Increasing the quantum yield of red fluorescent proteins using rational design

Antonia T. Pandelieva

Submitted to the Department of Chemistry and Biomolecular Sciences
of the Faculty of Science in partial fulfillment of the requirements
for the degree of Master of Science in Chemistry

University of Ottawa
Ottawa, ON, Canada

January 2016

Abstract

Monomeric red fluorescent proteins (RFPs) are used extensively for applications in molecular biology research, and are especially suited for whole body imaging applications due to their longer excitation and emission wavelengths, which are less damaging and penetrate deeper into animal tissue. However, these proteins suffer from reduced brightness compared to other fluorescent proteins, and require further engineering, which is often achieved through random methods, incurring large time and resource costs. Here we propose a rational design approach to improve the quantum yield of RFPs by reducing conformational variability of the chromophore. We engineered mRojoA, a mutant containing a π-stack involving Tyr197 and the chromophore phenolate, to include the P63F/H/Y mutations on its other side, by simultaneously mutating neighbouring positions 16, 143, and 163. The brightest mutants that we found in each library, mRojo-VYGV, mRojo-VFAV, and mRojo-VHSV, exhibited 1.8- to 2.4-fold increases in brightness, and quantum yield increases of up to 2.1-fold. In all three mutants, the increases in brightness were predominantly due to improvements in the quantum yield and not the extinction coefficient. Solving the crystal structures of two of these mutants along with a dim variant allowed us to strongly infer a link between rigidity of the chromophore and increased quantum yield. In addition, back-mutating position 63 in the highest quantum yield mutant, mRojo-VYGV, reversed the improvement in quantum yield, indicating that Y63 was the primary residue responsible for the improved brightness of the protein. Unfortunately, the mCherry-VYGV mutant did not achieve a similar increase in quantum yield or brightness. This is likely due to the lack of a second bulky aromatic residue at position 197, which is present in mRojoA. Nevertheless, this rational approach could be applied to some other RFPs whose chromophores exhibit increased conformational variability in order to further improve their brightness.

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Roberto Chica, for the incredible positive influence that he has had on my development as a scientist throughout the years. He has always pushed me to succeed above and beyond my expectations of myself, and has taken the time to make sure that I was equipped to excel in any challenge that I undertook. Thanks to his support, I have gained and built upon valuable skills in research, public speaking, and written communication that I know will be invaluable to me in my future career path.

Secondly, I would like to thank all of my past and present colleagues for helping to make every day in the lab more enjoyable by creating a fun and supportive work environment. Every one of them has helped me to solve problems or overcome challenges I encountered throughout my studies, and the friendly banter in the lab created a cheerful atmosphere even during days when the workload was heavy. I will be sorry to finally part ways with my fellow Biotechnology students, Matthew Eason and Adam Damry, and I am certain that they will both excel at anything they undertake in the future. I would also like to extend a special thank you to Jamie Davey, Guido Calderini, Miranda Baran, and Jenna McCann for the excellent work they did in support of this project. I could not have asked for better partners in crime.

Thirdly, a warm thank you to my family for their support throughout my studies in the form of encouragement, sound advice, home cooked meals, and late night pickups from the lab. Thank you as well to my wonderful boyfriend Patrick Morris, who kept me company and never let me slack when I had work to do. I can honestly say that I would not have accomplished as much as I have without all of their positive influence. Lastly, I would like to thank the NSERC Create program and the University of Ottawa for funding my Master's studies.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CFP - cyan fluorescent protein

EGFP - enhanced green fluorescent protein

FP - fluorescent protein

ESPT - excited state protein transfer

FRET - Förster resonance energy transfer

GFP - green fluorescent protein

H-bond – hydrogen bond

IPTG - isopropyl β-D-1-thiogalactopyranoside

Ni-NTA - nickel – nitrilotriacetic acid

RFP - red fluorescent protein

mFP - monomeric fluorescent protein

mRFP - monomeric red fluorescent protein

MPD - 2-methyl-2,4-pentanediol

PBS - phosphate buffered saline

PCR - polymerase chain reaction

PDB - protein data bank

QM/MM - quantum mechanics / molecular mechanics

SDS-PAGE - sodium dodecyl sulphate polyacrylamide gel electrophoresis

SOE - splicing through overlap extension

YFP - yellow fluorescent protein

1. Introduction

1.1. Structural and Historical Overview of Red Fluorescent Proteins from Anthozoa *Species*

In 1999, a red fluorescent protein (RFP) (λ_{em} > 570 nm) from *Discosoma* coral species in the *Anthozoa* class was the first of its kind to be cloned, allowing the expansion of the previously available yellow, green, and cyan fluorescent protein (FP) colour palette to proteins of emission wavelengths of up to 583 nm (*1*). These proteins, which are 220 – 250 amino acids in length and have ~30% sequence identity to *Aequorea victoria* green fluorescent protein (GFP), contain motifs that are homologous to each other and all GFP-like FPs. Each protein adopts a β-barrel structure composed of 11 β-strands with a distorted α-helix in the core containing a fluorescent chromophore species (Figure 1.1). The chromophore is formed through cyclization of three residues (Q66-Y67-G68 in *DsRed*) during post-translational modification, followed by oxygen-mediated dehydration and oxidation to form a conjugated system (*1, 2*) (Figure 1.2).

The first cloned and characterized RFP, DsRed (originally drFP583), is an obligate homotetramer (*2, 3*), and all RFPs found in nature to date have oligomeric structures (*4, 5*). In 2002, the monomeric protein mRFP1 was engineered from DsRed through eight rounds of random mutagenesis, sequentially disrupting the two interacting protein interfaces to create the non-fluorescent monomer mRFP0.1, and then recovering its fluorescence (*3*). While monomeric fluorescent proteins are desirable for their smaller size, fewer folding errors in fusion, and lesser tendency to aggregate in cells (*6, 7*), one source suggests that tetrameric proteins have increased quantum yields and brightness compared to their monomeric counterparts because of reduced non-radiative dynamic relaxation (*8, 9*). Shortly after the creation of mRFP1, the protein underwent

extensive directed evolution to produce the mFruits family of proteins, which are among the most widely used today (*4, 5, 10*).

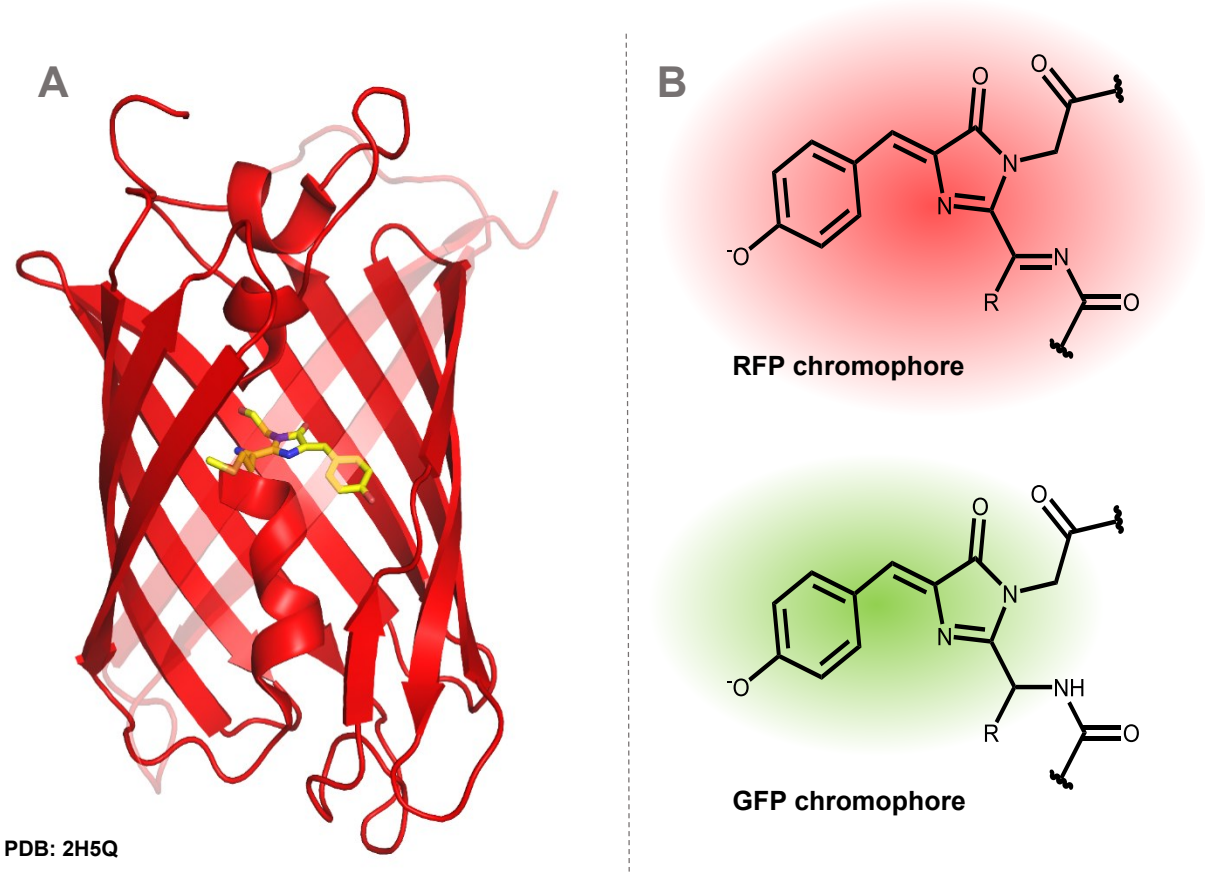

Figure 1.1. Structural characteristics of red fluorescent proteins (RFPs). A) Barrel structure of an RFP monomer with the chromophore shown in yellow. B) Comparison of the RFP and GFP anionic chromophore structures.

Other *Anthozoa* species, such as the sea anemone *Entacmaea quadricolor*, have produced related families of red fluorescent proteins. Two bright red tetrameric fluorescent proteins, eqFP578 and eqFP611, were cloned from this species (*11*). A fast-maturing, pH stable, and bright dimeric version of eqFP578 called TurboRFP was developed later (*11*), and served as a template

for creating Katushka, a variant with an emission wavelength (635 nm) that approached the far-red emission spectrum (≥ 650 nm), which is considered ideal for whole body imaging applications (*12*). Residues surrounding the chromophore of Katushka were added to a monomeric version of TurboRFP called TagRFP (*11*) to make mKate, a monomeric RFP with an identical emission wavelength, and a quantum yield of 0.28. mKate's brightness was equivalent to 45% of EGFP, making it one of the most widely used RFPs for whole animal imaging applications at the time (*5, 12*). Since the chromophores of this RFP family were also prone to *cis*-to-*trans* isomerization (*13*), leading to a decrease in quantum yield, the brighter variant mKate2 was later engineered to stabilize the brighter *cis* isomer, leading to an increase of quantum yield to 0.40 (14).

The proteins stemming from *Discosoma* species and *Entacmaea quadricolor* gave rise to a large palette of other monomeric RFPs with diverse Stokes shifts, brightnesses, pH sensitivities, photosensitivities, maturation rates, and excitation and emission wavelengths, including the far-red variants mPlum (*15*), and mNeptune (*16*). Despite the wide variety of mRFP variants already available, the quantum yields of this FP family remain smaller than those of FPs of other wavelengths (*4, 5, 17*). This represents a challenge in the useful application of these proteins *in lieu* of, as well as alongside their blue-shifted counterparts.

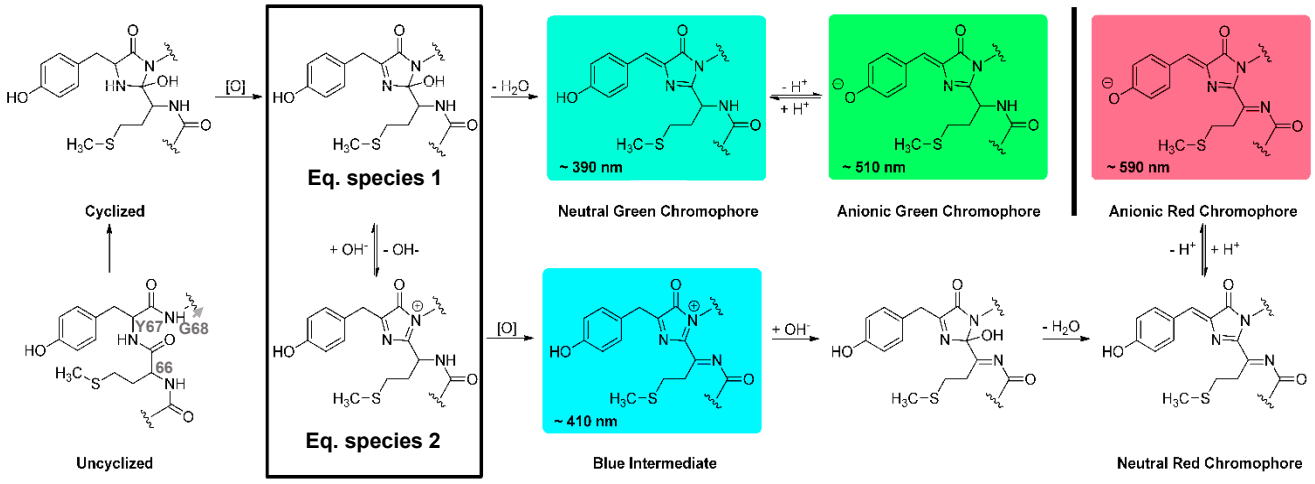

Figure 1.2. Maturation pathway of the green and red chromophores (*18*). The formation of a cyclized chromophore precursor is oxygen-dependent. This species (Eq. species 1) undergoes dehydration to irreversibly form the green chromophore. Alternatively, its cationic form (Eq. species 2) irreversibly forms the acylimine group through a second oxygen-catalyzed reaction. The resulting blue intermediate undergoes hydroxide addition followed by dehydration to form the imidazolinone group and complete the red chromophore. Only the anionic red chromophore, formed by the loss of the hydroxyphenyl hydrogen, is fluorescent.

Figure modified from: Moore, M. M *et al.* (2012) *PLOS One* **7**(12):e52463.doi:10.1371/journal.pone.0052463

1.2 Applications of RFPs

Since their discovery and cloning, fluorescent proteins have gained immense popularity in biochemistry research as genetically encodable tags (*12, 19, 20, 21*). The green fluorescent protein (GFP), which was the first of its kind to be cloned (*22*) and used for *in vivo* labelling (*19*), revolutionized the imaging of living systems and earned Osamu Shimomura, Roger Y. Tsien, and Martin Chalfie the Nobel Prize in Chemistry in 2008. When fused to other proteins, GFP-like

proteins of all colours can help to visualize their expression level (*19*), localization (*23, 24*), or interactions between proteins (*25, 26*). They can also be utilized to visualize subcellular structures. In 2014, fluorescent protein research resulted in a second Nobel Prize in Chemistry, which was awarded to Eric Betzig, Stefan W. Hell, and William E. Moerner for the development of super-resolution fluorescence microscopy, a technique that uses pattern illumination or photoswitchable fluorescent proteins to visualize intracellular structures with a greater resolution than Abbe's diffraction limit (*27*).

Fluorescent proteins have been engineered to act as biosensors. For example, pHRed, a mutant of the mRFP mKeima, is a protein that ratiometrically measures pH by a change in the ratio of excitation at 585 nm and 440 nm in the range of pH 5.5 – 9 (*28*). In addition to pH (*28, 29*), modified fluorescent proteins can detect changes in the concentration of Hg(II) (*30*) and reactive oxygen species (*31*), while fusions of fluorescent proteins with different protein domains have produced other sensors to detect H_2O_2 (*32*) and calcium concentration (*33, 34*). Fluorescent proteins with compatible optical windows can be used as part of Förster Resonance Energy Transfer (FRET) pairs, where energy transfer from the excited state of the donor protein chromophore serves to excite the acceptor protein chromophore, leading to emission at the acceptor's emission wavelength (λ_{em}). The emission intensity is inversely proportional to the sixth power of the distance between the donor and acceptor, providing the ability to measure the interactions of the two proteins and their fusion partners. For example, a FRET pair can be fused together by a known recognition sequence to measure the activity of a protease (*35*), as well as fused to a protein domain in order to detect conformational changes upon interaction with a specific molecule (*36, 37*).

In addition to the microbiological applications outlined above, the main advantage of RFPs over other members of the FP family is their facilitated use for the purpose of whole animal imaging. Their longer excitation and emission wavelengths are lower in energy, and therefore less damaging to animal cells and less prone to scattering in animal tissues (*12, 38*). More importantly, they are better suited for whole animal imaging because longer wavelength light (ideally 650 – 1100 nm) is not strongly absorbed by water, melanin, or hemoglobin, the strongest absorbers in these systems (*12*). Thus, the light is able to penetrate further, allowing the imaging of internal organs in mice or rats and requiring less intense illumination while providing a stronger, clearer signal (*4, 5, 12*). Far-red fluorescent proteins are defined as having emission wavelengths above 650 nm, although other RFPs are also used for whole animal imaging (*12, 14*). While higher wavelengths are strongly desirable, the brightest far-red monomeric fluorescent protein to date, mNeptune (λ_{em} = 650 nm), has a fluorescence efficiency (quantum yield) of 0.20, far lower than the brightest RFP (mApple, 0.49) (*39*) or proteins of other wavelengths, such as EGFP (0.60) (*40*).

1.3 Fluorescence and Brightness in RFPs

The *p*-hydroxybenzylidene-imidazolinone chromophore of all GFP-like fluorescent proteins is autocatalytically formed in the presence of O_2 from three residues located at positions 65 – 67 (GFP) or 66 – 68 (DsRed) of the internal α-helix (*4, 18, 40, 41)*. The DsRed numbering will be used in this work. The first residue can vary between different FPs, but Tyr67 and Gly68 are usually conserved. In naturally occurring proteins, residue 67 is always a tyrosine, but a histidine (*41-45*), tryptophan (*46-48*), and phenylalanine (*49*) have been artificially inserted at that position to create blue, cyan, and violet fluorescent proteins, respectively. In RFPs, the chromophore conjugation found in a green chromophore is extended (Figure 1.1B) through the formation of an acylimine group between residues 65 and 66 as a result of oxidation coupled to

trans-to-*cis* isomerization of the peptide bond (*50*) (Figure 1.2). The increase in conjugation causes a red-shift in the fluorescence of the chromophore. A branching point has been proposed in the maturation pathway of the red chromophore (*18*), where undesirable conditions can lead a cyclized precursor (Eq. species 1, Figure 1.2) to undergo dehydration to form the green species. Alternatively, the cyclized intermediate is in equilibrium with a precursor to the red chromophore, Eq. species 2, which can undergo oxidation of the F65-M66 peptide bond (mCherry) to form a blue intermediate before the fully mature red species. Since the formation of red chromophore requires a second O_2-mediated step, the proportion of green and red chromophores is dependent on the amount of available oxygen, as well as on surrounding residues that interact with the chromophore backbone (*18, 41*). Moreover, these two species cannot interconvert once formed, so the availability of oxygen during protein expression and maturation is crucial (*18, 41, 51*). While both the neutral and anionic green chromophore species are fluorescent, with different λ_{ex} but emission at the same λ_{em} following excited state proton transfer (ESPT), only the anionic species of the red chromophore is fluorescent (*18*). The red chromophore also undergoes *cis*- to - *trans* isomerization around the double bond between the phenolate and imidazolinone moieties in some proteins from *Entacmaea quadricolor*. The *cis* isomer is brighter, with brightness increasing at higher pH, while the *trans* isomer is dimmer with decreasing brightness at higher pH (*13, 14*). Irradiation with green light causes *trans*- to - *cis* isomerization of the chromophore, leading to a 10-20% increase in fluorescent brightness (*12*).

Brightness is a desirable property of fluorescent proteins, and significant effort has been made to engineer brighter variants of red fluorescent proteins (*3, 52-54, 10, 14*). Brightness is defined as the product of the extinction coefficient and the quantum yield (*1*):

$$Brightness = \varepsilon \cdot \varphi \qquad [1]$$

where ε is the extinction coefficient [$M^{-1}cm^{-1}$] and φ is the fluorescence quantum yield. The quantum yield is a measure of the efficiency of conversion of absorbed light into fluorescence (*55, 56*):

$$\varphi = \frac{number\ of\ emitted\ photons}{number\ of\ absorbed\ photons},\ 0 < \varphi < 1 \qquad [2]$$

The extinction coefficient is a constant in Beer's law, and a measure of the efficiency of absorption of light:

$$\varepsilon = \frac{A}{cl} \qquad [3]$$

where c is the concentration of chromophore [M], A is the absorbance of light at a given wavelength, and l is the path length [cm]. Since it is dependent on the concentration of fully formed red chromophore, the extinction coefficient of a protein is affected by chromophore maturation. On the other hand, the quantum yield is dependent on the propensity of the chromophore to undergo fluorescence instead of alternative pathways of relaxation (Figure 1.3, equation 4). Upon illumination of the chromophore, a ground state electron is excited to a higher electronic level and its associated vibrational states. In order for fluorescence to occur, the electron undergoes vibrational relaxation back to the lowest excited state vibrational level and then relaxes back to the ground state, emitting a photon. Since some energy was lost due to the vibrational relaxation, this photon is always lower in energy than the exciting photon. The energy difference between the ground and excited states determines the wavelength of light, with smaller energy gaps producing longer wavelengths. Since fluorescence is not the only pathway back to the ground state, the quantum yield equation can be written as follows:

$$\varphi = \frac{k_F}{k_F + k_{IC} + k_{EC} + k_{ISC} + k_{DISS} + k_{PREDISS}} \quad [4]$$

where each k value represents a rate constant : k_F - fluorescence, k_{IC} - internal conversion (energy transfer to matching vibrational levels not involving a spin change, e.g. singlet-singlet), k_{EC} - external conversion (energy transfer due to collisional deactivation), k_{ISC} - intersystem crossing (transfer of energy involving a spin change, e.g. singlet-triplet and sometimes leading to phosphorescence), k_{DISS} - dissociation (excitation to an energy level strong enough to break a bond), and $k_{PREDISS}$ - predissociation (relaxation to an energy level strong enough to break a bond) (*57*). All the rates except k_F represent non-radiative pathways for relaxation of an excited-state electron.

In RFPs in particular, the gap between the excited and ground states of the chromophore is smaller than for proteins exhibiting lower wavelength fluorescence (higher energy). This could represent an intrinsic propensity for non-radiative decay in red fluorescent proteins, through internal conversion of energy from an excited state vibrational level to a matching accessible vibrational level in the ground state. While the quantum yields of monomeric fluorescent proteins of lower wavelengths are generally > 0.5 (*4, 5*), with mTurquoise (λ_{em} = 474 nm) reaching 0.93 (*46*), most monomeric RFPs fall below this threshold (mCherry φ = 0.22 (*10*), mRaspberry φ = 0.15 (*15*), mPlum φ = 0.10 (*15*), mKate2 φ = 0.40 (*14*), mRuby2 φ = 0.38 (*53*)), with the highest quantum yield mRFP to date being mApple (φ = 0.49) (*39*). The extinction coefficients of most mRFPs are in the range of 40,000 – 100,000 $M^{-1}cm^{-1}$, similar to their blue, green, and yellow counterparts as well as to most commercially available small molecule fluorophores (*4, 5*). Therefore, quantum yield plays an important role in engineering the brightness of these proteins.

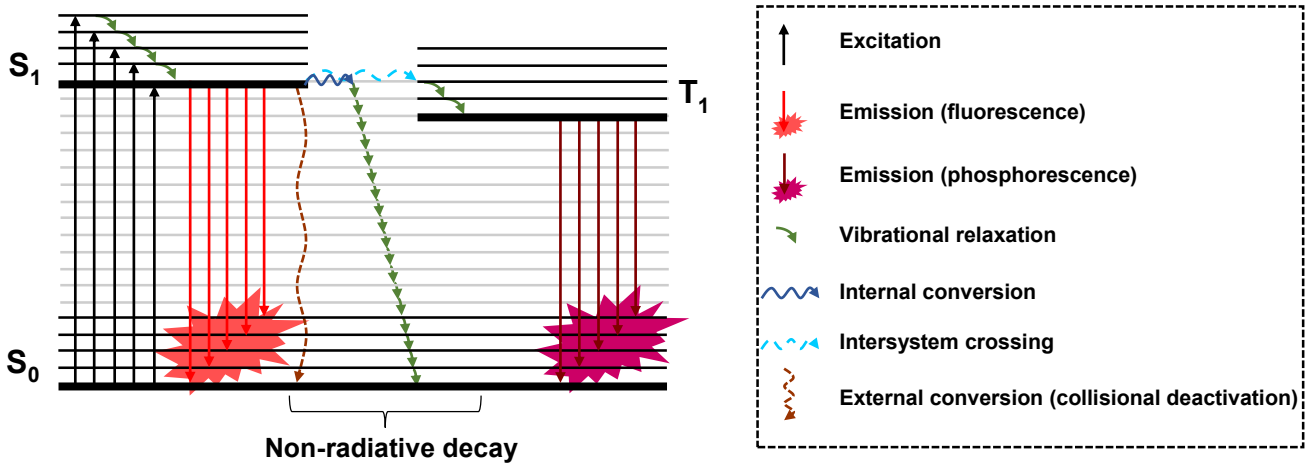

Figure 1.3. Jablonski diagram representing the radiative and non-radiative pathways for relaxation of an excited-state electron in a fluorophore.

Different schools of thought exist concerning the characteristics of RFP chromophores that are desirable for brightness. Planarity of the chromophore (*58* - *60*), dynamics (*9, 61, 62*), *cis*-to-*trans* isomerization (*13, 59, 12, 14*), and positioning of Lys70 above the chromophore (*63*) have all been hypothesized to play a role. In the mFruits family, higher twist and tilt angles of the chromophore phenolate group appear to correlate with decreased fluorescence quantum yield (*60*). This crystallographic study of different mFruits showed that mCherry (twist = 11.3°, tilt = 13.7°, φ = 0.22), mStrawberry at pH 9.5 (twist = 5.9°, tilt = 17.9°, φ = 0.31), and mOrange (twist = 1.5°, tilt = 10.3°, φ = 0.69) all have significantly higher tilt / twist angles and lower quantum yields than wild type GFP (twist = 0.0°, tilt = 3.7°, φ = 0.79) and DsRed (twist = 0.9°, tilt = 0.2°, φ = 0.79). The data for mStrawberry at pH 10.5 (twist = 0.4°, tilt = 4.6°, φ = 0.36), with lower angles but higher quantum yield than the protein at pH 9.5, appears to support the hypothesis that planarity of the chromophore is important for increased quantum yield. However, the angles seen in the mStrawberry chromophore at pH 10.5 are much closer to those of the DsRed chromophore than

mStrawberry at pH 9.5, while the quantum yield remains much lower than that of DsRed. This suggests that planarity may not be the only factor affecting quantum yield. Molecular dynamics (MD) simulations of pressure effects on mStrawberry and mCherry followed by quantum mechanics / molecular mechanics (QM/MM) modelling suggest that the quantum yield of FPs is dependent on the flexibility of the protein. More specifically, in more dynamic protein backbones such as mStrawberry, the range of motion of the chromophore and the instability of its hydrogen bonded network to surrounding residues negatively affect its quantum yield. According to this study, a protein with a sufficiently rigid backbone and chromophore would not benefit further from further rigidification (*9*). In a different MD study, the rotational freedom of the chromophores of seven yellow, blue, and green FPs in their respective environments was compared to their quantum yields (*61*). It was determined that dynamics of the chromophore correlated negatively with quantum yield, and that the simulated chromophore structures of brighter FPs were not necessarily coplanar. Given the contradictory nature of possible contributing factors found in the literature, more evidence is needed to ascertain the importance of each factor to the brightness of FP chromophores.

1.4. Case Studies - Engineering of Brightness in RFPs

RFPs are often engineered through directed evolution. In particular, the effect of a design on the brightness of the protein is difficult to predict rationally, since there are many degrees of freedom that need to be considered. The first monomeric RFP, mRFP1, was engineered sequentially by first breaking the interfaces between the monomers of the tetramer DsRed, and then recovering the fluorescence of the resulting monomer mRFP0.1. The process of creating mRFP1 from mRFP0.1 required screening six semi-random combinatorial libraries ranging in size from 2,300 to 6,000,000 mutants each, finally introducing a total of one deletion and 23 mutations

in order to obtain a quantum yield of 0.25 and an extinction coefficient of 44,000 $M^{-1}cm^{-1}$ (*3*). This project, while incurring a large time and resource cost, produced a useful monomeric fluorescent protein that was the precursor of the mFruits family, which remains one of the most useful mRFP families today. The mFruits family in turn was created from mRFP1 by screening millions of mutants during multiple rounds of directed evolution together with introducing random amino acids at specified positions, leading to a diversity of colours and brightnesses (*10*). The two original red variants in the family, mStrawberry and mCherry, both exhibited increased brightness compared to mRFP1 – mStrawberry through a 1.2x increase in quantum yield and 1.8x increase in extinction coefficient, and mCherry through a 1.4x increase in the extinction coefficient and a reduction in quantum yield to 0.9x. While mCherry is the result of two rounds of directed evolution together with smaller random libraries, mStrawberry underwent three additional rounds of directed evolution.

In a recent example, mKate2 was engineered from mKate, a monomeric RFP derived from the eqFP578 protein found in the sea anemone *Entacmaea quadricolor*. The chromophore of these proteins exists in two conformations (*13*), a *cis* and *trans* conformer of the double bond between the phenolate and imidazolinone moieties. While both conformations are fluorescent, the *cis*-conformation is the brighter of the two. The mKate2 mutant was engineered using saturation mutagenesis at position 165 to break a stabilizing hydrogen bonding interaction of the wild type serine to the *trans* conformation of the chromophore, thus favouring the *cis* isomer (*14*). This was followed by two rounds of directed evolution, at an average of seven mutations per 1000 bp, leading to mutations that improved the maturation efficiency. A total of five mutations caused an increase in quantum yield from 0.28 to 0.40 (1.4x) and a 2.0x increase in extinction coefficient, leading to an overall brightness increase of 2.8x.

A similar protein from the same organism, eqFP611, led to the development of the monomeric protein mRuby (λ_{em} = 605 nm, φ = 0.35, ε = 112,000 $M^{-1}cm^{-1}$). In turn, mRuby2 was engineered from mRuby for stabilization of the *cis* conformation of the chromophore, leading to a 1.1x increase in both quantum yield and extinction coefficient (*53*). This variant, engineered in concert with the GFP Clover as an optimized FRET pair, was one example of RFP brightness design by screening a small library of mutants. The authors identified residues near the chromophore and mutated them in order to stabilize the brighter *cis* isomer. The resultant protein has a total of four mutations compared to mRuby, which led to a 1.1x increase in brightness.

In a 2014 USPTO patent awarded to Glick, B. S. *et al,* a monomeric DsRed variant called DsRed.M1 was mutated by screening approximately 200,000 mutants in two rounds of homology-guided site-specific mutagenesis of more than a dozen positions surrounding the chromophore. This achieved the creation of three different red variants. The most red-shifted mutant, called mRed-595 for its excitation wavelength, had eighteen mutations which increased quantum yield from 0.17 to 0.31 (1.8x), increased the extinction coefficient 2.1x, and led to an overall 3.7x increase in brightness (*52*). Some mutations, like Q66M, M83L, L150M, F177V, S197I, and N203S, are common to mCherry, which also has a similar emission wavelength (λ_{em} = 610 nm). Interestingly, this mutant also has a 25 nm red shift (λ_{em} = 607 nm) relative to the parent DsRed.M1, making the quantum yield and brightness increases even more remarkable.

All of the examples highlighted in this section achieve increases in brightness through a greater or equal increase of extinction coefficient than quantum yield. These reflect changes in the efficiency of maturation, since the extinction coefficient is dependent on the amount of mature red chromophore present in solution. In a recent example, a red-shifted mutant of mKate, mNeptune, was subjected to saturation mutagenesis at eleven positions near the chromophore in order to

identify the quantum yield improving mutation S28H (*54*). The residues at this position participate in a water-mediated hydrogen bond with the acylimine oxygen atom of the chromophore, and the larger histidine residue was hypothesized to have a stronger hydrogen bond while simultaneously packing around the chromophore better, reducing its accessible vibrational states. This produced a protein with a quantum yield increase of 0.05 (1.3x) and a similar extinction coefficient to mNeptune, leading to an overall brightness increase of 1.3x. This case is an example of the improvement of quantum yield, as opposed to the extinction coefficient, by affecting the chromophore's environment through a relatively small library of a few hundred mutants. While the positions were selected rationally, the use of saturation mutagenesis and the necessity to maintain a small library size meant that only one position was mutated at a time. Therefore, the method did not take into account complementary mutations that could have worked together to produce even better results, or even the additive effect of more than one beneficial mutation being introduced at once. On the other hand, in the other examples besides mRuby2, the improved variants are the result of screening thousands to millions of mutants in multiple rounds through random or semi-random methods. Rational design of quantum yield–increasing mutations remains a significant challenge, and could prove a useful tool for engineering novel red and far-red FPs that can match the brightness found in green, yellow, and cyan fluorescent proteins.

