging tools for investigating various problems in cell biology and neuroscience.

For the development of FP-based sensors, rational design is typically first used to obtain a functional sensor prototype. Iterative directed evolution is then used in order to optimize the functionality of the prototype. Large libraries of FP-based biosensor mutants will be generated and be screened for desired sensor properties, such as the sensitivity, dynamic range, and sensor brightness. Library screening is an extremely important, yet labour intensive, step in the cycles of directed evolution. Conventional screening methods like random selection usually are not very efficient, as it requires a secondary screening step to extract the proteins and test their functionalities in a multi-well plate form. Some more technologically sophisticated screening techniques, such as fluorescence-activated cell sorting (FACS), require expensive and complicated instrumentation. Being able to screen

a large number of protein variants in the context of bacterial colonies would be greatly beneficial to FP-based biosensor development. One example of colony-based screening is to export protein candidates to the bacterial periplasmic space. This enables the partial control of the sensor candidate environment, which allows the screening of sensor response. Using this prescreening method, thousands of bacterial colonies expressing diversified genetically encoded sensors candidates were screened and a number of new  $\text{Ca}^{2+}$  biosensors were identified<sup>83</sup>. However, the size of the exporting protein and exportable amount of the protein are limited in this method<sup>172</sup>. Moreover, due to the buffering effect of growth medium, relatively large amount of analyte solution is often required.

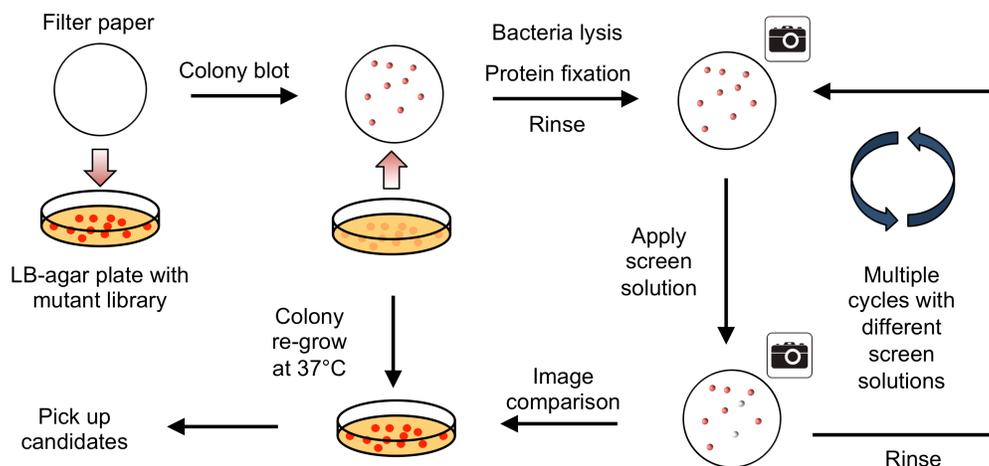
In the present work, a new screening method has been developed based on filter paper blotting method<sup>173, 174</sup>. This screening assay allows simple and efficient identification of potential FP-based sensors expressed in *E. coli*. Briefly, bacterial colonies with mutant libraries are blotted on filter paper, followed by an on-paper lysis performed via freeze-thaw cycles to lyse the cells and expose the expressed FPs. Fixative is then used to crosslink and fix proteins on the filter paper and prevent protein diffusion in the following steps. With fixed proteins, the filter paper can be treated with analyte solution. Thus, sensor candidates that are sensitive to specific analytes can be identified according to the fluorescence change. This screening method is validated using the existing analyte sensitive FPs including a pH-sensitive RFP and a  $\text{Cl}^-$  sensitive YFP. In addition, new  $\text{Ca}^{2+}$  sensitive RFPs have been identified through this screening method. These new

Ca<sup>2+</sup> sensors, with a binding site built into the RFP barrel<sup>71, 175</sup>, were found to have low-affinity Ca<sup>2+</sup> binding.

## **6.2 Results**

### **6.2.1 Development of a filter paper blot based assay for screening**

The procedure of this screening method is performed as follow. First, *E. coli* colonies expressing FP libraries are blotted on filter paper. Subsequently, on-paper lysis is performed via freeze-thaw cycles to expose the expressed FPs. Fixative such as formaldehyde solution is then used to crosslink and fix FPs on the filter paper. After the fixation, the filter paper can be rinsed or treated with multiple rounds of analyte solutions without substantial protein diffusion. Fluorescence images are taken before and after the treatment of the analyte solution. Sensor candidates that are sensitive to specific analytes can be identified according to the fluorescence intensity change from the comparison of images (Figure 6.1).

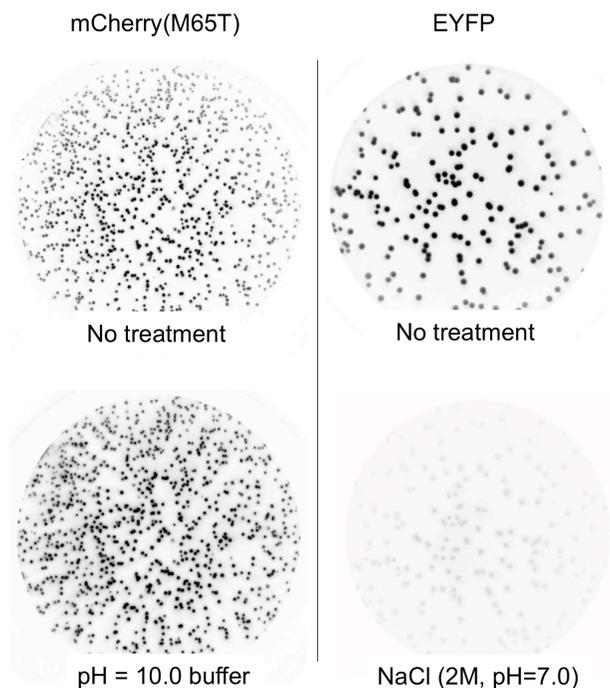


**Figure 6.1** Schematic presentation of the filter paper blot screening method procedure.

This screening method utilizes a simple and low-cost blotting method to transfer colony libraries onto a filter paper, on which the expressed proteins can be exposed, fixed and then screened for desired functionality as a sensing probe. In this screening method, the filter paper serves as the medium for the mutant protein library. In parallel, colony residues on the original plate can be re-grown and colonies of interest can be later picked up for DNA amplification and isolated. Therefore, the plate serves as the medium for the mutant DNA library. Since the colony patterns on both plate and paper are the same, when a protein mutant shows fluorescence change upon treating a specific analyte, the corresponding colony can be quickly identified on the plate. Further more, as the protein is fixed on the paper, it is also possible to perform multiple cycles of screening with different analytes. Compared to conventional screening methods, filter paper blot

screening is relatively time and cost efficient, with minimum instrumentation requirements. In terms of throughput, libraries with size of  $10^3$  can be screened in one single day, as opposed to the  $10^2$ -screening throughput by random selection.

This filter paper blot screening method was validated using a pH-sensitive variant (Met65Thr) of RFP mCherry as well as a chloride ion ( $\text{Cl}^-$ ) sensitive yellow FP EYFP<sup>68</sup>. After colony blotting and freeze-thaw lysis, the filter paper blotted with bacteria expressing mCherry M65T variant was treated with a pH 10 buffer. Fluorescence imaging before and after treatment revealed a substantial increase in red fluorescence after the buffer treatment. Similarly, the yellow fluorescence of EYFP-expressing colonies was quenched after the applying a solution with high concentration of  $\text{Cl}^-$  ( $\text{NaCl}$ , 2 M, pH=7.0) (Figure 6.2). These results demonstrated that the method could be used to differentiate the fluorescence changes upon the analyte solution treatment, for both fluorescence increases and decreases.



