Published in final edited form as:

Arch Biochem Biophys. 2009 June 1; 486(1): 27–34. doi:10.1016/j.abb.2009.03.016.

Facilitating Chromophore Formation of Engineered Ca²⁺ Binding Green Fluorescent Proteins

Angela N. Holder^{a,†}, April L. Ellis^{a,†}, Jin Zou^{a,*}, Ning Chen^a, and Jenny J. Yang^{a,*}

^a Department of Chemistry and Center for Drug Design and Advanced Biotechnology, Georgia State University, 50 Decatur Street, Atlanta, GA 30302, USA

Abstract

Green fluorescent protein (GFP) containing a self-coded chromophore has been applied in protein trafficking and folding, gene expression, and as sensors in living cells. While the “cycle3” mutation denoted as C3 mutation (F99S/M153T/V163A) offer the ability to increase GFP fluorescence at 37 °C, it is not clear whether such mutations will also be able to assist the folding and formation of the chromophore upon the addition of metal ion binding sites. Here, we investigate in both bacterial and mammalian systems, the effect of C2 (M153T/V163A) and C3 (F99S/M153T/V163A) mutations on the folding of enhanced GFP (EGFP, includes F64L/S65T) and its variants engineered with two types of Ca²⁺ binding sites: (1) a designed discontinuous Ca²⁺ binding site and (2) a grafted continuous Ca²⁺-binding motif. We show that, for the constructed EGFP variants, the C2 mutation is sufficient to facilitate the production of fluorescence in both bacterial and mammalian cells. Further addition of the mutation F99S decreases the folding efficiency of these variants although a similar effect is not detectable for EGFP, likely due to the already greatly enhanced mutation F64L/S65T from the original GFP, which hastens the chromophore formation. The extinction coefficient and quantum yield of purified proteins of each construct were also examined to compare the effects of both C2 and C3 mutations on protein spectroscopic properties. Our quantitative analyses of the effect of C2 and C3 mutations on the folding and formation of GFP chromophore that undergoes different folding trajectories in bacterial versus mammalian cells provide insights into the development of fluorescent protein-based analytical sensors.

Keywords

Green fluorescent protein; fluorescent imaging; protein stability; protein folding; chromophore formation; cycle 3 mutation

Introduction

Green Fluorescent Protein (GFP) from *Aequorea victoria* is a 238-residue protein that consists of an 11 stranded β -barrel wrapped around a single central helix. The protein assumes this “ β -can” structure and completely buries its chromophore, which is formed from the tripeptide segment -S65-Y66-G67- [1,2]. Enhanced GFP (EGFP) contains the F64L/S65T mutation

*To whom correspondence should be addressed: Jin Zou, Fax: 404-413-5551, Tel: 404-413-5521, chejz@langate.gsu.edu, Jenny J. Yang, Fax: 404-413-5551, Tel: 404-413-5520, chejy@langate.gsu.edu.

[†]Equal contribution

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

discovered by Heim et. al.[1,3], which causes a greatly improved brightness and four times faster posttranslational oxidation of the chromophore, decreasing the time necessary for a bright chromophore formation. This is one of the most frequently used fluorescent proteins until now. It is utilized as the base protein for our studies due to its brightness and quicker chromophore formation to the fully fluorescent state. GFP spontaneously forms the chromophore without the use of any external co-factors and does not exhibit interference from fused proteins. This advantage has made it a target for use in a variety of applications in protein trafficking and folding, gene expression, and as sensors in living cells. Several different types of sensors have been developed by directly modifying the signal of fluorescent proteins. Extensive efforts have been devoted towards further development of fluorescent proteins as colorimetric sensors to measure metal ion concentration and stimulus induced change, oxidation potentials, protein-protein interactions, and enzyme activities.

Ca^{2+} functions as a ubiquitous intracellular messenger that regulates many different cellular processes including fertilization, proliferation, secretion, metabolism, contraction, differentiation, and apoptosis. Ca^{2+} signalling pathways are widely proliferated throughout the cell and participate in all basic cellular functions with Ca^{2+} binding proteins (CaBPs) facilitating Ca^{2+} buffering, Ca^{2+} transports, enzyme activation and inhibition, and channel regulation [4–11]. The ability to monitor Ca^{2+} concentration change and understand cellular Ca^{2+} signalling pathways is essential to develop various powerful tools for both biotechnological setting and disease therapy. To achieve this objective, engineered Ca^{2+} binding GFPs have been applied as one of strategies. Our lab has developed several GFP-based Ca^{2+} binding proteins by grafting a continuous Ca^{2+} binding motif or designing a discontinuous Ca^{2+} binding site in EGFP using the common features of Ca^{2+} binding sites in proteins based on our CaBP knowledge [11–17].

While the Ca^{2+} binding of CaM EF-hand-III grafted EGFP has been published, we have not yet addressed the development of sensors with additional mutations to increase the fluorescent yield [17]. One of major problems encountered while engineering ligand binding sites in a fluorescent protein is the loss of fluorescent properties upon mutation. Inclusion of negatively charged ligands to construct a Ca^{2+} binding site often affects the chromophore formation and its charge state [1,2,18]. Analysis of this affect is important now to demonstrate the feasibility in designing and engineering Ca^{2+} binding sites in fluorescent proteins without loss of the desirable fluorescent properties. Partly due to the native environment of jellyfish, temperatures greater than 30 °C cause a decrease in the folding efficiency of wild type GFP during its manufacturing [1]. The “cycle3” mutation, denoted as C3 mutation (F99S/M153T/V163A), is among the first set of mutations from DNA shuffling to show increased fluorescence at 37 °C [19,20]. While there are extensive applications of these thermally stable fluorescent proteins, the mechanism for the folding and the key factors that control the chromophore formation are yet to be clearly elucidated. It is not apparent whether such mutation will also be able to assist the folding and formation of the chromophore upon addition of a continuous metal ion binding motif or discontinuous binding site, especially the ligand types necessary to form Ca^{2+} binding sites such as Asp and Glu [11–15]. It is also very important to know the differential effects of such modifications on bacterial and mammalian cells with different folding machinery. Therefore, incorporation of “folding” mutations and a quantification of their fluorescence increase will provide us with insight into the future development of fluorescent protein-based Ca^{2+} sensors for real time imaging of Ca^{2+} signalling.

