Biophysical characterization of natural and mutant fluorescent proteins cloned from zooxanthellate corals ☆,☆☆,★

Yi Sun[a,*,1], Edward W. Castner Jr.[b], Catherine L. Lawson[b], Paul G. Falkowski[a,*]

[a] *Environmental Biophysics and Molecular Ecology Program, Institute of Marine and Coastal Sciences, Rutgers University, 71 Dudley Road, New Brunswick, NJ 08901, USA*
[b] *Department of Chemistry and Chemical Biology, Rutgers University, 610 Taylor Road, Piscataway, NJ 08854, USA*

Received 18 May 2004; revised 9 June 2004; accepted 9 June 2004

Available online 2 July 2004

Edited by Richard Cogdell

Abstract 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. The spectroscopic properties of a green fluorescent protein cloned from *Meandrina meandrites* (mmGFP) resemble the commercially available GFP derived originally from the hydrozoan *Aequorea victoria* (avGFP). mmGFP is a monomeric protein with an excitation maximum at 398 nm and an emission maximum at 505 nm, a fluorescence lifetime of 3.10 ± 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.

Keywords: Green fluorescent proteins; Corals; Fluorescence lifetimes; Time-correlated single photon counting; Cnidarians

1. 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 [3]. 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. We present here 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 seven non-fluorescent chromoproteins have been isolated and spectroscopically characterized from Anthozoa [1,2,4,5]. 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 that all the GFPs are derived from a single common ancestor, but have diverged significantly within the phylum cnidaria.

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. The kinetics of protein folding and chromophore formation can be followed using fluorescence [6–9]. The excited state lifetimes for mature fluorescent proteins typically range from 3 to 4 ns. Their fluorescent yield depends on pH and temperature [10], 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 *A. victoria* have been carried out on mutants with characteristics more conducive to biotechnological applications, including enhanced fluorescence intensity [11,12], shifted emission wavelength [6,13,14], altered sensitivity to pH [15], and faster protein folding [16–18]. Recent studies have focused on altering the fluorescent and

☆ This work was supported by the Office of Naval Research (Award #: N00014-00-1-0795).
☆☆ The sequences reported in this paper have been deposited in the GenBank database with Accession Nos: AY056460 (mcCFP); AY155343 (mmGFP); AY155344 (mmGFP $_{S110N}$); AY362545 (mcRFP).
★ Positions of amino-acid residues mentioned in this paper correspond to *Aequorea victoria* GFP (Accession No. P42212).

* Corresponding authors. Fax: +1-732-932-4083.
E-mail address: falko@imcs.rutgers.edu (P.G. Falkowski).

[1] Present address: Waksman Institute, Rutgers University, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA.

Abbreviations: GFP, green fluorescent protein; DsRed, red fluorescent protein; RACE, rapid amplification cDNA ends; PCR, polymerase chain reaction; GdnHCl, guanidinium chloride; PAGE, polyacrylamide gel electrophoresis; TCSPC, time-correlated single photon counting

0014-5793/$22.00 © 2004 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/j.febslet.2004.06.043

biochemical properties of red fluorescent protein drFP583 (DsRed), including increasing the rate of "maturation" [19,20] and reducing oligomerization [21]. 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*, a GFP-like protein (mmGFP) from *Meandrina meandrites*, and a mutant derived from the latter.

2. Materials and methods

2.1. Cloning, expression and mutagenesis of target proteins

Samples of the stony corals, *M. cavernosa* and *M. 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 rapid amplification cDNA ends-polymerase chain reaction (RACE-PCR) as described by Matz et al. [1]. Gene-specific primers were designed as follows:

for mcCFP: 5′-GCGTCTTCTTCTGCATAACTGGACCACTG-GAGG-3′;

for mmGFP: 5′-TGGATTACAGGTCCATTGGCGGGAAA-GT-3′.

The anti-sense primer sequences came from Matz et al. [22]. The restriction endonuclease sites (*Bst*BI and *Bgl*II for mcCFP; *Hin*dIII and *Bst*BI 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× His-tag at the C-terminus. The recombinant plasmid was transfected into an *Escherichia 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 (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 (350 nm) over three generations. This process recovered a single mutant $mmGFP_{S110N}$. Both wild type and mutant mmGFPs with C-terminal His-tags were expressed in *E. coli* BL21 (DE3) (Novagen) in LB medium.

2.2. 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 with 100 μg/ml lysozyme. NaCl was added to the lysate to a final concentration of 100 mM. The preparation was cleared by centrifugation at $10\,000 \times g$ 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–nitrilotriacetic 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 guanidinium chloride (GdnHCl), refolding was induced by a 100-fold dilution of the total denatured protein (0.1 mg/ml) into ice-cold refolding buffer (25 mM Tris–HCl and 150 mM NaCl, pH 8.0). The development of fluorescence was monitored for the next 15 min.

