

Identification of a Fluorescent Protein from *Rhacostoma Atlantica*[†]

Michael R. Tota¹, Jeanna M. Allen¹, Jan O. Karolin², Chris D. Geddes² and William W. Ward^{*3}

¹Brighter Ideas, Inc., Metuchen, NJ

²Institute of Fluorescence and Department of Chemistry & Biochemistry, University of Maryland Baltimore County, Baltimore, MD

³Department of Biochemistry and Molecular Biology, Rutgers University, New Brunswick, NJ

Received 21 February 2016, accepted 2 June 2016, DOI: 10.1111/php.12609

ABSTRACT

We have cloned a novel fluorescent protein from the jellyfish *Rhacostoma atlantica*. The closest known related fluorescent protein is the *Phialidium* yellow fluorescent protein, with only a 55% amino acid sequence identity. A somewhat unusual alanine–tyrosine–glycine amino acid sequence forms the presumed chromophore of the novel protein. The protein has an absorption peak at 466 nm and a fluorescence emission peak at 498 nm. The fluorescence quantum yield was measured to be 0.77 and the extinction coefficient is 58 200 M⁻¹ cm⁻¹. Several mutations were identified that shift the absorption peak to about 494 nm and the emission peak to between 512 and 514 nm.

Abbreviations: BLAST, Basic Local Alignment Search Tool; cDNA, complementary deoxyribonucleic acid; CFP, cyan fluorescent protein; *E. coli*, *Escherichia coli*; EGFP, enhanced green fluorescent protein; FP, fluorescent protein; FRET, fluorescence resonance energy transfer; GFP, green fluorescent protein; HIC, hydrophobic interaction chromatography; LB, Luria–Bertani; OD, optical density; PCR, polymerase chain reaction; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; TFP, teal fluorescent protein; TPP, three-phase partitioning; Tris, Tris(hydroxymethyl)aminomethane; Tris-EDTA buffer, 10 mM Tris, 1 mM ethylenediaminetetraacetic acid; YFP, yellow fluorescent protein; β Gal, beta galactosidase.

INTRODUCTION

Fluorescent proteins (FPs) continue to be important reporters for imaging and cell-based assays. The advantage of FPs over other fluorescent reporters is their ability to form a fluorescent chromophore autocatalytically (1,2), which allows them to be used *in vivo* and in cell lines without the addition of external reagents. Mutagenesis has expanded the number of useful FPs from a pool of naturally occurring sequences (3,4). A variety of colors and properties are available for many tasks, such as sensors and fluorescence resonance energy transfer (FRET) applications (5). Improvements are continually being sought. For example, it has

been difficult to generate a sensitive far-red variant of the *Aequorea victoria* green FP (4,6). Improvements are also being made to cyan FPs by altering their fluorescent lifetimes to serve as better FRET donors (7,8). It would be desirable to increase the repertoire of FP sequences in the hope that novel properties could be discovered, from either native or mutated sequences.

Rhacostoma atlantica has been classified as a member of the order Leptomedusae and family Aequoreidae (9,10). The bioluminescent jellyfish is often seen on the coasts of the northeastern United States. We have observed them from the New Jersey to the North Carolina shore. Using a PCR-based approach, we have cloned a novel FP from *R. atlantica*. To our knowledge, there is no sequence information available for any other *Rhacostoma* protein. The closest related FP is the *Phialidium* yellow FP, with only a 55% amino acid sequence identity (11). The protein was expressed in *Escherichia coli* and purified to near homogeneity using three-phase partitioning (TPP) followed by hydrophobic interaction chromatography. The excitation and absorption spectra of the wild-type *Rhacostoma* FP are red shifted compared to the major excitation peak at about 397 nm in the native *Aequorea* FP, although the *Aequorea* FP also has a smaller peak at about 475 nm (12,13). The *Rhacostoma* FP may be a useful alternative to the cyan FP.

A combination of random and site-specific mutagenesis was utilized to generate red-shifted *Rhacostoma* FPs. Detection of mutants was accomplished by first generating a colorless *Rhacostoma* protein to serve as a template for random mutagenesis. Several red-shifted mutants were identified.

MATERIALS AND METHODS

Materials. DNA mini preps were prepared using a Zyppy plasmid miniprep kit. PCR products were purified using an Omega E.N.Z.A Cycle Pure Kit (#D6492-00). The thermocycler was a Minicycler from MJ Research. The DR-46B transilluminator was from Clare Chemical Research. The transilluminator has a peak excitation at about 450 nm and is viewed through a 555 nm long-pass filter.

cDNA cloning and mutagenesis. A specimen of *R. atlantica* was collected off the coast of New Jersey and frozen in liquid nitrogen. Total RNA was isolated with a MasterPure Complete DNA and RNA Purification Kit (Epicentre, Madison, WI) and a cDNA library was prepared with a SMARTer PCR cDNA synthesis kit (Takara Bio USA, Inc., Mountain View, CA). The RNA preparation and initial cDNA library were prepared by ACGT Inc. from supplied frozen samples. An Advantage-HF 2 PCR Kit (Takara Bio USA, Inc.) was used for amplifying DNA from the *Rhacostoma* cDNA library. The forward primer was HGGDRANNTH CCWGTWCCATGGBCWAC (Leptomedusae primer) used at 2 μ M and the reverse primer was GTGGTATCAACGCAGAGTACTTTTTTTTTT

*Corresponding author email: crebb@rci.rutgers.edu (William W. Ward)

[†]Part of the data in this paper was presented at the 37th Meeting of the American Society for Photobiology in June 2014 and in the patent application WO2014145554A2.