1.5. Strategies for Designing an Increased Quantum Yield

In order to improve the efficiency of emission of light in the chromophore, it is beneficial to review some common small fluorophores and the pathways through which they undergo quenching. Ethidium bromide, a fluorophore that undergoes a 21x increase in quantum yield upon binding to double stranded DNA, is protected from quenching by the surrounding solvent when it is inserted in to the hydrophobic region of the DNA strand (*64*). In the case of the RFP

chromophore, its location in the core of the protein already protects it from most of the surrounding solvent. In other cases, such as anthracene and its derivatives, quenching can occur through contact with molecular oxygen and nitric oxide (*65, 66*). While this has not been described in RFPs, some studies indicate that access to O_2 reduces their photostability (*39, 67*). In addition, studies in GFP-based sensors indicate that reducing the integrity of β barrel of the chromophore, which would lead to increased oxygen diffusion, also reduces the brightness by increasing the possibility of solvent quenching (*68*). Therefore, a design that would aim to reduce diffusion through openings in the protein barrel structure could lead to increases in both photostability and quantum yield. However, O_2 is required for the efficient maturation of all chromophores in this family, and RFPs in particular require additional oxygen for the formation of the red-shifting acylimine group (Figure 1.2) (*18*). Therefore, limiting its diffusion to the core of the β barrel risks stunting the maturation of the chromophore, and by extension, the brightness of the protein. Lastly, in the case of the dye auramine O, quantum yield increased at higher viscosity when the molecule was dissolved in different concentrations of glycerol or dextrose-glycerol-water mixtures, indicating that the loss of fluorescence efficiency was due to flexibility of the fluorophore (*69*). For coumarin and rhodamine dyes, the rigidification through covalent modification of groups that were dynamic in the excited state also increased the efficiency of fluorescence (*69-71*). Since there is significant evidence in the literature (*9, 72*) that rigidifying the chromophore of RFPs or the entire protein structure would help to increase the efficiency of fluorescence, targeting this approach appears most beneficial.

1.6. Description of the Project

1.6.1. Objective

Engineering brighter RFPs is desired. Barring a few cases (*53, 54*), brighter red fluorescent protein variants have been created primarily by achieving increases in the extinction coefficient

through directed evolution or other random methods that lead to large library sizes (*3, 10, 13, 52*). There are two drawbacks to this approach. Firstly, the screening of large libraries is costly in terms of time and effort. Directed evolution approaches often require screening thousands to millions of mutants in order to find a favourable mutation. Secondly, large library sizes necessitate the use of fluorescence intensity as a criterion for screening. This method does not allow for the differentiation of improvements in only one of the two parameters contributing to brightness: the quantum yield and the extinction coefficient (equation 1-3).

A rational approach would allow for faster engineering of brighter RFPs. This can be achieved by targeting either the quantum yield or the extinction coefficient. Mutations that improve the extinction coefficient are difficult to rationalize. In some cases, mutations stabilizing the more favourable *cis*-conformation of the chromophore are introduced to achieve this goal (*14*). However, once this becomes the dominant conformation, predicting mutations that target the electronic states of the chromophore in order to achieve more efficient absorbance is nearly impossible. Furthermore, the extinction coefficient calculation for red fluorescent proteins as performed in many cases (*1-3, 6, 7, 10-12, 14, 16-18*) contains an artifact that allows for improvements in maturation efficiency to be interpreted as an increase in brightness. This is because not all folded proteins contain mature red chromophore, while protein concentration is determined using methods that do not account for the presence of immature or green chromophore in the protein population. This has been addressed through the development of the dynamic difference method (*73*) (Chapter 5.3.2), but this is a recent improvement that has not yet been adopted as standard practice in related literature. The complications associated with the extinction coefficient prediction and calculation led us to explore the possibility of targeting the quantum yield instead.

The objective of the present work was to develop a **rational design approach** to increase the **quantum yield** of red fluorescent proteins. This approach could be used alone or in combination with directed evolution methods to maximize the engineering potential for brighter red fluorescent proteins.

1.6.2. Hypothesis

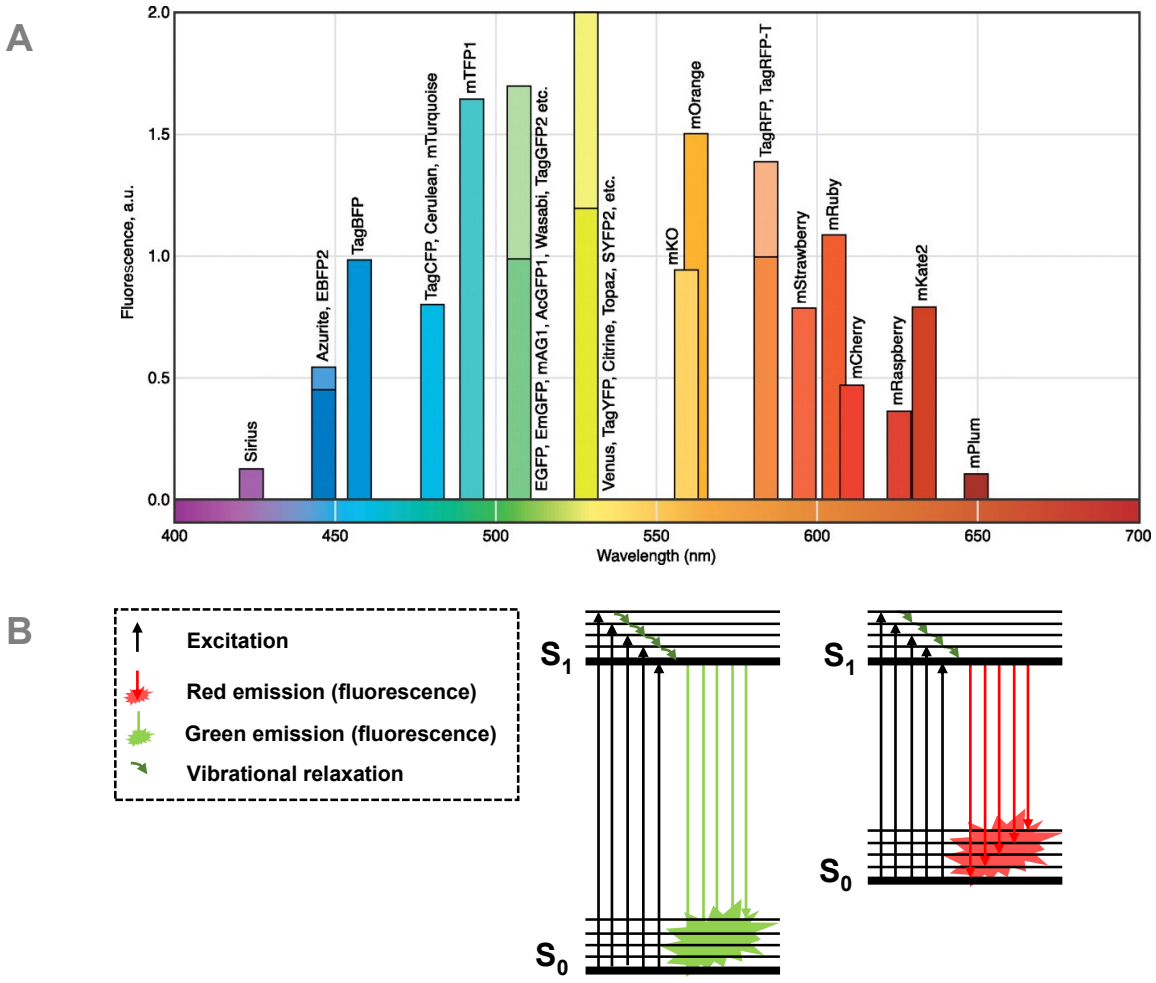

Figure 1.4. Monomeric fluorescent brightness and energy levels at different wavelengths. Relative brightness of various monomeric fluorescent proteins as a function of their emission wavelength (A) (*4*). The emission wavelength depends on the energy difference between the ground and the excited states of the chromophore (B).

Figure 1.4A modified from: Chudakov, D. M *et al.* (2010) *Physiol Rev* **90**(3). doi: 10.1152/physrev.00038.2009

It is known that red fluorescent proteins of longer emission wavelengths tend to exhibit lower brightness compared to their counterparts in the green – yellow FP range (*4*) (Figure 1.4).

This can be attributed to a smaller gap between the ground and excited states of the chromophore, which determines the energy of the emitted photon and therefore its wavelength (Figure 1.4).

In an ideal case, excitation of the electrons from $S_0 \rightarrow S_1$ electronic states would lead to rapid vibrational relaxation and a return to the ground electronic state accompanied by the emission of a photon (Figure 1.5A). However, if the gap between the two electronic states is sufficiently low, the accessible vibrational states of the ground and excited states of the chromophore could overlap (Figure 1.5B). This overlap provides a pathway for non-radiative decay through internal conversion from $S_1 \rightarrow S_0$ and vibrational relaxation to the lowest-energy ground vibrational state.

We hypothesized that we could predict mutations that would restrict the accessibility of the higher vibrational levels of the ground state by limiting the conformational flexibility of the chromophore (Figure 1.5C). This would reduce the chromophore's propensity for non-radiative decay through internal conversion, and improve its efficiency of fluorescence.

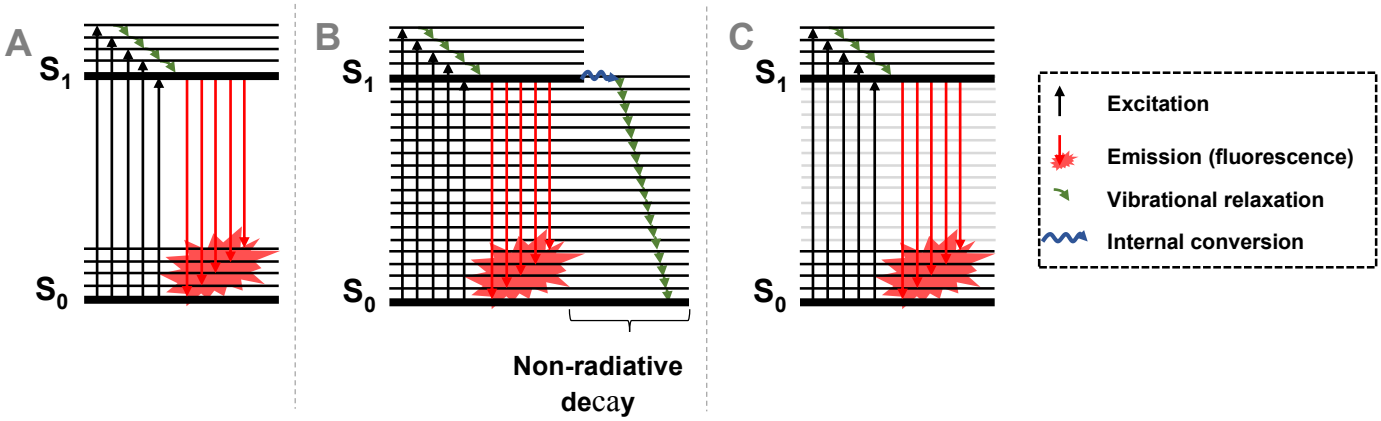

Figure 1.5. Jablonski diagrams representing the hypothesis. The ideal pathway to fluorescence (A), the case where internal conversion leads to non-radiative decay (B), and the case where internal conversion is prevented by limiting accessibility to higher-level vibrational states of the ground state S_0 (C).

1.6.3. Method

In order to test the hypothesis, we selected the mCherry mutant mRojoA (*74*). This protein, containing six mutations, was previously engineered for red-shifted fluorescence (Figure 1.6). The 22-nm red shift in emission is partially due to a π – stacking interaction with Y197 below the chromophore.

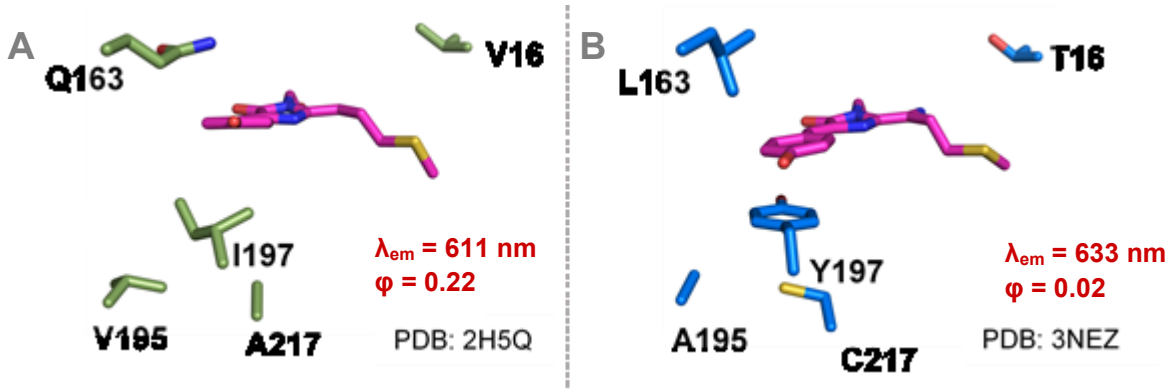

Figure 1.6. Residues surrounding the chromophore in mCherry (A) that were mutated to create mRojoA (B). One mutation, R125H, is not included in the figure since it is not located near the chromophore.

mRojoA was the ideal starting point for testing the hypothesis because of two important characteristics. The first is its low quantum yield of 0.02, which provides a low baseline for screening and allows for easier detection of brighter variants. The second is the existence of a bulky aromatic residue below the chromophore at position 197. The introduction of an additional aromatic residue above the chromophore would restrict its conformational flexibility by pinching the phenolate group between two bulky aromatics (Figure 1.7). In addition to being some of the largest amino acids, aromatics are ideal because of their planarity, which affords the side chain less flexibility while allowing for additional steric complementarity and a possibility of π –

stacking with the chromophore. Tyrosine, phenylalanine, and histidine were selected for inclusion at position 63.

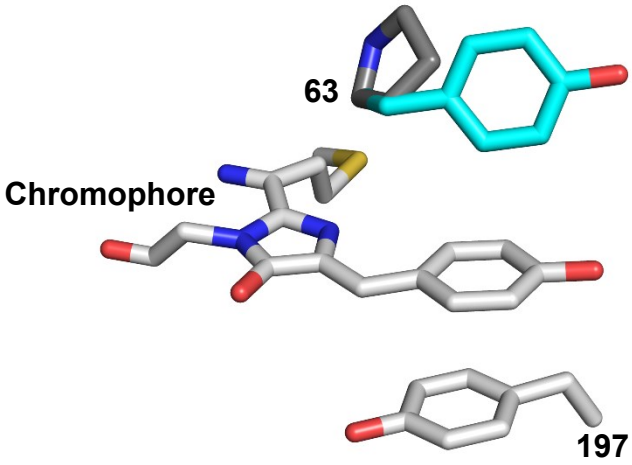

Figure 1.7. The chromophore and Y197 of mRojoA (white) with P63 (dark gray) and one of the proposed aromatic mutations at position 63 (P63Y) (cyan).

Since the single mutants did not result in mature chromophore formation (Figure 2.S1A), a small library was designed to help incorporate the aromatic mutations at position 63. The aromatic residues alone were found to disrupt chromophore maturation, since they were in close proximity to the chromophore and replacing a significantly smaller proline at position 63. Nearby positions that could clash with the aromatics were identified (Figure 1.8A) and mutated accordingly. The rationally designed libraries contain mutations at three positions surrounding position 63 and the chromophore (Figure 1.8B, Table 1.1). At position 143, which is proximal to position 63, the library introduced the three smallest amino acids (A, G, S) in order to accommodate the inclusion of the bulky aromatics. Position 163, which was also near the chromophore, was mutated to the smaller valine residue. Lastly, the T16V mutation was included in the library, since it had previously been shown to improve chromophore maturation (*18*). The parent protein residues were included in the library in order to help preserve function in the mutant population.

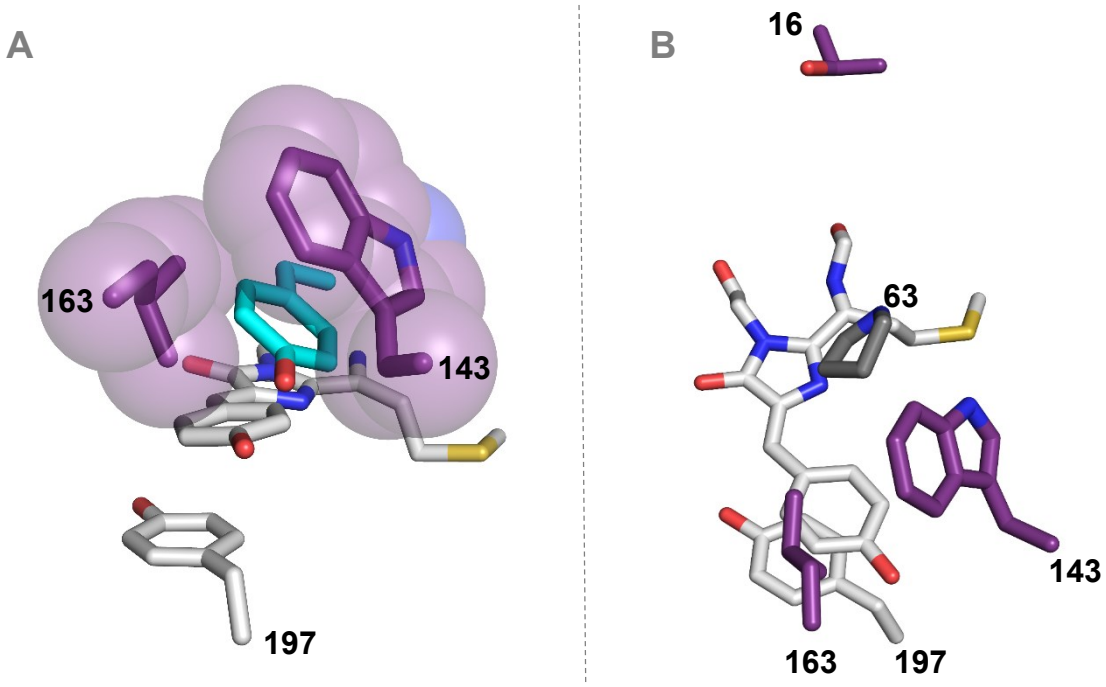

Figure 1.8. Residues identified for rational design (purple) in order to accommodate an aromatic at position 63. The chromophore and Y197 are shown in white, P63 in dark gray, and Y63 in cyan. Space-filling models (A) of W143 and L163 show the clash between these residues and the desired Y63 mutation. The residues selected for mutagenesis are T16, W143, and L163 (B).

Table 1.1. Rationally designed libraries of mRojoA for inclusion of aromatics at position 63.

	Residues at position				Library size
	16[1,2]	**63**	**143**	**163**	
mRojoA	T	P	W	L	-
Library 1	VT	H	AGSW	LV	16
Library 2	VT	Y	AGSW	LV	16
Library 3	VT	F	AGSW	LV	16
Mutagenesis codon	*GTT*	*CAT / TAC / TTC*	*KGG*	*STG*	-

[1] V16 was previously shown to improve maturation (*18*).
[2] Degenerate codons were not employed due to the impossibility of excluding unwanted mutations. Instead, the mutation was introduced individually and mixed with DNA containing the parent protein threonine to create the library.

A general overview of the method employed to make and screen the library is presented in Figure 1.9. The library was introduced into the mRojoA gene through the use of splicing through overlap extension (SOE) mutagenesis with the help of mutagenesis primers (Chapter 5.1, Table 5.1.1). The SOE products were ligated into expression vectors and transformed into *E. coli* cells in order to create cell stocks of the mutant library. Sequencing of fluorescent mutants allowed for the identification of combinations of mutations contributing to the greatest increases in brightness. The brightest mutants, and mutants containing related mutations, were characterized in detail in order to elucidate the relationships between the presence of certain amino acids and increases in quantum yield.

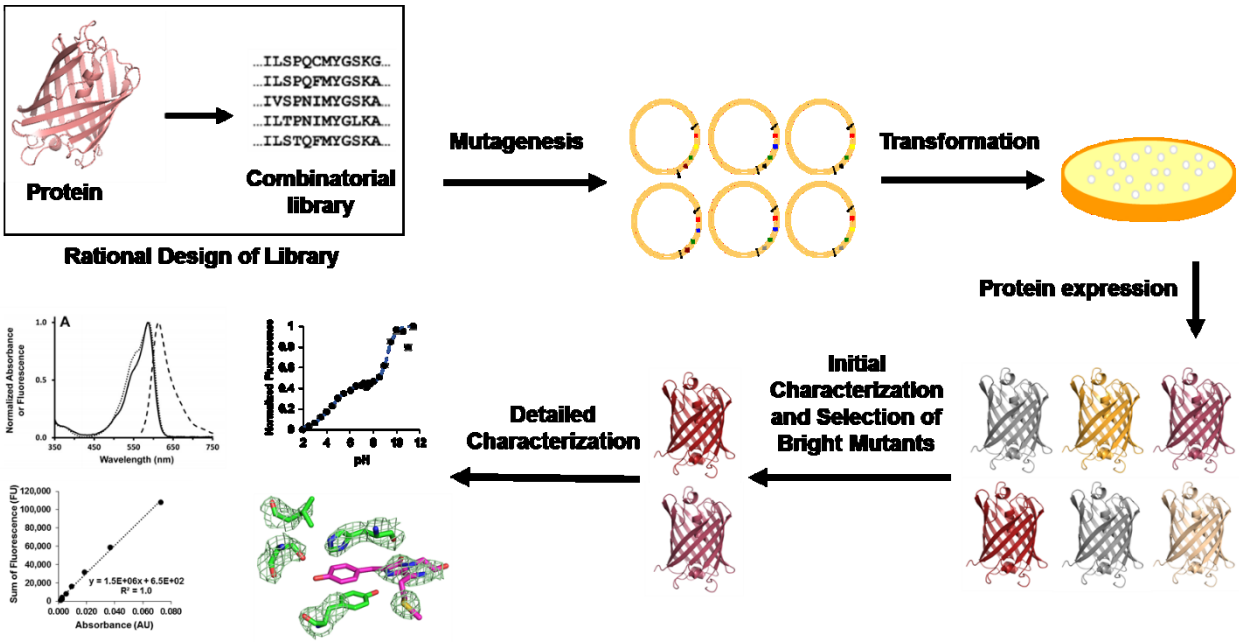

Figure 1.9. Flow chart of the steps used to create, screen, and characterize a small library of mutants.

1.7. References

1. Matz, M. (1999) Fluorescent proteins from nonbioluminescent anthozoa species (vol 17, pg 969, 1999), *Nature Biotechnology 17*(12), 1227-1227.

2. Baird, G., Zacharias, D., and Tsien, R. (2000) Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral, *Proceedings of the National Academy of Sciences of the United States of America 97*(22), 11984-11989. doi:10.1073/pnas.97.22.11984

3. Campbell, R., Tour, O., Palmer, A., Steinbach, P., Baird, G., Zacharias, D., and Tsien, R. (2002) A monomeric red fluorescent protein, *Proceedings of the National Academy of Sciences of the United States of America 99*(12), 7877-7882. doi:10.1073/pnas.082243699

4. Chudakov, D. M., Matz, M. V., Lukyanov, S., and Lukyanov, K. A. (2010) Fluorescent proteins and their applications in imaging living cells and tissues, *Physiological Reviews 90*(3), 1103-1163. doi:10.1152/physrev.00038.2009

5. Nienhaus, K., and Nienhaus, G. U. (2014) Fluorescent proteins for live-cell imaging with super-resolution, *Chemical Society Reviews 43*(4), 1088-1106. doi:10.1039/c3cs60171d

6. Fradkov, A., Verkhusha, V., Staroverov, D., Bulina, M., Yanushevich, Y., Martynov, V., Lukyanov, S., and Lukyanov, K. (2002) Far-red fluorescent tag for protein labelling, *Biochemical Journal 368*, 17-21. doi:10.1042/BJ20021191

7. Mizuno, H., Sawano, A., Eli, P., Hama, H., and Miyawaki, A. (2001) Red fluorescent protein from discosoma as a fusion tag and a partner for fluorescence resonance energy transfer, *Biochemistry 40*(8), 2502-2510. doi:10.1021/bi002263b

8. Bulavina, A., McDermott, G., Matveeva, E., and Savitsky, A. (2003) Oligomerization strongly influences the brightness of the DsRed fluorescence, *Genetically Engineered and Optical Probes for Biomedical Applications 4967*, 78-87. doi:10.1117/12.479830

9. Laurent, A. D., Mironov, V. A., Chapagain, P. P., Nemukhin, A. V., and Krylov, A. I. (2012) Exploring structural and optical properties of fluorescent proteins by squeezing: Modeling high-pressure effects on the mstrawberry and mcherry red fluorescent proteins, *Journal of Physical Chemistry B 116*(41), 12426-12440. doi:10.1021/jp3060944

10. Shaner, N., Campbell, R., Steinbach, P., Giepmans, B., Palmer, A., and Tsien, R. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from discosoma sp red fluorescent protein, *Nature Biotechnology 22*(12), 1567-1572. doi:10.1038/nbt1037

11. Merzlyak, E. M., Goedhart, J., Shcherbo, D., Bulina, M. E., Shcheglov, A. S., Fradkov, A. F., Gaintzeva, A., Lukyanov, K. A., Lukyanov, S., Gadella, T. W. J., and Chudakov, D. M. (2007) Bright monomeric red fluorescent protein with an extended fluorescence lifetime, *Nature Methods 4*(7), 555-557. doi:10.1038/NMETH1062

12. Shcherbo, D., Merzlyak, E. M., Chepurnykh, T. V., Fradkov, A. F., Ermakova, G. V., Solovieva, E. A., Lukyanov, K. A., Bogdanova, E. A., Zaraisky, A. G., Lukyanov, S., and Chudakov, D. M. (2007) Bright far-red fluorescent protein for whole-body imaging, *Nature Methods 4*(9), 741-746. doi:10.1038/NMETH1083

13. Pletnev, S., Shcherbo, D., Chudakov, D. M., Pletneva, N., Merzlyak, E. M., Wlodawer, A., Dauter, Z., and Pletnev, V. (2008) A crystallographic study of bright far-red fluorescent protein mKate reveals pH-induced cis-trans isomerization of the chromophore, *Journal of Biological Chemistry 283*(43), 28980-28987. doi:10.1074/jbc.M800599200

14. Shcherbo, D., Murphy, C. S., Ermakova, G. V., Solovieva, E. A., Chepurnykh, T. V., Shcheglov, A. S., Verkhusha, V. V., Pletnev, V. Z., Hazelwood, K. L., Roche, P. M., Lukyanov, S., Zaraisky, A. G., Davidson, M. W., and Chudakov, D. M. (2009) Far-red fluorescent tags for protein imaging in living tissues, *Biochemical Journal 418*, 567-574. doi:10.1042/BJ20081949

15. Wang, L., Jackson, W. C., Steinbach, P. A., and Tsien, R. Y. (2004) Evolution of new nonantibody proteins via iterative somatic hypermutation, *Proceedings of the National Academy of Sciences of the United States of America 101*(48), 16745-16749. doi:10.1073/pnas.0407752101

16. Lin, M. Z., McKeown, M. R., Ng, H., Aguilera, T. A., Shaner, N. C., Campbell, R. E., Adams, S. R., Gross, L. A., Ma, W., Alber, T., and Tsien, R. Y. (2009) Autofluorescent proteins with excitation in the optical window for intravital imaging in mammals, *Chemistry & Biology 16*(11), 1169-1179. doi:10.1016/j.chembiol.2009.10.009

17. Subach, F. V., Piatkevich, K. D., and Verkhusha, V. V. (2011) Directed molecular evolution to design advanced red fluorescent proteins, *Nature Methods 8*(12), 1019-1026. doi:10.1038/NMETH.1776

18. Moore, M. M., Oteng-Pabi, S. K., Pandelieva, A. T., Mayo, S. L., and Chica, R. A. (2012) Recovery of red fluorescent protein chromophore maturation deficiency through rational design, *Public Library of Science One 7*(12), e52463. doi:10.1371/journal.pone.0052463

19. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W., and Prasher, D. (1994) Green fluorescent protein as a marker for gene-expression, *Science 263*(5148), 802-805. doi:10.1126/science.8303295

20. Markwardt, M. L., Kremers, G., Kraft, C. A., Ray, K., Cranfill, P. J. C., Wilson, K. A., Day, R. N., Wachter, R. M., Davidson, M. W., and Rizzo, M. A. (2011) An improved cerulean fluorescent protein with enhanced brightness and reduced reversible photoswitching, *Public Library of Science One 6*(3), e17896. doi:10.1371/journal.pone.0017896

21. Toomre, D., and Bewersdorf, J. (2010) A new wave of cellular imaging, *Annual Review of Cell and Developmental Biology, Vol 26 26*, 285-314. doi:10.1146/annurev-cellbio-100109-104048

22. Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G., and Cormier, M. J. (1992) Primary structure of the aequorea-victoria green-fluorescent protein, *Gene 111*(2), 229-233. doi:10.1016/0378-1119(92)90691-H

23. Hoi, H., Shaner, N. C., Davidson, M. W., Cairo, C. W., Wang, J., and Campbell, R. E. (2010) A monomeric photoconvertible fluorescent protein for imaging of dynamic protein localization, *Journal of Molecular Biology 401*(5), 776-791. doi:10.1016/j.jmb.2010.06.056

24. Young, C. L., Raden, D. L., Caplan, J. L., Czymmek, K. J., and Robinson, A. S. (2012) Cassette series designed for live-cell imaging of proteins and high-resolution techniques in yeast, *Yeast 29*(3-4), 119-136. doi:10.1002/yea.2895

25. Hu, C. D., and Kerppola, T. K. (2003) Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis, *Nature Biotechnology 21*(5), 539-545. doi:10.1038/nbt816

26. Kerppola, T. K. (2008) Biomolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells, *Annual Review of Biophysics 37*, 465-487. doi:10.1146/annurev.biophys.37.032807.125842

27. Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., and Hess, H. F. (2006) Imaging intracellular fluorescent proteins at nanometer resolution, *Science 313*(5793), 1642-1645. doi:10.1126/science.1127344

28. Tantama, M., Hung, Y. P., and Yellen, G. (2011) Imaging intracellular pH in live cells with a genetically encoded red fluorescent protein sensor, *Journal of the American Chemical Society 133*(26), 10034-10037. doi:10.1021/ja202902d

29. Johnson, D. E., Ai, H., Wong, P., Young, J. D., Campbell, R. E., and Casey, J. R. (2009) Red fluorescent protein pH biosensor to detect concentrative nucleoside transport, *Journal of Biological Chemistry284*(31), 20499-20511. doi:10.1074/jbc.M109.019042

30. Jiang, T., Guo, D., Wang, Q., Wu, X., Li, Z., Zheng, Z., Yin, B., Xia, L., Tang, J., Luo, W., Xia, N., and Jiang, Y. (2015) Developing a genetically encoded green fluorescent protein mutant for sensitive light-up fluorescent sensing and cellular imaging of hg(II), *Analytica Chimica Acta 876*, 77-82. doi:10.1016/j.aca.2015.03.026

31. Ostergaard, H., Tachibana, C., and Winther, J. R. (2004) Monitoring disulfide bond formation in the eukaryotic cytosol, *Journal of Cell Biology 166*(3), 337-345. doi:10.1083/jcb.200402120

32. Belousov, V. V., Fradkov, A. F., Lukyanov, K. A., Staroverov, D. B., Shakhbazov, K. S., Terskikh, A. V., and Lukyanov, S. (2006) Genetically encoded fluorescent indicator for intracellular hydrogen peroxide,*Nature Methods 3*(4), 281-286. doi:10.1038/NMETH866

33. Tallini, Y. N., Ohkura, M., Choi, B. R., Ji, G. J., Imoto, K., Doran, R., Lee, J., Plan, P., Wilson, J., Xin, H. B., Sanbe, A., Gulick, J., Mathai, J., Robbins, J., Salama, G., Nakai, J., and Kotlikoff, M. I. (2006) Imaging cellular signals in the heart in vivo: Cardiac expression of the high-signal Ca2+ indicator GCaMP2, *Proceedings of the National Academy of Sciences of the United States of America 103*(12), 4753-4758. doi:10.1073/pnas.0509378103

34. Zhao, Y., Araki, S., Jiahui, W., Teramoto, T., Chang, Y., Nakano, M., Abdelfattah, A. S., Fujiwara, M., Ishihara, T., Nagai, T., and Campbell, R. E. (2011) An expanded palette of genetically encoded Ca2+ indicators, *Science 333*(6051), 1888-1891. doi:10.1126/science.1208592

35. Mahajan, N. P., Harrison-Shostak, D. C., Michaux, J., and Herman, B. (1999) Novel mutant green fluorescent protein protease substrates reveal the activation of specific caspases during apoptosis,*Chemistry & Biology 6*(6), 401-409. doi:10.1016/S1074-5521(99)80051-9

36. Baker, B. J., Mutoh, H., Dimitrov, D., Akemann, W., Perron, A., Iwamoto, Y., Jin, L., Cohen, L. B., Isacoff, E. Y., Pieribone, V. A., Hughes, T., and Knopfel, T. (2008) Genetically encoded fluorescent sensors of membrane potential, *Brain Cell Biology 36*(1-4), 53-67. doi:10.1007/s11068-008-9026-7

37. Green, H. M., and Alberola-Ila, J. (2005) Development of ERK activity sensor, an in vitro, FRET-based sensor of extracellular regulated kinase activity, *BMC Chemical Biology 5*, 1-Article No.: 1. doi:10.1186/1472-6769-5-1

38. Jobsis, F. F. (1977) Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters, *Science 198*(4323), 1264-1267. doi:10.1126/science.929199

39. Shaner, N. C., Lin, M. Z., McKeown, M. R., Steinbach, P. A., Hazelwood, K. L., Davidson, M. W., and Tsien, R. Y. (2008) Improving the photostability of bright monomeric orange and red fluorescent proteins, *Nature Methods 5*(6), 545-551. doi:10.1038/nmeth.1209

40. Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP), *Gene 173*(1), 33-38. doi:10.1016/0378-1119(95)00685-0

41. Strack, R. L., Strongin, D. E., Mets, L., Glick, B. S., and Keenan, R. J. (2010) Chromophore formation in DsRed occurs by a branched pathway, *Journal of the American Chemical Society 132*(24), 8496-8505. doi:10.1021/ja1030084

42. Ai, H., Shaner, N. C., Cheng, Z., Tsien, R. Y., and Campbell, R. E. (2007) Exploration of new chromophore structures leads to the identification of improved blue fluorescent proteins, *Biochemistry 46*(20), 5904-5910. doi:10.1021/bi700199g

43. Heim, R., and Tsien, R. Y. (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer, *Current Biology 6*(2), 178-182. doi:10.1016/S0960-9822(02)00450-5

44. Kremers, G., Goedhart, J., van den Heuvel, D. J., Gerritsen, H. C., and Gadella, T. W. J.,Jr. (2007) Improved green and blue fluorescent proteins for expression in bacteria and mammalian cells, *Biochemistry46*(12), 3775-3783. doi:10.1021/bi0622874

45. Mena, M. A., Treynor, T. P., Mayo, S. L., and Daugherty, P. S. (2006) Blue fluorescent proteins with enhanced brightness and photostability from a structurally targeted library, *Nature Biotechnology24*(12), 1569-1571. doi:10.1038/nbt1264

46. Goedhart, J., van Weeren, L., Hink, M. A., Vischer, N. O. E., Jalink, K., and Gadella, T. W. J.,Jr. (2010) Bright cyan fluorescent protein variants identified by fluorescence lifetime screening, *Nature Methods7*(2), 137-U74. doi:10.1038/NMETH.1415

47. Heim, R., Prasher, D. C., and Tsien, R. Y. (1994) Wavelength mutations and posttranslational autoxidation of green fluorescent protein, *Proceedings of the National Academy of Sciences of the United States of America 91*(26), 12501-12504. doi:10.1073/pnas.91.26.12501

48. Rizzo, M. A., Springer, G. H., Granada, B., and Piston, D. W. (2004) An improved cyan fluorescent protein variant useful for FRET, *Nature Biotechnology 22*(4), 445-449. doi:10.1038/nbt945

49. Tomosugi, W., Matsuda, T., Tani, T., Nemoto, T., Kotera, I., Saito, K., Horikawa, K., and Nagai, T. (2009) An ultramarine fluorescent protein with increased photostability and pH insensitivity, *Nature Methods 6*(5), 351-353. doi:10.1038/NMETH.1317

50. Wall, M. A., Socolich, M., and Ranganathan, R. (2000) The structural basis for red fluorescence in the tetrameric GFP homolog DsRed, *Nature Structural Biology 7*(12), 1133-1138.

