

# Creation of novel photochangable fluorescent protein through directed evolution

By

Pravin Vitthal Marathe

(LIFE09201304003)

Tata Memorial Centre, Mumbai

A thesis submitted to the

Board of Studies in Life Sciences

In partial fulfilment of requirements

For the Degree of

**DOCTOR OF PHILOSOPHY**

**OF**

**HOMI BHABHA NATIONAL INSTITUTE**

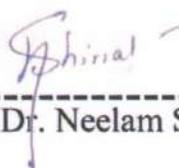


January, 2021

# Homi Bhabha National Institute

## Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Mr. Pravin Vitthal Marathe entitled “**Creation of novel photochangable fluorescent protein through directed evolution**” and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.



Chairperson – Dr. Neelam Shirsat

Date: 12.01.2021



Guide – Dr. Dibyendu Bhattacharyya

Date: 12.01.2021



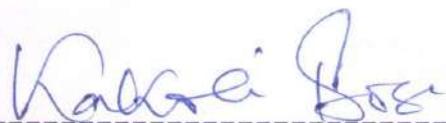
External Examiner – Dr. Debi Prasad Sarkar

Date: 12.01.2021



Member – Dr. Sanjeev Waghmare

Date: 12.01.2021



Member – Dr. Kakoli Bose

Date: 12.01.2021

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.

I hereby certify that I have read this thesis prepared under my direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 12.01.2021

Place: Navi Mumbai

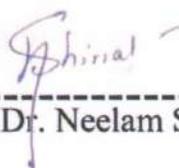


Dr. Dibyendu Bhattacharyya  
Guide

# Homi Bhabha National Institute

## Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Mr. Pravin Vitthal Marathe entitled “**Creation of novel photochangable fluorescent protein through directed evolution**” and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.



Chairperson – Dr. Neelam Shirsat

Date: 12.01.2021



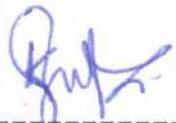
Guide – Dr. Dibyendu Bhattacharyya

Date: 12.01.2021



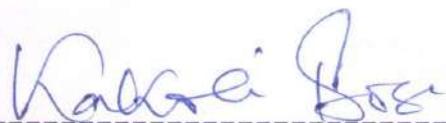
External Examiner – Dr. Debi Prasad Sarkar

Date: 12.01.2021



Member – Dr. Sanjeev Waghmare

Date: 12.01.2021



Member – Dr. Kakoli Bose

Date: 12.01.2021

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.

I hereby certify that I have read this thesis prepared under my direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 12.01.2021

Place: Navi Mumbai



Dr. Dibyendu Bhattacharyya  
Guide

## STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfilment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.



Pravin Vitthal Marathe

## DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.



**Pravin Vitthal Marathe**

## DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out under my supervision. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.



Dr.Dibyendu Bhattacharyya

Guide

## List of Publications arising from thesis

Pravin Marathe, Mahadeva Swamy, Deepak Nair, and Dibyendu Bhattacharyya.  
“mEosBrite are improved variants of mEos3.2 developed by semi- rational protein engineering”. Journal of fluorescence. April 2020 PMID: 32385659  
DOI: 10.1007/s10895-020-02537-8

### Conference/Symposium:

1. Poster presentation on topic entitled “**Bioengineering of fluorescent proteins through directed evolution**” at “Optics within Life Sciences (OWLS)” March 16 to 18, 2016-TIFR, Mumbai
2. Poster presentation on topic entitled “**Breaking the limits of light microscopy by creation of smart variants of mEos3.2 protein**” at “International conference of Cell Biology (ICCB)” held from January 27 to 31, 2018, Hyderabad
3. Oral presentation on topic entitled “**Creation of smart variants Photoconvertible fluorescent protein mEos3.2 through directed evolution**” at “Focus on Microscopy (FOM)” held from March 25 to 28, 2018, Singapore



Pravin Vitthal Marathe

**This thesis is dedicated to my beloved family,  
friends and teachers for their endless support  
and constant encouragement**



## Acknowledgements

I wish to take this opportunity to thank everyone who had been instrumental in making this long Journey a rewarding and truly an indelible one. Foremost, I extend my deepest gratitude to my Thesis research advisor Dr.Dibyendu Bhattacharyya, for introducing me to the field of protein Engineering, his guidance, encouragement, advice and constant support on countless occasions during my PhD tenure. Thank you for giving me an opportunity and freedom to expose myself to the world of research.

I would like to express my gratitude to Dr.Sudeep Gupta, Director, Dr. Prasanna Venktraman, and Dy.Director ACTREC for providing excellent infrastructure, encouragement and endless supports throughout tenure. I am also thankful to Dr. Shubhada Chiplunkar and Dr. Rajiv Sarin (former Director). My sincere thanks to the funding agencies, Sam-mystery and HBNI for supporting the international travel to present my work at an international conference.

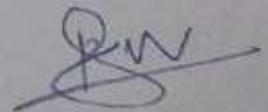
I am grateful to all my Doctoral Committee Chairpersons Dr.Neelam Shirsat (Chairperson), Dr.Sanjeev Waghmare and Dr. Kakoli Bose for their invaluable input, advice and support in shaping and refining my research work.

I would like to acknowledge all the members of Bhattacharyya Lab. I am thankful to Sanjay and Mansi for their excellent technical help. I thank all my colleagues, Madhura, Abira, Prasanna, Bhavik, Kaushal, Chumki, Ketakee, Naini, Sudeshna, Shreosi and Roma for their constant support and healthy discussions during last seven years. They really made difficult times easier for me with tons of jokes,

funny fights & laughter. Countless amazing food parties & tea sessions will always be cherished.

My Very special thanks to members of CIR, Imaging facility and EM facility staff, Uday Dandekar sir, Vaishali, Tanuja, Jayraj, Vinita Sawant, Siddhi. I am thankful to the Director's office (Mrs. Pritha Menon and Mrs. Lata Shelar), Academic office (Mrs. Sharvari Joshi, Maya Dolas), Administration, Accounts, Dispatch, Purchase, Stores and Photography department for their constant support throughout the tenure. Special appreciation also goes out to my friends Raghav, Rajashree, Ajay, Shalini, and all of my batchmates for their moral support, Memorable trips , birthday celebrations, Amazing food parties and creating cheerful environments throughout last seven years. I especially thank Rajashree, Roma, Raghav, Shreosi, Ajay, Shalini, Sudeshna, Naini and for their ever-present unconditional support for scientific and personal work.

I dedicate my warmest gratitude to all my family members for their love & constant faith in me, even when I couldn't explain to them what I was doing. They have always given me the latitude to develop my interest in whatever I'm passionate about and have always been happy and proud of my accomplishments.



**Pravin Vitthal Marathe**

## TABLE OF CONTENT

<b>Synopsis.....</b>	<b>19</b>
<b>List of Figures.....</b>	<b>33</b>
<b>List of Tables.....</b>	<b>35</b>
<b>Abbreviations.....</b>	<b>36</b>
<b>1. Introduction.....</b>	<b>37</b>
1.1 Background of thesis.....	39
<b>2. Review of literature.....</b>	<b>41</b>
2.1 Fluorescent proteins.....	43
2.2 FP is ancient metazoan gene.....	43
2.3 Biological functions of FP.....	44
2.4 Key characteristic of FP for its practical use.....	45
2.5. Basic applications of FP.....	48
2.6 GFP and its structure, Chromophore and its maturation.....	52
2.7 Spatial Structure and Diversity of Chromophores.....	57
2.8 Modern palette of FP.....	62
2. 9 PTFP and its types.....	68
2.10 Directed evolution.....	77
2.11. Eos fluorescent protein and its Directed evolution.....	84
<b>3. Aims and objectives.....</b>	<b>89</b>
3.1 Hypothesis.....	91
3.2 Objectives.....	91
3.3 Detailed objectives.....	91
3.4 Work done.....	92

<b>4. Materials and methods.....</b>	<b>93</b>
4.1: Molecular biology methods.....	95
4.1.1: Preparation of ultra-competent E. coli: .....	95
4.1.2: Bacterial Transformation.....	96
4.1.3: Plasmid DNA isolation.....	96
4.1.4: Agarose gel electrophoresis.....	98
4.1.5: Polymerase chain reaction (PCR) .....	99
4.1.6: Gene Cloning.....	101
4.1.7 Mutagenesis methods.....	106
4.2 Creation of mutant library and its screening.....	109
4.2.1. General methods and materials.....	109
4.2.2. Cloning of mEos3.2 gene.....	110
4.2.3. Mutant library creation.....	110
4.2.4. Replica plating of transformants.....	110
4.2.5. Mutant library screening.....	111
4.3. Mammalian expression vectors, Cells, and transfection.....	111
4.4: Live cell imaging and its parameters.....	112
4.4.1. Live cell imaging.....	112
4.5. Protein expression and purification protocol.....	113
4.5.1 Protocol for 250 ml culture.....	113
4.5.2 Ni-IDA column preparation.....	114

4.5.3	Lysate preparation.....	114
4.5.4	Column regeneration.....	114
4.5.5	Affinity purification using Ni-IDA column.....	114
4.5.6	Composition of buffers.....	115
4.6.	Biophysical studies of purified protein.....	115
4.6.1	Fast Pressure Liquid Chromatography (FPLC) analysis.....	115
4.6.2	Dynamic Light Scattering.....	116
4.6.3	SDS PAGE and Western Blotting.....	116
4.6.4	pH-dependent Performance of purified proteins.....	118
4.6.5	Maturation kinetics.....	119
4.6.6	Determination of the quantum yield.....	119
4.6.7.	Bacterial cytotoxicity.....	120
4.6.8.	Assay for cytotoxicity in HeLa cell by transient transfection.....	120
4.6.9.	Aggregation assay.....	121
4.6.10.	Photoconversion and Half-life measurement Experiments.....	122
<b>5.</b>	<b>Results</b> .....	<b>125</b>
5.1	Introduction.....	127
5.2	Results.....	130

5.2.1 Creation of mutant library of mEos3.2 gene using three different methods of mutagenesis.....	130
5.2.2 Microscopic screening of the entire mutant library revealed three brightness improved variants of mEos3.2 collectively named as 'mEosBrite'.....	131
5.2.3 Insilco protein models generation and its Ramachandran analysis.....	135
5.2.4 Expression and purification of mEosBrite variants.....	138
5.2.5 Excitation-emission spectra and maturation kinetics of mEosBrite variants is comparable to WT mEos3.2.....	139
5.2.6 mEosBrite variants shows improvement in quantum yield and inherent brightness.....	141
5.2.7 mEosBrite variants are monomeric in nature.....	144
5.2.8 The pH-dependent behaviour of WT mEos3.2 and mEosBrite variants is comparable.....	146
5.2.9 mEosBrite variants show lower cytotoxicity in bacterial cells. ....	147
5.2.10 mEosBrite variants show significantly low cytotoxicity as compared to mEos3.2 protein when expressed in mammalian cells.....	149
5.2.11 mEosBrite variants have lower protein aggregation tendency as compared to WT mEos3.2.....	151
5.2.12 mEosBrite variants have higher photoconversion efficiency.....	152
5.2.13 the half-life of mEosBrite variants is comparable to WT mEos3.2.....	153
5.2.14 mEosBrite variants performed better in Super resolution imaging.....	154
5.3 Discussion.....	157
<b>6. Summary and conclusion.....</b>	<b>161</b>
6.1 Summary and conclusion.....	163
6.2 Future Perspectives.....	166
<b>7. Bibliography.....</b>	<b>169</b>

<b>Appendix.....</b>	<b>195</b>
<b>Publication.....</b>	<b>209</b>

## LIST OF FIGURES

**Figure 2.1 Structure of GFP**

**Figure 2.2 Proposed mechanism for chromophore maturation**

**Figure 2.3 Expression of GFP in *E. coli*.**

**Figure 2.4 Expression of GFP in *C.elegans* larva.**

**Figure 2.5 Naturally occurring GFP-like protein chromophores structure and their pathway of maturation**

**Figure 2.6 Chemical structure of chromophore generated by mutating natural fluorescent proteins**

**Figure 2.7 Multicolour imaging (Brainbow)**

**Figure 2.8 Spectral diversity of available monomeric FPs**

**Figure 2.9 Molecular mechanism for Green to red photoconversion**

**Figure 2.10 Schematic outline of a directed evolution process**

**Figure 2.11 Screening methods for protein evolution**

**Figure 2.12 Tetrameric structure of wtEosFP**

**Figure 5.1 Sequence alignment of mEos3.2 with mClavGR2**

**Figure 5.2. Live cell imaging of mEosBrite variants shows that mEosBrite variants are much brighter than the WT mEos3.2**

**Figure 5.3 Insilco protein models of mEosBrite variants perfectly to WT mEos3.2**

**Figure 5.4 Ramachandran plot for Insilco protein models**

**Figure 5.5 SDS-PAGE of mEosBrite variants**

**Figure 5.6 Spectral properties and maturation kinetics of mEosBrite variants are comparable**

**Figure 5.7 mEosBrite variants have higher Quantum yield as compared to mEos3.2 protein**

**Figure 5.8 mEosBrite variants are monomeric even at higher concentration**

**Figure 5.9 pH dependent behaviour of mEosBrite variants is comparable to WT mEos3.2**

**Figure 5.10 mEosBrite variants are not toxic to bacterial cells even when overexpressed**

**Figure 5.11 mEosBrite variants are less cytotoxic in mammalian cells are compared to WT mEos3.2**

**Figure 5.12 mEosBrite variants show lesser protein aggregation**

**Figure 5.13 mEosBrite variants show faster photoconversion than WT mEos3.2**

**Figure 5.14 The half-life of mEosBrite variants is comparable to WT mEos3.2**

**Figure 5.15 mEosBrite variants perform better than WT mEos3.2 in SIM**

**Figure 5.16 mEosBrite variants perform better than WT mEos3.2 in PALM**

## LIST OF TABLES

**Table 2.1 Popular known signal motifs used to target subcellular compartments**

**Table 2.2 Comparison and summary of different methods of gene diversification**

**Table 4.1 Contents of PCR reaction mixture**

**Table 4.2 Conditions for Typical PCR cycle**

**Table 4.3 Contents of of digestion reaction**

**Table 4.4 Calculation for Typical Ligation Reaction**

**Table 4.5 Component of Ligation reaction**

**Table 4.6 Contents of PCR reaction for site-directed mutagenesis**

**Table 4.7 Cycling conditions for PCR for site-directed mutagenesis**

**Table 4.8 Components of EP-PCR**

**Table 4.9 Cycling conditions for EP-PCR**

**Table 4.10 Composition for 15% resolving gel of SDS PAGE**

**Table 4.11 Composition of a stacking SDS-PAGE gel**

**Table 5.1: Mutational replacement in mEosBrite variants**

**Table 5.2 Ramachandran analysis for mEos3.2 and mEosBrite Insilco protein models**

**Table 5.3 Average QY and inherent brightness of mEos3.2 and mEosBrite variants**

# Chapter 1

## Introduction



## 1.1. Background of Thesis

Fluorescent proteins (FPs) are extensively utilised in study of varied living systems. When encoded in frame with the protein of interest (POI), FPs can be used to study different physiological processes inside the cell like protein localization, its movement, protein ageing, turnover rate of protein, etc. (1). Today a wide range of fluorescent proteins are available due to researchers' extensive efforts in the study and discovery of naturally occurring fluorescent proteins. Bioengineering of these naturally available fluorescent proteins has further expanded the palette of fluorescent proteins.

Phototransformable fluorescent protein (PTFP) is one of the most exciting class of fluorescent proteins. They are well-known for their chromophore modification. Chromophore modification of PCFP happens upon exposure of UV light (2). One of the widely used subclass of PTFP is Photo convertible fluorescent proteins (PCFPs). PTFP have a distinct feature of non-reversible photoconversion from one emission spectra to another. PCFPs hold a great influence on confocal and Super-resolution microscopic techniques because of their better contrast, higher photostability and stochastic activation. mEos3.2 is one such well-known PCFP that shows photoconversion from Green emission spectra to red emission spectra. mEos3.2 was developed by multiple rounds of bioengineering from its primitive Eos protein(3)(4)(5).

The biophysical properties of a PCFP determines its performance in confocal and super-resolution microscopy. These biophysical properties involves brightness, maturation rate, signal to noise ratio, labeling density, photostability, oligomeric nature, pH stability, and on-off switching rate, etc. of PCFP(6)(7)(8). Most of Biophysical properties of PCFP mEos3.2 protein are already optimized. The only limitation of mEos3.2 is it's comparatively lower brightness for red (after photoconversion) spectra. This limitation of mEos3.2 protein restricts its use in high speed super-resolution microscopy. Our aim was to target this limitation of mEos3.2 protein while retaining its remaining biophysical fluorescent properties to make this protein a better candidate for high speed or live super-resolution microscopy.



# Chapter 2

## Review of literature



## **2.1 Fluorescent proteins**

Fluorescent proteins are a non-invasive tool that enable us to visualize molecules, structures and varied cellular processes. The most exciting feature of a FPs gene is that one can manipulate it using standard molecular biology tools thus allowing us to create fusion constructs, recombinant proteins, and even transgenic animals. Since FPs are genetic labels, there is no requirement of exogenous labeling, fixation or permeabilization to produce fluorescence signal. Fluorescent protein is expressed in 1:1 ratio to that of target protein which also facilitates quantitative imaging (9).

Fluorescent proteins belong to a homologous class and their size is also conserved i.e. 25 KDa irrespective of their source of origin. The basic structure of a fluorescent protein is conserved. It consists of extremely rigid “Beta – can” made of 11 beta sheets having a centrally located  $\alpha$ -helix (10). This centrally located  $\alpha$ - helix has the three amino acid domain called a chromophore. The amino acid residues and their position in chromophore are mostly conserved i.e. from 62-65 or 65-67. The chromophore undergoes three step maturation process which ultimately gives rise to fluorescence in the presence of oxygen. The FP chromophore is located deep inside the “Beta -can” which is very stable and thus protects the chromophore from solvents. This “Beta – can” is stable due to various noncovalent interactions that provide stability against denaturation and proteolysis (11) (12).

## **2.2 FPs are ancient metazoan gene**

FP gene is mostly found in marine animals which indicated that the primary function of fluorescent protein could be related to marine environment. First member of fluorescent protein family predicted to exist in early metazoan lineage between 500-1000 million years ago. Studies also shows that GFP could be the first and most ancestral member of this family. RFP and other fluorescent proteins originated later in more than one species as it required many chromophore modifications (17). This course of evolution in fluorescent proteins has been achieved in laboratory settings (18).

### 2.3 Biological functions of FP

Fluorescent proteins are one of the most widely studied proteins, still there is lot of ambiguity about its actual function in host organism. There are lot of assumptions, theories and explanations proposed time to time by scientist regarding the function of fluorescent proteins. According to different theories, the biological function of fluorescent protein can be categorised in two main classes 1. Photosynthetic modulation and 2. Optical communication. First function is mainly associated with reef building coral and their symbiotic associates. It varies from very high intensity to very low intensity. According to this all the reef building corals are restricted to photic zone as these reefs are directly dependent on the algal symbionts for photosynthesis. The intensity of light in this photic region therefore the role of coral becomes crucial as they act as important entity to control the light intensity and help symbiotic algal partner in adaptation and acclimatisation in this photic zone. In previous studies it was shown that FP increases the availability of light under low light conditions (19). While in recent study it was shown that FP acts as photo protective by dissipating excess energy at wavelengths of low photosynthetic activity. It was also shown that FPs protect the reef from mass bleaching during heat stress which helps to preserve overall biodiversity of reef building corals (20). There are reports suggesting that expression of fluorescent protein is upregulated in presence of high light especially in presence of blue light which is photosynthetically most relevant (21). While there are many counterarguments which scrap photosynthetic modulation function like fluorescent proteins which are present in corals are not strong enough to make any significant impact on photic zone of symbiotic algae (22). Another counter argument is, there are many Anthozoa species which do not have symbiotic association and still they show bright and multi-colored fluorescence. One of such Anthozoa species is *Corynactis californica*. Discovery of such non symbiotic anthozoan species highlights the adaptive ecological significance of fluorescence rather than light modulation for algal symbiont.

Visual communication is one of important aspect of fluorescence mostly in case of green fluorescence emitting organism like *Aequorea victoria* or sea pansy *Renilla reniformis* (23)(24). Green fluorescence in greenish costal water can be seen from long distance because of less attenuation. Many hydrozoans including medusa and siphonophores show fluorescence from the tentacles and oral appendages which suggest its role in pray attraction (25). Another alternative explanation for such localisation of fluorescence signal is that this retractable body part provides chance to produce flash of fluorescent signal in response to mechanical stimulus which could be useful in many circumstances (14). Sometimes the green fluorescence signal from tentacles and oral appendages could be useful to attract potential symbiont. This green fluorescence signal could acts as mate recognition signal, aposematic function to advertise the unpalatability of organism. Green fluorescence signal from corals sometimes acts a camouflage to protect the symbiont algae from herbivorous fishes. Induction of red FP and chromo protein in immunocompromised coral tissue showed the role of fluorescent protein in coral immunity (26)(27).

One of the most fascinating function of GFP which was discovered recently is its role as electron acceptor in presence of electron donor, including biologically relevant one (45). This electron acceptor behaviour is showed exclusively by FP which has natural Tyr66- based chromophore. Mutant with artificial Trp66- or His66 does not show similar behaviour.

## **2.4 Key characteristic of FP for its practical applications**

Natural diversity of FP served a strong molecular tool which has unlimited application in study of complex biological system. In this topic we have discussed numerous features or characteristics of FP which are crucial for their practical use. Importance of each and every characteristic of FP varies based upon the nature of experiment.

### **2.4.1 Brightness:**

Brightness is a product of extinction coefficient and quantum yield. Brightness of FP regulates its signal to noise ratio along with signal sensitivity. Brighter FP require low dose of

excitation light which reduces photo toxicity of FP, which is very crucial in living system. Brighter FP can be even expressed in lower concentration which limits its interference with tagged protein. Theoretical limit for brightness is 1. Brightness of FP can be improved by improving QY, EC or both. For brightness improvement different approach like random mutagenesis, site directed mutagenesis or somatic hyper mutation has been used from long time (28)(29). Improvement in brightness can compromise other fluorescent properties like its monomeric nature, pH and photo stability, etc. That's why validation of these characteristic is essential in brighter variants.

#### **2.4.2 Maturation Rate**

Chromophore maturation leads to the generation of fluorescence. Chromophore maturation of FP is a result of protein folding which includes distinct covalent modifications. Maturation of chromophore maturation is a rate limiting step in generation of fluorescence. Maturation of a FP can happen in few minutes, hours or days (30)(31). Maturation of FP depends on multiple factors like temperature, oxygen concentration, etc. FP with fast maturation is always preferable. Most of FP have maturation half-time from 40 min to 1–2 h, which is good enough to label cells, organelles, and proteins of interest and to perform various quantitative experiments. Maturation time of FP in invitro condition can vary from that of in vivo condition because of different turnover rate and expression efficiency. Diversity of technique employed in measurement of maturation rate make it difficult to compare the maturation rate of different FP.

#### **2.4.3 Photostability**

Higher photo stability of a FP is always preferable for longer, quantitative experiments. Higher photostability of the FP is because of its slower photobleaching. In presence of light, photobleaching causes irreversible destruction of a fluorophore. Photo stability of FP is mostly determined by chromophore surrounding b-barrel (Beta-can) which also ensures low cytotoxicity in naturally occurring FP. Both these features i.e low cytotoxicity and higher

photostability is important for FP for their natural functions in sea organisms. During optimisation of FP its higher photo stability should be preserved. Photostability of fluorescent protein depends on multiple factor like intensity or strength of excitation light, pulse of excitation light, wavelength of excitation light. Photostability can be better expressed in terms of half-life. Half-life is time taken by fluorescent protein to reach 50% of optimum fluorescent intensity from its optimum fluorescent intensity. Brighter FP always appear to be more photo stable because of its higher photon budget which is why it is always beneficial to use brighter FP for experiments. Temporary quenching (kindling), irreversible photo conversion of a FP also affects its photo stability (32)(33)(34)(35).

#### **2.4.4 Oligomeric Nature and Aggregation**

The FP used for labeling of POI must be monomeric. Oligomeric nature of tagged FP can hamper the localisation and normal function of POI. POI which is itself oligomeric, when labelled with oligomeric FP, it leads to the formation of aggregates (36)(37)(38). Most of the recently discovered fluorescent protein from anthozoa, copepoda, hydrozoa are mostly oligomeric (39)(40)(41)(30)(14). Oligomeric nature is also very common in red and orange fluorescent protein (42)(43)(14)(44). Extensive efforts have been made to monomerize red and orange fluorescent protein using rational site directed mutagenesis method. In some cases oligomeric FP perform better as compared to their monomeric counterpart in experiment which does not involve molecular tagging. In addition to oligomerisation, some FP also shows protein aggregation which causes cytotoxicity and difficulty in generating stable cell line and transgenic animals.

#### **2.4.5 pH Stability**

pH stability is also a key parameter for use of FP mostly in live cell experiments. Ideal Fluorescent protein should give optimum fluorescence within physiological pH range. pH stability is sometimes expressed in term of pKa, which is pH value at which fluorescent protein shows 50% of its optimum fluorescence. Most of fluorescent protein has pKa value

between 5 to 7. Typically fluorescence increases with pH up to a certain value. Further increase in pH leads to denaturation of protein because of which protein loses its fluorescence. Most of the fluorescent proteins show optimum fluorescence within a pH range of 8-9. Optimum pH range should fall within physiological pH range i.e. 6-10. pH stability is important aspect mostly in experiments which involve study of physiological process that often involves change in pH. In such a case, the performance of fluorescent protein can vary based upon the change in pH. Therefore pH stability is important aspect in quantitative experiments like ratio metric dual-or multicolor imaging and for FRET techniques. Change in fluorescence with respect to pH is sometimes used to study change in pH and vesicular transport in live cell.

## **2.5. Basic applications of FP**

FP has multiple applications in studying complex biological and non-biological system. The range of application of FP is continuously increasing. Here we have discussed few of the most prominent applications of FPs.

### **2.5.1 Protein Labeling**

In 1994 it was shown that when a gene of interest is labelled with FP and expressed in cell or organism, it can be used to localise POI (45). Protein labeling is one of the most important application of FP which is used to study different biological processes like protein expression, localization, translocations, interactions and degradation in living systems. Selection of FP as a label depends on its characteristic which we have discussed earlier like brightness, oligomeric nature, pH stability, maturation rate, photostability, etc. Expression of fusion construct (POI and labelled fluorescent protein) also varies in vivo depending upon transcription efficiency, mRNA stability, efficiency of translation, FP maturation rate, and stability of the protein chimera etc. Stability of the POI is very crucial in determining its turnover rate which is further responsible for the overall stability of fusion protein. The expression of fusion protein should be optimum enough in maintaining the balance between

ample fluorescence for imaging and lesser intervention with the biochemistry of a live cell. Each protein has different tolerance level inside the cell depending upon its localisation, function and also based upon the FP used as a tag. Expression levels are very crucial in case of multicolor imaging as overexpression of several FPs can severely affect the biology inside the cell. The expression of FPs should also be comparable for multicolor imaging to avoid error in result as dominantly expressing FP can give the different picture than reality. Transiently transfected cell always has this problem of uneven expression. For high reproducibility of quantitative result, generation of stable cell line is always beneficial.

### **2.5.2 Photobleaching Techniques**

Photobleaching of a FP, which is one of the problems in imaging experiment is useful to study mobility POI by means of tagged FP (46). Study of protein movement can be very useful to understand the functional activity and its interactions with other proteins. FRAP is one such most widely used technique which uses Photobleaching property of a FP. In FRAP we bleach small region of interest and subsequently track fluorescence recovery rate in the very same region. The fluorescence recovery rate is measured by rate of movement of unbleached molecule within photobleached patch from another region of cell.

### **2.5.3 Subcellular Localizations**

FPs can be used to label and study subcellular compartment, which help us to understand the subcellular structure and event happening inside them. There are different marker proteins already known which can be used to study such subcellular structures (Table.2.1). This marker protein can be helpful to labelled specific subcellular compartment.

	Targeting Signal Motifs	Source
Nucleus	COOH terminus: PKKKRKVEDA COOH terminus: DPKKKRKV COOH terminus: DPKKKRKVDPKKKRKVDPKKKRKVGSTGSR In general, various KR-rich sequences close to the COOH terminus can lead to more or less efficient protein localization to the nucleus or nuclear membrane	NLS NLS of SV40 TAg
Cytosol	COOH terminus: LALKLAGLDI	NES
ER lumen	NH <sub>2</sub> terminus: MLLSVPLLLGLLGLAAAD and COOH terminus: KDEL	Calreticulin
Mitochondrial matrix	NH <sub>2</sub> terminus: MSVLTPLLLRGLTGSARRLPVPRAKIHSLGDP Tandem repeat of this signal provides more reliable targeting	Cytochrome-c oxidase
Mitochondrial membrane	NH <sub>2</sub> terminus: MVGRNSAIAAGVCGALFIGYCIYFDRKRRSDPN NH <sub>2</sub> terminus: MAIQLRSLFPLALPGMLALLGWWFFSRKK	Tom20 DAKAP1a
Golgi lumen	NH <sub>2</sub> terminus: 81 amino acids of the human $\beta$ 1,4-galactosyltransferase	1,4-Galactosyltransferase
Golgi membrane	NH <sub>2</sub> terminus: MGNLKVAQEPGPPCGLGLGLGLGKQGPA	eNOS
Plasma membrane	NH <sub>2</sub> terminus: MGCIKSKRKNLNDGVDGMT Myristoylation and palmitoylation COOH terminus: KKKKSKTKCVM Farnesylation COOH terminus: KLNPPDESGPGCMSCKCVLS Farnesylation NH <sub>2</sub> terminus: MLCCMRRTKQVEKNDEDQKI Double palmitoylation	Lyn kinase K-ras4B v-Ha-Ras GAP-43
Peroxisomal matrix	COOH terminus: SKL COOH-terminal sequence XKL can also determine notable targeting to the peroxisomal matrix	

**Table 2.1 Popular known signal motifs used to target subcellular compartments (47)**

(Adapted from Amy E Palmer, *Current Opinion in Chemical Biology* 2008, 12:60–65)

#### 2.5.4 Timers

Timers are a class of fluorescent proteins which change fluorescence color with time. Timers are useful to study their temporal expression in retrospect. Timers have different maturation time which varies from minutes to day and that's why they can be used to study different cellular aspect having variety of time scale. DsRed-E5 was first timer FP reported in 2000 which shows green color after it synthesis which changes into red after few hours (48). DsRed-E5 is tetrameric in nature; some of the chromophore in this tetramer first mature to GFP like chromophore which give rise to green fluorescence. Green fluorescence dominates for some time until the slowly maturing red chromophore appears which receives excitation light from green chromophore as result of hyper efficient intratetrameric Forster Resonance Energy Transfer (49). Thus cell or tissue showing green fluorescence indicate their recent production, while cell or tissue showing red fluorescence indicate their production had happened several hours ago. As the tetrameric FPs are not preferred in many experiments, few monomeric timer like monomeric mCherry-based Timers named fast-FT, medium-FT,

and slow-FT have been developed. These timers can take few hours to a day to change its color from blue to red. Monomeric timers can concurrently monitor protein age as well as protein localization unlike tetrameric timers.

