NTNU – Trondheim
Norwegian University of
Science and Technology

Synthetic biology: Generation of new GFP variants with improved characteristics

Azra Fida

Biotechnology
Submission date: May 2015
Supervisor: Martin Frank Hohmann-Marriott, IBT
Co-supervisor: Rahmi lale, IBT
Eivind Almaas, IBT

Norwegian University of Science and Technology
Department of Biotechnology

NTNU

Synthetic Biology: Generating Fluorescent Proteins with improved characteristics

Norwegian University of Science and Technology, Institute of Biotechnology

Azra Fida

5/15/2015

Masters in Biotechnology (2 years)

Supervisor: Martin Frank Hohmann-Marriott, IBT

Co-Supervisor: Rahmi Lale, IBT

Abstract

This project aimed to create new fluorescent protein (FPs) variants using the GFPmut3b, EYFP and mRFP1 through fusing two FPs together. The proximity of these fused FPs was expected to result in altered light absorption and fluorescence emission spectra compared to the single FPs. In order to fuse the Fps, plasmids were generated by genetically fusing the coding sequences of two FPs into the expression vector pSB1AK3 with a lacI IPTG-inducible promoter. Gibson cloning and CPEC cloning were used to genetically link the FP. Sequence analysis of the generated constructs, however, indicated that two identical FPs were integrated into the expression vector.

Characterization of the linked FPs for light absorption and fluorescence emission showed that their spectral features were identical to the original FPs with absorption maxima at 501 nm, 514 nm, 584 nm and fluorescence emission peaks at 511 nm, 527 nm and 607 nm, for GFPmut3b, EYFP and mRFP1 respectively. The maturation of expressed single FPs and linked FPs was also characterized by inhibiting protein expression with the inhibitor chloramphenicol.

Sammendrag

Målet av prosjektet var å lage nye fluorescerende protein (FP) varianter ved å fusjonere to FPer (GFPmut3b, EYFP og mRFP) sammen. De fusjonerte FPene ble forventet å føre til endrede lys absorpsjon og fluorescens emisjonsspektre i forhold til de enkelte FPene. Fusjonerte FPene ble satt sammen i ekspresjonsvektoren pSB1AK3 under kontroll av en lacI-IPTG-induserbar promoter. Gibson og CPEC kloning metodene ble brukt for å koble de to FP genene. Sekvensanalyse av de genererte konstruksjonene indikerte at to identiske FPer ble integrert i ekspresjonsvektoren. Karakterisering av de fusjonerte FPene viste at de spektrale egenskaper var identiske med de opprinnelige enkelte FPene med absorpsjonsmaksima på 501 nm, 514 nm, 584 nm og fluorescens emisjonsmaksima på 511 nm, 527 nm og 607 nm, for GFPmut3b, EYFP og mRFP1. Stabilitet av uttrykte enkelte og fusjonerte FPene ble også karakterisert ved bruk av translasjon inhibitoren kloramfenikol.

Acknowledgements

Firstly, I am grateful to Almighty God for blessing me with good health and wellbeing that were necessary to complete this project. Then, I would like to thank my family for their kind support and encouragement.

Completion of this project was achieved under the supervision of my advisors, Associate Professor Martin Hohmann-Marriott and Rahmi Lale, who were there for me anytime I needed help. They encouraged me while I faced any failure during this project. I learned a lot during this entire project under their supervision. I would like to express my immense gratitude to them for their support as I would never have made it through without their guidance. I would also like to thank Thor Bernt Melø for helping me with Spectroscopy.

Apart, I am thankful to my friends especially Ewelina and other lab fellows as well as the staff of Biotechnology by providing a friendly environment in the lab.

Contents

List of Figures

List of Tables

Abbreviations

Amp	Ampicillin
CAP	Chloramphenicol
CPEC	Circular Polymerase Extension Cloning
DMSO	Dimethyl sulfoxide
EYFP	Enhanced Yellow Fluorescent Protein
FP	Fluorescent Protein
FIC	Fluorescence Intensity per Cell
FRET	Förster resonance energy transfer
GFP	Green Fluorescent Protein
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
PCR	Polymerase Chain Reaction
RFP	Red Fluorescent Protein

1 Introduction

1.1 Background

The history of FPs goes back to 1955 when different luminescent responses were studied by Davenport and Nicol using animals such as, *Aequorea, Halistaura, Phialidium* and *Stomotoca*. They described green fluorescence produced by the light organs of *Aequorea* for the first time (Davenport & Nicol, 1955). This green fluorescent substance was later on discovered to be a protein in 1962 as it was isolated as a by-product during the purification of Aequorin, a luminiscent substance found in *Aequorea* jellyfish. This protein was green in its precipitated form, so it was called `green protein` at that time (Osamu Shimomura, Johnson, & Saiga, 1962). In 1969, an important point was raised: The light emitting organism emits a green light as a whole, but the emission produced by the isolated reaction *in vitro* was blue. This puzzle was solved in 1971 when energy transfer was studied in light emitting ceolentrates and it was found that the bioluminiscence in the light emitting animals involves the energy transfer from an excited molecule, a calcium-activated protein to a second, molecule named the green fluorescent protein with emission at 508 nm. This calcium-activated protein was Aequorin discovered earlier (Morin & Hastings, 1971). So, upon activation Aequorin emits a blue light which is absorbed by GFP and produces green fluorescence.

The structure of the GFP chromophore was elucidated in 1979. The GFP was first heat-denatured, and then treated with papain to digest it. A fragment containing the chromophore was isolated from the digest and for the structure determination. The structure was deduced on the basis of an analogy with a model compound (O. Shimomura, 1979). This structure did not show any proper physical evidence and did not possess some important information such as, the amino acid glycine being part of the structure was missing, no quantitative data was present and the structure could not explain the spectroscopic properties of the chromophore. Therefore, a full amino acid sequence as well as structure of papain-derived chromophore with hexapeptide was proposed in 1992 (Cody, Prasher, Westler, Prendergast, & Ward, 1993).

After the determination of the structure of the GFP, a transforming discovery was that GFP could be cloned and expressed in living animals. The cloning of the GFP was achieved in 1992 for the first time, when the gene and cDNA for GFP were isolated for characterization, as well as for the determination of the evolutionary relationships in coelentrates bioluminiscence (Prasher, Eckenrode, Ward, Prendergast, & Cormier, 1992). Another major breakthrough regarding GFP is its expression in prokaryotic cells, such as *Escherichia coli* and eukaryotic cells, such as *Caenorhabditis elegans* which was performed in 1994. Expression of GFP in these organisms showed, that no exogenous substrates or cofactors are required for the fluorescence of GFP. GFP can be used to track gene expression and protein localization in living cells (Chalfie, Tu, Euskirchen, Ward, & Prasher, 1994).

At the time of their discovery, GFP and Aequorin did not find adaptation as tools, but they are widely used 40 years after their discovery. Aequorin is used as a calcium probe while GFP as a protein marker. The use of GFP and its colored mutants has contributed to discoveries in the life sciences. The people involved in the discovery and development of GFP were awarded Nobel Prize in Chemistry 2008.

1.2 Fluorescent Proteins

After the discovery of GFP, many GFP like proteins have been identified in other marine animals like *Hydroid polyps, medusa, scleractinian corals, sea anemones, sea pens, crustacea, maxillopoda, copepoda and pontellidae* etc (Chudakov, Lukyanov, & Lukyanov, 2005). The GFP family is comprised of homologous proteins that have an 11 stranded B barrel structure as shown in figure 1.1 that contains a chromophore formed by their internal amino acids (Miyawaki, Nagai, & Mizuno, 2003).

All these FPs have sizes between 220-240 amino acid residues, have beta-barrel shape and possess a chromophore located in the center. The GFP from *Aequorea victoria* consists of 238 amino acids having serine, tyrosine and glycine at positions 65, 66 and 67 respectively. The cyclization and oxidation of these three amino acids at positions 65-67 generates an internal p-hydroxybenzylideneimidazolidinone chromophore (see figure 1.1) which is

responsible for the GFP absorption band in the blue and fluorescence emission peak in the green (Heim, Prasher, & Tsien, 1994). The FPs do not require any external cofactor or enzymes to produce the chromophore but rather self-generate it through post translational self-modification of the three amino acids at position 65-67.

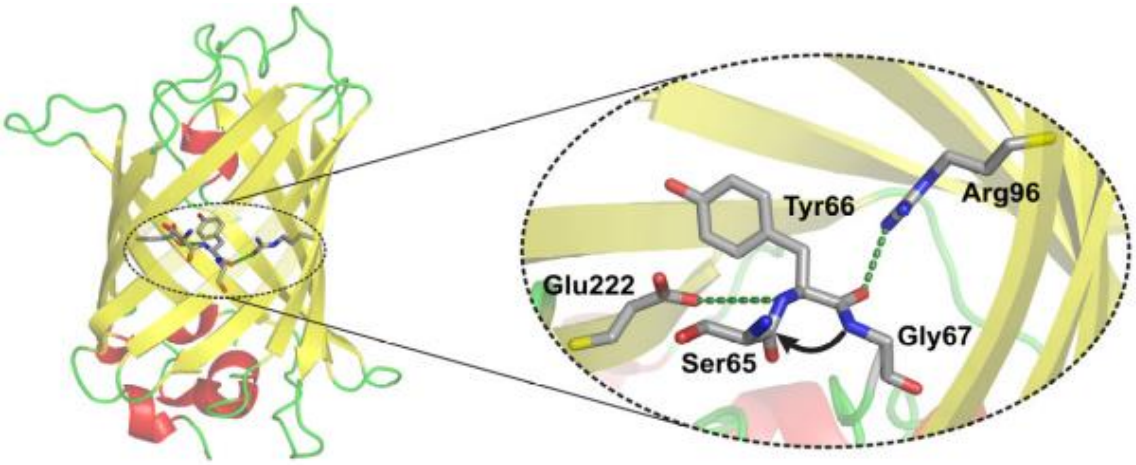


Figure 1.1 - The beta-barrel structure of FPs and the amino acid residues involved in the formation of the chromophore with Arg96 and Glu222 in surrounding (Pakhomov & Martynov, 2008).

1.2.1 Chromophore formation

The amino acid-derived chromophore is the common feature of all the FPs which is the integral part of the protein backbone. It makes the FPs useful as genetically encoded tools as this chromophore is produced auto catalytically.

The mechanism of chromophore formation involves the following steps, (1) Folding of the GFP into a native conformation (2) The formation of imidazoline, a five membered ring, due to backbone cyclization resulting from the nucleophilic attack of the amide of glycine at position 67 on the carbonyl group of the serine residue at position 65 (3) Dehydration (4) Oxidation which causes the dehydrogenation of α-β bond hydoxybenzyl sidenchain of tyrosine residue at position 66 to conjugate its aromatic group with the imidazolinone shown

in figure 1.2. The maturation of the chromophore results in a protein with visible absorbance and fluorescence (Tsien, 1998).

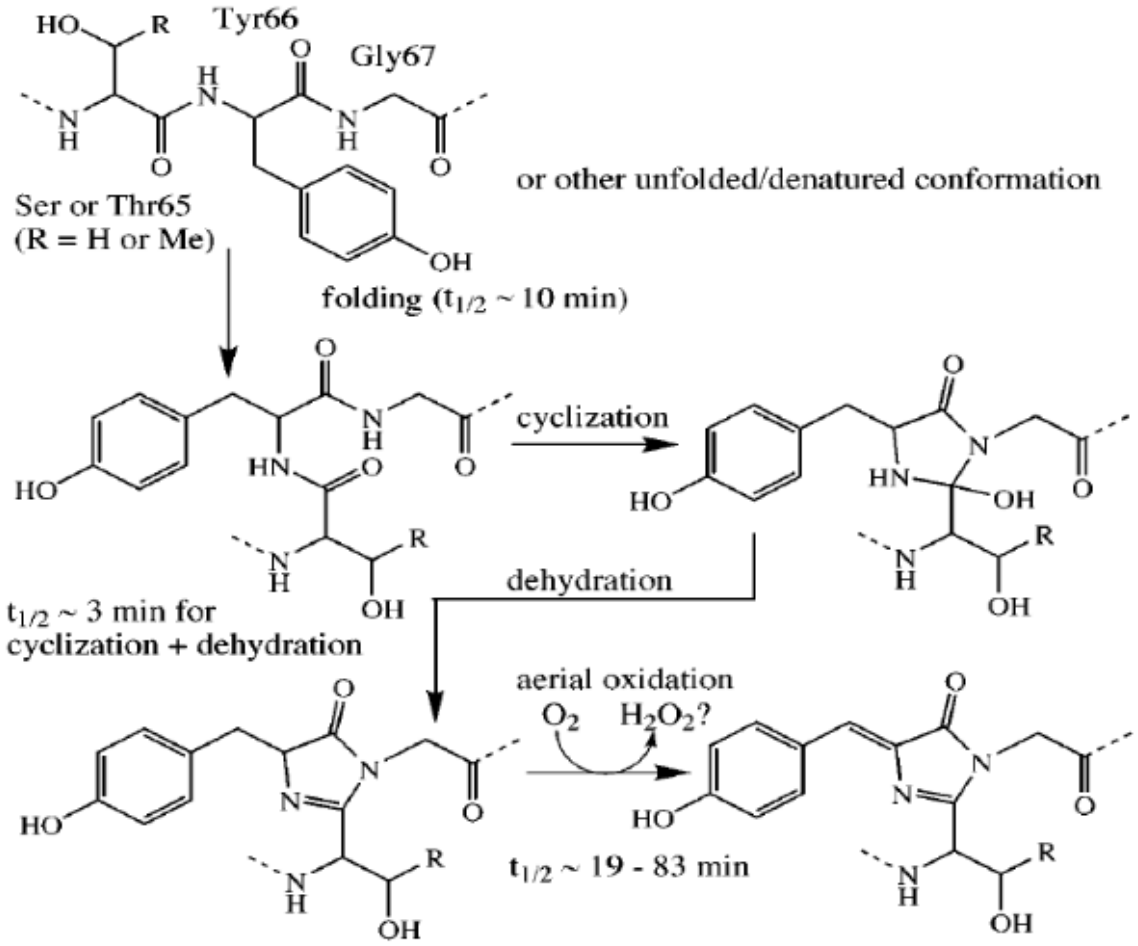


Figure 1.2 - Intra Molecular biosynthesis of GFP chromophore Mechanism proposed by (Cubitt et al., 1995)**:** The translated protein is cyclized, and remains soluble and non-fluorescent until oxidation occurs that dehydrogenates Tyr66 and generates a fluorophore (Tsien, 1998).

1.2.2 Diversity of chromophores

The three amino acids at position 65, 66 and 67 contribute in the formation of the chromophore, are a glycine at position 67 which is important for the chromophore formation,

a tyrosine lies at position 66 and is important for chromophore maturation by providing a suitable oxidative chemistry. This tyrosine also prevents unwanted side reactions like hydrolysis and backbone fragmentation. The tyrosine can be replaced by any aromatic amino acid emitting light in cyan-blue spectra, hence producing a chromophore that is mature and chemically different. Amino acid residues at position 65 have a high influence on the overall structure of the chromophore and are extremely varied among all the FPs (Stepanenko et al., 2011).

All FPs have the same chromophore but it is the modification in the protein environment which creates protein with different spectral structure. In case of red FPs longer emission of about 583 nm is produced because chromophore becomes physically larger due to an additional step of oxidation after chromophore synthesis. This additional oxidation step extends the conjugated Π system of the chromophore, thereby broadening the polypeptide backbone leading to a longer wavelength emission (Remington, 2006) as shown in figure 3. So, the color differences seen in different proteins are because of changes in the chromophore structure (Labas et al., 2002).

FP Kaede, forms a fluorescence protein family that is converted from green to red involves a different way of red chromophore formation involving illumination. This protein has histidine at position 65 and it matures to its green fluorescent form in the dark where the chromophore exists in equilibrium between protonated and deprotonated states. When illuminated with UV light, the protonated chromophore matures into its red form. Similarly, chemical modifications in the DsRed FP results in the formation of yellow-orange FPs (Chudakov, Matz, Lukyanov, & Lukyanov, 2010) (see figure 1.3).

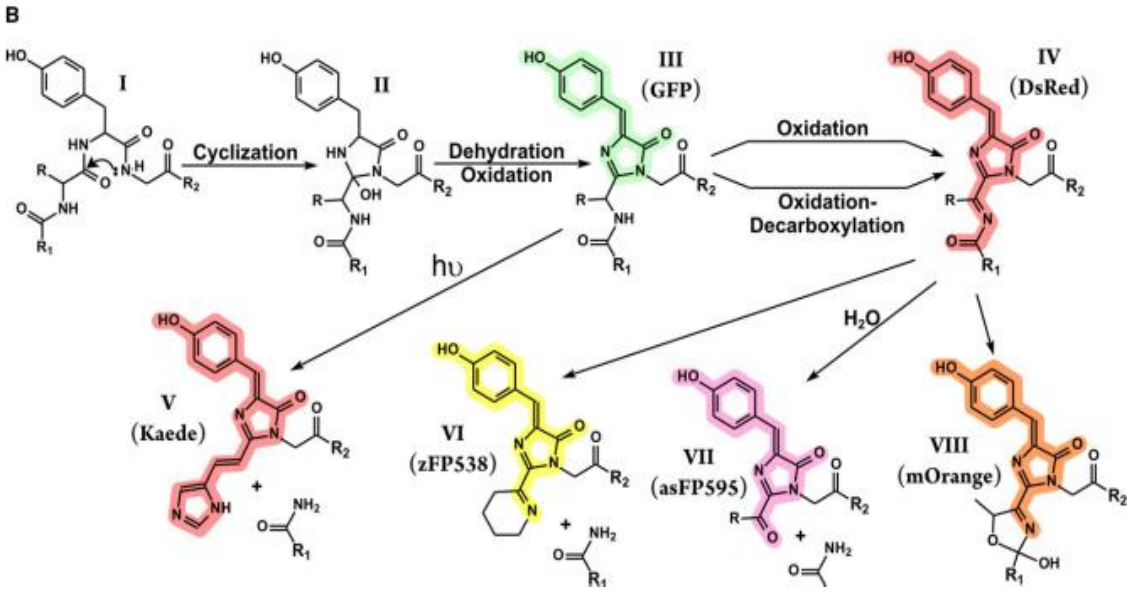


Figure 1.3 - Chromophore formation and natural diversity: After synthesis of the chromophore additional modifications can lead to diverse chromophore structure emitting lights of various colors and wavelength. An extra step of oxidation after GFP formation (Green) produces FP with a red color (Red). Hydrolysis of RFP leads to further changes in chromophore structure that produces FPs with different colors (shown as yellow, orange and purple) (Pakhomov & Martynov, 2008).

