Tuning the Properties of Protein-Based Biosensors by

Chromophore Replacement Strategies

by

Nazanin Assempour

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

Master of Science

Department of Chemistry
University of Alberta

Abstract

Optical imaging of membrane potentials has become an immensely popular approach to study neuronal activity. Over the past decades, a number of voltage indicators have been introduced including the voltage sensitive fluorescent protein (VSFP) and microbial rhodopsins. However, most of these sensors suffer from weak signal-to-noise ratios, low sensitivity and incompatibility for deep tissue imaging.

The second Chapter of this thesis describes our current efforts on improving microbial rhodopsinbased voltage indicators by replacing the retinal cofactor with a red shifted analogue. We synthesized the red-shifted retinal analogue **(S4)** with higher quantum yield and extinction coefficient compared to the natural retinal. Next, we reconstituted the ARCH and PROPS variants using the synthetic retinal and obtained the proteins with remarkably red-shifted absorbance and fluorescence spectra. This result demonstrates that chromophore replacement is promising strategy for modulating the spectral and voltage properties for more sensitive action potential detection.

In the third Chapter, we describe our attempts to extend the fluorescence modulation strategy to "empty" fluorescent protein barrels, by replacing the endogenous chromophore with a synthetic version **(HBDI)**. We expect that a chromophore replacement strategy could help provide new imaging tools with improved fluorescent properties. We performed some modifications on FP barrels with the aim of allowing the synthetic chromophore to insert into the empty barrel. Unfortunately, we could not obtain any improvements on fluorescent properties of FPs by this strategy.

Preface

This thesis is an original work by Nazanin Assempour. None of the text of this thesis has been previously published.

The mutated ARCH and PROPS genes used in Chapter 2 were developed and provided by my colleague Dr. Yongxin Zhao. The photocleavable protein (PhoCle) gene that is employed in Chapter 3 was engineered and provided by Wei Zhang. The empty cp-EGFP was engineered by Ahmed Abdelfattah and used in Chapter 3.

Acknowledgement

Last but not least, I would like to thank my parents and my lovely sister Yasamin for their support and encouragement. My especial thanks to my husband Reza Rezaei, he was always there cheering me up and stood by me through the good times and bad.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ε	Extinction Coefficient
φ	Quantum Yield
A	Alanine
Abs	Absorbance
ARCH	Archaerhodopsin
avGFP	*Aequorea Victoria* Green Fluorescent Protein
BCA	Bicinchoninic Acid
BFP	Blue fluorescent protein
B-PER	Bacterial Protein Extraction Reagents
BR	Bacteriorhodopsins
BSA	Bovine SerumAlbumin
CFP	Cyan Fluorescent Protein
ChR	Channelrhodopsin
cp-EGFP	Circularly Permuted- Enhanced Green Fluorescent Protein
cp-GFP	Circularly Permuted Green Fluorescent Protein
ddFP	Dimerization-Dependent Fluorescent Protein
D	Aspartic Acid
DIBALH	Di-Isobutyl Aluminum Hydride
DMPU	1,3-Dimethyl-3,4,5,6-tetrahydro-2-pyrimidinone
DMSO	Dimethyl Sulfoxide

DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Triphosphate
DPA	Dipicrylamine
EGFP	Enhanced Green Fluorescent Protein
E	Glutamic Acid
Em	Emission
Ex	Excitation
F	Phenylanine
FLaSH	Fluorecent Shaker
FP	Fluorescent Protein
FRET	Förster Resonance Energy Transfer
G	Glycine
GEVI	Genetically Encoded Voltage Indicators
GFP	Green Fluorescent Protein
GPCR	G Protein-Coupled Receptor
HBDI	4-Hydroxybenzylidene-1,2-dimethyl-imidazolinone
H	Histidine
HR	Halorhodopsins
hVoS	hybrid Voltage Sensor
HWE	Honor-Emmons-Wadsworth
hν	Light (Energy)
I	Isoleucine

J	Coupling Constant
K_d	Dissociation Constant
L	Leucine
M	Methionine
MBP	Maltose Binding Protein
µL	Microlitre
mL	Millilitre

Chapter 1: Introduction

1.1 Fluorescence

Fluorescence is the emission of light as the result of a three-stage process by fluorescent molecules (fluorophore). The Jablonski diagram [1] explains these stages by representing the electronic transitions in the fluorophore **(Figure 1.1)**.
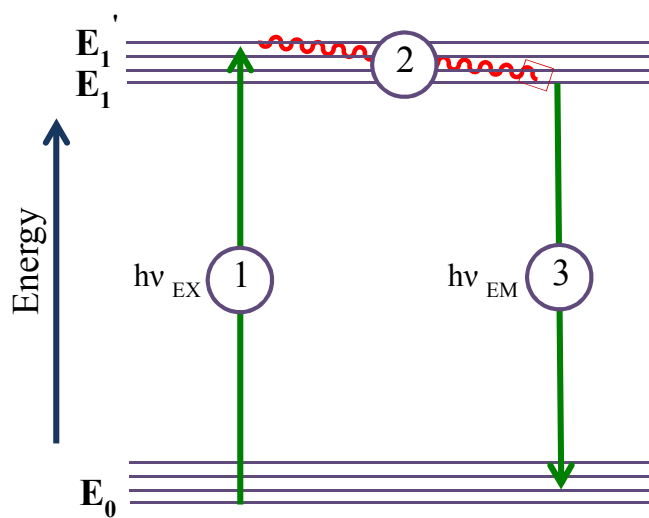
Figure 1.1 Simplified Jablonski energy diagram of a fluorescence event. 1) fluorescence excitation, 2) internal conversion, 3) fluorescence emission.

In the first stage, called *fluorescence excitation,* the fluorophore in the ground state (E_0) is excited to a high vibrational energy level in the first electronic singlet state (E_1) by absorbing a photon of energy ($h\nu_{EX}$) provided by an external light source (e.g., a laser). The molecule will stay in the excited state for a while (1-10 nanoseconds) and during this time (which is called fluorescence lifetime), the fluorophore interacts with the molecular

environment and loses some of its energy relaxing to the lowest vibrational energy level of S_1. This process is called *internal conversion*. Ultimately, the fluorophore returns to the ground state (E_0) by emitting a photon of energy ($h\nu_{EM}$) in the process of *fluorescence emission.*

The energy of the photon emitted in the third stage ($h\nu_{EM}$) is lower than the energy of the absorbed photon ($h\nu_{EX}$) and therefore, the emission peak shows up at longer wavelength compared to the excitation peak **(Figure 1.2)**. The wavelength or energy difference between these two peaks is called the *Stokes shift* and it is mainly due to the excited state internal conversion phenomenon: before releasing the absorbed energy as luminescence, the fluorophore loses a small amount of energy as thermal energy.
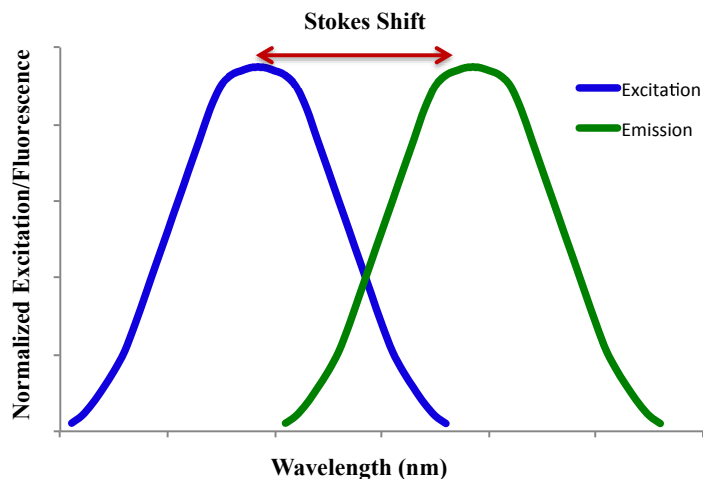
Figure 1.2 Excitation, emission and Stokes Shift.

There are three main factors that define the fluorescence characteristics of the molecules: Extinction coefficient (ε), Quantum yield (ϕ), and fluorescence lifetime (τ).

According to the Beer-Lambert Law the absorbance of a solution (A) at a specific wavelength is proportional to the concentration of the absorbing species (C), path length (L), and extinction coefficient giving the overall relationship $A = \varepsilon C L$. Since the absorbance is unitless, extinction coefficient has the unit of $M^{-1}cm^{-1}$. In terms of fluorescence, extinction coefficient is related to the ability of the fluorophore to absorb the light. The brightness of fluorescence is proportional to both this parameter and quantum yield.

Quantum yield is the emission efficiency of the fluorophore and can be described as the ratio of the number of photons emitted to the number of photons absorbed by the molecule. The quantum yield ranges between the value of zero to one and represents the probability of the fluorophore to emit a photon (fluorescence).

The fluorescence lifetime refers to the time that the fluorophore spends in the excited state before relaxing back to the ground state. For many fluorescent molecules, including *Aequorea victoria* GFP (avGFP), the lifetime is in the nanosecond range.

There are a large number of fluorescence-based techniques and applications that are commonly used in the life sciences. Examples of current fluorescence applications include detection of single nucleotide polymorphisms (SNP) [2], DNA sequencing, [3] and investigation of protein-protein and other macromolecular interactions [4]. Of most relevance to this thesis, fluorescence enables researchers to study a wide range of phenomena in cell biology and address fundamental questions related to cell physiology, when used with optical microscopes. The application of fluorescence microscopy to the study of gene expression and subcellular location of proteins in live cells was revolutionized by the advent of green fluorescent protein (GFP) [5].

1.2 *Aequorea Victoria* Green Fluorescent Protein

The avGFP was first discovered and isolated from the bioluminescent jellyfish *Aequorea*, by Shimomura *et al.* in 1962 [6]. The avGFP consists of 238 amino acids with a molecular mass of 26.9 kD, as first determined in 1978[7]. The sequence of avGFP remained unknown until 1992 when Prasher reported the cloning and expression of wild-type avGFP in bacteria [8]. The crystal structure of wild-type avGFP was determined at 1.9 Å resolution in 1996, and revealed that the protein consisted of eleven β-strands arranged as a β-barrel with an α-helix in the middle [9, 10] **(Figure 1.3)**.
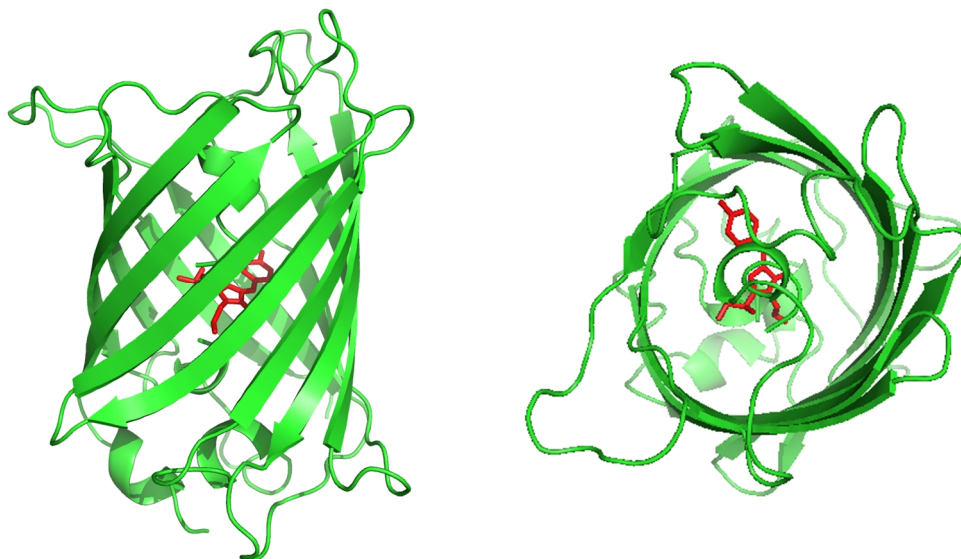
Figure 1.3 Three-dimensional structure of avGFP (Protein data base (PDB) ID: 1EMA) [9, 11].

The visible absorbance and fluorescence of avGFP arise from a post-translationally created *p*-hydroxybenzylidene-imidazolone chromophore in the middle of the central α-helix. The chromophore is formed via the cyclization of S65, Y66, and G67 residues followed by 1,2-dehydrogenation and oxidation of the tyrosine [12] **(Figure 1.4)**.

GFP absorbs light with peak maxima at 395 nm and 475 nm and emits green light at 508 nm with a quantum yield in the range of 0.72-0.85 [13]. The presence of two absorption peaks suggested that there are two different forms of the chromophore. In 1994 Heim *et al.* proposed that protonated and deprotonated forms of chromophore are responsible for the 395 nm and 475 nm peaks, respectively [13, 14]**(Figure 1.5)**.
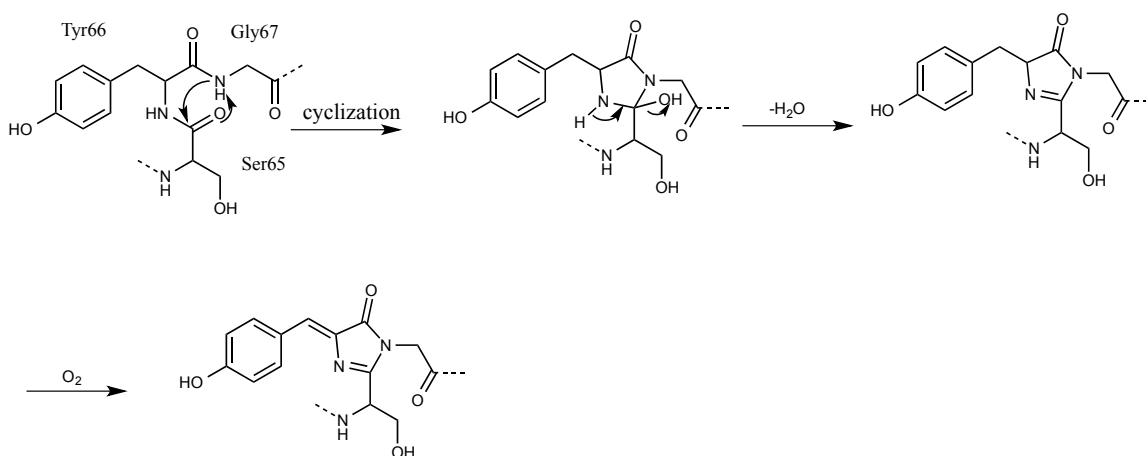
Figure 1.4 Mechanism of formation of avGFP chromophore proposed by Tsien [12].
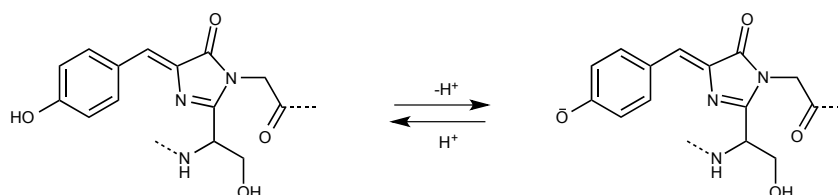
Figure 1.5 The equilibrium between the two ionization states of the avGFP chromophore.

The avGFP chromophore is surrounded by a protein environment that strongly influences the equilibrium between its two ionization states. Specifically, the

chromophore is part of a hydrogen bond network that can transfer a proton between the chromophore and the vicinal side chains of the protein including E222, H148, and T203. This hydrogen bond network is responsible for stabilizing the deprotonated state of the chromophore, which absorbs at 475 nm. On the other hand, the neutral form of the chromophore, is stabilized via electrostatic interactions and hydrogen bonding of the carboxylate of E222 with S65 and also with a bound water molecule that is hydrogen bonded with the hydroxyl group of Y66 and also oxygen atoms of S205 and T203. The presence of E222 is essential for the hydrogen-bonding network as the E222G mutation, which removes the hydrogen bond between E222 and S65 leads to suppression of the peak at 395 nm [13].

During the mid to late 1990s, GFP was rapidly adopted throughout biological science as an imaging tool. It has since been employed as a popular research tool in cellular and molecular biology because of its stability, relatively small size, and ease of use. It is mostly used as a fluorescent marker for protein localization in cells and organisms [15]. The ability of avGFP and its mutants to tolerate genetic fusion to other proteins of interest, and thereby label them such that they can be monitored using fluorescence methods, has made them versatile indicators of structure and function in cells.
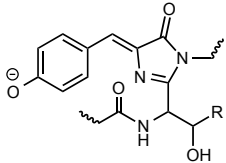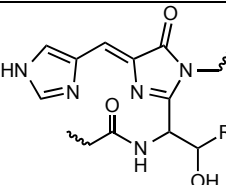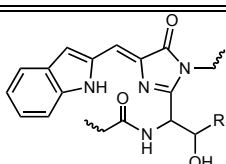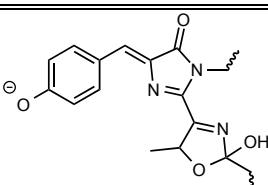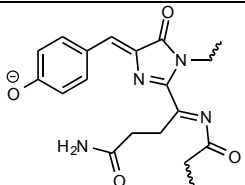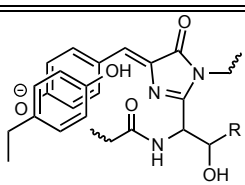
1.3 The Fluorescent Protein (FP) Color Palette

Although avGFP was soon recognized as a valuable tool to scientists, it had several limitations such as poor photostability, low brightness, and slow maturation. To address these problems, protein engineers turned to various mutagenesis strategies in order to

develop new variants with different properties and colors that range from blue to yellow [16]. The range of modifications can be performed include 1) replacement of amino acids in the vicinity of the chromophore leading to changes in the chromophore environment and properties; 2) modification of the chromophore structure by using different aromatic amino acids at position 66; 3) improvement protein folding; 4) improvement chromophore maturation; 5) alteration the oligomeric state. All of these changes and modifications happen due to amino acid substitutions resulting from intentional mutation of the gene that encodes the FP. Moreover, engineering monomeric FPs from marine coral *Discosoma striata* led to orange to red FPs, which represented a major advancement in FP development [11]. **Table 1.1** shows the chromophore structures of FP variants.

In 1995, Roger Tsien's laboratory studied the effect of mutation of the first amino acid in the avGFP chromophore, S65. This work led to the avGFP-S65T mutant that was a distinct improvement compared to the wild-type avGFP [17]. This mutation stabilized the hydrogen-bonding network in the ionized form of the chromophore resulting in a single absorption peak at 489 nm. This improved variant of avGFP was called Enhanced GFP (EGFP) and is among the brightest and most photostable of the *Aequorea*-based FPs. Blue FP (BFP) obtained from the conversion of the tyrosine residue at position 66 to histidine resulted in an absorption peak at 380 nm and emission at 448 nm [18]. This protein has a lower quantum yield and brightness and rapid photobleaching compared to avGFP and required more mutations for enhancement [19]. Cyan FP (CFP) is the result of substitution of tyrosine at the position of 66 with tryptophan that yielded a chromophore with the absorption maximum at 436 nm and emission at 485 nm [16].

7

Table 1.1 Chromophore structures of FP variants.

Name (Colour)	Chromophore structure	Key Mutation or Change Emission Wavelength
EGFP (Green)		S65T 500-520 nm
BFP (Blue)		Y66H 440-470 nm
CFP (Cyan)		Y66W 470-500 nm
OFP (Orange)		550-575 nm
RFP (Red)		Extra Double (C=N) bond 575-610 nm
YFP (Yellow)		Ser65Gly T203Y 520-550 nm

One of the red-shifted avGFP variants is yellow FP (YFP), which was constructed by introducing a tyrosine mutation on the β-barrel (position 203). The π-π interaction between the aromatic rings of chromophore and Y203 shifts the spectra to the longer excitation and emission wavelength [20]. However, YFP shows some deficiencies including poor photostability and expression at 37 °C and high pH sensitivity that encouraged more protein engineering efforts to improve it. Venus, mCitrine, and YPet are among the YFP variants with improved characteristics [21].

In 1999, Matz *et al*. found a red FP (called DsRed) in corals, opening the door to the discovery of many different FPs and chromophore structures [22]. DsRed protein has impressive brightness, stability against pH changes, and photobleaching. The chromophore structure of this protein has an extra double bond compared to the EGFP chromophore that leads to extension of conjugation system and creates a red-shifted chromophore [23] **(Table 1.1).**

The discovery and engineering of this broad range of FP variants has resulted in fluorescence emission spectral profiles that cover the entire visible light spectrum. Further efforts to improve FPs will facilitate applications such as multicolor imaging of differential gene expression, protein localization or cell fate.

1.4 Circularly Permuted Green Fluorescent Protein

The biological activity of macromolecules like DNA, RNA, and proteins depends on their tertiary structure (three dimensional arrangement of building blocks). Any change or rearrangement in this structure may lead to change in function, folding, and stability of these molecules.

Circular permutation of macromolecules is an engineering method to make an isomer of the molecule. To obtain the circularly permuted molecule the normal termini are linked genetically via a linker peptide and new termini are produced somewhere else along the backbone. In principle, this cleavage can occur at any site of the backbone. Therefore, each molecule has many structural possibilities [24]. Many studies on this type of engineering on proteins has revealed that introduction of new termini into secondary structures may be tolerated and often allows folding to a biologically active conformation. Among different circularly permuted proteins, avGFP and several other FP variants have been circularly permuted and studied. The N- and C-termini of avGFP are located in close proximity to each other at one end of the barrel **(Figure 1.6)**. For avGFP, introduction of new termini by circular permutation is well tolerated and the protein folds and forms its chromophore [25].
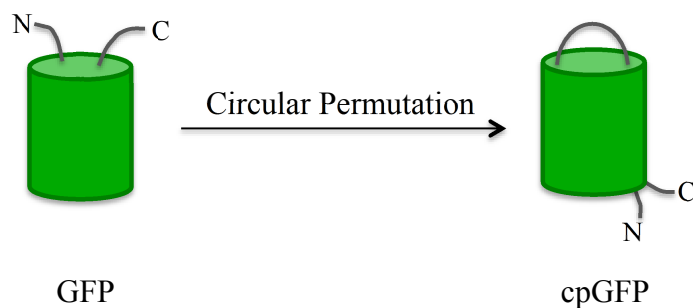
Figure 1.6 The representative structure of avGFP and circularly permuted avGFP (cp-GFP) [26].

Manipulation of the protein sequence by this method provides alternative N- and C-termini. Variants with altered position of termini can sometimes have advantages when it comes to creating fusion proteins, improving catalytic activities, and altering the ligand or

substrate binding affinity or enabling the design of new biocatalysts and biosensors [27, 28].

1.5 Synthesis of FP Chromophores

4-(*p*-hydroxybenzylidene)-5-imidazolone is the most representative structure of avGFP chromophore. This chromophore is formed via a self-sufficient post-translational modification that involves dehydration of S-Y-G to give an imidazolinone moiety, following by oxidation [12]. Other avGFP chromophore analogues are produced via related routes from similar tripeptides precursors. For example, the chromophore of the blue FP (BFP) is formed from S-H-G. The conformation of the GFP chromophore inside the barrel is restricted to the *Z*-form that enables it to exhibit fluorescence. However, a synthetic chromophore that is identical to a GFP chromophore in solution would isomerize leading to fluorescence quenching [29] **(Figure 1.7)**. Quenching could occur by other factors as well such as rotation around single bonds and solvent interactions (solvatochromism).
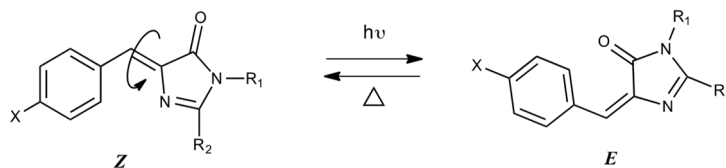
Figure 1.7 Isomerization of GFP-type chromophores.

Synthetic chromophore analogues make it possible for researchers to study the spectral characteristics of the chromophore, and gain insight into the mechanism of

formation, and the influence of the protein environment. By eliminating the protein barrel, the determination of photophysical characteristics of chromophores such as extinction coefficient has been simplified. Synthetic chromophores can also be used as biological probes by covalently attaching them to macromolecules like DNA and proteins [30, 31].

The most common method to synthesize the 4-arylidene-5-imidazolinones is based on the Erlenmeyer azlactone synthesis **(Figure 1.8)** [32]. This reaction occurs between *N*-acylglycine (**I**) and an arylaldehyde in the presence of sodium acetate and acetic anhydride, to produce a 4-arylidene-5-oxazolinone (**II**) [33]. This product can undergo a condensation with primary amines in the presence of a base to produce 4-arylidene-5-imidazolinone (**III**) [31, 34].
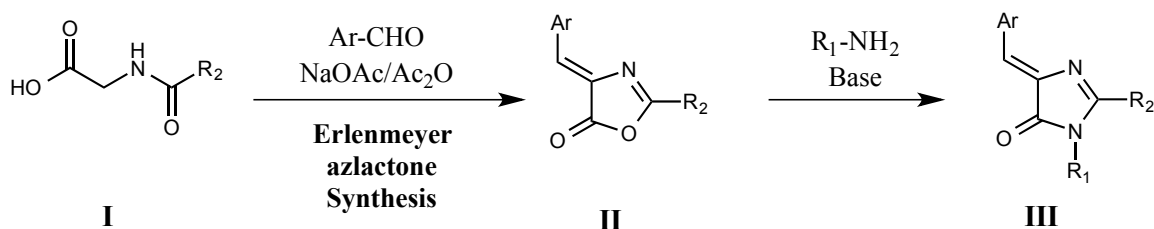
Figure 1.8 Synthesis of 4-arylidene-5-imidazolinones via Erlenmeyer azlactone reaction.

In addition to the Erlenmeyer azlactone procedure, there are other synthetic pathways that have been proposed for synthesizing more complex model chromophores [35-37]. GFP chromophore analogues are widely used as excellent fluorogenic probes for labeling proteins, RNA, and cellular metabolites, and thus development of new synthetic routes could facilitate these labeling strategies.

1.6 Biosensors

Biosensors are useful means for visualizing and monitoring complex biological activities occurring in different organisms and living cells. They contain a biochemical component (e.g., a protein) that is responsible for sensing a specific analyte or target. A physiochemical transducer is also necessary to convert the presence of a target to a measurable signal. The obtained signal is proportional to the concentration or activity of the target **(Figure 1.9)** [38].

