University of Alberta

Development of fluorogenic fluorescent protein heterodimers

by

Spencer C. Alford

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

Fall 2012
Edmonton, Alberta

Abstract

Fluorescent proteins (FPs) are indispensible biochemical tools. The concerted efforts of protein engineers have produced FPs spanning the visible colour spectrum. This wide variety of FPs has greatly facilitated the development of FP-based biosensors. However, researchers rely on relatively few fundamental biosensor design templates. Förster resonance energy transfer and bimolecular complementation are the principal FP-based technologies for live cell imaging of physiological events, such as changes in small molecule concentration, enzymatic activities, and protein-protein interactions. Although widely used, these techniques are often restrictive due to poor signal-to-noise ratios and irreversible sensing, respectively. Furthermore, examples of these biosensor strategies incorporating red FPs are limited.

In this thesis we describe our efforts to address this shortcoming in the area of FP-based biosensors. We developed a dimerization-dependent red FP (ddRFP) that serves as an alternative template for biosensor construction. The prototype ddRFP was engineered from a homodimeric variant of a *Discosoma* red FP. Through extensive directed evolution the homodimer was converted into a fluorogenic obligate heterodimer. The reversible changes in fluorescence intensity that result from association of the ddRFP monomeric constituents, or the irreversible decrease that accompanies dissociation of covalently linked partners following linker cleavage, provides a useful spectroscopic signal for biosensing applications. Specifically, we demonstrated that ddRFP is useful for detecting *in vitro* protein-protein interactions, as well as imaging changes in calcium ion concentration and activation of caspase-3 in live cells.

We also report the expansion of the ddFP colour palette through the development of green (ddGFP) and yellow (ddYFP) ddFP variants. These variants have several improvements relative to the ddRFP prototype including increased *in vitro* contrast and brightness for ddGFP, and a reduced pK_a for ddYFP. While their utility for some live cell imaging applications is restricted due to low dissociation constants, ddGFP proved to be a useful fluorescent label of intermembrane contact sites between the endoplasmic reticulum and the mitochondrial network.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

ε	extinction coefficient
Φ	quantum yield
n_H	Hill coefficient
Å	Angstrom
A	acceptor fluorescent protein
AC	refers to A and C protomeric subunits of DsRed
AbeGFP	*Abylopsis eschscholtzii* Green fluorescent protein
Asp	aspartate
Arg	arginine
A. victoria	*Aequorea victoria*
BFP	blue fluorescent protein
BiFC	bimolecular fluorescence complementation
β-ME	β-mercaptoethanol
Bn	barnase
bp	base pairs
B-PER	bacterial protein extraction reagent
BRET	Bioluminescent resonance energy transfer
BS	barstar
CalN	calnexin
CaM	calmodulin
Casp-3	caspase-3
cm	centimeter
CID	chemical inducer of dimerization
CFP	cyan fluorescent protein

cpGFP	circularly permuted green fluorescent protein
CyPet	cyan fluorescent protein for energy transfer
D	donor fluorescent protein
ddFP	dimerization-dependent fluorescent protein
ddFP-A	dimerization-dependent fluorescent protein copy A
ddFP-B	dimerization-dependent fluorescent protein copy B
ddGFP	dimerization-dependent green fluorescent protein
ddRFP	dimerization-dependent red fluorescent protein
ddYFP	dimerization-dependent yellow fluorescent protein
DEVD	aspartate-glutamate-valine-aspartate
DHFR	dihydrofolate reductase
DIC	differential interference contrast
DMEM	Dulbecco's modified eagle media
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates
DsRed	Discosoma species Red fluorescent protein
dTFP0.7	dimeric teal fluorescent protein variant 0.7
dTomato	dimer Tomato fluorescent protein
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol-bis(2-aminoethylether)-*N,N,N′,N′*-tetraacetic acid
ER	endoplasmic reticulum
FACS	fluorescence activated cell sorting
Farn	farnesyl
FKBP12	FK506 binding protein 12

fMtx	fluorescein-conjugated methotrexate
FP	fluorescent protein
FRAP	FK506 binding protein rapamycin-associated protein
FRET	Förster resonance energy transfer
FRB	FK506 binding protein- rapamycin binding
GDP	guanosine diphosphate
GFP	green fluorescent protein
Gln	glutamine
Gly	glycine
GTP	guanosine triphosphate
GTPase	guanosine triphosphatase
HeLa	cervical cancer cell line originating from Henrietta Lacks
HHBSS	HEPES-buffered Hank's balanced salt solution
His	histidine
Hrs	hours
HTS	high throughput screening
Ile	Isoleucine
IPTG	isopropyl β-D-thiogalactopyranoside
kb	kilobases
kDa	kilodalton
K_d	dissociation constant
λcI	lambda repressor protein
LB	Luria Bertani
mDHFR	murine dihydrofolate reductase
MAM	mitochondria-associated ER membrane
MBP	maltose binding protein

MBSU	Molecular Biology Services Unit
mg	milligram
mL	milliliter
mM	millimolar
mRNA	messenger ribonucleic acid
mTFP1	monomeric teal fluorescent protein version 1
mTomato	monomeric Tomato
Mtx-SLF	methotrexate-synthetic analogue of FK506
MW	molecular weight
ng	nanogram
nm	nanometer
nM	nanomolar
NTA	nitrilotriacetic acid
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
pDES	dual expression system (plasmid)
PH	pleckstrin homology
Phe	phenylalanine
PINTS	protein interaction mediated toxin silencing
PIP3	phosphatidylinositol (3,4,5)-triphosphate
pmol	picomolar
PPI	protein-protein interaction
RET	resonance energy transfer
RFP	red fluorescent protein

RBD	Ras binding domain
SASG	Serine-alanine-serine-glycine
Ser	serine
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPR	surface plasmon resonance
Sup	suppressor
TMP	trimethoprim
Tyr	tyrosine
µg	microgram
µM	micromolar
UV	ultraviolet
WT	wild type
YFP	yellow fluorescent protein
YPet	yellow fluorescent protein for energy transfer

Chapter 1 General Introduction[1]

1.1 Overview and Premise

Molecular fluorescence is an invaluable component of life science research and now more than ever, fluorescence-based techniques are widespread throughout the scientific community. Although it has affected many fields of study, fluorescence has dramatically impacted the field of real-time imaging of biomolecules in live cells. More specifically, 'genetically encoded fluorescence' has revolutionized the ability of researchers to monitor protein localization, bio-analyte flux, enzymatic activities, and the proximity or physical interaction of biomolecules.

The key molecules that enable genetically encoded fluorescence are fluorescent proteins (FPs). FPs can be defined in simple terms as chemical fluorophores; they absorb and emit visible light. This trait is conferred by an intrinsic chromophore that forms autonomously from a tripeptide encoded by the FP gene [1]. Therefore, the simple introduction of a gene encoding an FP into a live cell bestows that cell with the ability to fluoresce.

FPs are particularly useful when they are tethered to other proteins of interest. This is achieved by means of a simple genetic manipulation in which an FP-encoding gene is linked to a gene that codes for a target protein (Figure 1.1A). Expression of this gene fusion permits spatial tracking of the protein-of-interest (Figure 1.1B).

[1] A version of this chapter has been submitted for publication. Alford, S.C., et al. 2012 *Biology of the Cell*.

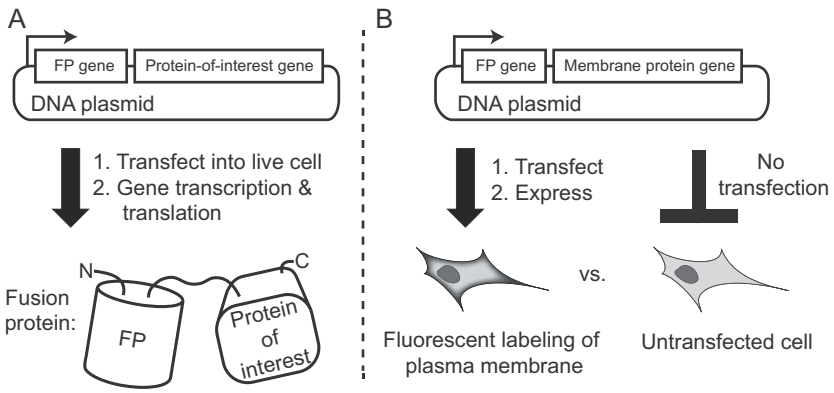


Figure 1.1. Schematic representation of FP reporter fusions. (A) An FP gene is linked to a gene encoding a protein-of-interest. The product is a fusion protein in which the protein-of-interest is tagged with the FP. (B) Example of an FP localization reporter targeted to the plasma membrane.

A significant problem with FP fusions is the oligomerization state of some FPs. Oligomerization is problematic because it may perturb the natural physiological function of its fused protein-of-interest by forming aggregates (Figure 1.2) [2-4]. These higher order structures may decrease cell viability or lead to erroneous localization patterns [3, 5-7]. Although some clever strategies have been used to cope with oligomerization [3, 8, 9], generating monomeric variants is the preferred solution (Figure 1.2) [10]. Hence, it is often essential during the directed evolution of FPs to incorporate monomerization strategies in which the protein-protein interfaces that mediate oligomerization are broken [9, 11]. Directed evolution is the cyclical process of generating a library of protein variants and screening for a desired phenotype [12]. In most cases, oligomerization proves to be crucial for development of fluorescence, and, therefore, directed evolution must be used to rescue fluorescence after interface disruption [9, 11]. The concept of oligomerization-dependent fluorescence is an intriguing phenomenon that has yet to be fully explored by protein engineers.

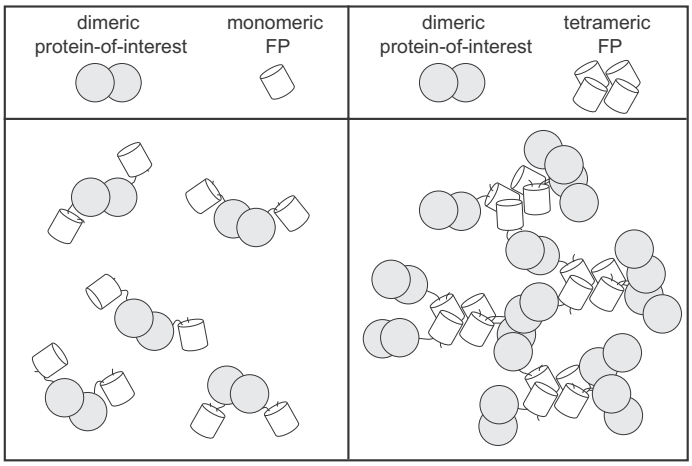


Figure 1.2. Monomeric FPs are preferable to oligomeric FPs as fusion tags. The left panel depicts a dimeric protein-of-interest tagged with a monomeric FP. The right panel shows the aggregates that result from tagging a dimeric protein-of-interest with a tetrameric FP.

The number of FP variants that have resulted from directed evolution of naturally occurring FPs is remarkable. In fact, there are so many different FPs available today it is now most appropriate to discuss them in terms of class and colour hue, rather than a simple non-descript reference to their emission wavelength [13]. Despite this impressive array of variants, researchers in the field of FP technology still rely on relatively few core strategies to design biosensors for live cell imaging applications. For example, the detection of protein-protein interactions (PPIs) in live cells using FP-based methods primarily relies on Förster resonance energy transfer (FRET) [14, 15] and bimolecular fluorescence complementation (BiFC) [16, 17] (these techniques will be discussed in sections 1.3.7.1 and 1.3.7.2, respectively).

For the research described in this thesis, we attempted to address this deficiency by using protein engineering to generate a new biosensor template. Our approach centers on utilizing oligomerization-dependent FP fluorescence as a beneficial attribute rather than a deleterious one. For the remainder of the

introductory chapter, we present a brief overview of the essential concepts necessary to address this challenge. Topics to be reviewed include: protein-complementation assays, the properties and chemistry of FPs, the design and applications of FP-based biosensors, and protein engineering strategies.

1.2 Detection of protein-protein interactions by protein-complementation

Since the ground-breaking description of the yeast two-hybrid assay (Figure 1.3) [18], many genetic systems have been developed to detect PPIs [19]. Applications of these assays include validating hypothesized protein interactions [20], screening genomic libraries for PPIs [21], dissecting the molecular complementarity of PPI interfaces [22], and screening for inhibitors of PPIs [23]. Commonly used two-hybrid style assays rely on the same general strategy: a PPI drives the refolding or reconstitution of an effector protein, or effector protein complex [24]. The effector may function directly as an assayable reporter (Figure 1.4), or alternatively, trigger a reporter output to indirectly indicate the occurrence of the PPI (as in the yeast-two hybrid assay) (Figure 1.3). Prominent examples of reconstitution-mediated reporter systems include assays based on dimerization of the LexA repressor [25], split FPs [26], split ubiquitin [27], split luciferase [28], split β-lactamase [29], and split dihydrofolate reductase (DHFR) [30, 31]; see [32] and [33] for reviews.

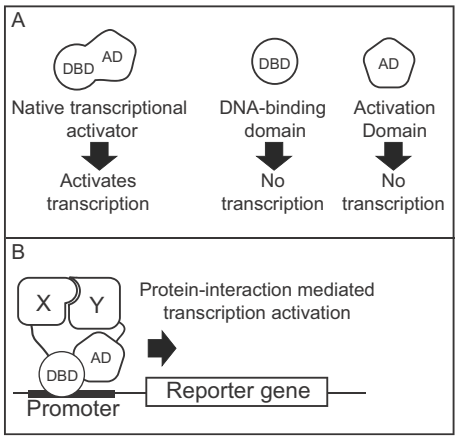


Figure 1.3. Schematic summary of the yeast-two hybrid assay. (A) A transcriptional activator protein contains a DNA-binding domain (DBD) and a transcription activation domain (AD). Independent DBDs and ADs are not capable of initiating transcription. (B) The fusion of the DBD and the AD to interacting protein partners, X and Y, mediates reconstitution of transcriptional activity and expression of a reporter gene.

Each of these strategies relies on genetic systems deployed in live organisms (summarized in Figure 1.4). The genes encoding for the proteins being tested (often designated as ‘bait’ and ‘prey’) are fused to the genes of specific reporter elements. The gene hybrids are then introduced (by transformation or transfection) into a suitable host organism to express the protein hybrids. Bacteria or yeast are typically used as hosts, but mammalian systems have also been utilized [19, 34]. Once expressed in the host, the reporter phenotype is assayed, or screened, to determine if a PPI occurred between the protein hybrids. Examples of reporter outputs, or phenotypes, include survival, fluorescence, or development of colour [32].

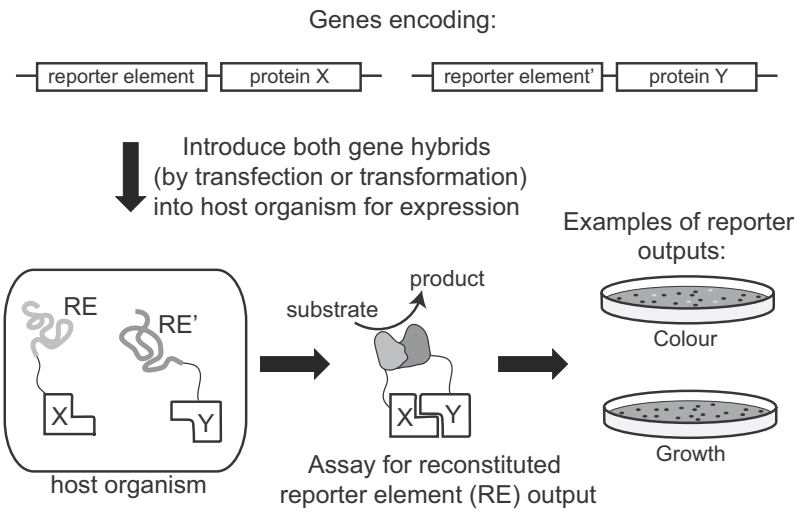


Figure 1.4. Schematic representation of a genetic system used for detection of protein-protein interactions. Gene hybrids of non-functional reporter element (RE) fragments and interacting proteins partners (X and Y) are introduced into a host organism. Interaction of X and Y reconstitutes the reporter and restores its function. The reporter output is then assayed to indicate the occurrence of the interaction between X and Y.

Reconstitution-based reporter systems utilize either genetic screens or selections. Screening involves analyzing a large population of variants, known as a library, for changes in phenotype [35, 36]. Therefore, this type of assay requires a suitable “spatially addressable format” or physical separation to distinguish one variant from another [37, 38]. For example, several hundred bacterial colonies can be screened for a desired phenotype (*e.g.*, colony colour or fluorescence intensity) on a single agar dish, or hundreds of cell lysates can be screened on a multi-well plate (*e.g.*, for enzyme activity). This technique is thus limited in regards to number of variants that can be practically screened [36]. Genetic selection assays, on the other hand, rely on selectable mutations that confer individual variants with a growth advantage under user-defined growth conditions [35]. Therefore, selection is particularly powerful because it permits single mutants exhibiting the selectable trait to be isolated from very

large libraries without the need to analyze each independent variant [36]. For this reason, genetic selections are often preferable to screens.

Genetic selection assays for PPIs generally resemble classic auxotroph complementation experiments [35]. That is, a metabolic deficiency, such as deletion of a gene whose protein product produces an essential metabolite, prevents growth of a host organism under nutrient-limiting conditions. Genetic mutants that exhibit metabolic-competency (*i.e.,* the ability to synthesize their own nutrients) have a growth advantage and are able to survive. This concept has been applied to detect PPIs using protein complementation assays (PCA) utilizing split metabolic enzymes and metabolically compromised host organisms (summarized in Figure 1.5).

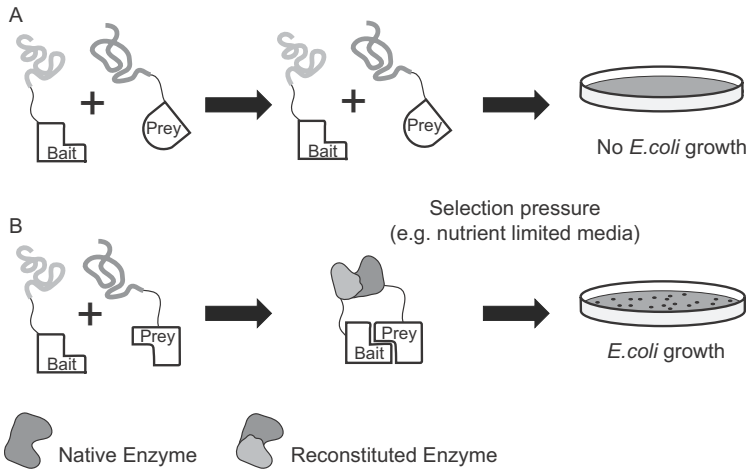


Figure 1.5. General strategy for protein complementation assays in *E. coli*. Bait/prey interactions drive the reconstitution of unfolded metabolic enzyme fragments (e.g. DHFR). The reassembled enzyme confers metabolic competency that is otherwise deficient under selective pressure conditions. The growth of the host organism is thus dependent on the interaction of bait and prey.

In PCA selection assays, split unfolded enzyme fragments are fused to bait and prey proteins that may or may not interact. If an interaction occurs between bait and prey, the split fragments reconstitute and adopt the native enzyme fold

to complement the metabolic deficiency and promote growth. Two examples of metabolic PCA selection assays include split DHFR [30, 31] and split cytosine deaminase [22, 39] (see [34] and [40] for additional examples), but the former is the most widely implemented metabolic PCA strategy and is used here to illustrate the concept.

In vivo DHFR PCA selection assays have been deployed for detecting PPIs in bacteria, yeast, plants, and mammalian cells [30, 31, 41, 42]. Metabolic nuances inherent to each organism type necessitate slightly different strategies to mediate selection. For the purpose of this thesis, we restrict our discussion to implementation of DHFR in bacteria.

DHFR catalyzes the synthesis of tetrahydrofolate, an essential precursor for amino acids, nucleotides, nucleotide bases, and vitamins [31]. The enzyme can be inhibited by a folic acid analogue, trimethoprim (TMP). It follows that *E. coli* cannot grow on nutrient-depleted minimal media in the presence of this drug. However, murine DHFR (mDHFR) can complement growth because its affinity for TMP is much lower than that of *E. coli* DHFR [30, 43]. DHFR PCA selection utilizes split mDHFR fragments fused to bait and prey proteins. If an interaction occurs, mDHFR is reconstituted and catalyzes the synthesis of essential metabolites. As a result, survival is permitted even in the context of TMP-inhibited *E. coli* DHFR [30].

In addition to metabolic selection, alternative direct reporter PCA assay strategies utilize genetic screens using split enzymes (although exogenous fluorogenic or colorimetric substrates are required), or split FPs. The latter approach is often referred to as bimolecular fluorescence complementation (BiFC) and will be reviewed in the discussion of FP-based biosensor strategies (section 1.3.7.2). DHFR and β-lactamase are examples of split enzyme

platforms for direct reporter PCA screens [29, 44-48]. These PCA screens can be deployed *in vitro* or *in vivo*. For example, cleavage of the *in vitro* chromogenic substrate nitrocefin by reconstituted β-lactamase causes a distinct colour change from yellow to red [29, 47]. Split β-lactamase can also be used *in vivo* in mammalian cells using a fluorogenic substrate called CCF2/AM (coumarin cephalosporin fluorescein/acetoxymethylated) [29, 48]. This particular substrate is advantageous in that it provides a ratiometric response [48]. As an alternative, reassembled DHFR and fluorescein-conjugated methotrexate (fMtx) may be used to detect the occurrence of PPIs in live cells [45, 46]. Methotrexate is a folate analogue that binds stoichiometrically to reassembled DHFR. PPI-driven DHFR reassembly can thus be detected using fMtx [45].

1.3 Fluorescent proteins

1.3.1 Primary sequence

The sequencing of the gene encoding green fluorescent protein (GFP), and its corresponding amino acid sequence, was the key breakthrough that enabled FP technology development. The isolated GFP gene encodes a 238 amino acid polypeptide of calculated mass ~27 kDa. The key feature of the sequence is a tripeptide (serine-65-tyrosine-66-glycine-67), which post-translationally forms a chromophore.

Strong amino acid conservation is observed at four positions in all FPs. These residues correspond to tyrosine-66, glycine-67, arginine-96, and glutamate-222 (GFP numbering) [49]. Although substitutions of these residues may be tolerated and confer different emission hues [50], they are essential to

fluorogenesis in FPs [51]. Glycine-67 is of particular importance and demonstrates absolute conservation [1, 52].

1.3.2 Chromophore formation

The chromophore of GFP and GFP-like proteins is formed in an autogenic process (Figure 1.6). The polypeptide chain inherently contains all the functionalities required for assembling the chromophore with the exception of molecular oxygen. The process involves four key steps: (1) pre-organization of chromophore-forming residues, (2) cyclization, (3) oxidation, and (4) dehydration. The precise mechanism of chromophore formation has remained an open question for many years due to conflicting results from multiple studies [53-59]. However, a series of experiments performed on the chromophore of DsRed (a well characterized *Discosoma* species red FP) provides the best current model for chromophore formation [60].

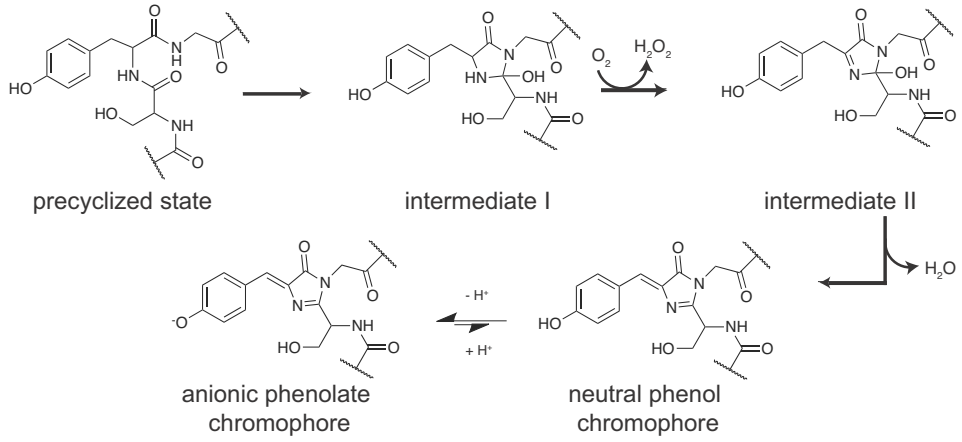


Figure 1.6. Mechanism of GFP chromophore formation.

Chromophore formation initiates with the generation of a reaction-competent state (Figure 1.6). A distortion of the polypeptide backbone positions the serine carbonyl carbon into close proximity to the glycine amide nitrogen in the precyclized state [61]. The glycinyl nitrogen attacks the adjacent serine

carbonyl to render an unstable imidazolinone ring (intermediate I) which is subsequently oxidized [60]. Some evidence suggests this oxidation proceeds through an α-enolate [55, 59]. The unstable ring product (intermediate II) is then quickly trapped by dehydration to form the phenol chromophore. The tyrosyl phenol is deprotonated to render the anionic chromophore [60].

The mechanism of DsRed chromophore formation begins on the GFP pathway, except a branch point occurring at the hydroxylated cyclic imine (intermediate II) leads to the red emitting chromophore [59, 60] (Figure 1.7). Intermediate II (Figures 1.6 and 1.7) is in rapid equilibrium with a cyclic imine (intermediate III), which can be oxidized to form a short lived blue intermediate (intermediate IV). The blue species undergoes irreversible hydroxylation (to form intermediate V) and is dehydrated to form the phenolic form of red chromophore [60]. Final deprotonation removes the phenolic proton to render the final anionic red emitting species.

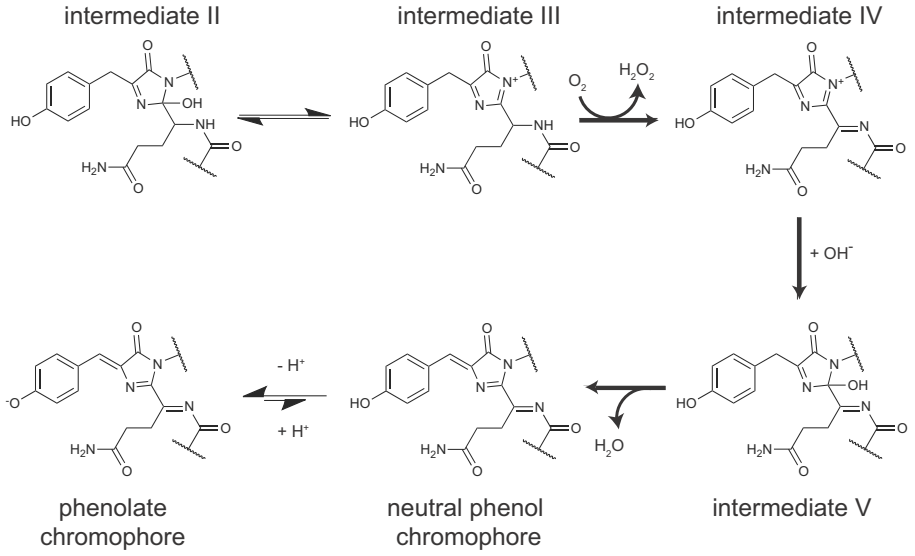


Figure 1.7. Mechanism of DsRed chromophore formation. A branch point occurs in the GFP chromophore formation pathway (Figure 1.6) at intermediate II, which initiates DsRed chromophore formation. Note that residue 65 in DsRed is glutamine as opposed to serine in GFP.

1.3.3 Three-dimensional structure

The optical attributes of FPs are ultimately conferred by the chromophore structure and environment, which is produced by their unique three-dimensional structure. In 1996, two groups reported the crystal structure of GFP [62, 63]. GFP adopts an eleven-stranded β-barrel can-like structure with a central helix penetrating its core. This central helix contains the tripeptide that forms the chromophore. The eleven β strands form a near perfect cylinder with a length of ~42 Å and cross-section of ~24 Å (Figure 1.8) [63]. For this reason, FPs are depicted schematically as cartoon cylinders in figures throughout this text. The chromophore is buried in the core of the barrel and is isolated from surrounding solvent. It is oriented approximately perpendicular to the longitudinal axis of the barrel [62, 63].

FPs originating in corals share a homologous three-dimensional structure with Aequorea victoria GFP. For example, the Anthozoan FP DsRed differs from GFP by an average root mean square of Cα atoms of only 1.9Å [64]. Given their highly homologous structures, it is somewhat surprising that coral FPs only exhibit a sequence identity with GFP in the range of 20% to 30% [49, 65].

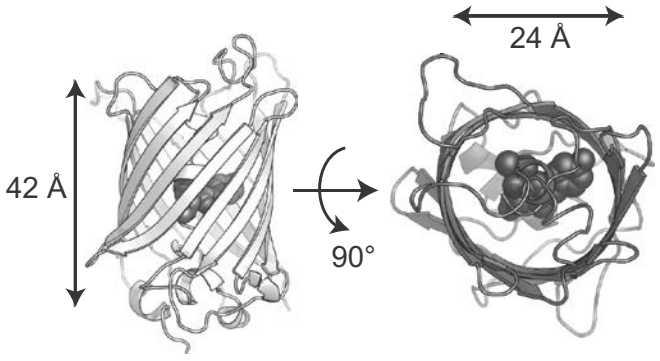


Figure 1.8. Three-dimensional structure of FPs. Shown is a rendering of the GFP β-barrel (PDB ID 1GFL [62]). Structure rendered with PyMol (www.pymol.org) [66].

1.3.4 Oligomeric Structure

All naturally occurring FPs demonstrate some extent of self-association or oligomerization, typically exhibiting dimerization or tetramerization [67]. *Aequorea victoria* GFP*,* however, is only weakly dimeric; it exhibits a weak propensity for dimerization with a K_d estimated to be 60 - 100 μM [68-70]. Indeed, in one of the two initial crystal structures of GFP, the protein was revealed to exist as a dimer [62]. Biochemically, the GFP dimer interface is largely mediated by hydrophilic contacts, with only three core hydrophobic amino acid contacts [62]. Where problematic for biological applications, this interface is readily broken by mutating alanine-206 to lysine [69].

The 1999 discovery of several FPs in non-bioluminescent Anthozoan reef corals enabled red-hued FPs to be used as biological markers [49]. Unfortunately, it was soon discovered that these proteins formed oligomers [71]. Various analyses carried out with the best-characterized Anthozoan FP, DsRed, illustrated an obligate tetrameric nature [71-73]. Furthermore, X-ray crystal structures clearly demonstrate the obligate tetrameric nature of Anthozoan FPs [64, 74-77].

The crystallographic studies of DsRed can be used to highlight the nature of tetramer formation. Its structure can be described as a dimer of homodimers in which each protomer has two chemically distinct interfaces (Figure 1.9). Each protomer has both a hydrophobic and hydrophilic interface [64, 74]. The hydrophobic interface is characterized by a cluster of hydrophobic amino acids encircled by several residues with polar side chains [64, 74]. The hydrophilic interface is characterized by polar amino acids that mediate hydrogen bonds and charged amino acids that form electrostatic ion-pair interactions across the interface. Interestingly, a hydrogen bond network appears to couple the

chromophores of the protomeric units forming this interface [64]. Finally, four residues at the C-termini of the protomers form a ‘clasp-like’ association [64, 74].

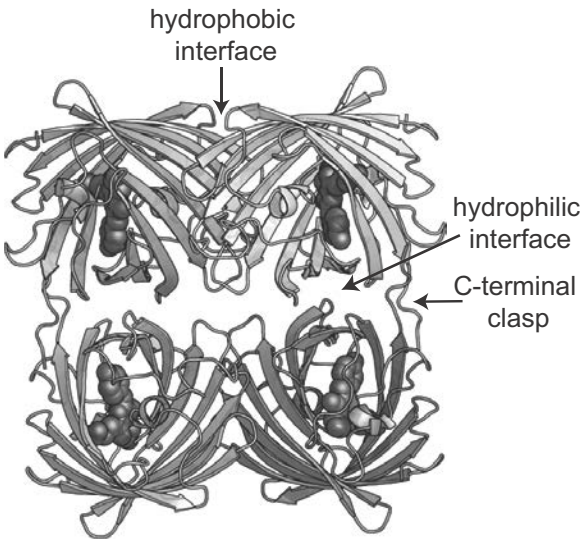


Figure 1.9. Obligate tetrameric structure of Anthozoan coral FPs. Shown is the crystal structure of DsRed (PDB ID 1GGX [64]). Structure rendered with PyMol (www.pymol.org) [66].

Sequence alignment analysis of Anthozoan FPs has shown interface residues are highly conserved [78], which further explains the frequent tetramerization observed in FPs of this order. The question thus arises as to why the FPs of coral form obligate tetramers. Oligomerization is clearly important to fluorescence development, as disruption of the interfaces by mutation perturbs folding and fluorescence intensity, chromophore maturation, and resistance to photobleaching [9, 79]. It has also been suggested that tetramerization evolved to increase photostability and thermotolerance, which would be beneficial to a reef coral living in environments with severe sun exposure [78].

1.3.5 Chromophore modulation and colour diversity

FPs are typically classified according to their remarkable colour diversity. Naturally occurring FPs from marine jellyfish and corals evolved to exhibit several different colours and protein engineers have elaborated on these templates to generate many additional colour hues. FPs can be subcategorized into six colour classes based on the emissions associated with their chromophores: blue FPs, 440-470 nm; cyan FPs, 470-500 nm; green FPs, 500-520 nm; yellow FPs, 520-550 nm; orange FPs, 550-575 nm; and red FPs 575-610 nm. Figure 1.10 shows chromophore structures of common FPs from each of the aforementioned classes.

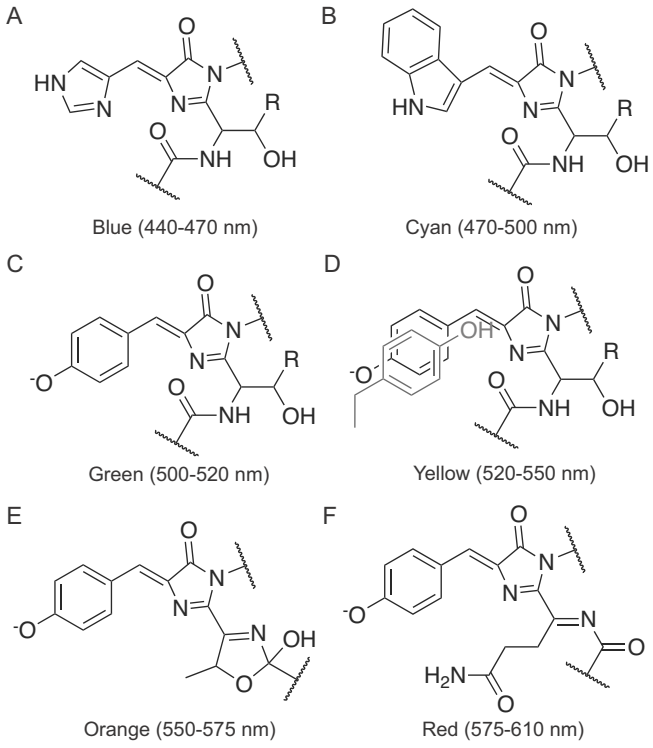
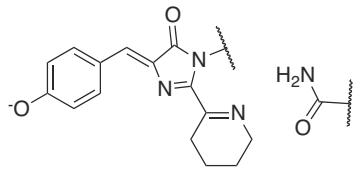


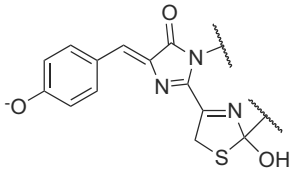
Figure 1.10. Chromophores of various FP colour classes. Indicated are the general colour classes and their corresponding emission wavelengths.

Several additional chemically distinct chromophores have been discovered in nature (e.g., zFP538) or engineered in the laboratory. Examples are illustrated in Figure 1.11.

Three ring chromophores:

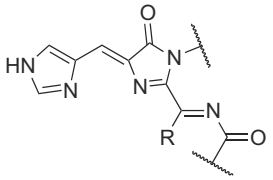


zFP538 (λ_{em} 538 nm)

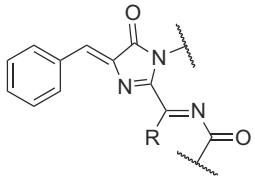


Kubashira orange (λ_{em} 559 nm)

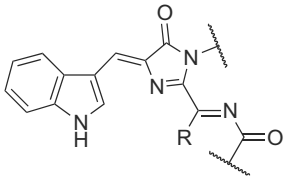
Tyr67His/Phe/Trp DsRed-derivatives:



mCherry Y67H (λ_{em} 464 nm)

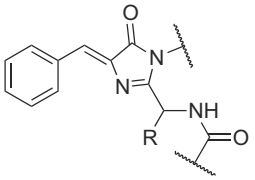


mBlueberry (λ_{em} 450-470 nm)

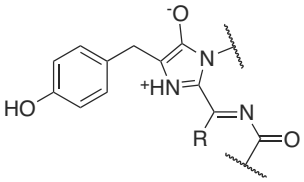


mHoneydew (λ_{em} 537,562 nm)

Alternative blue chromophores:



Sirius (λ_{em} 424 nm)



TagBFP (λ_{em} 456 nm)

Figure 1.11. Additional examples of chemically distinct FP chromophores. (Top row) zFP538 [80] and Kubashira orange [81] form three ring chromophores and exhibit yellow-to-orange emissions. zFP538 undergoes a further cleavage of the peptide backbone. (Middle row) Substitution of Tyr66 with His/Phe/Trp in the DsRed chromophore generates chromophores with blue-shifted emissions [82]. (Bottom row) Blue chromophores can also be generated by substitution of Tyr65Phe in the GFP chromophore [83] or stabilization of the neutral phenol DsRed chromophore [84]. Emission wavelengths are indicated in brackets.

As Figures 1.10 and 1.11 demonstrate, change in colour hue is often attributed to direct chemical modification of the chromophore. However, the alteration of non-covalent interactions between the chromophore and the buried

amino acid side chains in the vicinity of the chromophore may also change its emission wavelength. Perhaps the best illustration of this effect is observed in yellow FP (YFP). By mutating threonine-203 in GFP to hydrophobic residues, the protein is converted into a yellow-emitting species despite possessing a chromophore chemically identical to GFP (Figure 1.10D) [63]. The hydrophobic side chains are situated just above the phenolate moiety of the chromophore and engage in π-π stacking interactions, presumably decreasing the energy of the excited state and allowing more efficient electron delocalization. The result is an increase in emission wavelength.

The emission properties of FPs may also be modulated by further light-induced post-translational modification. In the *Trachyphyllia geoffroyi* coral FP, Kaede, a His-Tyr-Gly green-emitting chromophore is converted to a red-emitting species upon UV irradiation [85]. This green-to-red photoconversion is caused by the photo-induced cleavage of the peptide bond adjacent to the chromophore (Figure 1.12) [85, 86].

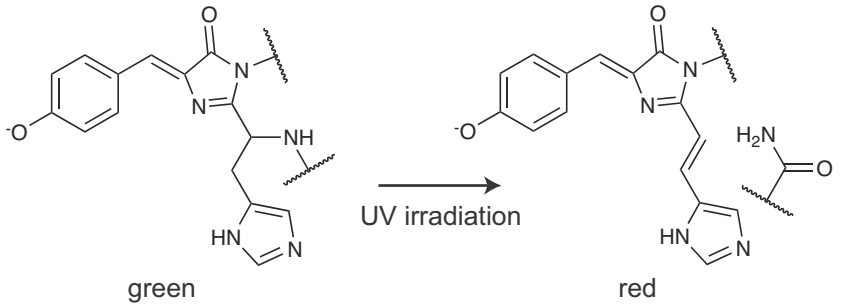


Figure 1.12. The Kaede FP green-emitting chromophore is photoconverted into a red-emitting chromophore upon UV exposure. Backbone cleavage accompanies the photoconversion.

It is clear that changing the hydrogen bond patterns and electrostatic interactions in the microenvironment of the chromophore can change its spectral properties. However, it has also been suggested that the electric field within the

barrel also plays a role [87]. The chromophore possesses an inherent dipole moment. Therefore, the relationship between the electric field inside the barrel and the energy difference between the ground state and excited state chromophore dipole also likely contributes to spectral tuning [87]. This could help to explain the spectral shifts in the mFruit series of FPs (DsRed derivatives) [88].

Another important consideration for FP colourization is the *E/Z* isomerization state of the oxidized tyrosyl α-β bond. The majority of FPs possess chromophores in the *Z* configuration (Figure 1.13) [2]. Rare examples of FPs possessing *E* configuration chromophores include eqFP611 [75], HcRed [89], as well as the engineered FP, Rtms5 [90]. Crystallographic studies have also revealed that *E/Z* photoisomerization is responsible for the photoreversible photobleaching observed in teal FP, where the *Z* isomer represents the fluorescent state and the *E* isomer represents the non-fluorescent state [91].

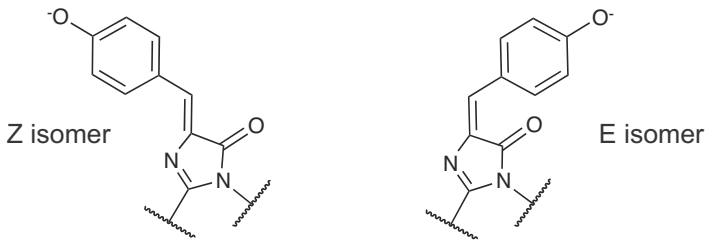


Figure 1.13. *E/Z* isomers of FP chromophores. Most FPs possess chromophores in the *Z* configuration.

1.3.6 Engineered variants of fluorescent proteins

Naturally occurring FPs evolved in jellyfish and coral to exhibit properties optimized for the specific habitats of those organisms. However, the biological systems used by researchers rarely correspond to their natural environments.

As a result, naturally occurring FPs often exhibit sub-optimal properties under the demanding experimental conditions used by researchers and much effort has been dedicated to optimize FPs for use in non-natural systems. Fortunately, FPs are particularly amenable to improvement by protein engineering on account of their relatively good expression and accumulation in *E. coli* [92, 93]. Fluorescence phenotypes can also be easily screened without the need for complex assays using purified proteins [9, 94]. Indeed, using some simple molecular biology techniques, the folding efficiency of GFP was quickly improved following the first demonstrations of its heterologous expression [95, 96]. Additionally, it was quickly discovered that GFP's spectral properties could be modulated for better overall protein performance; simple mutations could alter the excitation profile of GFP to increase its brightness [96, 97]. These types of improvements occur routinely during directed evolution of FPs. FPs are typically engineered to alter their colour hue, optimize spectral features (such as brightness, photostability, and pK_a), as well as to minimize their oligomerization propensity [9, 11, 83, 98, 99].

Although the discovery of FPs in Anthozoan reef corals [49] offered many new colours of variants, including red, their utility was hampered by their inherent propensity for tetramerization. However, guided by X-ray crystal structure information, these interfaces were proven to be breakable, albeit with the unfortunate consequence of loss of fluorescence [9]. The first Anthozoan derived monomeric RFP was derived from DsRed and required substantial mutagenesis to restore fluorescence after its interfaces were sequentially disrupted (Figure 1.14) [9]. This strategy has been used to engineer other monomeric variants of varying colour hue [11, 100].

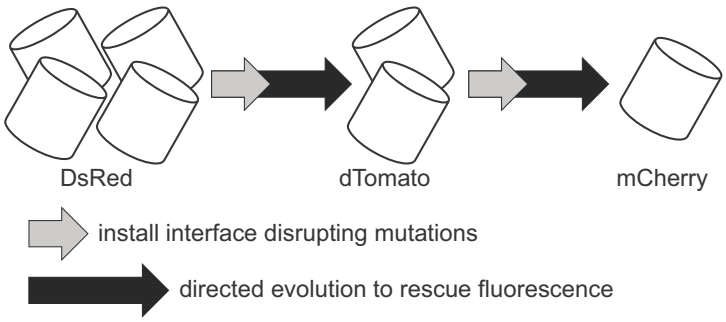


Figure 1.14. Directed evolution of monomeric FPs (e.g. mCherry) from obligate tetrameric FPs (e.g. DsRed). Shown is a cartoon representation of the disruption of the DsRed interface to produce a dimeric variant, called dimer Tomato (dTomato). The dTomato interface is then broken to ultimately produce a monomeric variant called mCherry.

Although monomeric FPs represent the popular choice for biological applications, tandem FPs have also proven useful. Tandem dimer FPs are generated by joining two copies of a dimeric FP gene with a linker sequence. The tandem dimer FP forms an intramolecular interface and behaves as a pseudomonomer of twice the size of a typical FP [2, 9]. Although physically larger, tandem dimer FPs exhibit enhanced brightness since they possess two chromophores. Examples include tandem dimer variants of dTomato and HcRed [9, 101].