1.3 Fluorescence and Absorbance

A bluish-green light is emitted from the margin of the umbrella of Aequorea jelly fish. This light is produced because of two closely related proteins: Aequorin and green fluorescent protein. Aequorin is 21.4 KDa protein which contains a calcium-binding apoprotein called apoaequorin, an organic substrate called coelentrazine and molecular oxygen. Green Fluorescent protein is 27 KDa containing chromophore which is a modified hexapeptide. When calcium binds to aequorin, coelentrazine is oxidized to coelentramide producing carbon dioxide, light and the blue fluorescent protein with fluorescence 470 nm, carbon dioxide and light. Upon adding GFP to the reaction mixture, light that appears bluish green similar to the one produced by the animal *in vivo* is generated with a fluorescence emission peak at 508 nm (Inouye & Tsuji, 1994; Osamu Shimomura et al., 1962).

FPs have specific excitation and emission spectra. GFP from Aequorea has the largest absorption peak at 395 nm which is three times higher than a smaller peak at 475 nm in amplitude. GFP gives an emission at 508 nm when excited at 395 nm (Cubitt et al., 1995), which shows that the chromophore may exist in two incontrovertible states. The 476 nm peak is because of the GFP molecules that are deprotonated while the 376 nm peak is from the molecules in protonated state (Tsien, 1998). Yellow FPs are produced by stacking GFP-like chromophore with Tyrosine at position 203 and have an excitation peak at 515 nm and emission peak at 530 nm. Red FPs have excitation peaks ranging between 560-590 nm and emission ranges between 580-610 nm. The absorption and fluorescence emission spectra of different FPs are shown in figure 1.4.

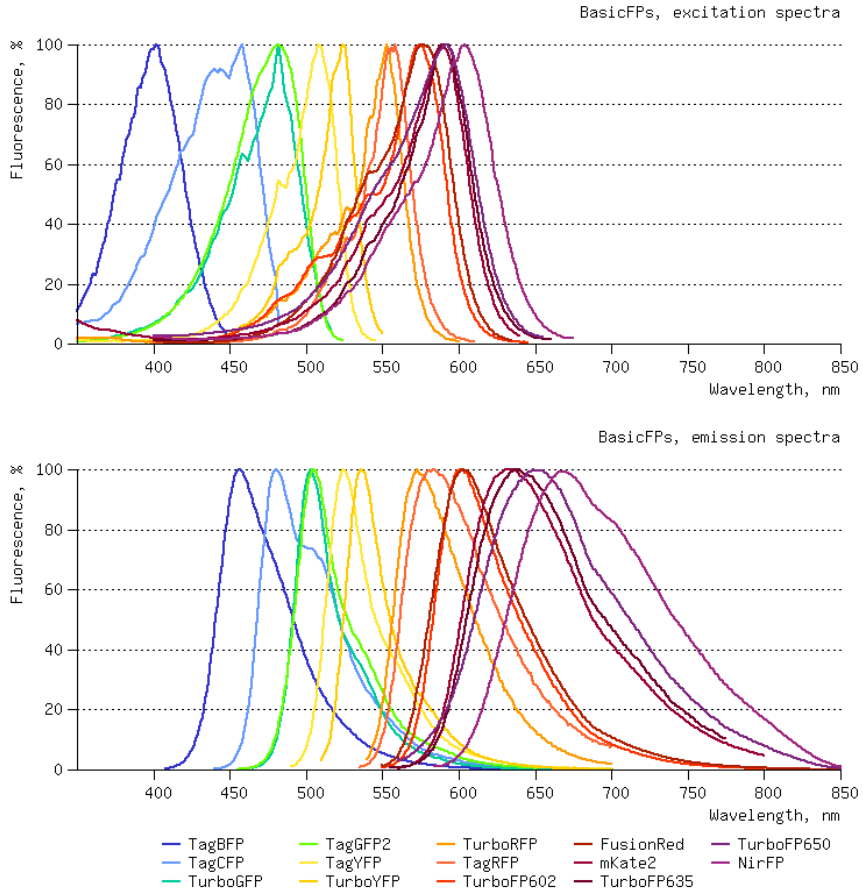


Figure 1.4 - Absorption and Emission peaks of different FPs (http://www.evrogen.com/products/basicFPs.shtml)**.**

1.4 Variants of GFP and red fluorescent protein

In order to provide distinct markers to track different cellular events at the same time and to describe the structural-functional relationship of cellular proteins, mutation of FP with modified excitation emission spectra are important. One of the drawbacks in using wild type GFP as a marker is the low fluorescence yield and long maturation times. Secondly, as discussed earlier, the wild type GFP gives two absorbance peaks, one at 375 nm with high amplitude and another at 475 nm with lower amplitude, the former being sensitive to photo bleaching while the latter is resistant (Chalfie et al., 1994). Mutating the wild type GFP can help to change the rate of chromophore formation hence fluorescence as well as the spectral properties of the protein based on the discovery that tyrosine at position 66 can be substituted with any aromatic amino acid (see figure 1.5). Therefore, an attempt was made to convert the excitation spectrum of GFP to a single peak at longer wavelength through random mutagenesis which resulted in different variants of wild type GFP due to single amino acid substitution (Heim et al., 1994).

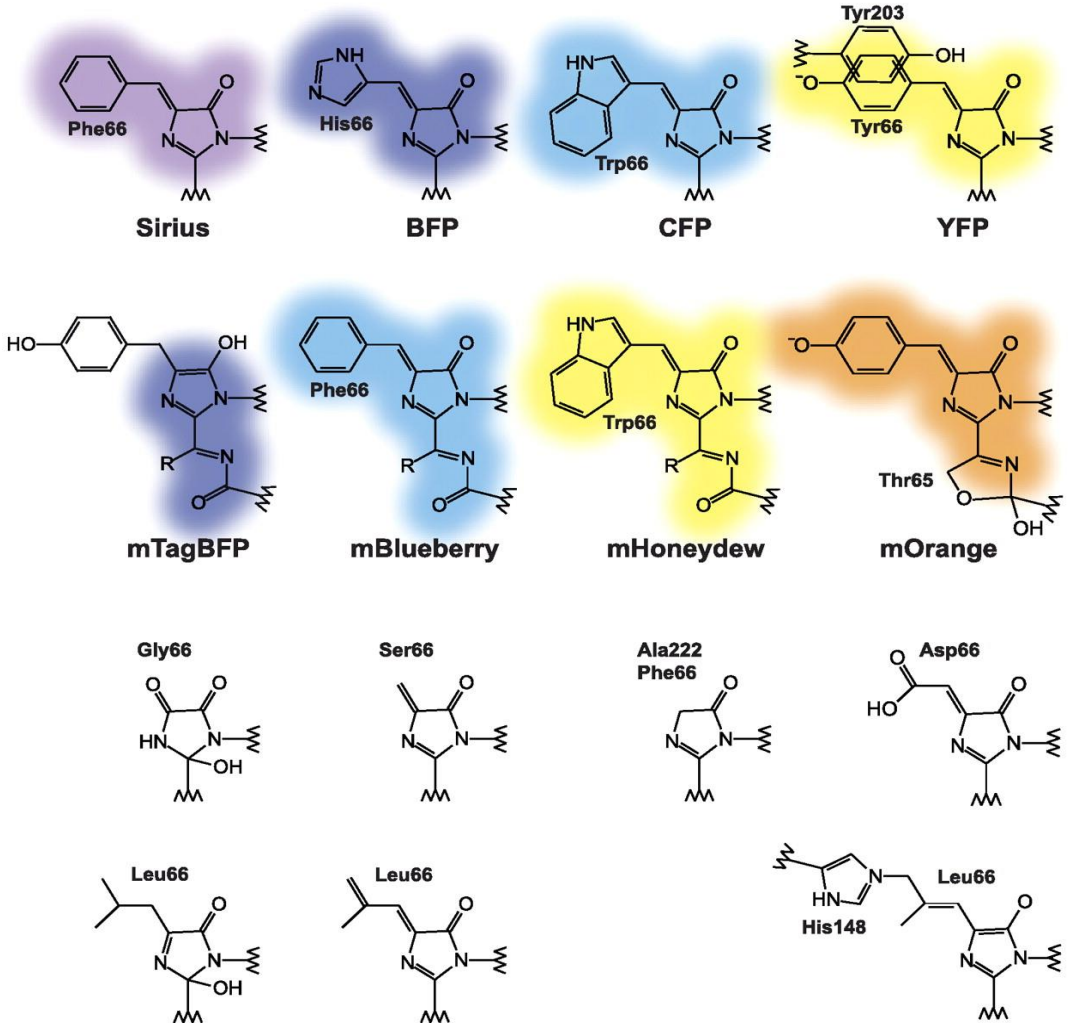
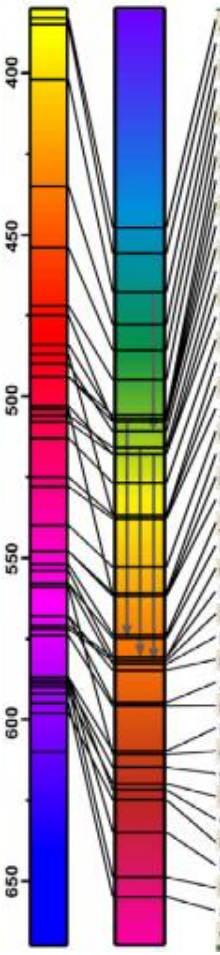


Figure 1.5 - Chromphore structure in FPs mutated artificially: Colors of the chromophore seen after replacing tyr at 66 aromatic amino acids like Phe, His and Trp in blue shifted chromophore (Top row). Middle row: blue-shifted modification of DsRed chromophore addition of Phe or Trp residues as substitute for Tyr66, or cyclization of Thr65 by mutation Gln65Thr. Bottom rows: Colorless GFP chromophore variants in Tyr66-substitued variants (Chudakov et al., 2010).

To produce FPs with improved characteristics different mutant FPs have been developed. The variants of these FPs are different from wild types regarding their spectral features (see figure 1.6). Mutant proteins of GFP subfamily cover an emission range between 440-527 nm and emit the light of blue, cyan green and yellow.

The wild type GFP was used to produce FPs of various colors like green, blue and cyan but the most important color that is red was still missing. The main reason to search for a red FP was to image the tissue from live animals because hemoglobin in blood absorbs infrared wavelength which is emitted in 650-900 nm range. A red fluorescent protein isolated from *Discosoma* species belongs to Ds Red family which includes orange, red and far red FPs with an emission range 550-650 nm (Pakhomov & Martynov, 2008). Ds red from corals had some properties that were not satisfactory to be ideal like, they were tetrameric, more orange than red and have a slow maturation. To improve the characteristics of Ds red 33 mutations were introduced and a monomeric red fluorescent protein (mRFP1) was created. This RFP is used to create more improved variants.



Protein	Absorbance maximum, nm	Fluorescence emission maximum, nm	Molar extinction coefficient, x10^{3} M^{-1}cm^{-1}	QY[a]	Subfamily/ Chromophore	References[b]
Azurite	384	450	26.2	0.55	GFP-Y66H	Mena et al., 2006
mKalama1	385	456	36	0.45	GFP	Ai et al., 2007
PS-CFP (cyan)[c]	402	468	34	0.16		Chudakov et al., 2004; Lukyanov et al., 2005
ECFP	434	477	26	0.4	GFP-Y66W	Heim et al., 1994; Patterson et al., 2001
amFP486	454	486	ND[f]	0.71	GFP	Matz et al., 1999; Henderson and Remington, 2005
MiCy	472	495	27.25	0.9		Karasawa et al., 2004
zFP506	494	506	ND	ND	GFP	Matz et al., 1999; Pakhomov and Martynov, 2007
EGFP	484	507	53	0.6	GFP	Heim et al., 1995; Patterson et al., 1997
GFP	475; 395	508	25	0.79	GFP	Ormo et al., 1996; Patterson et al., 1997
DendFP (green)[c]	492	508	90	0.65	GFP	Pakhomov et al., 2004; Gurskaya et al., 2006
PS-CFP (green)[c]	490	511	27	0.19		Chudakov et al., 2004; Lukyanov et al., 2005
EosFP (green)[c]	506	516	72	0.70	GFP	Wiedenmann et al., 2004; Nienhaus et al., 2005
Dronpa	503	518	95	0.85	GFP	Ando et al., 2004; Andresen et al., 2007
Kaede (green)[c]	508	518	98.8	0.80	GFP	Ando et al., 2002; Mizuno et al., 2003
EYFP	514	527	84	0.61	GFP	Ormo et al., 1996; Wachter et al., 1998
mHoneydew[d]	487	537	17	0.12		Shaner et al., 2004
PhiYFP	525	537	115	0.60		Shagin et al., 2004
zFP538	528	538	ND	ND	DsRed-related	Matz et al., 1999; Remington et al., 2005
mBanana	540	553	6	0.70		Shaner et al., 2004
mKO	548	561	51.6	0.60		Karasawa et al., 2004
mHoneydew[d]	504	562	17	0.12		Shaner et al., 2004
mOrange	548	562	71	0.69	DsRed-related	Shaner et al., 2004; Shu et al., 2006
z2FP574	552	574	ND	ND	DsRed-related	Labas et al., 2002; Pakhomov and Martynov, 2007
DendFP (red)[c]	557	575	35	0.68	Kaede	Pakhomov et al., 2004; Gurskaya et al., 2006
dTomato	554	581	69	0.69		Shaner et al., 2004
EosFP (red)[c]	571	581	41	0.55	Kaede	Wiedenmann et al., 2004; Nienhaus et al., 2005
Kaede (red)[c]	572	582	60.4	0.33	Kaede	Ando et al., 2002; Mizuno et al., 2003
DsRed	558	583	75	0.79	DsRed	Matz et al.,1999; Gross et al.,2000; Shaner et al.,2004
mTangerine	568	585	38	0.30		Shaner et al., 2004
asFP595 (KFP)	568	595	56.2	<0.001	DsRed-related	Lukyanov et al., 2000; Tretyakova et al., 2007
mStrawberry	574	596	90	0.29	DsRed	Shaner et al., 2004; Shu et al., 2006
mCherry	587	610	72	0.22	DsRed	Shaner et al., 2004; Shu et al., 2006
eqFP611	559	611	78	0.45	DsRed	Wiedenmann et al., 2002; Petersen et al., 2003
gtCP[e]	589	615	ND	chromo	DsRed	Gurskaya et al., 2001; Martynov et al., 2003
Rtms5[e]	592	620	80	chromo	DsRed	Prescott et al., 2003
cgCP[e]	587	622	ND	chromo	DsRed	Gurskaya et al., 2001; Pakhomov et al., 2006
mRaspberry	598	625	86	0.15		Wang et al., 2004
Katushka	588	635	65	0.34		Shcherbo et al., 2007
mPlum	590	649	41	0.10		Wang et al., 2004
aeCP-AQ143	595	655	90	0.04		Shkrob et al., 2005
cjBlue	610	ND	66.7	chromo	DsRed	Chan et al., 2006

Figure 1.6 - Spectral characteristics of wild type GFP and its mutants (Pakhomov & Martynov, 2008)

1.5 Applications of FPs

FPs are used to label cell organelles and proteins of interest as genetically encoded tracing molecules. They are also applicable in studying protein-protein interactions on the basis of Förster resonance energy transfer (FRET).FPs have been used in protein tracking by fusing them with the protein of interest to know about their functions on the basis of their location inside the cell. FPs are also used to label sub-cellular compartments to help in the identification of proteins limited to those compartments as well as the analysis of the compartments themselves. FPs can also be used to track embryogenesis and also the cell movement and progression during tumor development. FPs find use as sensors inside the cells as the fluorescence emission according to changes in the environment of the cell, such as pH or membrane voltage. FPs can also be used in Super Resolution Microscopy where they assist to image proteins with a higher resolution far better than conventional microscopy (Jung, 2012).

1.6 Techniques used to clone and characterize the FPs

In this thesis different features of FPs will be characterized, including expression, maturation and light absorption and fluorescence emission.

The Gibson cloning technique, which is an *in vitro* recombination system, was developed by Dr. Daniel Gibson in 2009. This technique has the ability to join and repair overlapping DNA molecules in single isothermal reaction. This reaction involves the activity of three enzymes that contribute in joining the DNA fragments to form a circular molecule. The first enzyme is T5 exonuclease and, removes nucleotides at 5` end of the double stranded DNA molecules. This leads to annealing of the complementary single stranded DNA overhangs. Then the gaps are filled by Phusion polymerase and the nicks are sealed by Taq DNA ligase (see figure 1.7).

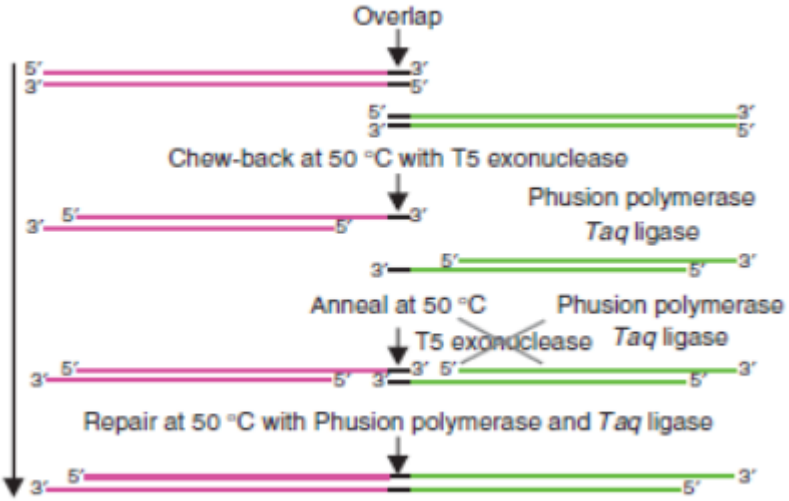


Figure 1.7 - An overview of the Gibson cloning showing the functions of the three enzymes involved: T5 exonuclease digests the 5`end of the double stranded DNA fragment, Phusion polymerase fills the gaps and T4 DNA ligase seals the nicks (Gibson et al., 2009).