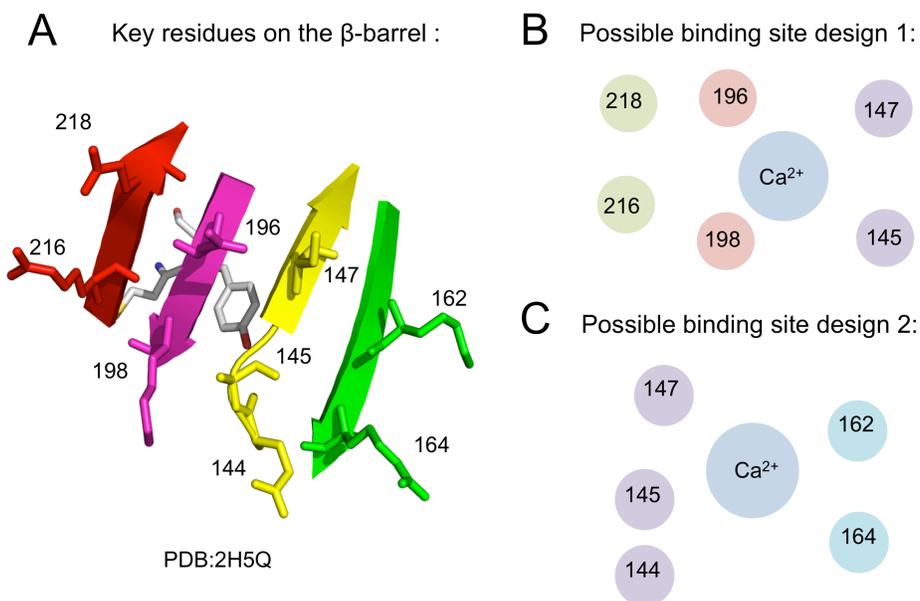
**Figure 6.2** Validation of the on filter fluorescence change. Fluorescence images were taken using the same filter set and exposure time for each protein sensor variant. Images are inverted for illustration purpose.

### 6.2.2 Library construction and $\text{Ca}^{2+}$ responsive variant screening

With the demonstration of the capability of detecting fluorescence changes using the filter paper blotting, we further explored the use of this approach for screening of *in vitro* FP-based sensor libraries. A strategy<sup>71</sup> was previously reported for designing and engineering a  $\text{Ca}^{2+}$  binding site into GFP through introducing charged residues, such as aspartic acid (D) or glutamic acid (E), on the barrel near chromophore. The resulted  $\text{Ca}^{2+}$ -sensitive variant, known as CatchER, was used

to detect the  $\text{Ca}^{2+}$  dynamics in high  $\text{Ca}^{2+}$  concentration cellular compartments such as endoplasmic reticulum (ER).

The same design strategy was employed for the RFP based  $\text{Ca}^{2+}$  sensors in this work. The scaffold of this designed  $\text{Ca}^{2+}$  sensitive RFP was based the monomeric RFP mApple. The binding site is designed to reside at the close proximity to the chromophore phenolate group. The rationale for this design was that, upon  $\text{Ca}^{2+}$  binding, the change on chromophore environment would lead to a change in the chromophore environment and a detectable fluorescence change. Accordingly, residues 144, 145, 147, 162, 164, 196, 198, 216, and 218 (numbering according to the mCherry crystal structure, PDB: 2H5Q) were selected for site-directed mutagenesis (Figure 6.3A). Two putative  $\text{Ca}^{2+}$  binding sites were designed to provide potential feasible binding geometries for  $\text{Ca}^{2+}$  ion (Figure 6.3B, 6.3C). Corresponding to these two possible  $\text{Ca}^{2+}$  binding site design, two mutant libraries were constructed.



**Figure 6.3**  $\text{Ca}^{2+}$  binding site design and library construction. A, residues for mutation library design. B and C, potential  $\text{Ca}^{2+}$  binding sites on the barrel of the FP.

Through the filter paper blotting screening method, we successfully identified two  $\text{Ca}^{2+}$  responsive variants in the library of  $\text{Ca}^{2+}$  binding site design 1 where residues 145, 147, 196, 198, 216, and 218 were mutated. These two variants were designated as Cas-RFP (C $\text{a}^{2+}$  sensitive red fluorescent protein) 0.1a and 0.1b. Cas-RFP 0.1a had three mutations to acid residues (D/E) on the barrel structure of the mApple protein, constructing a putative  $\text{Ca}^{2+}$  binding site. Upon  $\text{Ca}^{2+}$  binding, the fluorescence intensity of the Cas-RFP 0.1a increases by 2.5-fold with a dissociation constant ( $K_d$ ) of 1.03 mM. As for Cas-RFP 0.1b, it possessed an additional acid residue mutation R216E comparing to 0.1a, which might attribute

to the somewhat higher affinity for  $\text{Ca}^{2+}$  ( $K_d = 0.89$  mM). Cas-RFP 0.1a was then used as a template to further increase the fluorescence intensity change upon  $\text{Ca}^{2+}$  binding by site directed mutagenesis and random mutations. Coupled with the filter paper screening, Cas-RFP0.2 was identified with a 5-fold increase of red fluorescence emission with  $\text{Ca}^{2+}$  binding. But in the meantime, the  $\text{Ca}^{2+}$  binding affinity was significantly lowered to 1.59 mM. Due to the extremely low affinities to  $\text{Ca}^{2+}$ , Cas-RFPs were unable to detect the  $\text{Ca}^{2+}$  dynamics in the ER compartment of live mammalian cell (data not shown).

**Table 6.1** Characterization of  $\text{Ca}^{2+}$  responsive fluorescent proteins

<b>Protein</b>	<b>Mutations (to mApple)</b>	<b>Intensity fold-change</b>	<b><math>K_d</math> (mM)</b>	<b><math>n_H</math></b>
Cas-RFP0.1a	A145D, D196E, K198E	2.5×	$1.03 \pm 0.08$	1.1
Cas-RFP0.1b	A145D, D196E, K198E, R216E	2×	$0.89 \pm 0.02$	1.4
Cas-RFP0.2	F65L, A145D, I161A, K163G, D196E, K198E	5×	$1.59 \pm 0.02$	1.6

Whereas in the other library of 144, 145, 147, 162, and 164 mutation sites, no  $\text{Ca}^{2+}$  responsive variant were identified using the filter paper based screening method. This library was then subjected to random selection with multi-well plate test against  $\text{Ca}^{2+}$  solution. The result showed that none of the randomly picked variants had notable fluorescence changes upon  $\text{Ca}^{2+}$  solution treatment. This result suggested that the negative result from the filter paper blotting method was

valid. We speculate that the residues mutated in this library were not suitable for  $\text{Ca}^{2+}$  binding due the relative positions of the  $\beta$ -strands where these residues located.