Recognition of analyte   Transduction and conversion into signal   Signal
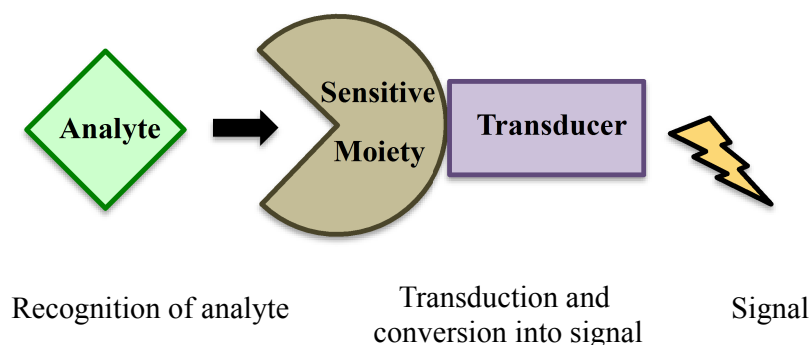
Figure 1.9 Schematic of a generic biosensor.

Over the years, biosensor technology has been developed and utilized in a variety of applications and analytical problems in medicine, the food industry, and environmental detect among them. The discovery of avGFP and its different variants prompted researchers to make FP-based biosensors. These efforts have led to the introduction of fully protein-based (and therefore fully genetically encoded) biosensors as useful tools for probing the cellular activities, conformational changes, and detection of biomolecules both *in vivo* and *in vitro* [39-41].

1.6.1 Fluorescent Protein (FP) Biosensors

Over the last two decades, FP-based biosensors have been utilized as genetically encoded biosensors to probe numerous biochemical and physiological activities in cells such as membrane potential and protein-protein interactions. They are employed for protein localization by fusing one or two FPs to the protein of interest and tracing its localization by fluorescence microscopy. A biochemical signal in the cell leads to a conformational change, which results in a change in fluorescence intensity or hue. FP-based biosensors have many advantages over the other biosensing platforms. For example they are constructed by using simple molecular biology technologies; the protein makes the fluorophore autonomously and so they can be introduced into living cells relatively non-invasively in the form of the DNA gene. These types of biosensors are classified in the following categories: 1) single FP-based biosensors [42]; 2) Förster resonance energy transfer (FRET)-based biosensors [43, 44]; and 3) split FP (complementation)-based biosensors. In this thesis, split FP-based biosensors are of the most relevance, so they are discussed below in additional detail.

1.6.1.1 Split Fluorescent Protein (complementation) Biosensors

The development of methods and tools that identifies the interactions between proteins and peptides is critical to interrogate and understand complex cellular processes, modifications, and their locations. In split protein reassembly assays (also called protein fragment complementation) [45], a reporter protein is split into two fragments and fused to possible interacting protein or peptide partners. If these partners have affinity for each other, they bring the two fragments into close proximity so that they associate, and the

reporter protein regains function. β-lactamase [46], dihyhdrofolate reductase [47], β-galactosidase [48], and GFP are the most commonly used split reporter proteins.

The ability to reconstitute GFP from two fragments was discovered first by Regan and coworkers in 2000 [49]. In a split GFP biosensor, two fusion proteins are modified such that each one is fused to a fragment of the GFP protein. The two separate split-GFP fragments are not fluorescent. If the two proteins interact together, they bring the split-GFP fragments close together, which reconstitute the fully fluorescent GFP **(Figure 1.10)**.

Split GFP fragments   Reconstituted GFP
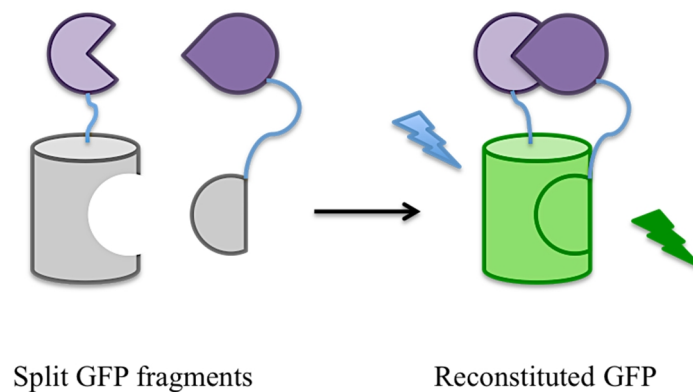
Figure 1.10 Split GFP reassembly for protein-protein interaction.

In addition to split GFP, other FPs have been engineered into split-FP based biosensors. These include YFP [50] [51], mRFP1 [52], and mCherry [53], which can also be used for detecting protein-protein interactions. The split FP-based system is widely utilized for detecting protein-protein interactions since it expresses, folds, and fluoresces in a large number of cell types. However, its irreversible complementation nature and instability of the FP fragments limits the application of this technology [54].

1.6.1.2 Dimerization-dependent FP (ddFP) Biosensors

As an alternative method to the split FPs and complementation-based biosensors, dimerization-dependent fluorescence protein (ddFP) biosensors were designed and reported by Alford and coworkers in 2012 [55, 56].
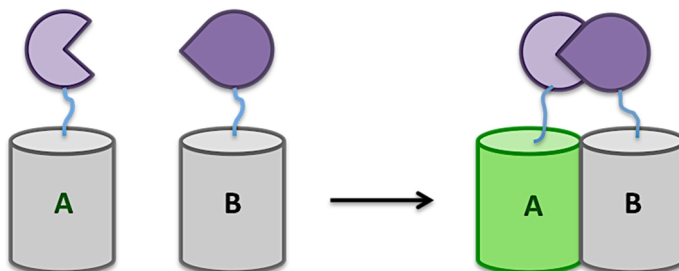
Figure 1.11 The ddFP-based complementation strategy.

In this strategy, two dark FP monomers are involved; one of them contains the quenched (protonated) chromophore in the monomeric form (**Figure 1.11**, copy A) and the second one has no chromophore (copy B). However, by associating these two monomers via fused partners, a heterodimer (AB) is formed non-covalently, which is now fluorescent because of a change in the chromophore environment in copy A, leading to deprotonation of the chromophore.

Further investigations on ddFP suggest that this type of indicator could be used for studying protein-protein interactions *in vitro* as well as Ca^{2+} dependent protein-protein interactions in live cells.

1.7 Rhodopsin

1.7.1 G protein-coupled receptor (GPCR)

G protein-coupled receptors constitute a large family of membrane proteins in the human genome and have three main domains: an extracellular protein domain, an intracellular protein domain, and a membrane-spanning segment connecting the two. These are also called seven-transmembrane receptors. The primary function of these proteins is transduction of extracellular signals to intracellular proteins (called G proteins). These signals are most typically soluble ligands that bind to the extracellular domain, though the rhodopsin family of GPCRs is sensitive to light. This binding (or absorbance of a photon) causes a conformational change in the receptor and dissociation of G proteins [57][58]. There are numerous types of GPCRs that can be classified in five families: Rhodopsin family (sensitive to light), the adhesion family, the taste family, the glutamate family, and the secretin (sensitive to hormones) family [59].

1.7.2 Animal Rhodopsin

Rhodopsin is a visual photoreceptor and is a member of the family of GPCRs that is contained in the light sensitive cells in the retina of the eye [60]. This photosensitive protein captures photons and converts them to an electrical signal that leads to vision. Rhodopsin is composed of the opsin protein that is covalently linked to the 11-*cis*-retinal chromophore through a Schiff base linkage to the amino group of the K296 at the side chain. Retinal is made from Vitamin A in the retina **(Figure 1.12)**.
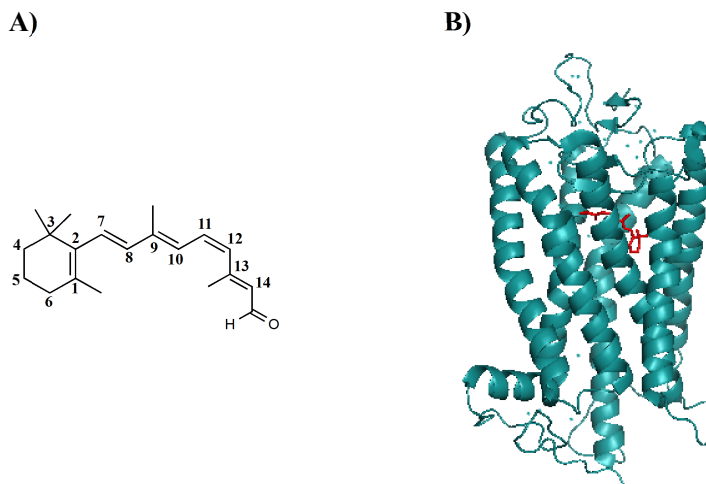
Figure 1.12 A) 11-*cis* retinal. B) Bovine rhodopsin with 11-*cis* attached retinal (PDB ID: 1U19) [61].

By absorption of light, the retinal chromophore undergoes isomerization to the all-*trans* form and causes a conformational change in rhodopsin that leads to activation of a G protein cascade inside the cell **(Figure 1.13)**.

11-cis-retinilidene   **All trans retinilidene**
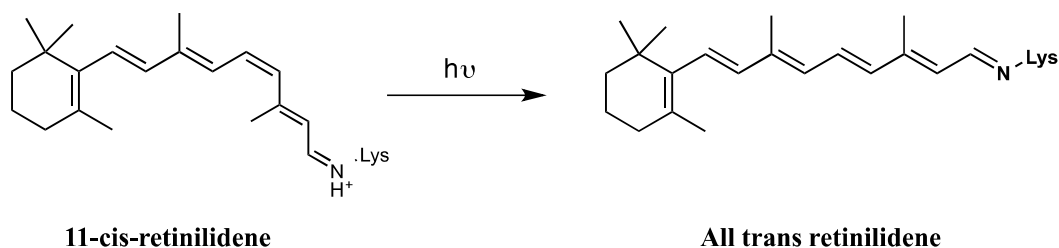
Figure 1.13 Animal retinal photo-isomerization.

The x-ray structure of rhodopsin combined with computational and experimental approaches was essential for understanding of GPCRs' structure and activation [58]. Rhodopsin is more suitable candidate for structural investigations compared to other

GPCRs because of its high stability under various conditions and its availability due to the large quantities present in the retina [62, 63].

1.7.3 Microbial Rhodopsin

Microbial rhodopsins were first found in archaea (a kingdom prokaryote microbe) in 1971 and initially they were called archaeal rhodopsins [64]. There is no similarity between the sequence of these types of rhodopsins and animal rhodopsins. However, both of them are seven transmembrane α-helical opsins with their C-terminus and N-terminus facing inside and outside of the cell, respectively **(Figure 1.14)**. The protein structures of microbial rhodopsins are designed to transfer the ions such that absorption of a photon leads to ion transportation across the membrane. The functions of both types (animal and microbial) are initiated by light absorption.

A)   B)
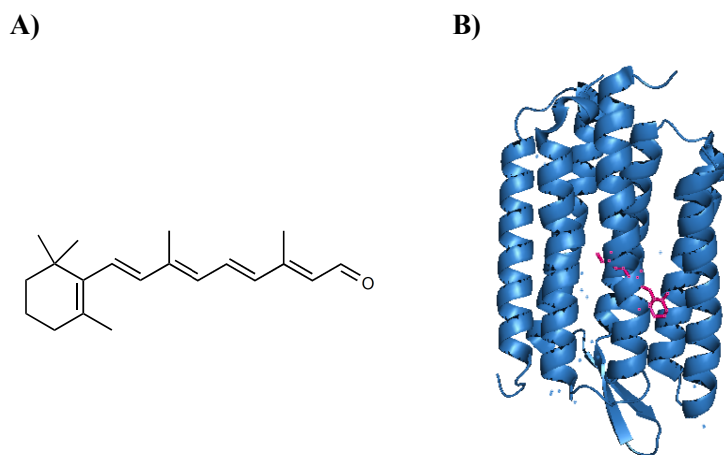
Figure 1.14 A) all-*trans* retinal. B) Bacteriorhodopsin with all-*trans* attached retinal (PDB ID: 1C3W) [68].

Microbial rhodopsins can be divided into several types including: 1) bacteriorhodopsin (BR); 2) halorhodopsin (HR); 3) channelrhodopsin (ChR); and 4) sensory rhodopsin (SR). BRs pump protons out of the cell whereas HRs pump chloride ions into the cell [65]. ChR was found in green algae and is responsible for depolarizing of the plasma membrane by acting as a light-gated cation channel upon light absorption [66]. SR on the other hand, is a light-sensor protein [67].

In microbial rhodopsins, retinal is utilized in the all-*trans* conformation (compared to 11-*cis* for animal rhodopsin) as the ground state that is covalently linked to the K216 side chain through a protonated Schiff base. Absorption of a photon by this chromophore results the isomerization from all-*trans* to 13-*cis* and consequently, conformational changes occur in the protein (**Figure 1.15**).

All trans retinilidene   **13-cis-retinilidene**
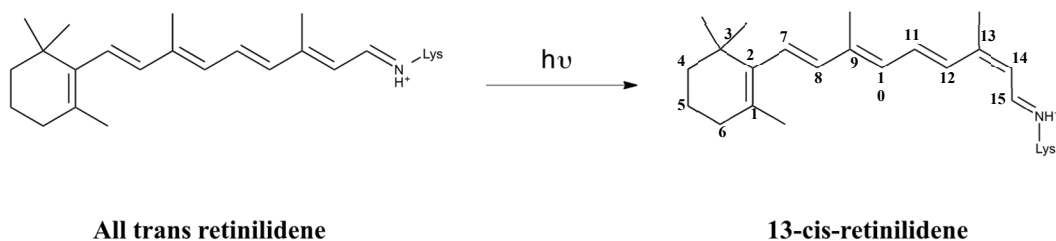
Figure 1.15 Microbial retinal photo-isomerization.

1.7.4 Synthesis of Retinal

Retinal, the aldehyde of vitamin A, is the photosensitive ligand of rhodopsin that can be found linked covalently to the opsin via a specific lysine residue in the binding site. In recent years, many efforts have been employed for synthesizing a variety of retinal analogues in order to understand and modify the spectral and photochemical properties of

rhodopsin. A large number of synthetic chromophore analogs have been incorporated into the binding site of the protein to analyze the chromophore-protein interactions [69, 70]. Substitution of retinal with different analogues may alter the protein absorption characteristics as well as enhance the performance for applications in molecular electronics [71]. Synthetic retinals have also been utilized for modulating the proton-transfer properties of BRs as well as acting as FRET acceptors [72]. There are different synthetic pathways to prepare a vast number of retinal analogues and the main experimental procedure is shown in **Figure 1.16**.

The first step of this procedure [74] is the olefination of the aldehyde with the anion of the phosphonate reagent. This is followed by the reduction of the nitrile group with di-isobutyl aluminum hydride (DIBALH). In the third and forth steps, olefination of the new aldehyde with a phosphonate and reduction of the nitrile group leads to production of the target aldehyde.
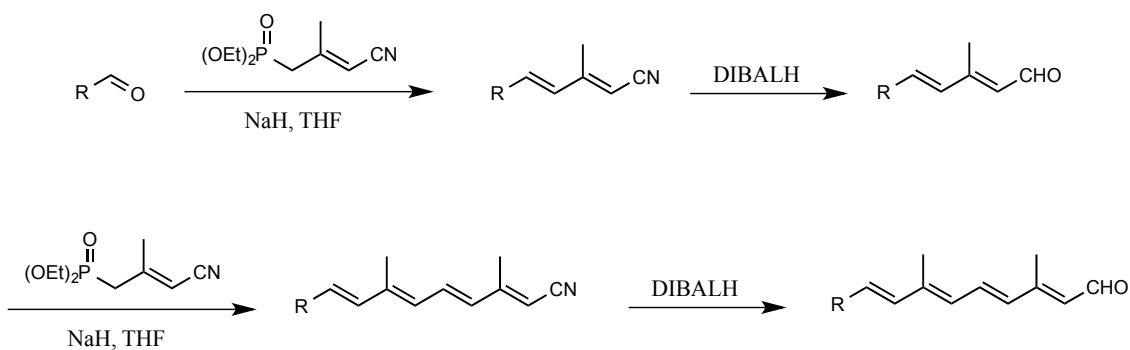
Figure 1.16 Synthesis of retinal and analogues [73].

1.8 Neuroimaging

1.8.1 Cell membrane potential

Cell membranes maintain a small voltage across the membrane in their normal (resting) state, that is called resting membrane potential (RMP). In this state, the inside of the cell is negative with respect to outside because of the different concentrations of ions (K^+, Na^+, Cl^-, and so on) on each side. If membrane potential becomes more negative, it is called hyperpolarization, while if it becomes less negative it is termed depolarization. The cell membrane contains protein channels that allow the K^+ and Na^+ ions to cross the membrane. Additionally, K^+/Na^+ pumps move Na^+ into the cell and K^+ out of the cell in order to maintain a stable RMP (**Figure 1.17)**.

The membrane potential is one of the important properties of living cells, since it modifies intercellular communications and the dynamics of transmembrane proteins. It regulates the activity of ion channels as well as G-protein coupled receptors and phosphatases. Measuring the accurate value or relative changes in membrane voltage is very useful for understanding cellular physiology.
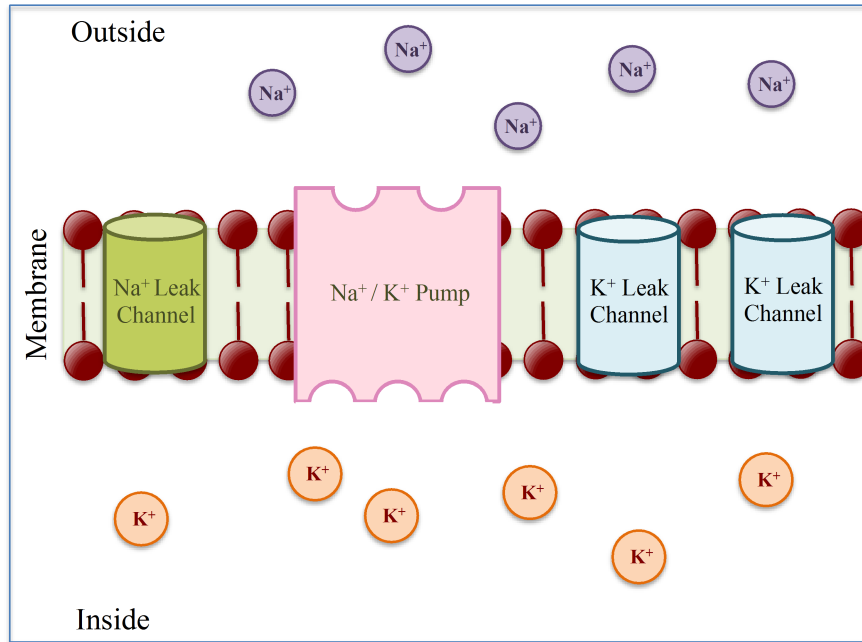
Figure 1.17 Cell membrane maintains a stable membrane potential.

1.8.2 Action Potential

The membrane potential of neurons in the resting state is -70 mV. An action potential is a process in neurons during which the membrane allows the positively charged ions to enter into the cell and negatively charged ions to leave the cell (depolarization). This process is stimulated by a sensory cell or another neuron, and causes the target cell to open the Na^+ channel.

If the potential changes from -70mV to greater than -55 mV **(Figure 1.18)**, an action potential is triggered as voltage gated channels open in the membrane and Na+ rushes into cell. When the membrane potential reaches +30 mV the K^+ channels open

allowing the potassium ions out of the cell. At a certain point, the Na^+ channels close and repolarization returns to the resting state (-70 mV). At this point, sodium and potassium pumps start transporting sodium out of the cell and potassium inside the cell, so that the cell is ready for the next action potential.
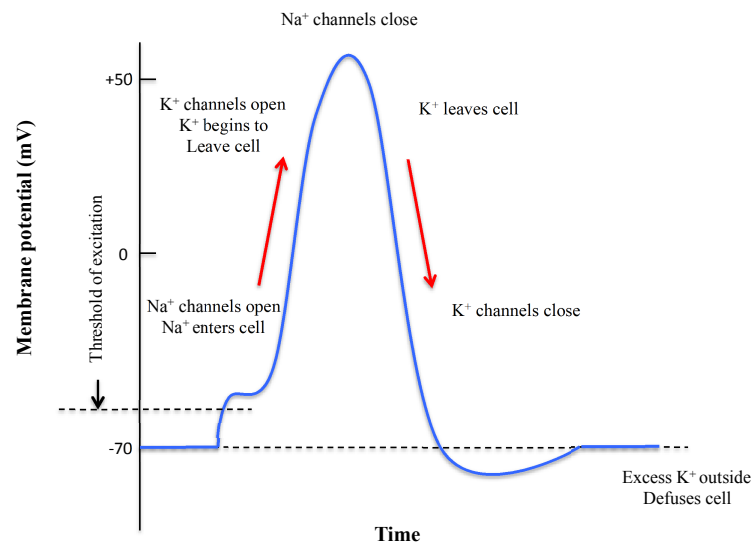
Figure 1.18 The action potential [75].

1.8.3 Neuron Action Potential Imaging Methods

Imaging of membrane potential of neurons is a promising approach to study the individual and collective dynamics of neurons [76]. The idea of optical measurement of neuronal activities was introduced first by Cohen *et al* in 1968 [77]. Subsequently, a variety of optical methods were developed based on organic based-dyes as indicators of Ca^{2+}, protein-dye interactions, FPs, and small molecules. In this thesis, two major methods for reporting the neurons activity have been considered: 1) FP-based voltage indicators; and 2) Microbial rhodopsin-based biosensors.

1.8.3.1 Genetically Encoded Voltage Indicators (GEVIs)

The ability to monitor the activity of neuron circuits is crucial for understanding brain function, and voltage imaging has provided many significant advances in this field. Soon after the molecular cloning of GFP and other spectral modified variants, researchers began working on the construction of genetically encoded voltage indicators for live cell labeling. The mechanism of voltage sensitivity of these types of indicators varies for different constructs. However, all of them must convert a biochemical change associated with an action potential into a fluorescence change [78] [79-81]. In the simplest case, the voltage reporter molecule undergoes a conformational change that leads to a change in its spectra [82].

There are several types of genetically encoded fluorescent voltage sensors. The earliest examples include FLaSH [83] and SPARC [84] that were constructed by insertion of reporter FPs into voltage-gated potassium and sodium channels, respectively. They work based on conformational changes within the voltage-sensing domain that is associated with modulation in fluorescence. Although this generation of voltage indicators was capable of showing voltage changes in membrane, it suffered from poor membrane targeting in cells and cytotoxicity that limited its applications in mammalian systems. Therefore, an improved generation of voltage indicators was evolved called voltage sensitive FPs (VSFP 1, 2, etc.). These GEVIs contain two proteins attached to a voltage sensitive domain of an ion channel or a transmembrane segment [85, 86] and are capable of monitoring electrical events of neurons. The slow kinetics and small signal

magnitudes are some of the issues of VSFP-type voltage indicators that limited their applications for imaging the neuronal activities.

There are yet other classes of GEVIs such as hybrid voltage sensors (hVoS) that employ a membrane-targeted FP. They generate voltage-dependent fluorescence changes in the presence of a synthetic voltage-sensing molecule, dipicrylamine (DPA) as the result of a voltage-dependent optical interaction between the two molecules [87] [88]. These voltage sensors are highly sensitive to voltage changes and could overcome the slow kinetic problem of previous classes. However, some disadvantages include the lack of control over the distribution of DPA in tissues and its toxicity in cells.

In recent years, a new generation of voltage indicators has been developed based on microbial rhodopsin after it was discovered that some rhodopsin variants exhibit weak fluorescence that responds to the membranes voltage changes with high speed and sensitivity [89, 90]. These indicators, including BR and Archaerhodopsin (ARCH), are membrane proteins that bind to retinal via a Schiff base.

1.8.3.2 Microbial Rhodopsin-based Biosensors

Voltage sensitive fluorescence in microbial rhodopsins arises from changes in protonation of the retinal Schiff base. The protonated state is fluorescent whereas the deprotonated state is not. Modulating the pK_a of the Schiff base, by changes in the electric field at the membrane controls the local electrochemical potential of protons and hence, the acid-base equilibrium **(Figure 1.19)** [91].

The primary application of microbial-rhodopsin based biosensors is the detection and recording of action potentials of mammalian neurons at first reported by Adam

Cohen's group in 2012 [89]. The Cohen group first developed the proteorhodopsin-based optical proton sensors (PROPS) that were functional in *E. coli.* However, PROPS did not work in eukaryotic cell. Therefore, they employed ARCH as another type of microbial rhodopsin that can be localized to the plasma membrane of mammalian cells. ARCH-based indicators could detect neuronal action potential with a high signal-to-noise ratio (SNR). However, further modifications (such as D95N) were necessary to improve their functions since they suffer from some limitations such as low voltage sensitivity [91].
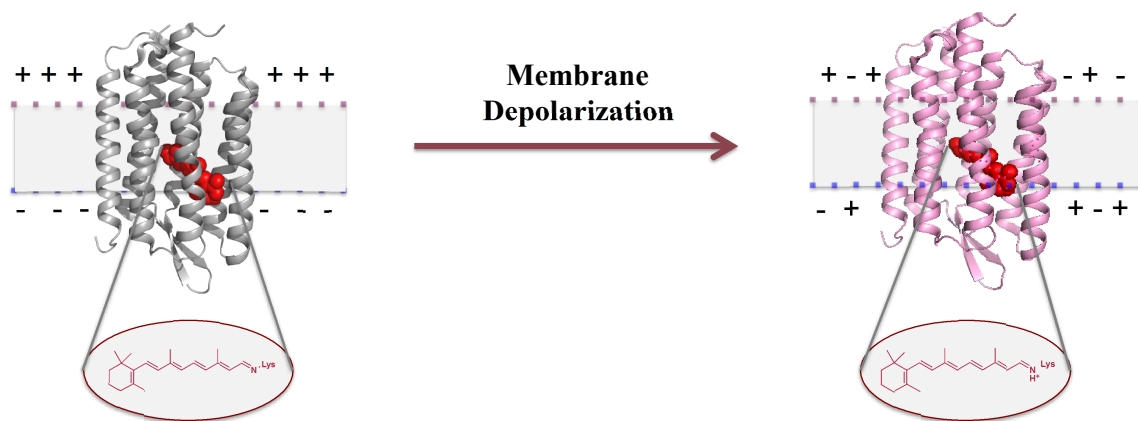
Figure 1.19 Model of microbial rhodopsin as a voltage sensor in which pH and membrane potential can alter the protonation of the Schiff base.

Although genetically encoded voltage indicators suffer from some particular issues and limitations, they are becoming increasingly accepted as an approach for optical interrogation of neuronal activities. Researchers have succeeded in using them to record the membrane potential in a variety of mammalian cell types, including neurons. To address their remaining weaknesses, considerable challenges need to be overcome including improvements in key properties including brightness, photostability, and kinetics.