The Gibson cloning system offers advantages over traditional cloning as large fragments can be combined in one step without depending on the availability of restriction sites (Gibson et al., 2009).

1.6.1 BioBricks

Biobricks were first introduced by Knight in 2003, that are standard biological parts enabling the design and construction of synthetic biological systems (Knight, 2003). The biobricks are designed to be incorporated in cells like *E.coli* to produce new biological systems. Each biobrick consists of a circular, double stranded vector with precisely defined flanking sequences called prefix and suffix located upstream and downstream respectively. The upstream sequence consists of EcoR1 and Xba1 restriction sites whereas; the downstream sequence consists of Spe1 and Pst1 restriction sites. A new biobrick with improved features can be designed if a fragment from one biobrick is assembling with the other, the latter being used as a backbone. This can be achieved by digesting one biobrick containing the insert with EcoR1 and Pst1 while the other with Spe1 and Pst1 which is the destination plasmid (Shetty,

Endy, & Jr, 2008). The destination plasmid and the insert are then ligated together (see figure 1.8).

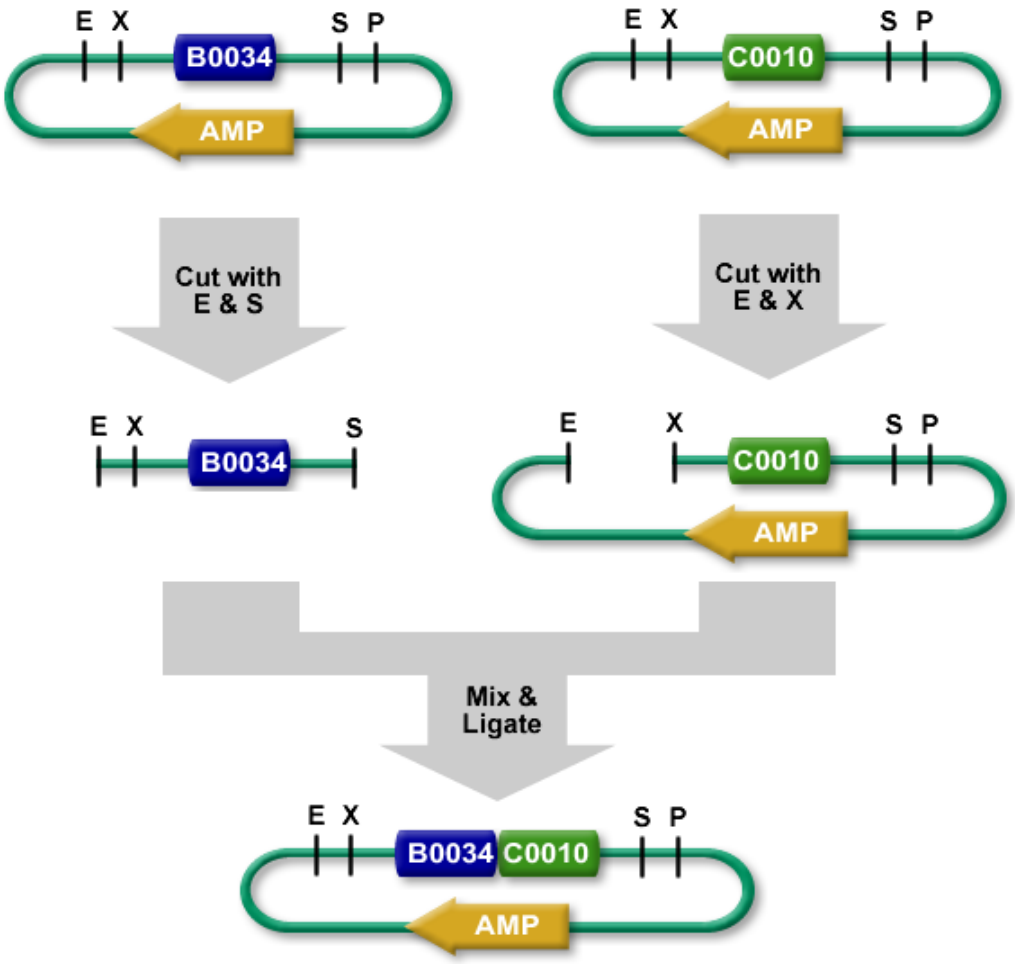


Figure 1.8 - An overview of the Biobrick cloning. The BioBrick has flanking restriction enzymes EcoRI and XbaI on left while SpeI and PstI on right. The BioBrick with C0010 (green) is the destination plasmid while B0034 (blue) is the fragment in BioBrick to be cut and inserted in the destination plasmid. Cutting the destination plasmid with EcoRI and XbaI creates a gap for the fragment that is cut with EcoRI and SpeI. The E sticky ends come together with E sticky ends and S sticky ends come together with X sticky ends of the destination plasmid (http://igem.org).

These Biobricks are the DNA sequences submitted to the Registry of standard biological parts (http://partsregistry.org/) by the students participating in The International Genetically Engineered Machine (iGEM) competition (http://igem.org). BioBricks are identified by a

specific code with a number preceded by Bba_. These parts may be classified as basic or composite parts and should have a well defined biological function. A nucletide sequence submitted directly is called the basic part while a combination of BioBrick parts in a defined order is called the composite part.

The iGEM Foundation is devoted to education, and advancement of synthetic biology, and the development of open community and collaboration.

1.6.2 Spectroscopy

FPs play important roles as biological markers but their optimization is totally dependent on understanding their molecular spectroscopy.

Luminescence is a phenomenon observed in light emitting species for example FPs which is the emission of light by a molecule in its excited state. Fluorescence is a type of Photoluminescence which is produced when a substance absorbs light of high energy (at a shorter wavelength) and emits a light of lower energy (at higher wavelength). The process of emission and absorption can be well understood by a diagram first presented by Aleksander Jablonski in 1930 (see figure 1.9), which shows that fluorescence is a three step process. In first step, a molecule is excited by absorbing photon of energy ($h\nu_{EX)}$. The molecule at this stage is in excited electronic singlet state S_1`. The excited molecule can lose its excitation by emitting a photon with an energy lower than the absorbed photon ($h\nu_{Em}$), thereby returning to its ground state (S_0) (Lichtman & Conchello, 2005).

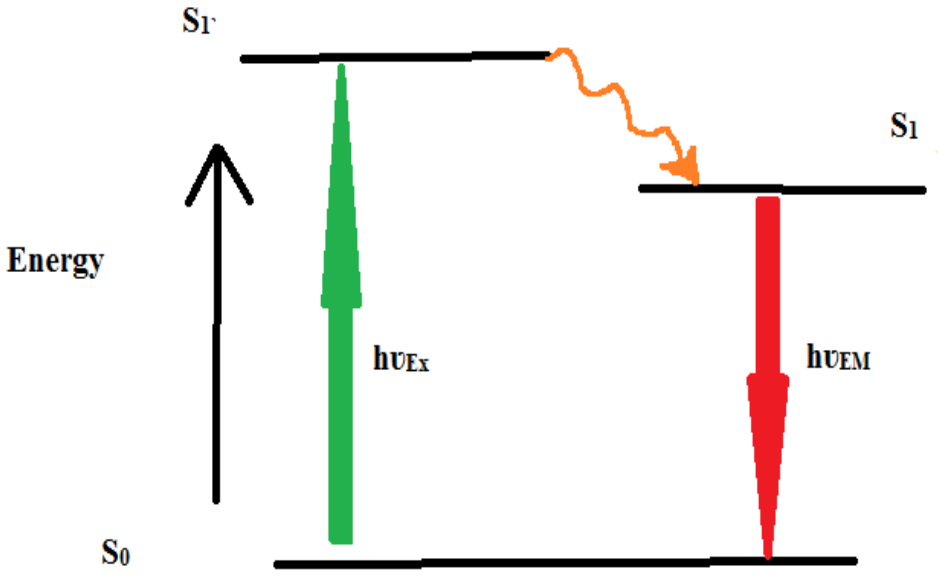


Figure 1.9 - The three steps of fluorescence described by Aleksander Jablonski. Energy ($h\upsilon_{EX}$) is absorbed by the molecule in ground state (S_0) and attains its excited state of the molecule (S_1`). Enegry ($h\upsilon_{EM}$) is emiited by this molecule and reaches the ground state.

Every fluorescent molecule has two characteristic fluorescence spectra: (1) The fluorescence excitation spectrum, which visualizes which photons cause fluorescence emission at a certain wavelength, (2) The emission spectrum which provides the relative intensity of radiations emitted when a fluorophore has been excited at a certain wavelength. The excitation spectrum measured by a fluorimeter is usually similar to absorption spectrum measured by a spectrophotometer (Lakowicz, 2007). The spectrophotometer used for absorption measurement in this study was U-3010 Hitachi spectrophotometer and the fluorimeter was Hobiba-Yvon Spex Fluorolog fluorometer.

1.6.3 Microplate Reader

Expression of FPs in prokaryotic organisms was carried out in 1994 for the first time. The expression of a fluorescent protein is self-directed and independent of cell type and location. Expression of the FPs is carried out by integrating the gene in an expression plasmid. The

expression can easily be captured by a visible signal in the form of fluorescence (Southward & Surette, 2002). Several instruments are capable of detecting the fluorescence signal, such as spectro(photo)meters, flow cytometers and automated micro plate readers. Plate readers enable to assess the presence of a mature FP in a cell population (Jong, Ranquet, Ropers, Pinel, & Geiselmann, 2010). Experiments can be performed to determine the maturation and the stability of FPs.

Microplate readers are very powerful instruments to acquire large amounts of data quickly, due to the employed multi-sample approach. The samples are loaded in the wells of the microplate and inserted in the instrument. The light source used is xenon flash lamp. The excitation wavelengths are selected as desired by using a monochromator which is transmitted to a mirror with a hole. The excitation light strikes the sample and the fluorescence emitted by the sample is transferred to the detector by the same mirror. In this way fluorescence emission can be recorded in many samples over extended periods of time (Lakowicz, 2007).

Microplate reader TECAN n200 pro was used in this study to characterize the expression, maturation and stability of FPs.

1.7 Aim of the project

The aim of this project is to generate FPs with novel spectral features by genetically combining two FPs. FPs have their own specific emission and excitation spectra but energy transfer between the fused FPs may result in novel spectral characteristics.

The coding sequences of single FPs and fused FP pairs will be inserted into an expression vector. These proteins will be characterized using spectroscopic techniques.

FPs both individual as well the fused variants will be expressed in *E.coli* cells to determine protein expression, maturation and stability This study aims to provide information about the FPs regarding their expression, stability, maturation and spectral properties.

2 Material and Methods

Work with microorganisms was carried out under standard microbial condition in a laminar flow hood. Deionized water was used for the preparation of solutions and autoclaved at 121 °C for 20 min. DNA samples, buffers and enzymes were stored at -20 °C.

2.1 Bacterial strain

Bacterial strain E.coli DH5α was used for transformation purposes. It was grown on LB agar plates with specific antibiotics or LB agar for overnight cultures. The growth was maintained at 37 °C.

2.2 Fluorescent Proteins

FPs used in this study are GFPmut3b, mRFP1 and EYFP that were obtained from the BioBrick registry with codes BBa_E0040, BBa_E1010 and BBA_E0030 respectively.

The Expression vector was also obtained from the registry with the code BBa_J04500. This Vector pSB1AK3 is 3409 bp long and has LacI promoter with ribosome binding site. LacI is IPTG inducible. The origin of replication pMB1 which pUC19 derived and has a high copy number (100-300 copies/cell). AmpR and KanR are the marker sites. The plasmid map is shown in the Figure 2.1.

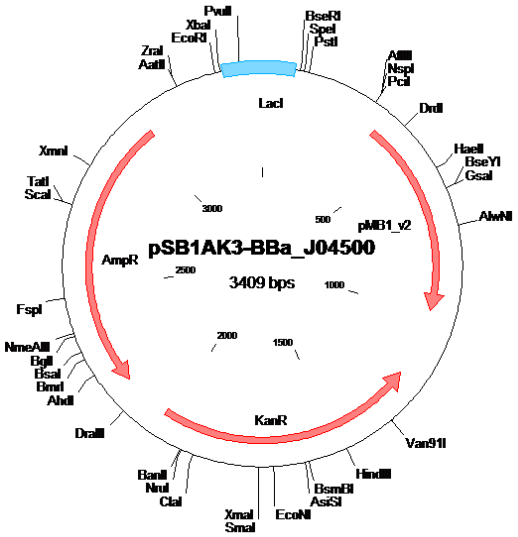


Figure 2.1 - Plamid Map of pSb1Ak3 with LacI IPTG inducible promoter marked as blue. The origin of replication (pMB1) and marker genes (kanR and AmpR) are shown as arrows marked pink.

2.3 Growth Media and stock solutions

In order to grow E.coli on the medium specific antibiotics are required to be added to the medium. So, stock solutions for the antibiotics were prepared first as shown in Table 2.1. Two types of media were used to the growth of E.coli i.e Lysogeny Broth (LB), both liquid and agar and other was SOC. The compositions for both the media are shown in Table 2.2.

Table 2.1 - Antibiotic Stocks solutions

Antibiotics	Working solution	Final soltuon	Solvent	Storage
Kanamycin	50 mg/ml	50 µg/ml	Water, filter sterilize	-20 °C
Ampicillin	50 mg/ml	50 µg/ml	Water, filter sterilize	-20 °C
Chloramphenicol	30 mg/ml	50 µg/ml	Ethanol	-20 °C

Table 2.2 - Growth Medium: Compositions of Lysogeny Broth, agar and SOC medium used for the growth of E.coli

Stock	Ingredients	Amount
LB (1L)	Tryptone	10 g
	Yeast extract	10 g
	NaCl	5 g
	Agar (For plates)	20 g
SOC solution (1L)	Bateriotryptone	20 g
	Yeast extract	5 g
	NaCl	0.584 g
	KCl	0.186 g

Water was added till the total volume is 1000 ml and autoclaved at 121 °C for 20 minutes.

The SOC solution after being autoclaved can be stored at 4 °C, but before using for transformation some molar solutions of $MgCl_2$, Glucose and $MgSO_4$ should be added to it to increase the transformation efficiency. The molar solutions are given in Table 2.3.

Table 2.3 - Solutions prepared for SOC medium

Solution	Working solution	Final soltuon	Storage
Glucose	1 M	0.05 mM	Autoclave and store at 4 °C
$MgCl_2$	1 M	0.1 mM	
$MgSO_4$	1 M	0.1 mM	

2.4 Competent cells preparation

Competent cells of E.coli DH5α were prepared. Before the competent cells preparation, some reagents are required that need to be prepared prior to competent cells preparation. The reagents required are given in Table 2.4.

Table 2.4 - Reagents for the preparation of E.coli DH5α competent cells

Reagent	Ingredients	Amount	Preparation
YB medium (500 ml)	Bactotryptone	10 g	
	Yeast extract	2.5 g	
	KCl	0.38 g	pH adjusted to 7.6 and Autoclaved
	$MgSO_4$ (sterile)	17 ml	
TfB1 Solution (500 ml)	Potassium Acetate	1.47 g	
	$MnCl_2$	4.95 g	
	RbCl	6.05 g	
	$CaCl_2$	0.74 g	
	Glycerol	15 ml	pH adjusted to 5.8 with 0.2 M acetic acid. Filter sterilized and stored at 4°C
TfBII solution (100 ml)	MOPs (100 mM)	10 ml	pH 7.0 with 0.2-1 M KOH
	$CaCl_2$	1.10 g	
	RbCl	0.12 g	
	Glycerol	15 ml	Autoclaved and stored at 4 °C.
SOC solution (1L)	Bateriotryptone	20 g	
	Yeast extract	5 g	
	NaCl	0.584 g	
	KCl	0.186 g	Autoclaved and 10 ml (1 M) $MgSO_4$ and 20 ml (1 M) Glucose were added afterwards.

A single colony of E.coli DH5α was inoculated in 20 ml of SOC medium and allowed to grow overnight in a shaking incubator at 37 °C. This culture was diluted by transferring to 300 mL pre-heated YB media and incubated till OD_{600} was 0.3-0.4. The cells were the incubated on ice for 5 min and transferred to pre-chilled Falcon tubes (15 mL). They were centrifuged at 4000 rpm at 4 °C for 10 min. Cells were resuspended in 15 mL chilled TfBI solution and immediately centrifuged at 4000 rpm at 4 °C for 10 min. The pellets were resuspended in 1 mL chilled TfBII soultion. Competent cells were frozen with dry ice immersed in ethanol and 100 µL aliquots were prepared and stored at -80 °C.

2.5 Transformation of the Competent cells (Heat shock method)

The transformation efficiency of the competent cells prepared need to be calculated by the transformation of the cells. The plasmid pUC-19 (100 pg/µL) with Amp resistance was used for transformation. Plasmid and the competent cells were thawed on ice for 10 min. Plamid (1 µL) was added to the competent cells and incubated on ice for 30 min. Cells were subjected to a heat shock at 42 °C for 42 sec and held on ice for 3 min. 900 µL of LB was added to the cells and incubated in a shaking incubator for 60 min at 37 °C. Cell culture (50 µL) was plated out on LB plates with ampicillin. The remaining culture was centrifuged at 8000 rpm for 60 sec and the pellet was resuspended in 100 µL LB and plated all of it on ampicillin plate. The plates were incubated overnight at 37 °C.

The transformation efficiency was calculated using the formula below:

$$\text{Transformation efficiency} = \frac{No.of\ colonies}{conc.of\ the\ plasmid} \times \frac{106\ pg}{\mu g} \times \frac{Volume\ of\ transformants}{volume\ of\ cells\ plated} \times dilution\ factor$$

The transformation efficiency of the cells turned out to be $8 \times 10^{7.}$

2.6 Transformation of the FPs

The Coding sequences for the FPs and the promoter were in the BioBricks and their location was found by using the iGEM registory (http://parts.igem.org/Protein_coding_sequences/Reporters). The FPs with their names, BioBrick codes and location are given in Table 2.5.

