Biophysical Characterization of Natural and Mutant Fluorescent Proteins Cloned from Zooxanthellate Corals*

Yi Sun\textsuperscript{1,2}, Edward W. Castner, Jr.\textsuperscript{3}, Catherine L. Lawson\textsuperscript{3}, and Paul G. Falkowski\textsuperscript{1¶}

\textsuperscript{1}Environmental Biophysics and Molecular Ecology Program, Institute of Marine and Coastal Sciences, Rutgers University, 71 Dudley Road, New Brunswick, NJ 08901, USA
\textsuperscript{2}Present address: Department of Chemical Biology, Ernest Mario School of Pharmacy, Rutgers University, 164 Frelinghuysen Road, Piscataway, NJ 08854, USA
\textsuperscript{3}Department of Chemistry and Chemical Biology, Rutgers University, 610 Taylor Road, Piscataway, NJ 08854, USA

Running title: Biophysical Properties of GFPs

\* This work was supported by the Office of Naval Research (Award #: N00014-00-1-0795).

\textsuperscript{a} The sequences reported in this paper have been deposited in the GenBank database with accession numbers: AY056460 (mcCFP); AY155343 (mmGFP); AY155344 (mmGFP \textsuperscript{S110N}); AY362545 (mcRFP). \textsuperscript{b} Amino acid positions mentioned in this paper correspond to \textit{Aequorea victoria} GFP (accession number: P42212).

¶ Corresponding Author: Paul G. Falkowski
Institute of Marine and Coastal Sciences, Rutgers University,
71 Dudley Road, New Brunswick, NJ 08901, USA.
Tel.: 1-732-932-6555;
Fax: 1-732-932-4083;
E-mail: falko@imcs.rutgers.edu

ABBREVIATIONS: GFP, green fluorescent protein; DsRed, red fluorescent protein; RACE, rapid amplification cDNA end; PCR, polymerase chain reaction; GdnHCl, guanidinium chloride; bp, base pair. PAGE, polyacrylamide gel electrophoresis; TCSPC, time-correlated single photon counting.
SUMMARY

Two novel colored fluorescent proteins were cloned and biophysically characterized from zooxanthellate corals (Anthozoa). A cyan fluorescent protein derived from the coral Montastrea cavernosa (mcCFP) is a trimeric complex with strong blue-shifted excitation and emission maxima at 432 and 477 nm, respectively. The native complex has a fluorescence lifetime of 2.66 ± 0.01 ns and an inferred absolute quantum yield of 0.385. A green fluorescent protein cloned from Meandrina meandrites (mmGFP) is homologous to the GFP from the hydrozoan Aequorea victoria (avGFP). This monomeric protein has an excitation maximum at 398 nm and an emission maximum at 505 nm, a fluorescence lifetime of 3.09 ± 0.01 ns and an absolute quantum yield of 0.645. Sequence homology with avGFP and the red fluorescent protein (DsRed) indicates that the proteins adopt the classic β-barrel configuration with 11 β-strands. The three amino acid residues that comprise the chromophore are QYG for mcCFP and TYG for mmGFP, compared with SYG for avGFP. A single point mutation, Ser-110 to Asn, was introduced into mmGFP by random mutagenesis. Denaturation and refolding experiments showed that the mutant has reduced aggregation, increased solubility and more efficient refolding relative to the wild type. Time-resolved emission lifetimes and anisotropies suggest that the electronic structure of the chromophore is highly dependent on the protonation state of adjoining residues.
INTRODUCTION

Green fluorescent proteins (GFPs) appear to be ubiquitous in marine cnidarians (1, 2). Unlike those in the jellyfish, *Aequorea victoria*, from which they were first isolated, the GFPs found in most other cnidarians, including zooxanthellate corals (Anthozoa), do not transfer energy to a terminal luminescent acceptor, nor are they energetically coupled to photosynthetic electron transport in the symbiotic algae. Rather, the GFPs in these organisms exhibit prompt fluorescence with no known biological function. The retention of these proteins in non-luminescent organisms poses fundamental questions about their evolutionary history, molecular structure, biophysical properties and biological function of these commercially exploited molecules. In this paper, we present the first biophysical and biochemical characterization of two native and one mutant colored fluorescent proteins cloned from two common zooxanthellate corals found in the Caribbean.

More than 20 fluorescent GFP-like proteins and 7 non-fluorescent chromoproteins have been isolated and spectroscopically characterized from Anthozoa (1-4). The maximum emission wavelength for fluorescent GFP-like proteins ranges from 483 to 593 nm. No truly blue or far-red fluorescent proteins have been found in nature (2). Comparative sequence homology analyses suggest all the GFPs are derived from a single common ancestor, but have diverged significantly within the phylum cnidaria. These proteins are not found in bilateral metazoans.