In the present study, we investigate both *in vitro* and *in vivo* folding of EGFP and its variants with two categories of Ca^{2+} binding sites. The effects of C2 (M153T/V163A) and C3 (F99S/M153T/V163A) mutations on the fluorescent intensity of proteins expressed at different temperatures and in both bacterial and mammalian cells were examined. We show that, for the constructed EGFP variants, the C2 mutation is sufficient to produce fluorescence in both

bacteria and mammalian cells. Further addition of the mutation F99S decreases the folding efficiency of these mutants although a similar effect is not detectable for EGFP. The extinction coefficient and quantum yield of purified proteins of each construct were also examined to compare the effects of both C2 and C3 mutations on protein spectroscopic properties. Results of our studies provide insights into the engineering of GFP for the development of fluorescent protein-based analytic sensors.

Materials and methods

Variant Constructs

The GFP variant EGFP-D2 with a discontinuous Ca^{2+} binding site (S2D, S86D, L194E) and C2 (M153T/V163A) and C3 (F99S/M153T/V163A) mutations were made through site-directed mutagenesis with PCR and turbo *pfu* (Stratagene) following the manufacturer's suggestions with EGFP (S65T, F64L, V22L, M218I, H231L) as the initial template. EGFP-G1 contains a continuous Ca^{2+} binding motif, CaMEF-hand-III, which was inserted by several rounds of PCR utilizing turbo *pfu*. The primers were designed in house and purchased through Sigma-Genosys. The linear DNA was ligated with T4 DNA ligase (Promega) following the manufacturer's instructions, and the circular DNA was transformed into DH5 α *E. coli* competent cells for DNA amplification. The variant DNA was verified by automated sequencing at the GSU core facility. The cDNA encoding the EGFP variants with BamH I and EcoR I restriction enzyme sites between the N and C terminals was subcloned into mammalian expression vector pcDNA3.1+, which utilizes a CMV promoter (Invitrogen).

Bacterial Expression and Purification

The proteins were expressed in the vector pet28a (EMD Biosciences) with an N-terminal 6x His-tag by BL21(DE3) *E. coli* and in LB-kanamycin (30 $\mu\text{g}/\text{mL}$). Expression was induced at an O.D.600 of 0.6 with 0.2 mM IPTG, and expression was allowed to continue for 22 hours before the cells were harvested by centrifugation. For these studies, the temperature was controlled at both 30 °C and 37 °C after induction. Three 1 ml samples were collected at time points throughout the expression, and centrifuged at 14 K rpm for 3 min. The cell pellets were resuspended in 1 ml of 10 mM Tris buffer at pH 7.4, and 200 μl was analyzed using a FLUOstar OPTIMA (BMG Labtech) plate reader with excitation filter of 460 nm and an emission filter at 510 nm to monitor the expression and folding of EGFP and its variants.

Protein purification was carried out using an Amersham-Pharmacia 5 mL HiTrap chelating HP column charged with nickel. The cell pellets were resuspended in 20 mM Tris, 10 mM NaCl, 0.1% Triton X-100, pH 8.8 and sonicated. The cellular debris was removed by centrifugation and the protein was loaded onto the prepared HiTrap column connected to an Amersham-Pharmacia AktaPrime FPLC. After washing to remove contaminant proteins, the protein of interest was eluted with an imidazole gradient. Contaminant imidazole was removed by dialysis, and the protein was further purified using a HiTrap Q ion-exchange column (Amersham) with a NaCl gradient at pH 8.0. Protein purity was verified by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

Mammalian Cell Culture

HeLa cells were grown on 60 mm culture dishes in Dulbecco's Modified Eagles Medium (DMEM, Sigma Chemical Co., St. Louis, MO) with 44 mM NaHCO_3 , pH 7.2, and supplemented with 10% (v/v) fetal bovine serum (FBS), 100 U/ml penicillin and 0.1 mg/ml streptomycin (Pen/Strep) at 37 °C with 5% CO_2 humidified incubation chamber. HeLa cells were grown to 80–90% confluency before transient transfection.

Plasmid DNA used for transfection was harvested from transformed *E. coli* (DH5- α) using QIAGEN's miniprep protocol (Qiagen). Each of nine GFP variants was individually and transiently transfected into HeLa cells with Lipofectamine-2000 (Invitrogen Life Technologies) and serum-free Opti-MEMI (Gibco Invitrogen Corporation) per the manufacturer's instructions. A typical transfection consisted of 1 or 2 μ g plasmid DNA with a ratio of DNA to Lipofectamine2000 between 1:1 and 1:3 (μ g/ μ l), dependent upon the protein construct. Protein expression was allowed to proceed for 48 and 72 h before inverted epifluorescence imaging. Control transfections with EGFP were performed in the same conditions as each construct.