### 6.3 Discussion

In this work, a general screening scheme for the development of FP-based biosensor as established, allowing the identification of analyte responsive FPs from large libraries of variants. The principle of this screening method relies on the bacterial colony transfer and lysis conducting on a filter paper. This filter paper blot screening method can be used to efficiently screen FP mutant libraries and identify potential sensors with desired functionality. The main advantages of our methodology are: (1) it is simple, cost effective, and requires no special instrumentation or specialties; (2) it has the potential to be universal for any *in vitro* evolution of FP-based biosensors; (3) it allows screening of a large number of variants; (4) it is sensitive and enables detection of small improvements; (5) any existing protein expression system can be used without prior DNA manipulations.

We demonstrated the utility of this screening method by developing a series of  $\text{Ca}^{2+}$  responsive RFPs. Following construction and screening of two libraries of potential  $\text{Ca}^{2+}$  binding RFPs, variants with  $\text{Ca}^{2+}$  binding affinities at mM level were successfully isolated. The somewhat lower affinities of the red sensors compared with their green counterpart CatchER<sup>71, 176</sup> hampered the applications in

live cell fluorescence imaging of ER  $\text{Ca}^{2+}$  dynamics. This affinity difference is probably due to the less favoured geometry of the surface residues on the FP. Although the structure alignment shows all the residues located in a similar setting, the mutations could lead to drastically different conformations. Another contributing reason might be that EGFP has a proton “wire” in close proximity to position 148<sup>154</sup>, but no equivalent wire has been reported in the RFP templates.

The uses of this filter paper blot screening method in the present work is but one example of the many possible applications of this simple and efficient screening technique. It is also well suited to screen for other single FP-based ion sensors, such as  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Mg}^{2+}$ , and  $\text{Cl}^-$  sensitive FPs. The future direction of this work will be using this robust screening method to develop more new FP-based sensors as useful tools for biological and medical fluorescence imaging.

## **6.4 Materials and methods**

### **6.4.1 Mutagenesis library construction and molecular cloning**

To construct  $\text{Ca}^{2+}$  sensitive RFP, pBAD-mApple was used as a template. Point mutations and randomizations of specific amino acid were introduced using QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent) as per manufacturer’s instructions. Oligonucleotides containing specific mutations were designed in the aid of Agilent online mutagenesis primer design program. Random mutagenesis was performed using error-prone PCR amplification with

*Taq* polymerase (Bio-Rad), the amplified PCR product was digested using FastDigest XhoI and HindIII restriction enzymes (Thermo Scientific) and then ligated to bacterial expression vector pBAD/His-B (Life technologies) using a T4 DNA ligase (Life technologies). Plasmid purification was performed using GeneJet plasmid miniprep kits (Thermo Scientific) according to the the manufacturers instructions, DNA sequencing was done using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosciences) and analyzed by the University of Alberta, Molecular Biology Services Unit.

#### **6.4.2 Filter paper based screen assay**

Electrocompetent *E. coli* (DH10B EletroMax, Life technologies) cells were transformed with mutated DNA library and then plated at different dilution levels on agar plates supplemented with 0.4 mg/ml ampicillin (Fisher Scientific) and 0.02% (wt/vol) L-arabinose (Alfa Aesar). Plates were incubated at 37°C overnight. Plates with colonies numbers in the range of 200 to 500 are used for screening. Colonies were blotted on to a filter paper (Waterman #5 or VWR #413) by overlaying the paper on the agar plate and then peeling off. The plates with residue colonies were incubated in 37°C for another 6 to 10 hours for colony re-growing. The filter paper with colonies was freezed in liquid nitrogen (-80C) for around 10 seconds and then thawed in room temperature for 1 min to lyse the bacteria. Freeze/thaw cycle was repeated for 2 times. Fixation of the protein was performed by soaking the filter paper into the fixation solution of 4% paraformaldehyde in phosphate buffered saline (pH = 7.2) for 5 min<sup>177</sup>. The filter

paper was first imaged using custom-built fluorescence imaging system described previously<sup>95</sup> with proper optical filter settings, then the paper is soaked into solutions with analyte. Colony fluorescence was imaged again after the treatment. By aligning two fluorescence images and compare the change of fluorescence intensity, analyte sensitive variants were identified; the corresponding colonies were picked up from the plate and then cultured for plasmid extraction and protein purification.

#### **6.4.3 Protein purification and *in vitro* characterization**

A single colony containing plasmid of desired variant was picked and grown overnight in a culture tube at 37°C with 5 ml LB medium supplemented with 0.1 mg/ml ampicillin. The 5 ml culture tube was inoculated into 500 ml of LB medium supplemented with 0.1 mg/ml ampicillin. After 6 hours of shaking at 37°C, arabinose was added to a final concentration of 0.02% (wt/vol) and continued shaking overnight. Bacteria were harvested by centrifuge at 10,000 rpm, 4°C for 10 min and the pellet is resuspended at 4°C in 10 mM Tris-Cl, 150 mM NaCl, pH 7.4 and then lysed using a cell disruptor. The proteins were purified from the supernatants using Ni-NTA (MCLAB) affinity chromatography. Proteins were eluted with 300 mM imidazole, which was then removed by buffer exchange with 10 mM PBS or 10mM Tris, pH 7.4 using centrifugal filter tubes (Amicon, MWCO 10,000). Proteins were stored at 4°C prior to spectral characterization.

The apparent  $K_d$  for  $\text{Ca}^{2+}$  was determined by mixing the purified protein with MOPS (3-(N-morpholino) propanesulfonic acid, 30mM, pH = 7.2) buffers containing various amount of  $\text{Ca}^{2+}$  ranging from 0.05 mM to 10mM. 2  $\mu\text{l}$  of the protein was mixed with 50  $\mu\text{l}$  of each  $\text{Ca}^{2+}$  buffer into a 384-well plate and the fluorescence emission was recorded using the plate reader. Each emission peak was integrated and plotted against the log of the calculated free  $\text{Ca}^{2+}$  concentration. Each  $\text{Ca}^{2+}$  titration curve was fit with an appropriate sigmoidal curve to determine the values of the  $K_d$  and the  $n_H$ .

## **Chapter 7 Conclusions and future work**

## 7.1 Summary of the thesis

In the past decade a series of increasingly advanced RFPs have been developed since the first monomerization of the natural RFP DsRed<sup>11</sup>. These enhanced RFPs<sup>12, 13, 27, 33, 34</sup> bring new colours to the existing FP spectrum and provided new possibilities to investigate biological problems. Additionally they can help visualize biology at all levels from single molecules to entire living organisms. Genetically encoded red fluorescent indicators have also been rapidly developed on the basis of these RFPs<sup>74, 83, 84, 86</sup>. These sensors have various capabilities to report biochemical changes and enzymatic activities in live cells along with their green counterparts. The rapid expansion of the RFP and RFP-based biosensor toolkit represents significant technical advances in the research methods of cell biology and neuroscience.

In this thesis, we described our engineering efforts in developing new RFPs and RFP based biosensors. These new RFPs include a number of RFPs increasing spectral diversity, a series of pH-sensitive RFPs and assays for screening new biosensors. The RFP templates chosen for engineering in this work were all mFruit RFPs such as mOrange, mApple, and mCherry. In Chapters 2 and 3, we describe the engineering of pH-sensitive RFPs with a large fluorescence intensity change or an obvious ratio change over the physiological pH range. In Chapters 4 and 5, we reported novel RFP variants with properties including chromism, long Stoke shift, and red-shifted excitation and emission. In Chapter 6, a screening

method utilizing filter paper colony blotting for RFP based biosensors was developed.