All GFP structures characterized to date are barrels consisting of 11 β-sheets encasing an internal, cyclized fluorescent chromophore derived from the oxidation of three amino acids. The formation of the chromophore requires molecular oxygen. Protein folding and chromophore formations can be followed kinetically from fluorescence properties (8, 20, 21, 22). The excited state lifetimes for mature fluorescent proteins typically range from 3 to 4 ns. Their fluorescent
yield depends on pH and temperature (19), and is related to bulk activation energies for the formation of the secondary and tertiary structures.

The inherently low quantum yield and slow maturation times of many naturally occurring GFPs has stimulated the isolation of natural, novel variants as well as mutants with altered biophysical properties. The studies on the luminescent properties of the GFP from *Aequorea victoria* have been carried out on mutants with characteristics more conducive to biotechnological applications, including enhanced fluorescence intensity (5, 6), shifted emission wavelength (7-9), altered sensitivity to pH (10), and faster protein folding (11-13). Recent studies have focused on altering the fluorescent and biochemical properties of red fluorescent protein drFP583 (DsRed), including increasing the rate of “maturation” (14, 15) and reducing oligomerization (16). Rather little is known about the folding mechanisms and related biochemical and biophysical properties of other non-bioluminescent GFP-like proteins.

The objectives of the present study were to identify the factors regulating the fluorescent properties of novel fluorescent proteins in corals in an attempt to understand their potential functional roles. We present results from two proteins cloned from non-bioluminescent zooxanthellate corals: a blue–shifted, or cyan, fluorescent protein (termed mcCFP) derived from *Montastrea cavernosa*, and a GFP-like protein (mmGFP) from *Meandrina meandrites*, and a mutant derived from the latter.
EXPERIMENTAL PROCEDURES

Cloning, Expression and Mutagenesis of Target Proteins

Samples of the stony corals, Montastrea cavernosa and Meandrina meandrites, were collected from waters around the Caribbean Marine Research Center at Lee Stocking Island, Bahamas. Total RNA was extracted from colored tissue with TRIzol (Life Technologies). Total cDNA was synthesized using the SMART cDNA Synthesis Kit (Clontech) following the manufacturer’s protocol. Both 3’ and 5’ end fragments coding for fluorescent protein genes were amplified by RACE-PCR as described by Matz et al. (1). Gene-specific primers were designed as follows:

for mcCFP: 5’-GCGTCTTCTTCTGCTATAACTGGACCACTGGAGG-3’;
for mmGFP: 5’-TGGATTACAGGTCCATTGGCGGGAAAGT-3’. The anti-sense primer sequences came from Matz et al. (17). The restriction endonuclease sites (BstBI and BglII for mcCFP; HindIII and BstBI for mmGFP) were introduced into the pCRII vector (Invitrogen) to yield a full length of target cDNA. The coding region was inserted in frame into the pBAD-TOPO expression vector (Invitrogen) with a 6 X His tag at the C-terminus. The recombinant plasmid was transfected into an E. coli host (One Shot TOP 10, Invitrogen), and protein expression was induced by adding 0.2% L-arabinose to the RM medium at 37°C.

Random mutagenesis was achieved for mmGFP using E. coli strain ES1578 (kindly provided by the E. coli Genetic Stock Center, Yale University), which introduces random mutations during plasmid replication. Colonies were selected for increased fluorescence by visual inspection under illumination by UV light (350nm) over three generations. This process recovered a single mutant mmGFP<sub>S110N</sub>. Both wild type and mutant mmGFPs with C-terminal His-tags were expressed in E. coli BL21 (DE3) (Novagen) in LB medium.
Protein Purification, Spectral Characterization and Refolding Tests

Affinity purification of His-tagged wild type and mutant mmGFPs from BL21 (DE3) was performed as follows: cells were suspended in 50 mM Tris HCl (pH 8.0) and lysed by 100 µg/ml lysozyme. NaCl was added to the lysate to a final concentration of 100 mM. The preparation was cleared by centrifugation at 10000 x g for 15 min, at 4°C. Cell debris was removed by filtration through a 0.22 µm syringe filter and applied to a column of nickel-nitritolriacetic acid (Ni-NTA) agarose (Qiagen) at room temperature. The elution buffer was 50 mM Tris HCl (pH 8.0), 100 mM NaCl and 100 mM imidazole. Protein concentrations were determined using a BCA Protein Assay Kit (PIERCE) with a plate reader (Spectra MAX Gemini XS, Molecular Devices Corporation). The purity of all protein preparations was at least 95% as verified by 12% SDS-PAGE visualized with coomassie blue stain.