1.3.7 Applications of FPs: genetically encoded biosensors

FPs are most commonly used as a reporters of protein localization and organelle structure. That is, any given FP can be used to tag a protein-of-interest and trace its localization in live cells. This represents a relatively simple strategy and only requires that the FP can be fused to a target protein without inducing its aggregation or mislocalization. Such fusions are readily generated by genetically tethering an FP-encoding gene to the gene encoding a protein-of-interest (Figure 1.1). The gene fusion is assembled in a plasmid suitable for expression in mammalian cells (or whichever host cell system is being used).

After the plasmid is transfected into cells, and gene expression is allowed to proceed -- usually for 12 to 48 hours -- the localization of the target protein can be determined using fluorescence microscopy. In addition to localization reporters, FPs are also widely used as reporters for transcriptional activation [2]. Such technology is often utilized to identify DNA-binding proteins or PPIs. In the former application, fluorescence is used as an optical readout of a transcription factor binding a promoter to initiate gene expression [102] (reviewed in [103]). For the latter application, potential interacting partners are tethered to components of a split transcription factor complex. If the proteins interact, transcription is initiated and the FP reporter is produced [104]. In addition to being useful as simple fluorescence reporters, FPs possess unique properties that enable more elaborate analyses, namely biosensing applications.

In general terms, a biosensor may be defined as a detection platform that utilizes biological recognition and a physical transducer to couple a recognition event to an assayable signal output [105]. Since biomolecular recognition regulates physiological behaviour at the level of the cell, the concept of biosensing lends itself to use by biochemically and biologically minded researchers. The sensitivity of fluorescence and its ability to be genetically encoded make FPs ideal for designing biosensors. Today, several biosensing platforms or strategies exist to assay physiological processes in real time. These strategies fall into three general classes: resonance energy transfer (RET) biosensors, complementation based biosensors, and single FP-based biosensors.

1.3.7.1 RET biosensors

Fluorescent protein based RET biosensors may be designed using bioluminescent (BRET) or fluorescent (FRET) platforms. FP acceptors are common to both strategies, but for BRET biosensors the donor energy is derived from a biologically catalyzed chemical reaction, whereas FRET biosensors utilize an FP as a source of donor energy [106, 107].

The general mechanism of BRET involves a chemiluminescent reaction (catalyzed by a bioluminescent protein) that provides energy to excite an acceptor FP, which in turn emits light [108, 109]. Therefore, no excitation light is required for BRET. An absolute requirement, however, is the spectral overlap of the donor emission with the absorbance spectrum of the acceptor (Figure 1.15) [108]. Further, the efficiency of BRET is a distance dependent phenomenon [108, 110]. It follows that BRET is used as a biosensing strategy for the proximity of biomolecules, typically proteins [106, 110].

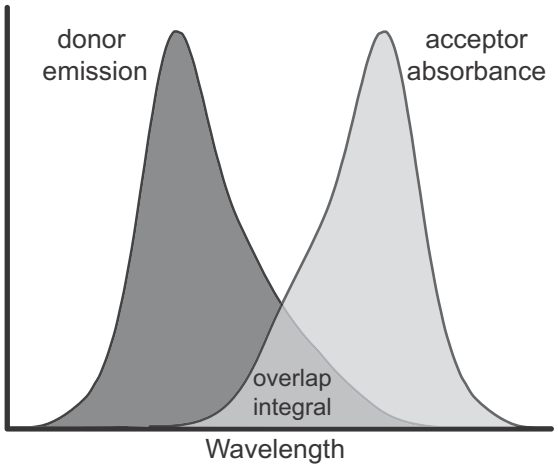


Figure 1.15. Schematic representation of RET spectral overlap. A necessity for RET processes is overlap of the donor emission spectrum with the acceptor absorbance spectrum.

In most BRET applications, *Renilla* luciferase is used to catalyze the oxidation of a substrate to generate the donor energy [110]. Depending on the cell permeable substrate used, typically coelenterazine (λ_{em} 480 nm) or DeepBlueC (λ_{em} ~395 nm), YFP or GFP may be used as acceptor FPs, respectively [110]. These types of BRET based biosensors have been used for many live cell applications, such as monitoring oligomerization of G-protein coupled receptors [111, 112] and transcription factor oligomerization [113]. Despite the advantage of not requiring excitation light, the utilization of FRET technology is far more common than the usage of BRET.

Genetically encoded FRET-based biosensors operate on a principle similar to BRET, with the exception that the energy donor is an FP requiring excitation by light. Again, FRET efficiency is dependent on distance and orientation of donor and acceptor fluorophores (Figure 1.16) [107, 108]. A major advantage to FRET (as well as BRET) technology is that measurements are inherently ratiometric. This permits the inherent calibration of quantitative microscopy experiments to minimize cell-to-cell variability in the expression level of biosensor. The prevailing choice for a FRET donor and acceptor pair is a Cyan FP (CFP) and YFP, but several other suitable pairs have been used in practice [114, 115].

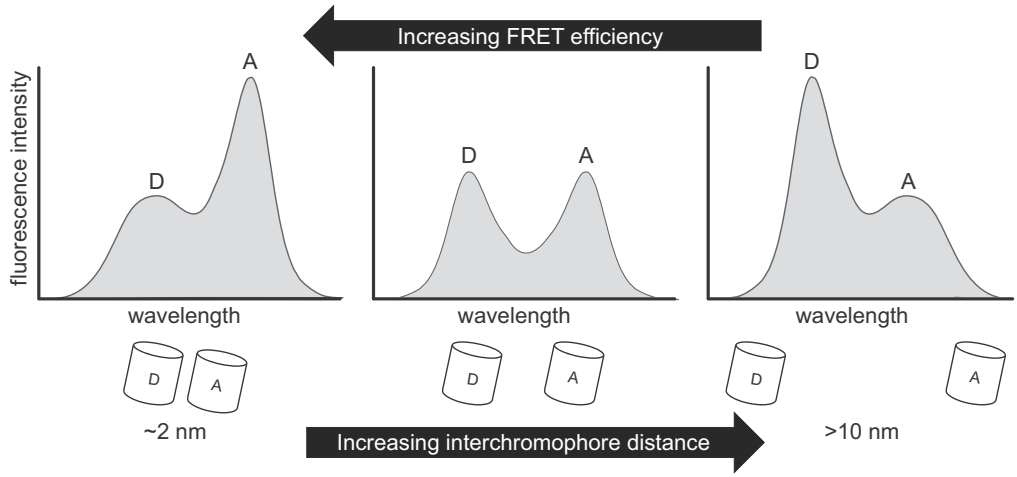


Figure 1.16. Schematic representation of emission spectra for a typical FRET-type biosensor. As the distance between donor (D) and acceptor (A) increases, the emission of the acceptor decreases.

Several FP-based FRET biosensor design strategies are commonly used. Almost every strategy relies on fusing molecular recognition domains to a pair of FPs that demonstrate an appropriate spectral overlap. A summary of popular design strategies is presented in Figure 1.17.

In the first strategy, biosensors are based on a small-molecule dependent PPI. These biosensors may be designed as intramolecular single polypeptide biosensors (Figure 1.17A), or alternatively as intermolecular biosensors (Figure 1.17B). This type is best known for the 'cameleon'-style biosensors of Ca^{2+} [116, 117]. These biosensors are designed by fusing a donor FP (e.g. CFP) to a Ca^{2+} binding module called Calmodulin, or CaM. A second fusion is generated in which an acceptor FP (e.g. YFP) is fused to a peptide, called M13, that forms a complex with Ca^{2+}-loaded CaM [116]. When deployed together these fusions provide robust sensitivity to Ca^{2+} concentration changes in live cells as the donor and acceptor FPs are brought into close proximity in a Ca^{2+}-dependent manner [116-118].

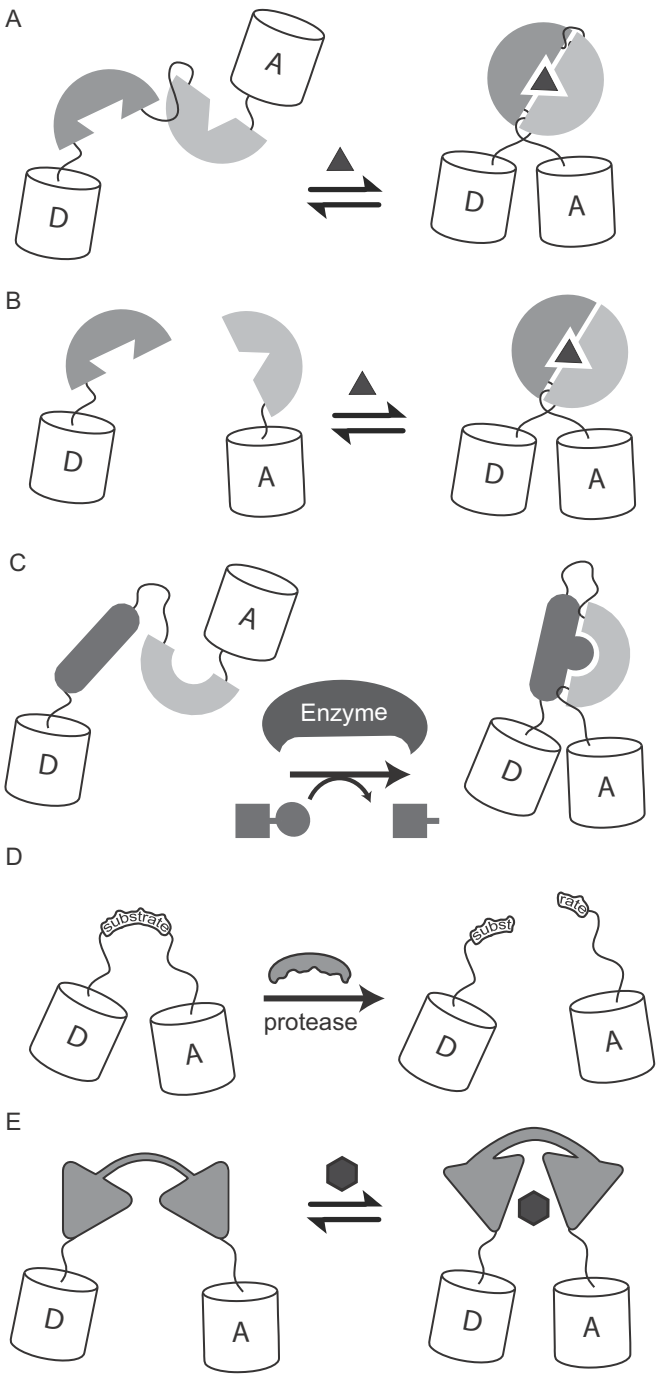


Figure 1.17. FP-based FRET biosensor design strategies. For all examples, D is the donor FP and A is the acceptor FP. (A) Intramolecular biosensors of a small molecule (grey triangle). (B) Intermolecular biosensors of small molecule (grey triangle). (C) Biosensors of enzymatic activity. The circle depicts a chemical functionality that is catalytically installed into a substrate domain (ellipse). (D) Protease biosensors. Donor and acceptor FPs are linked by a protease-labile substrate. (E) Allosteric 'clam-shell' biosensors for small molecules (grey hexagon).

The second strategy relies on a post-translational enzymatic modification to induce a conformational change. These biosensors utilize a substrate domain and molecular recognition domain that binds to the enzymatically modified substrate. The molecular recognition domain and substrate are linked together with donor and acceptor FPs at each end of the polypeptide (Figure 1.17C). The distance separating the donor and acceptor, as well as the relative orientation of their dipoles, changes in the bound state relative to the unbound state to produce a change in FRET efficiency. The strategy has been used to probe the activity of enzymes, such as kinases [119] or GTPases [120].

The third strategy is based on protease-substrate recognition and always operates with a protease-dependent loss-of-FRET response (Figure 1.17D). This type of biosensor was used in the first demonstration of FRET using FPs in which a single polypeptide comprised of blue FP (BFP) and GFP was joined by a trypsin-cleavable linker [121]. This particular biosensor was limited to *in vitro* cleavage. Genetically encoded proteolysis sensors have since been used to detect other proteases such as caspase-3 [11] or to screen for inhibitors of viral enzymes such as poliovirus 2A protease [122] in live cells.

The fourth FRET biosensor design strategy relies on a conformational change of a single protein domain upon binding a small molecule. Periplasmic binding proteins are typically utilized to engineer this class of biosensor. They are themselves a biosensing module that operate on a ‘clam-shell’ like mechanism whereby two protein domains, linked by a hinge, clamp down on a ligand [123, 124]. This changes the proximity and orientation of two genetically linked FPs, thus producing a change in FRET efficiency (Figure 1.17E). These biosensors have been engineered to respond to various analytes, such as citrate, maltose, and glucose [125-128].

The strategies described above are popular designs for FP-based biosensors, but do not represent an exhaustive list. Other interesting FRET biosensor designs have been developed for specific applications, such as ratiometric pH sensitive probes [129] or mechanical tension biosensors [130, 131], but these designs are less commonly utilized.

Although engineering strategies for FP-based biosensors typically rely on the non-homologous protein modules that are fused to FPs, the FP components themselves can be targeted to improve FRET efficiency [132-134]. Of particular relevance to this thesis are laboratory-optimized 'sticky' FRET pairs that promote intramolecular complex formation [134]. It is believed this FRET pair can form a weak heterodimer in which the chromophore dipoles are in a near optimal orientation for energy transfer. The sticky FRET partners, known as CyPet and YPet [134], were derived from CFP and YFP (both of which are descendants of GFP). It is therefore not surprising that superposition of the mutations in CyPet and YPet on the GFP structure correspond to residues at the GFP homodimer interface [135]. This optimized pair exhibits improved sensitivity and enhanced dynamic range relative to the conventional CFP and YFP FRET pair [134, 135].

A final note on FP-based FRET technology is the current trend to deploy FRET pairs in tandem to concurrently monitor independent physiological events in the same cell [114, 115, 133, 136]. Substantial care must be taken to choose appropriate FRET pairs to enable this multiparameter imaging [114]. Examples include concurrent monitoring of caspase-3 activity and Ca^{2+} dynamics [114, 133], as well as Src kinase activity in conjunction with matrix-metalloproteinase activity [115].

1.3.7.2 Bimolecular complementation biosensors

Complementation may be defined as a process in which two protein fragments assemble to form a whole protein. This concept has often been utilized to develop biosensors for PPIs and it relies on the artificial genetic fragmentation of proteins or enzymes [33]. The non-functional split fragments are fused to (potentially) interacting protein partners. If those proteins happen to interact, the proximity-induced reconstitution of the inert fragments restores the native activity of the split molecule and thereby indicates the presence of an interaction. This concept was introduced in section 1.2.

GFP was successfully split and developed into a complementation technology in 2000 [137] (Figure 1.18). Non-fluorescent N- and C-terminal GFP fragments were fused to leucine zipper domains that form functional dimers. The two FP fragments formed a fluorescent complex spectrally indistinguishable from its GFP parent molecule upon dimerization of leucine zipper domains [137]. FP complementation is also commonly referred to as bimolecular fluorescence complementation (or BiFC) [17].

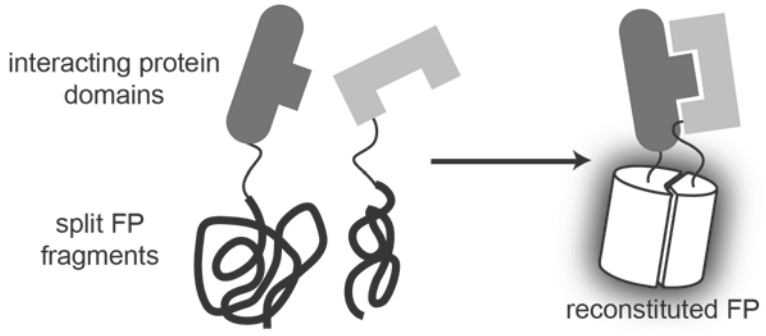


Figure 1.18. Bimolecular fluorescence complementation strategy. Non-fluorescent split FP fragments are irreversibly reconstituted in the presence of a protein-protein interaction (shaded domains) to restore fluorescence.

The primary use of FP complementation is to visualize PPIs or compartmentalization of interactions in live cells [17]. The strategy requires interacting partners be fused to each of the split FP fragments, as well as achieving sufficient expression levels of the fusions in an aerobic environment [138]. Interactions may be visualized in any cellular compartment provided the fused proteins can gain access to one another [138]. Split FP complementation technology has also been applied in several high-throughput strategies to identify protein-interaction partners [26, 139-143]. However, significant limitations to this technology include slow kinetics of fluorescence development and the irreversible nature of complementation [17].

The split FP complementation colour palette has now expanded beyond GFP. Several FPs of varying colour hue, from blue to far red, have now been split and used as complementation based biosensors: cerulean [144]; venus [144, 145]; red [146, 147]; and far red [148]. Furthermore, the availability of multiple colours has enabled the imaging of multiple PPIs occurring simultaneously in a cell [149, 150].

1.3.7.3 Single-FP biosensors

As their name suggests single-FP biosensors utilize a single FP that is spectrally responsive to its environment or an analyte-of-interest. The optical response may be an intensiometric change in the FP emission intensity or a change to its excitation or emission profile. Advantageously, the magnitude of an intensiometric change is typically greater for single-FP biosensors than that observed for FRET based biosensors [151]. Although the majority of single FP-based biosensors are intensiometric, ratiometric biosensors have also been reported [152].

Single-FP biosensors can be categorized into two sub-classes. The first class is single copy FPs with inherent biosensing capabilities. These FPs are not fused to any molecular recognition domains or biosensing modules, but instead possess unique properties that cause them to change their output signal in response to a change in their environment (Figure 1.19). These biosensors have been developed to exhibit sensitivities to pH [153], halides [154, 155], as well as redox potential [156-158]. The number of analytes or biological processes that can effectively elicit a change in the optical output of a single copy FP is limiting. However, by fusing FPs to extrinsic biosensing domains the range of potential analyte sensitivities increases [151].

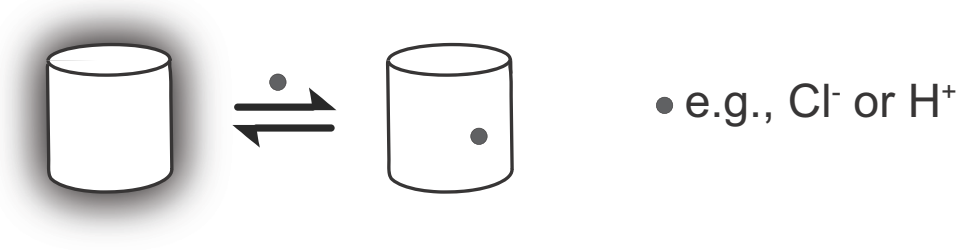


Figure 1.19. Schematic representation of a single-FP based intrinsic biosensor. The fluorescence emission of the FP changes in response to small ions, such as Cl^- or H^+.

The second class of single-FP biosensors utilizes a design in which FPs are fused to an extrinsic molecular recognition module capable of sensing an analyte-of-interest. The effect of the molecular recognition event is to confer, typically through a conformational change, a perturbation to the FP structure or chromophore microenvironment to change its optical output. These analyte-induced perturbations typically restrict solvent access and promote fluorescence [159]. To affect the chromophore in such a way, physical access of solvent to the chromophore is a requirement. Unfortunately, the native termini of FPs are

not sufficiently proximal to the chromophore [63] and thus the fusion of a molecular recognition element to the naturally occurring termini will not result in significant perturbation to the chromophore upon analyte binding. Therefore, additional engineering is required to position molecular recognition elements close to the chromophore.

Access to the chromophore can be enabled through circular permutation of FPs [160-162]. This technique joins the natural termini and installs new N- and C-termini in the vicinity of the chromophore (Figure 1.20). Fusion of molecular recognition elements to the new termini allows for perturbation of the chromophore with analyte-induced conformational changes. Examples include the well-known Ca^{2+} indicators, GCaMP and PeriCaM [160, 161].

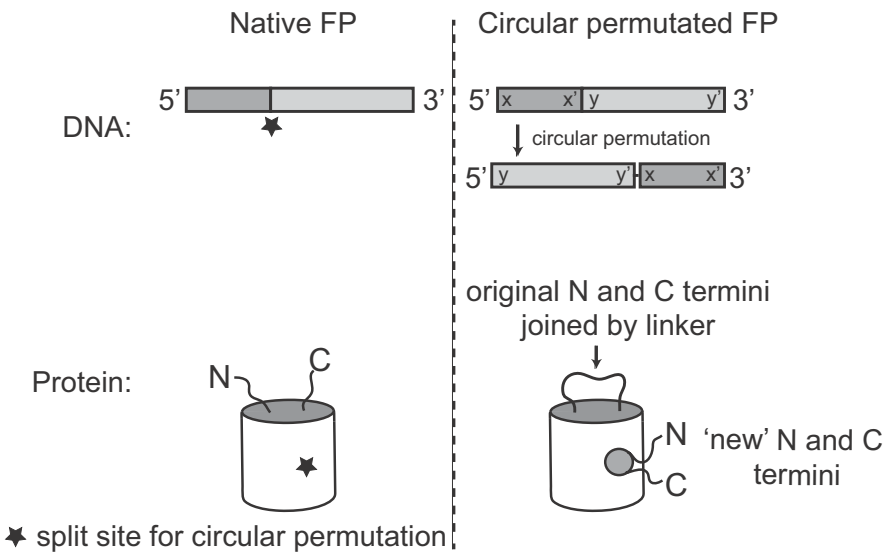


Figure 1.20. Fluorescent protein circular permutation strategy. Original N- and C-termini are joined by a linker and new termini are created in close proximity to the chromophore. Despite the change in topological connectivity, the 3D-structure is retained.

The prototypal biosensor of this kind is GCaMP and is based on a circularly permutated variant of GFP, called cpGFP (Figure 1.21). The molecular

recognition elements are the CaM domain and the M13 peptide, which are fused to the N- and C-termini of cpGFP (Figure 1.21). Crystal structures of GCaMP in the Ca^{2+}-loaded and Ca^{2+}-free states have revealed the mechanism of the intensiometric response to Ca^{2+} [159]. In the absence of Ca^{2+}, solvent has free access to the chromophore and quenches fluorescence. When Ca^{2+} is present, the conformational change associated with the binding of Ca^{2+}-loaded CaM to M13 effectively plugs the hole adjacent to the chromophore and restores fluorescence by preventing solvent access [159]. For many years these Ca^{2+} indicators were limited to green fluorescence, but a new series of Ca^{2+} indicators was recently engineered to include blue- and red-CaMP variants [152].

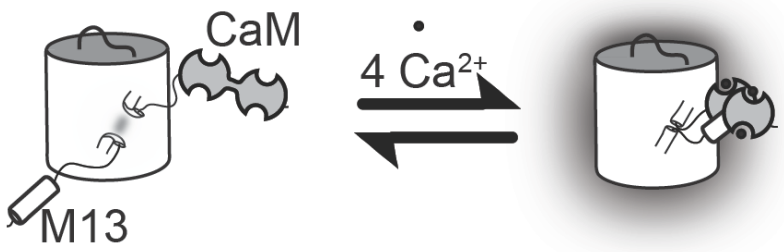


Figure 1.21. Single-FP based GCaMP style biosensor. In the absence of Ca^{2+}, fluorescence is quenched. In the presence of Ca^{2+}, the conformational change that accompanies the CaM/M13 interaction restores fluorescence.

1.4 Protein engineering

Protein engineering is the process by which purposeful changes are made to a polypeptide sequence to generate a protein with a new or altered function [163]. In many cases, new or improved functions may be rationally installed into target proteins using genetic manipulations that permit the site-specific alteration of protein templates. This strategy is almost always guided by

published crystallographic or structure-function studies of the progenitor protein or its homologues. Proteins engineered in this manner must be empirically assayed for phenotype to evaluate the success of the design. Alternatively, proteins are often engineered using a less stringent design whereby they are indiscriminately altered and subjected to a selection pressure. Using an appropriate screening method, proteins exhibiting desired phenotypes can be identified and isolated from a heterogeneous pool of candidates [164]. This process has been described as "…similar to that whereby genes with new functions are evolved by natural selection" [165].

With the advent of new recombinant methods researchers realized they could directly manipulate this artificial evolution process. An invaluable advance in this area was the development of site-directed mutagenesis [166]. This simple genetic technique enabled researchers to rapidly incorporate specific amino acids at specific sites in proteins. Protein engineers immediately adopted the method. The first examples used site-directed mutagenesis to mutate active site residues in enzymes [167, 168]. These early cases did not improve enzyme function, but clearly demonstrated researchers could now alter protein function by targeting any residue of their choosing for mutagenesis. In subsequent years, protein engineers used the technique to generate proteins with desired functional improvements. A prominent example is the industrially important protease, Subtilisin E, which was modified through site-directed mutagenesis for increased stability to oxidation and detergent [169, 170].

Today, protein engineering is largely achieved through what is termed directed evolution. The process utilizes molecular biology techniques coupled with genetic or phenotypic screens, to rationally, semi-rationally, or randomly generate altered proteins in the lab [12].

Directed evolution involves performing iterative cycles of generating a diverse gene library, introducing the library into a suitable host (e.g. *E. coli*), screening the library for variants exhibiting a desired phenotype, and selecting variants for the isolation of next generation gene templates (Figure 1.22). In that sense, researchers direct the evolution of the protein towards the desired phenotype [164]. The success of directed evolution depends largely on the extent of library diversity and the ability to effectively screen for and detect phenotypic improvements [171].

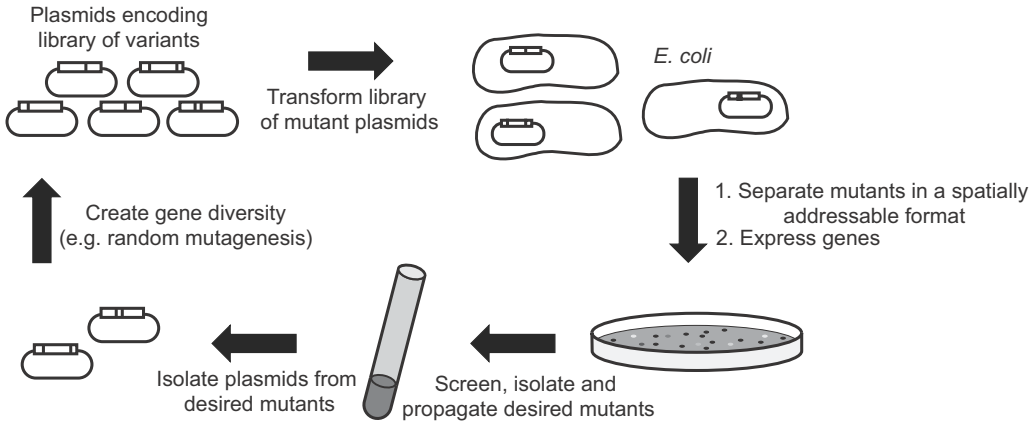


Figure 1.22. Schematic representation of a typical directed evolution strategy.

Two popular strategies are used to generate library diversity: mutagenesis and gene recombination. Mutagenesis is a cornerstone of directed evolution. Most often, random point mutagenesis is performed on a single gene template (or pool of templates) to generate libraries, but saturation site-directed mutagenesis is also commonly utilized [12]. Randomized gene libraries are created by amplifying genes using the polymerase chain reaction (PCR) with error prone reaction conditions (*i.e.,* using low fidelity or error prone polymerases and doping with Mn^{2+}) [172, 173]. Although some studies indicate

high error rates can be effective [174-176], the likelihood of acquiring improved variants decreases with increasing mutation rate due to accumulation of detrimental mutations [176]. Thus, the iterative approach used in directed evolution is limited to improvements made in small steps.

Gene recombination is a powerful strategy to overcome this limitation. This strategy assembles hybrid genes from several gene templates containing mutually exclusive beneficial mutations. The template sequences may be a pool of naturally occurring homologous genes (often used as an initial step in directed evolution [177-179]) or a pool of artificially mutagenized genes derived from a single progenitor [180]. These gene-shuffling strategies typically result in the assembly of hybrid genes possessing synergistic combinations of beneficial mutations [177, 180]. Two popular genetic techniques for shuffling genes include ‘sexual’ PCR (which entails gene fragmentation and reassembly by PCR [178, 181]) or staggered extension PCR (which uses abbreviated thermocycling conditions to recombine DNA sequences by PCR [180]). Many other genetic techniques have been used in various directed evolution strategies (reviewed in [12, 182]), but we have limited our discussion to the most popular methods.

A dutiful discussion of protein engineering and directed evolution must extend beyond the generation of library diversity. Of equal, if not more, importance is the screening strategy. Any screening strategy must link the library of gene sequences to the proteins they encode for, and in turn, link the proteins to their output ‘product’ [37, 183]. Furthermore, effective screening strategies should be highly tailored to the desired phenotype and sensitive over the target dynamic range [184]. Meeting these requirements often proves difficult and represents a bottleneck in many directed evolution strategies [183]. In particular, stringent phenotype selection is essential. The first law of directed

evolution has been suggested as “you get what you screen for” and where protein engineering efforts fail, this tenet has often been flawed [185]. Common strategies include screening colonies on agar growth media or crude cell lysates for the production of a fluorophore or a chromophore [183]. Many additional techniques have been utilized for directed protein evolution (reviewed in [37]), but a thorough analysis of each method is beyond the scope of this thesis.

Directed evolution has been applied to a wide variety of proteins, the bulk of which are enzymes. Important examples include DNA modifying enzymes (e.g. restriction endonucleases and polymerases) as well as industrially useful enzymes (e.g. proteases, cellulolytic enzymes, and lipases) (reviewed in [12]). FPs have also been extensively engineered by directed evolution [2, 9, 11, 13].

Protein engineering has thus far been discussed from a micro-scale directed evolution perspective (*i.e.,* small changes at the amino acid level). However, macromolecular protein engineering strategies (*i.e.,* involving alterations to secondary or tertiary structure elements) are also typically performed. This type of engineering entails circular permutation (previously summarized in Figure 1.20) as well as the design of allosteric or biosensing fusion proteins.

As a final remark, it must be acknowledged that designed proteins rarely function as optimally as protein engineers envision. Rationally designed proteins typically underperform or fail to perform at all [185]. Fortunately, when engineered proteins prove to be even remotely promising, directed evolution can be performed, which often leads to significant improvements [186] to the protein’s intended function (e.g. the evolution of Ca^{2+} biosensors [152] or the optimization of FRET biosensors [132, 133]).

1.5 Research Objectives

We propose the inherent tendency of coral FPs to oligomerize can be exploited to generate a new class of biosensor template. We will use directed evolution and semi-rational protein engineering to convert a dimeric red FP variant, called dTomato, into a two protein system in which low affinity quenched monomeric FPs can reversibly associate to form a bright fluorescent heterodimer. This strategy will require the discovery and engineering of a unique protein-protein interface whose formation will spontaneously mediate fluorogenesis.

Specifically, we hypothesize that a pair of chemically distinct FPs can be engineered to exhibit little-to-no fluorescence in their monomeric state, yet reversibly form a red fluorescent heterocomplex when brought into close proximity (Figure 1.23). This reversible fluorogenic heterodimerization will represent a novel FP-based technology that will extend the currently limited set of available FP-based platforms for biosensor design.

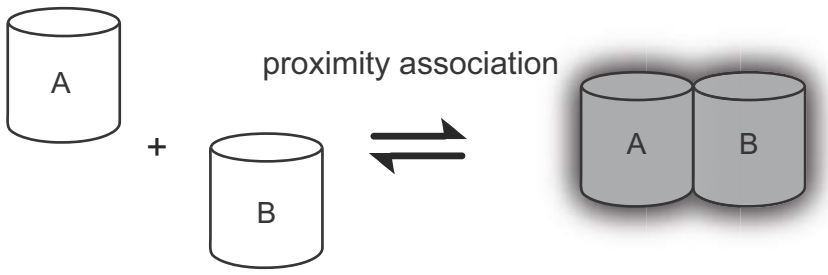


Figure 1.23. General concept of research proposed in this thesis. Two quenched FPs reversibly associate to form a bright fluorescent FP heterodimer.

The specific objectives of this thesis are (1) to develop an assay to screen for PPIs in *E. coli*, (2) to rescue the binding of a monomerized dTomato variant by engineering a heterodimeric interface, (3) to identify and characterize

fluorogenic red FP heterodimers, (4) to show the utility of fluorogenic heterodimers in several live cell biosensing applications, and (5) to generate alternate colour hues of fluorogenic heterodimers.

Chapter 2 Development of an *in vivo* selection assay for protein-protein interactions

2.1 Introduction

The finely orchestrated physiological processes performed by cells rely on molecular recognition and communication. At the molecular level, nearly all communication is mediated by non-covalent physical interactions between biomolecules. Therefore, assays to identify these interactions are important tools for researchers studying biological systems. Although understanding how proteins associate with lipids, carbohydrates, and DNA is essential, the majority of genetic assays developed to date are used to identify and characterize protein-protein interactions (PPIs) [19, 39, 187].

This chapter outlines our efforts to develop a genetic selection assay suitable for the identification of PPIs in *E. coli*. We wanted to determine if the formation of a protein toxin/inhibitor complex could be used as reporter of PPIs in a survival based PCA-like assay (Figure 2.1). This strategy differs from conventional PCA assays in that the toxin and inhibitor are pre-folded modules, as opposed to split unfolded enzyme fragments [24].

Bacillus amyloliquefaciens produces two proteins, Barnase (Bn) and Barstar (BS) that function as a toxin/inhibitor pair. The pair has been thoroughly studied with reports of extensive biochemical and crystallographic characterizations [188-198]. Bn, an extracellular ribonuclease that cleaves single stranded RNA [196], is comprised of 110 amino acid residues and its intracellular inhibitor, BS, is comprised of just 89 residues [194]. The pair forms a tight 1:1 complex with a

dissociation constant of 10^{-14} M [189, 197]. Mutagenesis studies have also revealed the key amino acids involved in catalysis and interface formation [190, 194, 196-199]. Notably, the potent ribonuclease activity of Bn renders its expression in *E. coli* lethal in the absence of BS expression [200-202].

We wanted to determine if the complexation of Bn and BS proteins could be used in a strategy for detecting PPIs in *E. coli*. We envisioned a system in which bait-Bn and prey-BS protein hybrids are co-expressed, and survival is contingent on the interaction of the bait and prey (Figure 2.1). In the absence of an interaction, Bn is uninhibited and promotes cell death. We designated this concept as protein interaction-mediated toxin silencing (PINTS).

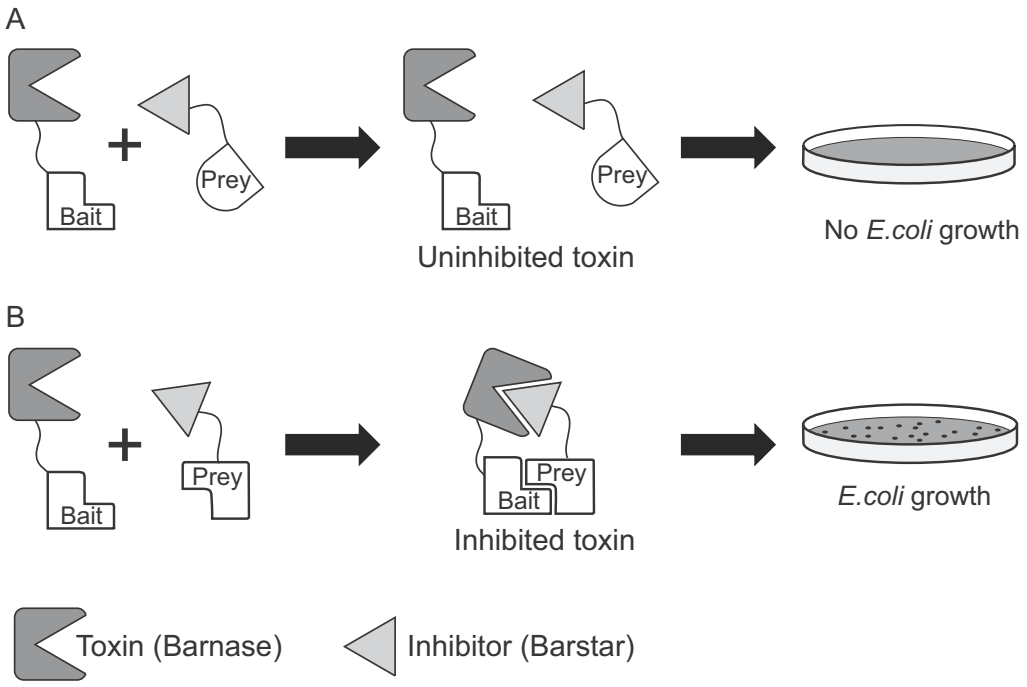


Figure 2.1. Proposed mechanism of protein interaction-mediated toxin silencing, or PINTS. (A) In the absence an interaction Bn is uninhibited and prevents growth. (B) In the presence of an interaction, Bn and BS form a complex, which inhibits Bn activity and permits growth.

In this work, we developed a PINTS assay that utilizes a reduced affinity Bn/BS pair as a survival reporter for PPIs in *E. coli.* We first validate the PINTS

concept using the established PPIs of oligomeric FPs, as well as the heterotypic interaction of two mammalian signaling proteins, Ras and the Ras-binding domain (RBD) of Raf-1 [203]. Furthermore, we tested whether the PINTS system could be used to rescue an interaction abolished by an interface disrupting mutation. Specifically, we screened a library of mutant Ras proteins against a binding deficient RBD mutant, R89L, and identified a Ras variant exhibiting rescued binding in live cells. These experiments showed that the PINTS strategy is useful for validating and identifying PPIs in *E. coli*.

2.2 Results and discussion

2.2.1 PINTS strategy and plasmid design

The development of a bacterial PINTS selection system necessitated the engineering of the Bn/BS pair in two regards. First, the toxicity of WT Bn had to be attenuated. Second, the affinity of the Bn/BS pair had to be reduced relative to that of WT.

Bn is a potent ribonuclease whose expression in *E. coli* is lethal without co-expression of the BS inhibitor [200, 202]. In our assay, the interactions between bait and prey proteins will be in equilibrium, where the bait-Bn fusion is expected to exist free from its interaction with prey at least part of the time. Therefore, WT Bn cannot be used on account of its lethality. Ideally, we would prefer to use a reduced activity Bn mutant that would permit growth in the presence of a small non-lethal fraction of the uncomplexed bait-Bn fusion.

In regards to the second criteria, a modified low affinity Bn/BS pair was required since complexation of the WT proteins is strongly favoured (K_d 10^{-14} M) [197]. Our goal was to select for growth only in those instances where bait and

prey proteins interact, but the high affinity of the WT interaction would confer survival independent of any bait/prey interaction (Figure 2.2A). Thus, a reduced affinity Bn/BS pair is essential. Several Bn/BS mutagenesis studies helped to guide us in selecting a suitable Bn/BS pair [190, 194, 196, 197].

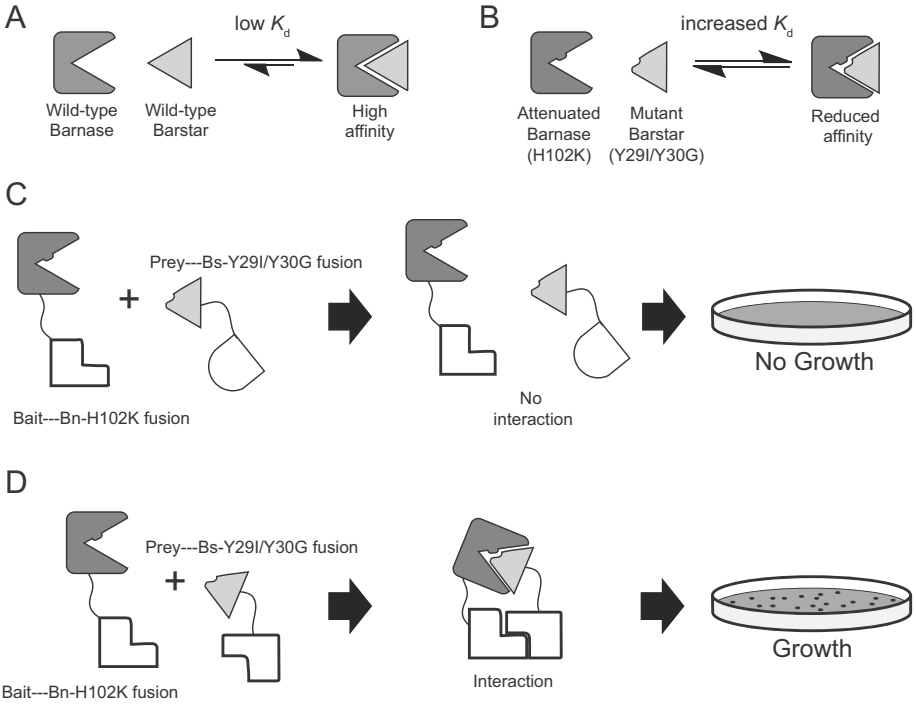


Figure 2.2. A low affinity Bn/BS pair is required for the PINTS assay. (A) The WT Bn/BS interaction is strongly favoured. (B) The Bn-H102K and BS-Y29I/Y30G pair exhibits a reduced affinity [199]. (C) In the absence of an interaction Bn-H102K is uninhibited and prevents growth. (D) When bait and prey proteins interact, BS-Y29I/Y30G inhibits Bn-H102K and growth is permitted.

In Bn, His-102 plays an important role in both catalysis and molecular recognition by BS. During RNA hydrolysis, His-102 is thought to act as the general acid [196]. It also contributes three hydrogen bonds to the Bn/BS interaction and its imidazole-bearing side chain is completely buried in a BS binding pocket [204]. Mutation of His-102 to lysine (*i.e.*, Bn-H102K) severely impairs, but does not eliminate the catalytic activity of Bn [199]. It follows that expression of Bn-H102K alone in *E. coli* is toxic, but to a lesser extent than WT

[199]. On account of its role in BS recognition, the H102K mutation increases the K_d between the pair to > 5 x10^{-5} M [199]. The binding affinity of Bn-H102K for BS can be rescued by extragenic suppressor mutations in BS such that co-expression of the Bn-H102K/mutant BS pair in *E. coli* confers survival [199]. One BS mutant in particular, Y29I/Y30G, has been reported to exhibit minimal ability to rescue survival [199]. Although the exact K_d has not been reported for this mutant Bn/BS pair, it can be estimated to be between 10^{-9} M and 10^{-5} M [199]. For our PINTS assay, we decided to use Bn-H102K and BS-Y29I/Y30G as a decreased affinity pair to mediate survival-based selection contingent on the interaction of bait and prey (Figure 2.2B, C, and D).

The PINTS principle requires concurrent expression of two gene fusions. We used a custom *E. coli* expression plasmid capable of expressing two genes under two different promoters (Figure 2.3). The bait-Bn-H102K fusion was cloned under a tightly regulated P_{BAD} promoter that allowed us to control the extent of toxin expression. The prey-BS-Y29I/Y30G fusion was cloned under a P_{Tac} promoter (Figure 2.3A). The genes under the P_{BAD} promoter and P_{Tac} promoters can be induced by L-arabinose and isopropyl β-D-thiogalactopyranoside (IPTG), respectively. For each fusion, a tandem $(glycine)_4$serine linker was installed to permit some flexibility and allow for both the bait/prey interactions as well as the Bn/BS interactions to occur by potentially increasing collisions between the fusion partners (see Materials and Methods for exact linker used). For convenient cloning of different bait and prey genes, the plasmid was designed to accept prey-encoding genes between *XhoI* and *EagI* restriction sites (Figure 2.3B) and bait-encoding genes between *EcoRI* and *KpnI* restriction sites (Figure 2.3C).

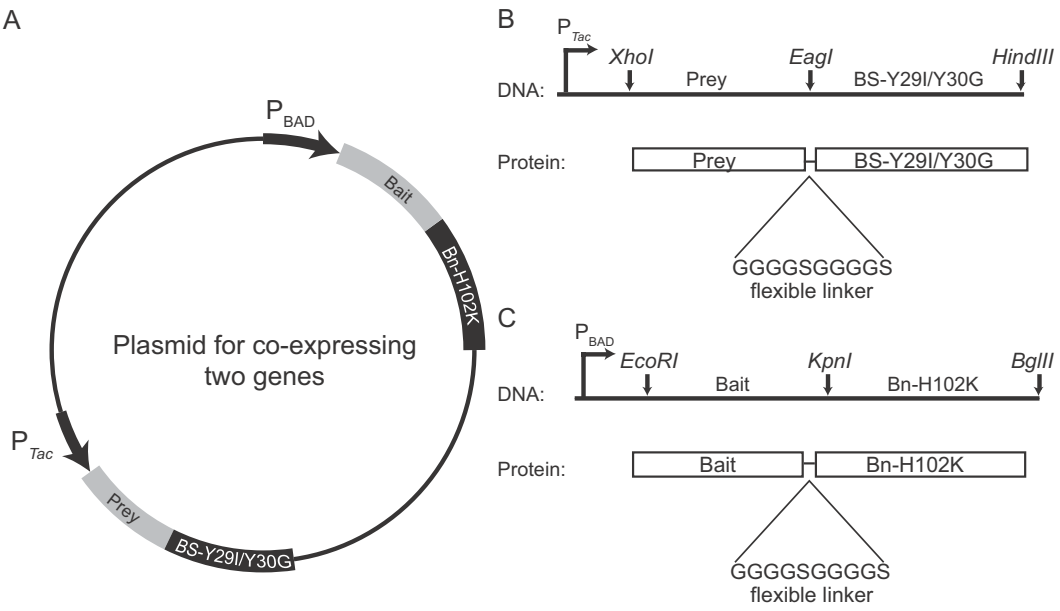


Figure 2.3. Expression system used for the design of PINTS assays. **(A)** Bacterial expression plasmid used for the expression of bait-Bn-H102K and prey-BSY29I/Y30G fusions in *E. coli*. (B) Expression cassette design for prey-BS-Y29I/Y30G fusion. (C) Expression cassette design for bait-Bn-H102K fusion.