Fluorescence microscopy imaging and its quantifications

An inverted epifluorescence microscope (Zeiss Axiovert 200) was utilized for fluorescence intensity screening *in vivo*, and images were obtained with an AxioCam 5 CCD camera and Zeiss Axiovision Rel 4.3 software. The microscope is equipped with a Xenon Arc Lamp, standard DAPI, FITC, and Texas Red filters, and transmitted light. For the fluorescence intensity measurements of the different protein constructs with each set of mutations, the 40x dry objective was utilized with FITC filters and exposure times ranging from 50 to 2000 ms. The images with exposure times allowing for fluorescence intensity within the dynamic range were utilized for data analysis. The fluorescence intensity measured in this time range was a linear function of the exposure time. AxioVision LE Rel. 4.3 software (AxioCam HRc) was used to quantify the fluorescence excited at 488 nm of the HeLa cells transfected with various GFP variants. Both area and mean fluorescence intensity of transfected cells ($n > 20$ cells per image) were measured and the total mean fluorescence intensity of cells in each imaged field was obtained with the calculation of Eq. (1):

$$F = \frac{\sum_{i=1}^n S_i F_i}{\sum_{i=1}^n S_i} \quad (1)$$

in which, F is the total mean fluorescence intensity excited at 488 nm of cells in each image, n is the number of fluorescent cells, S_i is the area of i^{th} fluorescent cell, and F_i is the mean fluorescent intensity excited at 480 nm of i^{th} fluorescent cell.

The total mean fluorescent intensity excited at 488 nm of the HeLa cells two days after transfection with EGFP-D2, EGFP-G1, or EGFP was used as a reference for the comparison of fluorescence intensity in the same group, respectively, according to Eq. (2):

$$F' = \frac{F}{F_0} \quad (2)$$

in which, the F' is the relative fluorescent intensity excited at 488 nm of the HeLa cells, F is the total mean fluorescence intensity excited at 488 nm of the HeLa cells, and F_0 is the total mean fluorescent intensity excited at 488 nm of the HeLa cells incubated for two days after transfection with EGFP-D2, EGFP-G1, or EGFP.

Measurement of ultra-violet (UV) and visible absorption spectrum

Spectroscopic properties of EGFP and its variants were measured using purified proteins by UV and visible absorption spectra with a Shimadzu UV and Visible Light Spectrophotometer

from 600 to 220 nm. The concentrations of the proteins were determined by the absorbance at 280 nm using the molar extinction coefficient of $21,890 \text{ M}^{-1}\text{cm}^{-1}$ calculated from the contribution from aromatic residues (1 Trp and 11 Tyr) (5500 and $1490 \text{ M}^{-1}\text{cm}^{-1}$ for Trp and Tyr, respectively). The extinction coefficient (398 nm or 490 nm) of the EGFP variants was obtained by Eq. (3):

$$\varepsilon_p = \varepsilon_{p,280\text{nm}} \left(\frac{A_p}{A_{p,280\text{nm}}} \right) \quad (3)$$

in which the ε_p is the extinction coefficient at 398 nm or 490 nm of EGFP variants, $\varepsilon_{p,280\text{nm}}$ is the extinction coefficient at 280 nm of EGFP variants, A_p is the absorption of EGFP variants at 398 nm or 490 nm, and $A_{p,280\text{nm}}$ is the absorption of EGFP variants at 280 nm. EGFP was used as a reference in the measurement of the extinction coefficients of the variants.

Fluorescence excitation and emission spectra

Spectroscopic properties of EGFP and its variants were also monitored with their fluorescence spectra, measured in a Fluorescence Spectrophotometer (Photon Technology International, Inc.) with a 1 cm path length quartz cell at room temperature and at $1 \mu\text{M}$ concentration in 10 mM Tris and 1 mM DTT (pH 7.4). The slit widths of 3 nm and 5 nm were used for excitation and emission, respectively. The quantum yield of EGFP variants with different excitation wavelengths was obtained with a calculation of equation Eq. (4):

$$\phi_p = \phi_r \left(\frac{A_r}{A_p} \right) \left(\frac{F_p}{F_r} \right) \left(\frac{n_p^2}{n_r^2} \right) \quad (4)$$

in which, ϕ_p is the relative quantum yield excited at 490 nm of EGFP variants, ϕ_r is the relative quantum yield excited at 490 nm of the reference sample, A_p is the absorption of EGFP variants at or 490 nm, A_r is the absorption of the reference sample at 490 nm, F_p is the integrated fluorescence intensity in the range of 500 nm to 600 nm excited at 490 nm of EGFP variants, F_r is the integrated fluorescence intensity in the range of 500 nm to 600 nm excited at 490 nm of the reference sample, n_p is the refractive index of EGFP variants, and n_r is the refractive index of the reference sample. EGFP was used as the reference sample in the measurement of quantum yield of EGFP variants.

Statistical analysis

Statistical analysis was performed with the software package Super ANOVA (Abacus Concepts, Berkeley, CA). Values were expressed as mean $\pm$ SEM. Control and treatment groups were compared by performing an analysis of variance (ANOVA). Fisher's Protected Least Significance Difference Test (Fisher's PLSD) was employed for post-hoc tests of statistical significance. Significance levels compared to day 1 are indicated as follows: * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

Results

Design of EGFP-based Ca^{2+} binding proteins

Two different types of Ca^{2+} binding sites were created in EGFP. Fig. 1A shows the design of EGFP-D2 containing a discontinuous Ca^{2+} binding site designed by our MetalFinder software based on common pentagonal bipyramidal geometry and chemical properties [14,15]. It is formed from five negatively charged ligand residues from sidechain carboxyl groups by mutated amino acids, S2D, L194E, and S86D, and the natural ligands of D82 and E5. Fig. 1B

shows the engineering of a continuous Ca^{2+} binding site EGFP-G1 by grafting the EF hand Ca^{2+} binding motif III of calmodulin at loop 9 between residues E172-D173 of EGFP [17]. In addition, to fulfil the required criteria for Ca^{2+} binding to have proper local Ca^{2+} binding geometric properties and charge arrangement, we also selected these Ca^{2+} binding sites based on the following additional criteria to assist chromophore formation. First, the site location and residue mutations should not abolish the chromophore synthesis or folding of the protein. Any residues that are conserved in fluorescent proteins and essential for protein structure and folding are not mutated. In addition, the location should be in a solvent-exposed region to have good accessibility to enable Ca^{2+} binding. Further, to avoid the drastic alteration of the protein folding and chromophore formation by introducing the charged Ca^{2+} ligand residues, a putative Ca^{2+} binding pocket with less required mutations was preferred. The effect of folding mutations on improvement of the fluorescence at both 30 °C and 37 °C was indicated with C2 and C3 mutations in the two types of GFP-based Ca^{2+} binding proteins (EGFP-D2 and EGFP-G1), respectively. The same mutations, C2 and C3, were also applied to EGFP.