Room temperature absorbance and fluorescence spectra were measured using an SLM-Aminco DW-2000 UV-VIS spectrophotometer, and an Aminco Bowman Series 2 luminescence spectrometer, respectively. Time-dependent protein denaturation and refolding tests were conducted by following the kinetics of the change in fluorescence intensity. After 30 min denaturation in 6 M GdnHCl, refolding was induced by a 100-fold dilution of the total denatured protein (0.1mg/mL) into ice-cold refolding buffer (25 mM Tris-HCl, 150 mM NaCl, pH 8.0). The development of fluorescence was monitored for the next 15 min.

Kinetic Fluorescence Measurements

Affinity purified proteins were immediately incubated under different conditions and their fluorescence maturation monitored over time. Oxygen level, pH, and temperature were the main parameters examined. Fluorescence measurements were performed with 395 nm excitation and 505 nm emission wavelengths.
Oxygen demand during protein maturation was assessed by comparing time dependent fluorescence intensity at different partial pressures of air versus nitrogen using the luminescence spectrometer. Protein samples were diluted into refolding buffer at a final concentration of 30 µg/mL at room temperature (295K), pH 8.0. The O₂ concentrations were kept as 260µM or < 1µM by bubbling of compressed air or pure N₂, respectively. Both temperature and pH dependent half-life maturation tests were conducted using a fluorescence plate reader (Spectra MAX Gemini XS). pH dependence was measured by diluting protein samples with refolding buffers ranging in pH from 3.0 to 12.0 in 96-well optical plates. Maturation was monitored at room temperature by following the change in fluorescence over 48 hrs at a protein concentration of 15 µg/mL. Temperature-dependence of maturation was measured using a custom-built 96-well temperature gradient aluminum block that fit into a plate reader. This block permitted continuous temperature control from 3 to 48°C for optical measurements. The fluorescence was monitored every 60 sec in the refolding buffer at a final protein concentration of 15 µg/mL at pH 8.0. Activation energies, \(E_a\), were calculated from the Arrhenius equation by regression analysis of the reaction rate constant (i.e., fluorescence yield) as a function of 1/T (K).

**Time-Resolved Emission Spectroscopy**

Details of the time-correlated single photon counting (TCSPC) instrument have been described previously (25-27). Recent modifications include replacement of the NIM-style electronics with a Becker and Hickl TCSPC data acquisition board, Model SPC-630, and incorporation of a quartz depolarizer (Optics for Research) directly in front of the spectrometer entrance slit. The depolarizer eliminated polarization bias from the spectrometer grating during fluorescence anisotropy measurements.
A Glan-Laser polarizer was placed in front of the sample to define vertical polarization for excitation, and a matched polarizer was used as an analyzer in the emission direction for vertical (parallel), magic angle (54.7 degrees from vertical) and horizontal (perpendicular) fluorescence polarization decays.

Fluorescence decays were measured at 395 nm (excitation) and 505 nm (emission). The TCSPC transients were measured to contain up to $2^{16}$-1 fluorescence counts for high dynamic range data, spanning 33 ns, with 4096 data points, for a resolution of 8.06 ps/channel. Buffer solutions were the same as the refolding buffer and were checked for fluorescence background; no transient emission at 505 nm was observed. Temporal response profiles for the TCSPC instrument were obtained by scattering excitation light at 395 nm from a freshly prepared aqueous suspension of non-dairy creamer. The instrument response was typically 40 ps fwhm. Protein concentrations were adjusted to obtain absorbance values in the range from 0.25 to 0.30 at 395 nm (10 mm path).

TCSPC transients were analyzed by a convolute-and-compare nonlinear least-squares program implemented in the multi-platform analysis program Igor (Version 4, Wavemetrics Inc., www.wavemetrics.com). In the analysis, this scattered light transient is convoluted with the fluorescence decay model function, comprised of a weighted sum of up to four exponential components for the rotation-independent data set ($I_{VM}$). Analysis of the anisotropies followed the global fitting procedure first outlined by A. J. Cross and G. R. Fleming (28). Rather than construct the calculated anisotropy, $r(t)$, the three transients ($I_{VM}$, $I_{VV}$, and $I_{VH}$) were fitted to the equations:

$$I_{VM} = K(t) \; ; \; \; \; I_{VV} = \frac{1}{3} K(t) \left[ 1 + 2 \, r(t) \right] \; ; \; \; I_{VH} = \frac{1}{3} K(t) \left[ 1 - r(t) \right]$$

(1)
\[ K(t) = \sum_i A_i \exp \left( -\frac{t}{\tau_{\text{fluor},i}} \right) \quad \text{and} \quad r(t) = 0.4 \sum_i r_i \exp \left( -\frac{t}{\tau_{\text{rot},i}} \right) \]  

(2)

where the fluorescence decay law is given by \( K(t) \), the orientation time correlation function is given by \( r(t) \), the amplitudes are \( A_{\text{fluor},i} \) and the \( r_i \) sum to unity.
RESULTS