2.2.2 PINTS Proof-of-Concept

Before we could use our assay for a practical purpose, we first needed to validate the suitability of the reduced affinity Bn-H102K/BS-Y29I/Y30G pair for PINTS. Since our lab's area of expertise is FP engineering, we turned to a set of FPs with well-known association states: dimeric teal fluorescent protein version 0.7 (dTFP0.7), and monomeric teal fluorescent protein version 1 (mTFP1) [11]. We predicted dimerization of dTFP0.7 fusions would be sufficient to negate cell death by driving the formation of the inhibited complex of BS-Y29I/Y30G and Bn-H102K. Conversely, the monomeric mTFP1 fusions would leave Bn-H102K unsequestered and compromise cell viability (Figure 2.4A).

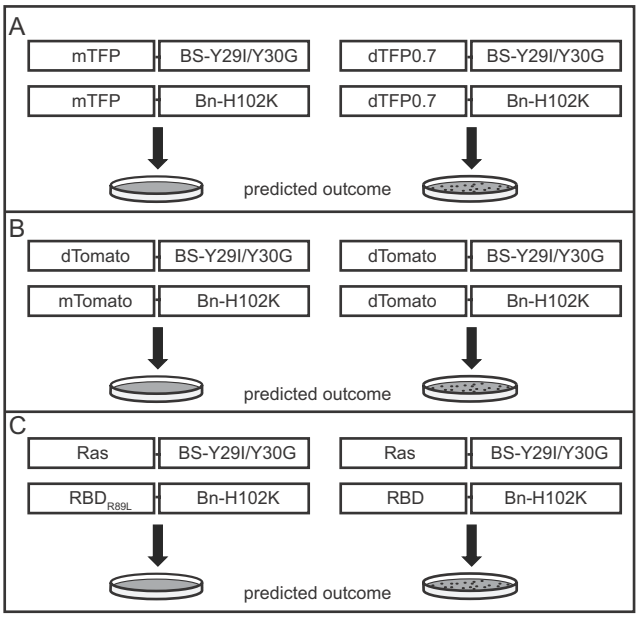


Figure 2.4. Schematic representation of PINTS constructs used in proof-of-concept assays. The predicted outcomes are shown for the following bait and prey pairs: (A) mTFP1 and dTFP0.7, (B) dTomato and mTomato, (C) Ras and RBD or RBD_{R89L}.

The dTFP0.7 or mTFP1 genes were fused to both Bn-H102K and BS-Y29I/Y30G genes in the PINTS plasmid (Figure 2.4A). Each plasmid was then plated onto growth media supplemented with L-arabinose and IPTG to induce expression of each FP gene hybrid. We fixed the IPTG concentration at 1 mM and supplemented the media with increasing concentrations of L-arabinose. At each concentration of L-arabinose we examined plates for colony growth after overnight incubation. Our goal was to find conditions such that the extent of toxin expression was lethal to cells when the mTFP1 gene fusions were expressed, but growth-permissive when the dTFP0.7 gene fusions were expressed. As a control, we also co-expressed dTFP0.7-Bn-H102K with an unrelated protein (*i.e.,* a FRET biosensor) to ensure that the dimeric state of the dTFP0.7 bait was not attenuating the toxin.

After overnight incubation we found that *E. coli* expressing the dTFP0.7 fusions survived on media supplemented with L-arabinose at 0.45%, while the mTFP1 fusions did not (Table 2.1). Co-expression of dTFP0.7-Bn-H102K with the unrelated protein was lethal at this L-arabinose concentration, indicating the survival phenotype was conferred by dTFP0.7-mediated complexation of Bn-H102K and BS-Y29I/Y30G.

Table 2.1. Results summary for the mTFP1- and dTFP0.7-mediated PINTS assay.

	Colony Growth		
% L- arabinose	dTFP-Bn-H102K unrelated protein[1]	dTFP-Bn-H102K dTFP-BS-Y29I/Y30G	mTFP1-Bn-H102K mTFP1-BS-Y29I/Y30G
0.00	+	+	+
0.15	+	+	+
0.30	-	+	-
0.45	-	+	-
0.60	-	-	-

[1] A FRET construct was expressed as an unrelated protein.
'+' indicates colony growth
'-' indicates no growth

We corroborated this result using another dimeric FP variant called dTomato [9] (Figure 2.4B); induction at 0.45% L-arabinose was growth permissive (Table 2.2). To further confirm that survival was conferred by the dimerization of FPs, we predicted growth could be prevented by installing monomerizing mutations into the dTomato gene fused to Bn-H102K (Figure 2.4B). The mutations H162K and A164R convert the dTomato protein into a monomer [9]. For the purposes of this report we designate this variant as monomeric Tomato, or mTomato (Figure 2.5).

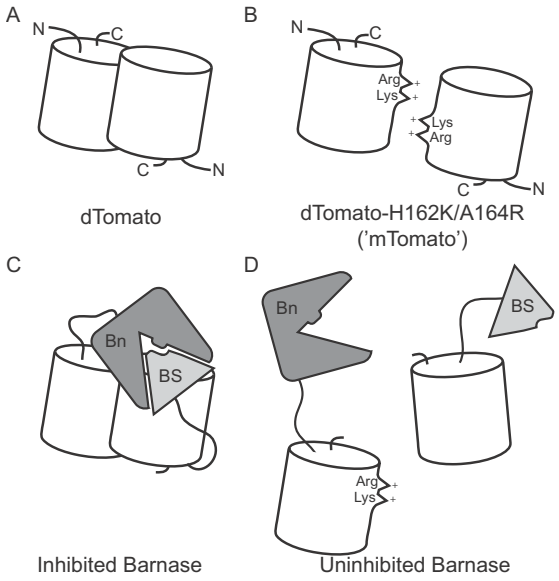


Figure 2.5. Schematic representation of dTomato association states and their utility for PINTS proof-of-concept assays. (A) Native dTomato homodimer. (B) The mutations H162K and A164R convert dTomato into a monomer (mTomato). (C) Dimerization of dTomato-Bn and dTomato-BS fusions inhibits the Bn toxin. (D) mTomato-Bn cannot dimerize with dTomato-BS which leaves Bn uninhibited. Although not depicted in (D), dTomato-BS would be present as a homodimer.

Co-expression of mTomato-Bn-H102K and dTomato-BS-Y29I/Y30G at 0.45% L-arabinose was not growth permissive (Table 2.2). This result suggested the PINTS principle was valid and dimerization of bait-Bn-H102K and prey-BS-Y29I/Y30G fusions was the driving force for survival.

Table 2.2. Results summary for the mTomato- and dTomato-mediated PINTS assay.

	Colony Growth		
% L-Arabinose	dTomato-Bn-H102K unrelated protein[1]	dTomato-Bn-H102K dTomato-BS-Y29I/Y30G	mTomato-Bn-H102K- dTomato-BS-Y29I/Y30G
0.00	+	+	+
0.15	+	+	+
0.30	-	+	-[2]
0.45	-	+	-
0.60	-	-	-

[1]A FRET construct was expressed as an unrelated protein
[2] few small pin prick colonies were present
'+' indicates colony growth
'-' indicates no colony growth

Although dimeric FP-mediated PINTS conferred a growth phenotype, surviving colonies displayed an aberrant morphology. Whereas typical *E. coli* form cream-coloured opaque colonies, surviving PINTS clones were slower growing (*i.e.,* smaller colony size) and exhibited a translucent morphology. After several days at room temperature, many of these clones developed an opaque focus at their center or periphery. These colonies were streaked on fresh growth media and after overnight incubation the original translucent phenotype was obtained. The reason for this phenotype is uncertain, but it is consistent with reports of slow unstable growth and foci development for Bn-H102K-expressing *E. coli* colonies [194, 199, 201]. We were concerned about the stability of the plasmids in these clones, but gene sequencing verified their integrity.

Despite the aberrant phenotype, the ability to obtain stable clones using our reduced affinity Bn/BS selection strategy appeared to validate the PINTS assay. However, we acknowledge the use of homodimeric FPs was not ideal due to the formation of bait/bait and prey/prey homodimers. We wanted to determine if the PINTS assay could detect more conventional heterotypic PPIs. Therefore, we performed an additional proof-of-concept assay using the interaction of Ras with the RBD of Raf-1 [203]. Ras is a small membrane-associated GTPase and Raf-1 is a serine/threonine kinase that mediates cell proliferation responses downstream of activated Ras [205, 206]. The Ras/RBD interaction has been used previously to validate a DHFR PCA strategy [22, 31].

In our PINTS dual expression plasmid, we fused the RBD to Bn-H102K and Ras to BS-Y29I/Y30G (Figure 2.4C). As a control we also made a PINTS plasmid encoding a binding deficient RBD mutant, R89L [207, 208] (Figure 2.4C). This mutation effectively abolishes the RBD/Ras interaction with an estimated K_d >

500 μM [209]. Further, the R89L mutation prevents growth in DHFR PCAs. Co-expression of Ras-Bn-H102K and RBD-BS-Y29I/Y30G fusions at 0.30% L-arabinose resulted in colony growth after overnight incubation at 37°C, while no growth was observed for co-expression with the RBD_{R89L} fusion (Figure 2.6). This result provides strong support for the PINTS concept and its suitability for detecting PPIs in *E. coli*.

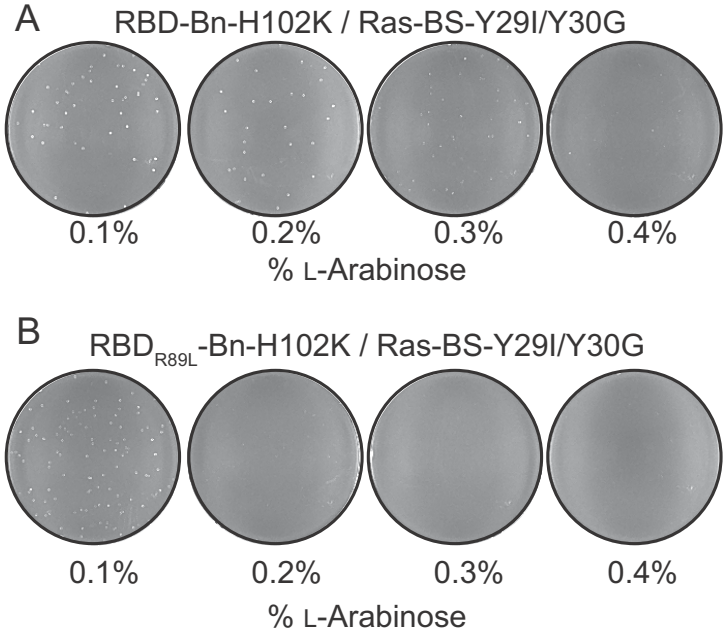


Figure 2.6. Validation of the PINTS assay using the Ras/RBD interaction. (A) The interaction of Ras with the RBD rescues *E. coli* growth. (B) The RBD R89L mutation abolishes the interaction with Ras and prevents *E. coli* growth. Images depict *E. coli* colonies growing on solid agar media containing L-arabinose at the designated concentrations.

2.2.3 Identification of Ras extragenic suppressors of RBD_{R89L}

Given our proof-of-principle results, the PINTS concept appeared promising as a strategy to pan libraries for interacting proteins. The RBD_{R89L} protein provided us with convenient bait. We predicted we could mutate Ras to rescue its binding to the RBD_{R89L}. Our plasmid encoding the RBD_{R89L}-Bn-H102K fusion provided us with a means to screen for extragenic suppressors by panning a Ras

interface mutant library. Intragenic suppressors have been identified within Raf-1 that can restore signaling of Raf-1_{R89L} mutants [210], but to our best knowledge, no Ras extragenic suppressors have been reported. An extragenic Ras suppressor could be a useful tool for biochemists studying Ras signaling.

We set out to generate a library of Ras interface mutants. Our objective was to randomize the Ras interface such that new residues would form a surface complementary to the R89L surface and restore the interaction. We turned to the X-ray crystal structure of the protein complex of Raf-1 RBD and $Rap1a_{GDP}$ [211]. Rap1a is a close homologue of Ras [211]. We targeted five Ras residues in the vicinity of RBD Arg-89 to randomize: Ile-21, Gln-25, Asp-38, Ser-39 and Tyr-40. Each of Asp-38, Ser-39, and Tyr-40 lies on an integral β-strand of the Ras/RBD interface. We selected Asp-38 and Tyr-40 for randomization since their side chains are within hydrogen bond distance (3-4 Å) of the RBD Arg-89 guanidinium moiety and Asp-38 forms a direct hydrogen bond with Arg-89 [211]. We randomized Ser-39 to permit conformational flexibility in the β-strand between the randomized flanking residues. Further, its carbonyl oxygen forms a hydrogen bond with the Arg-89 side chain in the RBD [211]. Ile-21 and Gln-25 are situated on a helix that is slightly more distal to Arg-89 (~ 6-10 Å). We chose to randomize these residues in the event that the R89L mutant displays its side chain in a different rotamer than Arg-89 in the native structure.

We recognize that the simultaneous randomization of five residues could potentially have deleterious effects on protein folding and function [164]. However, the survival selection mode of PINTS affords the screening of a suitably large number of variants to select for those rare variants that fold and function properly. The theoretical library contained 3.2×10^6 protein variants, but the library selection was limited to ~ 10^4 clones.

We assembled our randomized Ras suppressor library and cloned it in place of the WT Ras gene in the Ras-BS-Y29I/Y30G expression cassette. We co-expressed this library with the RBD_{R89L}-Bn-H102K fusion. We first plated *E. coli* transformed with our library of plasmids on growth media containing glucose (to tightly suppress Bn-H102K expression) and IPTG (to induce BS-Y29I/Y30G). From this analysis, we cautiously estimated that 2,000 clones were plated per 10 cm^2 dish. We then plated the library in six dishes supplemented with 0.3% L-arabinose and 1 mM IPTG. 10-20 colonies per dish survived after overnight incubation. Colonies varied in size and most displayed the translucent phenotype observed in our proof-of-concept experiments. However, 2-5 morphologically distinct colonies of larger size and increased opacity grew on each plate. The phenotypic differences led us to compare the restriction enzyme digestion patterns of the DNA plasmids for each colony type. Only the plasmids obtained from slower growing and translucent colonies demonstrated the anticipated restriction pattern (data not shown). The plasmids obtained from the large opaque colonies appeared to have undergone significant rearrangements; plasmids were smaller in size and did not demonstrate the predicted restriction patterns (see Appendix B, Figure B1 for supporting data from an independent assay). This result is consistent with previous studies that describe the DNA plasmids prepared from Bn-H102K-expressing colonies as rearrangement prone [194, 199]. We suggest that in rare instances (<0.1%), genotoxicity is a selective pressure that gives the bacteria with a rearranged plasmid a growth advantage. Fortunately, these clones are easy to distinguish from other colonies. We therein neglected large opaque colonies and only selected the more numerous translucent clones for characterization. We picked 16 clones for sequencing.

Sequence analysis showed no consensus for the residues randomized in Ras (Table 2.3). Given that these clones represent a limited sample set and the library was not screened to redundancy, the lack of a consensus sequence is reasonable. However, sequencing revealed nonsense mutations in Ras suppressors, #7 and #8 (Table 2.3). Since BS-Y29I/Y30G was absent, the reason for the growth of these clones was not certain. One possible explanation was that Bn was not active in these clones for some reason. The Bn genes for these clones were sequenced and variant #7 acquired an additional mutation, R87L. Arg-87 is situated at the Bn/BS interface and is essential to the ribonuclease activity of Bn [195, 212]; it is thought to contribute to active site formation and binding of the phosphate substrate [195]. Therefore, this mutation (coupled with the H102K mutation) likely crippled the ribonuclease activity of Bn and thus permitted colony growth in the absence of BS.

Table 2.3. Substitutions identified in potential Ras suppressors.

	Substitution in Suppressor[1]				
Ras suppressor	**Ile-21**	**Gln-25**	**Asp-38**	**Ser-39**	**Tyr-40**
1	T	H	M	P	G
2	S	P	C	N	R
3	P	A	V	R	A
4	S	L	V	W	W
5	Y	G	F	L	L
6	P	N	Q	R	G
7	STOP	--	--	--	--
8	STOP	--	--	--	--
9	L	W	T	R	L
10	T	P	A	N	G
11	R	R	N	P	N
12	I	S	N	L	W
13	S	R	K	D	T
14	Y	M	F	C	L
15	T	T	D	L	W
16	G	A	F	L	F

[1]Letters in table correspond to the standard single-letter designations for amino acids.

We cannot as yet attribute a cause to the growth of Ras suppressor #8. Given Bn-H102K expression appears to generate a genetically unstable environment, it is reasonable to suggest mutations may have occurred in the promoter region or other transcriptional activation elements upstream of the Bn-H102K gene, each of which may decrease Bn-H102K expression.

Since we observed a deleterious mutation in BnH102K for suppressor #7, we were concerned the other variants may have also acquired genetic defects. From a practical standpoint, sequencing each Bn gene was not desirable. Instead, we simply pooled the DNA from all potential suppressors, excised the Ras genes by *XhoI*/*EagI* restriction digest, and re-cloned them into the PINTS selection plasmid. If most of these Ras genes were false positives, then few colonies, if any, would grow after the genes were sub-cloned back into the original context of Bn-H102K. As a control, we re-cloned the WT Ras gene. After sub-cloning the suppressor library and transforming into *E. coli*, colony growth was dense on agar media, whereas re-cloning WT Ras into the same plasmid produced no colonies. Thus, there is a good likelihood genuine suppressors were contained in our small Ras library.

2.2.4 Validation of RBD_{R89L} extragenic suppressors in live cells

We wanted to confirm our PINTS-derived Ras suppressors were capable of binding the Raf-1 RBD_{R89L}. We decided to test their binding ability under physiological conditions in live cells. We arbitrarily chose four potential Ras suppressors (variants 1, 3, 4, and 5) from our limited sample set, but attempted to fairly represent the sequence diversity by picking variants that did not share any sequence identity at the randomized positions (Table 2.3). We installed the native C-terminal Ras lipidation sequence [213, 214] in each variant for

expression in mammalian cells. Expression of these genes should localize the Ras suppressors to the plasma membrane [213-215]. We also generated an RBD_{R89L}-mCherry fusion to use as a localization marker. We proposed if the interaction between the RBD_{R89L} and Ras was rescued, we would be able to visualize plasma membrane localization by virtue of the fluorescence signal of RBD_{R89L}-mCherry (Figure 2.7).

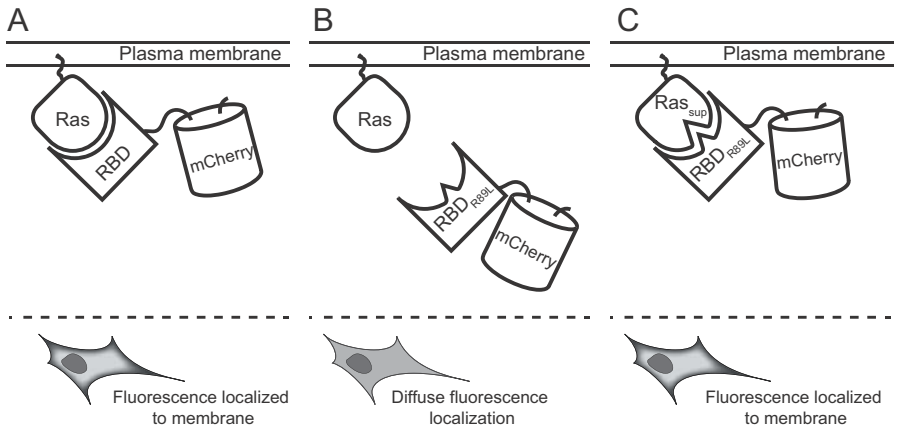


Figure 2.7. Predicted results for live cell demonstration of rescued binding of Ras suppressors. (A) The interaction of WT Ras and RBD will localize mCherry at membrane. (B) The RBD R89L mutation will prevent mCherry membrane localization. (C) Rescued binding of Ras suppressors to the RBD R89L mutant will localize mCherry at the membrane.

We first co-transfected HeLa cells with plasmids encoding WT Ras and WT RBD-mCherry. Cells were imaged for red fluorescence 24 hours post transfection. Fluorescent labeling was clearly distinguishable at the plasma membrane (Figure 2.8A). Conversely, when WT Ras and RBD_{R89L}-mCherry fusions were co-expressed, no membrane localization was visible; only diffuse fluorescence throughout the cytoplasm and nucleus was observed (Figure 2.8B). This result is consistent with the R89L mutation abolishing its interaction with Ras [207, 208]. We then tested our four selected Ras suppressors, but only suppressor #5 demonstrated membrane localization (Figure 2.8C). Importantly,

the localization was specific for the RBD_{R89L} variant; co-expression with WT RBD resulted in diffuse localization (Figure 2.8D). This result suggests that a specific and remodeled interface exists between Ras suppressor #5 and RBD_{R89L}. Unfortunately, Ras variants #1, 3, and 4 showed ubiquitous fluorescence labeling throughout the cytoplasm and nucleus (Figure 2.8 E-G). We repeated these experiments using Ras G12V variants, which render Ras locked in the guanosine triphosphate (GTP) bound state [216, 217]. Localization patterns were identical to those observed for WT Ras clones (Figure 2.9).

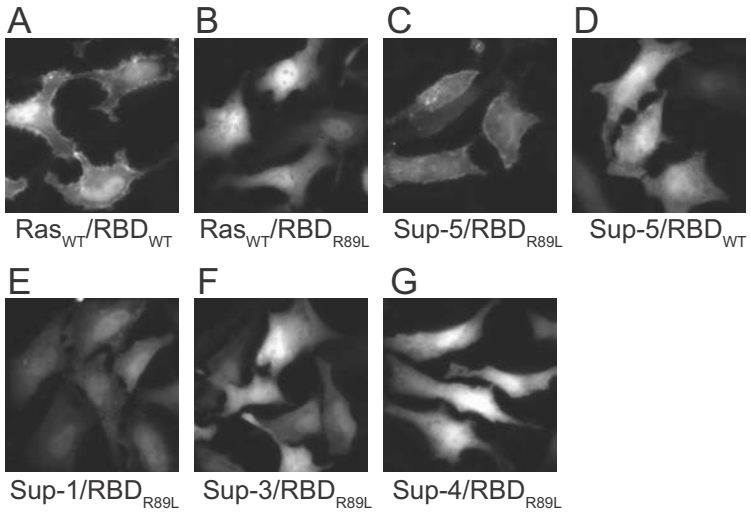


Figure 2.8. Live cell imaging of Ras/RBD interactions. HeLa cells were co-transfected with designated Ras and RBD-mCherry fusions. Plasma membrane localization indicates the occurrence of an interaction. (A) WT Ras binds RBD, but not (B) RBD_{R89L}. (C) Ras suppressor (Sup) #5 binds RBD_{R89L}, but not (D) WT RBD. (E to G) Ras suppressors #1, 3, and 4 do not show binding to RBD_{R89L}.

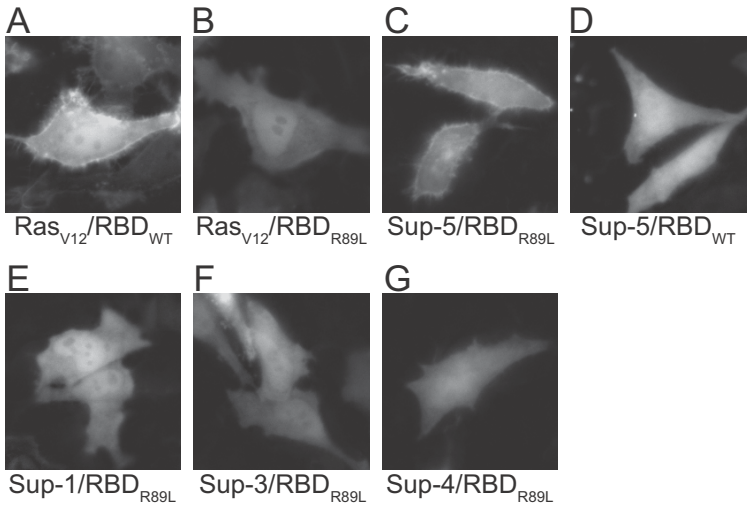


Figure 2.9. Live cell imaging of Ras-G12V/RBD interactions. HeLa cells were co-transfected with designated Ras and RBD-mCherry fusions. Plasma membrane localization indicates the occurrence of an interaction. Each Ras variant shown carries the G12V mutation. (A) WT Ras binds RBD, but not to (B) RBD_{R89L}. (C) Ras suppressor (Sup) #5 binds to RBD_{R89L}, but not to (D) WT RBD. (E to G) Ras suppressors #1, 3, and 4 do not bind to RBD_{R89L}.

2.2.5 Discussion of variation in PINTS clone phenotype

The growth of slow-growing translucent colonies was a hallmark characteristic observed in all PINTS assays performed in this work. The assessment of this phenotype, which we found to be associated with intact plasmids, is not trivial. Furthermore, the molecular determinants of colony opacity remain largely unknown [218]. However, it seems reasonable that colony opacity could be correlated with gene expression and protein production, where increased opacity is associated with elevated protein levels. This has been shown in at least some strains of *E. coli* [218]. We ascribe the decrease in colony opacity to the residual ribonuclease activity of Bn-H102K [201], which will decrease the RNA transcript load and, thus, reduced the accumulation of protein. However, several factors complicate this explanation.

First, Bn will catalyze hydrolysis of its own bait-toxin transcript, as well as the prey-BS transcript. Second, the amount of accumulated soluble prey fusion protein will depend on differences in expression level due to codon usage, folding efficiency, and protease stability. Together these factors dictate that different ratios of Bn-to-BS will exist in each colony during library selection. Not only will the ratio of accumulated proteins vary, but so will the K_d values for the bait/prey interactions. Therefore, a clone-specific Bn/BS equilibrium will determine the extent of genetic stability and toxicity for each colony. For example, efficient folding and slow turnover are advantageous traits that should favour the selection of highly soluble and protease resistant Ras variants. Unfortunately, poorly folded or unstable variants will leave the toxin uninhibited and could thus equally promote growth due to deleterious effects incurred through genetic instability. Therefore, care must be taken to validate growth events as the result of genuine PPIs. Due to the probable instability of PINTS clones, prolonged incubation times should be avoided. Furthermore, surviving clones should be propagated using glucose-supplemented growth media to suppress expression of Bn-H102K until false positives can be ruled out.

2.2.6 Discussion of false positives

The fact that three out of four potential Ras suppressors failed to show an interaction by our live cell imaging approach is disappointing. However, a previous DHFR PCA selection in *E. coli* reported just a single genuine PPI out of 30 potential binding partners picked from a large naïve antibody library panned against a protein antigen [219]. By comparison our 25% success rate is promising, but given the limited number of variants tested, we cannot confidently claim the proportion of survival events resulting from genuine PPIs as of yet.

We made a final attempt to characterize the three apparent false positives by retroactively sequencing the Bn-H102K genes in the original plasmids, but found no errors or truncations. Therefore, although we do not rule out genetic instability as a cause, we cannot directly attribute our false positives to this phenomenon. Instead, reasonable explanations for the apparent unspecific *E. coli* growth and live cell imaging result are the non-specific interaction of bait and prey or the occurrence of genuine, but very low affinity bait/prey interactions.

The induction conditions used for the *in vivo* PINTS selection likely results in high levels of gene expression. Protein overproduction could lead to aggregation and non-specific interactions of bait and prey. Indeed, varying the induction conditions in DHFR PCAs can significantly affect the survival of *E. coli* [219]. However, simple reduction of the induction levels in our system is complicated by the cross-talk of IPTG and L-arabinose inducible promoters [220].

An alternative explanation to non-specific bait/prey interactions is the occurrence of well folded, but very low affinity Ras variants. Single point mutations in Ras are sufficient to decrease its affinity for the RBD by greater than 100 fold (e.g. from 130 nM to 14 μM [208]). The Ras mutation D38A alone reduces the affinity 72 fold [203]. Given that we introduced five mutations into Ras to complement the RBD R89L mutation, it is plausible that we isolated suppressors with very low affinities. Low affinity interactions may still be able to confer survival through PINTS. Since BS-Y29I/Y30G was initially characterized as a minimal suppressor mutant for Bn-H102K [199], low affinity interactions may be just sufficient to breach the survival threshold maintained by BS-Y29I/Y30G. High expression levels (as described in section 2.2.5) of low affinity Ras variants would further perpetuate this phenomenon, as the Ras-BS fusions may be present at concentrations exceeding the K_d for the Ras/RBD_{R89L} interaction.

Interactions of this type are likely unavoidable and represent an inherent drawback to the PINTS principle. If this explanation is correct, our assay design provides no 'reward' for high affinity interactions.

2.3 Conclusion

In this chapter, we describe an assay for the detection of PPIs in *E. coli* based on protein-interaction mediated toxin silencing, or PINTS. Using two sets of FP variants, as well as Ras and RBD proteins, we show that the PINTS strategy is suitable for the survival-mediated identification of PPIs *in vivo*. Further, we were able to successfully isolate a Ras suppressor to RBD_{R89L}, illustrating the PINTS strategy is suitable for identifying interacting partners from large libraries and can be used to discover remodeled interfaces. Although stable clones can be selected, the use of a pre-folded Bn toxin module appears to generate a genetically unstable environment, which limits the broad applicability the PINTS system. Currently, the assay should be conducted with due vigilance and thorough characterization of selected clones to rule out false positives. Nonetheless, we believe that the PINTS assay can serve as a useful tool for biochemists and protein engineers alike.

2.4 Materials and Methods

2.4.1 Recombinant DNA techniques and reagents

All DNA manipulations including small scale preparation of plasmid DNA, restriction enzyme digestion, ligations, and agarose gel electrophoresis were performed according to Sambrook et al. [221]. Where indicated, overlap extension PCR was performed as described in Bessette et al. [222]. Oligonucleotide primers (oligos) were purchased from IDT technologies. Unique

numerical identifiers (e.g. Oligo 2.1) and corresponding sequences for all oligos are compiled in Table A1 (Appendix A).

All restriction enzymes and *Pfu* DNA polymerase were obtained from Fermentas/Thermo Scientific and T4 DNA ligase was obtained from Invitrogen/Life Technologies. The GeneJet® Gel Extraction or Plasmid Miniprep Kits (Fermentas/Thermo Scientific) were used according to the manufacturer's instructions. Ligations were transformed into electrocompetent DH10B *E. coli*. The plasmid pMT416, encoding Bn and BS [200], was obtained from Addgene. The plasmids pQE-Ras DHFR, pQE-Raf DHFR, and pQE-R89L were the kind gift of Dr. S. Michnick [31].

Standard PCR amplifications carried out using *Pfu* DNA polymerase were performed in volumes of 50 µl containing nuclease-free water, 1× reaction buffer, 200 µM dNTPs (Invitrogen), 200 nM forward and reverse oligos, 10 - 50 ng of template DNA, and 1.0 Unit of *Pfu* DNA polymerase. Typical cycling parameters were as follows: initial denaturation at 95°C for 50 seconds; 20 cycles of 95°C for 15 seconds, 55-60°C for 20 seconds, 72°C for 60 seconds per kb of gene target; final extension at 72°C for 4 minutes.

A custom *E. coli* dual expression plasmid (pDES) (credit to Andreas Ibraheem), containing an *XhoI*/*HindIII* polylinker under control of P_{Tac} and an *EcoRI*/*BglII* polylinker under control of P_{BAD}, was used for PINTS assay plasmid construction.

The general cloning strategy for each construct or library is outlined in Table 2.4. Destination plasmids, relevant oligos, and restriction enzyme sites used to clone designated constructs are provided in the table. Brief descriptions of cloning strategies are included in the methods where appropriate.

2.4.2 Sequencing

All sequencing was performed at the Molecular Biology Services Unit (MBSU) at the University of Alberta. All pcDNA3.1 (+) constructs were sequenced with a T7 forward primer (Oligo 2.1) or a BGH reverse primer (Oligo 2.2). Dual expression plasmid (pDES) constructs were sequenced with Oligo 2.3 (forward; P_{BAD} cassette), Oligo 2.4 (reverse; P_{BAD} cassette), Oligo 2.5 (forward; P_{Tac} cassette), and Oligo 2.6 (reverse; P_{Tac} cassette). Ras suppressor variants were sequenced with Oligo 2.28.

2.4.3 PINTS plasmid design

Bait and BnH102K proteins were joined with a 13-resdiue linker (AGTGGGGSGGGGS) and assembled by three-part ligation (Table 2.4). Constructs were designed with convenient restriction sites to permit swapping of genes as required. The bait gene was cloned between 5' *EcoRI* and 3' *KpnI* sites and the BnH102K gene was cloned between 5' *KpnI* and 3' *BglII* sites (Figure 2.2).

Prey and BS-Y29I/Y30G proteins were joined with a 12-residue linker (TAGGGGSGGGGS) and assembled by three-part ligation (Table 2.4). Constructs were designed with convenient restriction sites to permit swapping of genes as required. The prey gene was cloned between 5' *XhoI* and 3' *EagI* sites and the BS-Y29I/Y30G gene was cloned between 5' *EagI* and 3' *HindIII* sites. The BS-Y29I/Y30G gene was generated independently by overlap extension PCR (Table 2.4).

2.4.4 Ras and RBD PINTS Plasmid Design

To generate interface libraries in Ras, an overlap extension PCR method was used (Table 2.4). The amplified product was digested with *XhoI* and *EagI*

and ligated with *XhoI*- and *EagI*-treated PINTS plasmid encoding RBD_{R89L}-Bn-H102K.

2.4.5 Cloning of mammalian expression constructs

Prior to expression in mammalian cells, the native H-Ras C-terminal farnesylation sequence (KLRKLNPPDESGPGCMSCKCVLS) was added to wild type (WT) Ras and Ras suppressors. The overlap extension method was used on two pre-amplified templates to append the lipidation sequence (Table 2.4). The final product was treated with *XhoI* and *HindIII* enzymes and ligated into *XhoI*- and *HindIII*-treated pcDNA3.1 (+). Raf-mCherry fusions were also assembled by overlap PCR using two pre-amplified templates (Table 2.4). The final product was treated with *XhoI* and *HindIII* enzymes and ligated into *XhoI*- and *HindIII*-treated pcDNA3.1 (+). Oligos used in the above amplifications incorporated Kozak consensus sequences.

2.4.6 Live cell imaging

HeLa cells were maintained in Dulbecco's modified eagle medium (DMEM) with 10% fetal bovine serum at 37°C and 5% CO_2 according to standard procedures. Transient transfections were performed using Turbofect (Fermentas/Thermo Scientific). Designated expression plasmids (375 ng each) were transfected according to the manufacturer's instructions. Generally, 2 µl of transfection reagent was used per 1 µg of DNA transfected. Transfected cells were imaged 24 hours following transfection in HEPES-buffered Hank's balanced salt solution on an inverted Nikon Eclipse Ti microscope equipped with a 200W metal halide lamp (PRIOR Lumen) and a QuantEM: 512SC 16-bit cooled CCD camera (Photometrics).

Table 2.4. Cloning strategy for constructs described in Chapter 2.

		Amplification 1		Gene Assembly	
Construct	**Plasmid**	**Template**	**Oligos[1]**	**Oligos**	**RE Sites (5'/3')**
dTFP bait	pPINTS	dTFP0.7	2.7/2.8	n/a[2]	*EcoRI/ KpnI*
dTFP prey	pPINTS	dTFP0.7	2.11/2.12	n/a	*XhoI/ EagI*
BnH102K	pPINTS	pMT416	2.9/2.10	n/a	*KpnI/ BglII*
BS-Y29I/Y30G	pPINTS	pMT416	2.13/2.27 2.14/2.26	2.13/2.14	*EagI/ HindIII*
mTomato bait	pPINTS	dTomato	2.29/2.7 2.30/2.8	2.7/2.8	*EcoRI/ KpnI*
dTomato prey	pPINTS	dTomato	2.11/2.12	n/a	*XhoI/ EagI*
Ras prey	pPINTS	pQE-Ras DHFR	2.31/2.33	n/a	*XhoI/ EagI*
RBD bait	pPINTS	pQE-Raf DHFR	2.32/2.34	n/a	*EcoRI/ KpnI*
RBD R89L bait	pPINTS	pQE-R89L DHFR	2.32/2.34	n/a	*EcoRI/ KpnI*
Ras suppressor library	pPINTS	Ras-pPINTS	2.31/2.23 2.24/2.25	2.31/2.25	*XhoI/ EagI*
Ras-C-Lip	pcDNA3.1 (+)	Ras/RBD pPINTS pcDNA3.1 (+)	2.15/2.16 2.17/2.18	2.15/2.18	*HindIII/ XhoI*
Ras-sup# C-Lip	pcDNA3.1 (+)	Ras-sup # 1,3,4 or 5 pcDNA3.1 (+)	2.15/2.16 2.17/2.18	2.15/2.18	*HindIII/ XhoI*
RBD-mCherry	pcDNA3.1 (+)	Ras/RBD-pPINTS mCherry2	2.22/2.21 2.19/2.20	2.22/2.20	*HindIII/ XhoI*
RBD R89L-mCherry	pcDNA3.1 (+)	Ras/R89L-pPINTS mCherry2	2.22/2.21 2.19/2.20	2.22/2.20	*HindIII/ XhoI*
RasV12-C-Lip	pcDNA3.1 (+)	Ras-C-Lip	2.35/2.18	n/a	*HindIII/ XhoI*
RasV12-sup# C-Lip	pcDNA3.1 (+)	Ras-sup#1,3,4 or 5 C-Lip	2.35/2.18	n/a	*HindIII/ XhoI*

[1] Oligonucleotide sequences are listed in Appendix A (Table A1). [2] Not applicable

Chapter 3 Development of a fluorogenic red fluorescent protein heterodimer[2]

3.1 Introduction

Fluorescent protein (FP)-based biosensors that report on biochemical events in live cells typically operate on one of three principles: the modulation of FRET efficiency between two FPs of different hues [116], the modulation of the fluorescent intensity of a single FP through analyte-induced changes in the chromophore environment [161, 223], or the fluorogenic reconstitution of a single FP (also known as complementation) from two polypeptide chains [224]. Each of these design principles is associated with distinct advantages. For example, FRET-based biosensors benefit from being inherently ratiometric in their response and single FP-based biosensors often provide better signal-to-noise due to larger intensiometric responses. FP complementation uniquely generates an irreversible response and can be applied to detect interacting proteins in proteome-wide screens due to negligible background signal [141].

A shortcoming shared between these three design strategies is that they are most commonly implemented with engineered variants of *A. victoria* GFP and only rarely implemented with Anthozoan-derived RFPs [49]. While there have been reports of orange FP-RFP FRET pairs [115, 136], single RFP-based Ca^{2+}-indicators [152], and complementation of split RFPs [146, 147], such examples are relatively few in number. Among the contributing factors to the discrepancy between the popularity of GFP and RFP variants in biosensing applications is the

[2] A version of this chapter has been published. Alford, S.C., et al. 2012 *Chemistry & Biology*. *19*(3): 353-360.

poor sensitized emission of RFPs when used as FRET acceptors [114], and the challenge of engineering circularly permuted RFPs for use in single FP-based biosensors [225].

We reasoned that a new approach, which better harnesses the inherent properties of RFPs, could serve to stimulate development of a broader selection of RFP-based biosensors. One property that differentiates GFP from naturally occurring RFPs is the obligate tetrameric structure of the latter [71]. For many applications, the tetrameric structure is undesirable and substantial effort has gone into engineering dimeric and monomeric FPs [9, 11]. However, it is also apparent that the oligomeric structure of RFPs helps stabilize the chromophore in a conformation that favours bright fluorescence. This role is apparent when one compares the quantum yields (Φ) of dimeric (*i.e.*, dimer2 and dTomato; both 0.69) and monomeric (*i.e.*, mRFP1 and mCherry; 0.25 and 0.22, respectively) variants of DsRed [9, 49, 98]. This dimerization-dependent brightness suggested to us a new strategy for the creation of RFP-based biosensors.

Our idea was to develop an alternative biosensing system by engineering the oligomeric structure of RFPs to create a low affinity heterodimer that exhibits bright red fluorescence in the associated state and dim fluorescence in the dissociated state (Figure 3.1).

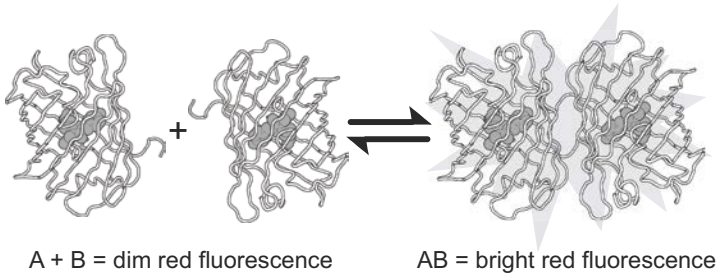


Figure 3.1. Proposal for engineering a low affinity fluorogenic heterodimer FP. An RFP heterodimer that is more brightly fluorescent in the associated state than in the dissociated state provides a new strategy for the construction of genetically encoded indicators.

We expected that two key challenges in engineering such a system would be: 1) to sufficiently decrease the affinity such that the dimer partners were not fully associated at intracellular concentrations of ~1-50 μM [226] and 2) to obtain high contrast between the associated and dissociated states. Here we describe our efforts to overcome these challenges and produce a fluorogenic RFP heterodimer that has proven useful in a series of representative biosensing applications.

3.2 Results and Discussion

3.2.1 Directed evolution and characterization

Our strategy for generating a fluorogenic RFP heterodimer was inspired by the well-established practice of introducing interface breaking mutations to produce a monomeric FP from a dimeric precursor [9, 11]. Accordingly, we created a monomeric and dimly fluorescent RFP variant (designated **A**) by introducing the H162K and A164R substitutions into dTomato [9]. We then set out to rescue dimer formation and fluorescence by creating a dTomato-derived partner (designated **B**) with surface modifications that complemented the modifications to **A**. To achieve this goal, three distinct library screening strategies were developed and implemented during an extensive process of directed evolution. For clarity we append a subscript number to **A** and **B** to designate the generation of a given variant. Thus, dTomato-H162K/A164R is designated as $\mathbf{A}_{0.1}$ and the most extensively optimized version of **A** is designated $\mathbf{A}_{1}$. A summary of the clones obtained during the directed evolution is provided in Appendix C (Table C1).

To identify a heterodimeric partner for $\mathbf{A}_{0.1}$, we developed a two-step assay that involved an image-based screen for red fluorescence of *E. coli* colonies

followed by an electrophoretic mobility shift analysis of protein extracts (Figure 3.2). A dual expression plasmid with two different promoters (P_{Tac} and P_{BAD}) was used to express $\mathbf{A}_{0.1}$ as a fusion with the 43 kDa *E. coli* maltose binding protein (MBP), and a dTomato-R149X/H162X/Y192X (where X = all 20 amino acids) gene library. These three residues were targeted due to their proximity to H162 and A164 across the hydrophilic dimer interface of tetrameric DsRed [74]. Colonies with the brightest red fluorescence were picked, cultured, and the protein extracts analyzed by SDS-PAGE under conditions that preserve high affinity oligomeric interactions of FPs [9, 71].

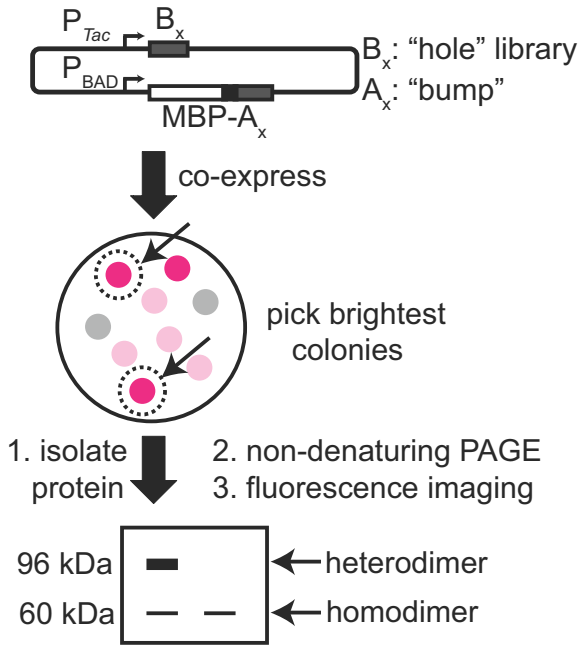


Figure 3.2. Schematic representation of electrophoretic mobility shift screen for heterodimeric pairs of RFPs.