Table 2.5 - FPs with their BioBrick codes

Fluorescent Protein	BB Code	Plasmid	Antibiotic resistance
Enhanced Yellow Fluorescent protein (EYFP)	BBa_E0030	pSB1AK3	Amp+Kan
Green Fluorescent protein (GFPmut3b)	BBa_E0040	pSB1A2	Amp
Red Fluorescent protein (mRFP1)	BBa_E1010	pSB1C3	Chl

These BioBricks were stored in iGEM plates in dried form in wells. To obtain the plasmid, 10 µL deionized water was added to the well with the desired BioBrick. The water on turning red is drawn out in an eppendorf tube. ABout 1-2 µL was used for transformation while the rest was stored at -20 °C.

The competent cells *E.coli* DH5α were transformed with these BioBricks as well as the Plasmid with the promoter and Ribosome binding site using heat shock transformation method.

The cells were plated out on the plates with their respective antibiotic plates and incubated overnight at 37 °C.

2.6.1 Colony PCR

The colonies appeared on the plates were confirmed for the BioBricks by colony PCR with BioBrick primers that are complementary to the prefix and suffix of the BioBricks. The sequences are given below:

BB_pre_fw gaattcgcggccgcttctagag

BB_suf_rv ctgcagcggccgctactagta

Colony PCR was performed using Taq DNA polymerase, and the annealing temperature for the primers was calculated using T_m calculator thermoscientific. The PCR reaction was prepared as shown in Table 2.6 using thermocycler conditions shown in Table 2.7.

Table 2.6 - PCR Reaction (50 µL) for Taq DNA polymerase

Reaction components	Volume	Final concentration
Sterile water	31 µL	_
10X Thermopol Reaction buffer	5 µL	1X
dNTPs (10 mM)	1 µL	200 µM
Primer Forward (10 mM)	1 µL	0.2 µM
Primer Reverse (10 mM)	1 µL	0.2 µM
Template DNA	variable	<1000 ng
Taq DNA polymerase	0.25 µL	1.25units/50 µL

Table 2.7 - Thermocycler conditions for Colony PCR using Taq DNA polymerase

Step	Temperature	Time
Initial denaturation	95 °C	30 sec
30 cycles	95 °C	15 sec
	53 °C	45 sec
	68 °C	1 min
Final Extension	68 °C	5 min
Hold	4-10 °C	Infinite

2.6.2 Gel Electrophoresis

The PCR products obtained after running the colony PCR were separated by Gel electrophoresis to see if the desired genes have been amplified.

Agarose gel 0.8 % (w/v) with 10000x Gel Green™ (Biotium, USA) was added in 1x TAE buffer (40 mM Tris, 20 mM acetic acid and 1 mM EDTA). The PCR products were added with 2 µL of 10x loading dye (0.25% bromophenol blue, 0.25% xylene cyanol FF and 30% glycerol) and a 5 µL DNA ladder was prepared for fragment length analysis by adding 3 µL sterile water, 1 µL loading dye and 1 µL DNA 1Kb GeneRuler™ DNA ladder (Thermo Scientific) in an eppendorf tube. The samples were loaded on the gel and run for 45 min at 85 Volts. The results were visualized on UV transilluminator and gel images were captured using a GelDoc (BioRad, USA).

2.6.3 Cell culturing

After the confirmation of the colonies being tranformed with the desired DNA, the cells were inoculated in a 5 mL LB with appropriate antibiotics and incubated overnight at 37 °C in a shaking incubator.

2.6.4 DNA purification

The DNA from the culture was purified by using Promega SV Wizard Miniprep Kit.

2.6.5 DNA Quantification

DNA purified from the culture was then quantified using NanoDrop.

2.7 BioBrick Cloning

The coding sequences for the fluorescent protein were incorporated in the expression vector to see their expression. This was done by removing the coding sequences from its plasmid with restriction digestion and ligating them in the expression vector.

2.7.1 Restriction digestion

The BioBricks with the coding sequences for the FPs were treated with XbaI and PstI, while expression vector was treated with SpeI and PstI. The buffer (Cutsmart) to be used was found by using NEB double digest finder. The Restriction digestion mixture was prepared by adding the following ingredients shown in Table 2.8.

Table 2.8 - Reaction Mixture for restriction digestion

Reaction components	Volume
Sterile water	16 µL
Buffer (Cutsmart)	2 µL
Insert (EYFP, GFPmut3B, mRFP1)	100 ng
Enzyme I (XbaI)	0.5 µL
Enzyme II (PstI)	0.5 µL

The same reaction mixture was prepared for the vector with enzymes mentioned above. Reaction mixtures were incubated at 37 °C for 1 hour. Then, the reaction mixtures with the insert were added with Calf Intestine Phosphatase (CIP) to prevent from religation. They were incubated again for 1 hour at 37 °C.

The restriction digests were purified using QIAquick® PCR purification kit (Qiagen, Germany) and quantified using Nanodrop.

2.7.2 Ligation

The inserts obtained were ligated with the vector in 3:1 ratio (Vector:Insert). The amount of the vector used was 5 µL and the volume of the inserts used was calculated using the formula given below:

$$\text{Insert (µL)} = \frac{\text{Vector (ng)} \times \text{sizeof the insert (bp)} \times \text{Vector(µL)}}{\text{Size of the vector (bp)} \times \text{insert(ng)}} \times 3$$

The Ligation mixture was prepared shown in Table 2.9.

Table 2.9 - Ligation Reactions.

Reaction components	LigationI	LigationII	LigationIII
Sterile water	6 µL	12 µL	11 µL
Buffer (T4 DNA ligase buffer)	2 µL	2 µL	2 µL
Insert	6 µL	-	6 µL
Vector	5 µL	5 µL	-
Ligase	1 µL	1 µL	1 µL

These ligation mixtures were incubated overnight at 16 °C. They were transformed in E.coli by heat shock method and plated out on plates with antibiotics appropriate for the vector. The plates were incubated at 37 °C overnight and observed the next day.

2.7.3 Confirmation of the ligation

The colonies were cultured in 5mL LB. 800 µL of the overnight cultures were added to 60 % gycerol and stored at -80 °C. The rest of the culture was purified for plasmid extraction by Miniprep and quantified. Restriction digestion was carried out using the enzymes XbaI ans SpeI with a cutsmart buffer. The restriction mixture was incubated at 37 °C for one hour. Loading dye (2 µL) was added to each tube and analysed by running on Gel electrophoresis at 85 V for 45 min.

2.8 Spectroscopy

The expression of the FPs (EYFP, GFPmut3b and mRFP1) by using spectroscopy to see their absorption and emission spectra. Before observing them, certain steps were performed.

2.8.1 Induction with IPTG

Overnight cultures of cells with FPs from the glycerol stock stored at -80 °C were diluted 1:100 and induced with 0.5 mM IPTG. These induced cultures were incubated at 37 °C for 3 hours in a shaking incubator. They were then prepared to be observed by spectroscopy by spinning down the cells with sterile water multiple time to wash away the medium.

2.8.2 *In vivo* Absorption

The absorption spectra for the FPs were determined using Hitachi U-3010 spectrophotometer with an integrating sphere (150 mm). Absorption spectra were run at a sampling interval of 1.00 nm and a scan speed of 300 nm/min to achieve a high resolution spectrum in the 400-800 nm regions. The spectra were normalized by subtracting the average of the tail (750-800 nm).

2.8.3 Fluorescence Emission

Fluorescence emission spectra were measured with a Triax-320 imaging spectrograph (Horiba). GFPmut3b, EYFP and mRFP1 were excited at 480 nm, 470 nm and 550 nm respectively and the fluorescence emission scan was detected from 500 to 650 nm.

2.9 Fluorescent protein expression measurements

E.coli DH5α with pSB1AK3 with fluorescent protein coding sequences ligated were inoculated from the glycerol stock frozen at -80 °C and incubated overnight at 37 °C in a shaking incubator.

2.9.1 Maturation of FPs

The overnight cultures were diluted 1 : 100 into LB with Kan and allowed to grow for 1 hour at 37 °C in a shaking incubator. For each of the culture OD_{600} as well as the fluorescence was measured. The culture were induced with 0.5 mM IPTG and incubated again for 1 hour. Each of the culture was sampled as 2 x 1 mL into 13 mL falcon tubes. For each culture, one was termed "control" and the other "inhibited" sample. CAP (30 mg/mL) was added to the inhibited sample to a final concentration 50 µg/mL. All of the samples were mixed rapidly by inverting the tubes 10 times. From each sample, 3 x 100 µL was transferred to separate wells on a Nunclon 96 Flat Bottom Black Polystyrol microtitre plate. Initial reading for OD_{600} and fluorescence was taken immediately after adding CAP. Fluorescence and OD were measured every 10 minutes for 2 hours.

2.10 Generating Variants of the FPs

The FPs under study were used to generate new variants by fusing two proteins together in an expression plasmid. They were combined together in way that the resulting insert has start codon of the first protein but no stop codon while no start codon for the second coding sequence but stop codon present. They were fused together with a linker that was 2 GGSGGS repeats. The proteins were fused together in the following combinations.

Start codon + GFPmut3b (Linker) EYFP + Stop codon

Start codon + EYFP (Linker) GFPmut3b + Stop codon

Start codon + GFPmut3b (Linker) mRFP1 + Stop codon

Start codon + mRFP1 (Linker) GFPmut3b + Stop codon

Start codon + mRFP1 (Linker) EYFP + Stop codon

Start codon + EYFP (Linker) mRFP1 + Stop codon

2.10.1 Plasmid Maps construction

The theoretical plasmid Maps for all the combinations were constructed using Clone manager and an online tool Vector Editor https://j5.jbei.org/VectorEditor/VectorEditor.html.

2.10.2 Gibson Cloning

The fragments to be assembled by Gibson cloning were needed to be amplified first. The primers were designed both for the backbone (pSB1AK3 used as expression vector) and for each of the coding sequence in such a way that they the forward primer of the first insert had overlap complementary to the backbone while the reverse had the linker as overhang complementary to the linker overhang of the forward primer of the second insert while its reverse primer had a complementary overlap with the backbone. The primers were designed using NEBuilder online tool. The primer sequences are given below.

Primers	Sequence (5`-3)	Length
pSB1AK3_rev	TATTTCTCCTCTTTCTCTAGTATGTG	26 bp
pSB1AK3_fwd	GTCCGGCAAAAAAACGGG	18 bp
yfp(1)_fwd	agagaaagaggagaaataATGGTGAGCAAGGGCGAG	36 bp
yfp(1)_rev	gccaccactaccgccCTTGTACAGCTCGTCCATGC	35 bp
gfp(1)_fwd	agagaaagaggagaaataATGCGTAAAGGAGAAGAAC	37 bp
gfp(1)_rev	gccaccactaccgccTTTGTATAGTTCATCCATGCC	36 bp
rfp(1)_fwd	agagaaagaggagaaataATGGCTTCCTCCGAAGACG	37 bp
rfp(1)_rev	gccaccactaccgccAGCACCGGTGGAGTGACG	33 bp
yfp(2)_rev	cgttttttgccggacTTATTACTTGTACAGCTCGTCCATG	41 bp
gfp(2)y_fwd	ctgtacaagggcggtagtggtggcagcggagggagtggcggtagcCGTAAAGGAGAAGAACTTTTC	66 bp
Gfp(2)_rev	cgttttttgccggacTTATTATTTGTATAGTTCATCCATG	41 bp
rfp(2)_rev	cgttttttgccggacTTATTAAGCACCGGTGGAG	35 bp
yfp(2)g_fwd	actatacaaaggcggtagtggtggcagcggagggagtggcggtagcGTGAGCAAGGGCGAGGAG	64 bp
yfp(2)r_fwd	ccggtgctggcggtagtggtggcagcggagggagtggcggtagcGTGAGCAAGGGCGAGGAG	65 bp
rfp(2)y_fwd	ctgtacaagggcggtagtggtggcagcggagggagtggcggtagcGCTTCCTCCGAAGACGTTATC	66 bp
gfp(2)r_fwd	ccggtgctggcggtagtggtggcagcggagggagtggcggtagcCGTAAAGGAGAAGAACTTTTC	65 bp
rfp(2)g_fwd	actatacaaaggcggtagtggtggcagcggagggagtggcggtagcGCTTCCTCCGAAGACGTTATC	65 bp

2.10.3 Polymerase Chain Reaction

All of the fragments were amplified with PCR using phusion polymerase. The PCR mixture was prepared and is given in Table 2.10.

Table 2.10 - PCR Reaction (50 µL) for phusion DNA polymerase.

Reaction components	**Volume**	**Final concentration**
Sterile water	To 50 µL	_
5X HF buffer	10 µL	1X
dNTPs (10 mM)	1 µL	200 µM
Primer Forward (10 mM)	2.5 µL	0.5 µM
Primer Reverse (10 mM)	2.5 µL	0.5 µM
DMSO	1.5 µL	3 %
Template DNA	variable	<1000 ng
Phusion polymerase	0.5 µL	1.0units/50 µL

Annealing temperature for the primers was calculated using Tm calculater NEB. The PCR condition used is shown in Table 2.11.

Table 2.11 - Thermocycler conditions dor PCR using Phusion DNA polymerase.

Step	Temperature	Time
Initial denaturation	98 °*C*	30 sec
30 cycles	98 °C	10 sec
	45-72 °C	10-30 sec
	72 °C	30 sec/Kb
Final Extension	72 °C	5 min
Hold	4-10 °C	Infinite

2.10.4 DpnI Digestion

The fragments amplified by PCR were digested with 2 µL DpnI to remove the methylated DNA that was not amplified. The sample was incubated at 37 °C for 1 hour.

2.10.6 PCR purification

PCR products after DpnI digestion were purified with QIAquick PCR purification kit (Qiagen, Düsseldorf Germany) to remove the unwanted enzymes and dNTPs.

2.10.7 Gibson Assembly

Before mixing the fragments together, the amount (in pmole) of each fragment to be added was calculated first. 100 ng of the Vector backbone was used, this amount was converted to pmoles using the following formula,

$$pmole = \frac{weight\ in\ ng \times 1000}{size\ of\ the\ fragment\ (bps) \times 650\ da}$$

According to the formula, the amount of vector was 0.05 pmole. The amount of fragments was also converted to pmole. Vector and fragments were added in a 1:3 ratio. The mixture for the Gibson assembly prepared is shown in Table 2.12.

Table 2.12 - Reaction Components for Gibson Assembly cloning

Reaction components	**Volume (2-3 fragments)**	**Positive control**
Sterile water	To 20 µL	To 20 µL
Gibson Assembly MasterMix	10 µL	10 µL
Fragment 1	0.02-0.5 pmol	-
Fragment 2	0.02-0.5 pmol	-
Vector Backbone	0.02-0.5 pmol	-
Positive control	-	10 µL

The mixtures were incuabted at 50 °C for 15 min.

2.10.8 Transformation

Gibson Assembly mix 2 µL was then transferred to *E.coli* by heat shock transformation. The cells were plated out on plates with respective antibiotics.

2.10.9 Diagnosis of Gibson cloning

To see if the Gibson cloning worked colony PCR with Taq polymerase was operated and the products were analyzed by running Gel electrophoresis.

2.10.10 Culturing

Colonies from the plates were cultured in LB overnight and used for plasmid extraction by Miniprep.

2.10.11 Restriction Digestion

The plasmids isolated were digested with XhoI to confirm if they have the desired insert.

2.10.12 Spectroscopy

The FPs were analyzed for their absorption and emission spectra through spectrophotometer.

2.10.13 Circular Polymerase Extension Cloning (CPEC)

CPEC based on (Quan & Tian, 2011) was another method used to clone all the three fragments together. It was performed by adding 100 ng of the linearized vector backbone with equimolar molar amounts of the FP fragments to a total volume of 25 µl. The assembly mixture was prepared in the following manner (table 2.13), amplified and visualized.

Tabel 1.13 - Reaction mixture for CPEC

Reaction components	**Volume**	**Final concentration**
Sterile water	To 25 µL	_
5X Phusion Reaction buffer	5 µL	5X
dNTPs (10 mM)	1 µL	200 µM
DMSO	0.75 µL	3 %
Vector backbone	Variable	<100 ng
Each fragment	Variable	Equimolar with backbone
Phusion polymerase	0.5 µL	1.0units/50 µL

3 Results

The aim of this project was to generate variants of FPs with altered optical characteristics. A strategy for obtaining new optical properties was to fuse two FPs together by a linker using molecular cloning techniques. Before cloning the two proteins together, the expression of FPs used in this study i.e. mRFP1, GFPmut3b and EYFP was analyzed by incorporating them in an expression vector (pSB1AK3) with a lacI promoter that could be induced by IPTG. The spectral features, maturation and stability of individual and fused FPs were characterized.

3.1 Transformation of the FPs

The FPs contained in the BioBricks plasmids were transformed into *E.coli* DH5α. The colonies obtained were confirmed to have the FPs by colony PCR and analysis of the PCR product on Gel electrophoresis as shown in figure 3.1.

Figure 3.1 - Gel Electrophoresis of the PCR products, obtained from colony PCR for transformants with FPs. DNA 1 Kb Marker, GFPmut3b, EYFP and mRFP1 (Starting from left to right).

3.2 Cloning of FPs

The FPs contained in the BioBricks plasmids were cut out with restriction enzymes XbaI and PstI, and ligated in pSB1AK3 the expression vector described in the material and methods section. The ligated products were transformed in *E.coli*.

The correct ligation was confirmed by digesting the plasmid with XbaI and PstI, resulting in an FP fragment of about 1000. These restriction enzyme digests were run on a gel and the results obtained are shown in the figure 3.2.