© 2016 The American Society of Photobiology

TTT used at 200 nm. All primers are listed from the 5' to 3' direction. The reaction was run at 94°C for 90 s followed by 26 cycles of 30 s at 94°C, 30 s at 60°C and 3 min at 72°C. The product was then reamplified for an additional 15 cycles. The PCR products were purified and ligated into a pGEM vector using T/A cloning and transformed into *E. coli* (XL10-Gold; Agilent Technologies, Santa Clara, CA). Resulting open reading frame DNA sequences were examined for GFP-like sequences using a BLASTP search (14). A DNA sequence corresponding to the C-terminal region of a potential FP was identified and used to design an additional set of primers. In order to provide a 5' template, the cDNA library was inserted into a pGEM vector. A PCR was performed on the ligation mixture using *Rhacostoma* primer 712 (TGGTAGTGCAGT CATTCCACATCAACG) and a primer corresponding to the pGEM vector just 5' of the incorporation site (TGACCATGATTACGCCAAGC TATTTAGGTG, BGAL primer). A second PCR was performed on the product using the BGAL primer and a nested primer (CGGCAGTGATG-TATTCCACCATCCTC, primer 642). Each primer was at 200 nm and the PCR was run as above for 26 cycles for each reaction. The product was cloned back into pGEM as described earlier.

The resulting 5' fragment of the DNA, corresponding to the N terminus of the novel FP, was then used to generate a new set of forward primers. The full-length sequence was obtained by a nested PCR on the original cDNA library. The first reaction primer set was primer 13 (CAGCGAGCGATACATCACACACACC) and primer 712. The second reaction used primer 28 (CACACACACCAAGAATTCAAAGTT TCC) and primer 712. Each reaction was run for 25 cycles as described earlier. The resulting PCR product was cloned into a pGEM vector as described and sequenced. In order to move the coding sequence into an expression vector, the DNA was amplified from the pGEM vector with the primers NDE-145 (CAAACCCATATGAGCACTGGAAAGACTGG) and KPN-145 (CGGGTACCTTATCCTCTTTATTATAAGCAGT GTC), thus introducing a *NdeI* site and a *KpnI* site at the 5' and 3' end, respectively, for the PCR product. The purified PCR product was digested with restriction enzymes and ligated (with T4 ligase) into a pBAD vector cut with the same enzymes. The pBAD vector harboring the *Rhacostoma* FP open reading frame DNA sequence was transformed into *E. coli* DH10 or NEB10 (New England Biolabs, Ipswich, MA) cells. Sequence comparisons were performed with ClustalX (15).

Mutant V204Y was created by site-directed mutagenesis using a primer overlap extension method (16). Wild-type *Rhacostoma* FP DNA was used as a template and amplified with the primer sets NDE-145 (GAG-GACTGTTTGATACTGAATATGGTGATAATGC, V204YR) and KPN-145 (GCATTATCACCATATTCAGTATCAAACAGTCTC, V204YF). Phusion polymerase (New England Biolabs) was used and template DNA was removed by treatment with *DpnI*. The two overhanging PCR products were combined and amplified with Phusion polymerase. The resulting PCR product was used as a megaprimer with the wild-type *Rhacostoma* sequence in pBAD as the template. The protein resulting from the mutant had little or no fluorescence. The V204Y was used as a template for two rounds of random mutagenesis using a GeneMorph II EZ clone Domain Mutagenesis Kit (#200552). The error prone PCR was performed using the NDE-145 and KPN-145 primers as directed in the kit to yield 4.5–9 mutations per kb. The first round of random mutagenesis yielded mutant V204Y/M221T and the second round yielded mutant T186N/V204Y/M221T. The T186N/V204Y/M221T was used as a template to generate a semirandom mutation targeted at Alanine 68. A PCR product was prepared with the primers CAACTATTTTAACCT-CACTGNNNTATGGAGTCAC (A68X) and KPN-145 using Phusion polymerase. The PCR product was used as a megaprimer.

Protein expression and purification. *Escherichia coli* was grown in a 3–5 mL starter culture of LB media containing carbenicillin (0.1 mg mL⁻¹). The starter culture was incubated for ~16 h (overnight) at 37°C with shaking. The overnight culture was centrifuged and collected cells were resuspended in fresh media. The collected cells were then diluted 50- to 200-fold in fresh LB media with carbenicillin. The resuspended cells grew for 3 h at 37°C. L-arabinose (0.1%) was added as an inducer and the cells were allowed to incubate for 22 h at room temperature (~21°C) with shaking. Cells were harvested by centrifugation for 30 min at 3000 g.

A TPP method (17,18) was used to purify the *Rhacostoma* FP for characterization. Using the conditions stated previously, 100–500 mL cultures were grown. For each 50 mL of culture, the cell pellet was resuspended in 25 mL of an aqueous solution of 1.6 M ammonium sulfate buffered with 50 mM Tris-HCl and 0.02% sodium azide at pH 8.0.