Characterization of Novel Fluorescent Proteins

Two GFP-like proteins, mcCFP and mmGFP, derived from *Montastrea cavernosa* and *Meandrina meandrites*, respectively, were cloned using RACE PCR strategies and spectroscopically characterized. The names for these new GFP-like proteins follow the nomenclature of other non-bioluminescent fluorescent proteins (1, 2) and the widely accepted *Aequorea victoria* GFP (*av*GFP). A BLAST protein sequence homology search revealed that mcCFP has 31% identity to *av*GFP and 51% identity to cFP 484 from *Clavularia* sp.; mmGFP has 34% identity to *av*GFP and 55% identity to DsRed (drFP583) from *Discosoma* sp. These results indicate mcCFP and mmGFP belong to the GFP protein family (1, 2). Based on sequence homology, the predicted fluorescence chromophore is Gln-62, Tyr-63, Gly-64 in mcCFP and Thr-60, Tyr-61, Gly-62 in mmGFP, compared with Ser-65, Tyr-66, Gly-67 in *av*GFP.

Three-dimensional structural models derived by homology (SWISSMODEL) show that both mcCFP and mmGFP polypeptides can readily adopt the β-barrel structures of *av*GFP (PDB # 1EMB) and DsRed (PDB # 1G7K), with 11 β-strands and two cap regions. Key secondary structural elements in DsRed are also observed in mmGFP. Two tryptophan residues, corresponding to DsRed positions 94 and 145, are found in the β-sheet of mmGFP at positions 87 and 137. Residues in the vicinity of the *av*GFP chromophore such as R\(^{69}\), E\(^{150}\), R\(^{96}\) and E\(^{222}\) are also found in both new GFP-like proteins. In the immediate the vicinity of the chromophore, T\(^{203}\) in *av*GFP and S\(^{196}\) in DsRed is replaced by H\(^{191}\) in mmGFP.

Based on the results from size-exclusion chromatography (SEC) and native PAGE, mcCFP is a trimer with an approximate monomeric molecular weight of 26 kDa (theoretically 25.8 kDa), and mmGFP is a monomer of approximately 31 kDa (expected 29.2 kDa).
trimeric and monomeric nature of mcCFP and mmGFP is further supported by TCSPC anisotropies (see Figure 5 and discussion below).

The excitation and emission spectra in vitro are shown in Figure 1. The fluorescence emission spectra were deconvoluted in the wavenumber domain with Gaussian least square fitting. The emission spectrum of mcCFP reveals a peak in the blue-green at 474 nm with a clear shoulder at 497 nm and a smaller component at 524 nm. The 474 component is blue shifted by 9 nm relative to other cyan-FPs reported to date (2). The mcCFP excitation maximum is at 431 nm with components at 388 and 453 nm. The measured fluorescence lifetime is 2.66 ± 0.01 ns at all emission wavelengths, which is consistent with excitation of a single chromophore.

The fluorescence spectrum of mm GFP is characterized by a sharp primary peak at 504 nm with shoulders at 521 and 539 nm. This peak and the shoulders correspond to features of the mcCFP emission at essentially the same wavelengths. The mmGFP emission can be populated by excitation from two bands. The broad main band is centered at 400 nm that has components centered at 383 and 411 nm. There is a small component at 494 nm.

Based on the fluorescence lifetimes, the quantum yields of fluorescence are estimated 0.385 for mcCFP and 0.645 for mmGFP. In avGFP the calculated quantum yield is 0.8 based on a maximum extinction coefficient of 27600 M⁻¹cm⁻¹.(29)

**Figure 1 here**

**Folding, Maturation and Kinetics of Wild Type and Mutant mmGFP**

We introduced random mutation into mmGFP by using *E. coli* strain ES1578. Two point mutations were obtained and one was further selected based on visual inspection of fluorescence intensity. The mutant, mmGFP<sub>S110N</sub>, has a single base mutation AGC to AAC, which changes
serine110 to asparagine (S110N). The second was a silent mutation at the amino acid position 56 (TCG to TCA). The S110N mutation is remote from the chromophore and predicted to reside in a hairpin loop of the beta-barrel protein (Figure 2). In vivo, the maturation time of mmGFP\textsubscript{S110N} is ~50% shorter: 20-24 h for the wild type, compared to ca. 12 h for the mutant at 25\textdegree C. Purification and characterization of the proteins revealed that per unit mass, fluorescence of the S110N mutant was about 30% increased relative to wild type.

**Figure 2 here**

We studied denaturation in GdnHCl and refolding by dilution into refolding buffer for both wild type and the S110N mutant by measuring the fluorescence intensities. The fluorescence intensity for different concentrations of GdnHCl is shown in Figure 3A. The refolding kinetics for both the wild type and mutant proteins can be fit to two-exponential rate models. Refolding efficiencies were determined by the recovery of fluorescence after 6M GdnHCl denaturation, observed at 20 min after refolding was initiated by dilution. The values obtained were 67% for wild type and 81% for S110N mutant (Figure 3B).

