

University of Alberta

Generation of a fluorescent protein-derived biosensor based on alternate frame folding

by

Ritesh K. Saini

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

Master of Science

Department of Chemistry

©Ritesh K. Saini

Fall 2012

Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

Abstract

To develop a single fluorescent protein (FP)-based scaffold that could potentially serve as a universal platform for biosensor creation, we have introduced the tandem alternate frame switching strategy. We genetically inserted a non-identical copy of 7th β -strand in *Aequorea victoria* greenFP(avGFP) to create protein conformational switch. By creating a series of variants, we demonstrated that equilibrium between conformations can be manipulated by substitution, deletion and insertion of residues. The chromophore absorption spectra unique to each conformation allowed the determination of bound equilibrium through independent component analysis. To verify the potential of this strategy, we strategically created calcium-ion (Ca^{2+}) and cyclic adenosine-monophosphate (cAMP) biosensors and evaluated the change in the fluorescence spectra corresponding to binding of the analyte. We performed preliminary study to explore switch generation in cpmCherry. We anticipate that this rational strategy for construction of genetically-encoded indicators could be applicable to a variety of FPs and sensing motifs.

Acknowledgement

First and foremost, I want to acknowledge the invaluable advice, support, mentorship and insights of my supervisor, Dr. Robert Campbell. He created the recombinant constructs and gave me enough room to work independently. The successful completion of this thesis was not possible without his help and support. I am indebted to him more than he would admit.

I am also thankful to Spencer Alford for teaching me the techniques; guiding me and helping me develop my background in chemistry; all lab mates and colleagues for their help and making the labwork an excellent experience. The acknowledgement would be incomplete without recognizing the help and support of staff of Department of Chemistry.

Beyond Chemistry, I want to acknowledge Jaya Nagpal and Suparna Sarkar for being a pleasant housemate. I also want to extend my deepest gratitude to Ryan Viases, Zahra Shire and other friends who kept encouraging me and giving me strength in all ups and downs of graduate school.

Finally, I owe my earnest appreciation to my family, for their continuous support, best wishes and immense faith in me.

Table of Contents

CHAPTER 1: INTRODUCTION	1
1.1 INTRODUCTION TO FLUORESCENT PROTEIN-BASED BIOSENSORS	2
1.2 INTRODUCTION TO FPS	3
1.3 DEVELOPMENT OF FP-BASED BIOSENSORS	6
<i>1.3.1 FRET-based biosensors</i>	<i>7</i>
<i>1.3.2 BiFC (Bimolecular fluorescence complementation)</i>	<i>9</i>
<i>1.3.3 Single FP-based biosensors</i>	<i>10</i>
1.3.3.1 Tandem fusion of MRE to FP.....	12
1.3.3.2 Insertion of an MRE into a FP.....	13
1.3.3.3 Insertion of FP in RE.....	14
1.3.3.4 Tandem fusion of MRE to circularly permuted FP.....	14
1.3.3.5 Insertion of MRE in circularly permuted FP.....	15
1.3.3.6 Insertion of circularly permuted FP in MRE.....	15
1.3.3.7 Summary of single FP-based design strategy.....	16
1.4 OVERVIEW OF EXISTING SWITCHING MECHANISMS FOR DEVELOPMENT OF SWITCH-BASED FP	17
<i>1.4.1 Rigid domain movement</i>	<i>19</i>
<i>1.4.2 Limited structural rearrangement</i>	<i>19</i>
<i>1.4.3 Global fold switching</i>	<i>21</i>
<i>1.4.4 Folding-unfolding switches</i>	<i>23</i>
1.5 EMPLOYMENT OF PROTEIN-BASED SWITCHES IN BIOSENSOR DESIGN	24
1.6 OBJECTIVES OF THE THESIS	27

CHAPTER 2: DEVELOPMENT AND CHARACTERIZATION OF AN AVGFP VARIANT CAPABLE OF TANDEM ALTERNATE FRAME SWITCHING.....	29
2.1 INTRODUCTION	30
2.2 MATERIALS AND METHODS	35
<i>2.2.1 General methods and materials</i>	<i>35</i>
<i>2.2.2 Construction of an avGFP encoding plasmid</i>	<i>37</i>
<i>2.2.3 Construction of control protein constructs</i>	<i>38</i>
<i>2.2.4 Protein expression and purification</i>	<i>38</i>
<i>2.2.5 In vitro spectroscopic analysis of protein constructs.....</i>	<i>40</i>
<i>2.2.6 Data analysis of protein constructs</i>	<i>40</i>
<i>2.2.7 Construction of Ca²⁺ and cyclic AMP biosensors.....</i>	<i>41</i>
<i>2.2.8 Characterization of biosensor constructs.....</i>	<i>41</i>
2.2.8.1 Calcium biosensor.....	41
2.2.8.2 Cyclic AMP biosensor	42
2.3 RESULTS AND DISCUSSION.....	43
<i>2.3.1 Development and evaluation of TAFS-based avGFP scaffold ...</i>	<i>43</i>
<i>2.3.2 Manipulating the bound equilibrium.....</i>	<i>49</i>
2.3.2.1 Insertion of residues	50
2.3.2.2 Deletion of residues	52
2.3.2.3 Mutation of residues	54
2.3.2.4 Rational manipulation of version 2 variants	56
<i>2.3.3 Insertion of binding modules to create biosensors</i>	<i>58</i>
2.3.3.1 Insertion of CAM/M13 from YC-Nano15 to create Ca ²⁺ biosensors.....	61
2.3.3.2 Insertion of CAP to create cyclic AMP biosensors	64
2.4 CONCLUSION.....	66

CHAPTER 3: PRELIMINARY STUDY TOWARDS DEVELOPMENT OF RFP-BASED AFF SCAFFOLD	68
3.1 INTRODUCTION	69
3.2 MATERIALS AND METHODS	72
<i>3.2.1 General methods and materials</i>	<i>72</i>
<i>3.2.2 Search and selection of the consensus motif</i>	<i>72</i>
<i>3.2.3 Site directed mutagenesis studies.....</i>	<i>73</i>
<i>3.2.4 Screening of library.....</i>	<i>74</i>
<i>3.2.5 Characterization and analysis of variants.....</i>	<i>75</i>
<i>3.2.6 Construction and analysis of biosensor scaffold with alternate frame 75</i>	
3.3 RESULTS AND DISCUSSION.....	76
<i>3.3.1 Library creation and screening.....</i>	<i>76</i>
<i>3.3.2 Generation of Ca²⁺ biosensor scaffold</i>	<i>81</i>
3.4 CONCLUSION.....	82
 CHAPTER 4: CONCLUSION AND FUTURE DIRECTIONS	84
 BIBLIOGRAPHY	90

LIST OF FIGURES

FIGURE 1-1 CARTOON REPRESENTATION OF AVGFP.....	4
FIGURE 1-2 REVISED MECHANISM OF CHROMOPHORE MATURATION	5
FIGURE 1-3 FRET-BASED BIOSENSOR DESIGN.....	8
FIGURE 1-4 CARTOON REPRESENTATION OF BiFC APPROACH.. ..	10
FIGURE 1-5 CARTOON ILLUSTRATION OF CONSTRUCTION APPROACHES OF SINGLE FP- BASED BIOSENSORS.....	12
FIGURE 1-6 CARTOON ILLUSTRATION OF CIRCULAR PERMUTATION.. ..	15
FIGURE 1-7 GLOBAL SWITCHING.	22
FIGURE 1-8 FOLDING-UNFOLDING EQUILIBRIUM OF CAM.....	24
FIGURE 1-9 CARTOON REPRESENTATION OF BARNASE-AFF DEVELOPED BY LOH <i>ET</i> <i>AL.</i> ⁶⁷	26
FIGURE 2-1 CARTOON REPRESENTATION OF THE AVGFP STRUCTURE.	32
FIGURE 2-2 CARTOON REPRESENTATION OF PREDICTED CONFORMATIONAL EQUILIBRIUM BETWEEN ALTERNATE FRAME FOLDING INVOLVING STRAND 7A (S7A) & STRAND 7B (S7B)	33

FIGURE 2-3 CARTOON REPRESENTATION OF INSERTED B-STRAND 7 (B) WITH HIS148ASP AND LINKER SUCCEEDING THE ORIGINAL B-STRAND 7 (A) WITH HIS148.....	44
FIGURE 2-4 NORMALIZED ABSORBANCE SPECTRA EXHIBITED BY THE VARIANTS WITH SINGLE MUTATION AT 148.....	44
FIGURE 2-5 FITTING AHBD AS A SUM OF THE AHBD AND AHBD REVEALS THAT THE SPECTRUM IS 97±2% AHBD AND 3±17% AHBD.	45
FIGURE 2-6 DESIGN OF V2 VARIANT.	57
FIGURE 2-7 CARTOON ILLUSTRATION OF BIOSENSOR CONSTRUCTION.....	60
FIGURE 2-8 Ca ²⁺ -BIOSENSOR TITRATION CURVE.	64
FIGURE 3-1 SCHEMATIC REPRESENTATION OF OVERALL FOLDING OF CP196 MCHERRY.	71
FIGURE 3-2 PYMOL RENDERED VIEW OF CONSENSUS MOTIF RESIDUES (AS LABELED) WITH RESPECT TO THE CHROMOPHORE (LABELED AS CRO). ¹⁴	77
FIGURE 3-3 DIAGRAM DEPICTING CONSTRUCT DESIGN.	82

LIST OF TABLES

TABLE 1-1 COMPARISON OF VARIOUS DESIGNS OF Ca^{2+} INDICATORS.	20
TABLE 2-1 CHARACTERIZATION OF V1 VARIANTS WITH INSERTIONS	51
TABLE 2-2 THERMODYNAMIC CHARACTERIZATION OF V1 VARIANTS WITH DELETIONS. `D' REPRESENTS DELETED AND LETTER CODE OF AMINO ACID DELETED IS SHOWN NEXT TO `D'.	52
TABLE 2-3 CHARACTERIZATION OF V1 VARIANTS WITH MUTATIONS	54
TABLE 2-4 CHARACTERIZATION OF V2 VARIANTS WITH MANIPULATIONS	57
TABLE 2-5 EVALUATION OF Ca^{2+} BIOSENSOR PERFORMANCE (EXCITATION RATIOMETRIC CHANGE)	62
TABLE 2-6 EVALUATION OF cAMP BIOSENSOR PERFORMANCE (EXCITATION RATIOMETRIC CHANGE)	65
TABLE 3-1 DISTRIBUTION OF RESIDUES ALLOWED AT EACH POSITION.....	79
TABLE 3-3 FLUORESCENCE ANALYSIS DATA OF DEPICTED VARIANTS.....	81
TABLE 3-2 SEQUENCES OF CORRESPONDING FRAMES DESIGNED TO GENERATE PEAK SHIFT.....	82

List of abbreviations

AFF	Artificial frame folding
ATP	Adenosine triphosphate
avGFP	<i>Aequorea victoria</i> green fluorescent protein
B-PER	Bacterial protein extraction reagent
Ca ²⁺	Calcium ion
CaM	Calmodulin
cAMP	Cyclic adenosine monophosphate
CAP	Catabolite activator protein
cGMP	Cyclic guanosine monophosphate
cp	Circularly permuted
CpmCherry	Circularly permuted monomeric cherry
CRP	cAMP receptor protein
dI	Deionized
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dsDNA	Double stranded deoxyribonucleic acid
EGTA	Ethyleneglycol tetraacetic acid
EGFP	Enhanced green fluorescent protein
FP	Fluorescent protein
FRET	Förster resonance energy transfer

IPTG	Isopropyl β -D-1-thiogalactopyranoside
K_d	Dissociation constant
K_{eq}	Equilibrium constant
LB	Luria-Bertani medium
M13	Myosin light-chain kinase target peptide
MBSU	Molecular biology services unit at the University of Alberta
MCS	Multiple cloning site
MOPS	4-morpholinepropanesulfonic acid
MRE	Molecular recognition element
NADH	Reduced nicotinamide adenine dinucleotide
Ni-NTA	Ni^{2+} nitrilotriacetate
PCR	Polymerase chain reaction
PDB	Protein data bank
SOB	Super optimal broth
TAFS	Tandem alternate frame switching
TRIS-HCl	tris(hydroxymethyl)aminomethane hydrochloride
UV	Ultraviolet light
wt	Wild-type
YC-Nano15	Yellow cameleon version 15 with nanomolar K_d

Chapter 1: Introduction

1.1 Introduction to Fluorescent Protein-based biosensors

Fluorescent protein (FP)-based biosensors^(*) are analytical probes with a biomolecular recognition element as receptor and an auto-fluorescent protein as transducer. FP-based biosensors are one of the most valued probes for live cell imaging and for elucidation of biological activities in living cells.¹ The conversion of FPs into bioanalytical probes is thus counted as one of the revolutionary advances in cell biology.

The construction of biosensors from FPs typically involves genetically fusing specific molecular recognition elements (MREs) to FPs. The simple fusion of a receptor and a FP, however, frequently modifies the native function of the receptor or the spectral properties of the FP.¹ The interplay between each recognition and transduction event changes with corresponding MREs and is unique to each biosensor. Thus, the design and optimization of FP-based biosensors is not trivial and achieving physiologically useful responses and sensitivity are not guaranteed.^{1,2} Accordingly, a wide number of design strategies have been employed for construction of biosensors targeting different analytes or processes. The design of biosensors is hindered by the lack of generic approach. The

^(*) As per IUPAC, "A biosensor is an integrated receptor-transducer device, which is capable of providing selective quantitative or semi-quantitative analytical information using a biological recognition element."⁸⁸

purpose of the thesis is to develop a versatile biosensor design strategy by way of engineering a generic FP scaffold.

1.2 Introduction to FPs

Aequorea victoria green FP (avGFP) was the first FP to be discovered and isolated by Osamu Shimomura from the *A. victoria* jellyfish.³ The primary sequence of avGFP first reported by Prasher *et al.* and the functional expression of avGFP in *C. elegans* by Chalfie *et al.* brought FPs to the widespread attention of the biological community.^{4,5} The insight into local chemistry in the vicinity of the chromophore and the unique β -barrel structure was revealed by the x-ray crystal structure, as shown in **Figure 1-1**.^{6,7} Roger Tsien and coworkers took the lead in rationally engineering and improving avGFP for use in fluorescence imaging applications.⁸⁻¹¹ For their groundbreaking work on the discovery and development of avGFP, Shimomura, Chalfie and Tsien were awarded the 2008 Nobel Prize in Chemistry.

The fluorescence of FPs is attributed to their intrinsic chromophores, derived from 3 primary amino acids, Ser65-Tyr66-Gly67.⁸ Chromophore formation in the FPs is a posttranslational modification that occurs only in the context of the folded protein.¹² A recently proposed 'universal' mechanism for chromophore formation of avGFP (and red variants to be discussed below) is depicted in **Figure 1-2**.¹³

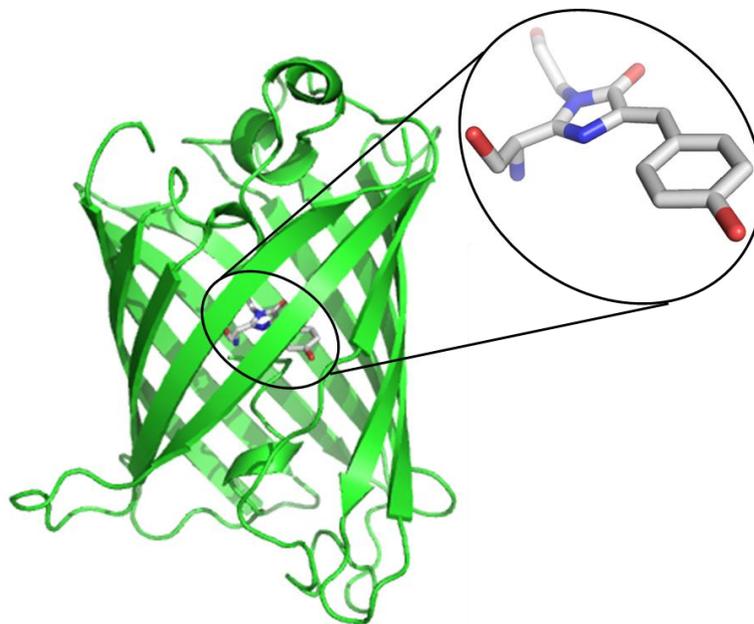


Figure 1-1 Cartoon representation of avGFP. The inset view is a stick representation of the intrinsic chromophore, derived from 3 primary amino acids Ser65-Tyr66-Gly67, responsible for its fluorescent properties. Figure 1.1 was rendered using Pymol (PDB ID: 1EMA)¹⁴

Protein structure in the vicinity of the chromophore plays a critical role in determining the final spectral outcome.^{6,15} Most famously, in the case of wild-type avGFP, an excited-state proton transfer pathway (ESPT) has an influence on the absorbance spectrum of the protein and relies on an H-bond wire that runs through the side chains of surrounding residues including His148 and Glu222.^{6,15} Single mutation Ser65Thr in wild-type avGFP leads to disruption of ESPT whereas additional mutation His148Asp to Ser65Thr wild-type avGFP retains ESPT, verifying the significance of single avGFP residue.^{6,15,16}

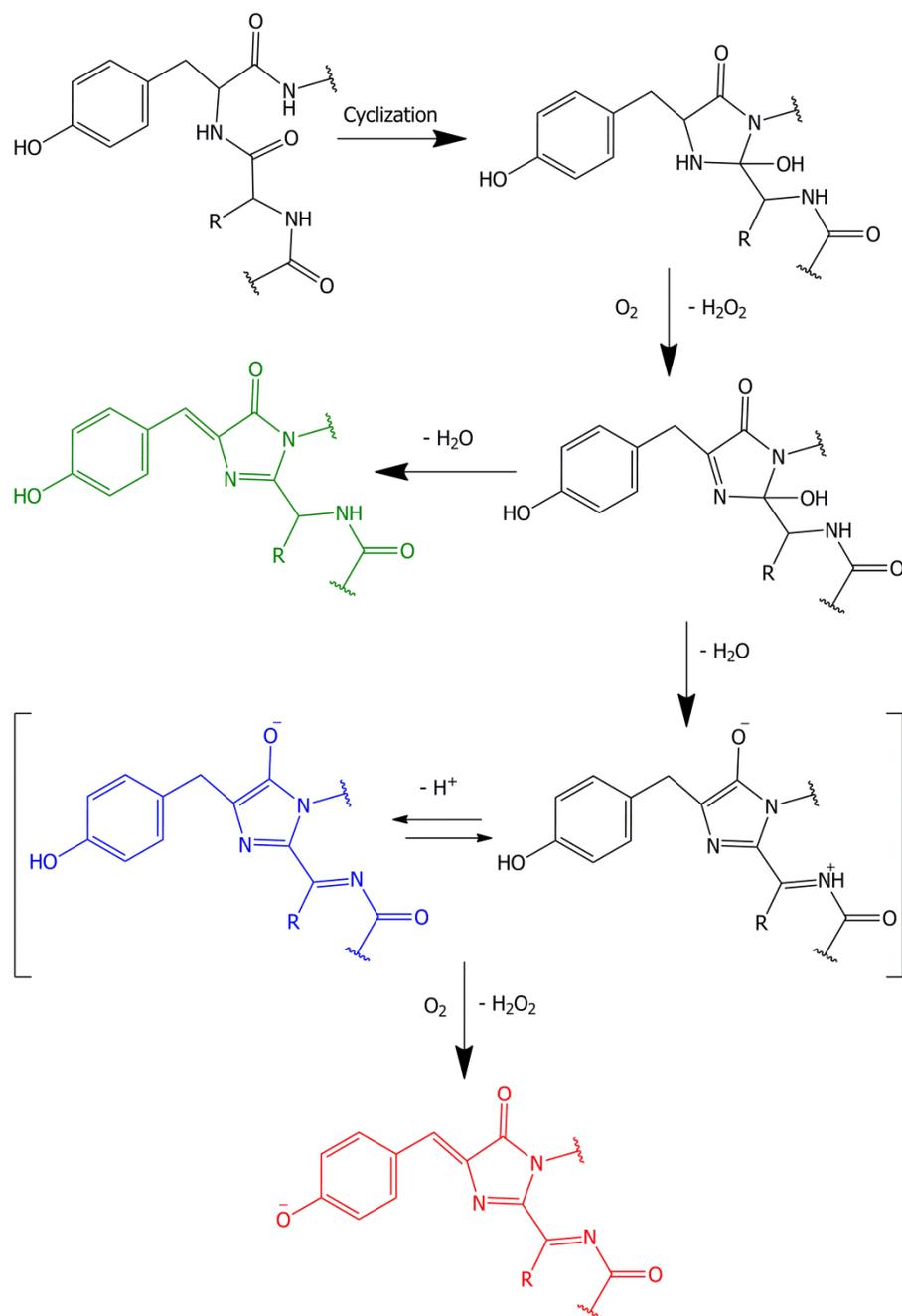


Figure 1-2 Revised mechanism of chromophore maturation. It involves a series of reactions including cyclization, oxidation and dehydration to form the green chromophore. Formation of the red chromophore involves dehydration of acylimine to form a blue intermediate followed by an additional oxidation step. Adapted from Bravaya *et al.*¹³

The capability of FPs to tolerate the modifications in their framework make them an attractive tool for live-cell imaging.¹⁷ In addition to spectral significance, several mutations are crucial in improving the properties of FPs. For instance, Ser65Thr avGFP, the improved avGFP variant which will be used in our study, possesses a six-fold stronger fluorescence; better photostability; four-fold faster maturation rate than avGFP and primary excitation at 490 nm compared to a major peak at 395 nm and a minor one at 475 nm for avGFP.⁹

The discovery and cloning of a red FP, called DsRed, from *Discosoma* coral was another important achievement given that red fluorescence is preferred for many imaging applications.¹⁸ Specifically, at red wavelengths there is decreased overlap with the absorption spectra of body fluids (*e.g.*, hemoglobin in blood) and lower cytotoxicity owing to no UV irradiation required for excitation.¹⁸ The expansion of the colour palette to red region also enabled researchers to undertake multicolour imaging of cells with a broader range of hues.¹⁹

1.3 Development of FP-based biosensors

The variety of FP candidates with different hues and properties; the insight into their structural data; and possible spectral modulation criteria provided us with choices for construction of biosensors to investigate

different cellular activities.²⁰ There are three primary approaches to design biosensors, including Förster resonance energy transfer (FRET)-based biosensors, single FP-based biosensors, and bimolecular complementation-based biosensors.

1.3.1 FRET-based biosensors

FRET is the radiationless energy transfer from a donor molecule in its electronic excited state to an acceptor molecule.²¹ It generates spectroscopic signals sensitive to molecular orientation and proximity of two fluorophores in the 1–10 nm range.^{11,21} FRET biosensors are essential tools in the biomolecular imaging field and are generically employed to monitor the proximity of biomolecules for elucidating intermolecular protein-protein (**Figure 1-3a**) or peptide-protein interactions; intramolecular conformational changes (**Figure 1-3b**); and enzymatic proteolysis (**Figure 1-3c**).²² The first FRET-based calcium ion (Ca^{2+}) biosensor, known as cameleon, utilized the conformational change of the binding domain on interaction with Ca^{2+} to bring the two FPs at favorable distance for FRET response.¹¹ This class of indicators will be discussed further in **Section 1.3.3**.

The FRET-based cameleon-type Ca^{2+} biosensors have proven to be powerful tools for investigating Ca^{2+} dynamics and neural activity in live cells.^{11,23,24} They are also the archetypes for the FRET-based design

strategy.²⁵⁻²⁹ Although FRET-based biosensors have been engineered to respond to a variety of analytes and enzyme activities, the signal changes with these biosensors are typically small relative to the most highly optimized cameleons.²⁵⁻²⁹

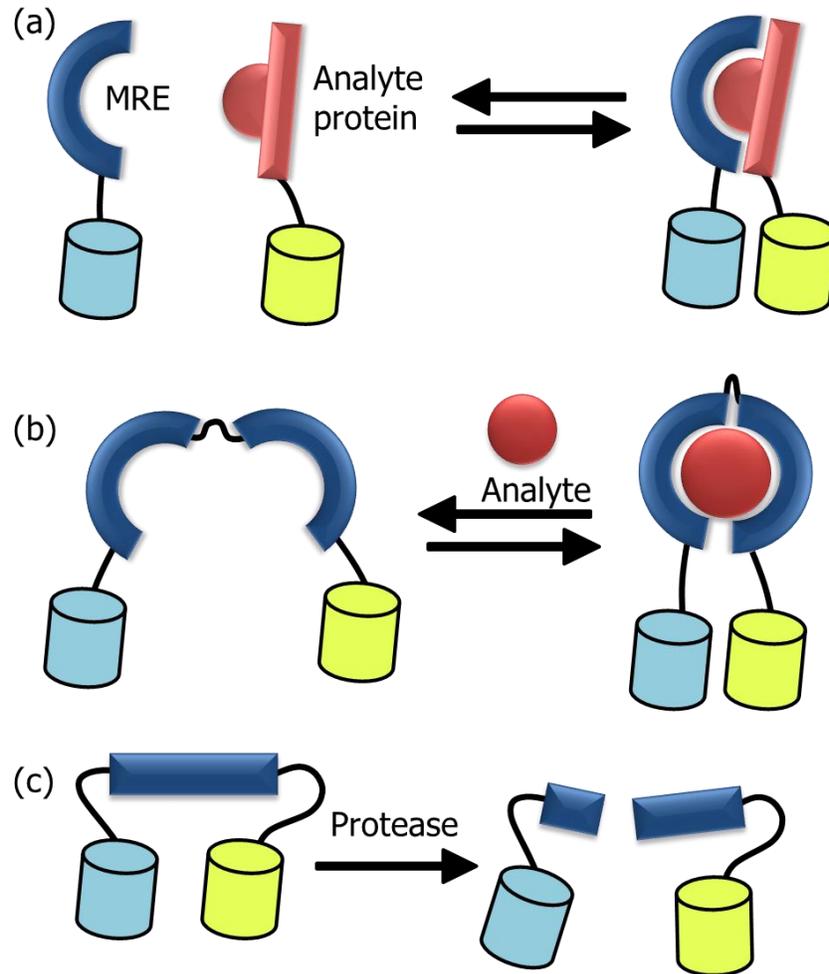


Figure 1-3 FRET-based biosensor design. (a) Analyte protein in red and MRE in blue are linked to donor and acceptor, respectively. Protein-protein interactions can be visualized in live cells by such tagging. (b) A conformation switching MRE between donor and acceptor FRET (c) A protease substrate sandwiched between donor and acceptor FRET pair. Figure 1-3 is adapted from Ibraheem *et al.*²²

Other limitations include uncertainty in the orientation of fluorophores needed for energy transfer and masking of energy transfer by free fluorophores.³⁰ For further information on FRET-based biosensors, the reader is directed to previously published reviews on this topic.^{2,22}

1.3.2 BiFC (Bimolecular fluorescence complementation)

BiFC employs fragments of a split FP, where each fragment is non-fluorescent by itself, to detect the interaction of two polypeptides each fused to one of the fragments as illustrated in **Figure 1-4**.³¹ Upon interaction of the polypeptides, the fragments are brought to suitable spatial proximity allowing for the reconstitution of an intact FP. Boxer *et al.* employed this approach by splitting a single β -strand from the remainder of a FP and studied the thermodynamics of complementation.³² Fan *et al.* developed a split mCherry-based BiFC system and employed it for demonstrating interaction between large T antigen and human p53 protein.³¹

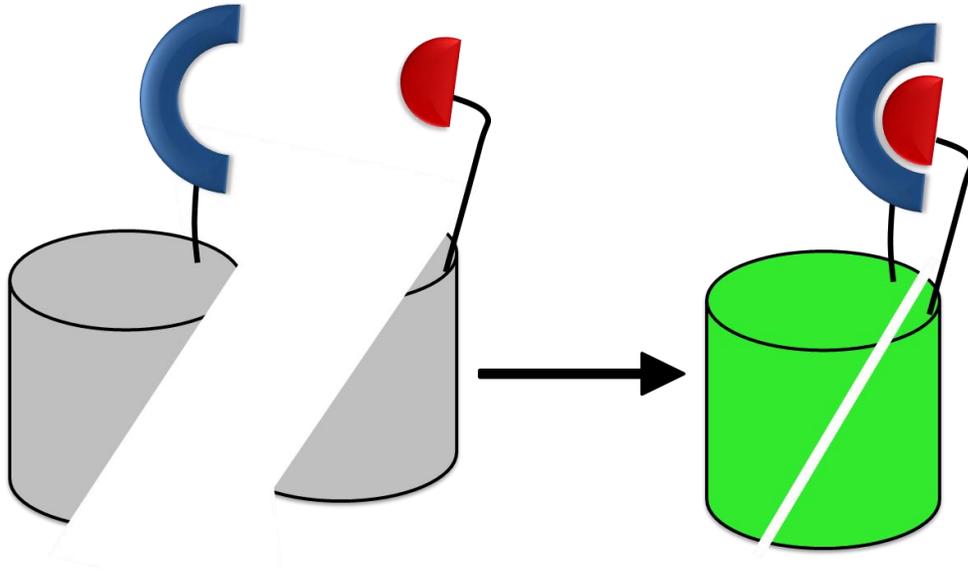


Figure 1-4 Cartoon representation of BiFC approach. Complementation of split FP fragments occurs when corresponding polypeptides, shown in blue and red, interact.

1.3.3 Single FP-based biosensors

A single FP-based biosensor utilizes only one fluorescent protein to generate a response to a given analyte. The single FP-based biosensors generally have the advantage of high signal-to-noise ratio. These biosensors have proven useful for the detection of small molecules such as cGMP, zinc ion, mercury ion and the investigation of small molecule messenger dynamics.^{20,33–35} The mode of action for single FP biosensors may be intrinsic or extrinsic³⁶.

Intrinsic biosensors rely on the inherent sensitivity of certain FPs to specific analytes (like Cl^- , nitrate) or to change in microenvironment.^{20,37}

Accordingly, avGFP variants have been used as intrinsic biosensors of pH

and halides.^{38,39} However, the range of target specificity that can be engineered or possessed by single FPs is limited.

Extrinsic biosensors are designed by fusing an FP and an extrinsic MRE in such a way that ligand binding to the MRE causes a change in the fluorescence profile of the FP.²⁰ Extrinsic biosensors can be constructed for a wide variety of targets and are usually sensitive to MRE conformational changes.¹ For small molecule sensing, extrinsic biosensors are promising templates to achieve versatility in the design strategy.

A number of design strategies have been employed for development of extrinsic biosensors. Much of this work has been driven by the development of biosensors for Ca^{2+} , as described in **Sections 1.3.3.1** through **1.3.3.6**. The range of strategies illustrates the efforts invested in designing the biosensors and success achieved in setting up the general principles for structural design. However, focus is needed to establish one universal strategy that could be adapted for a wide variety of biosensor specificities.

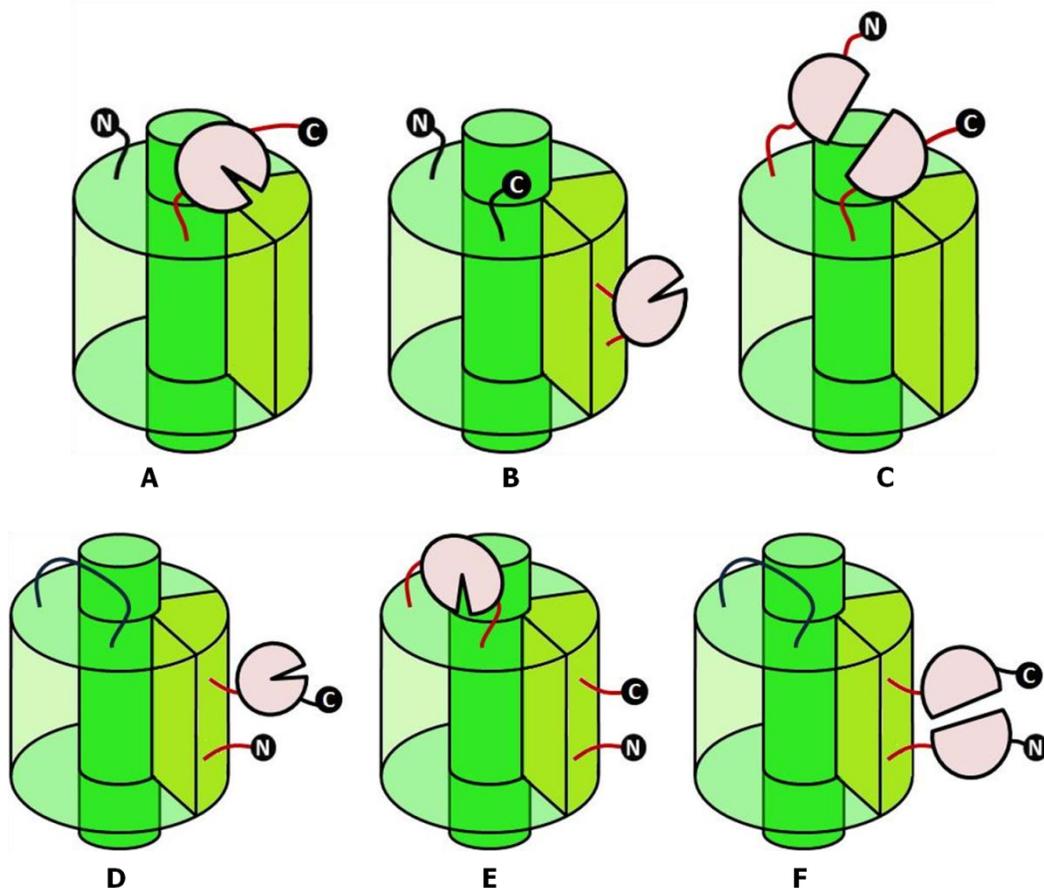


Figure 1-5 Cartoon illustration of construction approaches of single FP-based biosensors. Cylindrical slice in yellow is depicting a β -strand having circular permutation site. Cartoon in pink represents the MRE. A represents tandem fusion of MRE to FP; B represents insertion of MRE in an FP; C represents insertion of an FP in an MRE; D represents tandem fusion of MRE to a circularly permuted (cp) FP; E represents insertion of an MRE in cpFP; F represents insertion of cpFP in MRE. Figure 1-5 is adapted from Baird *et al.*¹²

1.3.3.1 Tandem fusion of MRE to FP

Fusion of the MRE to one of the termini of the FP is generally assumed to minimize potential destabilization of the FP compared to

insertion into an interior location or secondary structural element.⁴⁰ However, in this arrangement the binding event is typically not efficiently communicated to the FP chromophore given that there is little to no conformational change in the chromophore vicinity (**Figure 1-5A**). This approach is commonly employed in constructing fluorescent reporters or FRET-based indicators by fusing the MRE with acceptor and donor FP in a tandem fashion (**Table 1-1a**) as mentioned earlier in **Section 1.3.1**.