1.9 Research Objectives

Protein-associated chromophores can be formed intrinsically as in GFP, or extrinsically as in rhodopsin. The spectroscopic properties of the protein are mainly dictated by the structure of the chromophore and secondarily depend on the surrounding microenvironment of the proteins. Many new modified FP-based biosensors have mutations either within the chromophore-forming amino acids or in chromophore-interacting regions. These mutations, which are mostly discovered through directed evolution procedures, are focused on the enhancement of brightness and sensitivity of the protein-based biosensor. In this thesis, I describe our efforts to modulate the fluorescent properties and sensitivity of microbial rhodopsin by substituting all-*trans* retinal with a red-shifted merocyanine analogue that exhibits much brighter fluorescence. We have also attempted to extend this strategy to empty FP barrels by replacing the endogenous chromophore with a synthetic version. We expect that the chromophore replacement strategy could help to provide insight into the mechanism of chromophore formation, and possibly provide new imaging tools with improved fluorescent properties. Potentially, this chromophore replacement strategy can open new windows to detailed understanding the mechanism of chromophore formation. This insight into the molecular level fluorescence mechanism is helpful for modulating the characteristics of the chromophore to provide the most desirable properties in imaging tools.

In Chapter 2, I describe the synthesis of a red-shifted merocyanine retinal analogue and development of modified microbial rhodopsin in terms of fluorescence properties. In Chapter 3, I describe the synthesis of the 4-hydroxybenzylidene-1,2-dimethyl-

imidazolinone (**HBDI**) GFP chromophore analogue. I then go on to describe our efforts to insert this chromophore in the empty circularly permuted EGFP and mCitrine which have their intrinsic chromophores already removed by molecular biology techniques. Future directions and suggestions for improvement of these strategies are discussed in Chapter 4.

Chapter 2: Fluorescent Modulation of Microbial Rhodopsin-Based Biosensors by Retinal Replacement

2.1 Introduction

BR has gained protein engineers' attention since its discovery in *Halobacterium halobium* in 1971 [64] and various technical and optical applications have been explored since that time. BR has some specific properties that make it attractive from a protein engineering perspective. For instance, it has a high stability toward chemical and thermal degradation, and undergoes conformational changes induced by light. The molecular function of BR is as a light-driven proton pump, and the vast majority of investigations have focused on modifying this ability [92].

BR is weakly fluorescent, with an excitation maximum at 560 nm, an emission maximum at 750 nm, and a low quantum yield [93] that limits its application for fluorescence imaging by using standard microscopes. However in 2011, the group of Adam Cohen succeeded in imaging of a modified BR homologue. Specifically, they found that the D97N mutation of PROPS (a homologue of BR) exhibits fluorescence that varies depending on membrane voltage and therefore, could be utilized as a voltage indicator for recording the membrane potential in bacteria [90]. The pKa of the Schiff base between retinal and a lysine residue in the opsin protein varies as a function of voltage, and the change in the protonation state leads to a change in fluorescence. The main limitation of PROPS is that it cannot be functionally expressed in mammalian cells. Therefore, Cohen's lab used ARCH, a homologue that has similar properties to PROPS

and retains its voltage sensitivity when expressed in mammalian cells. Both PROPS and ARCH can be used as voltage sensors, but their applications are limited because of their dim fluorescence.

The objective of the first project described in thesis was to develop brighter fluorescent analogues of PROPS and ARCH that can be used as voltage sensors of membrane potential with high signal-to-noise ratio in bacteria and mammalian cells. The strategy to achieve this aim is reconstitution of opsin variants with synthetic retinal analogues that are more fluorescent than the natural retinal itself. To date, a vast number of retinal analogues have been reported [70, 73]. Some examples of these analogues that are brightly fluorescent and can be used to reconstitute BR and other homologues are shown in **Figure 2.1**.

I

$_{max}$, BR = 590 nm

III

$_{max}$, BR = 552 ,571, 725 nm

II

$_{max}$, BR = 662 nm

IV

$_{max}$, BR = 755 nm
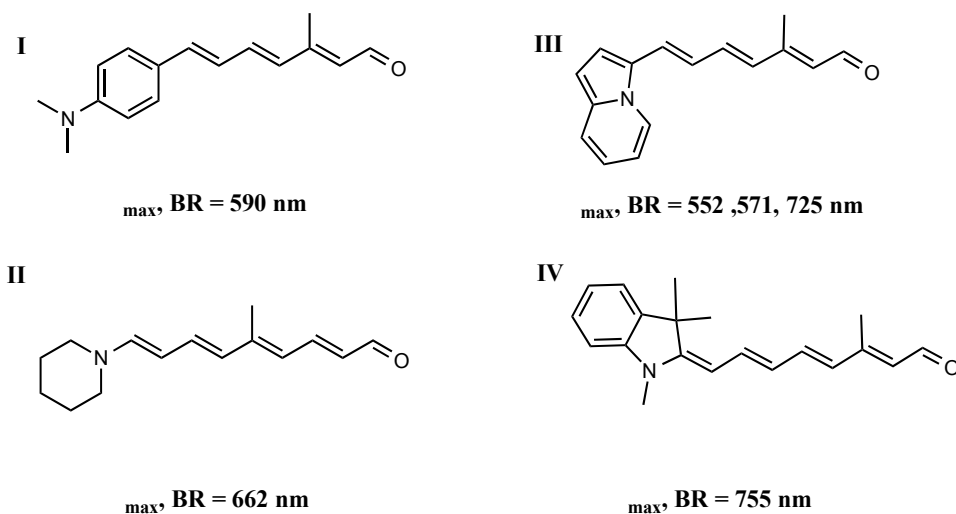
Figure 2.1 Structures of a selection of red-shifted retinal analogues and their emission maxima within the rhodopsin.

In this Chapter, I describe the synthetic pathway to produce merocyanine **(IV)** as a red-shifted retinal analogue. I also describe our efforts to reconstitute the ARCH and PROPS opsin for brighter fluorescence with this retinal chromophore in *E. coli*.

2.2 Results and Discussion

2.2.1 Synthesis of a Red-Shifted Merocyanine Retinal Analogue

My work started with synthesis of merocyanine **(IV)** as a red-shifted and more fluorescent retinal analogue that was already reported by Hoischen and co-workers [109]. Here after, merocyanine **(IV)** is called S_4. We chose this molecule because it has the most red-shifted emission wavelength among the others As mentioned in the first Chapter, the main procedure for synthesizing retinal analogues is based on the olefination of the starting aldehyde and extension of the conjugation system **(Figure 1.16)**. Therefore, in order to obtain the retinal analogue of interest **(IV)**, four steps should be performed starting with olefination of 2-(1,3,3-trimethylindolin-2-ylidene) acetaldehyde, S_0 **(Figure 2.2)**.
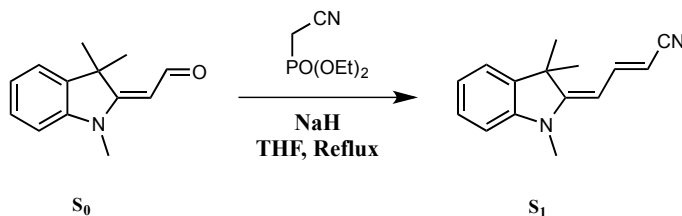
Figure 2.2 The first step of Merocyanine synthetic pathway based on HWE reaction.

Olefination of aldehydes and ketones with phosphonate carbanions as the nucleophile is done by a modified version of the Wittig reaction called the Honor-Wadsworth-Emmons (HWE) reaction [94]. The HWE reaction begins with deprotonation of the phosphonate to produce a carbanion. Nucleophilic addition of the carbanion to the aldehyde is followed by cyclization of the intermediate and subsequent elimination yields the alkene of interest **(Figure 2.3)**.
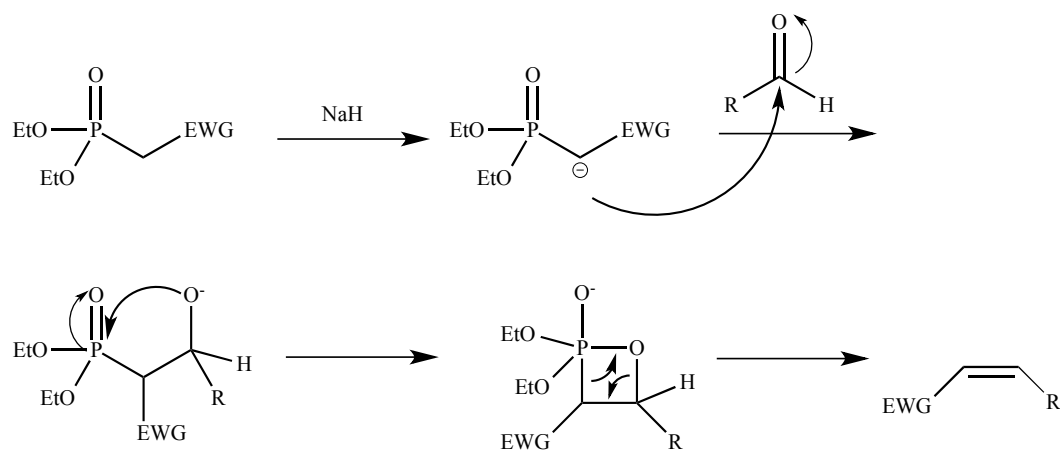
Figure 2.3 Honor-Emmons-Wadsworth (HWE) reaction mechanism (EWG: electron-withdrawing group).

In my effort to synthesize the target molecule, S_1 was produced by the reaction of S_0 with diethyl cyanomethyl phosphonate in the reflux condition for overnight in 48% yield **(Figure 2.2)**. The second step of this procedure is the reduction of the nitrile functional group of S_1 to an aldehyde by using diisobutylaluminum hydride (DIBALH) as the reducing reagent **(Figure 2.4)**.
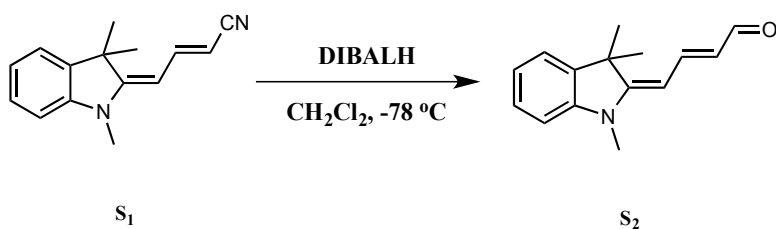
Figure 2.4 Reduction of nitrile group to aldehyde: the second step of merocyanine synthetic pathway.

The reduction reaction was performed at -78 $^{\circ}$C for 2 hours and the aldehyde S_2 was obtained with a yield of 60%. In order to avoid the formation of by-products, the

DIBALH was added at a 1:1.2 reactant/DIBALH stoichiometric ratio by dropwise over at least a 15-minute window.

In order to extend the conjugation of the molecule, two more steps are needed: HWE reaction on **S₂** to obtain **S₃**, and reduction of nitrile group to aldehyde to produce the final target molecule, **S₄ (Figure 2.5)**.
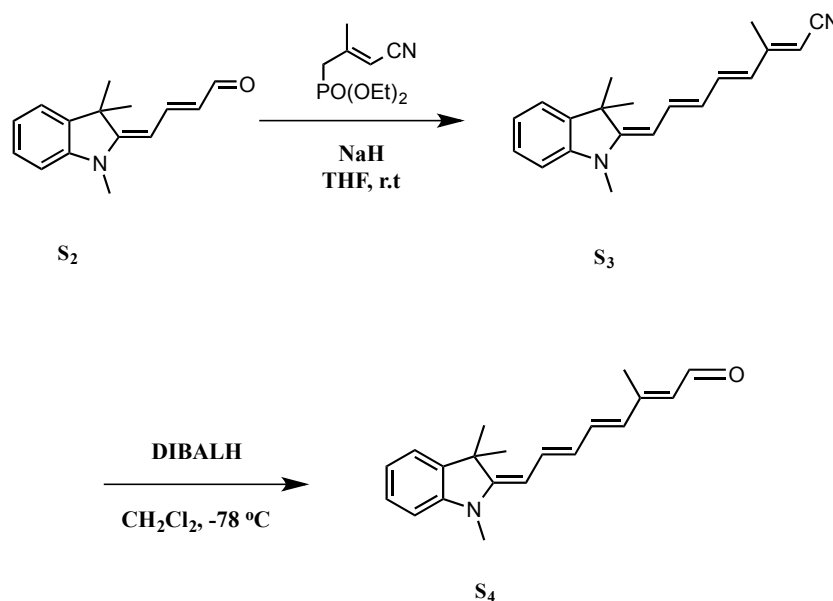
Figure 2.5 The third and forth steps of merocyanine synthesis.

By performing another HWE reaction on aldehyde **S₂** using diethyl (3-cyano-2-methylallyl) phosphonate at room temperature overnight, the new nitrile molecule **S₃** was obtained in 51% yield.

The final step to synthesize the target aldehyde **S₄** was very challenging and it took a lot of effort. In the first trials, the same procedure used for synthesizing **S₂** was followed but was not successful. The reaction either formed only by-products or resulted in starting materials recovered. Presumably, the nitrile moiety of **S₃** is less electrophilic compared to

the nitrile group of S_1. One reason might be because the longer conjugation chain of the molecule delocalizes the electrons better than shorter chains. Consequently, it decreases the positive character of the nitrile center and therefore, according to the mechanism of reduction of nitrile to aldehyde by DIBALH **(Figure 2.6)**, it is more difficult for the hydride ion to attack to the nitrile.

Figure 2.6 The mechanism of reduction of the nitrile functional group to aldehyde using DIBALH.

In order to overcome this problem, a variety of conditions including different solvents (CH$_2$Cl$_2$, Toluene, THF), reaction time, temperature (-78 $^\circ$C to r.t, -78 $^\circ$C to reflux, 0 $^\circ$C to r.t and 0 $^\circ$C to reflux), and also different proportions of DIBALH to S_3 (1:1, 3:1, 10:1 and 20:1) were tested, since different amounts of DIBALH lead to different products. For instance, increasing the amount of DIBALH results in the production of an amine instead of an aldehyde. However, none of these conditions led to production of the

desired target molecule. Therefore, we came up with an alternative solution and we decided to convert the nitrile group to an ester hoping to reduce the ester to an aldehyde.

For synthesizing the ester moiety, **ES₃** by the HWE reaction triethyl 3-methyl-4-phosphono-2-butenoate was used in place if the phosphonate regent **(Figure 2.7).** This procedure led to production of **ES₃** with a 52% yield.
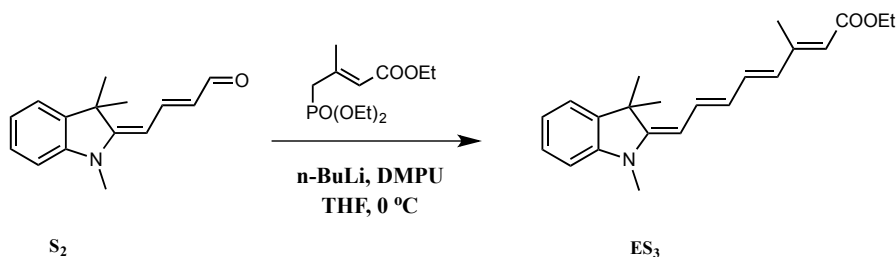
Figure 2.7 Synthesis of ester ES3 moiety from aldehyde **S₂**.

Once the ester molecule was obtained, the reduction step with DIBALH was performed and different conditions were tried. However, the obtained product was not the desired molecule of interest, **S4**. The mass spectrum of the crude solution indicated the presence of a molecule with the molecular weight of 295.2 g/mole and chemical formula of $C_{20}H_{25}NO$, which correlates with the alcohol moiety shown in **Figure 2.8**.
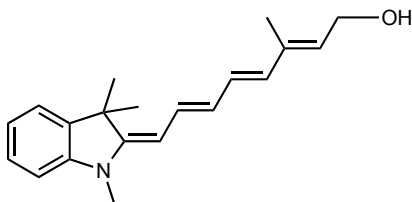
Figure 2.8 S4-OH, Unwanted product of reduction reaction of **ES₃**.

In order to overcome these problems and synthesizing the **S₄** molecule, I joined Professor Kazuya Kikuchi's group in Osaka University, Osaka, Japan as an exchange student in February 2014. With the advice and guidance of Dr. Toshiyuki Kowada (an assistant professor in this group) I successfully obtained the **S₄** aldehyde with a yield of 67%. The main procedure that we followed for this aim was almost the same as the procedure I had tried in Dr. Campbell's lab. However, a few minor differences in reaction conditions and workup details led to this achievement. These differences are described are here.

- For my first tries in Campbell's lab, we always used the anhydrous solvent whereas; in Kikuchi's lab we just used normal THF. Therefore, we now assume that the presence of water in this reaction must be critical. Although the second step of this reaction proceeded despite using the anhydrous solvent leaving us with a small mystery.

- In this reaction the minimum amount of DIBALH was used (1.2 of DIBALH to 1 of **S₃**) which is necessary to prevent the reaction proceeding directly to amine production rather than producing the aldehyde.

- The reaction was stopped and quenched after 80 minutes with saturated ammonium chloride solution since a longer reaction time would lead to unwanted products.

Although these points may not appear very crucial, they made a huge difference in the final result.

The **S₄** molecule, (2E,4E,6E)-3-methyl-8-((E)-1,3,3-trimethylindolin-2-ylidene)octa-2,4,6-trienal has a molecular weight of 293.2 g/mole. However, as shown in the NMR spectra of **S₄** (Appendix A), the sample is a mixture of isomers (3:1) and because of the close polarities of these two fractions the separation was not possible. In

the ^{1}H NMR spectrum the coupling constant (^3J) values of these isomers are 8.3 Hz and

8.1 Hz for peaks at 10.1 ppm and 10.2 ppm respectively, suggesting that the correlation

of H$_1$ and H$_2$ in both isomers are similar **(Figure 2.9)**.
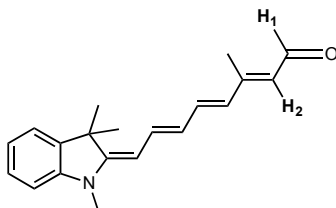
Figure 2.9 The correlation of aldehyde proton (H$_1$) and its vicinal proton (H$_2$) with the coupling constant ^3J value of 8 Hz.

On the other hand, since the ^3J values for all of the other vicinal protons in the

conjugated chain are from 11.2 to 15 Hz, we can conclude that all the double bonds of the

major isomer are *trans*. However; it is not easy to determine the conformation of the

minor isomer because it does not have sufficient SNR to disstinguish coupling constants.

2.2.2 Spectral Characterization of Merocyanine (S$_4$)

By confirming the **S$_4$** structure with ^{1}H, ^{13}C NMR, 2D NMR and mass spectrometry

(Appendix A), the spectral properties were characterized as well. **S$_4$** has an absorbance

maximum at 489 nm and an emission maximum at 689 nm with a 200 nm Stokes shift

(Figure 2.10A).

Compared to the original all-*trans* retinal that has the absorbance maximum at 392 nm

and an emission maximum at 442 nm (Stoke shift of 50 nm), **S$_4$** shows a remarkable red-

shifted emission maximum with a far larger Stokes shift. The absorbance and emission

spectra of all-*trans* retinal were obtained with the same condition as **S₄** and are shown in

(Figure 2.10B).
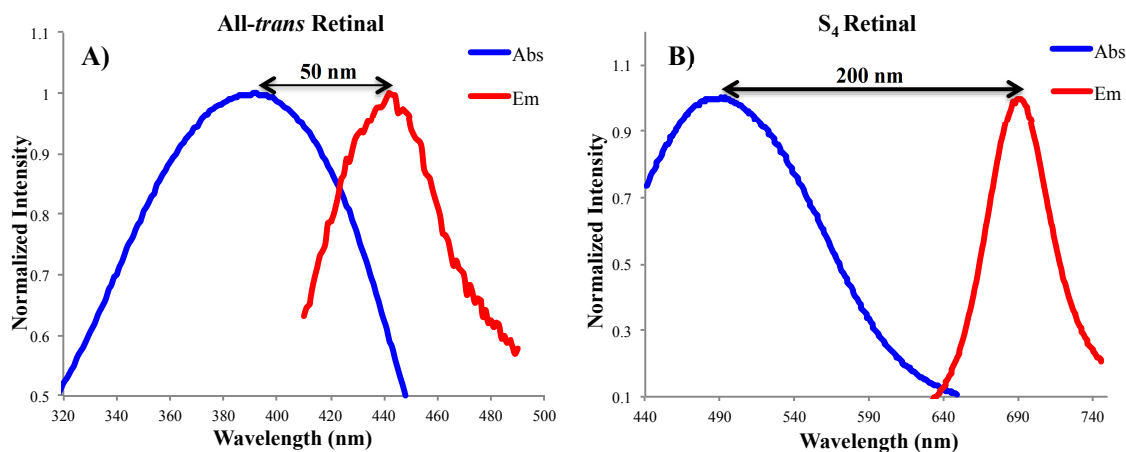
Figure 2.10 Normalized overlaid absorbance and emission spectra of A) all-*trans* retinal with 50 nm Stoke shift and B) **S₄** with 200 nm Stoke shift.

The comparison of the emission intensities of equal concentrations of **S₄** and all-*trans* retinal indicates a significant difference in terms of brightness of **S₄** over all-*trans* retinal **(Figure 2.11A)**.

As mentioned in Chapter 1, the brightness of a fluorophore is proportional its to quantum yield and extinction coefficient. The extinction coefficient values (ε) for all-*trans* retinal (28136 $M^{-1}cm^{-1}$) and **S₄** (47605 $M^{-1}cm^{-1}$) were obtained by measuring the absorptions in DMSO at 375 nm and 489 nm, respectively and were calculated based on the Beer-Lambert Law.
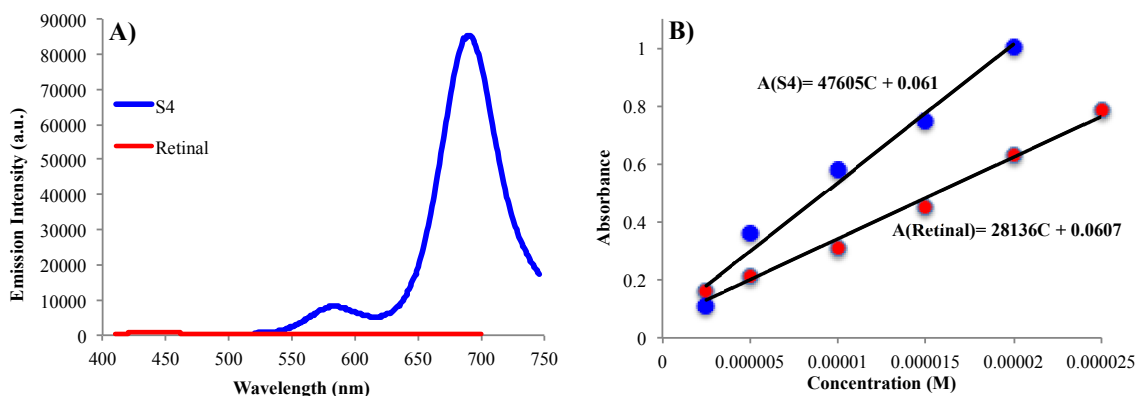
Figure 2.11 A) The comparison of S₄ and all-*trans* retinal emissions (excitation wavelength: 483 nm and 380 nm, respectively). B) The absorbance vs concentration data used for extinction coefficient calculation.

In addition, the quantum yield for **S₄** retinal, 0.058 ± 0.008 was measured using fluorescein (in 0.1 M NaOH) as the standard. In **Table 2.1** the optical information obtained for **S₄** and all-*trans*-retinal was summarized. Since the reported quantum yield for all-*trans* retinal is not obtained in the same solvent as used for **S₄**, the observed difference between the brightness of **S₄** and all-*trans* retinal **(Figure 2.11A)** does not correlate with information in **Table 2.1**.

Table 2.1 The optical information obtained for all-*trans*-retinal and S₄.

	$\lambda_{Max, Ex}$ (nm)	$\lambda_{Max, Em}$ (nm)	ε (M⁻¹cm⁻¹)	ϕ [a]
all-*trans*-retinal	392	442	28136	7×10^{-5} [a]
S₄	489	689	47605	0.058 ± 0.008

[a] The quantum yield value for all-*trans* retinal was reported in literatures in EPA (ether-isopenthane-ethanol 5:5:2) [95].

The **S₄** retinal fluorescence emission spectra recorded at 483 nm excitation at different pH values from 5-12 **(Figure 2.12)**. According to this figure, by decreasing the

pH, the absorbance peak is shifted to shorter wavelengths. Decreasing the pH also leads to an increase in the fluorescence, suggesting that the protonated form of **S4** molecule (has two functional groups; carbonyl and amine group) in the excited state is able to fluoresce more efficiently compared to the neutral form. This phenomenon is similar though opposite to GFP, where the anionic state is more fluorescent compared to neutral form. The competition between pH-sensitive non-radiative decay and pH-sensitive radiative decay for charged and neutral form of the fluorophore are different and leads to a different fluorescence quantum yield [96].
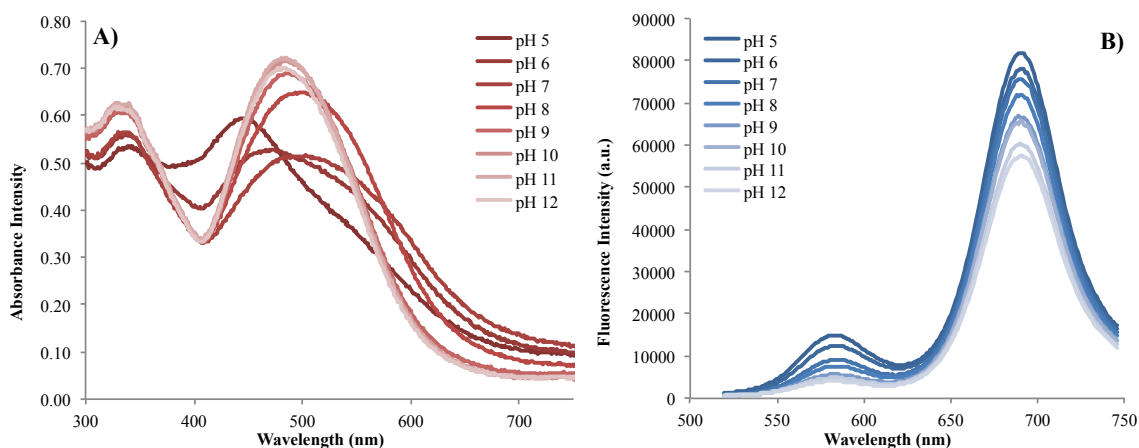
Figure 2-12 Dependence of A) absorption and B) fluorescence spectra (excited at 483nm) of S4 on pH.

On the other hand, by changing the pH, the interaction of the molecule with the environment and solvation capacities would change and that can also lead to different fluorescence emission by fluorophore.

2.2.3 Reconstitution of ARCH and PROPS using S_4 Retinal

In order to study the behavior of rhodopsin with synthetic retinal S_4, ARCH, and five mutants of ARCH, and also two variants of PROPS were examined. These variants were engineered by Dr. Yongxin Zhao, which are presented in **Table 2.2** [97].