Twenty-five milliliter of *t*-butanol was added to the 25 mL cell suspension and the mixture was shaken vigorously by hand for ~60 s. The sample was centrifuged in a 50 mL centrifuge tube at 2550 g in a table-top swinging bucket centrifuge for 20 min. As expected, three phases (layers) were produced. The top alcohol layer was aspirated away and the congealed, semisolid disk of aggregated protein, nucleic acids and cell debris was removed with a spatula. The bottom layer of ammonium sulfate solution (~20 mL) containing the FP was saved. Thirty milliliters of fresh *t*-butanol was added to the saved ammonium sulfate solution. The above steps (vigorous shaking followed by centrifugation) were repeated. After centrifugation, the alcohol (top) layer and ammonium sulfate (bottom) layer were removed, leaving a thin green disk containing the FP. A minimal volume of 1.6 M ammonium sulfate stock solution buffered to pH 8.0 was used to suspend and dissolve the green disk. The dissolved green disk was transferred to a 1.5 mL microcentrifuge tube and centrifuged for 10 min at 16 000 g. The solution was separated into four layers, and the fluorescent layer, containing dissolved FP, was pipetted out and saved. The recovered fluorescent sample was run over a Phenyl Sepharose FF column (16 × 70 mm = 14 mL) equilibrated in 1.6 M ammonium sulfate buffered to pH 8.0 with Tris-HCl. The column was extensively washed with equilibration buffer (about 16 bed volumes or until the eluted fractions are less than 0.076 optical density [OD] units at 277 nm). The column was then eluted isocratically with 0.8 M ammonium sulfate in 10 mM Tris-HCl pH 8.0, until the OD466 had dropped 100-fold from the absorbance peak value. Fractions with an absorbance ratio (A_{466}/A_{277}) of 1.65 or above were pooled. Pools usually average a ratio of ~1.76.

UV and visible spectroscopy. The molar extinction coefficient of the FP was determined by estimating the chromophore concentration under alkaline conditions. At neutral pH, *Rhacostoma* FP has an absorption maximum at 466 nm. But when denatured with 0.1 M NaOH that absorption maximum shifts to about 446 nm. Both *Aequorea* GFP and *Renilla* GFP show a similar shift to 448 nm with 0.1 M NaOH and, at that common wavelength, have an extinction coefficient of 44 100 M⁻¹ cm⁻¹ (19). All three proteins have a tyrosine in position number 2 within the chromophore. So, in the denatured state, the chromophore absorbance peak of all three of these proteins becomes independent of the different protein environments in which they find themselves in the native state. Knowing that *Rhacostoma* has nearly the same absorption maximum in the denatured state, one can calculate, by a simple proportion, the molar extinction coefficient of its chromophore at neutral pH. In order to determine the concentration, a sample was diluted in 10 mM Tris-HCl pH 8.0 and mixed with 1/100 volume of 10 M NaOH. The FP concentration was calculated based on a molar excitation coefficient of 44 100 M⁻¹ cm⁻¹ for the alkaline-denatured chromophore. Absorbance values for these samples were typically between 0.1 and 0.2 OD units. The quantum yield of the wild-type *Rhacostoma* FP was estimated by comparison to EGFP at similar optical densities at 450 nm, using a value of 0.6 for the EGFP quantum yield (13). Emission spectra of *Rhacostoma* FP and EGFP were recorded from 470 to 600 nm using a Gilford Fluoro IV spectrofluorometer and the integrated areas were compared. Data were collected in a 0.5 × 1.0 cm cuvette with the 0.5 cm path used for excitation. To avoid inner filter effects, sample absorbance at 450 nm was lower than 0.05 OD units (based on a 1 cm path). In some instances, the peak absorbance for EGFP at 488 nm was between 0.05 and 0.09 OD. As an added precaution, the fluorescence emission spectra for these samples were corrected for a small inner filter effect (20). The correction reduced the quantum yield by less than 3%. For mutant *Rhacostoma* proteins, the samples were excited at 465 nm and the emission collected between 476 and 600 nm.

The excitation–emission matrices (EEM) were recorded on a Fluorolog 4, Horiba, NJ, USA, equipped with polarizers and a Peltier temperature control unit. From recorded polarized spectra, I_{xm} , a sum, s , and a difference, d , matrix were constructed according to equations:

$$s(\lambda_{ex}, \lambda_{em}) = I_{vv}(\lambda_{ex}, \lambda_{em}) + 2gI_{vh}(\lambda_{ex}, \lambda_{em}) \quad (1a)$$

$$d(\lambda_{ex}, \lambda_{em}) = I_{vv}(\lambda_{ex}, \lambda_{em}) - gI_{vh}(\lambda_{ex}, \lambda_{em}) \quad (1b)$$

The first and second subscript, x and m , indicates the orientation of the excitation and emission polarizers, respectively, *i.e.* v = vertical and

h = horizontal. The wavelength setting of the excitation and emission monochromators are indicated by λ_{ex} and λ_{em} respectively. The g factor corrects for different transmission efficiencies of polarized light through the optical detection system, and was calculated from

$$g(\lambda_{\text{ex}}, \lambda_{\text{em}}) = \frac{I_{\text{hv}}(\lambda_{\text{ex}}, \lambda_{\text{em}})}{I_{\text{hh}}(\lambda_{\text{ex}}, \lambda_{\text{em}})} \quad (2)$$

The matrix representing the steady-state fluorescence anisotropy was finally constructed from

$$r(\lambda_{\text{ex}}, \lambda_{\text{em}}) = \frac{d(\lambda_{\text{ex}}, \lambda_{\text{em}})}{s(\lambda_{\text{ex}}, \lambda_{\text{em}})} \quad (3)$$

and color coded in the software package Origin from OriginLab.