51. Verkhusha, V. V., Chudakov, D. M., Gurskaya, N. G., Lukyanov, S., and Lukyanov, K. A. (2004) Common pathway for the red chromophore formation in fluorescent proteins and chromoproteins, *Chemistry & Biology 11*(6), 845-854. doi:10.1016/j.chembiol.2004.04.007

52. Glick, B. S., Strongin, D. E., Keenan, R., Strack, R. L., and Bhattacharyya, D. (2014) Red fluorescent proteins with enhanced bacterial expression, increased brightness and reduced aggregation. US8679749 B2

53. Lam, A. J., St-Pierre, F., Gong, Y., Marshall, J. D., Cranfill, P. J., Baird, M. A., McKeown, M. R., Wiedenmann, J., Davidson, M. W., Schnitzer, M. J., Tsien, R. Y., and Lin, M. Z. (2012) Improving FRET dynamic range with bright green and red fluorescent proteins, *Nature Methods 9*(10), 1005-+. doi:10.1038/NMETH.2171

54. Li, Z., Wang, D., Zhang, Z., Bi, L., Cui, Z., Deng, J., and Zhang, X. (2014) The S28H mutation on mNeptune generates a brighter near-infrared monomeric fluorescent protein with improved quantum yield and pH-stability, *Acta Biochimica Et Biophysica Sinica 46*(9), 802-809. doi:10.1093/abbs/gmu063

55. Bowen, E. J. (1954) Fluorescence quenching in solution and in the vapour state, *Transactions of the Faraday Society 50*(2), 97-102. doi:10.1039/tf9545000097

56. Parker, C. A., and Rees, W. T. (1960) Correction of fluorescence spectra and measurement of fluorescence quantum efficiency, *Analyst 85*(1013), 587-600. doi:10.1039/an9608500587

57. Haebig, J. E. (1967) Vibrational effects in radiationless transitions in aromatic molecules, *Journal of Physical Chemistry 71*(13), 4203-&. doi:10.1021/j100872a007

58. Henderson, J. N., and Remington, S. J. (2006) The kindling fluorescent protein: A transient photoswitchable marker, *Physiology 21*, 162-170. doi:10.1152/physiol.00056.2005

59. Quillin, M. L., Anstrom, D. M., Shu, X., O'Leary, S., Kallio, K., Chudakov, D. M., and Remington, S. J. (2005) Kindling fluorescent protein from anemonia sulcata: Dark-state structure at 1.38 A resolution, *Biochemistry 44*(15), 5774-5787. doi:10.1021/bi047644u

60. Shu, X., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y., and Remington, S. J. (2006) Novel chromophores and buried charges control color in mFruits, *Biochemistry 45(32*), 9639-9647. doi: 10.1021/bi0607731

61. Megley, C. M., Dickson, L. A., Maddalo, S. L., Chandler, G. J., and Zimmer, M. (2009) Photophysics and dihedral freedom of the chromophore in yellow, blue, and green fluorescent protein, *Journal of Physical Chemistry B 113*(1), 302-308. doi:10.1021/jp806285s

62. Tolbert, L. M., Baldridge, A., Kowalik, J., and Solntsev, K. M. (2012) Collapse and recovery of green fluorescent protein chromophore emission through topological effects, *Accounts of Chemical Research 45*(2), 171-181. doi:10.1021/ar2000925

63. Strongin, D. E., Bevis, B., Khuong, N., Downing, M. E., Strack, R. L., Sundaram, K., Glick, B. S., and Keenan, R. J. (2007) Structural rearrangements near the chromophore influence the maturation speed and brightness of DsRed variants, *Protein Engineering Design & Selection 20*(11), 525-534. doi:10.1093/protein/gzm046

64. Lepecq, J. B., and Paoletti, C. (1967) A fluorescent complex between ethidium bromide and nucleic acids - physical-chemical characterization, *Journal of Molecular Biology 27*(1), 87-&. doi:10.1016/0022-2836(67)90353-1

65. Ware, W. R., and Cunningham, P. T. (1965) Lifetime and quenching of anthracene fluorescence in the vapor phase, *The Journal of Chemical Physics 43*(11), 3826-3831.

66. Wirp, C., Gusten, H., and Brauer, H. D. (1996) Fluorescence quenching of meso-substituted anthracene derivatives by molecular oxygen, *Berichte Der Bunsen-Gesellschaft-Physical Chemistry Chemical Physics 100*(7), 1217-1225.

67. Chapagain, P. P., Regmi, C. K., and Castillo, W. (2011) Fluorescent protein barrel fluctuations and oxygen diffusion pathways in mCherry, *Journal of Chemical Physics 135*(23), 235101. doi:10.1063/1.3660197

68. Akerboom, J., Rivera, J. D. V., Guilbe, M. M. R., Malave, E. C. A., Hernandez, H. H., Tian, L., Hires, S. A., Marvin, J. S., Looger, L. L., and Schreiter, E. R. (2009) Crystal structures of the GCaMP calcium sensor reveal the mechanism of fluorescence signal change and aid rational design, *Journal of Biological Chemistry 284*(10), 6455-6464. doi:10.1074/jbc.M807657200

69. Oster, G., and Nishijima, Y. (1956) Fluorescence and internal rotation - their dependence on viscosity of the medium, *Journal of the American Chemical Society 78*(8), 1581-1584. doi:10.1021/ja01589a021

70. Drexhage, K. H., and Reynolds, G. A. (1974) New dye solutions for mode-locking infrared lasers, *Optics Communications 10*(1), 18-20. doi:10.1016/0030-4018(74)90093-5

71. Reynolds, G. A., and Drexhage, K. H. (1975) New coumarin dyes with rigidized structure for flashlamp-pumped dye lasers, *Optics Communications 13*(3), 222-225. doi:10.1016/0030-4018(75)90085-1

72. Helms, V., Straatsma, T. P., and McCammon, J. A. (1999) Internal dynamics of green fluorescent protein, *Journal of Physical Chemistry B 103*(16), 3263-3269. doi:10.1021/jp983120q

73. Kredel, S., Nienhaus, K., Oswald, F., Wolff, M., Ivanchenko, S., Cymer, F., Jeromin, A., Michel, F. J., Spindler, K.-D., Heilker, R., Nienhaus, G. U., and Wiedenmann, J. (2008) Optimized and Far-Red-Emitting Variants of Fluorescent Protein eqFP611, *Chem Biol 15*, 224-233.

74. Chica, R. A., Moore, M. M., Allen, B. D., and Mayo, S. L. (2010) Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries, *Proceedings of the National Academy of Sciences of the United States of America 107*, 20257-20262.

2. Brighter Red Fluorescent Proteins by Rational Design of Triple-Decker Motif (Manuscript)

(Published online on December 23, 2015)

Antonia T. Pandelieva,[1] Miranda J. Baran,[1] Guido F. Calderini,[1] Jenna L. McCann,[1] Véronique Tremblay,[2] Sabina Sarvan,[2] James A. Davey,[1] Jean-François Couture[2,3,*] & Roberto A. Chica[1,3,*]

Adapted with permission from:

Pandelieva, A. T., Baran, M. J., Calderini, G. F., McCann, J. L., Tremblay, V., Sarvan, S., Davey, J. A., Couture, J.-F. & Chica, R. A. (2015). Brighter red fluorescent proteins by rational design of triple – decker motif. *ACS Chemical Biology,* doi:10.1021/acschembio.5b00774

[1] Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, Ontario, K1N 6N5, Canada

[2] Ottawa Institute of Systems Biology, Department of Biochemistry, Microbiology and Immunology, University of Ottawa, Ottawa, Ontario, K1H 8M5, Canada

[3] Centre for Catalysis Research and Innovation, University of Ottawa, Ottawa, Ontario, K1N 6N5, Canada

2.0. Preface and Contribution Statement

This chapter contains a manuscript describing the majority of results obtained in the context of my master's research project, which help to support the hypothesis presented in section 1.6.2. Contributions of all co-authors to the project are detailed below:

Conception: The direction of this project was determined by my supervisor Dr. Chica, and we collaborated on the design of experiments and data analysis. Dr. Jean-François Couture designed crystallography experiments.

Writing: Chapter 2 was co-written by Dr. Chica, Dr. Couture, and me.

Experimental: Undergraduate students Miranda J. Baran and Guido F. Calderini helped to prepare mRojoA mutant libraries by performing some of the individual overlap extension PCR steps and by contributing to cloning; undergraduate student Jenna L. McCann prepared by mutagenesis the mRojo-VPGV, mRojo-VYWV, mRojo-TYWV and mRojo-VYWL mutants for individual characterization; Miranda J. Baran screened the P63F library and measured the spectral properties of one individual preparation of mRojo-VFAV; Véronique Tremblay taught me to set up the crystallography experiments, and fished and screened some of the crystals of mRojo-VYGV; Savina Sarvan fished and screened the mRojo-THSL crystal; James A. Davey analyzed computationally the geometry of triple-decker motifs found in the crystal structures; Dr. Jean-François Couture solved the crystal structures of mRojo-VYGV, mRojo-VFAV, and mRojo-THSL. All other experiments for mutagenesis, screening, protein expression and purification, spectroscopic characterization, pKa measurement, and crystallization were performed by me.

2.1. Introduction

Red fluorescent proteins (RFPs) from *Anthozoa* have found widespread application in biological research. They have been used as fusion tags to track proteins within cells (*1*), as partners in FRET pairs (*2*), and as signal transducers in biosensors (*3, 4*). Because of their longer emission wavelengths (> 570 nm), RFPs are particularly suited for whole-body imaging of research model animals, as red light is less toxic to cells and can penetrate deeper into tissues (*5, 6*). For all applications of RFPs, high brightness is desirable. However, a drawback of monomeric RFPs is that they typically display lower brightness than green and yellow fluorescent proteins, and their brightness tends to decrease as their emission wavelength increases (*5*). Thus, there is still a need to engineer monomeric RFPs to maximize their brightness at longer emission wavelengths.

RFP brightness is defined as the product of the molar extinction coefficient and quantum yield (*7*). The extinction coefficient relates to the efficiency with which the chromophore can absorb light. However, reported values of molar extinction coefficients are also dependent on maturation kinetics since they are directly proportional to the number of molecules that contain a mature chromophore in the RFP population in solution (*8*). Extinction coefficients of most monomeric RFPs are in the 40-100 mM^{-1} cm^{-1} range (*5, 9*), which is similar to that of *Aequorea victoria* GFP and its variants. On the other hand, quantum yield relates to the efficiency with which the chromophore emits fluorescence, and is defined as the number of photons emitted over the number of photons absorbed. Although monomeric RFPs can have quantum yields up to 0.49 (*10*), most display quantum yields lower than 0.30, making them significantly less bright than

monomeric green and yellow fluorescent proteins, which predominantly have quantum yields in the 0.60-0.80 range (*5, 9*).

To increase the brightness of monomeric RFPs, researchers have traditionally used directed evolution methods, whereby large libraries of random mutants (10^3-10^6) are screened using high-throughput fluorescence-based assays. Although brighter monomeric RFPs have been engineered using these approaches (*11-13*), multiple rounds of mutagenesis followed by high-throughput screening were required to achieve two- to three-fold increases in brightness. Unsurprisingly, much of the observed brightness increases are due to improvements in chromophore maturation that result in higher molar extinction coefficients (*11, 12*), and only modest quantum yield increases (< 0.10) are typically obtained. In addition, mutations found by these random approaches are located all over the RFP barrel structure, making it difficult to rationalize the effect that they have on quantum yield.

In this study, we developed a rational approach to increase the quantum yield of RFPs. We hypothesized that this could be achieved by rigidification of the chromophore, which would lead to reduced excited-state decay through non-radiative means (*14, 15*). To test our hypothesis, we generated mutants of mRojoA (*16*), a previously engineered RFP that contains a π-stacked Tyr residue at position 197, directly beneath the chromophore. These mRojoA mutants contain an additional aromatic residue (His, Phe, or Tyr) at position 63, directly above the chromophore. The best mutants identified displayed over three-fold increased quantum yield and brightness relative to mRojoA. X-ray crystallography of the mutants confirmed that the increased quantum yield is caused by chromophore rigidification, as evidenced by significantly reduced B-factors. The approach developed here could be applied to other fluorescent proteins in order to increase their quantum yield and overall brightness.

2.2. Results

2.2.1. Hypothesis and Design Strategy

Our goal was to develop a rational design approach to increase the quantum yield of monomeric RFPs in order to improve their brightness. As is the case with small molecule fluorophores, fluorescent protein chromophores can lose energy via radiationless internal conversion followed by vibrational relaxation, resulting in lower quantum yields. Since this process is more probable with a decreasing energy difference between the excited and ground singlet states, the quantum yields of monomeric RFPs tend to be lower than those of blue-shifted fluorescent proteins such as GFP and YFP. Thus, we postulated that we could increase the quantum yield of monomeric RFPs by restricting the conformational degrees of freedom of their chromophore, which would reduce internal conversion due to decreased overlap in vibrational states between the ground and excited singlet states.

To test this hypothesis, we attempted to rigidify the chromophore of a monomeric RFP by sandwiching it between two aromatic amino acids. These amino acids (Tyr, His, and Phe) were selected because their aromatic side chains are bulkier than those of aliphatic amino acids and are planar, which would help restrict conformational flexibility of the chromophore via improved steric complementarity and packing. In addition, aromatic amino acids can form π-stacking interactions, which could favor chromophore planarity, a property that has been linked to increased quantum yield in RFPs (*17*).

As a starting structure to test our hypothesis, we selected the mRojoA monomeric RFP that we previously engineered (*16*). mRojoA is a five-fold mutant of the monomeric RFP mCherry containing the V16T, Q163L, V195A, I197Y, and A217C (numbering according to mCherry

structure [PDB ID: 2H5Q] (*17*)) mutations. mRojoA displays a 22-nm emission wavelength bathochromic shift relative to its parent mCherry and a quantum yield of 0.02, ten times lower than that of the parent (*16*). This lower quantum yield is mostly caused by the presence of the π-stacking Y197 residue, which reduces quantum yield to 0.03 and also red-shifts the emission wavelength by 7 nm. mRojoA was selected as the test protein for our hypothesis because it already contains an aromatic residue that is π-stacked with the chromophore and because its low quantum yield can be further optimized, providing a good test case for our rational design approach.

2.2.2. Rational Design of Quantum-Yield Enhancing Mutations

Inspection of the crystal structure of mRojoA (PDB ID: 3NEZ (*16*)) shows that a second π-stacking interaction could be introduced at position 63, which is located directly above the chromophore (Fig. 2.1). In mRojoA, a Pro residue is found at this position. As a first step towards testing our hypothesis, we mutated P63 of mRojoA to the aromatic residues His, Phe, and Tyr. As shown on Figure 2.S1A, these mutations had a detrimental effect on chromophore maturation, resulting in folded proteins that do not contain mature chromophore. Although these single aromatic substitutions do not appear to prevent folding into soluble protein, they likely hinder the backbone conformational changes required for chromophore forming residues (M66-Y67-G68) to cyclize in the first step of the chromophore maturation mechanism (*18*). Given that P63 is located directly next to the large amino acids W143 and L163 (Fig. 2.1), it is likely that the bulky side chain of aromatic residues cannot be accommodated at position 63 without causing significant steric clashes that would distort the structure of the RFP.

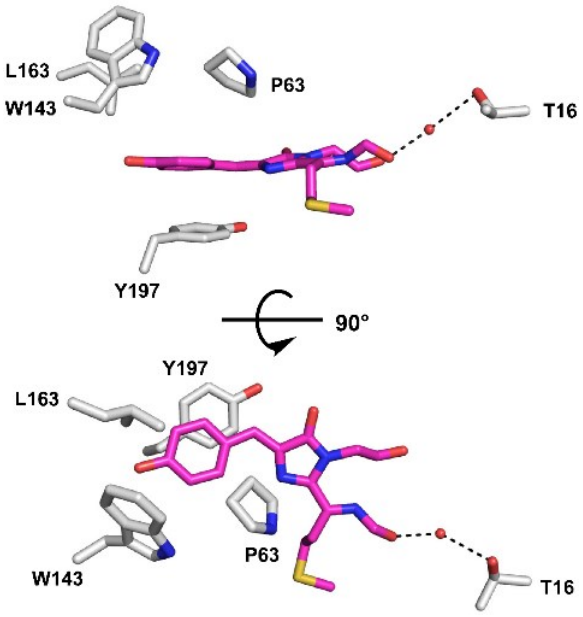

Figure 2.1. Crystal structure of mRojoA. The chromophore (magenta) and nearby residues (white) are shown in stick form. H-bonds are indicated by dashed lines.

In order to accommodate the His, Phe, and Tyr mutations at position 63, we prepared a small combinatorial library by mutating W143 to the smallest amino acids (Gly, Ala, and Ser) and L163 to the smaller amino acid Val. We also reverted the T16 residue present in mRojoA to Val, the residue found at this position in mCherry, in order to help with maturation. Indeed, we previously showed that the presence of an amino acid capable of hydrogen bonding (H-bonding) to the chromophore acylimine oxygen atom at this position is detrimental to efficient red chromophore maturation, presumably by restricting the conformational freedom of the peptide backbone required for the second oxidation reaction that generates the acylimine group to take place (*19*). In addition to the mutations listed above, we also included the wild-type residue at each position in our libraries. Thus, we generated three individual libraries containing 16 mutants each, for a total of 48 mutants.

2.2.3. Spectral Characterization of Mutants

Following screening, we identified the brightest mutant from each combinatorial library. These quadruple mutants, mRojo-VHSV, mRojo-VFAV, and mRojo-VYGV (Table 2.1), are named based on the amino-acid identity at designed positions. For example, the mRojo-VHSV mutant contains Val, His, Ser, and Val residue at positions 16, 63, 143, and 163, respectively. At pH 7.4, all mutants displayed emission wavelength hypsochromic shifts of 11-17 nm (Fig. 2.2) as well as approximately two-fold increased brightness relative to mRojoA. The observed brightness increases are mostly due to improved quantum yield (1.4- to 2.1-fold increases) and not extinction coefficient (<1.4-fold increases). This result indicates that the combinations of four mutations have a more important effect on quantum yield than on chromophore maturation and extinction coefficient, in agreement with our hypothesis. Of the three mutants, mRojo-VYGV had the highest quantum yield (0.050 ± 0.006), a 2.1-fold increase compared to mRojoA, and mRojo-VHSV was the brightest (236 % of mRojoA). We also measured the extinction coefficients and quantum yields of these mutants at pH 10, since the pH-profiles of these mutants showed two pKa values, one at approximately 4 and another at approximately 9 (Table 2.S1, Fig. 2.S2), similar to what is observed for mCherry (*17*). The identity of the second chromophore ionizable group causing the high pKa has not yet been determined conclusively. However, Shu et al. showed by x-ray crystallography that the chromophore imidazolinone N2 atom forms a hydrogen bonding interaction with the side chain of Glu215 in mCherry and mStrawberry (*17*). Loss of the proton exchanged between these groups at high pH has been proposed to result in hypsochromic shifts in absorbance wavelength, similar to those observed for our mRojoA mutants (Table 2.1). Interestingly, the quantum yields of the three quadruple mutants but not of mRojoA increased significantly at higher pH. At pH 10, mRojo-VYGV has the highest quantum yield (0.092 ± 0.005), which is over three-fold greater than

that of mRojoA (0.027 ± 0.005). Although the overall brightness of these mutants is 2.5- to 3.1-fold greater than that of mRojoA at pH 10, molar extinction coefficients did not vary significantly at higher pH, demonstrating that quantum yield increases are the main contributor to improved brightness.

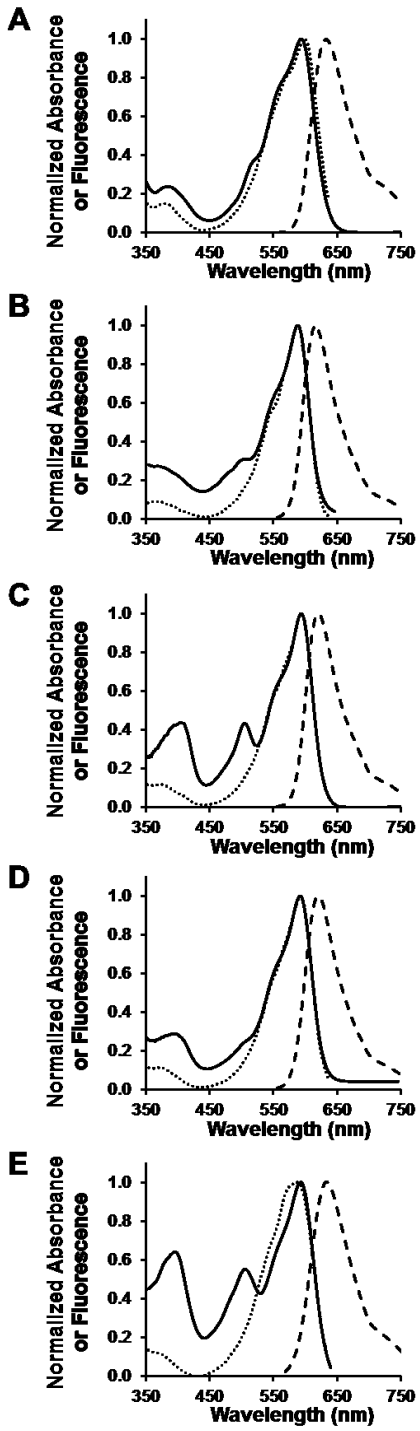

Figure 2.2. Fluorescence and absorption spectra of mRojoA and selected variants. Absorbance (solid), excitation (dots), and emission (dashes) are shown for mRojoA (A), mRojo-VYGV (B), mRojo-VFAV (C), mRojo-VHSV (D), and mRojo-THSL (E).

Table 2.1. Spectral properties of various RFPs.

Protein	Mutations				pH 7.4		pH 10		Extinction Coefficient (pH 7.4) ($mM^{-1}cm^{-1}$)	Extinction Coefficient (pH 10) ($mM^{-1}cm^{-1}$)	Quantum Yield (pH 7.4)	Quantum Yield (pH 10)	Relative Brightness pH 7.4 (% mRojoA)	Relative Brightness pH 10 (% mRojoA)
	16	63	143	163	λ_{ex} (nm)	λ_{em} (nm)	λ_{ex} (nm)	λ_{em} (nm)						
mRojoA	T	P	W	L	596±1	633±2	593±3	632±3	83±13	82±9	0.024±0.003	0.027±0.005	100±27	100±29
mRojo-THSL	-	H	S	-	596±2	627±1	575±2	625±1	55.3±0.2	56±4	0.027±0.003	0.042±0.005	76±17	106±26
mRojo-THSV	-	H	S	V	594±1	626±3	579±1	627±2	59±3	54±6	0.030±0.001	0.054±0.002	91±18	132±32
mRojo-VHSV	V	H	S	V	592±1	622±2	574±4	616±1	117±13	93±22	0.040±0.004	0.075±0.001	236±58	311±98
mRojo-TFAL	-	F	A	-	592±1	622±1	577±1	620±3	68±12	74±11	0.0215±0.0003	0.036±0.001	74±20	120±30
mRojo-TFAV	-	F	A	V	592±1	623±1	578±1	622±1	84±5	86±5	0.029±0.002	0.052±0.001	125±26	201±43
mRojo-VFAV	V	F	A	V	593±3	621±1	577±3	619±2	94±19	86±18	0.034±0.002	0.075±0.004	164±46	290±86
mRojo-TYGL	-	Y	G	-	590±2	624±1	572±3	621±2	64±5	60±3	0.031±0.002	0.043±0.004	99±21	117±27
mRojo-TYGV	-	Y	G	V	592±2	624±2	575±5	622±4	75.2±0.3	73±10	0.041±0.001	0.067±0.003	156±30	218±55
mRojo-VYGV	V	Y	G	V	591±2	616±1	571±3	608±4	72±10	60±17	0.050±0.006	0.092±0.005	183±49	247±89
mRojo-VYGL	V	Y	G	-	590±3	617±1	570±1	609±1	92±30	75±13	0.033±0.002	0.061±0.004	154±58	205±57
mRojo-VYWV	V	Y	-	V	-	-	-	-	-	-	<0.001	<0.001	-	-
mRojo-TYWV	-	Y	-	V	-	-	-	-	-	-	<0.001	<0.001	-	-
mRojo-VYWL	V	Y	-	-	-	-	-	-	-	-	<0.001	<0.001	-	-
mRojo-VPGV	V	-	G	V	589±4	616±1	564±5	608±1	77±15	73±16	0.016±0.003	0.06±0.01	61±19	186±64

All experiments were performed in triplicate using RFPs from a minimum of three independent protein preparations.

2.2.4. Contribution of Individual Mutations to observed Quantum Yield Increases

To evaluate the contribution of individual mutations to increased quantum yield, we prepared a series of variants of the highest quantum yield mutant mRojo-VYGV, where each mutation with the exception of P63Y was reverted back to the parent amino acid. As was the case with the P63Y single mutant, all double and triple mutants containing the P63Y mutation in combination with the parental W143 residue did not form mature chromophore (Fig. 2.S1B), suggesting that close proximity of these two bulky aromatic residues results in steric clashes that are detrimental to maturation. As expected, chromophore maturation is restored when the compensating W143G mutation is introduced next to Y63 (Table 2.1). Furthermore, all triple mutants containing the P63Y and W143G mutations (mRojo-TYGV and mRojo-VYGL) display an increased quantum yield relative to mRojoA (Table 2.1) but lower than that of quadruple mutant mRojo-VYGV, at both pH 7.4 and 10. Similar results are obtained with mRojoA mutants containing the P63H or P63F mutation, which require the W143S or W143A compensating mutation, respectively, to form mature chromophore. Interestingly, the T16V mutation that we introduced to improve chromophore maturation does not significantly improve molar extinction coefficient in these mutants, with the exception of the mRojo-VHSV mutant that has an approximately two-fold higher extinction coefficient than mRojo-THSV.

Because introduction of the P63Y substitution alone in mRojoA prevents chromophore maturation (Fig. 2.S1A), we prepared the mRojo-VPGV triple mutant in which the Y63 substitution is reverted back to the parental Pro residue in order to evaluate the effect on quantum yield of the P63Y substitution. This triple mutant displays a quantum yield decrease of approximately 0.03 relative to quadruple mutant mRojo-VYGV, at both pH 7.4 and 10. These data confirm that the Tyr residue at position 63 is the main contributor to increased quantum yield in

these mutants, presumably by helping to rigidify the chromophore via improved packing or π-stacking interactions.

2.2.5. Structural Characterization

While spectral characterization and mutagenesis data supports our hypothesis that sandwiching the chromophore between two aromatic residues results in increased quantum yield, it does not provide evidence of chromophore rigidification through π-stacking interactions. To evaluate whether the aromatic residue at position 63 was stacked with the chromophore, we solved the crystal structure of two of the three brightest mutants (mRojo-VFAV and mRojo-VYGV) as well as of the mRojo-THSL double mutant, since this variant contains the necessary aromatic residue for π-stacking but does not display an increased quantum yield relative to mRojoA. Data collection and refinement statistics are reported in Table 2.S2. We also attempted to crystallize mRojo-VHSV but were unable to obtain high-quality crystals. The three solved structures were obtained from crystals that were prepared at pH 5.5 (mRojo-THSL and mRojo-VYGV) and 7.4 (mRojo-VFAV). In both the mRojo-THSL and mRojo-VFAV mutants, no appreciable omit map density (contoured at 1.5σ) corresponding to the *p*-hydroxybenzylidene moiety of the chromophore was observed (Fig. 2.3A and B), suggesting that the chromophore is either dynamic or can adopt multiple conformations. Additionally, no density was observed for the phenolate group of Y197 in both of these mutants and for the isopropyl group of the L163 side chain of mRojo-THSL. This is in contrast with the structure of the highest quantum yield mutant mRojo-VYGV that shows clearly observable electronic density for all atoms of the chromophore and Y197 (Fig. 2.3C), suggesting that the chromophore and neighboring residues are more rigid. Interestingly, the average chromophore B-factors of mRojo-THSL and mRojo-VFAV are

significantly higher than the average B-factors of each protein (Table 2.S2), suggesting that the chromophore is indeed more dynamic than the average residue of the protein. This is not the case for mRojo-VYGV, whose chromophore B-factor is nearly-identical to that of the rest of the protein. Taken together, these observations strongly support that the chromophore of mRojo-VYGV is more rigid than that of mRojo-THSV and mRojo-VFAV, in agreement with the observed higher quantum yield of mRojo-VYGV.

Analysis of the crystal structure of mRojo-VYGV shows that the chromophore is in the *cis* conformation, which positions it between the side chains of the two Tyr residues found at positions 63 and 197 (Fig. 2.4). As expected, Y197 forms a face-to-face parallel π-stacking interaction with the chromophore phenolate moiety with an average centroid-to-centroid distance and angle between the two normals to the benzene rings of 3.7 ± 0.1 Å and 0 ± 3°, respectively. These values are similar to those found in mRojoA (*16*) and the yellow fluorescent protein Citrine (*20*), which have centroid-to-centroid distances of 3.9 Å and 3.6 Å, and angles of 8° and 6.1°, respectively. In the case of Y63, we observe a perpendicular edge-to-face interaction between its phenolate group and that of the chromophore. In this interaction, the centroid-to-centroid distance and the angle between the two normals to the benzene rings are 4.8 ± 0.1 Å and 71° ± 4°, respectively. These metrics are consistent with a weak edge-to-face π-stacking interaction between two Tyr in methanol that was calculated *in silico* (*21*) and with distances and angles between Tyr – Tyr interactions found in protein structures deposited in the Protein Data Bank (*22*).