## **7.2 Future directions**

### **7.2.1 Optimization of red fluorescent protein based pH sensors**

With the engineering of pHuji and pHlorina, we expect that these new pH sensors will be widely applied in cell biology research as the preferred second colour for pH-sensitive probes, enabling multicolour imaging of pH dynamics when coupled with currently available green pH sensors. Although pHuji is the most pH-sensitive RFP developed to date, it is still surpassed by its green counterpart superecliptic pHluorin<sup>102, 123</sup> in terms of pH sensitivity. Therefore, in the future, it would be beneficial to further improve the pH sensitivity of the currently available pH-sensitive RFPs. Another major critique for pHuji is its photoactivation problem, which could limit the applications of this sensor. Thus elimination of photoactivation would enhance its utility. The study of shyRFP in Chapter 4 could, in principle, provide some insight for fixing this problem since pHuji and shyRFP share the same template mApple.

### **7.2.1 Optimization of the new red fluorescent proteins**

The development of LSSmCherry and RDSmCherry offer new opportunities for multi-colour fluorescence imaging. However, spectral properties of these new

RFPs are still far from optimal in the aspects of fluorescent brightness and spectral properties. Future engineering efforts are required to optimize these new spectrally diversified RFPs, since RFP variants spectrally shifted to the near-infrared region are still poor, but in great demand; and a suitable LSS RFP for multi-photon excitation would also be extremely valuable. In addition, the widely used standard RFPs still fall short in certain applications when compared to the classic EGFP. One major problem of mRFP1 derived RFPs is that these RFPs tend to form aggregates when expressed in neural cells, therefore hampering their application in neuroscience. The Campbell lab has undertaken the mission of seeking an ideal RFP, and is still actively working on optimizing mCherry for neuronal research.

### **7.2.3 Application of filter paper based screening method**

The filter paper blot screening method developed in this thesis is a simple and efficient screening technique for FPs and FP based biosensors. It is well suited to screen for single FP-based ion sensors, such as  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Mg}^{2+}$ , and  $\text{Cl}^-$ . Although it is clear that real-context screening in live cells is a definite trend for the development of new genetically encoded sensors for cellular and neuroscientific research purposes<sup>100, 178</sup>. Filter paper based screening method can still be useful to serve as primary screening assay. It also should be able to find applications in the areas where extreme conditions need to be applied, for instance, for the engineering of pH resistant or thermostable FPs<sup>179</sup>.

### **7.3 Closing remark**

In the research work described in this thesis, we have successfully created new RFP-based biosensors and engineered RFP variants with distinctive spectral properties. For the past decade, extensive engineering effort has been invested to develop new RFP and RFP-based sensors. These RFP and RFP-based biosensors have become indispensable tools with ever-increasing importance. However, none of the FPs or sensors can be claimed to be flawless. Limitations emerged when these shiny new tools were put into previously uncharted research territories. These shortfalls might lead to experimental artifacts and the misinterpretation of the biological question under investigation. In addition, many RFP and RFP based sensors are reported to perform sub-optimally in transfected neural tissues or transgenic animals<sup>180</sup>. Therefore, continuing efforts from protein engineers will be required to overcome all these limitations and push the frontier forward once again.

## Reference

1. Shimomura, O., Johnson, F.H. & Saiga, Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea*. *J. Cell. Comp. Physiol.* **59**, 223-239 (1962).
2. 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* **111**, 229-233 (1992).
3. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. & Prasher, D.C. Green Fluorescent Protein as a Marker for Gene-Expression. *Science* **263**, 802-805 (1994).
4. Yang, T.T., Cheng, L. & Kain, S.R. Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. *Nucleic Acids Res.* **24**, 4592-4593 (1996).
5. Heim, R., Prasher, D.C. & Tsien, R.Y. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 12501-12504 (1994).
6. Matz, M.V. et al. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat. Biotechnol.* **17**, 969-973 (1999).
7. 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.* **97**, 11984-11989 (2000).

8. Cotlet, M. et al. Identification of different emitting species in the red fluorescent protein DsRed by means of ensemble and single-molecule spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14398-14403 (2001).
9. Gross, L.A., Baird, G.S., Hoffman, R.C., Baldrige, K.K. & Tsien, R.Y. The structure of the chromophore within DsRed, a red fluorescent protein from coral. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 11990-11995 (2000).
10. Yarbrough, D., Wachter, R.M., Kallio, K., Matz, M.V. & Remington, S.J. Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-Å resolution. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 462-467 (2001).
11. Campbell, R.E. et al. A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 7877-7882 (2002).
12. Shaner, N.C. et al. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* **22**, 1567-1572 (2004).
13. Shaner, N.C. et al. Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nat. Methods* **5**, 545-551 (2008).
14. Pletnev, S. et al. A crystallographic study of bright far-red fluorescent protein mKate reveals pH-induced cis-trans isomerization of the chromophore. *J. Biol. Chem.* **283**, 28980-28987 (2008).
15. Ormo, M. et al. Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392-1395 (1996).

16. Shu, X., Shaner, N.C., Yarbrough, C.A., Tsien, R.Y. & Remington, S.J. Novel chromophores and buried charges control color in mFruits. *Biochemistry* **45**, 9639-9647 (2006).
17. Verkhusha, V.V., Chudakov, D.M., Gurskaya, N.G., Lukyanov, S. & Lukyanov, K.A. Common pathway for the red chromophore formation in fluorescent proteins and chromoproteins. *Chem. Biol.* **11**, 845-854 (2004).
18. Day, R.N. & Davidson, M.W. The fluorescent protein palette: tools for cellular imaging. *Chem. Soc. Rev.* **38**, 2887-2921 (2009).
19. Strack, R.L., Strongin, D.E., Mets, L., Glick, B.S. & Keenan, R.J. Chromophore formation in DsRed occurs by a branched pathway. *J. Am. Chem. Soc.* **132**, 8496-8505 (2010).
20. Subach, F.V. & Verkhusha, V.V. Chromophore transformations in red fluorescent proteins. *Chem. Rev.* **112**, 4308-4327 (2012).
21. Shaner, N.C., Steinbach, P.A. & Tsien, R.Y. A guide to choosing fluorescent proteins. *Nat. Methods* **2**, 905-909 (2005).
22. Shaner, N.C., Patterson, G.H. & Davidson, M.W. Advances in fluorescent protein technology. *J. Cell Sci.* **120**, 4247-4260 (2007).
23. Lin, M.Z. et al. Autofluorescent proteins with excitation in the optical window for intravital imaging in mammals. *Chem. Biol.* **16**, 1169-1179 (2009).
24. Wiedenmann, J. et al. A far-red fluorescent protein with fast maturation and reduced oligomerization tendency from *Entacmaea quadricolor*

- (Anthozoa, Actinaria). *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11646-11651 (2002).
25. Merzlyak, E.M. et al. Bright monomeric red fluorescent protein with an extended fluorescence lifetime. *Nat. Methods* **4**, 555-557 (2007).
  26. Shemiakina, II et al. A monomeric red fluorescent protein with low cytotoxicity. *Nat. Commun.* **3**, 1204 (2012).
  27. Kredel, S. et al. mRuby, a bright monomeric red fluorescent protein for labeling of subcellular structures. *PLoS One* **4**, e4391 (2009).
  28. Lam, A.J. et al. Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat. Methods* **9**, 1005-1012 (2012).
  29. Wang, L., Jackson, W.C., Steinbach, P.A. & Tsien, R.Y. Evolution of new nonantibody proteins via iterative somatic hypermutation. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 16745-16749 (2004).
  30. 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.* **107**, 20257-20262 (2010).
  31. Shcherbo, D. et al. Bright far-red fluorescent protein for whole-body imaging. *Nat. Methods* **4**, 741-746 (2007).
  32. Shcherbo, D. et al. Far-red fluorescent tags for protein imaging in living tissues. *Biochem. J* **418**, 567-574 (2009).