2.3. 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 (295 K), pH 8.0. The O_2 concentrations were kept as 260 μM or <1 μM by bubbling of compressed air or pure N_2, 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 h 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 s 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).

2.4. Time-resolved emission spectroscopy

Details of the time-correlated single photon counting (TCSPC) instrument have been described previously [23–25]. 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° 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 non-linear 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 Cross and Fleming [26]. Rather than construct the calculated anisotropy, $r(t)$, and the three transients (I_{VM}, I_{VV} and I_{VH}) were fitted to the equations:

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

$$K(t) = \sum_i a_i \exp\left(-\frac{t}{\tau_{\text{fluor},i}}\right) \quad \text{and} \quad r(t) = 0.4 \sum_i r_i \exp\left(-\frac{t}{\tau_{\text{rot},i}}\right), \tag{2}$$

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.

3. Results

3.1. Characterization of novel fluorescent proteins

Two GFP-like proteins, mcCFP and mmGFP, derived from *M. cavernosa* and *M. 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 *A. victoria* GFP (avGFP) [27]. A BLAST protein sequence homology search revealed that mcCFP has 31% identity to avGFP and 51% identity to cFP 484 from *Clavularia* sp.; mmGFP has 34% identity to avGFP 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 the 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 avGFP.

Three-dimensional structural models derived by homology (SWISSMODEL) show that both mcCFP and mmGFP polypeptides can readily adopt the β-barrel structures of avGFP (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 avGFP chromophore, such as R^{69}, E^{150}, R^{96} and E^{222}, are also found in both new GFP-like proteins. In the immediate vicinity of the chromophore, T^{203} in avGFP and S^{196} in DsRed are replaced by H^{191} in mmGFP.

Based on the results from size-exclusion chromatography and native polyacrylamide gel electrophoresis (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). The mmGFP wild type and S110N mutant flourescence anisotropies may be consistent with the existence of at least two species, including monomer and oligomer of trimer or larger size (see Fig. 5 and discussion below).

The excitation and emission spectra in vitro are shown in Fig. 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 mmGFP is characterized by a sharp primary peak at 504 nm with a secondary maximum at

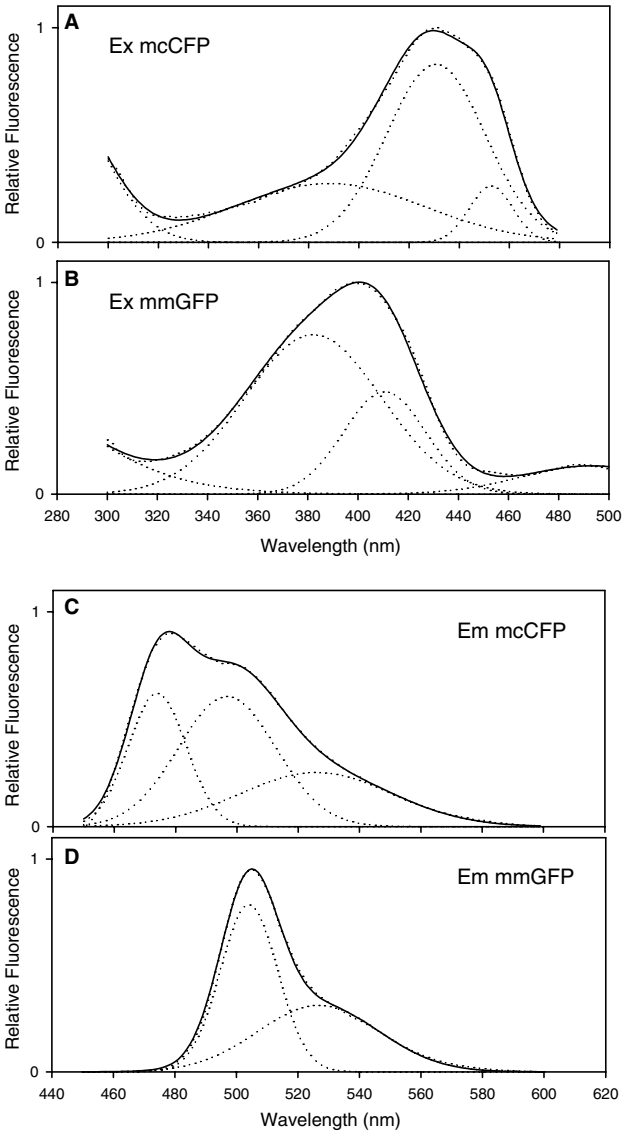

Fig. 1. Deconvolution analysis of the steady-state fluorescence excitation and emission spectra of mcCFP and mmGFP. Solid lines are for the measured excitation and emission spectra; dotted lines denote the Gaussian spectral components. (A) Fluorescence excitation spectra of mcCFP and (B) mmGFP, and (C) 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 and 477 nm, respectively.