Bacteria expression of EGFP-based Ca^{2+} binding proteins

To ensure that protein expression disparities were not responsible for the differences observed, each protein was compared qualitatively in cell extract via SDS-PAGE and found to be similar. Other work in our laboratory with purified protein suggests the expression efficiency of these proteins is similar with both fluorescent and non-fluorescent protein expressed at ~ 20 mg/L in bacteria and ~ 1 mg/L in mammalian cells. Moreover, while these proteins are grafted/ mutated to contain Ca^{2+} binding sites, the Ca^{2+} concentration during these measurements is constant at the physiologic cytosolic concentration of ~1 μM . Comparison of these engineered proteins is possible since they have their Ca^{2+} binding affinity with dissociation constant greater than 50 μM (shown in our published data[17] and unpublished data) and all of them are predominant in the apo-form under the cytosolic condition.

Nine proteins EGFP, EGFP-C2, EGFP-C3, EGFP-D2, EGFP-D2-C2, EGFP-D2-C3, EGFP-G1, EGFP-G1-C2, and EGFP-G1-C3 were first expressed in bacteria (BL 21) at both 30 and 37 °C to examine the differences in the chromophore formation by monitoring the fluorescence intensity at 510 nm (excited at 488 nm) using a FLUOstar OPTIMA (BMG Labtech) plate reader. Average intensities of nine proteins were taken in five time points throughout their expression. Fig. 2 lists the normalized average fluorescence intensities for 22 hrs after IPTG induction at both 30 and 37 °C, calculated using the method described in the Materials and Methods. Compared to EGFP at both 30 °C and 37 °C, the fluorescent intensities of both EGFP-D2 and EGFP-G1 are significantly low due to designing a discontinuous Ca^{2+} binding site and grafting a continuous Ca^{2+} binding motif in EGFP, respectively. The C2 and C3 mutations in EGFP-D2 resulted in 37- and 18-fold increases of its fluorescence intensity at 30 °C, respectively. Moreover, the fluorescence intensity increases (6- and 4-fold) were also observed with C2 and C3 mutations in EGFP-G1 at 30 °C. However, a similar fluorescence intensity increase was not observed with C2 and C3 mutations in EGFP at 30 °C, likely due to the already optimal folding properties of EGFP. The similar effect of C2 and C3 mutations on chromophore formation of both EGFP-D2 and EGFP-G1 was also observed at 37 °C although fluorescence intensity was decreased at this temperature compared to that at 30 °C. Taken together, the results shown in Fig. 2 indicate that the C2 constructs for both EGFP-D2 and EGFP-G1 surprisingly exhibit an increased fluorescence over the C3 variants and F99S actually interferes with the folding of the protein variants when applied to the M153T/V163A construct though it is not as low in fluorescence as the protein variants without C2 and C3 mutations.

Mammalian cell expression of EGFP-based Ca^{2+} binding proteins

The effect of C2 and C3 mutations on the expression of EGFP variants in mammalian cells was monitored using a fluorescence microscope. Fig. 3 illustrates the fluorescence microscopy

imaging of HeLa cells expressing EGFP-G1, EGFP-G1-C2, and EGFP-G1-C3 for 48 hr after transfection at both 30 °C and 37 °C. As shown in Fig. 3A–C, two days after transfection and expression at 30 °C, EGFP-G1 variant and its C2 and C3 mutations were expressed and folded well in the majority of HeLa cells as indicated by their strong fluorescence signals. However, Fig. 3D–E showed that EGFP-G1 lost its fluorescence signal at 37 °C, indicating that this temperature was not optimal to chromophore formation of EGFP-G1 in HeLa cells. In contrast, the addition of C2 and C3 mutations in EGFP-G1 resulted in the improvement of chromophore formation at 37 °C in HeLa cells.

Fig. 4 demonstrates a quantitative analysis of fluorescence intensity of HeLa cells (more than 20 cells per image) transfected with different series of EGFP-D2, EGFP-G1, and EGFP at both 30 °C and 37 °C. A low fluorescence intensity of HeLa cells transfected with EGFP-D2 was observed in Fig. 4A at both 30 °C and 37 °C. Compared to EGFP-D2, the C2 and C3 mutations resulted in 4.4- and 3.0- fold increases of its fluorescence intensity at 30 °C, respectively. The similar effect of C2 and C3 mutations on chromophore formation of EGFP-D2 was also observed at 37 °C though fluorescence intensity was decreased at this temperature compared to that at 30 °C. Moreover, the fluorescence intensity increases (11 and 9.9 folds) were also observed with C2 and C3 mutations in EGFP-G1 at 37 °C even though the same effect was not observed at 30 °C as shown in Fig. 4B. However, the similar effect of C2 and C3 mutations on fluorescence intensity of EGFP series proteins was not observed at both 30 °C and 37 °C as shown Fig. 4C. The results obtained in mammalian cells are consistent with those observed in *E. coli*.