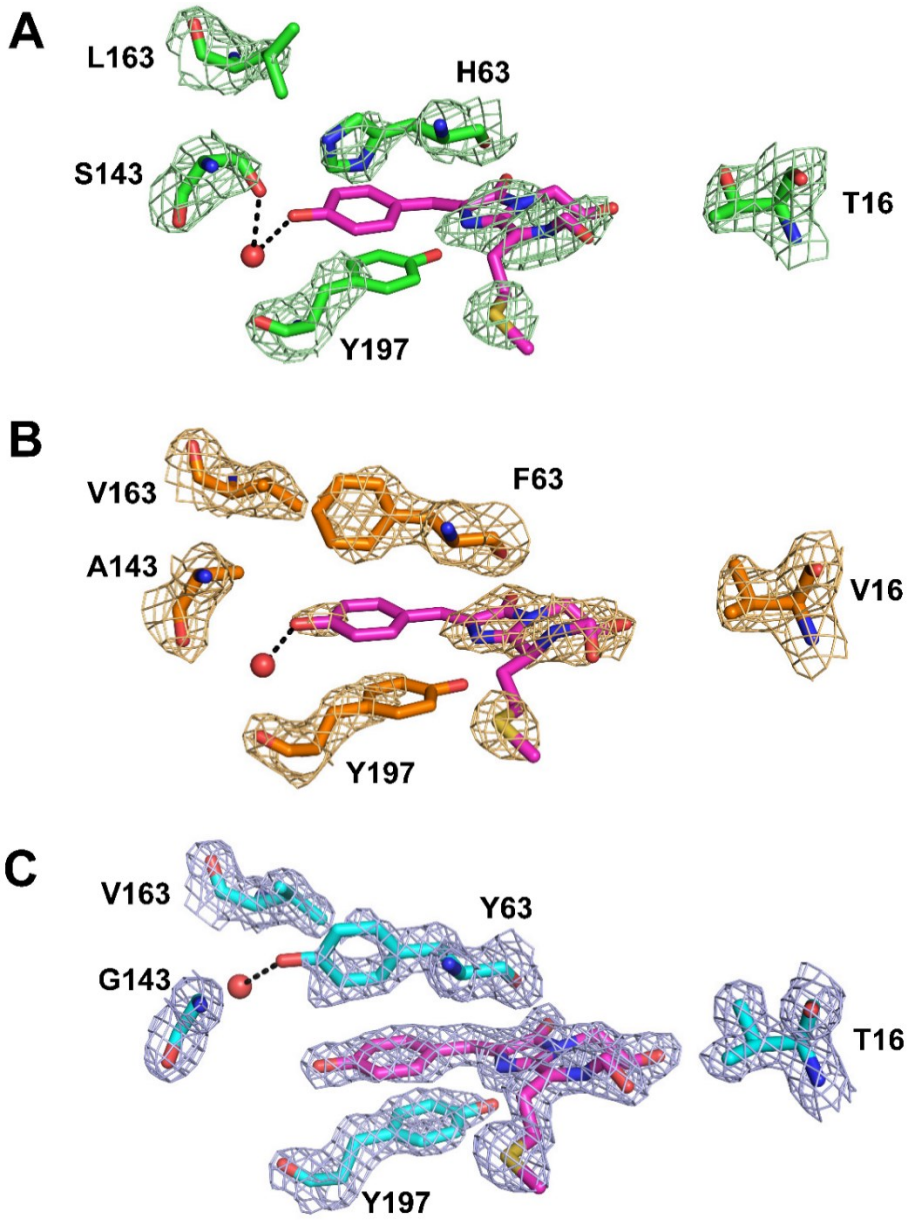

Figure 2.3. Crystal structures of mRojoA mutants. Electron densities of the chromophore (magenta) and surrounding residues of the crystal structures of mRojo-THSL (A, green), mRojo-VYGV (B, blue), or mRojo-VFAV (C, orange). 2mFo-DFc maps were contoured at 1.5σ.

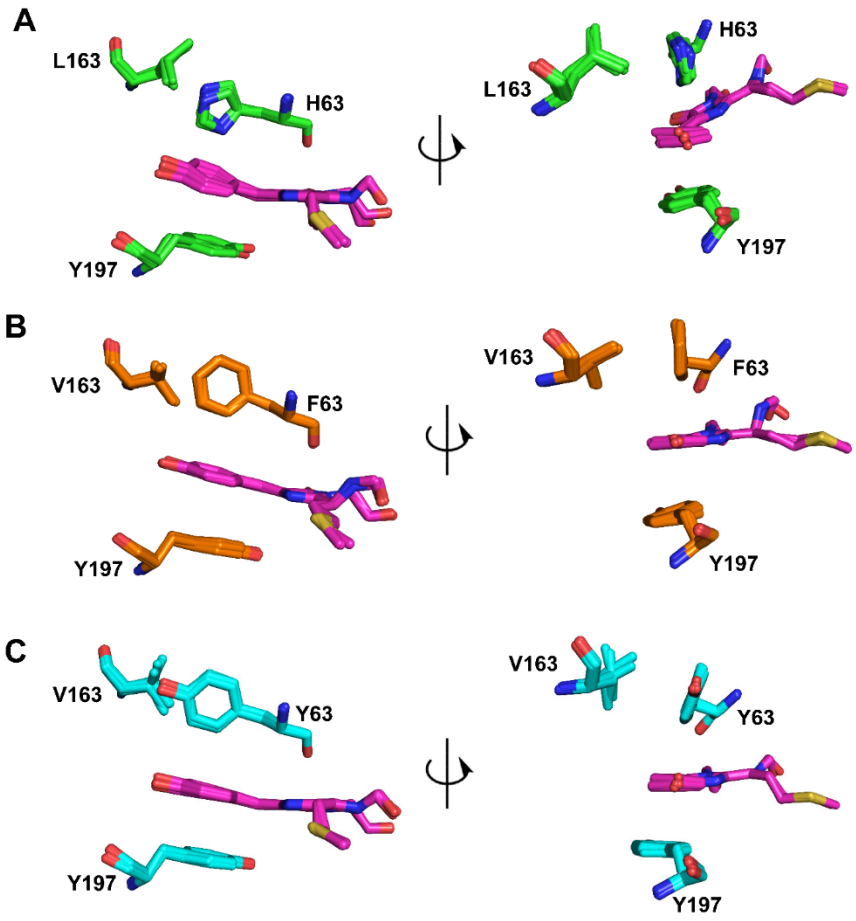

Figure 2.4. Triple-decker motif in mRojoA mutants. The chromophore (magenta) and surrounding residues of mRojo-THSL (A, green), mRojo-VYGV (B, blue), or mRojo-VFAV (C, orange) are shown in each structure's four crystallographic monomers.

2.3. Discussion

2.3.1. Comparison to Directed Evolution

Using rational design, we were able to increase the quantum yield and thereby brightness of monomeric RFPs by over three-fold. The greatest absolute increases in quantum yield that we achieved were for the mRojo-VYGV mutant, whose quantum yield was increased with respect to mRojoA by approximately 0.03 and 0.07 at pH 7.4 and 10, respectively. These quantum yield increases are comparable to what other researchers have achieved by directed evolution using random mutagenesis. Most notably, mStrawberry, a member of the mFruits family of monomeric RFPs that also includes mCherry, was created from mRFP1 by screening millions of mutants during multiple rounds of directed evolution, leading to a quantum yield increase of 0.04 and a 2.1-fold increase in brightness, which was mostly due to the protein's improved extinction coefficient (*11*). In a more recent example, mKate2 was engineered from mKate using saturation mutagenesis followed by two rounds of directed evolution. This process led to the identification of five mutations that cause a 1.4-fold increase in quantum yield and a two-fold increase in extinction coefficient, primarily by stabilizing the *cis* conformation of the chromophore (*12*). The absolute increase in quantum yield achieved in mKate2 relative to mKate has been measured by several groups to be in the 0.06-0.12 range (*12, 23, 24*). In contrast to these examples, our rational design approach involved the screening of fewer than 50 mutants, a library size that is several orders of magnitude smaller than those previously used to achieve comparable quantum yield increases in other RFPs. More importantly, the designed mutations did not significantly increase extinction coefficient in all but one case (mRojo-VSHV at pH 7.4), demonstrating that it is possible to increase the brightness of RFPs by increasing the quantum yield alone via the introduction of a

few mutations in the vicinity of the chromophore. While the mutants developed here are not as bright as other available monomeric RFPs, they represent promising templates for further directed evolution to improve their brightness.

2.3.2. Triple-Decker Motif

The structure of mRojo-VYGV reported here shows that the chromophore is sandwiched between two Tyr residues in a triple-decker motif. To our knowledge, triple-decker motifs of aromatic groups involving the chromophore have not yet been observed in natural or engineered fluorescent proteins. However, single π-stacking interactions with the chromophore exist in a number of natural (*25*) and engineered (*16, 26-28*) fluorescent proteins. In most cases, the π-stacking residue is found beneath the chromophore at a position analogous to mRojoA position 197. At this position, several aromatic amino acids can be found (Tyr, Phe, or His), and they form parallel π-stacking interactions with the chromophore in the crystal structures. One of the few examples where a π-stacking residue is found above the chromophore at a position analogous to mRojoA position 63 is cpYGFP, a yellow-green fluorescent protein from the marine copepod *Chiridius poppeii* (PDB ID: 2DD7 *(25)*). In cpYGFP, H52 makes a parallel stacking interaction with the chromophore, and this interaction has been observed to cause an emission wavelength bathochromic shift. A comparison of the cpYGFP structure with that of mRojo-THSL, which also contains a His residue above the chromophore, shows significantly different conformations of the aromatic amino-acid side chain. In mRojo-TSHL, the H63 side chain has χ_1 and χ_2 dihedral angles of -49.1° and -91.7°, respectively, corresponding to *gauche*$^-$ conformations. However, the side chain χ_1 and χ_2 dihedral angles of H52 in cpGFP are -67.6° and -171°, respectively. The *trans* conformation of the χ_2 dihedral in cpYGFP H52 positions the imidazole group in a parallel

orientation to the chromophore (angle between the two normals to the benzene rings of 12.8 ± 0.1 °). Although no density for the chromophore *p*-hydroxybenzylidene group was observed for mRojo-THSL, modeling of the chromophore structure shows that the phenolate plane would be perpendicular to that of the imidazole ring of H63 (Fig. 2.4A), with an angle between the two normals of 87 ± 12°.

In all cases where the chromophore of a fluorescent protein is stacked against an aromatic residue, the π-stacking interaction leads to a bathochromic shift in emission wavelength. A computational study that investigated the effect on emission wavelength of a triple-decker π-stacking motif between two Tyr and the GFP chromophore predicted an emission bathochromic shift of 25 nm relative to wild-type GFP (*29*). A similar bathochromic shift is not observed for the mRojoA mutants described here, which instead display hypsochromic shifts in fluorescence emission. This discrepancy may be due to differences in the three-dimensional arrangement of aromatic groups in both triple-decker motifs, which would result in significantly altered molecular orbital interactions, and thus, emission wavelength. Indeed, the theoretical triple-decker motif modelled by Grigorenko *et al.* (*29*) is composed of three aromatic rings, two from Tyr residues and one from the chromophore, that form parallel π-stacking interactions. This is in contrast with the triple-decker motif observed in the crystal structure of mRojo-VYGV, which is composed of a parallel face-to-face π-stacking interaction between the chromophore phenolate moiety and Y197, and a perpendicular edge-to-face interaction between the chromophore and Y63. It is unlikely that the latter interaction is a π-stacking interaction, given that the introduction of the P63Y mutation into the mRojo-VPGV mutant does not result in an emission wavelength shift, suggesting negligible molecular orbital overlap. Absence of a parallel π-stacking interaction between the chromophore and Y63 even in the presence of two aromatic groups in close proximity

demonstrates that the parallel triple-decker configuration is not favorable in the context of the mutants tested herein. Further engineering will be required to incorporate a triple-decker made of exclusively parallel π-stacking interactions in RFPs.

2.3.3. Emission Wavelength Hypsochromic Shifts

The rational design approach tested here aimed to increase the quantum yield of RFPs by rigidifying the chromophore, which would decrease radiationless decay by limiting internal conversion followed by vibrational relaxation. Radiationless decay can also be reduced by increasing the energy difference between ground and excited singlet states, which would decrease the probability of internal conversion and vibrational relaxation. Accordingly, mutations that create emission wavelength hypsochromic shifts in fluorescent proteins also tend to enhance quantum yield (*2, 11, 16*). It is thus possible that the increased energy difference between ground and excited singlet states in our blue-shifted mRojoA mutants contributes to the observed quantum yield increases. However, the mRojo-VPGV mutant has an emission wavelength identical to that of mRojo-VYGV at pH 7.4, but a quantum yield that is over three-fold lower than that of the brighter mutant (Table 2.1). This example demonstrates that it is not sufficient to blue-shift the emission wavelength of an RFP in order to increase its quantum yield, and that chromophore rigidification is a viable path to achieve higher quantum yields in RFPs. Indeed, it is known that oligomeric RFPs tend to have significantly higher quantum yields than monomeric variants (*7, 11, 30*), which has been proposed to result from stabilization of the chromophore in a conformation that favors bright fluorescence upon binding of additional subunits. This hypothesis is supported by observations from Alford and Campbell who showed that binding of another subunit to form a heterodimer leads to a quantum yield increase of approximately 0.05 in a dim monomeric RFP,

but does not shift its emission wavelength (*31*). The authors proposed that the quantum yield increase upon heterodimer formation results from a change to the chromophore environment that stabilizes it in a coplanar conformation by restricting its conformational degrees of freedom.

2.4. Conclusions

We developed a rational design approach to increase the quantum yield of a monomeric RFP. To do so, we sandwiched the chromophore between two aromatic amino acids, improving packing and restricting its conformational degrees of freedom. The approach developed here could potentially be applied to additional RFPs displaying already higher quantum yield, or to fluorescent proteins that emit light of other colors. Two potential candidates include YFP and cpYGFP from Copepod, both of which already contain an aromatic amino acid that is stacked against the chromophore. The mRojoA mutants developed here can also serve a starting points in directed evolution to further increase their quantum yield and overall brightness via the introduction of additional mutations. Finally, our mRojo-VYGV mutant could serve as input template to computational design in order to identify mutations that will stabilize the Y63 residue in a parallel orientation to the chromophore, giving rise to a triple-decker motif that may result in significantly red-shifted emission wavelength, a desirable property of RFPs.

2.5. Materials and Methods

2.5.1 Materials

All reagents used were of the highest available purity. Restriction enzymes and DNA-modifying enzymes were from New England Biolabs. Synthetic oligonucleotides were obtained from Integrated DNA Technologies and Eurofins MWG Operon, and Ni-NTA agarose resin was

obtained from Promega. Lysozyme was purchased from AMRESCO. All aqueous solutions were prepared using water purified with a Barnstead Nanopure Diamond system.

2.5.2. Mutagenesis

The mRojoA gene cloned into the pET11a plasmid (Novagen) via *Nde*I/*Bam*HI. mRojoA (*16*) served as a template for mutagenesis by overlap extension PCR (*32*) using Vent DNA polymerase (New England Biolabs). Briefly, external primers were used in combination with sets of complementary pairs of degenerate oligonucleotides (Table 5.1.1) containing the desired mutations in individual PCR reactions. The resulting overlapping fragments were gel-purified (Omega Bio-Tek) and recombined by overlap extension PCR. The resulting amplicons were digested with *Nde*I/*Bam*HI, gel-purified, and ligated into pET11a expression vector with T4 ligase (New England Biolabs). All constructs were verified by sequencing the entire reading frame.

2.5.3. Protein expression and purification for spectral characterization

Proteins were expressed in 500-mL cultures (Luria-Bertani Broth supplemented with 100 μg/mL ampicillin) of *E. coli* BL21-Gold(DE3) cells transformed with pET11a containing the mutated genes. Cultures were grown at 37 °C with shaking until an optical density of 0.6 at 600 nm was reached. At this point, protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and cells were incubated overnight at 16 °C with shaking. Following incubation, cells were harvested by centrifugation and lysed with an EmulsiFlex-B15 cell disruptor (Avestin). Proteins were purified by immobilized metal affinity chromatography according to the manufacturer's protocol. The eluted fractions were desalted by gel filtration using EconoPAC 10DG columns (Bio-Rad) into phosphate-buffered saline (PBS) at pH 7.4.

2.5.4. Spectroscopic characterization

Protein concentrations were quantified using the alkali denaturation method (*33*). Briefly, RFPs were alkali-denatured with an equal volume of 2 M NaOH. It is known that the alkali-denatured RFP chromophore converts to a GFP-like one, with extinction coefficient 44,000 M^{-1} cm^{-1} at 452 nm under these conditions. Absorption, emission, and excitation spectra were recorded in phosphate-buffered saline (PBS) (pH 7.4) and Britton-Robinson buffer (pH 10) (*34*) with a Tecan Infinite M1000 plate reader. Path lengths for each well were calculated ratiometrically using the difference in absorbance of PBS at 900 and 998 nm. Molar extinction coefficients at pH 7.4 and pH 10 were calculated using the dynamic difference method described by Kredel *et al.* (*8*) in order to account for the presence of molecules containing a GFP-like green chromophore in the RFP population in solution. Briefly, RFPs were denatured slowly by diluting them 60-fold in Britton-Robinson buffer (pH 11.4). During alkali denaturation, native green chromophore species denatures faster than native red species, giving rise to an isosbestic point in the absorbance spectrum (*8*). Following the apparition of this point, the absorbance increase corresponding to formation of alkali denatured chromophore and the absorbance decrease corresponding to degradation of native red chromophore were measured at various time intervals. With these absorbance values, a ratio of extinction coefficients for alkali denatured and native red chromophores was calculated. Relating the peak intensities of green and red chromophores in the absorption spectrum of each mature RFP sample at pH 7.4 or 10 to that of alkali denatured RFP at pH 13 (extinction coefficient = 44,000 M^{-1} cm^{-1} at λ = 452 nm) allowed for the calculation of an extinction coefficient for the native red chromophore only. For determination of quantum yields, the integrated fluorescence intensity of mutants of interest was compared with that of equally

absorbing samples of mCherry and mRaspberry (quantum yields 0.22 (*11*) and 0.15 (*35*), respectively) with excitation at 535 nm.

2.5.5. pKa measurements

pH titrations were performed using a series of Britton-Robinson buffers from pH 2 to 12. Proteins were diluted into these buffers to a final concentration of 0.5–1.0 μM. Fluorescence emission spectra (λ_{ex} = 540 nm) were measured at each pH value using a Tecan Infinite M1000 plate reader. Fluorescence intensities corresponding to the area under the curve were plotted as a function of pH. The following biphasic equation was fitted to the data using a least-squares calculation in order to calculate the pKa values for each protein:

$$F = \frac{H}{1 + 10^{(pKa_1 - pH)*(n_1)}} + \frac{1 - H}{1 + 10^{(pKa_2 - pH)*(n_2)}} \quad [5]$$

where *F* is the calculated fluorescence intensity at a given pH, *H* is the fluorescence intensity of the plateau at the left end of the curve, and *n* is the Hill coefficient.

2.5.6. Protein expression and purification for crystallography

Proteins purified as described above were allowed to mature in PBS (pH 7.4) in the dark at 4 °C for five days. The proteins were then concentrated through centrifugation (Pall Microsep Advance Centrifugal Device 10K), resuspended in 20 mM Tris-HCl pH 7.4 supplemented with 100 mM NaCl, and further purified through size exclusion chromatography (Bio-Rad ENrich SEC 650).

2.5.7. Crystallography

Dark blue crystals of mRojo-THSL were grown in 1 µL × 1 µL sitting drops with a precipitant solution of 125 mM Bis-Tris buffer, pH 5.5, and 21 % (w/v) polyethylene glycol 3350. Dark purple crystals of mRojo-VYGV were grown in 1 µL × 1 µL sitting drops with a precipitant solution of 125 mM Bis-Tris buffer, pH 5.5, 200 mM $MgCl_2$, and 25 % (w/v) polyethylene glycol 3350. Reddish purple crystals of mRojo-VFAV were grown in 1 µL × 1 µL sitting drops with a precipitant solution of 0.1 M HEPES pH 7.4 and 27 % (w/v) polyethylene glycol 3350. Large multi-crystalline chunks were prodded with a nylon loop to break off smaller shards for isolation and collection of diffraction data. Crystals were harvested in mother liquor supplemented with 10 % (mRojo-THSL and mRojo-VYGV) or 30 % 2-methyl-2,4-pentanediol and frozen in liquid nitrogen.

Data sets were collected at the 17-ID-D beamline of LS-CAT (Advanced Photon Source, Argonne National Laboratories). Data sets were indexed with XDS and scaled using aimless (*36, 37*). A molecular replacement solution was found using Phaser (*38*) and mRojoA structure as a search model (PDB ID: 3NEZ (*16*)). The model was completed using iterative rounds of refinement and model building using Buster and Coot, respectively (*39, 40*). Quality of the models was assessed using Molprobity (*41*).

2.6. Supplemental Information

Supplemental Information includes two figures and two tables.

Table 2.S1. pKa values for various RFPs.

Protein	pKa_1	n_1*	pKa_2	n_2*
mRojoA	**4.6**	**0.6**	**9.6**	**0.7**
mRojo-TFAL	6.1	0.4	9.9	1.8
mRojo-TFAV	4.3	0.5	9.3	1.1
mRojo-VFAV	4.4	0.5	9.2	1.4
mRojo-THSL	3.7	0.4	9.0	0.9
mRojo-THSV	3.2	0.7	9.3	0.9
mRojo-VHSV	4.3	0.6	9.5	1.1
mRojo-TYGL	4.5	0.5	8.9	1.1
mRojo-TYGV	3.9	0.6	9.3	0.8
mRojo-VYGV	4.3	0.7	9.1	1.1
mRojo-VYGL	4.1	0.5	9.1	1.0
mRojo-VPGV	4.2	0.6	9.5	1.0

* Hill coefficients obtained by fitting experimental data to the Henderson-Hasselbach equation using a least squares approach.

Table 2.S2. Crystallographic data and refinement statistics for mRojoA mutants

	mRojo-THSV	mRojo-VYGV	mRojo-VFAV
Data collection			
Space group	$P2_1$	$P2_1$	$P2_1$
Cell dimensions			
a, *b*, *c* (Å)	61.7, 91.5, 84.6	59.0, 108.1, 74.0	61.4, 94.7, 78.9
α, β, γ (°)	90.0, 108.7, 90.0	90.0, 105.2, 90.0	90.0, 98.5, 90.0
Resolution (Å)*	33.5 - 2.2 (2.3 – 2.2)	71.4 – 1.8 (1.9 -1.8)	78.07 - 2.0 (2.1 - 2.0)
R_{sym}	5.1 (43.1)	5.5 (52.1)	5.8 (55.5)
$I / \sigma I$	17.9 (3.6)	15.9 (3.1)	12.7 (2.4)
Completeness (%)	99.8 (99.9)	96.4 (95.9)	99.7 (100.0)
Redundancy	4.2 (4.2)	3.4 (3.5)	4.2 (4.1)
Molecules per asymmetric unit	4	4	4
Refinement			
Resolution (Å)	32.8 – 2.2	71.4 – 1.8	30.0 – 2.0
No. reflections	42608	85121	55078
R_{work} / R_{free}	18.2 /23.2	18.5/21.7	16.7/20.9
No. atoms			
Protein	6935	7022	6914
Chromophore	92	92	92
Water	216	551	425
B-factors (Å^2)			
Protein	58.3	22.3	32.9
Chromophore	90.5	22.4	47.4
Water	60.2	31.8	50.2
R.M.S. deviations			
Bond lengths (Å)	0.010	0.010	0.010
Bond angles (°)	1.20	1.16	1.14
Molprobity score	1.56	1.58	1.58

Each dataset was collected on a single crystal

* Highest resolution shell is shown in parentheses

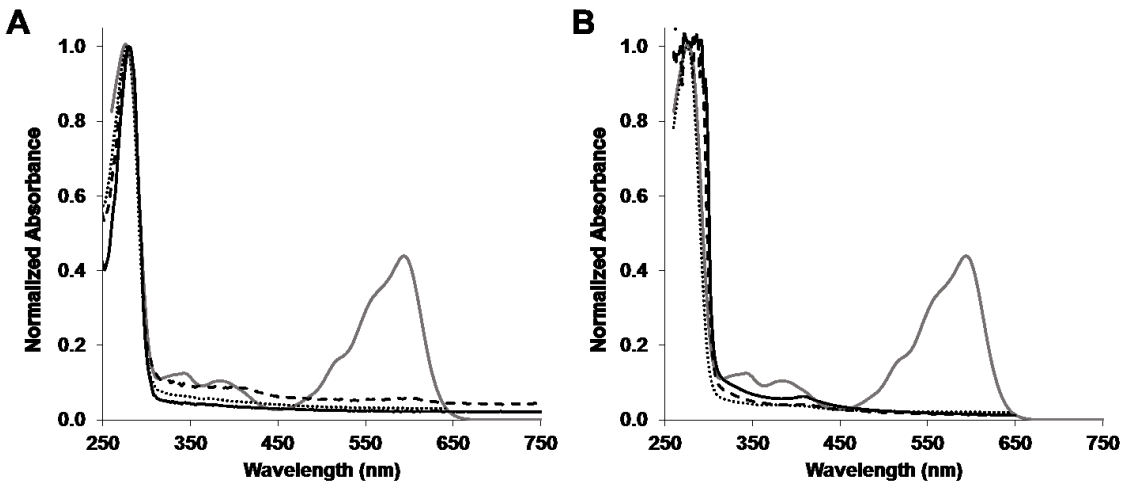

Figure 2.S1. mRojoA mutants containing aromatic residues at positions 63 and 143 do not form mature chromophore. **A)** Absorbance spectra of mRojoA (grey, solid), mRojoA-P63H (black, dashed), mRojoA-P63F (black, dotted), and mRojoA-P63F (black, solid), normalized to absorbance at 280 nm. **B)** Absorbance spectra of mRojoA (grey, solid), mRojo-VYWL (black, dashed), mRojo-TYWV (black, dotted), and mRojo-VYWV (black, solid), normalized to absorbance at 280 nm.

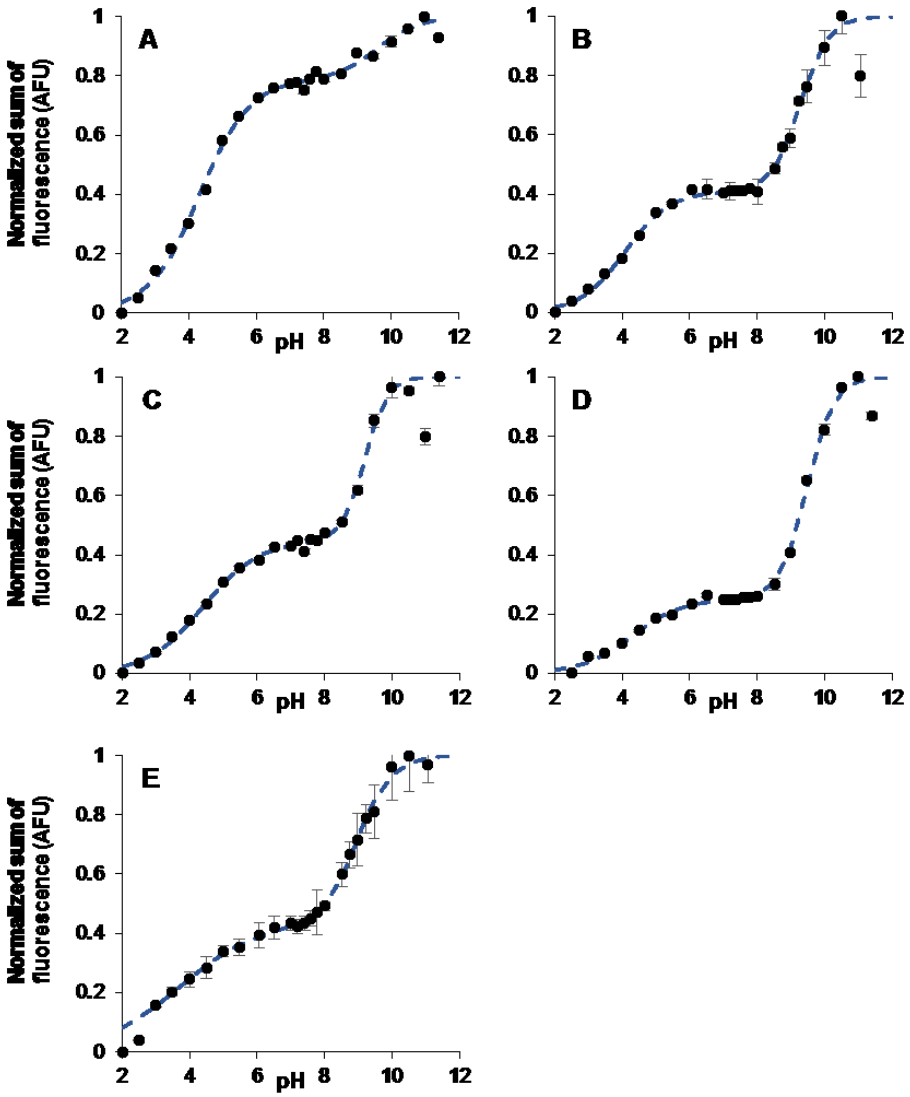

Figure 2.S2. pH profiles of selected variants. mRojoA (A), mRojo-VYGV (B), mRojo-VFAV (C), mRojo-VHSV (D), and mRojo-THSL integrated fluorescence is plotted as a function of pH. All measurements were performed in triplicate.

Accession numbers

The atomic coordinates and structure factors of mRojo-THSL, mRojo-VFAV, and mRojo-VYGV have been deposited in the Protein Data Bank (PDB) under accession codes 5H87, 5H88, and 5H89, respectively.

Author contributions

A.T.P., M.J.B., G.F.C., J.L.M., V.T., and S.S. performed the experiments. A.T.P., J.A.D., J.-F.C., and R.A.C. analyzed data. A.T.P., J.-F.C., and R.A.C. designed experiments and wrote the paper.

Acknowledgments

R.A.C. acknowledges grants from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Ontario Research Fund, and the Canada Foundation for Innovation. J-F.C. acknowledges an Early Researcher Award from the Ontario Ministry of Economic Development & Innovation, a Canada Research Chair, and grants from the Canadian Institutes for Health Research and NSERC. J.A.D. is the recipient of an Ontario Graduate Scholarship, and M.J.B. is the recipient of a NSERC Undergraduate Student Research Award.

2.8. References

1. Young, C. L., Raden, D. L., Caplan, J. L., Czymmek, K. J., and Robinson, A. S. (2012) Cassette series designed for live-cell imaging of proteins and high-resolution techniques in yeast, *Yeast 29*, 119-136.

2. Lam, A. J., St-Pierre, F., Gong, Y., Marshall, J. D., Cranfill, P. J., Baird, M. A., McKeown, M. R., Wiedenmann, J., Davidson, M. W., Schnitzer, M. J., Tsien, R. Y., and Lin, M. Z. (2012) Improving FRET dynamic range with bright green and red fluorescent proteins, *Nat Methods 9*, 1005-1012.

3. Zhao, Y., Araki, S., Wu, J., Teramoto, T., Chang, Y. F., Nakano, M., Abdelfattah, A. S., Fujiwara, M., Ishihara, T., Nagai, T., and Campbell, R. E. (2011) An expanded palette of genetically encoded Ca(2)(+) indicators, *Science 333*, 1888-1891.

4. Tantama, M., Hung, Y. P., and Yellen, G. (2011) Imaging intracellular pH in live cells with a genetically encoded red fluorescent protein sensor, *J Am Chem Soc 133*, 10034-10037.

5. Chudakov, D. M., Matz, M. V., Lukyanov, S., and Lukyanov, K. A. (2010) Fluorescent proteins and their applications in imaging living cells and tissues, *Physiol Rev 90*, 1103-1163.

6. Konig, K. (2000) Multiphoton microscopy in life sciences, *J Microsc 200*, 83-104.

7. Matz, M. V., Fradkov, A. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L., and Lukyanov, S. A. (1999) Fluorescent proteins from nonbioluminescent Anthozoa species, *Nat Biotechnol 17*, 969-973.

8. Kredel, S., Nienhaus, K., Oswald, F., Wolff, M., Ivanchenko, S., Cymer, F., Jeromin, A., Michel, F. J., Spindler, K.-D., Heilker, R., Nienhaus, G. U., and Wiedenmann, J. (2008) Optimized and Far-Red-Emitting Variants of Fluorescent Protein eqFP611, *Chem Biol 15*, 224-233.

9. Nienhaus, K., and Nienhaus, G. U. (2014) Fluorescent proteins for live-cell imaging with super-resolution, *Chem Soc Rev 43*, 1088-1106.

10. Shaner, N. C., Lin, M. Z., McKeown, M. R., Steinbach, P. A., Hazelwood, K. L., Davidson, M. W., and Tsien, R. Y. (2008) Improving the photostability of bright monomeric orange and red fluorescent proteins, *Nat Methods 5*, 545-551.

11. Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E., and Tsien, R. Y. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein, *Nat Biotechnol 22*, 1567-1572.

12. Shcherbo, D., Murphy, C. S., Ermakova, G. V., Solovieva, E. A., Chepurnykh, T. V., Shcheglov, A. S., Verkhusha, V. V., Pletnev, V. Z., Hazelwood, K. L., Roche, P. M.,

Lukyanov, S., Zaraisky, A. G., Davidson, M. W., and Chudakov, D. M. (2009) Far-red fluorescent tags for protein imaging in living tissues, *Biochem J 418*, 567-574.

13. Glick, B. S., Strongin, D. E., Keenan, R., Strack, R. L., and Bhattacharyya, D. (2014) Red fluorescent proteins with enhanced bacterial expression, increased brightness and reduced aggregation. US8679749 B2

14. Megley, C. M., Dickson, L. A., Maddalo, S. L., Chandler, G. J., and Zimmer, M. (2009) Photophysics and dihedral freedom of the chromophore in yellow, blue, and green fluorescent protein, *J Phys Chem B 113*, 302-308.