### **2.5.5 Cell and Tissue Labeling**

FPs can be used to label cell, tissue, organ and even whole organism when expressed under specific promoter in appropriate genetic context. This labeling of cells, tissue, organ and whole organism is helpful in field of immunology(50)(51), developmental biology(52)(53)(54), carcinogenesis(55)(56)(57)(58)(59), transplantology(60)(61)(62)(63), neurobiology(64)(65)(66) (67)(68). The major problem faced during fluorescent imaging of proteins, cells, and tissues within whole animals is the light absorption and scattering by melanin and haemoglobin (69). Longer wavelength excitation light reduces the impact of light absorption and light scattering. The optimum optical window which is most suitable for visualisation in live cell sample is between 650–700 and 1,100 nm (69). Considering this, the significance of bright far red and infrared fluorescent proteins has increased in fluorescent imaging of live cell imaging. Katushka is one of such important far red fluorescent protein with emission profile peaks at 635nm which extend far beyond 700nm (70). It has been observed that katushka shows far better brightness compared to other fluorescent protein in emission spectra beyond 650nm which gives it best sensitivity (71)(72)(73). Excitation maxima for all possible far red FPs is under 590nm. This excitation maxima needed to be pushed much more towards longer wavelength to make excitation by 633 laser possible, which will eventually shift emission spectra further toward red. Creation of brighter infra-red FP will be ideal solution which will allow us for whole body imaging, but again it is difficult task to develop FP with infrared emission spectra. Recently, abounding advanced techniques are invented, out of which Multispectral Optoacoustic Tomography (MSOT) showed very promising result in deep tissue imaging of live sample (74)

### **2.5.6 DNA and RNA Labeling**

For live cell studies, in vivo imaging of mRNA production, its localisation and dynamics is very useful. This technique allows real time mRNA labeling and it's tracking with use of FP. As mRNA and FP constructs are genetically encoded, labeling of mRNA through a fluorescent protein could be very useful to create stable cell lines and transgenic animal. The easiest way to do mRNA labeling is to fuse FP with RNA binding domain and then target this RNA binding domain by corresponding RNA motif which is fused to the target mRNA. The major drawback of this method is that a large amount of background signal is generated through unbound FP molecules. This large amount of background signal conceivably controlled by lesser FP expression (75). For DNA labeling, FP is fused to DNA binding domain e.g. Transcription factor. This DNA binding domain subsequently binds to introduced recombinant DNA molecule which has many specific binding sites (76)(77)(78).

### **2.6 GFP and its structure, Chromophore and its maturation**

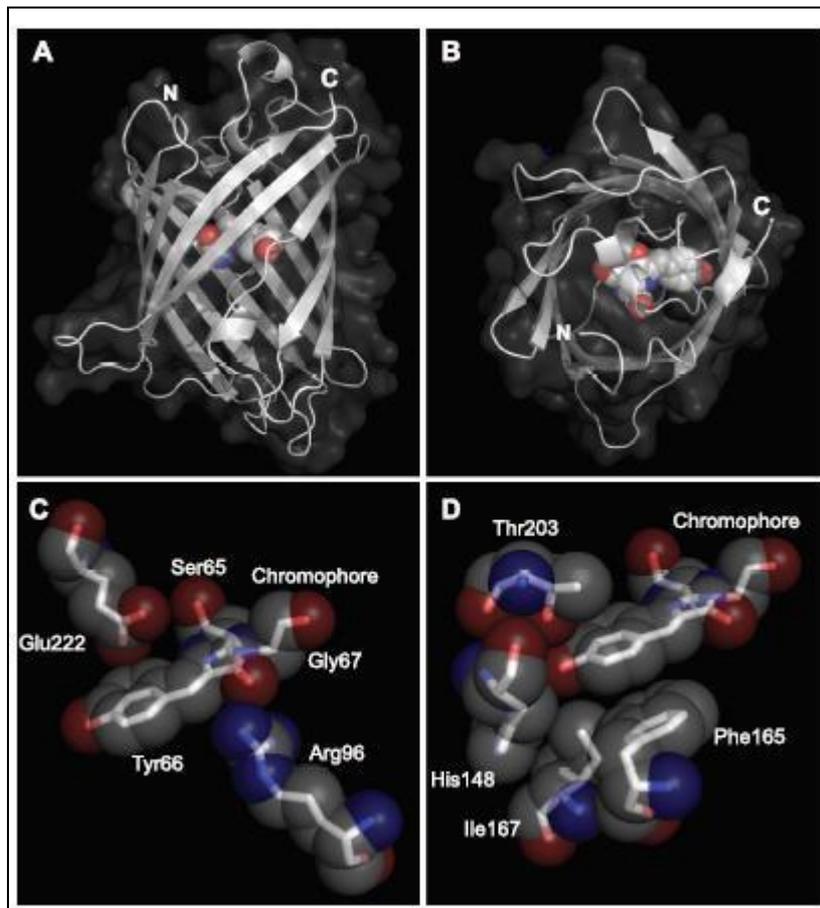
GFP is one of the most commonly used FP in the field of life sciences. GFP, a fully genetically encoded label, turned out to be unparallel tool which enables us the direct visualization of molecules, structures and processes in living cells and organisms. The most exciting thing about GFP gene is that it can be conveniently manipulated using standard molecular biology tools, which allowed us to create fusion constructs, recombinant proteins, and even transgenic organisms. GFP allowed us to label nearly every possible protein in living organisms that would otherwise be impossible to label using non genetically encoded fluorophores (79). GFP was discovered back in the early 1960s from *Aequorea victoria* jellyfish by Osamu Shimomura.

He was originally studying the blue-light-emitting bioluminescent protein called aequorin in which he came to know about another protein which was eventually named the green fluorescent protein (24). Aequorin is a Ca<sup>2+</sup>-activated photoprotein that excites GFP, making

GFP emit green light; aequorin itself has been used as a sensor of intracellular calcium (80).

However, aequorin is not an essential part of the green-light-producing reaction.

GFP consists of 11  $\beta$ -sheets that form the  $\beta$ -barrel structure. This  $\beta$ -barrel is threaded by an alpha helix running up the axis of the cylinder (Fig 2.1). This  $\beta$  barrel structure get interwoven and form cylindrical structure called  $\beta$  can. In 1979, Shimomura showed that GFP contains a set of 3 amino acids which absorbs and emits light. He called this structure as chromophore. This chromophore is present in  $\alpha$ -helix and is buried almost perfectly in the center of the cylinder (Fig.2.1). This “Beta-can” provides a perfect microenvironment for chromophore maturation.



**Figure 2.1 Structure of GFP**

**A. Structure of GFP  $\beta$ -barrel from side**

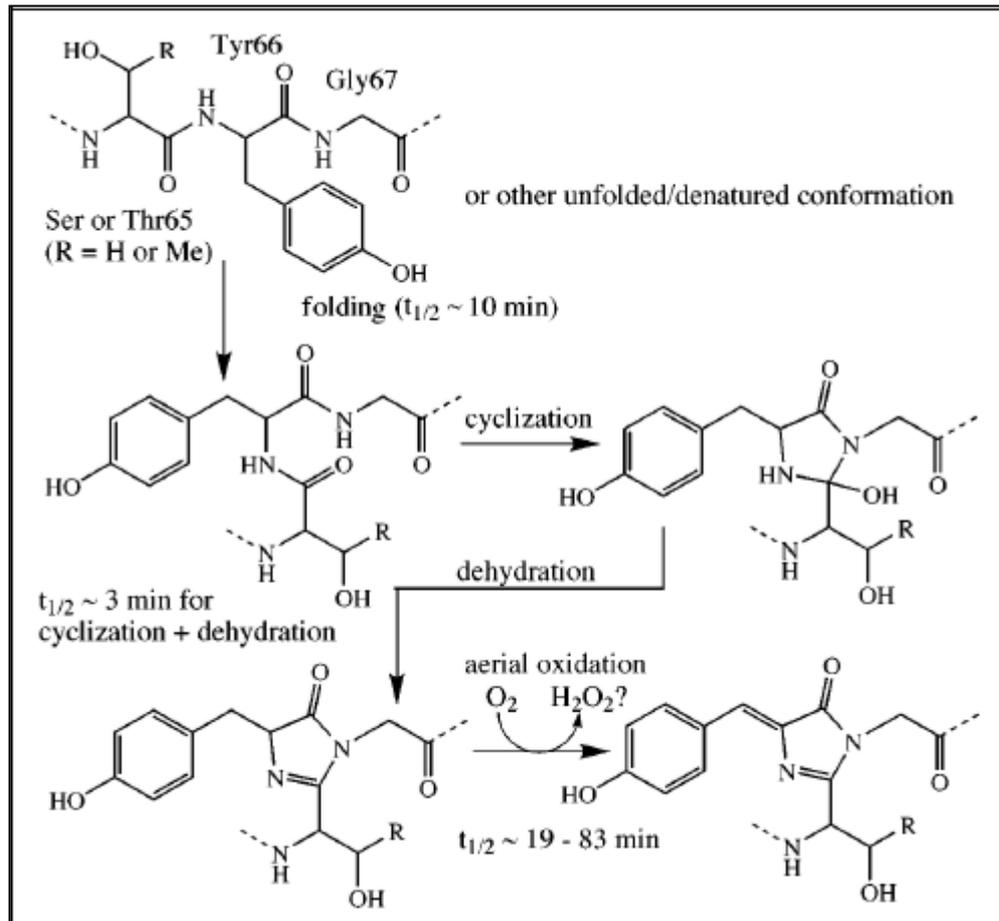
**B. structure of GFP  $\beta$ -barrel from top**

**C. Chromophore depicted in a spacefill presentation**

#### **D. GFP chromophore and selected nearby residues in sticks and semi-transparent spacefill representation.**

(Carbon atoms are shown in gray color, nitrogen atoms shown in blue color and oxygen atoms are shown in red color)

GFP is 25 KD protein having 238 amino acids, absorbs blue light (maximally at 395 nm with a minor peak at 470 nm) and emits green light (peak emission at 509 nm with a shoulder at 540 nm (81)(82). The chromophore of GFP is a p-hydroxy benzylidene imidazolinone formed from residues 65–67, which are Ser-Tyr-Gly. Fig 2.2 shows most widely accepted model for chromophore formation (82)(83)(84). Initially, GFP folds into native conformation. This is further followed by formation of imidazolinone due to the nucleophilic attack of Gly67 amide on the carbonyl of residue 65. This is followed by dehydration. Finally, molecular oxygen dehydrogenates the  $\alpha$ - $\beta$  bond of residue 66. This puts the aromatic group of Tyr66 into conjugation with the imidazolinone. Only after that, the chromophore acquires visible absorbance and fluorescence.



**Figure 2.2 Proposed mechanism for chromophore maturation (85)**

(Adapted from Roger Y. Tsien *Annu. Rev. Biochem.* 1998. 67:509–44)

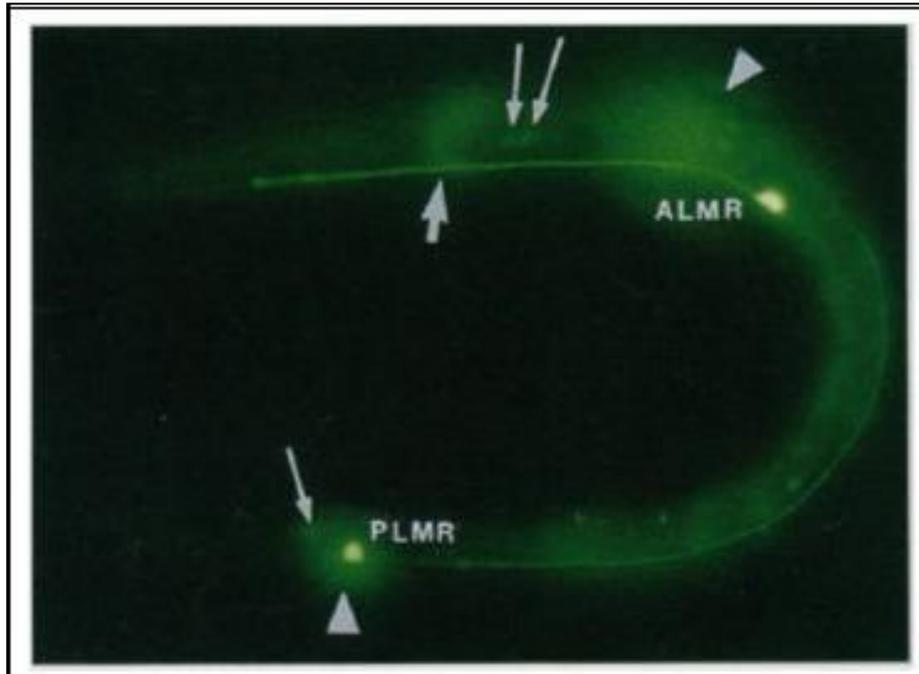
GFP was shown to express efficiently under the control of T7 promoter which results into easily detectable green fluorescence in bacteria (Fig 2.3). Green fluorescent bacteria were detected on plates that contained the inducer isopropyl- $\beta$ -D-thiogalactoside (IPTG) upon illumination with a long-wave ultraviolet (UV) light. These cells expressing GFP gene grew well in the continual presence of the inducer IPTG, This indicated that GFP is non-toxic to the cells (86).



**Figure 2.3 Expression of GFP in *E. coli*. The bacteria on the right side of the figure were transformed with GFP expression plasmid (86)**

**(Adapted from Douglas C. Prasher *Science*, New Series, Vol. 263, No. 5148, Feb 1994)**

In 1988, Martin Chalfie succeeded in introducing the GFP gene into the DNA of *Caenorhabditis elegans* (86). GFP was expressed by *C.elegans* cells, giving off green fluorescence without the addition of any extra components (Fig 2.4). It was also observed that expression of GFP protein did not cause any damage to cells upon its expression. Eventually it was found that GFP gene can be fused with genes of other proteins, opening-up the world of possibilities for study of various intracellular structures and mechanisms.



**Figure 2.4 Expression of GFP in *C.elegans* larva. Two touch receptor neurons (ALMR and PLMR) can be seen with their strongly fluorescing cell bodies (86)**

(Adapted from Douglas C. Prasher *Science*, New Series, Vol. 263, No. 5148 . Feb 1994)

## 2.7 Spatial Structure and Diversity of Chromophores

### 2.7.1 Structure

The structure of fluorescent protein is conserved having almost 220 to 240 a.a (25 KDa). Chromophore of a FP is buried and protected in the center of the  $\beta$  can. Chromophore residues and their position are mostly conserved in almost all FPs i.e. 65– 67. Chromophore of a FP is formed through distinct posttranslational modifications. The first residue of chromophore position 65 could differ while Tyr66 and Gly67 residues are mostly conserved in natural GFP-like proteins. Chromophore is the most sensitive part of fluorescent protein which is protected from external environment by means of Beta-barrel. Fluorescent protein structure is also stabilized by multiple non covalent interactions which provide thermal and chemical stability to fluorescent protein against denaturation (12)(11). Amino acids side chains, are buried deep into the FP protein molecule. These side chains of amino acids has

very important role in formation of chromophore as well as in determination of spectral characteristics of FP. Amino acid residues present in centre of beta strands play crucial role as they determines the chromophore surrounding region. Each beta strand controls chromophore in specific directions. Many residues have a very important role in normal functioning of fluorescent protein like Arg96 from strand 4, present in very close proximity of chromophore. Arg96 is also very crucial catalytic amino acid which is conserved amidst all Fluorescent proteins. Arg96 residue also involved in backbone cyclization of FP amid its maturation (87)(88)(89). While some amino acid residues are more resistant to mutations and does not change fluorescent properties significantly (90)(91)(92). The residues which are in the sidechains in contact with chromophore residue T66 control the properties like protonation state (anionic or neutral), polarization, rotational freedom (93)(94)(95)(96)(97)(98) and spatial conformation (Cis or Trans).

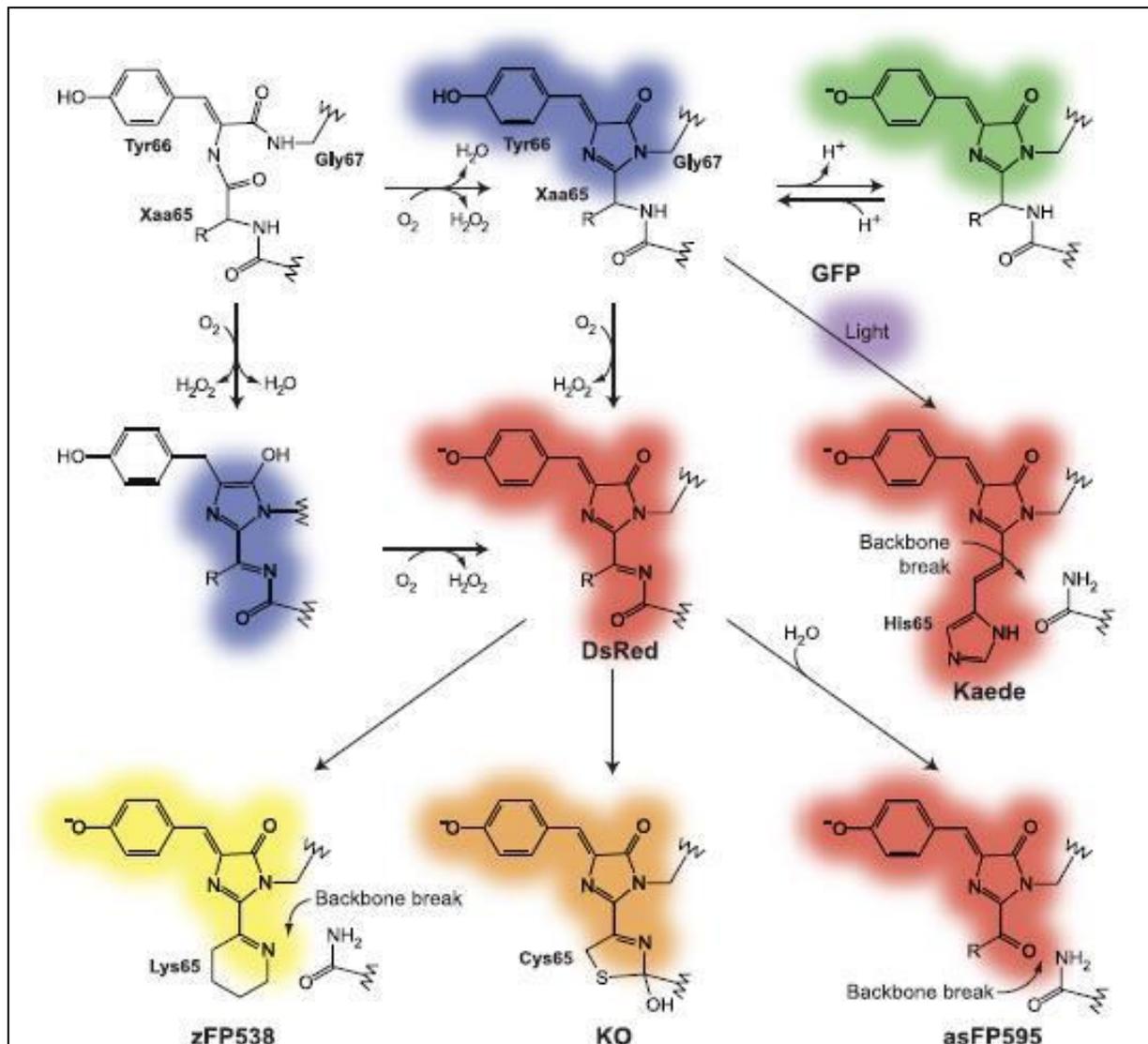
Mutations at these positions can significantly change the properties of fluorescent protein like change in excitation emission spectra, brightness of FP and its photoswitching behaviour. In the process of FP engineering, many useful mutational positions have been identified with their significance in different fluorescent properties. One of such mutation i.e. Thr203Tyr substitution in YFP which leads to a red-shift in excitation, Emission maxima (99). Photoswitching property of a photoconvertible FP, Dronpa is mostly controlled by positions 148, 165, 167, and 203 (100).

### **2.7.2 Natural diversity of chromophores**

GFP like fluorescent proteins which exist in nature among different organisms have been widely studied. However, even after studying the chromophore structure and its formation in these proteins in great detail, there are still disputes about the actual mechanism of chromophore maturation. GFP like proteins show great variation in their spectral properties like excitation emission spectra based upon their chromophore structure and chromophore surrounding environment, but their emission spectra is limited for up to 540nm. To further

red shift the emission spectra, GFP like chromophore needs to undergo covalent modifications. Two of such naturally occurring GFP like proteins with red emission spectra are DsRed and Kaede. DsRed chromophore is generated through oxidation of C $\alpha$ -N bond of chromophore amino acid 65. Oxidation of chromophore residue 65 leads to the extension of conjugated  $\pi$  system which causes a bathochromic shift of excitation-emission spectrum. Photoconvertible FP Kaede which shows change in emission spectra from green to red has different mechanism to give red fluorescence than that of DsRed. Kaede does not involve oxidation of residue 65 to give rise to red fluorescence, instead Kaede requires 405nm light irradiation to attain the red form. Kaede has histidine in chromophore at position 65. Initially, Kaede forms green chromophore which has equilibrium among protonated and deprotonated form. Irradiation to UV light leads to conversion of protonated chromophore into red form. Irradiation to UV light leads to cleavage of peptide between C $\alpha$  atom and amide of His65 which leads to formation of double bond between C  $\alpha$  and C $\beta$  of His65. Formation of double bond leads to extension of conjugated  $\pi$  system which causes bathochromic shift in excitation-emission spectra (Fig 2.5) Crystallographic studies reveal that chromophore has different conformational states along with diversity in its chemical structure. Most of FP chromophores exist in cis confirmation with few exceptions like part of red FPs, all the non-fluorescent GFP-like chromo proteins (101)(93)(102)(103). Most commonly found natural GFP like FP are of cyan, green, red and green to red Photoconvertible fluorescent protein. Excitation emission spectra for cyan fluorescent protein peaks are approximately 450nm and 485 nm, respectively (104). Cyan has GFP-like chromophore that interacts with nearby residues. Its chromophore also contains a buried water molecule in it because which its spectra is blue shifted (105)(106). While the chromophore of GFPs are in deprotonated form (107)(108). GFP shows red shifted spectra. It shows excitation at 480–510 nm, emission at 500–520 nm. Red fluorescent proteins have approximate Excitation emission spectra at 560–

580 and 570–610 nm, respectively (108). Photoconvertible FP Kaede has His65 in its chromophore and emission spectra peak around 620-630 nm (109).

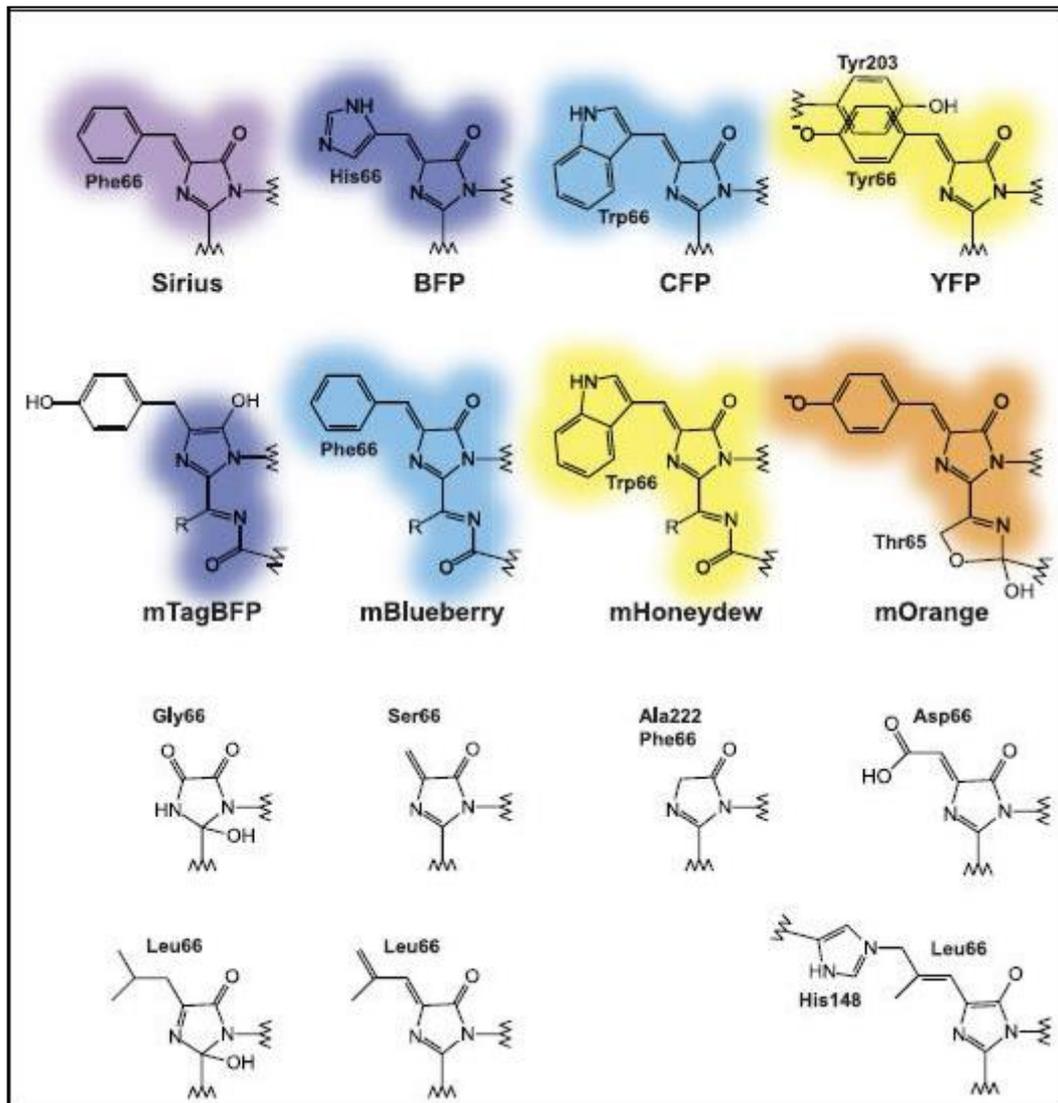


**Figure 2.5 Naturally occurring GFP- like protein chromophores structure and their pathway of maturation**

(Adapted from Konstantin A. Lukyanov, *Physiol Rev* 90: 1103–1163, 2010)

### **2.7.3 Artificial chromophores and colors**

Palette of fluorescent proteins has been increased by performing different mutagenesis experiments. Replacement of tyr66 residue of GFP chromophore with aromatic residues (phe, his, or trp) results in blue shift of fluorescence spectra. CFP, ECFP (110)(111), Cerulean(112), mTurquoise (113), TagCFP, EBFP (111)(114)(115), SBFP2(116), EBFP2 (117), Azurite (118) and FP Sirius (119) are few examples of such blue shifted FPs (Fig 2.6). Yellow fluorescent proteins were created from GFP by mutating the chromophore surrounding residue Thr203Tyr. This mutation caused red shift in emission spectra of GFP. mOrange is generated from DsRed by introducing Thr65 mutation which causes blue shift in emission spectra (120). Replacement of Tyr66 residue with that of nonaromatic residue causes loss of fluorescence in FPs (121)(122)(123)(124)(125). This extended fluorescent protein palette with varied excitation emission spectra enables us to understand complex biological system more efficiently.



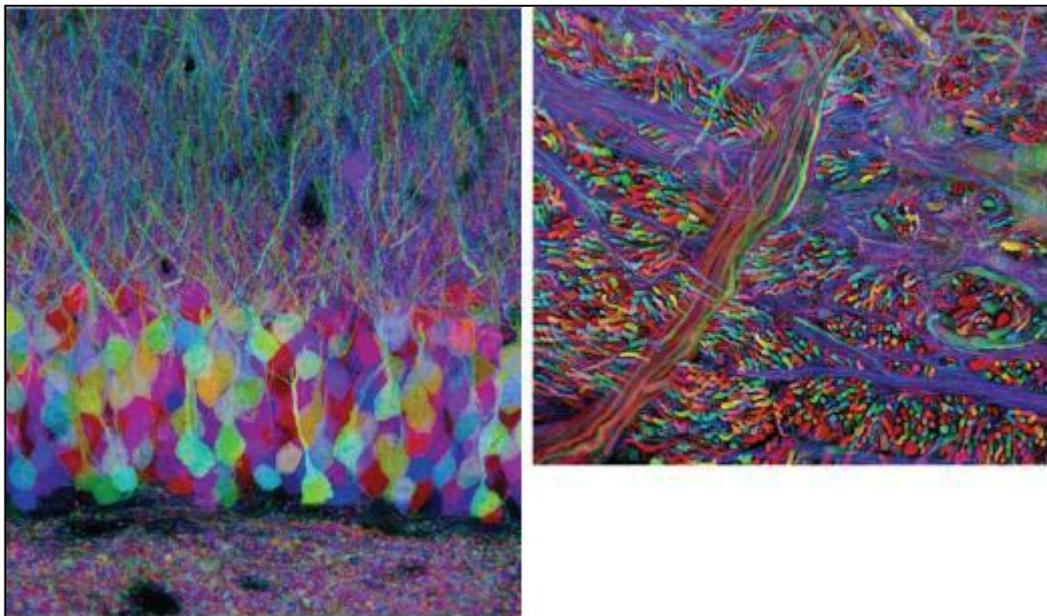
**Figure 2.6 Chemical structure of chromophore generated by mutating natural fluorescent proteins (1)**

(Adapted from Konstantin A. Lukyanov, *Physiol Rev* 90: 1103–1163, 2010)

## 2.8 Modern palette of FP

Continuous discovery and modification of available FPs gave us modern palette of FP which covers almost all the colors in visible spectra i.e. violet (emission peak at 424nm) to far-red (emission peak at 650 nm). This modern palette also includes smaller groups of FPs with large stokes shift. Such proteins which show larger stokes shift have special significance in

microscopy experiments. Near infrared is the only gap that remains to be filled in this palette. Near infrared FP with high brightness and high stability is useful in multicolor imaging and whole body labeling. Multicolor labeling with 6 fluorescent colors has been achieved by using this modern palette of FPs (126). Multicolor imaging is expanded up to 10 colors (Fig 2.7) by involving use of FP which has wide stokes shift like T-Sapphire, mKeima. It also involves use of RSFP like Dronpa (127), Padron (128), RsCherrys (129) etc. In this section we have discussed various monomeric fluorescent protein developed which proved to be good candidate for multicolor imaging.



**Figure 2.7 Multicolor imaging (Brainbow)**

(Adapted from Jeff W. Lichtman. Vol 450, November 2007, doi: 10.1038/nature06293)

### **2.8.1 Violet FP**

This group involves FPs which have excitation maxima at 355 nm. Emission maxima for violet FPs is 424 nm. Presence of Tyr66Phe mutations is the characteristic feature of violet fluorescent proteins (85)(82). Sirius was the first usable violet FP discovered, had emission maxima at 424 nm. Emission maxima of Sirius is lowest among known FPs (119). Sirius FP

can be used for experiment which require longer exposure or for study of acidic organelle because of its higher pH as well as photostability. This FP showed its utility in FRET studies in pair with blue cyan sensor. Low brightness and general toxicity by violet and UV light restricts the use of violet FPs. Cells also show auto fluorescence in violet spectrum of visible light which further reduces the usability of violet FPs (66). Brighter Sirius can be useful addition in palette of modern FP.

### **2.8.2 Blue FPs**

Blue FPs show excitation peaking at 380–400nm. Emission of blue FP peaks at 450 nm. EBFP was the only known fluorescent protein available in Blue FPs till date (115)(111)(114). EBFP is Tyr66His mutant variant of *Aequorea victoria* GFP. EBFP is good FRET candidate in pair with EGFP (111)(115)(130). Use of EBFP is limited by low brightness and low photo stability. Some improved Blue FPs have been developed from EBFP namely SBFP2 (116), Azurite(118), and EBFP2 (117). All these improved variants showed their utility as fusion protein and because of their monomeric nature. Azurite and EBFP2 are characterised by high photostability compared to EBFP. Recently developed TagBFP (131) showed higher photostability and higher brightness. This available palette of blue FP showed their utility in FRET, FCCS as well as multicolor imaging.

### **2.8.3 Cyan FPs**

Cyan FPs show excitation peaking at 430–460 and emission peaking at 480–490 nm. ECFP is one of the popular Cyan FPs for dual color imaging and FRET use with pair to YFPs. Cerulean (112), SCFPs (132), TagCFP, and mTurquoise (113) all these brighter cyan FPs with faster maturation rate were developed from ECFP through directed evolution. mTFP is one of such improved FP developed from tetrameric protein cFP484 (104). mTFP shows higher brightness and photostability compared to Cerulean. mTFP also has narrower emission spectra which reduces the crosstalk in multicolor imaging as well as FRET studies (133).

#### **2.8.4 Green FPs**

Green FPs show excitation peaking at 490 and emission peaking at 510 nm. This spectrum of fluorescent proteins includes many brighter, monomeric and photo stable FPs. EGFP is one of the first improved variant of GFP which has many optimised features. EGFP has one point mutation compared to GFP which allows it to remain monomeric even at higher concentration (507). There are many GFPs available with optimised fluorescent properties like Wasabi which shows higher brightness at expense of photostability and pH stability, Emerald (EmGFP) which shows faster maturation, super folder GFP that supports solubility of fusion protein although it has a higher tendency to get dimerize, TagGFP2 which shows high pH stability(134)(135)(136).