Heterodimeric proteins migrated slower than $\mathbf{B}_{0.1}$ homodimers during PAGE analysis (Figure 3.3A). Screening of ~20,000 colonies for fluorescence, and approximately 100 proteins by gel-shift, led to the identification of a pool of 25 variants ($\mathbf{B}_{0.1,\ pool}$) that exhibited some heterodimeric character (Figure 3.3B).

However, even the best of these variants existed as a mixture of the homodimeric and heterodimeric states (Figure 3.3B).

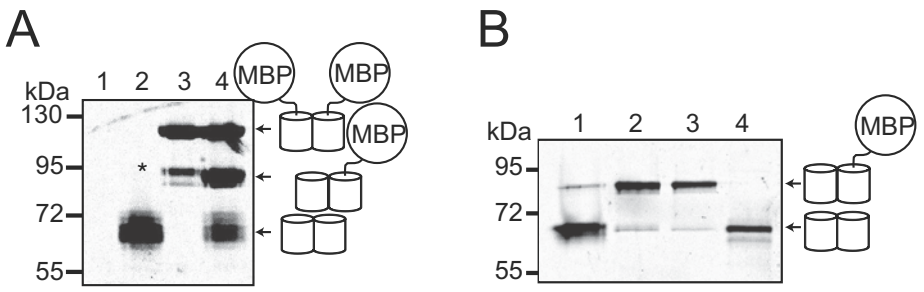


Figure 3.3. Characterization of homo- and heterodimeric structure by gel electrophoresis. The contrast of each whole image has been adjusted to emphasize weak bands. (A) Validation of the SDS-PAGE electrophoretic mobility shift screen for pairs of protein variants that exhibit heterodimeric character. Proteins were detected by in-gel imaging of red fluorescence. Lane 1, uninduced culture; lane 2, dTomato; lane 3, MBP-dTomato; lane 4, co-expressed dTomato and MBP-dTomato. Asterisk (*) indicates heterodimer resulting from proteolysis of the dTomato-MBP linker. (B) Image of the red fluorescence of a gel used for PAGE analysis of four representative variants (numbered 1-4) assayed by the gel mobility-shift strategy.

In an effort to engineer variants that existed primarily as heterodimers, we flipped the identity of two residues (R153 and E100) that form an ion-pair interaction across the interface (Figure 3.4). Specifically, we introduced R153E into $\mathbf{A}_{0.1}$ to produce $\mathbf{A}_{0.2}$ and E100R into a $\mathbf{B}_{0.1,\,pool}$ to create $\mathbf{B}_{0.2,\,pool}$. The R153E mutation had the additional benefit of increasing the amount of soluble protein produced for the **A** variant. We then created a second generation library from the template of $\mathbf{B}_{0.2,\,pool}$ by randomizing positions E160 and H162. Screening of this library led to the identification of $\mathbf{B}_{0.3}$ which was equivalent to dTomato-E100R/R149L/E160H/H162F/Y192G and, by PAGE analysis, predominantly formed a heterodimeric complex with $\mathbf{A}_{0.2}$. To further diminish the residual homodimeric character, we removed three residues that contribute to the dTomato interface (F224, L225, and Y226) [74] from the C-termini of $\mathbf{A}_{0.2}$ and

$\mathbf{B}_{0.3}$. The resulting proteins, $\mathbf{A}_{0.3}$ and $\mathbf{B}_{0.4}$, exclusively formed a heterodimer, but exhibited reduced fluorescent brightness.

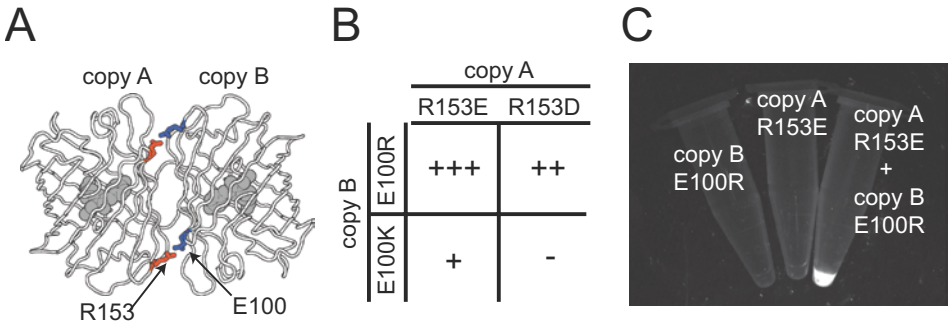


Figure 3.4. Flipping the identity of two residues in a salt-bridge across the dimer interface effectively converts a homodimeric protein into a heterodimeric protein. (A) Schematic overview of the effect of flipping the orientation of the ion-pair. (B) Relative fluorescence intensities of cultures co-expressing two dTomato genes with various combinations of potential salt-bridge forming substitutions. (C*)* Fluorescence image of cell pellets for cultures of *E. coli* expressing dTomato-E100R, dTomato-R153E, or both dTomato-E100R and dTomato R153E. We noted that R153E increased the amount of soluble protein produced, while E100R had no apparent effect on protein solubility.

We next turned to a colony-based assay in which replica plating was used to identify non-covalent **AB** pairs that were brightly fluorescent when both proteins were expressed, but dimly fluorescent when only one protein was expressed (Figure 3.5A). Briefly, we used the dual expression plasmid to express one randomly mutated partner under the IPTG-inducible P_{Tac} promoter, and one genetically-fixed partner under the L-arabinose-inducible P_{BAD} promoter. For library screening, transformed *E. coli* was plated on media permissive for expression of both partners (*i.e.*, supplemented with IPTG and L-arabinose). The brightest colonies were picked and manually arrayed onto two new plates: one with media for expression of both partners and one with media for expression of only the variable partner (*i.e.*, supplemented with IPTG and glucose). The red fluorescence of each plate was digitally imaged and then processed in pairs to determine the fluorogenic contrast for each colony. We performed several rounds

of screening in which $\mathbf{B}_{0.4}$-derived variants were randomized and $\mathbf{A}_{0.3}$ was held constant, followed by several rounds with $\mathbf{A}_{0.3}$-derived libraries and $\mathbf{B}_{0.5}$ held constant.

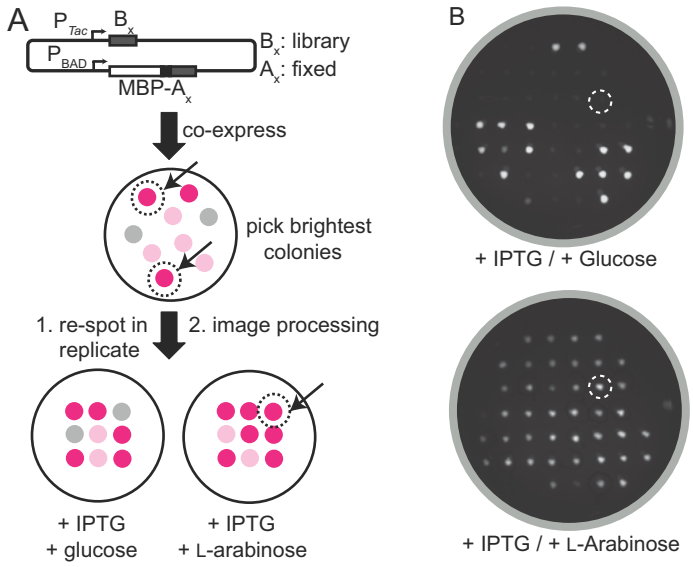


Figure 3.5. Replica-plate screening for fluorogenic and heterodimeric pairs of RFPs. (A) Schematic of replica plating strategy. (B) Fluorescence images of a pair of representative plates from the replica plating screen. Top plate contains IPTG and glucose. Bottom plate contains IPTG and L-arabinose. The faint shadows adjacent to many colonies are due to reflection off the bottom of the clear plastic dish. The dashed circle indicates a high contrast clone that was picked for further evolution.

This procedure led to our first fluorogenic dimerization-dependent RFP composed of $\mathbf{A}_{0.4}$ and $\mathbf{B}_{0.5}$ (ddRFP-$\mathbf{A}_{0.4}\mathbf{B}_{0.5}$). Characterization of ddRFP-$\mathbf{A}_{0.4}\mathbf{B}_{0.5}$ revealed that the FP partners exhibited a 5-fold increase in fluorescence upon dimerization both in *E. coli* (Figure 3.6A) and *in vitro* (Figure 3.6B).

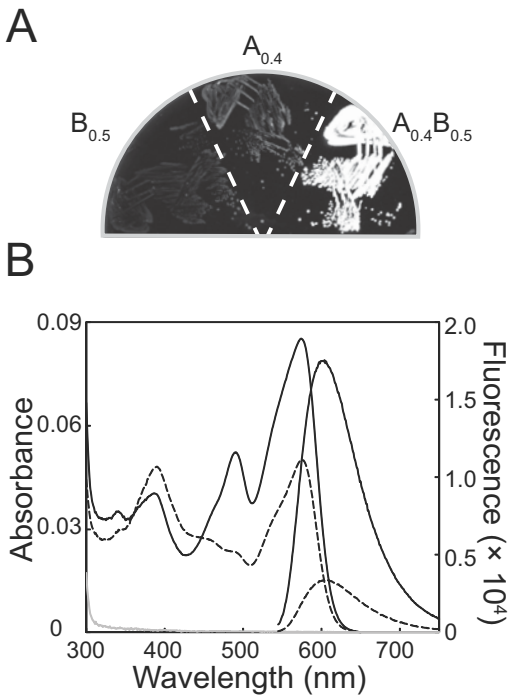


Figure 3.6. Characterization of ddRFP-$\mathbf{A}_{0.4}\mathbf{B}_{0.5}$. (*A*) Fluorescence image of *E. coli* expressing heterodimer constituents $\mathbf{A}_{0.4}$, $\mathbf{B}_{0.5}$ and tandem dimer $\mathbf{A}_{0.4}\mathbf{B}_{0.5}$. (*B*) Absorbance and fluorescence emission spectra for $\mathbf{A}_{0.4}$ (dashed lines; 10 μM), $\mathbf{B}_{0.5}$ (solid grey lines; 10 μM), and the equimolar mixture of $\mathbf{A}_{0.4}\mathbf{B}_{0.5}$ (solid black lines; 10 μM each).

To further improve the fluorogenic contrast of ddRFP-$\mathbf{A}_{0.4}\mathbf{B}_{0.5}$, we developed a third screening procedure that involved the use of proteolyzable tandem heterodimers (Figure 3.7) composed of **A** and **B** joined by a 23-residue trypsin-labile linker. Libraries of randomly mutated heterodimer genes were expressed in *E. coli* and the brightest colonies picked and cultured. The fluorescence of protein extracts was measured before and after treatment with trypsin, and clones with the greatest contrast were carried on to subsequent rounds. Application of this strategy for several rounds provided only moderate improvements in contrast, so we next attempted to rationally engineer improved contrast.

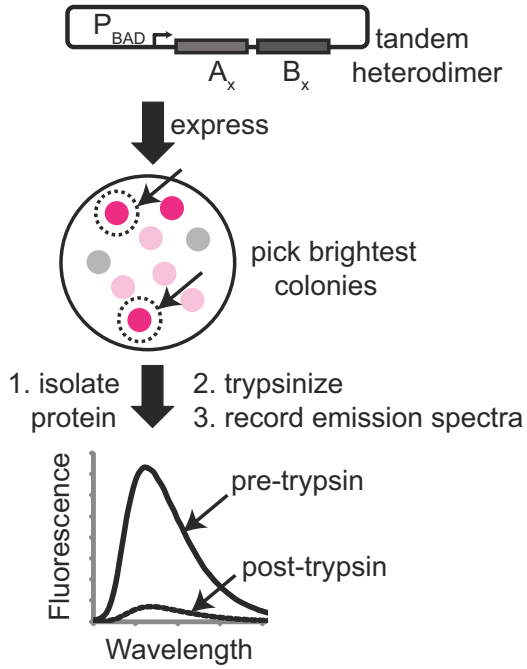


Figure 3.7. Tandem heterodimer proteolysis screening procedure used to identify high contrast fluorogenic heterodimeric RFPs.

Since the fluorescence contrast of ddRFP is limited by the residual fluorescence of **A** in the absence of **B** (Figure 3.6A and B), we reasoned that manipulating the intraprotein interactions of the chromophore in partner **A** could lead to variants with improved contrast. Accordingly, we introduced the S146A substitution into $\mathbf{A}_{0.4}$ to produce $\mathbf{A}_{0.5}$ which removed a hydrogen bond to the phenolate oxygen of the **A** chromophore [74]. From a library in which positions 144 and 145 of $\mathbf{A}_{0.5}$ were randomized, we picked clones that were most dimly fluorescent when expressed in the absence of **B**. The resulting $\mathbf{A}_{0.6,\,pool}$ was fused in tandem to a $\mathbf{B}_{0.5}$-derived gene library in which three residues (192, 194, and 222) in close proximity to residues 144-146 of the dimer partner were randomized [74]. To increase the K_d of the heterodimer, we also reverted E100R and installed R153E in the **B** partner such that two favourable arginine-glutamate electrostatic

interactions were replaced with unfavourable glutamate-glutamate contacts. The library was screened as tandem heterodimers and clones with substantially improved contrast were identified. The **A** and **B** partners of the clone that exhibited the highest contrast were designated $\mathbf{A}_1$ and $\mathbf{B}_1$.

Purified $\mathbf{A}_1$ and $\mathbf{B}_1$ exhibit a 10-fold increase in fluorescence upon dimerization (Figure 3.8) and an apparent K_d of 33 μM (Figure 3.9). This contrasts with the apparent K_d of 1.1 μM observed for $\mathbf{A}_{0.4}$ and $\mathbf{B}_{0.5}$ (Figure 3.9).

As with ddRFP-$\mathbf{A}_{0.4}\mathbf{B}_{0.5}$, interaction with $\mathbf{B}_1$ results in an instantaneous increase in the fluorescent brightness of $\mathbf{A}_1$. $\mathbf{A}_1$ alone has $\Phi = 0.026$ and $\varepsilon = 11{,}800\ M^{-1}\cdot cm^{-1}$ at pH 7.4 and these values increase to 0.074 and $48{,}300\ M^{-1}\cdot cm^{-1}$, respectively, upon formation of a heterodimer with $\mathbf{B}_1$. No visible wavelength absorbance or fluorescence could be detected for $\mathbf{B}_1$, even after several weeks at 4°C.

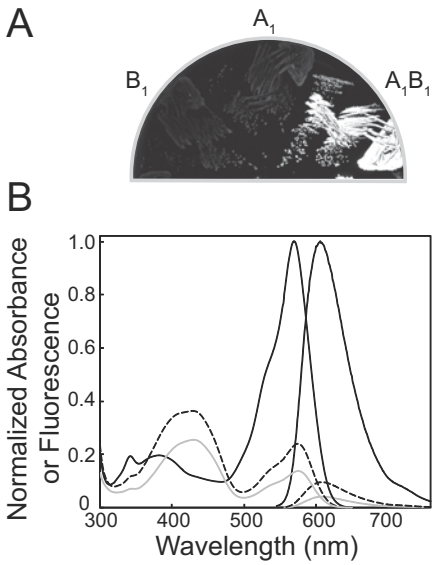


Figure 3.8. Characterization of ddRFP-$\mathbf{A}_1\mathbf{B}_1$. (A) Fluorescence image of *E. coli* expressing heterodimer constituents $\mathbf{A}_1$, $\mathbf{B}_1$, and tandem dimer $\mathbf{A}_1\mathbf{B}_1$. (B) Absorbance and fluorescence spectra of $\mathbf{A}_1$ (grey line; 20 μM) and tandem ddRFP-$\mathbf{A}_1\mathbf{B}_1$ (black line; 20 μM) before (solid line) and after (dashed line) treatment with trypsin. $\mathbf{B}_1$ has baseline absorbance and fluorescence.

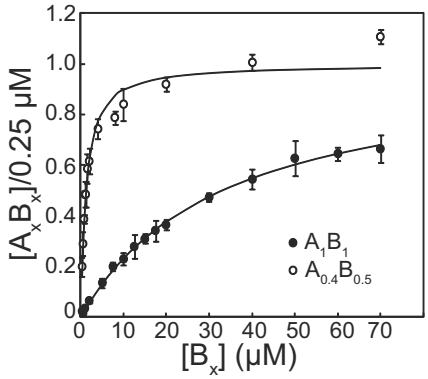


Figure 3.9. Saturation binding curves for fluorogenic RFP heterodimers. Binding curves are shown for $\mathbf{A}_{0.4}$ with $\mathbf{B}_{0.5}$ (white circles) and $\mathbf{A}_1$ with $\mathbf{B}_1$ (black circles). K_d values are 1.1 and 33 µM, respectively. Average of three independent experiments (± standard deviation).

Furthermore, boiling of purified $\mathbf{B}_1$ resulted in none of the backbone hydrolysis that is characteristic of DsRed-type proteins (including $\mathbf{A}_1$) with mature chromophores (Figure 3.10) [73]. We observed no limits to the solubility of $\mathbf{A}_1$ and $\mathbf{B}_1$; expression in *E. coli* gave yields of 10-80 mg per liter.

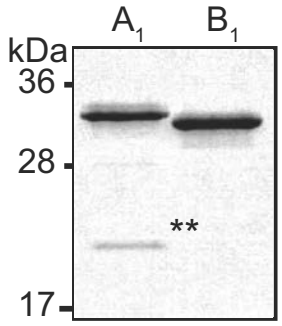


Figure 3.10. Coomassie stained SDS-PAGE of purified $\mathbf{A}_1$ and $\mathbf{B}_1$ variants. Samples were boiled in sample buffer prior to electrophoresis. The double asterisk (**) indicates the 19kDa product of chromophore hydrolysis routinely observed for fluorescent proteins [73]. The 11 kDa product stains too faintly to be observed by Coomassie stain.

Both $\mathbf{A}_1$ and the $\mathbf{A}_1\mathbf{B}_1$ complex exhibit relatively high apparent fluorescence pK_as (9.0 and 7.4, respectively) and thus their fluorescence is sensitive to physiologically relevant changes in pH (Figure 3.11). Size exclusion

chromatography confirmed that both $\mathbf{A}_1$ and $\mathbf{B}_1$ are monomeric in isolation, but form a heterodimer when mixed (Figure 3.12).

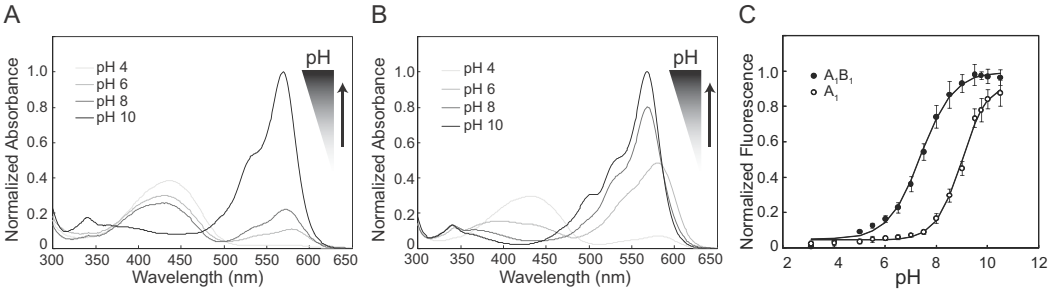


Figure 3.11. pH sensitivity of ddRFP-$\mathbf{A}_1\mathbf{B}_1$. (A) Absorbance spectra of $\mathbf{A}_1$ at various pH values. (B) Absorbance spectra of ddRFP-$\mathbf{A}_1\mathbf{B}_1$ at various pH values. (C) Fluorescence intensity of ddRFP-$\mathbf{A}_1\mathbf{B}_1$ and $\mathbf{A}_1$ alone as a function of pH. The apparent pK_as are 7.4 and 9.0, respectively. Fluorescence intensity data represents at least three independent experiments ± standard deviation.

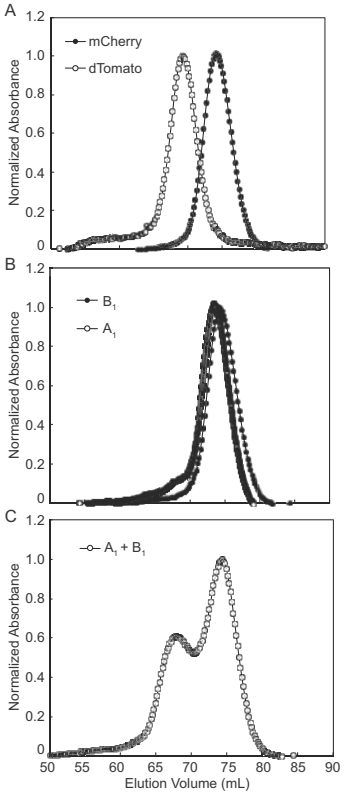


Figure 3.12. Determination of oligomerization state of $\mathbf{A}_1$ and $\mathbf{B}_1$ by size exclusion chromatography. Designated proteins were resolved on a HiLoad 16/60 Superdex 75 pg gel filtration column using 20 mM Tris-Cl, 200 mM NaCl, pH 7.5. (A) Elution profiles of monomeric mCherry and dimeric dTomato were used as oligomeric size standards. (B) Elution profiles of $\mathbf{A}_1$ and $\mathbf{B}_1$ (loaded at 30 μM). (C) Elution profile of an $\mathbf{A}_1$ plus $\mathbf{B}_1$ mixture (100 μM each).

Interestingly, we noted that $\mathbf{A_1B_1}$ undergoes a reversible reaction with β-mercaptoethanol (possibly a nucleophilic addition to the chromophore [227]) that results in loss of the red fluorescent state and the formation of a new species that absorbs at 402 nm and emits at 466 nm (Figure 3.13).

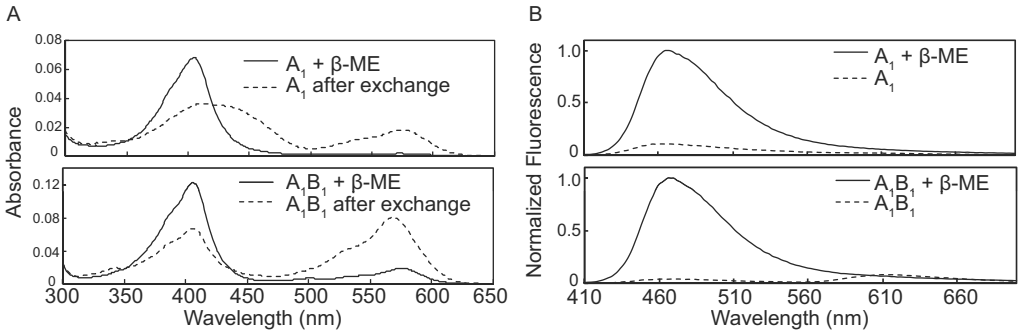


Figure 3.13. The effect of β-mercaptoethanol on ddRFP-$\mathbf{A_1B_1}$. (A) Absorbance spectra of $\mathbf{A_1}$ and $\mathbf{A_1B_1}$ (tandem heterodimer) treated with 2mM β-mercaptoethanol (β-ME) and following exchange into β-ME free buffer. (B) Emission spectra of $\mathbf{A_1B_1}$ (tandem heterodimer) and $\mathbf{A_1}$ alone before and after treatment with β-ME. Spectra were collected after excitation at 385nm.

3.2.2 *In vitro* and live cell applications

We next investigated the utility of ddRFP-$\mathbf{A_1B_1}$ in representative *in vitro* and live cell imaging applications. To determine whether ddRFP-$\mathbf{A_1B_1}$ could be used to detect protein-protein interactions (PPIs) *in vitro*, we turned to the rapamycin-dependent interaction of FK506 binding protein (FKBP) and the FKBP-rapamycin binding domain (FRB) [228]. We fused $\mathbf{A_1}$ and $\mathbf{B_1}$ to the C-termini of FRB and FKBP with a 43-residue unstructured linker to produce FRB-$\mathbf{A_1}$ and FKBP-$\mathbf{B_1}$. These fusions were expressed in *E. coli*, purified, and mixed at a concentration (0.5 μM each) that is well below the K_d for ddRFP-$\mathbf{A_1B_1}$. As shown in (Figure 3.14), addition of rapamycin gave a dose dependent increase in fluorescence. This result indicates that ddRFP-$\mathbf{A_1B_1}$ is useful for the detection of PPIs with dissociation constants well below its K_d.

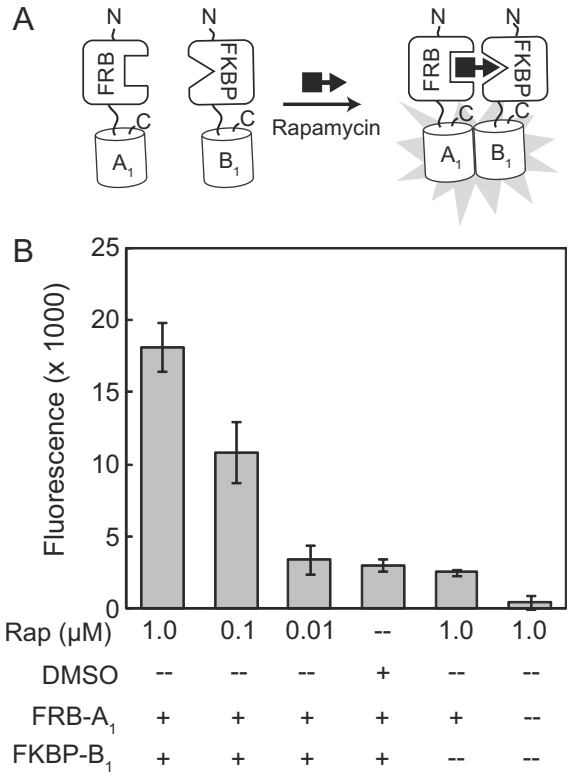


Figure 3.14. Fluorescence-based detection of the rapamycin-inducible interaction between FRB-$\mathbf{A}_1$ and FKBP-$\mathbf{B}_1$. (A) Schematic representation of the interaction. (B) Fluorescence in the absence and presence of rapamycin. Proteins were mixed together at 0.5 μM and rapamycin was added at the indicated concentrations. Dimethyl sulfoxide (DMSO) was added as a vehicle control where appropriate. The data shown represents the mean integrated emission peak intensities for three independent experiments (± standard deviation) and corrected for background (*i.e.*, assay buffer).

To determine if ddRFP-$\mathbf{A}_1\mathbf{B}_1$ could be employed to detect reversible PPIs in live cells, we borrowed the so-called 'split-cameleon' strategy in which the Ca^{2+}-induced interaction of calmodulin (CaM) and M13 peptide brings two FPs into close proximity [117]. Accordingly, mammalian expression plasmids were prepared in which $\mathbf{A}_1$ was fused to the N-terminus of CaM ($\mathbf{A}_1$-CaM) and $\mathbf{B}_1$ was fused to the C-terminus of the M13 peptide (M13-$\mathbf{B}_1$) (Figure 3.15A). HeLa cells were co-transfected with both plasmids and imaged at 24 hours post transfection.

Upon histamine treatment, we observed oscillations in the red fluorescence intensity (Figure 3.15B) that were consistent with the results obtained for previously reported Ca^{2+} indicators [117]. *In situ* calibration of the dynamic range was accomplished by treating cells with ionomycin/EGTA to deplete Ca^{2+} and ionomycin/Ca^{2+} to saturate CaM. These experiments revealed that the dynamic range for $\mathbf{A}_1$-CaM plus M13-$\mathbf{B}_1$ in cells was 5.7 ± 1.3 (N= 16). This dynamic range is diminished relative to the *in vitro* range for $\mathbf{A}_1$ and $\mathbf{B}_1$ alone, indicating that the partners are partially associated at the intracellular protein concentrations of these experiments.

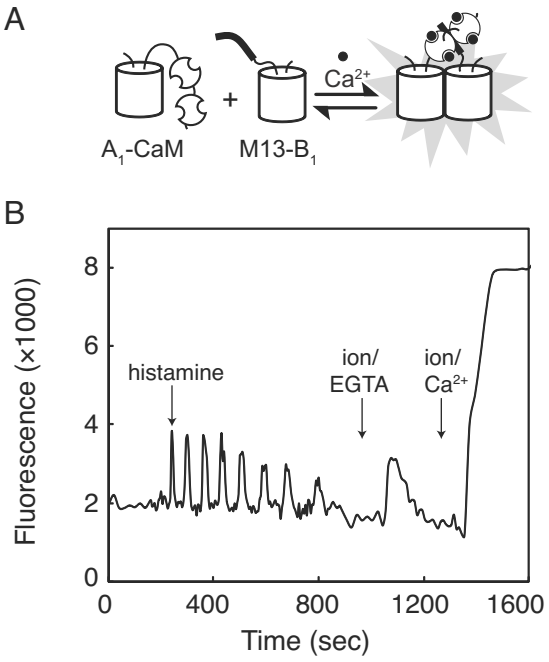


Figure 3.15. Imaging of Ca^{2+} dynamics in live cells using ddRFP-$\mathbf{A}_1\mathbf{B}_1$. **(**A**)** Schematic of transfected constructs $\mathbf{A}_1$-CaM and M13-$\mathbf{B}_1$ in HeLa cells. (B) Representative live cell trace of a transfected HeLa cell treated with histamine, followed by EGTA/ionomycin (ion) and Ca^{2+}/ionomycin. Red fluorescence was imaged as a function of time.

For a third representative application of ddRFP-$\mathbf{A}_1\mathbf{B}_1$, we envisioned a strategy for protease activity sensing in which $\mathbf{A}_1$ and $\mathbf{B}_1$ were fused in tandem with the well-characterized Aspartate-Glutamate-Valine-Aspartate (DEVD)

caspase-3 substrate sequence in the linker region [229]. As in our trypsinolysis assay, cleavage of the substrate should result in a loss of fluorescence as the partners dissociate, provided the concentration is well below the K_d (Figure 3.16A). As a non-proteolyzable control, we created an analogous construct with Serine-Alanine-Serine-Glycine (SASG) in place of DEVD. HeLa cells transfected with a mammalian expression plasmid harbouring the DEVD tandem heterodimer exhibited bright red fluorescence (Figure 3.16B). When these cells were treated with Tumor necrosis factor-α (TNF-α) and imaged through time, we reliably observed a decrease in red fluorescence intensity, with a dynamic range of 9.7 ± 1.9 (N=8), just prior to the cell shrinkage and blebbing associated with the end stages of apoptosis (Figure 3.16B). No comparable loss of red fluorescence signal was observed with the SASG control construct (Figure 3.17). These experiments confirm that the affinity of $\mathbf{A}_1$ and $\mathbf{B}_1$ is sufficiently low to allow protease detection, and imply that the brightness of the heterodimer is sufficient for imaging of intracellular concentrations that are well below the K_d of 33 μM.

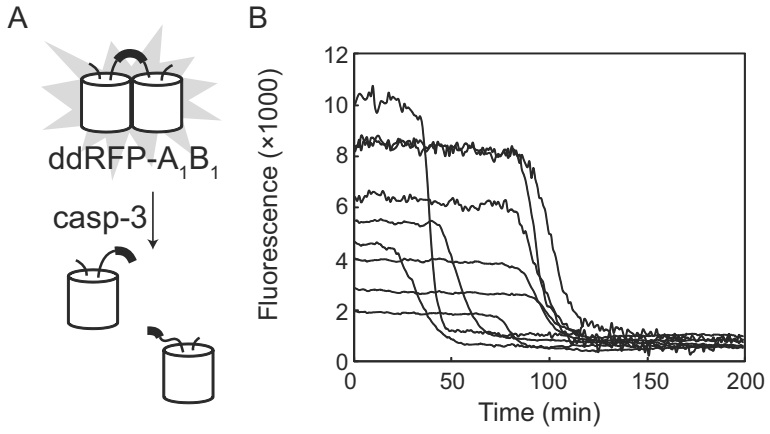


Figure 3.16. Imaging of caspase-3 (casp-3) activation in live cells using ddRFP-$\mathbf{A}_1\mathbf{B}_1$. (A) Schematic of transfected tandem heterodimer with a DEVD substrate in the linker region. (B) Transfected HeLa cells were treated with TNF-α (t = 0) to initiate apoptosis and red fluorescence was imaged as a function of time. Shown are several representative caspase-3 activity responses. The live cell traces depicted in Figure 3.16B are credited to Yidan Ding.

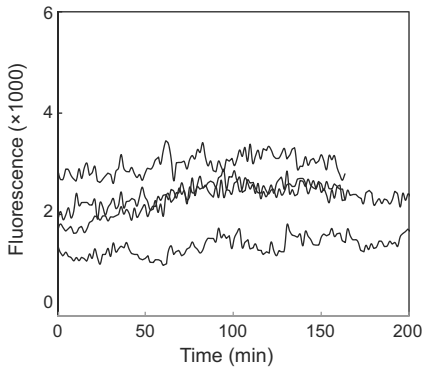


Figure 3.17. Imaging of caspase-3 activity with a ddRFP-$\mathbf{A}_1\mathbf{B}_1$ tandem heterodimer with a SASG non-cleavable substrate in the linker region. Transfected HeLa cells were treated with TNF-α (t = 0) to initiate apoptosis and red fluorescence was imaged as a function of time. Shown are several representative caspase-3 activity responses.

3.2.3 Discussion of fluorogenesis

Two factors contribute to the fluorogenesis of ddRFP-$\mathbf{A}_1\mathbf{B}_1$: pK_a modulation and Φ modulation. At physiological pH and in the absence of $\mathbf{B}_1$, the fully formed chromophore of $\mathbf{A}_1$ exists primarily in the protonated and non-fluorescent state due to its pK_a of 9.0. Upon interaction with $\mathbf{B}_1$, the chromophore environment of $\mathbf{A}_1$ is modified such that the pK_a is lowered to 7.4 and the equilibrium shifts toward the anionic fluorescent state. A similar mechanism has been proposed for a system in which a single domain antibody (nanobody) modulates the fluorescence of GFP [230]. An additional contribution to fluorogenesis is the increase in Φ from 0.026 in $\mathbf{A}_1$ to 0.074 in the $\mathbf{A}_1\mathbf{B}_1$ complex. We speculate that in free $\mathbf{A}_1$ the chromophore spends more time in non-coplanar conformations that favour excited-state deactivation by processes other than fluorescence [88]. Upon formation of the $\mathbf{A}_1\mathbf{B}_1$ complex, a modified chromophore environment stabilizes a coplanar conformation of the chromophore, and Φ increases. Further structural studies should shed light on the molecular details of the interface interactions that contribute to fluorogenesis.

3.2.4 Homology modeling of ddRFP-$\mathbf{A_1B_1}$

In total, the directed evolution of ddRFP- $\mathbf{A_1B_1}$ resulted in 35 amino acids substitutions relative to dTomato, in addition to three-residue truncations at the C-terminus of both $\mathbf{A_1}$ and $\mathbf{B_1}$ (Appendix D, Table D1). To gain a better understanding of how the mutations in ddRFP-$\mathbf{A_1B_1}$ contribute to its heterodimeric character and fluorogenesis, a homology model of ddRFP-$\mathbf{A_1B_1}$ was generated (Figure 3.18).

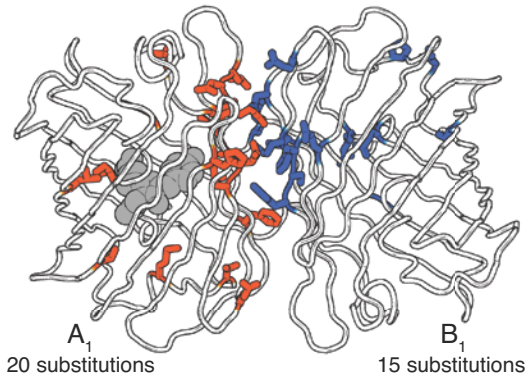


Figure 3.18. Structural model of ddRFP-$\mathbf{A_1B_1}$ based on the AC dimer of DsRed [74]. The amino acid side chains for all mutations in ddRFP-$\mathbf{A_1B_1}$ are shown in stick representation and the chromophore is rendered in space filling. The homology model depicted in Figure 3.18 is credited to Ahmed Abdelfattah. The image was generated with PyMOL (www.pymol.org) [66].

Inspection of this model reveals that $\mathbf{A_1}$ and $\mathbf{B_1}$ have six and four substitutions, respectively, at positions with side-chains directed towards the interior of the β-barrel. In $\mathbf{A_1}$, just three of these residues are in close proximity to the chromophore: the conservative I161L substitution, and the non-conservative S146A and K163G substitutions that are likely destabilizing the phenolate form of the chromophore due to loss of a hydrogen bond and a favourable electrostatic interaction, respectively. In contrast, the cavity that would normally harbour the chromophore of $\mathbf{B_1}$ has been dramatically remodeled with multiple non-

conservative substitutions that include K70E, Y120C, I161S, and E215G. The remodeled chromophore environment is consistent with the conclusion, supported by spectroscopic and biochemical evidence, that $\mathbf{B}_1$ does not form a chromophore.

The homology model of ddRFP-$\mathbf{A}_1\mathbf{B}_1$ also reveals that shape complementarity and electrostatic interactions across the dimer interface (Figure 3.19A) have been substantially modified relative to dTomato (which preserves the interface of DsRed) (Figure 3.19B).

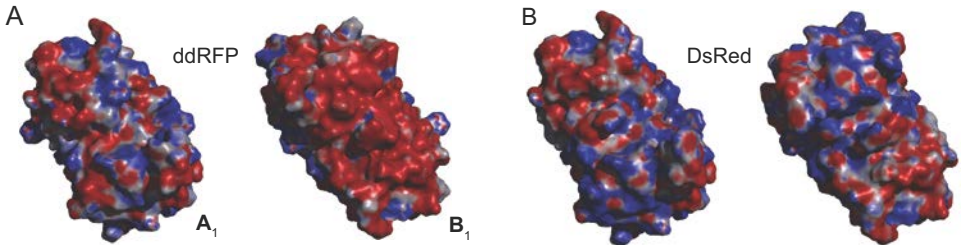


Figure 3.19. Surface representation of the ddRFP-$\mathbf{A}_1\mathbf{B}_1$ model coloured according to electrostatic potential. Shown are the two protein monomers separated and rotated 90° to expose the dimer interface. (A) Heterodimer interface of ddRFP-$\mathbf{A}_1\mathbf{B}_1$. (B) Homodimer interface of DsRed. The images depicted in Figure 3.19 are credited to Ahmed Abdelfattah.

Indeed, of the 19 substitutions in $\mathbf{A}_1$ and $\mathbf{B}_1$ with side chains directed to the exterior of the β-barrel, 15 are located in the interface. The homology model suggests that a principle determinant of the heterodimeric character is a new "bump-and-hole" interaction across the dimer interface. In the X-ray crystal structure of DsRed, the large side chain of R149 is positioned close to the small side chain of A164 on the dimer partner (Figure 3.20A) [74]. In ddRFP-$\mathbf{A}_1\mathbf{B}_1$, the A164R substitution of $\mathbf{A}_1$ is a "bump" that was introduced to destabilize the dTomato homodimer due to steric clashes and electrostatic repulsion with R149 of $\mathbf{B}_1$. In our model, the "hole" on the surface of $\mathbf{B}_1$ is formed by reorientation of

the side chain of R149M of $\mathbf{B}_1$ such that it is directed away from A164R of $\mathbf{A}_1$ and is filling an adjacent cavity created by Y194C of $\mathbf{B}_1$ (Figure 3.20B).

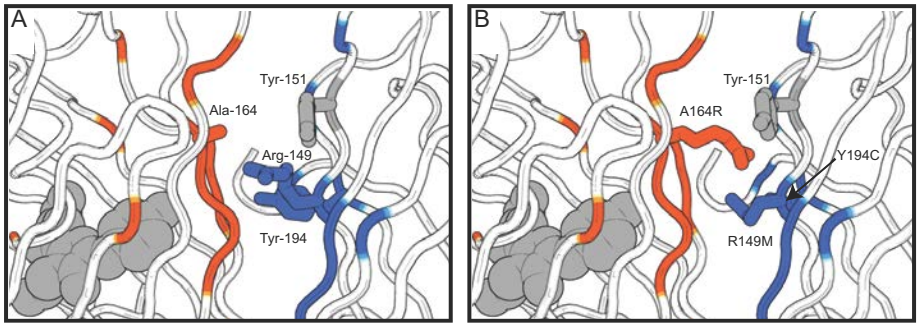


Figure 3.20. A structural model of ddRFP-$\mathbf{A}_1\mathbf{B}_1$ suggests the presence of a new 'bump-and-hole' interaction in the remodeled dimeric interface. (A) Modeled structure of DsRed AC interface from PDB ID 1G7K. (B) Modeled structure of the ddRFP-$\mathbf{A}_1\mathbf{B}_1$ interface. The **A** copy is on the left with substituted positions (in ddRFP-$\mathbf{A}_1\mathbf{B}_1$ relative to dTomato) coloured in orange, and the **B** copy is on the right, with substituted positions in blue. The images depicted in Figure 3.20 are credited to Ahmed Abdelfattah.

3.3 Conclusion

By employing a series of three distinct protein library screening strategies, we have engineered a fluorogenic RFP heterodimer (ddRFP-$\mathbf{A}_1\mathbf{B}_1$) with an associated state that is 10-fold brighter than the dissociated state. The convoluted evolutionary pathway that led to ddRFP-$\mathbf{A}_1\mathbf{B}_1$ reflects the practical challenge of simultaneously optimizing for at least five distinct, and possibly conflicting, properties: heterodimeric structure; minimal homodimeric structure; high-contrast fluorogenesis; high brightness in the associated state; and high K_d. Therefore, ddRFP-$\mathbf{A}_1\mathbf{B}_1$ represents a compromise, but it is the best compromise that we have identified to date.

Relative to the three well-established biosensing strategies, dimerization of $\mathbf{A}_1$ and $\mathbf{B}_1$ to form ddRFP-$\mathbf{A}_1\mathbf{B}_1$ is most closely analogous to FP complementation

[224]. A key difference between single FP complementation and ddRFP-$\mathbf{A}_1\mathbf{B}_1$ heterodimerization is that, while the fragments of a split FP are non-fluorescent, $\mathbf{A}_1$ retains 10% of the brightness of the complex. This residual fluorescence is likely to complicate efforts to use ddRFP-$\mathbf{A}_1\mathbf{B}_1$ in high-throughput screens for identification of interacting proteins [231].

With the development of ddRFP-$\mathbf{A}_1\mathbf{B}_1$, we have added a new entry to the short list of strategies (*i.e.*, FRET, insertion in a single FP, and FP complementation) that can be used for creation of FP-based biosensors. We have demonstrated the practicality of ddRFP-$\mathbf{A}_1\mathbf{B}_1$ by using it to detect PPIs, image a reversible Ca^{2+}-dependent PPI, and image protease activity. The primary advantage of ddRFP1-$\mathbf{A}_1\mathbf{B}_1$ relative to other strategies is that it provides a general means of creating red intensiometric biosensors with a reversible response. Potential disadvantages include the pH sensitivity and relative low brightness of ddRFP-$\mathbf{A}_1\mathbf{B}_1$, though we note that the brightness is comparable to the commercially available DsRed-monomer [232].

Future efforts to improve ddRFP-$\mathbf{A}_1\mathbf{B}_1$ should provide variants with decreased affinity and increased contrast. Until such variants become available, the most appropriate applications of ddRFP-$\mathbf{A}_1\mathbf{B}_1$ are as an alternative to intermolecular FRET and for protease activity sensing with tandem heterodimers. As we have done with the split cameleon-type Ca^{2+} sensor and the caspase-3 sensor, it should be relatively straightforward to convert existing FRET-based biosensors into ddRFP-$\mathbf{A}_1\mathbf{B}_1$-based biosensors. In these contexts, ddRFP-$\mathbf{A}_1\mathbf{B}_1$ will complement the existing repertoire of FP-based biosensing strategies, as it enables creation of spectrally distinct biosensors with an intensiometric and reversible red fluorescent signal.

3.4 Materials and Methods

3.4.1 Recombinant DNA techniques and reagents

All DNA manipulations including small scale preparation of plasmid DNA, restriction enzyme digestion, ligations, and agarose gel electrophoresis were performed according to Sambrook et al. [221]. Oligos were purchased from IDT Technologies. Unique numerical identifiers (e.g. Oligo 3.1) and corresponding sequences for all oligos are compiled in Table A1 (Appendix A). All restriction enzymes and *Pfu* DNA polymerase were obtained from Fermentas/Thermo Scientific and T4 DNA ligase was obtained from Invitrogen/Life Technologies. *Taq* polymerase was obtained from New England Biolabs. The GeneJet® Gel Extraction or Plasmid Miniprep Kits (Fermentas/Thermo Scientific) were used according to the manufacturer's instructions. Ligations were transformed into electrocompetent DH10B *E. coli*.