15. Laurent, A. D., Mironov, V. A., Chapagain, P. P., Nemukhin, A. V., and Krylov, A. I. (2012) Exploring structural and optical properties of fluorescent proteins by squeezing: modeling high-pressure effects on the mStrawberry and mCherry red fluorescent proteins, *J Phys Chem B 116*, 12426-12440.

16. Chica, R. A., Moore, M. M., Allen, B. D., and Mayo, S. L. (2010) Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries, *Proc Natl Acad Sci U S A 107*, 20257-20262.

17. Shu, X., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y., and Remington, S. J. (2006) Novel chromophores and buried charges control color in mFruits, *Biochemistry 45*, 9639-9647.

18. Strack, R. L., Strongin, D. E., Mets, L., Glick, B. S., and Keenan, R. J. (2010) Chromophore formation in DsRed occurs by a branched pathway, *J Am Chem Soc 132*, 8496-8505.

19. Moore, M. M., Oteng-Pabi, S. K., Pandelieva, A. T., Mayo, S. L., and Chica, R. A. (2012) Recovery of red fluorescent protein chromophore maturation deficiency through rational design, *PloS One 7*, e52463.

20. Barstow, B., Ando, N., Kim, C. U., and Gruner, S. M. (2008) Alteration of citrine structure by hydrostatic pressure explains the accompanying spectral shift, *Proc Natl Acad Sci U S A 105*, 13362-13366.

21. Chelli, R., Gervasio, F. L., Procacci, P., and Schettino, V. (2002) Stacking and T-shape competition in aromatic-aromatic amino acid interactions, *J Am Chem Soc 124*, 6133-6143.

22. McGaughey, G. B., Gagne, M., and Rappe, A. K. (1998) pi-Stacking interactions. Alive and well in proteins, *J Biol Chem 273*, 15458-15463.

23. Lin, M. Z., McKeown, M. R., Ng, H. L., Aguilera, T. A., Shaner, N. C., Campbell, R. E., Adams, S. R., Gross, L. A., Ma, W., Alber, T., and Tsien, R. Y. (2009) Autofluorescent proteins with excitation in the optical window for intravital imaging in mammals, *Chem Biol 16*, 1169-1179.

24. Strack, R. L., Hein, B., Bhattacharyya, D., Hell, S. W., Keenan, R. J., and Glick, B. S. (2009) A rapidly maturing far-red derivative of DsRed-Express2 for whole-cell labeling, *Biochemistry 48*, 8279-8281.

25. Suto, K., Masuda, H., Takenaka, Y., Tsuji, F. I., and Mizuno, H. (2009) Structural basis for red-shifted emission of a GFP-like protein from the marine copepod Chiridius poppei, *Genes to cells : devoted to molecular & cellular mechanisms 14*, 727-737.

26. Shkrob, M. A., Yanushevich, Y. G., Chudakov, D. M., Gurskaya, N. G., Labas, Y. A., Poponov, S. Y., Mudrik, N. N., Lukyanov, S., and Lukyanov, K. A. (2005) Far-red fluorescent proteins evolved from a blue chromoprotein from Actinia equina, *Biochem J 392*, 649-654.

27. Wachter, R. M., Elsliger, M. A., Kallio, K., Hanson, G. T., and Remington, S. J. (1998) Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein, *Structure 6*, 1267-1277.

28. Akerboom, J., Carreras Calderon, N., Tian, L., Wabnig, S., Prigge, M., Tolo, J., Gordus, A., Orger, M. B., Severi, K. E., Macklin, J. J., Patel, R., Pulver, S. R., Wardill, T. J., Fischer, E., Schuler, C., Chen, T. W., Sarkisyan, K. S., Marvin, J. S., Bargmann, C. I., Kim, D. S., Kugler, S., Lagnado, L., Hegemann, P., Gottschalk, A., Schreiter, E. R., and Looger, L. L. (2013) Genetically encoded calcium indicators for multi-color neural activity imaging and combination with optogenetics, *Front Mol Neurosci 6*, 2.

29. Grigorenko, B. L., Nemukhin, A. V., Polyakov, I. V., and Krylov, A. I. (2013) Triple-Decker Motif for Red-Shifted Fluorescent Protein Mutants, *J Phys Chem Lett 4*, 1743-1747.

30. Campbell, R. E., Tour, O., Palmer, A. E., Steinbach, P. A., Baird, G. S., Zacharias, D. A., and Tsien, R. Y. (2002) A monomeric red fluorescent protein, *Proc Natl Acad Sci U S A 99*, 7877-7882.

31. Alford, S. C., Abdelfattah, A. S., Ding, Y., and Campbell, R. E. (2012) A fluorogenic red fluorescent protein heterodimer, *Chem Biol 19*, 353-360.

32. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction, *Gene 77*, 51-59.

33. Gross, L. A., Baird, G. S., Hoffman, R. C., Baldridge, K. K., and Tsien, R. Y. (2000) The structure of the chromophore within DsRed, a red fluorescent protein from coral, *Proc Natl Acad Sci U S A 97*, 11990-11995.

34. Britton, H. T. S., and Robinson, R. A. (1931) CXCVIII.-Universal buffer solutions and the dissociation constant of veronal, *J Chem Soc (Resumed)*, 1456-1462.

35. Wang, L., Jackson, W. C., Steinbach, P. A., and Tsien, R. Y. (2004) Evolution of new nonantibody proteins via iterative somatic hypermutation, *Proc Natl Acad Sci U S A 101*, 16745-16749.

36. Kabsch, W. (2010) Xds, *Acta Crystallogr D Biol Crystallogr 66*, 125-132.

37. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution?, *Acta Crystallogr D Biol Crystallogr 69*, 1204-1214.

38. Zwart, P. H., Afonine, P. V., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., McKee, E., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., Storoni, L. C., Terwilliger, T. C., and Adams, P. D. (2008) Automated structure solution with the PHENIX suite, *Methods Mol Biol 426*, 419-435.

39. Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S. M., and Bricogne, G. (2004) Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT, *Acta Crystallogr D Biol Crystallogr 60*, 2210-2221.

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

41. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr D Biol Crystallogr 66*, 12-21.

3. Incorporation of Quantum–Yield–Enhancing Mutations into mCherry

3.1. Introduction

As shown in the previous chapter, we were able to increase the quantum yield and brightness of mRojoA by over 3-fold using rational design. While the quantum yield of mRojoA was only 0.024 ± 0.003, this protein provided a useful system to detect increases in brightness that may have been indistinguishable had a brighter parent protein been chosen for mutagenesis and screening. However, the significant improvements in quantum yield achieved in the mRojoA mutants were insufficient to create RFPs that were bright enough to effectively employ in live-cell imaging. In order to create a novel useful monomeric RFP using our approach, a parent protein with a higher quantum yield was necessary. We chose mCherry as a template to create such a mutant, since it is the parent protein from which mRojoA was engineered (*1*). We reasoned that the core of mCherry was sufficiently similar to the core of mRojoA to allow us to forego screening of a new library for the incorporation of aromatic residues at position 63 (Figure 3.1). Instead, we chose to introduce the P63Y, W143G, and Q163V mutations from mRojo-VYGV into the mCherry template, since mRojo-VYGV is the highest quantum yield mutant that we found in the mRojoA libraries (0.050 ± 0.006 at pH 7.4). Introduction of the V16 residue was not required as this residue is already present at this position in mCherry. In this chapter, we will describe the spectral characteristics of the mCherry-VYGV mutant.

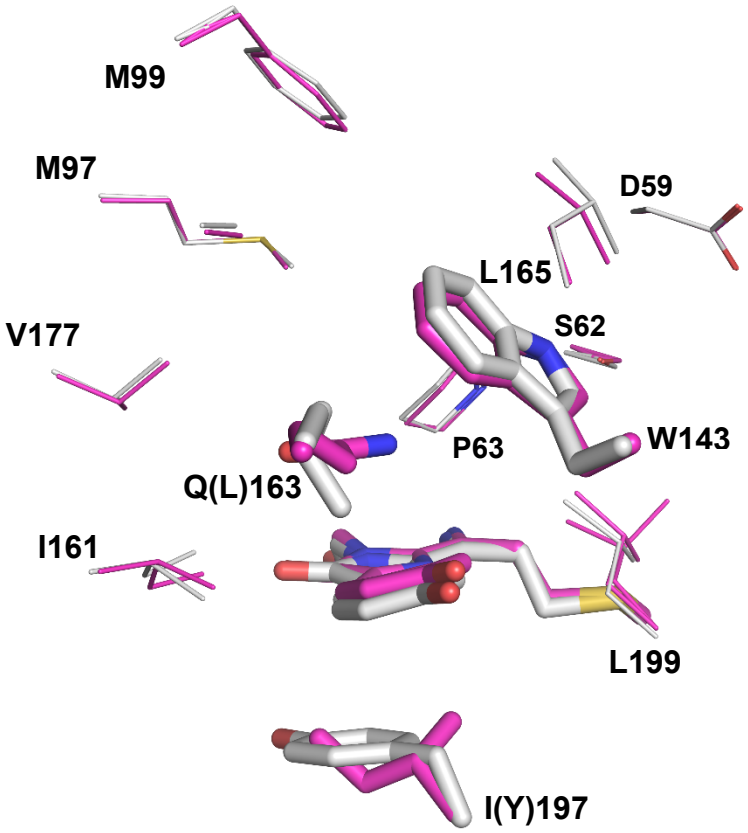

Figure 3.1. Overlay of core residues of mCherry (PDB ID: 2H5Q, magenta) and mRojoA (PDB ID: 3NEZ, white). Where the residue is not the same between the two proteins, the mRojoA residue is labelled in parentheses. Positions 143 and 163 (designed in mRojoA mutants), the chromophore, and position 197 (aromatic in mRojoA) are represented as sticks. Residues within 5 Å of position 143 or 163 and facing towards the designed positions are represented as lines. Position 16 is not pictured, since the V16 residue of mCherry did not need to be mutated to create mCherry-VYGV.

3.2. Results and Discussion

We successfully generated the mCherry-VYGV mutant (see Section 2.5.2 - Mutagenesis) using SOE mutagenesis (*2*) and the primers employed for mutagenesis of mRojoA (Table 5.1.1), with one exception. Since the DNA sequences of the *mCherry* and *mRojoA* genes differ in regions surrounding position 143, this required us to use primers specific to the *mCherry* sequence (Table 5.1.1). Following DNA sequencing of the complete reading frame, mCherry-VYGV was expressed, purified, and characterized in triplicate using mCherry and mRaspberry as standards for quantum yield determination.

The absorption spectrum of mCherry-VYGV (Figure 3.2B) contains a shoulder at approximately 500 nm that corresponds to the anionic GFP-like green chromophore, indicating that mCherry-VYGV matures less efficiently than mCherry (Figure 1.2) (*3*). A similar increase in the proportion of green to red chromophore was seen in the mRojoA mutants (Figure 2.2). Nevertheless, the ratio of red to green absorption peaks indicates that the majority of molecules in the RFP population in solution contain red chromophores, assuming similar extinction coefficients for each species. The mCherry-VYGV mutant does not have a significant shift in emission wavelength compared to mCherry (Table 3.1). However, there is an approximately 7-nm increase in the Stokes shift of this protein due to a decrease in excitation wavelength to 581 nm. In contrast, all excitation wavelengths of the mRojoA mutants remained within error to that of mRojoA, while the emission wavelength decreased by 10–15 nm, resulting in smaller Stokes shifts.

Table 3.1. Spectral properties of mCherry and mCherry-VYGV.

Protein	Mutations				pH 7.4		pH 10		Extinction Coefficient (pH 7.4) ($mM^{-1}cm^{-1}$)	Extinction Coefficient (pH 10) ($mM^{-1}cm^{-1}$)	Quantum Yield (pH 7.4)	Quantum Yield (pH 10)	Relative Brightness pH 7.4 (% mCherry)	Relative Brightness pH 10 (% mCherry)
	16	63	143	163	λ_{ex} (nm)	λ_{em} (nm)	λ_{ex} (nm)	λ_{em} (nm)						
mCherry	V	P	W	L	588 ± 1	609 ± 2	578 ± 3	603 ± 3	110 ± 13	90 ± 9	0.224 ± 0.004	0.351 ± 0.007	100 ± 34	100 ± 17
mCherry-VYGV	V	Y	G	V	581 ± 4	608 ± 1	558 ± 5	595 ± 1	112 ± 15	79 ± 16	0.19 ± 0.02	0.29 ± 0.07	85 ± 24	73 ± 20

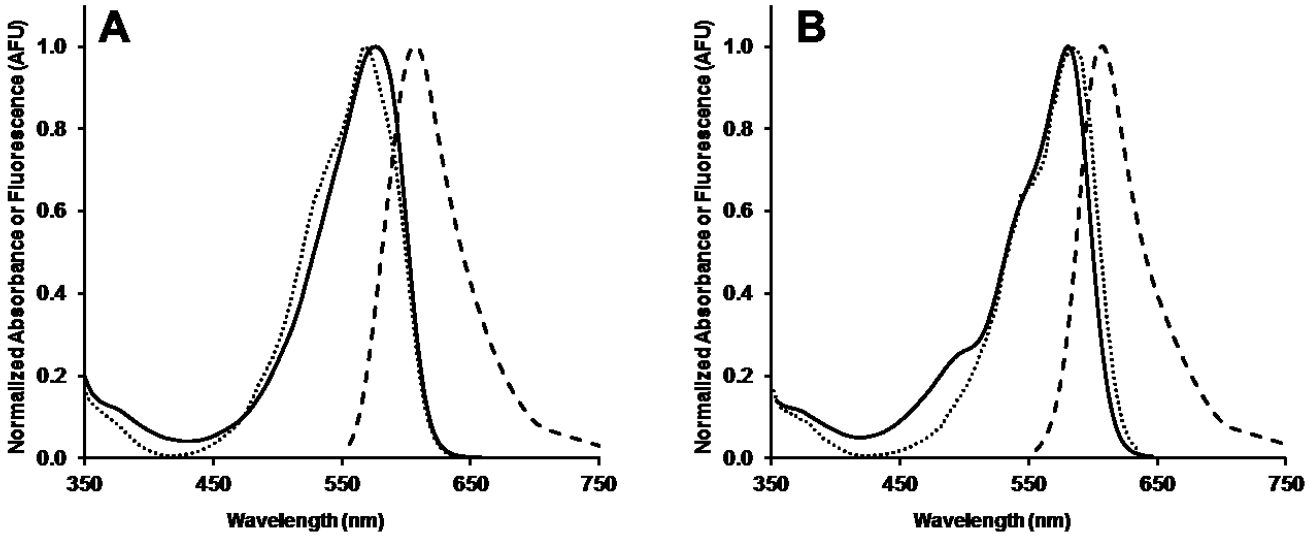

Figure 3.2. Fluorescence and absorption spectra of mCherry (A) and mCherry-VYGV (B). Normalized absorbance (solid), excitation (dots), and emission (dashes) at pH 7.4 are plotted.

The extinction coefficient of mCherry-VYGV remained the same as that of mCherry (Table 3.1) while the quantum yield decreased slightly to 0.19 ± 0.02 (quantum yield of mCherry is 0.224 ± 0.004 at pH 7.4), contrary to our objective. This resulted in a calculated relative brightness that was within error with that of the parent protein. Similar results were observed at pH 10. Similarly to the mRojoA mutants (Figure 2.S2, Table 2.S1), mCherry and mCherry-VYGV display a second pKa of approximately 9 (Figure 3.3, Table 3.2). This has previously been reported for mCherry, with a pKa value of 10.29 ± 0.04 (*4*). The nature of the species responsible for the higher pKa has not yet been conclusively identified. However, several studies (*4*-*7*) implicate the imidazolinone nitrogen atom in GFP - like chromophores as a second ionizable group. Shu *et al* have observed through crystallographic analysis a hydrogen bonding interaction of Glu215 with this group at higher pH in the mFruits family, which is thought to result in a blue shift similar to that observed in the mutants in Table 2.1 and Table 3.1 (*4*). The measured pKa values for mCherry and mCherry-VYGV differ by 0.03 – 0.04 pH units.

Our data show that insertion of the P63Y, W143G, and Q163V mutations into mCherry does not result in a mutant of higher quantum yield or brightness, unlike what is observed in mRojoA (Section 2). In order to understand the different effects that these mutations have on mCherry and mRojoA, we attempted to solve the crystal structure of mCherry-VYGV. Unfortunately, despite growing, fishing, and screening several crystals (Figure 3.4), none of them diffracted sufficiently to obtain useful data.

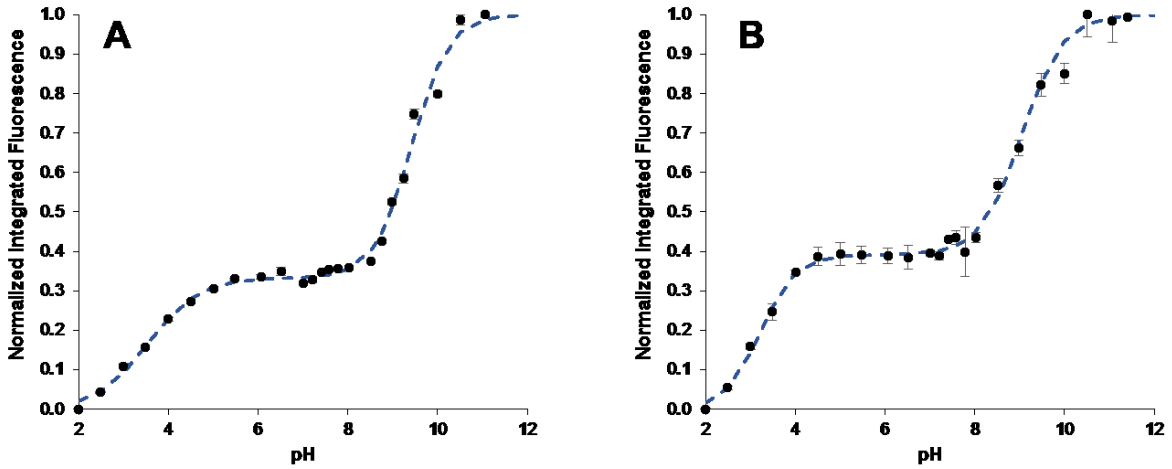

Figure 3.3. pH profiles of mCherry (A) and mCherry-VYGV (B). Integrated fluorescence is plotted as a function of pH. All measurements were performed in triplicate.

Table 3.2. pKa values for mCherry and mCherry-VYGV

Protein	pKa_1	n_1*	pKa_2	n_2*
mCherry	3.6	0.75	9.4	1.0
mCherry-VYGV	3.2	1.1	9.1	0.94

* Hill coefficients obtained by fitting experimental data to the Henderson-Hasselbach equation using a least squares approach.

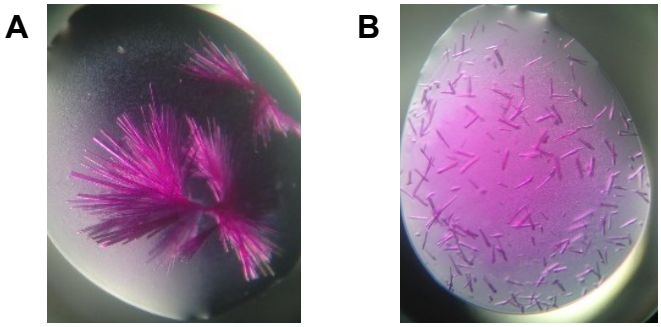

Figure 3.4. Photographs of mCherry-VYGV protein crystals. Crystals photographed in 2-µl sitting drops three weeks after setup of optimization plates. Branching needles (A) were predominantly observed. Some small (< 0.2-mm) single and branching crystals were also found (B). More detailed screening information is found in Table 5.5.1 and Figure 5.5.6.

Interestingly, the absence of an emission wavelength red shift in mCherry-VYGV suggests that Y63 is unlikely to be π-stacked with the chromophore phenolate group (*1*). The lack of a quantum yield increase in mCherry-VYGV relative to mCherry may be due to the fact that the mCherry chromophore is already as rigid as it can be, and that introduction of the P63Y, W143G, and Q163V mutations cannot rigidify the chromophore further (*8*). Another possibility is that the achievement of a significant increase in the rigidity of the chromophore could have failed because of the absence of Y197. This residue is present below the chromophore in mRojoA-VYGV (Figure 2.3, Figure 3.1). In mCherry-VYGV, I197 is significantly smaller and incapable of π - stacking to the phenolate of the chromophore. This difference in size and complementarity removes part of the triple decker motif that may be necessary for an effective restriction of the conformational freedom of the chromophore.

3.3. Conclusion

A triple mutant of mCherry was created that incorporates the mutations found in mRojo-VYGV. The P63Y, W143G, L163V mutations did not significantly affect the excitation and emission wavelengths or the pKa constants of the mutant at physiological pH. Unfortunately, they also did not result in an increase in quantum yield. Therefore, alternate approaches should be undertaken to engineer a brighter variant of mCherry. An attempt to rigidify the protein through mutations distant from the chromophore, which is currently being undertaken by other members of the Chica lab, could prove successful in increasing the quantum yield of this protein.

3.4. References

1. Chica, R. A., Moore, M. M., Allen, B. D., and Mayo, S. L. (2010) Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries, *Proceedings of the National Academy of Sciences of the United States of America 107*, 20257-20262.

2. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction, *Gene 77*, 51-59.

3. Moore, M. M., Oteng-Pabi, S. K., Pandelieva, A. T., Mayo, S. L., and Chica, R. A. (2012) Recovery of red fluorescent protein chromophore maturation deficiency through rational design, *Public Library of Science One 7*, e52463.

4. Shu, X., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y., and Remington, S. J. (2006) Novel chromophores and buried charges control color in mFruits, *Biochemistry 45*, 9639-9647.

5. Voityuk, A.A., Kummer, A.D., Michel-Beyerle, M.E., Rösch, N. (2001) Absorption spectra of the GFP chromophore in solution: comparison of theoretical and experimental results, *Chemical Physics 269*, 83-91.

6. Yan, W. and Xie, D. (2009) Protonation states and conformational flexibility of the red fluorescent protein chromophore, *Journal of Theoretical and Computational Chemistry 8*(6), 1117-1129.

7. Mironov, V.A., Bravaya, K.B., and Nemukhin, A.V. (2015) Role of zwitterions in kindling fluorescent protein photochemistry, *Journal of Physical Chemistry 119*, 2467-2474.

8. Laurent, A. D., Mironov, V. A., Chapagain, P. P., Nemukhin, A. V., and Krylov, A. I. (2012) Exploring structural and optical properties of fluorescent proteins by squeezing: Modeling high-pressure effects on the mStrawberry and mCherry red fluorescent proteins, *Journal of Physical Chemistry B 116*(41), 12426-12440. doi:10.1021/jp3060944

4. Conclusions and Perspectives

4.1. Summary of Project Findings

Engineering brighter RFPs is necessary to develop a functional and diverse palette of proteins with fluorescence intensities that are equivalent to one another, regardless of their emission wavelength. Due to an intrinsically smaller energy gap between the ground and excited states of the RFP chromophore, these proteins are more susceptible to non-radiative decay pathways of internal conversion coupled to vibrational relaxation, than their blue-shifted counterparts (Figure 1.4, Figure 1.5). The brightest monomeric red fluorescent protein to date (*1*) is 53% as bright as the brightest monomeric green fluorescent protein (*2*), highlighting the need to continue developing brighter RFPs.

The goal of our project was to develop a rational design approach to engineering a higher quantum yield RFP. We chose to target chromophore dynamics, since studies in small molecules (*3, 4)* and other fluorescent proteins (*5, 6*) have demonstrated the link between quantum yield and fluorophore rigidity. As a case study, we sought to increase the quantum yield of mRojoA, an RFP containing a π-stacking interaction between the chromophore and Y197. To do so, we designed small, combinatorial mutant libraries to accommodate aromatic amino acids (H, Y, F) at position 63, which is located on the other side of the chromophore (Figure 1.8, Table 1.1). These aromatic amino acids were chosen because of their larger size and rigid conformations, and the fact that they are capable of forming π-stacking interactions with the chromophore. The potential π-stacking interaction could help to limit the chromophore's conformational freedom, and to maintain its planarity. In addition to mutations needed to accommodate an aromatic amino acid at position 63,

the T16V mutation was introduced to improve maturation of the chromophore to the red species (*7*). This resulted in the formation of three individual libraries containing 16 mutants each.

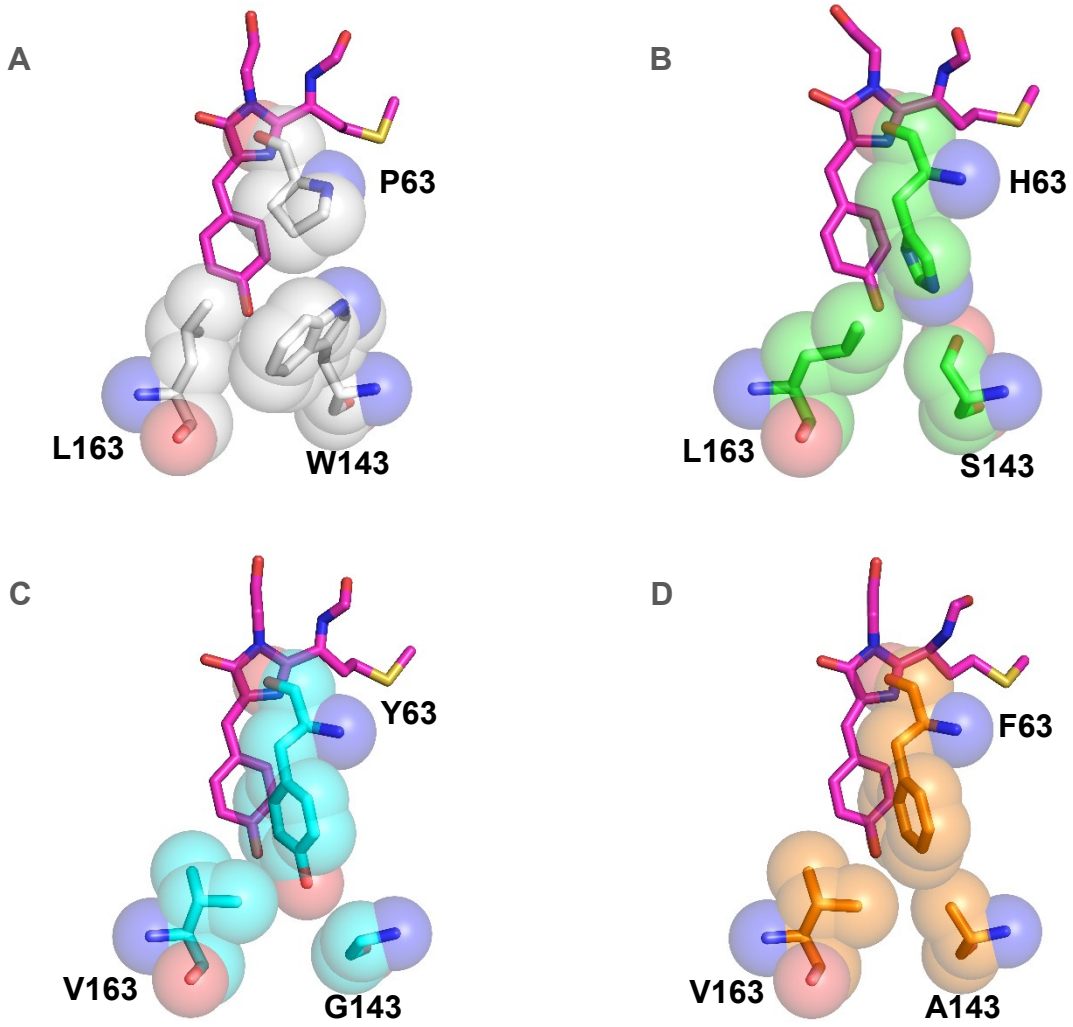

Figure 4.1.1. Crystal structures of mRojoA (A) (PDB ID: 3NEZ), mRojo-THSL (B), mRojo-VYGV (C), and mRojo-VFAV (D) showing complementary amino acids at design positions.

When the libraries were screened, combinations of complementary mutations at positions 63 and 143 were found: P63Y/W143G, P63F/W143A, and P63H/W143S (Figure 4.1.1). These combinations paired the amino acids with each other in an optimal configuration for their size and

polarity, and resulted in the brightest proteins within their respective library. The double mutants containing these mutations were all fluorescent, but only mRojo-TYGL had a significantly higher quantum yield than mRojoA. Triple mutants containing these mutations and a valine at positions 16 or 163 saw further increases in quantum yield to form significantly brighter variants, and the quadruple mutants mRojo-VYGV, mRojo-VFAV, and mRojo-VHSV were the brightest in their respective libraries, leading to an up to 2.1-fold increase in quantum yield or 2.4-fold increase in brightness at physiological pH.

We attempted to solve the crystal structures of the quadruple mutants and of the dimer mRojo-THSL double mutant in order to elucidate the relationship between the increased quantum yields and chromophore rigidity. We did not obtain quality crystals for mRojo-VHSV, but the other three mutant structures were solved at similar resolutions to each other. For the brightest mutant, mRojo-VYGV, strong electronic density around the chromophore and its surrounding residues allowed for the resolution of the structure in its entirety. B-factor analysis demonstrated that the highest quantum yield mutant (mRojo-VYGV) has the most rigid chromophore while the lowest quantum yield mutant (mRojo-THSL) has the least rigid chromophore compared to the B-factors of the whole proteins (Figure 4.1.2). Therefore, the chromophore was successfully rigidified in order to increase the quantum yield of mRojo-VYGV and, to some extent, mRojo-VFAV.

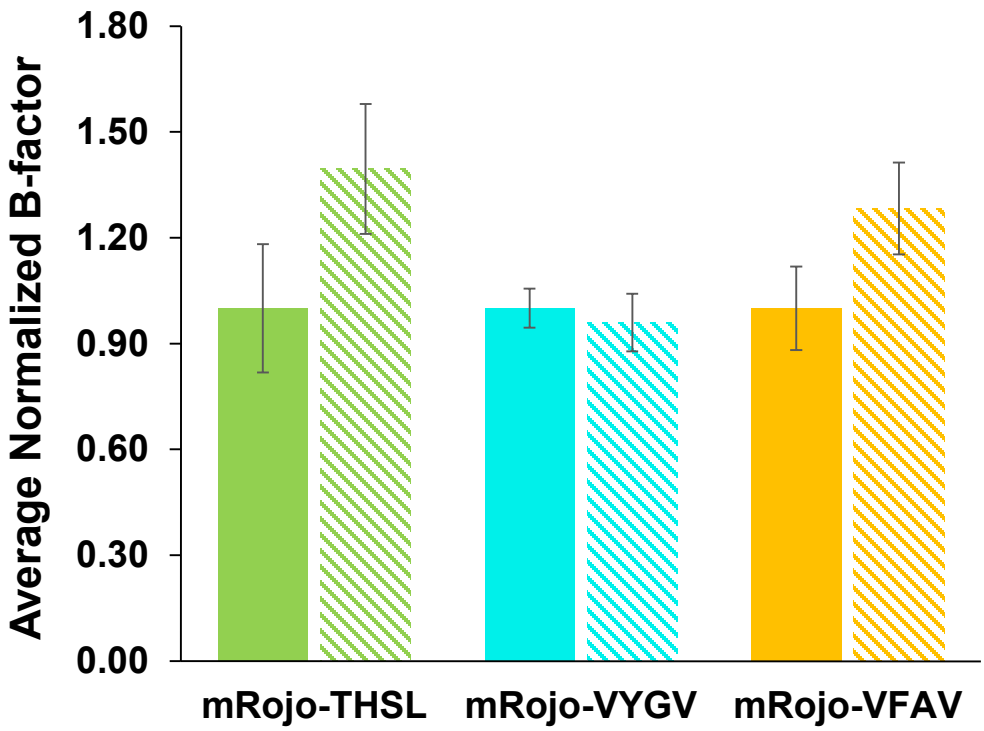

Figure 4.1.2. Normalized B-factors of mRojo-THSL (green), mRojo-VYGV (cyan), and mRojo-VFAV (orange), averaged across each protein's four crystallographic monomers. The B-factors of the entire protein (solid) and the chromophore (stripes), normalized to the whole protein B-factor, are plotted for each crystallized mutant of mRojoA.

4.2. Impact of this Work on Red Fluorescent Protein Engineering

In the past, directed evolution was often used independently or in conjunction with semi-random approaches in order to increase fluorescence intensity or recover it altogether (*8, 9, 10, 11*). Unfortunately, the random nature of directed evolution calls for the screening of thousands to millions of mutants in order to find beneficial mutations. Complementary mutations, such as the ones that were found at positions 63 and 143 using the rationally designed libraries in this project, have a lower probability of being found using directed evolution than a single mutation. Furthermore, the method presented in this work resulted in a total of four beneficial mutations in

the best variants. Finding the same mutants using a directed evolution approach would likely require multiple rounds of mutagenesis. Developing a rational approach to reliably engineer brightness decreases the need for high throughput methods and the cost required to produce useful mutants.