The locations of mutations in QuasAr1 modeled on ARCH-2 crystal structure are shown in **Figure 2.13**. This structure indicates that the P60S mutation is close to the Schiff base of retinal suggesting an important role of this mutation in modification of voltage sensitivity and brightness of the protein. Cohen's group studies also showed that D95N mutation in ARCH increases its voltage sensitivity [91]. However, the location of T80S and F161V mutations on the protein barrel suggests a low probability of direct impact on the photophysical properties of the protein. They are more likely to affect the protein folding.

Table 2.2 ARCH and PROP mutations.

Name	Mutations
ARCH	-
ARCH D95N	**ARCH** D95N
QuasAr1	**ARCH** P60S/T80S/D95H/D106H/F161V
QuasAr2	**ARCH** P60S/T80S/D95Q/D106H/F161V
2.14	**ARCH** D95H/D106H/P60S
2.18	**ARCH** D95H/D106H/F161V
PL4-98	**PROPS** S65F/D98N/F152L/K172E/N220D/I237V
PL8-41	**PROPS** S65F/D98M/ F152L/A116D/ P201S/I237V

According to Dr. Yongxin Zhao's investigations [97], T80S, F161V or P60S as an individual mutation with D95H ad D106H mutations does not improve the brightness of

the ARCH protein indicating the important effect of the combination of these three mutations on brightness improvement.

As the first step, these eight rhodopsin genes were transformed in *E. coli* and cultured in LB (Luria Broth)- media and then, **S₄** was added at the same time as adding inducer. Therefore, the proteins were expressed in the presence of **S₄** retinal. Among these variants, three of them (QuasAr2, PL4-98 and PL8-41) showed a remarkable red shift in excitation and emission compared to ARCH whereas the other proteins indicated the same behavior as ARCH. In **Figure 2.14**, the overlaid excitation and emission peaks for the improved variants as well as ARCH are presented.
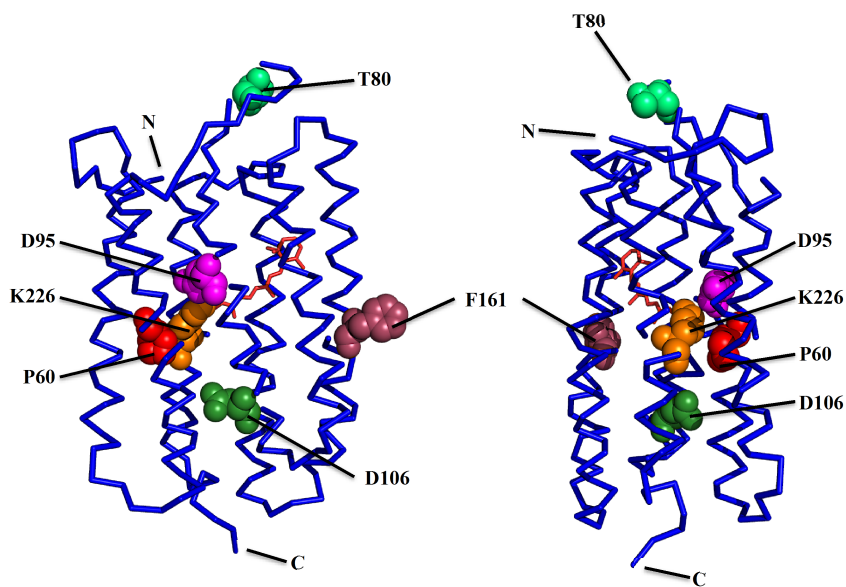
Figure 2.13 The modeled structure of QuasAr1 based on the homologous protein ARCH-2 (PDB: 2EI14) [98].

Comparison of the excitation and fluorescence emission maxima of QuasAr2- all-*trans*-retinal (590 nm and 715 nm, respectively) [97], with excitation and fluorescence emission maxima of QuasAr2-**S₄** (752 nm and 779 nm, respectively) indicates 162 nm shift of excitation and 64 nm red-shift of the emission.

As shown in **Table 2.2**, QuasAr1 and QuasAr2 (superior quality to ARCH) differ from wild-type ARCH by the mutations P60S, T80S, D106H, and F161V. In addition, QuasAr1 has the D95H mutation whereas QuasAr2 has D95Q. This mutation is likely the main reason of the difference in behavior of QuasAr1 and QuasAr2 towards the reconstitution with **S4** retinal. Glutamine (Q) is more compatible with the chromophore, either by being more flexible than histidine (H) or having a better shape or better interaction with the chromophore leading to brighter fluorescence.
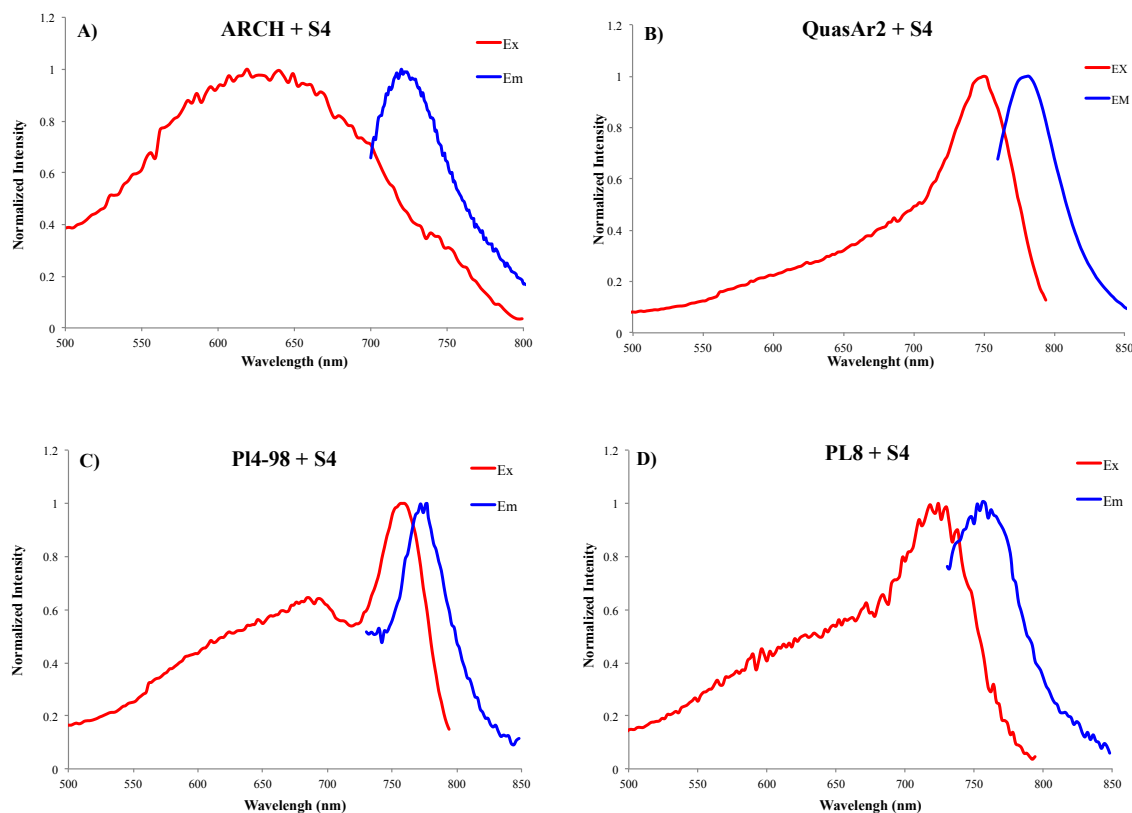
Figure 2.14 The excitation and emission spectra of the different ARCH variants reconstituted with the synthetic chromophore. A) ARCH with λMax, Ex= 640 nm and λMax, Em= 720 nm B) QuasAr2 with λMax, Ex= 752 nm and λMax, Em= 779 nm C) PL4-98 with λMax, Ex= 756 nm and λMax, Em= 772 nm and D) PL8-41 with λMax, Ex= 727 nm and λMax, Em= 762 nm.

The K_d values were obtained for all of these complexes by measuring the emissions with different concentrations of the chromophore. The obtained curves are shown in **Figure 2.15.**
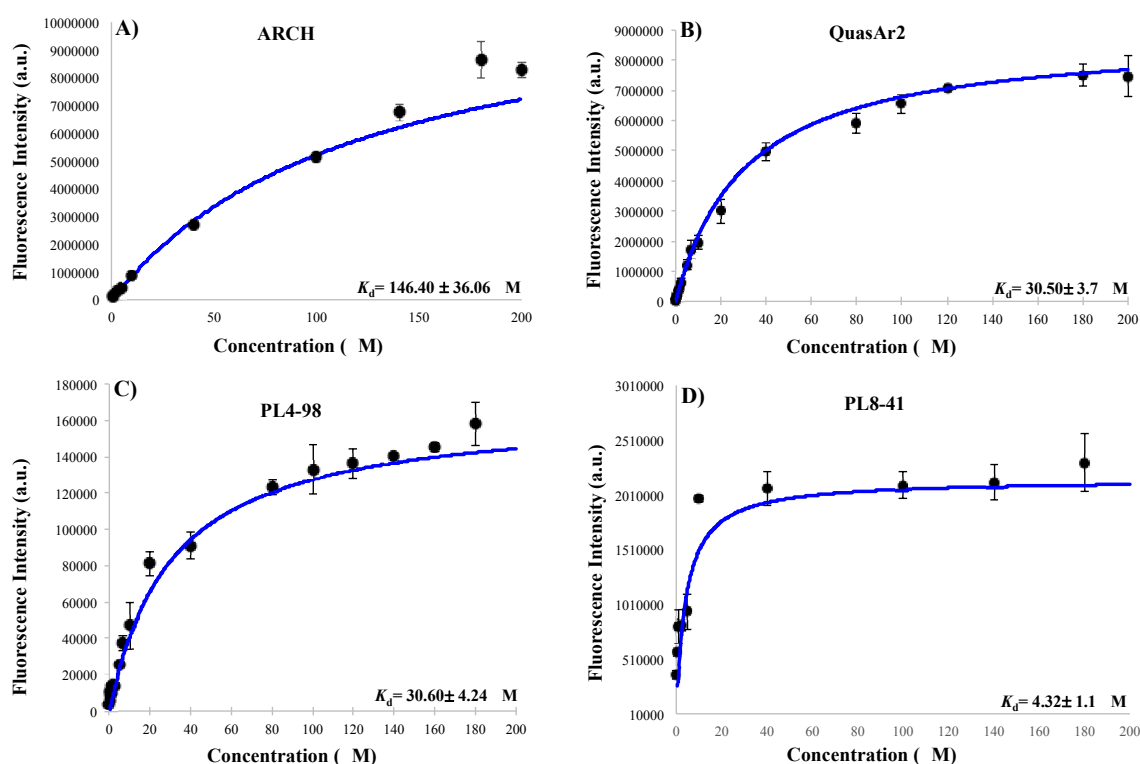
Figure 2.15 Saturation binding curve for A) ARCH, B) QuasAr2, C) PL4-98, and D) PL8-41. PL8-41 indicates the highest binding affinity.

According to these results, PL8-41 has the lowest K_d value (4.32 ±1.1 µM) and highest binding affinity to **S$_4$** retinal, while PL4-41 and QuasAr2 has lower binding affinity (K_d = 30.60 ± 4.24 µM and 30.50 ± 3.7 µM), respectively. **Table 2.3** summarizes the obtained information for the best variants and ARCH.

Table 2.3 The optical information and K_d values for the best variants with S_4.

	$\lambda_{Max, Ex}$ (nm)	$\lambda_{Max, Em}$ (nm)	Stokes Shift (nm)	K_d (µM)
ARCH	640	720	80	146.40 ± 36.06
QuasAr2	752	779	27	30.50 ± 3.7
PL4-98	756	772	16	30.60 ± 4.24
PL8-41	727	762	35	4.32 ± 1.1

Finally, the pH- dependence excitation curves for QuasAr2 and PL4-98 were obtained in pH interval 3.4 to 11.3 **(Figure 2.16)**. According to the data, the emission intensity at lower pHs is higher for QuasAr2 indicating that the protonated state of the Schiff base bond between retinal and lysine has a higher fluorescence. However, PL4-98 (PROPS variant) does not follow this trend and by increasing the pHs the fluorescence intensity is increased.
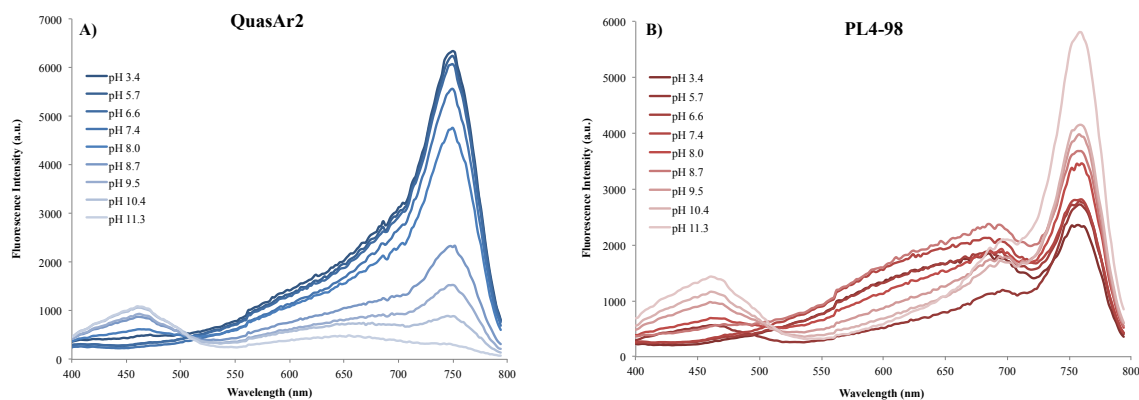
Figure 2.16 pH-dependence excitation curves for A) QuasAr2, and B) PL4-98.

2.3 Conclusion

In summary, we have synthesized a retinal analogue **S₄** with a remarkable red-shifted absorbance and fluorescence emission spectra with 200 nm stokes shift and also a higher extinction coefficient compared to all-*trans* retinal. This fluorophore was used to reconstitute the ARCH and PROPS to improve their photophysical properties. Among eight protein variants expressed in *E. coli,* QuasAr2 (ARCH mutant), PL4-98, and PL8-41 (PROPS mutants) showed a significant red shift in excitation (~100 nm shift) and fluorescence emission (~50 nm). QuasAr2 was the most red-shifted protein. These improved variants could be considered as suitable candidates as voltage sensors and cell imaging tools. As an attempt for imaging, neurons were transfected with plasmid containing QuasAr2 and synthetic retinal was added **S₄** and after incubation at 37 $^\circ$C for 1 hour, the cells were imaged. However, **S₄** penetrated into the mitochondria and imaging was not successful.

2.4 Materials and Methods

2.4.1 General Methods and Materials

All the organic compounds and solvents were purchased from Aldrich. 1D and 2D proton nuclear magnetic resonance spectra were obtained at 300 MHz, 400 MHz, and 500 MHz in CDCl₃ and coupling constants (*J*) are reported in hertz (Hz). Carbon nuclear magnetic resonance spectra were recorded at 100 or 126 MHz. The chemical shifts are reported on the δ scale (ppm) downfield from tetramethylsilane (0 ppm). The molecular weight of

synthetic products was determined by mass spectra obtained in positive mode of a high-resolution electrospray ionization mass spectroscopy.

The six ARCH variants and two PROPS genes were provided by Dr. Yongxin Zhao in pBAD plasmid. The MicroPulser cuvettes (Bio-Rad Inc.) and MicroPulser electroporator (Bio-Rad Inc.) were used from electroporation to transform the gene to bacteria. The absorption, excitation and emission spectra were acquired by using Beckman Coulter DU 800 UV/Vis spectrometer and 384-well microplate reader (Tecan safireII^{TM}).

2.4.2 Synthetic Procedures and Structural Characterizations

1,3,3-trimethylindolin-2-ylidene)but-2-enenitrile (S$_1$)
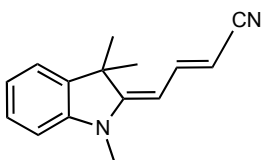
Under Argon gas at 0 $^\circ$C, NaH 60% (160 mg, 4 mmol) was added to 5 mL of THF and then diethyl cyanomethyl phosphonate (647 mg, 4 mmol) dissolved in 1mL THF was added dropwise. This solution was stirred for 1 hour to produce the phosphonate carbanion. The solution of aldehyde S$_0$ (603 mg, 3 mmol) was added dropwise over 15 min and the reaction was refluxed at 65 $^\circ$C overnight. Before working up the reaction, about 3 gr of ice was added in the reaction to cool it down, and 10 mL of saturated solution of NH$_4$Cl was used for quenching. The crude product was extracted by ethyl acetate and dried with sodium sulphate. The product was purified by flash column

chromatography using hexane: ethyl acetate (8:1) and an oily orange compound was obtained with 48% yield. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.36 (dd, J= 10.8, 12.8 Hz, 1H), 7.22 (ddd, J< 1, J= 1.2, 7.6, 8Hz, 1H), 7.16 (dd, J< 1, J=6.4, 1H), 6.94 (ddd, J< 1, J= 7.2, 8.4, 1H), 6.73 (d, J= 7.9, 1H), 5.7 (d, J= 12.8, 1H), 4.68 (dd, J< 1, J= 10.4, 1H), 3.22 (s, 3H), 1.55 (s, 6H); MS m/z calcd for C$_{15}$H$_{17}$N$_2$ (M+H$^+$) 225.13, found 225.1

1,3,3-trimethylindolin-2-ylidene)but-2-enal (S$_2$)
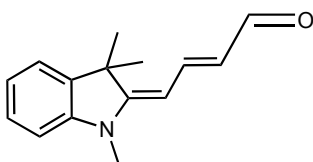
At -78 $^°$C under Argon, **S$_1$** (291.4, 1.3 mmol) was dissolved in CH$_2$Cl$_2$ and DIBALH (1.3 ml, 1.3 mmol) was added dropwise within 15 minutes. The reaction stirred for 3hours and then, warmed up to room temperature. 30 mL of NH$_4$Cl 5% was added to quench the reaction and stirred for 1 hour. The organic layer was separated and dried with sodium sulfate. The product was purified by flash column chromatography using hexane: ethyl acetate (3:1) and a dark yellow compound was obtained with yield of 60%. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.46 (d, J= 8.4, 1H), 7.72 (dd, J= 12.8, 13.2, 1H), 7.19-7.27 (m, 2H), 6.98 (dd, J= 6.4, 13.8, 1H), 6.77, (d, J= 7.9, 1H), 5.98 (dd, J= 8.2, 13.9, 1H), 5.55 (d, J= 12.4, 1H), 3.23 (s, 1H), 1.63 (s, 6H); ^{13}C NMR (100.40 MHz, CDCl$_3$) δ 192.26, 165.51, 150.12, 143.94, 138.96, 127.94, 123.08, 121.61, 121.43, 107.21, 94.86, 46.64, 29.34, 28.34; MS m/z calcd for C$_{15}$H$_{18}$N$_2$ (M+H$^+$) 229.1383, found 229.1378.

3-methyl-8-(1,3,3-trimethylindolin-2-ylidene)octa-2,4,6-trienenitrile (S₃)
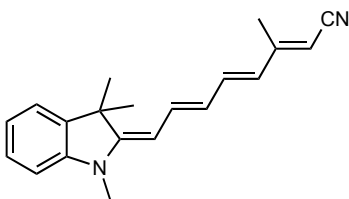
In 5 of mL THF, NaH (60% dispersion in mineral oil, 33.6 mg, 0.84 mmol) was added at 0 °C under argon and then, a solution of diethyl (3-cyano-2-methylallyl) phosphonate (221.7 mg, 0.84 mmol) in 1 mL of THF was added dropwise and the solution was stirred for 1 hour. **S₃** aldehyde (159 mg, 0.7 mmol) was dissolved in 3 mL of THF, added dropwise, and stirred at room temperature overnight. Quenching the reaction was performed by adding approximately 3 g of ice and 10 mL of saturated solution of NH_4Cl. The organic layer was extracted with ethyl acetate and washed by a brine solution. The product was purified by flash column chromatography using hexane: ethyl acetate (8:1) as the eluent and a red powder was obtained with 51% yield. The NMR spectrum indicates that the obtained compound is the mixture of isomers. ^1H NMR (400 MHz, CDCl₃) δ 7.17-7.12 (m, 2H), 7.04-6.98 (m, 1H), 6.86 (dd, J= 7.3, 7.3, 1H), 6.74-6.62 (m, 2H), 6.18-6.03 (m, 1H), 5.38 (dd, J= 10.5, 11.2, 1H), 3.20 (d, J= 5.5, 3H), 2.19 (s, 3H), 2.04 (s, 2H), 1.58 (s, 6H); MS *m/z* calcd for $C_{20}H_{22}N_2$ (M+H⁺) 291.1775, found 291.1778.

3-methyl-8-((*E*)-1,3,3-trimethylindolin-2-ylidene)octa-2,4,6-trienal (S₄)
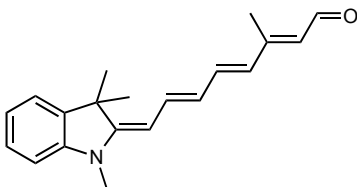
Under the argon atmosphere and at -78 °C, **S₃** (51 mg, 0.175 mmol) was added to 3 mL of CH₂Cl₂. DIBALH (0.18 ml, 0.18 mmol) was added dropwise and stirred for 80 minutes. By adding this reagent, the orange solution turned to dark red. Then, the reaction was warmed up to room temperature and quenched by saturated solution of NH₄Cl. The organic layer was separated and washed with brine solution. Sodium sulfate was used to dry the solution and the product was purified by flash column chromatography using hexane: ethyl acetate (3:1). The target molecule was dark red powder. The NMR spectra show trace amount of impurities. ^{1}H NMR (500 MHz, CDCl₃) δ 10.1 (d, J= 8.3, 1H), 7.22-7.16 (m, 2H), 7.08 (dd, J= 13.1, 13.2 1H), 6.96-6.88 (m, 2H), 6.64 (d, J= 7.9, 1H), 6.24 (d, J= 15.0, 1H), 6.16 (dd, J= 11.2, 14.1, 1H), 5.95 (d, J= 8.3, 1H), 5.42 (d, J= 12.3, 1H), 3.16 (s, 3H), 2.36 (s, 3H), 1.61 (s, 6H); ^{13}C NMR (126 MHz, CDCl₃) δ 190.81, 159.38, 155.76, 144.83, 138.92, 138.45, 135.55, 129.64, 127.85, 127.28, 125.76, 124.30, 121.64, 120.01, 106.25, 96.64, 45.86, 29.15, 28.34, 13.11; MS m/z calcd for C₂₀H₂₃NO (M+H⁺) 294.1780, found 294.1778.

Ethyl-3-methyl-8-((E)-1,3,3-trimethylindolin-2-ylidene)octa-2,4,6-trienoate (ES₃)
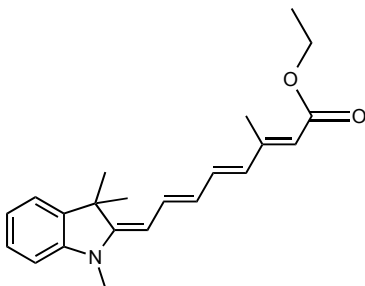
In 5 mL of THF, the solution of triethyl 3-methyl-4-phosphono-2-butenoate (0.2 ml, 0.75 mmol) in 1 mL of THF, DMPU (0.2 ml, 1.5 mmol) and n-BuLi (0.5 ml, 0.75 mmol) were

added at 0 $^{\circ}$C under N$_2$ gas. The reaction was stirred for 20 minutes and then the temperature was reduced to -78 $^{\circ}$C and **S$_2$** (113.5 mg, 0.5 mmol) dissolved in 6 mL of THF and added dropwise. The reaction was stirred for 3 hours and cooled to 0 $^{\circ}$C and stirred for overnight. The product was purified by flash column chromatography by using hexane: ethyl acetate (5:1) as the eluent with 52% yield. MS *m/z* calcd for C$_{22}$H$_{27}$NO$_2$ (M+H$^+$) 338.2037, found 338.2036.

3-methyl-8-((*E*)-1,3,3-trimethylindolin-2-ylidene)octa-2,4,6-trien-1-ol (S$_4$-OH)
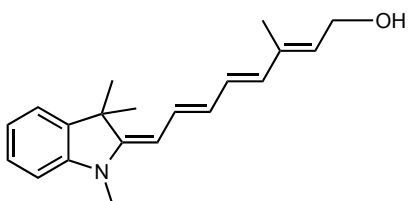
Under N$_2$ gas and at -78 $^{\circ}$C, **ES$_3$** (101.2 mg, 0.3 mmol) was dissolved in 5 mL of CH$_2$Cl$_2$, and DIBALH (0.5 ml, 0.5 mmol) was added dropwise. The reaction was stirred for 5 hours and quenched by saturated NH$_4$Cl after warming up the reaction to room temperature. The product was not obtained by column chromatography because of high polarity, however; the mass spectrometer of the crude product confirmed the molecular weight of **S$_4$-OH**. MS *m/z* calcd for C$_{20}$H$_{26}$NO (M+H$^+$) 297.2009, found 296.2012.

2.4.3 Merocyanine Retinal (S$_4$) Optical Characterizations

A stock solution of **S$_4$** and all-*trans* retinal with the concentration of 10 mM in DMSO was prepared and the absorption, excitation and emission spectra were obtained by

diluting it in different concentrations with PBS (phosphate buffered saline) or water. The extinction coefficient values were calculated by measuring the absorbance of different concentration of S_4 and all-*trans* retinal (2.5 μM to 25 μM) in DMSO and 489 and 375 nm, respectively. The pH-dependence absorbance and fluorescent emission curves of S_4 at the pH interval 5 to 12 recorded were by making the 40 μM solutions with Carmody buffer [99]. The absorption and emission of the samples in each pH were read three times and the average values were used to plot the curves.

2.4.4 Reconstitution of ARCH and PROPS Variants Using S_4

ARCH, five mutated variants of ARCH and two PROPS were transformed into ElectroMax DH10B *E. coli.* Transformations were plated on LB-agar media supplemented with ampicillin and 0.003% L-arabinose. For each gene, one colony was picked into 4 mL of LB-media containing 100 μg/mL ampicillin and incubated at 37 °C for overnight. To express the protein, 0.02% arabinose was added at the same time with adding the S_4 retinal and incubated at 37 °C for 3.5 hours. The cells were pelleted in a centrifuge and the liquids were removed. The pellets were resuspended with 150 μL of PBS. In a 384-well microplate, 50 μL of PBS and 10 μL of suspension of each gene was added and the absorption, excitation and emission data were recorded **(Figure 2.17)**.
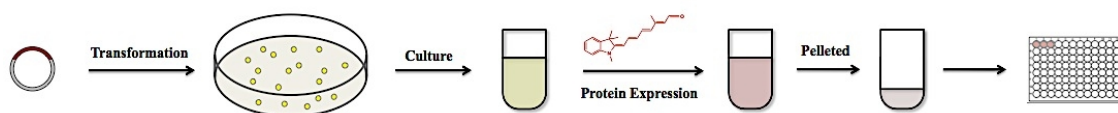
Figure 2.17 Schematic presentation of screening procedure.