Time-resolved fluorescence decays recorded in the magic angle, *i.e.* at 54.7° , were obtained on a Flurocube from Horiba, NJ, USA, using the time-correlated single photon counting technique. The samples were excited with a NanoLEDs, centered at 444 nm or 467 nm, and the emission collected through a monochromator on a TBX 04 photomultiplier unit. An instrumental response function was recorded from a scattering solution and the data subsequently analyzed the software package Das6 from Horiba. The data were fitted to a multiexponential decay model

$$i(t) = \sum_{i=1}^k \alpha_i \exp\left(-\frac{t}{\tau_i}\right) \quad (4)$$

where α_i and τ_i indicates the pre-exponential factor and the decay constant of component i respectively. The quality of the fit was justified by inspecting the weighted residuals and the reduced chi-square value that should be close to 1 for a good fit.

RESULTS

cDNA cloning of the wild-type *R. atlantica* FP

We have cloned a novel FP from *R. atlantica*. The strategy was to amplify the FP from a *Rhacostoma* cDNA library using primers designed from highly conserved regions of FPs from coelenterates in the Leptomedusae order. A degenerate primer was designed based on the amino acid stretch GKLPVPWPT in *A. victoria*. This highly conserved sequence includes a turn leading into the central helix of the protein. The degenerate primer was used in conjunction with a primer containing a poly T repeat to amplify potential FPs from a cDNA library of a *Rhacostoma* specimen collected at the New Jersey coast. The PCR products were sequenced and open reading frames were used in a BLAST search (14) against the GenBank protein database to identify GFP-like proteins. A partial sequence was used to design primers to amplify the full-length FP, as described in Materials and Methods section. The DNA sequence and deduced amino acid sequence are shown in Fig. 1.

Figure 2 shows a comparison of the *Rhacostoma* protein with several FPs from the order Leptomedusae. The original degenerate primer design was based on the nucleotide alignment of the cDNA from these Leptomedusae proteins. The closest related protein known to us was the *Phialidium* yellow protein (11) with a 55% overall amino acid sequence identity. The identity with *A. victoria* GFP was only 50%. The N-terminal region of the *Rhacostoma* FP sequence (amino acids 1–118) has a 58% identity compared to the N-terminal sequence of the *Aequorea* GFP and a 59% identity compared to the N-terminal sequence of the *Phialidium* FP. The C-terminal region (amino acids 119–238)

has a 42% identity with the *Aequorea* GFP sequence and a 50% identity with the *Phialidium* FP sequence. Thus, the *Rhacostoma* FP has a slightly greater diversity in the C-terminal region. It is interesting to note that the four members of the Aequoreidae family in Fig. 2 shared over 80% identity with each other, but none had more than 50% identity with *Rhacostoma* FP.

Protein expression and purification

The *Rhacostoma* FP was subcloned into a pBAD vector and expression was induced in *E. coli* with L-arabinose. The FP was isolated by TPP (17) followed by hydrophobic interaction chromatography (HIC). Following TPP of a representative preparation, the ratio of the absorbance at 466 nm divided by the absorbance at 277 nm was about 0.87. The 466/277 ratio can be used as an indicator of purification. After HIC purification, this ratio was typically about 1.76, although it ranged from 1.46 to 1.78. A sample, analyzed by SDS-PAGE (Fig. 3A), was shown to be purified to near homogeneity. Size exclusion chromatography indicated a molecular weight of 46 kDa (Fig. 3B). An absorption spectrum of the protein is shown in Fig. 4, showing a peak absorption at 466 nm. An additional gel-filtration step using Cl-Sepharose 6B was occasionally used, resulting in a protein with a 466/277 ratio of 1.80.

Fluorescent properties of the wild-type *Rhacostoma* FP

The fluorescence EEM recorded for *Rhacostoma* FP is presented in Fig. 5A. The graph was constructed from calculated sum curves, $s(\lambda_{\text{ex}}, \lambda_{\text{em}})$, according to Eq. (1a), and scaled to 1 at the dominant peak centered at $(\lambda_{\text{ex}}, \lambda_{\text{em}}) = (470, 498)$ nm. The weaker band, observed for excitation at 275 nm, is presumably due to energy transfer from the aromatic amino acids to the chromophore complex within *Rhacostoma* FP. Excitation coefficients and quantum yields are summarized in Table 1. The corresponding steady-state anisotropy matrix, calculated according to Eq. (1b), is shown in Fig. 5B. The steady-state anisotropy is constant, *i.e.* wavelength independent, and close to 0.33 over the main emission peak, indicating nearly parallel absorption and emission transition dipoles. The high anisotropy is consistent with a rigid chromophore and congruent excitation and emission dipoles. The emission observed at 500 nm, following excitation at 275 nm, may involve energy transfer processes from aromatic amino acids, causing depolarization of the emission, as is reflected by the low anisotropy indistinguishable from the background.

Figure 6 shows the dependence of fluorescence on temperature for the wild-type *Rhacostoma* FP. A fluorescence decrease of 26% was seen when the temperature was raised from 20 to 40°C.

Heating the protein to 60°C caused a 30% decrease in fluorescence which was not completely reversed by cooling. Steady-state fluorescence anisotropies were recorded for *Rhacostoma* FP following a heat and cool cycle between 5°C and 60°C, and the corresponding Perrin graphs are presented in Figs. 7 and 8. The limiting anisotropies are identical between the heat and cool cycle, $r_0 = 0.330$, however, the calculated protein radius is slightly smaller for the cool cycle (3.99 ± 0.15 nm) as compared to the heat cycle (4.61 ± 0.15 nm).