#### **2.8.5 Yellow FPs**

Yellow FPs show excitation peaking at 515 and emission peaking at 530 nm. EYFP is one of the first optimised yellow FP, which has lesser pH stability. Additionally, it is highly sensitive to halide ions. Citrine (137), Venus (138), Topaz (139), and TagYFP (140) are few improved yellow FPs. YPet is a brightest yellow FP. It is used in FRET studies in combination with cyan FP CyPet (141). But this pair shows homo and hetero dimerization tendency (142)

#### **2.8.6 Orange FPs**

Orange Fluorescent Proteins excitation maxima peaks around 550nm while emission maxima peaks around 560 nm. Orange and red spectra also possess abundance of fluorescent protein. Many fluorescent proteins have been optimised in this class. mKO and mOrange are two important Orange FPs developed from Anthozoa tetrameric FPs. Optimised version of these 2 FPs, i.e. mKO2 which shows faster maturation with low pH stability and mKO2 with higher photostability and slower maturation rate, are of great utility in protein labeling and many other applications.

#### **2.8.7 Red FPs**

Red FPs show excitation peaking at 560–590 and emission peaking at 580–610 nm. A large array of red fluorescent protein is available, out of which some also show emission in orange

spectra. TagRFP (143) is brightest red FP. mStrawberry and mRuby (144) shows high brightness but low photostability as compared to mCherry (145). Red fluorescent proteins makes good pair with yellow FP in FRET because of their high extinction coefficient. Although optimisation of RFPs is an ongoing process, the available palette of red FPs is capable of satisfying most of the applications.

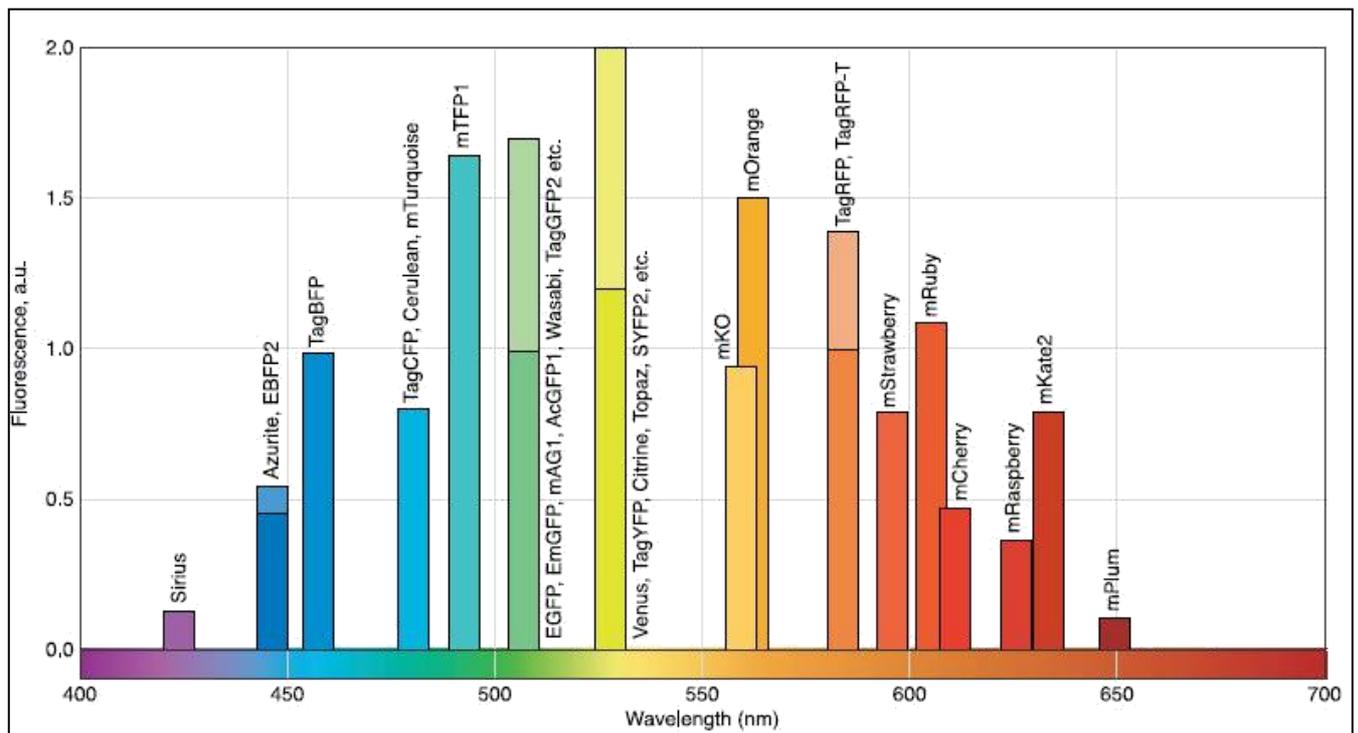
### **2.8.8 Far-red FPs**

Far-red FPs show excitation peaking at 590 and emission peaking at 630–650 nm. Far-red FPs show less background auto fluorescence, lesser photo toxicity and higher penetration capability in biological tissue sample. Far red FPs also have special significance in multicolor imaging and FRET. There are 4 Far-red FP namely mRaspberry (146), mPlum (146), mKate2 (147), and mNeptune (148). All these are monomeric Far-red FPs which are mostly used in imaging. Out of all these 4 Far-red FPs, mKate2 is brighter and more photostable, used for almost all applications including protein labeling. Another far red FP mNeptune is preferred in multicolor imaging in combination of orange FP. Scientist are continuously trying to expand the palette of Far-red FPs , but presence of brighter and photo stable Far-red FPs is not natural to fluorescent protein.

### **2.8.9 Fluorescent proteins with extended stokes shift**

Fluorescent proteins with extended stokes shift are characterised by difference of minimum 100nm between excitation and emission spectra. These fluorescent proteins comfortably used in combination with normal FP for fluorescence cross-correlation spectroscopy (FCCS) (149) as well as FRET studies. FPs with increased Stokes shift when used in combination with normal FP, it can show 2 color emissions by using single excitation wavelength. Similarly, we can also get single emission spectra using two different excitation wavelengths light. FPs with increased Stokes shift has protonated chromophore, which when exposed to excitation light undergoes excited state proton transfer which leads to transition of chromophore in charged form and this charged chromophore gives fluorescence emission of longer wavelength (150). Green FP Sapphire (151)(110), T-Sapphire (152), and yellow fluorescent protein mAmetrine (133) has

large stokes shift as they are excited by violet excitation spectra. Red FP mKeima shows largest stokes shift as it has excitation spectra peaked at 440nm and emission spectra peaked at 620nm. Most of the fluorescent properties of mKeima are suboptimal. So the optimised mKeima could be good option for multicolor imaging and FCCS study (149)(126). In conclusion, currently available fluorescent palette is almost sufficient (Fig 2.8) for most of applications. Still there is always scope of improvement in biophysical and biochemical properties of available fluorescent proteins which will ensure better use of these FP to understand the complex biological system.



**Figure 2.8 Spectral diversity of available monomeric FPs. Columns show positions of emission maxima and relative brightness of representative monomers(1)**

(Adapted from Konstantin A. Lukyanov, *Physiol Rev* 90: 1103–1163, 2010 )

### 2.8.10 Tandems

Monomeric nature of a fluorescent protein is important for their use in protein labeling. Still in some experimental setups, tandem dimers of fluorescent protein are used for protein labeling. Tandem dimer of a FP is generated by fusing 2 copies of dimeric fluorescent protein head to tail

with short and flexible linker protein (153)(154). Tandem FP forms intramolecular dimer which acts as a monomer of double size and it doesn't interact with other FP protein molecules. First tandem version was generated for HcRed (154) and DsRed (153) FP. TdTomato which is one of the most famous tandems shows higher brightness as well as higher photostability. Tandem variants have also been generated for Keima (149), red TurboRFP, far-red Katushka2 (147), RFP611 (144), RFP639 (144), and EosFP (155) FP. Tandem version of fluorescent protein have performed well in protein labeling. However, larger size of tandem can cause functional disruption of fused protein, therefore monomeric FP are always preferred in crucial protein labeling experiment (156).

### **2.8.11 Dimeric and Tetrameric FPs**

As discussed earlier, monomeric nature of a FP is very important characteristic in case of protein labeling. Other than protein labeling, monomeric nature of a FP is not absolutely necessary. Infact in many experiments, dimeric and tetrameric FP are preferred as they provide brighter and stable fluorescent signal. Several such FPs are popular for range of experiments. Here, we have mentioned some important dimeric and tetrameric FP:

1. Tetrameric FP ppluGFP2: It shows brighter and stable signal along with faster maturation rate. This protein forms needle shape aggregation within mammalian cells (157) under 48 hrs. of expression, which make this protein unsuitable for longer experiments.
2. Yellow FPs ZsYellow1 and TurboRFP: These proteins show bright and stable fluorescent signal and have faster maturation rate (14)
3. Red FP DsRed-Express2 and DsRed-Max: These show lower cytotoxicity in long term experiments (158)
4. Katushka (TurboFP635): It is useful in whole body imaging as it has highest emission spectra (147)(73).

### **2. 9 PTFP and its types**

Phototransformable fluorescent proteins is distinct class of FPs whose fluorescence changes on upon exposure to light of specific wavelength. Selective activation of PTFPs is possible

within area of interest. PTFPs are excellent FP label for proteins, cell organelles, cells and tissues. PTFPs provide higher signal to noise ratio, higher labeling and tracking precision which is helpful in confocal and super resolution imaging

### **2.9.1 Structural Basis of Photoactivation**

PTFPs Properties like its excitation emission spectra, extinction coefficient, and quantum yield; protonation state are strictly controlled by chromophore surrounding region. Cis-trans transition of chromophore along with conformational changes in chromophore surrounding residues which is a result of light induced photoconversion, changes the spectral properties of PTFP(100)(159)(160)(94)(161)(95). Cis-Trans isomerisation leads to reversible photo conversion in PTFP. Cleavage of protein backbone within chromophore or decarboxylation of the Glu222 causes irreversible photoconversion in PTFPs when exposed to excitation light (162)(163)(164)(165).

### **2.9.2 Key Properties of PTFPs**

Just like a normal fluorescent protein, the fluorescent properties like quantum yield, pH stability, maturation rate, excitation emission spectra, brightness, photostability, labeling density, absence of additional photoconversion, etc. are also crucial for PTFPs for their use in biological systems. PTFPs provide maximum contrast because the brightness of activated form is much higher than that of the initial form. Because of that PTFPs, generate higher signal to noise ratio (159). Higher the signal to noise ratio, better will be the signal. Intensity of light required to activate the molecule is also very crucial in case of PTFPs. In case needed light intensity is excessive, it will harm the cells, while if the prolonged activation is required then it will lengthen the experiment time and will not be suitable for fast track experiments. Contrasting to this, a very easy activation can cause undesirable photoconversion of PTFPs. So light intensity, continuity of light, light wavelength, protein mobility, zoom, irradiated field are also important factors for activation of PTFPs (166)(167)(168).

### 2.9.3 Irreversible PTFPs

Activation of irreversible PTFPs is one time event. Once activated the PTFP molecules could be bleached however its reversion to original state is impossible. Irreversible PTFPs either show activation from non-fluorescent form to their fluorescent form or conversion of fluorescent spectra from one color to another color. Mostly change in emission spectra occurs from cyan to green or green to red. Irreversible PTFPs are useful in different imaging experiments as they produce brighter signal for longer durations. Irreversible PTFPs are also useful in various photobleaching experiments like FRAP. There are numerous irreversible PTFPs that have been developed like PA-GFP, PS-CFP, PS-CFP2, Kaede, Dendra2, PAmRFP1, PAmCherry1, -2, and -3, mEos2, mEos3.2 etc.

PA-GFP is one of the first irreversible PTFPs reported. PA-GFP shows change from weak fluorescent form to bright green fluorescent form when exposed to activation light of 400nm. Uncontrolled activation is one of the drawback of PA-GFP because of its nature to easily get photoactivated. In spite of that, due to its monomeric nature, higher brightness and better contrast, it is one of the most important PTFPs (169)(170)(171)(172).

PS-CFP, PS-CFP2 are other PTFPs which show photoconversion from cyan to green fluorescent form when exposed to 405nm light. The initial cyan fluorescent form can be visualised without much photoconversion which help us to achieve better contrast. The overall brightness of activated PS-CFP, PS-CFP2 is comparatively lower to that of activated PA-GFP.

Kaede like proteins is another specific class of PTFPs which shows irreversible photoconversion from green fluorescent spectra to red fluorescent spectra upon exposure to UV light. When exposed to UV light Kaede like proteins undergoes cleavage of peptide bond of His65 which is first residue of chromophore. Cleavage of peptide bond causes extension of conjugated  $\pi$  system of Tyr66 into the imidazole ring of His65 causing a significant

Bathochromic shift (173)(174). Kaede provides better contrast and the green fluorescent form of Kaede can be visualised easily without any photoactivation. Kaede showed its utility in labeling cells, photobleaching experiments, protein tracking, superresolution experiments etc. (156)(175)(176). Dendra is also one of the popularly-used PTFPs which shows photoconversion when exposed to violet light as well as less harmful blue light.

Another important type of PTFPs is none to red PTFPs. PA-mRFP1, PAmCherry1, 2, and 3 (177) are examples of this class. These proteins provide higher brightness, better contrast and higher photostability which could be useful for 2 color PALM. IrisFP is a unique addition among PTFP developed from EosFP PTFP. IrisFP shows non reversible green to red photoconversion. Irreversible PTFPs can be used to perform diverse experiments. Irreversible PTFPs can produce the stable signal for longer time duration which can be used for tracking of POI. Irreversible PTFPs also can be used for organelle labeling, cell labeling (168)(178), to study protein interactions, protein degradation profile (179). Irreversible PTFPs also show their utility in superresolution microscopy.

#### **2.9.4 Reversible PTFPs**

Reversible PTFPs show increase or decrease in fluorescence intensity when expose to excitation light. Sometimes activation is from off state to fluorescent state (kindling). In other cases activation is from fluorescent state into off state (quenching). For controlled photoactivation, the intensity, wavelength and duration of irradiated light should be perfectly aligned to that of requirement. Known Reversible PTFPs are asFP595 and its mutant variants none to red kindling fluorescent protein (52)(100), Dronpa which is green to none quenchable Reversible PTFPs (180), mTFP0.7 that is cyan to none quenchable Reversible PTFPs (181), rsCherry and rsCherryRev which are none to red activable Reversible PTFPs (129). Reversible PTFPs because of their higher signal to noise ratio (182) can be used for tracking of rapid

protein movement (183). Reversible PTFPs can also be used for FRET studies, dual color imaging, super resolution studies, etc (184)(185).

### **2.9.5 PTFPs and Super resolution microscopy**

Optical microscopy resolution could achieve resolution up to a certain level called as diffraction limit of light. Because of the diffraction limit of light we cannot resolve two fluorescent protein molecule which are in a very close proximity. This phenomenon is called as point spread function (PSF). Numerical aperture of objective determines the PSF and thus resolution of optical microscopy is limited by PSF i.e. 200 nm laterally and 500nm axially. The objective lens with numerical aperture of  $<1.5$  can theoretically can achieve resolution upto 200nm laterally and 500nm axially.

Resolution achieved by optical microscopy is insufficient to study many subcellular structures and molecular interactions which demand the resolution at nanometerscale. Electron microscopy enables us to achieve resolution less than 100nm (186), but we cannot do live cell imaging in electron microscopy (EM). It is difficult to visualise the molecular interaction using EM because of its low labeling density.

In order to increase the resolution in light microscopic system, several new methods are developed like confocal microscopy, improvement in axial resolution by 2 objective lenses (187)(188), multiphoton excitation system (189). In spite of all these efforts, the axial resolution is not improved very significantly and light microscopy still restricted by lights diffraction limit. Superresolution microscopy has given a long awaited solution to these problems and allowed us to cross the limit of diffraction limit of light. Super resolution techniques either involve avoiding the diffraction limit significantly (near-field super-resolution imaging) or overcoming the diffraction limit of light (super-resolution imaging). Superresolution microscopy crosses diffraction limit by either modulating the molecular state

of fluorophore that narrows the PSF or by precisely locating the activated fluorophore molecule by processing of blurred images (175). Combination of both these approaches of superresolution microscopy along with use of PTFPs has revolutionised the fluorescence in vivo imaging through developing a method called PALM. There are different type of PALM like PALMIRA (190), FPALM (91)(113)(191), which efficiently cross the diffraction limit. PALM involves stochastic and gradual activation of PTFPs for which it involves the use of activation light with less power. This low power activation light is incapable of activating all PTFPs in one event. PALM involves multiple cycles causing the alternate activation and bleaching of the activated molecules. Stacks of images are produced from these cycles. These images are then processed in such way that each molecule can be localised precisely to determine its centre of emission. Localisation precision of activated molecule depends upon its quantum yield and photostability. PALM in ideal conditions allows us to achieve resolution unto 10nm precision. PA-GFP(56)(191), Dronpa (192) PS-CFP2 (192) mEos2 (156), Dendra2 (156), and PA-mCherry1 (177) are few important PTFPs molecules which are used in PALM

### **2.9.6 Photoconvertible fluorescent proteins**

Photoconvertible fluorescent proteins (PCFPs) belong to a unique sub class of Phototransformable fluorescent proteins as they show irreversible green to red photoconversion of emission spectra. Upon tagging a protein of interest with PCFPs, varied molecular dynamics and cellular events can be easily tracked. Newly synthesized proteins can be visualised and tracked upto its final destination using PCFP. They prove to be more suitable tags that can be used in various conventional and super-resolution imaging modalities, like wide field and confocal microscopy, structured illumination microscopy (SIM), and single-molecule localization microscopy (6,11,12). Kaede, KikGR, Eos and Dendra are few well known PCFPs. We have describe each one of them in detail below.

### **2.9.6.1 Dendra**

Dendra is a photoconvertible fluorescent protein that shows green to red photoconversion upon exposure to UV-violet spectral region. Dendra was developed from DendGFP obtained from octocoral *Dendronephthya* species. It was observed that DendGFP has histidine in its chromophore (His62-Tyr63-Gly64) like other photoconvertible fluorescent proteins. DendGFP also showed photoconversion from green to red form when irradiated with UV-Violet light. However, DendGFP was tetrameric in nature. Dendra was developed from DendGFP using protein engineering approach and was the first Photoconvertible fluorescent protein which simultaneously showed monomeric nature and maturation at 37°C(193). Excitation maxima for green and red fluorescent form of Dendra are 488nm and 556nm, respectively. Emission maxima for green and red fluorescent form of Dendra are 505nm and 575nm, respectively. Dendra has quantum yield 0.7 and 0.72 for its green and red emission spectra. Dendra has substantial high photostability, highly contrasting photoconversion and low phototoxic activation with 488nm light. Improved variants like Dendra2, moxDendra2, NijiFP, Dendra2-M159A, Dendra2-T69A were further developed from Dendra.

### **2.9.6.2 Kaede and KikGR**

Photoconvertible fluorescent protein Kaede was discovered in *Trachyphyllia geoffroyi*. Kaede showed yellow fluorescence along with green and red fluorescence. The chromophore of Kaede consists of tripeptide, His-Tyr-Gly. Kaede is tetrameric in nature. The QY of green and red form of Kaede is 0.88 and 0.33 respectively. It shows excitation maxima for green and red spectra at 508nm and 572 nm, while its emission maxima for green and red spectra are 518nm and 580nm, respectively (194).

Photoconvertible fluorescent protein KikGR was developed from Kaede by semi-rational mutagenesis of chromophore surrounding residues. KikGR showed more efficient

photoconversion in mammalian cells and its both green and red states are several folds brighter than Kaede. Green and red emission spectra of KikGR are largely separated. These separated green and red spectra enable more efficient detection of photoconversion (195).

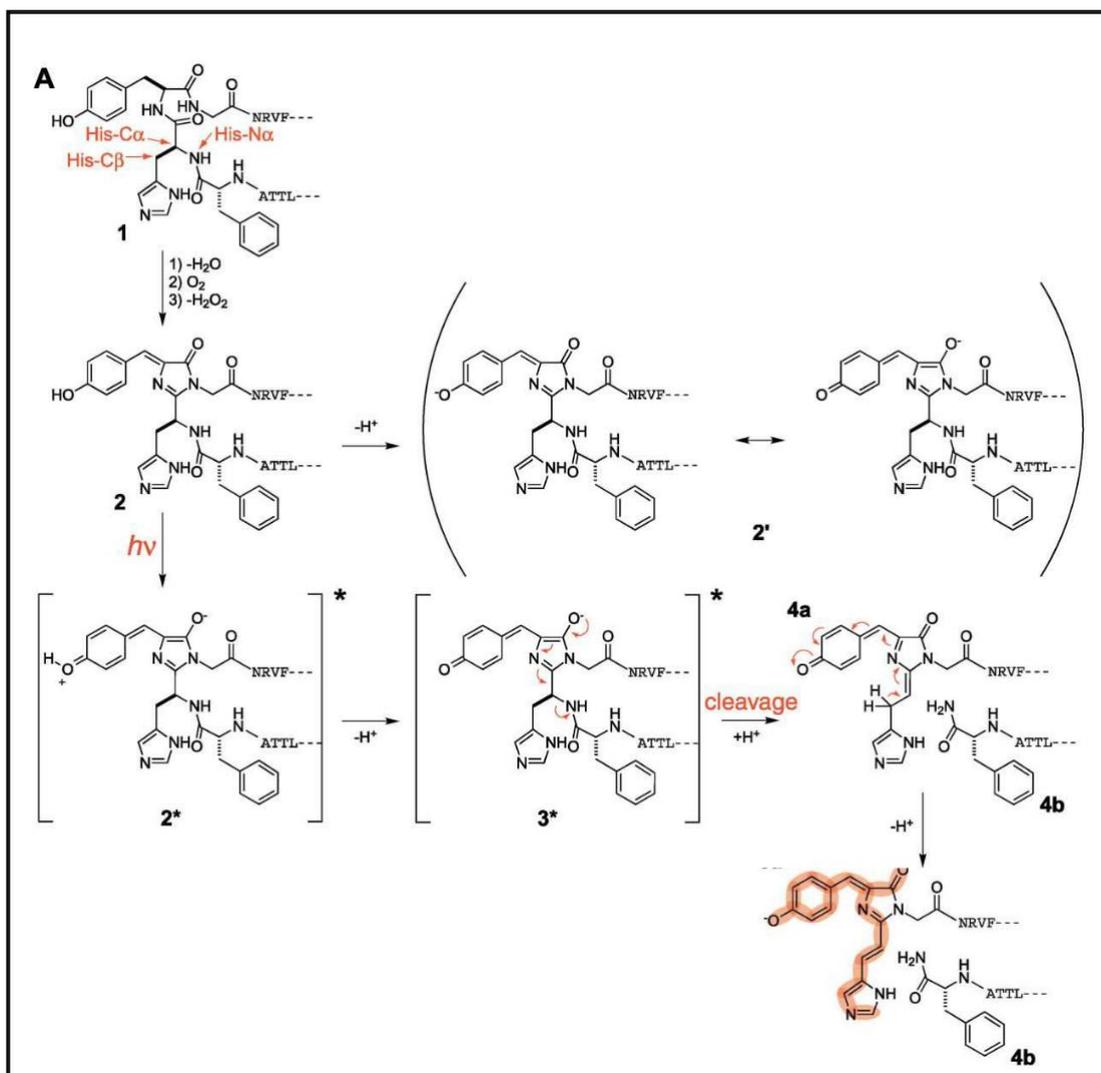
### **2.9.6.3 mClavGR2**

mClavGR2 is also a photoconvertible fluorescent protein which shows photoconversion from green to red spectra. mClavGR2 is developed from artificially synthesised PCFP mClavGR1, using combination of directed and random mutagenesis. mClavGR2 has 1.4 fold brighter after conversion red form compared to mClavGR1. mClavGR2 has 1.8 fold higher photoconversion contrast as compared to mClavGR1. mClavGR2 is completely monomeric in nature and shows augmented maturation of chromophore in *E.coli*. Excitation maxima for before and after photoconversion of mClavGR2 are 488nm and 566nm. Emission maxima for before and after photoconversion state of mClavGR2 are 504nm and 583nm, respectively. The QY for before and after photoconversion of mClavGR2 are 0.77 and 0.53, respectively. mClavGR2 showed utility in different biological studies like diffusion kinetics of membrane protein(196).

### **2.9.7 Molecular basis of green to red photoconversion**

All Photoconvertible fluorescent protein has conserved chromophore having sequence His-Tyr-Gly which is buried deep inside B barrel structure. The position of these amino acids is 62-64 which is also conserved. The first step in maturation of Photoconvertible FP is formation of green chromophore. This green chromophore is similar to that of *Aequorea victoria* derived GFP. Formation green chromophore requires properly folded protein and involves main chain cyclization event. We have shown formation and photo induced cleavage of chromophore of Photoconvertible FP kaede (Fig.2.9) (173). In green to red photoconversion along with chromophore amino acid i.e. Phe61, His62, Tyr63, amino acid Gly64 is also involved. First

nascent fluorescent protein is synthesised. This nascent fluorescent protein undergoes protein folding process which gives rise to a characteristic  $\beta$ -barrel structure. Further this protein is post-transnationally modified. The chromophore can be present either in neutral phenol form or anionic phenolate form. Local microenvironment as well as pH can change the equilibrium between these two forms. Species-1 molecule forms structure called 4-(p-hydroxybenzylidene)-5-imidazolinone which is species-2 molecule. The chromophore in species-2 molecule is neutral form and doesn't give any fluorescence. Deprotonation of hydroxyl group of Tyr63 results in a green-fluorescence emitting chromophore (2'). When these green fluorescence emitting neutral form chromophore exposed to UV light the excited state chromophore (2\*) releases proton to form the excited intermediate (3\*). Then cleavage occurs at the  $N\alpha$ - $C\alpha$  bond of His62 to eliminate a carboxamide group containing peptide (4b). The subsequent loss of a proton from His62C  $\beta$  gives a Trans double bond occurs between  $C\alpha$  and  $C\beta$  of His residue leading to the extension of the pi conjugation system of the imidazole ring in His (5a). When this modified chromophore is exposed to the red excitation light, it gives rise to red fluorescence. The mechanism of chromophore formation and its modification mechanism is conserved among PCFPs.



**Figure 2.9 Molecular mechanism for Green to red photoconversion (173)**

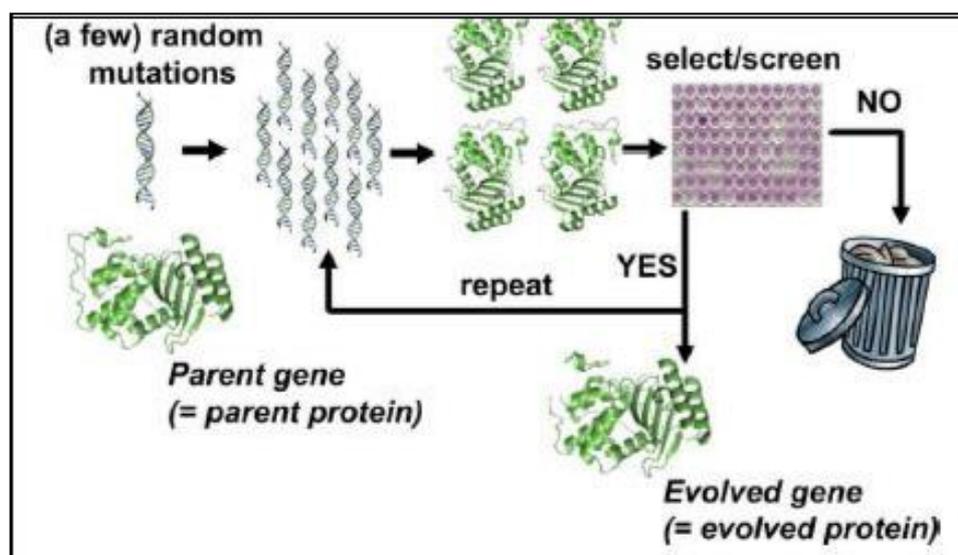
(Adapted from Atsushi Miyawaki, *Molecular Cell*, Vol. 12, 1051–1058, October, 2003)

### 2.10 Directed evolution

Natural evolution is one of the oldest biological process, existing on earth immediately after the emergence of life on earth. Genes are mutated which lead to the generation of improved proteins that help in better adaption to continuously challenging environment. For thousands of years, human has shaped the process of evolution by selectively breeding animals and plants with desired properties. Humans have evolved and optimised many enzymes and binding protein

unconsciously. Directed evolution of proteins is a man-made process based on molecular insights which speeds up the evolution process in laboratory setup. Directed evolution process relies on intended variations in protein sequence with certain level of randomness. This is followed by consciously designed selection and screening strategies.

Directed evolution is iterative process which involves steps like identification of starting state protein, diversification of its gene, its expression and screening (Fig 2.10). The process of gene diversification, expression and screening goes on again and again until a satisfactory and desired improvement is achieved in the targeted protein. Directed evolution is very well understood technique and widely used in chemical as well as pharmaceutical industries. Directed evolution has created many optimised enzymes with increased enzyme activity as well as higher tolerance to extreme enzymatic conditions.



**Figure 2.10. Schematic outline of a directed evolution process (197)**

**(Adapted from Frances H. Arnold, PNAS, June 16, 2009, vol. 10)**

The processes of directed evolution starts with the selection of a target gene of parent protein. This parent gene is then subjected to the process of diversification using

method like error-prone PCR or some similar technique. The library of mutant genes generated by gene diversification is then subjected for screening for the desired target property (e.g., improved brightness, portability of FP, improved activity, increased stability of enzymes). Those mutants which fail to show improvements in the screening are typically discarded, while the genes for the improved mutants are subjected as the parent genes for the next round of gene diversification and screening. This procedure is repeated until the evolved protein exhibits satisfactory improvement in target property (197).

Directed evolution is thought to be very effective in optimising the activities of genes and their products (proteins) which are the fundamentals of biology. The naturally occurring spontaneous mutation rate is insufficient to explore the desired variant of protein in time scale that is practical for laboratory evolution. Gene diversification techniques accelerate this process which helps us to get improved variants in shorter time scale. Gene diversification is the most important aspect of directed evolution followed by appropriate screening methods. The different methods of gene diversification and screening methods are described below (Table 2.2)

### **2.10.1 Random mutagenesis:**

Traditional random mutagenesis method involves the use of many mutagenesis agents, which randomly incorporates the mutations in template gene. This list of mutagenic agents involves alkylating compounds such as ethyl methane sulfonate (EMS), deaminating compounds such as nitrous acid, base analogues such as 2-aminopurine, and ultraviolet irradiation. These chemical mutagenic agents are less preferred in mutagenesis because of their biasness in mutational spectrum (198)(199)

Many mutator strain have also been used to randomly introduce mutations in template gene. These mutator strains contain deactivated DNA proof reading and DNA repairing enzymes like mutS, mutT and mutD (200)(201). But the problem with these mutator stains is that they not only mutate the library member but also induce deleterious mutations in the host genome.

These *in vivo* mutational methods offers low mutation rate and lack of control that is why *in vitro* mutagenesis methods like error prone PCR (ePCR) are preferred.

Error prone PCR involves the use of low fidelity DNA polymerase enzymes which generate point mutations during PCR amplification in gene of interest. Mutation rate can be further increased in Error prone PCR by increased magnesium concentrations, supplementation with manganese or the use of mutagenic dNTP analogues. These can reduce the base-pairing fidelity and increase the mutation rates. As the mutations accumulate in the DNA template in each cycles, the number of mutations can be increased by just increasing the number of PCR cycles. DNA polymerases used in epPCR can exhibit mutational biases. This mutational biasness can be countered by unbalanced dNTP concentrations and proprietary mixtures of polymerases (202)(203). Error prone PCR is very easy to follow and can give high mutation rate with fairly diverse mutational spectra.