1.3.3.2 Insertion of an MRE into a FP

Insertion of MRE in close proximity to the chromophore of a FP is an effective strategy to improve the communication of conformational changes between MREs and FPs. Camgaroo type Ca^{2+} indicators are the classic example of this class, with calmodulin (CaM) inserted at position 145 of FP (**Figure 1-5B**).¹²

Calmodulin (CaM) is a ubiquitous eukaryotic Ca^{2+} -binding protein, capable of reversibly binding Ca^{2+} . In the Ca^{2+} bound state, it engages a peptide called M13 leading to a change in conformation.⁴¹ The binding domains of all Ca^{2+} biosensors discussed in this thesis are derived from this CaM-M13 system.

1.3.3.3 Insertion of FP in RE

Insertion of FPs in close proximity to the active site of an MRE can also be employed in similar design as above (**Figure 1-5C**). This strategy can be employed if an appropriate insertion site in the MRE has previously been identified. Covalently fused proteins with small linkers facilitate mutual communication.⁴²

1.3.3.4 Tandem fusion of MRE to circularly permuted FP

Circular permutation (cp) is a rearrangement of a protein sequence, such that the native amino and carboxyl termini are genetically joined with a short polypeptide spacer and new termini are introduced at another place on the protein surface as illustrated in **Figure 1-6**.⁴³ Cp can be used to identify sites that are tolerant to insertion of other domains. Also the circularly permuted proteins differ from the non-permuted variant in terms of the spatial relationship between their termini and their chromophore, and thus their response differs from those biosensors described in **Section 1.3.3.2**.¹² The fusion of an MRE with the new termini of a cpFP is represented in **Figure 1-5D**. Single FP-based cGMP (cyclic 3',5'-guanosine monophosphate) indicators called FlincGs (fluorescent indicators of cGMP) were engineered with this approach.³³ cGMP binding region of PKGs (cGMP-dependent protein kinases) was fused to N-terminal of cpEGFP to engineer FlincGs.³³

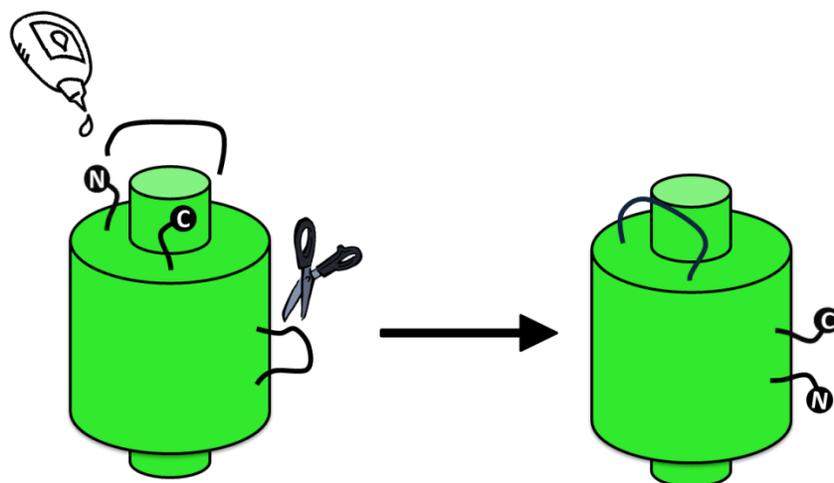


Figure 1-6 Cartoon illustration of circular permutation. The cartoon on left represents native FP with termini on top. Cartoon on right represents the cpFP with new termini at the side and the original termini are joined by a linker peptide. In practice, this conversion is achieved through the use of multiple molecular biology procedures.

1.3.3.5 Insertion of MRE in circularly permuted FP

The insertion of MREs in cpFPs is, in principle, a possible way of engineering biosensors (**Figure 1-5E**), though it is not clear if this strategy would have any advantages relative to the proven approaches described here.¹² To the best of our knowledge, this approach has not been previously reported.

1.3.3.6 Insertion of circularly permuted FP in MRE

Insertion of cpFP in the MRE is of particular significance when dealing with MRE of large size or one having multiple domains.¹² In this form even

small changes in the proximity of chromophore region and in the sequence around new termini are efficiently transmitted to enhance the FP's change in fluorescence.¹² CpFPs can be inserted either in native MREs or in permuted MREs near active sites (**Figure 1-5F**). G-CaMP type Ca^{2+} indicators engineered using this approach proved to be a great improvement over previous constructs.⁴⁴ These indicators transmit the interaction between CaM and M13 mediated by binding of Ca^{2+} to CaM into a conformational change.⁴⁵ The conformational change is then directly translated into a fluorescence change of the fused cpFP.⁴⁵ The mechanism underlying conformational change in G-CaMP will be discussed further in **Section 1.4.2**.

1.3.3.7 Summary of single FP-based design strategy

The G-CaMP type Ca^{2+} biosensors are the archetypes for a single FP-based design strategy.⁴⁴ Extensive optimization efforts have led to continued improvements in the G-CaMP type biosensors and the latest versions exhibit improved signal changes and affinities.⁴⁶⁻⁴⁹ If there is one area where these design strategies disappoint, it is the restriction of analyte sensitivity to Ca^{2+} and eventually it is the ease by which they can be developed for the detection of analytes other than Ca^{2+} .

It was long anticipated that efforts to adapt the single FP-type design to analytes other than Ca^{2+} would be facilitated with an atomic

resolution view of G-CaMP or pericam.^{44,50} Fortunately, the X-ray crystal structure of an optimized G-CaMP variant has now been reported.⁴⁵ One striking aspect of the molecular structure of G-CaMP is an almost complete replacement of β -strand 7 by an α -helical region of CaM.⁴⁵ It implies that the mechanism underlying the modulation of fluorescence profile in G-CaMP is closely associated with the particular structure of its MRE, CaM. This represents a mechanistic caveat that hinders the development of analogous biosensors for non-Ca²⁺ analytes. This reinforces the need for design strategies whose action and improvement is guided by unified mechanistic principles intrinsic to FP and not specific to any single type of MRE.

1.4 Overview of existing switching mechanisms for development of switch-based FP

Large-scale protein conformational changes induced by a binding event in a biological milieu can be thought of as conformational switches. Conformational changes are primarily mediated by changes in multiple weak interactions including van der Waals interactions, dipole-dipole interactions and hydrogen-bonds associated with the binding event.⁵¹ Protein conformational switches are proteins that undergo a substantial change in their 3-dimensional conformation in response to an input signal, such as binding of ligand (*e.g.*, binding to Ca²⁺), a change in environment

(*e.g.*, pH), or covalent modification (*e.g.*, a posttranslational modification such as phosphorylation).⁴⁰ Indeed, many of the proteins and enzymes that regulate biological activities in live cells can be thought of as switches that are modulated in an allosteric fashion.

The value of versatile and generic molecular engineering systems is becoming increasingly evident.^{1,52} The possibility to engineer a conformational switch in proteins, which may or may not possess any conformational change in its native forms, was expected to facilitate the development of generic MREs and FP scaffolds. Attempts were thus made to better understand the mechanism of action in naturally occurring switches to guide generic engineering of biosensors.^{40,42} The significant benefits offered by conformational switches for design of biosensors involve robustness and high specificity; time efficient and reversible action; versatility and adaptability for application in any protein; and a tunable equilibrium state with respect to target concentration providing quantitative information about the target.⁵¹

Understanding of natural protein switches may guide the development of biosensors that can transduce the conformation of a switch (on/off) into a measurable output signal. Based on the foundation of existing natural switches and associated mechanism, Loh *et al.* introduced artificial switches into proteins.⁵³ In **Sections 1.4.1** and

1.4.2, the mechanisms of naturally occurring conformational switches are examined.

1.4.1 Rigid domain movement

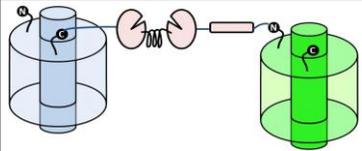
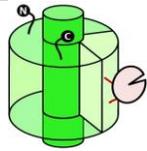
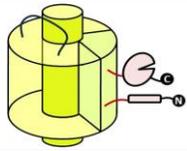
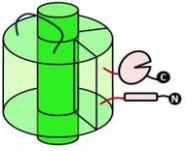
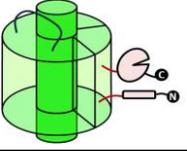
Displacement of two or more protein domains with respect to each other is often termed as rigid domain movement. Open-to-closed or hinge-type domain movements constitute this class of conformational switches.⁴⁰ Multidomain proteins that undergo rigid body movements are highly useful for designing biosensors as well as modulating the function of proteins. A class of bacterial receptors called periplasmic binding proteins (PBPs) exhibits these types of movements. PBPs are composed of two domains that tend to change conformation from an open-to-closed state upon ligand binding.⁵⁴ Consequently, PBPs have been used in FRET biosensor designs in which ligand binding triggers movement of the two fused FPs into close proximity (**Figure 1-4b**).⁵⁵

1.4.2 Limited structural rearrangement

Proteins that exhibit limited structural rearrangement represent another class of natural switches. These switches undergo localized conformational changes upon encountering the analyte. The Ca^{2+} -dependent interaction of CaM and M13 could be considered an example of this type of conformational change.¹¹ CaM-M13 assembly rearranges the

structure from a dumb-bell shape to compact globular shape upon Ca^{2+} binding which causes an efficient fluorescence change.⁵⁶

Table 1-1 Comparison of various designs of Ca^{2+} indicators.

S.No.	Class of indicator	Structure	Type of response	Underlying Mechanism
a	Cameleon ¹¹		Emission ratiometric	Binding induced conformational change in recognition element
b	Camgaroo ¹²		Intensiometric	Binding induced conformational change in FP
c	Pericam ⁵⁰		Flash/Inverse= Intensiometric Ratiometric = Excitation	Binding induced conformational change in RE and FP
d	GCaMP ^{44, 47, 48}		Based on mutations. May be Intensiometric / Ratiometric	Binding induced conformational change in RE and FP
e	GECO ⁴⁹		Based on mutations. May be Intensiometric / Ratiometric	Binding induced conformational change in RE and FP

The limited structural rearrangement-type conformational change of CaM-M13 system has served as design principle for most Ca^{2+} indicators including cameleon, pericam, G-CaMP and green fluorescent genetically encoded Ca^{2+} indicators for optical imaging (GECOs).^{11,12,23,44,49,50} For example, camgaroo, the first single FP-based Ca^{2+} indicator which has CaM inserted at YFP 145 position, utilizes the

limited structural rearrangement of the CaM-M13 system.¹² The change in the molecular environment of the chromophore induced by Ca²⁺ binding to CaM is assumed to be responsible for modulation of the fluorescent signal in camgaroo.¹² All FP-based Ca²⁺ indicators utilize the mechanism of limited structural rearrangement in certain way, as summarized in **Table 1-1**. It is noteworthy that if one can learn to introduce limited structural rearrangement in any protein, then development of new functional switches will be possible.

1.4.3 Global fold switching

Global fold switching is the third category of switches. It involves conversion between two unrelated folds of the peptide. Metamorphic proteins are representative of this class of switching. On encountering the ligand, these proteins unfold completely; ultimately refolding into an alternative structure.⁴⁰ Only a few instances of naturally occurring proteins are known to exhibit this type of switching including lymphotactin, Mad2 spindle checkpoint protein, and chloride intracellular channel 1 (CLIC1) protein.⁵⁷⁻⁵⁹ Lymphotactin exists in two equilibrating forms; a monomeric chemokine fold and a novel dimeric β -sandwich fold.⁶⁰ The equilibrium tends to shift completely from one form to the other on varying salt and temperature conditions.⁶⁰

In order to explore this strategy artificially, Orban *et al.* rationally designed a G_A/G_B system (**Figure 1-7**), representing 2 binding domains of *Streptococcus* protein G.⁵⁷ They demonstrated that G_B (IgG-binding), a 4β -strand + 1 α -helix fold, can be transformed into G_A (an albumin-binding), a 3α -helix fold, via a mutational pathway in which neither function nor native structure is completely lost.⁵⁷

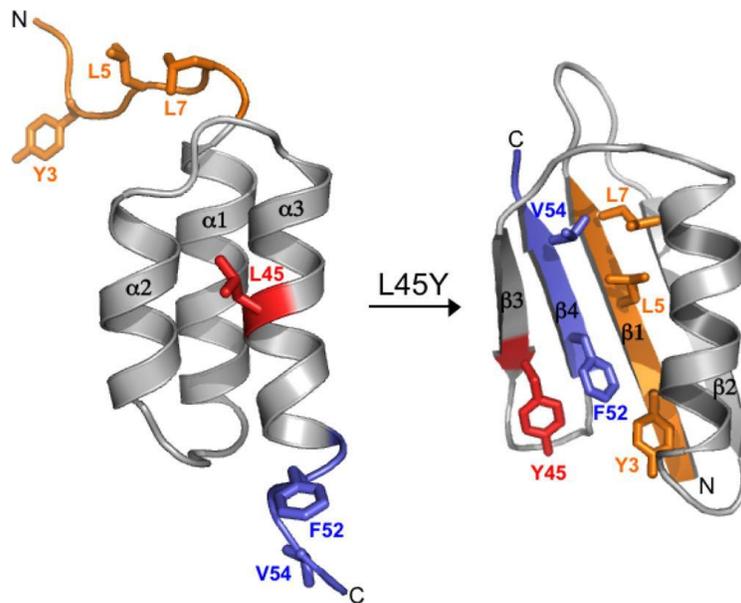


Figure 1-7 Global Switching. Alternate conformations of G_A ($3\times\alpha$) and G_B ($4\times\beta+1\times\alpha$) owing to single mutation L45Y. Figure is taken from Alexander *et al.*⁵⁷

This example suggests that artificial design of global switch-based sensor is theoretically plausible. If one can insert a partially overlapping sequence of residues with the specific mutation to induce global switch,

ligand bound/unbound forms would differ entirely in structure and would extensively affect fluorescence properties.

1.4.4 Folding-unfolding switches

Folding-unfolding switches represent the fourth class of switching mechanism. It is considered as one of the most effective switching mechanisms involving change in structure, dynamics, and charge distribution.⁴⁰ Such switches exist in equilibrium between an intrinsically flexible or locally disordered state of a protein and a well-structured fold state of the same protein.⁶⁰ In natural proteins, unfolded regions can often engage in multiple protein-protein interactions. Disordered and unfolded protein ensembles are suggested to be more capable of responding to changes in their environments and are better suited for signaling events.⁶¹ The probable benefits include increasing binding promiscuity, increasing the rate of macromolecular association, and offering multiple sites for post-translational modification.^{40,62-64}

Unfolded regions also known as disordered regions are surprisingly common in mammalian proteins. In fact, about 75% of mammalian signaling proteins are predicted to contain long disordered regions (>30 residues) and about half of all mammalian proteins are predicted to contain long disordered regions.⁶¹ Two unfolded segments might look equally disordered but their fold propensity may differ entirely.⁶⁰ The

energy accessibility for switching between folded/unfolded regions is assumed to be low accounting for quick and easy signal transduction.⁶⁰ The fold preference of unfolded regions thus manifests itself in a specific structural context.⁶¹ For instance, the four to eight amino acid linkers in CaM form a helix in the crystal structure, but is mostly unstructured in solution, indicating that functionally important disordered regions can be as short as a few residues (**Figure 1-8**).^{65,66}

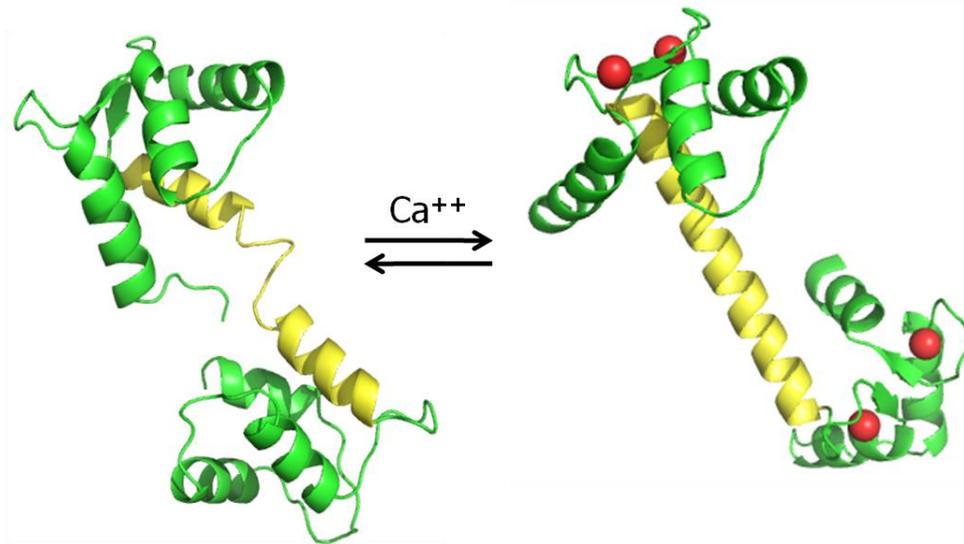


Figure 1-8 Folding-unfolding equilibrium of CaM. Intrinsically disordered region of CaM (left side, in yellow) in absence of Ca^{2+} ; converts to structured region on ligand binding (right side, in yellow). Red spheres indicate Ca^{2+} (PDB IDs: with Ca^{2+} - 3CLN; without Ca^{2+} 1CFD).¹⁴

1.5 Employment of protein-based switches in biosensor design

Based on the various mechanisms by which natural protein switches are known to operate, we can now consider a variety of potential

designs that could be used to achieve our goal of a generic biosensor. The first approach is to use a naturally occurring switch and redesign its binding site to support recognition of the desired target analyte. However, this involves introduction of selectivity, adjustment of affinity in complex chemical environment, and might require tedious evolution steps.⁵¹ The second approach is to harness existing proteins (either MREs, or FPs, or both) and re-engineer them to undergo switching upon target binding by either of the mechanisms discussed in **Section 1.3**. Inserting an additional sequence for global structure switch of chimeric biosensing protein or inserting an additional unfolded region to introduce binding induced folding in a highly cooperative two state process are possibly the best ways of designing generalized switch based biosensors.

Based on the foundation of naturally existing mechanism of switches, Loh *et al.* recently proposed a generic design strategy for introducing a conformational switch into any protein of interest.⁵³ The “alternate frame folding” (AFF) design duplicates a portion of the amino acid sequence, creating an additional “frame” of folding.⁵³ One frame represents the wild-type sequence that adopts the native fold structure. The second frame represents the alternate fold (AF), which effectively yields a circularly permuted protein. Both the frames compete for a shared

site in the globular protein and undergo folding in a mutually exclusive fashion (**Figure 1-9**).⁴⁰

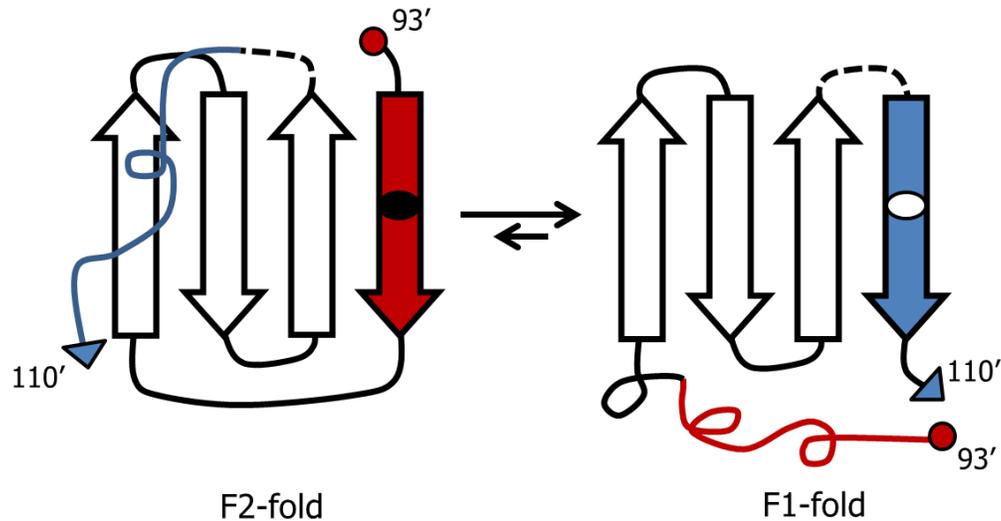


Figure 1-9 Cartoon representation of barnase-AFF developed by Loh *et al.*⁶⁷ Blue and red strands represent a duplicated sequence competing to interact with same strand. At any one time, one copy is folded and the other is unfolded. Both are representing protein ends (C-terminal in blue and N-terminal in red) as the construct is circularly permuted. **Figure** adapted from Mitrea *et al.*⁶⁷

Exploiting the AFF strategy to modulate the function of a protein requires a design strategy with 3 distinct steps. The first step is the identification of single amino acid position in the duplicated peptide region that is crucial for the function of the protein.⁵³ The second step is to ensure the ability of circular permutation in the binding protein.⁴² The third step involves inserting a partial sequence overlap with the mutated crucial amino acid in the circularly permuted termini of protein.⁵³ In the event of binding of ligand, the binding energy assisted folding of the less

stable frame (alternate fold usually) is triggered and accompanied with unfolding of the more stable fold (wild-type usually). Loh *et al.*, verified that in presence of ligand, the protein with an artificial frame will undergo structural remodeling to adjust the equilibrium between the wild-type and the alternate fold.⁵³

For an artificial switch to be effective, reversibility and inherent cooperativity are prerequisites. Furthermore, for constructs to be reversible, the wild-type fold and the AF fold should possess similar stabilities (usually, wild-type is the more stable one and the AF is the less stable one).⁵³ Stability of wild-type and AF fold equilibrium depends on temperature, pH and ligand binding.⁴² Under conditions of constant temperature and pH, the ligand binding solely dictates the thermodynamic stability and folding kinetics.⁴² Change in equilibrium could be linked to change in functional or detectable output implying the cooperativity of the switch.⁵²

1.6 Objectives of the thesis

In order to generalize the design strategy of biosensors for detection of varied biomolecules, a universal FP-based biosensor platform is needed. The recent report of the AFF strategy for introducing conformational switch in a binding domain inspired us to apply a similar approach to

avGFP. Our aim is to develop a biosensing platform in which modulation of spectral readout in response to the input signal will be highly predictable and comprehensible. Accordingly, we propose that a Tandem Alternate Frame Switching (TAFS) strategy could be an effective design for a universal FP-based biosensing platform.

In Chapter 2, we attempt to establish a foundation for the development and optimization of the TAFS mechanism as applied to avGFP. As the name suggests, we introduce into avGFP a “switch” that is composed of a repeated secondary structural element. The TAFS strategy involves limited remodeling of the FP structure through competitive binding of the repeated secondary structural element to the remainder of the protein fold. We present evidence in support of rational tunability of a TAFS based FP for any target given that the mechanism of action is rationally designed. We will demonstrate proof-of-concept applications of this strategy for creating biosensors for detection of the biomolecules cyclic adenosine monophosphate (cAMP) and Ca^{2+} .

In Chapter 3, we discuss the design and screening of three libraries of cpmCherry as a first step towards implementing the TAFS strategy with a red fluorescent protein. Future directions and suggestions for optimization of this strategy along with other potential applications will be discussed in Chapter 4.

Chapter 2:

Development and characterization of
an avGFP variant capable of tandem
alternate frame switching

2.1 Introduction

The ability to visualize various biological activities at subcellular resolution has been facilitated by combined advances in the fields of protein conformational switches and fluorescent protein (FP) engineering.^{36,40} To convert protein conformational changes into spectral signals that can be visualized using modern microscopy techniques, it is necessary to use protein engineering to introduce a fluorescent reporter that changes its fluorescence in response to the conformational change. For many live cell imaging applications, FPs are useful fluorescent reporters.⁶⁸ The primary advantage of using FPs is the biosensing construct (*i.e.*, the hybrid of the conformational switch and the FP) is proteinaceous and genetically encodable.⁴ Accordingly, the gene for the biosensing construct can be introduced into cells, tissues, or whole organisms and synthesized *in situ* with minimal perturbation to normal physiology.

As discussed in **Section 1.3**, there are a variety of protein switching mechanisms known, but only a subset of them have been applied to the construction of FP-based biosensors. In this chapter we describe our efforts to apply the AFF strategy to avGFP to generate a reporter module with a unified mechanism of action.

The idea of applying the AFF strategy to avGFP was inspired by previous reports that a specific mutation in the 7th β -strand of avGFP

produced a substantial change in the protein absorbance spectrum.⁶⁹ Wild-type avGFP has a histidine residue at position 148 (His148), which is located on the 7th β -strand as shown in **Figure 2-1**.⁷ His148 hydrogen bonds to the chromophore, stabilizing it in the anionic phenolate state.⁶⁹ In the phenolate state, the avGFP spectrum is defined by a strong absorbance peak with a maximum at 488 nm.¹⁵ This effect is particularly pronounced in the Ser65Thr variant of avGFP, where the equilibrium between the neutral phenol and anionic phenolate forms is heavily shifted to the latter.¹⁵ Mutation of His148 to aspartate (His148Asp) effectively destabilizes the phenolate form of the chromophore and shifts this equilibrium to the neutral phenol.^{16,70,71} Correspondingly, the absorbance maximum shifts to 400 nm. We expected that two copies of β -strand 7, one with His148 and one with His148Asp, could serve as the competing secondary structural elements in an AFF-like protein engineering strategy.⁵³

In the reported examples of AFF, the two competing secondary structural elements are positioned at the N- and C-termini of the protein (See **Figure 1-10, Section 1.5.1**).^{53,67,72} Such a strategy could be applied to avGFP by duplication of the first β -strand at the C-terminus or duplication of the 11th β -strand at the N-terminus. We have chosen to pursue a modified version of this strategy in which an internal secondary

structure element (*i.e.*, β -strand 7) is duplicated in tandem (**Figure 2-2**). To distinguish this strategy from the previously reported examples of AFF, we refer to it as Tandem Alternate Frame Switching (TAFS).

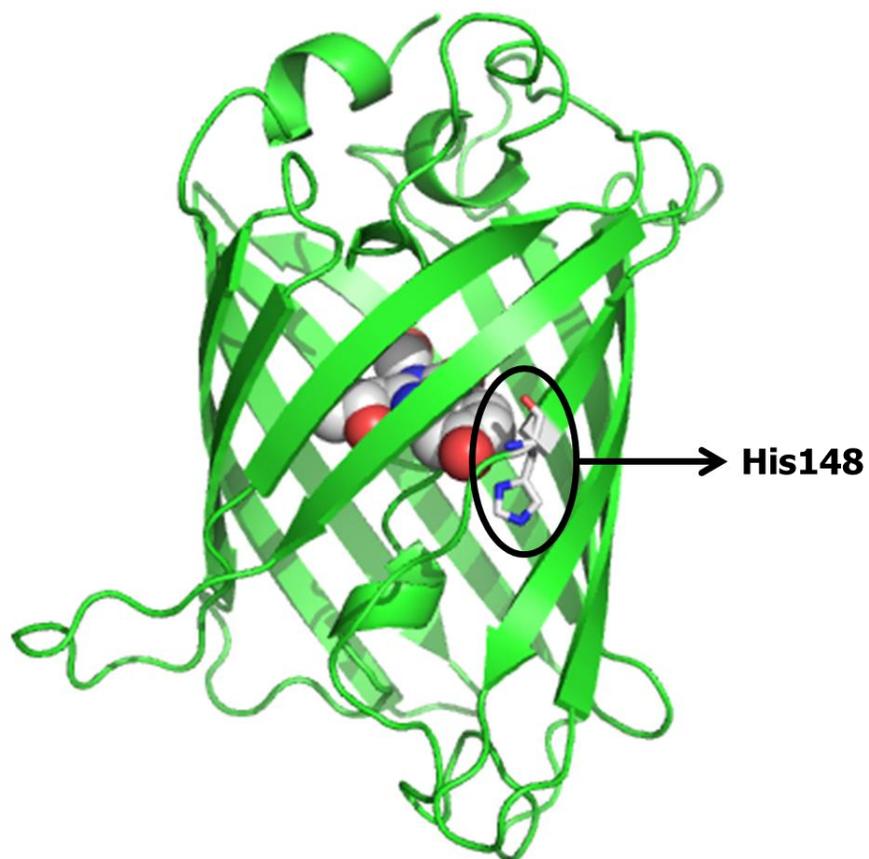


Figure 2-1 Cartoon representation of the avGFP structure. The His148 residue is shown in stick representation to illustrate its position relative to the chromophore (sphere representation). PDB ID: 1EMA.¹⁴

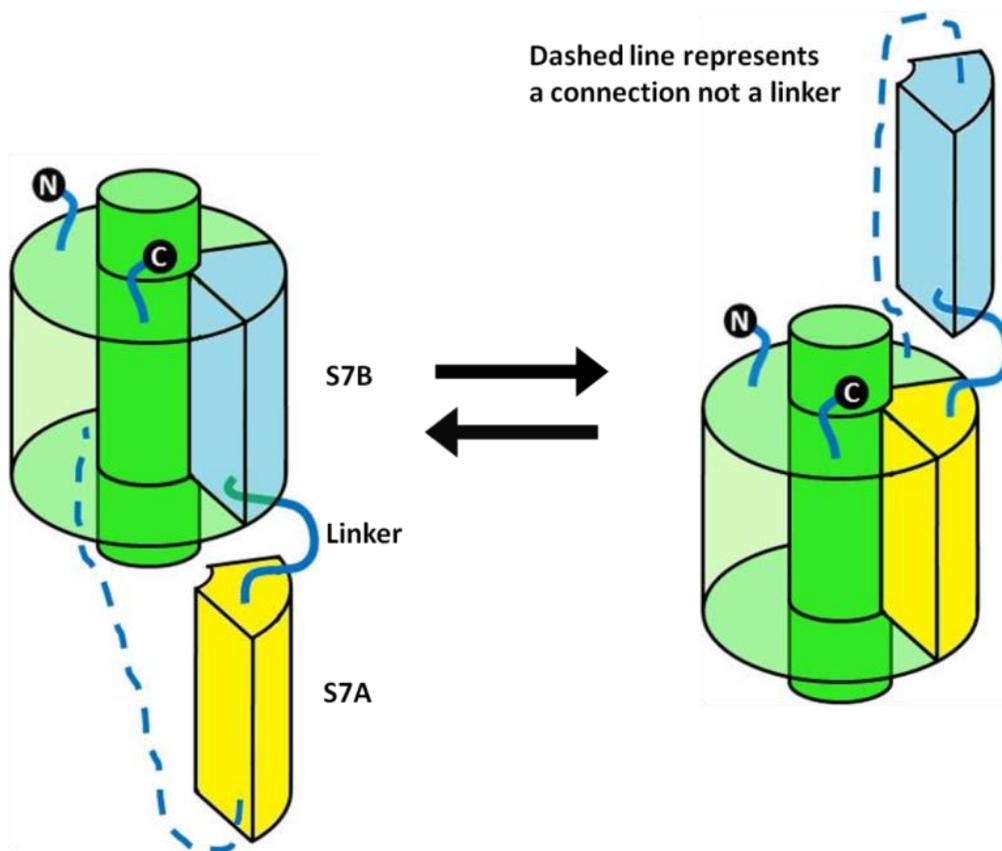


Figure 2-2 Cartoon representation of predicted conformational equilibrium between alternate frame folding involving strand 7A (S7A) & strand 7B (S7B). Different colours of the cylindrical slice are used to depict non-duplicated partially overlapping strand. The central cylinder is representing the α -helix region that contains the central chromophore. The cartoon style is adapted from Kent *et al.*⁷³

Application of the TAFS strategy to avGFP involves the genetic duplication of β -strand 7. We refer to the first (*i.e.*, N-terminal) copy as “Strand 7 copy A” (S7A) and the second copy as “Strand 7 copy B” (S7B) (**Figure 2-2**). Joining the two strands is a linker region that, at the DNA level, contains restriction sites that allow for eventual insertion of conformational switch domains between the repeated β -strands.

The *a priori* prediction of the effect of an inserted conformational switch domain on the equilibrium between bound S7A and bound S7B is difficult. Fortunately, the TAFS strategy could enable the equilibrium between the two bound states to be rationally tuned by the introduction of stabilizing or destabilizing modifications of either the first or second strand.⁴² Examples of such modifications could include mutation to amino acids with low β -sheet propensity or partial truncation of the strand.