33. Chu, J. et al. Non-invasive intravital imaging of cellular differentiation with a bright red-excitable fluorescent protein. *Nat. Methods* **11**, 572-578 (2014).
34. Morozova, K.S. et al. Far-red fluorescent protein excitable with red lasers for flow cytometry and superresolution STED nanoscopy. *Biophys. J.* **99**, L13-15 (2010).
35. Piatkevich, K.D. et al. Extended Stokes shift in fluorescent proteins: chromophore-protein interactions in a near-infrared TagRFP675 variant. *Sci. Rep.* **3**, 1847 (2013).
36. Kogure, T. et al. A fluorescent variant of a protein from the stony coral *Montipora facilitates* dual-color single-laser fluorescence cross-correlation spectroscopy. *Nat. Biotechnol.* **24**, 577-581 (2006).
37. Piatkevich, K.D. et al. Monomeric red fluorescent proteins with a large Stokes shift. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 5369-5374 (2010).
38. Yang, J. et al. mBeRFP, an improved large stokes shift red fluorescent protein. *PLoS One* **8**, e64849 (2013).
39. Piatkevich, K.D., Malashkevich, V.N., Almo, S.C. & Verkhusha, V.V. Engineering ESPT pathways based on structural analysis of LSSmKate red fluorescent proteins with large Stokes shift. *J. Am. Chem. Soc.* **132**, 10762-10770 (2010).
40. Shcherbakova, D.M., Hink, M.A., Joosen, L., Gadella, T.W. & Verkhusha, V.V. An orange fluorescent protein with a large Stokes shift for single-

- excitation multicolor FCCS and FRET imaging. *J. Am. Chem. Soc.* **134**, 7913-7923 (2012).
41. Zhang, J., Campbell, R.E., Ting, A.Y. & Tsien, R.Y. Creating new fluorescent probes for cell biology. *Nat. Rev. Mol. Cell Biol.* **3**, 906-918 (2002).
  42. Wu, B., Piatkevich, K.D., Lionnet, T., Singer, R.H. & Verkhusha, V.V. Modern fluorescent proteins and imaging technologies to study gene expression, nuclear localization, and dynamics. *Curr. Opin. Cell Biol.* **23**, 310-317 (2011).
  43. Berg, J., Hung, Y.P. & Yellen, G. A genetically encoded fluorescent reporter of ATP:ADP ratio. *Nat. Methods* **6**, 161-166 (2009).
  44. Cannon, M.B. & Remington, S.J. Re-engineering redox-sensitive green fluorescent protein for improved response rate. *Protein Sci.* **15**, 45-57 (2006).
  45. Dimitrov, D. et al. Engineering and characterization of an enhanced fluorescent protein voltage sensor. *PLoS One* **2**, e440 (2007).
  46. Yano, T. et al. A novel fluorescent sensor protein for visualization of redox states in the cytoplasm and in peroxisomes. *Mol. Cell Biol.* **30**, 3758-3766 (2010).
  47. Newman, R.H., Fosbrink, M.D. & Zhang, J. Genetically encodable fluorescent biosensors for tracking signaling dynamics in living cells. *Chem. Rev.* **111**, 3614-3666 (2011).

48. Subach, F.V., Piatkevich, K.D. & Verkhusha, V.V. Directed molecular evolution to design advanced red fluorescent proteins. *Nat. Methods* **8**, 1019-1026 (2011).
49. Shcherbakova, D.M., Subach, O.M. & Verkhusha, V.V. Red fluorescent proteins: advanced imaging applications and future design. *Angew. Chem. Int. Ed. Engl.* **51**, 10724-10738 (2012).
50. Ghosh, I., Hamilton, A.D. & Regan, L. Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein. *J. Am. Chem. Soc.* **122**, 5658-5659 (2000).
51. Wilson, C.G., Magliery, T.J. & Regan, L. Detecting protein-protein interactions with GFP-fragment reassembly. *Nat. Methods* **1**, 255-262 (2004).
52. Jach, G., Pesch, M., Richter, K., Frings, S. & Uhrig, J.F. An improved mRFP1 adds red to bimolecular fluorescence complementation. *Nat. Methods* **3**, 597-600 (2006).
53. Fan, J.Y. et al. Split mCherry as a new red bimolecular fluorescence complementation system for visualizing protein-protein interactions in living cells. *Biochem. Biophys. Res. Commun.* **367**, 47-53 (2008).
54. Keem, J.O., Lee, I.H., Kim, S.Y., Jung, Y. & Chung, B.H. Splitting and self-assembling of far-red fluorescent protein with an engineered beta strand peptide: application for alpha-synuclein imaging in mammalian cells. *Biomaterials* **32**, 9051-9058 (2011).

55. Chu, J. et al. A novel far-red bimolecular fluorescence complementation system that allows for efficient visualization of protein interactions under physiological conditions. *Biosens. Bioelectron.* **25**, 234-239 (2009).
56. Han, Y. et al. In vivo imaging of protein-protein and RNA-protein interactions using novel far-red fluorescence complementation systems. *Nucleic Acids Res.* **42**, e103 (2014).
57. Cabantous, S. et al. A new protein-protein interaction sensor based on tripartite split-GFP association. *Sci. Rep.* **3**, 2854 (2013).
58. Ai, H.W., Hazelwood, K.L., Davidson, M.W. & Campbell, R.E. Fluorescent protein FRET pairs for ratiometric imaging of dual biosensors. *Nat. Methods* **5**, 401-403 (2008).
59. Carlson, H.J. & Campbell, R.E. Genetically encoded FRET-based biosensors for multiparameter fluorescence imaging. *Curr. Opin. Biotechnol.* **20**, 19-27 (2009).
60. Goedhart, J., Vermeer, J.E., Adjobo-Hermans, M.J., van Weeren, L. & Gadella, T.W., Jr. Sensitive detection of p65 homodimers using red-shifted and fluorescent protein-based FRET couples. *PLoS One* **2**, e1011 (2007).
61. Piljic, A. & Schultz, C. Simultaneous recording of multiple cellular events by FRET. *ACS Chem. Biol.* **3**, 156-160 (2008).
62. Lindenburg, L.H., Hessels, A.M., Ebberink, E.H., Arts, R. & Merkx, M. Robust red FRET sensors using self-associating fluorescent domains. *ACS Chem. Biol.* **8**, 2133-2139 (2013).