Standard PCR amplifications were carried out using *Pfu* DNA polymerase. PCR reactions were performed in volumes of 50 µl containing nuclease-free water, 1× reaction buffer, 200 µM dNTPs (Invitrogen), 200 nM forward and reverse oligos, 10 - 50 ng of template DNA, and 1.0 Unit of *Pfu* DNA polymerase. Typical PCR cycling parameters were as follows: initial denaturation at 95°C for 50 seconds; 20 cycles of 95°C for 15 seconds, 55-60°C for 20 seconds, 72°C for 60 seconds per kb of gene target; final extension at 72°C for 4 minutes. Where indicated, overlap extension PCR was performed as described in Bessette et al. [222]. Error prone PCR amplifications were performed as described in Cirino et al. [173] using *Taq* polymerase. Gene shuffling amplifications were generally carried out according to Zhao et al. [180]. Typical cycling parameters for shuffling amplifications were as follows: initial denaturation at 95°C for 50 seconds; 65

cycles of 95°C for 15 seconds, 55°C for 20 seconds, 72°C for 10-12 seconds and a final extension at 72°C for 4 minutes.

The general cloning strategy for each construct or library described in this chapter is outlined in Table 3.1. Destination plasmids, relevant oligos, and restriction enzyme sites used to clone designated constructs are provided in the table. Brief descriptions of cloning strategies are included in the methods where appropriate.

3.4.2 Specialized expression plasmids

Two modified pBAD-HisB plasmids (Invitrogen/Life Technologies) were generated for sub-cloning individual heterodimer partners directly from tandem heterodimer fusions. The first was designated as pBAD5' and was generated using pBAD template and 5' phosphorylated oligos 3.1 and 3.2. The pBAD5' open reading frame (ORF) accommodates *XhoI*/*KpnI* 5' gene excisions from tandem heterodimer constructs. The second was designated as pBAD3' using pBAD-HisB template and 5' phosphorylated oligos 3.3 and 3.4. The pBAD3' ORF accommodates *KpnI*/*HindIII* 3' gene excisions from tandem heterodimer constructs. Amplified plasmids were ligated and transformed into electrocompetent DH10B *E. coli*.

The custom bacterial dual expression plasmid, designated as pDES (credit to Andreas Ibraheem), contains an *XhoI*/*HindIII* polylinker under control of P_{Tac} and an *EcoRI*/*BglII* polylinker under control of P_{BAD}. The ORF for each polylinker encodes a 5' poly-histidine tag derived from pBAD.

3.4.3 Sequencing

All sequencing was performed at the MBSU at the University of Alberta. Sequencing of constructs cloned in the pBAD or pBAD-modified plasmids were

performed with Oligo 2.3 (forward) or Oligo 3.0 (reverse). Constructs cloned in the pcDNA3.1 (+) (Invitrogen/Life Technologies) plasmid were sequenced using a T7 forward primer (Oligo 2.1) or a BGH reverse primer (Oligo 2.2). Dual expression plasmid (pDES) constructs were sequenced using Oligo 2.3 (forward; P_{BAD} cassette), Oligo 2.4 (reverse; P_{BAD} cassette), Oligo 2.5 (forward; P_{Tac} cassette), and Oligo 2.6 (reverse; P_{Tac} cassette).

3.4.4 Library screening by gel shift.

The custom dual expression plasmid, pDES, was used for the rescue-of-binding assay. The fusion of MBP and dTomato H162K/A164R with an intervening *KpnI* site was ligated into the P_{BAD} polylinker.

Random mutations at sites R149, H162, and Y192 were introduced into a second dTomato gene using overlap extension PCR [222] and megaprimer PCR techniques [233]. The gene product was ligated into the P_{Tac} polylinker and transformed into electrocompetent DH10B *E. coli*. Transformations were plated on LB-agar media supplemented with 400 µg/ml ampicillin, 1 mM IPTG, and 0.2% L-arabinose. *E. coli* colonies were screened with 535/50 nm excitation and visualization through 600 nm long pass goggles. Bright colonies were picked into LB media supplemented with 100 µg/ml ampicillin, 0.1 mM IPTG, 0.02% L-arabinose and incubated overnight. Crude protein extracts were obtained using B-PER Protein Extraction Reagent (Pierce/Thermo Scientific). Proteins were purified using Nickel-NTA (Nitrilotriacetic acid) (Agarose Bead Technologies). Briefly, B-PER extracts (100 µL per 1.5 mL of culture) were diluted 10 fold with phosphate-buffered saline (PBS; 0.2 M NaCl, 4.2 mM KCl, 12.7 mM Na_2HPO_4, 2.3 mM KH_2PO_4, pH 7.4) supplemented with 10 mM imidazole and incubated on ice with 15 µL of Nickel-NTA for 15-30 minutes. After gentle centrifugation the

supernatant was aspirated and beads were washed twice with PBS supplemented with 20 mM imidazole. Proteins were eluted with 20 µL PBS supplemented with 250 mM imidazole. Eluted proteins were prepared with 1× sample buffer (62.5 mM Tris-Cl pH 6.8, 2% SDS, 7.5% glycerol, 0.02% bromophenol blue). Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed (4% stacking gel/10% resolving gel) according to the method of Laemmli [234], except samples were not heated prior to electrophoresis. Gels were analyzed for in-gel fluorescence using digital imaging with appropriate filters (excitation 535/50 nm; emission 630/60 nm).

3.4.5 Library screening for fluorogenesis.

Error-prone PCR was performed using *Taq* polymerase as described by Cirino et al. [173]. *E. coli* was transformed with gene libraries in the P_{Tac} and P_{BAD} polylinker sites of the dual expression plasmid. Bright colonies were picked and replica plated in a regular grid pattern (see Appendix D; Figure D1) on one plate supplemented with 1 mM IPTG and 0.2% L-arabinose and one plate supplemented with 1 mM IPTG and 10 mM glucose. Following overnight incubation at 37°C, the red fluorescence of both plates was imaged and the ratio of intensities on the L-arabinose vs. glucose plates for each colony were determined using image processing macros. Clones with the highest ratios were pooled and used in the next round of mutagenesis.

3.4.6 Library screening with tandem heterodimers.

Tandem heterodimers were initially constructed by a three-part ligation strategy which provided a chimera of the form **A**-linker-**B** in *XhoI*/*HindIII* sites of pBAD/His B, where the linker was a 23-residue sequence (GHGTGSTGSGSSGTASSEDNNMA) that included a *KpnI* site. Bright colonies

were picked and cultured in LB supplemented with 100 µg/ml ampicillin and L-arabinose (0.02%) overnight at 37°C. Crude B-PER protein extracts were treated with trypsin (Sigma) at approximately 10 µg/ml for 30 minutes and emission spectra were acquired using a 96-well microplate reader (Tecan Safire2^{TM}). Contrast ratios were calculated as the integrated emission peak area of the non-trypsinized extract divided by the peak area of the trypsinized extract.

3.4.7 Protein purification

To obtain purified proteins, the genes for tandem ddRFP variants and monomeric partners were sub-cloned into pBAD and used to transform *E. coli*. Cultures at an optical density of 0.5-0.7 were induced with 0.02% L-arabinose and allowed to incubate a further 12-16 hrs with shaking at 37°C. Cell pellets were resuspended in 10 mM Tris-Cl, 0.5 M NaCl, 10 mM imidazole pH 8.0 and lysed using a cell disrupter (Constant Systems Ltd.). Cleared lysates were obtained by centrifugation at 13,000 rpm in 50 ml conical tubes for 50 minutes. Lysates were incubated with Nickel-NTA on ice with agitation for 30 minutes and then batch purified through a fritted column using a vacuum manifold. Protein-bound resin was washed several times with a total of 50 mL 5 mM Tris-Cl, 0.5 M NaCl, 20 mM imidazole pH 8.0. The final 10 mL of wash was performed by gravity flow. Proteins were eluted using 5 mM Tris-Cl, 0.5 M NaCl, 250 mM imidazole pH 8.0 and dialyzed into 5 mM Tris-Cl, 200 mM NaCl, pH 7.5.

3.4.8 pH sensitivity determination

pH sensitivity measurements were performed by incubating purified proteins in buffers of desired pH and acquiring emission spectra with a 96-well microplate reader. Briefly, a 1 µM solution of FP was prepared in 5 mM Tris-Cl, pH 7.5 and diluted 1:10 with a universal buffer of desired pH. A Britton Robinson universal

buffer was used to generate buffers of desired pH. A stock solution was prepared by mixing 50 mL each of 0.04 M H_3BO_3, 0.04 M CH_3COOH, and 0.04 M H_3PO_4. The pH was adjusted to the desired value by adding 0.2 M NaOH to the prepared stock solution. The pK_a was determined by fitting the experimental data to the equation:

$$F = A + B(1 + 10^{(pK_a - pH)n_H})^{-1}$$

Where F is fluorescence, A and B are variables that define the baselines, and n_H is the Hill coefficient.

3.4.9 Spectral feature determination

Spectra represented in this chapter were recorded with a DU-800 UV-visible spectrophotometer (Beckman) or a QuantaMaster spectrofluorimeter (Photon Technology International, Inc.). Absorbance measurements were made using a 1 cm quartz microcell cuvette. The alkaline chromophore denaturation method was used to determine extinction coefficient (ε) values [235]. dTomato was used as the reference for quantum yield (Φ) determination.

3.4.10 Determination of oligomerization state

To determine oligomerization state, purified recombinant proteins were resolved by gel filtration chromatography. Proteins were resolved over a HiLoad 16/60 Superdex 75 pg gel filtration column using an AKTA basic liquid chromatography system (GE Healthcare) with a multiple wavelength absorbance detector. Proteins were resolved with a 5 mM Tris-Cl, 200 mM NaCl, pH 7.5 mobile phase and 0.5 mL/min flow rate. Eluted fractions were monitored at 280 nm and 600 nm.

3.4.11 Dissociation constant determination

To determine the K_d of the purified recombinant ddRFP partners, an increasing amount of non-fluorescent ddRFP-**B** was mixed with a fixed amount of ddRFP-**A** (250 nM final) to generate ddRFP-**AB** complexes in 20 mM Tris-Cl, 100 mM NaCl and 0.5 mM EDTA, pH 7.5. The integrated fluorescence emission peaks recorded as a function of ddRFP-**B** concentration were used to generate saturation binding curves. Data was corrected for a background buffer blank.

Experimental data was fit using a modified Langmuir isotherm to account for ligand depletion:

$$[\boldsymbol{AB}] = \frac{([\boldsymbol{A}] + [\boldsymbol{B}] + \boldsymbol{K_d}) - \sqrt{([\boldsymbol{A}] + [\boldsymbol{B}] + \boldsymbol{K_d})^2 - \mathbf{4}[\boldsymbol{A}][\boldsymbol{B}]}}{\mathbf{2}}$$

3.4.12 *In vitro* ddRFP assay for rapamycin-mediated FKBP/FRB interaction

The genes encoding FRB and FKBP were fused to $\mathbf{A}_1$ and $\mathbf{B}_1$, respectively, to generate FRB-$\mathbf{A}_1$ and FKBP-$\mathbf{B}_1$ fusions. For both constructs, the linker sequence corresponded to residues 95 to 135 of phage λ repressor (source: *HindIII*-digested Lambda DNA, Invitrogen/Life technologies). The fusion proteins were expressed from pBAD/His B and soluble proteins were purified as described for the tandem ddRFP variants. Protein concentrations were determined by the Bicinchoninic acid method (Pierce/Thermo Scientific). Assay buffer was 10 mM Tris-Cl, 100 mM NaCl, pH 7.5. Proteins were mixed together to a final concentration of 0.5 μM and rapamycin (BD Biosciences) was added to the indicated final concentrations (1 μM, 100 nM, and 10 nM). Dimethyl sulfoxide (DMSO) was added as a vehicle control where appropriate. Reaction mixtures

were immediately analyzed to obtain emission spectra with a 96-well microplate reader.

3.4.13 Live cell imaging applications

The $\mathbf{A}_1$ and $\mathbf{B}_1$ copies in the DEVD tandem heterodimer caspase-3 were linked by the sequence GHGTHSTHSHSSHTASHDEVDGA or the same sequence with DEVD replaced with SASG. Kozak consensus sequences were introduced during cloning procedures. The gene sequence encoding the nuclear exclusion sequence (LALKLAGLDIGS) was appended to the 3' end of the gene [236]. For caspase-3 imaging, apoptosis was initiated by treatment with 100 ng/ml TNF-α (Sigma). Cells were imaged in HEPES-buffered Hank's balanced salt solution (HHBSS).

For the split cameleon-type biosensors, the genes encoding $\mathbf{A}_1$ and CaM as well as M13 and $\mathbf{B}_1$, were joined by the sequence GHGTGSTGSGSSTASSEDMA. The D3 variant of M13 was used to generate the biosensor [118]. The genes encoding the tandem heterodimer caspase-3 biosensors and both the CaM and M13 fusions were ligated into the *HindIII/XhoI* sites of pcDNA3 (+). Kozak consensus sequences were introduced during cloning procedures. For Ca^{2+} imaging, cells were imaged in HHBSS and consecutively treated with histamine (25 μM), EGTA (3 mM) with ionomycin (Life technologies 5 μM), and 10 mM $CaCl_2$ with 5 μM ionomycin. Experiments were performed as described in Palmer and Tsien, 2006 [237].

3.4.14 Mammalian cell culture and imaging

HeLa cells were maintained in DMEM with 10% fetal bovine serum at 37°C and 5% CO_2 according to standard procedures. Transient transfections were

performed using Turbofect (Fermentas/Thermo Scientific). Designated expression plasmids were transfected according to the manufacturer's instructions. For single plasmid transfections 750 ng plasmid DNA was used, and 375 ng of each plasmid was used for co-transfection of two plasmids. Generally, 2 µl of transfection reagent was used per 1 µg of DNA transfected. Transfected cells were imaged 24 hours following transfection in HHBSS on an inverted Nikon Eclipse Ti microscope equipped with a 200W metal halide lamp (PRIOR Lumen) and a QuantEM: 512SC 16-bit cooled CCD camera (Photometrics).

3.4.15 Homology Modeling

The homology modeling in this thesis is credited to Ahmed Abdelfattah. The homology model for ddRFP-$\mathbf{A}_1\mathbf{B}_1$ was constructed using the Rosetta fixed backbone design protocol [238] with the A and C chains of PDB ID 1G7K [74].

Table 3.1. Cloning strategy for constructs described in Chapter 3.

		Amplification 1		Gene assembly	
Construct	**Plasmid**	**Template**	**Oligos[1]**	**Overlap Extension Oligos**	**RE sites (5'/3')**
MBP (ΔBglII)	pDES	pMAL[2] pMAL	3.6/3.9 3.7/3.8	3.6/3.7	*EcoRI/ KpnI*
dTomato H162K/A164R	pDES	dTomato dTomato	3.10/3.13 3.11/3.12	3.12/ 3.13	*KpnI/ BglII*
192X megaprimer[3]	n/a[4]	dTomato	3.14/3.57	n/a	n/a
dTomato R149X/ H162X/Y192X	pDES	dTomato dTomato dTomato	3.35/3.15 3.14/3.18 3.17/ MP192X	3.35/ MP192X	*XhoI/ HindIII*
$B_{0.3}$	pDES	$B_{0.2,pool}$ (*XhoI/ HindIII*)	3.20/3.58 3.21/3.35	3.35/ 3.58	*XhoI/ HindIII*
$A_{0.3}$	pDES	$A_{0.2}$ (*KpnI/ BglII*)	3.12/3.58	n/a	*KpnI/ BglII*
$B_{0.4}$	pDES	$B_{0.3}$ (*XhoI/ HindIII*)	3.19/3.35	n/a	*XhoI/ HindIII*
td $A_{0.4}B_{0.5}$	pBAD	n/a	3.35/3.36 3.37/3.38	3-part ligation[5]	*XhoI/ HindIII*
$A_{0.6}$, pool dark	pBAD	$A_{0.5}$	3.53/3.57 3.54/3.60	3.57/3.60	*XhoI/ HindIII*
192X/194X/ 222X in B	pBAD	$B_{0.5}$	3.56/3.57 3.55/3.60	3.57/ 3.60	*XhoI/ HindIII*
FKBP-B_1	n/a	FKBP λ cI B_1	3.23/3.25 3.26/3.27 3.24/3.57	3.23/ 3.57	n/a
FRB-A_1	n/a	FRB λ cI A_1	3.22/3.29 3.30/3.31 3.32/3.57	3.22/ 3.57	n/a
FKBP-B_1 ΔHindIII	pBAD	FKBP-B_1	3.23/3.34 3.33/3.57	3.23/ 3.57	*XhoI/ HindIII*
FRB-A_1 ΔHindIII	pBAD	FRB-A_1	3.22/3.34 3.33/3.57	3.22/ 3.57	*XhoI/ HindIII*
A_1-CaM	pcDNA3.1 (+)	A_1-pBAD5' CaM	3.43/3.60 3.39/3.48	3-part ligation	*HindIII/ XhoI*
M13-B_1	pcDNA3.1 (+)	M13(D3) B_1-pBAD3'	3.49/3.42 3.60/3.44	3-part ligation	*HindIII/ XhoI*

A_1-DEVD-B_1	pcDNA3.1 (+)	A_1-pBAD5' B_1-pBAD3'	3.43/3.60 3.45/3.50	3-part ligation	*HindIII/ XhoI*
A_1-SASG-B_1	pcDNA3.1 (+)	A_1-pBAD5' B_1-pBAD3'	3.46/3.60 3.45/3.50	3-part ligation	*HindIII/ XhoI*

[1] Oligonucleotide primers are provided in Appendix A (Table A1). [2]The pMAL template was the kind gift of Dr. R. Ingham. [3]Mega primer (MP192X) amplified for use in subsequent overlap extension PCR. [4]Not applicable. [5]Amplified products were digested with appropriate restriction enzymes and assembled by 3-part ligation into the designated plasmids.

Chapter 4 Dimerization-dependent green and yellow fluorescent proteins[3]

4.1 Introduction

In Chapter 3, we described the engineering of a dimerization-dependent fluorescent protein, or ddFP. This technology expanded the range of design strategies (including FRET, complementation, and single FP-biosensors) for FP-based imaging tools. The concept of ddFP technology relies on the reversible binding of two chemically distinct dark FP monomers to form a fluorescent heterodimeric complex (Figure 4.1). The fluorogenic response associated with the FP/FP interaction provides a spectral signature that is correlated with an increase in the proximity or effective concentration of monomers. The prototype ddFP was a red fluorescent system (ddRFP) derived from a dimeric red FP called, dTomato [9]. While ddRFP did prove to be useful in a variety of applications, it did suffer from limited brightness (quantum yield of 0.074 and extinction coefficient of ~ 48,000 $M^{-1} \cdot cm^{-1}$), limited contrast (~10-fold), and was monochrome (only red).

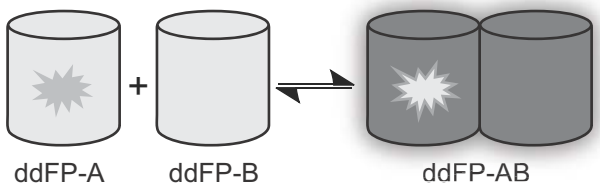


Figure 4.1. Dimerization-dependent fluorescent protein (ddFP) technology. Two monomeric FPs reversibly associate to form a fluorescent heterodimer. The starburst in ddFP-A corresponds to a mature chromophore.

[3] A version of this chapter has been accepted for publication. Alford, S.C., et al. 2012 *ACS Synthetic Biology*. DOI: 10.11021/sb300050j.

In this chapter, we describe our efforts to engineer new colour variants of ddFPs with improved brightness and contrast relative to ddRFP. Starting from a ddRFP template, we used a process of directed evolution to create green (ddGFP) and yellow (ddYFP) analogues of ddRFP. We explored the utility of ddGFP and ddYFP in live cell imaging applications and found while the low μM K_d of these ddFPs limits their utility in applications that involve free diffusion through the cytoplasm or within a single membrane, ddGFP is a useful reporter of mitochondria-endoplasmic reticulum (ER) contacts mediated by the mitochondrial-associated ER membrane (MAM).

We have adopted the convention of designating the monomers of a ddFP pair as copy A (e.g., ddGFP-A) and copy B (e.g., ddGFP-B). The ddFP-A partner possesses the preformed, but quenched chromophore, while the ddFP-B partner lacks a chromophore (Figure 4.1). We designate the non-covalent complex as AB (e.g., ddGFP-AB) and genetically fused tandem dimers with a 'td' rather than the 'dd' prefix (e.g., tdGFP-AB).

4.2 Results and Discussion

4.2.1 Directed evolution and characterization of ddGFP

DdRFP-**A_1B_1** was originally generated through an extensive process of directed evolution (Chapter 3). To generate new colour hues of ddFP we chose to modify ddRFP directly rather than repeat the directed evolution using a differently coloured FP starting template. Furthermore, ddRFP was itself engineered from a DsRed variant for which a variety of mutations have been reported to convert the red-emitting chromophore to a green-emitting one [71, 73,

239], as well as to yellow or orange emitting species' [98]. We first attempted to engineer a green-emitting ddFP using a ddRFP construct, ddRFP-**$A_1B_{1.1}$** (Appendix E, Table E1), that exhibits a modest increase in contrast (Figure 4.2; ~20-fold) and increased affinity (~2 μM; data not shown) relative to ddRFP-**A_1B_1**. We chose this template due to its improved contrast with the expectation that the dissociation constant could be rationally increased by introduction of repulsive electrostatic interactions once suitable hue-shifted variants were obtained.

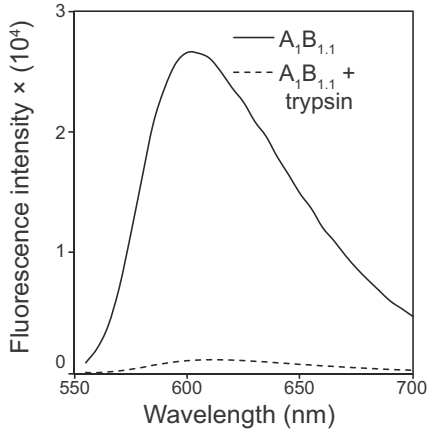


Figure 4.2. *In vitro* contrast of ddRFP variant **$A_1B_{1.1}$**. Shown is the emission spectrum of tandem dimer ddRFP-**$A_1B_{1.1}$** before (solid black line) and after (dashed line) trypsinolysis of an *E. coli* cell lysate. The calculated contrast between the associated and dissociated state is ~20 fold.

We initially introduced documented 'greenizing' mutations into the ddFP-A copy of the tandem dimer version of ddRFP-**$A_1B_{1.1}$**. Specifically, K83R, K70M [71, 73], V105A, and V71M [239] have been previously reported to convert DsRed into a green FP. In the context of ddRFP-**A_1B_1**, valine-71 corresponds to alanine-71, which arose during the evolution of dTomato from DsRed [9]. While the K83R and K70M mutations led to a complete loss of fluorescence, dim green fluorescence was observed with the A71M mutation independently or as a double

mutant with V105A. Absorbance spectra of purified proteins confirmed the presence of a green-absorbing species (Figure 4.3). The tdRFP-AB construct with mutations A71M or A71M/V105A in the A copy was used as the template for subsequent evolution.

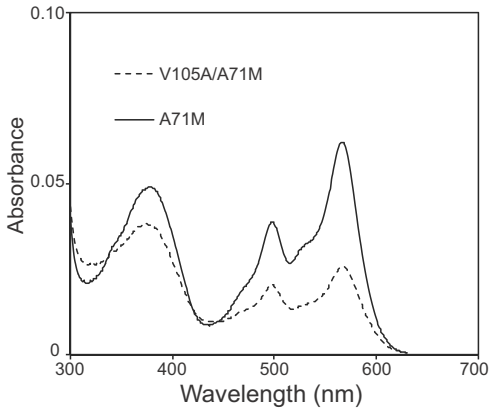


Figure 4.3. Absorbance profiles for the initial generations of ddGFP. Shown are the absorbance profiles of 5 μM purified V105A/A71M and A71M in 20 mM Tris-Cl, 100 mM NaCl, pH 7.5.

Multiple rounds of library creation by error-prone mutagenesis followed by screening for high green fluorescence and diminished red fluorescence were undertaken [11]. In each round, the brightest green variants were subjected to a secondary screen for contrast; the linker joining the two FPs in the tandem dimer was cleaved with trypsin and the variants that exhibited the largest decrease in intensities due to FP dissociation were identified. Continued evolution by random mutagenesis and intermittent gene shuffling was continued until no further improvements were achieved during a round of exhaustive library screening (Figure 4.4). The brightest green fluorescent variant (designated as ddGFP-AB) had twelve amino acid changes relative to ddRFP (Appendix E, Table E1). Eight of these mutations occur in ddGFP-A and four occur in ddGFP-B.

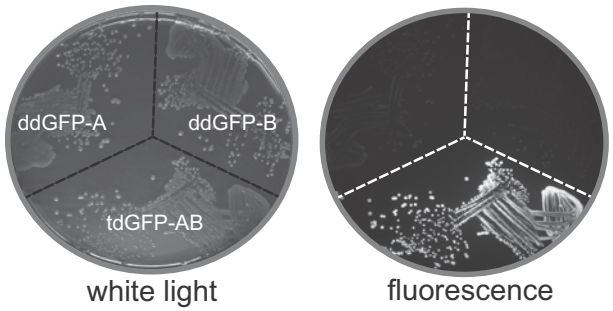


Figure 4.4. Recombinant expression of high contrast ddGFP-AB. *E. coli* expressing ddGFP-A, ddGFP-B, or tdGFP-AB. White light image is depicted on left and fluorescence image is depicted on right.

TdGFP-AB and its constituent monomers were purified to characterize their *in vitro* properties. The absorbance spectrum of tdGFP-AB is characterized by peaks centered at 380 nm and 493 nm (Figure 4.5) and the absence of peaks in the red region. Excitation of the 493 nm absorbing species results in a green emission centered at 508 nm (Figure 4.5). Excitation at the 380 nm peak results in weak (<0.1% relative to excitation of the 493 nm peak) emission peaks at 448 nm and 508 nm (Figure 4.5B).

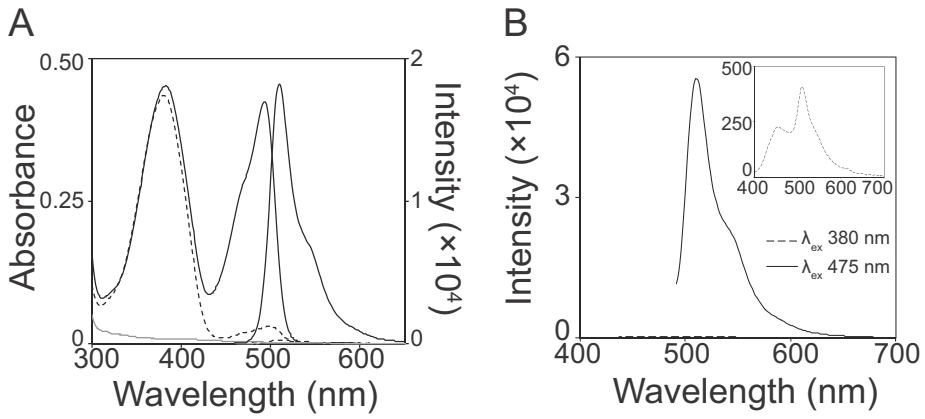


Figure 4.5. Spectral features for ddGFP-AB. (A) Absorbance and emission of ddGFP-A (black dashed line), ddGFP-B (grey line), and tdGFP-AB (solid black line). (B) Emission profiles for tdGFP-AB with excitation at 380 nm (inset) relative to excitation at 475 nm.

The absorbance spectra of ddGFP-A and tdGFP-AB exhibit pH dependence; the 380 nm species dominates at low pH, but is converted to the 493 nm species with increasing pH (Figure 4.6A,B). DdGFP-A and tdGFP-AB have pK_a values of 9.4 and 7.8, respectively (Figure 4.6C). These high pK_a values render the fluorescence of ddGFP-AB sensitive to physiologically relevant changes in pH.

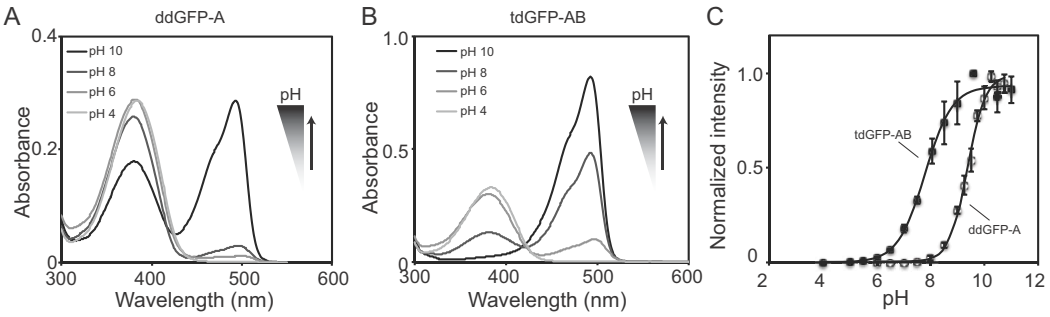


Figure 4.6. *In vitro* characterization of pH dependence for ddGFP-AB. Absorbance profiles for (A) ddGFP-A and (B) tdGFP-AB at designated pH values. (C) pH dependent emission profiles for ddGFP-A (open circles) and tdGFP-AB (filled circles). For (C) mean integrated emission peak intensities were normalized and plotted as a function of pH. pK_a values for ddGFP-A and tdGFP were determined to be 9.4 and 7.8, respectively. Error bars indicate ± standard deviation for at least three independent experiments.

To evaluate *in vitro* contrast, trypsinolysis assays were performed on purified recombinant tdGFP-AB and its monomers (Figure 4.7A,B). TdGFP-AB exhibited a ~60-fold increase in emission intensity relative to its monomers, ddGFP-A and ddGFP-B (Figure 4.7B). This contrast is a marked improvement relative to ddRFP-**A_1B_1** which exhibited only a 10-fold change in contrast between monomeric and heterodimeric states (Chapter 3). DdGFP-A demonstrates a Φ = 0.04 and an ε = of 2,000 $M^{-1}\cdot cm^{-1}$ at pH 7.4, but these values increase to 0.20 and 21,500 $M^{-1}\cdot cm^{-1}$ respectively for tdGFP-AB. Similar to ddRFP-B, ddGFP-B does not absorb or fluoresce (Figure 4.7A,B).

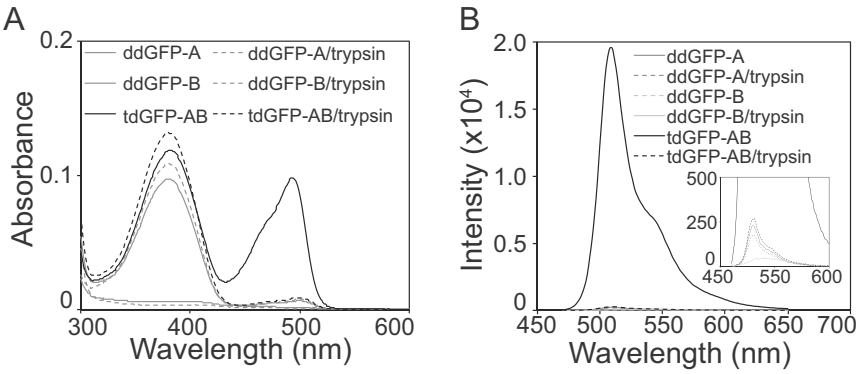


Figure 4.7. Trypsinolysis analysis of ddGFP-AB. (A) Absorbance and (B) emission profiles of tdGFP-AB and its monomers before and after trypsinolysis. Inset in (B) shows the low intensity fluorescent signal. Proteins were assayed at 5 μM in 20 mM Tris-HCl, 100 mM NaCl, 0.5 mM EDTA pH 7.4.

4.2.2 Directed evolution and characterization of ddYFP

Given the success of the general strategy for evolving a green-hued ddFP, we next attempted to engineer a yellow-hued ddFP by introducing the yellow-shifting mutation, M66C [98]. Installation of cysteine at this position produces a unique three-ring chromophore [81]. Similar to the initial steps in ddGFP-AB engineering, installation of M66C into tdRFP-AB decreased fluorescence intensity. Further directed evolution ultimately produced an effective ddYFP-AB variant with 19 mutations relative to ddRFP-$\mathbf{A_1B_1}$ (Appendix E, Table E1) and exhibited good contrast in *E. coli* and *in vitro* (Figures 4.8)**.**

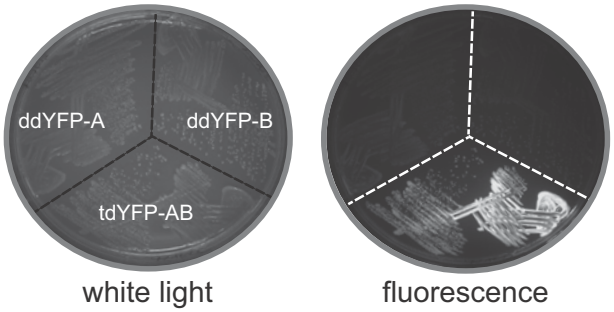


Figure 4.8. Recombinant expression of high contrast ddYFP-AB. *E. coli* expressing ddYFP-A, ddYFP-B, or tdYFP-AB. White light image is depicted on left and fluorescence image is depicted on right.

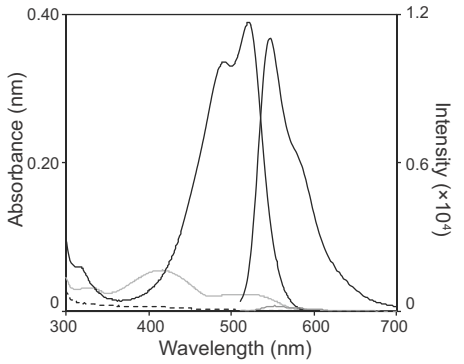


Figure 4.9. Spectral features for ddYFP-AB. (A) Absorbance and emission of ddYFP-A (grey line), ddYFP-B (black dashed line), and tdYFP-AB (solid black line). Spectra were obtained for purified proteins diluted to 10 μM in 20 mM Tris-Cl, 100 mM NaCl, 0.5 mM EDTA pH 7.4.

TdYFP-AB exhibits absorbance and emission peaks centered at 520 nm and 546 nm, respectively (Figure 4.9). Interestingly, a key mutation for achieving high contrast in ddYFP-AB was the A146S reversion mutation in the ddYFP-A copy, despite the fact that S146A was a key mutation for achieving reasonable contrast in ddRFP-A_1B_1 [240]. DdYFP-AB exhibits a 12-fold difference in emission intensity between the associated and unassociated states *in vitro* at pH 7.5 (Figure 4.10). DdYFP-A and tdYFP-AB exhibit ε values of 10,000 $M^{-1}\cdot cm^{-1}$ and 53,800 $M^{-1}\cdot cm^{-1}$ (at 520 nm, pH 7.5), and Φ values of 0.10 and 0.37, respectively. Although the brightness of this yellow variant is encouraging, its most promising feature is a reduced pK_a. At pH 7 ddYFP-A exhibits minimal absorbance at 516 nm (Figure 4.11A), whereas tdYFP-AB absorbs near maximally and the protonated chromophore species is absent, consistent with the measured pK_a of 6.4 (Figure 4.11B,C). This property should render ddYFP-AB relatively insensitive to physiologically relevant changes in pH.

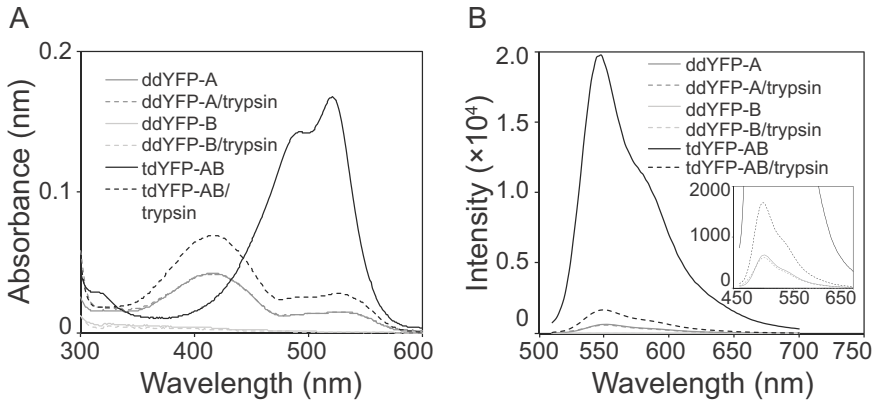


Figure 4.10. Trypsinolysis analysis of ddYFP-AB. (A) Absorbance and (B) emission profiles of tdYFP-AB and its monomers before and after trypsinolysis. Inset in (B) shows the low intensity fluorescent signal. Proteins were assayed at 5 μM in 20 mM Tris-HCl, 100 mM NaCl, 0.5 mM EDTA pH 7.4.

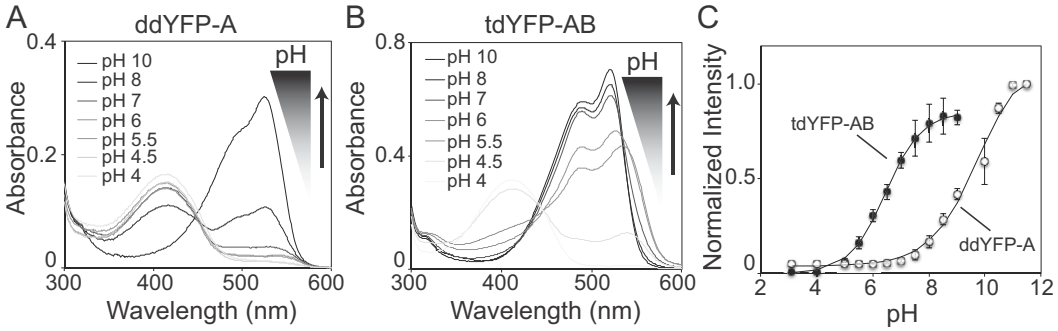


Figure 4.11. *In vitro* characterization of pH dependence for ddYFP-AB. Absorbance profiles for (A) ddYFP-A and (B) tdYFP-AB at designated pH values. (C) pH dependent emission profiles for ddYFP-A (open circles) and tdYFP-AB (filled circles). For (C) mean integrated emission peak intensities were normalized and plotted as a function of pH. pK_a values for ddYFP-A and ddYFP-AB were determined to be 9.8 and 6.4, respectively. Error bars indicate ± standard deviation for at least three independent experiments.

In vitro saturation binding experiments revealed that ddGFP-AB has a K_d of 9 μM (Figure 4.12A) and ddYFP-AB has a K_d of 14.5 μM (Figure 4.12B), which both represent a substantial decrease relative to the K_d of 33 μM for ddRFP-**A_1B_1** [241]. We speculated that the higher affinity could be attributed to the involvement of K153 in electrostatic interactions across the interface [74]. During the evolution of ddRFP-**A_1B_1** we identified this residue as a modulator of affinity

between the heterodimer partners. We therefore generated one further variant of ddGFP-AB by installing a K153E mutation in ddGFP-B. This K153E variant decreased Φ and ε values of tdGFP-AB to 0.17 and 17,600 $M^{-1}\cdot cm^{-1}$, respectively. Although K153E reduces the brightness by ~30% and contrast to ~37-fold, it increases the K_d to 27 μM (Figure 4.12A).

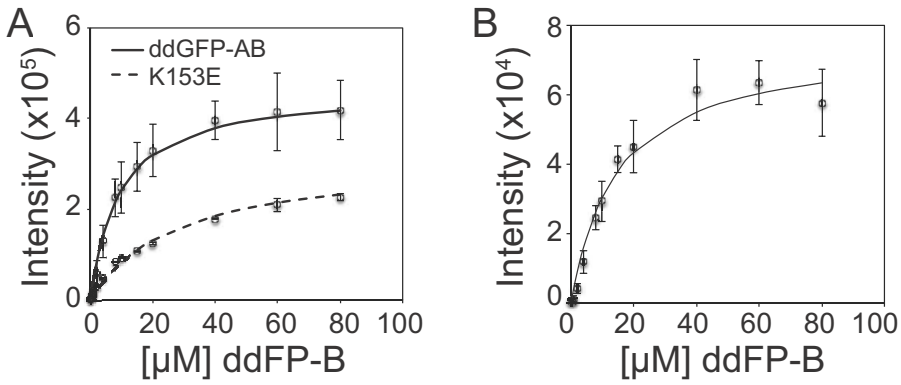


Figure 4.12. Saturation binding curves for fluorogenic ddGFP and ddYFP heterodimers. Saturation binding curves for (A) ddGFP-AB and ddGFP-AB + K153E (copy B) and (B) ddYFP-AB. K_d values are 9 μM, 27 μM, and 14.5 μM respectively. Error bars are ± standard deviation for at least three independent experiments.

4.2.3 Live cell applications

In Chapter 3 we demonstrated that ddRFP-$\mathbf{A_1B_1}$ is suitable as a template for construction of genetically encoded biosensors for monitoring fluctuations in Ca^{2+} concentration and caspase-3 activation in live cells. To assess whether our improved ddGFP-AB and ddYFP-AB variants could also serve as templates for caspase-3 biosensors, we introduced a caspase-3 substrate sequence into the linker joining the heterodimer partners as described in Chapter 3. Expression of this construct in mammalian cell culture revealed that tdGFP-AB generally gave brighter fluorescence than tdRFP, though the brightest cells exhibited only small changes in intensity following caspase-mediated cleavage during apoptosis.

Presumably, in these cells the concentration of tdGFP-AB is substantially greater than the ddGFP-AB K_d of 9 μM. In dim cells (expressing low levels of biosensor), we observed a 3.6 ± 1.0-fold (N=39) and 2.8 ± 0.6-fold (N=10) decrease in fluorescence following caspase-mediated cleavage of tdGFP-AB and tdGFP-AB with K153E in the B-copy, respectively (Figure 4.13A/B). The observed loss of fluorescence signal upon cleavage is reduced relative to the contrast observed in vitro. We speculate the concentration of monomers following cleavage is sufficient to cause non-covalent dimerization and thus the observed contrast is decreased. Indeed, we find that co-expression of ddGFP-A and ddGFP-B in HeLa cells results in diffuse cytoplasmic fluorescence (Figure 4.14E). For tdYFP-AB, caspase-3 activation traces were comparable in terms of kinetics and fold change to those obtained for tdGFP-AB-based biosensors (Figure 4.13C).

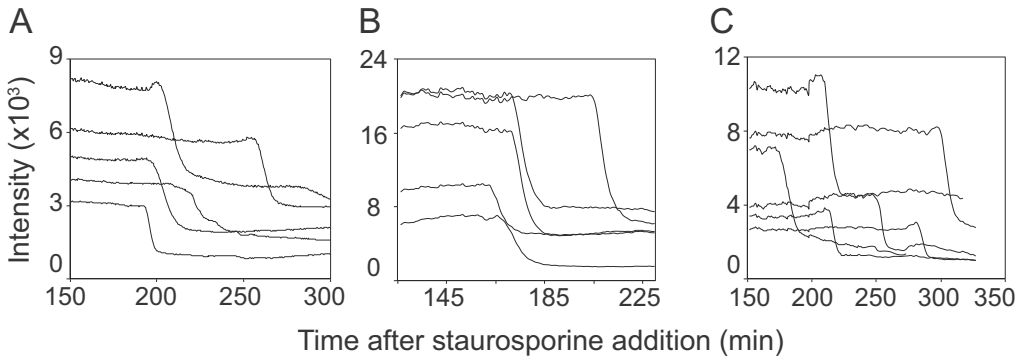


Figure 4.13. Caspase-3 activation assayed by ddFPs described in this work. Caspase-3 activity traces for (A) ddGFP-AB, (B) ddGFP-AB + K153E (B copy), and (C) ddYFP-AB. Transfected HeLa cells were treated with staurosporine (2 μM in HHBSS) and fluorescence was monitored over time. Representative traces for individual cells are depicted. The caspase-3 activity traces depicted in Figure 4.13 are credited to Yidan Ding.

Based on the results with the caspase-3 biosensors, it was apparent that the relatively high affinity of ddGFP-AB and ddYFP-AB would limit our ability to detect association or dissociation of proteins freely diffusing within the cytosol.

Accordingly, we explored the suitability of these tools for detection of protein proximity in cases where proteins are tethered to a membrane and thus not able to freely diffuse. We focused these efforts on ddGFP-AB due to its extremely high contrast, and have not yet pursued the use of ddYFP-AB in similar applications.