530 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 to be 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 27 600 $M^{-1}\,cm^{-1}$ [29].

3.2. Folding, maturation and kinetics of wild type and mutant mmGFP

We introduced random mutations into mmGFP using *E. coli* strain ES1578. Two point mutations were obtained and one

was selected for further study based on visual inspection of fluorescence intensity. The mutant, $mmGFP_{S110N}$, 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 is predicted to reside in a hairpin loop of the β-barrel protein (Fig. 2). In vivo, the maturation time of $mmGFP_{S110N}$ is ~50% shorter: 20–24 h for the wild type, compared to ca. 12 h for the mutant at 25 °C. Purification and characterization of the proteins revealed that per unit mass, fluorescence of the S110N mutant was about 30% higher relative to that of the wild type.

All fluorescence measurements used an excitation wavelength of 395 nm. Light at this wavelength can have an actinic effect that promotes chromophore maturation and is manifested by an increase in the absorption cross-section around 488 nm [28]. In our experimental protocol, however, the minor peak of wild type and mutant mmGFP did not change during the observational period because the excitation beam was attenuated to reduce actinic effects.

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 Fig. 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 6 M 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 (Fig. 3B).

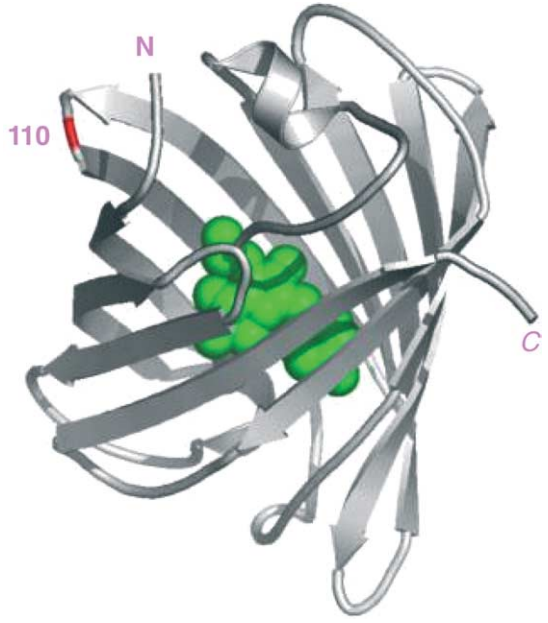

Fig. 2. *Meandrina meandrites* GFP homology model based on alignment with 1GGX and 1G7K using Swiss-PdbViewer 3.70b2 (http://www.expasy.ch/spdbv/). The homology model was initially made omitting chromophore residues in the mmGFP sequence and in 1GGX, 1G7K models. The chromophore and neighboring residues were patched in by hand afterwards. The figure was prepared using PyMOL version 0.82 (www.pymol.org). The position of residue 110 is indicated.

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 (Fig. 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 h compared with 3.48 h for the mutant. The fastest maturation rates for the wild type and mutant were at pH 9 and 9.5, respectively (Fig. 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 in vivo with *E. coli* cultures (data not shown). Analysis of the temperature-dependence of folding yields apparent activation energies of 16.6 and 18.8 kcal/mol for the wild type and mutant proteins, respectively (Fig. 4C).