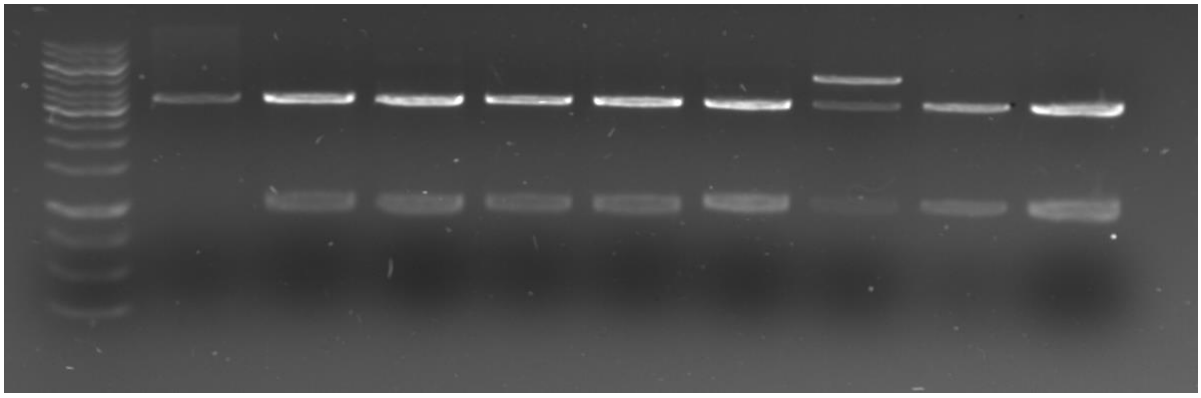


Figure 3.2 - Confirmation of the ligation after cloning FPs in expression vector by digesting with XbaI and PstI. Starting from right first is DNA ladder, positive control and restriction digests of the plasmids with FPs.

3.3 Spectroscopy

The overnight cultures of FPs were induced with 0.5 mM IPTG for 3 hours and washed with water before analyzing the cell pellets on a spectrophotometer. The induced FPs were compared with non-induced FPs. There was a slight difference between these two conditions as shown in figure 3.3. The induced samples showed more vivid coloration than the non-induced samples.

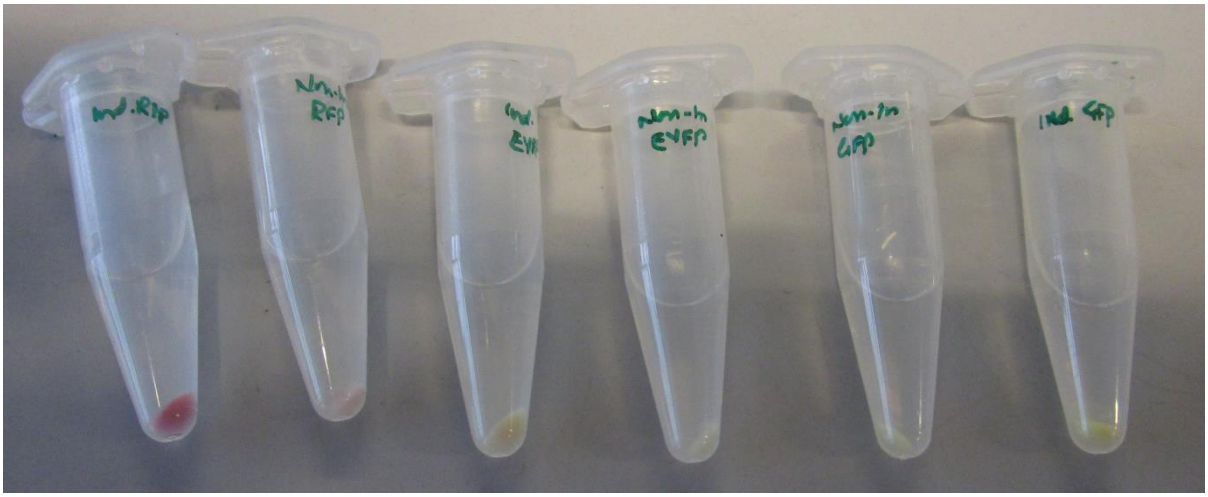


Figure 3.3 - Observation of FPs after induction with IPTG for 3 hours. From left Induced mRFP1, Non-induced mRFP1, Induced EYFP, Non-induced EYFP, Non-induced GFPmut3b and Induced GFPmut3b.

The absorption spectra for FPs (GFPmut3b, EYFP and mRFP1) were measured using the spectrophotometer. The absorption peak obtained for mRFP1 shown in figure 3.4 is similar to the one previously reported by Campbell and colleauges (Campbell et al., 2002). GFPmut3b showed an absorption peak consistent with previous literature (Cormack, Valdivia, & Falkow, 1996) and is shown in figure 3.5 while the absorption spectra of EYFP shown in figure 3.6 was also found similar to the one determined by Patterson and colleagues (Patterson, Day, & Piston, 2001).

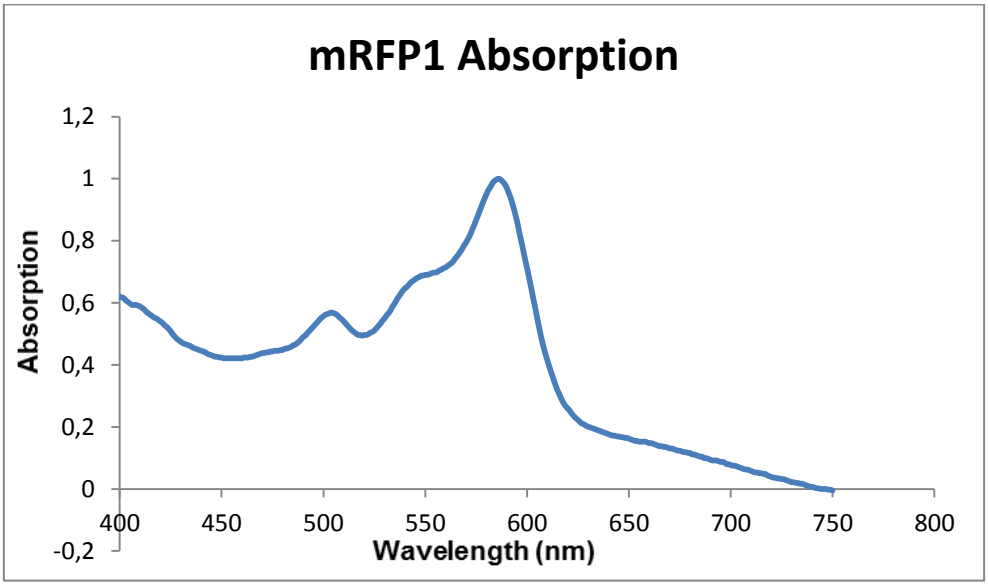


Figure 3.4 - Absorption spectrum of mRFP1 excited at 584 nm.

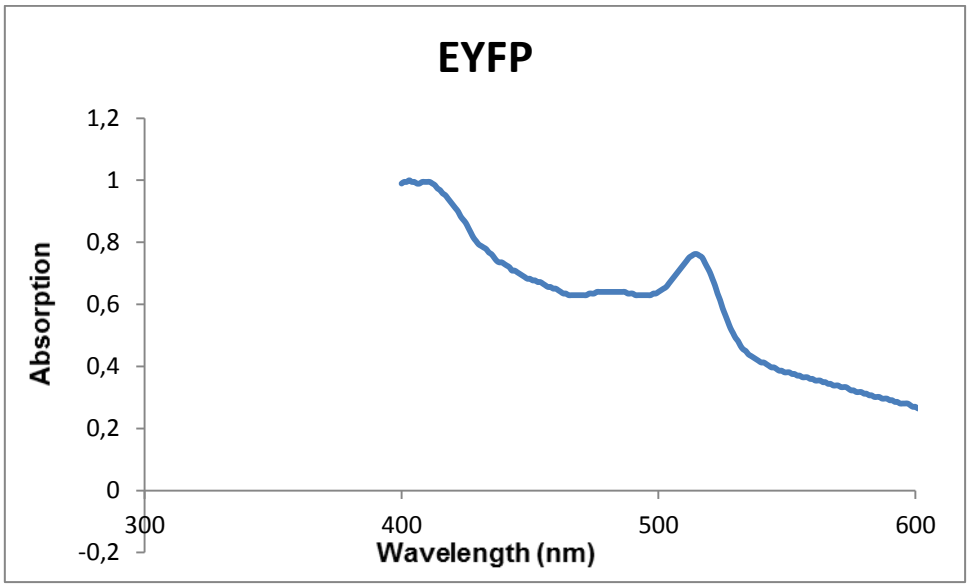


Figure 3.5 - Absorption Spectrum of EYFP excited at 514 nm

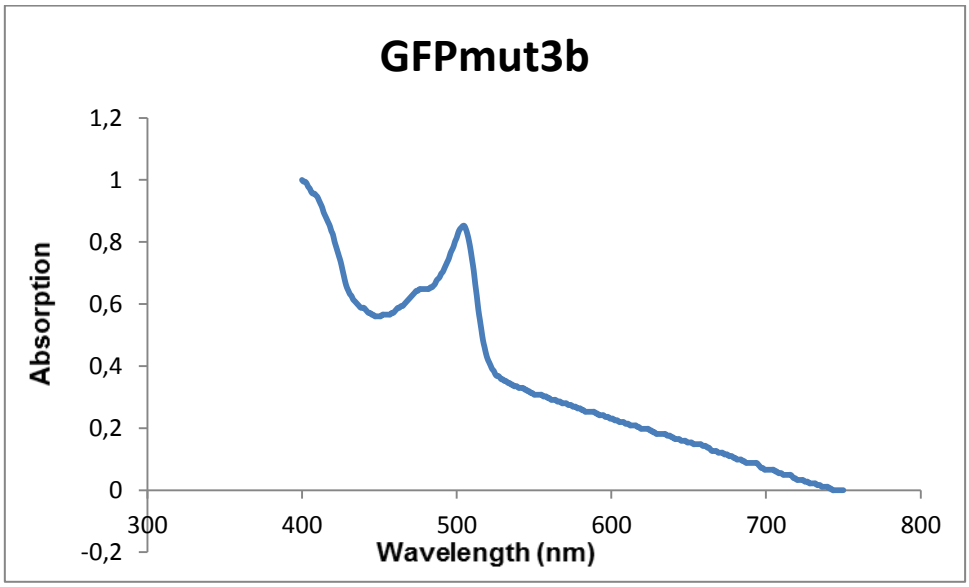


Figure 3.6 - Absorption spectrum of GFPmut3b excited at 501 nm.

The fluorescence emission spectra of these FPs were determined as well by exciting them at their respective peak absorption scanning from 450 nm to 700 nm. The emission spectrum for mRFP1 is shown in figure 3.7, for GFPmut3b in figure 3.8 while the emission spectrum for EYFP is shown in figure 3.9 respectively.

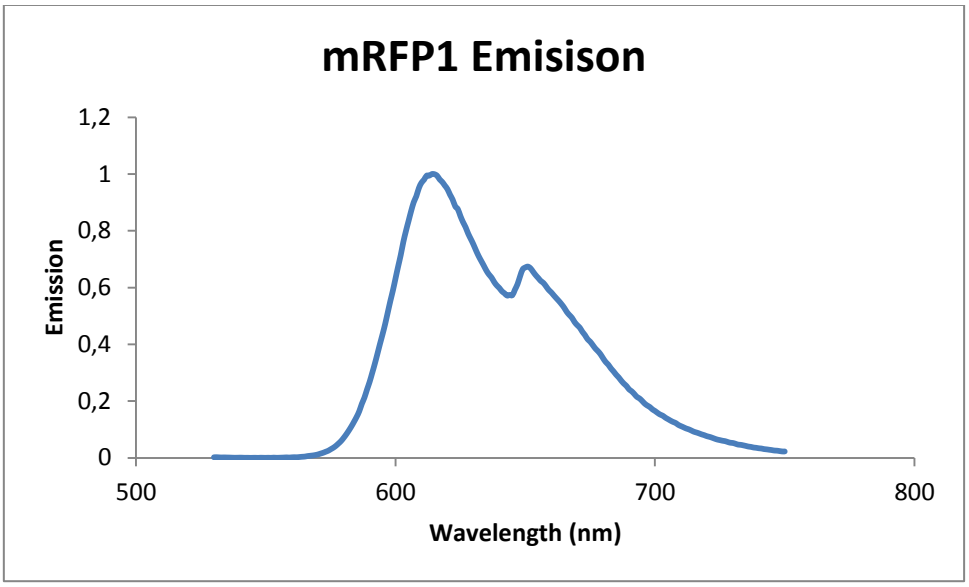


Figure 3.7 - Emission spectrum of mRFP1 excited at 580 nm scanning between 550 nm and 750 nm.

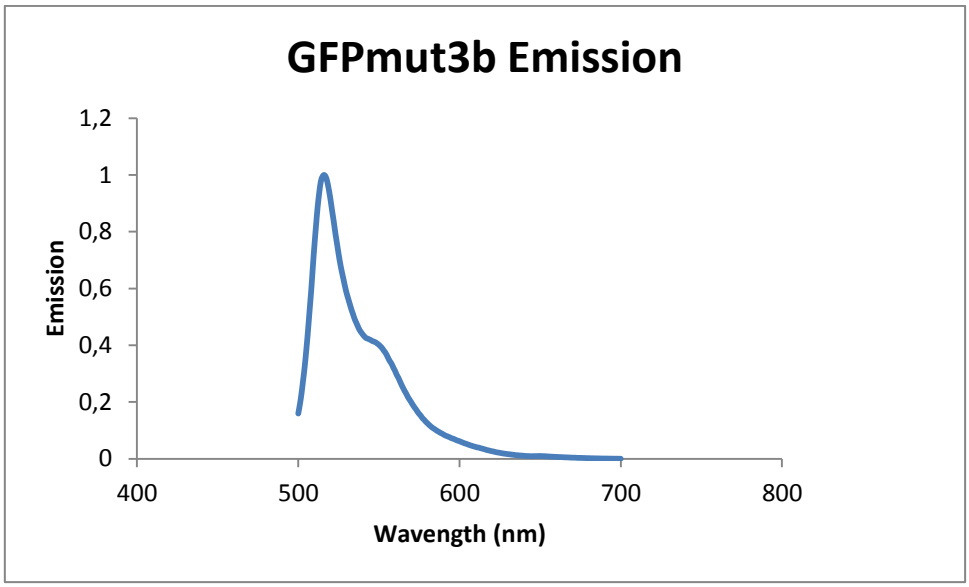


Figure 3.8 - Emission spectrum of GFPmut3b excited at 500 nm scanning between 550 nm and 750 nm.

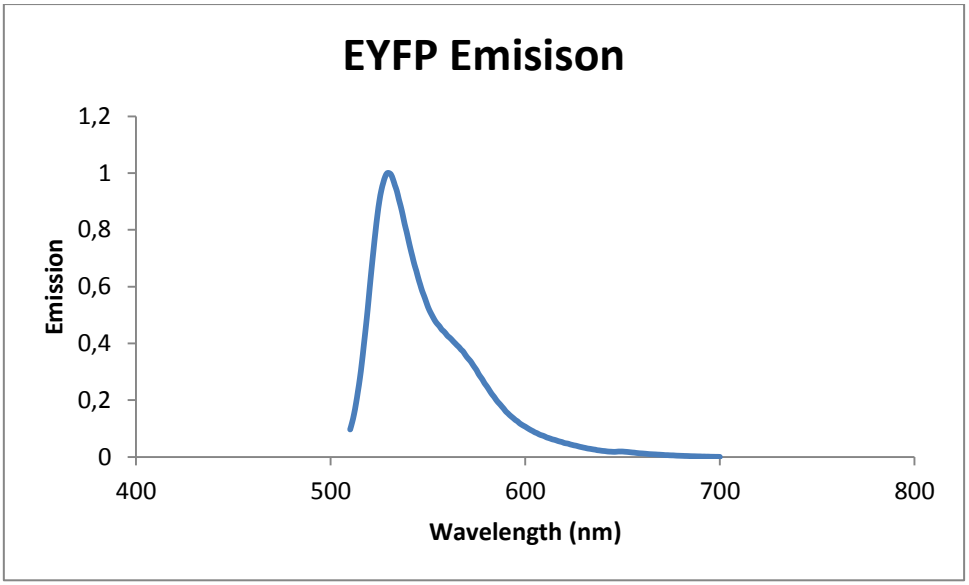
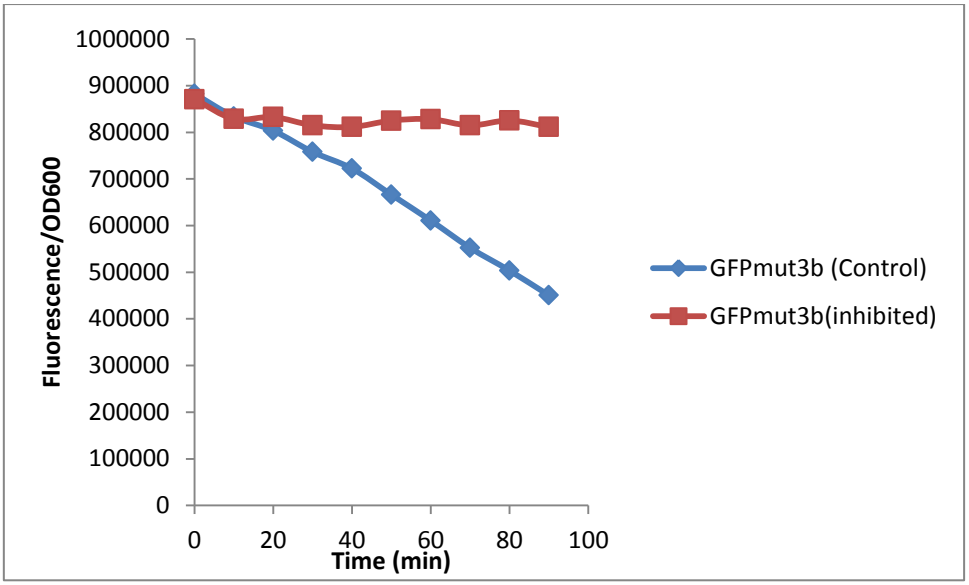
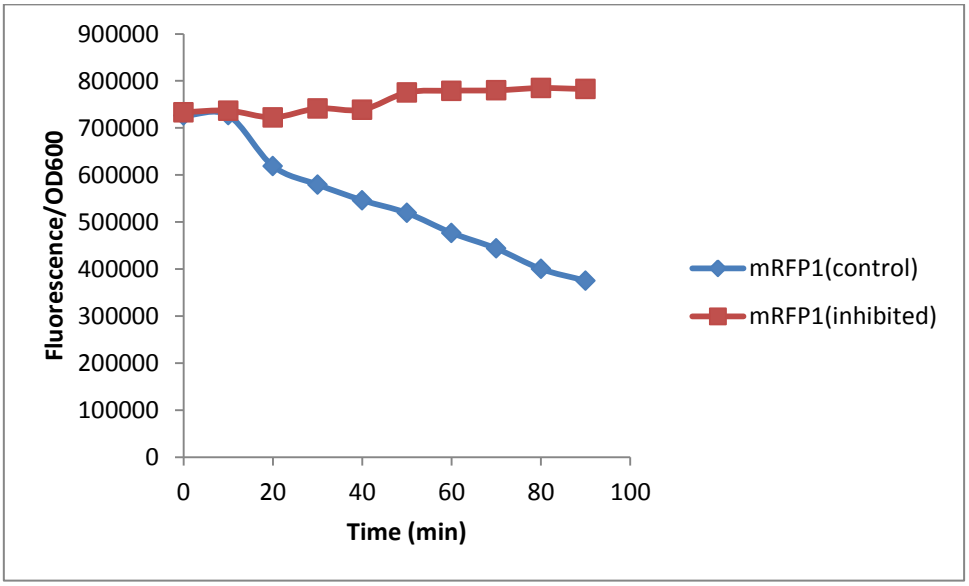


Figure 3.9 - Emission spectrum of EYFP excited at 514 nm scanning between 550 nm and 750 nm.