63. Lindenburg, L.H. et al. Quantifying Stickiness: Thermodynamic Characterization of Intramolecular Domain Interactions To Guide the Design of Forster Resonance Energy Transfer Sensors. *Biochemistry* **53**, 6370-6381 (2014).
64. Alford, S.C., Ding, Y., Simmen, T. & Campbell, R.E. Dimerization-dependent green and yellow fluorescent proteins. *ACS Synth. Biol.* **1**, 569-575 (2012).
65. Alford, S.C., Abdelfattah, A.S., Ding, Y. & Campbell, R.E. A fluorogenic red fluorescent protein heterodimer. *Chem. Biol.* **19**, 353-360 (2012).
66. Miesenbock, G., De Angelis, D.A. & Rothman, J.E. Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature* **394**, 192-195 (1998).
67. Jayaraman, S., Haggie, P., Wachter, R.M., Remington, S.J. & Verkman, A.S. Mechanism and cellular applications of a green fluorescent protein-based halide sensor. *J. Biol. Chem.* **275**, 6047-6050 (2000).
68. Galletta, L.J., Haggie, P.M. & Verkman, A.S. Green fluorescent protein-based halide indicators with improved chloride and iodide affinities. *FEBS Lett.* **499**, 220-224 (2001).
69. Hanson, G.T. et al. Investigating mitochondrial redox potential with redox-sensitive green fluorescent protein indicators. *J. Biol. Chem.* **279**, 13044-13053 (2004).

70. Dooley, C.T. et al. Imaging dynamic redox changes in mammalian cells with green fluorescent protein indicators. *J. Biol. Chem.* **279**, 22284-22293 (2004).
71. Tang, S. et al. Design and application of a class of sensors to monitor Ca<sup>2+</sup> dynamics in high Ca<sup>2+</sup> concentration cellular compartments. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 16265-16270 (2011).
72. Johnson, D.E. et al. Red fluorescent protein pH biosensor to detect concentrative nucleoside transport. *J. Biol. Chem.* **284**, 20499-20511 (2009).
73. Li, Y. & Tsien, R.W. pHTomato, a red, genetically encoded indicator that enables multiplex interrogation of synaptic activity. *Nat. Neurosci.* **15**, 1047-1053 (2012).
74. Tantama, M., Hung, Y.P. & Yellen, G. Imaging intracellular pH in live cells with a genetically encoded red fluorescent protein sensor. *J. Am. Chem. Soc.* **133**, 10034-10037 (2011).
75. Carlson, H.J., Cotton, D.W. & Campbell, R.E. Circularly permuted monomeric red fluorescent proteins with new termini in the beta-sheet. *Protein Sci.* **19**, 1490-1499 (2010).
76. Shui, B. et al. Circular permutation of red fluorescent proteins. *PLoS One* **6**, e20505 (2011).
77. Marvin, J.S. et al. An optimized fluorescent probe for visualizing glutamate neurotransmission. *Nat. Methods* **10**, 162-170 (2013).

78. Tantama, M., Martinez-Francois, J.R., Mongeon, R. & Yellen, G. Imaging energy status in live cells with a fluorescent biosensor of the intracellular ATP-to-ADP ratio. *Nat. Commun.* **4**, 2550 (2013).
79. St-Pierre, F. et al. High-fidelity optical reporting of neuronal electrical activity with an ultrafast fluorescent voltage sensor. *Nat. Neurosci.* **17**, 884-889 (2014).
80. Marvin, J.S., Schreiter, E.R., Echevarria, I.M. & Looger, L.L. A genetically encoded, high-signal-to-noise maltose sensor. *Proteins* **79**, 3025-3036 (2011).
81. Nakai, J., Ohkura, M. & Imoto, K. A high signal-to-noise Ca<sup>2+</sup> probe composed of a single green fluorescent protein. *Nat. Biotech.* **19**, 137-141 (2001).
82. Nagai, T., Sawano, A., Park, E.S. & Miyawaki, A. Circularly permuted green fluorescent proteins engineered to sense Ca<sup>2+</sup>. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3197-3202 (2001).
83. Zhao, Y. et al. An expanded palette of genetically encoded Ca<sup>2+</sup>(+) indicators. *Science* **333**, 1888-1891 (2011).
84. Wu, J. et al. Improved orange and red Ca<sup>2+</sup>/- indicators and photophysical considerations for optogenetic applications. *ACS Chem. Neurosci.* **4**, 963-972 (2013).
85. Wu, J. et al. Red fluorescent genetically encoded Ca<sup>2+</sup> indicators for use in mitochondria and endoplasmic reticulum. *Biochem. J* (2014).

86. Akerboom, J. et al. Genetically encoded calcium indicators for multi-color neural activity imaging and combination with optogenetics. *Front. Mol. Neurosci.* **6**, 2 (2013).
87. Jin, L. et al. Single action potentials and subthreshold electrical events imaged in neurons with a fluorescent protein voltage probe. *Neuron* **75**, 779-785 (2012).
88. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. & Pease, L.R. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51-59 (1989).
89. Wang, W.Y. & Malcolm, B.A. Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange (TM) site-directed mutagenesis. *BioTechniques* **26**, 680-682 (1999).
90. Hogrefe, H.H., Cline, J., Youngblood, G.L. & Allen, R.M. Creating randomized amino acid libraries with the QuikChange® multi site-directed mutagenesis kit. *BioTechniques* **33**, 1158-1165 (2002).
91. Yu, Y. & Lutz, S. Circular permutation: a different way to engineer enzyme structure and function. *Trends Biotechnol.* **29**, 18-25 (2011).
92. Arnold, F.H. & Georgiou, G. Directed evolution library creation. *Methods Mol. Biol.* **231** (2003).
93. Cirino, P.C., Mayer, K.M. & Umeno, D. in Directed Evolution Library Creation 3-9 (Springer, 2003).

94. Zhao, H., Giver, L., Shao, Z., Affholter, J.A. & Arnold, F.H. Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.* **16**, 258-261 (1998).
95. Ai, H.W., Baird, M.A., Shen, Y., Davidson, M.W. & Campbell, R.E. Engineering and characterizing monomeric fluorescent proteins for live-cell imaging applications. *Nat. Protoc.* **9**, 910-928 (2014).
96. Shemiakina, I.I. et al. A monomeric red fluorescent protein with low cytotoxicity. *Nat. Commun.* **3**, 1204 (2012).
97. Zhao, Y. et al. Microfluidic cell sorter-aided directed evolution of a protein-based calcium ion indicator with an inverted fluorescent response. *Integr. Biol. (Camb.)* **6**, 714-725 (2014).
98. Ibraheem, A., Yap, H., Ding, Y. & Campbell, R.E. A bacteria colony-based screen for optimal linker combinations in genetically encoded biosensors. *BMC Biotechnol.* **11**, 105 (2011).
99. Belal, A.S., Sell, B.R., Hoi, H., Davidson, M.W. & Campbell, R.E. Optimization of a genetically encoded biosensor for cyclin B1-cyclin dependent kinase 1. *Mol Biosyst* **10**, 191-195 (2014).
100. Wardill, T.J. et al. A neuron-based screening platform for optimizing genetically-encoded calcium indicators. *PLoS One* **8**, e77728 (2013).
101. Chen, T.W. et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295-300 (2013).