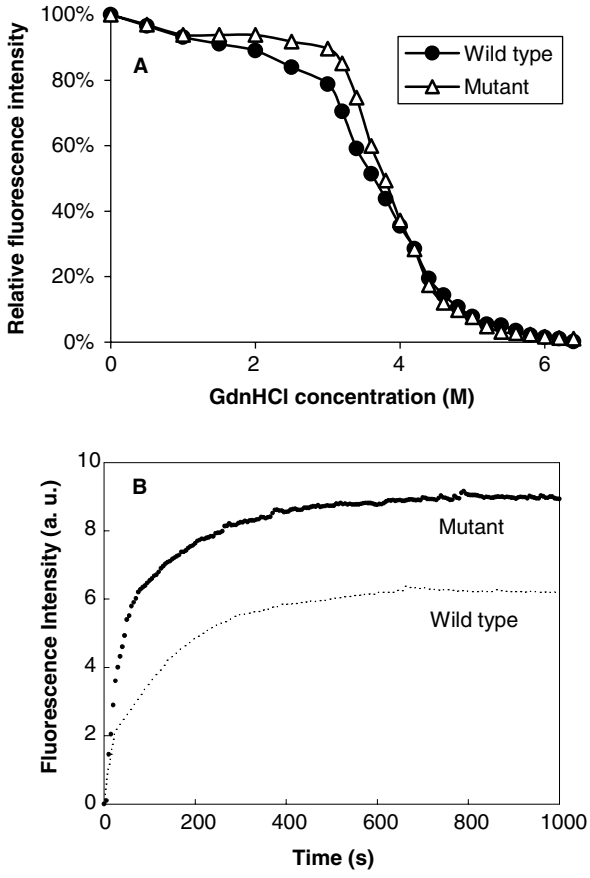

Fig. 3. (A) GdnHCl-induced unfolding transition curves of wild type and S110N mutant mmGFP. Fluorescence was measured 30 min after dilution into GdnHCl with excitation at 395 nm and emission at 505 nm, pH 8.0, 21 °C. The protein concentration was 2.5 μg/ml. (B) Fluorescence recovery of wild type and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

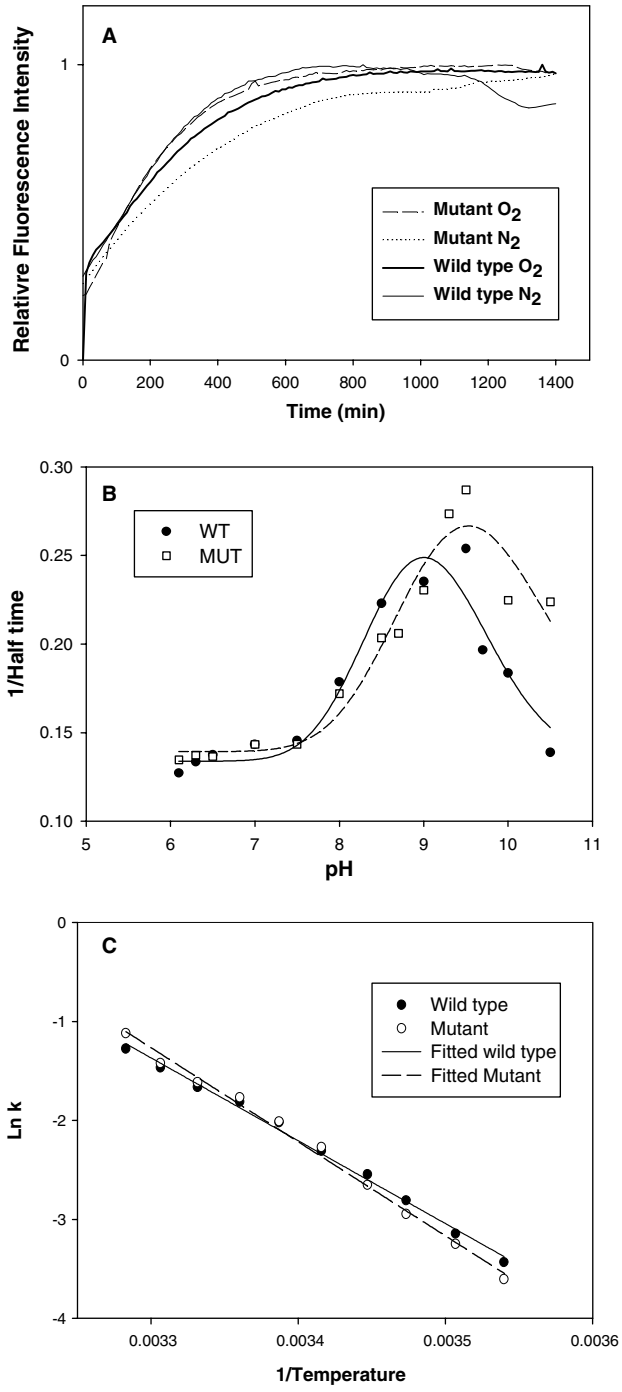

Fig. 4. (A) Time-dependent fluorescence under N_2 and O_2 incubation: fluorescence was measured with excitation at 395 nm and emission at 505 nm in refolding buffer: 25 mM Tris, 150 mM NaCl, pH 8.0, 23 °C and 30 μg/ml protein. (B) pH-dependent half-life maturation. Fluorescence was measured with excitation at 395 nm, emission at 505 nm, 23 °C, and a final protein concentration of 15 mg/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 and 150 mM NaCl, pH 8.0, at a final protein concentration of 15 μg/ml.

3.3. 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 (Fig. 5). The amplitudes and lifetimes are given in Table 1.