To determine if ddGFP-AB could be used to detect protein-protein proximity or association at the plasma membrane of mammalian cells, we fused ddGFP-A to a pleckstrin homology (PH) domain and installed a farnesylation substrate sequence on ddGFP-B [242, 243]. We expected that the PH domain would recruit ddGFP-A to the membrane through association with inositol phospholipids [244] (Figure 4.14) and lipidated ddGFP-B would be tethered directly to the membrane (Figure 4.14A). When each fusion was expressed independently in HeLa cells, we detected intensities consistent with cell autofluorescence (Figure 4.14B,C). However, co-expression of the membrane-targeted heterodimer partners resulted in prominent fluorescent labeling of the plasma membrane (Figure 4.14D). When a cytosolic ddGFP-B was expressed rather than the farnesylated version, weak fluorescence was observed throughout the cell (Figure 4.14E) with no increase in intensity at the plasma membrane. These results demonstrate that ddGFP-AB was efficiently reconstituted only when both the A and B monomers were recruited to the plasma membrane.

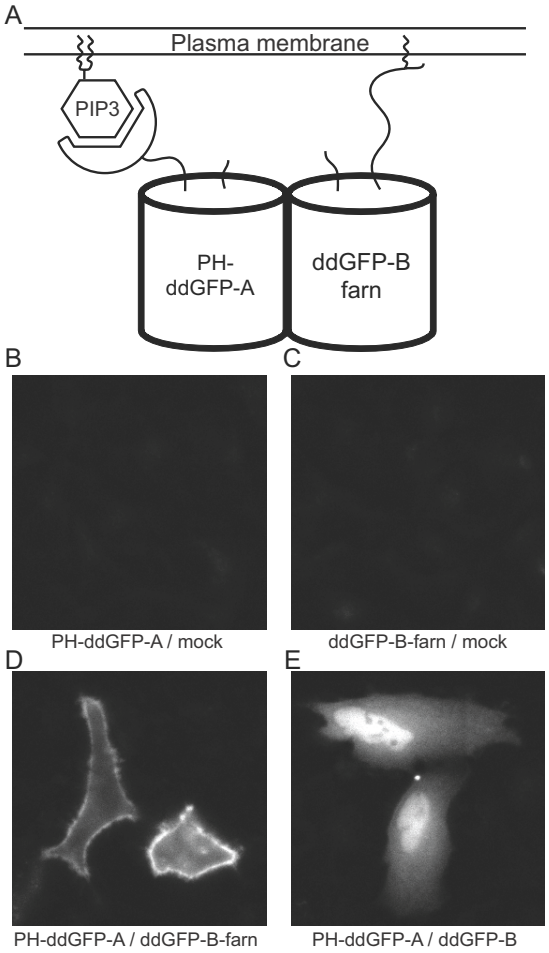


Figure 4.14. Proximity-mediated labeling of the plasma membrane with ddGFP in HeLa cells. (A) Strategy for plasma membrane ddGFP labeling using PH-ddGFP-A and farnesylated (Farn) ddGFP-B. PIP3; phosphatidylinositol (3,4,5) triphosphate. (B) Transfection of PH-ddGFP-A or (C) ddGFP-B-Farn alone does not result in detectable fluorescence. (D) Prominent green fluorescent labeling at the plasma membrane is seen when both PH-ddGFP-A and ddGFP-B-Farn are co-transfected. (E) Plasma membrane labeling is lost when ddGFP-B alone is expressed with PH-ddGFP-A.

We next tested the applicability of ddGFP-AB for labeling of the MAM-mediated mitochondria-ER endomembrane contact sites [245]. Our strategy was to co-express ddGFP-A or -B chimeras in which each partner was fused to either the mitochondrial protein translocase of outer membrane-20 (Tom20) or the ER protein, calnexin (Figure 4.15A). Tom20 is an outer membrane protein of the mitochondria that functions as a receptor during protein import and calnexin is a

Ca^{2+} binding chaperone that localizes to the MAM and regulates Ca^{2+} homeostasis [246-248]. We fused ddGFP-A to the C-terminus of calnexin and ddGFP-B to the C-terminus of Tom20 such that the ddGFP-AB partners are presented on the cytosolic face of the organelles (Figure 4.15A). When each gene fusion was expressed independently, no fluorescence labeling was detected above autofluorescence (Figure 4.15B,C). However, when both fusions were transfected we observed prominent green fluorescence localized in the perinuclear region of cells (Figure 4.15D). The perinuclear localization was defined by discrete labeling at points of mitochondria-ER contact (Figure 4.15D) consistent with previous reports of MAM staining [249, 250].

To confirm that we were observing genuine MAM-mediated contact sites, we performed a number of control experiments. First, co-transfection with mCherry targeted to either the mitochondria or ER demonstrated that the MAM-associated ddGFP-AB green fluorescence co-localized with both the mitochondria and the ER (Figure 4.15E,F). Second, to confirm that MAM labeling was dependent on the proximity association of Tom20 and calnexin rather than the result of non-specific ddGFP-AB aggregation, we utilized Tom20 mutants that have been reported to delocalize from the mitochondria [251]. Expression of these Tom20 mutants as mCherry fusions confirmed that these chimeras redistributed to the cytoplasm and golgi (Figure 4.16A-I). Fusion of these Tom20 mutants to ddGFP-B and co-expression with calnexin-ddGFP-A resulted in a loss of green fluorescent labeling of MAM interactions (Figure 4.17A,B).

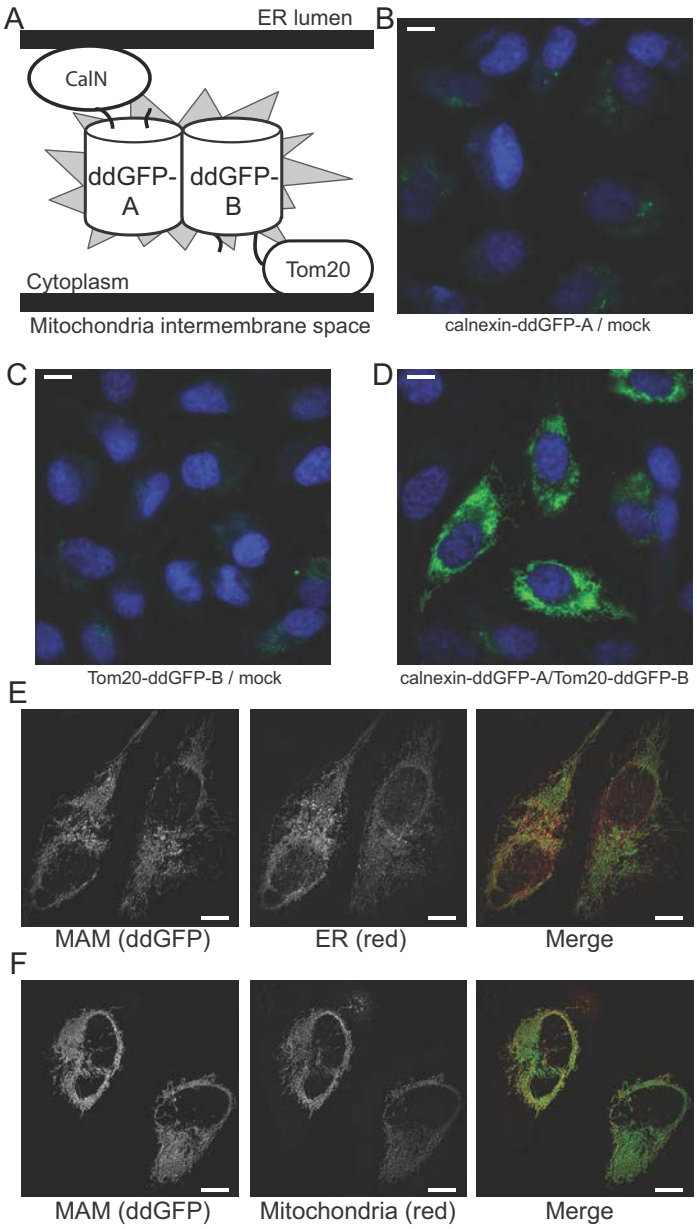


Figure 4.15. DdGFP labeling of MAM in HeLa cells. (A) MAM labeling strategy. (B) Calnexin-ddGFP-A alone. (C) Tom20-ddGFP-B alone. (D) Co-expression of Calnexin-ddGFP-A and Tom20-ddGFP-B. DdGFP fluorescence co-localizes with (E) ER and (F) mitochondria; left panel is ddGFP fluorescence, middle panel is red fluorescence (mCherry ER or mitochondria markers), right panel is merge. Cell nuclei (blue) are stained with DAPI. Images depicted in E and F were acquired on an Olympus IX-81 motorized spinning disc confocal microscope. Scale bars, 10 μm. Some of the images in Figure 4.15 are credited to Yidan Ding.

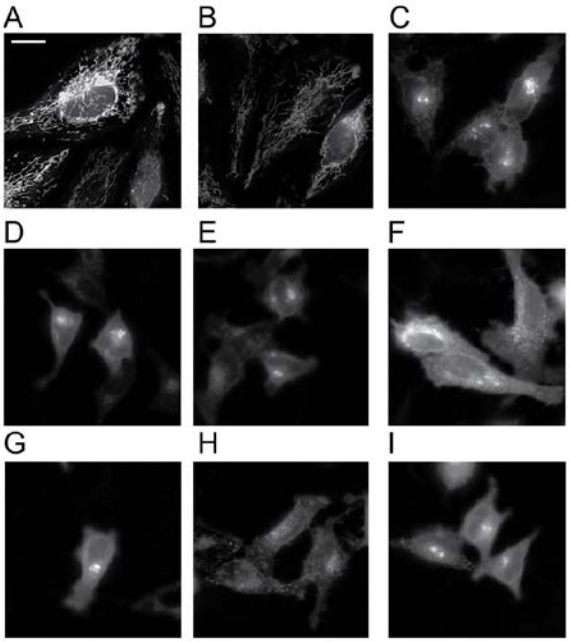


Figure 4.16. Tom20 truncation mutants delocalize from the mitochondria in HeLa cells. (A and B) Native Tom20-mCherry fusions localize to the mitochondrial network. (C) mCherry-golgi localization control. (D to F) Tom20 (Δ25-31)-mCherry and (G to I) Tom 20 (Δ34-51; K27S/R28S/R29S)-mCherry delocalize to exhibit a random distribution or golgi localization. Transfected cells were imaged 24 hours following transfection. Scale bar, 20 μm. Some of the images in Figure 4.16 are credited to Yidan Ding

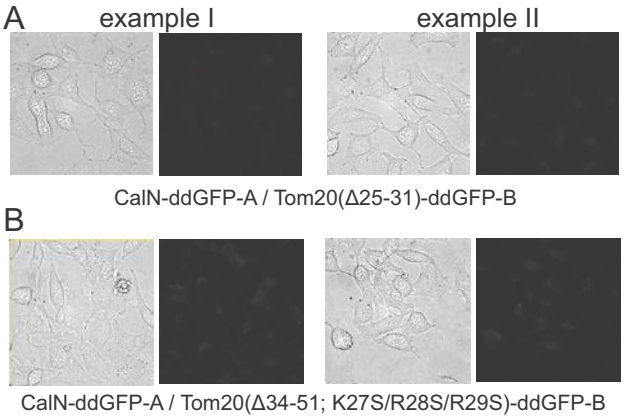


Figure 4.17. Delocalization of Tom20 prevents fluorescent labeling of the MAM. Two representative examples of HeLa cells co-transfected with calnexin (CalN)-ddGFP-A and delocalized Tom20 truncation-ddGFP-B fusions. For each example, left panels show DIC acquisitions and right panels show fluorescence acquisition. Depicted in row (A) are examples for the Tom20 (Δ25-31)-ddGFP-B truncation and in row (B) are examples for the Tom20 (Δ34-51; K27S/R28S/R29S)-ddGFP-B truncation. Some of the images in Figure 4.17 are credited to Yidan Ding.

4.3 Conclusion

In conclusion, we now have a series of ddFPs which possess a range of favourable features for live cell imaging: medium affinity and contrast (ddRFP), increased brightness and high *in vitro* contrast (ddGFP-AB), and increased brightness and reduced pK_a (ddYFP-AB). This ddFP series highlights both the potential and the challenges of using ddFPs as genetic parts for live cell biosensing. Specifically, each of the tools developed to date fall short of being an ideal ddFP which would be characterized by low affinity, low pK_a, high contrast, and high brightness.

4.4 Materials and Methods

4.4.1 Recombinant DNA techniques and reagents

All DNA manipulations including small scale preparation of plasmid DNA, restriction enzyme digestion, ligations, and agarose gel electrophoresis were performed according to Sambrook et al. [221]. Oligos were purchased from IDT Technologies. Unique numerical identifiers (e.g. Oligo 4.1) and corresponding sequences for all oligos are compiled in Table A1 in Appendix A. All restriction enzymes and *Pfu* DNA polymerase were obtained from Fermentas/Thermo Scientific and T4 DNA ligase was obtained from Invitrogen/Life Technologies. *Taq* polymerase was obtained from New England Biolabs. The GeneJet® Gel Extraction or Plasmid Miniprep Kits (Fermentas/Thermo Scientific) were used according to the manufacturer's instructions. Ligations were transformed into electrocompetent DH10B *E. coli*.

Standard PCR amplifications were carried out using *Pfu* DNA polymerase. PCR reactions were performed in volumes of 50 µl containing nuclease-free water, 1× reaction buffer, 200 µM dNTPs (Invitrogen), 200 nM forward and reverse oligos, 10 - 50 ng of template DNA, and 1.0 Unit of *Pfu* DNA polymerase. PCR reactions were generally performed according to the following cycling parameters: initial denaturation at 95°C for 50 seconds; 20 cycles of 95°C for 15 seconds, 55-60°C for 20 seconds, 72°C for 60 seconds per kb of gene target; final extension at 72°C for 4 minutes. Where indicated, overlap extension PCR was performed as described in Bessette et al. [222]. Error prone PCR amplifications were performed as described in Cirino et al. [173] using *Taq* polymerase. Gene shuffling amplifications were generally carried out according to Zhao et al. [180]. For gene shuffling amplifications, the following cycling parameters were used: initial denaturation at 95°C for 50 seconds; 65 cycles of 95°C for 15 seconds, 55°C for 20 seconds, 72°C for 10-12 seconds and a final extension at 72°C for 4 minutes.

The general cloning strategy for each construct or library is outlined in Table 4.1. Destination plasmids, relevant oligos, and restriction enzyme sites used to clone the designated constructs are provided in the table. Brief descriptions of cloning strategies are included in the methods where appropriate.

Two modified pBAD plasmids, pBAD5' and pBAD3' were generated for sub-cloning individual heterodimer partners directly from tandem heterodimer fusions. The design of these plasmids was described in Chapter 3.

4.4.2 Wavelength selection and library screening with tandem heterodimers.

Library screening was performed using the trypsinolysis assay described in

Chapter 3. New tandem heterodimers were initially constructed by sub-cloning 5' genes, containing the hue-shifting mutations, as fusions with ddRFP-B to provide a chimera of the form A-linker-B. The full-length fusion was cloned between *XhoI/HindIII* sites of pBAD/His B and the individual A and B genes were linked with a sequence encoding a 23-residue sequence (GHGTGSTGSGSSGTASSEDNNMA) that included a *KpnI* site. Ligations were transformed into electrocompetent DH10B *E. coli* and plated to LB agar media supplemented with 400 µg/ml ampicillin and 0.02% L-arabinose. Bright colonies were picked and cultured in LB media supplemented with 100 µg/ml ampicillin and L-arabinose (0.02%) overnight. Crude B-PER protein extracts were treated with trypsin at approximately 10 µg/ml for 30 minutes and emission spectra were acquired using a 96-well microplate reader. Contrast ratios were calculated as the integrated emission peak area of the non-trypsinized extract divided by that of the trypsinized extract.

Hue-shifted emission wavelengths were selected using our custom-built colony imaging system; variants exhibiting favourable emission wavelength ratios (e.g. green-to-red) were confirmed by trypsinolysis assay and carried forward for further mutagenesis.

4.4.3 Protein purification

To produce recombinant proteins, the genes encoding tdGFP and tdYFP were cloned into pBAD/His-B (Invitrogen). The monomeric constituents were subcloned in pBAD5' (ddFP-A) or pBAD3' (ddFP-B). Plasmids were transformed into electrocompetent DH10B *E. coli*. Shaker cultures at an optical density of 0.5-0.7 were induced with 0.02% L-arabinose and allowed to incubate a further 12-16 hrs. Cell pellets were resuspended in 10 mM Tris-HCl, 0.5 M NaCl, 10 mM

imidazole pH 8.0 and lysed using a cell disrupter (Constant Systems Ltd.). Cleared lysates were obtained by centrifugation at 13,000 rpm in 50 ml conical tubes for 50 minutes. Lysates were incubated with Nickel-NTA on ice with agitation for 30 minutes and then batch purified through a fritted column using a vacuum manifold. Protein-bound resin was washed several times with a total of 50 mL 5 mM Tris-HCl, 0.5 M NaCl, 20 mM imidazole pH 8.0. The final 10 mL of wash was performed by gravity flow. Proteins were eluted using 5 mM Tris, 0.5 M NaCl, 250 mM imidazole pH 8.0 and dialyzed into 20 mM Tris-HCl, 100 mM NaCl, 0.5 mM EDTA pH 7.4.

4.4.4 pH sensitivity determination

pH sensitivity measurements were performed by incubating purified proteins in buffers of desired pH and acquiring emission spectra with a 96-well microplate reader (Tecan Safire2^{TM}). A 1 μM solution of FP was prepared in 5 mM Tris-Cl, pH 7.5 and diluted 1:10 with a universal buffer of desired pH. Britton Robinson universal buffer was used to generate buffers of desired pH. A stock solution was prepared by mixing 50 mL each of 0.04 M H_3BO_3, 0.04 M CH_3COOH, and 0.04 M H_3PO_4. The pH was adjusted to the desired value by adding 0.2 M NaOH to the prepared stock solution. The pK_a was determined as described in Chapter 3.

4.4.5 Spectral feature determination

All spectra represented in this chapter were recorded with a DU-800 UV-visible spectrophotometer (Beckman) or a QuantaMaster spectrofluorimeter (Photon Technology International, Inc.). Absorbance measurements were made using a 1 cm quartz microcell cuvette. The alkaline chromophore denaturation method was used to determine ε values [235]. Enhanced GFP (Φ = 0.60) or

mCitrine (Φ = 0.76) was used as the reference for quantum yield determination for ddGFP-AB and ddYFP-AB, respectively.

4.4.6 Dissociation constant determination

To determine the K_d of the purified recombinant ddGFP partners, an increasing amount of non-fluorescent ddGFP-B was mixed with a fixed amount of ddGFP-A (250 nM final) to generate ddGFP-AB complexes in 20 mM Tris-HCl, 100 mM NaCl, and 0.5 mM EDTA, pH 7.5. The integrated fluorescence emission peaks recorded as a function of ddGFP-B concentration were used to generate saturation binding curves. Experimental data was fit using a modified Langmuir isotherm as described in Chapter 3.

4.4.7 Live cell applications

Caspase-3 biosensors were designed as described in Chapter 3. The genes encoding the biosensors were generated by three-part ligation as outlined in Table 4.1. For caspase-3 imaging, apoptosis was initiated by treatment with 2 μM staurosporine and imaged in HEPES-buffered Hank's balanced salt solution (HHBSS).

For membrane-targeting fusions the human Akt Pleckstrin homology (PH) domain (residues 5-108) was fused at the N-terminus of ddGFP-A with an eight residue linker (GSSGTASS). The K-Ras polybasic farnesylation substrate sequence, SKDGKKKKKKSKTKCVIM [242, 252], was installed at the C-terminus of ddGFP-B. Each gene fusion was cloned into pcDNA3.1 (+) (Invitrogen) between *HindIII* and *XhoI* restriction sites.

Tom20 and ddGFP-B were fused with a short linker (GTASSEDNNMA) to render Tom20-ddGFP-B (credit to Yidan Ding). Calnexin and ddGFP-A were

fused together using the linker GGASGGSGSGPV to render calnexin-ddGFP-A (credit to Yidan Ding). The following Tom20 delocalization mutants were generated: Tom20 (Δ25-51) and Tom20 (Δ34-51; K27S/R28S/R29S) [251]. Each of these truncations was in turn fused to ddGFP-B.

Localization markers were mCherry-mito-7 (mitochondria), mCherry-SiT-N-15 (golgi), and ER-mCherry (mCherry fused to calreticulin signal peptide and a KDEL retention sequence). Localization markers were the kind gift of Michael Davidson (NHMFL, The Florida State University).

4.4.8 Mammalian cell culture and imaging

HeLa cells were maintained in DMEM supplemented with 10% fetal bovine serum and glutamax (Invitrogen) at 37°C and 5% CO_2. Transient transfections of pcDNA3.1 (+) expression plasmids were performed using Turbofect (Fermentas/Thermo Scientific). For single plasmid transfections 750 ng plasmid DNA was used, and 375 ng of each plasmid was used for co-transfection of two plasmids. Generally, 2 µl of transfection reagent was used per 1 µg of DNA transfected. In control experiments, mock transfections were performed with the pcDNA3.1 (+) to control for the amount of DNA transfected. Twenty-four hours post-transfection, cells were imaged in HHBSS. Imaging was performed on an inverted Nikon Eclipse Ti microscope equipped with a 200W metal halide lamp (PRIOR Lumen) and a QuantEM: 512SC 16-bit cooled CCD camera (Photometrics). Colocalization data was collected using a spinning-disc confocal Olympus IX-81 motorized microscope. Green and red excitations were conducted with 50mW 491nm and 561nm pumped diode lasers. The two lasers were coupled to the spinning disc confocal head (CSU10; Yokogawa) mounted with a Sedat dichroic mirror (Semrock). The lasers were processed with appropriate

filter sets (Semrock) to capture fluorescence images with an EMCCD (C9100-13, Hamamatsu), driven by Perkin Elmer's Volocity software.

Table 4.1. Cloning strategy for DNA constructs described in Chapter 4.

		Amplification 1		Gene assembly	
Construct	**Vector**	**Template**	**Oligos[1]**	**Overlap Extension oligos**	**RE sites (5'/3')**
tdRFP-A71M	pBAD td**A**$_1$**B**$_{1.1}$	A_1-pBAD5' A_1-pBAD5'	4.8/4.1 4.7/4.2	4.7/4.8	*XhoI/KpnI*
tdRFP-A71M/V105A	pBAD td**A**$_1$**B**$_{1.1}$	A_1-pBAD5' A_1-pBAD5' A_1-pBAD5'	4.7/4.4 4.3/4.2 4.1/4.8	4.7/4.8	*XhoI/KpnI*
tdRFP-M66C	pBAD td**A**$_1$**B**$_{1.1}$	A_1-pBAD5' A_1-pBAD5'	4.5/4.7 4.6/4.8	4.7/4.8	*XhoI/KpnI*
A-DEVD-B[2]	pcDNA 3.1 (+)	A (pBAD5') B (pBAD3')	3.43/3.60 3.45/3.50	3-part ligation	*HindIII/ XhoI*
PH-ddGFP-A	pcDNA 3.1 (+)	Akt ddGFP-A	4.14/4.16 4.15/4.13	4.14/4.13	*HindIII/ XhoI*
ddGFP-B-Farn	pcDNA 3.1 (+)	ddGFP-B pcDNA3.1 (+)	4.9/4.12 4.11/4.10	4.9/4.10	*HindIII/ XhoI*

[1] Sequences for oligonucleotide primers are listed in in Appendix A (Table A1).

[2] Generic 'A' and 'B' designations refer to ddFP-A and ddFP-B for each of ddGFP, ddGFP (K153E), and ddYFP.

Chapter 5 Conclusions and Future Directions

FPs facilitate the production of stunning and eye-catching cellular images. Scientists, however, see more than just 'pretty colours'. FPs are transformative imaging and biosensing tools that allow researchers to peer at the inner workings of living cells and make observations in real time. The many different colours now available are used in a large variety of applications throughout the biological research community. However, despite the quick expansion of the colour palette and its widespread use, FP technology remains a burgeoning field. The development of new tools is required.

Today the most prevalent use of FPs remains as optical reporters of subcellular localization. Researchers use FPs to trace the position and, in some cases, the movement of proteins of interest. More advanced uses of FPs are, by comparison, few in number. Advanced applications utilize FPs as biosensors for dynamic cellular processes, such as monitoring the flux of small molecules, detecting PPIs, and reporting on enzymatic reactions. Although reported examples of FP-based biosensors are increasingly numerous, there are a limited number of design platforms available (*i.e.,* FRET, BiFC, and single-FP biosensors). It is perhaps for this reason that the use of FP-based biosensors is still dwarfed by reports of FP localization markers. The general goal of this work was to address this deficiency by developing a new template for FP-based biosensor design.

In this thesis, we described the development of two technologies: a protein interaction-mediated toxin silencing (PINTS) assay for detecting PPIs in *E. coli*

and dimerization-dependent FPs. The two technologies were linked by our intent to use the PINTS strategy to identify FP heterodimers from interface libraries of a dimeric RFP known as dTomato [9]. However, in the process of optimizing the PINTS assay, we discovered an alternate screening method to detect RFP heterodimers and proceeded with the latter. In this concluding chapter, we present a brief summary of results and our general perspectives for each technology. Prospects for future applications are also discussed.

5.1 PINTS assay: General perspectives

In Chapter 2 we described the development of a novel *in vivo* selection assay for identifying PPIs. In this assay, survival-based selection is governed by a low affinity Bn/BS pair and operates on the principle of PPI-mediated inhibition of a cytotoxic protein. The system utilizes a single plasmid, two gene expression platform to express both bait-Bn and prey-BS hybrids. Bait proteins are fused to attenuated Bn-H102K and prey proteins are fused to BS-Y29I/Y30G. The decreased K_d of this Bn/BS pair renders Bn free and lethal in the absence of a bait/prey interaction, but sequestered and innocuous when an interaction occurs.

We reported several results to support the use of PINTS as a suitable strategy for identifying PPIs. Two fundamental experiments were essential to validate our design. The first was the inhibition of growth by co-expression of mTomato as prey with dTomato bait, whereas expression of dTomato as both prey and bait was growth permissive. Similarly, the interface-disrupting mutation, R89L, prevented growth of *E. coli* co-expressing Ras and RBD_{R89L}, whereas co-expressing the WT proteins permitted growth. PPI interface disruption

experiments of this type are essential to *in vivo* selection assay development [34, 40].

Our PINTS assay differs from PCA-based *in vivo* selection strategies in a few regards. Our method utilizes a single plasmid system, whereas conventional PCA strategies rely on two plasmids. Co-transformation of two plasmids into *E. coli* is typically inefficient [253], and can therefore limit the library diversity that can be effectively explored. Comparatively, our single plasmid system is limited only by the transformation efficiency of the *E. coli* strain used. An additional difference is the use of folded Bn and BS proteins, rather than unfolded split protein fragments used in PCA strategies [24]. The use of unfolded domains can lead to unspecific *E. coli* growth through aggregation of overexpressed proteins [19, 219]. Theoretically, these aggregation artifacts should have minimal effect in the PINTS assay, but such problems likely affect *in vivo* selection strategies universally (as our results may reflect).

In our PINTS selection, the Bn 'executioner' appeared to generate genotoxic stress. Two key observations suggest genetic instability, including growth of large opaque colonies whose plasmids were characterized by rearrangements and the acquisition of a Bn activity-disrupting mutation in a single clone. However, the colonies exhibiting Bn-induced genotoxic stress were rare and should not limit the use of the PINTS assay for simple validation of PPIs.

We believe optimization of the assay could further reduce the frequency of genotoxic artifacts. For example, to preclude the occurrence of growth caused by plasmid rearrangements, additional antibiotic resistance markers could be incorporated in the PINTS plasmid. Large DNA rearrangements that truncate Bn would likely affect the antibiotic resistance genes as well. Plating libraries on

media supplemented with multiple antibiotics would thus prevent growth of rearrangement-affected clones.

Despite the drawbacks associated with Bn, we believe the assay is promising and could be further characterized to solidify its stake as a useful *in vivo* selection method for PPIs. The results from our Ras suppressor selection highlight the potential utility of PINTS. Despite examining only a small sample set of potential binders, we were able to isolate a genuine Ras suppressor that exhibited heteroselectivity for RBD_{R89L} in live cells. We suggest the Ras suppressor selection experiments be repeated with a more thorough statistical characterization to generate accurate numeric values for parameters such as survival event frequency, false positive rate, and 'hit' rate for genuine suppressors. If the data obtained supports our preliminary results, the PINTS assay may be a useful tool to screen genomic libraries for interacting partners of desired protein targets.

We further suggest *in vitro* binding assays be performed to verify potential PPIs between Ras variants and RBD_{R89L}. Surface plasmon resonance (SPR) assays performed with purified Ras variants and the RBD_{R89L} would be well suited for these validations. SPR would not only confirm the occurrence of PPIs, but also provide kinetic data in the form of binding constants. It would be especially interesting to test our apparent false positive clones to see if they are actually low affinity interacting partners as we proposed in section 2.2.6. An alternative method relevant to this thesis could utilize the ddRFP system. Potential Ras suppressors could be fused to ddRFP-$\mathbf{A}_1$ and the RBD_{R89L} could be fused ddRFP-$\mathbf{B}_1$. Simple mixing of the purified protein fusions could indicate the occurrence of an interaction by measuring fluorescence intensity relative to appropriate controls. Dissociation constants could also be determined using this

method. Each of these techniques would require a GTP loading step for the Ras variants, since, the RBD binds only to the Ras-GTP conformer [254].

5.2 PINTS assay: Future applications

A relatively straightforward future application for the PINTS assay is adaptation for use in mammalian systems. Mammalian proteins often exhibit limited solubility in bacteria (e.g. [255]) which limits the use of *E. coli* based PPI screens. Further, a limited number of *in vivo* selection assays are available for mammalian systems [19, 24, 34]. Given the toxicity of Bn to mammalian cells [256] the PINTS system could provide a rapid means to select for mammalian PPIs. Furthermore, the assay could be deployed in a high throughput manner to screen chemical libraries for therapeutically significant inhibitors of PPIs.

PPI assays can be used to pan genomic libraries with the objective of identifying binding partners for a given bait protein-of-interest. These types of assays are unquestionably valuable for piecing together protein interaction networks and assembling interaction networks [257]. However, an underappreciated utility of these assays is their potential as vehicles to discover new protein interfaces and engineer novel interactions [219, 253, 258-261]. Indeed, the design and discovery of heterospecific PPIs is a major objective for protein engineers [163, 262]. Given our discovery of a Ras suppressor for RBD_{R89L} from a modestly large library, we believe the utility of the PINTS assay could be extended to screen larger libraries for proteins with novel binding specificities.

We envision the assay could be modified to include several cycles of serial dilution and bacterial outgrowth in liquid media. Over many cycles this strategy

will select for clones exhibiting a growth advantage. Since growth is contingent on PPIs, high affinity interactions should be enriched with increasing cycle number. We have observed a growth rate dependence reflecting (assumed) homodimerization affinities of dTomato variants, H162K, A164R, and the double mutant H162K/A164R. We predict the homodimerization affinity of these variants follows an H162K > A164R > H162K/A164R hierarchy. Co-expression of these variants as Bn-H102K fusions with dTomato-BS-Y29I/Y30G shows growth-dependence (on agar media) that correlates with this prediction. These results seem to contradict the results obtained for our Ras suppressor screen in which the survival of false positives indicates a lack of reward for genuine binders versus unspecific interactions. Nonetheless, if the result observed for our dTomato affinity variants is valid, the PINTS assay could be used to enrich for high affinity binders using the iterative culture dilution and outgrowth strategy described above. This strategy would necessitate the absence of rearrangement-prone clones. If this proved to be a problem, plating to agar growth media in between each cycle could be performed to manually 'filter out' these clones.

Even without enrichment for high affinity binders, our ability to identify a Ras suppressor to RBD_{R89L} in a single round of *in vivo* selection shows the PINTS concept is suitable for identifying novel PPIs. The prevalence of false positives will likely, however, preclude the screening of naïve libraries. Engineered non-natural PPIs could be useful tools for studying biochemical pathways, creating scaffolds with novel binding specificities, or even have therapeutic applications.

5.3 DdFP technology: General perspectives

The primary purpose of the research presented in this thesis was to develop a new class of biosensor design strategy, which we designated as dimerization-dependent FPs. This technology is characterized by two dim or non-fluorescent monomeric FPs that reversibly associate to form a fluorescent heterocomplex. Using a red FP model (ddRFP) we illustrated how this strategy can be applied for detection of PPIs *in vitro*, as well as fluctuations in small molecule concentrations and enzymatic activities in live cells. Furthermore, we have extended the colour palette of the ddFP family to include both green and yellow hued variants.

The dimerization-dependent FP platform described in Chapters 3 and 4 represents a technology in its infancy. However, before proceeding with its further development one must carefully consider whether the improvements in optical properties will substantiate the effort required to achieve them. An ideal ddFP system will incorporate several features: fluorogenicity, high contrast, a requirement for monomeric constituents, low affinity, and low pK_a. Given the directed evolution tenant, “you get what you screen for” [185], the engineering of an ideal variant that incorporates all of these components may not be possible and a compromise of features may be the only feasible outcome. Indeed, each variant of the ddFP family described in this thesis represents a compromise; certain attributes were gained at the expense of others.

At a minimum, the current colour palette of ddFPs could be expanded using a similar strategy to that employed for making green and yellow variants. In addition to providing more colour options for researchers to choose from, the engineering of new colour hues can lead to fortuitous discoveries of desired optical properties. For example, the directed evolution of ddYFP led to the

important discovery of a variant exhibiting a reduced pK_a. This particular feature is highly desirable for live cell imaging and careful effort should be invested to preserve this phenotype in subsequent generations of ddFPs. This discovery shows engineering new colour variants may be worthwhile solely on the basis for potential to discover clones that exhibit desirable spectral features.

Logical targets for expansion of the colour palette include blue and far-red variants. Similar to the strategy described in Chapter 4, ddRFP is likely the best starting template to evolve these colour hues since blue and far-red 'chromo-shifting' mutations have already been reported for dsRED-derived FPs. For example, mBlueberry variants (λ_{em} ~450-470 nm) were derived from mCherry2 [82] and a far red protein, mPlum, (λ_{em} 649 nm) was evolved from mRFP [174]. Additional examples of blue-shifting and far red-shifting mutations have also been documented [84, 263, 264]. This strategy may have a limited probability of success, however. The 'literature-mining and residue grafting' approach to discover additional ddFP hues may not be suitable on account of the extensively remodeled amino acid landscape of ddRFP relative to dsRED. Indeed, this approach failed in our first attempt to 'greenize' ddRFP; well-characterized mutations [71] that convert the dsRED chromophore to a green-emitting species failed to generate green variants in the context of ddRFP. The evolution of additional hue-shifted ddFP variants may necessitate the investigation of alternative templates. For example, to obtain blue-shifted variants, specifically cyan or teal hues, dimeric teal FP 0.7 [11] is a promising candidate. Similar to dTomato, monomerizing mutations have been documented for this FP [11]. Therefore, a directed evolution strategy akin to that used for ddRFP would likely render a teal-hued ddFP.

Another intriguing template is *A. victoria* GFP. The rationale for choosing GFP as a template is its inherent propensity for low affinity homodimerization. The engineering of low affinity partners is of the utmost importance to ddFP technology and therefore, homodimeric GFP is an enticing template. Adding further credence to this approach is the FRET donor/acceptor pair known as CyPet and YPet [134]. These GFP-derived FPs were evolved to be 'sticky' for the purpose of optimizing FRET efficiency between partners. The mutations responsible for the 'stickiness' are situated at the GFP homodimer interface. Therefore, we envision remodeling the GFP interface residues surrounding the chromophore to generate a fluorogenic pair exhibiting a high K_d. We acknowledge engineering fluorogenicity into this system represents a significant challenge. Nonetheless, developing a fluorogenic variant that exhibits a K_d similar to that of GFP (>100 μM) would represent a major advancement in ddFP technology.

Finally, a recently reported Hydrozoan GFP-like FP, called abeGFP may also represent a promising template [265]. The discovery of abeGFP marked the first documented incidence of a naturally occurring FP concatenate [265]. We speculate it forms a tandem dimer with a complementary interface, as opposed to a beads-on-a-string structure, but no evidence exists to support this claim as of yet. AbeGFP was discovered in the siphonophore jellyfish, *Abylopsis eschscholtzii,* and its two domains are chemically distinct, sharing only 59% sequence identity [265]. The FP domains are joined by a 19 amino acid linker [265], which makes the tandem remarkably similar to the artificially engineered tdTomato, as well as tandem dimer ddRFP [9, 240]. Similar to our ddFPs, the N-terminal abeGFP domain exhibits minimal fluorescence in *E. coli* and expression of the C-terminal domain results in no detectable fluorescence [265]. However,

the naturally occurring tandem fusion exhibits strong fluorescence; the brightness of the tandem exceeds that of enhanced GFP (Φ = 0.66 and ε = 168,000 $M^{-1} \cdot cm^{-1}$) [265]. With a few reasonable assumptions, it appears as though abeGFP is a naturally occurring fluorogenic heterodimer and may represent an ideal template for engineering future generations of ddFPs. Furthermore, abeGFP partners share only 59% sequence identity, whereas the ddRFP partners share 87% sequence identity. We interpret this as encouraging since there is more sequence diversity to explore through directed evolution, which could lead to brighter variants of ddFPs with optimized attributes.

We believe further directed evolution of ddFPs to generate alternate colour hues is a relatively low-risk endeavour. However, such efforts could be described as semi-rational at best and would likely lead to ddFPs variants that exhibit a compromise of attributes. The major limitation to performing a more rational evolution is the lack of crystallographic data on ddRFP. Currently, the DsRed structure [64, 74] is the only model we can use to guide our evolution, but the ddRFP sequence has diverged greatly from this ancestor. Therefore, to make improvements to the optical properties of ddFPs, obtaining crystal structures of the heterodimer, as well as the free unbound monomers, will be important. High-resolution structures would surely provide insight into the mechanism of fluorogenesis and the chemistry of the chromophore microenvironment. Resolution of these details will permit targeted design strategies to guide the evolution of ddFPs toward desired phenotypes. Specifically, the molecular determinants of pK_a modulation would be invaluable for directed evolution of reduced pK_a variants. Despite the apparent need for crystal structures, we acknowledge pursuing such experiments may depend on the research community's demand for improved variants. If ddFP technology is embraced,

pursuing crystals will almost certainly be warranted. In the interim, however, alternative strategies could be used in an attempt to engineer optimized ddFP variants.

The utility of future ddFP generations will largely depend on generating low affinity variants that maintain high fluorogenic contrast. A possible strategy to evolve for these properties is presented in Figure 5.1 and relies on fluorescence activated cell sorting, or FACS, and the use of small chemical inducers of dimerization, or CIDs. We envision a system where each ddFP monomer is fused to a protein domain that forms a CID-dependent ternary complex (Figure 5.1A). For example, DHFR and FKBP modules form a heterocomplex with the molecule, Mtx-SLF [266]. Alternatively, the FRB and FKBP domains form a complex with rapamycin [46, 267].

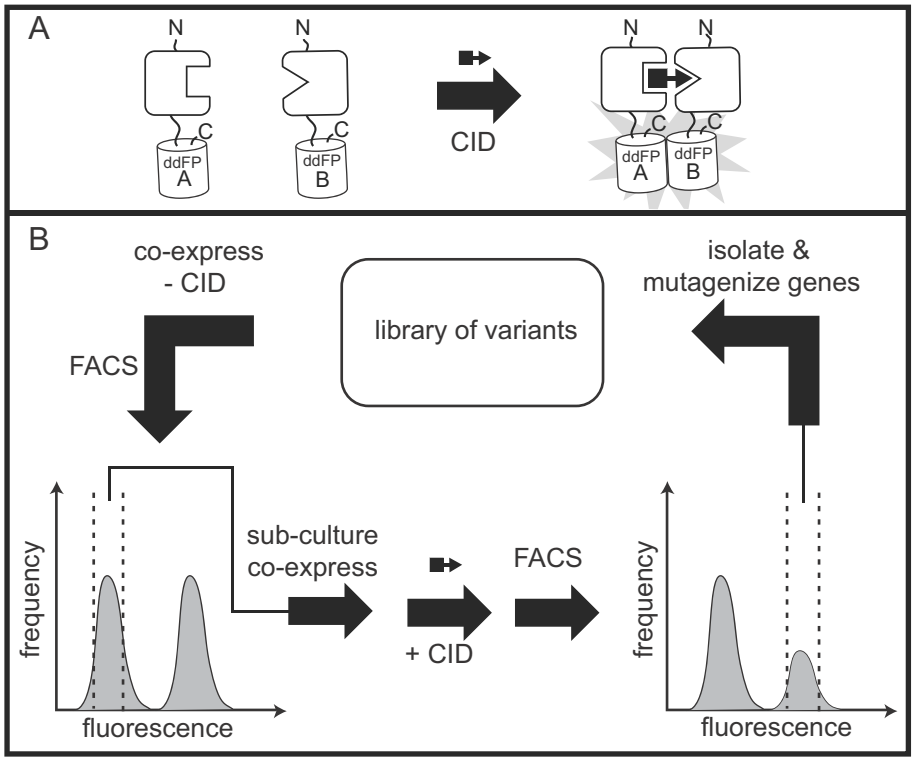


Figure 5.1. Schematic of FACS-based strategy for the discovery of low affinity high contrast fluorogenic heterodimers. (A) DdFP partners are fused to CID-dependent dimerization domains. Addition of the CID results in development of

fluorescence. (B) Iterative retention of the dim cell population in the absence of the CID, followed by retention of the fluorescent cell population in the presence the CID selects for low affinity high contrast ddFPs. The cell populations (shaded in grey) between the dashed lines represent cells to be retained for further directed evolution.

Mutagenized libraries of ddFP-A and ddFP-B copies could first be screened independently to isolate clone pools with low basal fluorescence. These A- and B-copy libraries could then be panned against each other by co-expressing them using a low copy number plasmid. The low copy number plasmid would ideally keep the intracellular concentrations of the expressed proteins low [268], preventing aggregation caused by high expression levels.

The fusion proteins would first be expressed in the absence of a CID, and sorted by FACS with retention of the non- or dim-fluorescent cell population (Figure 5.1B). The non-fluorescent clones could then be grown in the presence of a CID and again be subject to sorting by FACS. In this instance, the population of cells exhibiting the greatest fluorescence would be collected (Figure 5.1B). We anticipate repeating this cycle several times would lead to the discovery of high contrast and low affinity fluorogenic partners.

5.4 DdFP technology: Future applications

In this work we have demonstrated several uses of ddFP technology. Here we present a few additional possibilities for applied uses of ddFPs. These suggestions aside, it is our hope the research community will embrace this technology, and find new and exciting ways to utilize this template for biosensor design.

In our description of ddRFP in Chapter 3, we proposed protease sensing was particularly suitable for ddFP technology. In our example, we showed how

ddRFP could be used to sense caspase-3 activation in live cells. Specifically, we introduced a well-known caspase-3 substrate sequence into a flexible linker sequence joining the ddRFP partners. It is conceivable that the ddRFP template could be used to probe additional caspase-3 substrates. Indeed, ddRFP has subsequently been used to probe a predicted caspase-3 substrate in a mammalian transcription factor protein (unpublished personal communication; Dr. R. Ingham, University of Alberta). Using a ddRFP-style protease sensor design, a proposed caspase-3 substrate in the transcription factor was used to link ddRFP partners. Poor transfection efficiency has thus far limited imaging, but the products of *in vitro* caspase-3 cleavage were detected by immuno-blotting using an antibody raised against RFP (unpublished personal communication; Dr. R. Ingham, University of Alberta). A control design with a mutated substrate site did not show cleavage products. This result highlights the versatility of ddRFP as a template for protease sensing, insofar as tolerating different substrate sequences. It follows that additional substrates could be tested for caspase-3 susceptibility using the ddRFP platform.

Aside from protease sensing, we believe another promising area where ddFP technology could be useful is in BiFC-style assays. Although in its current state (limited brightness and contrast) it is unlikely to usurp BiFC approaches, ddFPs may prove more suitable in some cases on account of their reversible nature. BiFC is an irreversible phenomenon [17, 269]. The reversible nature of ddFPs, at least in the case of ddRFP, will allow researchers to probe dynamic interactions inside live cells [240]. A further advantage is the high solubility of the ddFP monomers, whereas BiFC fragments often suffer from poor solubility [270], limiting their general applicability.