Fluorescence lifetimes for wild-type *Rhacostoma* FP are shown in Table 2. The fluorescence decays are well described by

```

-----|-----|-----|-----|-----|
1 ATGAGCACTGGAAAGACTGGTAAAATGCTCTTCCAACAAGAGATTTCCTTT 50
1 M S T G K T G K M L F Q Q E I P F 17

-----|-----|-----|-----|-----|
51 CATCGTGTTCATTAGATGGTGAAGTTGAGGGAGAAAATATTTGGTGTTCAGAG 100
18 I V S L D G E V E G E I F G V R G 34

-----|-----|-----|-----|-----|
101 GGGAAGGATATGGAGATGCTACCATTGGTAAGATAGACATCACCTATCAT 150
35 E G Y G D A T I G K I D I T Y H 50

-----|-----|-----|-----|-----|
151 TGTATCACCGGGAAATTGCCAGTACCATGGCCAACCTATTTTAACTCACT 200
51 C I T G K L P V P W P T I L T S L 67

-----|-----|-----|-----|-----|
201 GGCCTATGGAGTCACATGTTTTGCGAAAATATCCCGAAAATGTCAACGATT 250
68 A Y G V T C F A K Y P E N V N D F 84

-----|-----|-----|-----|-----|
251 TCTTTAAAGATTGTATGCCTGAAGGCTACGTGCAGGAGAGGACTATCTCG 300
85 F K D C M P E G Y V Q E R T I S 100

-----|-----|-----|-----|-----|
301 TTTGAAGGTGAAGGCGTCTATAAGACACGAGCAGAAGTCACTTACGAAAG 350
101 F E G E G V Y K T R A E V T Y E S 117

-----|-----|-----|-----|-----|
351 TGGAAGTGTGTACAACAGAGTCCAATTGACTGGCTCTGGCTTCAAGAGAA 400
118 G T V Y N R V Q L T G S G F K R N 134

-----|-----|-----|-----|-----|
401 ATGGGAACATCCTAGCCAAGAAATTGGAATTCAATTTCAATCCAAGTTGC 450
135 G N I L A K K L E F N F N P S C 150

-----|-----|-----|-----|-----|
451 AGTTATGTTCTTCCAGACGCAGAGAACAATGGAATAAACCTTGTCTTTAA 500
151 S Y V L P D A E N N G I N L V F K 167

-----|-----|-----|-----|-----|
501 ACAGGTGCACAATATCGTTGGAGGTGATTTTATTATTGGCGAGCAGCATC 550
168 Q V H N I V G G D F I I G E H D Q 184

-----|-----|-----|-----|-----|
551 AGCAAACCAGGCCCATTTGGCAAGGGTCCGGACGCCCTCCCGCATTATCAC 600
185 Q T R P I G K G P D A L P H Y H 200

-----|-----|-----|-----|-----|
601 CATATTCAGGTTCAAACAGTCCTCTCAAAAAGACCCTGAGGAACCCAGAGA 650
201 H I Q V Q T V L S K D P E E P R D 217

-----|-----|-----|-----|-----|
651 CAATATGAGGATGGTGAATACATCACTGCCGTTGACTGCGACTGCTT 700
218 N M R M V E Y I T A V D C D T A Y 234

-----|-----|
701 ATAATAAAGGAGGATAA 717
235 N K G G * 238

```

Figure 1. Nucleotide and deduced amino acid sequence of the *Rhacostoma* fluorescent protein open reading frame.