Spectroscopic properties of EGFP-based Ca²⁺ binding proteins

Spectroscopic analysis of the optical properties such as extinction coefficients and quantum yields were performed with purified proteins. Both UV-visible absorption and fluorescence spectra of each purified protein were measured with different protein concentrations. Then, extinction coefficients and quantum yields of each protein were calculated with Eq (3) and (4) shown above, respectively. EGFP was used as a reference in the measurement of both extinction coefficients and quantum yields of the variants. Fig. 5 shows the visible absorbance and fluorescence emission spectra of EGFP-D2, EGFP-G1, and EGFP at pH 7.4. As shown in Fig. 5A, a major absorbance peak at 488 nm and a minor absorbance peak at 398 nm were observed in the visible spectra of EGFP, indicating that the anionic state of the chromophore was the main form in EGFP. A fluorescence emission peak at 510 nm excited at 488 nm was shown in EGFP fluorescence spectrum (Fig. 5B). The spectroscopic properties of EGFP and its variants including both extinction coefficients and quantum yields at 398 and 488 nm are summarized in Table 1[1,21,22]. The Ca²⁺-binding site formation of EGFP-D2 with three mutated ligands S2D, L194E and S86D and two natural ligands D82 and E5 resulted in a large decrease of visible absorption at both 398 and 488 nm as observed in Fig. 5A. Compared to EGFP, for example, the extinction coefficient at 488 nm of EGFP-D2 was decreased from 56 mM⁻¹ cm⁻¹ to 9.3 mM⁻¹ cm⁻¹. Concurrently, the fluorescence emission peak at 510 nm was also largely decreased in its fluorescence spectrum (Fig. 5B); although the quantum yield of EGFP-D2 was almost the same with that of EGFP. Strikingly, both C2 and C3 mutations in EGFP-D2 re-produced the major absorbance peak at 488 nm and minor absorbance peak at 398 nm similar to that of EGFP, as indicated in Table 1. Additionally, the visible absorption spectrum of EGFP-G1 showed a slight increase at 398 nm and a decrease at 488 nm, compared to EGFP shown in Fig. 5A. The extinction coefficient at 488 nm of EGFP-G1 was decreased from 56 mM⁻¹ cm⁻¹ to 28 mM⁻¹ cm⁻¹ and the extinction coefficient at 398 nm was increased from 5.1 mM⁻¹ cm⁻¹ to 9.2 mM⁻¹ cm⁻¹ compared to EGFP. Many previous reports demonstrated that both neutral and anionic states of chromophore exist in normal GFP [1,2]. The neutral state of chromophore exhibits its maximum absorbance at 397 nm and the anionic state of chromophore exhibits its maximum absorbance at 490 nm. The interactions of

chromophore and its environment residuals including His148, Thr203, Ser205, and Glu222 contribute to the equilibrium between neutral state and anionic state of chromophore. The results shown above in EGFP-G1 indicate that grafting a continuous Ca^{2+} binding motif induced a transfer from the anionic state to the neutral state of the chromophore, pushing the equilibrium from the anionic phenolate to the neutral phenol. However, as shown in Table 1, no effect from the C2 and C3 mutations on extinction coefficient and quantum yield of EGFP-G1 was observed. In addition to EGFP-G1, the results in Table 1 also showed that the C2 and C3 mutations did not affect the optical properties of EGFP, indicating the relative distribution of anionic-neutral states of the chromophore was not altered by the addition of folding mutations.

Discussion

Development of EGFP-based Ca^{2+} sensors with new approaches

There is a strong need to develop Ca^{2+} sensors capable of real-time quantitative Ca^{2+} measurements in specific subcellular environments [23–28]. In order to develop Ca^{2+} sensors without utilizing natural Ca^{2+} binding proteins to overcome the limitation of perturbing Ca^{2+} signalling network, our lab has developed several GFP-based Ca^{2+} binding proteins with two different approaches: (1) grafting a continuous Ca^{2+} -binding motif and (2) designing a discontinuous Ca^{2+} binding site in EGFP using the common features of Ca^{2+} -binding sites in proteins [11–17]. Grafting CaM loop III with flanking EF-helices (helix-loop-helix Ca^{2+} binding motif) at Glu172-Asp173 of EGFP, EGFP-G1 resulted in the formation of a protein with an increased extinction coefficient at 398 nm ($\epsilon_{398\text{nm}}$) and decreased extinction coefficient at 488 nm ($\epsilon_{488\text{nm}}$) compared to that of EGFP-wt shown in Table 1. Thus, the grafting of a Ca^{2+} binding motif at Glu172-Asp173 (position 1) in EGFP significantly shifts the population of the chromophore from the anionic state to the neutral state, suggesting that Glu172-Asp173 of EGFP is a chromophore sensitive location [17]. As another approach we have designed Ca^{2+} binding site using direct mutations (S2D, S86D, and L194E) in EGFP, EGFP-D2. Ca^{2+} binding was observed in this engineered EGFP (data not shown). Unfortunately, the formation of the Ca^{2+} binding site resulted in a large loss (more than 99% compared to that of EGFP shown in Fig. 2) in fluorescence when expressed in bacteria at both 30 and 37 °C.