3.4 Maturation of FPs

The FPs maturation experiment was performed by terminating the production of proteins with and without the addition of chloramphenicol as described in the materials and methods section. Fluorescence and OD_{600} was measured every ten minutes for 3 hours after the addition of chloramphenicol to one sample. Fluorescence per cell for both samples decreased. The OD for the untreated control sample increased while it started to decrease 20 minutes after the addition of chloramphenicol. All of the three FPs showed this response pattern. Fluorescence divided by OD gave higher values for the inhibited sample compared to the control sample as shown in figure 3.10


1000000
900000
800000
700000
600000
500000
400000
300000
200000
100000
0
Fluorescence/OD600
0
20
40
60
80
100
Time (min)
GFPmut3b (Control)
GFPmut3b(inhibited)


900000
800000
700000
600000
500000
400000
300000
200000
100000
0
Fluorescence/OD600
0
20
40
60
80
100
Time (min)
mRFP1(control)
mRFP1(inhibited)

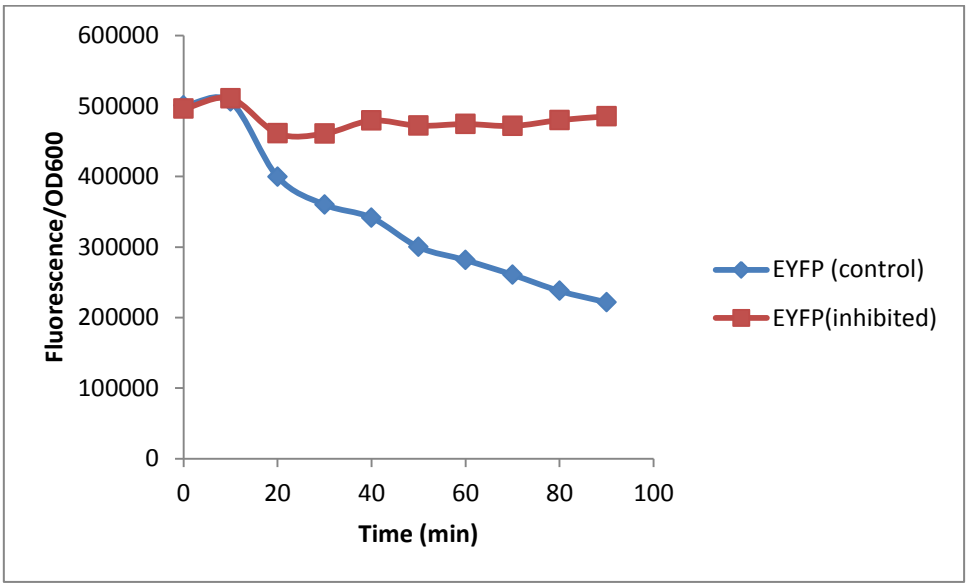

600000
500000
400000
300000
200000
100000
0
Fluorescence/OD600
0
20
40
60
80
100
Time (min)
EYFP (control)
EYFP(inhibited)

Figure 3.10 - Fluorescence divided by OD of the FPs after treatment with CAP (a) GFPmut3b (b) mRFP1 (c) EYFP which decreases in case of the control samples while it increases for the sample with CAP added.

3.5 Generation of Variants of FPs

The Variants of the FPs were designed using an online tool called Vector Editor. The Maps of the Protein variants are given in figure 3.11.

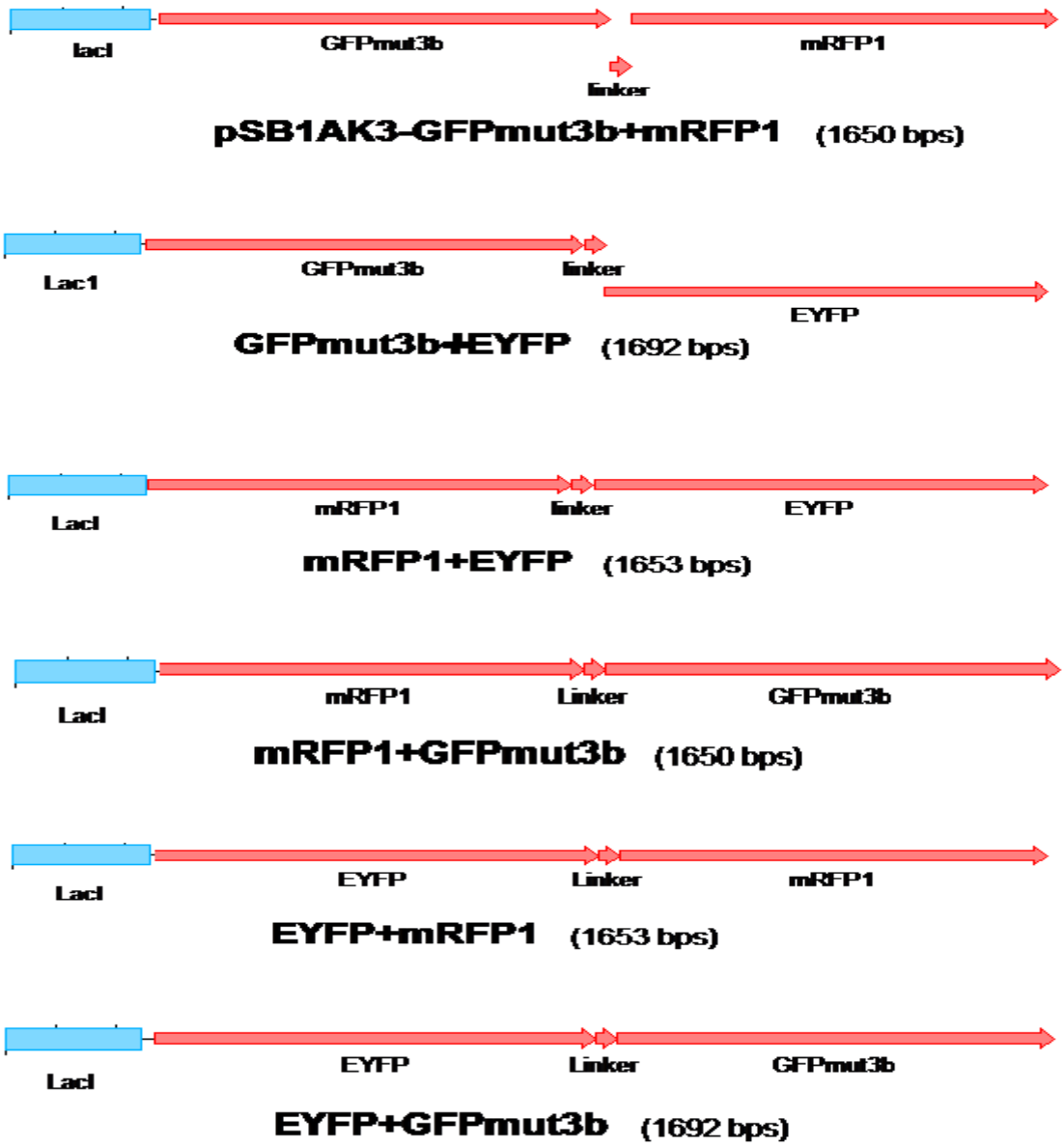


Figure 3.11 - Plasmid constructs of the variants of FPs containing the coding sequences of the two FPs in different combinations. The Promoter region (LacI) is marked blue while the two genes are shown by the two arrows (pink) with a linker in between them.

3.6 Gibson Cloning

In order to clone the fragments together, their PCR amplification was required. All coding sequences of the FP fragments as well as the vector backbone were amplified using the primers mentioned in the material and methods section. The fragments were amplified in such a way that the fragment of the first FP has overhangs complementary to the vector backbone on one side while being complementary to the overhangs of the second FP fragment, which also has an overhang complementary to the backbone. A diagram of the cloned fragments is shown in the figure 3.12.

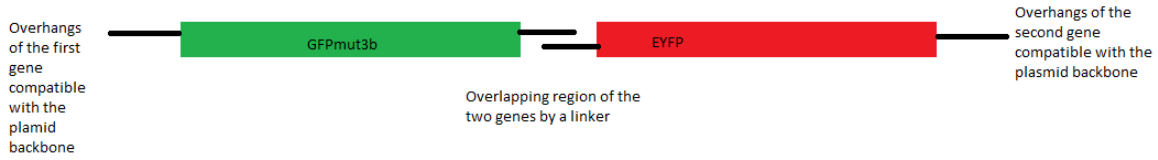


Figure 3.12 - FP variant with two genes fused with the overlapping sequences in between. The two genes have overhangs that are sequences that are complementary to the plasmid backbone.

The PCR amplified product for GFPmut3b was analyzed by gel electrophoresis. The expected band was about 750 bps. The figure 3.13 shows successful amplification of the GFPmut3b gene.

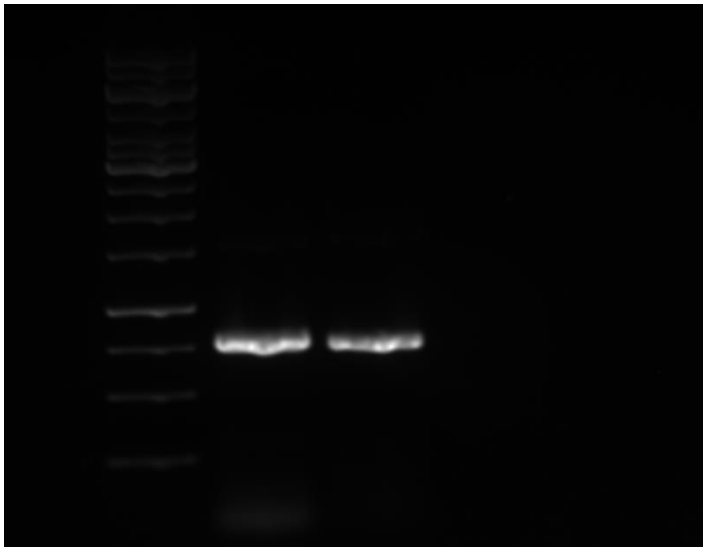


Figure 3.13 - The amplified product of GFPmut3b approximately 750 bp visualized on Gel. Starting from left is DNA 1 Kb ladder followed by the two GFPmut3b PCR products.

The PCR product of EYFP analyzed by gel electrophoresis. EYFP was successfully amplified by the primers used and is shown in the figure 3.14.

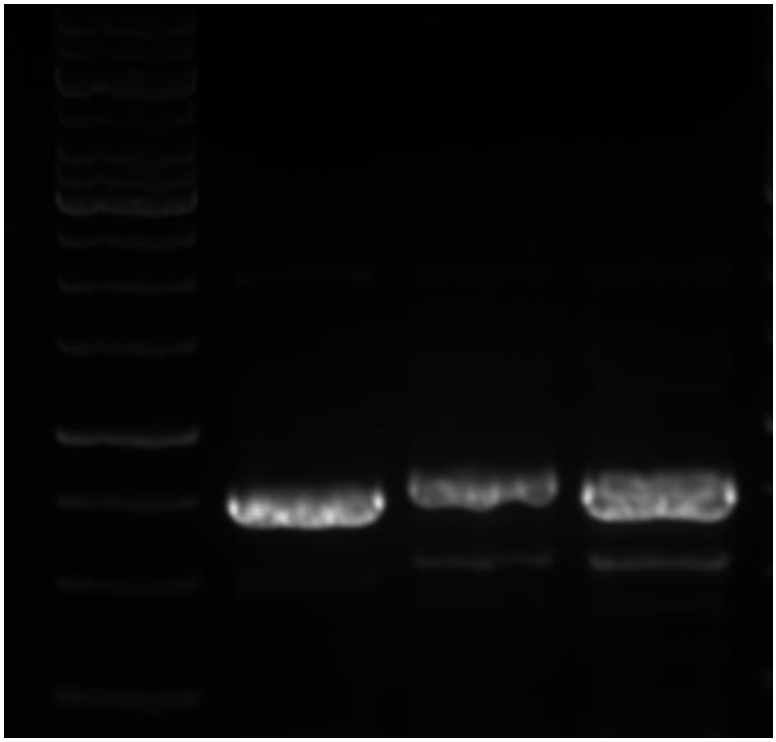


Figure 3.14 - PCR product of EYFP visualized on agarose gel. Starting from left DNA ladder 1 Kb followed by three EYFP PCR products of about 750 bp.

Figure 3.15 shows the PCR product of mRFP1 amplified successfully and run on an agarose gel.

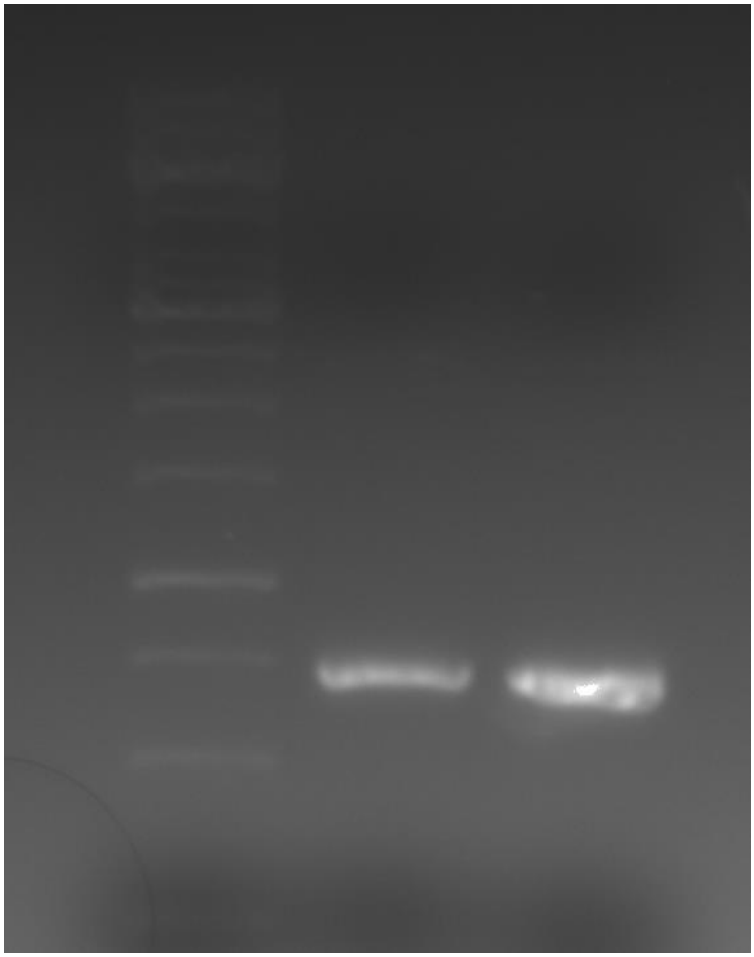


Figure 3.15 - PCR product of mRFP1 visualized on agarose gel. Left to right: DNA 1 Kb ladder, mRFP1 in the following 2 lanes.

The expression vector was also amplified by PCR and run on the Gel. The expected product was 3300 bps as shown in figure 3.16.

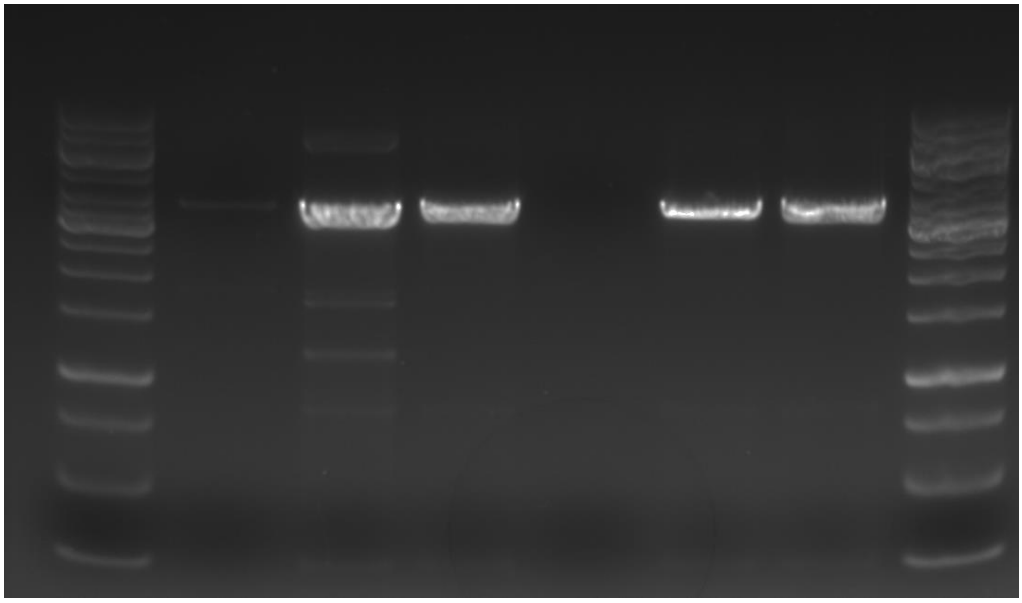


Figure 3.16 - PCR product of the plasmid pSB1AK3 of about 3300 bp visualized on an agarose gel. Left to right: DNA ladder 1 Kb, empty lane showing no amplified product, amplified product of the plasmid in lane 2,3,5 and 6.

The fragments were cloned together by Gibson cloning and transformed in *E.*coli. Both red and colorless colonies were observed as shown in figure 3.17.