Hypothetically, an inserted conformational switch domain in its apo state could shift the equilibrium far towards the state in which S7A was bound to the protein. In the best-case scenario, an induced change in the conformational switch domain (*e.g.*, resulting from ligand binding) would shift the equilibrium towards the state in which S7B was bound to the protein. The corresponding change in absorbance (and fluorescence) would make such a construct an effective sensor. In the worst-case scenario, an induced change in the conformational switch domain would shift the equilibrium even further towards the state in which S7A was bound to the protein. For example, if the equilibrium started out with 99% of the protein in the S7A bound state (1% S7B bound), and this shifted to 99.9% of the protein in the S7A bound state (0.1% S7B bound), the change in absorbance or fluorescence would only be $\sim 1\%$ of the total

signal intensity. Such a change would, at best, be barely indistinguishable from the noise.

To convert such a construct into a useful sensor, it should be possible to use genetic modification to destabilize the S7A bound state and stabilize the S7B bound state. Rational use of such modifications could tune the equilibrium to a point where the conformational switch produces a substantial and measurable change in fluorescence. In this modified case, there is now a ~10% change in the absorbance and fluorescence signal, a change that could be practically useful for live cell biosensing applications.

The potentially universal nature of the TAFS strategy, along with the built in mechanism for tuning the response, makes this a very promising strategy for the construction of FP-based sensor constructs. In this Chapter we describe our efforts to validate and apply the TAFS strategy to FPs.

2.2 Materials and methods

2.2.1 General methods and materials

Primers were designed and analyzed using OligoAnalyzer3.1 (www.idtdna.com) and ordered from Integrated DNA Technology. dNTPs were purchased from Invitrogen and *Pfu* Polymerase from Fermentas.

Unless otherwise indicated, polymerase chain reactions (PCRs) were performed following the supplied Fermentas protocols in Bio-Rad thermal cyclers. DNA fragments of interest were separated from crude PCR mixtures using agarose gel electrophoresis and stained by ethidium bromide. Subsequently, DNA bands were visualized on a trans-illuminator under UV excitation and extracted from gel slices using GeneJET™ gel extraction kit (Thermo Scientific).

DNA digestion was carried out with FastDigest® restriction endonucleases (Fermentas) at 37°C for 2-6 hours. Subsequently, digested DNA was electrophoresed and extracted as described above. The concentration of the digested DNA was determined using a Nanodrop1000 spectrophotometer (Thermo Scientific). DNA ligation was performed with T4 DNA ligase (Invitrogen) in reaction mixture with vector-to-insert ratio of 1:6 at 16°C overnight.

Electrocompetent *E. coli* strain JM109(DE3) (Promega) was used for plasmid propagation and recombinant protein production. Electroporation was performed using 0.2 cm MicroPulser cuvettes (Bio-Rad) and MicroPulser electroporator (Bio-Rad). The transformed cells were plated on LB-agar media plates containing 0.02% ampicillin. Single colonies were picked and cultured in liquid LB media while shaking (200-210 rpm) at 28°C.

Plasmid DNA was isolated using in-house plasmid miniprep protocol described in Appendix IX. DNA sequencing reactions were performed using BigDye® (Applied Biosciences) terminator cycle sequencing kits. Ethanol precipitation was performed to clean up the sequencing reactions before they were submitted to the Molecular Biology Services Unit at the University of Alberta (MBSU) for analysis. Sequencing chromatograms were visualized using Chromas Lite 2.01 (developed by Technelysium) and sequence-alignment was performed using ClustalW (European Bioinformatics Institute).

All chemicals and reagents were purchased from Fisher Scientific, unless otherwise specified.

2.2.2 Construction of an avGFP encoding plasmid

A non-cp GFP-encoding gene was reassembled from G-GECO1.0, (evolved for high folding efficiency and increased dynamic range).⁴⁹ Primers RC001 and 5'BH1(GAT)kzkgfp1; RC002 and GFP238EcoR13' were used to PCR amplify the segments 1-153 and 144-238 respectively using KOD-Plus-DNA polymerase (TOYOBO). The PCR products were digested and sequential ligation in pRSET B plasmid. This reassembled avGFP variant contained an in-line duplication of β -strand 7 (residues N144-Tyr145-N146-S147-H148-N149-V150-Y151-I152-M153K) connected by a 7-residue linker (TSSGGTS, including Kpn1 and Xho1 restriction sites).

Dr. Campbell created many of these constructs in the Fall of 2010 during his Sabbatical research in the lab of Dr. Takeharu Nagai. Plasmids were shipped to the University of Alberta as dried spots on filter paper. The labeled spots were then cut and suspended in 50 µl deionized (dI) water in an Eppendorf tube. Centrifugation of tubes then yielded the plasmid in the supernatant. Transformation of plasmid in JM109(DE3) was performed and recombinant proteins were expressed.

2.2.3 Construction of control protein constructs

PCR amplification of the C-terminal section of reassembled gene was done with primers RB7_ECORI3_R and RB7_KPN1_F (Schematic representation in Appendix II). The amplified gene was then sub-cloned in digested parent plasmid to generate the gene for the control protein with S7B copy removed (will be referred to as BR, further in the thesis). Similarly, Armv_Xho1_r and BH1b7F primers were used to PCR amplify the N-terminal section and sub-cloned in corresponding parent plasmid to generate AR (S7A removed) control protein.

2.2.4 Protein expression and purification

pRSET B (Invitrogen) was used for recombinant protein production. Nucleotide sequences encoding proteins of interest were placed downstream of the T7 promoter and in-frame with the 5' nucleotide

sequence encoding a hexa-histidine (6xHis) tag. Electrocompetent *E. Coli* JM109(DE3) was transformed with the plasmid of interest and over-expression of the 6xHis-tagged protein of interest was achieved by virtue of leaky expression without addition of IPTG. A transformant colony from a freshly plated bacterial culture was used to inoculate 25 ml SOB media supplemented with ampicillin to final concentrations of 0.04%. The culture was incubated for 48 hrs at 24°C or 28°C and 210 rpm shaking speed (Innova 4330 shaker, New Brunswick Scientific).

Bacterial cells were pelleted by centrifugation at 5500 rpm in a Beckman A-10 rotor for 10 min at 4 °C. Pellets were suspended in B-PER for lysis and agitated for ~10 min. To separate proteins from cell debris, centrifugation for 15 min at 10,000 rpm was performed using a Beckmann Rotor centrifuge. Equivalent volume of binding buffer (50 mM Tris-HCl, 8 gL⁻¹ sodium chloride, 0.2 gL⁻¹ potassium chloride, 10 mM imidazole, pH 7.1) and Ni-NTA beads (Qiagen) were added to the cleared cell lysate to capture the 6xHis-tagged proteins. After 60 min of incubation at 4 °C, the beads were collected in disposable plastic Polyprep columns (BioRad) using a vacuum manifold. Beads were extensively washed with wash buffer (50 mM Tris-HCl, 8 gL⁻¹ sodium chloride, 0.2 gL⁻¹ potassium chloride, 10 mM imidazole, pH 7.1) and the recombinant protein was subsequently eluted with elution buffer (50 mM Tris-HCl, 8 gL⁻¹ sodium

chloride, 0.2 gL⁻¹ potassium chloride, 500 mM imidazole, pH 7.1). To minimize proteolytic degradation of recombinant proteins, all steps were performed at 4 °C.

2.2.5 *In vitro* spectroscopic analysis of protein constructs

In vitro absorbance spectra were acquired in quartz cuvettes in a UV-vis spectrophotometer (Beckman Coulter). Absorbance of elution buffer without protein was used as the blank. Absorption spectra for each recombinant and control protein were acquired from 250 nm – 650 nm, with 1 nm step and 1200 nm/min scan speed. The spectra were then background corrected and normalized.

2.2.6 Data analysis of protein constructs

The concentration of S7A-bound and S7B-bound FP conformations was determined statistically by analyzing simulated absorption spectra created by summing experimental control protein spectra at various ratios. The simulated curves were fit using the same method used for experimental absorption spectra. The Solver tool of Microsoft Excel was used to perform independent component analysis of normalized control protein spectra to yield best fit to corresponding TAFS-based FP's spectrum. The % ratio of each control protein to yield the best fit of the simulated curve was used to calculate the relative concentrations of

corresponding frame conformations and used to calculate equilibrium constants (K_{eq}) and free energy changes between the two states (ΔG).

To obtain error estimates, an identical independent component analysis was performed on control samples prepared by mixing known proportions of control proteins. The error in predicting the correct proportions of corresponding components was used to yield the equation for error calculation.

2.2.7 Construction of Ca²⁺ and cyclic AMP biosensors

We inserted the binding domain of YC-Nano15 (a gift from Dr. Takeharu Nagai) between the two restriction sites in linker using standard protocols of digestion and ligation.²⁴ Similarly, we inserted the gene for *E. coli* catabolite activator protein (CAP), as a binding domain for cyclic AMP, between the two restriction sites.⁷⁴ Dr. Campbell created these biosensor constructs during the Fall of 2010 during his Sabbatical research in the lab of Dr. Takeharu Nagai.

2.2.8 Characterization of biosensor constructs

2.2.8.1 Calcium biosensor

The *in vitro* fluorescence assay was performed in Corning 384-well plates (COS384fb). Reactions were buffered with 50 mM Tris-HCl, 8 g L⁻¹

sodium chloride, 0.2 g L⁻¹ potassium chloride, 10 mM imidazole, and 10 mM EGTA at pH 7.1. Emission wavelength was set to 525 nm and fluorescence excitation profiles were collected from 300 nm to 500 nm wavelength. Spectra were collected using a TECAN Safire2 microplate. Spectra for the apo state were acquired in Ca²⁺-free buffer. The bound state was characterized with addition of 10 mM CaCl₂ to corresponding buffered solutions of biosensor.

For determination of dissociation constant (K_d) of Ca²⁺-biosensors, Ca²⁺-titration studies were performed. Different concentrations of Ca²⁺-buffer were prepared using reciprocal dilution method. The biosensor constructs were then evaluated for change in excitation ratio from apo state (buffered in 10 mM EGTA, 100 mM KCl, 30 mM MOPS, pH 7.1 and 20 °C) to Ca²⁺ bound state (buffered in 10 mM EGTA, 100 mM KCl, 30 mM MOPS, pH 7.1 and 20 °C; 1 mM Ca²⁺-EGTA and dilutions thereof).

2.2.8.2 Cyclic AMP biosensor

The *in vitro* fluorescence assay was performed in Corning 384-well plates with the TECAN Safire2 microplate reader. The biosensor was buffered with 50 mM Tris-HCl, 8 g L⁻¹ sodium chloride, 0.2 g L⁻¹ potassium chloride, 10 mM imidazole at pH 7.1. For the bound state analysis, 10 mM cyclic

AMP was added to the buffer. The above mentioned procedure was used for spectral characterization.

2.3 Results and discussion

2.3.1 Development and evaluation of TAFS-based avGFP scaffold

To evaluate the impact of tandemly fused non-duplicated strand (as depicted in **Figure 2-3**), we introduced the His148Asp substitution into either copy A or copy B. The His148Asp substitution stabilizes the protonated form of the chromophore and shifts the absorbance maximum from 488 nm (with His148) to 407 nm (with His148Asp).^{16,69,71} We compared the absorbance spectra of the variants ADBH, ADBD, AHBD and AHBH vs. the variants with single copy, as depicted in **Figure 2-4** (where AXBY residues at 148th position of copy A and copy B, respectively). The absorbance profile of each variant was observed to be unique. Additionally, the variants with His148 (abbreviated by H) in copy A showed absorbance peak close to 400 nm and those with Asp148 in copy A (abbreviated by D) close to 480 nm. We speculated that absorbance profile of each variant is symbolic of its conformational equilibrium position.

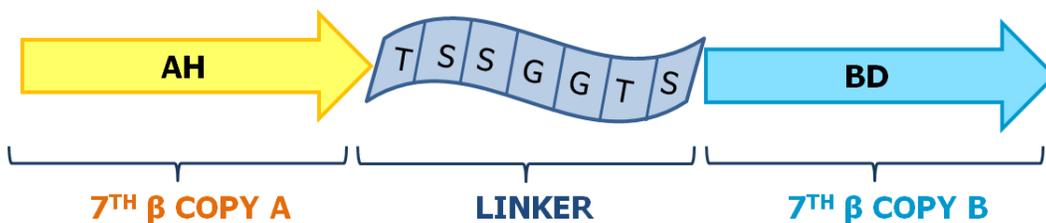


Figure 2-3 Cartoon representation of inserted β -strand 7 (B) with His148Asp and linker succeeding the original β -strand 7 (A) with His148. The whole protein sequence is represented in Appendix I.

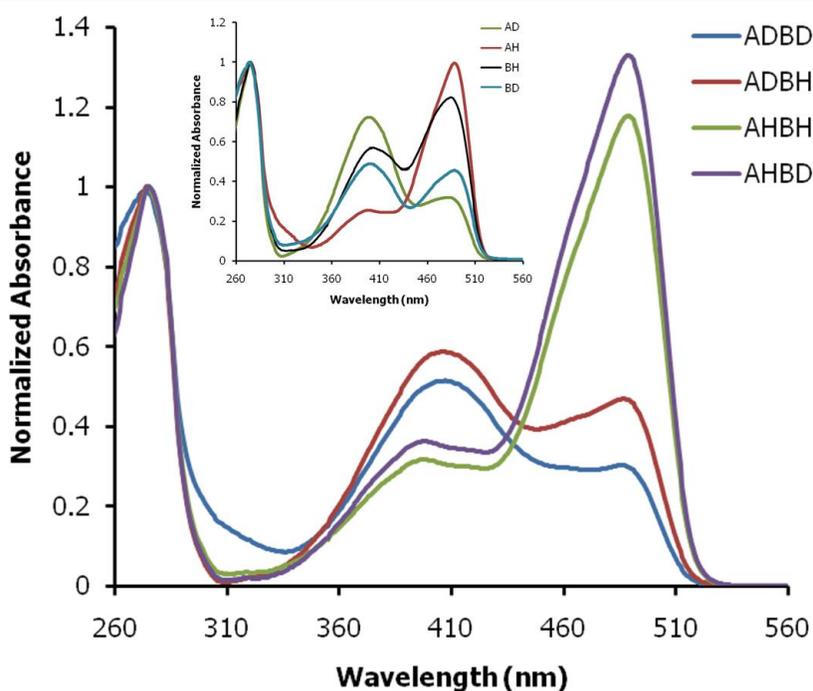


Figure 2-4 Normalized absorbance spectra exhibited by the variants with single mutation at 148. The inset shows the spectra exhibited by control proteins.

Based on the principle of alternate frame folding, at any time, only one copy will interact with the chromophore and integrate in the protein structure as β -strand and the other copy will presumably exist as a disordered loop.⁵³ In essence, the two copies of the β -strand are

competing for a single binding site on the protein. Under these conditions, the resultant absorbance spectra represent the proportional sum of individual spectrum of either copy A or copy B bound at equilibrium.

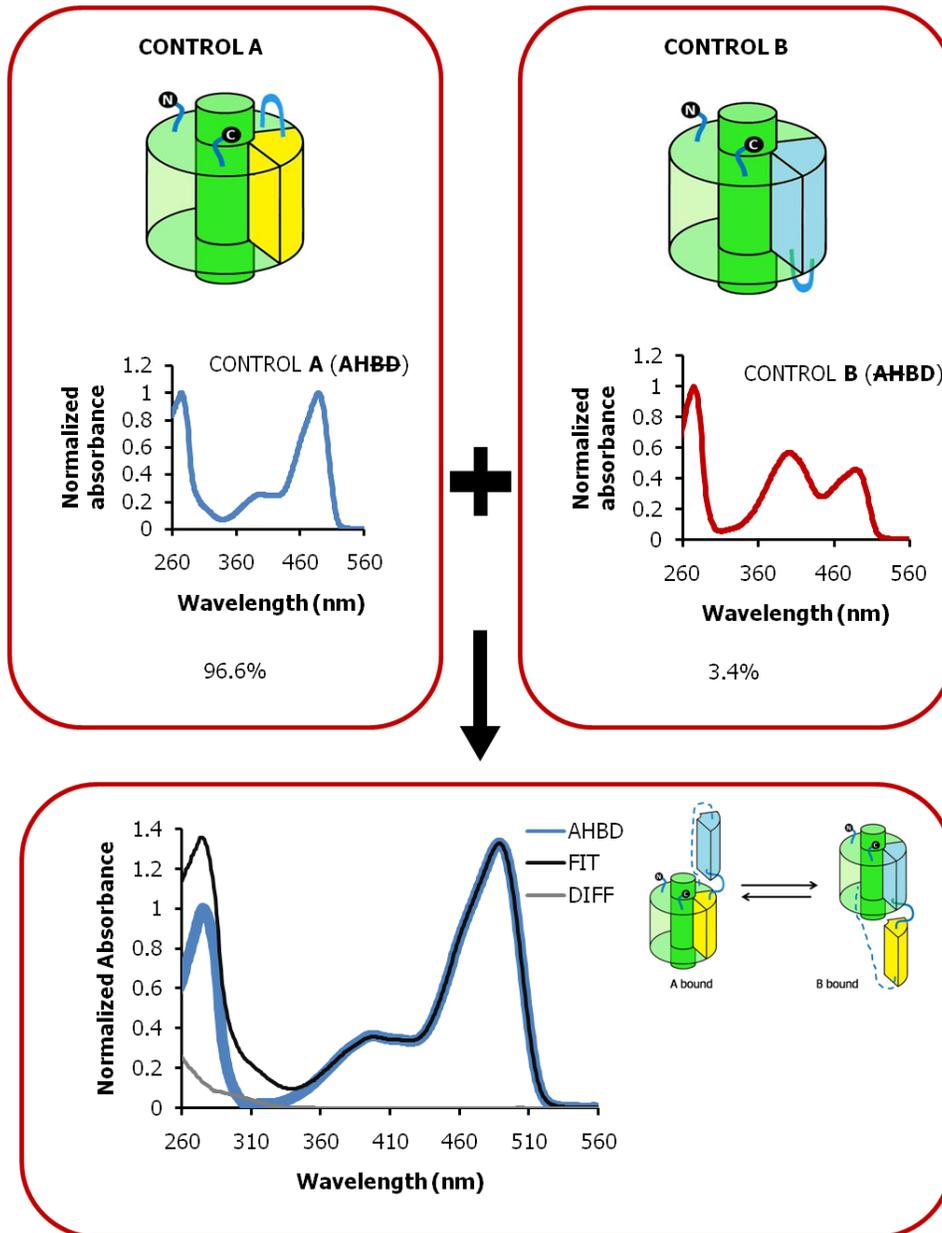


Figure 2-5 Fitting AHBD as a sum of the AHBD and AHBD reveals that the spectrum is $97 \pm 2\%$ AHBD and $3 \pm 17\%$ AHBD.

In order to quantify the contribution of each bound copy in producing the resultant spectrum of protein, we created a series of 4 control protein constructs designated as $AHBD$, $AHBD$, $ADBH$ and $ADBH$, where the strikethrough refers to the removed copy. The bound % was determined by analysis of simulated absorbance spectra created by summing experimental control protein spectra at various ratios. This analysis revealed $AHBD$ existed with $97 \pm 2\%$ of copy A bound and $3 \pm 17\%$ copy B bound, as shown in **Figure 2-5**. In contrast, $ADBH$ existed with $89 \pm 3\%$ A bound and $11 \pm 16\%$ B bound at equilibrium. This result leads us to conclude that, all other factors being the same, the first β -strand (copy A) forms a more stable complex with the protein than does the second β -strand (copy B).

The percentage abundances representing the relative equilibrium concentrations of each form were used to calculate the equilibrium constant *i.e.*, $K_{eq} = \% \text{ B-bound} / \% \text{ A-bound}$. The change in free energy of binding (ΔG) corresponding to swapping of the strand was determined using the equilibrium constant, $\Delta G = -R T \ln K_{eq}$ where $R = 8.3145 \text{ J K}^{-1} \text{ mol}^{-1}$ (universal gas constant) and T (temperature) = 296 K ($23 \text{ }^\circ\text{C}$). Converting the equilibrium constants to ΔG values (for the equilibrium $A \text{ bound} \rightleftharpoons B \text{ bound}$), we found that $AHBD$ is associated with ΔG of $8 \pm 4 \text{ kJ mol}^{-1}$ while $ADBH$ has a value of $5 \pm 2 \text{ kJ mol}^{-1}$. Increase in ΔG (owing to

mutation in FP) is associated with increase in A bound form or decrease in B bound form (Chart 2-1).

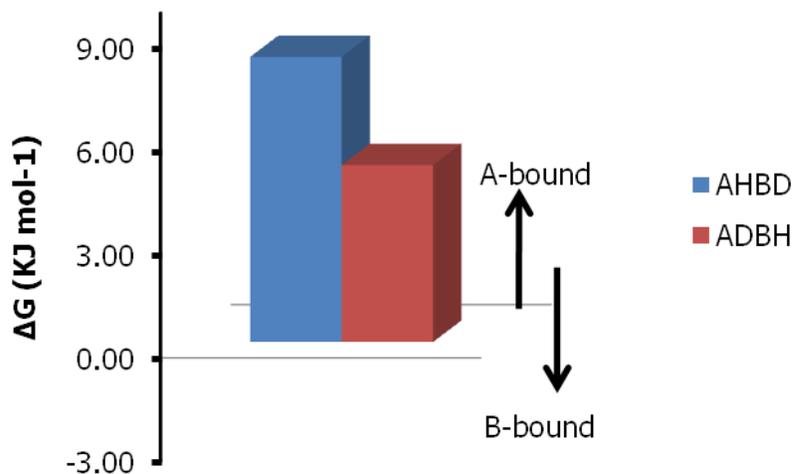


Chart 2-1 Change in ΔG and $A \rightleftharpoons B$ with respect to mutation at 148.

It should be noted that, although these two strands do make many identical interactions with the protein, the two states are not identical. The key difference between the two states is the location of the disordered loop that is composed of the β -strand sequence that is not bound. When copy A is bound to the protein, this loop occurs at the C-terminal end of β -strand 7. When copy B is bound to the protein, this loop occurs at the N-terminal end of β -strand 7. It is apparent from these results that there is a larger energetic penalty associated with having the loop at the N-terminal end than at the C-terminal end of β -strand 7.

We conclude that a strand with His148 forms a more stable complex with the protein than does a strand with Asp148. Accordingly, if strand A has His148 and strand B has Asp148, strand A becomes more favored to bind to the protein. Likewise, if strand A has Asp148 and strand B has His148, strand B becomes more favored to bind to the protein. The $\Delta\Delta G$ for the AHBD to ADBH conversion is $5\pm 2 \text{ kJ mol}^{-1} - 8\pm 4 \text{ kJ mol}^{-1} = -3\pm 7 \text{ kJ mol}^{-1}$. That is, the state with B bound is stabilized by $3\pm 7 \text{ kJ mol}^{-1}$ due to the combined effect of destabilizing strand A with His148Asp and stabilizing strand B with Asp148His. Since both interactions are identical, we estimate the energetic cost of a single His148Asp substitution as $-2\pm 3 \text{ kJ mol}^{-1}$.

All equilibrium constants and thermodynamic energy values are for the equilibrium A bound \rightleftharpoons B bound and so $\Delta\Delta G$ values for modifications are added when the modification is in the A copy and subtracted when it is in the B copy. For example, if the His148Asp mutation was introduced into the B copy of ADBH, the new ΔG would be $5\pm 2 \text{ kJ mol}^{-1} - (-2\pm 3 \text{ kJ mol}^{-1}) = 7\pm 6 \text{ kJ mol}^{-1}$. This increase in ΔG implies that the A bound \rightleftharpoons B bound equilibrium has shifted towards the A bound state, as would be expected when a destabilizing mutation is introduced into the B copy. It should be noted that this value of $7\pm 6 \text{ kJ mol}^{-1}$ is also the energetic cost of having the loop at the N-terminal end (strand B bound) rather than at

the C-terminal end (strand A bound) of β -strand 7. If the A bound and B bound states were truly identical, the equilibrium value would be 1 and the ΔG would be 0 kJ mol⁻¹.

With this proof of concept experiment, we confirmed that it was possible to adjust the equilibrium between the two bound strands through manipulation of the amino acid sequence. Our next goal was to investigate whether further perturbations of the equilibrium between the two bound states could be achieved. Specifically, we sought to find a set of modifications that could allow us to tune the bound equilibrium from having close to 100% strand A bound to close to 100% strand B bound. With the ability to achieve such tuning, we should be able to get the equilibrium close to 1, such that the change in absorbance spectra or excitation ratio following any perturbation to the equilibrium would be close to maximal.

2.3.2 Manipulating the bound equilibrium

We explored the possibility of further manipulating the equilibrium between A and B through the use of rational insertions, deletions, and substitutions. The ability to manipulate this ratio would ultimately enable us to tune biosensors for more maximum signal change. We started with the exploration of manipulation possibilities in AHBD repeat construct and all variants were grouped in V1 (version 1) for simplicity. Further, fruitful

manipulations were used in designing a second generation ADBH scaffold and the resulting group of modifications was named as V2 (version 2).

2.3.2.1 Insertion of residues

We anticipated that insertion of residues before β -7 copy A would destabilize the interaction of copy A with the chromophore and would shift the equilibrium towards binding of copy B and vice versa. Accordingly, the spectral characterization of variants with insertion was performed. The absorbance spectra were not a perfect fit when fitted against the controls AHBD and AHBD, which suggested that these manipulations may have altered the original spatial interaction of the 148th residue with the chromophore. In this case, the use of our previously made control constructs would be inappropriate since they would not accurately reflect the spectra of the bound states.

Subsequently, additional control constructs were created and characterized such that the control A had just B copy deleted but retained the linker and inserted residues before A, depicted for simplicity as (I4)AHBD. This procedure tended to provide very good fits of the experimental absorbance spectra. The following **Table 2-1** further provides the details.

Table 2-1 Characterization of V1 variants with insertions

S. No.	NAME	CARTOON	K_{eq}	ΔG
1	AHBD		≤ 0.2	8 ± 4
2	AHBD(I4)		≤ 0.2	10 ± 6
3	(I4)AHBD		≤ 0.9	1 ± 1
4	(I4)AHBD(I4)		≥ 99	≤ -11
5	(I4)AHBD(I8)		≥ 99	≤ -11
6	(I4)AHBD(I6)		≥ 99	≤ -11

I4 prefixed to AHBD represents insertion of 4 residues prior to copy A and Ix (where x could be 4, 6 or 8) suffixed to AHBD represents insertion of x residues after copy B. Unit for ΔG is kJ mol^{-1} .

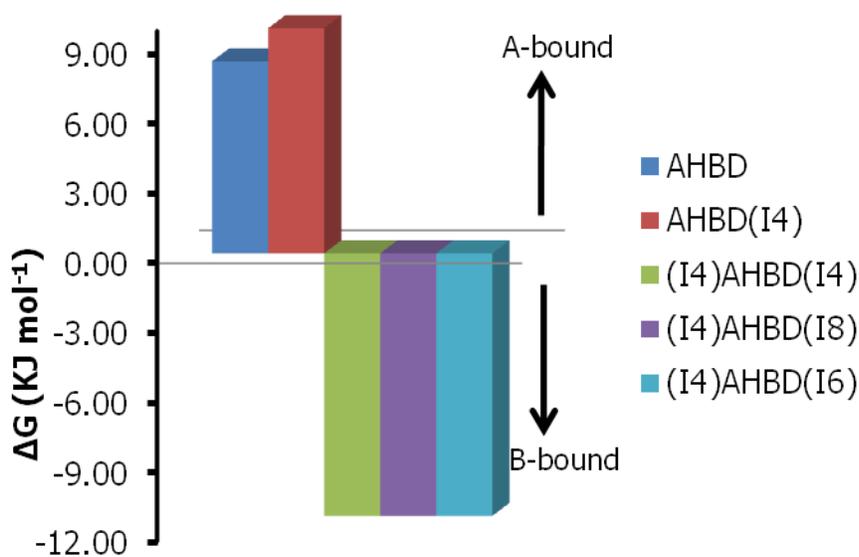


Chart 2-2 Change in ΔG and $A \rightleftharpoons B$ with respect to insertions

The results with insertions before A or after B revealed the overall equilibrium as depicted in entries 2 and 3 in **Table 2-1**. The insertion after S7B was revealed to destabilize B-bound and stabilize the A-bound form. Similarly, the insertion before S7A was observed to destabilize A-

bound or stabilize the B-bound form. However, when insertions were introduced at both ends the equilibrium changed quite dramatically to favor having S7B bound (entries 4, 5, and 6 in **Table 2-1**). Notably, increasing the length of the insertion after S7B from 4 to 6 or 8 residues had a slight and less intuitive destabilizing effect on the S7B bound form (Chart 2-2).

2.3.2.2 Deletion of residues

Based on our initial set of experiments, it was apparent that binding of S7A is favored over S7B. Thus, with the aim of destabilizing the binding of S7A and shifting the equilibrium towards unity, we decided to delete a few C-terminal residues of S7A, starting from 153rd to 151st. The needed constructs were likewise created and fitted against appropriate controls.

Table 2-2 Characterization of V1 variants with deletions.

S. No.	NAME	CARTOON	K _{eq}	ΔG
1	AHBD		≤0.20	8±4
2	(dK)AHBD		≤0.25	7±3
3	(dIK)AHBD		≥99	≤-11
4	(dYIK)AHBD		≥99	≤-11

'd' represents deleted and single letter code of deleted amino acid is shown next to 'd'. Unit for ΔG is kJ mol⁻¹.

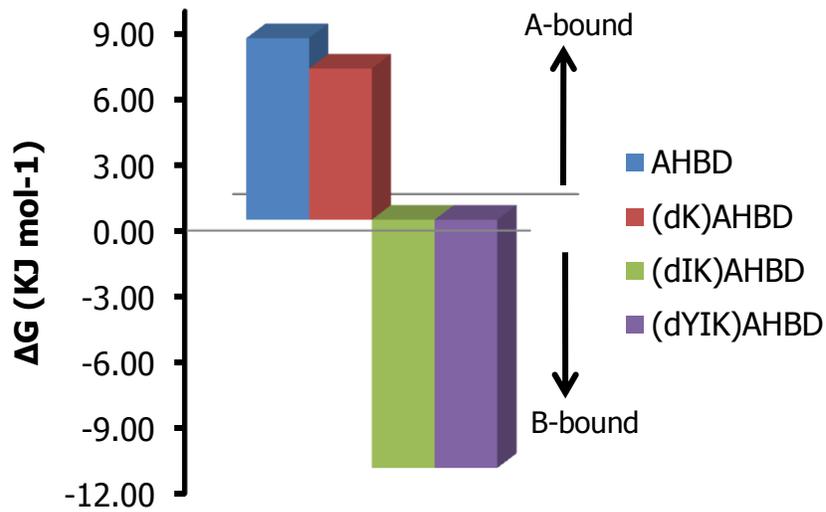


Chart 2-3 Change in ΔG and $A \rightleftharpoons B$ with respect to deletions

From the results it was revealed that deletion of two (specifically Ile152-Lys153) or three residues (specifically Tyr151-Ile152-Lys153) highly destabilizes the A bound form from $\geq 97\%$ to 0% (Chart 2-3). The control protein (dYIK)AHBD corresponding to control construct with three residue deletion, exhibited no long wavelength absorbance peak, suggesting that three or more residue deletion from within the strand is detrimental to protein integration and folding. In summary, deletion of two residues from the C-terminal end of S7A has a significant impact on altering the equilibrium and it is worthwhile to study them in further constructs. Deletion of 3 residues from the C-terminal end of S7B appears to interfere with chromophore formation and so this modification is less likely to be useful in future constructs.

2.3.2.3 Mutation of residues

Residues on either side of Asp148, that is Ser147 and Asn149, were mutated to alanine. Alanine, a small residue with a high α -helix propensity, was expected to destabilize the bound form of the strand into which it was introduced.⁶⁷ Accordingly, Ser147Ala and Asn149Ala in S7B were suspected to reduce the binding affinity of respective strand and thus destabilize its binding. **Table 2-3** gives a comparison of the effect of these mutations.

Table 2-3 Characterization of V1 variants with mutations

S. No.	NAME	CARTOON	K _{eq}	ΔG
1	(dIK)AHBD		≥ 99	≤ -11
2	(dIK)AHBD(S147A)		20±16	-7±1
3	(dIK)AHBD(S147A,N149A)		63±58	-10±2
4	(dYIK)AHBD		≥ 99	≤ -11
5	(dYIK)AHBD(S147A)		≥ 99	≤ -11
6	(dYIK)AHBD(S147A,N149A)		≥ 99	≤ -11
7	(I4)AHBD		1	1±1
8	(I4)AHBD(S147A)		2±1	-1±1
9	(I4)AHBD(S147A,N149A)		10±6	-6±1

S147A is Serine 147 mutated to alanine and N149A is aspartamine 149 mutated to alanine. Unit for ΔG is kJ mol^{-1} .

The introduction of point mutations into S7B, had little effect on the equilibrium (Chart 2-4). However, an effect was more apparent when these mutations were introduced in the context of the I4 modified S7A. Though it was surprising that the effect was the opposite of what we had predicted. The $\Delta\Delta G$ for the Ser147Ala mutation is $\leq -11 \text{ kJ mol}^{-1} - (-7 \pm 1 \text{ kJ mol}^{-1}) \leq -4 \pm 1 \text{ kJ mol}^{-1}$. That is, Ser147Ala in S7B destabilizes the bound state of S7B by $-4 \pm 1 \text{ kJ mol}^{-1}$, effectively shifting the equilibrium. The $\Delta\Delta G$ for the Asn149Ala mutation is $-7 \pm 1 \text{ kJ mol}^{-1} - (-10 \pm 2 \text{ kJ mol}^{-1}) = 3 \pm 3 \text{ kJ mol}^{-1}$. This mutation in S7B appears to counter intuitively stabilize the bound form of S7B. Due to the relatively subtle effect of the Asn149Ala mutation on shifting the equilibrium to the S7B-bound state, this mutation was included in later designs.