Improvement of GFP folding and chromophore formation

GFP chromophore fluorescence may be affected by many factors including the folding trajectory of the protein and interactions of the chromophore with surrounding amino acids. The loss of fluorescent properties is a common problem encountered while engineering ligand binding sites in a fluorescent protein [17,29–33]. As shown in Figs 2–4, the fluorescent intensities of both EGFP-D2 and EGFP-G1 are significantly decreased due to designing a discontinuous Ca^{2+} binding site and grafting a continuous Ca^{2+} binding motif in EGFP, respectively. Moreover, our previous work has also shown that the addition of a continuous Ca^{2+} binding loop at Asn144-Tyr145 of EGFP results in the loss of chromophore formation [17]. Numerous efforts have been made to identify important mutations called “folding” mutations to improve maturation of GFP in both bacterial and mammalian systems, especially at 37 °C [3,19,20,34–37]. Cubitt et al. summarized four groups of “folding” mutations: (1) mutations close to the chromophore, including five (S65A, -G, -C, -T, and -L) actually within the chromophore, two (F64L and S72A) that are close to it in the central α -helix, and others (Y145F, I167T, T203Y, and S205T) that are partially or completely buried; (2) mutations distant from the chromophore and in a buried location (V163A); (3) surface-located “folding” mutations (F99S, M153T, and S175G); and (4) surface-located “folding” mutations that also lie close to the chromophore (S147P and N149K) [38]. The “cycle3” mutation denoted as C3 was performed with (1) the substitution of two surface hydrophobic residues with hydrophilic ones (F99S and M153T) and (2) another substitution of hydrophobic residues with a less

hydrophobic one (V163A) in wild-type GFP [19,20]. Thus, C3 mutation in wild-type GFP decreases the protein tendency toward aggregation during folding and makes the protein more soluble than wild-type GFP. However, the effect of C3 mutation on the protein stability was not observed because it affects the stability of both the native and the unfolded states to the same extent. The results described above in this paper suggest that C2 mutation is sufficient to facilitate the formation of GFP chromophore upon designing a discontinuous Ca^{2+} binding site and grafting a continuous Ca^{2+} -binding motif in EGFP.

Comparison of protein manufacture in bacterial and mammalian cells

The manufacturing of a protein is a multistep process including transcription of a genome, translation of messenger RNA, and folding of newly synthesized peptide with chaperone systems in both prokaryotic and eukaryotic cells. The different cell environments and chaperone mechanisms are possible factors affecting the manufacture of EGFP and its variants in *E. coli* and HeLa cells which were applied in this study. The cell growth and protein expression are generally much faster in *E. coli* compared to that in HeLa cells. Moreover, DnaK/DnaJ, the first major chaperone system in bacteria, binds the newly synthesized unfolding polypeptides to prevent nascent chains against misfolding and aggregation after they are synthesized in the ribosome of bacteria. Then, GroEL/ES, the secondary major chaperone system in bacteria, receives the polypeptides in a GroEL/ES-dependent transfer step and mediates their folding to native state although the transfer to GroEL/ES is not necessary for all proteins in bacteria [39–42]. Partially due to the rapid expression of proteins in bacteria, proteins with reduced folding capability may not fold properly and, therefore, may be shunted to inclusion bodies.

In contrast, nascent polypeptide associated complex (NAC), a protein factor, may bind to short newly synthesized unfolding polypeptides (less than 30 amino acids) even before the newly synthesized unfolding polypeptides are bound with heat-shock proteins 40/70 (hsp40/70), the first major chaperone system in eukaryotic cells. The major function of hsp40/70 is to prevent the misfolding and aggregation of unfolding polypeptides following their synthesis in the ribosome of eukaryotic cells. NAC may form an external wall of ribosome-exit-pathway and play a critical role in deciding the protein folding locations, cytoplasm or endoplasmic reticulum, in eukaryotic cells, distinguishing the chaperone mechanism from bacteria. Another difference between prokaryotic and eukaryotic cells is that eukaryotic proteins are not released into the cytoplasm before they perform final folding with the eukaryotic chaperone CCT, the secondary major chaperone system in eukaryotic cells [42–47]. Compared to prokaryotic cells, thus, eukaryotic cells are more suitable to the manufacture of mutant proteins with reduced folding capacity due to disruption of sidechain and backbone interaction although the production rate is not high.

Mutant proteins for various purposes are highly likely to undergo alternate or reduced folding as the mutations disrupt sidechain and backbone contacts present in the native protein. The reduced tertiary structure packing is not always easily observed or is only observed as a reduction in protein yield. A protein like GFP, where the proper fold is absolutely necessary to form the chromophore, is more readily perceived to be disrupted. Compared to chromophore fluorescence of EGFP, for example, both direct mutation of native ligands (EGFP-D2) (less than 1%) and grafting a continuous Ca^{2+} binding motif (EGFP-G1) in a certain location of GFP (less than 10%) causes a large reduction in chromophore fluorescence when proteins were expressed in bacteria (Fig. 2). The incorporation of negatively charged sidechains and grafting an EF-hand Ca^{2+} binding motif likely reduces the tertiary packing of the protein, causing the loss of chromophore formation. Mammalian cells have better folding capabilities than bacteria for expression and packing of protein tertiary structure. Our results shown in Fig. 4 clearly

suggest that EGFP-D2 and EGFP-G1 exhibits more observable fluorescence when expressed in mammalian cells than when expressed in bacteria.

Application of C2 and C3 mutations in developing EGFP-based Ca^{2+} sensors

To improve the folding of the mutant protein, thereby improving the chromophore fluorescence, the well-known cycle3 mutation was incorporated and studied for chromophore fluorescence effects. C2 mutation was successfully applied to improve the protein expression in both bacteria and mammalian cells shown above in this study. At both 30 and 37 °C, the chromophore was observed to better form with the inclusion of the C2 mutation. The molar extinction coefficient of EGFP-D2-C2 at 488 nm is nearly equal that of EGFP, suggesting that the C2 mutations nearly abolish the detrimental effects of the negatively charged sidechains of the designed calcium binding site. This effect is likely due to better tertiary packing of the protein caused by the C2 mutations. Although inclusion of all 3 mutations with F99S (EGFP-D2-C3) does also increase the fluorescence versus EGFP-D2 without any of the folding mutations, the F99S mutation is unfavorable overall with a decrease in fluorescence observed as compared to the fluorescence of EGFP-D2-C2. Therefore, our study suggests that in engineering metal binding proteins based on EGFP, especially those which require charged residues, adding the C2 mutation is beneficial to facilitate the formation of chromophore and protein folding both in mammalian and bacterial systems.