Another advantage of rational design for the engineering of RFP brightness is its potential for engineering brighter proteins while targeting a specific parameter. Past engineering approaches have typically led to the identification of mutations that improve RFP brightness by increasing the extinction coefficient to a greater extent than the quantum yield (*8, 10, 11*). This was either achieved by stabilizing the chromophore in its brighter *cis* conformation using rationally-designed mutations (*8, 12*) or by isolating mutations that increase chromophore maturation efficiency using directed evolution approaches (*10, 11*). Since traditional methods for quantifying protein concentration cannot distinguish molecules containing immature or mature red chromophore in the RFP population in solution, improved molar extinction coefficients obtained by directed evolution prior to the development of the dynamic difference method in 2009 (*13*) tend to reflect improved maturation efficiency. Screening for increased fluorescence intensity alone during directed evolution does not allow one to isolate variants whose brightness is increased due to an improvement in a single specific property such as quantum yield, extinction coefficient or maturation. This work presents a rational design approach for increasing the quantum yield of monomeric RFPs that already contain a bulky or aromatic residue below the chromophore. This method is an improvement over random mutagenesis approaches not only because of the significant decrease in the size of the library that needs to be screened, but because it targets the quantum yield of the protein specifically. Of the characterized mutants (Table 2.1), the extinction coefficient increased significantly only in mRojo-VHSV (1.4x), while the quantum yield of this

mutant increased 1.7–fold. The highest quantum yield increase (2.1 x) was achieved in mRojo-VYGV, without significantly affecting the extinction coefficient of the protein. Therefore, this project achieved increases in brightness by targeting the parameter that cannot be engineered specifically through random approaches and is less likely to be affected by them (*8, 11, 12*). The rational design approach that was developed could be used alone or in conjunction with design strategies that increase the extinction coefficient, maximizing the potential for engineering bright fluorescent protein variants.

4.3. Future Work

Although we were able to successfully increase the quantum yield of RFPs using rational design, the resulting variants are not as bright as other known RFPs, limiting their applicability in various biological imaging experiments. Furthermore, attempting to replicate the success of the mutants found in Chapter 2 by creating mCherry-VYGV did not yield a mutant with an increased quantum yield or brightness. There are three ways to address this going forward. In the first approach, the designed mutations could be introduced into fluorescent proteins displaying high brightness that already contain a π-stacking residue on one side of the chromophore. For example, YFP, which has Y203 on one side of the chromophore (Figure 4.3.1A), or cpYGFP, which has H52 on the other, could be good candidates for a proof of concept (*14*).

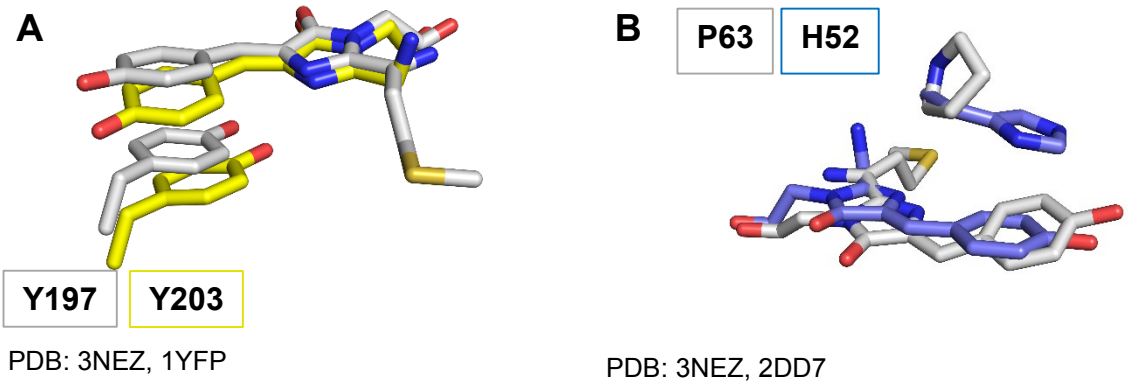

Figure 4.3.1. mRojoA (white) overlay with potential candidates for rational design of triple-decker motif. YFP (A, yellow) and cpYGFP (B, blue) are aligned with mRojoA and shown as sticks.

In the second approach, the brightness of the mRojoA mutants that we engineered could be further improved by directed evolution to produce useful variants. Directed evolution could be used to identify mutations that improve chromophore maturation and extinction coefficient, resulting in higher brightness. Screening could be performed using fluorescence activated cell sorting (FACS), enabling libraries containing up to 10^7 mutants to be screened for increased brightness.

As a third approach to improving upon the findings of this study, rational design could be attempted to rigidify the overall barrel structure of the RFP via the introduction of mutations away from the chromophore. To identify potential mutations that could rigidify the protein backbone, a survey of past studies on the correlation between rigidity and mutagenesis of different residues across the protein backbone would be required (*15, 16, 17*). As the mapping of useful mutations on these proteins grows and our understanding of their structure/function relationships improve, mutagenesis of the backbone to rigidify the protein core should become the preferred path to

achieving improvements in quantum yield, since mutating residues near the chromophore poses the risk of interfering with other properties of the protein such as excitation and emission wavelengths (*15, 18*). In the present work, all of the mutations caused a blue shift in fluorescence, which is undesirable in the context of engineering red and far-red fluorescent proteins for whole animal imaging. An example of engineering an increase in quantum yield through rigidification of the protein backbone was provided by Goedhart *et al*, who created mTurquoise2, a cyan fluorescent protein that has the distinction of having the highest quantum yield ($\phi = 0.93$) reported for a monomeric fluorescent protein to date (*17*). One of the approaches employed when engineering this protein involved identifying residues found on the β barrel of the protein that had low electron density in the crystal structure of the parent protein as well as high structural deviations from the structures of several related proteins, which could suggest flexibility (Figure 4.3.3A). They identified β-strand 7 as a dynamic part of the barrel structure, and pinpointed I146 as a dynamic residue. The I146F mutation at this site introduced a larger side chain, which increased the number of van der Waals interactions with neighbouring residues and stabilized the β strand. A similar approach could be beneficial for the mRojoA mutants with known crystal structures, and could be used to identify residues of low electron density or ones whose backbone conformation differs among the mutants. The crystal structures could also be analyzed using molecular dynamics simulations in order to identify dynamic residues for mutagenesis. Rigidification of the entire protein structure would ultimately lead to rigidification of the chromophore, and therefore an increase in quantum yield.

4.4. References

1. Merzlyak, E. M., Goedhart, J., Shcherbo, D., Bulina, M. E., Shcheglov, A. S., Fradkov, A. F., Gaintzeva, A., Lukyanov, K. A., Lukyanov, S., Gadella, T. W. J., and Chudakov, D. M. (2007) Bright monomeric red fluorescent protein with an extended fluorescence lifetime, *Nature Methods 4*(7), 555-557. doi:10.1038/NMETH1062

2. Shaner, N., Lambert, G. G., Chammas, A., Ni, Y., Cranfill, P. J., Baird, M. A., Sell, B. R., Allen, J. R., Day, R. N., Israelsson, M., Davidson, M. W., and Wang, J. (2013) A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum, Nature Methods 10*, 407-409. doi:10.1038/nmeth.2413

3. Drexhage, K. H., and Reynolds, G. A. (1974) New dye solutions for mode-locking infrared lasers, *Optics Communications 10*(1), 18-20. doi:10.1016/0030-4018(74)90093-5

4. Oster, G., and Nishijima, Y. (1956) Fluorescence and internal rotation - their dependence on viscosity of the medium, *Journal of the American Chemical Society 78*(8), 1581-1584. doi:10.1021/ja01589a021

5. Chu, J., Haynes, R. D., Corbel, S. Y., Li, P., Gonzalez-Gonzalez, E., and Burg, J. S. (2014) Non-invasive intraviral imaging of cellular differentiation with a bright red-excitable fluorescent protein, *Nature Methods 11*(5), 572-578. doi:10.1038/nmeth.2888

6. Shaner, N. C., Lin, M. Z., McKeown, M. R., Steinbach, P. A., Hazelwood, K. L., Davidson, M. W., and Tsien, R. Y. (2008) Improving the photostability of bright monomeric orange and red fluorescent proteins, *Nature Methods 5*(6), 545-551. doi:10.1038/nmeth.1209

7. Moore, M. M., Oteng-Pabi, S. K., Pandelieva, A. T., Mayo, S. L., and Chica, R. A. (2012) Recovery of red fluorescent protein chromophore maturation deficiency through rational design, *Public Library of Science One 7*(12), e52463. doi:10.1371/journal.pone.0052463

8. Shaner, N., Campbell, R., Steinbach, P., Giepmans, B., Palmer, A., and Tsien, R. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from discosoma sp red fluorescent protein, *Nature Biotechnology 22*(12), 1567-1572. doi:10.1038/nbt1037

9. Campbell, R., Tour, O., Palmer, A., Steinbach, P., Baird, G., Zacharias, D., and Tsien, R. (2002) A monomeric red fluorescent protein, *Proceedings of the National Academy of Sciences of the United States of America 99*(12), 7877-7882. doi:10.1073/pnas.082243699

10. Glick, B. S., Strongin, D. E., Keenan, R., Strack, R. L., and Bhattacharyya, D. (2014) Red fluorescent proteins with enhanced bacterial expression, increased brightness and reduced aggregation. US8679749 B2

11. Shcherbo, D., Murphy, C. S., Ermakova, G. V., Solovieva, E. A., Chepurnykh, T. V., Shcheglov, A. S., Verkhusha, V. V., Pletnev, V. Z., Hazelwood, K. L., Roche, P. M., Lukyanov, S., Zaraisky, A. G., Davidson, M. W., and Chudakov, D. M. (2009) Far-red

fluorescent tags for protein imaging in living tissues, *Biochemical Journal 418*, 567-574. doi:10.1042/BJ20081949

12. Lam, A. J., St-Pierre, F., Gong, Y., Marshall, J. D., Cranfill, P. J., Baird, M. A., McKeown, M. R., Wiedenmann, J., Davidson, M. W., Schnitzer, M. J., Tsien, R. Y., and Lin, M. Z. (2012) Improving FRET dynamic range with bright green and red fluorescent proteins, *Nature Methods 9*(10), 1005-+. doi:10.1038/NMETH.2171

13. Kredel, S., Nienhaus, K., Oswald, F., Wolff, M., Ivanchenko, S., Cymer, F., Jeromin, A., Michel, F. J., Spindler, K.-D., Heilker, R., Nienhaus, G. U., and Wiedenmann, J. (2008) Optimized and Far-Red-Emitting Variants of Fluorescent Protein eqFP611, *Chemistry & Biology 15*, 224-233.

14. Suto, K., Masuda, H., Takenaka, Y., Tsuji, F. I., and Mizuno, H. (2009) Structural basis for red-shifted emission of a GFP-like protein from the marine copepod chiridius poppei, *Genes Cells 16*(6), 727-737. doi: 10.1111/j.1365-2443.2009.01305.x.

15. Laurent, A. D., Mironov, V. A., Chapagain, P. P., Nemukhin, A. V., and Krylov, A. I. (2012) Exploring structural and optical properties of fluorescent proteins by squeezing: Modeling high-pressure effects on the mstrawberry and mcherry red fluorescent proteins, *Journal of Physical Chemistry B 116*(41), 12426-12440. doi:10.1021/jp3060944

16. Seifert, M. H. J., Georgescu, J., Ksiazek, D., Smialowski, P., Rehm, T., Steipe, B., and Holak, T. A. (2003) Backbone dynamics of green fluorescent protein and the effect of histidine 148 substitution, *Biochemistry 42*(9), 2500-2512. doi:10.1021/bi026481b

17. Goedhart, J., von Stetten, D., Noirclerc-Savoye, M., Lelimousin, M., Joosen, L., Hink, M. A., van Weeren, L., Gadella, T. W. J. Jr., ande Royant, A. (2012) Structure – guided evolution of cyan fluorescent proteins towards a quantum yield of 93%, *Nature Communications 3*:751. doi:10.1038/ncomms1738

18. Wannier, T. M., and Mayo, S. L. (2014) The structure of a far-red fluorescent protein, AQ143, shows evidence in support of reported red-shifting chromophore interactions, *Protein Science 23*(8), 1148-1153. doi:10.1002/pro.2498

5. Appendix

5.1. DNA Construction

The mutants were constructed using splicing through overlap extension (SOE) mutagenesis (*1*) (Figure 5.1.1). In this method, a template DNA molecule undergoes several PCR reactions in order to introduce a desired mutation into a specific position on the DNA strand. The mutations are contained on mutagenesis primers that anneal to the DNA strand around the target position. Two PCR reactions, each with a mutagenesis primer and a flanking primer that anneals to the end of the DNA of interest, create two fragments that overlap and contain the desired mutation. A third PCR reaction, sometimes called a combination PCR, is performed using the two fragments in a 1:1 ratio as the template in addition to the two flanking primers. This yields a combined fragment that contains the mutated residue.

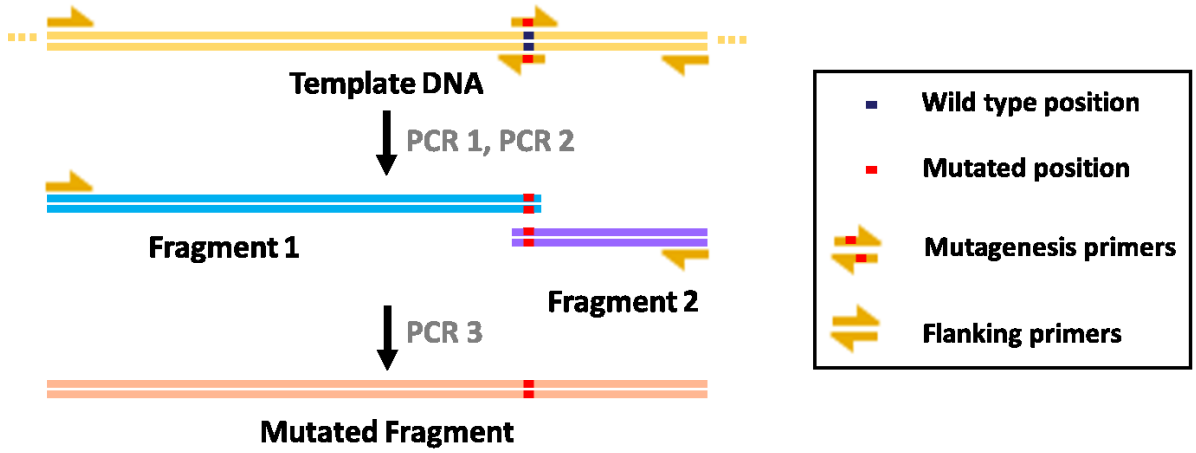

Figure 5.1.1. Diagram explaining SOE mutagenesis. The desired mutation (red) replaces the wild type codon (blue) at the target position by undergoing PCR reactions with mutagenesis primers that contain the desired mutation(s).

The flanking primers can contain restriction sites for inclusion of the mutated DNA fragment into a plasmid. The primers *Nde*I-His6-mFruits-F and mCherry-*Bam*HI-R (Table 5.1.1) were used as forward and reverse flanking primers, respectively, for the mutagenesis reactions of all the proteins in this project, since they anneal to parts of the DNA strand that are outside of the coding region of the DNA. These flanking primers contain restriction sites for the enzymes *Nde*I and *Bam*HI. *Nde*I-His6-mFruits-F also encodes a polyhistidine tag for protein purification using Ni-NTA affinity chromatography.

The mutagenesis primers are listed in Table 5.1.1. Where possible, degenerate codons were used to introduce multiple mutations at the same site. Degenerate codons contain an equimolar mix of nucleotides at one or more positions, allowing for the simultaneous introduction of more than one mutation. In Table 5.1.1, K corresponds to a mix of G and T, M corresponds to a mix of A and C, and S corresponds to a mix of C and G.

Figure 5.1.2 contains examples of the mutagenesis performed at two of the target positions. Few extra bands are observed, indicating that the primers annealed to the correct sites on the DNA and amplification was successful. After each PCR reaction, gel extraction of the desired bands was performed to ensure purity of the DNA for the subsequent step and avoid contamination from the template DNA. Following digestion with *Nde*I and *Bam*HI-HF restriction enzymes, the fully mutated DNA fragments, at a size of 752 bp, were ligated into double digested pET11a vector.

Table 5.1.1. Mutagenesis primers used for library construction and individual mutant preparation.

Primer name	Sequence
RojoA-16GTT-F[1]	AGTTCATGCGCTTCAAG**GTT**CACATGGAGGGCT
RojoA-16AAC-R[1]	AGCCCTCCATGTG**AAC**CTTGAAGCGCATGAACT
RojoA-63CAT-F	TGGGACATCCTGTCC**CAT**CAGTTCATGTACGGC
RojoA-63ATG-R	GCCGTACATGAACTG**ATG**GGACAGGATGTCCCA
RojoA-63TAC-F	TGGGACATCCTGTCC**TAC**CAGTTCATGTACGGC
RojoA-63GTA-R	GCCGTACATGAACTG**GTA**GGACAGGATGTCCCA
RojoA-63TTC-F	TGGGACATCCTGTCC**TTC**CAGTTCATGTACGGC
RojoA-63GAA-R	GCCGTACATGAACTG**GAA**GGACAGGATGTCCCA
RojoA-63CCG-F	TGGGACATCCTGTCC**CCG**CAGTTCATGTACGGC
RojoA-63CGG-R	GCCGTACATGAACTG**CGG**GGACAGGATGTCCCA
RojoA-143KGG-F	GAAGAAGACCATGGGC**KKG**GAGGCCAGCTCC
RojoA-143CCM-R	GGAGCTGGCCTC**CMM**GCCCATGGTCTTCTTC
RojoA-143GGT-F	GAAGAAGACCATGGGC**GGT**GAGGCCAGCTCC
RojoA-143ACC-R	GGAGCTGGCCTC**ACC**GCCCATGGTCTTCTTC
RojoA-163STG-F	GAAGGGCGAGATCAAG**STG**AGGCTGAAGCTGAAGGA
RojoA-163CAS-R	TCCTTCAGCTTCAGCCT**CAS**CTTGATCTCGCCCTTC
RojoA-163GTT-F	GAAGGGCGAGATCAAG**GTT**AGGCTGAAGCTGAAGGA
RojoA-163AAC-R	TCCTTCAGCTTCAGCCT**AAC**CTTGATCTCGCCCTTC
Cherry-143GGT-F[2]	GAAGAAGACCATGGGC**GGT**GAGGCCTCCTCCG
Cherry-143ACC-R[2]	CGGAGGAGGCCTC**ACC**GCCCATGGTCTTCTTC
NdeI-His6-mFruits-F (Flanking)	ACACCACATATGGGTCATCACCATCACCATCACGGTGTGAGCAAGGGCGAGGAG
mCherry-BamHI-R (Flanking)	TGGTGTGGATCCTTACTTGTACAGCTCGTCCAT

[1]Degenerate codons were not employed due to the impossibility of excluding unwanted mutations. Instead, the mutation was introduced individually and mixed with DNA containing the parent protein threonine to create the library.

[2]Unless specifically included in the table, mutagenesis primers specific to mCherry were not used. The mRojoA sequence is identical to that of mCherry at the positions in question.

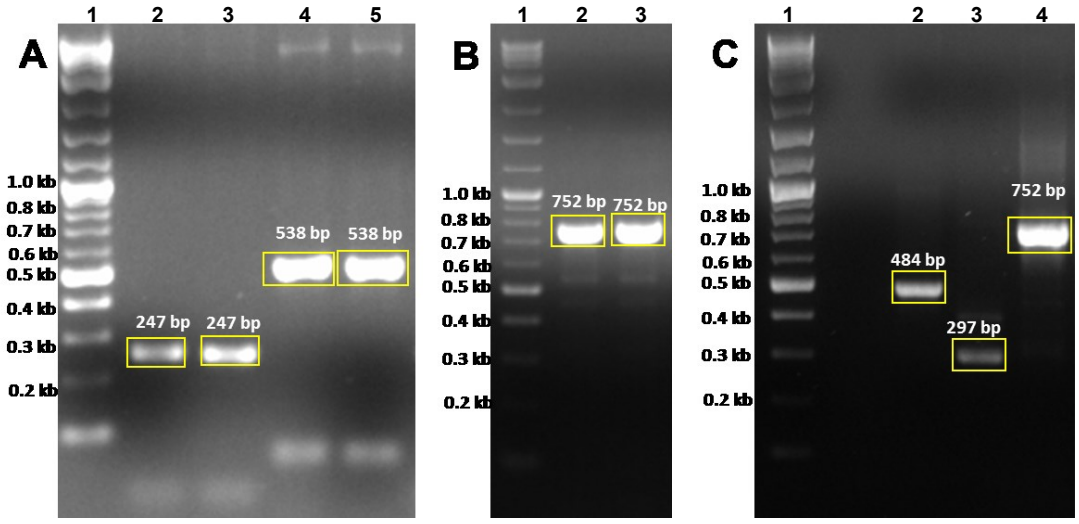

Figure 5.1.2. Mutagenesis results. Ethidium bromide agarose gels (1%) representing mutagenesis of mRojoA at positions 63 (duplicate of: A – intermediate mutagenesis fragments, B – full-length combination mutagenesis fragment) and 143 (2,3 – intermediate mutagenesis fragments, 4 – combination fragment) (C). Excised bands are indicated by yellow squares. Well 1 of each gel contains 2 – log DNA ladder (0.1 – 10.0 kb) from New England Biolabs.

5.2. Protein Purification

The fluorescent proteins all contained a polyhistidine (His) tag for purification using Ni-NTA resin (Promega). Cells were grown using the conditions outlined in section 2.5.3. Purification was performed according to the manufacturer's protocol. Briefly, after homogenizing the cells in 100 mM phosphate buffer at pH 8 supplemented with 5 mM imidazole, 8 mg lysozyme (Fisher Scientific) and 2 µL benzonase nuclease (VWR), the cell debris (*pellet*) were separated from the soluble portion (*supernatant*) through centrifugation at 18,000 g. Following a one-hour incubation of the supernatant in a column with 1 mL of Ni-NTA resin (Promega), the liquid was allowed to drain (*flow-through*), leaving the resin with attached His-tagged protein. Two wash steps, one with 10 mM imidazole (*wash 1*) and one with 20 mM imidazole (*wash 2*), were performed to elute proteins that were non-specifically bound. Finally, the His-tagged RFP was eluted with 4 ml of 100 mM phosphate buffer at pH 7.4 with 250 mM imidazole (*eluted*). A final desalting step, performed using gel filtration (EconoPAC 10DG columns, Bio-Rad), exchanged the imidazole-containing elution buffer for 5 mL of phosphate-buffered saline (PBS) at pH 7.4 (*desalted*). Figure 5.2.1 is an SDS-PAGE gel of an mRojoA purification, where each fraction was heated at 100 °C for 2 minutes prior to loading on the gel. The red box surrounds the band in each well that corresponds to mRojoA, which has a molecular weight of 28 kDa. While the pellet contains a large amount of the protein, the majority remains soluble in the supernatant to allow for its purification. Some protein is present in the flow-through, indicating that the soluble fraction contained enough protein to fully saturate the amount of resin that was used. A minimal amount of mRojoA was washed away during Wash 1 and 2, and other proteins were washed away during this step. Lastly, the eluted fraction contains three bands – a strong band at approximately 30 kDa corresponding to mRojoA, a very faint band at around 66 kDa that may be a weak dimer of mRojoA, and a third

band at approximately 20 kDa. This band is one of two that are commonly observed during SDS-PAGE electrophoresis of RFPs after boiling of the sample. It corresponds to cleavage of the C=N bond of the acylimine group in the red chromophore (*2*). A second band resulting from this cleavage, which is at approximately 8 kDa, is not always visible in gels containing 10% acrylamide due to its small size.

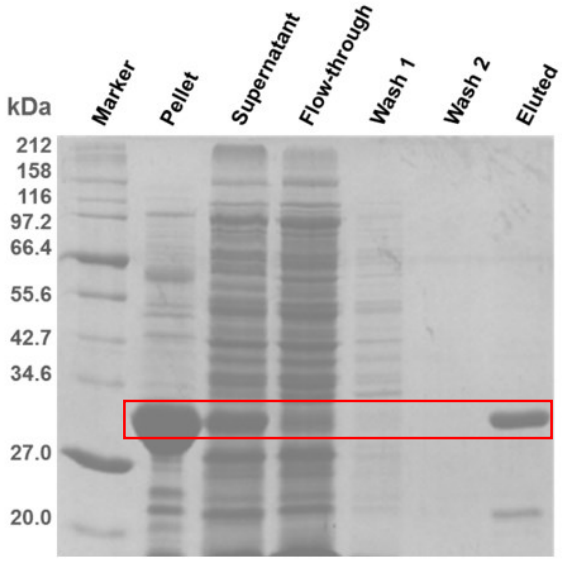

Figure 5.2.1. Example SDS-Page gel of mRojoA purification. The gel contains 10% acrylamide. mRojoA, outlined in red, has a molecular weight of approximately 28 kDa. The band at approx. 20 kDa corresponds to a cleavage product of heat denaturation during gel preparation (*5*).

Purification in the manner described above, using 500 ml of cell culture, results in protein preparations containing 1–5 mg of protein containing mature chromophore. This is a sufficient amount to perform a characterization of the quantum yield and extinction coefficient of the protein, to obtain a pH profile, or to undergo further purification through size exclusion chromatography to eventually set up one screening plate (96 wells) or 3–5 optimization plates (24 wells) for crystallography.

All the proteins that were characterized in detail were the result of a minimum of three separate purifications following this protocol. A representative gel of the desalted fractions of three standards and three mutants is shown in Figure 5.2.2. This gel contains 15% acrylamide, allowing for better resolution of the two lower bands corresponding to the two fragments obtained by cleavage of the acylimine bond found in RFP containing mature red chromophore. These bands are more prominent in the three standards than in the mutants, indicating a higher proportion of red chromophore, as opposed to green or uncyclized / immature chromophore, which do not contain the labile acylimine group. Very faint bands are observed at approximately 60 kDa for mRojo-VYGV and mRojo-TYGL, probably corresponding to a small proportion of protein dimers. The desalted fractions did not contain bands at other molecular weights, indicating that the protein preparations were free of contaminating proteins.

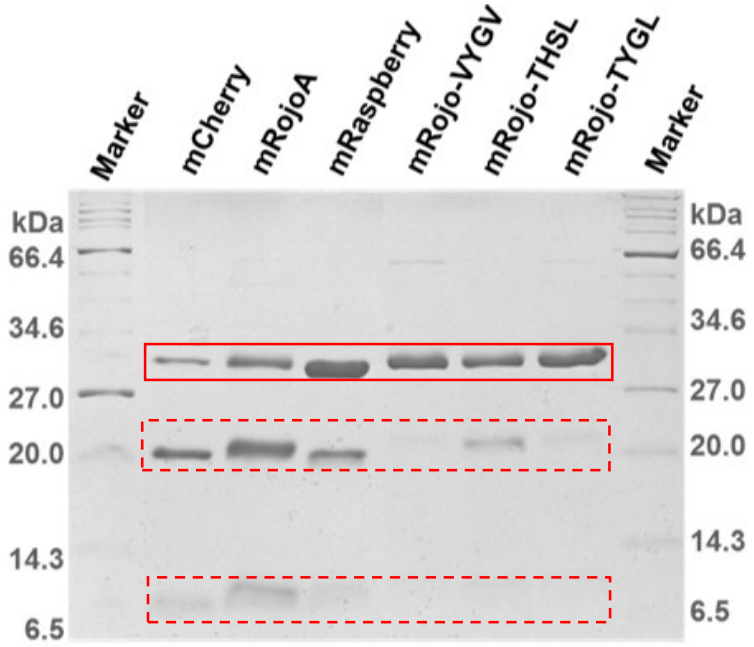

Figure 5.2.2. Example SDS-PAGE gel of purified standards and mutants. Gel containing 15% acrylamide. All proteins are at 26 – 28 kDa (solid red square). Bands at approximately 20 kDa and 7 kDa (red square, dashed) correspond to red chromophore-containing protein that was cleaved during the heat denaturation process.

5.3. Characterization of Fluorescent Properties

5.3.1. Quantum Yield Calculation

The quantum yields of the proteins were calculated by comparison with two standard proteins of known quantum yield – mCherry (φ = 0.22) (*3*) and mRaspberry (φ = 0.15) (*4*). The standard proteins must absorb and emit at similar wavelengths to the unknown, so two red fluorescent proteins were chosen. The quantum yield is the efficiency of conversion of absorbed photons into emitted photons (equation 2). Therefore, assuming identical detector sensitivity (i.e. gain), the following equation can be used to calculate the quantum yield of an unknown using a standard of known quantum yield:

$$\varphi_u = \varphi_s \left(\frac{A_s}{A_u}\right)\left(\frac{E_u}{E_s}\right)\left(\frac{I_s}{I_u}\right)\left(\frac{n_u^2}{n_s^2}\right) \quad [6]$$

where φ_u and φ_s are the quantum yields of the unknown and standard, respectively; A_u and A_s are their respective absorbances at the excitation wavelength; E_u and E_s are their respective integrated fluorescence emission intensities; I_u and I_s are the incident powers of illuminating light for the two samples; and n_u and n_s are the refractive indices of the solutions (*5*).

In order to ensure that the samples are at the correct concentration to avoid self-quenching of the fluorophores, the absorbance was maintained below 0.1. As verification, graphs of integrated fluorescence as a function of absorbance at different protein concentrations were plotted (Figure 5.3.1). These plots were linear, indicating no self-quenching. They also provided a better estimate of the quantum yield, since they contained more than one data point. The slope of the linear fit applied to this data set provides a numerical measure of the efficiency of fluorescence. In order to convert this number to a quantum yield, a comparison of the slopes of the unknown to a standard

of known quantum yield is necessary. Assuming that the incident light is of the same wavelength and power, the fluorescence is detected with the same gain, and the proteins are in solutions of equal refractive indices, the following equation can be applied to the calculation of the quantum yield:

$$\varphi_u = \varphi_s \left(\frac{m_u}{m_s}\right) \quad [7]$$

where m_u and m_s are the slopes of the linear equation fitted to the data points of the unknown and standard proteins, respectively. This equation was applied to the calculation of the quantum yields for all proteins in Table 2.1, with mCherry and mRaspberry as standards. The data was collected in triplicate for each protein preparation, and at least three preparations were made for each protein in order to obtain an average quantum yield. Figure 5.3.1 shows the data fits of single preparations of the two fluorescence standards, mRojoA, and the brightest mutant in each library. They were calculated from 6–9 data points obtained by a 0.5× serial dilution of each protein sample and blanked using the absorbance and fluorescence of the buffer that was used. Each fit had an R^2 of approximately 1.0, indicating an excellent fit of the data to this equation.

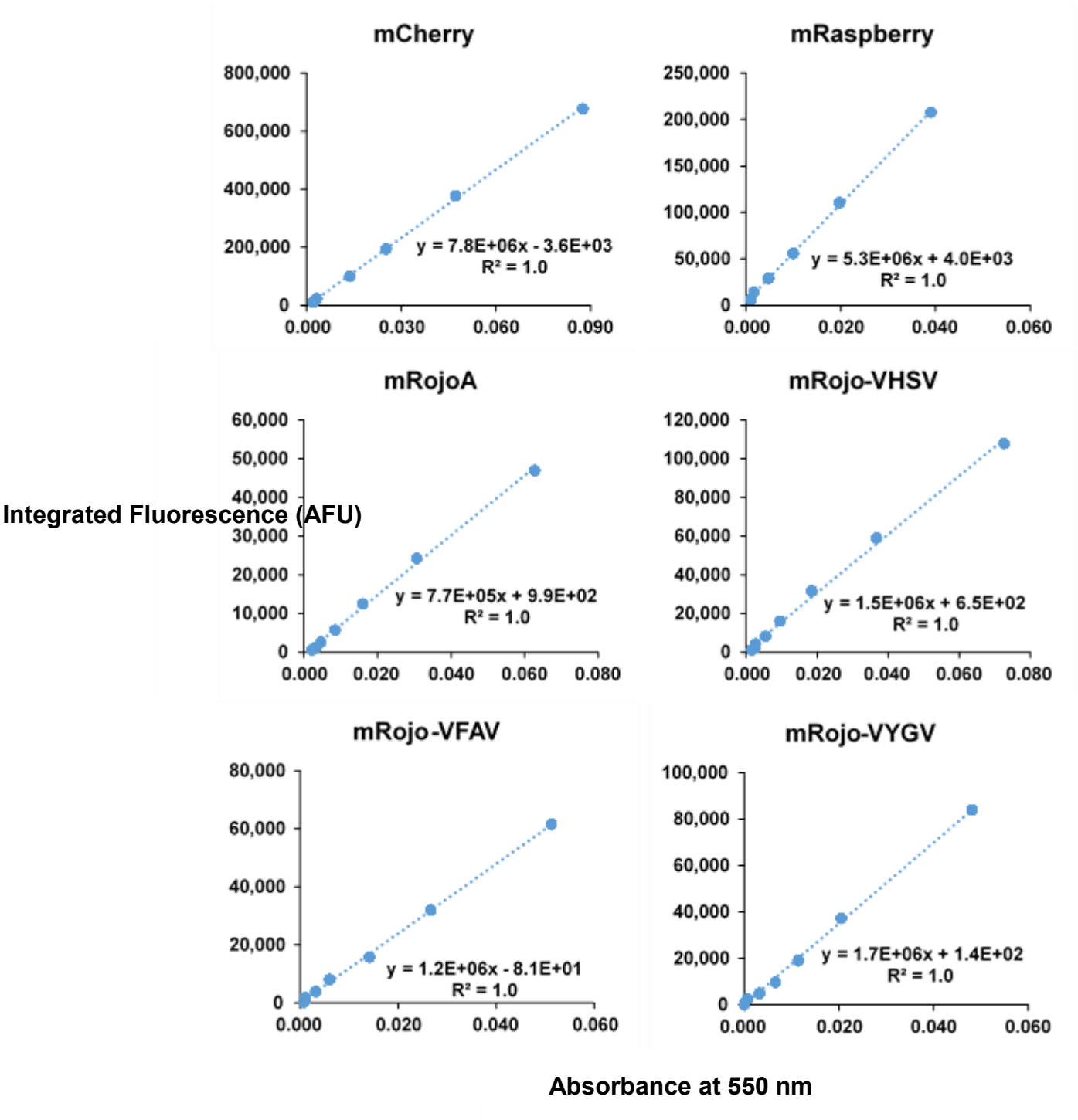

Figure 5.3.1. Integrated fluorescence of various RFPs as a function of absorbance at 550 nm. The slope of the regression of each graph was used to calculate the protein's quantum yield.