2.4.5 K_d Measurements for ARCH, QuasAr2, PL4-98 and PL8-41

In order to measure the K_d values and plotting the saturation curve for ARCH and the best protein variants (QuasAr2, PL4-98, and PL8-41) the same protocol was followed, so that twenty different concentrations of **S4** (0.2 μM to 200 μM) added to 4 mL of LB cultures of rhodopsin expressing *E. coli*. Emission spectra were obtained at excitation wavelength of 710 nm and all saturation binding curves were fitted using michaelis-menten equation (Originlab 9.1).

2.4.6 pH Titration of Reconstituted of ARCH and PROPS Variants

One colony was picked from each of the eight palates into 4 mL LB- media supplemented with 100 μg/mL ampicillin and incubated for overnight at 37 °C. The subcultured media was transformed to 50 mL LB- media including ampicillin and incubated for 3 more hours at 37 °C. The proteins were expressed by adding 0.2% arabinose at the same time with adding **S4** retinal and incubated for 3.5 hours at the same temperature. The solutions were centrifuged for 6 minutes at 8000 rpm, resuspended with 25 mL PBS and centrifuged for 6 minutes again. The pellets were resuspended with 6 mL tris- buffered Saline (TBS). The suspensions were sonicated for 5 minutes to break the cells. The resulting solutions were divided into 10 tubes (500 μL in each) and centrifuged. The liquids were discarded and the pellets were resuspended in 200 μL of 1% beta-octylthioglucoside (OTG) and vortexed for 2 minutes. They were centrifuged again and this time the liquids were collected and the emission and excitation spectra were obtained by SafireII™ plate reader.

2.4.7 Extinction Coefficient Calculation

The absorbance of **S₄** and all-*trans*-retinal solutions in DMSO in five different concentrations were measured (at 489 nm and 375 nm, respectively) with Beckman Coulter DU 800 UV/Vis spectrometer and plotted separately versus the concentrations in Molar unit. The slopes of the obtained lines are presenting the extinction coefficient value.

2.4.8 Quantum Yield Calculation

The absorbance and fluorescence emission of **S₄** and fluorescein solutions in 0.1 M NaOH at five different concentrations were obtained. The concentration of solutions where chosen in such a way to have the absorbance of lower than 0.1. Then, the integrated fluorescence emission intensity was plotted versus absorbance. The quantum yield values were obtained according the following formula:

$$\phi_x = \phi_{ST} \left[\frac{SL_x}{SL_{ST}} \right] \left[\frac{\eta_x^2}{\eta_{ST}^2} \right]$$

Where the subscripts ST and X stand for standard and test, respectively. ϕ is the quantum yield, *SL* is the slope from the plot, and η is refractive index of the solvent.

Chapter 3: Attempted Insertion of a Synthetic GFP Chromophore into an Empty FP Barrel

3.1 Introduction

Studies of the GFP chromophore have shown that it exhibits only weak fluorescence when it is outside the barrel structure of the GFP. There are two suggested reasons for this phenomenon. First, the GFP chromophore is a planar molecule because of the conjugated π system. Second, the GFP chromophore is protected from quenching by water and solutes, and it is also restricted to a *Z*-form inside the protein barrel. **(Figure 3.1)** [29, 100-102]

Z (Bright form)   **E** (Dark form)
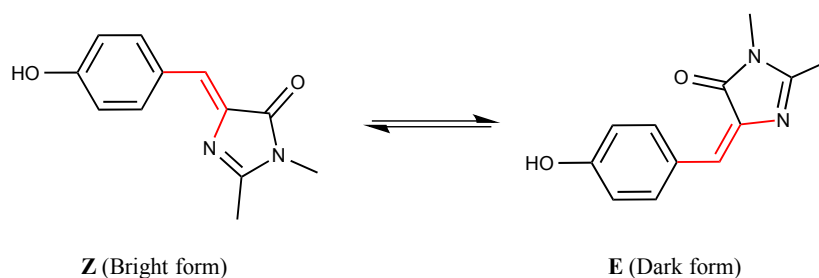
Figure 3.1 Isomerization of GFP chromophore structure. The *Z*-form is fluorescent while the *E*-form is not.

In addition to mentioned reasons, Dong and co-workers hypothesized that the intermolecular hydrogen bonding may lead to quenching GFP chromophore as well. Although this is in contrast with the fact that the hydrogen bond between the chromophore and the amino acids of the protein results the protein to fluoresce.

The objective of the second project of my thesis was to introduce a synthesized GFP chromophore into cp-EGFP or cp-mCitrine (YFP) barrels that have their

chromophore removed. If successful, this approach could provide a new method for labeling of intracellular proteins using unnatural GFP-type chromophores with improved properties. For example, one could imagine introducing synthetic chromophores with further red shifted fluorescence or high photostability. The empty protein barrels are engineered by circular permutation at position 65 and the three amino acids responsible for chromophore formation are deleted **(Figure 3.2)**. If the synthetic GFP chromophore successfully inserted into the barrel, we would expect to see an increase in fluorescence.
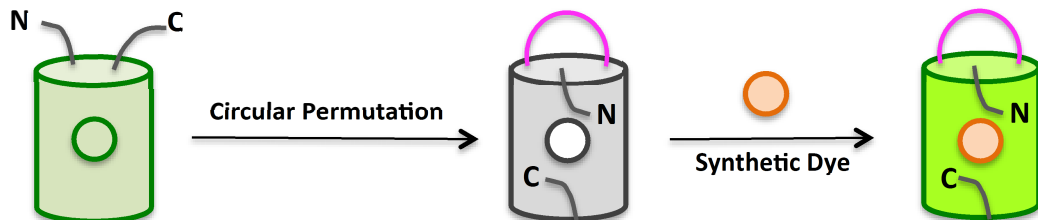
Figure 3.2 Introducing the synthetic chromophore inside the circularly permuted protein barrels without chromophore.

In this Chapter, I describe the synthesis of the **HBDI** chromophore as well as expressing of the empty cp-EGFP and cp-mCitrine followed by our efforts to insert the chromophore inside of these barrels. This includes engineering the libraries of slightly different FP barrels that could potentially better accommodate the chromophore and shortening of the central alpha helix in an effort to make room for the synthetic chromophore to insert in the barrel.

3.2 Results and Discussion

3.2.1 Synthesis of 4-Hydroxybenzylidene-1,2-Dimethyl-Imidazolinone (HBDI) GFP Chromophore

To start with, I synthesized the simplest form of the GFP chromophore; **HDBI**. This molecule was synthesized via two steps starting with Erlenmeyer azlactone synthesis. In this reaction, *N*-acetyl glycine **I** was reacted with 4-hydroxybenzaldehyde and sodium acetate in acetic anhydride to form the oxazolinone moiety **II (Figure 3.3)**.
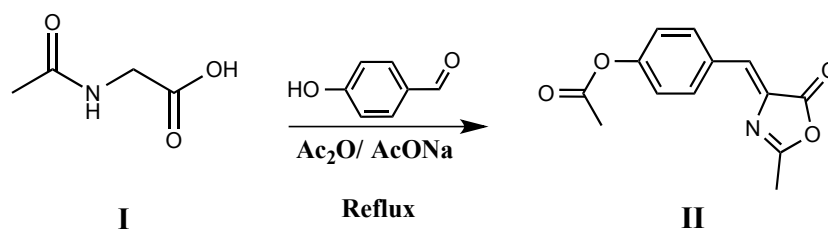
Figure 3.3 The first step of **HBDI** synthetic route.

In this reaction, *N*-acetyl glycine **I** first undergoes a cyclization reaction to generate azlactone **Ia** *in situ* following by dehydration. Next, an aldol reaction is carried out by formation of enolate anion and the nucleophilic addition of the enolate to 4-hydroxybenzaldehyde. Finally, another dehydration and hydroxyl protection produces the oxazolinone **II (Figure 3.4)**.

During the second step, nucleophilic ring opening of oxazolinone **II** with methylamine in presence of base (K$_2$CO$_3$) gave **HBDI** with molecular weight of 216.1 g/mol **(Figure 3.5)**.
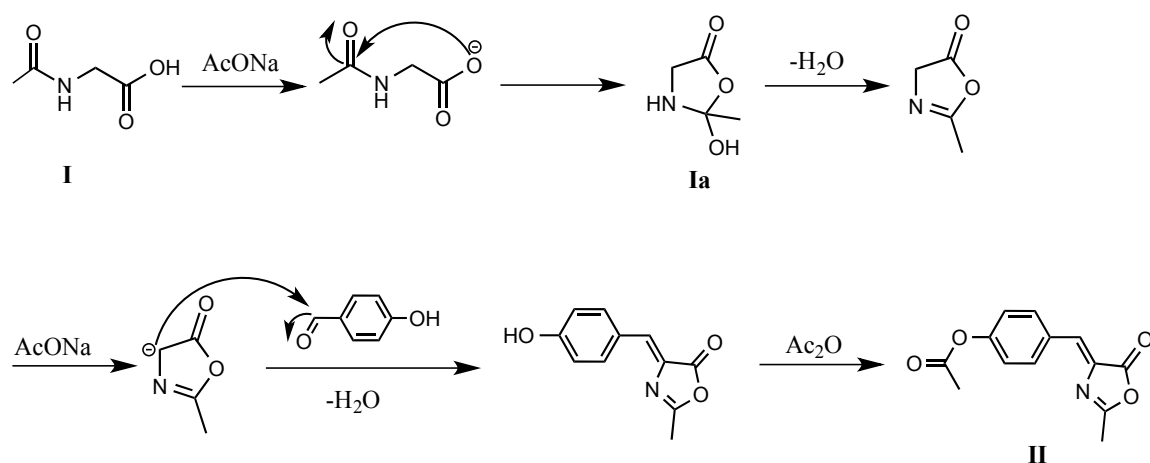
Figure 3.4 Mechanism of the first step of **HBDI** synthetic pathway.
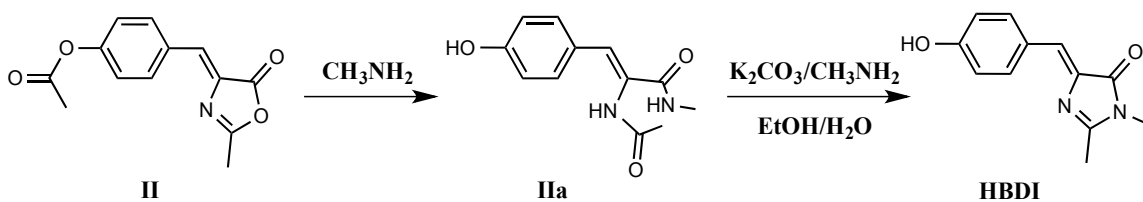
Figure 3.5 The second step of **HBDI** synthetic pathway.

3.2.2 Spectral Characterization of HBDI

After confirming the **HBDI** structure with standard characterization methods, including mass spectrometry (Appendix A), the optical properties of **HBDI** were obtained in water. **HBDI** has an absorbance maximum at 368 nm in acidic pH and 425 nm in basic pH. It also has the emission maxima at 415 nm (excited at 368 nm) and 490 nm (excited at 425 nm). **Figure 3.6** shows the obtained pH-dependence absorbance and emission spectra of **HBDI** excited at 368 nm and 425 nm.
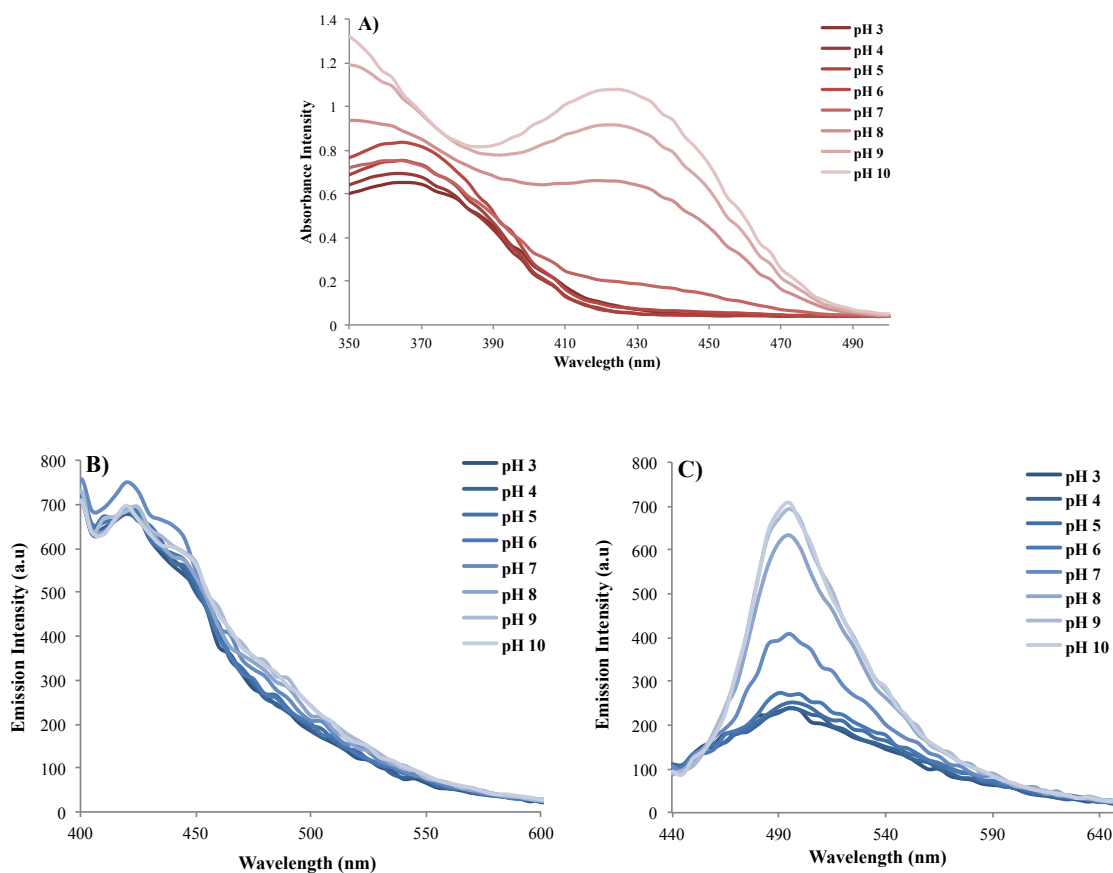
Figure 3.6 HBDI optical spectra measured in water; A) pH-dependence absorbance spectrum, B) pH dependence emission spectrum, excited at 368 nm, C) pH dependence emission spectrum, excited at 425 nm.

As mentioned in the introduction of this Chapter, the GFP chromophore weakly fluoresces when it is outside of the protein barrel because of its conformational freedom and quenching by water and therefore, the intensity of the emission peaks of **HBDI** are very low.

These results are consistent with previous literature reports for the same compound [103]. The extinction coefficient and pKa values have also been reported and are shown in **Table 3.1**.

Table 3.1 p*K*a and ε for **HBDI**[103].

pKa		ε (10^3 M^{-1}cm^{-1}) in water		
Anion/ Neutral	Neutral/Cation	Anion at 425 nm	Neutral at 370 nm	Cation at 393 nm
7.8	2.7	32.4	23.9	24.8

3.2.3 Development of cp-EGFP and cp-mCitrine without the Chromophore

Circularly permuted (cp) proteins can be generated in the research lab using standard molecular biology techniques. By making circularly permuted proteins, the N- and C-termini are moved to new locations and a genetically encoded linker connects the old N- and C-termini. We can choose the location of these new termini by designing the right primers to split the gene into two parts and then linking the old N- and C-termini with a linker.

In this project, we chose to make cp-EGFP and cp-mCitrine and remove the chromophores at the same time. (My colleague Ahmed Abdelfattah engineered the empty cp-EGFP and I performed the same procedure to obtain empty cp-mCitrine). In the first step, we designed two primers (2 and 3 in **Figure 3.7**) to copy the gene via two PCR reactions in such a way as produce two DNA fragments with no DNA coding for the chromophore (TYG residues for EGFP and GYG residues for mCitrine). We also designed the other two primers (1 and 4 in **Figure 3.7**) to pair with primers 2 and 3, respectively. Primers 1 and 4 include DNA for 6 histidines (called a His-Tag) to act as a linker and connect the original N- and C-termini (showed by pink color in **Figure 3.7**).

In this second step **(Figure 3.7B)** two split parts are overlapped via the complementary His-Tag DNA sequence and then primers (2 and 3) are used to copy the gene to produce the cp- EGFP and cp-mCitrine without any chromophore.
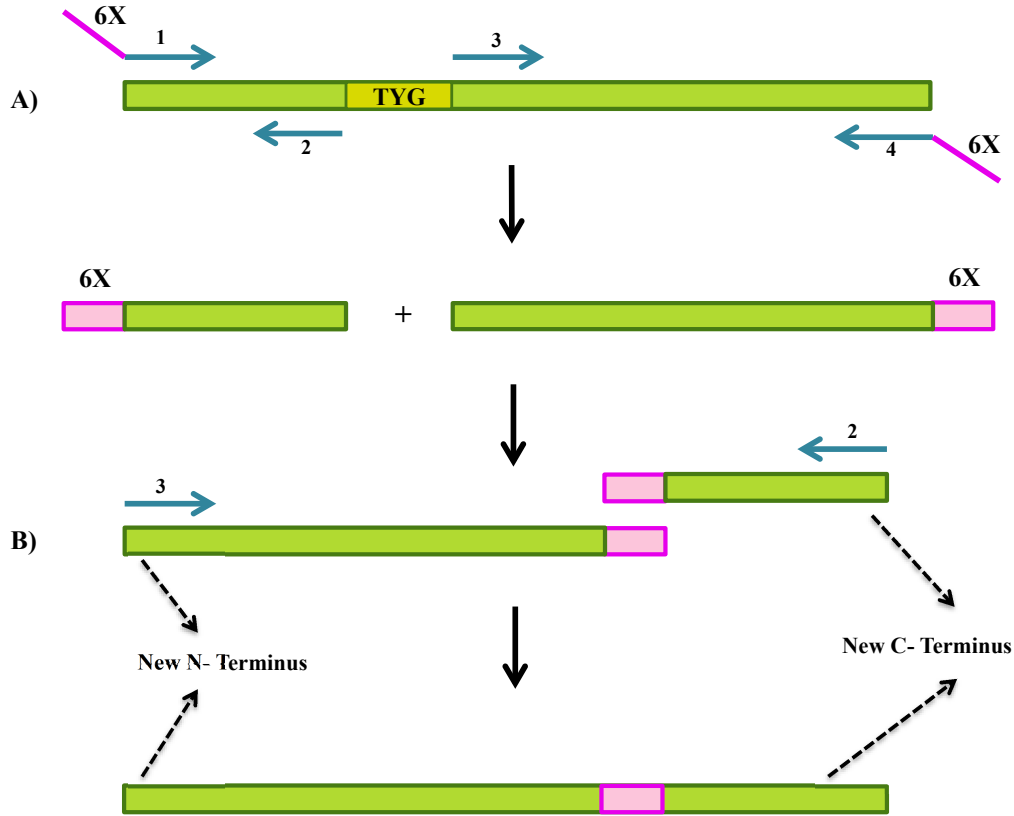
Figure 3.7 Schematic description of circular permutation of EGFP and mCitrine and removing the chromophore. The primers are shown with blue and His-Tag linkers with pink. The EGFP chromophore is indicated by yellow.

Next, the obtained genes were digested with appropriate restriction enzymes and were ligated to the pBAD vector using T4 enzyme and used to transform into *E. coli* cells by electroporation.

After confirming the gene sequences by DNA sequencing, both cp-EGFP and cp-mCitrine proteins were expressed in E. coli in LB-media. Purification of the proteins provided the "non-FPs" with no chromophore and new N- and C-termini the

chromophore was originally located. Their molecular weights (~28 kDa) were confirmed by SDS-polyacrylamide gel electrophoresis (PAGE) **(Figure 3.8)**. The protein concentrations were obtained by BCA protein assay kit.
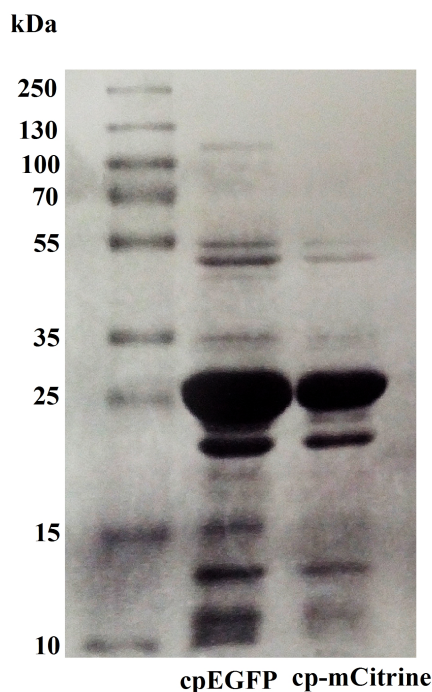
Figure 3.8 SDS-PAGE analysis shows cp-EGFP and cp-mCitrine with Mw ≈ 27 kDa.

3.2.4 Attempting to Introduce HBDI into the Empty FP Barrel

3.2.4.1 Mixing HBDI with Proteins

In the first attempt to try inserting **HBDI** into the empty barrels, eight different concentrations of **HBDI** were mixed with equal amount of proteins in TBS buffer at the pH 7 and 8. The minimum and maximum concentrations of **HBDI** in solutions were 10 μM and 100 μM, respectively. The emission spectra were obtained with the excitation wavelengths of 368 nm, 425 nm, and 460 nm. We were hoping to mimic wild-type GFP's

fluorescence, which has an emission peak at 520 nm and is excited with both 400 nm and 480 nm wavelengths. If the 520 nm fluorescence emission increased upon binding of our chromophore, we would have seen it in one of our scans mentioned above. But no specific changes, at 520 nm nor at any of other scanned wavelength were observed compared to the **HBDI** and empty FP barrels **(Figure 3.9)**.
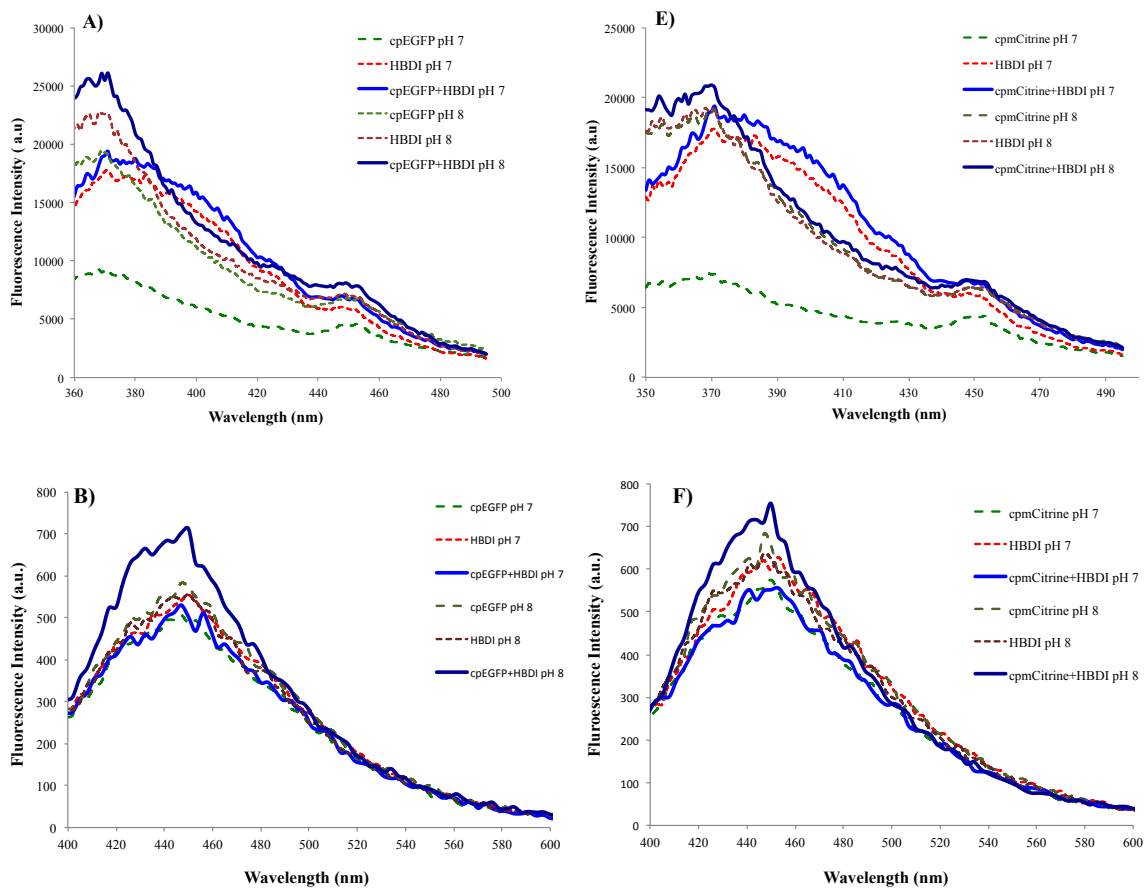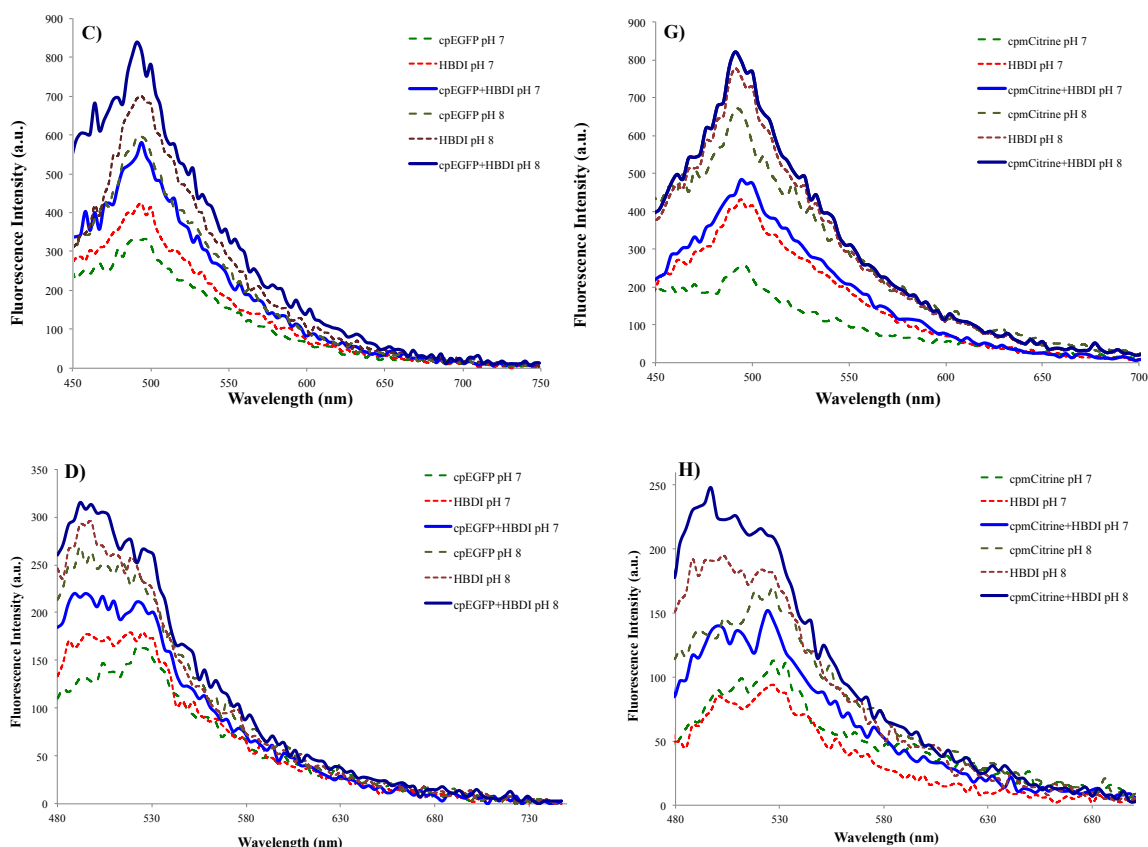
Figure 3.9 Excitation [emission wavelength= 530 nm (A & E)] and emission spectra [excited at 368 nm (B & F), 425 nm (C & G), and 460 nm (D & H)] obtained from introducing **HBDI** to cp-EGFP (A to D) and cp-mCitrine (F to H). The final concentration of **HBDI** is 100 μM.