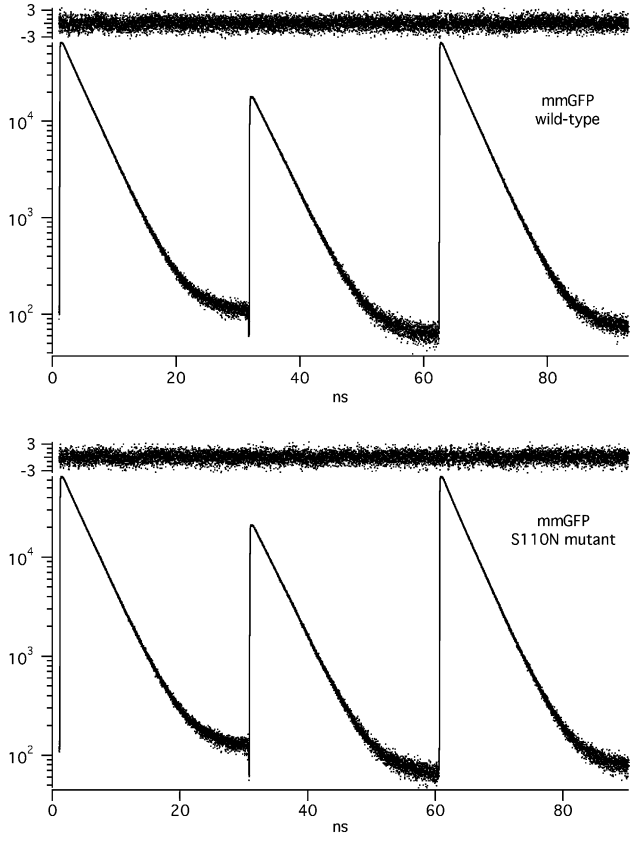

Fig. 5. Fluorescence polarization anisotropies for mmGFP (top) and mmGFP-S110N (bottom). Vertically polarized laser excitation was at 395 nm and emission was detected at 505 nm for three polarization angles: magic angle (VM), horizontal (VH), and vertical (VV), graphed from left to right, respectively. Convolute-and-compare analysis was done using a global fit to Eqs. (1) and (2).

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 that was the same as both wild type and the S110N mutant. To our knowledge, the minor (<5%) lifetime component of 709 and 624 ps for wild type and S110N mutant has not been previously reported.

In an unconstrained global fit to the fluorescence lifetime and anisotropy functions using the equations given in (1) and (2), best fits are obtained with a double-exponential anisotropy function. The dominant components of the anisotropy decays are much longer than the fluorescence lifetime, so that the estimated error is >25%. These long anisotropy decay components are consistent with trimer or larger oligomers. The short components of the fluorescence anisotropy have values of 3.9 and 4.7 ns for the wild type and S110N mutants, with relative amplitudes of 12.1% and 23.1%, respectively. These values do not allow an assignment of monomers with certainty; the fact that they are different to within error, and have amplitudes that differ by a factor of 2 also calls into question whether the anisotropy is

Table 1
Fluorescence lifetime and anisotropy decay parameters

	a_1 (%)	τ_1 (ns)	a_2 (%)	τ_2 (ns)	r_1 (%)	τ_{r1} (ns)	r_{r2} (%)	τ_{r2} (ns)	χ^2
MmGFP (wild type)	95.24	3.10	4.76	0.709	87.9	103	12.1	3.9	1.059
MmGFP (S110N)	95.49	3.11	4.41	0.624	76.9	68.6	23.1	4.7	1.061

derived from monomer reorientation or inefficient Förster energy transfer between protein subunits in trimeric or higher order oligomers.

4. Discussion

4.1. 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 so far in nature. The biological function of these proteins remains unknown. Although they have large UV cross-sections, their role as photoprotective pigments has not been clearly demonstrated. Our results provide the first biophysical characterization of a small subset of these proteins.

4.2. Spectral characterization of fluorescent proteins

In nature, the emission spectra of whole *M. cavernosa* coral exhibit peaks at 480, 510 and 580 nm (the latter with a shoulder at 630 nm) when excited at 432 nm. An un-rooted neighbor-joining phylogenetic tree was constructed to examine the relationship between the proteins responsible for these emission signals (Fig. 6). 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. 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 [29], i.e., from red to green, reverse color mutations should also be possible.

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 those observed for GFPs, are often assigned to dual emissive states, with the lower energy bathochromic 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.