**Figure 3 here**

We further tracked both wild type and mutant mmGFP maturation in vitro using time dependent fluorescence measurements as function of oxygen, pH and temperature. When the proteins were incubated in air or nitrogen, there was no significant difference in the rate of change of fluorescence intensity between the wild type and mutant proteins during the folding process (Figure 4A). The pH dependence of maturation was examined from pH 3 to 12. Folding
did not occur below pH 4 or above pH 11. The half-life for maturation at pH 9.5 for wild type mmGFP was 3.94 hours compare to 3.48 hours for the mutant. The fastest maturation rates for the wild type and mutant were at pH 9 and 9.5, respectively (Figure 4B). Temperature-dependence of maturation was tested from 3 to 48 °C and no folding was obtained above 40 °C. The optimum folding temperature for both wild type and mutant mmGFPs was 23 °C. The same optimum maturation temperature was confirmed \textit{in vivo} with \textit{E. coli} cultures (data not shown). Analysis of the temperature-dependence of folding yields apparent activation energies of 16.6 kcal/mole and 18.8 kcal/mole for the wild type and mutant, respectively (Figure 4C).

\textbf{Figure 4 here}

\textit{Time-Resolved Emission Spectroscopy for Wild Type and Mutant mmGFP}

Time-resolved emission spectroscopy was employed to characterize the wild type and S110N mmGFPs. The vertical excitation/magic-angle emission fluorescence decay is devoid of orientation information, and excellent fits were obtained for a double-exponential model function (Figure 5). The transients showed a nearly single-exponential decay at the peak of the emission spectrum with a dominant lifetime of 3.1 ns and carrying > 95 % of the peak emission amplitude. In an unconstrained global fit to the fluorescence lifetime and anisotropy functions using the equations given in (1) and (2), precisely the same amplitude and lifetime parameters resulted (Table 1). The faster lifetime component of about 600 ps with 5 % amplitude has not been previously reported to our knowledge. All of the fit parameters obtained were identical within error for both mmGFP and mmGFP_{S110N}.
In the global analysis, the orientation time correlation function is the best fit to a single exponential decay with a 7.5 ns time constant. Significant amplitude remains after this decay, indicating the presence of a very long-lived component in the orientational diffusion dynamics. The 7.5 ns time constant is consistent only with the reorientation time scale for a GFP monomer, and the large value of the residual anisotropy $r_\infty$ (equivalent to a steady-state fluorescence anisotropy value) indicates that dimers, trimers, or tetramers are not present, but rather that some fraction of the mmGFP aggregates into much larger structures.

**Figure 5 here**

**Figure 6 here**

**Table 1 here**
DISCUSSION

Color Diversity of Fluorescent Proteins in Zooxanthellate Corals

There is a wide variety of colored fluorescent proteins in zooxanthellate corals; thirteen GFP-like proteins have been cloned or identified in *M. cavernosa* alone (Matz et al. and our present study). All GFP-like proteins can be broadly classified into four-color types: green, yellow, orange-red, and non-fluorescent purple-blue. The green-type is further divided into cyan, with maximum emission at 483-486 nm, and truly green with emission at 499-518 nm. mcCFP is the most blue-shifted GFP-like protein found in nature. The biological function of these proteins remains unknown. Although they have high UV cross-sections, their role as photoprotective pigments has not been clearly demonstrated. Our results provide the first biophysical characterization a small subset of these proteins.

Spectral Characterization of Fluorescent Proteins

In nature, the emission spectra of whole *M. cavernosa* coral exhibits peaks at 480, 510 and 580 nm (the latter with a shoulder at 630 nm) when excited at 432 nm. A neighbor-joining phylogenetic tree was developed to examine the relationship between the proteins responsible for these emission signals (Figure 7). Based on the predicted 3D structural model, the first of the three amino acids forming the chromophore is different among these mcFPs and thus is the most important for the variability in color. It has been hypothesized that the color difference among different fluorescent proteins should be comparable within a species, but not for different species. From the nucleotide sequence analysis, the ancestral protein is derived from a gene encoding a green fluorescent chromophore. Indeed, green is the predominant color for most of the fluorescent corals (31). Some corals, e.g. *Discosoma* sp., have not retained the original green color. However the derived fluorescent proteins (e.g. RFP) can be easily converted back to
green in vitro (2, 30). It remains to be elucidated what environmental conditions select for the mutations or their biological functions. Since color-converted mutants are characterized by a shift from longer to shorter wavelengths (30), i.e. from red to green, reverse color mutations should also be possible.