An interesting biological application of ddFP technology could be imaging of dynamic communication between neurons. Split FP complementation was recently used to label neuron-to-neuron contacts and synapses [271, 272]. Due to the inherent irreversibility of BiFC, this technique is limited to fluorescently labeling single synaptic events. DdRFP could be an attractive alternative since it may permit more dynamic analyses. In theory, ddRFP could trace neuron-to-neuron contact and axon withdrawal [273]. In unrelated experiments, we have already demonstrated that ddRFP can function on the surface of live cells; co-expression of ddRFP monomers on the surface of the same cell results in fluorescent labeling of the cell periphery (see Appendix E, Figure E1). This result, although preliminary, is an encouraging indicator that such a strategy could work with additional optimization to better orient the ddFP partners on the cell surface of neurons.

Finally, assuming ddFP technology can be improved upon, we suggest it is potentially suitable for high-throughput screening (HTS) of PPIs or inhibitors thereof. Its application to HTS strategies would make ddFP technology an invaluable tool for drug discovery applications in that it could be deployed in distinct roles. Firstly, ddFPs could serve as the primary screening platform for inhibitors of PPIs, where the absence of signal would indicate an inhibitor. Secondly, ddFPs could be used as a quick method to validate ‘hits’ obtained from better-established high-throughput inhibitor screens. We must emphasize, however, that ddFP technology is not currently suitable for HTS due to its medium-to-high affinity and the residual fluorescence associated with the A-copy. If these two drawbacks are improved, the technique could be powerful in HTS of potential drugs.

5.5 Concluding remarks

In summary, the research presented in this thesis described our efforts to develop a reversible dimerization-dependent fluorescent protein biosensing strategy. With this purpose in mind we designed and validated an assay suitable for detecting protein-protein interactions in *E. coli*. Although this assay was not ultimately used in the directed evolution of ddFPs, it was proven useful for the identification of a novel Ras/RBD interface. Using a series of genetic screens and extensive directed evolution we were successful at engineering a new heterodimeric red FP interface from a homodimeric dTomato progenitor. The red FP heterodimer is comprised of quenched monomeric FP constituents that reversibly associate with development of fluorescence. DdRFP was subsequently shown to be useful for live cell intensiometric detection of small molecules and enzymatic activities. We further engineered the system to extend ddFP technology to include green and yellow variants. We believe ddFPs mark a significant addition to the FP toolkit and will facilitate the design of useful live cell biosensors.

Appendices

Appendix A: Oligonucleotide sequence supplement

Table A1. List of oligonucleotides used in this work

OLIGO #	Lay ID	SEQUENCE (5' TO 3')
2.1	T7_seq_F	TAATACGACTCACTATAGGG
2.2	BGH_seq_R	TAGAAGGCACAGTCGAGG
2.3	SA_BAD_F	ATGCCATAGCATTTTTATCC
2.4	SA_seqR1	ACTCAGGAGAGCGTTCAC
2.5	SA_tac_F	GACATCATAACGGTTCTGGC
2.6	SA_SeqR2	GTTAGCAGCCGGATCTCAC
2.7	ecoFP_F	CGGAATTCATGGTGAGCAAGGGCGAGG
2.8	kpnFP_R	CCGGTACCCTTGTACAGCTCGTCCATGCCTG
2.9	kpnBn_F	GCCGGTACCGGCGGTGGCGGTAGCGGTGGCGG TGGCAGCGCACAGGTTATCAACACGTTTGACGG GGTTGCGG
2.10	bglBnH102K R	GAAGATCTTTATCTGATTTTTGTAAAGGTCTGATA CTTGTCCGTTG
2.11	Xho_FP_F	CAGGCTCGAGCATGGTGAGCAAGGGCGAGGAG ACC
2.12	EagFP_R	TTGACCCGGCCGTCTTGTACAGCTCGTCCATGC CTGCC
2.13	'BS_F'	TTGACCCGGCCGGCGGTGGCGGTAGCGGTGGC GGTGGCAGCATGAAAAAAGCAGTCATTAACGGG GAACAAATCAG
2.14	'BS_R'	GAGCTAAGCTTTTAAGAAAGTATGATGGTGATGT CGCAGCC
2.15	HinRasF	TTTAAGCTTGCCACCATG ACAGAATACAAGCTCGTTG
2.16	Ras_myrPC_ R	CAGCCGGGGCCACTCTCATCAGGAGGGTTCAGC TTCCGCAGCTTGTGCTGCCGGATCTCACGC
2.17	Ras_myrpc_ F	TGGCCCCGGCTGCATGAGCTGCAAGTGTGTGCT CTCCTGACTCGAGTCTAGAGGGCCCGT
2.18	pcD_R	GAGTGGCACCTTCCAGGGTCAAGG
2.19	Raf_mCher_ F	GGATGGGAGCAGTGGGCATGGCATGGTGAGCAA GGGCGAG
2.20	Xho_Cher_R	AAACTCGAGTTACTT GTACAGCTCGTCCATG
2.21	Raf_mCher_ R	ACCATGCCATGCCCACTGCTCCCATCCAGGAAAT CTACTTG
2.22	hinRaf_F	TTTAAGCTTGCC ACCATGCATGGGAGC
2.23	Ras_SupR	GGTCGTATTCGTCCACAAAATGGTTMNNGATCAG CTGMNNGGTCAGCGCACTCTTGC
2.24	Ras_SupF	CATTTTGTGGACGAATACGACCCCACTATAGAGN NKNNKNNKCGGAAGCAGGTGGTCATTGATGGGG AGACG
2.25	Cas2_R	GCAGCAGCCAACTCAGCTTCCTTTC
2.26	Y29I/Y30G_ F	GCTTGCCCTTCCGGAAATCGGCGGTGAAAACCT GGACGC
2.27	Y29I/Y30G_	GCGTCCAGGTTTTCACCGCCGATTTCCGGAAGG

	R	GCAAGC
2.28	BS_seq_R	CCGTAGTATTCCGGAAGGGCAAGC
2.29	Mono_F	GAAGGGCGAGATCAAACAGCGCCTGAAGCTGAA GGACG
2.30	Mono_R	CGTCCTTCAGCTTCAGGCGCTGTTTGATCTCGCC CTTC
2.31	xho Ras F	CAGGCTCGAGCATGACAGAATACAAGCTCGTTGT TGTTGGC
2.32	EcoRafF	CGGAATTCATGCATGGGAGCAACACTATC
2.33	Eag Ras_R	TTGACCCGGCCGTTCCACTGCCGTGCTGCCGGA TCTC
2.34	Kpn_Raf_R	GCCGGTACCACTGCTCCCATCCAGGAAATCTAC TTG
2.35	HinV12_F	TTTAAGCTTGCCACCATGACAGAATACAAGCTCG TTGTTGTTGGCGCCGTGGGTGTGGGC
3.1	pkpnF	GACGGTACCTAACTGTTTTGGCGGATGAGAGAA G
3.2	pKpnR	AGATCTCGAGCTCGGATCCTTATCGTCATCG
3.3	P3'ORF_F	AGATGGTACCGCTCGGATCCTTATCGTCATCGTC
3.4	P3'ORF_R	GACGACAAGCTTTAACTGTTTTG GCGGATGAGA GAAG
3.5	SA_BAD_R	GATTTAATCTGTATCAGG
3.6	MBP_dBgl_F	GCTGATTTATAACAAAGACCTGCTGCCGAACC
3.7	MBP_dBgl_R	GGTTCGGCAGCAGGTCTTTGTTATAAATCAGC
3.8	ecoMBP_F	GCGAATTCATGAAAATCGAAGAAGGTAAACTGG
3.9	kpnMBP_R	GCCGGTACCTGAAATCCTTCCCTCGATCCC
3.10	Mono_F	GAAGGGCGAGATCAAACAGCGCCTGAAGCTGAA GGACG
3.11	Mono_R	CGTCCTTCAGCTTCAGGCGCTGTTTGATCTCGCC CTTC
3.12	Kpn_FP_F	GCCGGTACCGGCGGTGGCGGTAGCGGTGGCGG TGGCAGCATGGTGAGCAAGGGCGAGGAGACC
3.13	Blg_FP_R	GAAGATCTTTACTTGTACAGCTCGTCCATGCCTG CC
3.14	Sat_Y191_F	GCAACTGCCCGGCNNKTACTACGTGGACAC
3.15	Sat_Y149_R	GGAGGCCTCCCAGCCCATGGTCTTC
3.16	Sat 149_F	CCATGGGCTGGGAGGCCTCCACCGAGNNKCTGT ACCCCCG
3.17	Sat_H162_F	CCCGCGACGGCGTGCTGAAGGGCGAGATCNNK CAGGCCCTGAAGC
3.18	SatH162_R	CCCTTCAGCACGCCGTCGCGGGGGTACAGG
3.19	Tom_dFLY_ R	TTCCAAGCTTTTACTTGTACAGCTCGTCCATGCC CAGGTGGTGGCGGCCCTCGGAGCGC
3.20	E160_F	GCGACGGCGTGCTGAAGGGCNNKATCBHYCAGG CCCTGAAGC
3.21	E160_R	CTTCAGCACGCCGTCGCGGG
3.22	Xho_FRB_F	TCCGAGCTCGAGAATGGCTTCTAGAATCCTCTG
3.23	Xho_FKBP_ F	TCCGAGCTCGAGAGGA GTG CAG GTG GAA ACC ATC
3.24	cI_KpnHetB2 _F	GCGGTACCATGGTGAGCAAGAGCGAGGAG
3.25	FKBP_ci_R	CTTCTAAGTGACGGATAACTAGTTTCCAG
3.26	Ci_linkF	GTTATCCGTCACTTAGAAGTGAGTATG
3.27	cI-link-FP_R	GCTCACCATGGTACCGCTTACCCATCTCTCCGC

3.28	Kpn_FP_F2	GCGGTACCATGGTGAGCAAGGGC
3.29	FRB-cI R	CTTCTAAGTGACGGCTTTGAGATTCGTCGG
3.30	FRB-cI_F	CGACGAATCTCAAAGCCGTCACTTAGAAGTGAG
3.31	cI-link-FP_R2	CATGCCGGCCGTGCTTACCCATCTCTCCGC
3.32	Eag_FP_F2	GGTAAGCACGGCCGGCATGGTGAGCAAGGGC
3.33	dHinci_F	GGCAGGGATGTTCTCACCTAAACTTAGAACCTTTACC
3.34	dhinCI_R	AGGTGAGAACATCCCTGCC
3.35	xhoFPF	GATCCGAGCTCGAGAATGGTGAGCAAG
3.36	2A_tHetR	GGCGGTACCGGAGCTGCCGCTGCCGGTGCTGCCGGTGCCATGCCCCAGGTGGTGGCGGCCCTCG
3.37	3-2_tHETF	TCCGGTACCGCCTCCTCCGAGGACAACAACATGGCCACCATCAAAGAGTTCATGCGCTTC
3.38	Hin_DFLY_R	TTCCAAGCTTTTACTTGTACAGCTCGTCCATGCCCAGGTGGTGGCGGCCCTCGGAGCGC
3.39	Kpn_CaM_F	TCCGGTACCGCCTCCTCCGAGGACAACAACATGGCCGACCAACTGACAGAAGAGCAGATTGC
3.40	HinCaM_R	CAGCCAAGCTTTTACTTTGCTGTCATCATTTGTAC
3.41	xhoM13_F	GATCCGAGCTCGAGCGGAAAGAGGCGCTGGCAGAAAACAGG
3.42	kpnM13_R	GGCGGTACCGGAGCTGCCGCTGCCGGTGCTGCCGGTGCCATGCCCCAGTGCCCCGGAGCTGGAGATC
3.43	hin_FP_A_F	CAGCCAAGCTTGCCACCATGGTGAGCAAGAGCGAG
3.44	xho_FP_B_R with NES	TCGAGCTCGAGTTAGCTCCCGATATCAAGACCTGCTAATTTCAAGGCTAACTTGTACCGCTCGTCCATGCC
3.45	kpnDEVD_FP_F	GCTCCGGTACCGCCTCCGGCGATGAGGTGGATGGAGCCACCATCAAAGAGTTCATG
3.46	kpnSASG_FP_F	GCTCCGGTACCGCCTCCGGCAGCGCTAGCGGAGGAGCCACCATCAAAGAGTTCATG
3.47	hinM13_F2	CAGCCAAGCTTGCCACCATGAAGAGGCGCTGGCAG
3.48	xhoCaM_NES_R	TCGAGCTCGAGTTAGCTCCCGATATCAAGACCTGCTAATTTCAAGGCTAACTTTGCTGTCATCATTTGTAC
3.49	hinM13_F2	CAGCCAAGCTTGCCACCATGAAGAGGCGCTGGCAG
3.50	xhohetB_NES_R	TCGAGCTCGAGTTAGCTCCCGATATCAAGACCTGCTAATTTCAAGGCTAACTTGTACCGCTCGTCCATGCC
3.51	E160X/H162L_F	GCGACGGCGTGCTGAAGGGCNNKATCTTGCAGGCCCTGAAGC
3.52	E160X_R	CTTCAGCACGCCGTCGCGGG
3.53	E1444/A145/S146_F	GCAGAAGAAGACCTTGGGCTGGNNKNNKGSCACCGAGCGCCTGTACCCCGAAG
3.54	144-146_R	CCAGCCCAAGGTCTTCTTCTGCATTAC
3.55	192/194X_R	GGTGTAGTCCTCGTTGTGGGAGGTGATGTCCAGCTTGGTGTCCACMNNGTAMNNGCCGGGCAGTTGC
3.56	222xf	CAACGAGGACTACACCATCGTGGAACAGTACGGGCGCTCCGAGGGCCGCCACNNKCTGGGCATGGAC

3.57	badr	GGC GCT ACG GCG TTT CAC TTC TGA G
3.58	HinFP_R	TTTAAGCTTTTACTTGTACAGCTCGTCCATG
3.59	BgldFLY	TTCCAGATCTTTACTTGTACAGCTCGTCCATGCC CAGGTGGTGGCGGCCCTCGGAGCGC
3.60	Upst	GTCCACATTGATTATTTGCACGGCGTCACACTTT GCTATGCC
4.1	V105_F	GACGGCGGTCTGGCGACCGTGACTCAGG
4.2	V105_R	CCTGAGTCACGGTCGCCAGACCGCCGTC
4.3	A_71M_F	GTACGGCTCCAAGATGTACGTGAAGCACC
4.4	A_71M_R	GGTGCTTCACGTACATCTTGGAGCCGTAC
4.5	M65C_F	GTCCCCACAGATCTGCTACGGCTCCAAGG
4.6	M65C_R	CCTTGGAGCCGTAGCAGATCTGTGGGGAC
4.7	Upst	GTCCACATTGATTATTTGCACGGCGTCACACTTT GCTATGCC
4.8	badr	GGC GCT ACG GCG TTT CAC TTC TGA G
4.9	MVSKG mamF	TTTAAGCTTGCCACCATGGTGAGCAAGAGCGAG GAGACCATCAAAGAGTTCATGCGC
4.10	PCD_R	GAGTGGCACCTTCCAGGGTCAAGG
4.11	CVIM_F	GAAAAAGAAGTCAAAGACAAAGTGTGTAATTATG TAACTCGAGTCTAGAGGGCCCG
4.12	lip_R	CTTTGTCTTTGACTTCTTTTTCTTCTTTTTACCATC TTTGCTCTTGTACCGCTCGTCCAT
4.13	GhetAMamR	TTTCTCGAGCTAGGTGCTGCCGGTGCCATGCCC CAG
4.14	HinPHF	TTTAAGCTTGCCACCATGGACGACGTGGCCATC GTGAAGGAGGGCTGG
4.15	PHGHetAF	AGCTCCGGTACCGCCTCCTCCATGGCGAGCAAG AGCGAGGAGGTC
4.16	PH_GHET_A R	GGAGGAGGCGGTACCGGAGCTCCCGTCGGCCA CCGTCTGGATG

Appendix B: Analytical restriction digest supplement

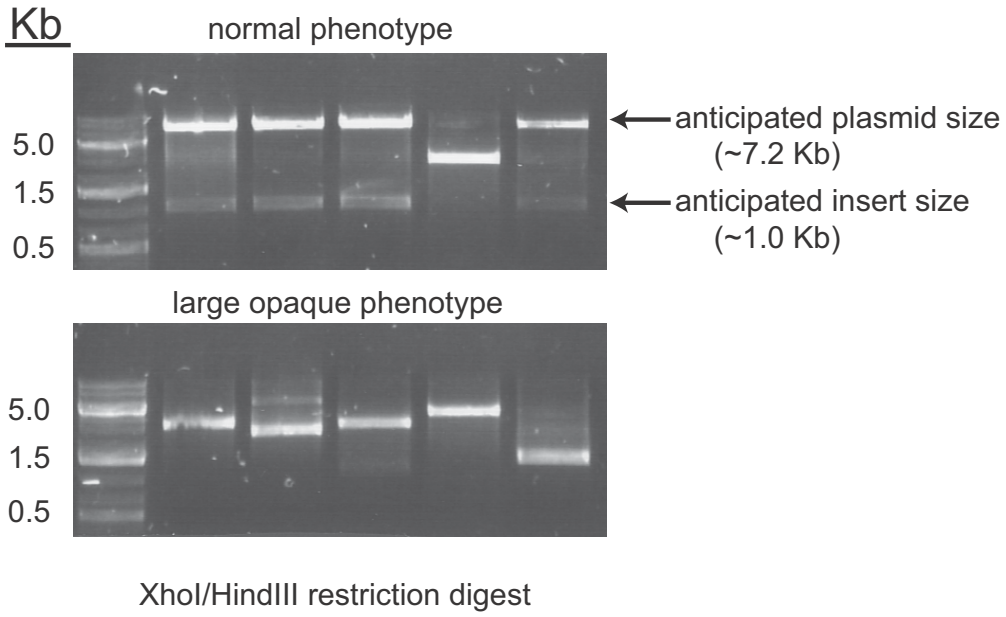


Figure B1. Analytical restriction digests for plasmids obtained from a PINTS library screen. Top panel shows *XhoI*/*HindIII* digests of plasmids obtained from cultures inoculated with normal translucent colonies. Lower panel show *XhoI*/*HindIII* digests of plasmids obtained culture inoculated with large opaque colonies. Digestion with restriction enzymes *XhoI* and *HindIII* should excise an FP-BS-Y29I/Y30G gene fusion of size ~1000 bp and leave a plasmid backbone of ~7.2 Kb. Plasmids derived from colonies exhibiting the normal phenotype exhibit the predicted restriction pattern (upper panel). DNA plasmids obtained from large opaque clones exhibit a restriction pattern consistent with significant DNA rearrangements and truncations (lower panel); no inserts are excised at the predicted sizes and the plasmid backbones are smaller in size than predicted.

Appendix C: Data supplement for Chapter 3

Table C1. Substitutions in new variants described in Chapter 3.

Progenitor protein (rationally modified) or source library	Partner protein for library screening	Resulting protein (or pool)[1] name	Substitutions in resulting protein[2]
dTomato	NA[3]	$A_{0.1}$	H162K, A164R
dTomato with R149X, H162X, Y192X	$A_{0.1}$	$B_{0.1,pool}$	pool of 25 best variants
$A_{0.1}$	NA	$A_{0.2}$	R153E, H162K, A164R
$B_{0.1, pool}$	NA	$B_{0.2, pool}$	$B_{0.1, pool}$ with E100R
$B_{0.2, pool}$ with E160X	$A_{0.2}$	$B_{0.3}$	E100R, R149L, E160H, H162F, Y192G
$A_{0.2}$	NA	$A_{0.3}$	R153E, H162K, A164R, Δ224-226
$B_{0.3}$	NA	$B_{0.4}$	E100R, R149L, E160H, H162F, Y192G, Δ224-226
$B_{0.4}$ with random mutation[4]	$A_{0.3}$	$B_{0.5}$	V7T, K70E, K74R, S86P, E100R, Y120C, R149M, E160H, I161S, H162Y, K178E, Y192G, E215G, Δ224-226
$A_{0.3}$ with random mutation	$B_{0.5}$	$A_{0.4}$	G4S, M18L, F65I, I79V, K83M, N98H, R153E, H162K, A164R, H172L, Y173N, Y194F, Δ224-226
$A_{0.4}$	NA	$A_{0.5}$	G4S, M18L, F65I, I79V, K83M, N98H, S146A, R153E, H162K, A164R, H172L, Y173N, Y194F, D200G, Δ224-226
$A_{0.5}$ with E144X, A145X	NA	$A_{0.6, pool}$	pool of dark variants

$B_{0.5}$ with Y192X, Y194X, H222X, R100E, R153E	$A_{0.6,\ pool}$	$A_{1.0}$	G4S, M18L, F56I, Y79I, K83M, N98H, S111T, M141L, E144D, A145Y, S146A, R153E, I161L, H162L, Q163G, A164R, H172L, Y173N, Y194F, D200G, Δ224-226
		$B_{1.0}$	V7T, K70E, K74R, S86P, Y120C, R149M, R153E, E160H, I161S, H162Y, K178E, Y192D, Y194C, E215G, H222R, Δ224-226

[1]A pool refers to a selection of improved variants that were picked but not sequenced or individually characterized. [2]Substitutions are relative to dTomato. [3]Not applicable. This indicates that the change was made rationally or that the library was screened for dark variants in the absence of a partner protein. [4]Random mutations accumulated during several rounds of error-prone PCR amplification and library screening.

Appendix D: Replica plate supplement

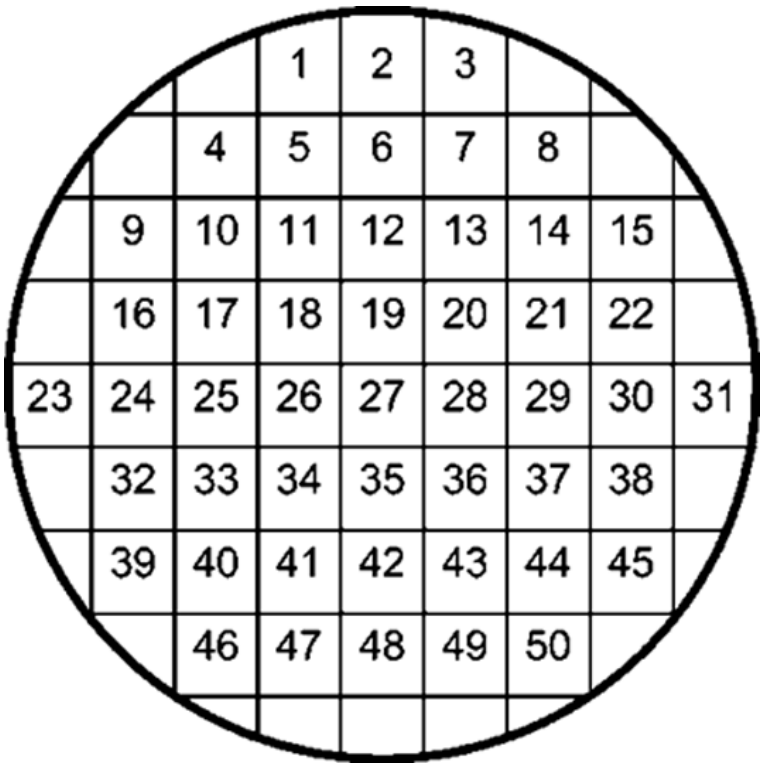


Figure D1. Replica plate grid used for replica plating *E. coli* colonies. Colonies were spotted at vertices to simplify image alignment with image processing macros. Relevant to Chapter 3 (Figure 3.5).

Appendix E: Data supplement for Chapter 4

Table E1. Substitutions in new variants described in Chapter 4.

	Substitution[1]	
	ddFP-A	**ddFP-B**
ddRFP-**$A_1B_{1.1}$**	none	F87L, G89V, C120Y, E153K, Y214F, G215E, M225K[2]
ddGFP	V2A, Q64L, A71M, V107A, T111S, D154E, F177S, G200D	F65I, E70K, M124I, E153K
ddYFP	R17H, M22V, A44T, M66C, A71V, M83I, A146S, L161I, G200D	K15Q, F65I, E70K, K92R, I119V, M124I, E153K, L174P, D116V, D200V

[1] Substitutions listed are relative to ddRFP-**A_1B_1**.

[2] This numbering reflects new numbering specific to ddRFP-A_1B_1, which had its C-terminal parent dTomato residues truncated (Δ224-226) and replaced with the GFP C-terminal sequence, GMDELYK

Appendix F: DdRFP cell surface labeling supplement

Figure F1. Cell surface labeling with ddRFP-$\mathbf{A}_1\mathbf{B}_1$. HeLa cells expressing $\mathbf{A}_1$ alone (left panel), $\mathbf{B}_1$ alone (middle panel), or both $\mathbf{A}_1$ and $\mathbf{B}_1$ (right panel). Genes encoding the ddRFP constituents, $\mathbf{A}_1$ and $\mathbf{B}_1$ were cloned into the mammalian expression plasmid, pDisplay (Invitrogen), and transfected into HeLa cells using standard procedures described in this thesis (e.g. as described in Chapter 2, section 2.4.6). Red fluorescence was imaged 24 hrs following transfection. (Credit to Yidan Ding for images depicted in Figure F1).

References

1. Tsien, R.Y. The green fluorescent protein. *Annu. Rev. Biochem.*, 1998. **67**: p. 509-44.

2. Chudakov, D.M., M.V. Matz, S. Lukyanov, and K.A. Lukyanov. Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol. Rev.*, 2010. **90**(3): p. 1103-63.

3. Lauf, U., P. Lopez, and M.M. Falk. Expression of fluorescently tagged connexins: a novel approach to rescue function of oligomeric DsRed-tagged proteins. *FEBS Lett.*, 2001. **498**(1): p. 11-15.

4. Mizuno, H., A. Sawano, P. Eli, H. Hama, and A. Miyawaki. Red fluorescent protein from Discosoma as a fusion tag and a partner for fluorescence resonance energy transfer. *Biochemistry*, 2001. **40**(8): p. 2502-10.

5. Bulina, M.E., V.V. Verkhusha, D.B. Staroverov, D.M. Chudakov, and K.A. Lukyanov. Hetero-oligomeric tagging diminishes non-specific aggregation of target proteins fused with anthozoa fluorescent proteins. *Biochem. J.*, 2003. **371**: p. 109-114.

6. Qin, H., Q. Shao, D.J. Belliveau, and D.W. Laird. Aggregated DsRed-tagged Cx43 and over-expressed Cx43 are targeted to lysosomes in human breast cancer cells. *Cell Commun. Adhes.*, 2001. **8**(4-6): p. 433-439.

7. Yanushevich, Y.G., D.B. Staroverov, A.P. Savitsky, A.F. Fradkov, N.G. Gurskaya, M.E. Bulina, K.A. Lukyanov, and S.A. Lukyanov. A strategy for the generation of non-aggregating mutants of Anthozoa fluorescent proteins. *FEBS Lett.*, 2002. **511**(1-3): p. 11-14.

8. Gavin, P., R.J. Devenish, and M. Prescott. An approach for reducing unwanted oligomerisation of DsRed fusion proteins. *Biochem. Biophys. Res. Commun.*, 2002. **298**(5): p. 707-713.

9. Campbell, R.E., O. Tour, A.E. Palmer, P.A. Steinbach, G.S. Baird, D.A. Zacharias, and R.Y. Tsien. A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.*, 2002. **99**(12): p. 7877-82.

10. Zhang, J., R.E. Campbell, A.Y. Ting, and R.Y. Tsien. Creating new fluorescent probes for cell biology. *Nat. Rev. Mol. Cell Biol.*, 2002. **3**(12): p. 906-18.

11. Ai, H.W., J.N. Henderson, S.J. Remington, and R.E. Campbell. Directed evolution of a monomeric, bright and photostable version of Clavularia cyan fluorescent protein: structural characterization and applications in fluorescence imaging. *Biochem. J.*, 2006. **400**(3): p. 531-40.

12. Yuan, L., I. Kurek, J. English, and R. Keenan. Laboratory-directed protein evolution. *Microbiol. Mol. Biol. Rev.*, 2005. **69**(3): p. 373-92.

13. Day, R.N. and M.W. Davidson. The fluorescent protein palette: tools for cellular imaging. *Chem. Soc. Rev.*, 2009. **38**(10): p. 2887-921.

14. Carlson, H.J. and R.E. Campbell. Genetically encoded FRET-based biosensors for multiparameter fluorescence imaging. *Curr. Opin. Biotechnol.*, 2009. **20**(1): p. 19-27.

15. Ibraheem, A. and R.E. Campbell. Designs and applications of fluorescent protein-based biosensors. *Curr. Opin. Chem. Biol.*, 2010. **14**(1): p. 30-6.

16. Hu, C.D., A.V. Grinberg, and T.K. Kerppola. Visualization of protein interactions in living cells using bimolecular fluorescence complementation (BiFC) analysis. *Curr. Protoc. Cell Biol.*, 2006. **Chapter 21**: p. Unit 21 3.

17. Kerppola, T.K. Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. *Annu. Rev. Biophys.*, 2008. **37**: p. 465-87.

18. Fields, S. and O. Song. A novel genetic system to detect protein-protein interactions. *Nature*, 1989. **340**(6230): p. 245-6.

19. Lievens, S., I. Lemmens, and J. Tavernier. Mammalian two-hybrids come of age. *Trends Biochem. Sci.*, 2009. **34**(11): p. 579-588.

20. Yamazaki, M., Y. Zhang, H. Watanabe, T. Yokozeki, S. Ohno, K. Kaibuchi, H. Shibata, H. Mukai, Y. Ono, M.A. Frohman, and Y. Kanaho. Interaction of the small G protein RhoA with the C terminus of human phospholipase D1. *J. Biol. Chem.*, 1999. **274**(10): p. 6035-6038.

21. Lalonde, S., A. Sero, R. Pratelli, G. Pilot, J. Chen, M.I. Sardi, S.A. Parsa, D.Y. Kim, B.R. Acharya, E.V. Stein, H.C. Hu, F. Villiers, K. Takeda, Y. Yang, Y.S. Han, R. Schwacke, W. Chiang, N. Kato, D. Loque, S.M. Assmann, J.M. Kwak, J.I. Schroeder, S.Y. Rhee, and W.B. Frommer. A membrane protein/signaling protein interaction network for Arabidopsis version AMPv2. *Front. Physiol.*, 2010. **1**: p. 24.

22. Ear, P.H. and S.W. Michnick. A general life-death selection strategy for dissecting protein functions. *Nat. Methods*, 2009. **6**(11): p. 813-U45.

23. Barreto, K., V.M. Bharathikumar, A. Ricardo, J.F. DeCoteau, Y. Luo, and C.R. Geyer. A genetic screen for isolating "lariat" Peptide inhibitors of protein function. *Chem. Biol.*, 2009. **16**(11): p. 1148-57.

24. Michnick, S.W., P.H. Ear, E.N. Manderson, I. Remy, and E. Stefan. Universal strategies in research and drug discovery based on protein-fragment complementation assays. *Nat. Rev. Drug Discovery*, 2007. **6**(7): p. 569-582.

25. Dmitrova, M., G. Younes-Cauet, P. Oertel-Buchheit, D. Porte, M. Schnarr, and M. Granger-Schnarr. A new LexA-based genetic system for monitoring and analyzing protein heterodimerization in Escherichia coli. *Mol. Gen. Genet.*, 1998. **257**(2): p. 205-212.

26. Wilson, C.G., T.J. Magliery, and L. Regan. Detecting protein-protein interactions with GFP-fragment reassembly. *Nat. Methods*, 2004. **1**(3): p. 255-262.

27. Johnsson, N. and A. Varshavsky. Split ubiquitin as a sensor of protein interactions in vivo. *Proc. Natl. Acad. Sci. U. S. A.*, 1994. **91**(22): p. 10340-4.

28. Paulmurugan, R. and S.S. Gambhir. Monitoring protein-protein interactions using split synthetic renilla luciferase protein-fragment-assisted complementation. *Anal. Chem.*, 2003. **75**(7): p. 1584-1589.

29. Galarneau, A., M. Primeau, L.E. Trudeau, and S.W. Michnick. beta-Lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein-protein interactions. *Nat. Biotechnol.*, 2002. **20**(6): p. 619-622.

30. Pelletier, J.N., F.X. Campbell-Valois, and S.W. Michnick. Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. *Proc. Natl. Acad. Sci. U. S. A.*, 1998. **95**(21): p. 12141-12146.

31. Remy, I., F.X. Campbell-Valois, and S.W. Michnick. Detection of protein-protein interactions using a simple survival protein-fragment complementation assay based on the enzyme dihydrofolate reductase. *Nat. Protoc.*, 2007. **2**(9): p. 2120-5.

32. Villalobos, V., S. Naik, and D. Piwnica-Worms. Current state of imaging protein-protein interactions in vivo with genetically encoded reporters. *Annu. Rev. Biomed. Eng.*, 2007. **9**: p. 321-349.

33. Shekhawat, S.S. and I. Ghosh. Split-protein systems: beyond binary protein-protein interactions. *Curr. Opin. Chem. Biol.*, 2011. **15**(6): p. 789-97.

34. Michnick, S.W., I. Remy, F.X. Campbell-Valois, A. Vallee-Belisle, and J.N. Pelletier. Detection of protein-protein interactions by protein fragment complementation strategies. *Methods Enzymol.*, 2000. **328**: p. 208-30.

35. Madigan, M.T., J.M. Martinko, and J. Parker, *Brock biology of microorganisms*. 9th ed2000, Upper Saddle River, NJ: Prentice Hall. xix, 991, 53 p.

36. Shuman, H.A. and T.J. Silhavy. The art and design of genetic screens: Escherichia coli. *Nat. Rev. Genet.*, 2003. **4**(6): p. 419-31.

37. Lin, H.N. and V.W. Cornish. Screening and selection methods for large-scale analysis of protein function. *Angew. Chem. Int. Ed. Engl.*, 2002. **41**(23): p. 4403-4425.

38. Aharoni, A., A.D. Griffiths, and D.S. Tawfik. High-throughput screens and selections of enzyme-encoding genes. *Curr. Opin. Chem. Biol.*, 2005. **9**(2): p. 210-216.

39. Michnick, S.W., P.H. Ear, C. Landry, M.K. Malleshaiah, and V. Messier. A toolkit of protein-fragment complementation assays for studying and dissecting large-scale and dynamic protein-protein interactions in living cells. *Methods Enzymol.*, 2010. **470**: p. 335-68.

40. Michnick, S.W. Protein fragment complementation strategies for biochemical network mapping. *Curr. Opin. Biotechnol.*, 2003. **14**(6): p. 610-617.

41. Shibasaki, S., K. Sakata, J. Ishii, A. Kondo, and M. Ueda. Development of a yeast protein fragment complementation assay (PCA) system using dihydrofolate reductase (DHFR) with specific additives. *Appl. Microbiol. Biotechnol.*, 2008. **80**(4): p. 735-43.

42. Subramaniam, R., D. Desveaux, C. Spickler, S.W. Michnick, and N. Brisson. Direct visualization of protein interactions in plant cells. *Nat. Biotechnol.*, 2001. **19**(8): p. 769-72.

43. Appleman, J.R., N. Prendergast, T.J. Delcamp, J.H. Freisheim, and R.L. Blakley. Kinetics of the Formation and Isomerization of Methotrexate Complexes of Recombinant Human Dihydrofolate-Reductase. *J. Biol. Chem.*, 1988. **263**(21): p. 10304-10313.

44. Remy, I., G. Ghaddar, and S.W. Michnick. Using the beta-lactamase protein-fragment complementation assay to probe dynamic protein-protein interactions. *Nat. Protoc.*, 2007. **2**(9): p. 2302-6.

45. Remy, I., I.A. Wilson, and S.W. Michnick. Erythropoietin receptor activation by a ligand-induced conformation change. *Science*, 1999. **283**(5404): p. 990-3.

46. Remy, I. and S.W. Michnick. Clonal selection and in vivo quantitation of protein interactions with protein-fragment complementation assays. *Proc. Natl. Acad. Sci. U. S. A.*, 1999. **96**(10): p. 5394-9.

47. Wehrman, T., B. Kleaveland, J.H. Her, R.F. Balint, and H.M. Blau. Protein-protein interactions monitored in mammalian cells via complementation of beta -lactamase enzyme fragments. *Proc. Natl. Acad. Sci. U. S. A.*, 2002. **99**(6): p. 3469-74.

48. Zlokarnik, G., P.A. Negulescu, T.E. Knapp, L. Mere, N. Burres, L. Feng, M. Whitney, K. Roemer, and R.Y. Tsien. Quantitation of transcription and

clonal selection of single living cells with beta-lactamase as reporter. *Science*, 1998. **279**(5347): p. 84-8.

49. Matz, M.V., A.F. Fradkov, Y.A. Labas, A.P. Savitsky, A.G. Zaraisky, M.L. Markelov, and S.A. Lukyanov. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat. Biotechnol.*, 1999. **17**(10): p. 969-73.

50. Heim, R., D.C. Prasher, and R.Y. Tsien. Wavelength Mutations and Posttranslational Autoxidation of Green Fluorescent Protein. *Proc. Natl. Acad. Sci. U. S. A.*, 1994. **91**(26): p. 12501-12504.

51. Craggs, T.D. Green fluorescent protein: structure, folding and chromophore maturation. *Chem. Soc. Rev.*, 2009. **38**(10): p. 2865-75.

52. Delagrave, S., R.E. Hawtin, C.M. Silva, M.M. Yang, and D.C. Youvan. Red-shifted excitation mutants of the green fluorescent protein. *Biotechnology (N Y)*, 1995. **13**(2): p. 151-4.

53. Barondeau, D.P., C.D. Putnam, C.J. Kassmann, J.A. Tainer, and E.D. Getzoff. Mechanism and energetics of green fluorescent protein chromophore synthesis revealed by trapped intermediate structures. *Proc. Natl. Acad. Sci. U. S. A.*, 2003. **100**(21): p. 12111-6.

54. Cubitt, A.B., R. Heim, S.R. Adams, A.E. Boyd, L.A. Gross, and R.Y. Tsien. Understanding, improving and using green fluorescent proteins. *Trends Biochem. Sci.*, 1995. **20**(11): p. 448-55.

55. Barondeau, D.P., J.A. Tainer, and E.D. Getzoff. Structural evidence for an enolate intermediate in GFP fluorophore biosynthesis. *J. Am. Chem. Soc.*, 2006. **128**(10): p. 3166-8.

56. Zhang, L., H.N. Patel, J.W. Lappe, and R.M. Wachter. Reaction progress of chromophore biogenesis in green fluorescent protein. *J. Am. Chem. Soc.*, 2006. **128**(14): p. 4766-72.

57. Rosenow, M.A., H.A. Huffman, M.E. Phail, and R.M. Wachter. The crystal structure of the Y66L variant of green fluorescent protein supports a cyclization-oxidation-dehydration mechanism for chromophore maturation. *Biochemistry*, 2004. **43**(15): p. 4464-72.

58. Rosenow, M.A., H.N. Patel, and R.M. Wachter. Oxidative chemistry in the GFP active site leads to covalent cross-linking of a modified leucine side chain with a histidine imidazole: Implications for the mechanism of chromophore formation. *Biochemistry*, 2005. **44**(23): p. 8303-8311.

59. Pouwels, L.J., L. Zhang, N.H. Chan, P.C. Dorrestein, and R.M. Wachter. Kinetic isotope effect studies on the de novo rate of chromophore formation in fast- and slow-maturing GFP variants. *Biochemistry*, 2008. **47**(38): p. 10111-22.

60. Strack, R.L., D.E. Strongin, L. Mets, B.S. Glick, and R.J. Keenan. Chromophore formation in DsRed occurs by a branched pathway. *J. Am. Chem. Soc.*, 2010. **132**(24): p. 8496-505.

61. Barondeau, D.P., C.J. Kassmann, J.A. Tainer, and E.D. Getzoff. Understanding GFP posttranslational chemistry: structures of designed variants that achieve backbone fragmentation, hydrolysis, and decarboxylation. *J. Am. Chem. Soc.*, 2006. **128**(14): p. 4685-93.

62. Yang, F., L.G. Moss, and G.N. Phillips, Jr. The molecular structure of green fluorescent protein. *Nat. Biotechnol.*, 1996. **14**(10): p. 1246-51.

63. Ormo, M., A.B. Cubitt, K. Kallio, L.A. Gross, R.Y. Tsien, and S.J. Remington. Crystal structure of the Aequorea victoria green fluorescent protein. *Science*, 1996. **273**(5280): p. 1392-5.

64. Wall, M.A., M. Socolich, and R. Ranganathan. The structural basis for red fluorescence in the tetrameric GFP homolog DsRed. *Nat. Struct. Biol.*, 2000. **7**(12): p. 1133-8.

65. Lukyanov, K.A., D.M. Chudakov, A.F. Fradkov, Y.A. Labas, M.V. Matz, and S. Lukyanov. Discovery and properties of GFP-like proteins from nonbioluminescent anthozoa. *Methods Biochem. Anal.*, 2006. **47**: p. 121-38.

66. Schrodinger, LLC, *The PyMOL Molecular Graphics System, Version 1.3r1*, 2010.

67. Shagin, D.A., E.V. Barsova, Y.G. Yanushevich, A.F. Fradkov, K.A. Lukyanov, Y.A. Labas, T.N. Semenova, J.A. Ugalde, A. Meyers, J.M. Nunez, E.A. Widder, S.A. Lukyanov, and M.V. Matz. GFP-like proteins as ubiquitous metazoan superfamily: evolution of functional features and structural complexity. *Mol. Biol. Evol.*, 2004. **21**(5): p. 841-50.

68. Zeng, W., H.E. Seward, A. Malnasi-Csizmadia, S. Wakelin, R.J. Woolley, G.S. Cheema, J. Basran, T.R. Patel, A.J. Rowe, and C.R. Bagshaw. Resonance energy transfer between green fluorescent protein variants: Complexities revealed with myosin fusion proteins. *Biochemistry*, 2006. **45**(35): p. 10482-10491.

69. Zacharias, D.A., J.D. Violin, A.C. Newton, and R.Y. Tsien. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science*, 2002. **296**(5569): p. 913-6.

70. Phillips, G.N., Jr. Structure and dynamics of green fluorescent protein. *Curr. Opin. Struct. Biol.*, 1997. **7**(6): p. 821-7.

71. Baird, G.S., D.A. Zacharias, and R.Y. Tsien. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proc. Natl. Acad. Sci. U. S. A.*, 2000. **97**(22): p. 11984-9.

72. Vrzheshch, P.V., N.A. Akovbian, S.D. Varfolomeyev, and V.V. Verkhusha. Denaturation and partial renaturation of a tightly tetramerized DsRed protein under mildly acidic conditions. *FEBS Lett.*, 2000. **487**(2): p. 203-8.

73. Gross, L.A., G.S. Baird, R.C. Hoffman, K.K. Baldridge, and R.Y. Tsien. The structure of the chromophore within DsRed, a red fluorescent protein from coral. *Proc. Natl. Acad. Sci. U. S. A.*, 2000. **97**(22): p. 11990-5.

74. Yarbrough, D., R.M. Wachter, K. Kallio, M.V. Matz, and S.J. Remington. Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-A resolution. *Proc. Natl. Acad. Sci. U. S. A.*, 2001. **98**(2): p. 462-7.

75. Petersen, J., P.G. Wilmann, T. Beddoe, A.J. Oakley, R.J. Devenish, M. Prescott, and J. Rossjohn. The 2.0-A crystal structure of eqFP611, a far red fluorescent protein from the sea anemone Entacmaea quadricolor. *J. Biol. Chem.*, 2003. **278**(45): p. 44626-31.

76. Pletneva, N.V., S.V. Pletnev, D.M. Chudakov, T.V. Tikhonova, V.O. Popov, V.I. Martynov, A. Wlodawer, Z. Dauter, and V.Z. Pletnev. Three-dimensional structure of yellow fluorescent protein zYFP538 from Zoanthus sp at the resolution 1.8 angstrom. *Russ. J. Bioorg. Chem.*, 2007. **33**(4): p. 390-398.

77. Pletneva, N.V., V.Z. Pletnev, Shemiakina, II, D.M. Chudakov, I. Artemyev, A. Wlodawer, Z. Dauter, and S. Pletnev. Crystallographic study of red fluorescent protein eqFP578 and its far-red variant Katushka reveals opposite pH-induced isomerization of chromophore. *Protein Sci.*, 2011. **20**(7): p. 1265-74.

78. Verkhusha, V.V. and K.A. Lukyanov. The molecular properties and applications of Anthozoa fluorescent proteins and chromoproteins. *Nat. Biotechnol.*, 2004. **22**(3): p. 289-96.

79. Sacchetti, A., V. Subramaniam, T.M. Jovin, and S. Alberti. Oligomerization of DsRed is required for the generation of a functional red fluorescent chromophore. *FEBS Lett.*, 2002. **525**(1-3): p. 13-19.

80. Remington, S.J., R.M. Wachter, D.K. Yarbrough, B. Branchaud, D.C. Anderson, K. Kallio, and K.A. Lukyanov. zFP538, a yellow-fluorescent protein from Zoanthus, contains a novel three-ring chromophore. *Biochemistry*, 2005. **44**(1): p. 202-12.