The samples were incubated at room temperature for 24 hours and the emission spectra were measured. Yet no increases in fluorescence were observed.

The whole experiment was repeated and this time, the samples were incubated at 37 °C for 2 hours after mixing the proteins with fluorophore and the emission spectra were obtained. The results were same as yet again no increase in fluorescence was observed.

3.2.4.2 Denaturing the Protein by Boiling and Mixing with HBDI

In an attempt to refold the protein in the presence of HBDI, cp-EGFP and cp-mCitrine were boiled until they unfolded and began to precipitate. Then the same solutions with equal concentrations as mentioned in **3.2.4.1** were made and the emission spectra were obtained during the same condition and similar results (as shown in **Figure 3.9)** were observed.

3.2.4.3 Denaturing the Protein Using Sodium Cholate Surfactant

Surfactants are partly hydrophilic and partly hydrophobic compounds that can be used as protein denaturing agent. They can disrupt the membranes and break the intermolecular and intramolecular protein-protein interactions since the proteins are also held in shape partially through hydrophobic and hydrophilic interactions. Some types of surfactants therefore attack the proteins and can unfold them, whereas others tend not to cause unfolding of proteins.
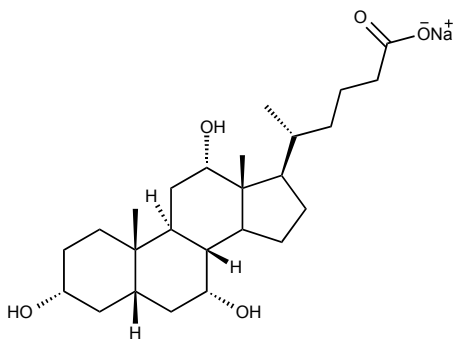
Figure 3.10 The schematic chemical structure of sodium cholate consisting of a steroid backbone, three hydroxyl and one carboxylate group.

Sodium cholate is a water-soluble surfactant (known as a detergent) that consists of a rigid steroid backbone with hydroxyl groups and a carboxylate group (**Figure 3.10**). It is commonly used for proteins and lipid isolation and protein denaturation [103, 104].

We attempted to unfold the proteins by using sodium cholate as the denaturing agent and then, after adding the **HBDI** chromophore, refold the protein. In order to refold the denatured protein, the denaturation agent must be removed and for this aim, detergent absorbing biobeads were utilized [105].

Proteins were mixed with sodium cholate and stirred for 2 hours. Then three different concentrations of **HBDI** in TBS at pH 7 and 8 were added to the denatured proteins and incubated for 1 hour at room temperature. Finally, sodium cholate was removed by using the biobeads and emission spectra were obtained. The results indicated that this method did not lead to insertion of **HBDI** into the protein barrels, either.

3.2.4.4 Addition of HBDI during Protein Expression

As an alternative method, we decided to add **HBDI** to the culture of transformed *E. coli* cells at the same time as adding the inducer (arabinose). By this method, the proteins would have access to the chromophore when they are being expressed in the cell. Therefore, the cells expressing cp-EGFP and cp-mCitrine genes were grown in LB-media containing antibiotic overnight. **HBDI** (final concentration of 100 μM) and arabinose were then added and the cells were induced for 8 hours to allow for protein expression and folding. The cells were pelleted and compared with the control pellets (that had no chromophore added) in terms of fluorescence. Unfortunately, no specific changes in fluorescence emission of the expressed proteins were observed.

3.2.4.5 Randomizing the First (N-terminus) and Last (C-terminus) Amino Acids

Since the N- and C-termini of cp-EGFP and cp-mCitrine proteins are located flanking the now missing chromophore **(Figure 3.7)**, this brought to mind the idea to randomize the amino acids at these positions. These changes could potentially change the pocket size around the deleted chromophore and increase the affinity of the synthetic chromophore for binding the pocket. Therefore, we decided to randomize the amino acids either at the N-terminus, C-terminus or N- and C-termini together, and screen the resulting libraries for proteins that could bind **HBDI**.

After ligating the randomized genes into the plasmid we explored two different methods to add **HBDI** to the cells: 1) transformation and plating the cells on agar plate and spraying the synthetic chromophore on them; and 2) Adding **HBDI** directly to the agar plates. Then, the brightest colonies were picked and grown in LB containing inducer and the proteins were expressed, purified and the fluorescence was checked **(Figure 3.11)**.

We screened colonies on the agar plate and tried picking the brighter colonies. Unfortunately, although some colonies did appear slightly brighter than others, the picked colonies did not show any changes in fluorescence when we expressed and purified the protein for testing with the chromophore. Retesting these picked colonies in the context of bacterial colonies also did not indicate any increase in fluorescence. **Table 3.2** indicates the mutations obtained for picked variants. Among the picked colonies, there was no colony that had the mutations at both termini.
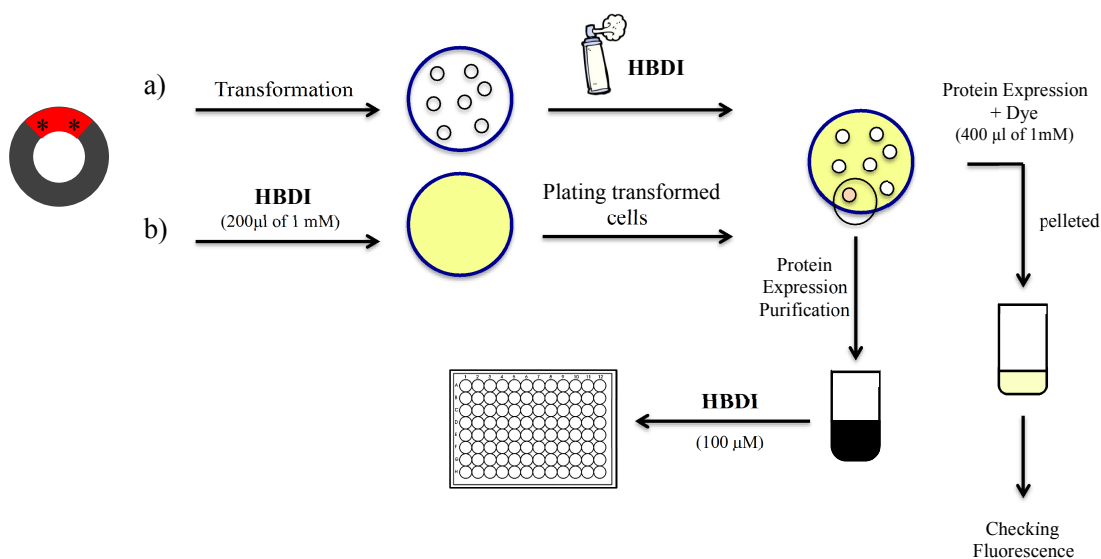
Figure 3.11 Schematic procedures of trying to insert **HBDI** in randomized proteins.

Table 3.2 The obtained mutations and N- and C-termini of cp-EGFP and cp-mCitrine**.**

	N-terminus	C-terminus	N-, C-Terminus
cp-EGFP	V1G	L250A, L250H	-
Cp-mCitrine	L1G, L1T	F250I, F250 T	-

3.2.4.6 Error Prone PCR

As yet another approach to try to evolve FP barrels that would accommodate the synthetic chromophore, error prone PCR reaction was performed on cp-EGFP and cp-mCitrine genes using a low fidelity or error prone polymerase to insert random mutations into the whole gene and create a library for screening. The brighter colonies were picked and examined using the same methods employed in the previous section. However, no variants showed a change in fluorescence when the dye was added to purified proteins.

3.2.4.7 H148G Mutation

As we can see in the 3D structure of GFP and its variants, histidine 148 is located in the barrel close to the hole where we expected the synthetic chromophore to enter **(Figure 3.12)**. This brought to mind the idea to increase the size of the entrance by mutating His148 to glycine. The mutation at this position has previously been used to increase Cl$^-$ entry into the FP barrels [106]. Due to the circular permutation, this location can also be numbered as His80.
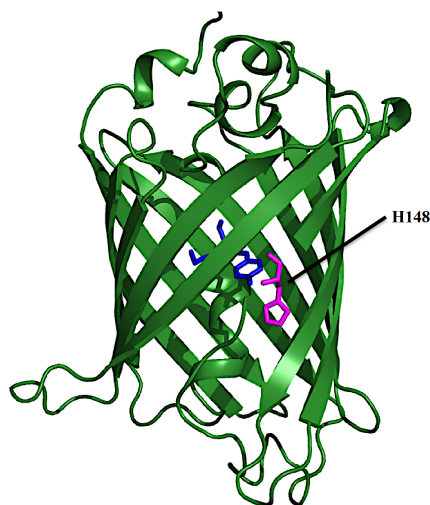
Figure 3.12 3D structure of GFP (PDB ID: 1GFL, [10]) showing the location of His148.

The H80G mutation (H148G in GFP) was introduced by a Quickchange reaction on cp-EGFP and cp-mCitrine and the mutation was confirmed by sequencing the obtained genes. Inserting **HBDI** into the mutated barrels was attempted using the procedure described in **3.2.4.5** but unfortunately, no change in fluorescence was observed **(Figure 3.13)**.
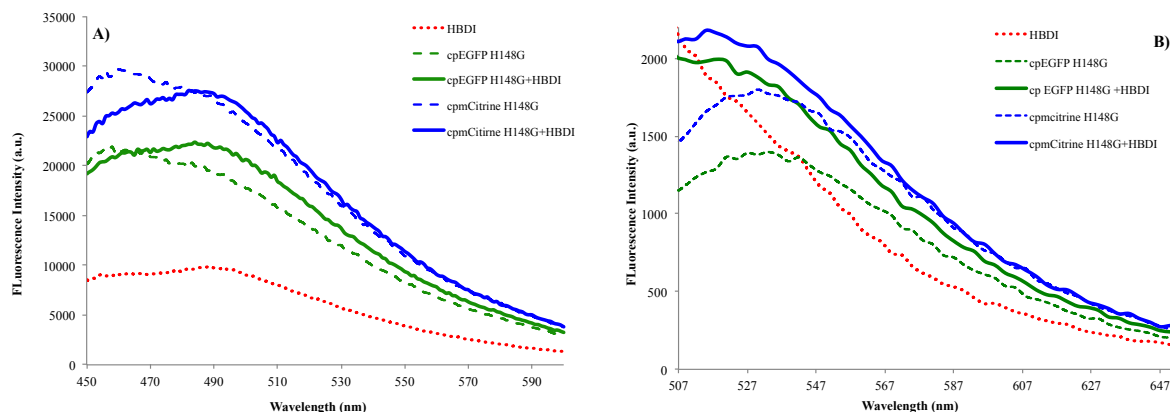
Figure 3.13 The fluorescence emission spectra of H148G mutants Excited at A) 400 nm, B) 460 nm. The final concentration of **HBDI** is 100 μM.

3.2.4.8 Truncation of the Helix

By circular permutation of the FPs and removal of the chromophore, the alpha helix inside the β-barrel had effectively been cut into two portions. To insert **HBDI** into the barrel we decided to remove these portions either partially (truncation from N- or C-terminus) or completely (truncation from both termini). Therefore, there would be a large pathway available for the synthetic chromophore to enter the barrel from either end, and hopefully fluoresce **(Figure 3.14)**.

The truncation was performed by removing eight amino acids from the N-terminus and/or eleven amino acids from C-terminus. Accordingly, a total of 6 genes were constructed corresponding to N-, C-and N- and C-terminal deletions for both cp-EGFP and cp-mCitrine. Sequencing was used to confirm the correct constructs had been built. Next, we tried to express these proteins and insert **HBDI** into them.

71
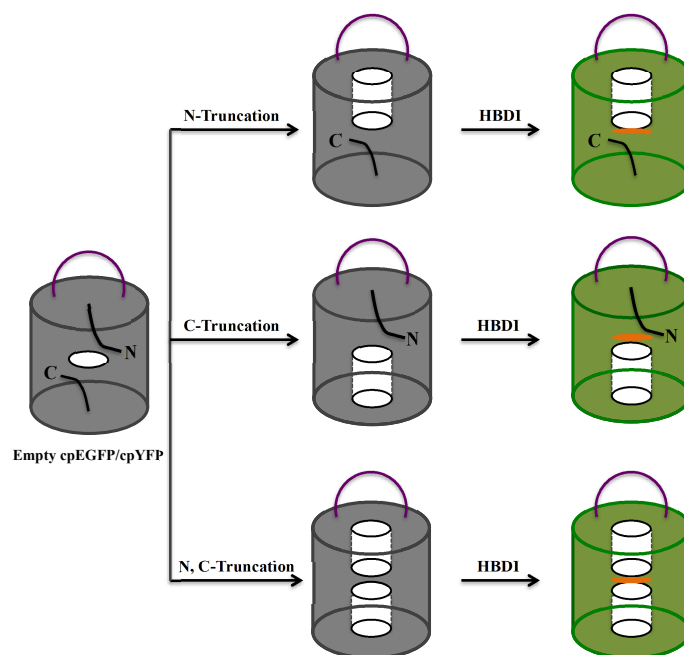
Figure 3.14 The schematic procedure of truncation of FP helix. The amino acids are truncated either from one of the N- or C-terminus or from both termini.
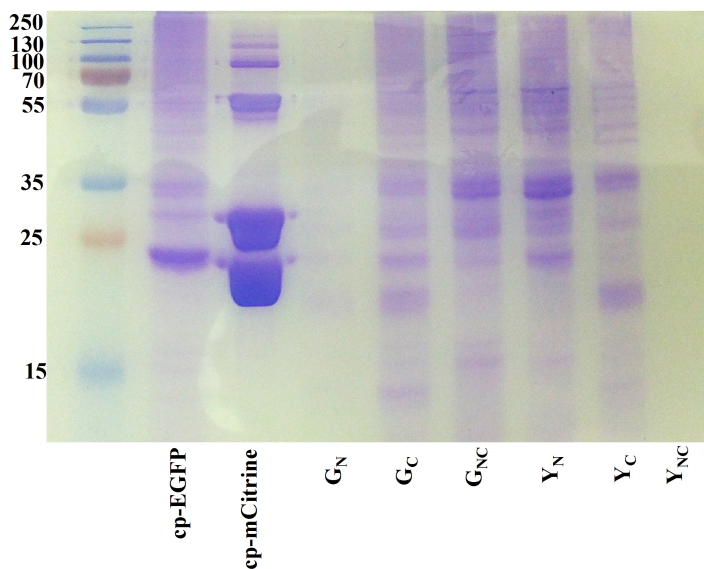
Figure 3.15 The SDS-PAGE gel of cp-EGFP and cp-mCitrine as the controls and six truncated variants indicating that truncated variants are not expressed in bacteria.

The SDS-PAGE gel that was run for cp-EGFP, cp-mCitrine (as the controls) and truncated variants **(Figure 3.15)** indicates that the truncated variants did not express or fold well in *E. coli*. Therefore, we can draw the conclusion that the presence of the helix is necessary for these proteins to express and fold and even one portion of this helix is not enough for expression. An important conclusion to draw from these results is that the protein is likely folding into the expected conformation, with the two fragments of the split central helix inserted into either end of the barrel.

3.2.4 Attempting to Introduce HBDI into the Photocleavable Protein Barrel

Recently, my colleague Wei Zhang developed a photocleavable protein (called PhoCle) from cp-mMaple that spontaneously and irreversibly dissociates into two fragments. Since PhoCle is an engineered homologue of GFP, it has all the advantages of other FPs including low cytotoxicity, robust expression in most of tissues and organisms and tolerance to genetic modifications and circularly permutations.

By circular permutation of the green form of mMaple, the new termini are located close to the chromophore. Violet light illumination interestingly leads to an elimination reaction that cleaves the main chain of the polypeptide inside the barrel. The elimination reaction forms a double bond that results in elongation of the conjugated system of the chromophore to create a new red fluorophore **(Figure 3.16)**.
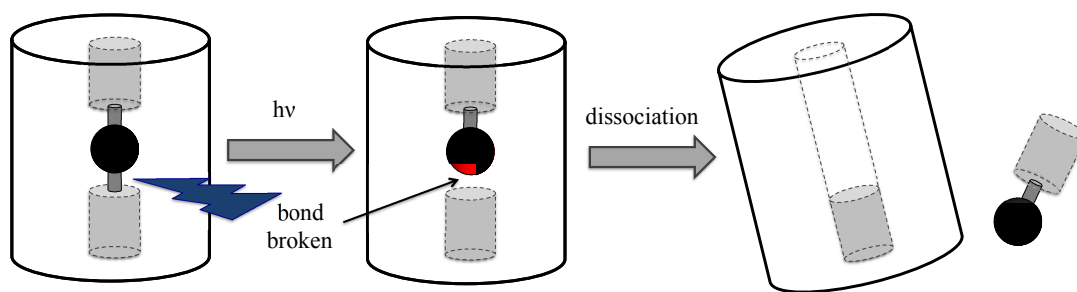
Figure 3.16 Schematic presentation of PhoCle cleavage and dissociation.
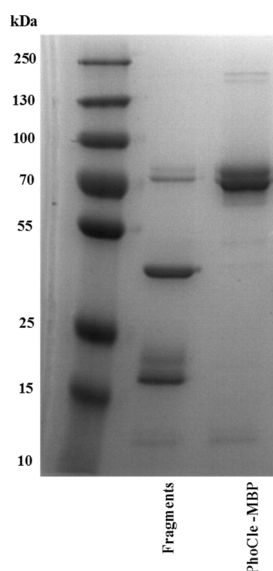
Figure 3.17 The SDS-PAGE analysis of PhoCle obtained by Wei Zhang, showing the PhoCle fused to MBP at ~65 kDa, and the fragments ~45 kDa and ~20 kDa.

Figure 3.17 is showing the SDS-PAGE analysis of eluting fractions indicating the non-cleaved PhoCle fused to maltose binding protein (MBP) at ~65 kDa, and fragments that obtained after cleavage at ~45 kDa and ~20 kDa.

Because our lab had available the cleaved PhoCle fragment that has ejected its chromophore-containing peptide, we decided to attempt to introduce the synthetic

chromophore **HBDI** into the PhoCle barrel and observe the changes in fluorescence **(Figure 3.18 A)**. Although, it is not yet been demonstrated whether this barrel is still open after the chromophore–containing peptide has been ejected or if it has collapsed because of the hydrophobicity of the inside of the barrel.

The PhoCle protein was expressed and cleaved by shining the light on it for 20 minutes and the two cleaved fragments were separated and purified. Then, **HBDI** was added to the cleaved fragment (that had ejected the chromophore-containing peptide) and the fluorescence spectrum was obtained. Unfortunately, we did not observe any specific increase in fluorescence of this protein **(Figure 3.18 B)**.
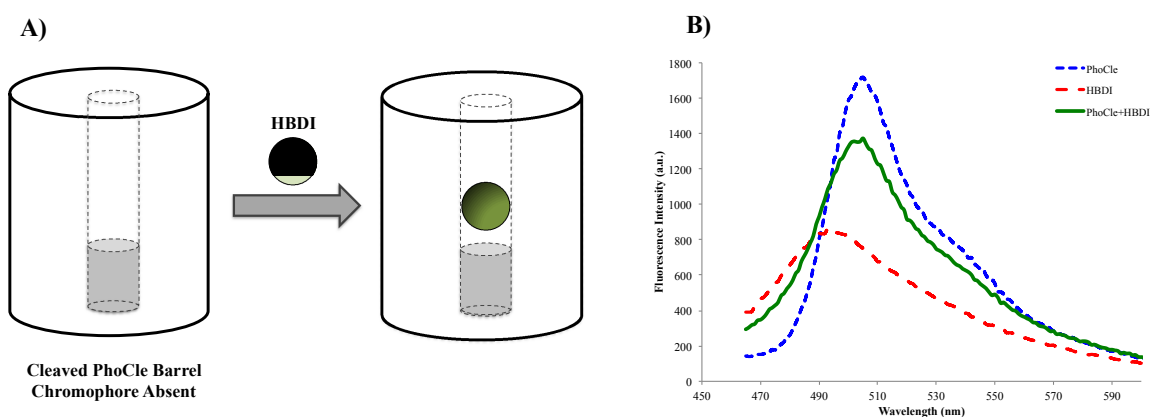
Figure 3.18 A) Schematic procedure of introduction of **HBDI** to cleaved PhoCle fragments. B) The fluorescence spectrum obtained by adding **HBDI** to PhoCle does not indicated any increase in fluorescence of the protein. The final concentration of **HBDI** is 100 μM.

The measurements were repeated after 2 hours and 24 hours incubation the mixture of protein and **HBDI** at 37 °C and the same results were obtained.

In this experiment, we do not expect to see any fluorescence for cleaved PhoCle fragments since the chromophore-containing peptide has already been rejected. However,

as we can see in the SDS-PAGE gel **(Figure 3.17)**, the cleavage of PhoCle is not complete, as a small portion of this protein still possesses the green chromophore and so the observed green fluorescence that arises from it.

According to the obtained results, addition of **HBDI** leads to a decrease in the fluorescence intensity. Since the concentration of **HBDI** that has been added is quite high, we suspect that this loss of fluorescence is due to an inner filter effect. Essentially, **HBDI** is absorbing some of the excitation light, and also some of the emitted photons, and thus decreasing the observed fluorescence.

3.3 Conclusion

In this Chapter, we described the synthetic procedure for producing the GFP chromophore **HBDI**, and the development of the circular permutations of EGFP and mCitrine with no chromophore. We then describe our efforts to insert **HBDI** into the empty protein barrels. We have demonstrated that it is challenging or even impossible to introduce a synthetic GFP chromophore into the cp GFP protein that lacks the chromophore. We attempted this experiment using a wide variety of approaches, and yet did not observe any increase in the fluorescence intensity. We even attempted to modify the protein by changing the amino acids at new N- and C-termini, yet this did not help, either. Unfolding and refolding of the protein barrel and H148G mutation were also examined as potential methods for insertion of **HBDI**, but these approaches did not cause any increase in fluorescence. We also showed that cp-EGFP and cp-mCitrine are not able to express or fold after the truncation of a few amino acids from either of the termini. Finally, we attempted to employ the cleaved PhoCle fragment in our project as the

protein barrel to insert **HBDI** into it. However, once again we did not observe any increase in fluorescence like the previous efforts.

All in all, we could not achieve any improvements in fluorescence emission by adding the synthetic chromophore to empty FP protein barrels. Since during these attempts no specific changes either in the fluorescence intensity or wavelength were observed, it is highly likely that the chromophore is not able to insert the empty barrel. However, this does not allow us to draw this conclusion with a high degree of certainty. The other possibility is that **HBDI** is able to go inside but the binding is not strong enough to hold it or it stays inside but the quantum yield is below our threshold of detection. Even if the chromophore binds inside the barrel, the solvent should also be able to enter and consequently, quench the fluorescence.

3.4 Materials and Methods

3.4.1 General Methods and Materials

All the organic compounds and solvents were purchased from Aldrich. 1D proton nuclear magnetic resonance spectra were obtained by at 300 MHz, 400 MHz and 500 MHz NMR field strengths, coupling constants (J) are reported in hertz (Hz). ^{13}C NMR spectra were recorded at 100 or 126 MHz. The chemical shifts are reported on the δ scale (ppm) downfield from tetramethylsilane (0 ppm). The molecular weight of synthetic products was determined by mass spectra obtained by a high-resolution electrospray spectrometer in positive ion mode.

The primers and synthetic DNA oligonucleotides were bought from Integrated DNA Technologies (IDT Inc.). Deoxyribonucleotide triphosphate (dNTP) solutions used in PCRs were purchased from Invitrogen. OligoAnalyzer (IDT website) was used for melting temperature (Tm) calculation and prediction of properties of designed polynucleotides. The restriction enzymes and *Pfu* DNA polymerase were obtained from Fermentas/Thermo Scientific and T4 DNA was obtained from Invitrogen/Life Technologies. Polymerase chain reactions (PCRs) were performed using either Taq DNA polymerase (Invitrogen Inc.) or *Pfu* DNA polymerase (Fermentas Inc). The PCR and restriction digestion products were purified by DNA agarose gel electrophoresis performed in Tris-acetate-EDTA (TAE) buffer and extracted by GeneJET gel extraction kit or QIAquick gel extraction kit. Ethidium bromide (Bio-Rad Inc.) was added to agarose gel as the fluorescent dye label for visualizing DNA bands by UV lamp.