<i>Aequorea victoria</i>	MSK...GEE ^L FTGVV ^P ILV ^E LD ^G VD ^V NG ^H HK ^F SVS ^G EGEG ^D ATY ^G KGLTLKFIC ^T TGK	52
<i>Aequorea coerulescens</i>	MSK...GAEL ^F FTGVV ^P ILIEL ^N GD ^V NG ^H HK ^F SVS ^G EGEG ^D ATY ^G KGLTLKFIC ^T TGK	52
<i>Aldersladia magnificus</i>	MSK...GAEL ^F FTGIV ^P ILIEL ^N GD ^V HG ^H HK ^F SVK ^G EGEG ^D ATY ^G KLEIKFV ^C TGK	52
<i>Aequorea macrodactyla</i>	MSK...GEE ^L FTGIV ^P VLI ^E LD ^G VD ^V HG ^H HK ^F SVR ^G EGEG ^D ADY ^G KLEIKFIC ^T TGK	52
<i>Obelia sp. MH-2011</i>	MS...SAGAL ^F TNKI ^P YV ^T E ^L EG ^D V ^N G ^M K ^F TIH ^G K ^G T ^G D ^A ST ^G HIEAKYV ^C TS ^G E	53
<i>Clytia gregaria</i>	MTALTEGAK ^L FEKEI ^P YIT ^E LE ^G D ^V EG ^M K ^F I ^I K ^G E ^G T ^G D ^A TT ^G TIKAKYI ^C TG ^D	55
<i>Phialidium sp. SL-2003</i>	MSS...GAL ^F HGKI ^P YV ^V E ^M E ^G N ^V D ^G H ^T F ^S IR ^G K ^G Y ^G D ^A SV ^G KVDAQFI ^C TG ^D	52
<i>Rhacostoma atlantica</i>	MSTGKT ^G KM ^L FQ ^Q EI ^P FIV ^S L ^D G ^E V ^E G ^E I ^F G ^V R ^G E ^G Y ^G D ^A TI ^G KIDITYH ^C IT ^G K	55
<i>Rhacostoma mutant 3</i>	MSTGKT ^G KM ^L FQ ^Q EI ^P FIV ^S L ^D G ^E V ^E G ^E I ^F G ^V R ^G E ^G Y ^G D ^A TI ^G KIDITYH ^C IT ^G K	55
<i>Rhacostoma mutant 4</i>	MSTGKT ^G KM ^L FQ ^Q EI ^P FIV ^S L ^D G ^E V ^E G ^E I ^F G ^V R ^G E ^G Y ^G D ^A TI ^G KIDITYH ^C IT ^G K	55
<i>Aequorea victoria</i>	L ^P V ^P W ^P T ^L V ^T T ^F S ^Y G ^V Q ^C F ^S R ^Y P ^D H ^M K ^Q H ^D F ^F K ^S A ^M P ^E G ^Y V ^Q E ^R T ^I F ^F K ^D D ^G N ^Y K	107
<i>Aequorea coerulescens</i>	L ^P V ^P W ^P T ^L V ^T T ^F S ^Y G ^V Q ^C F ^S R ^Y P ^D H ^M K ^Q H ^D F ^F K ^S A ^M P ^E G ^Y I ^Q E ^R T ^I F ^F K ^D D ^G N ^Y K	107
<i>Aldersladia magnificus</i>	L ^P V ^P W ^P T ^L V ^T T ^F S ^Y G ^V Q ^C F ^A R ^Y P ^E H ^M K ^M H ^D F ^F K ^S A ^M P ^E G ^Y I ^Q E ^R T ^I F ^F Q ^D D ^G K ^Y K	107
<i>Aequorea macrodactyla</i>	L ^P V ^P W ^P T ^L V ^T T ^F S ^Y G ^I Q ^C F ^A R ^Y P ^E H ^M K ^M N ^D F ^F K ^S A ^M P ^E G ^Y I ^Q E ^R T ^I F ^F Q ^D D ^G K ^Y K	107
<i>Obelia sp. MH-2011</i>	I ^P V ^P W ^A T ^L V ^S T ^M C ^Y G ^V Q ^C F ^A K ^Y P ^S H ^I K ^{..} D ^F Y ^K S ^A M ^P E ^G Y ^I Q ^E R ^T I ^S F ^E G ^D G ^V Y ^K	106
<i>Clytia gregaria</i>	L ^P V ^P W ^A T ^I L ^S S ^L S ^Y G ^V F ^C F ^A K ^Y P ^R H ^I A ^{..} D ^F F ^K S ^T Q ^P D ^G Y ^S Q ^D R ^I I ^S F ^D N ^D G ^Q Y ^D	108
<i>Phialidium sp. SL-2003</i>	V ^P V ^P W ^S T ^L V ^T T ^L T ^Y G ^A Q ^C F ^A K ^Y G ^P E ^L K ^{..} D ^F Y ^K S ^C M ^P E ^G Y ^V Q ^E R ^T I ^T F ^E G ^D G ^V F ^K	105
<i>Rhacostoma atlantica</i>	L ^P V ^P W ^P T ^I L ^T S ^L A ^Y G ^V T ^C F ^A K ^Y P ^E N ^V N ^{..} D ^F F ^K D ^C M ^P E ^G Y ^V Q ^E R ^T I ^S F ^E G ^E G ^V Y ^K	108
<i>Rhacostoma mutant 3</i>	L ^P V ^P W ^P T ^I L ^T S ^L A ^Y G ^V T ^C F ^A K ^Y P ^E N ^V N ^{..} D ^F F ^K D ^C M ^P E ^G Y ^V Q ^E R ^T I ^S F ^E G ^E G ^V Y ^K	108
<i>Rhacostoma mutant 4</i>	L ^P V ^P W ^P T ^I L ^T S ^L A ^S Y ^G V ^T C ^F A ^K Y ^P E ^N V ^N ..D ^F F ^K D ^C M ^P E ^G Y ^V Q ^E R ^T I ^S F ^E G ^E G ^V Y ^K	108
<i>Aequorea victoria</i>	T ^R A ^E V ^K F ^E G ^D T ^L V ^N R ^I E ^L K ^G I ^D F ^K E ^D G ^N I ^L G ^H K ^L E ^Y N ^Y N ^S H ^N V ^Y I ^M A ^D K ^Q K ^N G ^I K	162
<i>Aequorea coerulescens</i>	S ^R A ^E V ^K F ^E G ^D T ^L V ^N R ^I E ^L T ^G T ^D F ^K E ^D G ^N I ^L G ^N K ^M E ^Y N ^Y N ^A H ^N V ^Y I ^M T ^D K ^A K ^N G ^I K	162
<i>Aldersladia magnificus</i>	T ^R A ^E V ^K F ^E G ^D T ^L V ^N R ^I E ^L K ^G M ^D F ^K E ^D G ^N I ^L G ^H K ^L E ^Y S ^Y N ^S H ^N V ^Y V ^M A ^D K ^P N ^N G ^L K	162
<i>Aequorea macrodactyla</i>	T ^R G ^E V ^K F ^E G ^D T ^L V ^N R ^I E ^L K ^G M ^D F ^K E ^D G ^N I ^L G ^H K ^L E ^Y N ^F N ^S H ^N V ^Y I ^M P ^D K ^A N ^N G ^L K	162