Conclusions

The effect of C2 (M153T/V163A) and C3 (F99S/M153T/V163A) mutations on both *in vitro* and *in vivo* EGFP folding was investigated in reference to the design of two types of Ca^{2+} binding sites including (1) designing a discontinuous Ca^{2+} binding site and (2) grafting a continuous Ca^{2+} -binding motif in EGFP. The results show that, for the constructed EGFP variants, C2 mutation is sufficient to facilitate the production of fluorescence in both bacterial and mammalian cells. Further addition of the mutation F99S decreases the folding efficiency of these mutants although a similar effect is not detectable for EGFP. The extinction coefficient and quantum yield of purified proteins of each construct were also examined to compare the effects of both C2 and C3 mutations on protein spectroscopic properties. Our quantitative analyses of the effect of these folding mutants on the fluorescence intensity that undergo different folding trajectories in bacteria versus mammalian cells provide molecular insight into the development of analytic sensors based on fluorescent proteins.

Acknowledgments

We would like to thank Dan Adams, Michael Kirberger, and Nancy Huang for their critical reviews of this manuscript and helpful discussions and other members of Dr. Yang's group for their helpful discussions and suggestions. This work is supported in part by the following sponsors: NIH GM 62999-1, GM-70555 to JJY and GSU Molecular Basis of Disease Pre-doctoral Fellowship to NC.

References

1. Tsien RY. Annu Rev Biochem 1998;67:509–544. [PubMed: 9759496]
2. Zimmer M. Chem Rev 2002;102:759–781. [PubMed: 11890756]
3. Heim R, Cubitt AB, Tsien RY. Nature 1995;373:663–664. [PubMed: 7854443]
4. Berridge MJ, Bootman MD, Lipp P. Nature 1998;395:645–648. [PubMed: 9790183]
5. Bers DM. Nature 2002;415:198–205. [PubMed: 11805843]
6. Schafer BW, Heizmann CW. Trends in Biochemical Sciences 1996;21:134–140. [PubMed: 8701470]
7. Song LS, Guatimosim S, Gomez-Viquez L, Sobie EA, Ziman A, Hartmann H, Lederer WJ. Ann N Y Acad Sci 2005;1047:99–111. [PubMed: 16093488]

8. Nagamune K, Moreno SN, Chini EN, Sibley LD. *Subcell Biochem* 2008;47:70–81. [PubMed: 18512342]
9. Berridge MJ. *Cell Calcium* 2002;32:235–249. [PubMed: 12543086]
10. Corbett EF, Michalak M. *Trends Biochem Sci* 2000;25:307–311. [PubMed: 10871879]
11. Yang, JJ.; Yang, W. Calcium Binding Proteins, *Encyclopedia of inorganic chemistry*. Vol. 2. Bruce, King R., editor. John Wiley & Sons, Ltd; 2003.
12. Ye Y, Lee HW, Yang W, Shealy S, Yang JJ. *J Am Chem Soc* 2005;127:3743–3750. [PubMed: 15771508]
13. Yang JJ. *Calcium Binding Proteins* 2006;1:65–66.
14. Yang W, Wilkins AL, Li S, Ye Y, Yang JJ. *Biochemistry* 2005;44:8267–8273. [PubMed: 15938616]
15. Yang W, Wilkins AL, Ye Y, Liu ZR, Li SY, Urbauer JL, Hellings HW, Kearney A, van der Merwe PA, Yang JJ. *J Am Chem Soc* 2005;127:2085–2093. [PubMed: 15713084]
16. Zou J, Ye Y, Welshhans K, Lurtz M, Ellis AL, Louis C, Rehder V, Yang JJ. *J Biotechnol* 2005;119:368–378. [PubMed: 15935502]
17. Zou J, Hofer AM, Lurtz MM, Gadda G, Ellis AL, Chen N, Huang Y, Holder A, Ye Y, Louis CF, Welshhans K, Rehder V, Yang JJ. *Biochemistry* 2007;46:12275–12288. [PubMed: 17924653]
18. Kremers GJ, Goedhart J, van den Heuvel DJ, Gerritsen HC, Gadella TW Jr. *Biochemistry* 2007;46:3775–83. [PubMed: 17323929]
19. Crameri A, Whitehorn EA, Tate E, Stemmer WP. *Nat Biotechnol* 1996;14:315–319. [PubMed: 9630892]
20. Fukuda H, Arai M, Kuwajima K. *Biochemistry* 2000;39:12025–12032. [PubMed: 11009617]
21. Ward WW. *Methods Biochem Anal* 2006;47:39–65. [PubMed: 16335709]
22. Shaner NC, Campbell RE, Steinbach PA, Giepmans BN, Palmer AE, Tsien RY. *Nat Biotechnol* 2004;22:1567–72. [PubMed: 15558047]
23. Tallini YN, Ohkura M, Choi BR, Ji G, Imoto K, Doran R, Lee J, Plan P, Wilson J, Xin HB, Sanbe A, Gulick J, Mathai J, Robbins J, Salama G, Nakai J, Kotlikoff MI. *Proc Natl Acad Sci U S A* 2006;103:4753–4758. [PubMed: 16537386]
24. Reiff DF, Ihring A, Guerrero G, Isacoff EY, Joesch M, Nakai J, Borst A. *J Neurosci* 2005;25:4766–4778. [PubMed: 15888652]
25. Kotlikoff MI. *J Physiol* 2007;578:55–67. [PubMed: 17038427]
26. Mank M, Santos AF, Drenth S, Mrcic-Flogel TD, Hofer SB, Stein V, Hendel T, Reiff DF, Levelt C, Borst A, Bonhoeffer T, Hubener M, Griesbeck O. *Nat Methods* 2008;5:805–811. [PubMed: 19160515]
27. Rochefort NL, Konnerth A. *Nat Methods* 2008;5:761–762. [PubMed: 18756196]
28. Mank M, Griesbeck O. *Chem Rev* 2008;108:1550–1564. [PubMed: 18447377]
29. Abedi MR, Caponigro G, Kamb A. *Nucleic Acids Res* 1998;26:623–630. [PubMed: 9421525]
30. Doi N, Yanagawa H. *FEBS Lett* 1999;453:305–307. [PubMed: 10405165]
31. Topell S, Hennecke J, Glockshuber R. *FEBS Lett* 1999;457:283–289. [PubMed: 10471794]
32. Baird GS, Zacharias DA, Tsien RY. *Proc Natl Acad Sci U S A* 1999;96:11241–11246. [PubMed: 10500161]
33. Chiang CF, Okou DT, Griffin TB, Verret CR, Williams MN. *Arch Biochem Biophys* 2001;394:229–235. [PubMed: 11594737]
34. Delagrave S, Hawtin RE, Silva CM, Yang MM, Youvan DC. *Biotechnology (N Y)* 1995;13:151–154. [PubMed: 9634755]
35. Cormack BP, Valdivia RH, Falkow S. *Gene* 1996;173:33–38. [PubMed: 8707053]
36. Heim R, Tsien RY. *Curr Biol* 1996;6:178–182. [PubMed: 8673464]
37. Battistutta R, Negro A, Zanotti G. *Proteins* 2000;41:429–437. [PubMed: 11056031]
38. Cubitt AB, Woollenweber LA, Heim R. *Methods Cell Biol* 1999;58:19–30. [PubMed: 9891372]
39. Bukau B, Horwich AL. *Cell* 1998;92:351–366. [PubMed: 9476895]
40. Mayer MP, Bukau B. *Biol Chem* 1998;379:261–268. [PubMed: 9563820]
41. Frydman J. *Annu Rev Biochem* 2001;70:603–647. [PubMed: 11395418]