4.3. 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 [12]. 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 (Fig. 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 [7], 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 [7]. Renaturation kinetic profiles from different states of denatured mmGFP varied. Refolding did not occur for the mmGFP and S110N mutant after 48 h of 6 M 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 [30]. Based on the decay of the far-UV circular dichroism signal, the molecule lost secondary and tertiary structure when the temperature was increased [9]. If the temperature is increased above 37 °C, mmGFP irreversibly aggregates into oligomers.

4.4. Kinetics and spectroscopy of fluorescent proteins

The electronic structure of the chromophore in GFPs has been examined by electroabsorption spectroscopy [31] and femtosecond fluorescence experiments [32] to understand the

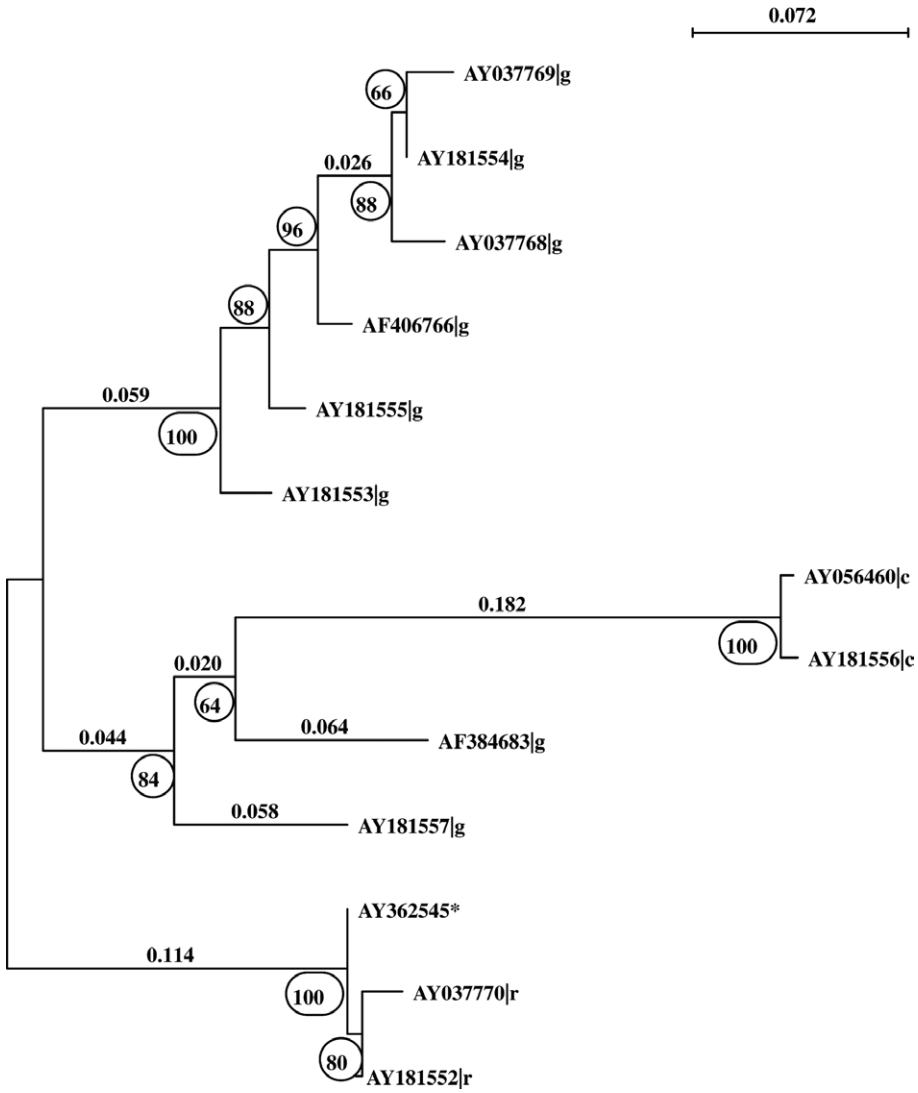

Fig. 6. The neighbor-joining phylogenetic tree was constructed by Phylo_win. The amino acids sequences derived from *M. cavernosa* are from NCBI GenBank (shown by accession number) and were aligned using Phylip 3.6. Genetic distance on branch length was provided and circled numbers indicate the number of times a branch appeared out of 100 bootstrap replicates. All 13 mcFPs were grouped into two clades: red and green. *Cloned from red fluorescent tissue.