81. Kikuchi, A., E. Fukumura, S. Karasawa, H. Mizuno, A. Miyawaki, and Y. Shiro. Structural Characterization of a Thiazoline-Containing Chromophore in an Orange Fluorescent Protein, Monomeric Kusabira Orange. *Biochemistry*, 2008. **47**(44): p. 11573-11580.

82. Ai, H.W., N.C. Shaner, Z. Cheng, R.Y. Tsien, and R.E. Campbell. Exploration of new chromophore structures leads to the identification of improved blue fluorescent proteins. *Biochemistry*, 2007. **46**(20): p. 5904-10.

83. Tomosugi, W., T. Matsuda, T. Tani, T. Nemoto, I. Kotera, K. Saito, K. Horikawa, and T. Nagai. An ultramarine fluorescent protein with increased photostability and pH insensitivity. *Nat. Methods*, 2009. **6**(5): p. 351-353.

84. Subach, O.M., I.S. Gundorov, M. Yoshimura, F.V. Subach, J. Zhang, D. Gruenwald, E.A. Souslova, D.M. Chudakov, and V.V. Verkhusha. Conversion of red fluorescent protein into a bright blue probe. *Chem. Biol.*, 2008. **15**(10): p. 1116-24.

85. Mizuno, H., T.K. Mal, K.I. Tong, R. Ando, T. Furuta, M. Ikura, and A. Miyawakil. Photo-induced peptide cleavage in the green-to-red conversion of a fluorescent protein. *Mol. Cell*, 2003. **12**(4): p. 1051-1058.

86. Hayashi, I., H. Mizuno, K.I. Tong, T. Furuta, F. Tanaka, M. Yoshimura, A. Miyawaki, and M. Ikura. Crystallographic evidence for water-assisted photo-induced peptide cleavage in the stony coral fluorescent protein kaede. *J. Mol. Biol.*, 2007. **372**(4): p. 918-926.

87. Drobizhev, M., S. Tillo, N.S. Makarov, T.E. Hughes, and A. Rebane. Color Hues in Red Fluorescent Proteins Are Due to Internal Quadratic Stark Effect. *J. Phys. Chem. B*, 2009. **113**(39): p. 12860-12864.

88. Shu, X., N.C. Shaner, C.A. Yarbrough, R.Y. Tsien, and S.J. Remington. Novel chromophores and buried charges control color in mFruits. *Biochemistry*, 2006. **45**(32): p. 9639-47.

89. Wilmann, P.G., J. Petersen, A. Pettikiriarachchi, A.M. Buckle, S.C. Smith, S. Olsen, M.A. Perugini, R.J. Devenish, M. Prescott, and J. Rossjohn. The 2.1 angstrom crystal structure of the far-red fluorescent protein HcRed: Inherent conformational flexibility of the chromophore. *J. Mol. Biol.*, 2005. **349**(1): p. 223-237.

90. Prescott, M., M. Ling, T. Beddoe, A.J. Oakley, S. Dove, O. Hoegh-Guldberg, R.J. Devenish, and J. Rossjohn. The 2.2 A crystal structure of a pocilloporin pigment reveals a nonplanar chromophore conformation. *Structure*, 2003. **11**(3): p. 275-84.

91. Henderson, J.N., H.W. Ai, R.E. Campbell, and S.J. Remington. Structural basis for reversible photobleaching of a green fluorescent protein homologue. *Proc. Natl. Acad. Sci. U. S. A.*, 2007. **104**(16): p. 6672-7.

92. Inouye, S. and F.I. Tsuji. Aequorea Green Fluorescent Protein - Expression of the Gene and Fluorescence Characteristics of the Recombinant Protein. *FEBS Lett.*, 1994. **341**(2-3): p. 277-280.

93. Wu, C.F., H.J. Cha, G. Rao, J.J. Valdes, and W.E. Bentley. A green fluorescent protein fusion strategy for monitoring the expression, cellular location, and separation of biologically active organophosphorus hydrolase. *Appl. Microbiol. Biotechnol.*, 2000. **54**(1): p. 78-83.

94. Bevis, B.J. and B.S. Glick. Rapidly maturing variants of the Discosoma red fluorescent protein (DsRed) (vol 20, pg 83, 2002). *Nat. Biotechnol.*, 2002. **20**(11): p. 1159-1159.

95. Crameri, A., E.A. Whitehorn, E. Tate, and W.P. Stemmer. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat. Biotechnol.*, 1996. **14**(3): p. 315-9.

96. Cormack, B.P., R.H. Valdivia, and S. Falkow. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*, 1996. **173**(1 Spec No): p. 33-8.

97. Heim, R., A.B. Cubitt, and R.Y. Tsien. Improved green fluorescence. *Nature*, 1995. **373**(6516): p. 663-4.

98. Shaner, N.C., R.E. Campbell, P.A. Steinbach, B.N. Giepmans, A.E. Palmer, and R.Y. Tsien. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat. Biotechnol.*, 2004. **22**(12): p. 1567-72.

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

100. Karasawa, S., T. Araki, M. Yamamoto-Hino, and A. Miyawaki. A Green-emitting fluorescent protein from Galaxeidae coral and its monomeric version for use in fluorescent labeling. *J. Biol. Chem.*, 2003. **278**(36): p. 34167-34171.

101. Fradkov, A.F., V.V. Verkhusha, D.B. Staroverov, M.E. Bulina, Y.G. Yanushevich, V.I. Martynov, S. Lukyanov, and K.A. Lukyanov. Far-red fluorescent tag for protein labelling. *Biochem. J.*, 2002. **368**: p. 17-21.

102. Zaslaver, A., A. Bren, M. Ronen, S. Itzkovitz, I. Kikoin, S. Shavit, W. Liebermeister, M.G. Surette, and U. Alon. A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. *Nat. Methods*, 2006. **3**(8): p. 623-628.

103. Southward, C.M. and M.G. Surette. The dynamic microbe: green fluorescent protein brings bacteria to light. *Molec. Microbiol.*, 2002. **45**(5): p. 1191-1196.

104. Shioda, T., S. Andriole, T. Yahata, and K.J. Isselbacher. A green fluorescent protein-reporter mammalian two-hybrid system with extrachromosomal maintenance of a prey expression plasmid: Application to interaction screening. *Proc. Natl. Acad. Sci. U. S. A.*, 2000. **97**(10): p. 5220-5224.

105. Lim, D.V., J.M. Simpson, E.A. Kearns, and M.F. Kramer. Current and developing technologies for monitoring agents of bioterrorism and biowarfare. *Clin. Microbiol. Rev.*, 2005. **18**(4): p. 583-607.

106. Xia, Z. and J. Rao. Biosensing and imaging based on bioluminescence resonance energy transfer. *Curr. Opin. Biotechnol.*, 2009. **20**(1): p. 37-44.

107. Campbell, R.E. Fluorescent-protein-based biosensors: modulation of energy transfer as a design principle. *Anal. Chem.*, 2009. **81**(15): p. 5972-9.

108. Lakowicz, J.R., Principles of fluorescence spectroscopy. 3rd ed. 2006, New York: Springer. xxvi, 954.

109. Jones, K., F. Hibbert, and M. Keenan. Glowing jellyfish, luminescence and a molecule called coelenterazine. *Trends Biotechnol.*, 1999. **17**(12): p. 477-81.

110. Pfleger, K.D.G. and K.A. Eidne. Illuminating insights into protein-protein interactions using bioluminescence resonance energy transfer (BRET). *Nat. Methods*, 2006. **3**(3): p. 165-174.

111. Pfleger, K.D.G. and K.A. Eidne. Monitoring the formation of dynamic G-protein-coupled receptor-protein complexes in living cells. *Biochem. J.*, 2005. **385**: p. 625-637.

112. Hollins, B., S. Kuravi, G.J. Digby, and N.A. Lambert. The c-terminus of GRK3 indicates rapid dissociation of G protein heterotrimers. *Cell. Signalling*, 2009. **21**(6): p. 1015-1021.

113. Germain-Desprez, D., M. Bazinet, M. Bouvier, and M. Aubry. Oligomerization of transcriptional intermediary factor 1 regulators and interaction with ZNF74 nuclear matrix protein revealed by bioluminescence resonance energy transfer in living cells. *J. Biol. Chem.*, 2003. **278**(25): p. 22367-22373.

114. Ai, H.W., K.L. Hazelwood, M.W. Davidson, and R.E. Campbell. Fluorescent protein FRET pairs for ratiometric imaging of dual biosensors. *Nat. Methods*, 2008. **5**(5): p. 401-3.

115. Ouyang, M., H. Huang, N.C. Shaner, A.G. Remacle, S.A. Shiryaev, A.Y. Strongin, R.Y. Tsien, and Y. Wang. Simultaneous visualization of protumorigenic Src and MT1-MMP activities with fluorescence resonance energy transfer. *Cancer Res.*, 2010. **70**(6): p. 2204-12.

116. Miyawaki, A., J. Llopis, R. Heim, J.M. McCaffery, J.A. Adams, M. Ikura, and R.Y. Tsien. Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. *Nature*, 1997. **388**(6645): p. 882-7.

117. Miyawaki, A., O. Griesbeck, R. Heim, and R.Y. Tsien. Dynamic and quantitative Ca2+ measurements using improved cameleons. *Proc. Natl. Acad. Sci. U. S. A.*, 1999. **96**(5): p. 2135-40.

118. Palmer, A.E., M. Giacomello, T. Kortemme, S.A. Hires, V. Lev-Ram, D. Baker, and R.Y. Tsien. Ca2+ indicators based on computationally redesigned calmodulin-peptide pairs. *Chem. Biol.*, 2006. **13**(5): p. 521-30.

119. Zhang, J. and M.D. Allen. FRET-based biosensors for protein kinases: illuminating the kinome. *Mol. Biosyst.*, 2007. **3**(11): p. 759-765.

120. Aoki, K. and M. Matsuda. Visualization of small GTPase activity with fluorescence resonance energy transfer-based biosensors. *Nat. Protoc.*, 2009. **4**(11): p. 1623-1631.

121. Heim, R. and R.Y. Tsien. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Biol.*, 1996. **6**(2): p. 178-82.

122. Hwang, Y.C., J.J.H. Chu, P.L. Yang, W. Chen, and M.V. Yates. Rapid identification of inhibitors that interfere with poliovirus replication using a cell-based assay. *Antiviral Res.*, 2008. **77**(3): p. 232-236.

123. Dwyer, M.A. and H.W. Hellinga. Periplasmic binding proteins: a versatile superfamily for protein engineering. *Curr. Opin. Struct. Biol.*, 2004. **14**(4): p. 495-504.

124. Quiocho, F.A. and P.S. Ledvina. Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Molec. Microbiol.*, 1996. **20**(1): p. 17-25.

125. Fehr, M., W.B. Frommer, and S. Lalonde. Visualization of maltose uptake in living yeast cells by fluorescent nanosensors. *Proc. Natl. Acad. Sci. U. S. A.*, 2002. **99**(15): p. 9846-9851.

126. Fehr, M., S. Lalonde, D.W. Ehrhardt, and W.B. Frommer. Live Imaging of glucose homeostasis in nuclei of COS-7 cells. *J. Fluoresc.*, 2004. **14**(5): p. 603-609.

127. Deuschle, K., S. Okumoto, M. Fehr, L.L. Looger, L. Kozhukh, and W.B. Frommer. Construction and optimization of a family of genetically encoded metabolite sensors by semirational protein engineering. *Protein Sci.*, 2005. **14**(9): p. 2304-2314.

128. Ewald, J.C., S. Reich, S. Baumann, W.B. Frommer, and N. Zamboni. Engineering Genetically Encoded Nanosensors for Real-Time In Vivo Measurements of Citrate Concentrations. *Plos One*, 2011. **6**(12).

129. Esposito, A., M. Gralle, M.A.C. Dani, D. Lange, and F.S. Wouters. pHlameleons: A Family of FRET-Based Protein Sensors for Quantitative pH Imaging. *Biochemistry*, 2008. **47**(49): p. 13115-13126.

130. Stabley, D.R., C. Jurchenko, S.S. Marshall, and K.S. Salaita. Visualizing mechanical tension across membrane receptors with a fluorescent sensor. *Nat. Methods*, 2012. **9**(1): p. 64-U172.

131. Grashoff, C., B.D. Hoffman, M.D. Brenner, R.B. Zhou, M. Parsons, M.T. Yang, M.A. McLean, S.G. Sligar, C.S. Chen, T. Ha, and M.A. Schwartz. Measuring mechanical tension across vinculin reveals regulation of focal adhesion dynamics. *Nature*, 2010. **466**(7303): p. 263-U143.

132. Ibraheem, A., H. Yap, Y. Ding, and R.E. Campbell. A bacteria colony-based screen for optimal linker combinations in genetically encoded biosensors. *BMC Biotechnol.*, 2011. **11**: p. 105.

133. Ding, Y.D., H.W. Ai, H. Hoi, and R.E. Campbell. Forster Resonance Energy Transfer-Based Biosensors for Multiparameter Ratiometric Imaging of Ca(2+) Dynamics and Caspase-3 Activity in Single Cells. *Anal. Chem.*, 2011. **83**(24): p. 9687-9693.

134. Nguyen, A.W. and P.S. Daugherty. Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nat. Biotechnol.*, 2005. **23**(3): p. 355-60.

135. Vinkenborg, J.L., T.H. Evers, S.W.A. Reulen, E.W. Meijer, and M. Merkx. Enhanced sensitivity of FRET-based protease sensors by redesign of the GFP dimerization interface. *Chembiochem*, 2007. **8**(10): p. 1119-1121.

136. Piljic, A. and C. Schultz. Simultaneous recording of multiple cellular events by FRET. *ACS Chem. Biol.*, 2008. **3**(3): p. 156-60.

137. Ghosh, I., A.D. Hamilton, and L. Regan. Antiparallel leucine zipper-directed protein reassembly: Application to the green fluorescent protein. *J. Am. Chem. Soc.*, 2000. **122**(23): p. 5658-5659.

138. Kerppola, T.K. Visualization of molecular interactions by fluorescence complementation. *Nat. Rev. Mol. Cell Biol.*, 2006. **7**(6): p. 449-56.

139. Morell, M., A. Espargaro, F.X. Aviles, and S. Ventura. Study and selection of in vivo protein interactions by coupling bimolecular fluorescence complementation and flow cytometry. *Nat. Protoc.*, 2008. **3**(1): p. 22-33.

140. Kojima, T., S. Karasawa, A. Miyawaki, T. Tsumuraya, and I. Fujii. Novel screening system for protein-protein interactions by bimolecular fluorescence complementation in Saccharomyces cerevisiae. *J. Biosci. Bioeng.*, 2011. **111**(4): p. 397-401.

141. Remy, I. and S.W. Michnick. A cDNA library functional screening strategy based on fluorescent protein complementation assays to identify novel components of signaling pathways. *Methods*, 2004. **32**(4): p. 381-8.

142. Ding, Z., J. Liang, Y. Lu, Q. Yu, Z. Songyang, S.Y. Lin, and G.B. Mills. A retrovirus-based protein complementation assay screen reveals functional

AKT1-binding partners. *Proc. Natl. Acad. Sci. U. S. A.*, 2006. **103**(41): p. 15014-9.

143. Remy, I. and S.W. Michnick. Regulation of apoptosis by the Ft1 protein, a new modulator of protein kinase B/Akt. *Mol. Cell. Biol.*, 2004. **24**(4): p. 1493-504.

144. Shyu, Y.J., H. Liu, X. Deng, and C.D. Hu. Identification of new fluorescent protein fragments for bimolecular fluorescence complementation analysis under physiological conditions. *Biotechniques*, 2006. **40**(1): p. 61-6.

145. Ohashi, K., T. Kiuchi, K. Shoji, K. Sampei, and K. Mizuno. Visualization of cofilin-actin and Ras-Raf interactions by bimolecular fluorescence complementation assays using a new pair of split Venus fragments. *Biotechniques*, 2012. **52**(1): p. 45-50.

146. Jach, G., M. Pesch, K. Richter, S. Frings, and J.F. Uhrig. An improved mRFP1 adds red to bimolecular fluorescence complementation. *Nat. Methods*, 2006. **3**(8): p. 597-600.

147. Fan, J.Y., Z.Q. Cui, H.P. Wei, Z.P. Zhang, Y.F. Zhou, Y.P. Wang, and X.E. Zhang. Split mCherry as a new red bimolecular fluorescence complementation system for visualizing protein-protein interactions in living cells. *Biochem. Biophys. Res. Commun.*, 2008. **367**(1): p. 47-53.

148. Chu, J., Z. Zhang, Y. Zheng, J. Yang, L. Qin, J. Lu, Z.L. Huang, S. Zeng, and Q. Luo. A novel far-red bimolecular fluorescence complementation system that allows for efficient visualization of protein interactions under physiological conditions. *Biosens. Bioelectron.*, 2009. **25**(1): p. 234-9.

149. Grinberg, A.V., C.D. Hu, and T.K. Kerppola. Visualization of Myc/Max/Mad family dimers and the competition for dimerization in living cells. *Mol. Cell. Biol.*, 2004. **24**(10): p. 4294-308.

150. Hu, C.D. and T.K. Kerppola. Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat. Biotechnol.*, 2003. **21**(5): p. 539-45.

151. Frommer, W.B., M.W. Davidson, and R.E. Campbell. Genetically encoded biosensors based on engineered fluorescent proteins. *Chem. Soc. Rev.*, 2009. **38**(10): p. 2833-41.

152. Zhao, Y., S. Araki, J. Wu, T. Teramoto, Y.F. Chang, M. Nakano, A.S. Abdelfattah, M. Fujiwara, T. Ishihara, T. Nagai, and R.E. Campbell. An expanded palette of genetically encoded Ca(2) indicators. *Science*, 2011. **333**(6051): p. 1888-91.

153. Miesenbock, G., D.A. De Angelis, and J.E. Rothman. Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature*, 1998. **394**(6689): p. 192-5.

154. Wachter, R.M., D. Yarbrough, K. Kallio, and S.J. Remington. Crystallographic and energetic analysis of binding of selected anions to the yellow variants of green fluorescent protein. *J. Mol. Biol.*, 2000. **301**(1): p. 157-71.

155. Jayaraman, S., P. Haggie, R.M. Wachter, S.J. Remington, and A.S. Verkman. Mechanism and cellular applications of a green fluorescent protein-based halide sensor. *J. Biol. Chem.*, 2000. **275**(9): p. 6047-50.

156. Hanson, G.T., R. Aggeler, D. Oglesbee, M. Cannon, R.A. Capaldi, R.Y. Tsien, and S.J. Remington. Investigating mitochondrial redox potential with redox-sensitive green fluorescent protein indicators. *J. Biol. Chem.*, 2004. **279**(13): p. 13044-53.

157. Dooley, C.T., T.M. Dore, G.T. Hanson, W.C. Jackson, S.J. Remington, and R.Y. Tsien. Imaging dynamic redox changes in mammalian cells with green fluorescent protein indicators. *J. Biol. Chem.*, 2004. **279**(21): p. 22284-93.

158. Ostergaard, H., A. Henriksen, F.G. Hansen, and J.R. Winther. Shedding light on disulfide bond formation: engineering a redox switch in green fluorescent protein. *EMBO J.*, 2001. **20**(21): p. 5853-5862.

159. Akerboom, J., J.D. Rivera, M.M. Guilbe, E.C. Malave, H.H. Hernandez, L. Tian, S.A. Hires, J.S. Marvin, L.L. Looger, and E.R. Schreiter. Crystal structures of the GCaMP calcium sensor reveal the mechanism of fluorescence signal change and aid rational design. *J. Biol. Chem.*, 2009. **284**(10): p. 6455-64.

160. Baird, G.S., D.A. Zacharias, and R.Y. Tsien. Circular permutation and receptor insertion within green fluorescent proteins. *Proc. Natl. Acad. Sci. U. S. A.*, 1999. **96**(20): p. 11241-6.

161. Nagai, T., A. Sawano, E.S. Park, and A. Miyawaki. Circularly permuted green fluorescent proteins engineered to sense Ca2+. *Proc. Natl. Acad. Sci. U. S. A.*, 2001. **98**(6): p. 3197-202.

162. Topell, S. and R. Glockshuber. Circular permutation of the green fluorescent protein. *Methods Mol. Biol.*, 2002. **183**: p. 31-48.

163. Park, S.J. and J.R. Cochran, Protein engineering and design. 2010, Boca Raton: CRC Press. xiv, 402.

164. Bloom, J.D. and F.H. Arnold. In the light of directed evolution: Pathways of adaptive protein evolution. *Proc. Natl. Acad. Sci. U. S. A.*, 2009. **106**: p. 9995-10000.

165. Campbell, J.H., J.A. Lengyel, and J. Langridge. Evolution of a second gene for beta-galactosidase in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.*, 1973. **70**(6): p. 1841-5.

166. Hutchison, C.A., 3rd, S. Phillips, M.H. Edgell, S. Gillam, P. Jahnke, and M. Smith. Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.*, 1978. **253**(18): p. 6551-60.

167. Winter, G., A.R. Fersht, A.J. Wilkinson, M. Zoller, and M. Smith. Redesigning enzyme structure by site-directed mutagenesis: tyrosyl tRNA synthetase and ATP binding. *Nature*, 1982. **299**(5885): p. 756-8.

168. Sigal, I.S., B.G. Harwood, and R. Arentzen. Thiol-beta-lactamase: replacement of the active-site serine of RTEM beta-lactamase by a cysteine residue. *Proc. Natl. Acad. Sci. U. S. A.*, 1982. **79**(23): p. 7157-60.

169. Estell, D.A., T.P. Graycar, and J.A. Wells. Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation. *J. Biol. Chem.*, 1985. **260**(11): p. 6518-21.

170. von der Osten, C., S. Branner, S. Hastrup, L. Hedegaard, M.D. Rasmussen, H. Bisgard-Frantzen, S. Carlsen, and J.M. Mikkelsen. Protein engineering of subtilisins to improve stability in detergent formulations. *J. Biotechnol.*, 1993. **28**(1): p. 55-68.

171. Hibbert, E.G. and P.A. Dalby. Directed evolution strategies for improved enzymatic performance. *Microb. Cell Fact.*, 2005. **4**: p. 29.

172. Beckman, R.A., A.S. Mildvan, and L.A. Loeb. On the Fidelity of DNA-Replication - Manganese Mutagenesis In vitro. *Biochemistry*, 1985. **24**(21): p. 5810-5817.

173. Cirino, P.C., K.M. Mayer, and D. Umeno. Generating mutant libraries using error-prone PCR. *Methods Mol. Biol.*, 2003. **231**: p. 3-9.

174. Wang, L., W.C. Jackson, P.A. Steinbach, and R.Y. Tsien. Evolution of new nonantibody proteins via iterative somatic hypermutation. *Proc. Natl. Acad. Sci. U. S. A.*, 2004. **101**(48): p. 16745-9.

175. Wang, C.L., D.C. Yang, and M. Wabl. Directed molecular evolution by somatic hypermutation. *Protein Eng. Des. Sel.*, 2004. **17**(9): p. 659-64.

176. Drummond, D.A., B.L. Iverson, G. Georgiou, and F.H. Arnold. Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins. *J. Mol. Biol.*, 2005. **350**(4): p. 806-816.

177. Chapman, S., C. Faulkner, E. Kaiserli, C. Garcia-Mata, E.I. Savenkov, A.G. Roberts, K.J. Oparka, and J.M. Christie. The photoreversible fluorescent protein iLOV outperforms GFP as a reporter of plant virus infection. *Proc. Natl. Acad. Sci. U. S. A.*, 2008. **105**(50): p. 20038-20043.

178. Stemmer, W.P.C. Rapid Evolution of a Protein in-Vitro by DNA Shuffling. *Nature*, 1994. **370**(6488): p. 389-391.

179. Crameri, A., S.A. Raillard, E. Bermudez, and W.P. Stemmer. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature*, 1998. **391**(6664): p. 288-91.

180. Zhao, H.M., L. Giver, Z.X. Shao, J.A. Affholter, and F.H. Arnold. Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.*, 1998. **16**(3): p. 258-261.

181. Stemmer, W.P.C. DNA Shuffling by Random Fragmentation and Reassembly - in-Vitro Recombination for Molecular Evolution. *Proc. Natl. Acad. Sci. U. S. A.*, 1994. **91**(22): p. 10747-10751.

182. Sen, S., V. Venkata Dasu, and B. Mandal. Developments in directed evolution for improving enzyme functions. *Appl. Biochem. Biotechnol.*, 2007. **143**(3): p. 212-23.

183. Aharoni, A., A.D. Griffiths, and D.S. Tawfik. High-throughput screens and selections of enzyme-encoding genes. *Curr Opin. Chem. Biol.*, 2005. **9**(2): p. 210-216.

184. Turner, N.J. Directed evolution drives the next generation of biocatalysts. *Nat. Chem. Biol.*, 2009. **5**(8): p. 567-573.

185. Schmidt-Dannert, C. and F.H. Arnold. Directed evolution of industrial enzymes. *Trends Biotechnol.*, 1999. **17**(4): p. 135-6.

186. Chica, R.A., N. Doucet, and J.N. Pelletier. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. *Curr. Opin. Biotechnol.*, 2005. **16**(4): p. 378-384.

187. Suter, B., S. Kittanakom, and I. Stagljar. Interactive proteomics: what lies ahead? *Biotechniques*, 2008. **44**(5): p. 681-691.

188. Hartley, R.W. Barnase and barstar: two small proteins to fold and fit together. *Trends Biochem. Sci.*, 1989. **14**(11): p. 450-4.

189. Buckle, A.M., G. Schreiber, and A.R. Fersht. Protein-protein recognition: crystal structural analysis of a barnase-barstar complex at 2.0-A resolution. *Biochemistry*, 1994. **33**(30): p. 8878-89.

190. Schreiber, G. and A.R. Fersht. Energetics of protein-protein interactions: analysis of the barnase-barstar interface by single mutations and double mutant cycles. *J. Mol. Biol.*, 1995. **248**(2): p. 478-86.

191. Hartley, R.W. Barnase-barstar interaction. *Methods Enzymol.*, 2001. **341**: p. 599-611.

192. Martin, C., V. Richard, M. Salem, R. Hartley, and Y. Mauguen. Refinement and structural analysis of barnase at 1.5 A resolution. *Acta Crystallogr. Sect. D: Biol. Crystallogr.*, 1999. **55**(Pt 2): p. 386-98.

193. Martin, C., R. Hartley, and Y. Mauguen. X-ray structural analysis of compensating mutations at the barnase-barstar interface. *FEBS Lett.*, 1999. **452**(3): p. 128-32.

194. Hartley, R.W. Directed mutagenesis and barnase-barstar recognition. *Biochemistry*, 1993. **32**(23): p. 5978-84.

195. Meiering, E.M., M. Bycroft, and A.R. Fersht. Characterization of Phosphate Binding in the Active-Site of Barnase by Site-Directed Mutagenesis and Nmr. *Biochemistry*, 1991. **30**(47): p. 11348-11356.

196. Mossakowska, D.E., K. Nyberg, and A.R. Fersht. Kinetic characterization of the recombinant ribonuclease from Bacillus amyloliquefaciens (barnase) and investigation of key residues in catalysis by site-directed mutagenesis. *Biochemistry*, 1989. **28**(9): p. 3843-50.

197. Schreiber, G. and A.R. Fersht. Interaction of barnase with its polypeptide inhibitor barstar studied by protein engineering. *Biochemistry*, 1993. **32**(19): p. 5145-50.

198. Schreiber, G., C. Frisch, and A.R. Fersht. The role of Glu73 of barnase in catalysis and the binding of barstar. *J. Mol. Biol.*, 1997. **270**(1): p. 111-22.

199. Jucovic, M. and R.W. Hartley. Protein-protein interaction: a genetic selection for compensating mutations at the barnase-barstar interface. *Proc. Natl. Acad. Sci. U. S. A.*, 1996. **93**(6): p. 2343-7.

200. Hartley, R.W. Barnase and barstar. Expression of its cloned inhibitor permits expression of a cloned ribonuclease. *J. Mol. Biol.*, 1988. **202**(4): p. 913-5.

201. Jucovic, M. and R.W. Hartley. In-Vivo System for the Detection of Low-Level Activity Barnase Mutants. *Protein Eng.*, 1995. **8**(5): p. 497-499.

202. Paddon, C.J. and R.W. Hartley. Expression of Bacillus amyloliquefaciens extracellular ribonuclease (barnase) in Escherichia coli following an inactivating mutation. *Gene*, 1987. **53**(1): p. 11-9.

203. Herrmann, C., G.A. Martin, and A. Wittinghofer. Quantitative analysis of the complex between p21ras and the Ras-binding domain of the human Raf-1 protein kinase. *J. Biol. Chem.*, 1995. **270**(7): p. 2901-5.

204. Guillet, V., A. Lapthorn, R.W. Hartley, and Y. Mauguen. Recognition between a bacterial ribonuclease, barnase, and its natural inhibitor, barstar. *Structure*, 1993. **1**(3): p. 165-76.

205. Moodie, S.A. and A. Wolfman. The 3rs of Life - Ras, Raf and Growth-Regulation. *Trends Genet.*, 1994. **10**(2): p. 44-48.

206. Dhillon, A.S., S. Hagan, O. Rath, and W. Kolch. MAP kinase signalling pathways in cancer. *Oncogene*, 2007. **26**(22): p. 3279-3290.

207. Fabian, J.R., A.B. Vojtek, J.A. Cooper, and D.K. Morrison. A Single Amino-Acid Change in Raf-1 Inhibits Ras Binding and Alters Raf-1 Function. *Proc. Natl. Acad. Sci. U. S. A.*, 1994. **91**(13): p. 5982-5986.

208. Block, C., R. Janknecht, C. Herrmann, N. Nassar, and A. Wittinghofer. Quantitative structure-activity analysis correlating Ras/Raf interaction in vitro to Raf activation in vivo. *Nat. Struct. Biol.*, 1996. **3**(3): p. 244-51.

209. Kiel, C., L. Serrano, and C. Herrmann. A detailed thermodynamic analysis of ras/effector complex interfaces. *J. Mol. Biol.*, 2004. **340**(5): p. 1039-58.

210. Cutler, R.E. and D.K. Morrison. Mammalian Raf-1 is activated by mutations that restore Raf signaling in Drosophila. *EMBO J.*, 1997. **16**(8): p. 1953-1960.

211. Nassar, M., G. Horn, C. Herrmann, A. Scherer, F. Mccormick, and A. Wittinghofer. The 2.2-Angstrom Crystal-Structure of the Ras-Binding Domain of the Serine Threonine Kinase C-Raf1 in Complex with Rap1a and a Gtp Analog. *Nature*, 1995. **375**(6532): p. 554-560.

212. Hoi, H., N.C. Shaner, M.W. Davidson, C.W. Cairo, J. Wang, and R.E. Campbell. A monomeric photoconvertible fluorescent protein for imaging of dynamic protein localization. *J. Mol. Biol.*, 2010. **401**(5): p. 776-91.

213. Prior, I.A. and J.F. Hancock. Compartmentalization of Ras proteins. *J. Cell Sci.*, 2001. **114**(9): p. 1603-1608.

214. Apolloni, A., I.A. Prior, M. Lindsay, R.G. Parton, and J.F. Hancock. H-ras but not K-ras traffics to the plasma membrane through the exocytic pathway. *Mol. Cell. Biol.*, 2000. **20**(7): p. 2475-2487.

215. Willumsen, B.M., A. Christensen, N.L. Hubbert, A.G. Papageorge, and D.R. Lowy. The P21 Ras C-Terminus Is Required for Transformation and Membrane Association. *Nature*, 1984. **310**(5978): p. 583-586.

216. Seeburg, P.H., W.W. Colby, D.J. Capon, D.V. Goeddel, and A.D. Levinson. Biological Properties of Human C-Ha-Ras 1 Genes Mutated at Codon-12. *Nature*, 1984. **312**(5989): p. 71-75.

217. Scheffzek, K., M.R. Ahmadian, W. Kabsch, L. Wiesmuller, A. Lautwein, F. Schmitz, and A. Wittinghofer. The Ras-RasGAP complex: Structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science*, 1997. **277**(5324): p. 333-338.

218. Barik, S. Relationship between opacity of transformed E-coli colonies and over-expression of the recombinant transcript. *Biotechniques*, 1997. **22**(1): p. 112-118.

219. Koch, H., N. Grafe, R. Schiess, and A. Pluckthun. Direct selection of antibodies from complex libraries with the protein fragment complementation assay. *J. Mol. Biol.*, 2006. **357**(2): p. 427-41.

220. Lee, S.K., H.H. Chou, B.F. Pfleger, J.D. Newman, Y. Yoshikuni, and J.D. Keasling. Directed evolution of AraC for improved compatibility of arabinose- and lactose-inducible promoters. *Appl. Environ. Microbiol.*, 2007. **73**(18): p. 5711-5715.

221. Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular cloning : a laboratory manual*. 2nd ed1989, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory.

222. Bessette, P.H., M.A. Mena, A.W. Nguyen, and P.S. Daugherty. Construction of designed protein libraries using gene assembly mutagenesis. *Methods Mol. Biol.*, 2003. **231**: p. 29-37.

223. Nakai, J., M. Ohkura, and K. Imoto. A high signal-to-noise Ca(2+) probe composed of a single green fluorescent protein. *Nat. Biotechnol.*, 2001. **19**(2): p. 137-41.

224. Kerppola, T.K. Visualization of molecular interactions using bimolecular fluorescence complementation analysis: characteristics of protein fragment complementation. *Chem. Soc. Rev.*, 2009. **38**(10): p. 2876-86.

225. Carlson, H.J., D.W. Cotton, and R.E. Campbell. Circularly permuted monomeric red fluorescent proteins with new termini in the beta-sheet. *Protein Sci.*, 2010. **19**(8): p. 1490-9.

226. Martin, B.R., B.N. Giepmans, S.R. Adams, and R.Y. Tsien. Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. *Nat. Biotechnol.*, 2005. **23**(10): p. 1308-14.

227. Dong, J., F. Abulwerdi, A. Baldridge, J. Kowalik, K.M. Solntsev, and L.M. Tolbert. Isomerization in fluorescent protein chromophores involves addition/elimination. *J. Am. Chem. Soc.*, 2008. **130**(43): p. 14096-8.

228. Chen, J., X.F. Zheng, E.J. Brown, and S.L. Schreiber. Identification of an 11-kDa FKBP12-rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue. *Proc. Natl. Acad. Sci. U. S. A.*, 1995. **92**(11): p. 4947-51.

229. Xu, X., A.L. Gerard, B.C. Huang, D.C. Anderson, D.G. Payan, and Y. Luo. Detection of programmed cell death using fluorescence energy transfer. *Nucleic Acids Res.*, 1998. **26**(8): p. 2034-5.

230. Kirchhofer, A., J. Helma, K. Schmidthals, C. Frauer, S. Cui, A. Karcher, M. Pellis, S. Muyldermans, C.S. Casas-Delucchi, M.C. Cardoso, H. Leonhardt, K.P. Hopfner, and U. Rothbauer. Modulation of protein properties in living cells using nanobodies. *Nat. Struct. Mol. Biol.*, 2010. **17**(1): p. 133-8.

231. Jackrel, M.E., A.L. Cortajarena, T.Y. Liu, and L. Regan. Screening libraries to identify proteins with desired binding activities using a split-GFP reassembly assay. *ACS Chem. Biol.*, 2010. **5**(6): p. 553-62.

232. Strongin, D.E., B. Bevis, N. Khuong, M.E. Downing, R.L. Strack, K. Sundaram, B.S. Glick, and R.J. Keenan. Structural rearrangements near the chromophore influence the maturation speed and brightness of DsRed variants. *Protein Eng. Des. Sel.*, 2007. **20**(11): p. 525-34.

233. Barik, S., *Site-Directed Mutagenesis In Vitro by Megaprimer PCR*, 1996. p. 203-215.

234. Laemmli, U.K. Cleavage of Structural Proteins during Assembly of Head of Bacteriophage-T4. *Nature*, 1970. **227**(5259): p. 680-&.

235. Ward, W.W., *Properties of the Coelenterate green-fluorescent proteins*. Bioluminescence and Chemiluminescence: Basic Chemistry and Analytical applications, ed. a.D.W.M. M. De Luca1981, New York: Academic Press.

236. Wen, W., J.L. Meinkoth, R.Y. Tsien, and S.S. Taylor. Identification of a signal for rapid export of proteins from the nucleus. *Cell*, 1995. **82**(3): p. 463-73.

237. Palmer, A.E. and R.Y. Tsien. Measuring calcium signaling using genetically targetable fluorescent indicators. *Nat. Protoc.*, 2006. **1**(3): p. 1057-1065.

238. Kuhlman, B., G. Dantas, G.C. Ireton, G. Varani, B.L. Stoddard, and D. Baker. Design of a novel globular protein fold with atomic-level accuracy. *Science*, 2003. **302**(5649): p. 1364-8.

239. Terskikh, A.V., A.F. Fradkov, A.G. Zaraisky, A.V. Kajava, and B. Angres. Analysis of DsRed Mutants. Space around the fluorophore accelerates fluorescence development. *J. Biol. Chem.*, 2002. **277**(10): p. 7633-6.

240. Alford, S.C., A.S. Abdelfattah, Y. Ding, and R.E. Campbell. A fluorogenic red fluorescent protein heterodimer. *Chem. Biol.*, 2012. **19**(3): p. 353-60.

241. Grefen, C., P. Obrdlik, and K. Harter. The determination of protein-protein interactions by the mating-based split-ubiquitin system (mbSUS). *Methods Mol. Biol.*, 2009. **479**: p. 217-33.

242. Cadwallader, K.A., H. Paterson, S.G. Macdonald, and J.F. Hancock. N-terminally myristoylated Ras proteins require palmitoylation or a polybasic domain for plasma membrane localization. *Mol. Cell. Biol.*, 1994. **14**(7): p. 4722-30.

243. Hancock, J.F., K. Cadwallader, H. Paterson, and C.J. Marshall. A CAAX or a CAAL motif and a second signal are sufficient for plasma membrane targeting of ras proteins. *EMBO J.*, 1991. **10**(13): p. 4033-9.

244. Kavran, J.M., D.E. Klein, A. Lee, M. Falasca, S.J. Isakoff, E.Y. Skolnik, and M.A. Lemmon. Specificity and promiscuity in phosphoinositide binding by pleckstrin homology domains. *J. Biol. Chem.*, 1998. **273**(46): p. 30497-508.

245. Lynes, E.M. and T. Simmen. Urban planning of the endoplasmic reticulum (ER): how diverse mechanisms segregate the many functions of the ER. *Biochim. Biophys. Acta.*, 2011. **1813**(10): p. 1893-905.

246. Lynes, E.M., M. Bui, M.C. Yap, M.D. Benson, B. Schneider, L. Ellgaard, L.G. Berthiaume, and T. Simmen. Palmitoylated TMX and calnexin target to the mitochondria-associated membrane. *EMBO J*, 2011. **31**(2): p. 457-70.

247. Hayashi, T., R. Rizzuto, G. Hajnoczky, and T.P. Su. MAM: more than just a housekeeper. *Trends Cell Biol.*, 2009. **19**(2): p. 81-8.

248. Baker, M.J., A.E. Frazier, J.M. Gulbis, and M.T. Ryan. Mitochondrial protein-import machinery: correlating structure with function. *Trends Cell Biol.*, 2007. **17**(9): p. 456-64.

249. Lynes, E.M., M. Bui, M.C. Yap, M.D. Benson, B. Schneider, L. Ellgaard, L.G. Berthiaume, and T. Simmen. Palmitoylated TMX and calnexin target to the mitochondria-associated membrane. *EMBO J*, 2011.

250. De Vos, K.J., G.M. Morotz, R. Stoica, E.L. Tudor, K.F. Lau, S. Ackerley, A. Warley, C.E. Shaw, and C.C. Miller. VAPB interacts with the mitochondrial protein PTPIP51 to regulate calcium homeostasis. *Hum. Mol. Genet.*, 2011.

251. Kanaji, S., J. Iwahashi, Y. Kida, M. Sakaguchi, and K. Mihara. Characterization of the signal that directs Tom20 to the mitochondrial outer membrane. *J. Cell Biol.*, 2000. **151**(2): p. 277-88.

252. Hancock, J.F., A.I. Magee, J.E. Childs, and C.J. Marshall. All ras proteins are polyisoprenylated but only some are palmitoylated. *Cell*, 1989. **57**(7): p. 1167-77.

253. Pelletier, J.N., K.M. Arndt, A. Pluckthun, and S.W. Michnick. An in vivo library-versus-library selection of optimized protein-protein interactions. *Nat. Biotechnol.*, 1999. **17**(7): p. 683-690.

254. Gorman, C., R.H. Skinner, J.V. Skelly, S. Neidle, and P.N. Lowe. Equilibrium and kinetic measurements reveal rapidly reversible binding of ras to raf. *J. Biol. Chem.*, 1996. **271**(12): p. 6713-6719.

255. Scheer, J.M., J.A. Wells, and M.J. Romanowski. Malonate-assisted purification of human caspases. *Protein Expression Purif.*, 2005. **41**(1): p. 148-153.

256. Prior, T.I., S. Kunwar, and I. Pastan. Studies on the activity of barnase toxins in vitro and in vivo. *Bioconjugate Chem.*, 1996. **7**(1): p. 23-29.

257. Parrish, J.R., K.D. Gulyas, and R.L. Finley, Jr. Yeast two-hybrid contributions to interactome mapping. *Curr. Opin. Biotechnol.*, 2006. **17**(4): p. 387-93.

258. Arndt, K.M., J.N. Pelletier, K.M. Muller, T. Alber, S.W. Michnick, and A. Pluckthun. A heterodimeric coiled-coil peptide pair selected in vivo from a designed library-versus-library ensemble. *J. Mol. Biol.*, 2000. **295**(3): p. 627-39.

259. Arndt, K.M., J.N. Pelletier, K.M. Muller, A. Pluckthun, and T. Alber. Comparison of in vivo selection and rational design of heterodimeric coiled coils. *Structure*, 2002. **10**(9): p. 1235-48.

260. Campbell-Valois, F.X., K. Tarassov, and S.W. Michnick. Massive sequence perturbation of a small protein. *Proc. Natl. Acad. Sci. U. S. A.*, 2005. **102**(42): p. 14988-93.

261. Jackrel, M.E., R. Valverde, and L. Regan. Redesign of a protein–peptide interaction: Characterization and applications. *Protein Sci.*, 2009. **18**(4): p. 762-774.

262. Binz, H.K. and A. Plückthun. Engineered proteins as specific binding reagents. *Curr. Opin. Biotechnol.*, 2005. **16**(4): p. 459-469.

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

264. Chica, R.A., M.M. Moore, B.D. Allen, and S.L. Mayo. Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries. *Proc. Natl. Acad. Sci. U. S. A.*, 2010. **107**(47): p. 20257-62.

265. Hunt, M.E., C.K. Modi, G.V. Aglyamova, D.V. Ravikant, E. Meyer, and M.V. Matz. Multi-domain GFP-like proteins from two species of marine hydrozoans. *Photochem. Photobiol. Sci.*, 2012.

266. Althoff, E.A. and V.W. Cornish. A bacterial small-molecule three-hybrid system. *Angew. Chem. Int. Ed. Engl.*, 2002. **41**(13): p. 2327-30.

267. Banaszynski, L.A., C.W. Liu, and T.J. Wandless. Characterization of the FKBP center dot Rapamycin center dot FRB ternary complex. *J. Am. Chem. Soc.*, 2005. **127**(13): p. 4715-4721.

268. Lerner, C.G. and M. Inouye. Low Copy Number Plasmids for Regulated Low-Level Expression of Cloned Genes in Escherichia-Coli with Blue

White Insert Screening Capability. *Nucleic Acids Res.*, 1990. **18**(15): p. 4631-4631.

269. Hu, C.D., Y. Chinenov, and T.K. Kerppola. Visualization of interactions among bZip and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol. Cell*, 2002. **9**(4): p. 789-798.

270. Ottmann, C., M. Weyand, A. Wolf, J. Kuhlmann, and C. Ottmann. Applicability of superfolder YFP bimolecular fluorescence complementation in vitro. *Biol. Chem.*, 2009. **390**(1): p. 81-90.

271. Feinberg, E.H., M.K. VanHoven, A. Bendesky, G. Wang, R.D. Fetter, K. Shen, and C.I. Bargmannl. GFP reconstitution across synaptic partners (GRASP) defines cell contacts and Synapses in living nervous systems. *Neuron*, 2008. **57**(3): p. 353-363.

272. Kim, J., T. Zhao, R.S. Petralia, Y. Yu, H.C. Peng, E. Myers, and J.C. Magee. mGRASP enables mapping mammalian synaptic connectivity with light microscopy. *Nat. Methods*, 2012. **9**(1): p. 96-U139.

273. Walsh, M.K. and J.W. Lichtman. In vivo time-lapse imaging of synaptic takeover associated with naturally occurring synapse elimination. *Neuron*, 2003. **37**(1): p. 67-73.