42. Fink AL. *Physiol Rev* 1999;79:425–449. [PubMed: 10221986]
43. Braig K. *Curr Opin Struct Biol* 1998;8:159–165. [PubMed: 9631288]
44. Carrascosa JL, Llorca O, Valpuesta JM. *Micron* 2001;32:43–50. [PubMed: 10900379]
45. Jacob E, Horovitz A, Unger R. *Bioinformatics* 2007;23:i240–i248. [PubMed: 17646302]
46. Horovitz A, Willison KR. *Curr Opin Struct Biol* 2005;15:646–651. [PubMed: 16249079]
47. Zhang X, Beuron F, Freemont PS. *Curr Opin Struct Biol* 2002;12:231–238. [PubMed: 11959502]

Figure 1. Model structure of Ca²⁺ binding fluorescent protein with the addition of EGFP-D2 (A) by computational design and EGFP-G1 (B) by grafting the EF-hand motif III from calmodulin at position 172. The model structures are constructed with PyMOL v0.99 based on 1EMA.pdb.

Figure 2.

Quantitative comparison of EGFP and its variant expression in *E. coli* bacteria (BL21-DE3) at 30° C and 37° C, respectively. Fluorescent intensity at 510 nm of different cell pellets was monitored with a FLUOstar OPTIMA (BMG Labtech) plate reader for 22 hrs after 200 mM IPTG induction. $\lambda_{\text{ex}} = 488 \text{ nm}$.

Figure 3.

Comparison of fluorescence microscope imaging of HeLa cells. The imaging was performed two days (48 hrs) after HeLa cells were transfected with EGFP-G1, EGFP-G1-C2, and EGFP-G1-C3 and expressed at both 30° C and 37° C. The exposure time was 200 ms. $\lambda_{ex} = 488$ nm.

Figure 4.

Quantitative expression comparison of EGFP-D2 series (A), EGFP-G1 series (B), and EGFP series (C) in mammalian cells (HeLa) at 30° C and 37° C, respectively. Fluorescent intensity at 510 nm of different cell pellets was obtained with fluorescence microscopy imaging for 2 days after transfection. $\lambda_{ex} = 488$ nm.

Figure 5.

Visible absorbance (A) and fluorescence (B) spectra of EGFP (○), EGFP-D2 (Δ), and EGFP-G1 (□). The measurements were performed in 10 mM Tris and 1 mM DTT (pH 7.4). The protein concentrations were 2 μ M for both absorbance and fluorescence experiments. The length of the cell was 1 cm. The fluorescence experiments were performed with a slit width of 1 nm for both excitation and emission. $\lambda_{\text{ex}} = 488$ nm.

Table 1

Spectroscopic properties of EGFP-D2, EGFP-G1, and EGFP and their C2/C3 variants

Protein	Extinction Coefficient (ϵ), $\text{mM}^{-1} \text{cm}^{-1}$		Quantum Yield 488 nm
	398 nm	488 nm	
EGFP-D2	1.3	9.3	0.56
EGFP-D2-C2	5.5	53	0.53
EGFP-D2-C3	5.5	56	0.51
EGFP-G1	9.2	28	0.59
EGFP-G1-C2	11	27	0.68
EGFP-G1-C3	9.9	28	0.69
EGFP	5.1	56	0.60
EGFP-C2	7.2	56	0.58
EGFP-C3	6.7	56	0.63