<i>Obelia sp. MH-2011</i>	T ^R A ^M V ^T Y ^E R ^G S ^I Y ^N R ^V T ^L T ^E N ^F K ^K D ^G H ^I L ^R K ^N V ^A F ^Q C ^L P ^S I ^L Y ^I L ^P D ^T V ^N N ^G I ^R	161
<i>Clytia gregaria</i>	V ^K A ^K V ^T C ^E N ^G T ^L Y ^N R ^V T ^V K ^G T ^G F ^K S ^N G ^N I ^L G ^M R ^V L ^Y H ^S P ^P H ^A V ^Y I ^L P ^D R ^K N ^G M ^K	163
<i>Phialidium sp. SL-2003</i>	T ^R A ^E V ^T F ^E N ^G S ^V Y ^N R ^V K ^L N ^G Q ^Q F ^K K ^D G ^H V ^L G ^K N ^L E ^F N ^F T ^P H ^C L ^Y I ^W G ^D Q ^A N ^H G ^L K	160
<i>Rhacostoma atlantica</i>	T ^R A ^E V ^T Y ^E S ^G T ^V Y ^N R ^V Q ^L T ^G S ^G F ^K R ^N G ^N I ^L A ^K K ^L E ^F N ^F N ^P S ^C S ^Y V ^L P ^D A ^E N ^N G ^I N	163
<i>Rhacostoma mutant 3</i>	T ^R A ^E V ^T Y ^E S ^G T ^V Y ^N R ^V Q ^L T ^G S ^G F ^K R ^N G ^N I ^L A ^K K ^L E ^F N ^F N ^P S ^C S ^Y V ^L P ^D A ^E N ^N G ^I N	163
<i>Rhacostoma mutant 4</i>	T ^R A ^E V ^T Y ^E S ^G T ^V Y ^N R ^V Q ^L T ^G S ^G F ^K R ^N G ^N I ^L A ^K K ^L E ^F N ^F N ^P S ^C S ^Y V ^L P ^D A ^E N ^N G ^I N	163
<i>Aequorea victoria</i>	V ^N F ^K I ^R H ^N I ^E D ^G ..S ^V Q ^L A ^D H ^Y Q ^Q N ^T P ^I G ^{..} D ^G P ^V L ^L P ^D N ^H Y ^L S ^T Q ^S A ^L S ^K D ^P N ^E ..	213
<i>Aequorea coerulescens</i>	V ^N F ^K I ^R H ^N I ^E D ^G ..S ^V Q ^L A ^D H ^Y Q ^Q N ^T P ^I G ^{..} D ^G P ^V L ^L P ^D N ^H Y ^L S ^T Q ^S T ^L S ^K D ^P N ^E ..	213
<i>Aldersladia magnificus</i>	V ^N F ^K I ^R H ^N I ^E G ^G ..G ^V Q ^L A ^D H ^Y Q ^T N ^V P ^L G ^{..} D ^G P ^V L ^L P ^I N ^H Y ^L S ^T Q ^T A ^I T ^K D ^P N ^E ..	213
<i>Aequorea macrodactyla</i>	V ^N F ^K I ^R H ^N I ^E G ^G ..G ^V Q ^L A ^D H ^Y Q ^T N ^V P ^L G ^{..} D ^G P ^V L ^I P ^I N ^H Y ^L S ^T Q ^T A ^I S ^K D ^R N ^E ..	213
<i>Obelia sp. MH-2011</i>	V ^E F ^N Q ^V Y ^D I ^E G ^E I ^{..} E ^K L ^V T ^K C ^S Q ^M N ^R P ^L A ^E S ^A A ^V H ^I P ^R Y ^H H ^I S ^K T ^H K ^L S ^K D ^L D ^E ..	214
<i>Clytia gregaria</i>	I ^E Y ^N K ^A F ^D V ^M G ^G ..G ^H Q ^M A ^R H ^A F ^N K ^P L ^G ..A ^W E ^E D ^Y P ^L Y ^H H ^L T ^V W ^T S ^F G ^K D ^P D ^D	215
<i>Phialidium sp. SL-2003</i>	S ^A F ^K I ^M H ^E T ^T G ^S K ^E D ^F I ^V A ^D H ^T Q ^M N ^T P ^I G ^{..} G ^P V ^H V ^P E ^Y H ^I T ^Y H ^V T ^L S ^K D ^V T ^D ..	213
<i>Rhacostoma atlantica</i>	L ^V F ^K Q ^V H ^N I ^V G ^G ..D ^F I ^I G ^E H ^D Q ^Q T ^R P ^I G ^{..} K ^G P ^D A ^L P ^H Y ^H H ^I Q ^V T ^V L ^S K ^D P ^E E ^{..}	214
<i>Rhacostoma mutant 3</i>	L ^V F ^K Q ^V H ^N I ^V G ^G ..D ^F I ^I G ^E H ^D Q ^Q N ^R P ^I G ^{..} K ^G P ^D A ^L P ^H Y ^H H ^I Q ^Y Q ^T V ^L S ^K D ^P E ^E ..	214
<i>Rhacostoma mutant 4</i>	L ^V F ^K Q ^V H ^N I ^V G ^G ..D ^F I ^I G ^E H ^D Q ^Q N ^R P ^I G ^{..} K ^G P ^D A ^L P ^H Y ^H H ^I Q ^Y Q ^T V ^L S ^K D ^P E ^E ..	214
<i>Aequorea victoria</i>	K ^R D ^H M ^V L ^L E ^F V ^T A ^A G ^I T ^H G ^M D ^E L ^Y K	238
<i>Aequorea coerulescens</i>	K ^R D ^H M ^I Y ^F E ^F V ^T A ^A A ^I T ^H G ^M D ^E L ^Y K	238
<i>Aldersladia magnificus</i>	T ^R D ^H M ^V F ^L E ^F F ^T A ^C G ^I T ^H G ^M D ^E L ^Y ..	237
<i>Aequorea macrodactyla</i>	T ^R D ^H M ^V F ^L E ^F F ^S A ^C G ^H T ^H G ^M D ^E L ^Y K	238
<i>Obelia sp. MH-2011</i>	R ^R D ^H M ^C L ^V E ^V V ^K A ^V D ^L D ^T Y ^Q	234
<i>Clytia gregaria</i>	E ^T D ^H L ^N I ^V E ^V I ^K A ^V D ^L E ^T Y ^R	235
<i>Phialidium sp. SL-2003</i>	H ^R D ^N M ^S L ^V E ^T V ^R A ^V D ^C R ^K ...TYL	234
<i>Rhacostoma atlantica</i>	P ^R D ^N M ^R M ^V E ^Y I ^T A ^V D ^C D ^T A ^Y N ^K G ^G ..	238
<i>Rhacostoma mutant 3</i>	P ^R D ^N M ^R T ^V E ^Y I ^T A ^V D ^C D ^T A ^Y N ^K G ^G ..	238
<i>Rhacostoma mutant 4</i>	P ^R D ^N M ^R T ^V E ^Y I ^T A ^V D ^C D ^T A ^Y N ^K G ^G ..	238