Figure 7 here

Table 2 lists the absorption and emission maxima and Stokes shifts for several GFPs, including mcCFP and mmGFP. The same Stokes shift is observed for both the native and mutant mmGFP samples, and is very similar to that reported for GFP from A. victoria. The mcCFP spectral properties are most similar to those observed for Discosoma striata DsRed (dsFP483). Large Stokes shifts such as observed for GFPs are often assigned to dual emissive states, with the lower energy bathochromatic emission arising from a proton transfer in the excited state. Such a state is presumed not to be operative in the DsRed and mmCFP proteins because of the substantially smaller Stokes shifts. The variability in emission spectra and Stokes shifts suggests that fluorescence signatures can be “tailored” by mutation within coral species, yet there is no obvious correspondence between the spectral signature and patterns of distribution.

Table 2 here

Maturation and Folding of Fluorescent Proteins

The exploitation of colored fluorescent proteins in biotechnological applications has led to closer examination of factors controlling the expression of the genes in heterologous systems,
fluorescence spectra and yields, as well as quaternary structures. The slow maturation of GFP-like fluorescent proteins is a major problem for heterologous expression (6). Folding requires three steps. The first is formation of the β-barrel. Chromophore formation is a sequential process, which includes two kinetic steps: cyclization followed by oxidation. Once folding is complete, the tripeptide chromophore motif is buried in the central helix of GFP. Comparing refolding kinetics de novo (Figure 3B), a lag occurred for the folding of mmGFP for samples that were directly purified from inclusion bodies. It is believed that oxidation is the rate-limiting step in chromophore formation (20), whereas the β-barrel formation for mmGFP is the fast step. However, our observation of increased refolding rate in the S110N mutant suggests that torsional flexibility within the hairpins connecting the β-sheets is another potentially important factor limiting protein folding. Mutations at such positions are readily accessible.

Folding pathways could differ depending on the nature of the starting material. Proteins from inclusion bodies may require steps such as proline isomerizations to reach the native conformations, and such steps may not be necessary if mature GFP is used for denaturation-renaturation reactions (20). Renaturation kinetic profiles from different states of denatured mmGFP varied. Refolding did not occur for the mmGFP and S110N mutant after 48 hours of 6M GdnHCl incubation. A likely explanation is that after complete denaturation, the protein is more vulnerable to misfolding and aggregation. The renaturation samples exposed to GdnHCl for 30 min should not have required the oxidation step because the Thr-Tyr-Gly tripeptide chromophore in mmGFP remained intact.

The optimum maturation temperature for both wild type and mutant mmGFP ranges from 22 to 24°C. No maturation occurred above 45°C. The folding of mmGFP can occur when samples are held at –20°C and even at –80°C. The half-life maturation rate below 0°C remains
unknown. Modest increases in temperature can profoundly decrease maturation efficiency of GFPs (18). Based on the decay of the far-UV circular dichroism signal, the molecule lost secondary and tertiary structure when the temperature was increased (22). If the temperature is increased (above 37°C), mmGFP irreversibly aggregates into oligomers.

**Kinetics and Spectroscopy of Fluorescent Proteins**

The electronic structure of the chromophore in GFPs has been examined by electroabsorption spectroscopy (32) and femtosecond fluorescence experiments (33) to understand the overlapping absorption and emission spectral bands. The overlapping absorption and emission bands are both assigned to singlet transitions between S₀ to S₁ states, but arising from different ground states- a neutral and an anionic form. The higher energy absorption is assigned to the neutral form of the chromophore, whereas the red-shifted absorption is assigned to the anion, formed by deprotonation of the phenolic proton. Using femtosecond fluorescence, Boxer and co-workers proposed that the higher energy neutral excited state could undergo a rapid excited state proton transfer to form the anion in several picoseconds (33).

Meech et al. studied a model of the GFP chromophore, p-hydroxybenzylidene-imidazolidinone, in alcohol solutions (34-37), which allowed them to explore several key features of the GFP photophysics. First, they suggested that if the chromophore is not in a rigid environment, torsional dynamics between the phenol and imidazole rings cause a rapid relaxation from the excited state via internal conversion from S₁ to S₀, which drastically reduces the fluorescence quantum yield (34). This nonradiative relaxation pathway has a negligible energy barrier at room temperature (35). From these results, they concluded that functional GFPs with strong emission yields must have the chromophore locked into an environment that is rigid on a time scale exceeding the fluorescence lifetime. They proposed that the non-
fluorescent mutants of GFPs may simply have a slightly different protein conformation surrounding the chromophore, permitting a slight degree of flexibility that activates the internal conversion process as the primary excited state relaxation pathway. Their results hold true for both the neutral and phenolate forms of the chromophore (32-37).