### **2.10.2 Focused mutagenesis strategies:**

The structure of many proteins have been characterised at such resolution that we can easily target the residues involved in functional aspect like substrate binding, catalytic activity, etc. Use of synthetic oligonucleotide which involves one or more degenerative codon corresponding to target residues is the most straightforward approach in focused mutagenesis strategies. Focused mutagenesis strategy only involves those mutations or substitution which have a potential to give desired results. Simultaneous introduction of multiple mutations can allow us to access combination of mutation which can have epistatic interactions. Phylogenetic analysis of homologues proteins can provide sufficient clues about potential beneficial mutations. Molecular modelling can also be used in finding about such potential beneficial mutations (204). Algorithms such as Rosetta, which calculate free energies based on steric clashes, hydrophobic packing, hydrogen bonding and electrostatic interactions, can also be useful to improvise the target protein properties (205).

Gene diversification can also be achieved by homologous as well as non-homologous recombination.

Category	Examples and methods	Random or focused?	<i>In vivo</i> or <i>in vitro</i> ?	Advantages	Disadvantages
Chemical mutagenesis <sup>11,12</sup>	EMS, nitrous acid, ultraviolet irradiation and bisulfite	Random	<i>In vitro</i> and <i>in vivo</i>	Dose-dependent mutation rates	Low mutation rates; uneven mutational spectrum; hazardous chemicals
Mutator strains	XL1-red <i>E. coli</i> <sup>16</sup> , mutagenesis plasmid (PACE) <sup>122</sup> and yeast orthogonal replication <sup>18</sup>	Random	<i>In vivo</i>	Easy to use	Low mutation rates; uneven mutational spectrum
epPCR	Taq supplemented with Mg <sup>2+</sup> , Mn <sup>2+</sup> and/or unequal dNTPs <sup>23</sup> ; proprietary enzyme mixes (Mutazyme) <sup>24</sup>	Random	<i>In vitro</i>	Permits high mutation rates; easy to use commercial formulations; relatively even mutational spectrum	Random mutagenesis at the nucleotide level but does not evenly sample amino acid codon space
Site-directed saturation mutagenesis	NNK and NNS codons (where N can be any of the four nucleotides, K can be G or T, and S can be G or C) on mutagenic primers <sup>30</sup>	Focused	<i>In vitro</i>	Fully samples amino acid repertoire; focus on functionally relevant residues increases library quality	Requires structural or biochemical knowledge; excess of inactive clones within simultaneous saturation libraries
Computational strategies for high-quality library design	Rosetta design and computationally guided libraries <sup>35</sup> , ISOR <sup>36</sup> , consensus design <sup>31</sup> , REAP <sup>32</sup> and SCHEMA <sup>50</sup>	Focused	<i>In vitro</i>	Can create small libraries pre-enriched for functional variation by natural selection and/or <i>in silico</i> filtering	Requires structural data or phylogenetic data
Homologous recombination	DNA shuffling <sup>37</sup> , family shuffling <sup>45</sup> , StEP <sup>46</sup> , RACHITT <sup>38</sup> , NExT <sup>39</sup> , heritable recombination <sup>46</sup> , ADO <sup>43</sup> and synthetic shuffling <sup>42</sup>	NA	<i>In vitro</i> or <i>in vivo</i>	Can identify beneficial combinations of mutations or eliminate passenger mutations; can also shuffle sequences of orthologous proteins to repurpose functional diversity from nature	Rely heavily on sequence homology; evolved clones and natural orthologues can be divergent in nucleotide sequence
Non-homologous recombination	ITCHY <sup>48</sup> , SHIPREC <sup>47</sup> , NRR <sup>49</sup> , SISDC <sup>51</sup> and overlap extension PCR <sup>52,53</sup>	NA	<i>In vitro</i>	Capable of shuffling distantly related sequences, permuting order and combinatorial gene fusions	Highly effective in niche applications but challenging to implement as a general strategy

ADO, assembly of designed oligonucleotides; *E. coli*, *Escherichia coli*; EMS, ethyl methanesulfonate; epPCR, error-prone PCR; ISOR, incorporating synthetic oligonucleotides via gene reassembly; ITCHY, incremental truncation for the creation of hybrid enzymes; NA, not applicable; NExT, nucleotide exchange and excision technology; NRR, non-homologous random recombination; PACE, phage-assisted continuous evolution; RACHITT, random chimeragenesis on transient templates; REAP, reconstructed evolutionary adaptive path; SHIPREC, sequence homology-independent protein recombination; SISDC, sequence-independent site-directed chimeragenesis; StEP, staggered extension process.

**Table 2.2: Comparison and summary of different methods of gene diversification (206)**

(Adapted from David R. Liu , Nature Reviews Genetics, June 2015)

### 2.10.3 Screening methods for protein evolution:

There are 4 basic methods for screening in protein evolution (Fig.2.11). Method of screening in protein evolution can be customised on the basis of targeted property.

#### 2.10.3.1 Screening of spatially separated variant:

In this method of screening, clonally isolated variants are screened as colony using either liquid or solid media. While in case of calorimetric or fluorescent reporters, microtiter plate

reader can be used. Alternatively, methods like chromatography, mass spectrometry or nuclear magnetic resonance (NMR) can be used to screen the lysate for product formation.

### **2.10.3.2 High throughput screening using flow cytometry**

Measurement of fluorescence from individual cell and its separation into distinct subpopulation compartment can be achieved by Fluorescence-activated cell sorting (FACS).

### **2.10.3.3 Yeast display techniques**

Yeast display techniques enable us to perform FACS screening of protein–protein interactions, bond formation and peptide bond cleavage.

### **2.10.3.4 Screening the artificial cell-like compartments**

Using double emulsions or with polyelectrolyte shells in vitro compartments are formed. DNA, translated proteins and fluorogenic substrates can be entrapped in this invitro compartment which can be further subjected to fluorescence-activated sorting of functional variants.

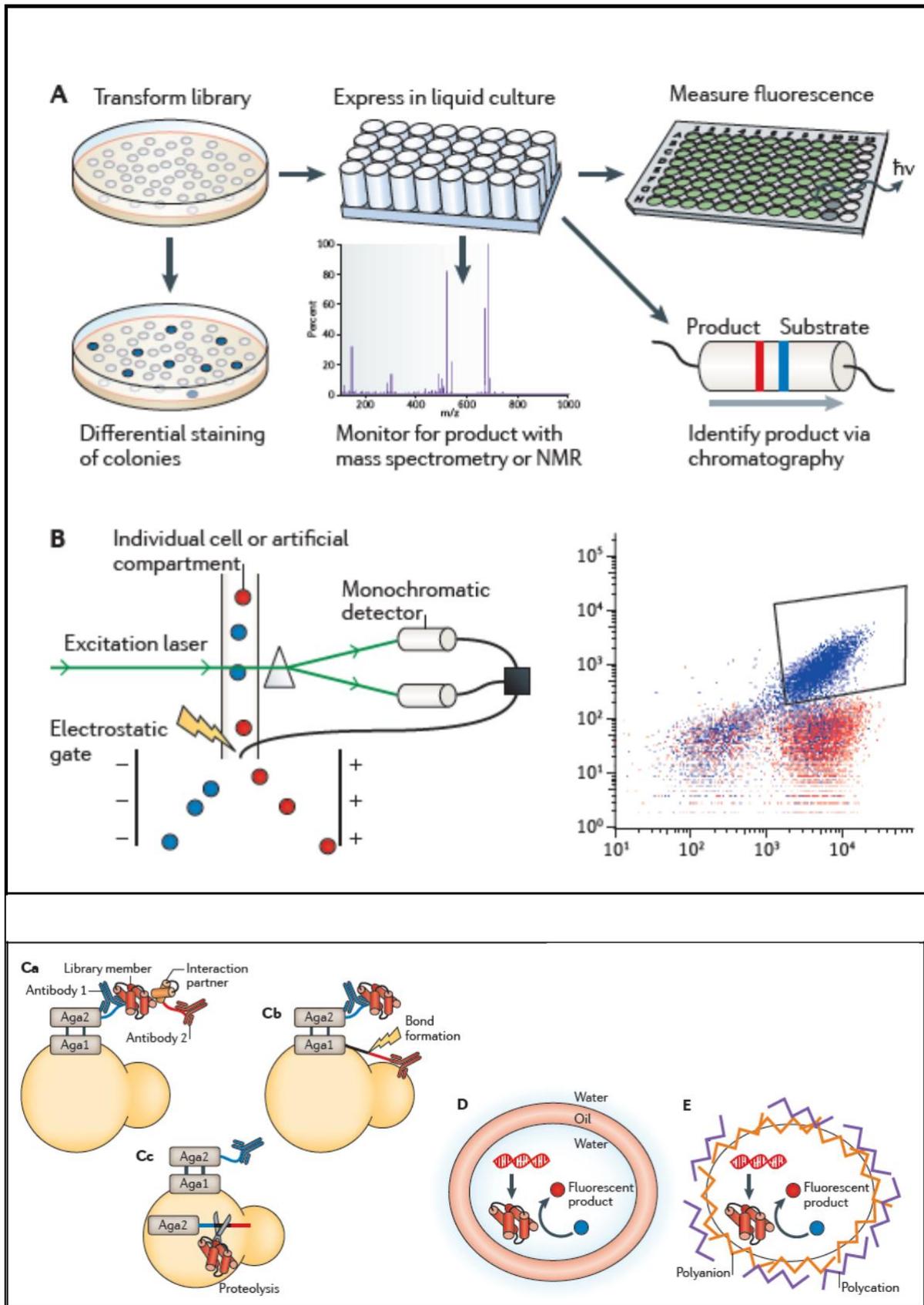


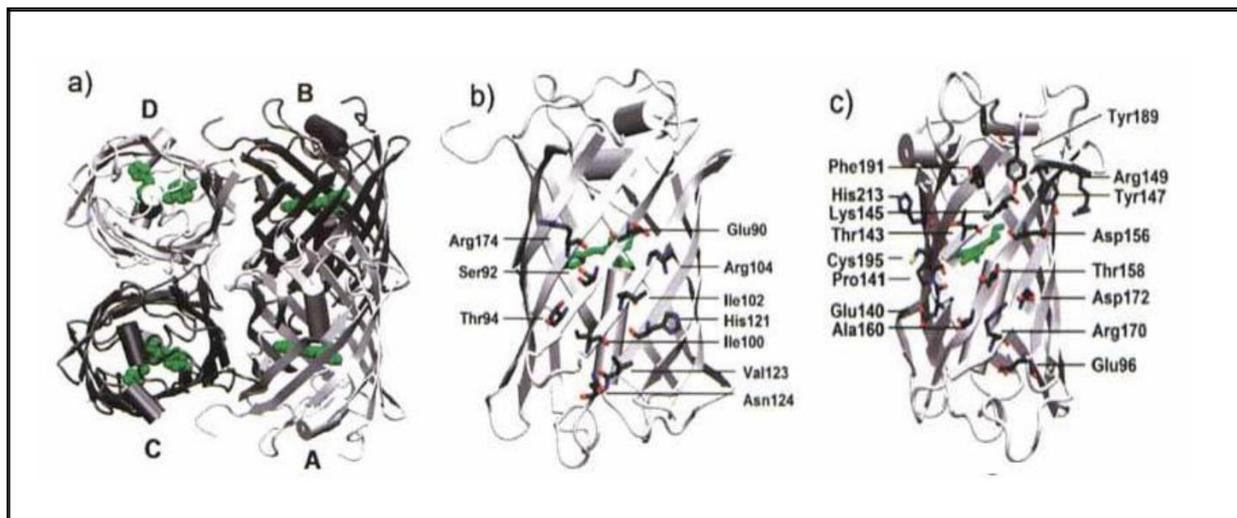
Figure 2.11 Screening methods for protein evolution

(Adapted from David R. Liu, Nature Reviews Genetics, June 2015)

**A. Screens of spatially separated variants, B. High throughput screening using flow cytometry, C. Yeast display techniques, D and E. Screening artificial cell-like compartments (206).**

### **2.11. Eos fluorescent protein and its Directed evolution**

Eos FP was first discovered in stony coral, *Lobophyllia hemprichii*. Eos FP is a Photoconvertible fluorescent protein which shows change in emission spectra i.e. green emission spectra into red emission spectra upon UV light exposure. Eos FP is tetrameric in nature, matures at 30°C, shows photobleaching in multiple steps and has comparatively low brightness (207). Eos FP consists of around 226 amino acid residues and has a molecular weight of 26.8 kDa. Eos FP shows around 84% sequence similarity to another Photoconvertible fluorescent protein i.e. Kaede. Chromophore of Eos FP is made up of three residues HYG, in which Histidine is important for its photoconversion property. Replacement of Histidine with that of M, S, T, and L yielded brightly green FPs with absence of photoconversion (207). Tetrameric structure of Eos FP consists of dimer of dimers (174). Each of this subunit has characteristic  $\beta$  barrel structure made of 11 beta sheets with central alpha helix. Each of this subunit has contact with 2 other subunits. The “Beta-can” of A chain is in antiparallel orientation with B chain similarly C chain is in antiparallel orientation to D, which result in mesh like layering of beta-sheets in contact area (Fig.2.12). Identical residues are present for both subunit in A/B interface which is stabilized by  $\pi$ -stacking interactions between the H121 imidazole side chains (distance 3.2 (Å)) and hydrophobic interactions between I100, I102 and V123.



**Figure 2.12 Tetrameric quaternary structure of wtEosFP**

(Adapted from G. Ulrich Nienhaus *Photochemistry and Photobiology*, 2006, 82: 351-358)

(a) Subunits are depicted by ribbon diagrams; chromophores in each subunit are depicted in green. (b) View of the A/B interface, with the residues forming the interface displayed in stick representation. (c) View of the A/C interface (155).

The green chromophore of Eos FP consist of 4-(p-hydroxybenzylidene)-5-imidazolinone moiety which is formed from its tripeptide residues histidine-tyrosine-glycine. Red chromophore is made up of structure 2-[(1E)-2-(5-imidazolyl) ethenyl] 4-(p-hydroxybenzylidene)-5-imidazolinone chromophore (208). Crystal structure of EosFP suggests that glutamate E212 acts as the base for proton abstraction from H62-C $\beta$  which is very crucial for photoconversion. Replacement of this glutamate by glutamine causes the loss of photoconversion ability in EosFP.

EosFP has been subjected to multiple steps of directed evolution. In first step of directed evolution of Eos FP, its oligomerisation problem has been resolved. Despite the high stability of tetrameric EosFP, its monomerisation was simple. Single point mutations were introduced

at each A/B and A/C interface to achieve functional dimers. V123 was replaced by hydrophilic threonine residue to break A/B interface, while T158 was replaced by histidine to disrupt the A/C interface. Combining both these mutations i.e. V123T-T158H resulted into functional monomer (3). This resultant protein was named monomeric Eos (mEos).

In next step of Eos FP optimisation, its maturation which earlier happened at 30°C was changed to 37°C. Because of the maturation of Eos protein at 30°C, it was unsuitable for mammalian cell use. For this optimisation process, side grafting strategy was followed. It was found that Eos protein was very much homologous to photoswitchable fluorescent protein Dronpa, which is very bright, photostable, monomeric at matured at 37°C. Structural alignment was performed between Eos and Dronpa FP to find the divergent amino acid positions in Dronpa which contribute to its monomeric nature and thermostability. From this study 28 positions of amino acids were selected for mutagenesis and in result three mutations (N11K, E70K, H74N) were found to rescue fluorescence at 37°C in bacterial cell. Surprisingly, none of these mutations was present at dimer-dimer interface. All these mutations were found to be improving the secondary structure preference. N11K, E70K mutations were found to improve electrostatics. Another mutation His121 to Tyr was found to be helping in protein folding at

37°C, optimising the  $\beta$ -strand preference and disfavours dimerization. These four mutations turned Eos protein into a potential green to red photoconvertible fluorescence protein which matured at 37°C. They named this protein as mEos2 protein (209). However, mEos2 protein still had a tendency to oligomerize when expressed at higher concentrations. This problem was resolved in next step of optimisation.

In this step of optimisation, mEos3.1 and mEos3.2 variants were generated from mEos2 protein. mEos3.1 and mEos3.2, have I102N, I157V, H158E, Y189A and I102N, H158E, Y189A mutations, respectively, relative to mEos2 sequence. It was found that mEos3.1 and mEos3.2 variants showed improvement in maturation, photon budget pH stability, labeling density, brightness etc (5). Most of the drawbacks of Eos protein have been resolved in

mEos3.1 and mEos3.2, except its low brightness for after photoconversion spectra. In the most recent study mEos4.1 and mEos4.2 protein variants were developed from mEos2 protein. mEos4.1 and mEos4.2 protein variants are resistant to chemical fixation and works well even in heavily fixed (0.5–1% OsO<sub>4</sub>) samples. mEos4.1 and mEos4.2 protein variants showed their utility in electron microscopy and correlative SRM (210).

# **Chapter 3**

## **Aims and Objectives**



### **3.1 Hypothesis**

mEos3.2 protein is a one of the widely known Photoconvertible fluorescent protein which shows irreversible photoconversion from green to red emission spectra upon exposure of UV light. The most primitive form of mEos3.2 is Eos protein that was first discovered in stony coral *Lobophyllia hemprichii* (3). This EosP had many limitations for its use in mammalian cells, like its tetrameric nature, maturation at 30°C and lower brightness. Multiple rounds of directed

evolution in EosP led to the creation of mEos3.2 protein which was monomeric in nature and matured at 37°C (3)(5). Although, most of the biophysical properties of mEos3.2 protein are good enough to perform different cell biology experiments; its relatively lower brightness for red (after photoconversion) spectra which limits its application in high-speed SRM. We hypothesized that this limitation of mEos3.2 protein could be targeted using semirational protein engineering approach in order to create the brightness improved variants of this protein. We believe that, these brightness improved variants would serve as a better tool for confocal and super-resolution microscopy.

### **3.2 Objectives:**

- I. To create brightness improved variants of mEos3.2 protein
- II. Validate the performance of improved variants into mammalian cells and their Biophysical characterisation

### **3.3 Detailed objectives:**

#### **Objective 1: To create brightness improved variants of mEos3.2 protein**

- A. Creation of a mutant library of mEos3.2 protein
- B. Screening of entire mutant library to get the improved variant

#### **Objective 2: Validate the performance of improved variants into mammalian cells and their Biophysical characterisation**

- A. Validate the performance of improved variants of mEos3.2 protein into mammalian cells using confocal microscopy
- B. Purification and spectral characterisation of improved variants of mEos3.2 protein
- C. Insilco modelling of improved variants of mEos3.2 protein
- D. Biophysical characterization of the improved variants of mEos3.2 protein
- E. Validate the performance of improved variants of mEos3.2 protein in superresolution microscopy setup PALM and SIM

### 3.4 Work done:

The detailed description of the work carried out under the above mentioned objectives are presented as a chapter 5 under following heading (chapter 5 and 6).

**Chapter 5:** mEosBrite are improved variants of mEos3.2 developed by semirational protein engineering.

Chapter 5 consists of Introduction, Results, and Discussion.

**Chapter 6:** Summary and conclusion of both the objectives

# Chapter 4

## Materials and Methods



## **4.1: Molecular biology methods**

**Host strain:** *E. coli* DH5 $\alpha$

**Luria-Bertani (LB) (HiMedia) medium:** Luria Broth powder (20g) was first dissolved in 800 ml deionized milliQ (D/W) and then final volume was adjusted to 1 litre using milliQ. The media was sterilized by autoclaving. For making LB-agar plates, 20g bacteriological grade agar powder was added into 1 litre LB media and sterilized by autoclaving. Average 40 ml sterile LB-agar media was poured in 90 mm sterile plates.

### **Antibiotics:**

Functional Concentrations of Ampicillin and Kanamycin used in LB-media is 50 $\mu$ g/ml and 30 $\mu$ g/ml, respectively.

### **4.1.1: Preparation of ultra-competent *E. coli*:**

High transformation efficiency can be achieved by using *E. coli* cells with higher competency which can be helpful in cloning. Ultra-competent DH5 $\alpha$  was made for its use in transformation of recombinant/routine plasmid vectors.

**Super-optimal broth (SOB):** SOB was prepared by dissolving 2% Bactopectone (HiMedia), 0.5% yeast extract (HiMedia), 10mM NaCl (Merck), 2.5mM KCL (Merck), 10mM MgCl<sub>2</sub>, in milliQ, and autoclaved for sterilization.

**Super-optimal catabolite (SOC) media:** SOC was prepared by adding filter sterilized 2M glucose and autoclaved 2M MgCl<sub>2</sub> to 98 ml of sterile SOB.

**Transformation buffer (TB):** Transformation buffer was prepared by adding 10 mM PIPES (Sigma), 15mM CaCl<sub>2</sub>, 250mM KCl, 55mM MnCl<sub>2</sub> into 100 ml of D/W. Transformation buffer pH was adjusted to 6.7 using 5N KOH. Finally Transformation buffer was sterilized by passing it through 0.2  $\mu$  membrane filter.

**Protocol for ultra-competent *E. coli* preparation:** First the *E. coli* DH5 $\alpha$  cells were streaked on LB agar plate and the plate was incubated overnight at 37<sup>0</sup>C. A single isolated colony of *E. coli* from this LB agar plate was inoculated into 250 ml SOB medium and further incubated in

refrigerated shaker incubator with 200 RPM at 18°C until OD<sub>600</sub> reached to 0.6. Next, the culture bearing flask was incubated on ice for 10 minutes and then spun in refrigerated centrifugation machine at 2500g (3500 RPM) for 10 minutes at 4°C. The pellet we get after centrifugation was then resuspended very gently in 80 ml of ice cold transformation buffer and again kept on ice for another 10 minutes. Mixture was then spun in centrifugation machine at 2500 x g (3500 RPM) for 10 minutes at 4°C. The cells from the pellet were again resuspended gently in 20 ml of ice cold transformation buffer and incubated on ice for next 10 minutes. 7% (1.4 ml) of DMSO was added to the cell suspension in transformation buffer and mixed gently by pipetting up and down. 100µL aliquot of cells was prepared in Eppendorf vials which were snap frozen by using liquid nitrogen and stored at -80°C.

#### **4.1.2: Bacterial Transformation (211)**

For bacterial transformation, ultra-competent cells (100µL aliquots) were taken which were stored at -80°C and subjected for thawing on ice. DNA sample of around 10µl (50-100 ng) was added to 100µl thawed competent cells. Cells were incubated on ice for 10 minutes. Heat shock was given to the cells in eppendorf at 42°C in water-bath for 45 seconds. The cells were immediately placed on ice for next 5 minutes. After that 5 mins, 200µl of ice cold SOC medium was added to the vial aseptically. Next, the cells were incubated at 37°C in shaking incubator at 180 RPM for 20 minutes. Finally cells were plated on LB agar plate containing appropriate antibiotic and plate was incubated at 37°C for 12-16 hours for colonies to appear.

#### **4.1.3: Plasmid DNA isolation**

Different methods were used to isolate Plasmid DNA.

##### **4.1.3.1: QIAprep Spin Miniprep method**

Reagent-Qiagen miniprep kit, as per the manufacturer's protocol.

QIAprep Spin Columns has a unique silica membrane which is capable of binding up to 20 µg DNA. Binding of DNA is assisted by high concentration of chaotropic salt and elution of bound DNA was performed by using small volume of low-salt elution buffer.

In first step, single colony of E.coli cell expressing plasmid of interest was inoculated in 10ml of LB-amp/ LB-kan media and incubated it at 37°C for 12-16 hours at 180 RPM in bacterial shaker incubator. The bacterial culture was then transferred into a 15ml tube and subjected for centrifugation at room temperature for 5 minutes at 5000 RPM. The pellet obtained was vortexed briefly and then resuspend in 250µl of Buffer P1 (RNase A has been added to Buffer P1) and transferred to a microcentrifuge tube. 250µl of buffer P2 was added into tube and inverted it for 4–6 times. Immediately after inverting tube for several times, 350µl of buffer N3 was added and the tube was again inverted for 4–6 times. The lysate was subjected for centrifugation at 13,000 RPM for 10 minutes. After centrifugation clear supernatant from the tube was very carefully transferred to QIAprep spin column. This column was centrifuged for 30–60 seconds at low speed. Flow through was discarded. In next step, 500µl of buffer PE was added in QIAprep spin column to wash the column and centrifuged at 13000 RPM for 1 minute. Residual buffer PE was removed by wiping the column from outside. Then column was placed in a dry tube place and spin at 13000 RPM for 2 minutes to remove residual wash buffer. Then, QIAprep column was place in a clean 1.5 ml microcentrifuge tube. 50µl Buffer EB (pre-warmed at 65°C) was added to the centre of the QIAprep spin column, kept it for 2 minutes. Finally, the tube was centrifuged for 2 minutes at 14000 RPM to elute Plasmid DNA.

#### **4.1.3.2: Plasmid DNA isolation using TELT buffer**

TELT method is used for regular screening for positive clones in the cloning experiments.

This is a quick and cost-effective method for isolating plasmid DNA

TELT buffer preparation [50mM Tris-Cl (Sigma) pH7.5, 62.5mM EDTA (Fischer Scientific) pH8, 0.4% Triton X100(Sigma), 2.5M LiCl (Sigma)], Lysozyme (Sigma) (50mg/ml), 70% ethanol, Absolute alcohol (Merck), TE buffer.

The single isolated bacterial colony of E.coli was inoculated into 1.5ml LB-antibiotic (Amp/Kana) media and incubated at 37°C, for 12-16 hours at 200 RPM. The bacterial culture was then subjected for centrifugation at 14000 RPM for 1 minute at 4°C. After the centrifugation, supernatant was discarded while the cell pellet was resuspended in 150µl

TELT buffer and vortexed briefly. 5.7µl lysozyme (Stock 50mg/ml) was added to the same vial and mixed well. The vial was incubated on ice for 1 minute. Next, the vial was kept in boiling water bath for 1 minute. Immediately after that the vial was placed on ice for 10 minutes. After that, the vial is subjected for centrifugation at 4°C with 15000 RPM for 10 minutes. After centrifugation supernatant was collected in a new vial. Next, 330µl ice chilled absolute alcohol is added in vial and incubated it at 80°C for 30 minutes. Vial was subjected for another round of centrifugation at 4°C with 15000 RPM for 10 minutes. In next step 200µl chilled 70% ethanol was added for washing the DNA pellet and again centrifuged it at 15000 RPM at 4°C for 5min. The pellet was dried by removing all the remaining alcohol and re-suspended it in 20µl TE buffer

#### **4.1.4: Agarose gel electrophoresis**

Analysis and preparation of DNA molecules was performed by using Agarose gel electrophoresis method. DNA fragments can be separated on the basis of their size using agarose gels of different concentrations.

#### **Different reagent used in agarose gel electrophoresis**

Ethidium bromide 0.5 µg/ml

Sodium Borate (SB) buffer: 10mM NaOH pH 8.5 adjusted with boric acid for 1X SB buffer.

6X Gel loading dye: 1.2ml glycerol, 1.2ml 0.3mM EDTA, 300µl of 20% SDS, 160 µl of 0.5% Bromophenol blue stock, nuclease free water to make up volume to 10ml.

The percentage of agarose gel used depends upon the size of DNA fragment needed to be resolved. For preparation of agarose gel, Agarose powder was weighted as per requirement of percentage of gels for example, to make a 0.8% agarose gel, 0.48g of agarose powder was added in a glass flask, to which 60ml of 1X sodium borate (SB) buffer was added. The agarose powder in the mixture was melted by microwaving for 2 minutes. Boiling agarose mixture was allowed to cool down. Once it cooled down, ethidium bromide (to visualize DNA) was added to this mixture at a final concentration of 1µg/ml (stock 10mg/ml) and

mixed well without creating bubbles. The final mixture was poured into the gel tray and comb was placed properly to create the wells. The comb was removed once the gel solidified. 1X SB buffer (running buffer) was poured into to the tank containing agarose gel. DNA sample (Plasmid DNA, PCR fragments, restriction digestion fragments, ligated DNA) was diluted with 6X gel loading dye (to make a final concentration of 1X). The size of target DNA fragments can be understood by running Standard 1Kb or 100bp ladders parallel. DNA bands were visualized using gel documentation system.

#### **4.1.5: Polymerase chain reaction (PCR)**

Specific DNA amplification of the sequence of interest from a template (plasmid DNA /cDNA) can be achieved by PCR technique with the help of two oligonucleotide primers which bind to the opposite strands in a sequence-specific manner. The extension of these primers at 3' end can be performed by using thermostable DNA polymerase. High fidelity DNA polymerase enzyme Phusion was used for PCR amplification process.

1. A typical PCR reaction mixture includes the following additive

	<b>Components</b>	<b>Final concentration</b>
1	H <sub>2</sub> O	To make up the volume
2	5X buffer HF/GC	1X
3	10mM dNTP mixture	200 $\mu$ M
4	Forward primer	0.5 $\mu$ M
5	Reverse primer	0.5 $\mu$ M
6	Template DNA	50ng (Plasmid DNA)/100ng (cDNA)
7	DNA polymerases	0.02 U/ $\mu$ l

**Table 4.1 Contents of PCR reaction mixture**

All the reagents and samples were kept in ice for thawing. Once reagents and sample get thawed, they were added as per the order stated in the table above to create reaction mixture. This reaction mixture in PCR tube was then subjected for short spin. PCR tube having reaction mixture in them were then quickly transferred to the thermocycler preheated to the denaturation temperature (98°C) so as to start the reaction.

Typical PCR cycle:

Cycle step	Temperature	Time	Cycles
Initial denaturation	98 °C	3 minutes	1
Denaturation	98 °C	30 seconds	
Annealing	Lower $T_m+3$ °C	30 seconds	30-34
Extension	72 °C	60 seconds/Kb	
Final extension	72 °C	8 minutes	1
Final hold	4 °C	forever	

**Table 4.2 Typical PCR cycle**

III. The PCR product was checked on an agarose gel.

#### **4.1.6: Gene Cloning (212)**

In first step of gene cloning, the plasmid and foreign DNA (which has gene of interest) were cleaved with one or more RE in order to get blunt/cohesive ends. In the second step, the gene of interest which was cleaved by one or more restriction enzyme, was inserted into digested plasmid by means of method called ligation. In third step, this ligated heterogeneous mix was transformed into a suitable bacterial host to propagate the clones. Finally, the resulting transformed recombinant clones were screened by Restriction digestion method to confirm the recombinant clone. Cloning of DNA fragment in plasmid vector can be executed by Different strategies like PCR based cloning, sticky end based directional cloning, etc.

##### **4.1.6.1: Restriction Digestion**

Restriction enzymes or restriction endonucleases has unique capability to cut the template DNA at specific site. This property of Restriction enzymes has been utilized in molecular cloning methods. The preparative and analytical restriction digestion reaction involves following components:

<b>Components</b>	<b>Preparative</b>	<b>Analytical</b>
Plasmid DNA	1 $\mu$ g	100ng
H <sub>2</sub> O	To make up the volume to	To make up the volume to
10X buffer	5 $\mu$ l	1 $\mu$ l
BSA	If required	If required
Enzyme	5U (1 $\mu$ l)	1U (0.2 $\mu$ l)

**Table 4.3 Content of digestion reaction**

**Digestion reaction has been set in following steps**

1. All the above mentioned components of digestion reaction were added in a micro centrifuge tube.
2. Enzyme should be added in the end.
3. The tube was briefly vortexed followed by short spin.
4. Micro centrifuge tube containing digestion reaction was incubated at 37°C for 2-4 hours in a water bath (or, at specifically recommended temperature for a particular enzyme).
5. For molecular cloning 1 $\mu$ l alkaline phosphatase (FastAP) (NEB) was added in the reaction tube for vector preparation and incubated it for next 1 hour (Alkaline Phosphatase removes the 5'-phosphate groups of DNA from both the termini of the digested vector so as to avoid the self-ligation of the vector).
6. Digested DNA fragments were analysed on an agarose gel.

**4.1.6.2: Purification of restriction digested DNA or PCR product**

For cloning it is very important to purify digested DNA fragments (either vector or insert) to remove nucleotides, primers, enzymes, mineral oil, salts, agarose, ethidium bromide, and other impurities from DNA samples.

Nucleotide removal kit (Qiagen), Gel extraction kit (Sigma)

For all cloning procedures QIA quick Nucleotide Removal Kit was used to remove impurities from DNA sample. Columns of QIA quick Nucleotide Removal Kit have a silica membrane assembly which is capable to bind DNA in high-salt buffer and elution of this silica membrane bound DNA been done with prewarmed water.