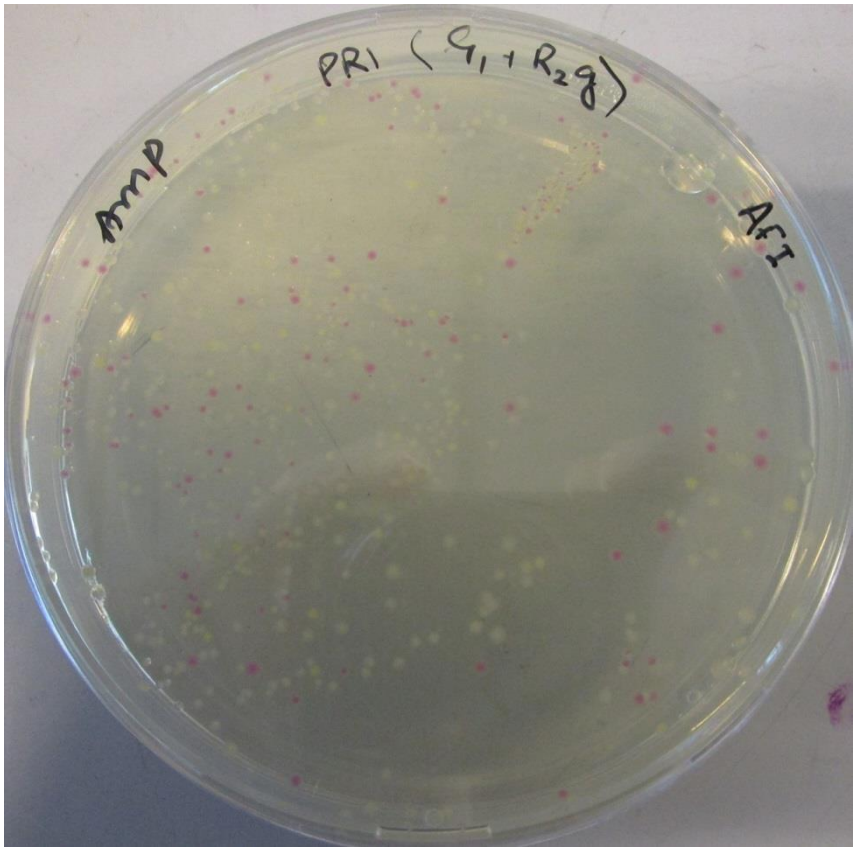


Figure - 3.17 - Colonies observed after transforming the Gibson cloning product in *E.coli* cells.

To obtain two genetically linked FP proteins the constructs were cloned into the vector backbone using Gibson cloning. In order to see if the Gibson was successful colony PCR was used, amplifying the desired insert using the forward primer of the first FP and reverse primer of the second FP. The result obtained is shown in the figure 3.18. As the two primers used for colony PCR screening were not optimized for amplifying the vector insert, a restriction digest of the constructed plasmid was performed

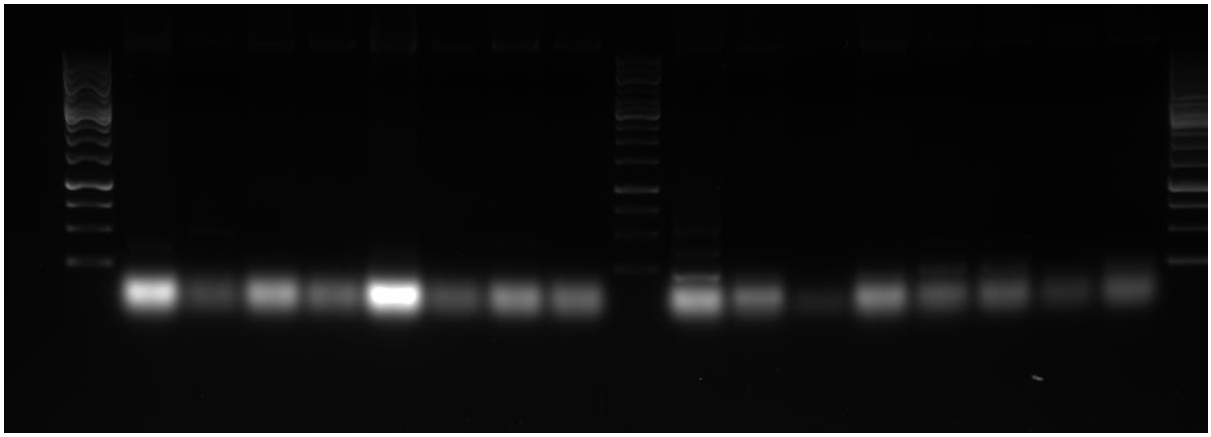


Figure 3.18 - Colony PCR after Gibson Assembly cloning. The Gel does not show any fragment amplified.

The restriction digest indicated that the construction of a fused FP was successful. The digest with enzyme XhoI resulted in two fragments one was 3500 bp while the other was approximately 1000bp (see figure 3.19). This restriction pattern indicates the presence of two fused FPs. It was therefore decided to try spectral analysis of the strains that potentially carry plasmids carrying the two genetically linked FPs.

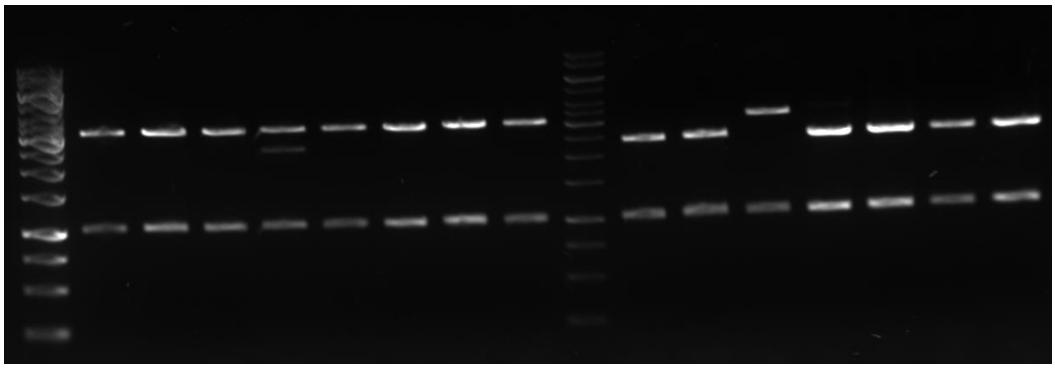


Figure 3.19 - Restriction Digestion with XhoI. Left to right: DNA ladder 1 Kb, GFP+RFP,RFP+GFP, GFP+YFP, YFP+GFP, YFP+RFP, RFP+YFP (lane 1-8), DNA ladder 1 Kb, Construct with single insert (lane 10.11). The constructs from lane 1 to 8 show two fragments of desired sizes when digested.

3.7 Spectroscopy

The absorption and fluorescence emission characteristics of individual FP and potentially fused FP were analyzed (Figure 3.20 and 3.21) by a spectrophotometer and a fluorimeter respectively. Spectra obtained for the protein variants with two FPs showed same features as single proteins; no new spectral features in the two FP variants were observed. Therefore, the plasmids were sent for sequencing and the sequencing results confirmed that the two inserts were not cloned. The plasmid contained either of the two insert.

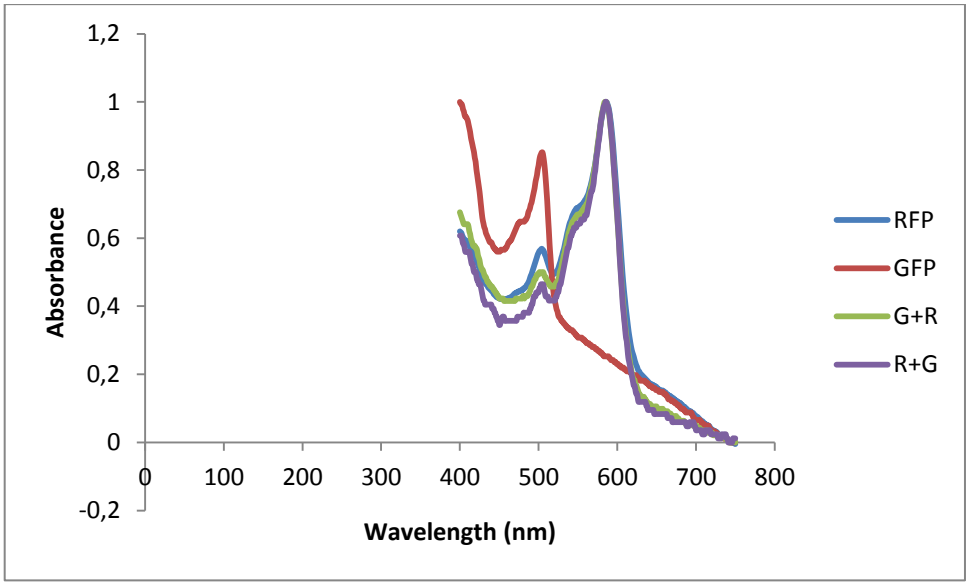


Figure 3.20 – Absorption Spectra of the FPs both with single FP genes and fused FP genes. mRFP1 (blue) and GFP (Red) are the constructs with single genes. G+R (green) and R+G (purple) are the constructs with the fused genes of both GFPmut3b and mRFP1. No difference is seen in the absorption spectra of G+R and R+G and RFP.

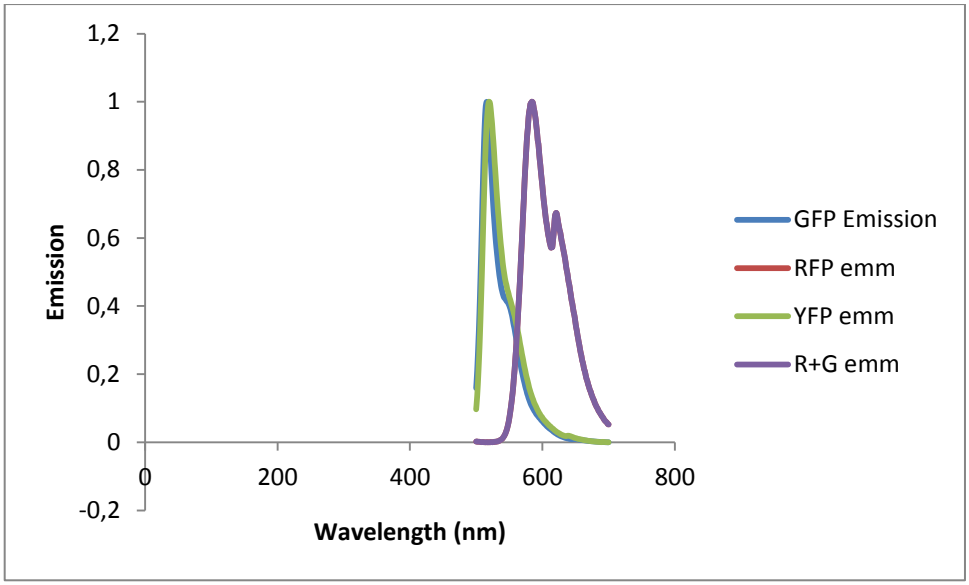


Figure 2.21 - Emission Spectra of the FPs both with single FP genes and fused FP genes. mRFP1 (blue) and GFP (Red) are the constructs with single genes. G+R (green) and R+G (purple) are the constructs with the fused genes of both GFPmut3b and mRFP1. No difference is seen in the absorption spectra of G+R and R+G and RFP.

The cloning procedure was repeated with modifications. In this case the two fragments of the FPs were first fused using the same Gibson Assembly protocol as described in the materials and methods but without the backbone. This product was then amplified by PCR using forward primer of the first fragment and reverse primer of the second fragment. The results obtained as shown in the figure 3.22 gave a desired fragment of approximately 1400 bp which confirmed that the two fragments are cloned together.

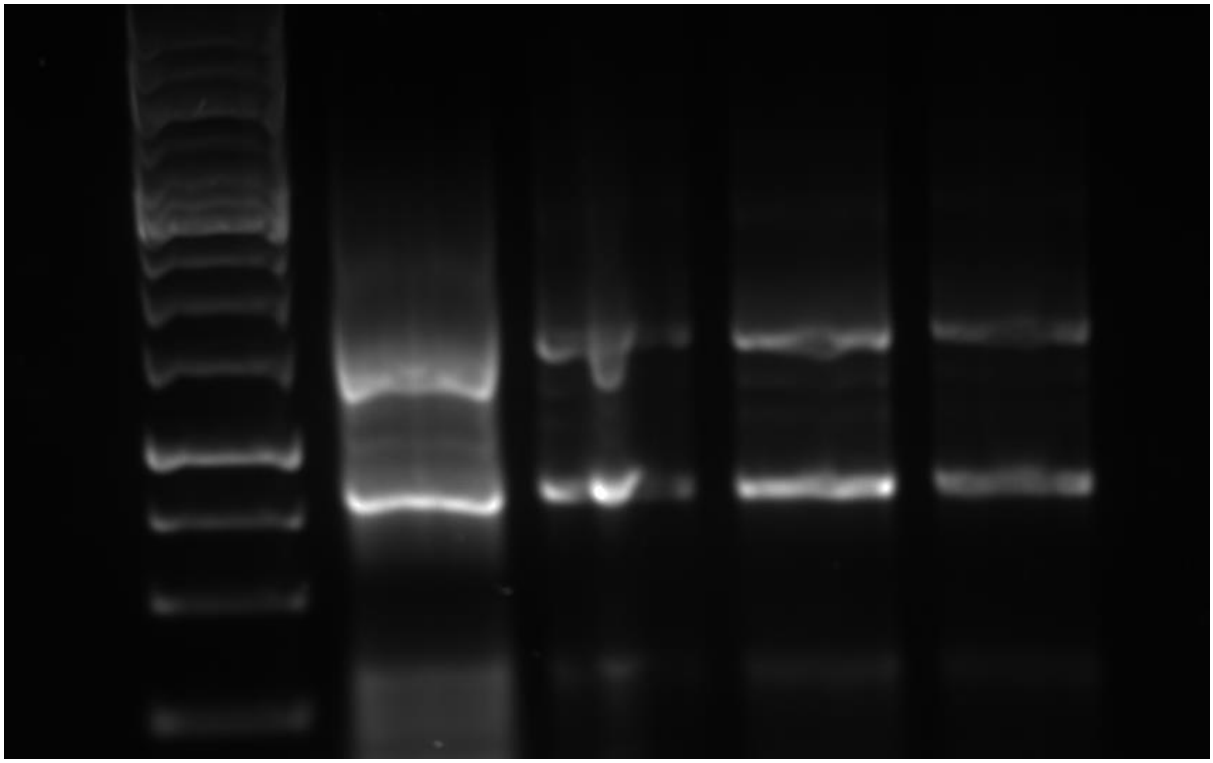


Figure 3.22 - PCR product of Green and Yellow FP cloned together by Gibson Assembly cloning. The two genes combined gave a fragment of about 1400 bp that can be seen in the figure.

These fragments were then cloned with the backbone using the Gibson Assembly protocol and transformed which lead to no growth on the plates. However, no colonies carrying the insert were obtained,

3.8 Circular Polymerase Extension Cloning (CPEC)

To obtain the desired genetically-linked FPs, the Gibson cloning procedure was repeated twice more; however the results obtained were not satisfactory. Therefore another cloning method called CPEC was used to clone the two FP fragments together. In this technique all of the fragments were amplified together by a PCR without using any primers. The backbone plasmid was amplified both with the fragments fused as a single fragment as well as with the individual fragments. The amplified product was run and visualized on agarose gel as shown in the figure 3.23. The vector backbone was used as a control to compare results and the agarose gel picture shows the largest band seen in the lane with all fragments cloned together was the same size as the plasmid backbone which means that the fragments have not been assembled. Hence, CPEC did not provide a successful cloning method.

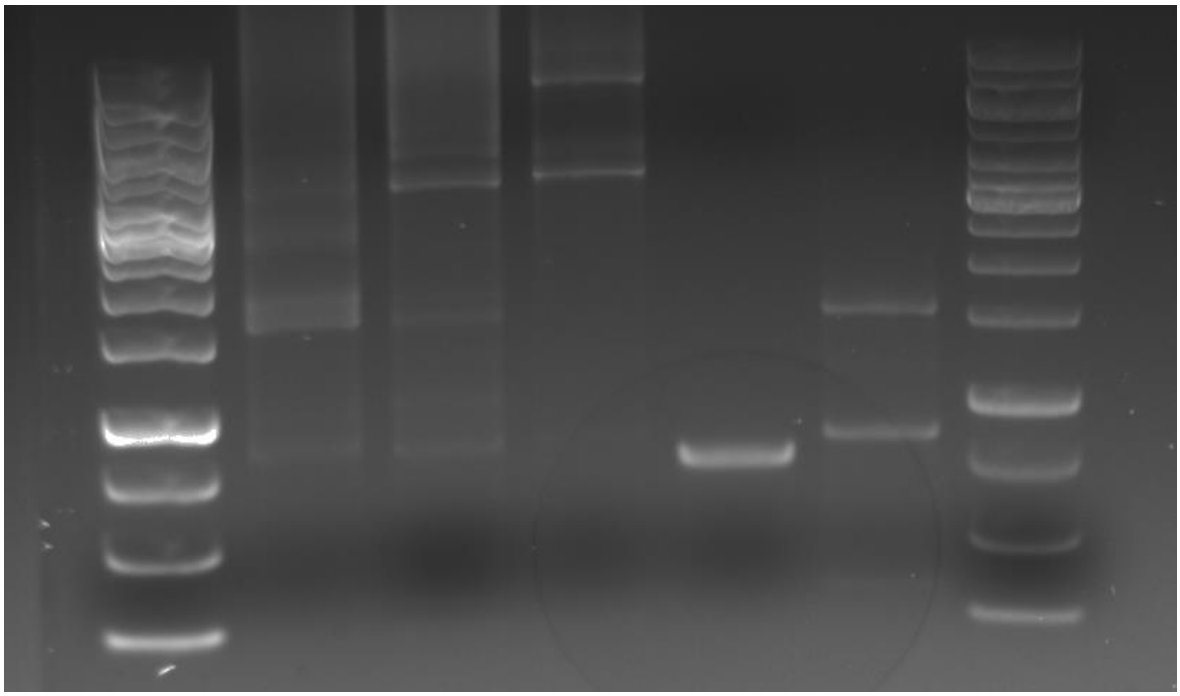


Figure 3.23 - PCR product after CPEC cloning with DNA ladder 1 Kb in lane 1, vector backbone and fused GFP+YFP in lane 2, vector backbone without any additional fragment in lane 3 and 4. Lane 5 shows the GFP fragment and lane 6 GFP+YFP fragment (used as control).