102. Sankaranarayanan, S., De Angelis, D., Rothman, J.E. & Ryan, T.A. The use of pHluorins for optical measurements of presynaptic activity. *Biophys. J.* **79**, 2199-2208 (2000).
103. Tsuboi, T. & Rutter, G.A. Multiple forms of "kiss-and-run" exocytosis revealed by evanescent wave microscopy. *Curr. Biol.* **13**, 563-567 (2003).
104. Yudowski, G.A., Puthenveedu, M.A. & von Zastrow, M. Distinct modes of regulated receptor insertion to the somatodendritic plasma membrane. *Nat. Neurosci.* **9**, 622-627 (2006).
105. Balaji, J. & Ryan, T.A. Single-vesicle imaging reveals that synaptic vesicle exocytosis and endocytosis are coupled by a single stochastic mode. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 20576-20581 (2007).
106. Gandhi, S.P. & Stevens, C.F. Three modes of synaptic vesicular recycling revealed by single-vesicle imaging. *Nature* **423**, 607-613 (2003).
107. Jullie, D., Choquet, D. & Perraiss, D. Recycling endosomes undergo rapid closure of a fusion pore on exocytosis in neuronal dendrites. *J. Neurosci.* **34**, 11106-11118 (2014).
108. Merrifield, C.J., Perraiss, D. & Zenisek, D. Coupling between clathrin-coated-pit invagination, cortactin recruitment, and membrane scission observed in live cells. *Cell* **121**, 593-606 (2005).
109. Taylor, M.J., Perraiss, D. & Merrifield, C.J. A high precision survey of the molecular dynamics of mammalian clathrin-mediated endocytosis. *PLoS Biol.* **9**, e1000604 (2011).

110. An, S.J., Grabner, C.P. & Zenisek, D. Real-time visualization of complexin during single exocytic events. *Nat. Neurosci.* **13**, 577-583 (2010).
111. Ishii, K., Hirose, K. & Iino, M. Ca<sup>2+</sup> shuttling between endoplasmic reticulum and mitochondria underlying Ca<sup>2+</sup> oscillations. *EMBO Rep.* **7**, 390-396 (2006).
112. Chen, Y. et al. Rab10 and myosin-Va mediate insulin-stimulated GLUT4 storage vesicle translocation in adipocytes. *J. Cell Biol.* **198**, 545-560 (2012).
113. Gautier, A. et al. How to control proteins with light in living systems. *Nat. Chem. Biol.* **10**, 533-541 (2014).
114. Akemann, W. et al. Imaging neural circuit dynamics with a voltage-sensitive fluorescent protein. *J. Neurophysiol.* **108**, 2323-2337 (2012).
115. Kneen, M., Farinas, J., Li, Y. & Verkman, A.S. Green fluorescent protein as a noninvasive intracellular pH indicator. *Biophys. J.* **74**, 1591-1599 (1998).
116. Sniegowski, J.A., Phail, M.E. & Wachter, R.M. Maturation efficiency, trypsin sensitivity, and optical properties of Arg96, Glu222, and Gly67 variants of green fluorescent protein. *Biochem. Biophys. Res. Commun.* **332**, 657-663 (2005).
117. Schweizer, F.E. & Ryan, T.A. The synaptic vesicle: cycle of exocytosis and endocytosis. *Curr. Opin. Neurobiol.* **16**, 298-304 (2006).

118. Gonzalez-Gonzalez, I.M., Jaskolski, F., Goldberg, Y., Ashby, M.C. & Henley, J.M. Measuring membrane protein dynamics in neurons using fluorescence recovery after photobleach. *Methods Enzymol.* **504**, 127-146 (2012).
119. Ganesan, S., Ameer-Beg, S.M., Ng, T.T., Vojnovic, B. & Wouters, F.S. A dark yellow fluorescent protein (YFP)-based Resonance Energy-Accepting Chromoprotein (REACH) for Forster resonance energy transfer with GFP. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4089-4094 (2006).
120. Carmody, W.R. Easily prepared wide range buffer series. *J. Chem. Educ.* **38**, 559 (1961).
121. Casey, J.R., Grinstein, S. & Orłowski, J. Sensors and regulators of intracellular pH. *Nat. Rev. Mol. Cell Biol.* **11**, 50-61 (2010).
122. Bencina, M. Illumination of the spatial order of intracellular pH by genetically encoded pH-sensitive sensors. *Sensors* **13**, 16736-16758 (2013).
123. Miesenbock, G. Synapto-pHluorins: genetically encoded reporters of synaptic transmission. *Cold Spring Harb. Protoc.* **2012**, 213-217 (2012).
124. Bizzarri, R., Serresi, M., Luin, S. & Beltram, F. Green fluorescent protein based pH indicators for in vivo use: a review. *Anal. Bioanal. Chem.* **393**, 1107-1122 (2009).
125. Bizzarri, R. et al. Development of a novel GFP-based ratiometric excitation and emission pH indicator for intracellular studies. *Biophys. J.* **90**, 3300-3314 (2006).

126. Scharnagl, C., Raupp-Kossmann, R. & Fischer, S.F. Molecular basis for pH sensitivity and proton transfer in green fluorescent protein: protonation and conformational substates from electrostatic calculations. *Biophys. J.* **77**, 1839-1857 (1999).
127. Chatteraj, M., King, B.A., Bublitz, G.U. & Boxer, S.G. Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 8362-8367 (1996).
128. Brejc, K. et al. Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 2306-2311 (1997).
129. Stoner-Ma, D. et al. Observation of excited-state proton transfer in green fluorescent protein using ultrafast vibrational spectroscopy. *J. Am. Chem. Soc.* **127**, 2864-2865 (2005).
130. Agmon, N. Proton pathways in green fluorescence protein. *Biophys. J.* **88**, 2452-2461 (2005).
131. Mahon, M.J. pHluorin2: an enhanced, ratiometric, pH-sensitive green fluorescent protein. *Adv. Biosci. Biotechnol.* **2**, 132-137 (2011).
132. Kogure, T., Kawano, H., Abe, Y. & Miyawaki, A. Fluorescence imaging using a fluorescent protein with a large Stokes shift. *Methods* **45**, 223-226 (2008).
133. Henderson, J.N. et al. Excited state proton transfer in the red fluorescent protein mKeima. *J. Am. Chem. Soc.* **131**, 13212-13213 (2009).

134. Violot, S., Carpentier, P., Blanchoin, L. & Bourgeois, D. Reverse pH-dependence of chromophore protonation explains the large Stokes shift of the red fluorescent protein mKeima. *J. Am. Chem. Soc.* **131**, 10356-10357 (2009).
135. Miksa, M., Komura, H., Wu, R., Shah, K.G. & Wang, P. A novel method to determine the engulfment of apoptotic cells by macrophages using pHrodo succinimidyl ester. *J. Immunol. Methods* **342**, 71-77 (2009).
136. Bien-Ly, N. et al. Transferrin receptor (TfR) trafficking determines brain uptake of TfR antibody affinity variants. *J. Exp. Med.* **211**, 233-244 (2014).
137. Day, J.H. Thermochromism. *Chem. Rev.* **63**, 65-80 (1963).
138. Dürr, H. & Bouas-Laurent, H. Photochromism: Molecules and Systems: Molecules and Systems. (Gulf Professional Publishing, 2003).
139. Seeboth, A., Lotzsch, D., Ruhmann, R. & Muehling, O. Thermochromic polymers--function by design. *Chem. Rev.* **114**, 3037-3068 (2014).
140. Berkovic, G., Krongauz, V. & Weiss, V. Spiropyrans and Spirooxazines for Memories and Switches. *Chem. Rev.* **100**, 1741-1754 (2000).
141. Maafi, M. Useful spectrokinetic methods for the investigation of photochromic and thermo-photochromic spiropyrans. *Molecules* **13**, 2260-2302 (2008).
142. Bhoo, S.H., Davis, S.J., Walker, J., Karniol, B. & Vierstra, R.D. Bacteriophytochromes are photochromic histidine kinases using a biliverdin chromophore. *Nature* **414**, 776-779 (2001).