The protocol for nucleotide removal is as follows:

5 volumes of PN Buffer was added to 1 volume of the reaction sample and mixed homogeneously. The mixture was then transferred in a QIA quick spin column and this spin column is placed in the 2ml collection tube. Quick spin column containing collection tube was centrifuged for 1 min at 6000 RPM, flow-through was discarded. 600µl of PE buffer was then added to the column and centrifuged for 1 min at 6000 RPM, flow through was discarded. The column was wiped from outside to remove any residual PE buffer. Then column was placed in a dry tube and spin at 13000 RPM for 2 minutes to remove any residual wash buffer. Next, the column was placed in a clean 1.5 ml microcentrifuge tube. 50µl pre-warmed (at 50°C) autoclaved water was added to the centre of the column, kept it for 2 minutes, and centrifuged for 2 minutes at 14000 RPM to elute pure DNA. The concentration of the pure DNA can be increased by freezing the DNA tube at -80°C for 20 minutes, once frozen, the DNA was concentrated by doing Speed-Vac at 4°C ( until the volume reduces)

#### 4.1.6.3: Purification of DNA fragments from agarose gel

For cloning, sometimes digested DNA fragments (either vector or insert) or PCR product, has to be run on agarose gel followed by isolation and purification of this DNA fragments from agarose gel. This purified DNA fragments from agarose gel purification are further used to set up ligation reaction. GenElute Gel Extraction Kit (Sigma) was used to purify DNA fragment from agarose

gels. Agarose gel containing DNA band of interest was placed in a gel doc machine under UV light to visualize DNA. DNA band was cut from the gel using sharp scalpel which is pre-sterilized using 70% alcohol (as much as excess agarose was removed to increase the yield). DNA band containing agarose gel was cut into smaller pieces and placed in an Eppendorf tube. 3 volume of the Gel Solubilization Solution was added to the Eppendorf tube. (300 ml of Gel Solubilization Solution added for every 100mg of agarose gel). The mixture was incubated at 60°C for 10-15 minutes with intermittent vortexing. In the meantime, 500 ml of the column preparation solution was added to the binding column and centrifuged for 1 minute. Flow through was discarded. 1 gel volume of 100% isopropanol was then added to the tube and mixed homogeneously. The solubilized gel solution mixture was then added to the binding column and centrifuged it for 1 minute at 6000rpm. The flow-through liquid was discarded. 600µl of PE buffer was added to the column and centrifuged for 1 min at 6000 RPM. Flow through was discarded, column was wiped from outside to remove any residual buffer PE. Column was then placed in a dry tube and spin at 13000 RPM for 2 minutes to remove any residual wash buffer. Next, QIAprep column was placed in a clean 1.5 ml microcentrifuge tube 50µl pre-warmed (at 50°C) autoclaved water was added to the center of the column, kept it for 2 minutes, and centrifuged for 2 minutes at 14000 RPM to elute pure DNA. The concentration of the pure DNA was increased by freezing the DNA tube at -80°C for 20 minutes, once frozen, the DNA was concentrated by doing Speed-Vac at 4°C (until the volume reduces)

#### 4.1.6.4: Ligation reaction

In ligation, the phosphodiester bond is established between a 5'- phosphate termini and a 3'-hydroxyl group of two different DNA fragments (i.e. vector and insert DNA) with the help of DNA ligase enzyme.

A typical ligation reaction is set as follows:

<b>Components</b>	<b>Molar ratio</b>	<b>Concentration(<math>\mu\text{g}/\mu\text{L}</math>)</b>	<b>Length (in base pair)</b>	<b>Max Volume</b>
Vector	1	$x\mu\text{g}/\mu\text{L}$	Size in bp	Available vol.
Insert	3	$y\mu\text{g}/\mu\text{L}$	Size in bp	Available vol.

**Table 4.4 Calculation for Typical Ligation Reaction**

Concentration of purified vector and insert DNA fragments was measured. The typical ratio of vector: insert used in ligation reaction was 1:3 which can vary depending upon the size of either vector or insert. The amount of vector and insert fragment required to achieve 1:3 molar ratio was calculated as per the above mentioned table or using in any insilico ligation calculator.

The components of the ligation reaction were added as mentioned in the following table:

<b>Components of the ligation</b>	<b>Total volume (10<math>\mu\text{L}</math>)</b>
Nuclease-free water	To make up the volume
10X T4DNAligase buffer	1 $\mu\text{L}$
Vector	As calculated from the above table
Insert	As calculated from the above table
T4 DNA Ligase	200U

**Table 4.5 Component of Ligation reaction**

A positive control (any plasmid DNA of same concentration to check if transformation worked) and a negative control (ligation mixture without insert fragment) were always kept. Ligation reactions were kept for incubation at 22°C for 2-4 hours or at 16°C for overnight. All the three-reaction mixtures were transformed in *E. coli* cells and screened for positive clones.

#### **4.1.6.5: Screening of recombinant bacterial clones**

Screening of recombinant clone has been performed to check the presence of the specific insert. If the test plate contains more number of colonies as compared to the negative control plate, then only we can proceed for clone screening. In the ideal cloning conditions, negative control plate should not be having any colonies.

Method:

Replica plating has been done for all the transformants on appropriate LB-Amp or LB-Kan plates. Each transformant colony was inoculated in 1.5ml antibiotic containing LB broth in Eppendorf tubes individually. These Eppendorf tubes were incubated at 37°C for 12-16 hours at 200 RPM. (Each clone was given a specific number for documentation). Following day, the plasmid DNA was isolated using TELT buffer protocol. Restriction digestion reaction was set for the clones along with vector control DNA in which one restriction enzyme site is present in the vector DNA and another enzyme site is present in the insert DNA to confirm the presence of an insert in the final clone. Digested DNA fragments were analysed in agarose gel electrophoresis to check the band pattern for positive clones.

#### **4.1.6: Mutagenesis methods**

##### **4.1.6.1 Quick change /Site directed mutagenesis (213)**

Quick change mutagenesis method is utilized to introduce point mutation in a gene of interest with the help of high fidelity polymerase enzyme Pfu Turbo. For quick change mutagenesis, primers are diluted in 1:10 proportion with ddH<sub>2</sub>O eg. 5µl of each forward and reverse primer (from 100µM stock) diluted into 45µl of ddH<sub>2</sub>O. Template DNA concentration should be within range of 40-60ng/µl.

The Quick change mutagenesis method reaction was set up as follows:

Components	Volume( $\mu$ l)
H2O	15.3
10X buffer for PfuTurbo	2
10mM dNTPs	0.4
Primer mix(1:10dilution)	0.4
Template DNA(40ng/ $\mu$ l)	1
PfuTurbo polymerases	0.4 (20U)

**Table 4.6 Contents of PCR reaction for site-directed mutagenesis**

Cycling conditions for quick change mutagenesis:

Cycle step	Temperature	Time	Cycles
Initial denaturation	95 °C	30 seconds	1
Denaturation	95 °C	30 seconds	
Annealing	55 °C	60 seconds	18
Extension	68 °C	2min/Kb	
Final extension	68 °C	8 minutes	1
Final hold	4 °C	forever	

**Table 4.7 Cycling conditions for PCR for site-directed mutagenesis**

#### 4.1.6.2. Random mutagenesis (EP-PCR) (214)

Error-prone PCR (EP-PCR) method was used for introducing random mutations into a defined segment of DNA. It is possible to mutagenize an entire gene or merely a segment

of a gene using EP-PCR reaction. The average number of mutations per DNA fragment can be controlled as a function of the number of EP-PCR cycles performed. In Error-prone PCR, the MnCl<sub>2</sub> was added immediately before initiating the thermal cycling reaction. When the thermal cycling reaction reached the first annealing step, then Taq DNA polymerase was added into reaction.

1. A typical EP-PCR reaction mixture includes the following additive

	<b>Components</b>	<b>Final concentration</b>
1	H <sub>2</sub> O	To make up volume
2	100 mM Tris-Cl, pH 8.3	10 mM
3	2 M KCl	50 mM
4	25 mM dCTP	1 mM
5	25 mM dTTP	1 mM
6	25 mM dGTP	0.2 mM
7	25 mM dAATP	0.2 mM
8	100 μM Forward primer	2 μM
9	100 μM Reverse primer	2 μM
10	200 pg/μl template DNA	20 pg/μl
11	25 mM MnCl <sub>2</sub>	0.5 mM
12	5 U/ μl Taq DNA polymerase	0.05 U/μl

**Table 4.8 Components of EP-PCR**

Cycling conditions for random mutagenesis are as follows:

Cycle step	Temperature	Time	Cycles
Initial denaturation	94 °C	3 minutes	1
Denaturation	94 °C	30 seconds	
Annealing	Lower T <sub>m</sub> +3	30 seconds	12 cycles
Extension	72 °C	60 seconds/Kb	
Final extension	72 °C	8 minutes	1
Final hold	4 °C	forever	

**Table 4.9 Components of EP-PCR**

The PCR product was checked on an agarose gel.

#### **4.1.6.3. Saturational mutagenesis (215)**

Saturational mutagenesis is fast simple and efficient method used for the substitution of predetermined protein sites against all twenty possible amino acids at once. For this Saturational mutagenesis involves use of degenerate synthetic oligonucleotides as primers. The reaction components and PCR conditions used for saturational mutagenesis are similar to Site-directed mutagenesis.

## **4.2. Creation of mutant library and its screening:**

### **4.2.1. General methods and materials**

All synthetic DNA oligonucleotides used for cloning and library construction were purchased from Sigma Aldrich (St. Louis, Missouri, United States). PCR and restriction digestion products were purified using the Qiagen nucleotide removal kits (Hilden, Germany) as described above. Restriction enzymes were purchased from New England Biolabs (Ipswich, Massachusetts, United States). Sequencing reactions were analyzed at ACTREC, DNA sequencing facility Unit. Antibiotics ampicillin and chloramphenicol purchased from Sigma Aldrich (St. Louis, Missouri, United States). LB media purchased from HiMEDIA

(Mumbai,India). Chemicals and reagents were purchased from Sigma Aldrich (St. Louis, Missouri, United States) and HIMEDIA (Mumbai,India)

#### **4.2.2. Cloning of mEos3.2 gene**

A forward primer having a NcoI restriction site and a reverse primer having a NotI restriction site were used to amplify the cDNA. This cDNA is then digested and subsequently ligated into a similarly digested iGFP-pQE81XN vector. The iGFP in the vector was popped out and replaced with mEos3.2 gene.

#### **4.2.3. Mutant library creation**

Mutant library was generated by employing PCR based methods. The three methods of mutagenesis were used: Random mutagenesis, Saturational mutagenesis and Site-directed mutagenesis.

Using error-prone PCR, randomly mutated mEos3.2 libraries were created in random mutagenesis. Saturational Mutagenesis was performed at 157 and 158 positions. For this we used was using overextension PCR (5). Overextension PCR involves use of degenerative primers. A mutant library of size 1200 mutants for 2 mutational positions was generated considering the degeneracy of codons. Site-directed mutagenesis was performed to target six residue of mEos3.2. Insilico studies suggested that these 6 residue of mEos3.2 could be crucial for its brightness. These six mutations were introduced in mEos3.2 gene with varied combinations.

This whole mutant library was transformed into Ultra-competent E. coli strain DH5 alpha, Plated on LB agar plates supplemented with ampicillin (50µg /ml) . Incubation of plates were done for 14 hr at 37 °C.

#### **4.2.4. Replica plating for screening of transformants**

The replica plating is performed to select the transformants on the antibiotic containing selective media plate. For replica plating grids and lines were made on a fresh antibiotic containing selective media plate and numbers were put on those grids. As much as a colony

from transformation plate (after colony appears) can be replica plated on a new plate to screen for positive transformants. Replica plate was incubated at the 37°C in bacterial incubator for 24 hours.

#### **4.2.5. Mutant library screening**

##### **4.2.5.1 Preparing the cells**

*E. coli* colonies expressing the mutant libraries of mEos3.2 were grown on 10-cm Petri dishes. In first step, every mutant colony from our Mutant library was inoculated into LB Ampicillin media and allowed to grow for overnight.

The glass bottom plate were coated with polylysine: 50  $\mu$ L polylysine was added onto the dish and kept for 15 mins. Then washed thoroughly with dH<sub>2</sub>O and dried. Just before imaging, the cells were adhered to the dish: 50  $\mu$ L from your culture onto the dish, wait for 10 mins, and rinsed gently several times with NSD by pipetting. 2-3 mL of fresh NSD was left in the dish.

Screening for brightness improved mutants was done using confocal on microscope Leica SP8 (Wetzlar, Germany) microscope which is equipped with a Photoactivation scanner emitting light of wavelength 405nm. The before photoconvesion images of cells (green channel) were acquired using 488nm argon-ion laser and after Photoconvesion spectra (red channel) with a helium-neon 543nm laser. For this the detector slit settings of 510–540 nm and 590–630nm were used, respectively. Mutants which showed the highest fluorescence intensities in green as well as red channel were chosen. These chosen mutants with improved fluorescence intensities were subjected for further studies.

#### **4.3. Mammalian expression vectors, Cells and transfection**

WT mEos3.2 and all three mEosBrite variant gene fragment were amplified with a 5' primer having restriction site *Nhe*I and a 3' primer with restriction site *Bgl* II. The PCR product were purified, digested with *Nhe*I and *Bgl*III and ligated into similarly digested mammalian vector pmGFP-actin. In this cloning, GFP gene was replaced by mEos3.2 and mEosBrite variants. The WT and all three mEosBrite- $\beta$ -actin fusion constructs were transfected in U2OS cell line.

This cell line was maintained on Dulbecco's modified Eagle's medium (Thermo Fisher Scientific Bengaluru, India) which is supplemented with 10% heat-inactivated fetal calf serum (Thermo Fisher Scientific Bengaluru, India). Cells were incubated at 37 °C and 5% CO<sub>2</sub>. Plasmid DNA was prepared by Qiagen Plasmid Miniprep Kit (Qiagen Hilden, Germany) in accordance with the provided protocol, used for transfection. Transfection was performed by using transfection reagent Lipofectamine 2000 (Thermo Fisher Scientific Bengaluru, India). Cells for transfection and imaging were cultured on 35-mm glass-bottom culture dishes. The transfection was performed as per manufacturer protocol. For transfection 3 µg of plasmid DNA, 3 µl of the transfection reagent were used. Cells were incubated for two and half an hour with the transfection reagent. After 24 hr media was changed. Imaging was performed after 48 hr of transfection.

#### **4.4: Live cell imaging and its parameters**

##### **4.4.1. Live cell imaging**

Leica sp8 microscope was used for Laser scanning confocal imaging of WT mEos3.2 and mEosBrite-β-actin expressing U2OS cells. This microscope has Photoactivation (SIM) scanner which can emit light of 405 nm wavelength. A 488-nm argon-ion laser (Melles Griot Rochester, New York) was used for imaging of before Photoconvesion green spectra and a helium-neon 543-nm laser was used for imaging of after Photoconvesion red spectra, using detector slit settings of 510–540 nm and 590–630 nm, respectively. Adobe Leica LAS software along with image J and Photoshop were used for image processing.

##### **Imaging parameters**

Live cell imaging using Leica SP8 imaging platform.

- 63x objective (NA=1.4 or higher)
- Frame size: 512 x 512 (width x height)
- Zoom: 4

- Pixel size: 60-70 nm
- XY scan direction: bidirectional, phase = 3.15,
- Pinhole: 1.2 AU
- Bit depth: 8-bit is almost always adequate
- Laser settings for green and red fluorescence channels:
  - 488: 3-10%, HyD, collection window = 495-550 nm, gain = 400-500
  - 561: 3-10%, HyD, collection window = 575-750 nm, gain = 400-500
  - 405: 10% laser power, for photoconversion from Green to Red

#### **4.5 Protein expression and purification protocol**

The plasmids expressing mEos3.2 and mEosBrite variants were separately transformed into Rosetta2DE3 strain. Transformed cells were grown in LB media having Kanamycin (50ug/mL) and Chloramphenicol (34ug/mL).

##### **Protocol for 250 ml culture**

The starter culture was prepared by inoculating an isolated colony in 2.5 ml LB media with 50µg/ml ampicillin. In next step, the starter culture was inoculated into LB media in 1:100 ratio (i.e. 2.5 ml starter culture in 250 ml media), added 50µg/ml ampicillin (125µl) and incubated the culture at 37°C for 2 and half hours. After approximately 2 and half hours, OD<sub>600</sub> was checked, if it reached to 0.6, then culture was induced by using 0.1 mM IPTG (25µl from 1M stock IPTG). After induction, the culture was incubated overnight at 18°C with 200RPM in the shaker incubator (minimum 16 hours incubation needed after induction). Flasks were taken out after overnight 18°C incubation. Harvested media was then subjected for centrifugation in bottles for 30 mins at 4°C, 5000 rpm. The culture harvested was either directly used for purification or be stored at -80°C for later use.

#### **4.5.1 Ni-IDA column preparation:**

Before using the column for purification, 20% ethanol was eluted from the column because these columns are stored in 20% ethanol. 4-5 washes of autoclaved distilled water were then given. Column beads were stripped using Stripping buffer (1 column volume i.e., ~ 10 ml). Next, one autoclaved distilled water wash was given. One column volume 100mM NiSO<sub>4</sub> solution was passed through the column. One more autoclaved distilled water wash was given and column was equilibrated in Binding buffer for one and half hour.

#### **4.5.2 Lysate preparation:**

Culture pellet was resuspended in 10 ml of lysis buffer + 10 µl (1X) of protease inhibitor cocktail, and transferred in Sorvall tube. Cells were sonicated for approximately 3-4 cycles of 1 min with 1 min interval between consecutive cycles. After sonication, the lysate was subjected for centrifuge for 55 minutes at 11,900 rpm, 4°C. Supernatant (Lysate) was then collected in another tube.

#### **4.5.3 Column reanartion:**

Ni- IDA Agarose beads were washed with 3 Column Volume (CV) of Stripping Buffer, which strips the metal ions from the agarose. Resin was washed with 3-5 CV volumes D/W. Nickel beads were generated by giving a Ni<sup>2+</sup> wash with 1 volume 50mM NiSO<sub>4</sub>.6H<sub>2</sub>O.

#### **4.5.4 Affinity purification using Ni-IDA column:**

Supernatant was loaded on the column and gently mixed by inverting the column for several times. Supernatant was then kept for binding for one and half hour. Post-load was collected after binding period of one and half hour. Next, the column was washed with 20mM sodium phosphate buffer with 100mM sodium chloride. Washes with buffers containing varying concentration of Imidazole (10mM to 500mM) was given and elutes were collected in tubes. (Each buffer gradient of 15 ml volume was passed through the column. Each elute was collected in a tube of 3 ml volume). Concentration of each elution fraction was measured

using Bradford method, which is helpful to understand the optimum Imidazole concentration required for elution buffer.

Based on the Bradford coloration, elutes were loaded on SDS-PAGE. Pure fractions of protein were kept for dialysis in pre chilled dialysis buffer (1X sodium phosphate buffer +10% glycerol + 1mM DTT). The pure samples having OD 0.5-1 were snap frozen for future use.

#### **4.5.6. Composition of buffers:**

##### **4.5.6.1. Dialysis buffer pH 8.0:**

20mM Sodium phosphate, 100mM Sodium chloride, 10% Glycerol

##### **4.5.6.2. Binding buffer pH 8.0:**

10mM Imidazole, 20mM Sodium phosphate, 100mM Sodium chloride, 8M Urea, 1% Glycerol

##### **4.5.6.3. Elution buffer pH 8.0:**

1M Imidazole, 20mM Sodium phosphate, 100mM Sodium chloride, 8M Urea, 1% Glycerol

##### **4.5.6.4. Lysis buffer:**

20mM Sodium phosphate, 100mM Sodium chloride, 8M Urea, 0.1% TritonX, pH 8.0, Protease inhibitor.

##### **4.5.6.5 Column storage:**

After the elution, column was washed with 5 CV volume of D/W, it was allowed to drain.

The resin was stored in 20% ethanol at 4°C.

#### **4.6. Biophysical studies of purified protein**

##### **4.6.1. Fast Pressure Liquid Chromatography (FPLC) analysis**

FPLC study was done using a HiLoad 16/60 Superdex 75 pg column on an AKTAbasic liquid chromatography system (GE Healthcare Chicago, Illinois, United States). In the FPLC study,

the protein buffer was replaced from sodium phosphate to Phosphate Buffer Saline ( pH 7.4).

The working concentration of proteins were 3 mg/ml

#### **4.6.2. Dynamic Light Scattering (DLS)**

Before the DLS experiment, all the samples were filtered with 0.45 µm filter. DLS was executed using 70µl of FPLC purified fractions for mEos3.2 and mEosBrite variants. Hydrodynamic radius of all the proteins was calculated using DynaProNanoStar, Wyatt Technology.

#### **4.6.3. SDS PAGE and Western Blotting (216)(217)**

##### **4.6.3.1 Protein Estimation**

Protein estimation was done using Bradford's reagent as per manufacturer's protocol. BSA (1mg/ml stock) was used as standard in Protein estimation experiment.

1ml (1:4 diluted) of Bradford reagent was added in 5µl of lysate and standard protein for determination of protein concentration. Samples were incubated for 10 minutes at room temperature. The optical density of the samples was measured at 595nm along with blank and standard. Curve were plotted using standards value.

##### **4.6.3.2: SDS-PAGE**

SDS-PAGE techniques separate proteins on the basis of their size and charge. The components of SDS-PAGE are:

30% Acrylamide solution preparation : 29g Acrylamide and 1g Bis-acrylamide(USB) were dissolved in 80 ml distilled water and kept on a magnetic stirrer overnight (O/N) at room temperature; The volume was of solution was made up to 100 ml and filtered by passing it through 0.45 µm filter. Filtered Acrylamide solution was stored in a dark bottle at 4°C.

Preparation of 6X sample loading buffer: 50mM Tris.Cl (pH 6.8), 10% glycerol, 2% SDS, 1% β- mercapto-ethanol (BME) 0.1 % bromophenol blue.

Preparation of Electrophoresis buffer: 25mM Tris base, 250 mM Glycine (pH 8.3) and 0.1% SDS

The composition of 15% resolving gel are as follows

Component	Volume for 10ml (15%)
H <sub>2</sub> O	2.4 ml
30% Acrylamide mix	5 ml
1.5M Tris pH 8.8	2.5 ml
20% SDS	50 µl
10% APS	100 µl
TEMED	10µl

**Table 4.10 Composition for 15% resolving gel of SDS PAGE**

The components for used making stacking gel are:

Component	Volume for 4ml (6%)
H <sub>2</sub> O	2.7ml
30% Acrylamide mix	0.8 ml
1M Tris pH 6.8	0.5 ml
20% SDS	20µl
10% Ammonium persulphate	40µl
TEMED	4µl

**Table 4.11 Composition of a stacking SDS-PAGE gel**

#### 4.6.3.3. Wet transfer of proteins on PVDF membrane

The wet transfer method is used to transfer proteins which are separated on SDS-PAGE gel onto PVDF membrane for further analysis by immune blotting method.

High Glycine transfer buffer: 0.1M Tris, 0.19M Glycine, 20% methanol, 0.04% SDS.

The resolving gel was removed from electrophoresis assembly and rinsed gently in water so as to remove excess of SDS. Then it was immersed in transfer buffer for 10 min. PVDF membrane was activated by soaking it in methanol for 1 minute and then it was immersed in transfer buffer. The gel and membrane were then placed in between pieces of filter paper and

fiber sheets in the transfer cassette and immersed in the transfer apparatus, with the gel towards the negative electrode. Electroblotting at 300 mA was continued for 3 hrs at 4°C.

#### **4.6.3.4: Western Blotting**

Western blotting is an analytical technique which is used to detect the presence of either native or denatured proteins present on PVDF membrane using protein-specific antibody.

**Preparation of Tris-buffered saline (TBS):** 150/500 mM NaCl, 20 mM Tris (pH 7.4);

Tris-buffered saline with Tween20 (TBS-T): 1X TBS + 0.1 % Tween20;

**Blocking agent:** 5% or 3% BSA in 1X TBS;

The membrane was first blocked with either 0.3% BSA in TBST (for  $\beta$  Tubulin) or in 0.5% BSA in TBST (for His-Tag antibody) at room temperature for 2-3 hour. Next the membrane was incubated with primary antibody (diluted in 1% BSA, TBST) for 1 hour [for His-tag (1:5000 dilutions)] at room temperature or for overnight in the slow rocker. The membrane was Washed for 3 times in TBST. A time duration of 10 mins was given for each wash in the high speed rocker. The membrane was next incubated with secondary antibody anti-mouse HRP of 1:5000 dilution in 0.5% BSA for 1 hr at room temperature. After the secondary antibody treatment, 3 washes of TBST were given for 10 mins each in the high-speed rocker. The signal was then detected by enhanced chemiluminescence (ECL+), by incubating the blot with a detection reagent for 5 minutes which is followed by exposure to X-ray film and blot development.

#### **4.6.4 pH-dependent Performance of purified proteins (196)**

For determination of pH dependence behaviour of mEos3.2 and all three mEosBrite variants, all these purified protein samples (1mg/ml) present in PBS were diluted in ratio of 1:50 with a series of pH-adjusted citrate saline (pH <8) and sodium phosphate (pH >8) buffers in a 96-well black clear-bottom plate (Corning Safire2 plate reader was used for fluorescence measurement .

#### 4.6.5 Maturation kinetics (196)

*E. coli* cells (BL21) which are expressing mEos3.2 and mEosBrite variants were cultured in LB broth containing Ampicillin (50mg/ml) and Chloramphenicol (34mg/ml) for overnight. The culture was then diluted to an OD600 of 0.6, sealed with a rubber septum, and incubated for next 1 hr to allow residual oxygen consumption. 0.1mM IPTG was then added via a syringe needle to induce expression of the FPs. The culture was then incubated in shaker incubator for next 4.5 h at 37 °C. Next the grown cultures were then transferred to an ice bath for 10 min and maintained at 4 °C. Cells were centrifuged at 8000 rpm for 2 min. Cell pellets were then lysed and lysates were incubated at room temperature for 12 min. The lysates were then centrifuged at 12,500 rpm for 5 min, and the supernatant was diluted five times into PBS (pH 7.4). Finally, the fluorescence maturation was monitored at 37 °C using a Safire2 plate reader (Tecan).

#### 4.6.6. Determination of the quantum yield (218)

##### 4.6.6.1. Sample Preparation

For all quantum yield experiments, potential inner filter effects were minimised by using solutions with low absorbance readings (0.0–0.04). A series of six standard solutions was prepared by diluting the standards in either ethanol (Rhodamine) or 0.1 M NaOH (Fluorescein). Similarly series of six fluorescent protein solutions was made by diluting the each purified protein sample in PBS.

##### 4.6.6.2. Measurement of Fluorescence

For each standard and protein sample dilution, integrated fluorescence intensity was measured using Safire2 plate reader (Tecan). Next, for each sample, integrated fluorescence intensity was plotted against absorbance which ideally should show a linear relationship.

##### 4.6.6.3. Quantum Yield Calculations

The quantum yield of the fluorescent protein ( $\Phi_{FP}$ ) was calculated using slope of the linear fit for the standards, mSTD, using following equation:

$$\Phi_{FP} = \Phi_{STD} \left( \frac{m_{FP}}{m_{STD}} \right) \left( \frac{\eta_{FP}^2}{\eta_{STD}^2} \right)$$

where  $\phi_{STD}$  is the quantum yield of the standard,

$m_{FP}$  is the slope of the linear fit for the integrated fluorescence intensity of the fluorescent protein as a function of absorbance

$n_{2FP}$  and  $n_{2STD}$  are the refractive indices of the fluorescent protein and the standard solutions,

Respectively.

#### **4.6.7. Bacterial cytotoxicity (219)**

LB + ampicillin plates were prepared, one without IPTG (repressing conditions) and one with 1 mM IPTG (derepressing conditions). Concentrations of pQE81XN-based mEos3.2 and mEosBrite protein expression plasmids were adjusted to 1 ng/ $\mu$ L using sterile H<sub>2</sub>O. These plasmids were transformed into competent DH5 $\alpha$  cells and transformation mixture was spread uniformly on these plates which are incubated at 37°C for 12–15 hrs. Colony sizes for repressing versus derepressing conditions were compared.

#### **4.6.8. Assay for cytotoxicity in HeLa cell by transient transfection (219)**

##### **4.6.8.1. Preparation of cells and Transfection:**

HeLa cells were cultured in a 100-mm dish and 50–70% confluence was achieved. Culture medium was removed by aspiration. Cells were washed with 10 mL sterile PBS. PBS was removed and the HeLa cells were dissociated by incubation with 2 mL of Trypsin-EDTA solution at 37°C for 4 min. resuspension of these cells were done in 5 mL of culture medium at 37°C. Cell suspension was added to a 15-mL conical tube and centrifuged it for 5 min at 2,000g. Supernatant was discarded and cell pellet was resuspended in 5 mL of fresh culture medium. 0.5–1 x 10<sup>5</sup> cells were added to each well of several 24-well plates. “No DNA” transfection was kept as control. Cells were allowed to grow for 16–24 h to attain ~30% confluence. mEos3.2 as well as mEosBrite variants expression plasmids were diluted to concentration of 400 ng/ $\mu$ L. Sterile H<sub>2</sub>O was used as “No DNA” control. Transfection

mixture was prepared by adding 2  $\mu$ L of plasmid DNA +2  $\mu$ L of DMEM -/- medium in one tube while in another tube we took 2  $\mu$ L of Lipofectamine 2000 +2  $\mu$ L of DMEM -/- medium. For each protein, the plasmid mixture was combined with the Lipofectamine mixture, mixed gently and incubated in RT for 20 minutes. Further this transfection mixture is added in designated culture wells. Cells were incubated for 24 h after which Lipofectamine-containing medium was removed and replaced with fresh culture medium.

#### **4.6.8.2. Sample preparation for flow cytometry**

HeLa cells expressing mEos3.2 and mEosBrite variants were analysed by flow cytometry at daily intervals for 5 days, beginning at 24 h post-transfection. To prepare cells for analysis on a given day, the culture medium was removed, cells were washed with 1 mL of PBS. The PBS was then replaced with 200  $\mu$ L of Trypsin- EDTA and the cells were incubated at 37°C for 4 min. Trypsin-EDTA was removed and the cells were resuspended in 2 mL of culture medium. Flow cytometry was performed with the suspended cells using suitable laser lines and emission filters. Fluorescent cells were gated based on the “No DNA” control sample. Flow cytometry data was analysed using with FlowJo software.

#### **4.6.9. Aggregation assay (219)**

E. coli colonies individually expressing mEos3.2 and mEosBrite fluorescent protein were inoculated into 175  $\mu$ l of LB + ampicillin media and grown for 10–12 h at 37°C in shaker incubator. Cultures were then centrifuged for 5 min at 3,000g. Supernatant were removed. Cell Pellet were resuspended in 100  $\mu$ L lysis buffer and further lysed by controlled sonication. These lysates were centrifuged for 5 min at 3,000g. Supernatants were transferred to a black 96-well plate. Remaining pellets was resuspended in 100 lysis buffer and transferred in another black 96-well plate. The fluorescence signals were quantified from both, the cell pellet and supernatant for each protein using a 96-well plate fluorescent plate reader.

Use of a fixed gain that is high enough to produce strong signals without overloading the detectors is very important. Aggregation value for each protein was determined by calculating the percentage of the signal found in the pellet and supernatant.

#### **4.6.10. Photoconversion and Half-life measurement Experiments**

##### **4.6.10.1 Sample preparation:**

WT mEos3.2 and mEosBrite variants (1 $\mu$ g/ $\mu$ l each) were immobilized in 0.5% agarose solution on a coverglass and the agarose drops were mounted on to the ludin chamber (life imaging Service Basel, Switzerland). This chamber was then mounted in an customized IX83 Olympus microscope. This microscope is equipped with 405nm, 488nm and 561 nm lasers (Gatca, France). A 488nm excitation laser and a GFP filter cube (Semrock Rochester, New York) were used to achieve the Green or the before photoconversion form of mEos3.2 and mEosBrite variants. While, the red or the after photoconversion forms of proteins was achieved by point illumination of 405nm and 561nm by wide -field illumination. For all the measurements, the 405nm laser and 561nm laser were maintained at 5 mW and 500mW at the output of the objective.

##### **4.6.10.2 Calculation of Photoconversion efficiency:**

The Equation used for the calculation of photoconversion efficiency is:

$$(F_i - F_0) / F_0$$

Where;  $F_i$  represents intensity fluorescence emission observed in  $i$ th frame.