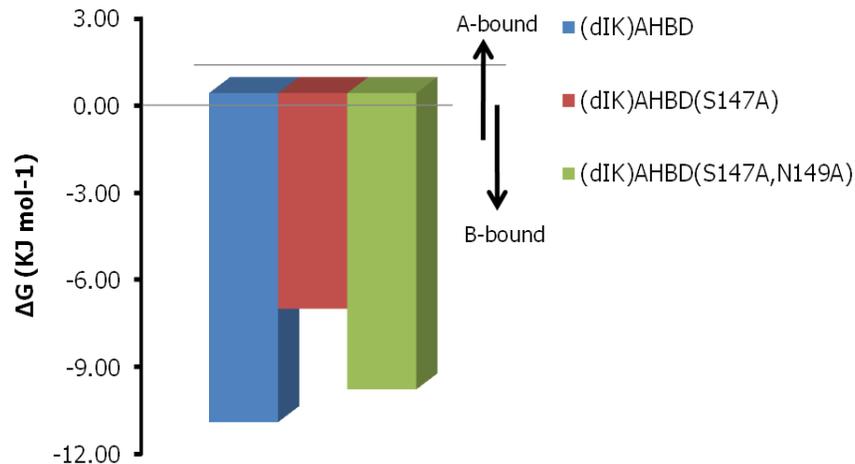


Chart 2-4 Change in ΔG and $A \rightleftharpoons B$ with respect to deletions

2.3.2.4 Rational manipulation of version 2 variants

The “version-2” series of variants were designed based on the foundation of results from manipulation of version-1 variants. The main goal was to get highly distinguishable change in spectral signature with respect to a small perturbation in equilibrium. To achieve this we introduced various modifications in order to bring the bound equilibrium closer to unity.

Since the bound equilibrium inherently favors the S7A-bound state by $7 \pm 6 \text{ kJ mol}^{-1}$, all other factors being the same, we sought to destabilize this state. One way of achieving this was to introduce the destabilizing His148Asp ($-2 \pm 3 \text{ kJ mol}^{-1}$) into S7A and leave His148 in S7B (the opposite of the version-1 series). Two other modifications were the deleting of two residues (dIK) from the C-terminus of S7A and two residues (one Ser from the linker region and an Asn) from the N-terminus of S7B. Our results had suggested that the dIK modification would have little effect on the equilibrium. We assumed the same to be true of the deletion of Asn from the N-terminus of S7B, but we did later add it back to test this assumption (**Figure 2-6**).

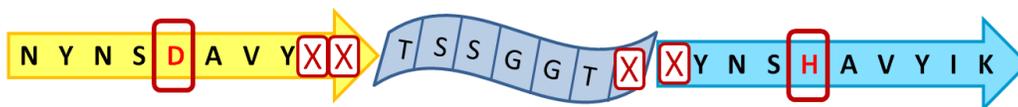


Figure 2-6 Design of V2 variant. The residues in red are 148th that is D (Asp) for copy A (yellow) and H (His) for copy B (cyan) respectively. The crossed residues were deleted. The 6-residue linker is in blue in middle.

Using the version-2 scaffold, we introduced several of the more promising modifications identified during studies of our version-1 scaffold. Results of the version-2 manipulation study are illustrated in **Table 2-4**. We predict that the variant associated with ΔG closer to unity are preferably better choices for generation of biosensors. Accordingly, V2_dYT_iN144 with ΔG of 4.37 kJ mol^{-1} is so far the best variant we have identified.

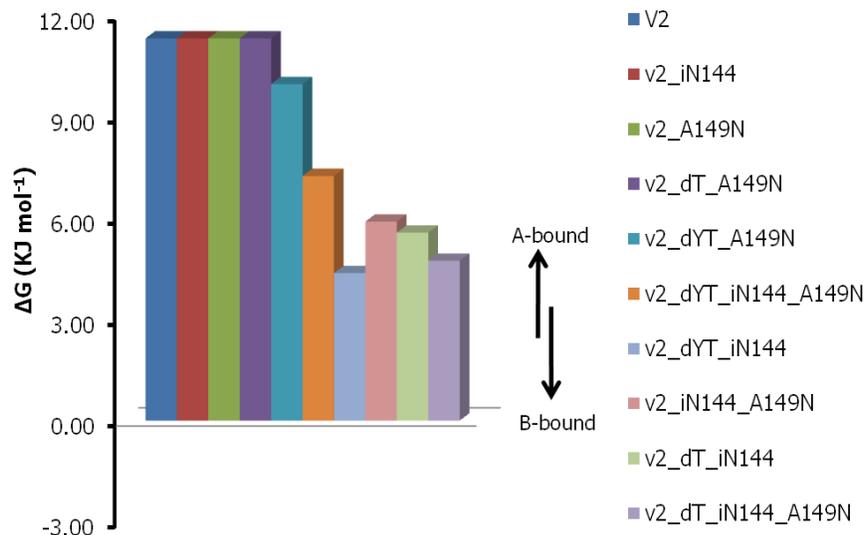


Chart 2-4 Change in ΔG and $A \rightleftharpoons B$ with respect to mutations.

Table 2-4 Characterization of V2 variants with manipulations

S. No.	NAME	CARTOON	K_{eq}	ΔG
1	v2		≤ 0.01	≥ 11
2	v2_iN144		≤ 0.01	≥ 11
3	v2_A149N		≤ 0.01	≥ 11
4	v2_dT_A149N		≤ 0.01	≥ 11
5	v2_dYT_A149N		≤ 0.20	10 ± 6
6	v2_dYT_iN144_A149N		≤ 0.23	7 ± 4
7	v2_dYT_iN144		≤ 0.36	4 ± 2
8	v2_iN144_A149N		≤ 0.28	6 ± 3
9	v2_dT_iN144		≤ 0.28	6 ± 3
10	v2_dT_iN144_A149N		≤ 0.34	5 ± 2

v2 represents version 2 variants. iN144 denote insertion of aspartamine at position 144; 'd' denote deletion; N149A is aspartamine 149 mutated to alanine. Unit for ΔG is kJ mol^{-1} .

2.3.3 Insertion of binding modules to create biosensors

Having constructed a series of variants with various ratios of the two copies of β -strand 7 bound to the protein, we next attempted to exploit this design for the construction of molecular biosensors. We anticipated that insertion of a binding domain between strands A and B (*i.e.*, insertion into Kpn1 and Xho1 restriction sites encoded within the linker region), could produce a protein in which a ligand-induced conformational change in the binding domain would alter the ΔG for the equilibrium $A \text{ bound} \rightleftharpoons B \text{ bound}$ and produce a ratiometric change in the absorbance and fluorescence excitation spectra. The basic biosensor construct is illustrated in **Figure 2-7**.

Ideally, the $\Delta\Delta G_{A\rightleftharpoons B}$ associated with ligand binding would be as large as possible, though 22.6 kJ mol^{-1} is sufficient for producing the maximum ratiometric change as the protein transitions between states with $>99\%$ A bound to $> 99\%$ B bound. For $\Delta\Delta G_{A\rightleftharpoons B} < 22.6 \text{ kJ mol}^{-1}$, at least one of the two states would need $\Delta G_{A\rightleftharpoons B}$ to be in the range of -11.3 to $+11.3 \text{ kJ mol}^{-1}$ for a spectral change to be observed. One could envision a variety of scenarios that would follow from insertion of a binding domain between strands A and B and it is entirely possible that the insertion of a binding domain could shift the equilibrium such that $\Delta G_{A\rightleftharpoons B} > 11.3 \text{ kJ mol}^{-1}$ or $< -11.3 \text{ kJ mol}^{-1}$ for both the ligand-bound and ligand-free states of the binding domain. In such cases the ability to adjust the equilibrium using rational changes such as those described above, could be used to bring the $\Delta G_{A\rightleftharpoons B}$ for at least one of the two states into the realm where changes in $\Delta G_{A\rightleftharpoons B}$ are manifested as a change in the absorbance spectrum.

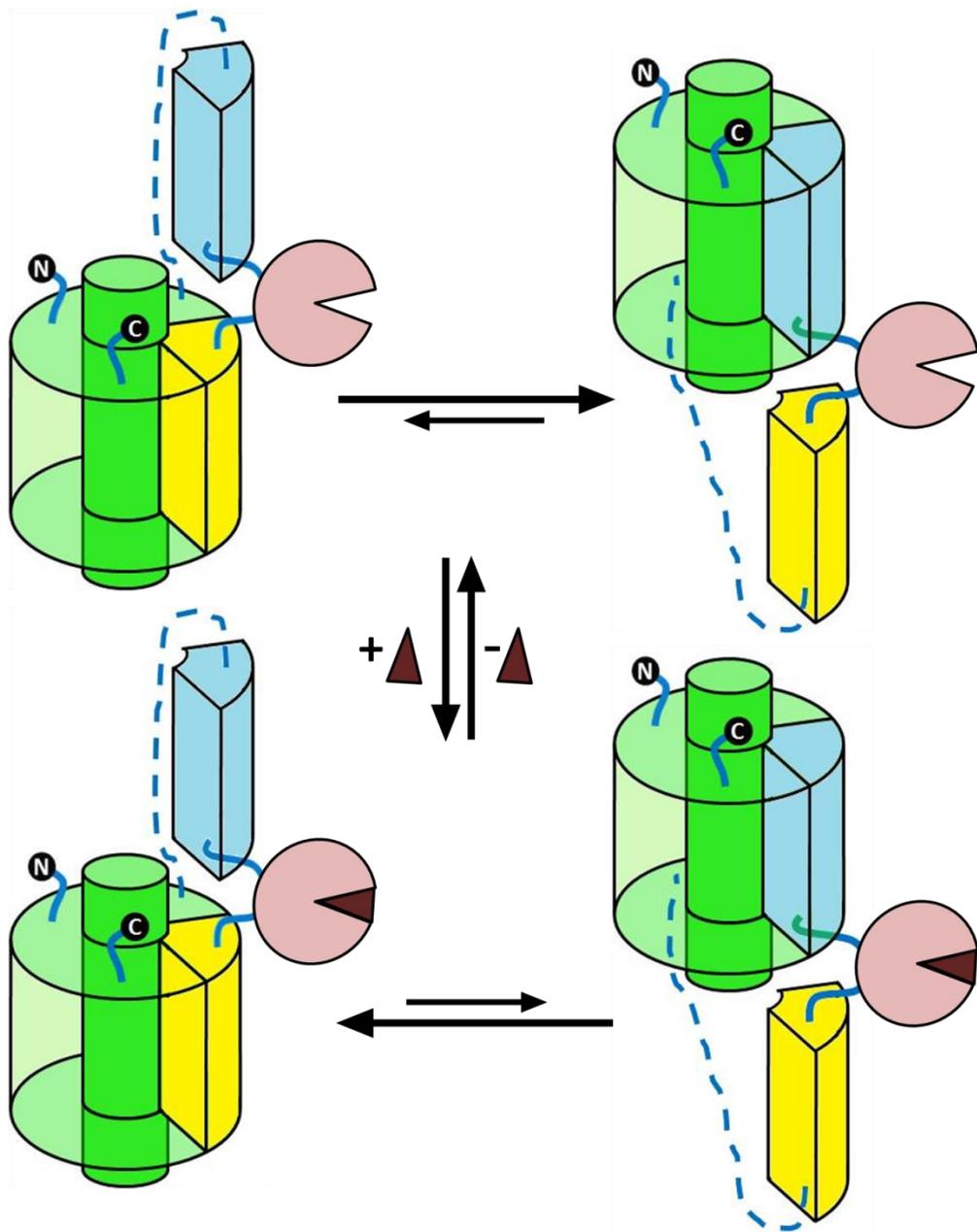


Figure 2-7 Cartoon illustration of biosensor construction. Light pink circular entity represents the binding domain and dark red triangular entity depicts the ligand. The larger arrow symbolizes the more favored reaction direction.

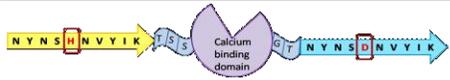
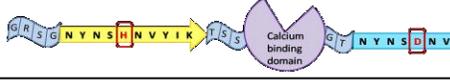
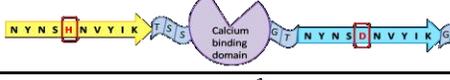
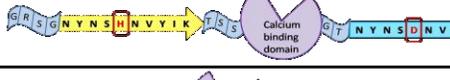
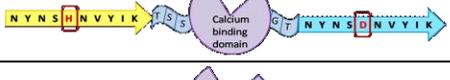
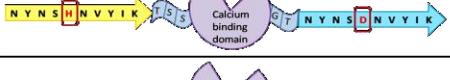
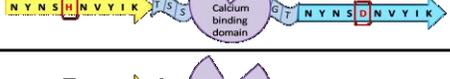
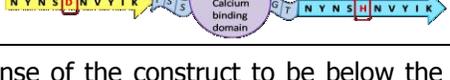
2.3.3.1 Insertion of CAM/M13 from YC-Nano15 to create Ca²⁺ biosensors

To evaluate the validity of β -repeat scaffold as biosensor we created a Ca²⁺ biosensor using the optimized β -repeat variants by inserting the Ca²⁺ binding domain derived from the yellow cameleon variant known as YC-Nano15. This binding domains consists of CaM and M13 linked by a GGGGS linker, and has a K_d for Ca²⁺ of 15.0 nM.²⁴ Titration of the biosensor with varying concentrations of Ca²⁺ and corresponding analysis of the fluorescence spectra was performed, and the K_d for the TAFS-based biosensor construct was ascertained to be 7.1 nM (**Figure 2-8**). The buffer preparation for titration was done using reciprocal dilution method. The biosensor constructs were then evaluated for change in excitation ratio from the apo state (buffered in 10 mM EGTA, 100 mM KCl, 10 mM MOPS, pH 7.1 and 20 °C) to the Ca²⁺ bound state. On evaluation of spectral characteristics, the Ca²⁺-sensing variant appeared to switch from having predominantly strand B bound in the absence of Ca²⁺ to predominantly strand A bound in the presence of Ca²⁺ (Appendix VA). The excitation ratio change (495 nm to 405 nm) was determined and is provided in **Table 2-5**. Specifically, upon adding Ca²⁺, we detected a green fluorescence emission intensity decreases when excited at 480 nm (the absorbance peak associated with strand B in the version-2 series), and an emission intensity increase when excited at 400

nm (the absorbance peak associated with strand A in the version-2 series). To gain additional evidence for or against the proposed mechanism, we constructed additional variants with mutations in either the first or second strand. It is evident that V2_dT_iN144 based biosensor gave the best performance of all variants used, as shown in Appendix VA.

Table 2-5 Evaluation of Ca^{2+} biosensor performance (excitation ratiometric change)

(A) Biosensor construct using V1-variants

S. No.	NAME	CARTOON	CONSTRUCTS WITH $\geq 9\%$ RATIO CHANGE IN PRESENCE OF Ca^{2+}
1	AH(CaM/M13)BD		X
2	(i4)AH(CaM/M13)BD		X
3	AH(CaM/M13)BD(i4)		X
4	(i4)AH(CaM/M13)BD(i4)		X
5	(dK)AH(CaM/M13)BD		X
6	(dIK)AH(CaM/M13)BD		X
7	(dYIK)AH(CaM/M13)BD		X
8	AD(CaM/M13)BH		X

Red 'X' denote the response of the construct to be below the 9% threshold ratiometric change on calcium binding.

(B) Biosensor construct using V2-variants

S. No.	NAME	CARTOON	CONSTRUCTS WITH $\geq 9.0\%$ RATIO CHANGE IN PRESENCE OF Ca^{2+}
1	v2_in144_A149N(CaM/M13)		X
2	v2(CaM/M13)		✓ (16.2%)
3	v2_delT(CaM/M13)		✓ (14.9%)
4	v2_delT_in144(CaM/M13)		✓ (90.9%)
5	v2_delT_in144_A149N(CaM/M13)		X
6	v2_delYT_in144(CaM/M13)		X
7	v2_delYT_in144_A149N(CaM/M13)		X
8	v2_delYT(CaM/M13)		X
9	v2_delYT_V150A(CaM/M13)		X
10	v2_in144(CaM/M13)		✓ (29.6%)
11	v2_delT_in144(CaM/M13)		X
12	v2_delYT_in144(CaM/M13)		X

Red 'X' denote the response of the construct to be below the 9% threshold ratiometric change on calcium binding and green tick mark denote response above the threshold.

The dissociation constant (K_d) of Ca^{2+} -biosensor, was determined to be 7.1 nM as depicted in the **Figure 2-8**.

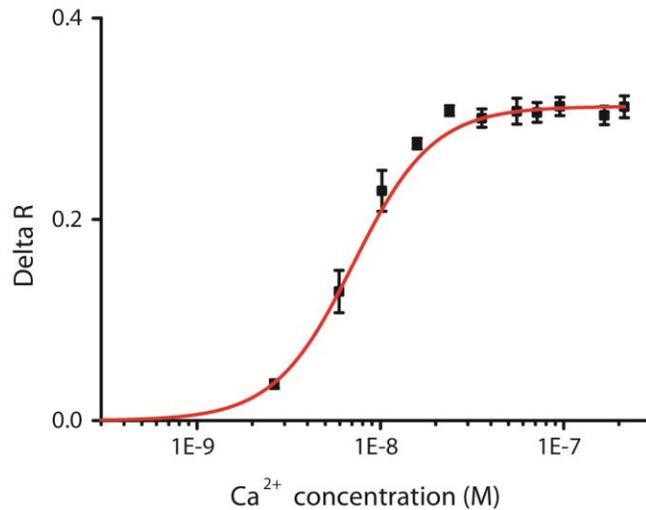


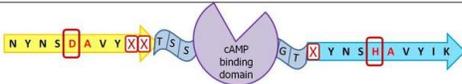
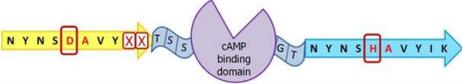
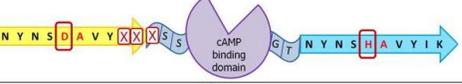
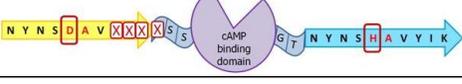
Figure 2-8 Ca^{2+} -biosensor titration curve. Delta R on Y-axis denotes (difference of ratio of excitation intensity at 400 nm to 480 nm for apo and Ca^{2+} bound state) divided by minimum of the ratio. On X-axis, concentration is drawn on logarithmic scale. Fitted curve for the averaged data of three independent measurements is presented.

2.3.3.2 Insertion of CAP to create cyclic AMP biosensors

To evaluate the scope of the biosensing applicability of the TAFS β -repeat scaffold, we created a cyclic AMP biosensor by inserting CAP (catabolite activator protein) as the binding domain. cAMP is a ubiquitous secondary messenger regulating multitude of tasks in every organism. cAMP is the allosteric effector of CAP also known as cAMP receptor protein (CRP) which exists as a dimer composed of two identical subunits of 209 residues long. We used single subunit of CAP that retains the cAMP

binding affinity but loses the ability to bind to DNA. Each subunit folds into two domains: a large N-terminal, cAMP binding domain and a small C-terminal, DNA binding domain.

Table 2-6 Evaluation of cAMP biosensor performance (excitation ratiometric change)

S. No.	NAME	CARTOON	CONSTRUCTS WITH $\geq 9.0\%$ RATIO CHANGE IN PRESENCE OF cAMP
1	v2(CAP)		✓ (11.0%)
2	v2_iN144(CAP)		✓ (9.0%)
3	v2_delT_iN144(CAP)		✓ (10.9%)
4	v2_delYT_iN144(CAP)		✓ (11.0%)

Green tick mark denotes response of construct to be equal to or above the threshold in presence of cAMP.

Evaluation of the cAMP biosensor performance was carried out by analyzing the change in absorbance spectra in transitioning from the apo state (buffered in 50 mM Tris-HCl, pH 7.1 at 20 °C) to the cAMP bound state (10 mM cAMP). The results are displayed in the **Table 2-6** and spectra are included in Appendix VB. Although the sensitivity of biosensor is poor, the activity is enough to suggest the generic use of the scaffold. The relatively low sensitivity may be attributed to rigid domain type movement of CAP subunits upon cAMP binding as compared to limited

structural rearrangement type change of CaM-M13 assembly. Thus, a longer linker between CAP and scaffold might improve the performance.

2.4 Conclusion

In an attempt to create a generic FP reporter, we employed a modified alternate frame folding strategy to introduce an artificial switch in avGFP. We demonstrated that insertion of a partially duplicated β -7 strand introduces swapping mechanism between the two frames. We ascertained the mutually exclusive nature of equilibrium between the two frames, with only one frame folding to occupy the β -barrel at any time and unfolding only at expense of other. We further manipulated the equilibrium between S7A and S7B through the use of rational insertions, deletions, and substitutions. The ability to manipulate this ratio is, ultimately, what would enable us to rationally tune biosensors for maximum signal change. We evaluated the effect of destabilization on the equilibrium and envision that deletions of more than two residues has a drastic effect akin to insertion of 4 residues whereas point mutations can serve to slightly change the equilibrium.

We verified that the absorption spectra of the variants are representative of the equilibrium between them and can be fitted against control protein spectra using independent component analysis providing

the equilibrium constant and free energy change of binding. We also created biosensors of Ca^{2+} and cyclic AMP from optimized variants and proved the universal applicability of the FP switch.

Chapter 3:
Preliminary study towards
development of RFP-based AFF
scaffold

3.1 Introduction

Multicolour fluorescence imaging is an important tool for real time in vivo studies of biological processes. Accordingly, expansion of the FP-based biosensor colour palette is vital for discrimination of different cell types, various biomolecules and biological activities. To fulfill the purpose, a number of FPs with excitation and emission profiles spanning the visible spectrum have been developed.^{10,75} Red fluorescent proteins (RFPs) have excitation and emission spectra closest to the infrared “optical window” where absorbance by hemoglobin, melanin and water is at a minimum.⁷⁶ By means of low scattering and reduced autofluorescence, RFPs are preferred probes for deep tissue imaging with high resolution.⁷⁷

The unique fluorescence signature of RFPs is attributed to different primary sequence and the chromophore chemistry compared to avGFP. The chromophore of DsRed involves post-translational extension of the conjugated system of the GFP chromophore by an additional N-acylimine moiety as described in **Section 1.2** and **Figure 1-2**.¹³ In addition, the chromophore of RFPs tend to remain in anionic state owing to low pK_a unlike GFP chromophore.⁷⁸

The chromophore chemistry classifies FPs into two super families named as green-emitting family and red-shifted family with an extended chromophore.⁷⁹ Having a red-shifted counterpart would not only serve as

a colour alternative to avGFP TAFS scaffold but also act as bench for tailoring other red-shifted variants for real time *in vivo* imaging in whole organism context, where green family variants disappoint.

The prerequisites for the application of alternate frame switching (AFS) strategies, which include the TAFS as well as the AFF strategy, to any FP include: 1) an insertion / permutation site in the vicinity of the chromophore; 2) a tolerance to flexibility of residues near the insertion site; and 3) the identification of a single crucial residue in the selected frame which influences the spectral signature of FPs.⁴² Once these criteria are met, alternate frames can be designed utilizing the protocols of genetic engineering. Evaluation of switching among the alternate frames and development and optimization of the scaffold can then be achieved centered on the principles of TAFS strategy as described in **Chapter 2**.

mCherry is one of the most highly optimized monomeric derivatives of DsRed isolated from coral; offering longer wavelengths with excitation and emission wavelength at 587 and 610 respectively, higher photostability, faster maturation and excellent pH resistance.⁷⁵ In addition, a variety of circular permutation sites in mCherry have been identified and positions expected to have tolerance to insertions have been identified in the 10th β -strand.^{80,81} The general structure of cpmCherry is displayed in the **Figure 3-1**.

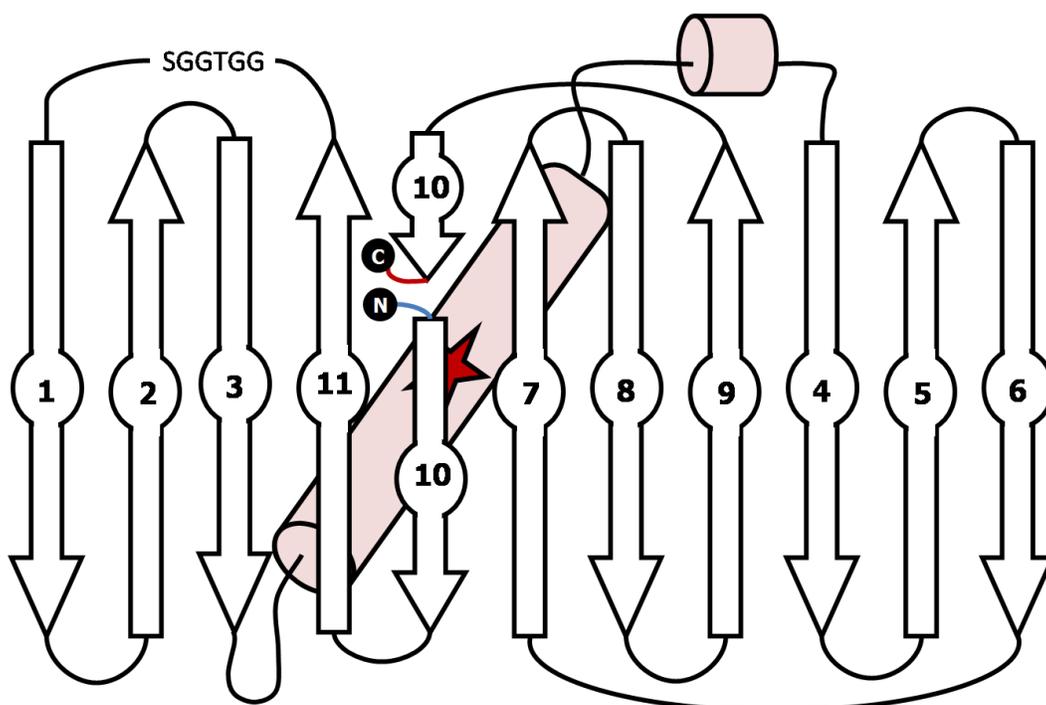


Figure 3-1 Schematic representation of overall folding of cp196 mCherry. The split β -10 strand is shown like that for clarity, but in reality the short portion is likely not forming a β -strand.

cp196mCherry was selected as primary template on which to develop a red fluorescent analogue of the avGFP TAFS scaffold, described in **Chapter 2** of this thesis. The translated gene sequence of cp196 mCherry is included in Appendix VI. Our ultimate goal is to introduce a conformational switch in mCherry that will provide a new biosensor scaffold and offer increased variety of choices for multicolour imaging applications. However, as a first step towards this goal we needed to identify colour-shifted variants that differ by only a few mutations that are located on a single strand of the protein. This would allow us to design a

variant with a duplicated β -strand and a spectral signature that could be used to experimentally determine if the first or second copy was bound to the protein.

In this chapter, we will discuss construction and screening of a library to assess the contribution of the Phe-Ile-Lys-Leu-Asp sequence motif located in the 10th β -strand (at the N-terminus of cp196 mCherry) in tuning the fluorescence hue. Use of primers with randomized codons to create a site directed library and screening of large libraries facilitated by a custom built colony imaging system will be described. The significance of single crucial amino acid residue as a prerequisite of frame-switching strategies will be ascertained.

3.2 Materials and methods

3.2.1 General methods and materials

The general procedures, protocols, buffers, chemicals, enzymes and laboratory equipment were described in chapter 2 in detail.

3.2.2 Search and selection of the consensus motif

The sequence of the cp196mCherry was compared using ClustalW for sequence homology with other FPs. The RCSB site's sequence motif feature (<http://www.rcsb.org/pdb/search/advSearch.do>) in advanced

search interface was used to blast the motif sequence and confirm the consensus among various FPs.

3.2.3 Site directed mutagenesis studies

Forward primers were designed to randomize three residues of the motif with an NNK codon, where N represents equimolar concentration of A, C, G or T while K represents equimolar concentration of G or T. Three primers were designed to cover all the possible combinations. Primer 1 (FPF_Xho_KLD_LIB) was for randomization of residues 198–200⁽⁺⁾; FPF_Xho_FIK_2 for randomization of residues 196–198 and FPF_Xho_IKL_3 for randomization of residues 197–199. Reverse primer HIN_FPR_D were paired with forward primer FPF_Xho_FIK_2, FPF_Xho_KLD_LIB and FPF_Xho_IKL_3 to amplify the template cp196 mCherry and generate the libraries #1, 2, and 3 respectively. PCR amplification of cp196 mCherry template was done using a standard protocol supplemented with 5 µl DMSO for a 60 µl PCR reaction. PCR products were gel purified and digested with restriction enzymes XhoI and HindIII and subcloned into the pBAD/His B vector (Invitrogen). The plasmids were transformed in electrocompetent *E. coli* strain ElectroMAX DH10B™ (Invitrogen) using standard protocols as described in the previous chapter. *E. coli* were spread on LB-agar supplemented with 0.2%

⁽⁺⁾ All the residue numbers are corresponding to positions in native mCherry sequence.

L-arabinose and 0.02% ampicillin in disposable polystyrene Petri dishes (100 mm x 15 mm). Plates were incubated at 37 °C overnight and were subsequently transferred to 4 °C.

3.2.4 Screening of library

The library screening was performed using in-house imaging system⁸². Images of fluorescent colonies were captured by imaging system. For excitation, a 175W xenon-arc lamp (Sutter Instrument Company, Novato, CA) was employed. Wavelength selection was achieved using band-pass filter in a filter-wheel (Sutter). A Retiga 1300i 12-bit CCD camera (QImaging, Burnaby, BC), placed immediately behind the emission filter, was used to capture images of the fluorescent intensities of the screened colonies. Positioning of excitation and emission filter-wheels, image capturing and processing were automated by custom macros run using ImagePro Plus (Media Cybernetics Inc., Silver Spring, MD). Three images using the following combination of excitation and emission filters were captured for each plate: 'first image' was the null/ background image; 'second image' using 510-560 nm excitation and 600-660 nm emission filters; and 'third image' using 540-580 nm excitation and 600-660 nm emission filters. Values of fluorescent intensities of the images for plates were exported to Microsoft Excel program and ratios of emission were calculated. Colonies with diverse ratios were propagated and their plasmid DNA was extracted and sequenced.

3.2.5 Characterization and analysis of variants

Selected sequenced variants from each library were cultured in LB media supplemented with 0.2% L-arabinose and 0.04% ampicillin for overnight at 37°C with shaking at 225 rpm. Bacterial cells were pelleted by centrifugation at 7500 rpm for 2 min at 4 °C. B-PER II (Pierce) was used to extract the soluble proteins from the cell pellet as per the manufacturer's protocol. Excitation wavelength scan from 400 – 600 nm and fluorescence emission scan from 590 – 690 nm were using TECAN Safire2 microplate reader. Excitation and emission scanning was repeated for Ni-NTA bead purified selected protein variants using Photon Technology International (PTI) MP1 fluorescence system.

3.2.6 Construction and analysis of biosensor scaffold with alternate frame

The R-GECO1 plasmid bearing the CaM and M13 gene segments flanking the mApple gene was used as parent plasmid (refer to Appendix VIII for sequence).⁴⁹ Introduction of partially overlapping sequence of copy of 10th β -strand at the C-termini was achieved using PCR amplification of pBAD plasmid encoding the corresponding cp196mCherry variant. Reverse primers 3-5 to 3-8 with MluI and forward primer 3-9 with SacI restriction site (Appendix I), were employed in this experiment.

cp196mCherry with alternate frame was sub-cloned in R-GECO1 plasmid between the M13 and CAM gene segments. The resultant biosensor constructs were expressed in LB-media and the B-PER extract was evaluated on a TECAN Safire2 microplate reader.

3.3 Results and discussion

3.3.1 Library creation and screening

We initially chose cp196mCherry (circularly permuted at 196 monomer Cherry) as the template FP which provided an established insertion site in vicinity of chromophore.^{80,81} As the crystal structure of cpmCherry has not been solved, the exact positioning and orientation of residues of cp termini is not known.

Consensus sequence search facilitated the identification of residues in the 10thβ-strand that are highly conserved and might have functional relevance to the fluorescence hue. BLAST searches indeed yielded F/D/N IKLD (representing Phe/Asp/Asn196-Ile197-Lys198-Leu199-Asp200) as a strongly conserved sequence in 10thβ-strand of cpmCherry which is the N-terminal strand in permuted version. The 197th (Ile) and 199th (Leu) residues are oriented towards the chromophore and are buried deep in hydrophobic core of the protein as depicted in **Figure 3-2**.