Table 2
Peak wavelengths for absorption and emission are given for seven previous literature values (1) along with calculate Stokes shift

Source of FP		λ_{max} (nm)		Stokes shift
		Absorption	Emission	$\Delta\lambda$ (nm)
Anemonia majano	amFP486	458	486	28
Zoanthus sp.	zFP506	496	506	10
	zFP538	528	538	10
Discosoma striata	dsFP483	443	483	40
Discosoma sp., ''red''	drFP583	558	583	25
Clavularia sp.	cFP484	456	484	28
Aequorea victoria	GFP	397	509	112
Montastrea cavernosa	mcCFP (our data)	432	477	45
Meandrina meandrites	mmGFP (our data)	398	505	107

overlapping absorption and emission spectral bands. The overlapping absorption and emission bands are both assigned to singlet transitions between S_0 and S_1 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 [32]

proposed that the higher energy neutral excited state could undergo a rapid excited state proton transfer to form the anion in several picoseconds.

Meech et al. [33–36] studied a model of the GFP chromophore, *p*-hydroxybenzylidene-imidazolidinone, in alcohol solutions, 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 S_1 to S_0, which drastically reduces the fluorescence quantum yield [33]. This non-radiative relaxation pathway has a negligible energy barrier at room temperature [34]. 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 [31–36].

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 [37]. 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 [10,13,38–41]. 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 and correlates 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 [37]. 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 [37]. The S110N mutant in mmGFP is not close to these sites based on the 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.

4.5. Fluorescence lifetimes

The mmGFP fluorescence lifetimes are quite similar to the values for both GFP and DsRed. The faster, smaller amplitude components of the fluorescence anisotropies are 3.9 ns (12.1%) and 4.7 ns (23.1%). It is possible that these should be assigned to monomeric protein reorganization, but more likely that these indicate not the presence of monomers, but inefficient energy transfer between oligomers. The work of Meech and co-workers [33–36] conclusively demonstrates that the chromophore must be rigidly bound within the protein, or the fluorescence will be strongly quenched. Hence, it is not plausible to assign the faster reorientational dynamics of 3.9 and 4.7 ns to restricted reorientation of the chromophore within the protein. For example, Heikal et al. [42] reported that tetrameric DsRed had a reorientation time constant of 53 ns, and that Citrine, a yellow fluorescent monomeric protein mutant, had a time reorientation constant of 16 ns. Thus, the dominant amplitude reorientational dynamics of >60 ns for both wild type and mutant *mmGFP* most likely arise from reorientation of chromophores rigidly bound within their respective monomeric units, and comprising an aggregate of three or more protein subunits.

5. Conclusions

The two naturally occurring colored fluorescent proteins we cloned from zooxanthellate corals are a small subset of a 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. The unique fluorescence properties of mcCFP and the sequence homology of $mmGFP_{S110N}$ with DsRed suggest that 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 No. N00014–00–1–0795 to PGF.

References

[1] Matz, M.V., Fradkov, A.F., Labas, Y.A., Savitsky, A.P., Zaraisky, A.G., Markelov, M.L. and Lukyanov, S.A. (1999) Nat. Biotechnol. 17, 969–973.
[2] Labas, Y.A., Gurskaya, N.G., Yanushevich, Y.G., Fradkov, A.F., Lukyanov, S.A. and Matz, M.V. (2002) Proc. Natl. Acad. Sci. USA 99, 4256–4261.
[3] Mazel, C.H., Lesser, M.P., Gorbunov, M.Y., Barry, T.M., Farrel, J.H., Wyman, K.D. and Falkowski, P.G. (2003) Limnol. Oceanogr. 48, 402–411.
[4] Dove, S.G., Hoegh-Guldberg, O. and Ranganathan, S. (2001) Coral Reefs 19, 197–204.
[5] Tu, H., Xiong, Q., Zhen, S., Zhong, X., Peng, L., Chen, H., Jiang, X., Liu, W., Yang, W., Wei, J., Dong, M., Wu, W. and Xu, A. (2003) Biochem. Biophys. Res. Commun. 301, 879–885.
[6] Heim, R., Prasher, D.C. and Tsien, R.Y. (1994) Proc. Natl. Acad. Sci. USA 91, 12501–12504.
[7] Reid, B.G. and Flynn, G.C. (1997) Biochemistry 36, 6786–6791.
[8] Inouye, S. and Tsuji, F.I. (1994) FEBS Lett. 351, 211–214.
[9] Bokman, S.H. and Ward, W.W. (1981) Biochem. Biophys. Res. Commun. 101, 1372–1380.
[10] Kneen, M., Farinas, J., Li, Y. and Verkman, A.S. (1998) Biophys. J. 74, 1591–1599.
[11] Heim, R., Prasher, D.C. and Tsien, R.Y. (1995) Nature 373, 663–664.
[12] Yang, T., Cheng, L. and Kain, S.R. (1996) Nucleic Acids Res. 24, 4592–4594.
[13] Ormo, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y. and Remington, S.J. (1996) Science 273, 1392–1395.
[14] Creemers, T.M.H., Lock, A.J., Subramaniam, V., Jovin, T.M. and Völker, S. (2000) Proc. Natl. Acad. Sci. USA 97, 2974–2978.