$F_0$  is the mean fluorescence intensity calculated from 10 continuous frames before the Photoactivation.

#### **4.6.10.3 Measurement of Half-Life:**

The half-life value of mEos3.2 and mEosBrite variants protein molecules was measured as time taken for bleaching from its optimum fluorescence intensity to its half. The graph of baseline corrected fluorescence intensity was plotted for mEos3.2 and mEosBrite variants proteins against time and the half-life was calculated from

# Chapter 5

## Results



## 5.1 Introduction

Fluorescent proteins enable us to understand varied intracellular mechanisms of the living cell in a non-invasive manner. By labeling the protein of interest with fluorescent epitope tags, one can visualize, and track its dynamics inside the cells. GFP which was the first fluorescent protein discovered, is still one of the most commonly used fluorescent proteins for imaging of biological samples (82)(139)(220). By mutating the chromophore-forming and chromophore surrounding residues of GFP protein, a large color palette of fluorescent proteins has been generated (221). Color palette of fluorescent proteins was further expanded by the discovery of GFP-like proteins (222). Ds-Red is an example of such GFP-like proteins, which shows emission spectra of red color. Ds-Red FP showed very unique feature of partial photoconversion in which its emission spectra changes from green to red upon exposure of UV light. Further research showed that Photoconversion of Ds-Red FP is associated with its chromophore modifications (208).

Just like Ds-Red-FP, one separate class of fluorescent proteins i.e. Phototransformable fluorescent protein (PTFPs) undergoes chromophore modification upon exposure of 405nm wavelength light. PTFPs are further categorised into three subclasses namely photoactivable, photoswitchable, and photoconvertible fluorescent proteins (2). Out of all three classes, Photoconvertible fluorescent proteins (PCFPs) undergo a unique type of chromophore modification in which their emission spectra irreversibly changes from green to red when exposed to UV light(2). PCFPs have been utilized to study different biological aspects like molecular dynamics of the tagged protein, visualizing and tracking the movement of newly synthesized proteins to its final destinations, various cellular events, etc (44)(223). PCFPs Due to their higher photostability, better contrast, stochastic activation, PCFPs have become popular tags for both, conventional (wide field and confocal microscopy) as well as super-resolution imaging modalities (structured illumination microscopy) and single-molecule

localization microscopy (156)(208)(224). Kaede, Eos, Dendra, mClavGR2 are few of the popularly known PCFPs which are all derived from class anthozoa (3)(208)(223).

The chromophore of PCFPs is conserved and composed of a tripeptide His-Tyr-Gly. This chromophore can be present either in neutral phenol or anionic phenolate form. The equilibrium between these two forms could change based upon its local microenvironment and pH. The anionic form gives green fluorescence as it absorbs the green excitation light. Neutral phenol form strongly absorbs UV light and undergoes unconventional cleavage between amide nitrogen and alpha carbon of His residue with effective beta-elimination. Beta-elimination is followed by the formation of double bond between C $\alpha$  and C $\beta$  of His residue which causes extension of the pi conjugation system of the imidazole ring in His, thus forming the modified chromophore, 2-[(1E)-2-(5-imidazolyl)ethenyl]-4-(p-hydroxybenzylidene)-5-imidazolinone (208)(225). The modified chromophore upon exposure to red excitation light gives red fluorescence. All PCFPs have this conserved mechanism of chromophore formation and its modification.

One of the commonly used PCFP is mEos3.2 which is generated by multiple steps of bioengineering on its most primitive form Eos. The photoconvertible fluorescent protein Eos was first discovered in stony coral *Lobophyllia hemprichii*(3). EosFP was tetrameric in nature and stably expressed only at 30°C. For application of EosFP in a biological system, it required several improvements. Monomeric Eos (mEosFP) was developed by introducing two point mutations V123T and T158H into EosFP. mEos2 FP which showed functional maturation at 37°C was developed by introducing 12 amino acid flexible linker region between the two protein subunits of mEos FP (4). Although mEos2 FP was monomeric in nature, it formed aggregates at higher concentrations, limits its role as a fusion protein. Zhang *et al.* developed rationally designed improved versions mEos3.1 and mEos3.2 which were monomeric, photostable, brighter, with higher photon budget, and higher label density (5).

biophysical properties like brightness, maturation rate, labeling density, photostability, oligomeric nature, pH stability, and on-off switching rate etc. of a PCFP, determine its performance in confocal and in SRM (6) (8) (196). For mEos3.2 protein, most of its biophysical properties are already improved excluding its relatively low brightness for red (after photoconversion) spectra. In this project decided to target this limitation of mEos3.2 protein using semi-rational protein engineering method.

Following semi-rational protein engineering approach we developed three improved variants of mEos3.2 protein which showed improvement in brightness for both green and red emission spectra. We named this improved variant as mEosBrite V1.1, mEosBrite V1.2 and mEosBrite V1.3. The improvement in brightness of mEosBrite variants was validated in mammalian cell line using confocal microscopy and superresolution microscopy. Detailed biophysical characterisation revealed that mEosBrite variants are brighter, truly monomeric, have lower cytotoxicity, lower protein aggregation tendency and higher photoconversion efficiency as that of WT mEos3.2. All these improved biophysical properties makes mEosBrite variants preferred candidates over the WT mEos3.2 protein for live-cell imaging studies

## 5.2 Results

### 5.2.1 Creation of a mutant library for mEos3.2 gene through varied mutagenesis methods

Different biophysical properties of PCFP like brightness, maturation rate, labeling density, photostability, oligomeric nature, pH stability and on-off switching rate, etc. determines its performance in diffraction-limited microscopy, as well as super-resolution microscopy(6)(8). Many of these biophysical properties of mEos3.2 protein has been already optimised, however its relatively lower brightness is one of the limitations for its utility in High-speed Super-resolution microscopy. To overcome this limitation mEos3.2 protein we decided to follow semirational protein engineering approach. In this approach primarily three different methods mutagenesis were used namely; random mutagenesis, Saturational mutagenesis and site-directed mutagenesis. A previous report on mEos3.2 protein indicated that its two amino acid residues at position 157 and 158 could be very crucial for its brightness (18). By means of Saturational mutagenesis using degenerative primers, we targeted these amino acid positions of mEos3.2. Saturational mutagenesis allowed the substitution of these amino acid positions against all 19 possible amino acids at once. In next method of mutagenesis to expand mEos3.2 mutant library, we first aligned mEos3.2 to another Photoconvertible fluorescent protein mClavGR2, which is a brighter variant generated by directed evolution of mCavGR1 (Fig.5.1). From the alignment study, we spotted six potential amino acid residues of mEos3.2 which might be crucial for its brightness. We thought that targeting these six amino acid residues of mEos3.2 protein could improve its brightness. These six amino acid residues were targeted in diverse combinations using site-directed mutagenesis. Next we also followed an unbiased way generating mutations anywhere in the mEos3.2 gene to further increase the mutant library. For that, random mutagenesis approach was followed using error-prone PCR. This complete library of mEos3.2 protein exceeding 1000 plus mutants was maintained in *E.Coli* cells.



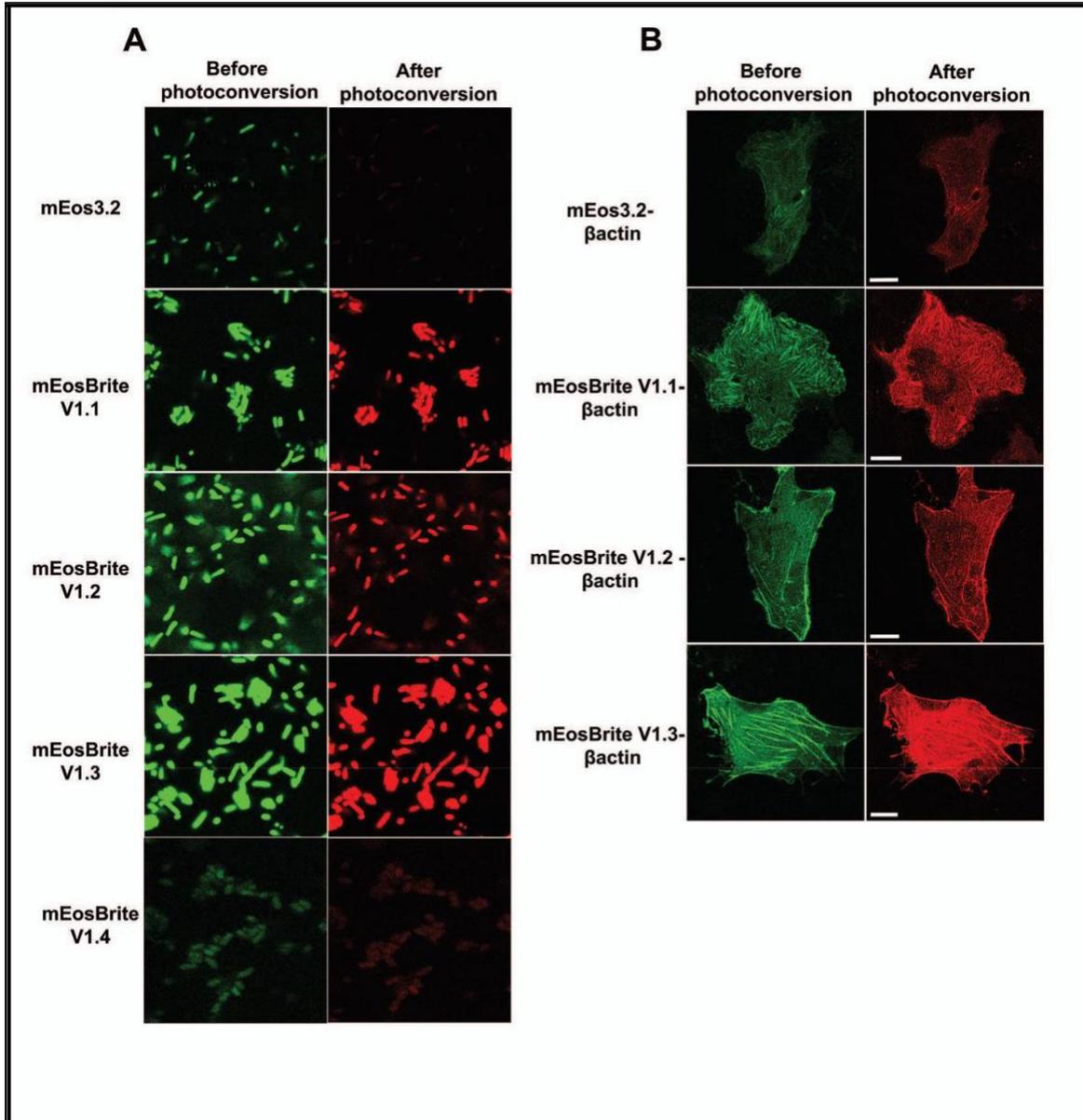
in LB-amp medium and incubated it for 16 hrs. Cells attached to glass bottom plate using polylysine and then subjected for microscopy screening. The improvement in their brightness for green (before photoconversion) and red emission spectra (after photoconversion) was monitored. After screening of whole mutant library, we found four variants of mEos3.2 protein which showed significant increase in its brightness of green and red emission spectra as compared to mEos3.2 protein. These brighter mEos3.2 protein variants were conjointly named as mEosBrite; mEosBrite V1.1, mEosBrite V1.2, mEosBrite V1.3, mEosBrite V1.4 (Fig. 5.2A). All these variants appeared brighter before and after photoconversion as compared to WT mEos3.2. Next, sequence analysis was done to understand the mutational replacement in these variants. In this sequence analysis we discovered the subsequent mutations: mEosBrite V1.1 (K37I, A69T, D73E), mEosBrite V1.2 (A69T, D73E, G111E), mEosBrite V1.3 (K37I, A69T, D73E, G111E) and mEosBrite V1.4 (I157F, E158I) (Table 5.1)

<b>Mutant number</b>	<b>Mutational replacement</b>
mEosBrite V1.1	(K37I, A69T, D73E)
mEosBrite V1.2	(A69T, D73E, G111E)
mEosBrite V1.13	(K37I, A69T, D73E, G111E)
mEosBrite V1.4	(I157F, E158I)

**Table 5.1: Mutational replacement in mEosBrite variants**

Improved brightness of mEosBrite variants was validated by their expression in the mammalian system (Fig. 4.2B). For that mEos3.2 and each of the four mEosBrite variants tagged  $\beta$ -actin constructs were created and transfected into U2OS cells. After 48 hrs. Live-cell confocal imaging was done. In confocal imaging we noticed that the cells expressing mEosBrite V1.1  $\beta$ -actin, mEosBrite V1.2  $\beta$ -actin and mEosBrite V1.3- $\beta$ -actin has

significantly brighter  $\beta$ -actin network than those of mEos3.2- $\beta$ -actin expressing cells for both green and red emission spectra (Fig. 5.2B). However, mEosBrite V1.4 does not showed any improvement in the brightness of the  $\beta$ -actin network as compare to mEos3.2 protein. This study confirmed that first three mEosBrite variants are much brighter than mEos3.2 protein and could be efficiently expressed and visualized in mammalian cells. So, for further study and biophysical characterization, we settled to go ahead with the first three mEosBrite variants (mEosBrite V1.1, mEosBrite V1.2, mEosBrite V1.3).



**Figure 5.2. Live cell confocal imaging of mEosBrite variants shows that mEosBrite variants are much brighter than the WT mEos3.2**

A) Confocal microscopy of *E.coli* cells expressing WT mEos3.2 as well as all the three mEosBrite variants from the mutant library. Bacterial cells from mutant library were independently grown in LB Ampicillin media overnight and then attached to glass bottom plate using polylysine. Live cell confocal microscopic screening was done using Leica SP8 microscope equipped with a Photoactivation scanner emitting at 405 nm. Left-hand panel of images shows green fluorescence images of *E.coli* cells captured before photoconversion. Right-

hand panel of image shows red fluorescence images of cells captured after subjecting them to Photoconvesion using 405nm laser.

B) U2OS cells expressing WT and mEosBrite variant- $\beta$ -actin fusion constructs were subjected for live cell imaging using Leica SP8 microscope. WT and mEosBrite- $\beta$ -actin fusion constructs were transfected in U2OS cells, the cells were cultured on 35-mm glass-bottom culture dishes. Imaging was performed after 48hrs of transfection for before as well as after photoconversion emission spectra at 63x magnification. Scale bar =10 $\mu$ m.

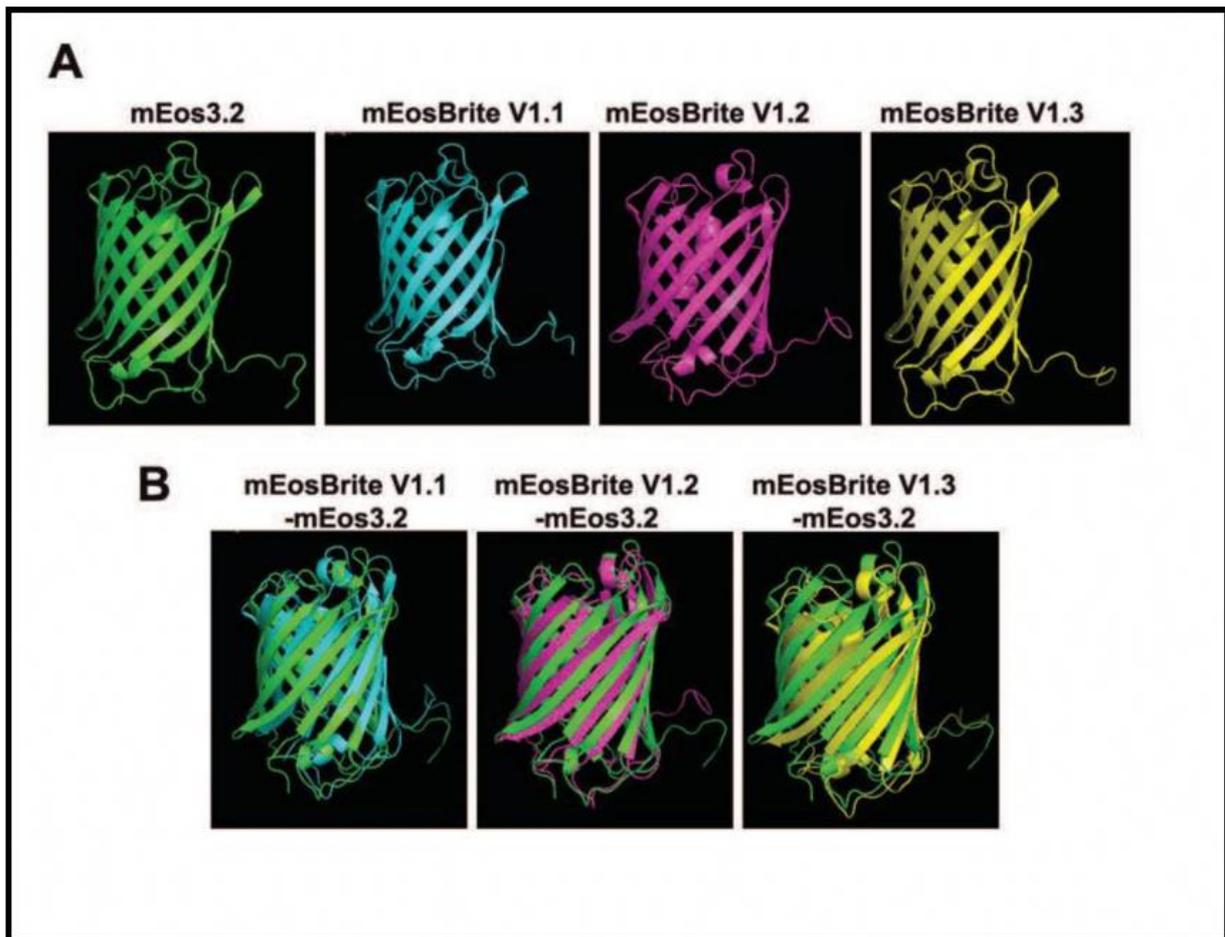
### **5.2.3 Insilco protein models generation and its Ramachandran analysis**

Before proceeding for Biophysical characterization, insilco protein models were generated for WT mEos3.2 and for all the mEosBrite variants by employing Phyre2 protein modelling server (Fig. 5.3A). Using Procheck server, Ramachandran plots were obtained for each protein mode (Fig. 5.4) l. Ramachandran plot's analysis (Table 5.2) suggests that, the percent of residues in the favoured region of mEos3.2 and mEosBrite V1.1, mEosBrite V1.2, mEosBrite V1.3 Insilco protein model is 94.2, 96, 95.5 and 95.5, respectively. The percent of residues in the allowed region of mEos3.2, mEosBrite V1.1, mEosBrite V1.2 and mEosBrite V1.3 Insilco protein model is 4.9, 4, 3.6, and 2.7, respectively. The percent of residues in the outlier region of mEos3.2, mEosBrite V1.1, mEosBrite V1.2 and mEosBrite V1.3 Insilco protein model is 0.9, 0, 0.9, and 1.8, respectively (Table 5.2). In conclusion, Ramachandran study advocated that there is no noteworthy changes in percent residues of favoured, allowed, and outlier region of Insilco protein model of mEosBrite variants compared to mEos3.2 protein model. This Insilco studies result indicated that structure of all mEosBrite variants protein would be stable despite possessing all shortlisted mutations.

<b>Name of variant</b>	<b>% of residue in favoured region</b>	<b>% of residue in allowed region</b>	<b>% of residue in outlier region</b>
mEos3.2	94.2	4.9	0.9
mEosBrite V1.1	96	4	0
mEosBrite V1.2	95.5	3.6	0.9
mEosBrite V1.3	95.5	2.7	1.8

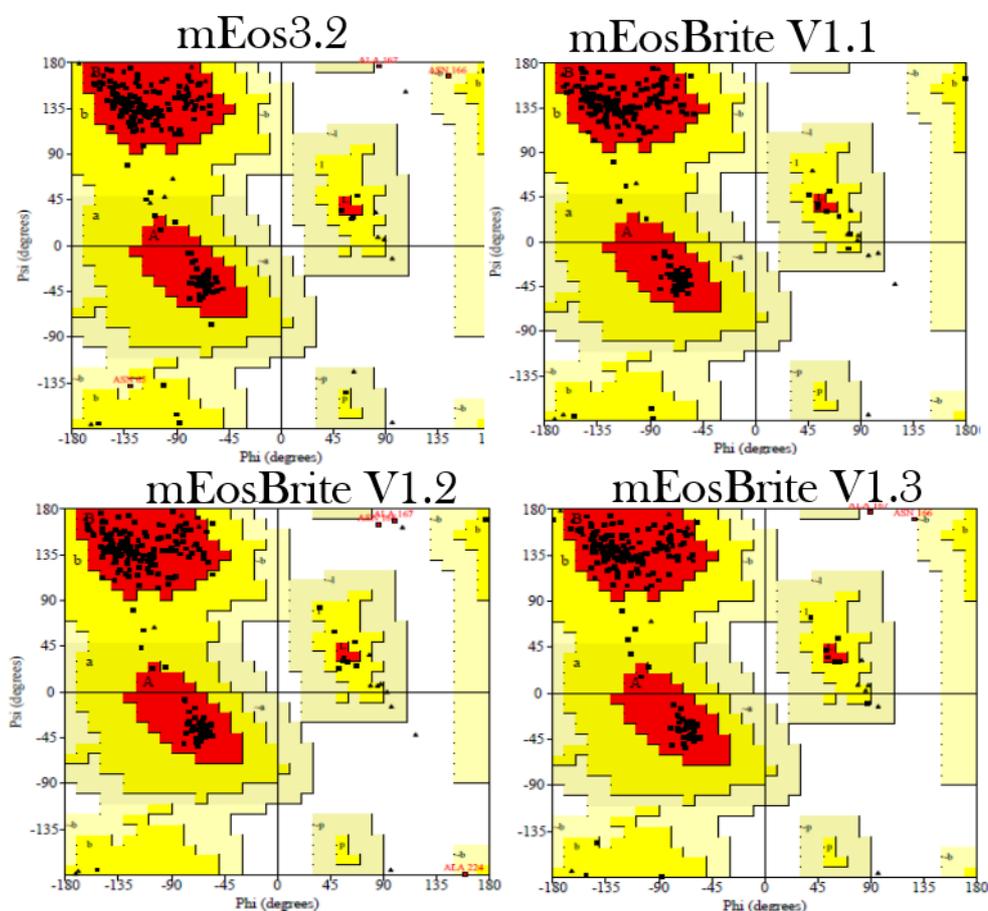
**Table 4.2 Ramachandran analysis for mEos3.2 and mEosBrite Insilco protein models**

The insilico protein model of WT mEos3.2 was then aligned along with each of mEosBrite variant model. In this alignment study, it was found that WT mEos3.2 protein model aligned perfectly with the all mEosBrite variants models (Fig.5.3B), thusly indicating that these mutations will not be changing the structure of mEosBrite variants significantly.



**Figure 5.3 In silico protein models of mEosBrite variants perfectly to WT mEos3.2**

A) Protein models of WT mEos3.2 protein and mEosBrite V1.1, mEosBrite V1.2, mEosBrite V1.3 were generated by employing Phyre2 protein modelling server. B) In-silico alignment of WT mEos3.2 protein model with each of mEosBrite variant model.

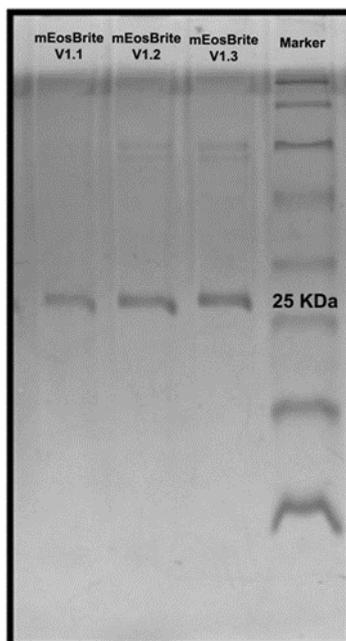


**Figure 5.4** Ramachandran plot for Insilco protein models

Ramachandran plot for Insilco protein models of WT mEos3.2 protein and mEosBrite V1.1, mEosBrite V1.2, mEosBrite V1.3 are generated by employing Procheck server.

### 5.2.4 Expression and purification of mEosBrite variants

To study the biophysical properties of mEosBrite variants, His-tagged WT mEos3.2 and mEosBrite variants were overexpressed in BL21 (DE3) cells and purified using nickel affinity chromatography. Optimum protein induction was achieved by using 0.1 mM IPTG, whilst for elution, 100mM imidazole was used in the elution buffer. Expression of proteins was confirmed by SDS-PAGE (Fig.5.5). We further put through these purified proteins sample for biophysical characterization.



**Figure 5.5 SDS-PAGE of mEosBrite variants**

Purification of His6-tagged mEosBrite variant proteins was performed by nickel affinity column method and analysed by SDS-PAGE (15%). All the purified protein variants gave a band just above 25KDa.

### **5.2.5 Excitation-emission spectra and maturation kinetics of mEosBrite variants is comparable to WT mEos3.2**

#### **5.2.5.1 Excitation-emission spectra:**

To find out the effect of incorporated mutations on excitation-emission spectra of mEosBrite variants, the excitation-emission spectra for WT mEos3.2 and mEosBrite variants was measured using fluorescence plate reader. For this study 1 mg/ml purified protein was used. It was found that the excitation maxima before photoconversion for WT mEos3.2, mEosBrite V1.1, mEosBrite V1.2 and mEosBrite V1.3 proteins were 491, 499, 497 and 498 nm, respectively while the emission maxima were 509, 510, 509 and 512 nm, respectively. The after photoconversion excitation and emission maxima of WT mEos3.2 and all mEosBrite variants was 560 nm and 575 nm, respectively (Fig.5.6A). From This study we came to know that the excitation and emission maxima of mEosBrite variants is comparable to that of WT

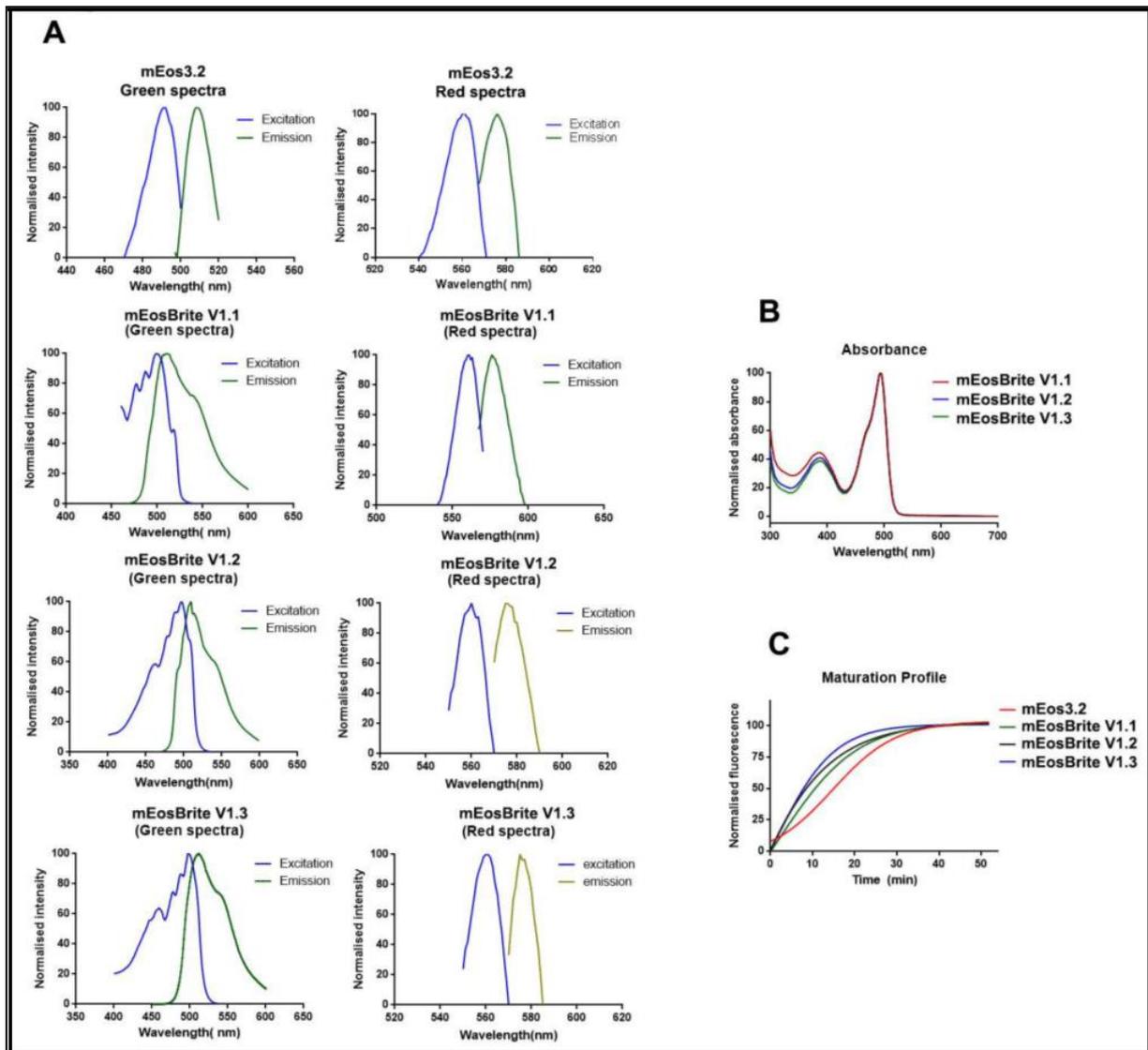
mEos3.2 protein for before photoconversion spectra, while it is exactly same for after photoconversion spectra.

#### **5.2.5.2 Absorbance spectra:**

1 mg/ml purified protein for each WT mEos3.2 and mEosBrite variants was used for this study using fluorescence plate reader. The absorption maxima of the mEosBrite variants exhibited the similar pattern of double peaks, i.e., one smaller peak at 385nm and one larger peak at around 493nm as that of WT mEos3.2. The smaller absorption peak is attributable to the neutral phenol form of the chromophore while larger absorption peak is attributable to its anionic phenolate form (Fig.5.6B).

#### **5.2.5.3 Fluorescence Maturation kinetics:**

The maturation profile of a fluorescent protein is the time taken by the protein to achieve fluorescent form once it is synthesized. Maturation profile of WT mEos3.2 and the mEosBrite variants was monitored at 37°C using a fluorescence plate reader. In the fluorescence maturation profile, we quantified the fluorescence against time. From our result we came to know that the maturation process for mEosBrite variants slightly faster as compared to WT mEos3.2. (Fig.5.6C)



**Figure 5.6 Spectral properties and maturation kinetics of mEosBrite variants are comparable**

A) Excitation and Emission fluorescence spectra of WT mEos3.2 and the mEosBrite variants before photoconversion spectra (green form) and after photoconversion spectra (red form). B) Absorbance spectra and C) Fluorescence maturation profile. All the three studies were done at 37°C by employing fluorescent plate reader.