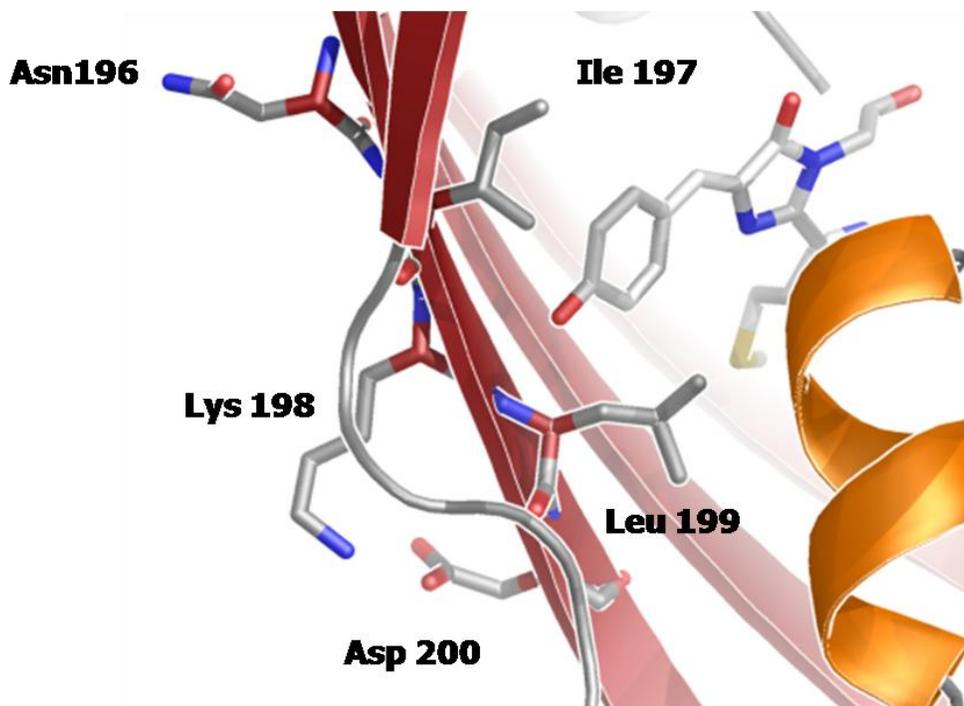


Figure 3-2 PyMol rendered view of consensus motif residues (as labeled) with respect to the chromophore (labeled as CRO).¹⁴

Site directed mutagenesis with NNK randomization was performed on 3 consecutive residue groups of consensus motif. A collection of 3 residues was chosen to reduce the library size (relative to randomization of 5 residues altogether). We could have also chosen to perform saturation mutagenesis at each residue individually, but this approach would not take into account the possible influence of adjacent residues (particularly those with solvent-exposed side chains) on the position and precise orientation of the buried residues. The library size with NNK randomization of a given codon is given by 32^n , where n is the number of residues mutated. In our case library size was 32^3 or 32768.

Three libraries were screened and a broad distribution of fluorescent brightnesses (as well as non-fluorescent colonies) was observed. The ratio of fluorescent to non-fluorescent colonies was taken as indicative of functional significance of corresponding residues. Residue at 197 position has been suggested in literature to affect the fluorescence emission of RFPs.^{10,77,82,83} Screening was thus designed to verify the significance of corresponding residues at 197th position.

The in-house custom imaging system was utilized and functionally significant variants were selected based on their fluorescence ratios of emission corresponding to each excitation.⁸² Firstly, excitation at 510-560 nm is expected to be optimal for blue-shifted variants while excitation at 540-580 nm for red-shifted variants; in both cases, higher emission intensity is expected when excited by optimal excitation and vice versa. Images were processed using ImageProPlus and data was exported to Microsoft Excel sheet and analyzed.

Variants were selected to represent the distribution of observed emission ratios. Sequencing of selected variants depicted a wide trend of allowed residues. The % abundance of each amino acid observed at particular position of selected functional library variants are depicted in **Table 3-1** corresponding to Library #1, 2, and 3 sequenced mutants.

Table 3-1 Distribution of residues identified at each position.

Library # 1				Library # 2				Library # 3			
Amino acid	p196	p197	p198	Amino acid	p198	p199	p200	Amino acid	p197	p198	p199
A	1.75	0	5.26	A	6.52	0	4.34	A	0	0	0
C	0	0	3.5	C	6.52	6.52	4.34	C	0	6.25	2.08
D	1.75	0	1.75	D	0	0	8.69	D	0	2.08	2.08
E	1.75	0	7.02	E	6.52	0	2.17	E	0	4.16	0
F	21.05	0	1.75	F	4.34	0	4.34	F	2.08	8.33	2.08
G	3.5	1.75	0	G	2.17	0	4.34	G	2.08	4.16	4.16
H	12.28	0	0	H	4.34	0	2.17	H	4.16	4.16	0
I	1.75	94.74	1.75	I	0	0	13.04	I	64.58	2.08	0
K	0	0	22.8	K	6.52	0	4.34	K	0	33.33	0
L	17.54	0	5.26	L	13.04	80.43	10.86	L	0	12.5	83.33
M	5.26	0	0	M	2.17	6.52	4.34	M	0	0	4.16
N	3.5	0	0	N	2.17	0	0	N	0	2.08	0
P	0	0	0	P	0	0	2.17	P	0	2.08	0
Q	0	0	7.02	Q	6.52	0	2.17	Q	0	2.08	0
R	1.75	0	10.53	R	2.52	0	4.34	R	2.08	6.25	0
S	1.75	0	10.53	S	13.04	0	4.34	S	0	6.25	0
T	12.28	1.75	3.5	T	2.17	0	2.17	T	6.25	2.08	0
V	1.75	1.75	8.77	V	10.86	4.34	6.52	V	16.66	4.16	2.08
W	1.75	0	1.75	W	4.34	0	6.52	W	0	0	0
Y	10.52	0	0	Y	4.34	2.17	10.86	Y	0	0	0

Column 1st is indicating single amino acid codes. Column 2nd to 4th depicting % abundance of amino acid residues identified at positions 196 to 200 referred as p196 to p200, respectively.

Library #1 with randomization of Asn196-Ile197-Lys198, provided the following insight: 1) the 196th and 198th positions, natively having Phe196 and Lys198 pointing outwards are promiscuous and can tolerate a wide variety of residues; 2) the 197th position was found to be conserved to a greater extent with allowance of only four residues, Ile/Gly/Val/Thr, as illustrated in **Table 3-1**. Gly and Val are both smaller aliphatic amino acids that might not allow perfectly compact packing however, are not detrimental to FP folding or function. The presence of residues like Tyr at 197th position have been quoted in literature to be significant in shifting the emission to longer wavelength side, however, we did not observed

these in the library.^{10,77,82,83} We predicted that use of larger bandpass filters in screening of the libraries might have been inefficient to detect minor changes in fluorescence intensity. Use of NNK codon is also predicted to create large bias against peptides containing single codon amino acids.⁸⁵⁻⁸⁷

Screening of Library #2 with randomization of Lys198-Leu199-Asp200 revealed similar trends. The 198th and 200th positions were observed to allow a large variety of residues whereas only a few residues were allowed at the 199th position, as expected. Residues like Cys, Met and Tyr were observed at Leu199 and further characterized to check their potential impact on fluorescence emission wavelength. Variety of sequences which were predicted to have some impact on chromophore properties were propagated for further analysis. However, the spectral analysis did not reveal any major changes in the emission wavelength.

Library #3 with randomization of Ile197-Lys198-Leu199, lead to the identification of variants with buried aromatic residues like Phe (at position 197 and 199) and His (at 197). These substitutions could be expected to contribute to a red-shift of emission spectra owing to π - π stacking of their aromatic rings with phenolate of the chromophore. Indeed the fluorescence emission from variants with His at 197 were confirmed to facilitate red-shift of the emission peak (from 609 nm to 618 nm). The

sequences of all the screened variants corresponding to each library are included in Appendix VII. The fluorescence analysis of all selected variants further ascertained identification of mutations contributing to shifted emission peak. The **Table 3-2** depicts the best variants that we identified.

Table 3-2 Fluorescence analysis data of selected variants.

Code	196	197	198	199	200	Ex	Em
3_88	E	I	V	L	D	584	604
7_51	Y	I	L	L	D	585	604
Parent	F	I	K	L	D	584	609
4_23	F	T	H	L	D	585	614
7_30	F	T	K	L	D	584	615
5_21	F	H	K	L	D	584	618
7_155	L	T	L	L	D	585	618

Name of column 2nd to 6th (i.e., 196 to 200) denote amino acid position in mCherry corresponding to motif residues. Column 2nd to 6th is showing single amino acid code representation of identified motif residues. Column 7th and 8th depict excitation and emission wavelength (in nm) at which peak intensity was observed.

3.3.2 Generation of Ca²⁺ biosensor scaffold

In order to create a biosensor switch scaffold, we created eight different constructs with partially overlapping 10-residue long β -strand regions (designated β -10b) at the C-terminus (**Figure 3-3**). We additionally mutated the native β -10a strand to confer greater relative shift in emission peak, such that one strand would cause a red-shift while the alternate strand would cause a blue-shift when bound to the

remainder of the protein. The comparative motif residues designed in each strand are depicted in **Table 3-3**. The constructs were sub-cloned in place of FP region of R-GECO1 to yield a Ca²⁺-biosensor scaffold. Unfortunately, none of the corresponding protein products showed any appreciable fluorescence and thus further development of this system will be required.

Table 3-3 Sequences of corresponding frames designed to generate peak shift.

CODE	B-10a	B-10b
R1	LTLLD	EIVLD
R2	LTLLD	YILLD
R3	YILLD	LTLLD
R4	YILLD	FHKLD
R5	EIVLD	LTLLD
R6	EIVLD	FHKLD
R7	FHKLD	EIVLD
R8	FHKLD	YILLD

B-10a and B-10b columns depict single amino acid code representation of motif residues used in the construct in copy a and b, respectively.



Figure 3-3 Diagram depicting construct design.

3.4 Conclusion

We used fluorescence based screening of randomized libraries to identify mCherry variants, with mutations in a single β -strand that differ in emission maximum by as much as 14 nm. The development of these variants may serve as a primary step in the development of a TAFS (as well as AFF) design strategy. These results may also contribute to our understanding of possible roles for proximal residues in affecting chromophore properties. It is unlikely that the variants discovered here will prove optimal for the development of an RFP TAFS or AFF scaffold, since the colour differences are relatively small and would be difficult to detect on a fluorescence microscope that operates using band pass filters. We suggest that future efforts might focus on other strand positions that are known to interact with the chromophore.

Chapter 4:

Conclusion and future directions

The reagentless and generally reversible nature of FP-based biosensors makes them practically ideal reporters for real time biosensing and imaging applications. However, the challenge of rationally designing functional biosensors is currently preventing biologists from harnessing FP-based biosensors to their full potential. In an attempt to address this challenge, we attempted to develop a conformational switch-based design strategy in FPs.

Having reviewed the different types of conformational switching mechanisms intrinsic to naturally occurring signaling proteins, we hypothesized that a folding-unfolding type switching mechanism could be applied to FPs to convert them into simple switches. Accordingly, we adapted the principles of the previously reported AFF strategy to introduce a folding-unfolding type switch into avGFP. An additional β -strand with partially overlapping sequence was introduced in avGFP to generate a potential alternate frame for protein folding. In contrast to the previously reported examples of the AFF strategy, we tandemly fused the two copies of the target β -strand via a short linker. We have designated our modified AFF strategy as a 'Tandem Alternate Frame Switching' strategy.

The 'Tandem Alternate Frame Switching' strategy provides a potential new route for generating universally applicable reporters from avGFP. To test whether the FP construct with tandem alternate frames

exists in conformational equilibrium between the two possible states, we rationally generated 28 closely related constructs with tandem alternate frame introduced after β -7 strand. We employed insertion, deletion and mutation of residues in the frame to shift the equilibrium towards either one state or the other. In each case, we determined the relative proportions of the two states by using independent component analysis of absorbance spectrum. The fitting of absorbance data was done against control protein spectra that were created by intentionally deleting one of the two possible folding frames from the appropriate construct.

We verified that insertion of 4 residues and deletion of two or more residues from bound/stabilized strand can drastically destabilize the binding of respective strand. However, deletion of 1 residue and point mutation of residues to alanine (S147A in particular) is capable of mildly tuning the equilibrium by destabilizing the binding of respective strand to some extent. The combinations of these modifications can be fruitful in achieving the close to ideal equilibrium state.

To generate potential Ca^{2+} and cAMP biosensors, we fused an appropriate binding domain between the two frames. We evaluated excitation ratiometric response of the biosensors on binding of ligand. We thus provided proof-of-concept for a switch-based FP reporter that can

potentially serve as universal platform for the generation of new biosensors.

In an attempt to generate a similar RFP reporter switch, to complement the avGFP-based switch in multicolour *in vivo* imaging, we tried to employ the AFF strategy to RFP reporter. We explored the β -10 strand of cp196mCherry for introduction of alternate frame switch. Site directed mutagenesis and imaging system-based colony screening resulted in identification of trend of amino acids allowed. Pairs of variants with differences of up to 15 nm in emission peak maximum were identified. The presence of aromatic residues at 197th position was verified to cause a red-shift in emission peak, speculated to be due to a π - π stacking interaction of these aromatic side chains with the phenolate group of the chromophore. However, there is still a possibility of attaining even better change in readout. Consequently, alternate frame comprising of sequence capable of yielding emission peak 15 nm distant from the native frame were introduced at other termini. Efforts to use these variants in an AFF-type strategy for Ca^{2+} biosensing were unsuccessful.

There is a high demand for a generic FP reporter strategy wherein any MRE could be simply and rationally fused into a particular location to generate a biosensor. However progress towards this goal has been slow and uniformly disappointing to date. The work described in this thesis

provides some new principles for development of switch-based biosensors employing single FP as the generic switch-based reporter. While we believe these results are very promising, we do foresee a number of challenges that would be needed to overcome in order to provide improved generic FP reporter designs. The first challenge is the further generation of TAFS-compatible modifications that could provide predictable experimental control over the binding equilibrium and shift $\Delta G_{A \rightleftharpoons B}$ towards values that would offer the greatest chance of observing an analyte dependent change in ratio. We expect that, even using the modification explored in this thesis, there are opportunities for better design of the TAFS scaffold. For example, it is now apparent to us that including the Asn149Ala substitution in S7A of the version-2 design was a mistake. Version-2 was designed by Dr. Campbell in Fall of 2010 prior to the construction of the appropriate control constructs used for fitting the data. It was only after the spectra were analyzed by fitting with the appropriate controls that it was realized this mutation had a stabilizing effect. Future work will involve the creation of a version-3 design with the Asn149Ala mutation in the S7B.

The second challenge is the identification and testing of a greater number of binding domains that undergo large changes in their conformation upon binding to the analyte. A third challenge which will

eventually enable multicolour imaging is that these strategies will need to be extended to FPs of different colours, with red fluorescence being the most desirable second colour.

Once all aspects of generic FP TAFS strategy are established and employed to develop variants with both green and red emission, these indicators will be used for the *in vivo* real time monitoring of biochemistry happening in live cells. Such studies will further our understanding of the most basic mechanisms of cell biology and ultimately lead to improved treatments for disease.

Bibliography

1. Tainaka, K. *et al.* Design strategies of fluorescent biosensors based on biological macromolecular receptors. *Sensors* **10**, 1355–76 (2010).
2. Campbell, R. E. Fluorescent-protein-based biosensors: modulation of energy transfer as a design principle. *Analytical chemistry* **81**, 5972–9 (2009).
3. Shimomura, O., Johnson, F. H. & Saiga, Y. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea. *Journal of Cellular and Comparative Physiology* **59**, 223–239 (1962).
4. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–05 (1994).
5. Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G. & Cormier, M. J. Primary structure of the Aequorea victoria green-fluorescent protein. *Gene* **111**, 229–233 (1992).
6. Ormö, M. *et al.* Crystal structure of the Aequorea victoria green fluorescent protein. *Science* **273**, 1392–95 (1996).
7. Yang, F., Moss, L. G. & Phillips, G. N. The molecular structure of green fluorescent protein. *Nature biotechnology* **14**, 1246–51 (1996).
8. Cubitt, A. B. *et al.* Understanding, improving and using green fluorescent proteins. *Trends in biochemical sciences* **11**, 448–55 (1995).
9. Heim, R., Cubitt, A. B. & Tsien, R. Y. Improved green fluorescence. *Nature* **373**, 663–64 (1995).
10. Lin, M. Z. *et al.* Autofluorescent proteins with excitation in the optical window for intravital imaging in mammals. *Chemistry & biology* **16**, 1169–79 (2009).
11. Miyawaki, A. *et al.* Fluorescent indicators for Ca²⁺ based on green fluorescent proteins and calmodulin. *Nature* **388**, 882–87 (1997).

12. Baird, G. S., Zacharias, D. A. & Tsien, R. Y. Circular permutation and receptor insertion within green fluorescent proteins. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 11241–46 (1999).
13. Bravaya, K. B., Subach, O. M., Korovina, N., Verkhusha, V. V. & Krylov, A. I. Insight into the common mechanism of the chromophore formation in the red fluorescent proteins: the elusive blue intermediate revealed. *Journal of the American Chemical Society* **134**, 2807–14 (2012).
14. The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.
15. Brejc, K. *et al.* Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 2306–11 (1997).
16. Shu, X. *et al.* Ultrafast excited-state dynamics in the green fluorescent protein variant S65T/ H148D. 1. mutagenesis and structural studies. *Biochemistry* **46**, 12005–013 (2007).
17. Miyawaki, a Mechanisms of protein fluorophore formation and engineering. *Current Opinion in Chemical Biology* **7**, 557–562 (2003).
18. Matz, M. V. *et al.* Fluorescent proteins from nonbioluminescent Anthozoa species. *Nature biotechnology* **17**, 969–73 (1999).
19. Stepanenko, O. V. *et al.* Modern fluorescent proteins: from chromophore formation to novel intracellular applications. *BioTechniques* **51**, 313–27 (2011).
20. Zhang, J., Campbell, R. E., Ting, A. Y. & Tsien, R. Y. Creating new fluorescent probes for cell biology. *Nature reviews. Molecular cell biology* **3**, 906–18 (2002).
21. Jares-Erijman, E. a & Jovin, T. M. FRET imaging. *Nature biotechnology* **21**, 1387–95 (2003).
22. Ibraheem, A. & Campbell, R. E. Designs and applications of fluorescent protein-based biosensors. *Current opinion in chemical biology* **14**, 30–6 (2010).

23. Miyawaki, A., Griesbeck, O., Heim, R. & Tsien, R. Y. Dynamic and quantitative Ca²⁺ measurements using improved cameleons. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 2135–40 (1999).
24. Horikawa, K. *et al.* Spontaneous network activity visualized by ultrasensitive Ca(2+) indicators, yellow Cameleon-Nano. *Nature methods* **7**, 729–32 (2010).
25. Kraynov, V. S. *et al.* Localized Rac Activation Dynamics Visualized in Living Cells. *Science* **290**, 333–37 (2000).
26. Dormann, D., Weijer, G., Parent, C. A., Devreotes, P. N. & Weijer, C. J. Visualizing PI3 kinase-mediated cell-cell signaling during Dictyostelium development. *Current Biology* **12**, 1178–88 (2002).
27. Ai, H., Hazelwood, K. L., Davidson, M. W. & Campbell, R. E. Fluorescent protein FRET pairs for ratiometric imaging of dual biosensors. *Nature Methods* **5**, 401–03 (2008).
28. Kolossov, V. L. *et al.* Development of a high-dynamic range, GFP-based FRET probe sensitive to oxidative microenvironments. *Experimental biology and medicine* **236**, 681–91 (2011).
29. Komatsu, N. *et al.* Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Molecular biology of the cell* **22**, 4647–56 (2011).
30. Grynkiewicz, G., Poenie, M. & Tsien, R. Y. A new generation of Ca²⁺ indicators with greatly improved fluorescence properties. *The Journal of biological chemistry* **260**, 3440–50 (1985).
31. Fan, J.-Y. *et al.* Split mCherry as a new red bimolecular fluorescence complementation system for visualizing protein-protein interactions in living cells. *Biochemical and biophysical research communications* **367**, 47–53 (2008).
32. Do, K. & Boxer, S. G. Thermodynamics, kinetics, and photochemistry of β -strand association and dissociation in a split-GFP system. *Journal of the American Chemical Society* **133**, 18078–81 (2011).
33. Nausch, L. W. M., Ledoux, J., Bonev, A. D., Nelson, M. T. & Dostmann, W. R. Differential patterning of cGMP in vascular smooth

- muscle cells revealed by single GFP-linked biosensors. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 365–70 (2008).
34. Mizuno, T., Murao, K., Tanabe, Y., Oda, M. & Tanaka, T. Metal-ion-dependent GFP emission in vivo by combining a circularly permuted green fluorescent protein with an engineered metal-ion-binding coiled-coil. *Journal of the American Chemical Society* **129**, 11378–83 (2007).
 35. Chapleau, R. R., Blomberg, R., Ford, P. C. & Sagermann, M. Design of a highly specific and noninvasive biosensor suitable for real-time in vivo imaging of mercury (II) uptake. *Protein Science* **17**, 614–22 (2008).
 36. Frommer, W. B., Davidson, M. W. & Campbell, R. E. Genetically encoded biosensors based on engineered fluorescent proteins. *Chemical Society reviews* **38**, 2833–41 (2009).
 37. Wachter, R. M. & Remington, S. J. Sensitivity of the yellow variant of green fluorescent protein to halides and nitrate. *Current biology* **9**, R628–9 (1999).
 38. Jayaraman, S., Haggie, P., Wachter, R. M., Remington, S. J. & Verkman, a S. Mechanism and cellular applications of a green fluorescent protein-based halide sensor. *The Journal of biological chemistry* **275**, 6047–50 (2000).
 39. Miesenböck, G., Angelis, D. A. D. & Rothman, J. E. Visualizing secretion and synaptic transmission with pH- sensitive green fluorescent proteins. *Nature* **394**, 192–95 (1998).
 40. Ha, J.-H. & Loh, S. N. Protein conformational switches: from nature to design. *Chemistry - A European Journal* **18**, 7984–99 (2012).
 41. Porumb, T., Yau, P., Harvey, T. S. & Ikura, M. A calmodulin-target peptide hybrid molecule with unique calcium-binding properties. *Protein engineering* **7**, 109–15 (1994).
 42. Stratton, M. M. & Loh, S. N. Converting a protein into a switch for biosensing and functional regulation. *Protein science: a publication of the Protein Society* **20**, 19–29 (2011).

43. Yu, Y. & Lutz, S. Circular permutation: a different way to engineer enzyme structure and function. *Trends in biotechnology* **29**, 18–25 (2011).
44. Nakai, J., Ohkura, M. & Imoto, K. A high signal-to-noise Ca²⁺ probe composed of a single green fluorescent protein. *Nature biotechnology* **19**, 137–41 (2001).
45. Akerboom, J. *et al.* Crystal structures of the GCaMP calcium sensor reveal the mechanism of fluorescence signal change and aid rational design. *The Journal of biological chemistry* **284**, 6455–64 (2009).
46. Ohkura, M., Matsuzaki, M., Kasai, H., Imoto, K. & Nakai, J. Genetically encoded bright Ca²⁺ probe applicable for dynamic Ca²⁺ imaging of dendritic spines. *Analytical chemistry* **77**, 5861–69 (2005).
47. Tallini, Y. N. *et al.* Imaging cellular signals in the heart in vivo: Cardiac expression of the high-signal Ca²⁺ indicator GCaMP2. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 4753–8 (2006).
48. Tian, L. *et al.* Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature methods* **6**, 875–81 (2009).
49. Zhao, Y. *et al.* An expanded palette of genetically encoded Ca²⁺ indicators. *Science* **333**, 1888–91 (2011).
50. Nagai, T., Sawano, a, Park, E. S. & Miyawaki, a Circularly permuted green fluorescent proteins engineered to sense Ca²⁺. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 3197–202 (2001).
51. Vallée-Bélisle, A. & Plaxco, K. W. Structure-switching biosensors: inspired by Nature. *Current opinion in structural biology* **20**, 518–26 (2010).
52. Plaxco, K. W. & Soh, H. T. Switch-based biosensors: a new approach towards real-time, in vivo molecular detection. *Trends in Biotechnology* **29**, 1–5 (2011).

53. Stratton, M. M., Mitrea, D. M. & Loh, S. N. A Ca²⁺-sensing molecular switch based on alternate frame protein folding. *ACS chemical biology* **3**, 723–32 (2008).
54. Sack, J. S., Saper, M. a & Quioco, F. a Periplasmic binding protein structure and function. Refined X-ray structures of the leucine/isoleucine/valine-binding protein and its complex with leucine. *Journal of Molecular Biology* **206**, 171–91 (1989).
55. Medintz, I. L. & Deschamps, J. R. Maltose-binding protein: a versatile platform for prototyping biosensing. *Current opinion in biotechnology* **17**, 17–27 (2006).
56. Tanaka, T. Calmodulin-Dependent Calcium Signal Transduction. *Japan Journal of Pharmacology* **46**, 101–07 (1987).
57. Alexander, P. a, He, Y., Chen, Y., Orban, J. & Bryan, P. N. A minimal sequence code for switching protein structure and function. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 21149–54 (2009).
58. Luo, X. *et al.* The Mad2 spindle checkpoint protein has two distinct natively folded states. *Nature structural & molecular biology* **11**, 338–45 (2004).
59. Littler, D. R. *et al.* The intracellular chloride ion channel protein CLIC1 undergoes a redox-controlled structural transition. *The Journal of biological chemistry* **279**, 9298–305 (2004).
60. Bryan, P. N. & Orban, J. Proteins that switch folds. *Current opinion in structural biology* **20**, 482–8 (2010).
61. Dunker, a K., Silman, I., Uversky, V. N. & Sussman, J. L. Function and structure of inherently disordered proteins. *Current opinion in structural biology* **18**, 756–64 (2008).
62. Tompa, P. Unstructural biology coming of age. *Current opinion in structural biology* **21**, 419–25 (2011).
63. Wright, P. E. & Dyson, H. J. Linking Folding and Binding. *Current opinion in structural biology* **19**, 31–38 (2009).

64. Babu, M. M., van der Lee, R., de Groot, N. S. & Gsponer, J. Intrinsically disordered proteins: regulation and disease. *Current opinion in structural biology* **21**, 432–40 (2011).
65. Yoon, J. H. & Jhon, M. S. Molecular dynamics studies of the Ca²⁺ binding effect on calmodulin. *Journal of Molecular Structure* **295**, 193–201 (1993).
66. Babu, Y. S., Bugg, C. E. & Cook, W. J. Structure of calmodulin refined at 2.2 Å resolution. *Journal of molecular biology* **204**, 191–204 (1988).
67. Mitrea, D. M., Parsons, L. S. & Loh, S. N. Engineering an artificial zymogen by alternate frame protein folding. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 2824–9 (2010).
68. Shaner, N. C., Patterson, G. H. & Davidson, M. W. Advances in fluorescent protein technology. *Journal of cell science* **120**, 4247–60 (2007).
69. Elsliger, M. a, Wachter, R. M., Hanson, G. T., Kallio, K. & Remington, S. J. Structural and spectral response of green fluorescent protein variants to changes in pH. *Biochemistry* **38**, 5296–301 (1999).
70. Shi, X. *et al.* Ultrafast Excited-State Dynamics in the Green Fluorescent Protein Variant S65T / H148D. 2. Unusual Photophysical Properties. *Biochemistry* **46**, 12014–25 (2007).
71. Leiderman, P. *et al.* Ultrafast excited-state dynamics in the green fluorescent protein variant S65T/H148D. 3. Short- and long-time dynamics of the excited-state proton transfer. *Biochemistry* **46**, 12026–36 (2007).
72. Radley, T. L., Markowska, A. I., Bettinger, B. T., Ha, J.-H. & Loh, S. N. Allosteric Switching by Mutually Exclusive Folding of Protein Domains. *Journal of Molecular Biology* **332**, 529–536 (2003).
73. Kent, K. P. & Boxer, S. G. Light-activated reassembly of split green fluorescent protein. *Journal of the American Chemical Society* **133**, 4046–52 (2011).

74. Sharma, H., Yu, S., Kong, J., Wang, J. & Steitz, T. a Structure of apo-CAP reveals that large conformational changes are necessary for DNA binding. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 16604–9 (2009).
75. Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nature biotechnology* **22**, 1567–72 (2004).
76. Subach, O. M. *et al.* A photoswitchable orange-to-far-red fluorescent protein, PSmOrange. *Nature methods* **8**, 771–7 (2011).
77. Subach, F. V., Piatkevich, K. D. & Verkhusha, V. V. Directed molecular evolution to design advanced red fluorescent proteins. *Nature methods* **8**, 1019–26 (2011).
78. Chica, R. a, Moore, M. M., Allen, B. D. & Mayo, S. L. Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 20257–62 (2010).
79. Shu, X., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y. & Remington, S. J. Novel chromophores and buried charges control color in mFruits. *Biochemistry* **45**, 9639–47 (2006).
80. Li, Y., Sierra, A. M., Ai, H.-W. & Campbell, R. E. Identification of sites within a monomeric red fluorescent protein that tolerate peptide insertion and testing of corresponding circular permutations. *Photochemistry and photobiology* **84**, 111–9 (2008).
81. Carlson, H. J., Cotton, D. W. & Campbell, R. E. Circularly permuted monomeric red fluorescent proteins with new termini in the beta-sheet. *Protein science: a publication of the Protein Society* **19**, 1490–9 (2010).
82. Cheng, Z. & Campbell, R. E. Assessing the structural stability of designed beta-hairpin peptides in the cytoplasm of live cells. *Chembiochem: a European journal of chemical biology* **7**, 1147–50 (2006).
83. Wachter, R. M., Elsliger, M. a, Kallio, K., Hanson, G. T. & Remington, S. J. Structural basis of spectral shifts in the yellow-

- emission variants of green fluorescent protein. *Structure* **6**, 1267–77 (1998).
84. Strack, R. L. *et al.* A rapidly maturing far-red derivative of DsRed-Express2 for whole-cell labeling. *Biochemistry* **48**, 8279–81 (2009).
 85. Cwirla, S. E., Peters, E. a, Barrett, R. W. & Dower, W. J. Peptides on phage: a vast library of peptides for identifying ligands. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 6378–82 (1990).
 86. Scott, J. K. & Smith, G. P. Searching for peptide ligands with an epitope library. *Science* **249**, 386–90 (1990).
 87. Tominaga, T. & Hatakeyama, Y. Determination of Essential and Variable Residues in Pediocin PA-1 by NNK Scanning. *Applied and Environmental Microbiology* **72**, 1141–47 (2006).
 88. Thevenot, D. R., Toth, K., Durst, R. A. & Wilson, G. S. Electrochemical biosensors: recommended definitions and classification. *Biosensors & bioelectronics* **16**, 121–31 (2001).