5.3.2. Extinction Coefficient Calculation

The extinction coefficient is traditionally calculated using Beer's law, which is seen in equation 3 of section 1.3 and can be rewritten as follows:

$$\frac{A}{l} = \varepsilon c \qquad [8]$$

Plotting the path-length corrected integrated absorbance of a chromophore (A / l, cm^{-1}) as a function of chromophore concentration (c) provides a linear fit whose slope is the extinction coefficient (ε). The drawback of this approach is that commonly used methods of protein concentration determination do not factor in all three maturation states of the chromophore: uncyclized, green, and red (Figure 1.2). Therefore, any calculated extinction coefficient of the red species would be greatly underestimated through the assumption that all the proteins matured to the red species.

Most calculations of fluorescent protein concentrations therefore use an alternative approach called the alkali denaturation method (*2*), which enables the quantification of FP molecules containing chromophores. In this method, equal volumes of protein solution and 2.0 N NaOH are mixed, resulting in cleavage of both the mature red and green chromophore species and the formation a denatured green species, with peak absorbance at 452 nm. This alkali denaturation product has a known extinction coefficient of 44,000 $M^{-1}cm^{-1}$. Therefore, reading the absorbance at 452 nm of alkali-denatured chromophore (pH 13) allows for the determination of its concentration through Beer's law (equation 3, 7). This concentration is then used instead of protein concentration determined by the Bradford (*6*), bicinchoninic acid (*7*), or other typical protein assays, in equation 8. The drawback of this approach is that alkali denaturation does not

differentiate between molecules that contain a green or red chromophore in the RFP population. While RFPs containing uncyclized chromophore and other protein impurities are not detected through this method, RFP molecules containing a green chromophore are. Therefore, the extinction coefficient of the red chromophore remains underestimated.

In order to address this inaccuracy, the dynamic difference method was developed by Kredel *et al* (*8*). This method takes advantage of the green alkali denaturation product of known extinction coefficient at 452 nm, denoted by $\varepsilon_{g,13}$ (44,000 $M^{-1}cm^{-1}$). Native green chromophore, with an absorbance peak at approximately 510 nm, typically denatures before the red chromophore. Achieving alkali denaturation in a slow, measurable manner, is done through the use of a buffer at lower pH (typically 11.6 – 12.6) (Table 5.3.1). Britton – Robinson buffer, consisting of 0.04 M H_3BO_3, 0.04 M H_3PO_4, and 0.04 M CH_3COOH titrated with NaOH to various pH values, was used in this project because of its broad pH range (*9*). Slow denaturation of the green species is complete after an isosbestic point is reached between the green and red peaks (Figure 5.3.2A, blue circle), i.e. all subsequent absorbance spectra have the same absorbance at this wavelength. This indicates the presence of only native red chromophores, which are also in the process of denaturation. Equation 9 represents the extinction coefficient calculation of a protein undergoing optimal slow denaturation at alkaline pH between 11.6 - 12.6:

$$\varepsilon_{r,7.4} = \varepsilon_{g,13}\left(\frac{\Delta A_{r,11.6-12.6}}{\Delta A_{g,11.6-12.6}}\right)\left(\frac{A_{g,11.6-12.6}}{A_{g,13}}\right)\left(\frac{A_{r,7.4}}{A_{r,11.6-12.6}}\right) \quad [9]$$

$$= (44{,}000\ M^{-1}cm^{-1})\left(\frac{\Delta A_{r,11.6-12.6}}{\Delta A_{g,11.6-12.6}}\right)\left(\frac{A_{g,11.6-12.6}}{A_{g,13}}\right)\left(\frac{A_{r,7.4}}{A_{r,11.6-12.6}}\right)$$

Where each absorbance corresponds to native red chromophore (r) or green alkali denaturation species (g) at a given pH, and ΔA values represent the difference in peak absorbance

between two points in time during slow alkali denaturation (Figure 5.3.2A). The determination of the extinction coefficient of the red chromophore species at pH 7.4 ($\varepsilon_{r,7.4}$) is done by equating it to $\varepsilon_{g,13}$ through the calculation of the ratios of several absorbance peaks at key pH values. Extinction coefficient changes in two ways: through the change in species identity, e.g. hydrolysis of red chromophore at high pH to create a green species with absorbance at 452 nm; and through the change in pH, which also affects the absorption wavelength of the red chromophore. The slow alkali denaturation of red chromophore allows for the calculation of a ratio between the absorbance peaks of the green and red species at a single alkaline pH ($\Delta A_{r,11.6\text{-}12.6}$ / $\Delta A_{g,11.6\text{-}12.6}$) (Figure 5.3.2A, equation 9). As the red species disappears, an equal amount of denatured green species is formed. An ideally slow denaturation allows for the resolution of several intermediate absorption spectra of partially denatured chromophore, in order to obtain an average of the ratios of red to green absorbance calculated between each pair of time points.

Table 5.3.1. Optimal pH for slow denaturation of RFP mutants and $\Delta A_{r,11.6\text{-}12.6}$ / $\Delta A_{g,11.6\text{-}12.6}$.

Protein	$\Delta A_{r,11.6\text{-}12.6}$ / $\Delta A_{g,11.6\text{-}12.6}$	Optimal pH
mRojoA	2.0 ± 0.3	11.6
mRojo-THSL	1.2 ± 0.1	11.8
mRojo-THSV	1.4 ± 0.3	11.8
mRojo-VHSV	1.34 ± 0.04	11.6 - 11.8
mRojo-TFAL	1.8 ± 0.7	11.8 - 12.0
mRojo-TFAV	1.9 ± 0.6	11.8 - 12.0
mRojo-VFAV	1.86 ± 0.05	11.8
mRojo-TYGL	1.5 ± 0.3	11.6
mRojo-TYGV	1.34 ± 0.08	11.6
mRojo-VYGV	1.5 ± 0.1	11.4 - 11.6
mRojo-VYGL	1.266 ± 0.003	11.8
mRojo-VPGV[1]	1.065	11.6
mCherry	2.42 ± 0.05	12.6 - 12.9
mCherry-VYGV	2.2 ± 0.4	12.6 - 12.9

[1] The ratio was only measured once for this protein.

The ratio $\Delta A_{r,11.6\text{-}12.6}$ / $\Delta A_{g,11.6\text{-}12.6}$ (Table 5.3.1, Figure 5.3.2A) links the extinction coefficients of the red and green denatured species by providing a measure of their relative absorbance peaks. However, each of these species absorbs differently at different pH. Therefore, the ratio of peak absorbance of denatured green species at pH 13 and 11.6 – 12.6 is calculated ($A_{g,11.6\text{-}12.6}$ / $A_{g,13}$) (Figure 5.3.2B, equation 9). Finally, the ratio of the absorbance peaks of the red chromophore at pH 11.6 – 12.6 and the desired pH (7.4 or 10 in this work) is calculated ($A_{r,7.4}$ / $A_{r,11.6\text{-}12.6}$) (Figure 5.3.2B, equation 9) in order to provide the final link between $\varepsilon_{r,7.4}$ and $\varepsilon_{g,13}$. One approximation was used in this work. As opposed to red absorbance at pH 11.6 – 12.6, red absorbance at a lower non-denaturing alkaline pH (e.g. pH 11.2 instead of 11.6 for mRojoA) was measured, and it was assumed that the error between the two absorbance values was negligible. This was done, since the time required to mix the protein into the buffer, load into the plate reader, and begin reading the absorbance was long enough that some red chromophore at the higher pH had already denatured, rendering the measurement of the initial red peak impossible.

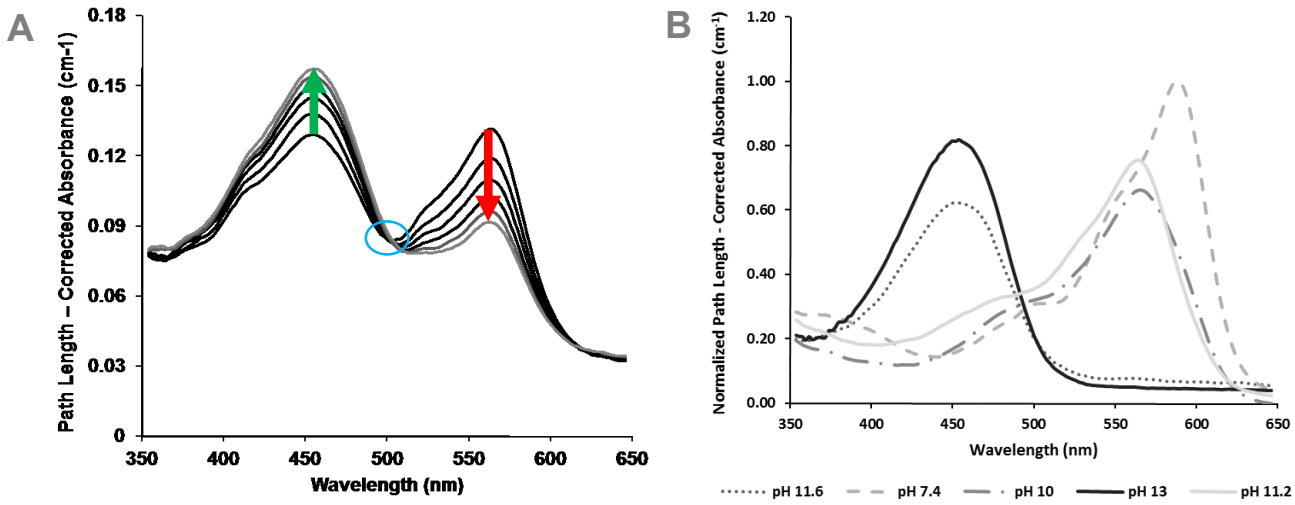

Figure 5.3.2. mRojo-VYGV extinction coefficient determination using dynamic difference method. (A) Path length – corrected absorbance spectra of the slow alkali denaturation (pH 11.6) of mRojo-VYGV red chromophore into a green species absorbing at 452 nm, after attaining the isosbestic point (blue circle). Red and green arrows represent the decrease of red and the increase of green peaks with time, respectively. This ratio of red : green was used to link the extinction coefficients of the two species. Spectra were read at 4.2 minute intervals and smoothed by a simple moving average spanning 8 nm. Peak values were subtracted for each consecutive pair of spectra in order to calculate the ratio $\Delta A_{r,11.6\text{-}12.6}$ / $\Delta A_{g,11.6\text{-}12.6}$. (B) Path length – corrected absorbance spectra of the protein at different pH, normalized to the highest red peak (pH 7.4). Alkali – denatured protein at pH 13 (solid, black) has a known extinction coefficient of 44,000 $M^{-1}cm^{-1}$ at 452 nm. Alkali – denatured protein at pH 11.6 (dotted, grey), attained by allowing slow denaturation to come to completion, absorbs to a lesser extent at the same wavelength. Fully intact red chromophore at an alkaline pH is represented by pH 11.2 (solid, grey), since at pH 11.6 denaturation occurs before the absorbance spectrum can be read. Fully intact red chromophore at pH 7.4 (dashed, grey) and 10 (dashed / dotted, grey) allow the final calculation of extinction coefficient at these pH values (equation 9). (*8*)

5.4. Absorbance and Fluorescence Changes at Different pH

The emission spectra (λ_{ex} = 530 nm) at different pH conditions of the best mRojoA mutants are shown in Figure 5.4.1, while Figure 5.4.2 shows the same spectra for mCherry and mCherry-VYGV. For all proteins, an increase in pH is accompanied by an increase in fluorescence intensity and a red – shift of emission wavelength until the first plateau of the pH profile (Figure 2.S2, Figure 3.3). The spectrum at pH 7.4 (yellow) is part of this plateau. Increasing the pH further results in a blue – shift of fluorescence and an increase in brightness, until the alkaline conditions begin to degrade the chromophore (typically after pH 11 for the mRojoA mutants and after pH 12 for mCherry and mCherry-VYGV). The emission spectra beyond this pH continue to blue – shift, but fluorescence intensity decreases due to chromophore degradation. Similar effects on absorbance and emission spectra have been observed in mCherry and mStrawberry (*10*), where the blue shifts were attributed to the Q215 residue, which H-bonds with the imidazolinone nitrogen of the chromophore. X-ray crystallography of these proteins at different alkaline pH indicate that at pH < 10, Q215 is protonated, participating in a hydrogen bonding interaction with the imidazolinone group. Above this pH, Q215 is deprotonated and interacts with surrounding water molecules instead. This corresponds to the higher pKa values that were measured for the proteins in this paper (Table 2.S1, Table 3.2), supporting the hypothesis that Q215 is responsible for the observed blue shifts at pH above 7.5 – 8.

Beyond the first plateau of the pH profiles (Figure 2.S.4, Figure 3.3), the mRojoA mutants exhibit blue shifts in absorbance of 14 – 25 nm between pH 7.4 and 10 (Table 2.1). However, only mRojo-VHSV, mRojo-VYGL, and mRojo-VYGV experience significant blue – shifts in their fluorescence emission peaks (6 nm, 8 nm, and 8 nm, respectively) (Table 2.1). Interestingly, the portion of the mRojoA profile above pH 7.4 experiences only a slight increase in fluorescence

intensity with increasing pH, which is shown by the lack of significant increases in this protein's quantum yield and extinction coefficient. The absorbance and emission peaks of this protein do not change in this pH range, in contrast to its mutants. Similarly to mRojo-VYGV, mCherry-VYGV experiences a 23-nm blue – shift in absorbance and a 13-nm blue shift in fluorescence between pH 7.4 and pH 10. This is in contrast to mCherry, which only blue – shifts by 10 nm and 6 nm in absorbance and fluorescence, respectively, maintaining a similar Stokes shift at the two pH levels. Since every mutant that was engineered in the present study differs from its parent in terms of the magnitude of the blue shifts in absorbance, and since every mutant contains an aromatic amino acid at position 63, it is likely that the presence of this residue is responsible for this effect. While histidine and tyrosine have ionisable groups, phenylalanine does not. Therefore, the effect of these mutations is likely due to the electronic configuration of these aromatic residues and their interaction with the chromophore phenolate and imidazolinone groups. Analyzing shifts in the emission wavelengths gives some clues concerning the nature of the different ionization effects at varying pH. While mutants containing a F residue at position 63 never exhibit a blue shift in fluorescence emission at pH 10, some mutants containing either a H or Y residue at this position do display an emission wavelength blue shift. In particular, all mRojoA mutants containing the P63H or P63Y mutation in combination with the T16V mutation experience emission blue shifts in emission at high pH. The T16V mutation introduces a hydrophobic residue (V) near the chromophore backbone, which replaces a polar amino acid (T) that forms a water-mediated hydrogen bond to the chromophore acylimine oxygen (11). This likely results in delocalization of charge that affects the chromophore's interaction with Q215. More crystallographic analysis would be preferable in order to observe the effect of a high pH on the T16V mutation combined with an aromatic at position 63. This was attempted for the proteins in

this study, but was unsuccessful (Chapter 5.5). In addition, QM / MM studies of the proteins could help to explain their electronic configurations at varying pH conditions.

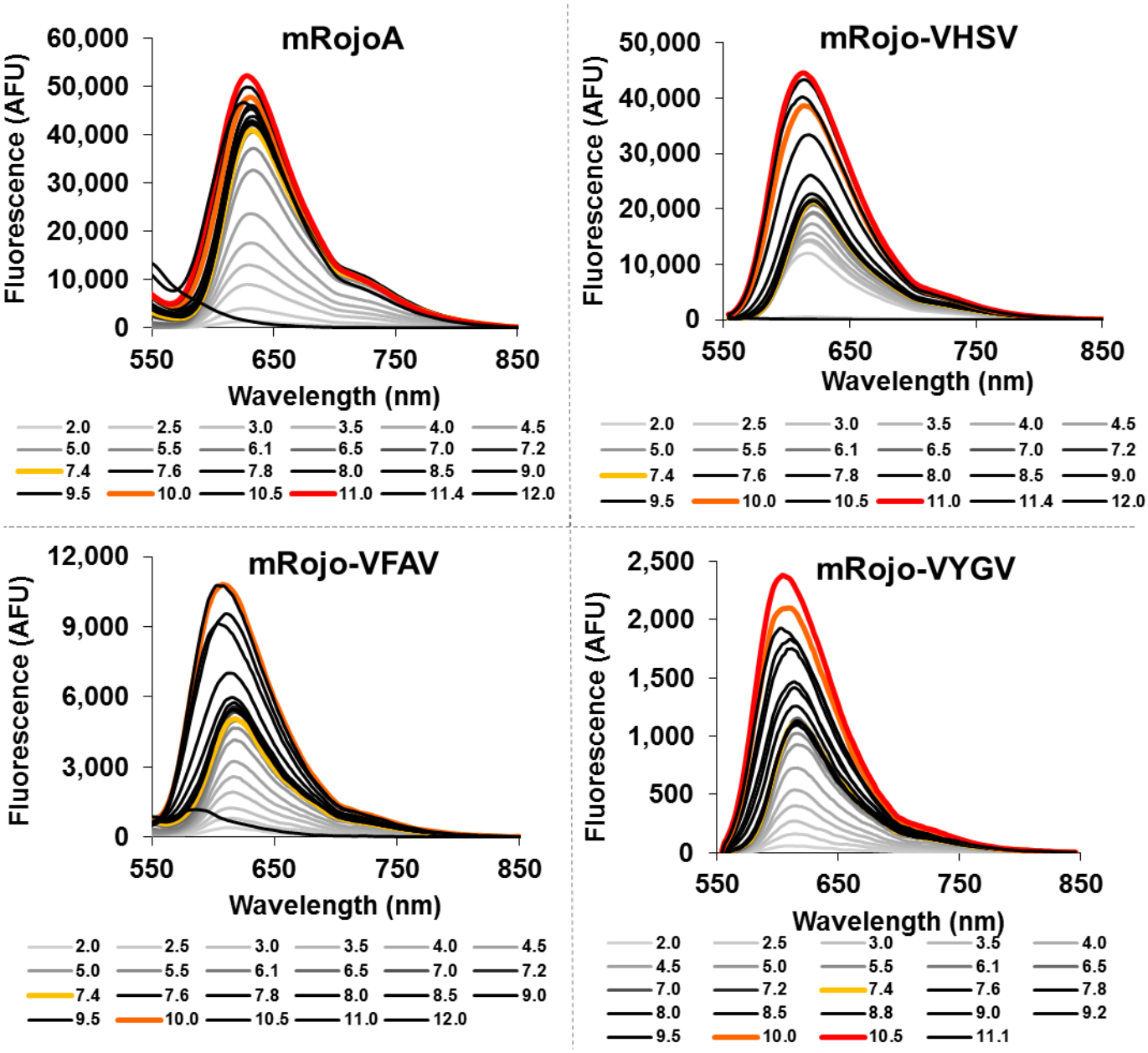

Figure 5.4.1. Emission spectra of mRojoA mutants at various pH. Shading is darker at higher pH. The spectrum with the highest fluorescence is coloured red, the spectrum at pH 10.0 is coloured orange, and the spectrum at pH 7.4 is coloured yellow.

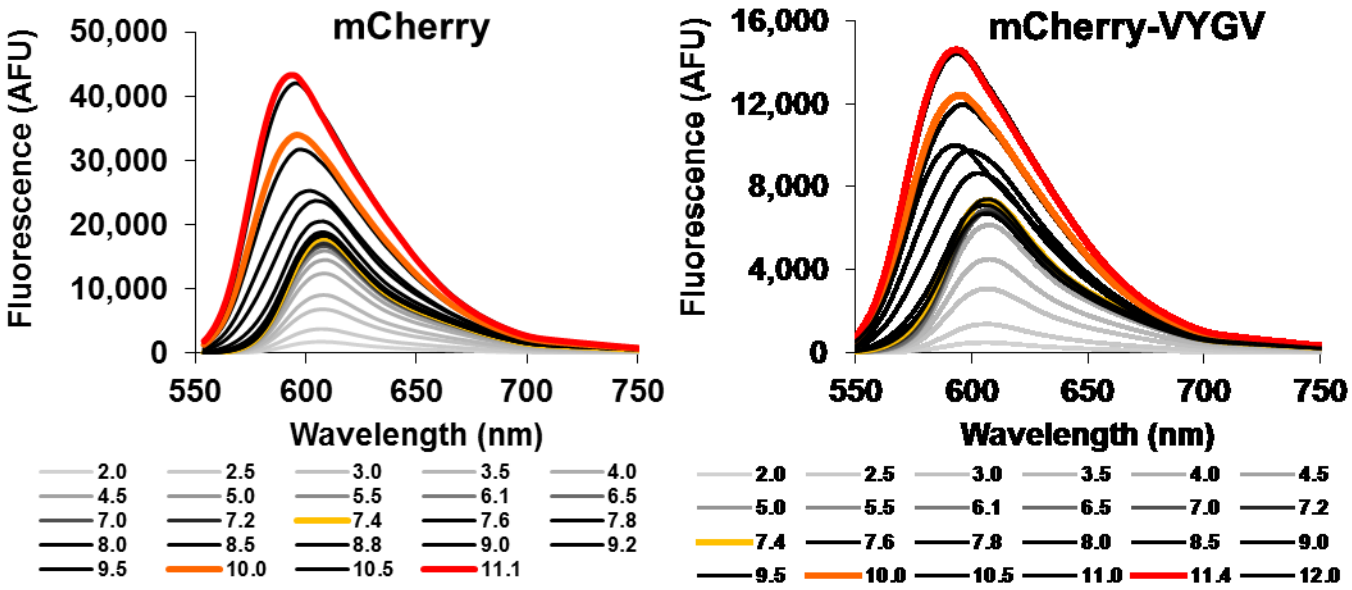

Figure 5.4.2. Emission spectra of mCherry and mCherry-VYGV at various pH. Shading is darker at higher pH. The spectrum with the highest fluorescence is coloured red, the spectrum at pH 10.0 is coloured orange, and the spectrum at pH 7.4 is coloured yellow.

5.5. Optimization of Crystallization Conditions

Crystallization was attempted for mRojo-VYGV, mRojo-VFAV, mRojo-VHSV, mRojo-THSL, and mCherry-VYGV. The proteins differ from each other and mRojoA (*11*) by 2 – 7 mutations, therefore it was possible that similar conditions could lead to crystal growth for at least a subset of these mutants. For this reason, when new crystallization plates were set up for each mutant, conditions that were noted as beneficial to crystal growth in the crystallized mutants were taken into account. The plate numbering in this section is in chronological order, according to when they were first set up.

Each protein underwent at least one commercially provided 96–well screen in sitting drops. All screens were performed by creating sitting drops using 1 μL of precipitant solution and 1 μL of protein in 20 mM Tris – HCl pH 7.4, 100 mM NaCl. **Plate # 1** was set up with mRojo-VYGV, with the protein solution at a concentration of 2.8 mg/ml and 7.5 mg/ml and precipitants found in the Wizard EB-JCSGP-B screen (Rigaku). mRojo-VYGV was also screened using the MCS6-1 screen (Microlytic) in **plate # 3**, at 2.6 mg/ml and 5.1 mg/ml. **Plate # 2** screened mRojo-THSL at 7 mg/ml and 15 mg/ml using the Index HT^{TM} – HR2-134 (Hampton Research) screen. mCherry-VYGV was screened using the Index HT^{TM} – HR2-134 (Hampton Research) screen at a concentration of 5.8 mg/ml (**plate # 10**), while mRojo-VFAV underwent the same screen at 6.2 mg/ml and 19 mg/ml (**plate # 11**). These screens identified 0.1 M Bis-Tris pH 5.5, 25% polyethylene glycol (PEG) 3,350 and 0.2 M $MgCl_2$, 0.1 M Bis-Tris pH 5.5, 25% PEG 3,350 as the best conditions to attempt optimization. A precipitant containing 0.2 M ammonium acetate, 0.1 M Bis-Tris pH 6.5, and 25 % PEG 3,350 was previously used to crystallize mRojoA (*10*). This precipitant solution was contained in two of the three brands of screens used for the mutants, and

resulted in crystal growth in plate 2 only. Therefore, the other two conditions were optimized instead.

In addition to the results of the initial screens, the desired pH of the crystallized proteins was taken into account. Ideally, the precipitant solution of a crystal structure should be at pH 7.4 or 10, because all of the spectral characterization experiments were performed at these two pH values. Since neither the conditions identified by the screens nor the previously found mRojoA crystal structure were at one of these two pH values, another RFP crystallization study was used to identify conditions for optimization that were closer to pH 7.4 and 10. The study, which involved crystallography of mKate at several different pH values, led to the identification of 0.1 M HEPES pH 7.0, 1 M LiCl, 20% PEG 6,000 and 0.1 M bicine pH 9.0, 1 M LiCl, 20% PEG 6,000 as additional candidates for crystal optimization (*12*). All the optimization plates (Figures 5.5.1 – 5.5.6) were set up with the goal of slightly modifying the four identified precipitant conditions in order to obtain the largest crystals with the best diffraction and separation. Most optimization plates contained some crystal growth, and clusters of needles were observed in all the wells of two plates (Table 5.5.1). Most of the crystals were coloured as expected for mature RFP protein, indicating that the chromophore was intact. Bleached crystals were not considered for fishing and screening.

Initially, no screen was used for mRojo-VHSV, since the other screens provided evidence for favourable optimization conditions that were common to all the mutants. Following unsuccessful optimization, fishing, and screening of potential crystals (Figure 5.5.5, Table 5.5.1), an Index HT^{TM} – HR2-134 (Hampton Research) screen was also set up for mRojo-VHSV at concentrations of 7 mg/ml and 35 mg/ml (**plate # 39**). None of the precipitants in this screen

resulted in crystal growth that was deemed promising for further optimization within the time frame of the study.

The observed crystal growth of all the proteins resulted in overlapping or clustered crystals (Figure 5.3.7). These require careful handling during fishing in order to separate individual chunks without cracking the crystal. While they complicate the fishing process, crystal structures from clustered crystals have been obtained for both mRojoA and mKate (*9, 10*). Therefore, fishing and screening was still performed for the mutants. In addition to this, mRojo-VFAV was chosen for an attempt to grow crystals with better separation. Plate 31 was set up with hanging drops and the conditions found in plate 25, which varied types and concentrations of PEG and crystal growth in all the wells (Figure 5.5.4). In plate 32, seeding of mRojo-VFAV was attempted (Figure 5.5.4). A crystal from plate 27 / well B4 was fished, crushed through vortexing with a ceramic bead, and added to the precipitant solution in several dilutions. The resultant microcrystals could provide scaffolds for crystal growth at a separation dictated by the dilution of the crushed crystal. Unfortunately, the crystals that were grown using both of these approaches were more fragile than the ones grown using sitting drops with no seeding. The crystals in plates 31 and 32 also overlapped, requiring additional manipulation during the fishing process. None of these crystals were used for screening.

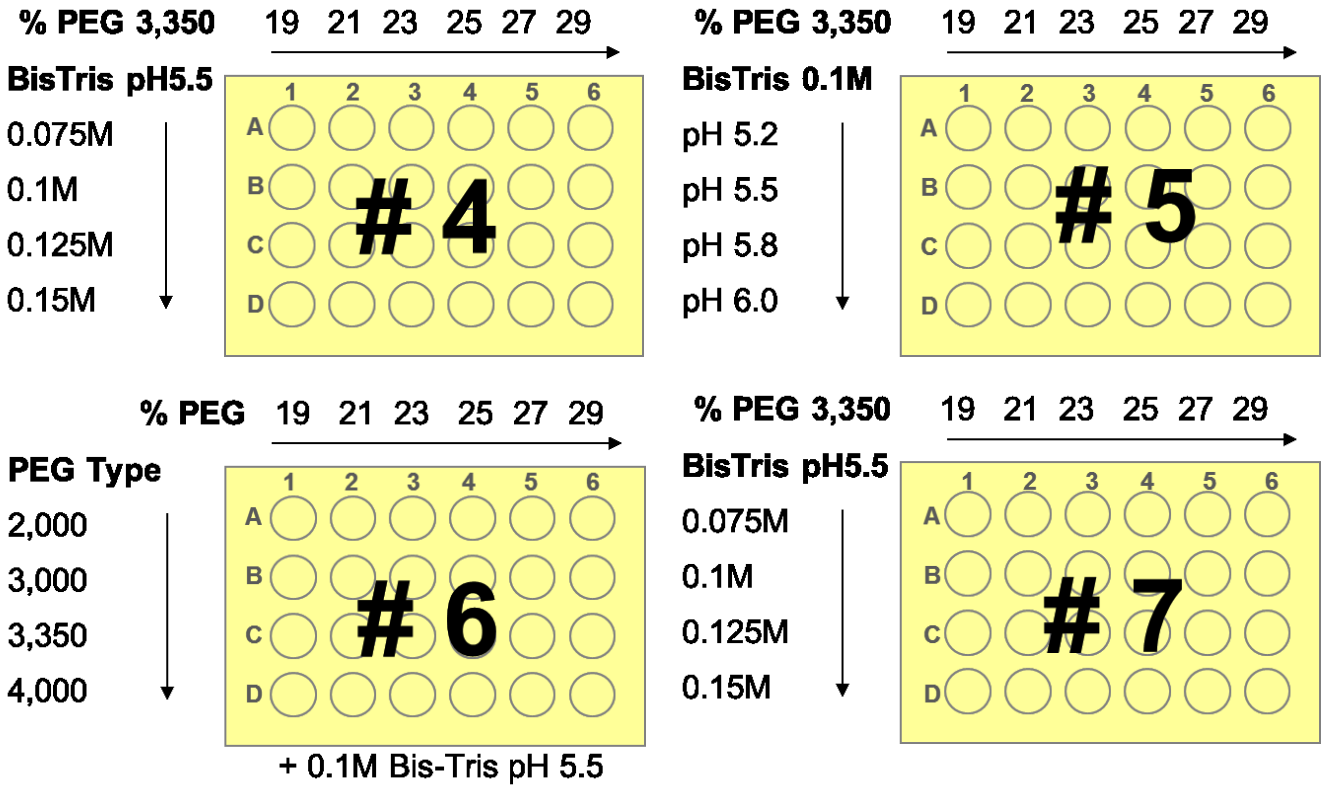

Figure 5.5.1. Optimization plates set up for crystallization of mRojo-THSL. Protein was suspended in 20 mM Tris – HCl pH 7.4, 100 mM NaCl. 1 µl protein solution at 16 mg / ml (# 4 – 6) or 10 mg / ml (# 7) was added to 1 µl of precipitant solution and the crystals were grown in sitting drops. The plates were kept at room temperature.

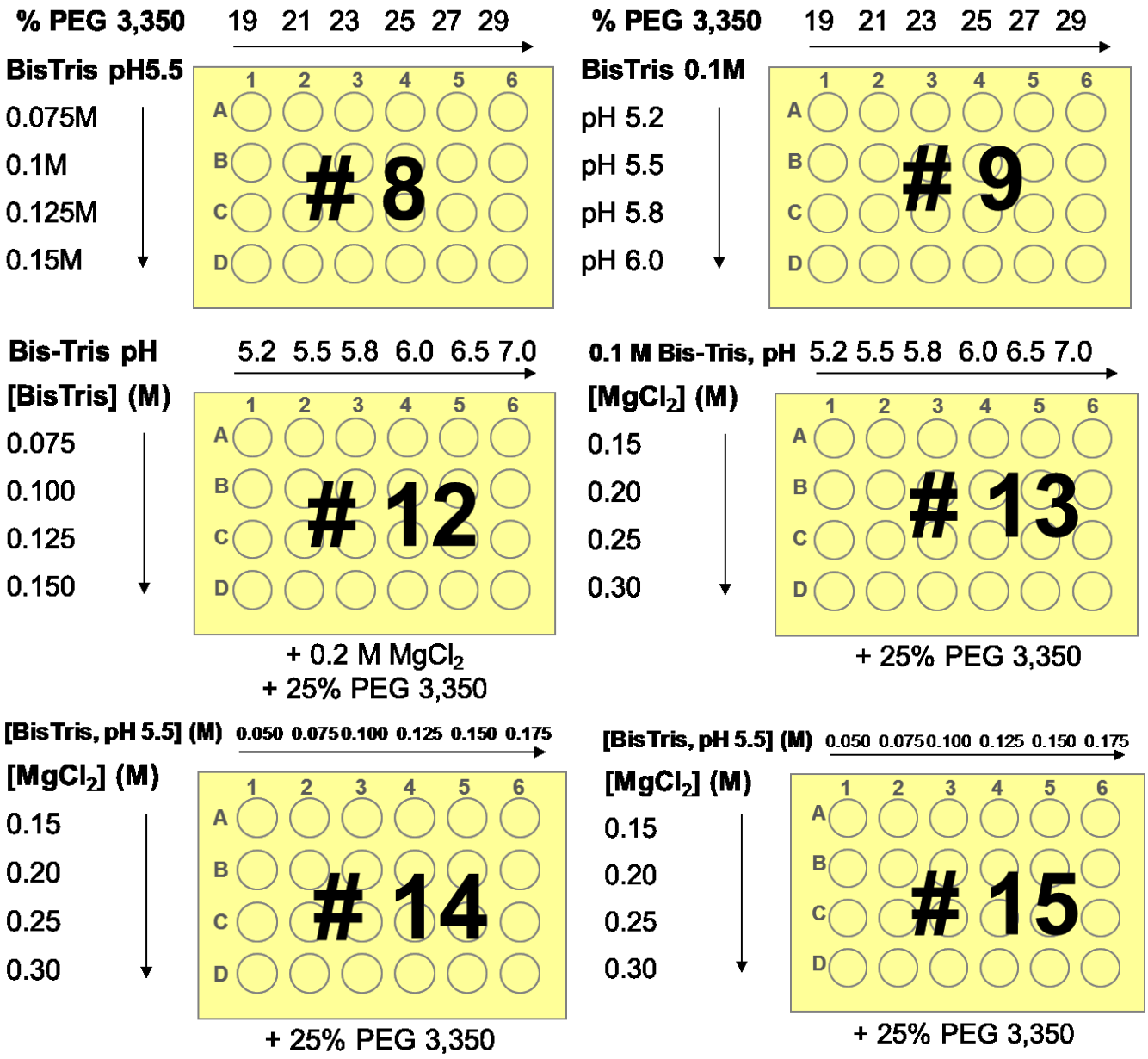

Figure 5.5.2. Optimization plates set up for crystallization of mRojo-VYGV. Protein was suspended in 20 mM Tris – HCl pH 7.4, 100 mM NaCl. 1 µl protein solution at 32 mg / ml (# 8), 16 mg / ml (# 9), 17 mg / ml (# 12 - 14), or 10 mg / ml (# 15) was added to 1 µl of precipitant solution and the crystals were grown in sitting drops. The plates were kept at room temperature.