### 5.2.6 mEosBrite variants shows improvement in quantum yield and inherent brightness

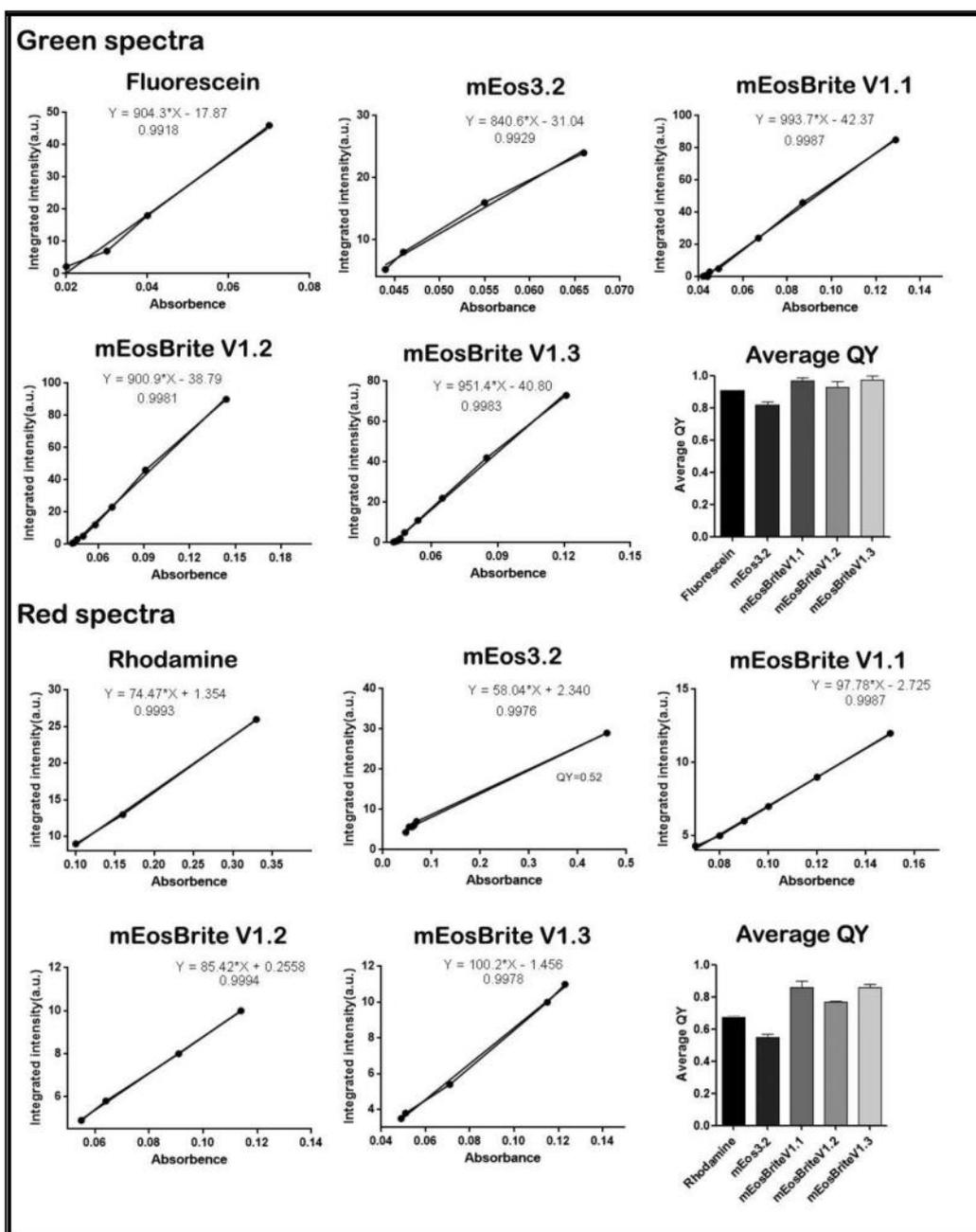
Quantum yield (QY) of a fluorescent protein is used for quantification of its fluorescence by using well-known standards. For green emission spectra, Fluorescein is used as a standard,

while for red emission spectra, Rhodamine is used as a standard. We have used the previously described method for QY measurements (218). The quantum yield value we got for green and red emission spectra of WT mEos3.2 protein is same as that of earlier reported value, i.e., 0.84 and 0.55 respectively (5) (Fig.5.7). When we studied the average QY of mEosBrite variants for green emission spectra we found that mEosBrite V1.1 and mEosBrite V1.3 showed maximum improvement i.e. 0.13 units while mEosBrite V1.2 showed increment in quantum yield by 0.09 units (Table 5.3). In case of red emission spectra, mEosBrite V1.1, mEosBrite V1.2 and mEosBrite V1.3 showed 0.31, 0.22 and 0.30 unit increments in average quantum yield, respectively (Table 5.3).

Next, the Inherent brightness of WT and the mEosBrite variants was calculated and compared. The Inherent brightness of Fluorescent Protein is a product of quantum yield (QY) and extinction coefficient (EC). Extinction coefficient of all three mEosBrite variants and WT mEos3.2 protein was calculated using online extinction coefficient calculator ExPASy. EC of WT mEos3.2 and the mEosBrite variants were remain unchanged i.e., 63400 for green state and 32200 for red state. Our results showed that the inherent brightness of mEosBrite V1.1, mEosBrite V1.2, and mEosBrite V1.3 for green spectra is 15%, 10% and 15% higher than WT mEos3.2 protein, respectively, while the inherent brightness of mEosBrite V1.1, mEosBrite V1.2, mEosBrite V1.3 proteins for red spectra is 55%, 39 % and 54% higher than WT protein, respectively (Table 5.3)

Name of protein	Green spectra		Red Spectra	
	Average QY	Inherent brightness	Average QY	Inherent brightness
mEos3.2	0.84	52.9	0.55	17.7
mEosBrite V1.1	0.97	61.1	0.86	27.6
mEosBrite V1.2	0.93	58.59	0.77	24.7
mEosBrite V1.3	0.97	61.1	0.85	27.3

**Table 5.3 average QY and inherent brightness of mEos3.2 and mEosBrite variants**



**Figure 5.7 mEosBrite variants have higher Quantum yield as compared to mEos3.2 protein**

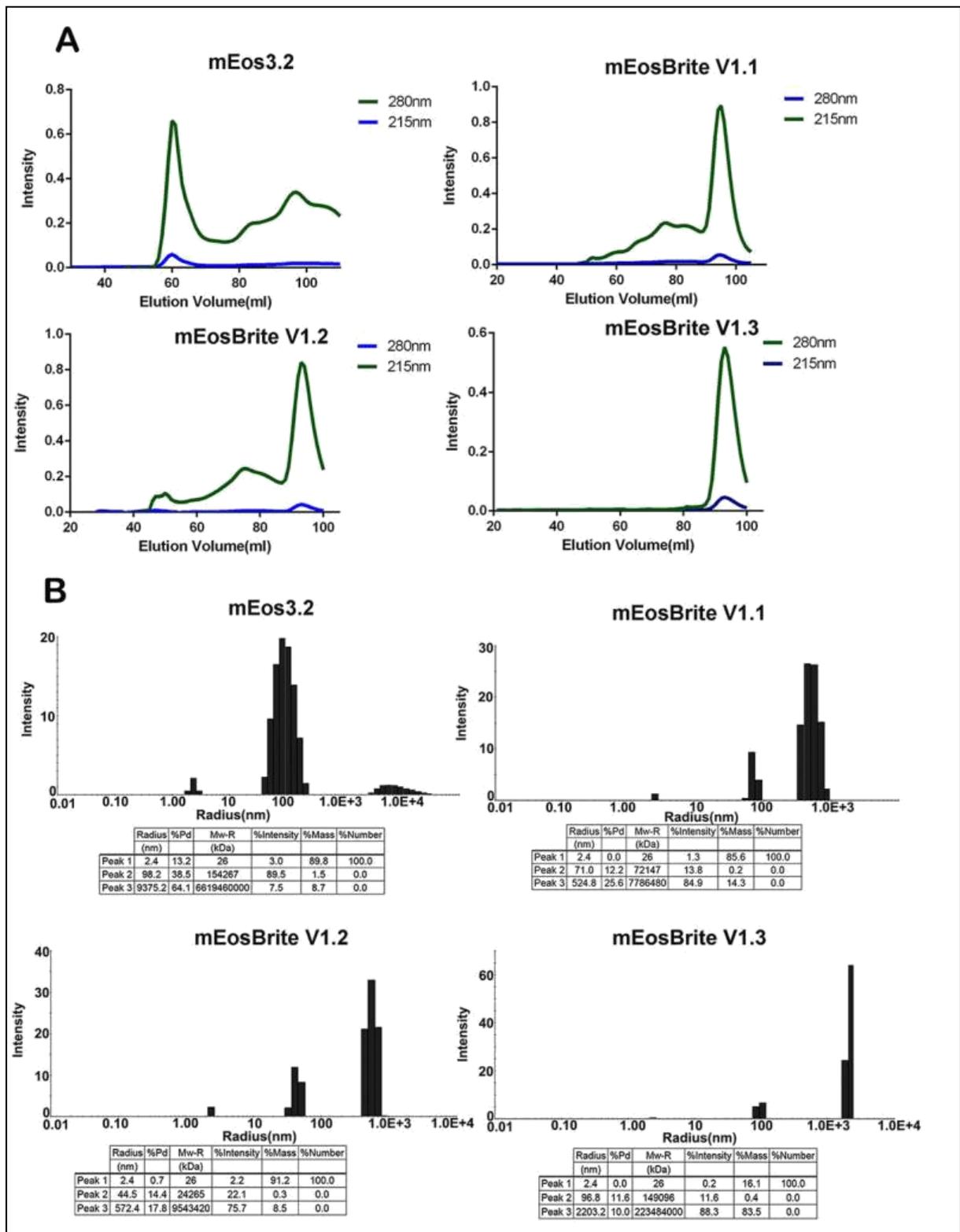
Integrated fluorescence intensity is plotted against absorbance for WT and mEosBrite variants. Fluorescein and Rhodamine were used as standards for quantum yield measurement. QY was calculated using formula  $Q_{STD}(m_{FP}/m_{STD})(\eta^2_{FP}/\eta^2_{STD})$  formula (218); where  $m_{FP}$  is slope of the linear fit for integrated fluorescence intensity of respective protein as a function of absorbance and  $m_{STD}$  slope of the linear fit for integrated fluorescence intensity

of standard as a function of absorbance,  $\eta_{FP}$  and  $\eta_{STD}$  are the refractive indices of the proteins and standard solutions.

### **5.2.7 mEosBrite variants are monomeric in nature**

We studied the oligomerization status of the mEosBrite variants and WT mEos3.2 protein using Fast Protein Liquid Chromatography (FPLC). Functional concentration of 3 mg/ml protein was used for this experiment. Standards were run on the FPLC column which showed that protein with molecular weight of 26KDa elutes at elution volume 93 ml, whilst protein with molecular weight of 52KDa elutes at elution volume 60 ml. When WT mEos3.2 protein was run on same column, it gave two peaks at elution volume 60 ml and 93 ml. Peak at elution volume 60 ml indicates the presence of some dimeric form species along with monomeric form of protein. When we ran mEosBrite variants on same column under exactly similar conditions, we got single sharp peak at elution volume 93 ml while there is no peak at elution volume 60 ml (Fig.5.8A). This result indicates that at higher concentration, the mEosBrite variants maintains its monomeric nature whilst the mEos3.2 protein get partially dimerize.

Dynamic Light Scattering (DLS) assay was also conducted for WT mEos3.2 and the mEosBrite variants using FPLC purified samples. The co-relation graph and fast decay time of WT and all the mEosBrite variants suggested that the mean radius fall within the range of proteins that is between 1 to 100  $\mu$ s. The estimated molecular weight of all the mEosBrite variants was around 26 KDa, which is similar to the molecular weight of monomeric mEos3.2 protein. The percent polydispersity of mEosBrite variants is much less than 20% suggesting a homogenous nature of protein solutions. However, for WT mEos3.2, the percent polydispersity was much higher than mEosBrite variants but still less than 20% indicating that some higher order oligomeric species could be present (Fig.5.8B).



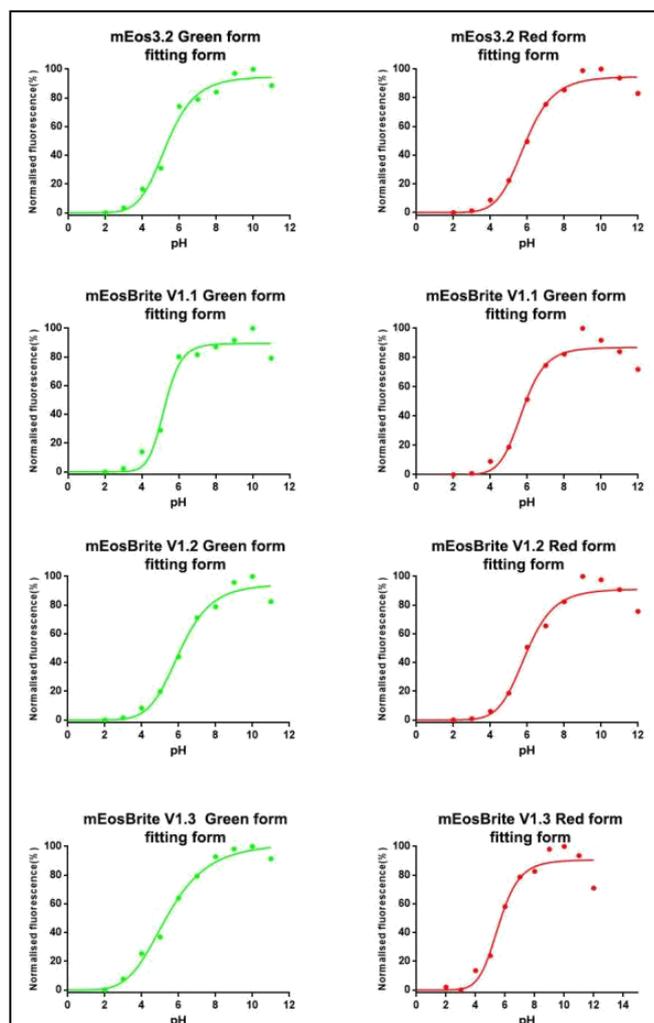
**Figure 5.8 mEosBrite variants are monomeric even at higher concentration**

A) FPLC elution profile of WT mEos3.2 and the mEosBrite variants. 3mg/ml each of purified WT mEos3.2 and mEosBrite variants were injected on HiLoad 16/60 Superdex 75 pg column to determine its oligomerization status.

B) DLS study for homogeneity of protein sample: Intensity (%) vs. radius (nm) plot of WT mEos3.2 and mEosBrite variants showing minimal polydispersity contributed by molecules having estimated molecular weight~26 kDa and monodisperse population with absolute % number.

### **5.2.8 The pH-dependent behaviour of WT mEos3.2 and mEosBrite variants is comparable**

The pH-dependent behaviour of WT mEos3.2 and the mEosBrite variants was studied in order to understand the effect of introduced mutations on their fluorescence intensity with respect to pH. For this study we individually diluted equal concentrations of purified proteins samples in the equal volume of pH buffers ranging from pH scale 2 to 12 and after that we measured their fluorescence intensity. Fluorescence intensity versus pH graph was plotted for WT and all the mEosBrite variant proteins for both i.e. green and red emission spectra (Fig.5.9). In this study we found that mEosBrite variants show optimum fluorescence intensity in pH range between pH 6 to 10 which is identical to that of WT mEos3.2 protein.

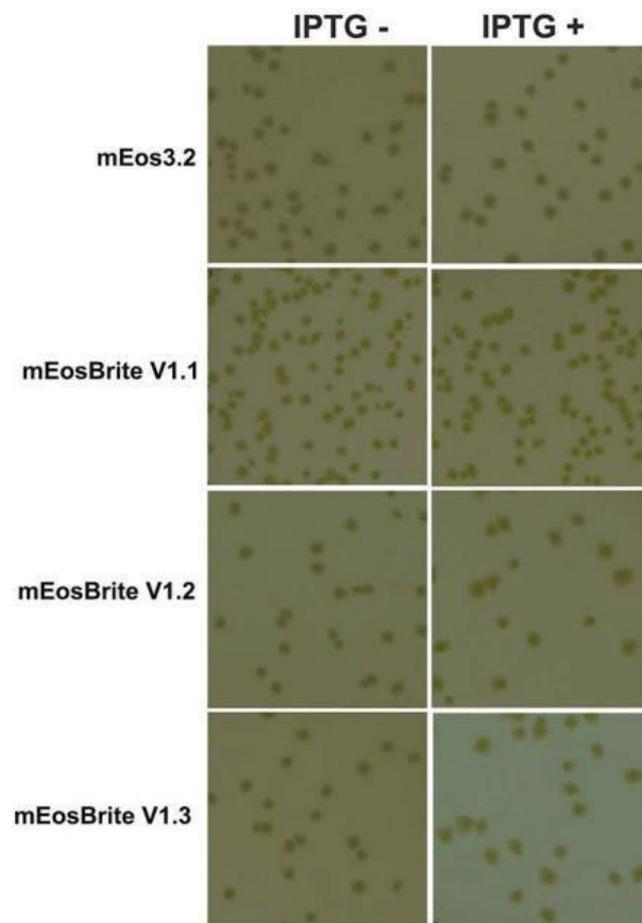


**Figure 5.9** pH dependent behaviour of mEosBrite variants is comparable to WT mEos3.2 Fluorescence intensity was plotted against pH (2 to 12) for green and red forms of WT as well as mEosBrite variants. Fluorescence intensities were measured using fluorescence plate reader.

### 5.2.9 MEosBrite variants show lower cytotoxicity in bacterial cells

Based on the biochemical nature of a fluorescent protein, it can be cytotoxic to cells (226)(219)(227). There is no known history of cytotoxicity for WT mEos3.2. We thereby decided to check whether any of the three mEosBrite variants show cytotoxicity in *E.coli* cell due to the incorporated mutations. During protein purification, we had earlier used 0.1mM IPTG concentration for the expression of all the proteins. At 0.1mM concentration of IPTG,

the cells expressing the protein had normal morphology and does not show any indication of cytotoxicity. To further understand toxic nature of mEosBrite variants in bacterial cells, we overexpressed the protein using higher concentration, i.e., 1mM of IPTG. Previous studies about fluorescent protein cytotoxicity suggested that if a protein is toxic to the cell, then its overexpression hinders the growth of bacterial cells and the colony size is significantly smaller. Our cytotoxicity results showed that even when the bacterial cells were induced with a higher concentration of IPTG, the colony size of the cells expressing the mEosBrite variants appeared almost similar to that of non-induced cells (Fig. 5.10), which indicates unusually low cytotoxicity of mEosBrite variants.

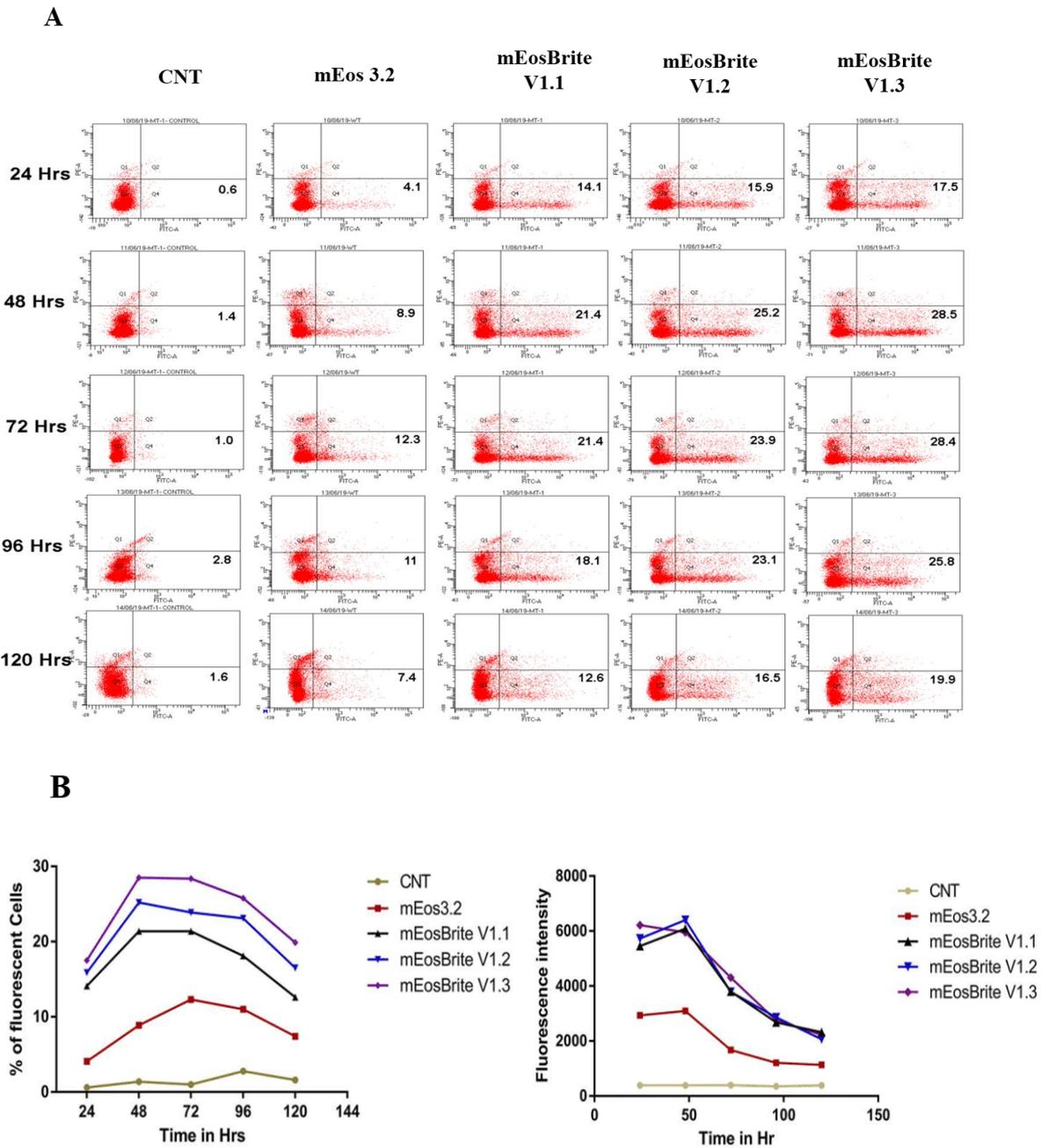


**Figure 5.10 mEosBrite variants are not toxic to bacterial cells even when overexpressed**  
Cytotoxicity of WT and the mEosBrite variants was compared by over-expressing these proteins in *E.coli* cells. 1mM IPTG was used to achieve overexpression of these proteins. The

colony size of *E.coli* cells before and after IPTG induction was compared to assay the toxicity of mEosBrite variants.

#### **5.2.10 mEosBrite variants show significantly low cytotoxicity as compared to mEos3.2 protein when expressed in mammalian cells**

The cytotoxicity of WT mEos3.2 protein and all the three mEosBrite variants was compared in mammalian cells. For this study we transiently transfected the HeLa cells with mammalian pCMV vectors independently expressing WT and the mEosBrite variants and monitored the number of fluorescent cells and their average brightness for 120 hrs using Flow Cytometry (fig.4.11A) If any of the mEosBrite variants is toxic to the cells, its overexpression is expected to cause a preferential loss in the number of cells (219)(228)(228). Loss of the cells will ultimately result into less number of fluorescent cells as well as lower average brightness. In our results it was observed for WT mEos3.2 and all 3 mEosBrite variants, highest number of fluorescent cells occurred between the time period of 48 to 72 hours while optimum brightness occurred at 48 hours (Fig.5.11B). Our result shows that the percentage of fluorescent cells expressing mEosBrite variants was almost two times higher as compared to WT mEos3.2 expressing HELA cells. The mEosBrite expressing HeLa cells also show much higher average brightness as compared to WT mEos3.2 expressing HeLa cells. These results suggests that all the three mEosBrite variants showed much lower cytotoxicity as compared to WT mEos3.2 when expressed in the mammalian cells.



**Figure 5.11 mEosBrite variants are less cytotoxic in mammalian cells are compared to WT mEos3.2**

- A) FACS profile of Hela cells expressing WT mEos3.2 protein and all the three mEosBrite variants over the period of 120 hrs.
- B) The percentage of fluorescent cells and average fluorescent intensity of WT mEos3.2 protein and mEosBrite variants was plotted against time.

### 5.2.11 mEosBrite variants have lower protein aggregation tendency as compared to WT mEos3.2

Fluorescent protein with a higher aggregation tendency is known to alter the normal localization of the tagged protein, affect the protein-protein interactions, and interfere with the protein sorting and intracellular trafficking. Thus, an aggregating fluorescent protein never gets priority in microscopy related cell biology experiment. To understand the aggregation nature of our proteins, the *E.coli* cells independently expressing WT and the mEosBrite variants were grown upto 0.6 OD, then induced with 0.1mM IPTG, lysed and centrifuged. The percent fluorescence for each protein in the pellet and in the supernatant was measured. A higher percent of fluorescence in the pellet indicates a higher aggregation and lower solubility tendency of fluorescent protein. WT mEos3.2 protein, showed more than 20% of the fluorescence in the pellet. While mEosBrite V1.1, mEosBrite V1.2, and mEosBrite V1.3 showed much lower fluorescence in pellet, i.e., 4.2%, 6.06% and 7.74%, respectively (Fig.5.12). Lower percent fluorescence in pellet of mEosBrite variants indicates that they have much lower protein aggregation tendency as compared to the WT mEos3.2.

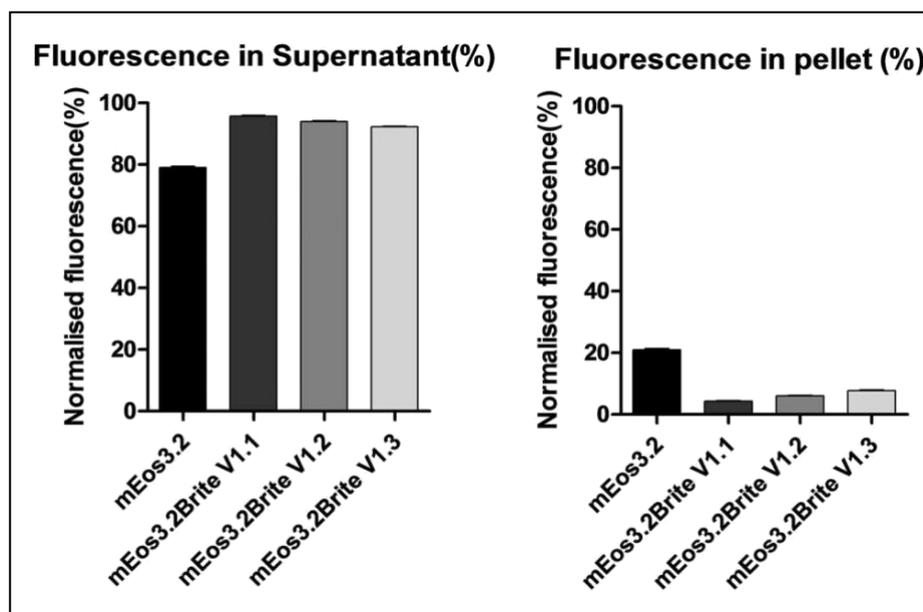
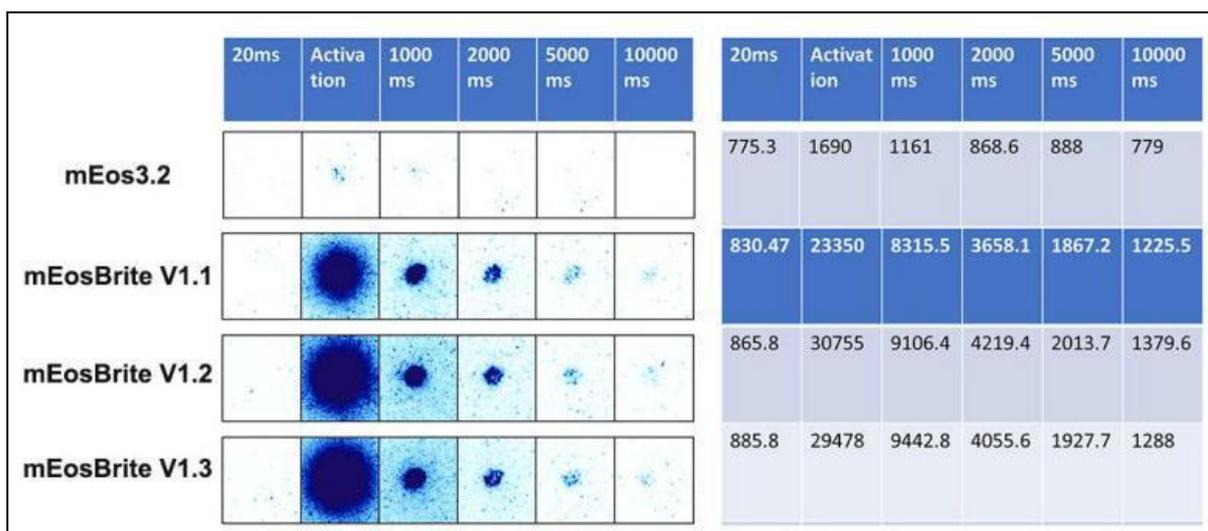


Figure 5.12 mEosBrite variants show lesser protein aggregation

WT and all 3 mEosBrite variants were expressed in BL21 cells. Cells were grown till 0.6 OD, then induced with 0.1mM IPTG and again allowed to grow for next 16 hrs. Then finally the cells were lysed and centrifuged. Fluorescence was measured for pellet and supernatant of each sample and a graph of Normalized fluorescence of pellet and supernatant was plotted.

#### **5.2.12 mEosBrite variants have higher photoconversion efficiency**

Higher photoconversion efficiency of Photoconvertible fluorescent proteins have great a significance in high-speed super-resolution imaging. For photoconversion efficiency study, we first entrapped 1mg/ml each of WT and the mEosBrite variants within agarose drops and subjected it for photoconversion. To achieve photoconversion of this entrapped mEos3.2 and mEosBrite variants protein molecule, the field of view was continuously illuminated with 561nm along with a single flash of 405 nm. The increase in intensity of (red fluorescence emission (after photoconversion) can be directly correlated to the photoconversion efficiency as all the variants are immobilized at the same concentration. These photoconverted molecules were acquired continuously over a period of 10 seconds at streaming mode with a frame rate of 50 Hz. In result it was found that mEosBrite V1.1, mEosBrite V1.2, mEosBrite V1.3 protein showed almost 23, 29 and 28 fold higher photoconversion efficiency, as compared to WT mEos3.2 (Fig.5.13). Higher photoconversion efficiency of mEosBrite variants makes them excellent tool for single molecule and intensity fluctuation based super resolution microscopy experiments.

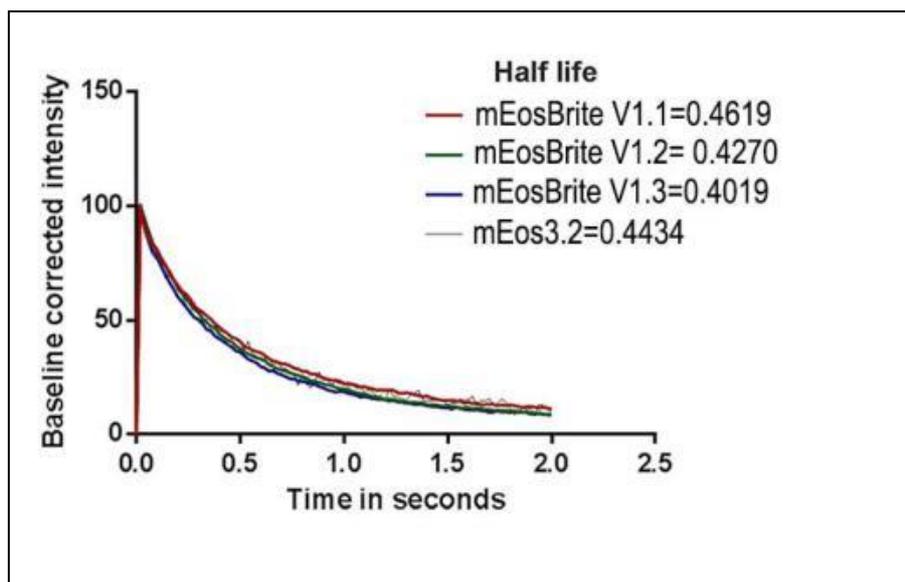


**Figure 5.13 mEosBrite variants show faster photoconversion than WT mEos3.2**

An assay system was developed to measure photoconversion efficiency, wherein 1 mg/ml each of WT and all mEosBrite variant proteins were separately immobilized within agarose drop on a coverglass and mounted on customized IX83 Olympus microscope. The photoconversion of these agarose entrapped molecules was done by employing 405nm laser. We tracked these photoconverted molecules for next 10 seconds. Using the equation  $(F_i - F_0)/F_0$ , efficiency of photoconversion was calculated; where  $F_i$  is the intensity of fluorescence emission observed in  $i$ th frame while  $F_0$  is the mean fluorescence intensity calculated from 10 continuous frames before the Photoactivation. Scaling for all the images is same.

### 5.2.13 The half-life of mEosBrite variants is comparable to WT mEos3.2

Half-life measurement was done in in-vitro condition by entrapping the protein molecules of WT and mEosBrite variants in the agarose drops and subjecting them for photoconversion. The time taken by the protein molecules to bleach from their optimum fluorescence intensity to half of this intensity was tracked. In result, we noticed the half-life of mEosBrite V1.1, mEosBrite V1.2, and mEosBrite V1.3 is 0.46s, 0.42s, and 0.40s, respectively, which is very much comparable to the half-life of WTmEos3.2 protein, i.e., 0.44s (Fig.5.14). From these results we came to know that mutations present in mEosBrite variants do not affect their half-life in in-vitro conditions.