[15] Miesenbock, G., De Angelis, D.A. and Rothman, J.E. (1998) Nature 394, 192–195.
[16] Fukuda, H., Arai, M. and Kuwajima, K. (2000) Biochemistry 39, 12025–12032.
[17] Makino, Y., Amada, K., Taguchi, H. and Yoshida, M. (1997) J. Biol. Chem. 272, 12468–12474.
[18] Ward, W.W. and Bokman, S.H. (1982) Biochemistry 19, 4535–4540.
[19] Baird, G.S., Zaccharias, D.A. and Tsien, R.Y. (2000) Proc. Natl. Acad. Sci. USA 97, 11984–11989.
[20] Terskikh, A.V., Fradkov, A.F., Zaraisky, A.G., Kajava, A.V. and Angres, B. (2002) J. Biol. Chem. 277, 7633–7636.
[21] Campbell, R.E., Tour, O., Palmer, A.E., Steinbach, P.A, Baird, G.S., Zacharisa, D.A. and Tsien, R.Y. (2002) Proc. Natl. Acad. Sci. USA 99, 7877–7882.
[22] Matz, M., Shagin, D., Bogdanova, E., Britanova, O., Lukyanov, S., Diatchenko, L. and Chenchik, A. (1999) Nucleic Acids Res. 27, 1558–1560.
[23] Shirota, H. and Castner Jr., E.W. (2000) J. Chem. Phys. 112, 2367.
[24] Castner Jr., E.W., Kennedy, D. and Cave, R.J. (2000) J. Phys. Chem. A 104, 2869.
[25] Frauchiger, L., Shirota, H., Uhrich, K.E. and Castner Jr., E.W. (2002) J. Phys. Chem. B 106, 7463–7468.
[26] Cross, A.J. and Fleming, G.R. (1984) Biophys. J. 46, 45–56.
[27] Ward, W.W. (1998) in: Green Fluorescent Protein: Properties, Applications, and Protocols (Chalfie, M. and Kain, S., Eds.), pp. 45–75, Wiley–Liss, New York.
[28] Patterson, G.H. and Lippincott-Schwartz, J. (2002) Science 297, 1873–1877.
[29] Gurskaya, N.G., Savitsky, A.P., Yanushevich, Y.G., Lukyanov, S.A. and Lukyanov, K.A. (2001) BMC Biochem. 2, 6.
[30] Tsien, R.Y. (1998) Annu. Rev. Biochem. 67, 509–544.
[31] Bublitz, G., King, B.A. and Boxer, S.G. (1998) J. Am. Chem. Soc. 120, 9370–9371.
[32] Chattoraj, M., King, B.A., Bublitz, G.U. and Boxer, S.G. (1996) Proc. Natl. Acad. Sci. 93, 8362–8367.
[33] Webber, N.M., Litvinenko, K.L. and Meech, S.R. (2001) J. Phys. Chem. B 105, 8036–8039.
[34] Litvinenko, K.L., Webber, N.M. and Meech, S.R. (2001) Chem. Phys. Lett. 346, 47–53.
[35] Mandal, D., Tahara, T., Webber, N.M. and Meech, S.R. (2002) Chem. Phys. Lett. 358, 495–501.
[36] Litvinenko, K.L., Webber, N.M. and Meech, S.R. (2003) J. Phys. Chem. A 107, 2616–2623.
[37] Scharnagl, C., Raupp-Kossmann, R. and Fischer, S.F. (1999) Biophys. J. 77, 1839–1857.
[38] Brejc, K., Sixma, T.K., Kitts, P.A., Kain, S.R., Tsien, R.Y., Ormö, M. and Remington, S.J. (1997) Proc. Natl. Acad. Sci. USA 94, 2306–2311.
[39] Palm, G.J., Zdanov, A., Gaitanaris, G.A., Stauber, R., Pavlakis, G.N. and Wlodawer, A. (1997) Nat. Struct. Biol. 4, 361–365.
[40] Elsliger, M.-A., Wachter, R.M., Hanson, G.T., Kallio, K. and Remington, S.J. (1999) Biochemistry 38, 5296–5301.
[41] VanThor, J.J., Pierik, A.J., Nugteren-Roodzant, I., Xie, A. and Hellingwerf, K.J. (1998) Biochemistry 37, 16915–16921.
[42] Heikal, A.A., Hess, S.T., Baird, G.S., Tsien, R.Y. and Webb, W.W. (2000) Proc. Natl. Acad. Sci. USA 97, 11996–12001.