Appendix I: Primer sequences

S. No.	Name of Primer	Sequence (5' - 3')
1	RB7_ECORI 3_R	GCAGAATTCTTACTTGTACAGCTCGTCCATG
2	RB7_KPN1_ F	GGCGGTACCGCCGACGAGCAGAAGAAC
3	Armv_Xho1 _r	GCCGCTCGAGGTGTACTCCAGCTTGTGCCCCAGGATG TTGCC
4	BH1b7F	ATTGGATCCCACCATGGTGAGCAAGGGCGAGGAG
5	Prsetb_seq _f	TAATACGACTCACTATAGGG
6	Rsetb_r_se q	CCTTTCGGGCTTTGTTAGCAGCCGGAT
7	RS_BAD_F	ATGCCATAGCATT TTTATCC
8	RS_BAD_R	GATTTAATCTGTATCAGG
9	FPR_Hin_C P196	CAGCCAAGCTTTCCTTCTGTCTCAGTTGGTCACC

10	FPF_Xho_I KL_3	TCCGAGCTCGAGCATGTCCAGCTCCGGGGCACTGGTC GGTTTTNNKNNKNNKGACATCACCTCACACAACGAG
11	FPF_Xho_K LD_LIB	TCCGAGCTCGAGAAGCTCGAGCATGTCCAGCTCCGGG GCACTGGTCGGTTTTATCANNKNNKNNKATCACCTCAC ACAACGAGGAC
12	FPF_Xho_FI K_2	TCCGAGCTCGAGCATGTCCAGCTCCGGGGCACTGGTC GGTNNKNNKNNKTTGGACATCACCTCACACAACGAG
13	EIV_b10b_ mlu1_R	GAATTCTCACTCTTCTGTCAGTTGGTCACGCGTATGG CTGGTAATATCCAGGACGATTTGACGTTGTAAGCGC CGGGCAG
14	HKL_b10b_ mlu1_R	GAATTCTCACTCTTCTGTCAGTTGGTCACGCGTATGG CTGGTAATATCCAGCTTGTGAAAGACGTTGTAAGCGC CGGGCAG
15	YIL_b10b_ mlu1_R	GAATTCTCACTCTTCTGTCAGTTGGTCACGCGTATGG CTGGTAATATCCAGCAGGATATAGACGTTGTAAGCGC CGGGCAG
16	LTL_b10b_ mlu1_R	GAATTCTCACTCTTCTGTCAGTTGGTCACGCGTATGG CTGGTAATATCCAGCAGGGTCAGGACGTTGTAAGCGC CGGGCAG
17	FPMACF_Sa cl_1	TCCGAGGAGCTCCATGTCCAGCTCCGGGGCACTGGTC GGT
18	5'BH1(GAT) kzkgfp1	ATTGGATCCCACCATGGTGAGCAAGGGCGAGGAG
19	GFP238Eco R13'	CATGGACGAGCTGTACAAGTAAGAATTCTGC
20	REC001	GCCGCTCGAGGTTTTAATGTACACATTGTGGCTATTA TAGTTGTA CTCCAGCTTGTGC

21	REC002	ACCTCGAGCGGCGGTACCAGCAATTACAACCTCGCATA ACGTCTATATCAAGGCCGAC
22	A- H148D_Xho 1	GCCGCTCGAGGTTTTAATGTACACATTGTCGCTATTA TAGTTGTA
23	B- H148D_Kpn 1	CGGCGGTACCAGCAATTACAACCTCGGATAACGTCTAT ATCAAG
24	4AH_BglII_ F	GGC AGA TCT GGG AAC TAT AAT AGC CAC AAT GTG TAC
25	4A_BglII_R	CCC AGA TCT GCC GTA CTC CAG CTT GTG C
26	4B_Sac1_F	GGG AGC TCG GGC GCC GAC GAG CAG AAG AAC
27	4BD_Sac1_ R	GCC CGA GCT CCC CTT GAT ATA GAC GTT ATC CGA GTT GTA ATT G
28	AdelK	GCC GCT CGA GGT AAT GTA CAC ATT GTG GCT ATT ATA G
29	AdelIK	GCC GCT CGA GGT GTA CAC ATT GTG GCT ATT ATA GTT G
30	AdelYIK	GCC GCT CGA GGT CAC ATT GTG GCT ATT ATA GTT GTA C
31	Xho1_CaM_ F	AAAACCTCGAGCGACCAACTGACAGAAGAGCAG

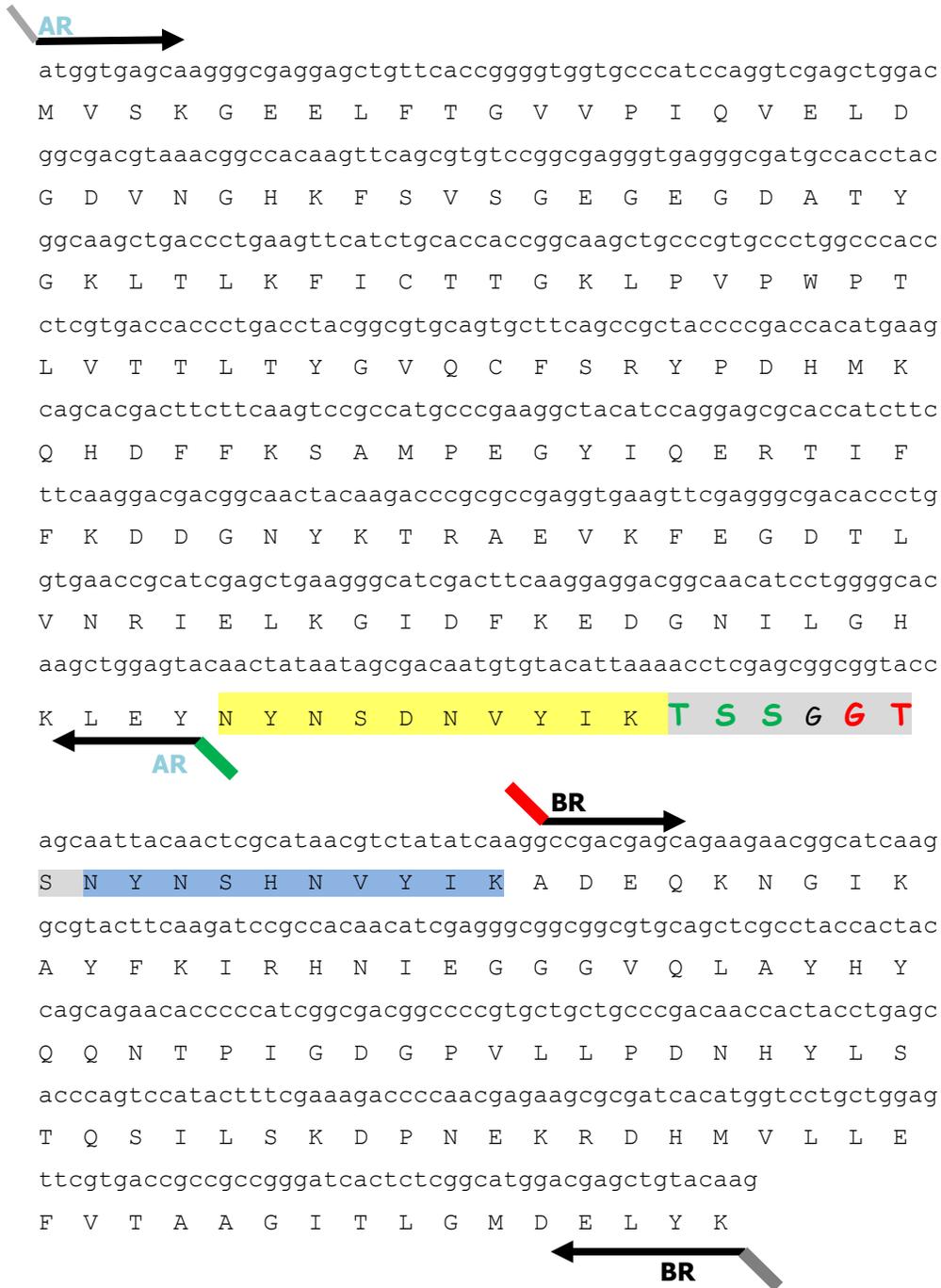
32	Kpn1_M13_R	TGCTGGTACCCAGTGCCCCGGAGCTG
33	Kpn1_BD_S147A	GGCGGTACCAGCAATTACAACGCGGATAACGTCTATA TCAAGGCCGAC
34	Kpn1_BD_S147A_N149A	GGCGGTACCAGCAATTACAACGCGGATGCCGTCTATA TCAAGGCCGAC
35	RC011	ATGTAGACGTTGTGCGAGTTGTACACAGCGTCGCTAT TATAGTTGTA
36	RC012	ACTCGCACAACGTCTACATCTCGAGCGGCGGTACCAA GGCCGACGAGCAG
37	RC013	GTTGTAGGTACCGCCGCTCGAGGTGTACACAGCGTC GCTATTATAG
38	RC014	GAGCGGCGGTACCTACAACCTCGCACGCCGTCTACATC AAGGCCGACG
39	Xho1_CAP_F	AAAACCTCGAGCGTGCTTGGCAAACCGCAAAC
40	Kpn1_CAP_R	TGCTGGTACCACGAGTGCCGTAAACGACG
41	v2_delT_Xho_RV	CCGCTCGAGTACACAGCGTCGCTATTATAGTTG
42	v2_delYT_Xho_RV	CCGCTCGAGACAGCGTCGCTATTATAGTTG

43	v2_V150A_ delYT_Xho_ RV	CCGCTCGAGGCAGCGTCGCTATTATAGTTG
44	v2_B_A149 N	CACCTCGAGCGGCGGTACCTACAACCTCGCACAACGTC TACATCAAGGCC
45	v2_B_iN144	CACCTCGAGCGGCGGTACCAATTACAACCTCGCACGCC G
46	v2_B_A149 N_iN144	CACCTCGAGCGGCGGTACCAATTACAACCTCGCACAAC GTC

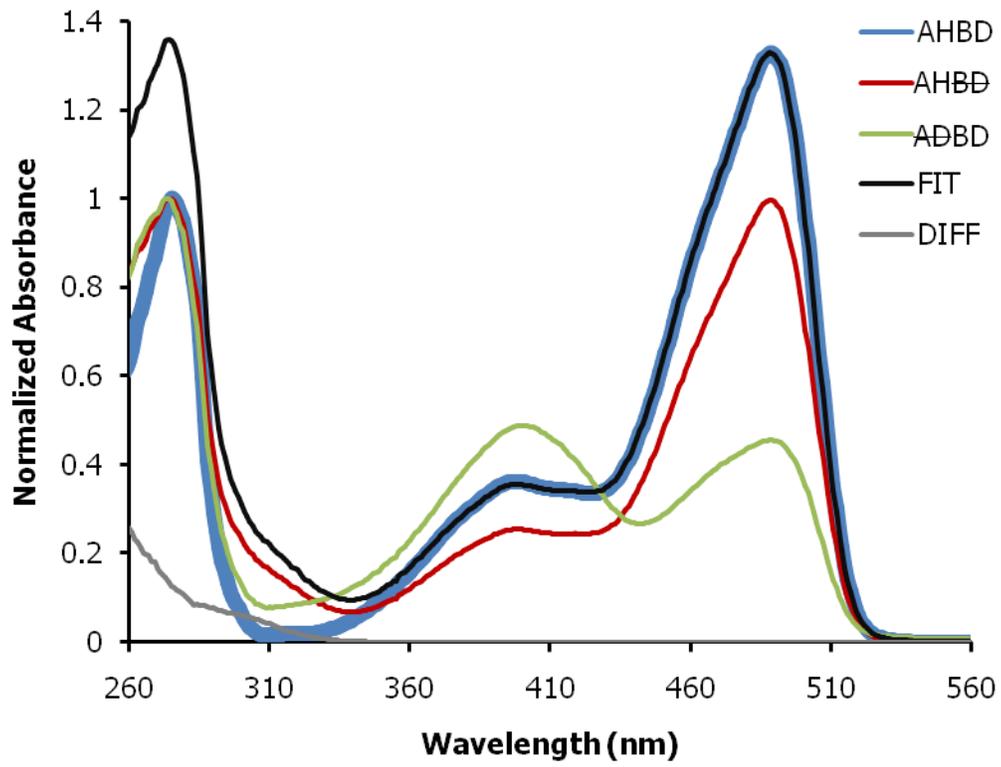
Appendix II: Construction strategy

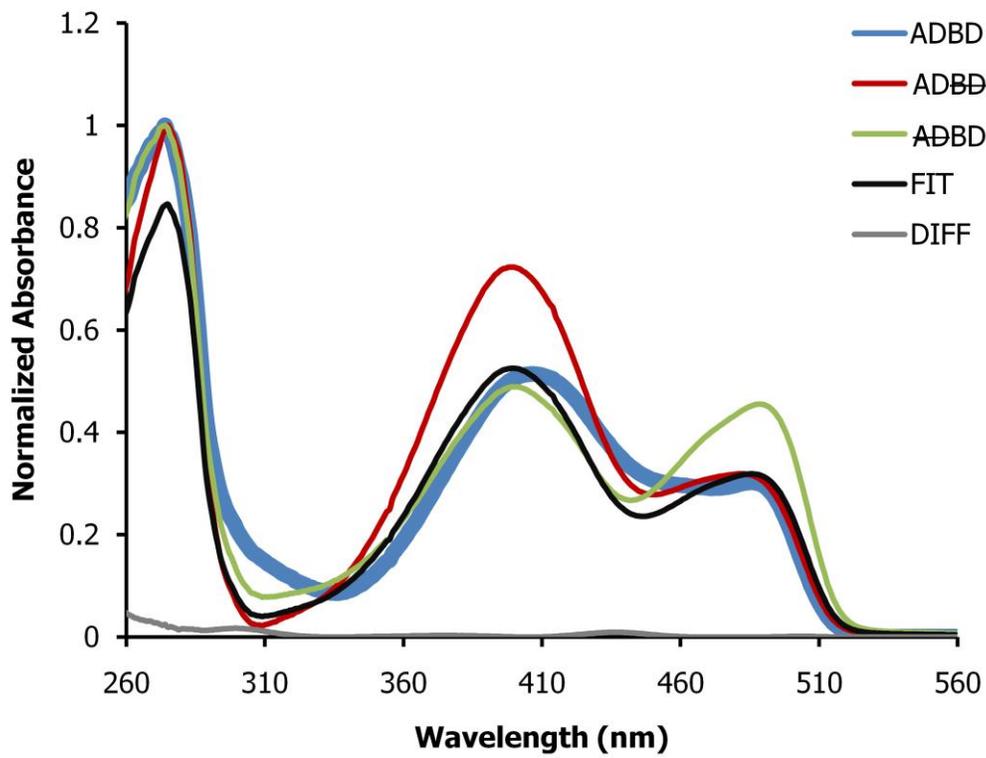
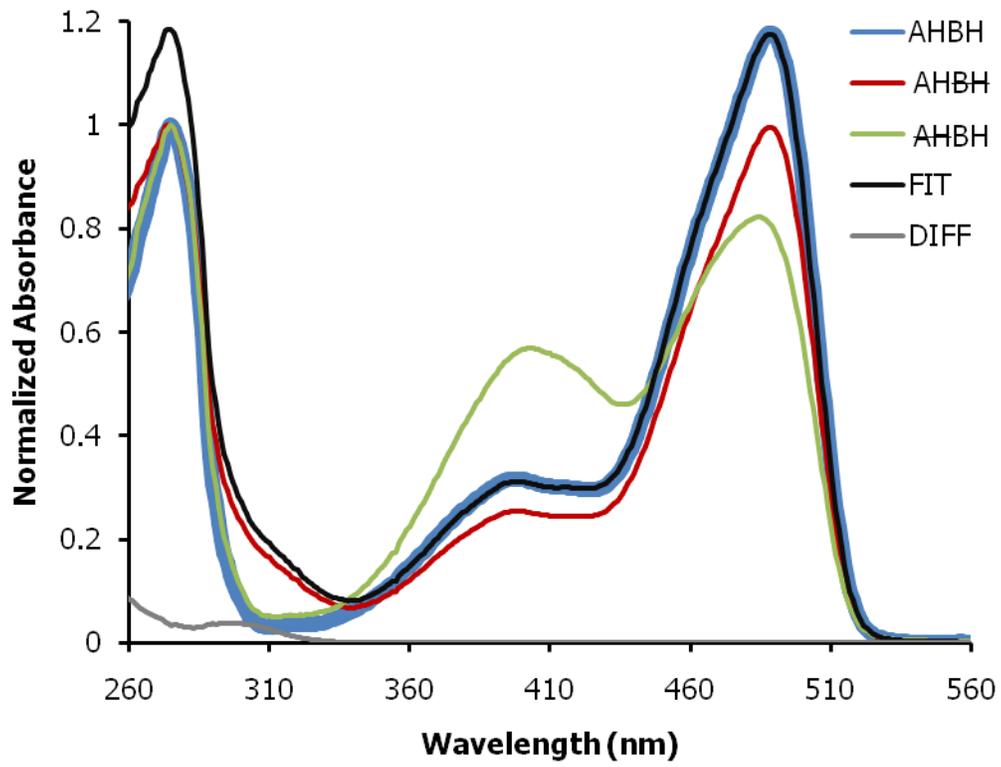
Sequence of (CS1) ADBH TAFS:

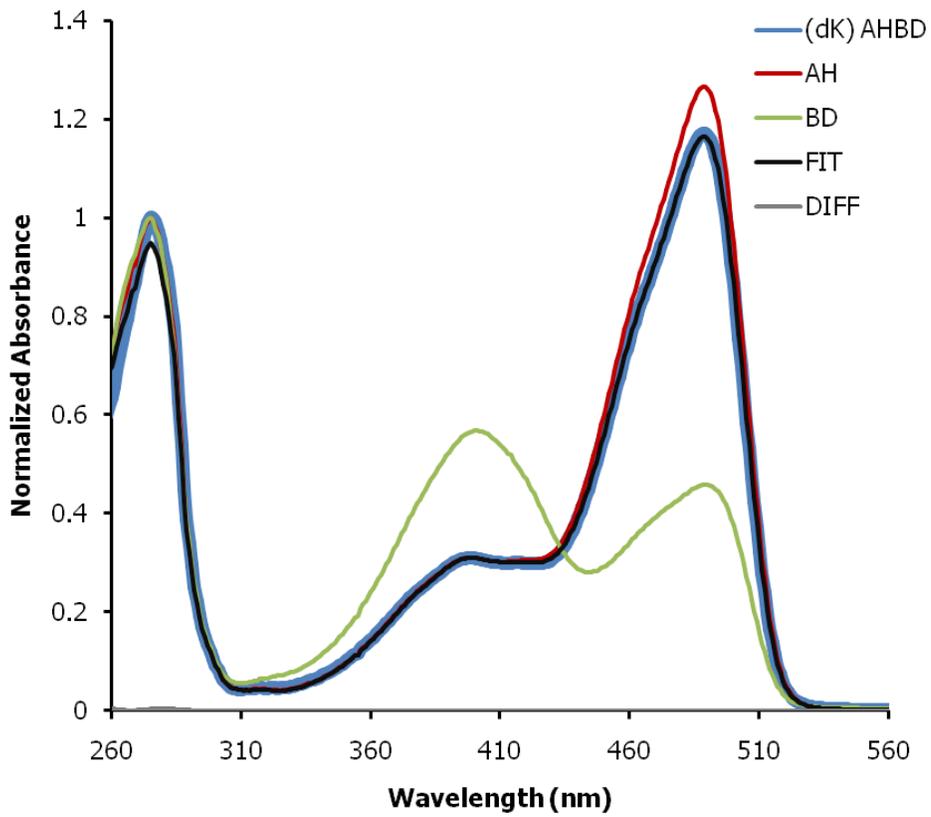
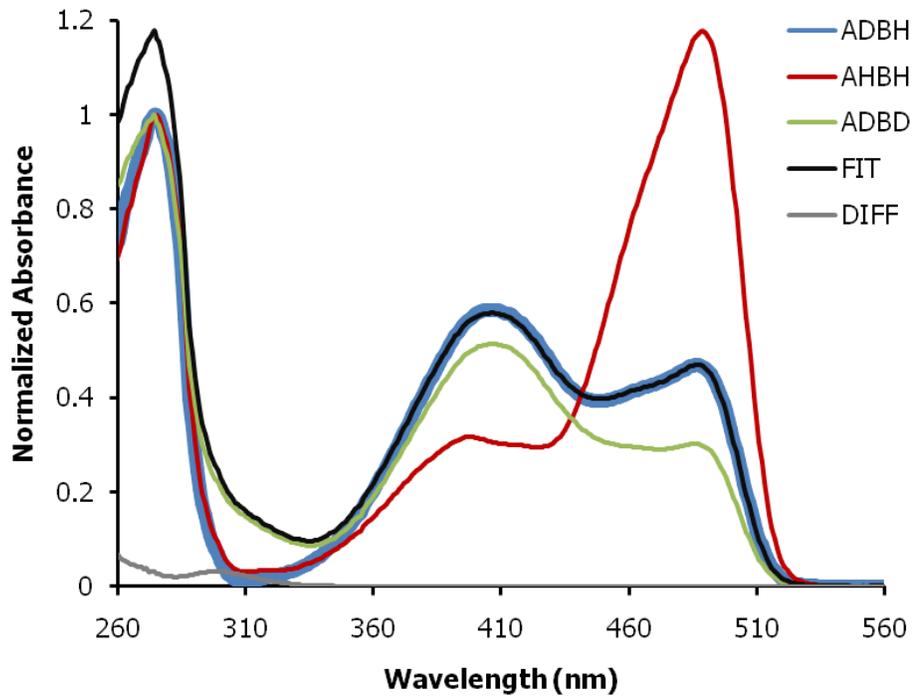
- Native S7A Frame  AR = S7A removal
- Linker with restriction sites  **TSS** = XhoI; **GT** = KpnI
- Alternate S7B Frame  BR = S7B removal

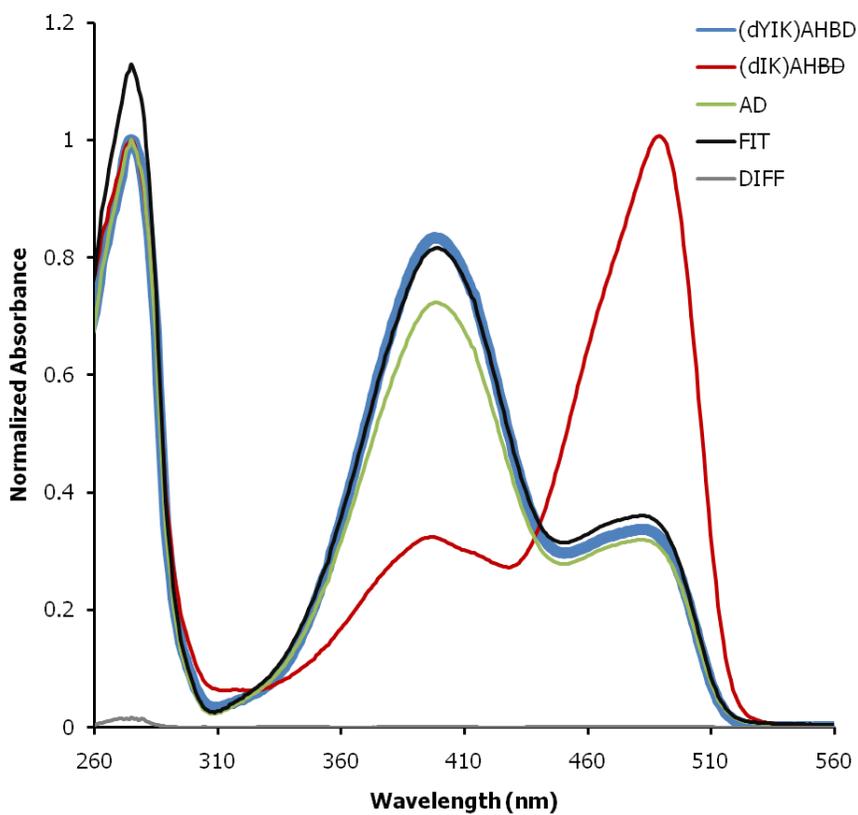
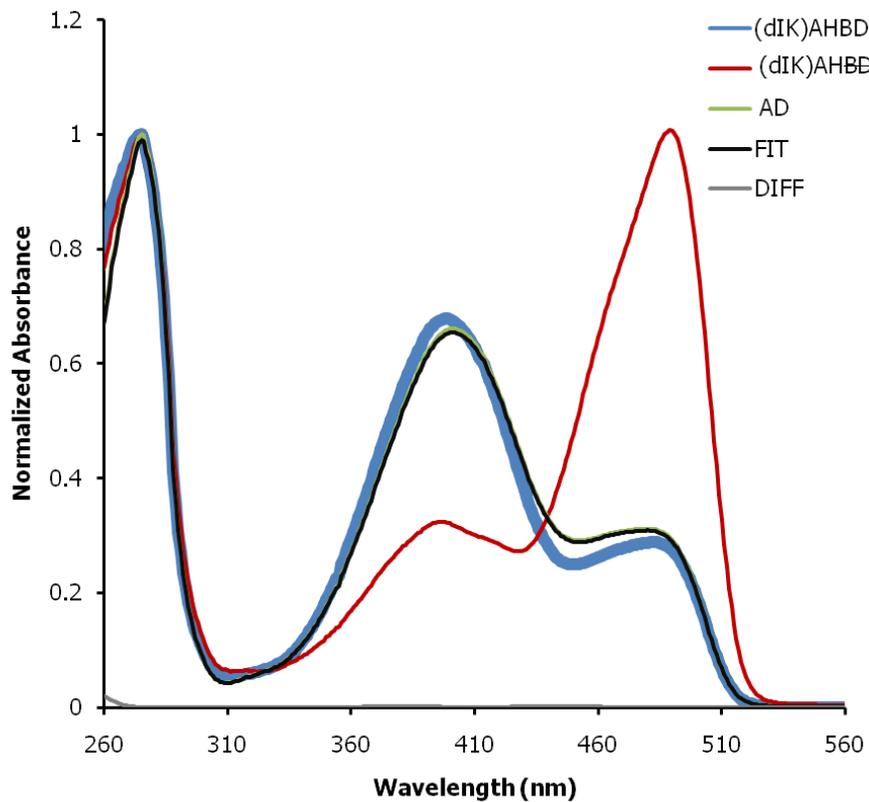


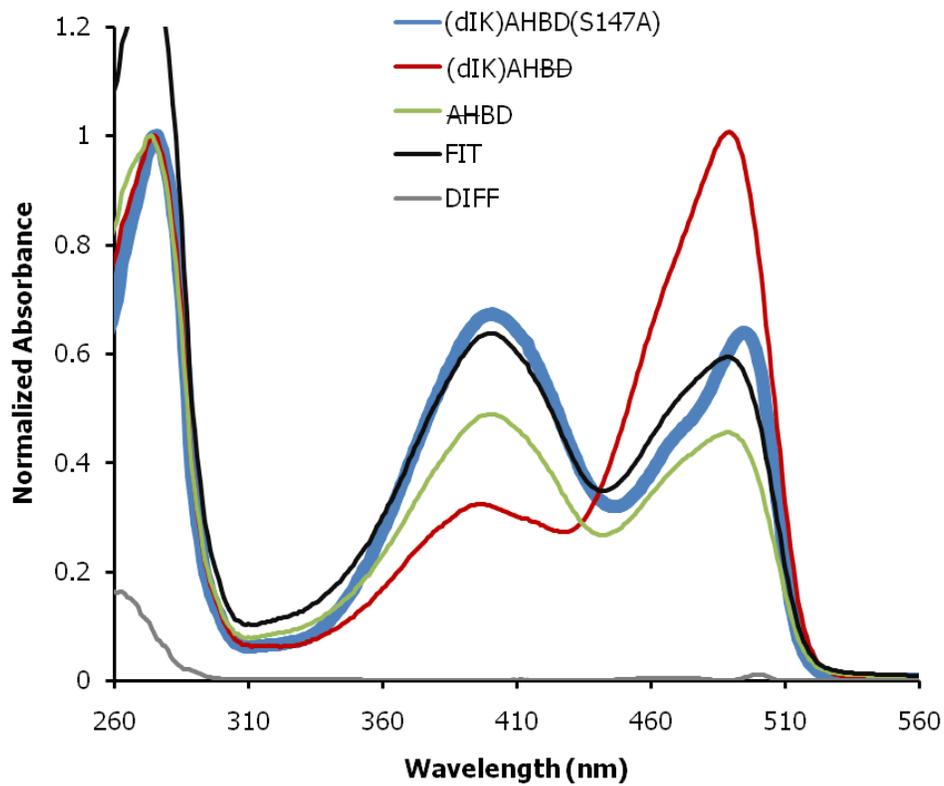
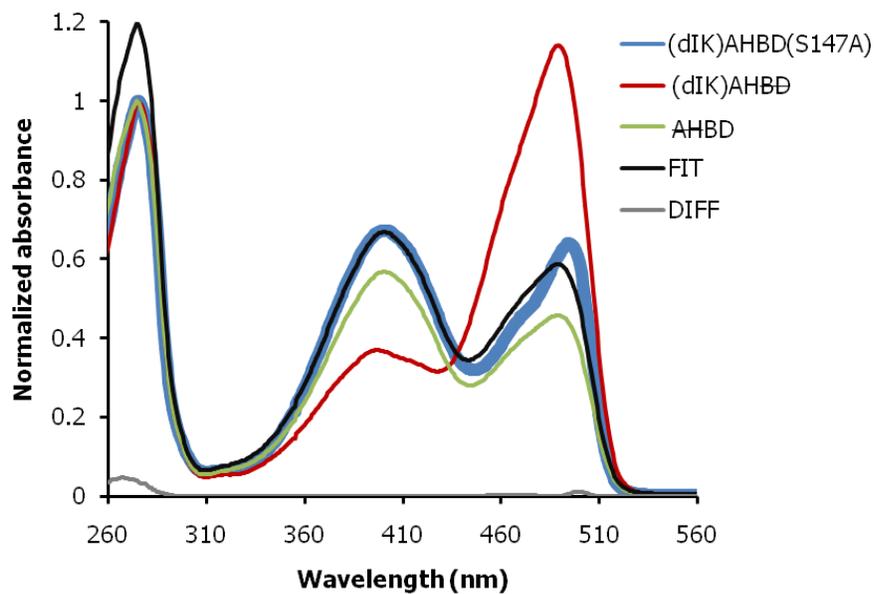
Appendix III: Fitting Curves

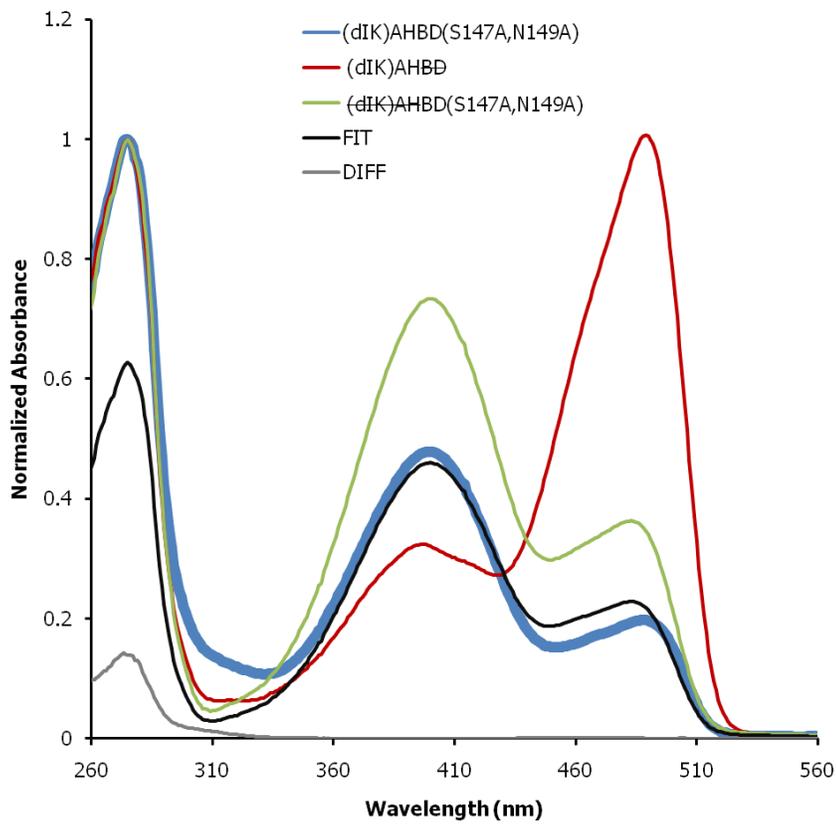
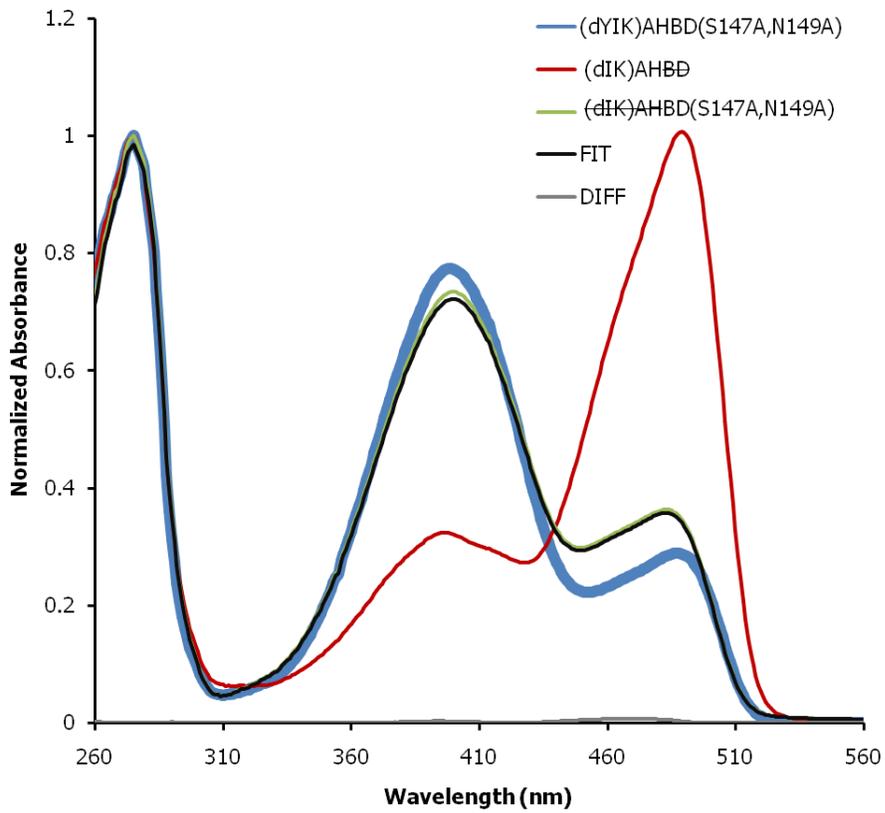