The protonation state of the tyrosylhydroxyl group of the chromophore is responsible for the pH sensitivity of GFP based on the data from X-ray crystallography, ultrafast optical spectroscopy, and site-directed mutation (38). Reorientations of hydroxyl dipoles and solvation by buried water in the immediate surroundings of the chromophore are likely responsible for the ultrafast pH-dependent properties of the fluorescent proteins (7, 19, 39-42). The behavior of key residues such as T203 and S65 seems to be important for the properties of avGFP. \( pK_a \) shifts due to chemical modification of aligning residues T203 and S65, correlate the orientation of T203 with the protonation state of the chromophore. This followed the response of the hydrogen bond net to ionization changes, and increased acidity of the chromophore in the excited state (38). The internal proton transfer is coupled to the presence of a buried water near the chromophore; translational movement or exchange of W22 will be additional determinants for the on/off behavior of GFP. E222 acts as proton acceptor in the ground and excited states. H148 may have additional degrees for orientation to the chromophore (38). The S110N mutant in mmGFP is not close to these sites based on alignment of the predicted model. The slope and peak of half-life maturation at pH 6 to 11 were not significantly different between the wild type and mutant, although the \( pK_a \) was shifted higher for the mutant.

**Fluorescence lifetimes**

The mmGFP fluorescence lifetimes are quite similar to the values for both GFP and DsRed. The single-exponential character of \( r(t) \), the longer 7.5 ns fluorescence anisotropy,
indicates that unlike dimeric GFP mutants and tetrameric DsRed, rapid energy transfer does not occur for mmGFP. The 7.5 ns fluorescence anisotropy decay time constant for mmGFP can be assigned to diffusive reorientation of a chromophore rigidly held within the protein interior, and a protein that is clearly a monomer, not a dimer or higher oligomer. For example, the work of Heikal, et al. (24) showed that tetrameric DsRed had a reorientation time constant of 53 ns, and that Citrine, a yellow fluorescent mutant, had a time reorientation constant of 16 ns. The observed 7.5 ns time constant for the fluorescence anisotropy decay is consistent with the observed monomeric volume of other β-barrel fluorescent protein structures.

CONCLUSIONS

The two naturally occurring colored fluorescent proteins we cloned from zooxanthellate corals are a small subset of large protein family that appears to be ubiquitous in symbiotic cnidarians. While the physiological role of these proteins remains unknown, the biophysical and biochemical data presented here suggest a great potential for natural and introduced genetic alteration of spectral variability. For example, dot-plot analysis between mmGFP, avGFP and DsRed (data not shown) reveal high sequence homology between mmGFP to DsRed, and significantly less homology to avGFP. The unique fluorescence properties of mcCFP and the sequence homology mmGFP_{S110N} with DsRed suggest there are alternative strategies for designing variant of GFP-like proteins with different spectra and enhanced fluorescence yields for biotechnology applications.
ACKNOWLEDGEMENTS

We thank the staff of the Caribbean Marine Research Center at Lee Stocking Island for support during field experiments. We thank Kevin Wyman and Drs. Maxim Gorbunov and Yoram Gerchman for technical support and Drs. Michail Matz and Peter C. Kahn for discussions. This work was supported by the Office of Naval Research through grant N00014-00-1-0795 to PGF.

REFERENCES


FIGURE LEGENDS

Figure 1. Deconvolution analysis of the steady-state fluorescence excitation and emission spectra of mcCFP and mmGFP. (A) Fluorescence excitation spectra of mcCFP and (B) mmGFP, and (C) emission spectra of mcCFP and (D) mmGFP. Excitation spectra were recorded at 398 nm for mmGFP and 432 nm for mcCFP. The maximum emission spectra for mmGFP and mcCFP were recorded at 505 nm and 477 nm, respectively.

Figure 2. *Meandrina meandrites* GFP homology model based on alignment with 1GGX and 1G7K using Swiss-PdbViewer 3.70b2 ([http://www.expasy.ch/spdbv/](http://www.expasy.ch/spdbv/)). The homology model was initially made omitting chromophore residues in the mmGFP sequence and in 1GGX, 1G7K models. The chromophore and neighboring residues were patched in by hand afterwards. The figure was prepared using PyMOL version 0.82 (www.pymol.org). (submitted as an electronic figure)

Figure 3.
A. GdnHCl-induced unfolding transition curves of wild type and S110N mutant mmGFP. Fluorescence was measured 30 min after dilution into GdnHCl with excitation at 395nm, emission at 505nm, pH 8.0, 21°C. The protein concentration was 2.5 µg/ml.

B. Fluorescence recovery of wild type and mutant mmGFP. 30 min after denaturation in 6 M GdnHCl, renaturation was initiated by 100-fold dilution into refolding buffer: 25 mM Tris, 150 mM NaCl, pH 8.0, 23°C, final protein concentration 2.5µg/mL.
A. Time dependent fluorescence under N₂ and O₂ incubation: fluorescence was measured with excitation at 395nm and emission at 505nm in refolding buffer: 25 mM Tris, 150 mM NaCl, pH 8.0, 23°C, 30 µg/ml protein.