143. Hampp, N. Bacteriorhodopsin as a Photochromic Retinal Protein for Optical Memories. *Chem. Rev.* **100**, 1755-1776 (2000).
144. Chudakov, D.M. et al. Kindling fluorescent proteins for precise in vivo photolabeling. *Nat. Biotechnol.* **21**, 191-194 (2003).
145. Ando, R., Mizuno, H. & Miyawaki, A. Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting. *Science* **306**, 1370-1373 (2004).
146. Stiel, A.C. et al. 1.8 Å bright-state structure of the reversibly switchable fluorescent protein Dronpa guides the generation of fast switching variants. *Biochem. J* **402**, 35-42 (2007).
147. Henderson, J.N., Ai, H.W., Campbell, R.E. & Remington, S.J. Structural basis for reversible photobleaching of a green fluorescent protein homologue. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 6672-6677 (2007).
148. Stiel, A.C. et al. Generation of monomeric reversibly switchable red fluorescent proteins for far-field fluorescence nanoscopy. *Biophys. J.* **95**, 2989-2997 (2008).
149. Subach, F.V. et al. Red fluorescent protein with reversibly photoswitchable absorbance for photochromic FRET. *Chem. Biol.* **17**, 745-755 (2010).
150. Andresen, M. et al. Structure and mechanism of the reversible photoswitch of a fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 13070-13074 (2005).

151. Habuchi, S. et al. Photo-induced protonation/deprotonation in the GFP-like fluorescent protein Dronpa: mechanism responsible for the reversible photoswitching. *Photochem. Photobiol. Sci.* **5**, 567-576 (2006).
152. Pletnev, S., Subach, F.V., Dauter, Z., Wlodawer, A. & Verkhusha, V.V. A structural basis for reversible photoswitching of absorbance spectra in red fluorescent protein rsTagRFP. *J. Mol. Biol.* **417**, 144-151 (2012).
153. Gayda, S., Nienhaus, K. & Nienhaus, G.U. Mechanistic insights into reversible photoactivation in proteins of the GFP family. *Biophys. J.* **103**, 2521-2531 (2012).
154. Leiderman, P., Huppert, D. & Agmon, N. Transition in the temperature-dependence of GFP fluorescence: from proton wires to proton exit. *Biophys. J.* **90**, 1009-1018 (2006).
155. Faro, A.R. et al. Low-temperature chromophore isomerization reveals the photoswitching mechanism of the fluorescent protein Padron. *J. Am. Chem. Soc.* **133**, 16362-16365 (2011).
156. Sarkisyan, K.S. et al. Tryptophan-based chromophore in fluorescent proteins can be anionic. *Sci. Rep.* **2**, 608 (2012).
157. Mizuno, H. et al. Light-dependent regulation of structural flexibility in a photochromic fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 9227-9232 (2008).
158. Kalninsh, K.K. Structure and thermochromism of spiropyrans. Triplet mechanism of the thermal cleavage closure of the pyran ring. *J. Struct. Chem.* **39**, 642-650 (1998).

159. Fernandez-Suarez, M. & Ting, A.Y. Fluorescent probes for super-resolution imaging in living cells. *Nat. Rev. Mol. Cell Biol.* **9**, 929-943 (2008).
160. Don Paul, C. et al. Phanta: a non-fluorescent photochromic acceptor for pcFRET. *PLoS One* **8**, e75835 (2013).
161. Petersen, J. et al. The 2.0-Å crystal structure of eqFP611, a far red fluorescent protein from the sea anemone *Entacmaea quadricolor*. *J. Biol. Chem.* **278**, 44626-44631 (2003).
162. Ai, H.W., Olenych, S.G., Wong, P., Davidson, M.W. & Campbell, R.E. Hue-shifted monomeric variants of *Clavularia* cyan fluorescent protein: identification of the molecular determinants of color and applications in fluorescence imaging. *BMC Biol.* **6**, 13 (2008).
163. Subach, O.M. et al. Conversion of red fluorescent protein into a bright blue probe. *Chem. Biol.* **15**, 1116-1124 (2008).
164. Carlson, H.J. & Campbell, R.E. Mutational analysis of a red fluorescent protein-based calcium ion indicator. *Sensors* **13**, 11507-11521 (2013).
165. Tripathi, N.K., Sathyaseelan, K., Jana, A.M. & Rao, P.V.L. High Yield Production of Heterologous Proteins with *Escherichia coli*. *Def. Sci. J.* **59**, 137-146 (2009).
166. Balbás, P. & Lorence, A. Recombinant Gene Expression: Reviews and Protocols. (Humana Press, 2004).

167. Costantini, L.M., Fossati, M., Francolini, M. & Snapp, E.L. Assessing the tendency of fluorescent proteins to oligomerize under physiologic conditions. *Traffic* **13**, 643-649 (2012).
168. Zhang, M. et al. Rational design of true monomeric and bright photoactivatable fluorescent proteins. *Nat. Methods* **9**, 727-729 (2012).
169. Subach, O.M., Cranfill, P.J., Davidson, M.W. & Verkhusha, V.V. An enhanced monomeric blue fluorescent protein with the high chemical stability of the chromophore. *PLoS One* **6**, e28674 (2011).
170. Frommer, W.B., Davidson, M.W. & Campbell, R.E. Genetically encoded biosensors based on engineered fluorescent proteins. *Chem. Soc. Rev.* **38**, 2833-2841 (2009).
171. Depry, C., Mehta, S. & Zhang, J. Multiplexed visualization of dynamic signaling networks using genetically encoded fluorescent protein-based biosensors. *Pflugers Arch.* **465**, 373-381 (2013).
172. Thestrup, T. et al. Optimized ratiometric calcium sensors for functional in vivo imaging of neurons and T lymphocytes. *Nat. Methods* **11**, 175-182 (2014).
173. Cornvik, T. et al. Colony filtration blot: a new screening method for soluble protein expression in *Escherichia coli*. *Nat. Methods* **2**, 507-509 (2005).
174. Ben-David, A., Shoham, G. & Shoham, Y. A universal screening assay for glycosynthases: directed evolution of glycosynthase XynB2(E335G)

- suggests a general path to enhance activity. *Chem. Biol.* **15**, 546-551 (2008).
175. Wang, X., Kirberger, M., Qiu, F., Chen, G. & Yang, J.J. Towards predicting Ca<sup>2+</sup>-binding sites with different coordination numbers in proteins with atomic resolution. *Proteins* **75**, 787-798 (2009).
  176. Wang, Z.M., Tang, S., Messi, M.L., Yang, J.J. & Delbono, O. Residual sarcoplasmic reticulum Ca<sup>2+</sup> concentration after Ca<sup>2+</sup> release in skeletal myofibers from young adult and old mice. *Pflugers Arch.* **463**, 615-624 (2012).
  177. Nybo, K. GFP imaging in fixed cells. *BioTechniques* **52**, 359-360 (2012).
  178. Hochbaum, D.R. et al. All-optical electrophysiology in mammalian neurons using engineered microbial rhodopsins. *Nat. Methods* **11**, 825-833 (2014).
  179. Aliye, N., Fabbretti, A., Lupidi, G., Tsekoa, T. & Spurio, R. Engineering color variants of green fluorescent protein (GFP) for thermostability, pH-sensitivity, and improved folding kinetics. *Appl. Microbiol. Biotechnol.* **98**, 1-12 (2014).
  180. Cai, D., Cohen, K.B., Luo, T., Lichtman, J.W. & Sanes, J.R. Improved tools for the Brainbow toolbox. *Nat. Methods* **10**, 540-547 (2013).