**Figure 5.14 The half-life of mEosBrite variants is comparable to WT mEos3.2**

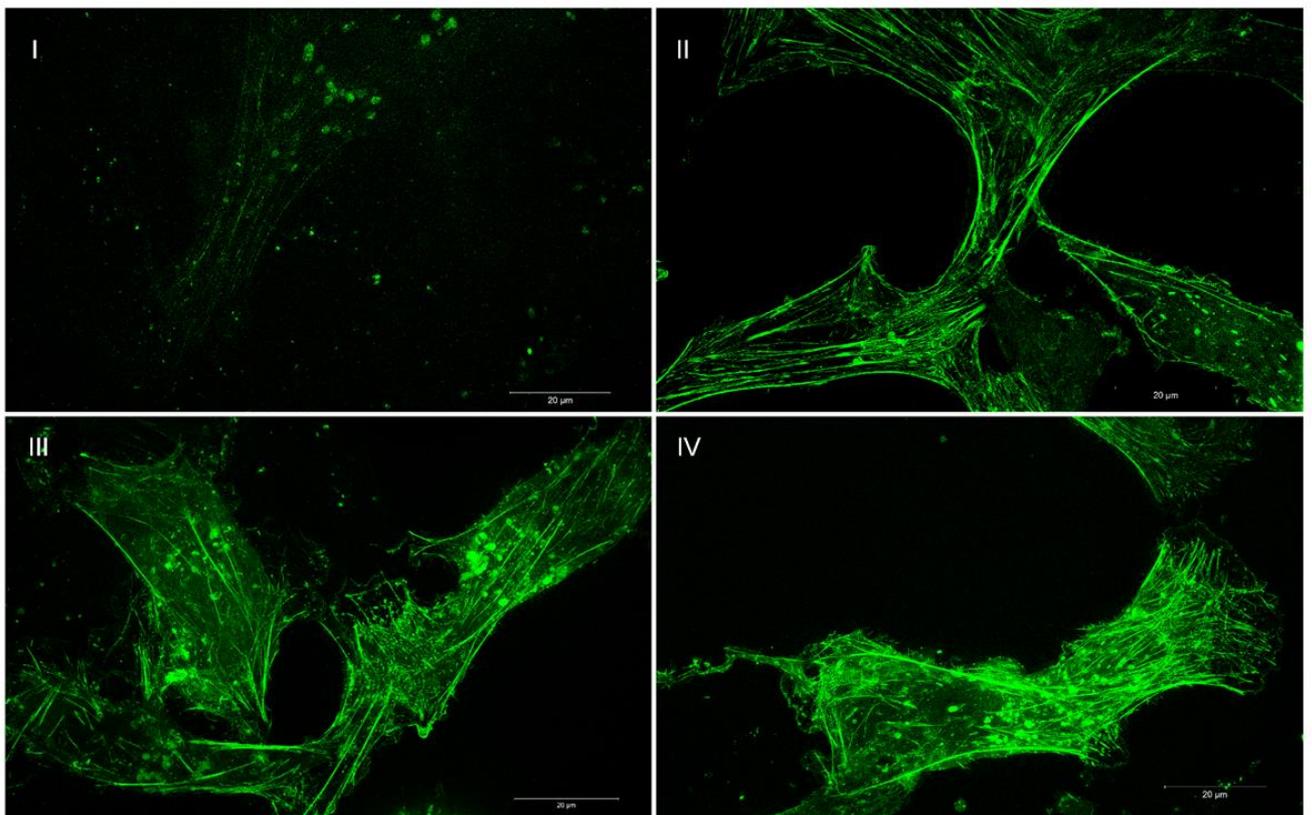
The agarose entrapped protein molecules of WT and all mEosBrite variant proteins were placed on coverglass and mounted on customized IX83 Olympus microscope. The entrapped protein molecules were then photoconverted. After photoconversion, the time taken by mEosBrite variants and WT mEos3.2 protein molecules to bleach from their optimum fluorescence intensity to half of this intensity was tracked. The baseline corrected fluorescence intensity graph was plotted for WT and all three mEosBrite variants with respect to time.

#### **5.2.14 mEosBrite variants performed better in Super resolution imaging**

Traditional microscopy can achieve resolution only up to certain limits i.e. 200nm because of the diffraction limit of light. Depending on the type of Superresolution microscopy, this barrier of diffraction limit can be crossed thus enabling us to achieve resolution less than 200nm. Superresolution microscopy crosses the diffraction limit either modulating the molecular state of fluorophore that reduces the PSF or by precisely locating the activated fluorophore molecule (stochastic activation)(175) . For this superresolution microscopy involves the use of Phototransformable fluorescent protein. After validating the performance of mEosBrite variants into mammalian system using confocal system we decided to validate

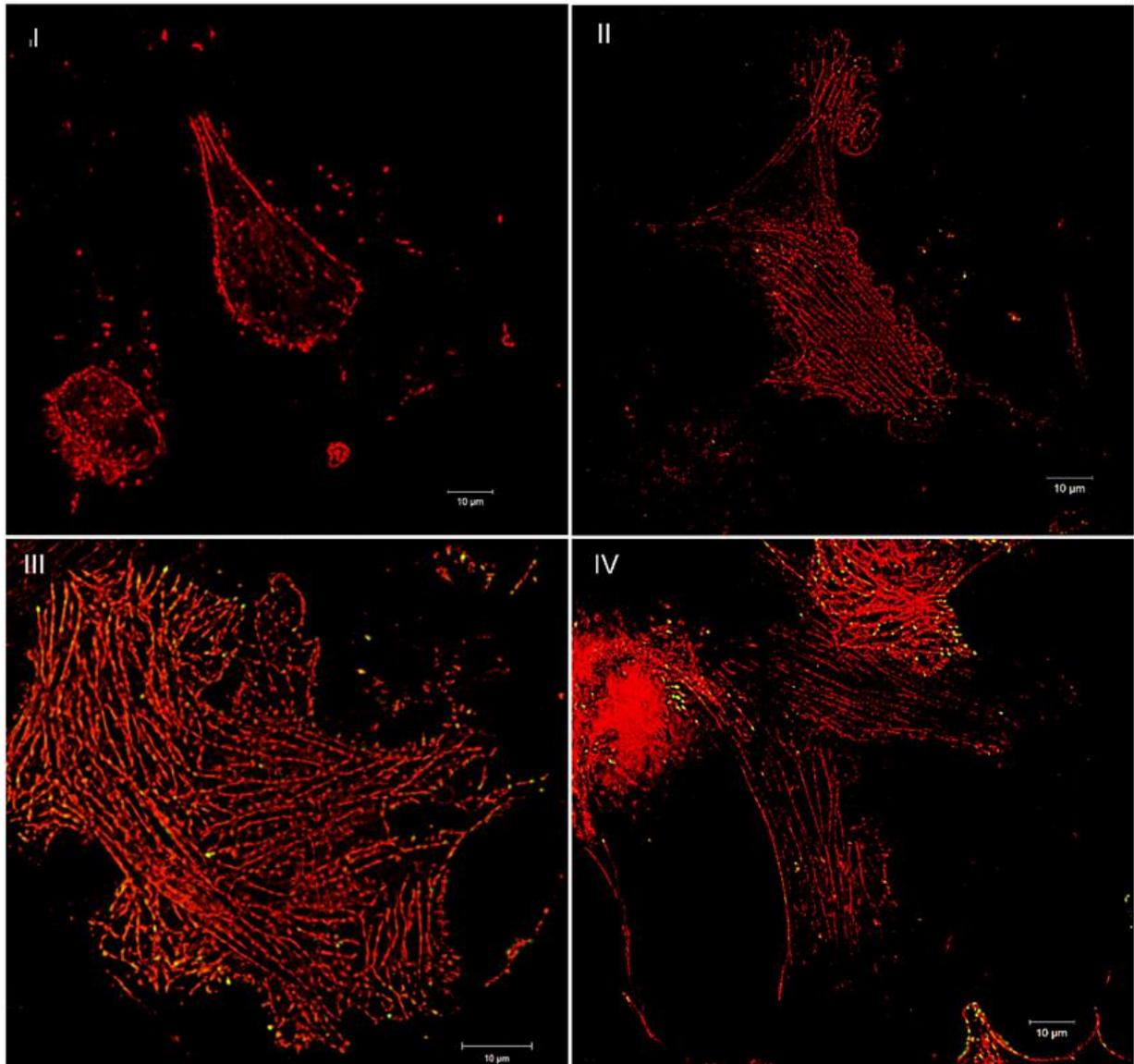
their performance using super resolution microscopy setup. For this we used structured illumination microscopy (SIM) and photoactivated localisation microscopy (PALM).

Structured illumination microscopy is capable of achieving much higher resolution (up to 120nm) compared to traditional confocal microscopy in both the lateral and axial dimensions while PALM in ideal conditions is capable of achieving resolution upto 20 nm. To check the performance of mEosBrite variants in SIM and PALM, the  $\beta$ -actin tagged constructs of WT mEos3.2 and all three mEosBrite variants were transfected into U2OS cells and subjected for live cell imaging. In result we found that the cells expressing mEosBrite V1.1, mEosBrite V1.2 and mEosBrite V1.3- $\beta$ -actin appear to have significantly brighter  $\beta$ -actin network as compared to the cells expressing WT mEos3.2- $\beta$ -actin. Also, mEosBrite variants tagged  $\beta$ -actin network appear to be better resolved as compared to mEos3.2 tagged  $\beta$ -actin network in both SIM (Fig.5.15) and PALM images (Fig.5.16). This study confirms that the mEosBrite variants are much brighter and better suitable for super resolution studies.



**Figure 5.15 mEosBrite variants perform better than WT mEos3.2 in SIM**

Structured illumination microscopy images of U2OS cells expressing (I) mEos 3.2- $\beta$ -actin (II) mEosBriteV1.1- $\beta$ -actin (III) mEosBriteV1.2- $\beta$ -actin (IV) mEosBriteV1.3- $\beta$ -actin



**Figure 5.16 mEosBrite variants perform better than WT mEos3.2 in PALM**

PALM images of U2OS cells expressing mEos3.2 protein and mEosBrite tagged  $\beta$ -actin (I) mEos3.2- $\beta$ -actin (II) mEosBriteV1.1- $\beta$ -actin (III) mEosBriteV1.2- $\beta$ -actin (IV) mEosBriteV1.3- $\beta$ -actin



# Discussion



mEos3.2 is a well-known photoconvertible fluorescent protein, whose nearly all the biophysical properties are reasonably improved excluding its relatively low brightness for red spectra. In this study, we targeted this limitation of mEos3.2 protein by employing method called semi-rational protein engineering. As a result, we found three improved variants, which are brighter than the WT mEos3.2 protein for both i.e. green and red emission spectra. These improved variants were named as mEosBrite variants. We validated the improved brightness of mEosBrite variants by their expression in U2OS cells which is followed by confocal imaging of U2OS cells. We also standardised the protein purification process for mEosBrite variants.

Next, we subjected them for invitro biophysical characterisation to comprehend the outcome of the incorporated mutations on its remaining fluorescent protein properties. First we performed quantum yield (QY) measurements for mEosBrite variants and WT mEos3.2 protein. Results showed that there is significant improvement in QY of green as well red emission spectra of mEosBrite variants. QY results of mEosBrite variants supported our previous results about improvement in brightness. mEosBrite variants with improved brightness could be good candidate for confocal and superresolution microscopy as its low level expression will be enough to perform imaging experiment. Further, Biophysical characterisation studies suggested that all the three mEosBrite variants are monomeric in nature, which suggests their utility in localisation studies. Single-molecule studies revealed that mEosBrite variants showed higher photoconversion efficiency in comparison to the WT mEos3.2 protein. Higher photoconversion efficiency of mEosBrite variants will be useful in super resolution imaging as it ensures faster imaging of sample and lesser exposure to harmful UV light. Out toxicity assay in mammalian cells showed that mEosBrite variants are lesser toxic in comparison to WT mEos3.2. Lesser toxicity of mEosBrite variants in mammalian cells suggest that mEosBrite variants could well tolerated in mammalian cells even when overexpressed. Protein Aggregation studies showed that mEosBrite variants form fewer protein aggregates as compared to WT mEos3.2 protein.

mEosBrite variants with low protein aggregation tendency could be excellent candidate for protein localisation studies. Our results of Structure Illumination Microscopy (SIM) and Photo Activated Localisation Microscopy studies revealed that our mEosBrite variants performed much better in superresolution microscopy in comparison to WT mEos3.2 protein. To conclude, we believe that mEosBrite variants with all these entire improved biophysical characteristic could be ideal candidates for future confocal and live-cell superresolution imaging studies.

# Chapter 6

## Summary and conclusion



## 6.1 Summary and conclusion

Photoconvertible fluorescent proteins are a type of Phototransformable fluorescent proteins which show change in emission spectra from one color to another when exposed to UV light and excitation light. Photoconvertible fluorescent proteins perform better in live cell superresolution microscopy techniques because of their properties like higher brightness, higher photostability, better contrast, stochastic activation etc. Dendra, mClavGR2, mEos3.2, kaede are few of the well-known Photoconvertible fluorescent proteins. In this project, we worked on mEos3.2 PCFP that shows green to red photoconversion on exposure to 405nm light.

mEos3.2 protein has many of its properties optimized like monomeric nature, photostability, maturation rate and maturation at 37°C. However, its brightness for after photoconversion spectra (red form) is less and has a scope of improvement. So in this project we decided to improve the brightness of mEos3.2 protein especially for red emission spectra after photoconversion, while retaining its remaining biophysical properties. To achieve this goal, we used semirational protein engineering approach.

Semirational protein engineering approach involves site directed and random mutagenesis methods. By using this approach, we created a mutant library of mEos3.2 protein that consisted of 1000+ mutants. Each of these mutants from mutant library was screened for improved brightness using confocal microscopy. Screening of the entire mutant library revealed 4 mutants which showed improvement in brightness for green and red emission spectra in comparison to the native mEos3.2 protein. We named these brightness improved variants as mEosBrite variants V1.1, mEosBrite variants V1.2, mEosBrite variants V1.3, mEosBrite variants V1.4.

As these improved mEosBrite variants are to be ultimately used in mammalian system, we validated their performance in mammalian cells. For that, we tagged  $\beta$  actin with WT mEos3.2

as well as all 4 mEosBrite variants and transfected the constructs into U2OS cell line. We then subjected these cells for confocal imaging. Confocal imaging data showed that cells expressing mEosBrite V1.1, mEosBrite V1.2 and mEosBrite V1.3- $\beta$ -actin appear to have significantly brighter  $\beta$ -actin network than that of the cells expressing WT mEos3.2- $\beta$ -actin. This improvement in brightness was found in both spectra i.e green emission spectra (before photoconversion spectra) and red emission spectra (after photoconversion spectra) of the protein. However, mEosBrite V1.4 failed to show such improvement and hence the first three mEosBrite variants were selected for further studies.

In vitro biophysical characterisation assays for mEosBrite variants were then conducted to comprehend the outcome of introduced mutation on remaining fluorescent properties of the protein. We measured the quantum yield value for WT mEos3.2 and mEosBrite variants. Quantum yield value of fluorescent protein tells about its brightness. Higher the quantum yield, brighter is the fluorescent protein. In our studies we found that mEosBrite variants have much higher quantum yield in comparison to WT mEos3.2 protein, which supported our in vivo confocal microscopy data. Thus, mEosBrite variants with improved QY can be good candidate for future confocal and superresolution microscopy experiments. We also studied the excitation emission spectra and maturation profile of mEosBrite variants and found that both of these properties of mEosBrite variants are very much comparable to that of mEos3.2 protein. We also subjected the mEosBrite variants for FPLC and DLS analysis to understand their oligomeric nature. Results from both these assays suggest that all the three mEosBrite variants retained their monomeric nature even when at higher concentration whilst the WT mEos3.2 protein partially dimerized. Monomeric nature of our mEosBrite variants makes them highly suitable for protein localisation studies. The assays for bacterial and mammalian cytotoxicity suggested that mEosBrite variants are much less cytotoxic in comparison to the WT mEos3.2. Comparatively lesser cytotoxicity of mEosBrite variants will give them upper

hand in the future imaging experiments. Higher photoconversion efficiency of mEosBrite variants ensures faster imaging with lesser UV rays exposure to cells during the photoconversion thus making them suitable candidates for high speed Super-resolution microscopy. Next we performed aggregation assay to understand the aggregation tendency of mEosBrite variants in which we found that mEosBrite variants display much less protein aggregation in comparison to WT mEos3.2 protein. Lesser aggregation tendency of mEosBrite variants could be useful for protein trafficking and protein localisation study. Lower aggregation might also be one of the potential reasons for the lower cytotoxicity of the mEosBrite variants.

We checked the photoconversion efficiency of mEosBrite variants and WT mEos3.2 protein. In our studies we found that mEosBrite variants have much higher photoconversion efficiency in comparison to the WT mEos3.2 protein.

We also checked the performance of mEosBrite variants using two different Super-resolution set ups i.e., SIM and PALM. In both the set ups, we found that the  $\beta$ -actin network tagged with the mEosBrite variants was not only brighter but also appeared to be better resolved in comparison to the WT mEos3.2 tagged  $\beta$ -actin network. We believe that the addition all the three mEosBrite variants to the toolbox of engineered FPs would be great asset, especially in live cell superresolution imaging experiments.

**Following are salient features of this research:**

- a. We have developed three improved variant of mEos3.2 protein through semirational protein engineering approach. We named them as mEosBrite V1.1, mEosBrite V1.2 and mEosBrite V1.3.
- b. mEosBrite variants showed significant improvement in their brightness for green spectra (before photoconversion) as well as red emission spectra (after photoconversion).
- c. FPLC and DLS studies suggest that mEosBrite variants are strictly monomeric even at higher concentration.

- e. Excitation-emission spectra, maturation profile of mEosBrite variants remained unchanged as that of mEos3.2 protein.
- f. Optimum pH range for mEosBrite variants is between 6-12 which is similar to that of mEos3.2 protein.
- g. mEosBrite variants are much lesser cytotoxic when expressed in mammalian cell in comparison to mEos3.2 protein.
- h. mEosBrite variants showed much higher photoconversion efficiency which could be useful for high speed superresolution imaging.
- i. mEosBrite variants have low tendency to form aggregates which could be beneficial in protein trafficking and protein localisation studies.
- j. Half-life of mEosBrite variants remain unchanged in comparison to mEos3.2 protein

## **6.2 Future Perspectives:**

In our studies we have shown that the introduced mutations in mEosBrite variants make them. Not only that, these mutations have also impacted other biophysical properties of protein. mEosBrite variants have higher quantum yield, are monomeric in nature, less cytotoxic, form lesser protein aggregates, show faster photoconversion as compared to mEos3.2 protein. Further, we would like to understand the details of the impact of introduced mutation on the structure of mEosBrite variants. For that we would setup extensive crystal trials for our mEosBrite variants.

We have shown that all the mEosBrite variants perform better than the WT mEos3.2 protein in confocal as well as superresolution microscopy (PALM and SIM) experiments. In future, we would like to test the performance of these mEosBrite variants on other superresolution microscopy setup like Stochastic Optical Reconstruction Microscopy (STORM).

We would also like to introduce similar sets of mutations in the remaining Photoconvertible fluorescent proteins Dendra and Kaede as our Sequence alignment study suggests that similar amino acid positions like mEos3.2 are present in these PCFPs that can be targeted. Next we

will check whether these mutation can give same improvement in the biophysical properties of these proteins or not?

Similarly, we would also like check the effect of same mutations of mEosBrite variants on the other types of Phototransformable fluorescent protein i.e Photoactivable and Photoswitchable fluorescent proteins if the amino acid positions are conserved among them.

## 6.1 Summary and conclusion

Photoconvertible fluorescent proteins are a type of Phototransformable fluorescent proteins which show change in emission spectra from one color to another when exposed to UV light and excitation light. Photoconvertible fluorescent proteins perform better in live cell superresolution microscopy techniques because of their properties like higher brightness, higher photostability, better contrast, stochastic activation etc. Dendra, mClavGR2, mEos3.2, kaede are few of the well-known Photoconvertible fluorescent proteins. In this project, we worked on mEos3.2 PCFP that shows green to red photoconversion on exposure to 405nm light.

mEos3.2 protein has many of its properties optimized like monomeric nature, photostability, maturation rate and maturation at 37°C. However, its brightness for after photoconversion spectra (red form) is less and has a scope of improvement. So in this project we decided to improve the brightness of mEos3.2 protein especially for red emission spectra after photoconversion, while retaining its remaining biophysical properties. To achieve this goal, we used semirational protein engineering approach.

Semirational protein engineering approach involves site directed and random mutagenesis methods. By using this approach, we created a mutant library of mEos3.2 protein that consisted of 1000+ mutants. Each of these mutants from mutant library was screened for improved brightness using confocal microscopy. Screening of the entire mutant library revealed 4 mutants which showed improvement in brightness for green and red emission spectra in comparison to the native mEos3.2 protein. We named these brightness improved variants as mEosBrite variants V1.1, mEosBrite variants V1.2, mEosBrite variants V1.3, mEosBrite variants V1.4.

As these improved mEosBrite variants are to be ultimately used in mammalian system, we validated their performance in mammalian cells. For that, we tagged  $\beta$  actin with WT mEos3.2

as well as all 4 mEosBrite variants and transfected the constructs into U2OS cell line. We then subjected these cells for confocal imaging. Confocal imaging data showed that cells expressing mEosBrite V1.1, mEosBrite V1.2 and mEosBrite V1.3- $\beta$ -actin appear to have significantly brighter  $\beta$ -actin network than that of the cells expressing WT mEos3.2- $\beta$ -actin. This improvement in brightness was found in both spectra i.e green emission spectra (before photoconversion spectra) and red emission spectra (after photoconversion spectra) of the protein. However, mEosBrite V1.4 failed to show such improvement and hence the first three mEosBrite variants were selected for further studies.

In vitro biophysical characterisation assays for mEosBrite variants were then conducted to comprehend the outcome of introduced mutation on remaining fluorescent properties of the protein. We measured the quantum yield value for WT mEos3.2 and mEosBrite variants. Quantum yield value of fluorescent protein tells about its brightness. Higher the quantum yield, brighter is the fluorescent protein. In our studies we found that mEosBrite variants have much higher quantum yield in comparison to WT mEos3.2 protein, which supported our in vivo confocal microscopy data. Thus, mEosBrite variants with improved QY can be good candidate for future confocal and superresolution microscopy experiments. We also studied the excitation emission spectra and maturation profile of mEosBrite variants and found that both of these properties of mEosBrite variants are very much comparable to that of mEos3.2 protein. We also subjected the mEosBrite variants for FPLC and DLS analysis to understand their oligomeric nature. Results from both these assays suggest that all the three mEosBrite variants retained their monomeric nature even when at higher concentration whilst the WT mEos3.2 protein partially dimerized. Monomeric nature of our mEosBrite variants makes them highly suitable for protein localisation studies. The assays for bacterial and mammalian cytotoxicity suggested that mEosBrite variants are much less cytotoxic in comparison to the WT mEos3.2. Comparatively lesser cytotoxicity of mEosBrite variants will give them upper

hand in the future imaging experiments. Higher photoconversion efficiency of mEosBrite variants ensures faster imaging with lesser UV rays exposure to cells during the photoconversion thus making them suitable candidates for high speed Super-resolution microscopy. Next we performed aggregation assay to understand the aggregation tendency of mEosBrite variants in which we found that mEosBrite variants display much less protein aggregation in comparison to WT mEos3.2 protein. Lesser aggregation tendency of mEosBrite variants could be useful for protein trafficking and protein localisation study. Lower aggregation might also be one of the potential reasons for the lower cytotoxicity of the mEosBrite variants.

We checked the photoconversion efficiency of mEosBrite variants and WT mEos3.2 protein. In our studies we found that mEosBrite variants have much higher photoconversion efficiency in comparison to the WT mEos3.2 protein.

We also checked the performance of mEosBrite variants using two different Super-resolution set ups i.e., SIM and PALM. In both the set ups, we found that the  $\beta$ -actin network tagged with the mEosBrite variants was not only brighter but also appeared to be better resolved in comparison to the WT mEos3.2 tagged  $\beta$ -actin network. We believe that the addition all the three mEosBrite variants to the toolbox of engineered FPs would be great asset, especially in live cell superresolution imaging experiments.

**Following are salient features of this research:**

- a. We have developed three improved variant of mEos3.2 protein through semirational protein engineering approach. We named them as mEosBrite V1.1, mEosBrite V1.2 and mEosBrite V1.3.
- b. mEosBrite variants showed significant improvement in their brightness for green spectra (before photoconversion) as well as red emission spectra (after photoconversion).
- c. FPLC and DLS studies suggest that mEosBrite variants are strictly monomeric even at higher concentration.

- e. Excitation-emission spectra, maturation profile of mEosBrite variants remained unchanged as that of mEos3.2 protein.
- f. Optimum pH range for mEosBrite variants is between 6-12 which is similar to that of mEos3.2 protein.
- g. mEosBrite variants are much lesser cytotoxic when expressed in mammalian cell in comparison to mEos3.2 protein.
- h. mEosBrite variants showed much higher photoconversion efficiency which could be useful for high speed superresolution imaging.
- i. mEosBrite variants have low tendency to form aggregates which could be beneficial in protein trafficking and protein localisation studies.
- j. Half-life of mEosBrite variants remain unchanged in comparison to mEos3.2 protein

## **6.2 Future Perspectives:**

In our studies we have shown that the introduced mutations in mEosBrite variants make them. Not only that, these mutations have also impacted other biophysical properties of protein. mEosBrite variants have higher quantum yield, are monomeric in nature, less cytotoxic, form lesser protein aggregates, show faster photoconversion as compared to mEos3.2 protein. Further, we would like to understand the details of the impact of introduced mutation on the structure of mEosBrite variants. For that we would setup extensive crystal trials for our mEosBrite variants.

We have shown that all the mEosBrite variants perform better than the WT mEos3.2 protein in confocal as well as superresolution microscopy (PALM and SIM) experiments. In future, we would like to test the performance of these mEosBrite variants on other superresolution microscopy setup like Stochastic Optical Reconstruction Microscopy (STORM).

We would also like to introduce similar sets of mutations in the remaining Photoconvertible fluorescent proteins Dendra and Kaede as our Sequence alignment study suggests that similar amino acid positions like mEos3.2 are present in these PCFPs that can be targeted. Next we

will check whether these mutation can give same improvement in the biophysical properties of these proteins or not?

Similarly, we would also like check the effect of same mutations of mEosBrite variants on the other types of Phototransformable fluorescent protein i.e Photoactivable and Photoswitchable fluorescent proteins if the amino acid positions are conserved among them.

## Thesis Abstract

**Name of the Student:** Mr. Pravin Vitthal Marathe

**Enrolment no:** LIFE09201304003

**Thesis Title:** "Creation of novel photochangable fluorescent protein through directed evolution"

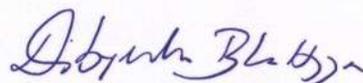
mEos3.2 is a well-known photoconvertible fluorescent protein with reasonably optimized biophysical properties except its relatively low brightness. In this study, we targeted this limitation of mEos3.2 protein by employing method called semi-rational protein engineering. As a result, we found three improved variants, which are brighter than the WT mEos3.2 protein for both i.e. green and red emission spectra. These improved variants were named as mEosBrite variants. We validated the improved brightness of mEosBrite variants by their expression in U2OS cells which is followed by confocal imaging of U2OS cells. We also standardised the protein purification process for mEosBrite variants.

Next, we subjected them for invitro biophysical characterisation to comprehend the outcome of the incorporated mutations on its remaining fluorescent protein properties. First we performed quantum yield (QY) measurements for mEosBrite variants and WT mEos3.2 protein. Results showed that there is significant improvement in QY of green as well red emission spectra of mEosBrite variants. QY results of mEosBrite variants supported our previous results about improvement in brightness. mEosBrite variants with improved brightness could be good candidate for confocal and superresolution microscopy as its low level expression will be enough to perform imaging experiment.

Further, Biophysical characterisation studies suggested that all the three mEosBrite variants are monomeric in nature, which suggests their utility in localisation studies. Single-molecule studies revealed that mEosBrite variants showed higher photoconversion efficiency in comparison to the WT mEos3.2 protein. Higher photoconversion efficiency of mEosBrite variants will be useful in super resolution imaging as it ensures faster imaging of sample and lesser exposure to harmful UV light. Out toxicity assay in mammalian cells showed that mEosBrite variants are lesser toxic in comparison to WT mEos3.2. Lesser toxicity of mEosBrite variants in mammalian cells suggest that mEosBrite variants could well tolerated in mammalian cells even when overexpressed. Protein Aggregation studies showed that mEosBrite variants form fewer protein aggregates as compared to WT mEos3.2 protein. mEosBrite variants with low protein aggregation tendency could be excellent candidate for protein localisation studies. Our results of Structure Illumination Microscopy (SIM) and Photo Activated Localisation Microscopy studies revealed that our mEosBrite variants performed much better in superresolution microscopy in comparison to WT mEos3.2 protein. To conclude, we believe that mEosBrite variants with all these entire improved biophysical characteristic could be ideal candidates for future confocal and live-cell superresolution imaging studies.



Mr. Pravin Marathe



Dr. Dibyendu Bhattacharyya (Guide)

Forwarded through:



Dr. Sorab N Dalal, 11/2/2024  
Chairperson, Academic &  
Training Programme, ACTREC

(Dr. Sorab N. Dalal)

Chairperson, Academic & Training Programme  
Tata Memorial Center, ACTREC,  
Nagar, Navi Mumbai - 410210



Dr. S.D. Banavali,  
Dean (Academics)  
T.M.C.

PROF. S. D. BANAVALI, MD  
DEAN (ACADEMICS)  
TATA MEMORIAL CENTRE  
MUMBAI - 400 012.

## Thesis Highlight

**Name of the Student:** Mr. Pravin Vitthal Marathe

**Name of the CI/OCC:** TMC-ACTREC

**Enrolment no:** LIFE09201304003

**Thesis Title:** “Creation of novel photochangable fluorescent protein through directed evolution”

**Discipline:** Life Sciences

**Sub-Area of Discipline:** Photoconvertible fluorescent protein mEos3.2 and its directed evolution

**Date of Viva voce:** 12-01-2021

mEos3.2 is a well-known photoconvertible fluorescent protein with reasonably optimized biophysical properties except its relatively low brightness. In this study, we targeted this limitation of mEos3.2 protein by employing method called directed evolution. In result, we found three improved variants, which are brighter than the WT mEos3.2 protein for its green and red emission spectra. These improved these improved variants were named as mEosBrite variants V1.1, V1.2 and V1.3. We validated the improved brightness of mEosBrite variants in mammalian cells by expressing them in U2OS cells followed by their confocal imaging.

Next, we subjected mEosBrite variants for invitro biophysical characterisation to comprehend the outcome of the incorporated mutations on its remaining fluorescent protein properties. Quantum yield studies showed that there is significant improvement in QY of green and red emission spectra of mEosBrite variants compared to mEos3.2 protein. FACS and DLS studies suggested that all the three mEosBrite variants are monomeric in nature, which suggests their utility in localisation studies. Single-molecule studies revealed that mEosBrite variants showed higher photoconversion efficiency in comparison to the WT mEos3.2 protein. Higher photoconversion efficiency of mEosBrite variants will be useful in super resolution imaging as it ensures faster imaging of sample with lesser exposure to harmful UV light.

Toxicity assay of mEosBrite variants in mammalian cells showed that mEosBrite variants are lesser toxic in comparison to WT mEos3.2. This assay suggested that mEosBrite variants could be well tolerated in mammalian cells even when overexpressed. Protein Aggregation studies showed that mEosBrite variants form fewer protein aggregates as compared to WT mEos3.2 protein which makes them excellent candidate for protein localisation studies. Superresolution SIM and PALM microscopy studies revealed that our mEosBrite variants performed much better in superresolution microscopy in comparison to WT mEos3.2 protein. In conclusion, we believe that the addition of all the three mEosBrite variants to the toolbox of engineered FPs would be great asset, especially in live cell superresolution imaging experiments.