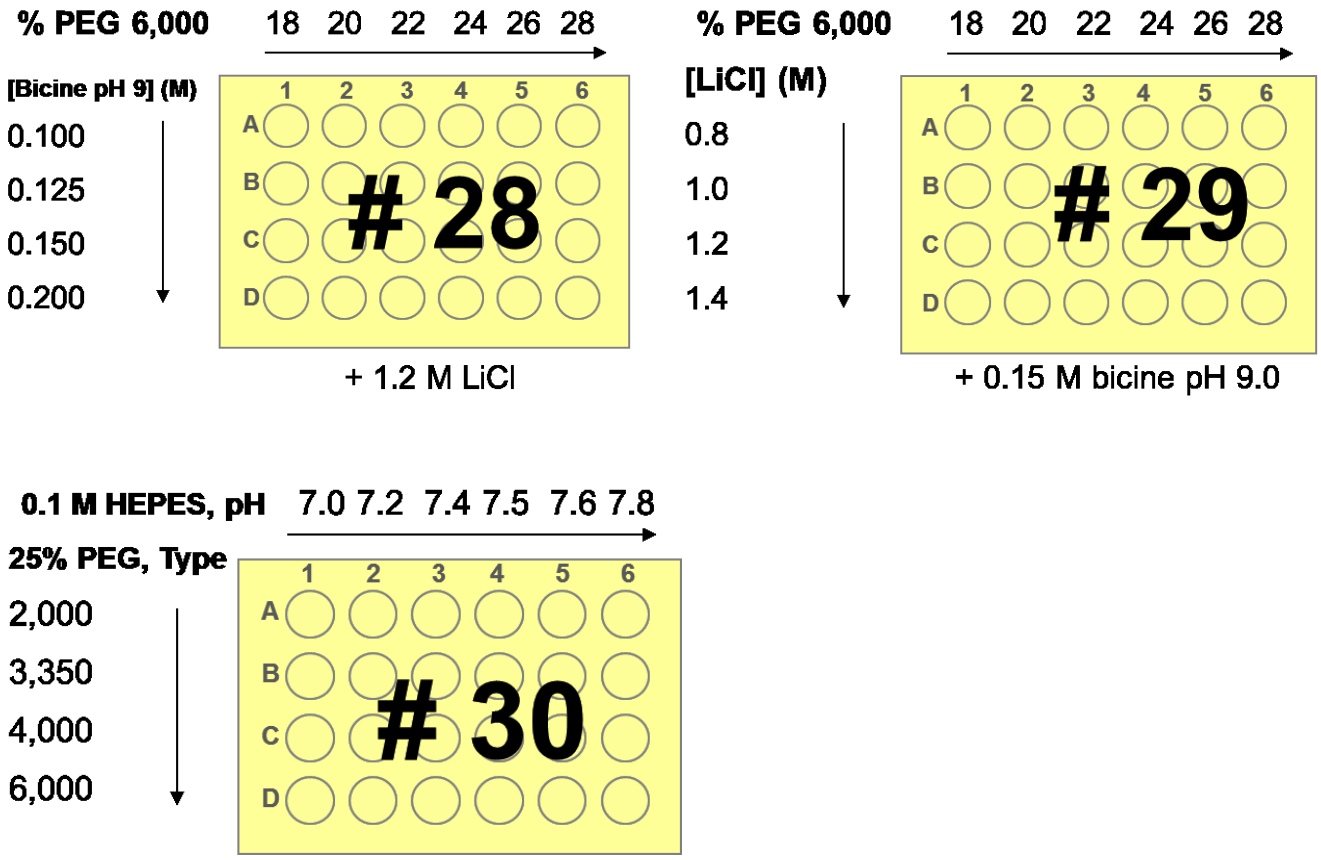

Figure 5.5.2. cont'd Optimization plates set up for crystallization of mRojo-VYGV. Protein was suspended in 20 mM Tris – HCl pH 7.4, 100 mM NaCl. 1 µl protein solution at 11 mg / ml was added to 1 µl of precipitant solution and the crystals were grown in sitting drops. The plates were kept at room temperature.

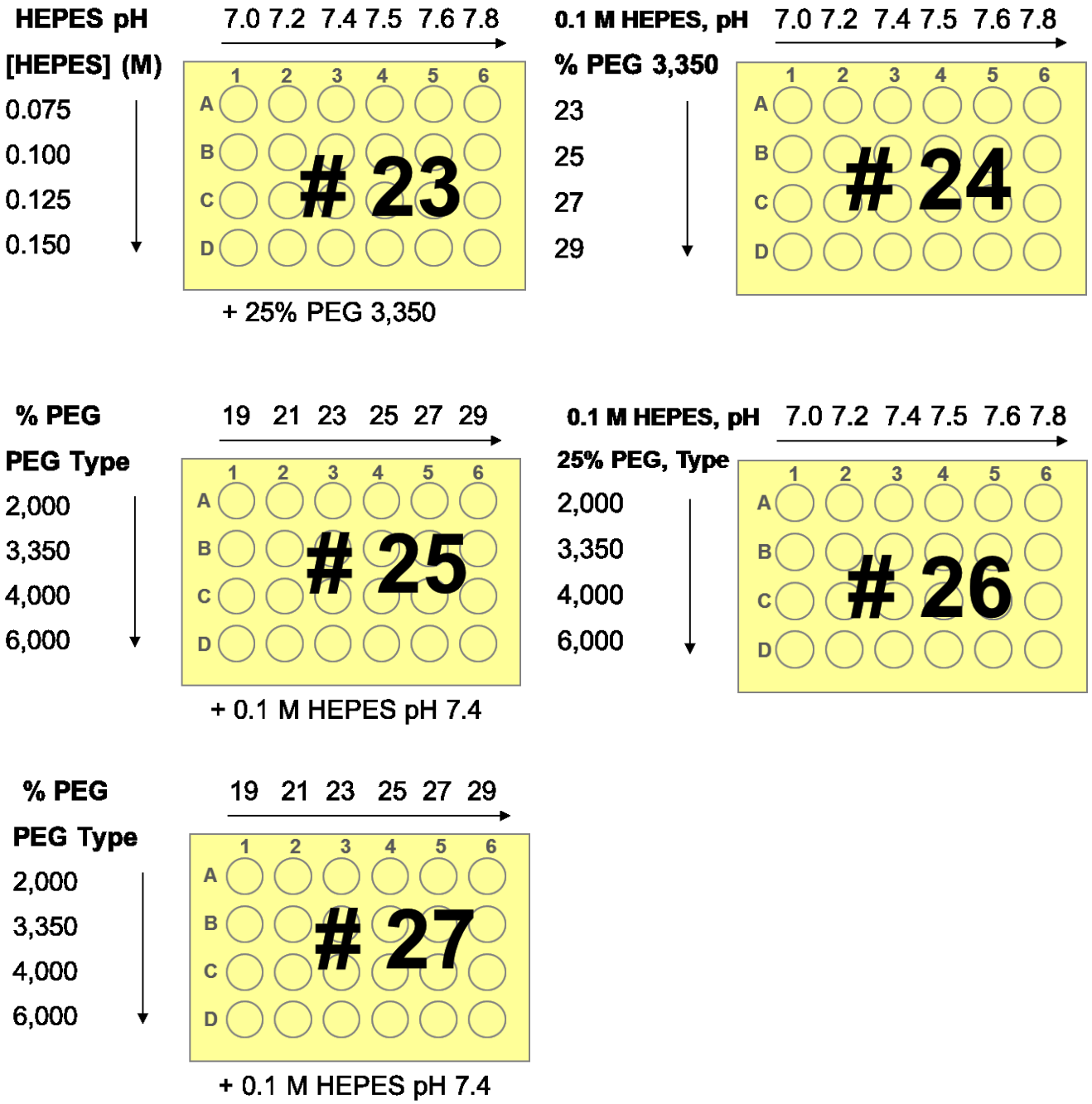

Figure 5.5.3. Optimization plates set up for crystallization of mRojo-VFAV. Protein was suspended in 20 mM Tris – HCl pH 7.4, 100 mM NaCl. 1 µl protein solution at 23 mg / ml (# 23 – 26) or 14 mg / ml (# 27) was added to 1 µl of precipitant solution and the crystals were grown in sitting drops. The plates were kept at room temperature.

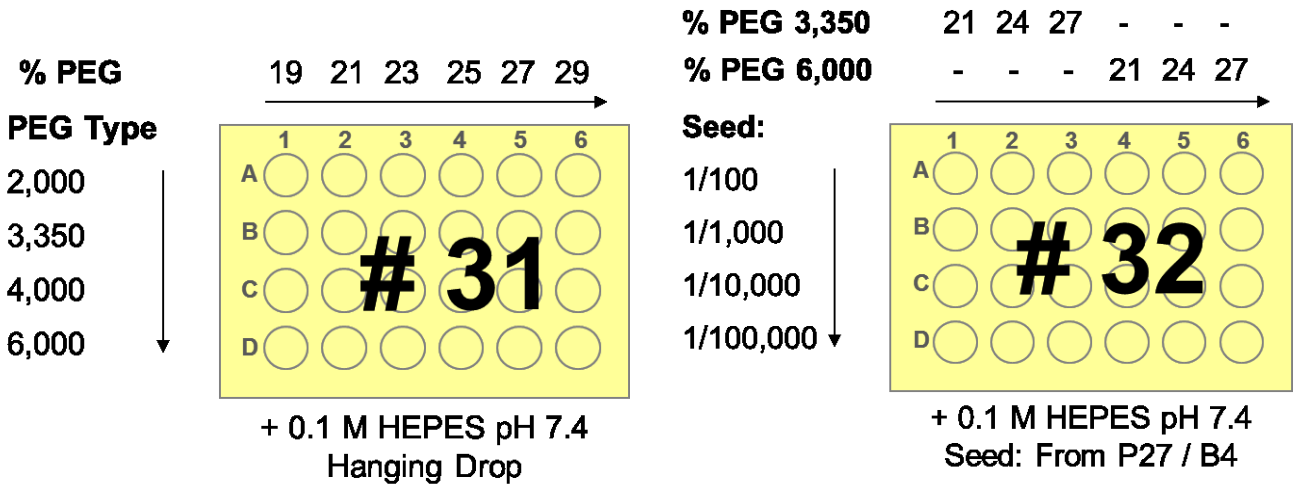

Figure 5.5.4. Special optimization plates set up for crystallization of mRojo-VFAV. Protein was suspended in 20 mM Tris – HCl pH 7.4, 100 mM NaCl. 1 µl protein solution at 4.2 mg / ml (# 23 – 26) or 14 mg / ml (# 27) was added to 1 µl of precipitant solution. The crystals on plate 31 were grown in hanging drops. The crystals in plate 32 were grown using a seeding technique. The seed was taken from plate 27, well B4 (P27 / B4) (crystal growth as cluster of purple needles), and added to the precipitant solution in varying degrees of dilution. The plates were kept at room temperature.

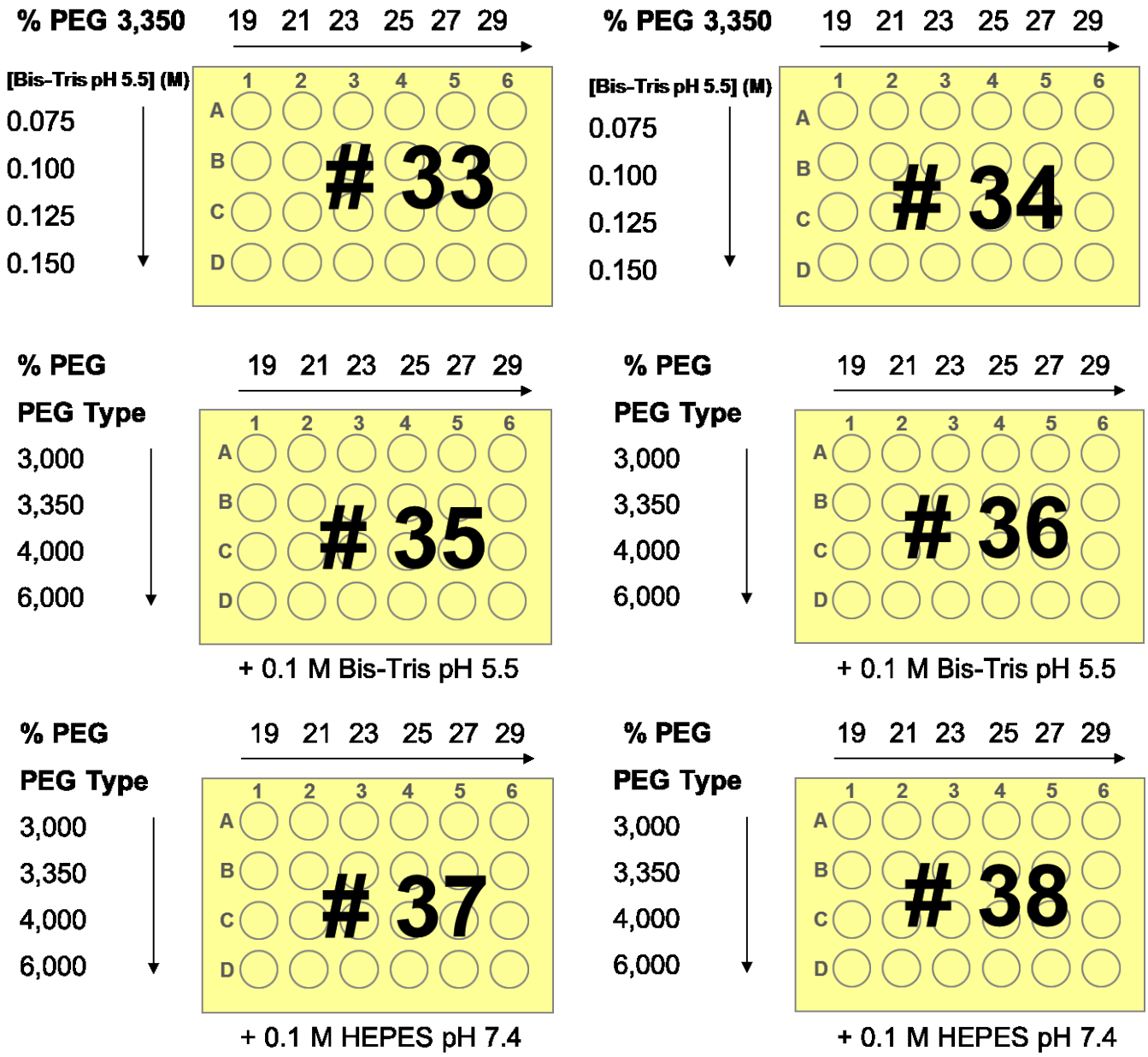

Figure 5.5.5. Optimization plates set up for crystallization of mRojo-VHSV. Protein was suspended in 20 mM Tris – HCl pH 7.4, 100 mM NaCl. 1 µl protein solution at 35 mg / ml (# 33, 35, 37) or 7.0 mg / ml (# 34, 36, 38) was added to 1 µl of precipitant solution and the crystals were grown in sitting drops. The plates were kept at room temperature.

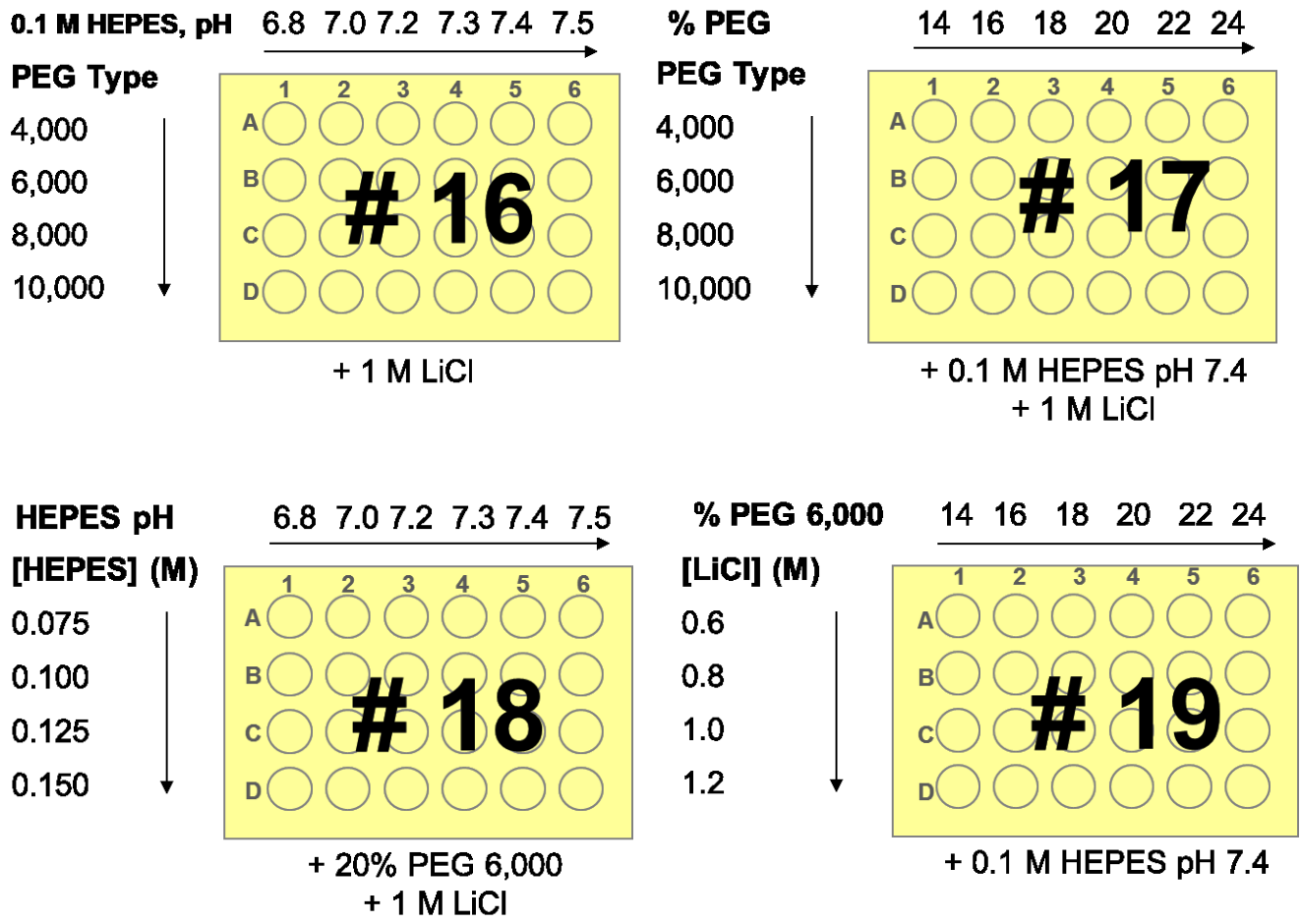

Figure 5.5.6. Optimization plates set up for crystallization of mCherry-VYGV. Protein was suspended in 20 mM Tris – HCl pH 7.4, 100 mM NaCl. 1 µl protein solution at 34 mg / ml was added to 1 µl of precipitant solution and the crystals were grown in sitting drops. The plates were kept at room temperature.

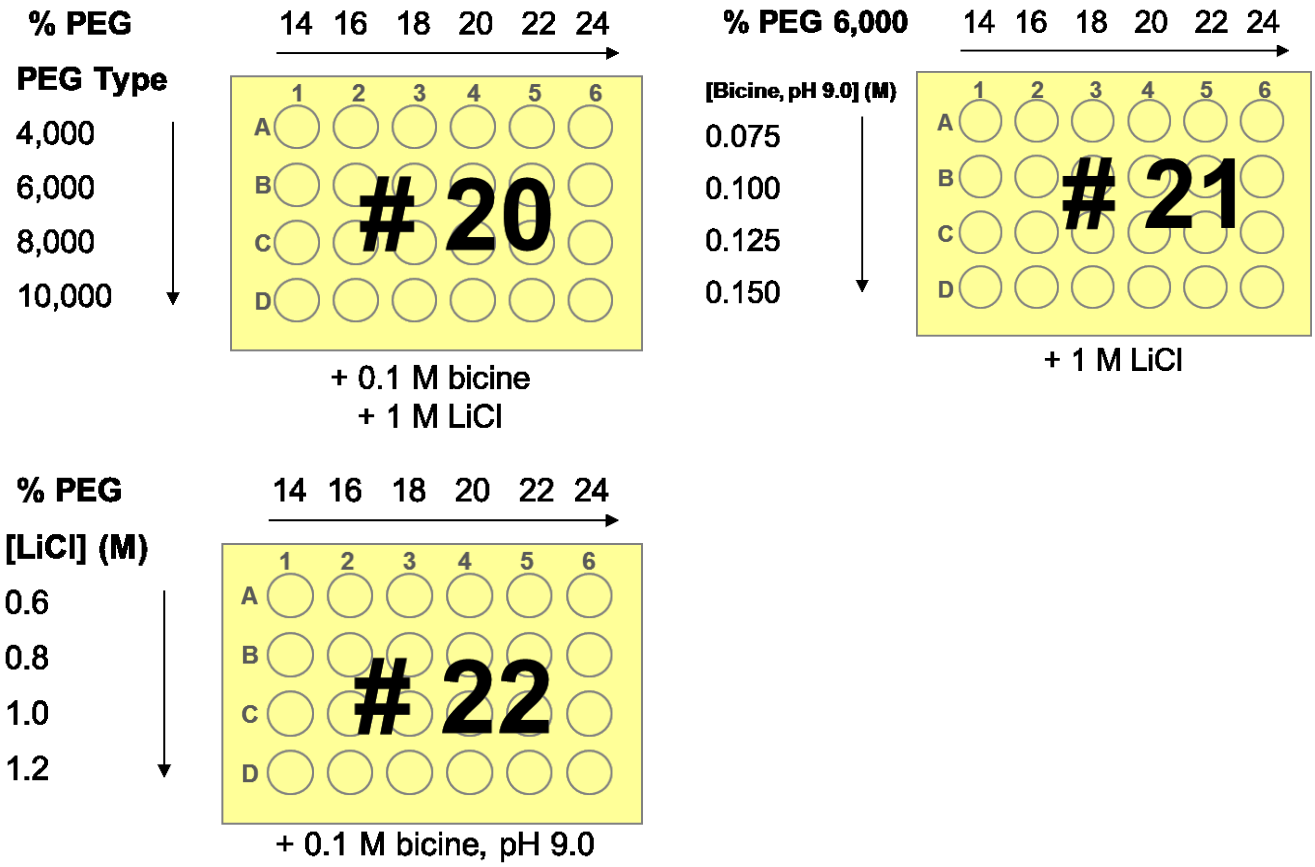

Figure 5.5.6. cont'd Optimization plates set up for crystallization of mCherry-VYGV. Protein was suspended in 20 mM Tris – HCl pH 7.4, 100 mM NaCl. 1 µl protein solution at 34 mg / ml was added to 1 µl of precipitant solution and the crystals were grown in sitting drops. The plates were kept at room temperature.

Table 5.5.1. Wells with crystal growth in screening and optimization plates of mutant RFPs.

Protein	Plate	Spheruloids	Needles	Plates	Small crystals	Large crystals	Bleached
mRojo-VYGV	1	A2, A5	A10, B8, C1, C5, D6, G1, G2, H7, H8	-	H11	-	-
	3	A7, C6, C9, C10, E4	E4	H9	E2, E9	-	-
	8	-	C2	A2, A3, B2, D4	B3	-	-
	9	-	B5	C3	-	A3	-
	12	-	*B5, B6*, *C5*, *C6, D5, D6*	A3, B4, D4	-	***C2***, *C3*	-
	13	-	*A6*	*A5*, B4, *B5*, *C5, C6*, D6	-	A2, *A3*	-
	14	-	-	A2, *C2*	-	*A1, B2*	-
	15	-	-	A1	-	*A2*, *B3*	-
	28	-	-	-	-	-	-
	29	-	-	-	-	-	-
	30	-	B1, B2, B3, B4, B5, B6	C1, C6	-	-	-
mRojo-THSL	2	D11	D1, D4, D6, D7,D8, D10, E7, F7, F11, F12, G7, G12, H1	H2	D7, E9, G4, G8	-	-
	4	C3	A2	D2	-	-	-
	5	-	A2, D1, D2, D3	-	A3, B2, C3	***C2***, D4	-
	6	-	C2, D4	-	-	-	-
	7	-	D2, D3	-	D4	-	-

All crystals had a colour typical of mature RFP chromophore, unless noted in the "Bleached" column. All crystal growth was in clusters, ranging from a few crystals to large tree – like growths. Wells that were fished for crystals are in *italics*. Crystals leading to crystal structures are in ***bold italics***.

Table 5.5.1. cont'd Wells with crystal growth in screening and optimization plates of mutant RFPs.

Protein	Plate	Spheruloids	Needles	Plates	Small crystals	Large crystals	Bleached
mRojo-VFAV	11	B5, B11, D3, E11, H8	A3	D5, D8	C11	-	-
	23	-	*C6, D3, D4*	-	A3, B2	A1, A2, B1, *C4*	-
	24	-	*C1, C4, B5*	-	*B4*	A1	-
	25	-	All (***B5***, *C5*)	-	-	-	-
	26	-	All (*C1, D2*)	-	-	-	-
	27	-	B1, *B2*, B3, B4, B5, B6, D1, D2, D3, D4, D5	C1, C2, C3, C4, C5	-	-	-
	31	-	B4, B5, C2, C3, C4, D3, D4	B2, B3, D1, D2	-	B6	A1 - A6
	32	-	A2, B1, B2, B3, B6, C2, C3, C4, C5, D5, D6	A1, C1, D1, D2, D3, D4	B4	B5	-
mRojo-VHSV	33	-	-	-	-	-	-
	34	-	-	-	-	-	-
	35	-	-	B5, C4, C5	-	-	-
	36	-	-	-	-	-	-
	37	-	B4, *B5*, *C3, C4*	B2, *C2, C5*, *D2, D3, D4*	*D4*	-	-
	38	-	-	D4, D5	-	-	-
	39	H8	D5, D6, D7, D8, E9, F12, G1, H11, H12	F10, F11, *G3*, G4, G6, G7, G8, G11, G12, H1	-	-	-

All crystals had a colour typical of mature RFP chromophore, unless noted in the "Bleached" column. All crystal growth was in clusters, ranging from a few crystals to large tree – like growths. Wells that were fished for crystals are in *italics*. Crystals leading to crystal structures are in ***bold italics***.

Table 5.5.1. cont'd Wells with crystal growth in screening and optimization plates of mutant RFPs.

Protein	Plate	Spheruloids	Needles	Plates	Small crystals	Large crystals	Bleached
mCherry-VYGV	10	A3, C9, C11, C12, D1, D12, F3	B4, G11	*D5*, *G10*	-	-	-
	16	-	A1, *C1*, D1, D2, D3, D4	-	-	-	D1 - D6
	17	-	-	C6	D3, *D4*	-	D3, D5, D6
	18	-	A2, B1, D1	-	*B2, C1*	-	-
	19	-	-	-	-	-	-
	20	-	-	-	-	-	-
	21	-	-	D6	-	-	-
	22	-	-	D6	-	-	-

All crystals had a colour typical of mature RFP chromophore, unless noted in the "Bleached" column. All crystal growth was in clusters, ranging from a few crystals to large tree – like growths. Wells that were fished for crystals are in *italics*. Crystals leading to crystal structures are in ***bold italics***.

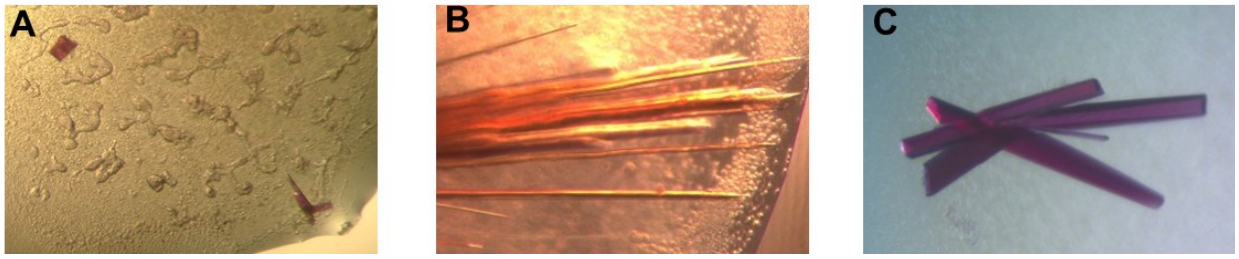

Figure 5.5.7. Photographs of protein crystals. mRojo-VHSV (A), mRojo-VFAV (B), and mRojo-VYGV (C) photographs taken in sitting drop solutions on the day of crystal harvest.

5.6. References

1. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction, *Gene 77*, 51-59.

2. Gross, L. A., Baird, G. S., Hoffman, R. C., Baldridge, K. K., and Tsien, R. Y. (2000) The structure of the chromophore within DsRed, a red fluorescent protein from coral, *Proceedings of the National Academy of Sciences of the United States of America 97*, 11990-11995.

3. Shaner, N., Campbell, R., Steinbach, P., Giepmans, B., Palmer, A., and Tsien, R. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from discosoma sp red fluorescent protein, *Nature Biotechnology 22*(12), 1567-1572. doi:10.1038/nbt1037

4. Wang, L., Jackson, W. C., Steinbach, P. A., and Tsien, R. Y. (2004) Evolution of new nonantibody proteins via iterative somatic hypermutation, *Proceedings of the National Academy of Sciences of the United States of America 101*(48), 16745-16749. doi:10.1073/pnas.0407752101

5. Williams, A. T. R., Winfield, S. A., and Miller, J. N. (1983) Relative fluorescence quantum yields using a computer – controlled luminescence spectrometer, *Analyst 108*, 1067 - 1071. doi: 10.1039/AN9830801067

6. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein – dye binding, *Analytical Biochemistry 72*(1-2), 248 – 254. doi: 10.1016/0003-2697(76)90527-3

7. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) Measurement of protein using bicinchoninic acid, *Analytical Biochemistry 150*(1), 76 – 85. doi: 10.1016/0003-2697(85)90442-7

8. Kredel, S., Nienhaus, K., Oswald, F., Wolff, M., Ivanchenko, S., Cymer, F., Jeromin, A., Michel, F. J., Spindler, K.-D., Heilker, R., Nienhaus, G. U., and Wiedenmann, J. (2008) Optimized and Far-Red-Emitting Variants of Fluorescent Protein eqFP611, *Chemistry & Biology 15*, 224-233.

9. Britton, H. T. S., and Robinson, R. A. (1931) CXCVIII.-Universal buffer solutions and the dissociation constant of veronal, *Journal of the Chemical Society (Resumed)*, 1456-1462.

10. Shu, X., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y., and Remington, S. J. (2006) Novel chromophores and buried charges control color in mFruits, *Biochemistry 45(32*), 9639-9647. doi: 10.1021/bi0607731

11. Chica, R. A., Moore, M. M., Allen, B. D., and Mayo, S. L. (2010) Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries,

Proceedings of the National Academy of Sciences of the United States of America 107, 20257-20262.

12. Pletnev, S., Shcherbo, D., Chudakov, D. M., Pletneva, N., Merzlyak, E. M., Wlodawer, A., Dauter, Z., and Pletnev, V. (2008) A crystallographic study of bright far-red fluorescent protein mKate reveals pH-induced cis-trans isomerization of the chromophore, *Journal of Biological Chemistry 283*(43), 28980-28987. doi:10.1074/jbc.M800599200