Electrocomponent *E. coli* stain ElectroMAX DH10B™ (Invitrogen Inc.) was used for plasmid transformation. Electroporation was done by using MicroPulser cuvettes (Bio-Rad Inc.) and the MicroPulser electroporator (Bio-Rad Inc.). The plasmid DNA was isolated and purified by GeneJET Plasmid Miniprep Kit (Fermentas). Sequencing reactions were analyzed at University of Alberta Molecular Biology Service Units (MBSU) or at University Core DNA Services (UCDNA) at the University of Calgary. Sequencing chromatograms were read by FintchTV 1.4.0 and sequencing alignments were performed by ClustalW. The optical spectra were recorded by using 384-well and 96-well microplate reader (Tecan SafireII™).

3.4.2 Synthetic Procedures and Structural Characterizations

(Z)-4-((2-methyl-5-oxooxazol-4(5H)-ylidene)methyl)phenyl acetate (II)
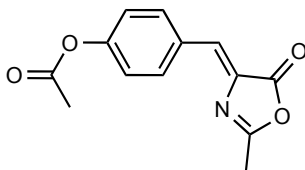
Under N_2 atmosphere, *N*-acetyl glycine **I** (240 mg, 2 mmol) was added to Ac_2O (2.5 ml) and stirred for 10 minutes at room temperature. A white suspension formed. Sodium acetate (100 mg, 1.26 mmol) was added to this solution and stirred for another 25 minutes. Then, 4-hydroxybenzaldehyde (300 mg, 2.5 mmol) was added and refluxed for 4 hours. The white suspension solution turned to a yellow solution. After 4 hours, 10 mL of distilled water and a brown precipitation was formed, immediately. The precipitate was filtered and washed with cold 95% ethanol and recrystallized by ethanol. The obtained product was yellow crystal. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.12 (d, *J*= 2.1), 8.10 (d, *J*= 2.1), 7.19 (d, *J*= 2.1), 7.17 (d, *J*= 2.1), 7.11 (s, 1H), 2.40 (d, *J*= 0.6), 2.32 (s, 1H); MS *m/z* calcd for $C_{13}H_{12}NO_4$ (M+H[+]) 246.0761, found 246.0756.

4-Hydroxybenzylidene-1,2-dimethyl-imidazolinone (HBDI)

(Or (Z)-5-(4-hydroxybenzylidene)-2,3-dimethyl-3,5-dihydro-4H-imidazol-4-one)
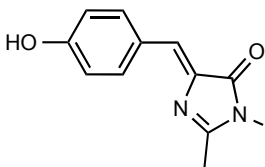
In 4 mL of ethanol, oxazolinone (98.5 mg, 0.4 mmol) was dissolved and K_2CO_3 (85 mg, 0.61 mmol) and methylamine (0.18 ml, 2.05 mmol) were added, respectively. The solution was stirred at room temperature for 10 minutes and then, it was heated at reflux temperature for 4 hours. The yellow solution turned red at once. After 4 hours the reaction was cooled down to room temperature and was filtered and the filtrate was collected. The solution was removed under reduced pressure. However, we were unable to purify the compound by column chromatography and the crude mixture was used for all experiments. MS *m/z* calcd for $C_{12}H_{13}N_2O_2$ (M+H$^+$) 217.0972, found 217.0969.

3.4.3 PCR Amplification, Digestion and Ligation Methods

Standard PCR amplifications carried out using *Pfu* (or Taq) DNA polymerase were performed in volumes of 50 μL containing nuclease-free water, *Pfu* (or Taq) buffer containing $MgSO_4$, 200 μM dNTPs (Invitrogen), forward and reverse primers, template DNA, and 1.0 unit of *Pfu* (or Taq) DNA polymerase. Typical cycling parameters were as follows: initial denaturation at 95 $^{\circ}$C for 3 minutes; 95 $^{\circ}$C for 20 seconds; 20 seconds for annealing temperature (depends on the primers Tm, generally it is 3-5 $^{\circ}$C lower than primers Tm); elongation temperature of 72 $^{\circ}$C as for 60 to 120 seconds (depends on the DNA template length) and final elonngation at 72 $^{\circ}$C for 4 minutes.

The genes of interest were digested using restriction enzymes (2 μL for each), FastDigest restriction endonuclease buffer (5 μL), water and FastDigest restriction endonucleases. The digestion reactions were carried out at 37 $^{\circ}$C for 4 hours.

In order to ligate the digested gene to the plasmid, the ligation reaction was carried out at 37 $^{\circ}$C for 2 hours using T4 DNA ligase (Invitrogen).

3.4.4 Transformation Method

The pBAD/His vector containing the gene encoding the protein of interest was added to *E. coli* strain DH10B competent cells, and the mixture was transferred to the cuvette to transform the DNA into the cells using the electroporator. 1 mL of LB-media with no antibiotics was added to the cuvette and gently mixed with the pipette. All the cells were transferred to a 1.5 mL tube and incubated for 20 minutes at 37 $^{\circ}$C. 1 μL of transformed cell were added to an agar plate and incubated at 37 $^{\circ}$C for overnight.

3.4.5 Protein Purification and Characterization

A single colony was picked into a 4 mL culture that was grown overnight (37 $^{\circ}$C, 240 rpm) and then, diluted into 0.5 L of LB medium supplemented with ampicillin. This culture was grown (37 $^{\circ}$C, 225 rpm) until it reached an OD of 0.6, induced by addition of 0.02% L-arabinose, and cultured for 48 hours at 28 $^{\circ}$C. Cells were pelleted by centrifugation and lysed by cell disruptor (Constant Systems). All proteins were purified by Ni-NTA chromatography using by collecting he beads in plastic Polyprer columns (Bio-Rad) elusion buffer (containing imidazole) and the buffer was exchanged with TBS buffer (pH 7.5). The purified proteins were analyzed by SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis). BCA protein assay kit (Pierce) was used to determine the protein concentrations.

In order to purify the proteins in a fast way, B-PER was used to lyse the pelleted cells so that the pellets were resuspended in B-PER and agitated for 10 minutes. To separate the proteins from cell debris, centrifugation for 15 minutes at 10,000 rpm was

performed. Ni-NTA beads were added to the lysate and separated from the liquid to capture the his-tagged proteins. After 1 hour of shaking at 4 °C, the beads were collected in disposable plastic columns. The other purification steps are same as method mentioned above.

3.4.6 Denaturation/Reconstitution of Proteins using Sodium Cholate

For every 1mmol of protein 2 mmol of sodium cholate (Mw= 430.55 g/mol) is needed [107]. Therefore, 3 mg of sodium cholate was added to 100 μM solutions of cp-EGFP and mCitrine (350 μL) and stirred for 2 hours. Then, the solutions were split to twelve portions, six of them were mixed by TBS (pH 7 and 8) and used as controls whereas other six parts were mixed with TBS (pH 7 and 8) and three different concentrations of **HBDI** solution (50, 500 and 1000 μM), respectively. The biobeads were added to the solutions and stirred for 3 hours. The biobeads were separated by centrifuge and the proteins were collected by pipette.

Chapter 4: Conclusion and Future Directions

4.1 Fluorescent Modulation of Microbial Rhodopsin-Based Biosensors by Retinal Replacement: Conclusion

In Chapter 2, we described the fluorescence modulation of ARCH and PROPS using the S_4 retinal analogue. First, we reported the synthetic pathway to produce S_4 as well as its optical characterizations. This S_4 fluorophore is classified as a merocyanine dye with red-shifted excitation and emission peaks (489 nm and 689 nm, respectively) and 1.7-fold large extinction coefficient compared to all-*trans*-retinal (the natural retinal in microbial rhodopsins). These properties make the S_4 retinal analogue a promising candidate to be used for reconstitution of rhodopsins in the context of colonies in *E. coli* and improvement of rhodopsin-based voltage indicators in terms of brightness and sensitivity.

ARCH and five mutated variants of ARCH and also two PROPS variants, which were developed by my colleague Dr. Yongxin Zhao, were examined for substitution of the chromophore. In this procedure, we expected to encounter one problem with the competition of endogenous retinal for the binding site of our engineered opsin variants. However, addition of the synthetic retinal at our lowest concentration (0.2 μM) showed an improvement in fluorescence.

In this procedure, the engineered variants were expressed in *E. coli* in presence of S_4 retinal and among them, QuasAr2 (ARCH), PL4-98, and PL8-41 (PROPS) indicated significant red-shifted excitation and emission and lower K_ds compared to the other variants.

4.2 Fluorescent Modulation of Microbial Rhodopsin-Based Biosensors by Retinal Replacement: Future Directions

One of the future directions for this project is synthesizing a series of promising red-shifted retinals and using them in the reconstitution of rhodopsins in order to achieve a modulated absorbance and emission maxima, and also brighter fluorescence.

To further improve the brightness of the rhodopsin with the synthetic chromophores, one option is to create libraries by randomizing residues in close proximity to the retinal binding site and more distant residues found to be important during the directed evolution of ARCH and PROPS. In addition, directed evolution by performing error prone PCR of the complete gene and treating them with the retinal analogue may lead to finding the variants with brighter fluorescence.

The most straightforward future application of this research is expression of improved variants in mammalian cells and employing them as indicators of membrane potentials. Since it has been recognized that ARCH and PROPS exhibit weak fluorescence, it limits the ability of researchers to directly image ARCH and PROPS fluorescence using standard microscopy. However, reconstitution of these proteins with improved fluorescence properties could provide them the imaging tools required for enhanced sensitivity.

4.3 Attempting to Insert the Synthetic GFP Chromophore into the FP Barrel: Conclusion

In Chapter 3, the synthetic procedure to produce GFP chromophore was described followed by our attempts to insert this chromophore into the circularly permuted empty FP barrel in order to recover and perhaps improve its fluorescence. We examined several modifications to allow the chromophore to enter the empty barrel such as randomizing the first and last amino acids next to the chromophore pocket, error prone PCR to find favorable mutations for insertion, the H148G mutation to increase the size of the side entrance though the barrel, and also truncation of amino acids from the N- and/or C-terminus to attempt to build an insertion channel for the chromophore. However, we could not detect any improvements in fluorescence or even insertion by these methods. We also used cleaved PhoCle fragment as the empty barrel to insert **HBDI** into it but the same disappointing result was obtained. These negative results however, do not allow us to conclusively state that the chromophore does not enter the protein barrel. There are different possibilities that lead to the obtained results. One possibility is that the chromophore is able to enter the barrel but the quantum yield is too low or solvent quenches the fluorescence. It is also possible that the chromophore can go into the barrel but the binding is weak and it exits again resulting in a below detection threshold population of bound chromophores.

4.4 Attempting to Insert the Synthetic GFP Chromophore into the FP Barrel: Future Directions

There are several directions that can be recommended for this research. First, we can modify the pocket size of the chromophore by deleting one amino acid from each terminus (shortening) so that there will be a larger room for chromophore. We can also substitute other amino acids in the entrance such as Phe165, Thr203, Ser147, and Asn149 to ease the insertion of the chromophore into the barrel.

Truncation experiments with cp-EGFP and cp-mCitrine can also be modified by removing one amino acid at a time. In other words, we can truncate fewer amino acids (compared to ten) and hopefully the protein will be able to express.

An alternative is use other methods to denature the FP barrel like guanidine or urea methods [108]. This might allow the chromophore to enter the unfolded empty barrel, and then refold around it.

Finally, we may synthesize other derivatives of GFP chromophore with different size and functional groups at interaction sites to increase the binding potential with the FP barrel. **Figure 4.1** shows the chemical structure of two derivative examples.
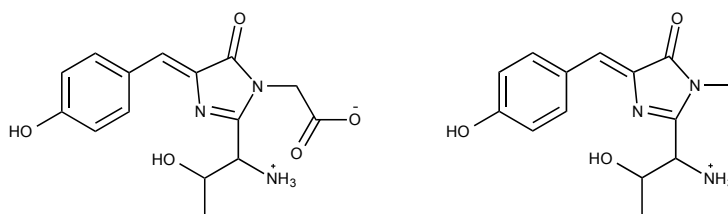
Figure 4.1 The chemical structures of two GFP chromophore derivatives.

Chromophore replacement strategy could help provide insight into the mechanism of the chromophore formation by proteins. It may lead to improved protein-based biosensors and possibly provide new imaging tools with improved fluorescent properties.

Appendix A

Selected NMR and Mass Spectra

NMR peak labels (ppm): 7.390, 7.364, 7.359, 7.332, 7.261, 7.241, 7.238, 7.219, 7.203, 7.199, 7.170, 7.154, 6.956, 6.940, 6.937, 6.919, 6.743, 6.723, 5.724, 5.692, 4.693, 4.667, 3.223, 1.555, 0.000

Expanded region labels: 7.390, 7.364, 7.359, 7.332, 7.261, 7.241, 7.238, 7.222, 7.219, 7.203, 7.199, 7.170, 7.154, 6.956, 6.940, 6.937, 6.919, 6.743, 6.723, 5.724, 5.692, 4.693, 4.667

Integrations (expanded): 0.89, 0.79, 0.92, 0.99, 0.97, 0.98, 6.63

Integrations (full): 0.89, 0.79, 0.92, 0.99, 0.97, 0.98, 1.00, 3.01, 6.63

S1

MS peaks (m/z): 193.1, 225.1, 247.1, 271.2

Counts vs. Mass-to-Charge (m/z)

S1

89

MS Spectrum

Cpd 1: C15H17NO: + Scan (1.9-2.1 min, 11 scans) 12070318.d Subtract

x10 5

| | |
250.1196
(M+Na)+

228.1378
(M+H)+

S2

Counts vs. Mass-to-Charge (m/z)

MS Spectrum Peak List

m/z	Calc m/z	Diff(ppm)	z	Abund	Formula	Ion
228.1378	228.1383	-1.98	1	80079	C15 H18 N O	(M+H)+
250.1196	250.1202	-2.43		252449	C15 H17 N Na O	(M+Na)+

--- End Of Report ---

S3

CN

91
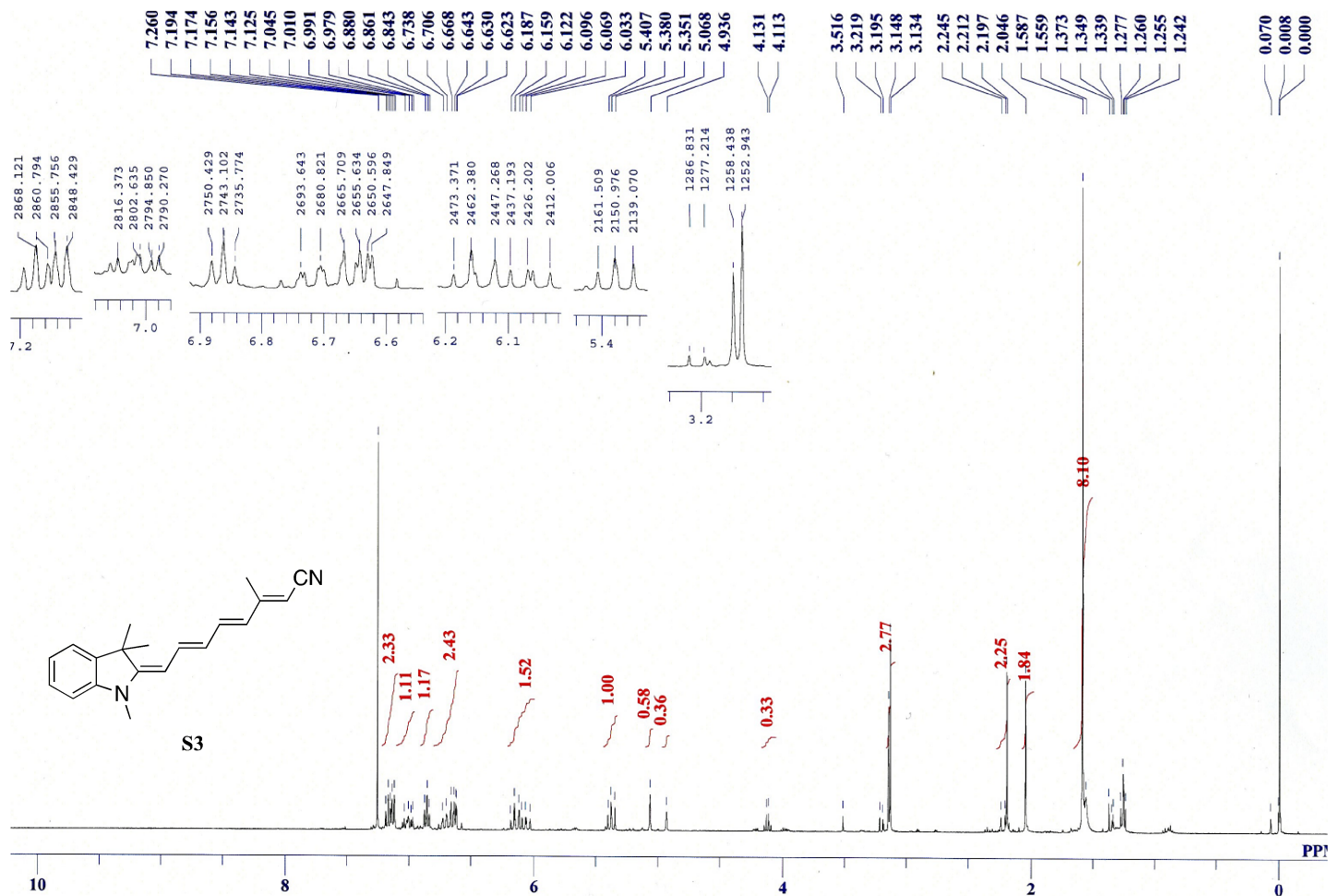
MS Spectrum

x10 4 | Cpd 1: C20H22N2: + Scan (2.5-3.2 min, 44 scans) 12072608.d Subtract

290.1775
M*+

291.1844
(M+H)+

292.1881
(M+H)+

3 -

2.5 -

2 -

1.5 -

1 -

0.5 -

0 -

290   290.5   291   291.5   292   292.5   293

Counts vs. Mass-to-Charge (m/z)
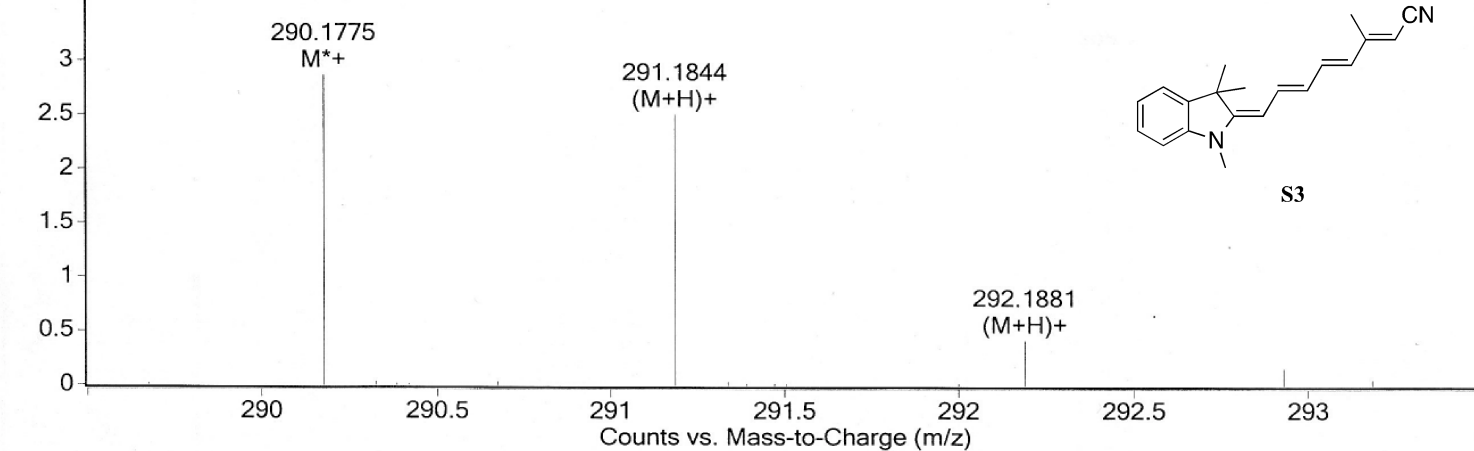
MS Spectrum Peak List

m/z	Calc m/z	Diff(ppm)	z	Abund	Formula	Ion
290.1775	290.1778	-0.86	1	28678	C20 H22 N2	M*+
291.1844	291.1856	-4.2	1	24936	C20 H23 N2	(M+H)+

--- End Of Report ---

S3

S4

10   9   8   7   6   5   4   3   2   1   0

92

190.809
189.818

159.382
155.766
144.833
139.296
138.920
138.455
135.554
135.460
129.648
127.839
127.278
125.760
124.302
124.275
121.639
121.577
120.063
120.018
106.282
106.256
96.692
96.642
77.314
77.059
76.805
45.862
45.834
29.154
28.347
21.257
13.110

S4

200 180 160 140 120 100 80 60 40 20

S4

F1 (ppm)

F2 (ppm)

10 9 8 7 6 5 4 3 2

S4

95

(ref. to CDCl3 @ 7.26/77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe   Pulse Sequence: gHMQC

S4

F1 (ppm)

F2 (ppm)

Fragmentor Voltage 150

Source Type Esi

x10 4

+ Scan (0.67-0.74 min, 5 scans) 14070905.d Subtract

179.0170

294.1851
(M+H)+

Counts vs. Mass-to-Charge (m/z)

Formula Calculator Results

Formula	Ion Species	Mass	Calc. Mass	m/z	Calc. m/z	Diff (mDa)	Diff(ppm)	DBE	Ion	Score
C20 H23 N O	C20 H24 N O	293.1778	293.178	294.1851	294.1852	0.16	0.54	10	(M+H)+	97.25

97

x10 3 | Cpd 1: C22H27NO2: + Scan (1.2-1.7 min, 31 scans) 12112709.d Subtract

337.2037
M*+

2.25
2
1.75
1.5
1.25
1
0.75
0.5
0.25
0

338.2081
M*+

339.2090
M*+

ES3

337.2 337.4 337.6 337.8 338 338.2 338.4 338.6 338.8 339 339.2 339.4
Counts vs. Mass-to-Charge (m/z)

MS Spectrum Peak List

m/z	Calc m/z	Diff(ppm)	z	Abund	Formula	Ion
337.2037	337.2036	0.11	1	2309	C22 H27 N O2	M*+

--- End Of Report ---

200   Esi

x10 5 | + Scan (1.3-1.4 min, 8 scans) 13012423.d Subtract

278.1900

3
2.5
2
1.5
1
0.5
0

149.0337
179.0445
213.0055

296.2012
(M+H)+

S4- OH

110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340
Counts vs. Mass-to-Charge (m/z)

Formula Calculator Element Limits

Element	Min	Max
C	15	25
H	20	30
N	0	2
O	0	2

Formula Calculator Results

Formula	Ion Species	m/z	Calc. m/z	Diff (mDa)	Diff (ppm)	DBE	Ion	Score
C20 H25 N O	C20 H26 N O	296.2012	296.2009	-0.33	-1.12	9	(M+H)+	76.7

98

8.127
8.126
8.121
8.120
8.109
8.109
8.104
8.103

7.260
7.259
7.195
7.190
7.189
7.178
7.173
7.172
7.113
7.113

2.406
2.405
2.322

8.0   7.5   7.0   6.5   6.0   5.5   5.0   4.5   4.0   3.5   3.0   ppm

8.127
8.126
8.121
8.120
8.109
8.109
8.104
8.103

7.260
7.259
7.195
7.190
7.189
7.178
7.173
7.172
7.113
7.113

.6   8.4   8.2   8.0   7.8   7.6   7.4   7.2   7.0   6.

x10 4 | + Scan (1.0-1.3 min, 16 scans) 13041605.d Subtract

246.0756
(M+H)+

268.0576
(M+Na)+

Counts vs. Mass-to-Charge (m/z)
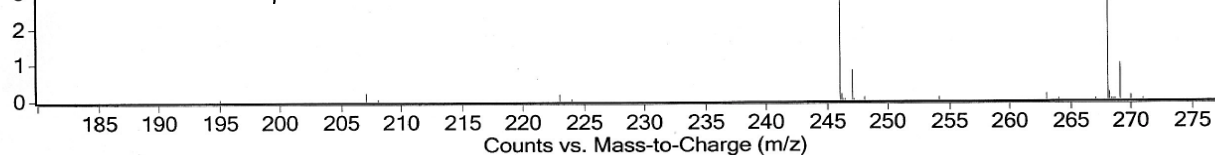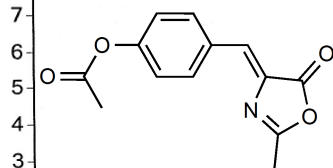
Formula Calculator Element Limits

Element	Min	Max
C	5	20
H	5	20
N	0	2
O	2	6

Formula Calculator Results

Formula	Ion Species	m/z	Calc. m/z	Diff (mDa)	Diff (ppm)	DBE	Ion	Score
C13 H11 N O4	C13 H12 N O4	246.0756	246.0761	0.47	1.91	9	(M+H)+	88.33
C13 H11 N O4	C13 H11 N Na O4	268.0576	268.058	0.43	1.74	9	(M+Na)+	90.76

--- End Of Report ---

MS Spectrum

x10 5 | Cpd 1: C12H12N2O2: - Scan (0.4 min) 13050732.d Subtract

248.0919

215.0822
(M-H)-

233.0928
(M+OH)-

Counts vs. Mass-to-Charge (m/z)
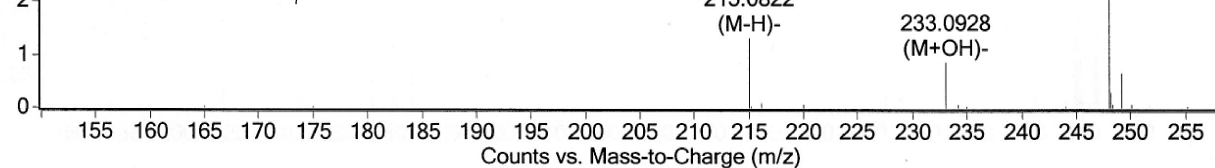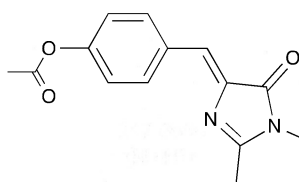
MS Spectrum Peak List

m/z	Calc m/z	Diff(ppm)	Abund	Formula	Ion
215.0822	215.0826	-1.84	132459	C12 H11 N2 O2	(M-H)-
233.0928	233.0932	-1.39	83990	C12 H13 N2 O3	(M+OH)-