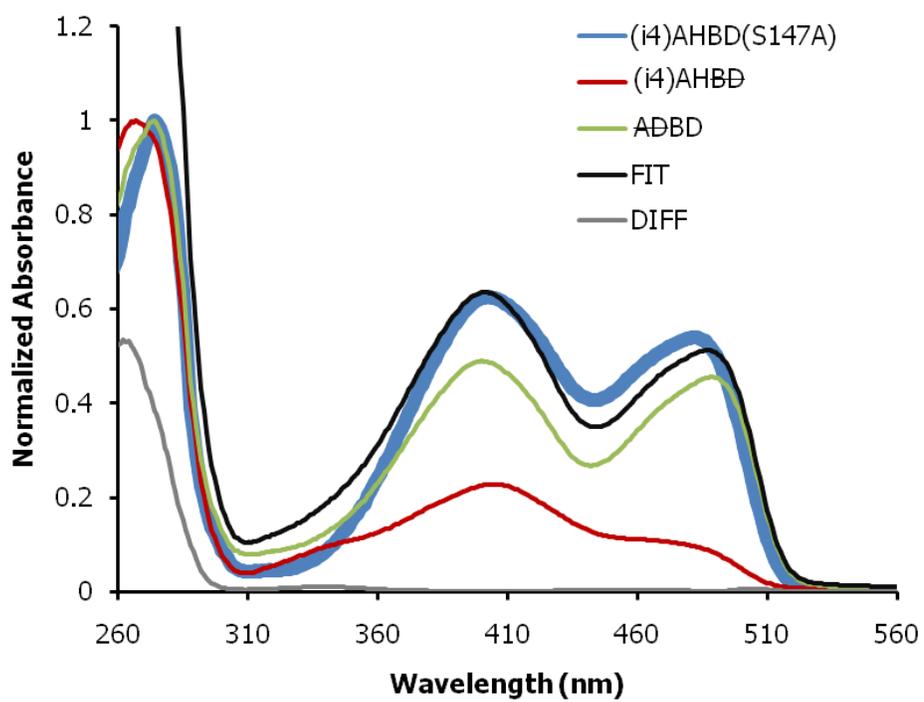
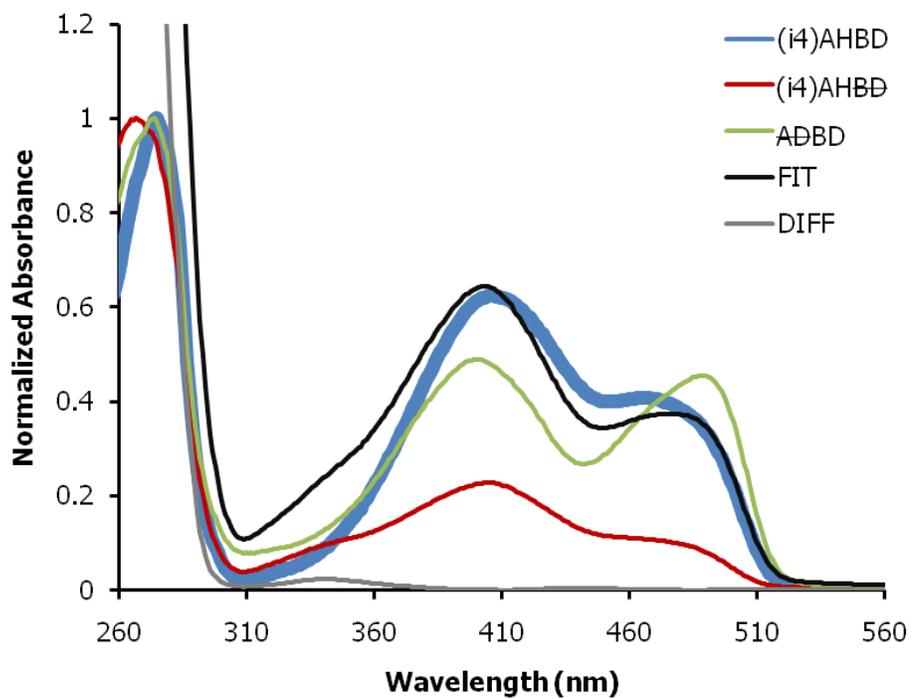


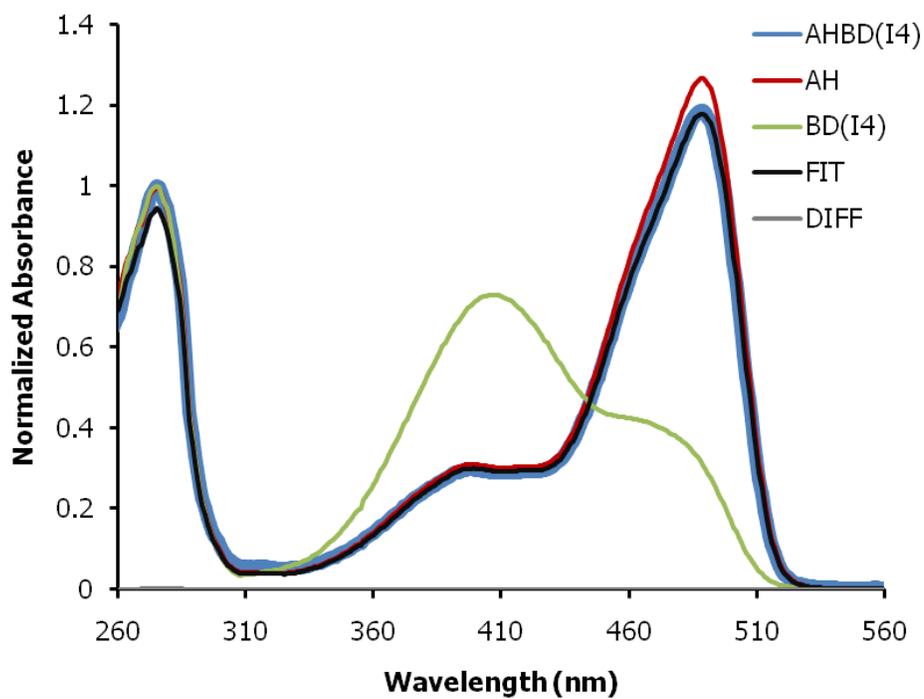
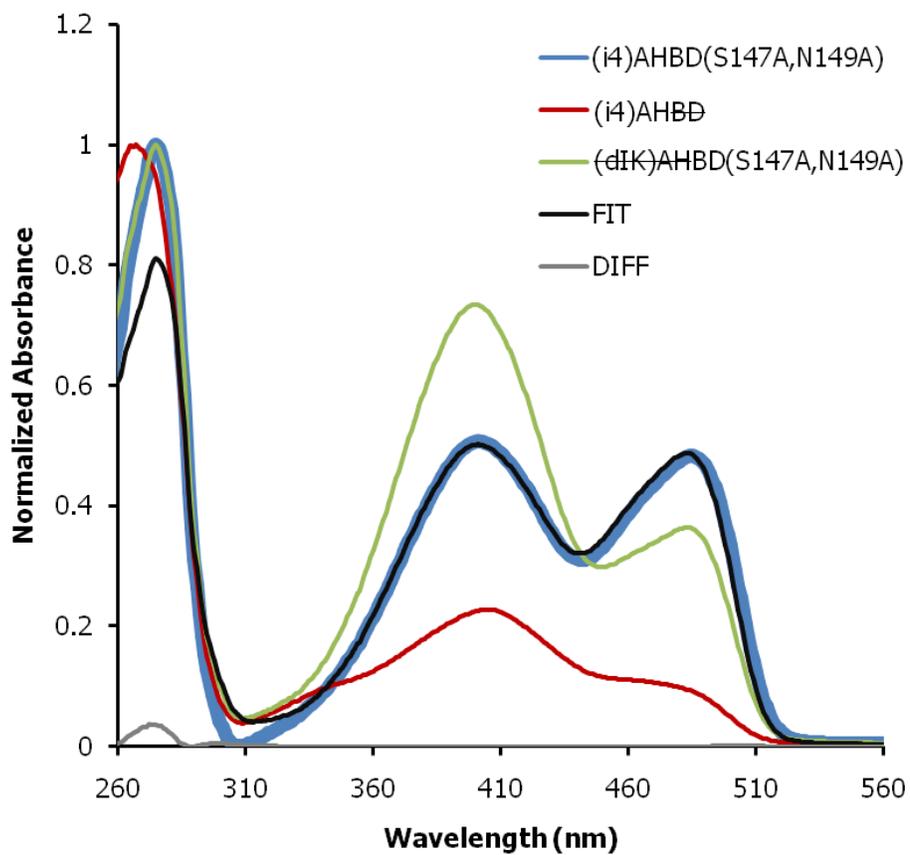


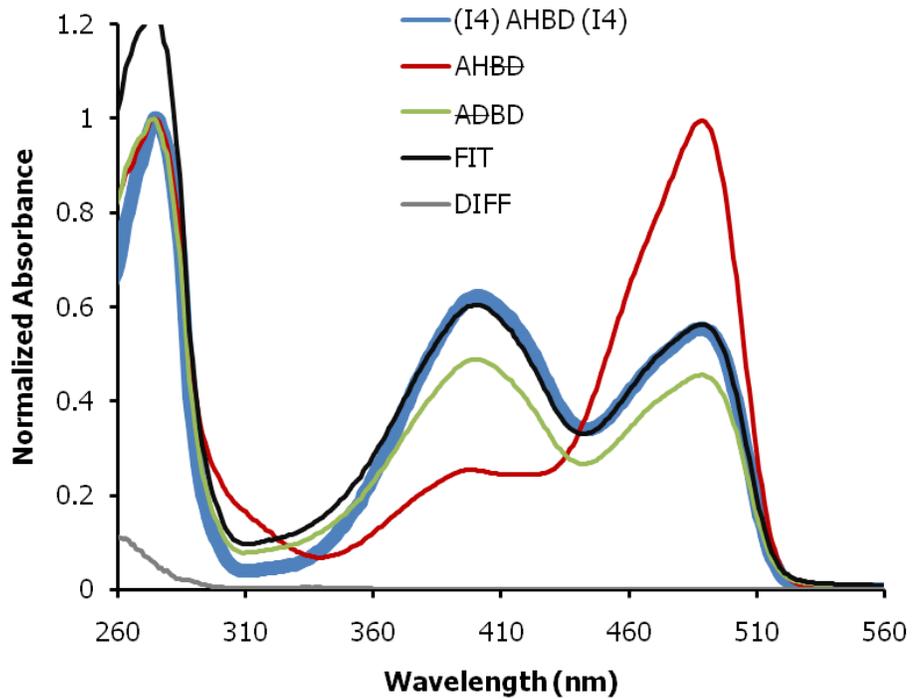
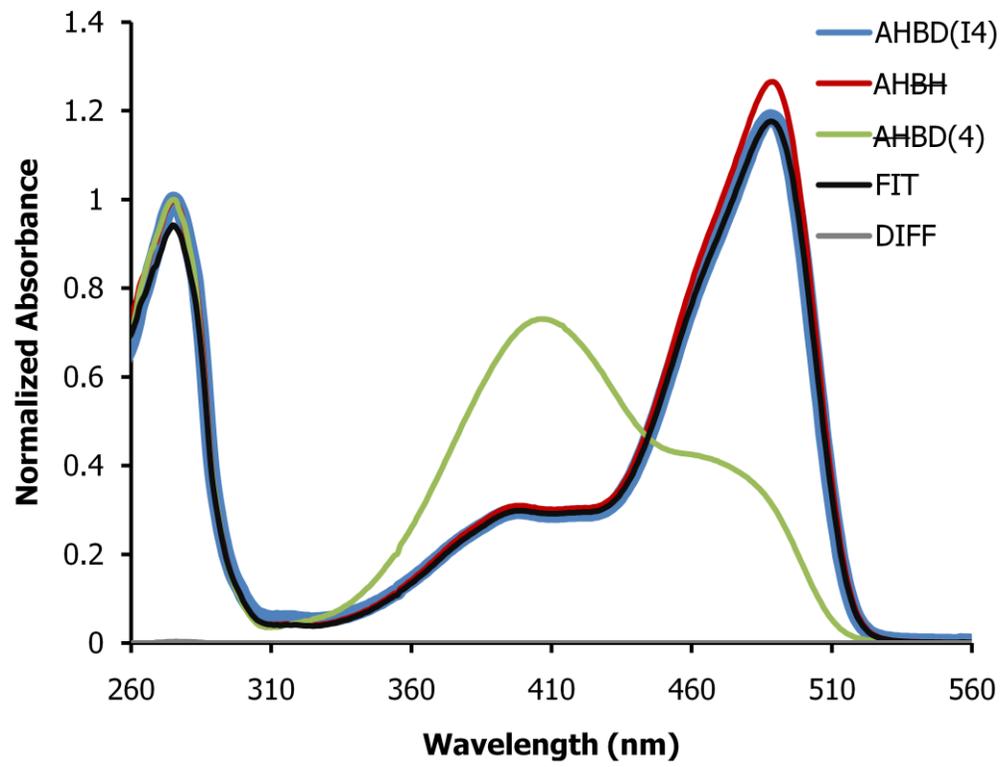


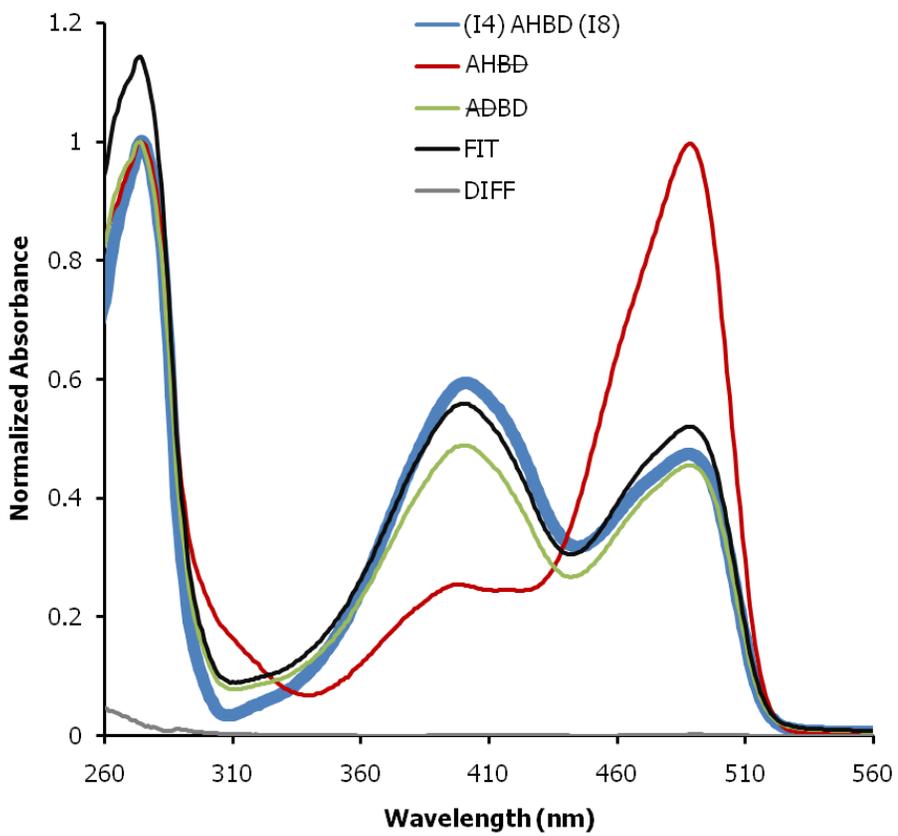
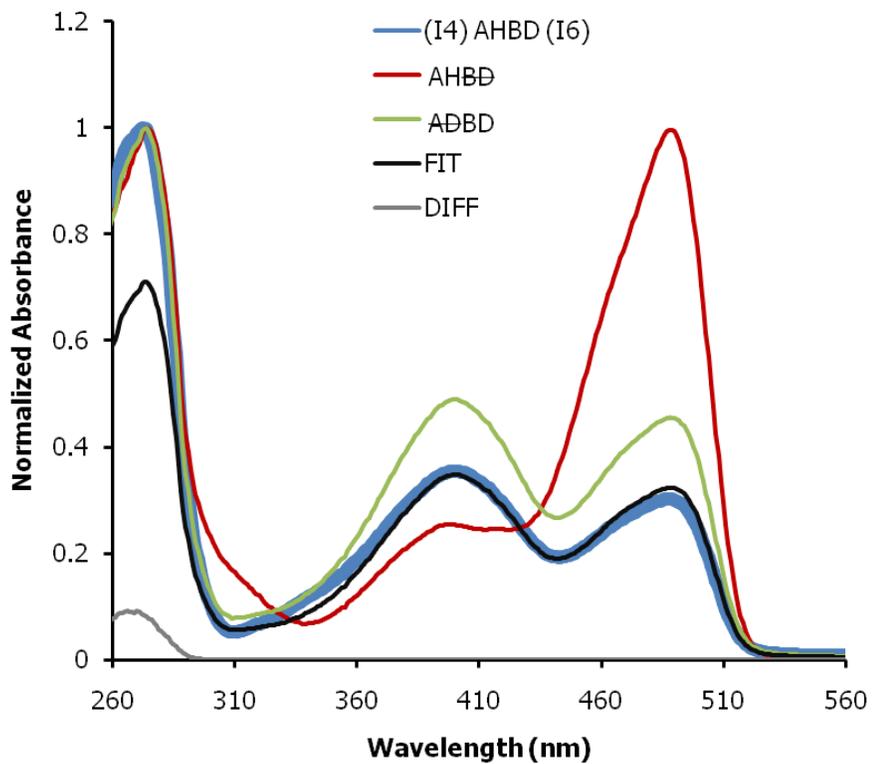


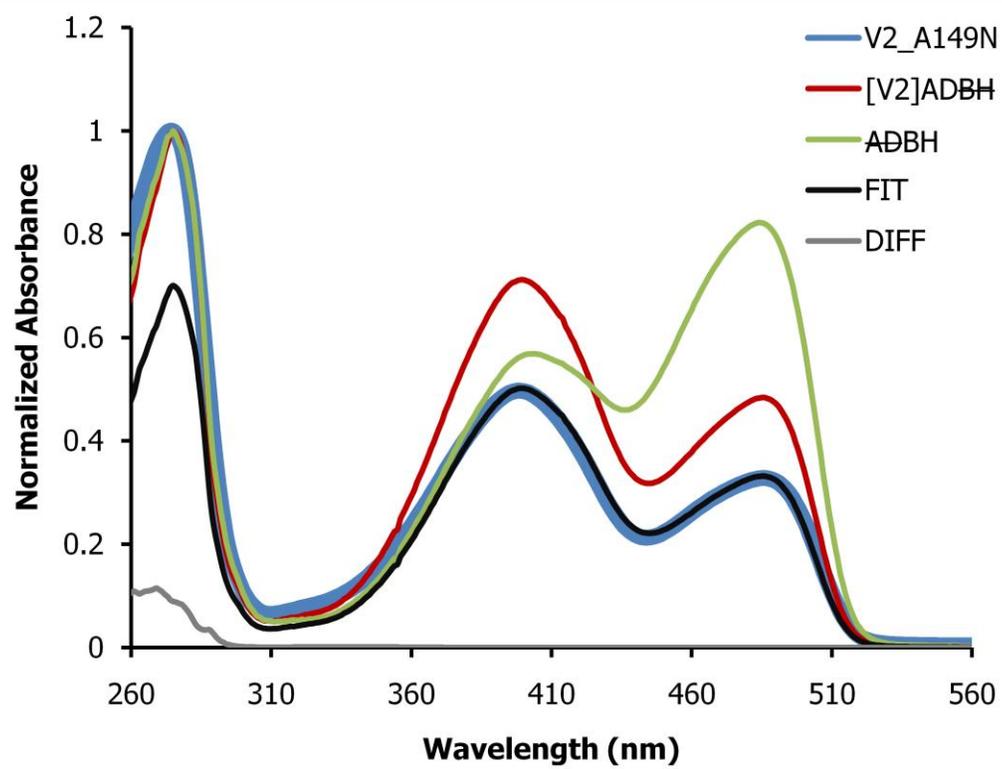
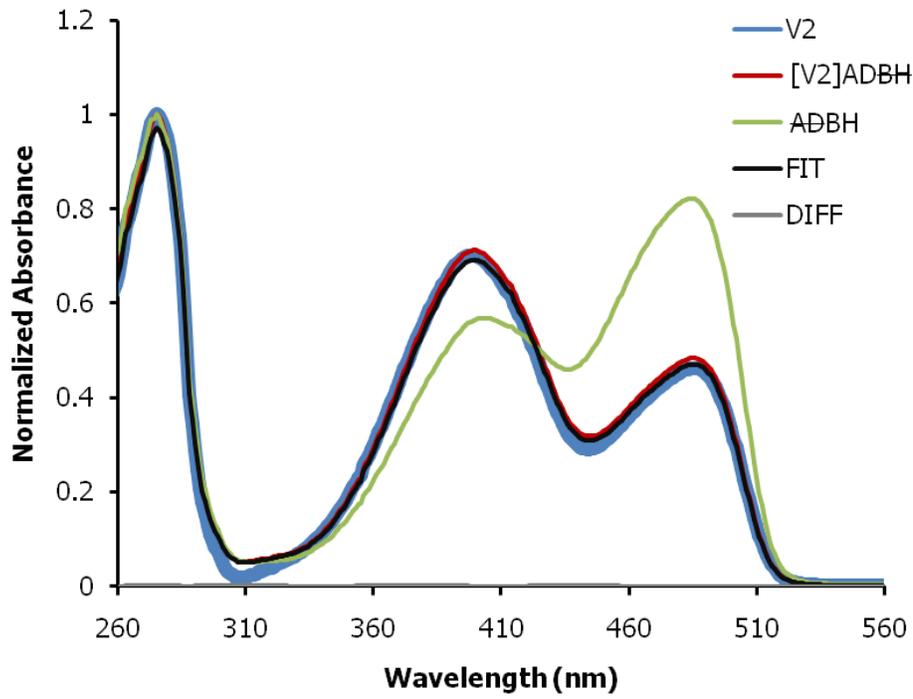


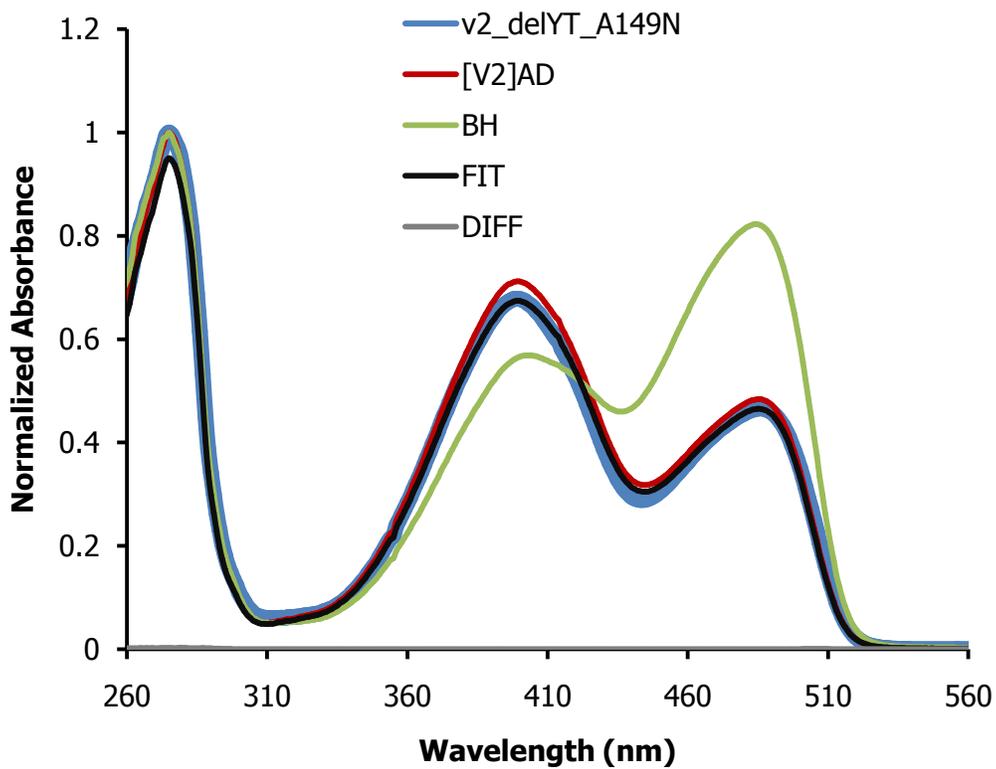
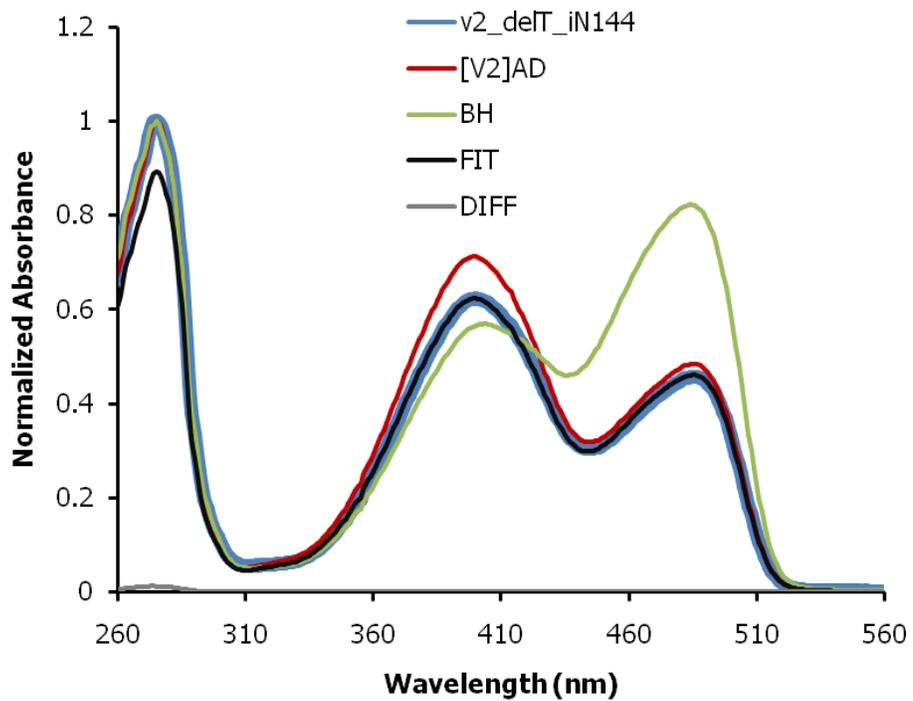


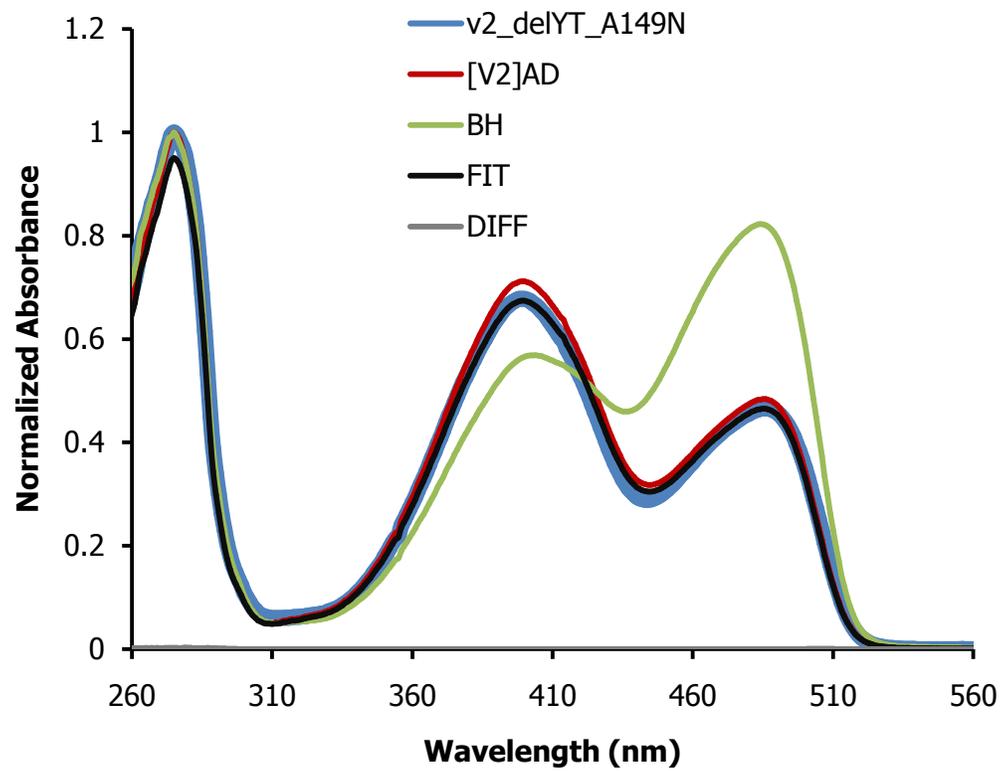
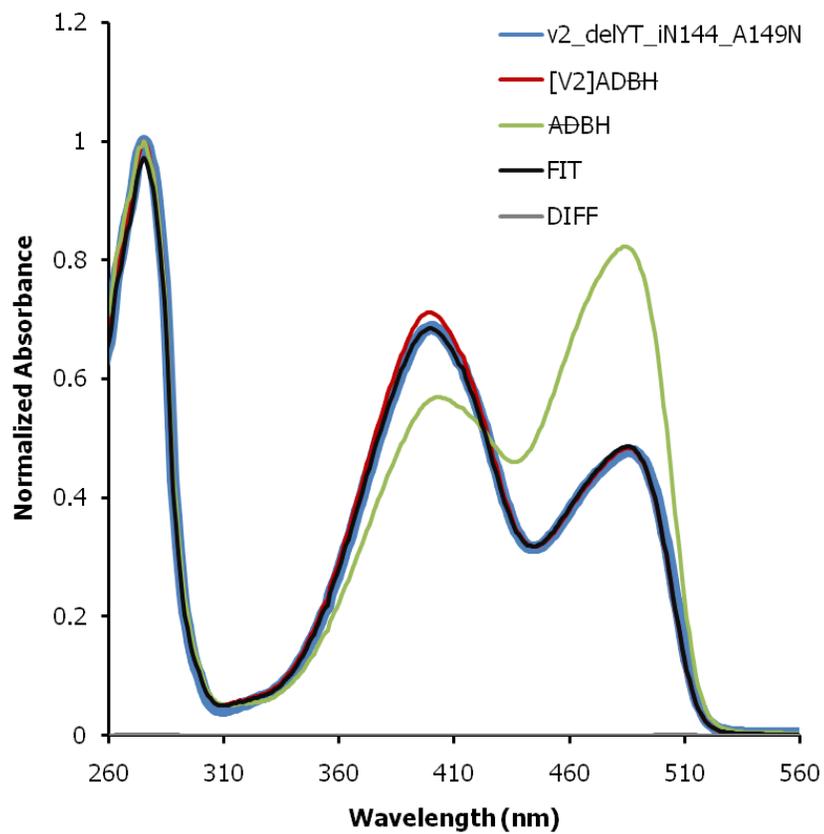


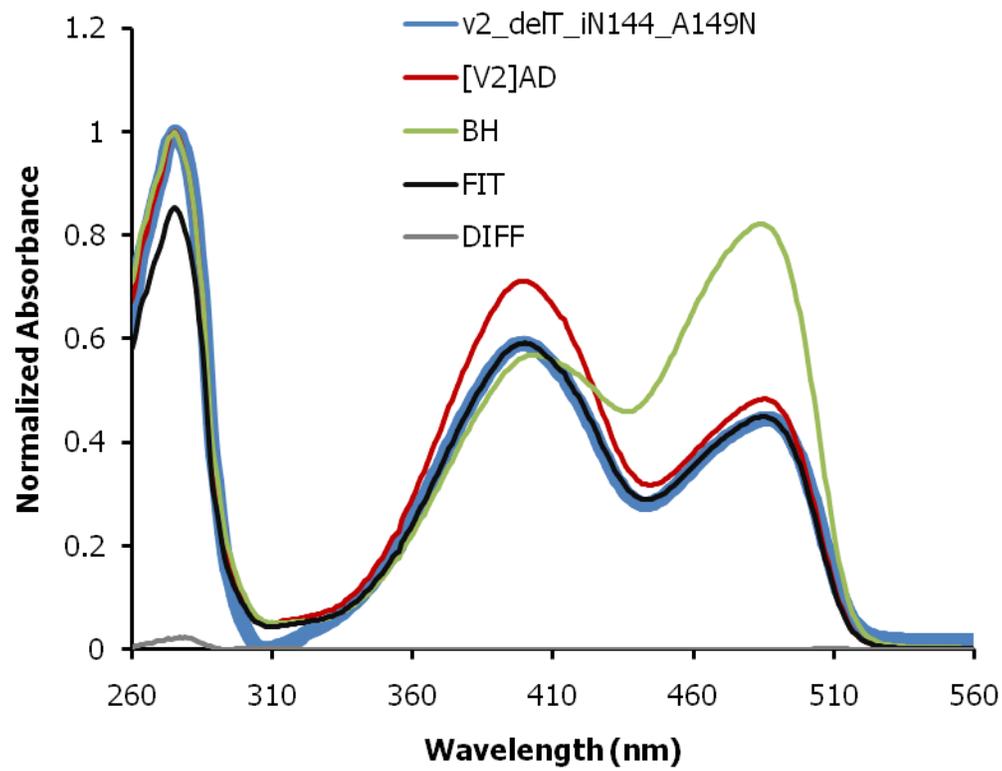


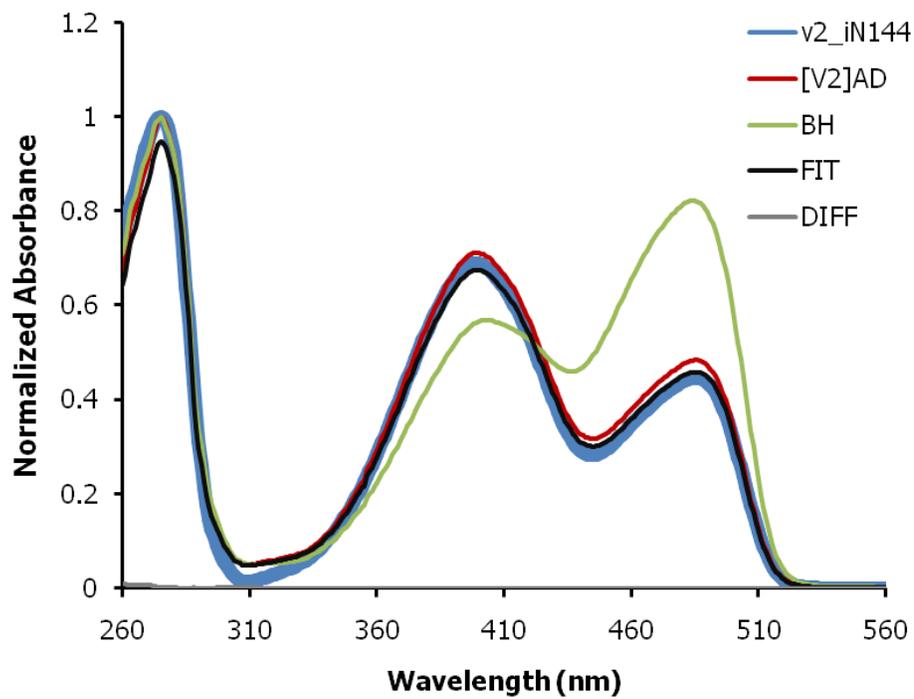
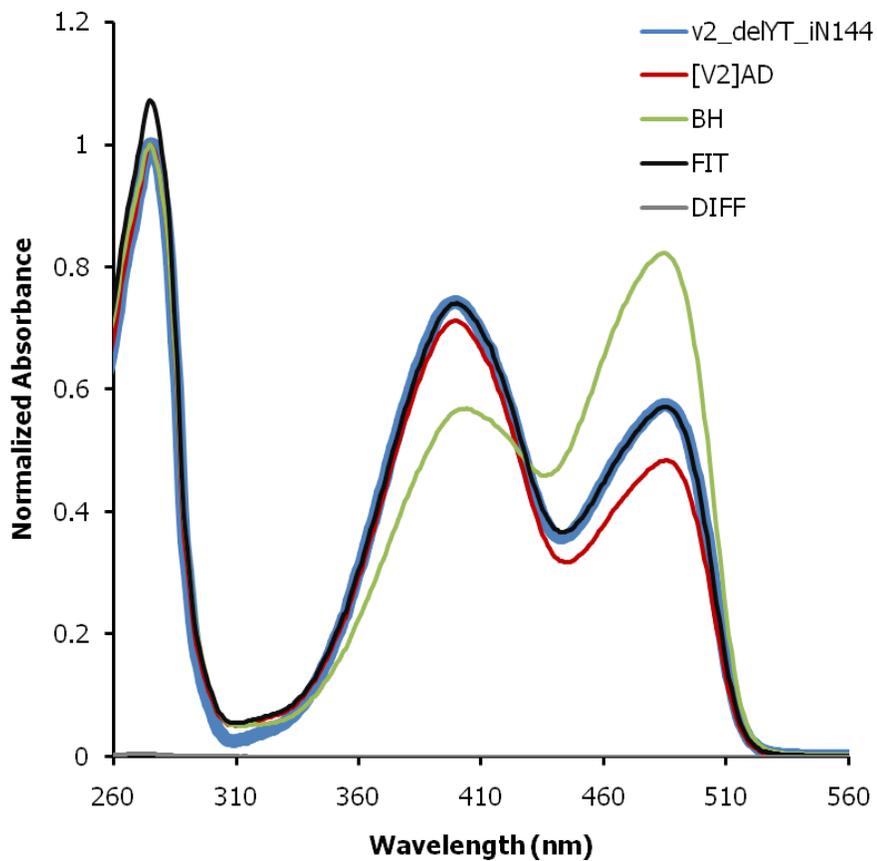