B. pH-dependent half-life maturation. Fluorescence was measured with excitation at 395nm, emission at 505nm, 23°C, and a final protein concentration of 15mg/mL.

C. Kinetics of temperature dependence fluorescence for wild type and mutant mmGFP. Fitted plot of 1/temperature (K). Fluorescence was recorded using a fluorescence plate reader in refolding buffer: 25 mM Tris, 150 mM NaCl, pH 8.0 at a final protein concentration of 15µg/ml.

Figure 5. TCSPC emission decays for mmGFP (top) and mmGFP-S110N (bottom). Laser excitation was at 395 nm, and emission was detected at the magic angle (VM) at 505 nm. Raw data are shown in blue, with the convolute-and-compare nonlinear least-squares fit shown in red. Instrument response function measured with a scattering solution is shown in black. Weighted fit residuals are shown at the top, indicating the excellent quality of the fit.

Figure 6. Fluorescence polarization anisotropies for mmGFP (top) and mmGFP-S110N (bottom). Raw data shown in blue; the best fit to a 2-exponential lifetime decay and single-exponential reorientation time correlation function is shown in red. Instrument temporal response shown in black, and was measured using a scattering solution. Convolute-and-compare analysis was done using a global fit to equations (1) and (2).

Figure 7. The neighbor-joining phylogenetic tree was constructed by Phylip 3.6. All 13 mcFPs were grouped into three clades: red, cyan and green, and green. The amino acids sequences
derived from *Montastrea cavernosa* are from NCBI GenBank (shown by accession number). Numbers at nodes denote the quartet-puzzling support values (1,000 puzzling attempts). (Bar = 0.1 replacements/sites). * Cloned from red fluorescent tissue.
Figure 1

A

Ex mcCFP

B

Ex mmGFP

C

Em mcCFP

D

Em mmGFP
Figure 2
Figure 3

A

B
Figure 4

A

![Graph showing relative fluorescence intensity over time for different conditions.]

B

![Graph showing pH vs. 1/Half time for wild type and mutant.]

C

![Graph showing Ln k vs. 1/Temperature for wild type and mutant.]

32
Figure 7

[Genetic tree diagram with labeled nodes and branch lengths.]
Table 1

Fluorescence lifetime and anisotropy parameters for analysis of TCSPC data shown in Figures 5 and 6

<table>
<thead>
<tr>
<th>Sample</th>
<th>$A_1$</th>
<th>$\tau_1$ (ns)</th>
<th>$A_2$</th>
<th>$\tau_2$ (ns)</th>
<th>$\tau_{rot}$ (ns)</th>
<th>$r_\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>mmGFP</td>
<td>0.955</td>
<td>3.09</td>
<td>0.045</td>
<td>0.657</td>
<td>7.52</td>
<td>0.296</td>
</tr>
<tr>
<td>S109N</td>
<td>0.958</td>
<td>3.10</td>
<td>0.042</td>
<td>0.586</td>
<td>7.52</td>
<td>0.234</td>
</tr>
</tbody>
</table>
Table 2

Peak wavelengths for absorption and emission are given for seven previous literature values (1) along with calculate Stokes shift.

<table>
<thead>
<tr>
<th>Source of FP</th>
<th>$\lambda_{\text{max}}$ (nm)</th>
<th>$\lambda_{\text{max}}$ (nm)</th>
<th>Stokes shift $\Delta\lambda$ (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absorption</td>
<td>Emission</td>
<td></td>
</tr>
<tr>
<td><em>Anemonia majano</em></td>
<td>amFP486</td>
<td>458</td>
<td>486</td>
</tr>
<tr>
<td><em>Zoanthus sp.</em></td>
<td>zFP506</td>
<td>496</td>
<td>506</td>
</tr>
<tr>
<td></td>
<td>zFP538</td>
<td>528</td>
<td>538</td>
</tr>
<tr>
<td><em>Discosoma striata</em></td>
<td>dsFP483</td>
<td>443</td>
<td>483</td>
</tr>
<tr>
<td><em>Discosoma sp., “red”</em></td>
<td>drFP583</td>
<td>558</td>
<td>583</td>
</tr>
<tr>
<td><em>Clavularia sp.</em></td>
<td>cFP484</td>
<td>456</td>
<td>484</td>
</tr>
<tr>
<td><em>Aequorea victoria</em></td>
<td>GFP</td>
<td>397</td>
<td>509</td>
</tr>
<tr>
<td><em>Montastrea cavernosa</em></td>
<td>mcCFP (our data)</td>
<td>432</td>
<td>477</td>
</tr>
<tr>
<td><em>Meandrina meandrites</em></td>
<td>mmGFP (our data)</td>
<td>398</td>
<td>505</td>
</tr>
</tbody>
</table>