--- End Of Report ---

Appendix B

<u>Selected Amino Acid Sequences</u>

Appendix B

Selected Sequences: Chapter 2

ARCH 3	QAGFDLLNDGRPETLWLGIGTLLMLIGTFYFIARGWGVTDKEAREYYAITI LVPGIASAAYLAMFFGIGVTEVELASGTVLDIYYARYADWLFTTPLLLLDL ALLAKVDRVTIGTLIGVDALMIVTGLIGALSKTPLARYTWWLFSTIAFLFVL YYLLTSLRSAAAKRSEEVRSTFNTLTALVAVLWTAYPILWIVGTEGAGVV GLGIETLAFMVLDVTAKVGFGFVLLRSRAILGETEAPEPSAGADASAAD
PL4-98 (PROPS)	MKLLLILGSVIALPTFAAGGGDLDASDYTGVSFWLVTAALLASTVFFFVER DRVSAKWKTSLTVFGLVTGIAFWHYMYMRGVWIETGDSPTVFRYINWLL TVPLLICEFYLILAAATNVAGSLFKKLLVGSLVMLVFGYMGEAGIMAAWP ALIIGCLAWVYMIYELWAGEGESACNTASPAVQSAYNTMMYIIIFGWAIYP VGYFTGYLMGDGGSALNLDLIYNLADFVNKILFGLVIWN VAVKESSNATG
PL8-41 (PROPS)	MKLLLILGSVIALPTFAAGGGDLDASDYTGVSFWLVTAALLASTVFFFVER DRVSAKWKTSLTVFGLATGIAFWHYMYMRGVWIETGDSPTVFRYIMWLL TVPLLICEFYLILADATNVAGSLFKKLLVGSLVMLVFGYMGEAGIMAAWP ALIIGCLAWVYMIYELWAGEGKSACNTASPAVQSAYNTMMYIIIFGWAIYS VGYFTGYLMGDGGSALNLNLIYNLADFVNKILFGLVIWN VAVKESSNATG

Selected Sequences: Chapter 3

EGFP TYG: chromophore	MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT GKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFF KDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNV YIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHY LSTQSKLSKDPNEKRDHMVLLEFVTAAGITLGMDELYK
cp-EGFP (No chromophore)	VQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKI RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDH MVLLEFVTAAGITLGMDELYKGGTGHHHHHHGTGGMVSKGEELFTGVVP ILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTL
mCitrine GYG: chromophore	MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT GKLPVPWPTLVTTFGYGLMCFARYPDHMKQHDFFKSAMPEGYVQERTIFF KDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNV YIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHY LSYQSKLSKDPNEKRDHMVLLEFVTAAGITLGMDELYK
cp-mCitrine (No chromophore)	LMCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKI RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSKLSKDPNEKRDH MVLLEFVTAAGITLGMDELYKGGTGHHHHHHGTGMVSKGEELFTGVVP ILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTF
cp-EGFP (No chromophore) V1G	GQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKI RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDH MVLLEFVTAAGITLGMDELYKGGTGHHHHHHGTGGMVSKGEELFTGVVP ILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTL
cp-EGFP (No chromophore) V250A	VQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKI RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDH MVLLEFVTAAGITLGMDELYKGGTGHHHHHHGTGGMVSKGEELFTGVVP ILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTA

cp-EGFP (No chromophore) L250H	VQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKI RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDH MVLLEFVTAAGITLGMDELYKGGTGHHHHHHGTGGMVSKGEELFTGVVP ILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTH
cp-mCitrine (No chromophore) L1G	GMCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKI RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSKLSKDPNEKRDH MVLLEFVTAAGITLGMDELYKGGTGHHHHHHGTGMVSKGEELFTGVVP ILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTF
cp-mCitrine (No chromophore) L1T	TMCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKI RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSKLSKDPNEKRDH MVLLEFVTAAGITLGMDELYKGGTGHHHHHHGTGMVSKGEELFTGVVP ILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTF
cp-mCitrine (No chromophore) F250I	TMCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKI RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSKLSKDPNEKRDH MVLLEFVTAAGITLGMDELYKGGTGHHHHHHGTGMVSKGEELFTGVVP ILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTI
cp-mCitrine (No chromophore) F250T	TMCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKI RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSKLSKDPNEKRDH MVLLEFVTAAGITLGMDELYKGGTGHHHHHHGTGMVSKGEELFTGVVP ILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTT
cp-EGFP (No chromophore) H148G	VQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG DTLVNRIELKGIDFKEDGNILGHKLEYNYNSGNVYIMADKQKNGIKVNFKI RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDH MVLLEFVTAAGITLGMDELYKGGTGHHHHHHGTGGMVSKGEELFTGVVP ILVELDGDVNGHKF SVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTL
cp-mCitrine (No chromophore) H148G	LMCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG DTLVNRIELKGIDFKEDGNILGHKLEYNYNSGNVYIMADKQKNGIKVNFKI RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSKLSKDPNEKRDH MVLLEFVTAAGITLGMDELYKGGTGHHHHHHGTGMVSKGEELFTGVVP ILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTF
cp-EGFP (No chromophore) N- truncation	DHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIEL KGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSV QLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFVTA AGITLGMDELYKGGTGHHHHHHGTGGMVSKGEELFTGVVPILVELDGDV NGHKFSVSGEG EGDATYGKLTLKFICTTGKLPVPXPTLVTTL
cp-EGFP (No chromophore) C-truncation	VQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVXFXG DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKI RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDH MVLLEFVTAAGITLGMDELYKGGTGHHHHHHGTGGMVSKGEELFTGVXP ILVXLDGDX NXHKFXVSGEGXGDATYXKLXLXFICTTGKL
cp-EGFP (No chromophore) N-, C-truncation	DHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIEL KGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSV QLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFVTA AGITLGMDELYKGGTGHHHHHHGTGGMVSKGEELFTGVVPILVELDGDV NGHKFSVSGEGE GDATYGKLTLKFICTTGKL
cp- mCitrine (No chromophore) N-truncation	DHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIEL KGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSV QLADHYQQNTPIGDGPVLLPDNHYLSYQSKLSKDPNEKRDHMVLLEFVTA AGITLGMDELYKGGTGHHHHHHGTGGMVSKGEELFTGVVPILVELDGDV NGHKFSVSGEG EGDATYGKLTLKFICTTGKLPVPWPTLVTTL

cp- mCitrine (No chromophore) C-truncation	GMCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKI RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSKLSKDPNEKRDH MVLLEFVTAAGITLGMDELYKGGTGHHHHHHGTGGMVSKGEELFTGVVP ILVELDGDVNGHK FSVSGEGEGDATYGKLTLKFICTTGKL
cp- mCitrine (No chromophore) N-, C-truncation	DHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIEL KGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSV QLADHYQQNTPIGDGPVLLPDNHYLSYQSKLSKDPNEKRDHMVLLEFVTA AGITLGMDELYKGGTGHHHHHHGTGGMVSKGEELFTGVVPILVELDGDV NGHKFSVSGEGE GDATYGKLTLKFICTTGKL
PhoCle	MVIPDYFKQSFPEGYSWERSMTYEDGGICIATNDITMEEDSFINKIHFKGTN FPPNGPVMQKRTVGWEVSTEKMYVRDGVLKGDVKMKLLLKGGSHYRCD FRTTYKVKQKAVKLPDYHFVDHRIEILSHDKDYNKVKLYEHAVARNSTDS MDELYKGGSGGMVSKGEETITSVIKPDMKNKLRMEGNVNGHAFVIEGEGS GKPFEGIQTIDLEVKEGAPLPFAYDILTTAFHYGNRVFTK YPR

Appendix C

Selected Primer DNA Sequences

Appendix C

Primers Sequences: Chapter 3

cp-EGFP (No chromophore)	Fw-1: CATCATCATCATCATCATGGTACTGGTGGTATGGTGAGCAAGGGCGAGGAGCTG
	Re-2: TTTAAGCTTTTACAGGGTGGTCACGAGGGTGGGCCAGGGCAC
	Fw-3: ACGTTCCATGGTGCAGTGCTTCAGCCGCTACCCCGAC
	Re-4: ACCACCAGTACCATGATGATGATGATGACCAGTACCACCCTTGTACAGCTCGTCCA
cp-mCitrine (No chromophore)	Fw-1: GAGCTGTACAAGCATCATCATCATCATGTGAGCAAGGGCGAGGAGCTG
	Re-2: GGTAAGCTTAGAAGGTGGTCACGAGGGTGGG
	Fw-3: CACCCATGGATGCTGATGTGCTTCGCCCGCTAC
	Re-4: ATGATGATGATGATGCTTGTACAGCTCGTCCATGCC
cp-EGFP 1st/Last a.a mutation	Fw-1: ACGTTCCATGGNNCAGTGCTTCAGCCGCTACCCCGAC
	Re-2: TTTAAGCTTTTAMNNGGTGGTCACGAGGGTGGGCCAGGGCAC
cp-mCitrine 1st/Last a.a mutation	Fw-1: CTGCCATGGNNATGTGCTTCGCCCGCTAC
	Re-2: GGTAAGCTTAMNNGGTGGTCACGAGGGTGGG
cp-EGFP cp-mCitrine H148G	Fw-1: GCTGGAGTACAACTACAACAGCGGCAACGTCTATATCATGGCCG
	Re-2: CGGCCATGATATAGAGTTGCCGCTGTTGTAGTTGTACTCCAGC
cp-EGFP cp-mCitrine N-, C-, Truncation	Fw-1: CTGCCATGGACCACATGAAGCAGCACGACTTCTTCAAGTCCG
	Re-2: CGCAAGCTTTTACAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTC

References

1.   Frackowiak, D., News and View: The Jablonski diagram. *J Photoch Photobio B,* **1988**, 2, 399.

2.   Piatek, A.S., et al., Molecular beacon sequence analysis for detecting drug resistance in Mycobacterium tuberculosis. *Nat Biotechnol*, **1998**, 16(4), 359.

3.   Prober, J.M., et al., A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science,* **1987**, 238(4825), 336.

4.   Alvarez-Curto, E., Applications of fluorescence and bioluminescence resonance energy transfer to drug discovery at G protein coupled receptors. *Anal Bioanal Chem*, **2010**, 398(1), 167.

5.   Johnson, W.L., Straight, A. F., Fluorescent protein applications in microscopy. *Method Cell Biol,* **2013**, 114(99), 399.

6.   Shimomura, O., et al., Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea. *J Cell Compar Physl,* **1962**, 59(3), 223.

7.   Prendergast, F.G., Mann, K. G., Chemical and physical properties of aequorin and the green fluorescent protein isolated from Aequorea forskålea. *Biochemistry-US*, **1978**, 17(17), 3448.

8.   Prasher, D.C., et al., Primary Structure of the Aequorea-Victoria Green-Fluorescent Protein. *GENE*, **1992**, 111(2), 229.

9.   Ormo, M., et al., Crystal structure of the Aequorea victoria green fluorescent protein. *Science,* **1996**, 273(5280), 1392.

10.   Phillips, G.N., et al., The molecular structure of green fluorescent protein. *Nat Biotechnol*, **1996**, 14(10), 1246.

11.   Baird, G.S., et al., Biochemistry, Mutagenesis, and Oligomerization of DsRed, a Red Fluorescent Protein from Coral. *PNAS USA,* **2000**, 97(22), 11984.

12.   Cubitt, A.B., et al., Understanding, improving and using green fluorescent proteins. *Trends Bioch Sci*, **1995**, 20(11), 448.

13. Brejc, K., et al., Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein. *P Natl Acad Sci USA,* **1997**, 94(6), 2306.

14. Heim, R., et al., Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *PNAS USA,* **1994**, 91(26), 12501.

15. March, J.C., et al., Biotechnological applications of green fluorescent protein. *Appl Microbiol Biot,* **2003**, 62(4), 303.

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

17. Heim, R., el at., Improved green fluorescence. *Nature*, **1995**, 373(6516), 663.

18. Dezz Ropp, J., et al., Aequorea Green Fluorescent Protein: Simultaneous Analysis of Wild-Type and Blue-Fluorescing Mutant by Flow Cytometry. *Cytometry*, **1996**, 24(3), 284.

19. Yang, T.T., et al., Improved fluorescence and dual color detection with enhanced blue and green variants of the green fluorescent protein. *J Biol Chem,* **1998**, 273(14), 8212.

20. Griesbeck, O., et al., Reducing the environmental sensitivity of yellow fluorescent protein - Mechanism and applications. *J Biol Chem,* **2001**, 276(31), 29188.

21. Nagai, T., et al., A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat Biotechnol,* **2002**, 20(1), 87.

22. Matz, M.V., et al., Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat Biotechnol,* **1999**, 17(10), 969.

23. Gross, L.A., et al., The Structure of the Chromophore within DsRed, a Red Fluorescent Protein from Coral. *PNAS USA,* **2000**, 97(22), 11990.

24. Pan, T., Uhlenbeck, O. C., Circularly permuted DNA, RNA and proteins--a review. *GENE,* **1993**, 125(2), 111.

25. Topell, S., et al., Circularly permuted variants of the green fluorescent protein. *FEBS Lett*, **1999**, 457(2), 283.

26. Baird, G.S., et al., Circular Permutation and Receptor Insertion within Green Fluorescent Proteins. *PNAS USA,* **1999**, 96(20), 11241.

27.   Yu, Y., Lutz. S., Circular permutation: a different way to engineer enzyme structure and function. *Trends biotechnol,* **2011** 29(1), 18.

28.   Cheltsov, A.V., et al., Circular permutation of 5-aminolevulinate synthase. Mapping the polypeptide chain to its function. *J Biol Chem*, **2001**, 276(22), 19141.

29.   Dong, J., et al., Isomerization in Fluorescent Protein Chromophores Involves Addition/Elimination. *JACS,* **2008**, 130(43), 14096.

30.   Clark, T.B., et al., Synthesis and Optical Properties of Two-Photon Absorbing GFP-type Probes. *J Phys Chem C*, **2011**, 115(15) ,7331.

31.   Ivashkin, P.E., et al., Synthesis and properties of chromophores of fluorescent proteins. *Russ J Bioorg Chem*, **2009**, 35(6), 652.

32.   Chandrasekhar, S., Srimannarayana. M., The Erlenmeyer synthesis with a thioazlactone. *JOC*, **2009**, 290.

33.   El-Mekabaty, A., Erlenmeyer Azlactones. Synthesis, Reactions and Biological Activity. *ChemInform,* **2013**, 44(34).

34.   Lee. C., et al., Facile synthesis of 4-arylidene-5-imidazolinones as synthetic analogs of fluorescent protein chromophore. *Tetrahedron*, **2012**, 68(29), 5898.

35.   Baldridge, A., Efficient Synthesis of New 4-Arylideneimidazolin-5-ones Related to the GFP Chromophore by [2+3] Cyclocondensation of Arylideneimines with Imidate Ylides. *Synthesis*, **2010**, 41(14), 2424.

36.   Guella, G., et al., Isolation, Synthesis and Photochemical Properties of Almazolone, a New Indole Alkaloid from a Red Alga of Senegal. *ChemInform*, **2006**, 37(41), 1165.

37.   Baranov, M.S., et al., A synthetic approach to GFP chromophore analogs from 3-azidocinnamates. Role of methyl rotors in chromophore photophysics. *Chem Commun* , **2013,** 49(51), 5778.

38.   Turner, A., Biosensors: Fundamentals and Applications, **1989**.

39.   Wouters, F.S., et al., Imaging biochemistry inside cells. *Trends Cell Biol*, **2001**, 11(5), 211.

40.   Zhang, J., et al., Creating New Fluorescent Probes for Cell Biology. *Nat Rev Mol Cell Bio*, **2002**, 3(12), 906.

41. Fernández-Suárez, M., Ting, A. M., Fluorescent probes for super-resolution imaging in living cells. *Nat Rev Mol Cell Bio,* **2008**, 9(12), 929.

42. Nausch, L.W.M., et al., Differential patterning of cGMP in vascular smooth muscle cells revealed by single GFP-linked biosensors. *PNAS USA,* **2008**, 105(1), 365.

43. Xia, Z., Rao, J., Biosensing and imaging based on bioluminescence resonance energy transfer. *Curr Opin Biotech*, **2009**, 20(1), 37.

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

45. Shekhawat, S.S., Ghosh, I., Split-protein systems: beyond binary protein-protein interactions. *Curr Opin Chem Bio,* **2011**, 15(6), 789.

46. Galarneau, A., et al., β-Lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein-protein interactions. *Nat Biotechnol*, **2002**, 20(6), 619.

47. Pelletier, N., et al., Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. *P Natl Acad Sci USA,* **1998**, 95(21), 12141.

48. Rossi, F., et al., Monitoring protein-protein interactions in intact eukaryotic cells by beta-galactosidase complementation. *P Natl Acad Sci USA,* **1997**, 94(16), 8405.

49. Ghosh, I., et al., Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein. *JACS,* **2000**, 123(23), 5658.

50. Shyu, Y.J., et al., Identification of new fluorescent protein fragments for biomolecular fluorescence complementation analysis under physiological conditions. *Biotechniques,* **2006**, 40(1), 66.

51. Hu, C.D., et al., Visualization of interactions among bZip and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol Cell,* **2002**, 9(4), 789.

52. Jach, G., et al., An improved mRFP1 adds red to bimolecular fluorescence complementation. *Nat Methods,* **2006**, 3(8), 597.

53. Fan, J.-Y., et al., Split mCherry as a new red bimolecular fluorescence complementation system for visualizing protein-protein interactions in living cells. *Bioch Bioph Res Co*, **2008**, 367(1), 47.

54.   Wilson, C.G.M., et al., Detecting protein-protein interactions with GFP-fragment reassembly. *Nat Methods,* **2004**, 1(3), 255.

55.   Alford, S.C., et al., Dimerization-Dependent Green and Yellow Fluorescent Proteins. *ACS Synth Biol*, **2012**, 1(12), 569.

56.   Alford, S.C., et al., A Fluorogenic Red Fluorescent Protein Heterodimer. *Chem Biol,* **2012**, 19(3), 353.

57.   Kroeze, W.K., G-protein-coupled receptors at a glance. *J Cell Sci,* **2003**, 116(24), 4867.

58.   Kobilka, B.K., et al., *G* protein coupled receptor structure and activation. *BBA-Biomembranes,* **2007**, 1768(4), 794.

59.   Hofmann, K.P., et al., A G protein-coupled receptor at work: the rhodopsin model. *Trends Biochem Sci,* **2009**, 34(11), 540.

60.   Menon, S.T., et al., Rhodopsin: structural basis of molecular physiology. *Physiol Rev*, **2001**, 81(4), 1659.

61.   Okada, T., et al., The retinal conformation and its environment in rhodopsin in light of a new 2.2 angstrom crystal structure. *J Mol Biol,* **2004**, 342(2), 571.

62.   Krebs, A., et al., The three-dimensional structure of bovine rhodopsin determined by electron cryomicroscopy. *J Biol Chem,* **2003**, 278(50), 50217.

63.   Schertler, G.F.X., *Structure of rhodopsin. EYE*, **1998**, 12, 504.

64.   Oesterhelt, Stoeckenius, W., D., Rhodopsin-like protein from the purple membrane of Halobacterium halobium. *Nature,* **1971**, 233(39), 149.

65.   Klare, J.P., et al., Microbial rhodopsins: scaffolds for ion pumps, channels, and sensors. *Results Probl Cell Differ*, **2008**, 45, 73.

66.   Nagel, G., et al., Channelrhodopsin-1:A Light-Gated Proton Channel in Green Algae. *Science,* **2002**, 296(5577), 2395.

67.   Jung, K.H., et al., Demonstration of a sensory rhodopsin in eubacteria. *Mol Microbiol*, **2003**, 47(6), 1513.

68.   Ni, L., et al., Human liver mitochondrial aldehyde dehydrogenase: Three-dimensional structure and the restoration of solubility and activity of chimeric forms. *Protein Sci*, **1999**, 8(12), 2784.

69.   López, S., et al., Synthesis of N-Heteroaryl Retinals and their Artificial Bacteriorhodopsins. *Chembiochem,* **2005**, 6(11), 2078.

70.   Barachevsky, V.A., et al., Properties of photochromic retinals. *Dyes Pigments,* **2012**, 92(2), 831.

71.   Bayley, H., et al., Light-driven proton translocation by bacteriorhodopsin reconstituted with the phenyl analog of retinal. *J Biol Chem,* **1981**, 256(8), 3797.

72.   Hota, P.K., et al., Bacteriorhodopsin analogue from indolic chromophores. *Lett Org Chem*, **2007**, 4(4), 300.

73.   Laptev, A.V., et al., Spiropyran Analogs of Retinal: Synthesis and Study of Their Photochromic Properties. *High Energ Chem+*, **2008**, 42(7), 601.

74.   Rein, T., et al., Synthesis of polyconjugated aldehydes using a new Horner--Wadsworth--Emmons reagent. *Acta Chem Scand B*, **1988**, 42(8), 569.

75.   Carlson, N.R., Foundations of physiological psychology. *Allyn and Bacon,* **1992**.

76.   Peterka, D.S., et al., Imaging Voltage in Neurons. *Neuron*, **2011**, 69(1), 9.

77.   Cohen, L.B., et al., Light scattering and birefringence changes during nerve activity. *Nature,* **1968**, 218(5140), 438.

78.   Nagai, T., et al., Circularly Permuted Green Fluorescent Proteins Engineered to Sense Ca 2+. 2001, *P Natl Acad Sci USA,* **2001** , 98(6), 3197.

79.   Nakai, J., et al., A high signal-to-noise Ca2+ probe composed of a single green fluorescent protein. *Nat Biotechnol,* **2001**, 19(2), 137.

80.   Miesenbock, G., et al., Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature,* **1998**, 394(6689), 192.

81.   Marcaggi, et al., P., Optical Measurement of mGluR1 Conformational Changes Reveals Fast Activation, Slow Deactivation, and Sensitization. *P Natl Acad Sci USA,* **2009**, 106(27),11388.

82.   Villalba-Galea, C.A., et al., Letter: Charge Movement of a Voltage-Sensitive Fluorescent Protein. *Biophys J*, **2009**, 96(2) , L19.

83.   Siegel, M.S., Isacoff, Y., A Genetically Encoded Optical Probe of Membrane Voltage. *Neurons,* **1997**, 19(4), 735.

84.   Ataka, K., Pieribone. V. A., A genetically targetable fluorescent probe of channel gating with rapid kinetics. *Biophysl J*, **2002**, 82(1), 509.

85.   Sakai, R., et al., Design and characterization of a DNA-encoded, voltage-sensitive fluorescent protein. *Eur J Neurosci*, **2001**, 13(12), 2314.

86.   Lundby, A., et al., Engineering of a Genetically Encodable Fluorescent Voltage Sensor Exploiting Fast Ci-VSP Voltage-Sensing Movements. *PLoS*, **2008**, 3(6), 1.

87.   Wang, D.S., et al., Hybrid voltage sensor imaging of electrical activity from neurons in hippocampal slices from transgenic mice. *J Neurophysiol*, **2012**, 108(11), 3147.

88.   Chanda, B., et al., A hybrid approach to measuring electrical activity in genetically specified neurons. *Nat Neurosci,* **2005**, 8(11), 1619.

89.   Kralj, J.M., et al., Optical recording of action potentials in mammalian neurons using a microbial rhodopsin. *Nat Methods*, **2012**, 9(1), 90.

90.   Kralj, J.M., et al., Electrical spiking in Escherichia coli probed with a fluorescent voltage-indicating protein. *Science*, **2011**, 333(6040), 345.

91.   Maclaurin, D., et al., Mechanism of voltage-sensitive fluorescence in a microbial rhodopsin. *P Natl Acad Sci USA*, **2013**, 111(15), 5939.

92.   Hampp, N., Bacteriorhodopsin as a photochromic retinal protein for optical memories. *Chem Rev,* **2000**, 100(5), 1755.

93.   Kouyama, T., et al., Research Article: Excited-state dynamics of bacteriorhodopsin. *Biophys. J*, **1985**, 47(1), 43.

94.   Maryanoff, B.E., et al., NMR rate study on the Wittig Reaction of 2,2-dimethylpropanal and trebutylidenephosphorane. *Tetrahedron Lett,* **1989**, 30(11), 1361.

95.   Becker, R.S., Kouyama, T., The visual process: photophysics and photoisomerization of model visual pigments and the primary reaction. *Photochem Photobiol*, **1988**, 48(3), 369.

96.   Yoriko, A., Akiyama, H., pH-dependent fluorescence spectra, lifetimes, and quantum yields of firefly-luciferin aqueous solutions studied by selective-excitation fluorescence spectroscopy. *Jpn J Appl Phys*, **2010**, 49(11), 117002.

97.   Hochbaum, D.R., et al., All-optical electrophysiology in mammalian neurons using engineered microbial rhodopsins. *Nat Methods,* **2014**, 11(8), 825.

98.   Yoshimura, K., Kouyama, T., Structural role of bacterioruberin in the trimeric structure of archaerhodopsin-2. *J Mol Biol*, **2008**, 375(5), 1267.

99.   Trainor, K., An Easily Prepared Wide Range Buffer Series. *Aust Scie Teach J*, **1971**, 17(4), 107.

100.   Nifosì, R., Tozzini, V., Cis–trans photoisomerization of the chromophore in the green fluorescent protein variant E^{2}GFP: A molecular dynamics study. *Chem Phys*, **2006**, 323 (2), 358.

101.   Jian, D., et al., Activation and tuning of green fluorescent protein chromosome emission by alkylsubstituent-mediated crystal packing. *J Phys Chem B*, **2008,** 112(46), 662.

102.   Ikejiri, M., et al., Design and Concise Synthesis of a Novel Type of Green Fluorescent Protein Chromophore Analogue. *Org Lett,* **2012**, 14(17), 4406.

103.   Conyard, J., et al., Chemically Modulating the Photophysics of the GFP Chromophore. *J Phys Chem B*, **2011**, 115(6), 1571.

104.   Simões, S.I., et al., The effect of cholate on solubilisation and permeability of simple and protein-loaded phosphatidylcholine/sodium cholate mixed aggregates designed to mediate transdermal delivery of macromolecules. *Eur J Pharm Biopharm,* **2004**, 58, 509.

105.   Seddon, A.M., et al., Review: Membrane proteins, lipids and detergents: not just a soap opera. *BBA - Biomembranes,* **2004**, 1666(1-2), 105.

106.   Zhong, S., et al., *A Genetically-Encoded YFP Sensor with Enhanced Chloride Sensitivity, Photostability and Reduced pH Interference Demonstrates Augmented Transmembrane Chloride Movement by Gerbil Prestin (SLC26a5). PLoS,* **2014**, 9(6), 1.

107.   Tzitzilonis, C., et al., Detergent/Nanodisc Screening for High-Resolution NMR Studies of an Integral Membrane Protein Containing a Cytoplasmic Domain. *PLoS,* **2013**, 8(1), 1.

108.   Monera, O.D., et al., Protein denaturation with guanidine-hydrochloride or uea provides a different estimate of stability depending on the contribution of electrostatic interactions. *Protein Sci*, **1994**, 3(11), 1984.

109.   Hoischen. D., et al., Merocyanines as extremely bathochromically absorbing chromophores in the halobacterial membrane protein bacteriorhodopsin. *Angew. Chem,* **1997**, 36 (15), 1630.