3 Discussion

The main goal of this project was to produce FP variants using GFPmut3b, EYFP and mRFP1 by combining two of the proteins together. The main objective was to analyze these variants for their spectroscopic features. The new FP variants obtained were expected to have new spectral characteristics based on two emission and excitation spectra representing two fused proteins. Before assembling the two proteins together, the expression of FPs individually was observed to characterize their spectral features, maturation and stability.

Cloning and characterization of individual FPs

Single FPs in the current study were expressed by incorporating them in the expression vector pSB1AK3 with LacI IPTG-inducible promoter and expressed in *E.coli* DH5α cells. The transformed colonies did not initially show any color, however in case of mRFP1 red color appeared about an hour after the cells were taken out of the incubator. Since no inducer (IPTG) was used, this indicates the leaking expression of the LacI promoter. Studies have shown that mRFP1 matures ten times faster than other FPs and shows a bright red color in living cells (Campbell et al., 2002). On the other hand GFPmut3b and EYFP could be identified by fluorescence emission after induction with IPTG but no color was visible to the naked eye without inducer.

Maturation experiments were carried out on individual FPs by inhibiting the protein production with chloramphenicol (CAP). CAP inhibits proteins production by deactivating an enzyme peptidyl transferase, which is an important part of ribosome needed for translation. The fluorescence measured immediately after adding CAP was similar in all samples including the control samples without CAP. In the Cap-treated samples fluorescence intensity per cell (FIC) started to decrease about 20 minutes after CAP addition (figure 3.10). The FIC signals increase after the addition of CAP can be attributed to the maturation of the FPs chromophore. This FIC was maintained for 3 hours in cells treated with CAP indicating stability of the expressed and matured FPs. While the total fluorescence intensity of the control samples increased with time along with the increase in cell population, the FIC decreased more than the FIC in the CAP-treated samples.

The decrease in FIC for the non-inhibited samples is unexpected and may be due to reabsorption of the fluorescence signal by the high number of FPs in the sample. The decrease in FIC increases with higher cell OD. One reason could be that cells in the untreated samples degrade, which could release the FP in the medium, and lead to FP degradation. Another possibility for the decrease in FIC of the untreated sample could be quorum sensing, which may causes changes in the pattern of gene expression in response to cell population density (Fuqua, Parsek, & Greenberg, 2001).

All FPs have a short maturation time and slow degradation. There is a short increase, indicating a small amount of FP that is still maturing in case of mRFP1 and EYFP followed by a slow decrease, indicating degradation (figure 3.10). But in case of GFPmut3b no increase in the peak intensity was observed which indicates that GFPmut3b started to degrade sooner as compared to mRFP1.

Construction and characterization of potential FP fusions

After confirming the expression of individual FPs, the next step was fusing the FPs together by Gibson Assembly cloning. Colonies on selective medium were obtained after cloning but when these colonies were analyzed for their spectroscopic features, no new spectral features were observed compared to individual FPs. The constructs (suspected of carrying two FPs) with mRFP1 gave colonies with red color that had the absorption and emission spectra similar to individual mRFP1. Some colonies showed no apparent change in color, were also analyzed by spectroscopic techniques, but they did not show the presence of any FP. The spectroscopic data indicates that cloning did result in the expected constructs. Analysis of the colored colonies by colony restriction digests indicated that two FPs were linked in the constructs. However, sequencing plasmids of the colored colonies confirmed spectroscopic data, that no new fused FPs had been generated.

In order to figure out, which step in the cloning procedure was not working, a series of tests were conducted. First it was confirmed that the two fragments (including linker) could be assembled correctly using Gibson cloning, because this fragment (FP-linker-FP) could be successfully amplified by PCR. However, cloning of these confirmed fragments into the vector backbone using Gibson cloning did not succeed after three additional attempts. Therefore another cloning method called CPEC was used to assemble all the fragments

together. No positive results were obtained using the CPEC procedure as well. In both cloning strategies, cloning of the fused FP fragment into the plasmid backbone could not be accomplished.

What can be changed to obtain the desired constructs? The cloning procedure could be repeated using a different plasmid backbone. The Fusion FPs may not be stable in E. coli, because due to the similarity of the two FP genes, DNA recombination in E coli resulted in the gene recombination. This recombination would likely result in hybrid sequences having the characteristics of both the proteins. But this is less likely to occur as FP fusion proteins have been reported previously (Evers, van Dongen, Faesen, Meijer, & Merkx, 2006).

The inducible expression in case of individual FPs worked and provided useful data regarding their expression and maturation.

5 Conclusion

The primary object of this project was to generate variants of FPs with new spectral characteristics. It was planned to genetically fuse two individual FPs with a linker. The sequences of suitable FPs were chosen form databases including BioBrick registry. A cloning strategy using Gibson cloning was developed to construct a vector-based inducible protein expression system in *E. coli*.

The spectral characteristics and maturation and proteins stability for individual FPs were assessed in this study. Individual FPs were cloned into a IPTG-inducible expression vector. Correct construction of the vector was verified by colony PCR, restriction enzyme digestion and sequencing. These constructs could be induced by IPTG and were used to analyze FP maturation and stability.

Unfortunately, fused GFP variants were not constructed using different cloning approaches.

These new FP variants can be used as fluorescent markers for a wide range of studies. The spectral features of these new FP variants are based on fluorescence resonance energy transfer (FRET) between donor FPs and acceptor FPs which would lead to a larger separation between excitation and fluorescence. Studying these new FP variants can provide an opportunity to learn more about Förster excitation transfer to support theoretical studies. These novel proteins could prove important markers for organisms that are heavily pigmented, such as photosynthetic organisms. Once generated these new FPs could also be further developed by circularization with inteins (Tavassoli & Benkovic, 2007), so more stable FPs can be produced. Generating new FP variant is an exciting field of study that could provide new opportunities for science.

References

Campbell, R. E., Tour, O., Palmer, A. E., Steinbach, P. A., Baird, G. S., Zacharias, D. A., & Tsien, R. Y. (2002). A monomeric red fluorescent protein. *Proceedings of the National Academy of Sciences*, *99*(12), 7877–7882. http://doi.org/10.1073/pnas.082243699

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

Chudakov, D. M., Lukyanov, S., & Lukyanov, K. A. (2005). Fluorescent proteins as a toolkit for in vivo imaging. *Trends in Biotechnology*, *23*(12), 605–613. http://doi.org/10.1016/j.tibtech.2005.10.005

Chudakov, D. M., Matz, M. V., Lukyanov, S., & Lukyanov, K. A. (2010). Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiological Reviews*, *90*(3), 1103–1163. http://doi.org/10.1152/physrev.00038.2009

Cody, C. W., Prasher, D. C., Westler, W. M., Prendergast, F. G., & Ward, W. W. (1993). Chemical structure of the hexapeptide chromophore of the Aequorea green-fluorescent protein. *Biochemistry*, *32*(5), 1212–1218. http://doi.org/10.1021/bi00056a003

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

Cubitt, A. B., Heim, R., Adams, S. R., Boyd, A. E., Gross, L. A., & Tsien, R. Y. (1995). Understanding, improving and using green fluorescent proteins. *Trends in*

Biochemical Sciences, *20*(11), 448–455. http://doi.org/10.1016/S0968-0004(00)89099-4

Davenport, D., & Nicol, J. a. C. (1955). Luminescence in Hydromedusae. *Proceedings of the Royal Society of London B: Biological Sciences*, *144*(916), 399–411. http://doi.org/10.1098/rspb.1955.0066

Evers, T. H., van Dongen, E. M. W. M., Faesen, A. C., Meijer, E. W., & Merkx, M. (2006). Quantitative Understanding of the Energy Transfer between Fluorescent Proteins Connected via Flexible Peptide Linkers. *Biochemistry*, *45*(44), 13183–13192. http://doi.org/10.1021/bi061288t

Fuqua, C., Parsek, M. R., & Greenberg, E. P. (2001). REGULATION OF GENE EXPRESSION BY CELL-TO-CELL COMMUNICATION: Acyl-Homoserine Lactone Quorum Sensing. *Annual Review of Genetics*, *35*(1), 439–468. http://doi.org/10.1146/annurev.genet.35.102401.090913

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–345. http://doi.org/10.1038/nmeth.1318

Heim, R., Prasher, D. C., & Tsien, R. Y. (1994). Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences*, *91*(26), 12501–12504.

Inouye, S., & Tsuji, F. I. (1994). Aequorea green fluorescent protein: Expression of the gene and fluorescence characteristics of the recombinant protein. *FEBS Letters*, *341*(2–3), 277–280. http://doi.org/10.1016/0014-5793(94)80472-9

Jong, H. de, Ranquet, C., Ropers, D., Pinel, C., & Geiselmann, J. (2010). Experimental and computational validation of models of fluorescent and luminescent reporter genes in bacteria. *BMC Systems Biology*, *4*(1), 55. http://doi.org/10.1186/1752-0509-4-55

Jung, G. (2012). *Fluorescent Proteins II: Application of Fluorescent Protein Technology*. Springer Science & Business Media.

Knight, T. (2003). *Idempotent Vector Design for Standard Assembly of Biobricks*.

Labas, Y. A., Gurskaya, N. G., Yanushevich, Y. G., Fradkov, A. F., Lukyanov, K. A., Lukyanov, S. A., & Matz, M. V. (2002). Diversity and evolution of the green fluorescent protein family. *Proceedings of the National Academy of Sciences*, *99*(7), 4256–4261. http://doi.org/10.1073/pnas.062552299

Lakowicz, J. R. (2007). *Principles of Fluorescence Spectroscopy*. Springer Science & Business Media.

Lichtman, J. W., & Conchello, J.-A. (2005). Fluorescence microscopy. *Nature Methods*, *2*(12), 910–919. http://doi.org/10.1038/nmeth817

Miyawaki, A., Nagai, T., & Mizuno, H. (2003). Mechanisms of protein fluorophore formation and engineering. *Current Opinion in Chemical Biology*, *7*(5), 557–562. http://doi.org/10.1016/S1367-5931(03)00097-8

Morin, J. G., & Hastings, J. W. (1971). Energy transfer in a bioluminescent system. *Journal of Cellular Physiology*, *77*(3), 313–318. http://doi.org/10.1002/jcp.1040770305

Pakhomov, A. A., & Martynov, V. I. (2008). GFP Family: Structural Insights into Spectral Tuning. *Chemistry & Biology*, *15*(8), 755–764. http://doi.org/10.1016/j.chembiol.2008.07.009

Patterson, G., Day, R. N., & Piston, D. (2001). Fluorescent protein spectra. *Journal of Cell Science*, *114*(5), 837–838.

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

Quan, J., & Tian, J. (2011). Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries. *Nature Protocols*, *6*(2), 242–251. http://doi.org/10.1038/nprot.2010.181

Remington, S. J. (2006). Fluorescent proteins: maturation, photochemistry and photophysics. *Current Opinion in Structural Biology*, *16*(6), 714–721. http://doi.org/10.1016/j.sbi.2006.10.001

Shetty, R. P., Endy, D., & Jr, T. F. K. (2008). Engineering BioBrick vectors from BioBrick parts. *Journal of Biological Engineering*, *2*(1), 1–12. http://doi.org/10.1186/1754-1611-2-5

Shimomura, O. (1979). Structure of the chromophore of Aequorea green fluorescent protein. *FEBS Letters*, *104*(2), 220–222. http://doi.org/10.1016/0014-5793(79)80818-2

Shimomura, O., Johnson, F. H., & Saiga, Y. (1962). Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea. *Journal of Cellular and Comparative Physiology*, *59*(3), 223–239. http://doi.org/10.1002/jcp.1030590302

Southward, C. M., & Surette, M. G. (2002). The dynamic microbe: green fluorescent protein brings bacteria to light. *Molecular Microbiology*, *45*(5), 1191–1196. http://doi.org/10.1046/j.1365-2958.2002.03089.x

Stepanenko, O. V., Stepanenko, O. V., Shcherbakova, D. M., Kuznetsova, I. M., Turoverov, K. K., & Verkhusha, V. V. (2011). Modern fluorescent proteins: from chromophore formation to novel intracellular applications. *BioTechniques*, *51*(5), 313–314, 316, 318 passim. http://doi.org/10.2144/000113765

Tavassoli, A., & Benkovic, S. J. (2007). Split-intein mediated circular ligation used in the synthesis of cyclic peptide libraries in E. coli. *Nature Protocols*, *2*(5), 1126–1133. http://doi.org/10.1038/nprot.2007.152

Tsien, R. Y. (1998). The Green Fluorescent Protein. *Annual Review of Biochemistry*, *67*(1), 509–544. http://doi.org/10.1146/annurev.biochem.67.1.509

Appendix

Sequences of the FPs:

The sequences of the FPs (GFPmut3b, EYFP and mRFP1) used in this study were retrieved using iGEM registry. The Sequences used are given below.

>BBa_E0030 (EYFP) Part-only sequence (723 bp)

atggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaacggc
cacaagttcagcgtgtccggcgagg
gcgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggc
ccaccctcgtgaccaccttcggcta
cggcctgcaatgcttcgcccgctaccccgaccacatgaagctgcacgacttcttcaagtccgccatgcccgaagg
ctacgtccaggagcgcaccatcttc
ttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatcgag
ctgaagggcatcgacttcaaggagg
acggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaagcaga
agaacggcatcaaggtgaacttcaa
gatccgccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagcagaacacccccatcggcgacgg
ccccgtgctgctgcccgacaaccac
tacctgagctaccagtccgccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtg
accgccgccgggatcactctcggca
tggacgagctgtacaagtaataa

>BBa_E0040 Part-only sequence (720 bp)

atgcgtaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaatgggcac
aaattttctgtcagtggagagggtg
aaggtgatgcaacatacggaaaacttacccttaaatttatttgcactactggaaaactacctgttccatggccaa
cacttgtcactactttcggttatgg
tgttcaatgctttgcgagatacccagatcatatgaaacagcatgacttttt caagagtgccatgcccgaaggtta
tgtacaggaaagaactatatttttc
aaagatgacgggaactacaagacacgtgctgaagtcaagtttgaaggtgataccc ttgttaatagaatcgagtta
aaaggtattgattttaaagaagatg
gaaacattcttggacacaaattggaatacaactataactcacacaatgtatacatcatggcagacaaacaaaaga
atggaatcaaagttaacttcaaaat
tagacacaacattgaagatggaagcgttcaactagcagaccattatcaacaaaatactccaattggcgatggccc
tgtccttttaccagacaaccattac
ctgtccacacaatctgccctttcgaaagatcccaacgaaaagagagaccacatggtccttcttgagtttgtaaca
gctgctgggattacacatggcatgg
atgaactatacaaataataa

>BBa_E1010 Part-only sequence (706 bp)

atggcttcctccgaagacgttatcaaagagttcatgcgtttcaaagttcgtatggaaggttccgttaacggtcac
gagttcgaaatcgaaggtgaaggtg
aaggtcgtccgtacgaaggtacccagaccgctaaactgaaagttaccaaaggtggtccgctgccgttcgcttggg
acatcctgtccccgcagttccagta
cggttccaaagcttacgttaaacacccggctgacatcccggactacctgaaactgtccttcccggaaggtttcaa
atgggaacgtgttatgaacttcgaa
gacggtggtgttgttaccgttacccaggactcctccctgcaagacggtgagttcatctacaaagttaaactgcgt
ggtaccaacttcccgtccgacggtc

cggttatgcagaaaaaaaccatgggttgggaagcttccaccgaacgtatgtacccggaagacggtgctctgaaag
gtgaaatcaaaatgcgtctgaaact
gaaagacggtggtcactacgacgctgaagttaaaaccacctacatggctaaaaaaccggttcagctgccgggtgc
ttacaaaaccgacatcaaactggac
atcacctcccacaacgaagactacaccatcgttgaacagtacgaacgtgctgaaggtcgtcactccaccggtgct
taataacgctgatagtgctagtgta
gatcgc

Sequence of the fused FP (GFPmut3b+mRFP1)

The constructed variants with fused FPs were sent for sequencing and the sequencing result showed that either of the two FPs incorporated in the vector backbone. Example sequence of one of the suspected fused FP is given below with mRFP1 highlighted red.

gattactatnaaaataggcgtannangaggcagaatttcagataaaaaaaatccttagctttcgctaaggatgat
ttctggaattcgcggccgcttctagagcaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgca
gctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcatt
aggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataacaatttcaca
catactagagaaagaggagaaatactagatggcttcctccgaagacgttatcaaagagttcatgcgtttcaaagt
tcgtatggaaggttccgttaacggtcacgagttcgaaatcgaaggtgaaggtgaaggtcgtccgtacgaaggtac
ccagaccgctaaactgaaagttaccaaaggtggtccgctgccgttcgcttgggacatcctgtccccgcagttcca
gtacggttccaaagcttacgttaaacacccggctgacatcccggactacctgaaactgtcctccggaaggttt
caaatgggaacgtgttatgaacttcgaagacggtggtgttgttaccgttacccaggactcctccctgcaagacgg
tgagttcatctacaaagttaaactgcgtggtaccaacttcccgtccgacggtccggttatgcagaananaaccat
gggttgggaagcttccaccgaacgtatgtacccggaagacggtgctctgaaaggtgaaatcaaaatgcgtctgaa
actgaaagacggtggtcactacgacgctgaagttaaaaccacctacatggctaanaaaccggttcagctgccggg
tgcttacaaaaccgacatcaaactggacatcacctcccacaacgaagactacaccatcgttgaacagtacgaacg
tgctgaaggtcgtcactccaccggtgcttaataacgctgatagtgctagtgtanntcgctactagtagcggccgc
tgcagtccggcaaaaaagggcaaggtgtcaccaccctgccctttttctttaaaaccgaaaaaattacttccngtt
atgcaggcttcctcgctcactgactcgctgcgctcggtcgttcggntgcggnnagcggtatcanntcnnncaa