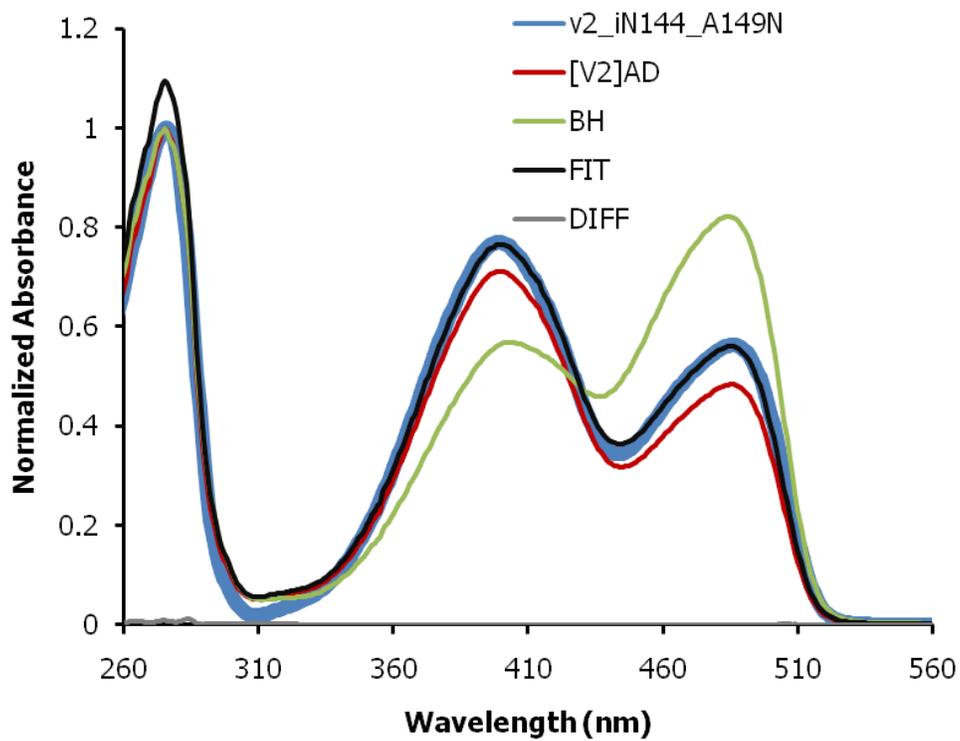
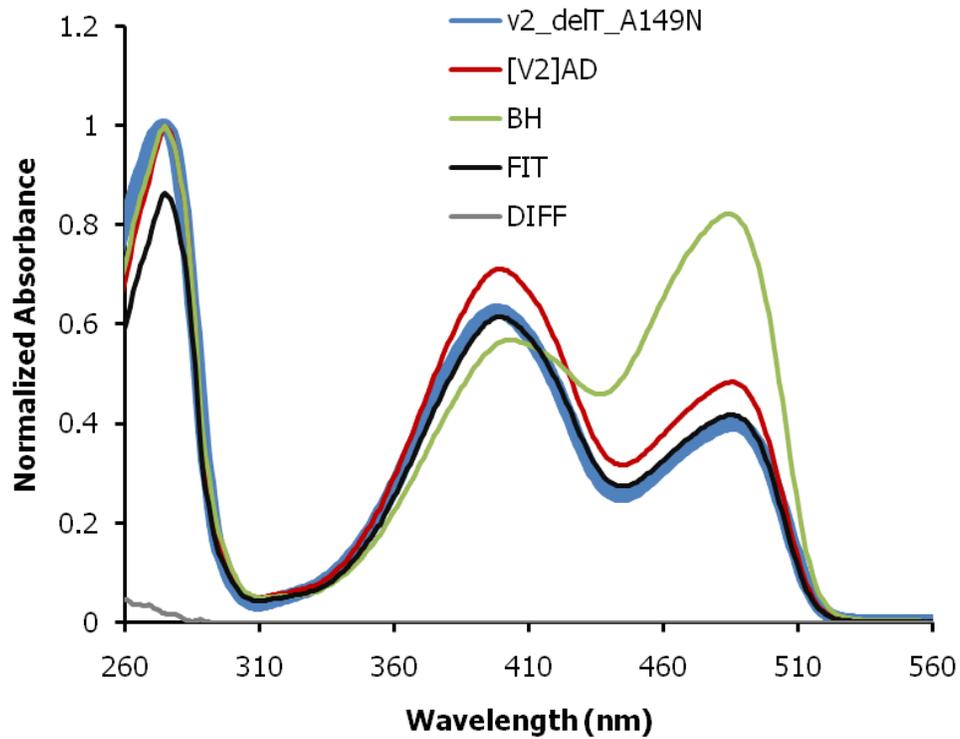








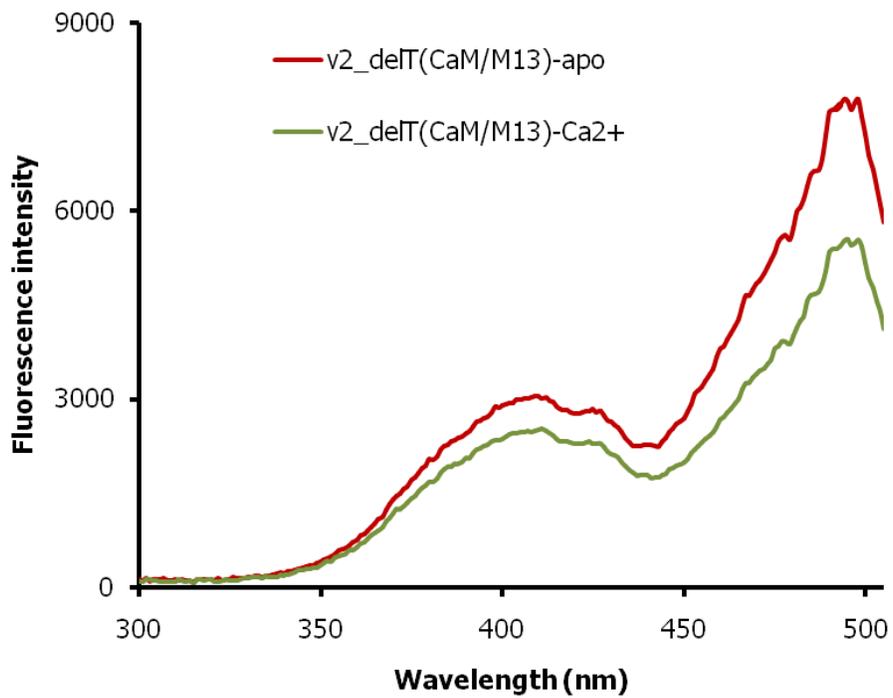
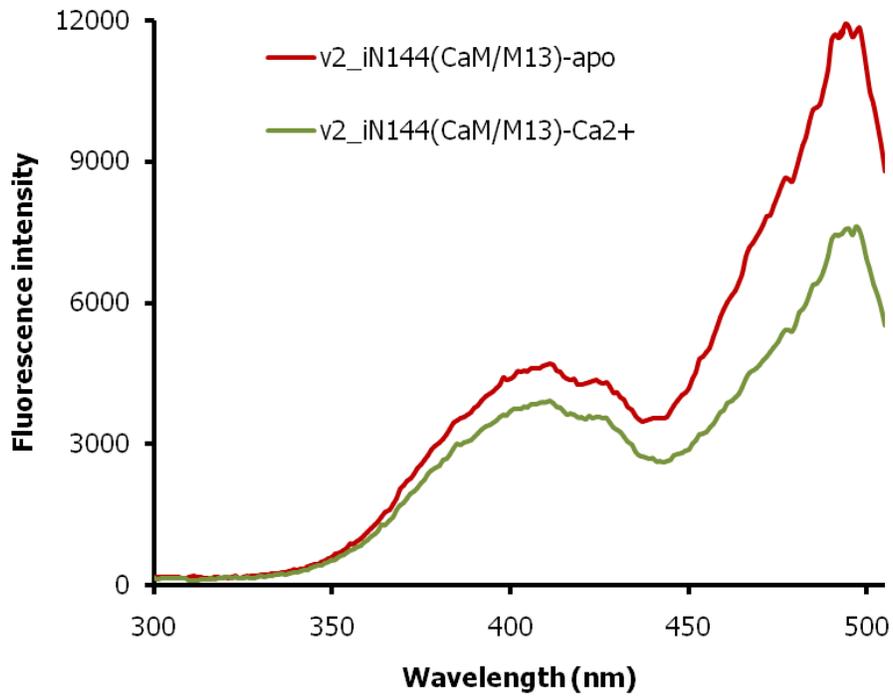


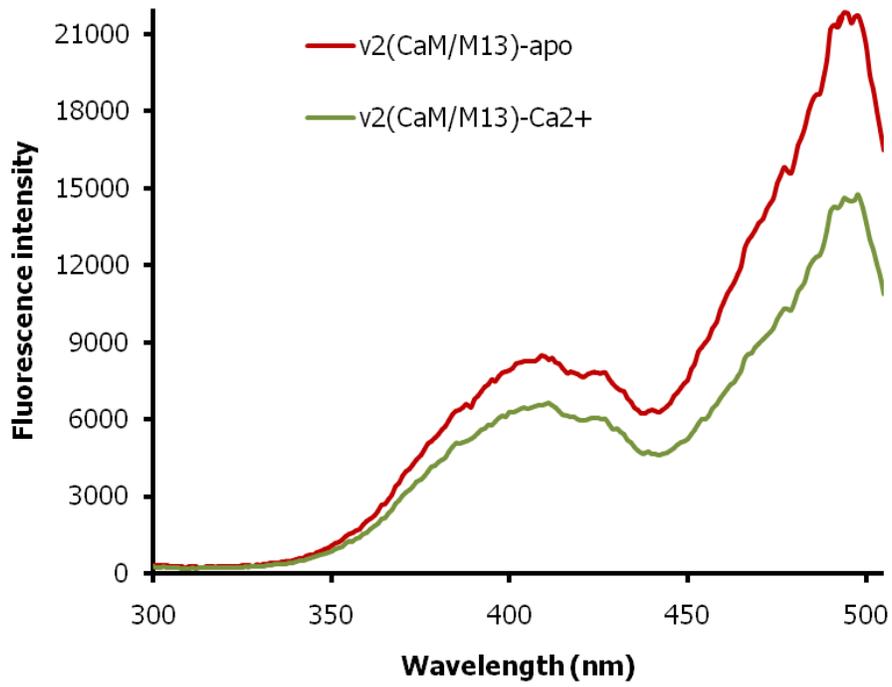
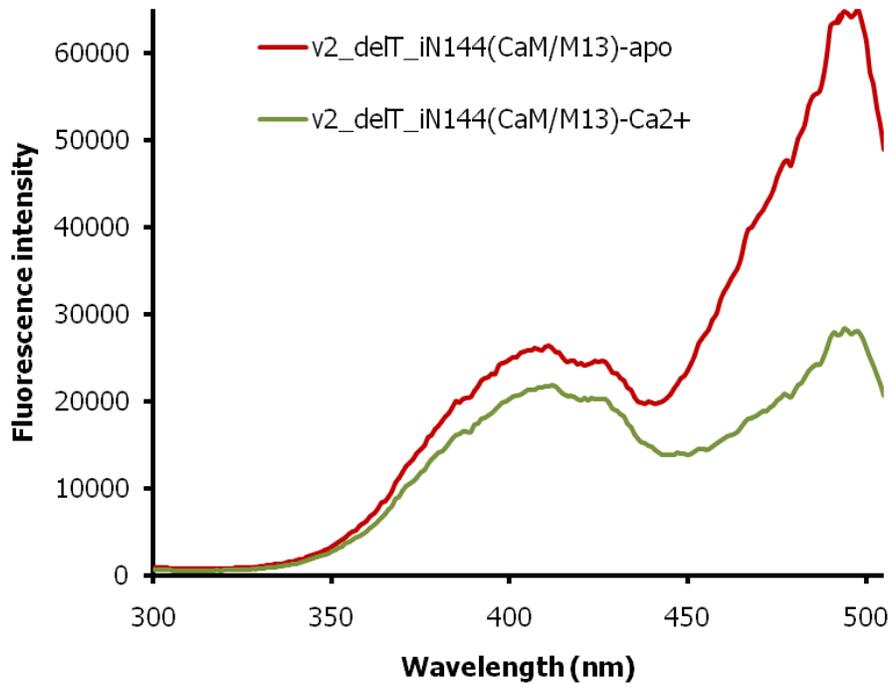


Appendix IV: MRE sequence

MRE	SEQUENCE	SOURCE
CaM/M13	KTSSDQLTEEQIAEFKEAFSLFDKDGDTITTKELGT VMRSLGQNPTEAELQDMINEVDADGNGTIYFPEFL TMMARKMKDSEEEIREFRVFDKDGNGYISAAE LRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYE EFVQMMTAKGGGGSKRRWKKNFIAVSAANRFKKIS SSGALGTS	CaM/M13 portion of YC- Nano15
CAP	VLGKPQTDPTLEWFLSHCHIHKYPKSTLIHQGEKA ETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGEL GLFEEGQERSAWVRAKTACEVAEISYKKFRQLIQVN PDILMRLSAQMARRLQVTSEKVGNLAFDVTGRIAQ TLLNLAKQPDAMTHPDGMQIKITRQEIGQIVGCSRE TVGRILKMLEDQNLISAHGKTIVVYGTR	E. coli CRP: 3HIF:A PDBID

Appendix V (A): Ca²⁺ biosensor analysis

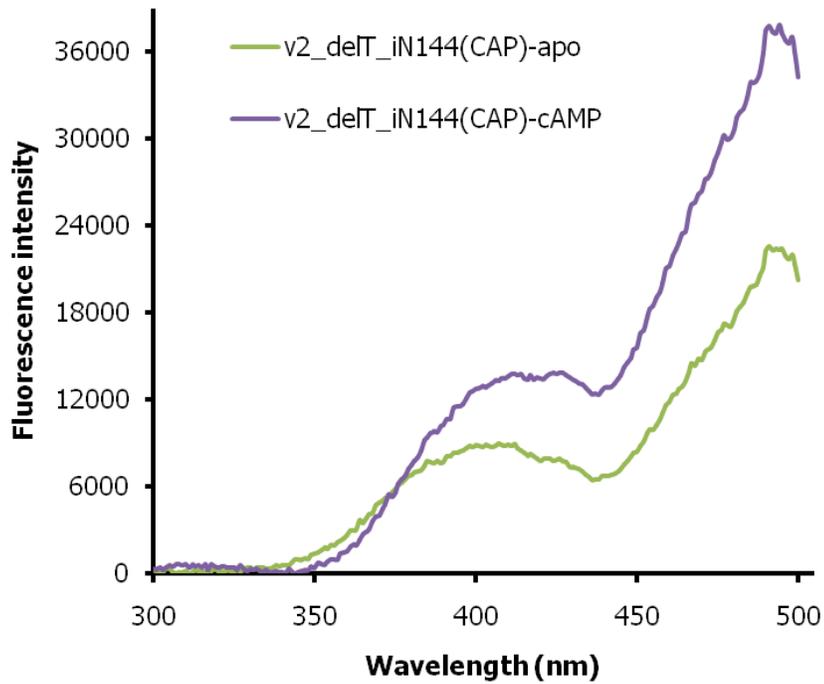
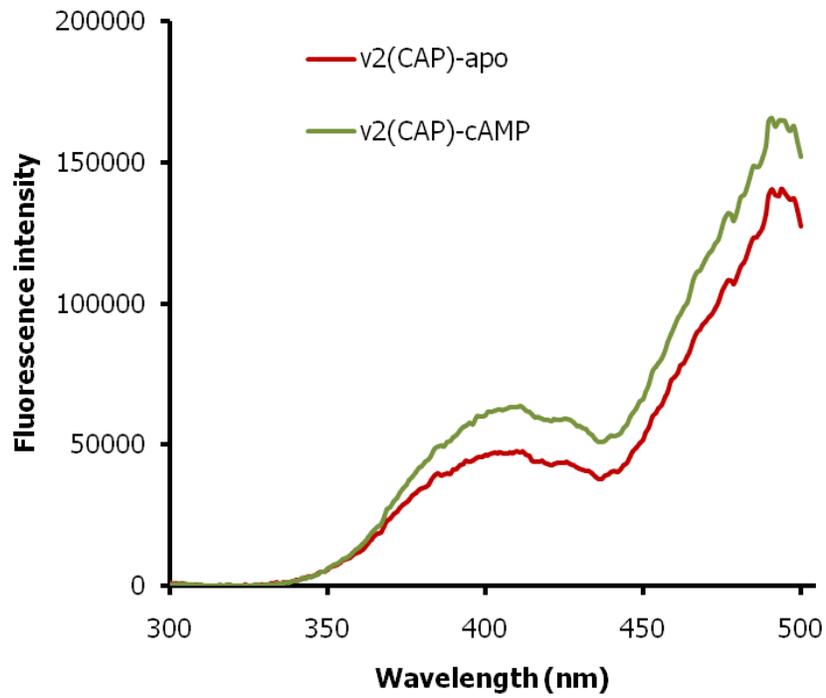




Characterization of Ca²⁺ biosensor constructs

S. No.	NAME	STATE	K _{eq}	ΔG	ΔΔG
1	v2(CaM/M13)	without Ca ²⁺	1.21±0.60	-0.47±0.99	1.09±1.92
		with Ca ²⁺	0.78±0.36	0.62±0.93	
2	v2_delT(CaM/M13)	without Ca ²⁺	1.09±0.52	-0.20±0.96	4.50±2.80
		with Ca ²⁺	0.17±0.19	4.30±1.84	
3	v2_delT_iN144(CaM/M13)	without Ca ²⁺	1.13±0.55	-0.31±0.97	1.00±1.90
		with Ca ²⁺	0.75±0.35	0.69±0.93	
4	v2_iN144(CaM/M13)	without Ca ²⁺	1.45±0.80	-0.91±1.08	2.32±2.06
		with Ca ²⁺	0.56±0.28	1.41±0.98	

Appendix V (B): cAMP biosensor analysis



Appendix VI: sequence of cp196mCherry

agctcgagcatgtccagctccggggcactggctcggttttatcaagttggacatcacctca
S S S M S S S G A L V G **F I K L D*** I T S
cacaacgaggactacaccatcggtggaacagtacgaacgcgcccagggtccgcccgtccacc
H N E D Y T I V E Q Y E R A E V R R S T
ggcggcatggacgagctgtacgagggggcactggcggttccatggtgagcaagggcgtg
G G M D E L Y E G G T G G S M V S K G V
gaggataacatggccttcatcaaggagttcatgcgcttcaagggtgcacatggagggctcc
E D N M A F I K E F M R F K V H M E G S
gtcaacggccacgagttcgagatcgagggcgagggcgagggccgcccctacgagggcacc
V N G H E F E I E G E G E G R P Y E G T
caggccgccaagctgaaggtgaccaaggggtggccccctgccccttcgctgggacatcctg
Q A A K L K V T K G G P L P F A W D I L
tcccctcagctcatgtacggctccaaggcctacgtgaagcaccgcccgcacatccccgac
S P Q L M Y G S K A Y V K H P A D I P D
tactggaagctgtccttccccgagggcttcaagtgggagcgcgtgatgaacttcgaggac
Y W K L S F P E G F K W E R V M N F E D
ggagggcgtggtgaccgtgaccaggactcctcctgcaggacggcgagttcatctacaag
G G V V T V T Q D S S L Q D G E F I Y K
gtgaagctgcgcggcaccaacttccccctccgacggccccgtaatgcagaagaagaccatg
V K L R G T N F P S D G P V M Q K K T M
ggctgggaggcctcctccgagcgggtgtaccccgaggacggcgctctgaagggcgagatc
G W E A S S E R L Y P E D G A L K G E I
aagcagaggctgaagctgaaggacggcggccactacgacgctgaggtcaagaccacctac
K Q R L K L K D G G H Y D A E V K T T Y
aaggccaagaagcctgtgcagctgcccggcgcttacaacgcccgggtgaccaactgaca
K A K K P V Q L P G A Y N A G G D Q L T
Gaagagtgagaattcgaagct (*** Consensus motif**)
E E - E F E A

Appendix VII: Screened colony sequences

LIBRARY # 1					
#	196	197	198	199	200
L1_1	F	I	V	C	C
L1_2	F	I	V	C	I
L1_3	F	I	A	C	T
L1_4	F	I	K	L	Q
L1_5	F	I	L	L	P
L1_6	F	I	K	L	D
L1_7	F	I	R	L	E
L1_8	F	I	Q	L	Y
L1_9	F	I	E	L	W
L1_10	F	I	S	L	R
L1_11	F	I	A	L	H
L1_12	F	I	S	L	L
L1_13	F	I	N	L	I
L1_14	F	I	A	L	D
L1_15	F	I	C	L	G
L1_16	F	I	V	L	D
L1_17	F	I	Q	L	A
L1_18	F	I	C	L	L
L1_19	F	I	M	L	L
L1_20	F	I	S	L	Y
L1_21	F	I	V	L	V
L1_22	F	I	L	L	K
L1_23	F	I	S	L	L
L1_24	F	I	W	L	D
L1_25	F	I	V	L	F
L1_26	F	I	H	L	W
L1_27	F	I	C	L	S
L1_28	F	I	L	L	C
L1_29	F	I	S	L	S
L1_30	F	I	F	L	V
L1_31	F	I	G	L	I
L1_32	F	I	L	L	I
L1_33	F	I	F	L	A
L1_34	F	I	Y	L	M
L1_35	F	I	L	L	Y
L1_36	F	I	E	L	G
L1_37	F	I	K	L	W
L1_38	F	I	S	L	I
L1_39	F	I	T	L	I
L1_40	F	I	H	L	R
L1_41	F	I	Q	M	L
L1_42	F	I	Y	M	V
L1_43	F	I	L	M	Y
L1_44	F	I	E	V	K
L1_45	F	I	S	V	M
L1_46	F	I	W	Y	F

LIBRARY # 2

#	196	197	198	199	200
L2_1	L	I	T	L	D
L2_2	F	I	L	L	D
L2_3	Y	I	V	L	D
L2_4	H	I	R	L	D
L2_5	H	I	E	L	D
L2_6	T	I	Q	L	D
L2_7	N	I	S	L	D
L2_8	N	I	E	L	D
L2_9	T	I	S	L	D
L2_10	W	I	S	L	D
L2_11	H	I	R	L	D
L2_12	Y	I	T	L	D
L2_13	Y	I	C	L	D
L2_14	F	I	C	L	D
L2_15	F	I	K	L	D
L2_16	F	I	A	L	D
L2_17	S	I	K	L	D
L2_18	F	I	S	L	D
L2_19	A	I	K	L	D
L2_20	L	I	V	L	D
L2_21	E	I	V	L	D
L2_22	H	I	Q	L	D
L2_23	D	I	S	L	D
L2_24	F	I	A	L	D
L2_25	M	I	S	L	D
L2_26	R	I	E	L	D
L2_27	L	I	I	L	D
L2_28	T	I	Q	L	D
L2_29	F	I	Q	L	D
L2_30	M	I	K	L	D
L2_31	T	I	K	L	D
L2_32	F	I	K	L	D
L2_33	L	I	R	L	D
L2_34	F	I	D	L	D
L2_35	L	I	V	L	D
L2_36	F	I	K	L	D
L2_37	Y	I	K	L	D
L2_38	M	I	K	L	D
L2_39	H	I	K	L	D
L2_40	F	I	K	L	D
L2_41	H	I	F	L	D
L2_42	L	I	R	L	D
L2_43	T	I	E	L	D
L2_44	T	I	K	L	D
L2_45	H	I	R	L	D
L2_46	L	T	L	L	D
L2_47	F	I	K	L	D
L2_48	G	V	V	L	D
L2_49	G	G	R	L	D
L2_50	Y	I	L	L	D
L2_51	L	I	W	L	D
L2_52	T	I	A	L	D

LIBRARY # 3

#	196	197	198	199	200
L3_1	F	I	K	L	D
L3_2	F	I	C	L	D
L3_3	F	I	K	L	D
L3_4	F	T	K	L	D
L3_5	F	G	L	V	D
L3_6	F	V	V	F	D
L3_7	F	I	K	L	D
L3_8	F	I	E	M	D
L3_9	F	T	C	L	D
L3_10	F	I	L	M	D
L3_11	F	V	T	L	D
L3_12	F	V	S	L	D
L3_13	F	H	P	G	D
L3_14	F	I	K	L	D
L3_15	F	I	N	L	D
L3_16	F	I	K	L	D
L3_17	F	H	K	L	D
L3_18	F	I	K	L	D
L3_19	F	I	S	L	D
L3_20	F	I	K	L	D
L3_21	F	R	L	D	D
L3_22	F	V	E	L	D
L3_23	F	I	L	L	D
L3_24	F	I	K	L	D
L3_25	F		G	G	D
L3_26	F	I	L	L	D
L3_27	F	I	K	L	D
L3_28	F	I	S	L	D
L3_29	F	I	H	L	D
L3_30	F	I	K	L	D
L3_31	F	T	H	L	D
L3_32	F	I	L	L	D
L3_33	F	I	K	L	D
L3_34	F	I	K	L	D
L3_35	F	V	C	L	D
L3_36	F	I	I	L	D
L3_37	F	I	V	L	D
L3_38	F	I	K	L	D
L3_39	F	I	R	L	D
L3_40	F	I	F	L	D
L3_41	F	V	R	L	D
L3_42	F	V	F	L	D
L3_43	F	I	Q	C	D
L3_44	F	F	D	L	D
L3_45	F	I	K	L	D
L3_46	F	I	G	L	D
L3_47	F	V	R	L	D
L3_48	F	I	F	L	D

Appendix VIII: R-GECO1 sequence

( M13;  mApple;  CAM)

gctttttatcgcaactctctactgtttctccatacccgtttttgggctaacaggannn
A F Y R N S L L F L H T R F L G - Q X X
ntaacatgggttttaaagaggagaaagggtcatgaacaataacgatctctttcaggcatca
X T M G L K R R K V M N N N D L F Q A S
cgtcggcggttttctggcacaactcggcagcttaaccgacgccgggacgctggggccgtca
R R R F L A Q L G S L T D A G T L G P S
ttgttaacgccgacgtgcgactgcgggcgcaagcggcgactgacgcttctagaggttct
L L T P R R A T A A Q A A T D A S R G S
catcatcatcatcatcatggatggctagcatgactggaggacagcaaagggtcgggat
H H H H H H G M A S M T G G Q Q M G R D
ctgtacgacgatgatgataaggatctcgccacaatggtcgactcttccgtcgtaagtgg
L Y D D D D K D L A T M V D S S R R K W
aataaggcaggtcacgcagtcagagctataggtcggctgagctcaccggtggtttccgag
N K A G H A V R A I G R L S S P V V S E
cggatgtacccccgaggacggcgccctgaagagcgagatcaagaaggggctgaggctgaag
R M Y P E D G A L K S E I K K G L R L K
gacggcgggccactacgccgaggtcaagaccacctacaaggccaagaagcccgtgacg
D G G H Y A A E V K T T Y K A K K P V Q
ctgccccggcgctacatcgctgacatcaagttggacatcggtgcccacaacgaggactac
L P G A Y I V D I K L D I V S H N E D Y
accatcggtggaacagtgccaacgcgccgagggcgccactccaccggcgcatggacgag
T I V E Q C E R A E G R H S T G G M D E
ctgtacaagggaggtacaggcgggagctctggtgagcaagggcgaggaggataacatggcc
L Y K G G T G G S L V S K G E E D N M A
atcatcaaggagttcatgcgcttcaaggtgcacatggagggctccgtgaacggccacgag
I I K E F M R F K V H M E G S V N G H E
ttcgagatcgagggcgagggcgagggcgccctacgaggcctttcagaccgctaagctg

F E I E G E G E G R P Y E A F Q T A K L
aaggtgaccaaggggtggccccctgcccttcgcctgggacatcctgtcccctcagttcatg
K V T K G G P L P F A W D I L S P Q F M
tacggctccaaggcctacattaagcaccagccgacatccccgactacttcaagctgtcc
Y G S K A Y I K H P A D I P D Y F K L S
ttccccgagggcttcaggtgggagcgcgtgatgaacttcgaggacggcggcattattcac
F P E G F R W E R V M N F E D G G I I H
gttaaccaggactcctcctgcaggacggcgtattcatctacaaggtgaagctgcgcggc
V N Q D S S L Q D G V F I Y K V K L R G
accaacttcccccccgacggccccgtaatgcagaagaagaccatgggctgggaggctacg
T N F P P D G P V M Q K K T M G W E A T
cgtgaccaactgactgaagagcagatcgcagaatttaaagaggctttctcctatttgac
R D Q L T E E Q I A E F K E A F S L F D
aaggacggggatgggacgataacaaccaaggagctggggacggtgatgcggtctctgggg
K D G D G T I T T K E L G T V M R S L G
cagaacccccacagaagcagagctgcaggacatgatcaatgaagtagatgccgacggtgac
Q N P T E A E L Q D M I N E V D A D G D
ggcacattcgacttccctgagttcctgacgatgatggcaagaaaaatgaatgacacagac
G T F D F P E F L T M M A R K M N D T D
agtgaagaggaaattagagaagcgttccgcgtgtttgataaggacggcaatggctacatc
S E E E I R E A F R V F D K D G N G Y I
ggcgcagcagagcttcgccacgtgatgacagaccttggagagaagttaacagatgaggag
G A A E L R H V M T D L G E K L T D E E
gttgatgaaatgatcagggttagcagacatcgatggggatgggtcaggtaaactacgaagag
V D E M I R V A D I D G D G Q V N Y E E
tttgtacaaatgatgacagcgaagtagaa
F V Q M M T A K -

Appendix IX: In-house Plasmid miniprep procedure

1. Harvest 1.5-3.0 ml bacterial cells by centrifuging 5 min, 7,500-10,000 rpm.
2. Discard supernatant (aspirate or pour off) and resuspend pellet in 150 uL of Solution I by vortexing.
3. Add 150uL Solution II and invert 5 times to mix.
4. Add 150uL Solution III and invert 5 times to mix.
5. Add 150uL chloroform and mix.
6. Spin 15,000rpm for 5 min, 4°C
7. Transfer 400uL of top layer to new tube.
8. Add 800 uL 100% EtOH and mix by inverting.
9. Spin 15,000rpm for 5 min, 4°C and discard supernatant.
10. Add 500uL 70% EtOH and quick spin.
11. Air dry for 15 minutes by leaving tubes on bench.
12. Redissolve pellets in 50uL sterile water.

Solution 1:

25 ml 1 M Tris-Cl (pH 8.0) (=50 mM)
10 ml 0.5 M EDTA (=10 mM)
50 mg RNase A (=100 ug/ml)
distilled water to 500 ml

Solution II

4 g NaOH (= 0.2 N)
5 g SDS (= 1%)
distilled water to 500 ml

Solution III

120 ml Acetic acid (= 2 M)
294.5 g Potassium acetate (= 3 M)
distilled water to 1 litre