Figure 2. Sequence alignments of selected Leptomedusae fluorescent proteins. The amino acid sequences used were of fluorescent proteins from *Aequorea victoria* (gi|1169893), *Rhacostoma atlantica* (from the sequence in Fig. 1), *Aequorea coerulescens* (gi|34421676), *Aldersladia magnificus* (gi|183186849), *Aequorea macrodactyla* (gi|18175254), *Obelia sp. MH-2011* (gi|342221051), *Clytia gregaria* (gi|298257355) and *Phialidium sp. SL-2003* (gi|40365351). Introduced gaps are shown by dots. Identical residues are highlighted in purple, residues conserved at greater than 75%, but less than 100%, are highlighted in blue. The sequences for *Rhacostoma* mutant 3 (T186N/V204Y/M221T) and mutant 4 (A68S/T186N/V204Y/M221T) are also shown, mutated residues are highlighted in red. The figure was prepared using the TEXshade program (30).

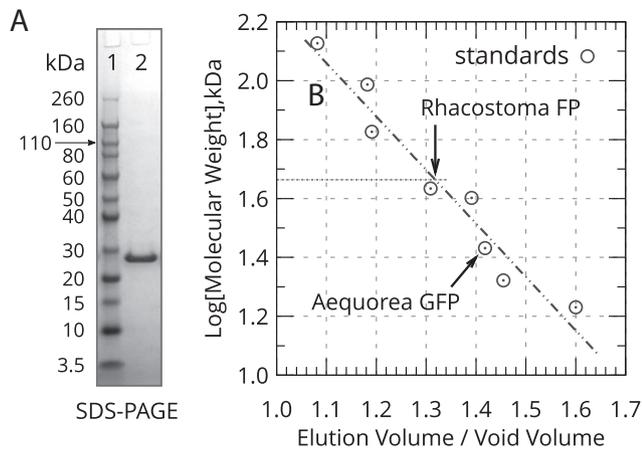


Figure 3. Molecular weight determination of the *Rhacostoma* fluorescent protein. (A) SDS gel of *Rhacostoma* fluorescent protein. Lane 1, Novex sharp prestained protein standard molecular weight markers. Lane 2, 2 μ g of purified wild-type *Rhacostoma* fluorescent protein. The proteins were heated with NuPage reducing agent (Invitrogen #NP0009) at 70°C for 10 min and run on a NuPAGE Novex 4–12% Bis-Tris Gel with 1% MES running buffer and NuPAGE antioxidant (Invitrogen #NP0323BOX) for 35 min at 200V constant and at an average current of 100 mA per gel. The gel was stained with SimplyBlue SafeStain (Invitrogen #LC6060) and captured by a Cannon PowerShot digital camera. The image was converted to black and white. (B) Size exclusion chromatography of the *Rhacostoma* FP. Molecular weight analysis was performed on a BioRad HPLC column (BioSil SEC-125–300 mm \times 7.8 mm). Buffer composition was 100 mM NaCl, 50 mM NaPO₄, 0.02% NaN₃—prefiltered through a 0.22 μ m ultra filter. The pH of the buffer was maintained at 6.5. Flow rate was 0.5 mL min⁻¹. The *Rhacostoma* FP weight was estimated to be 46 kDa from a linear fit of the log of the standards as a function of the elution volume from the SEC column. The molecular weight standards used were myoglobin (17 kDa), soybean trypsin inhibitor (21 kDa), *Aequorea* GFP (27 kDa), soybean peroxidase (40 kDa), ovalbumin (43 kDa), BSA monomer (67 kDa), phosphorylase b (97 kDa) and BSA stable dimer (134 kDa).

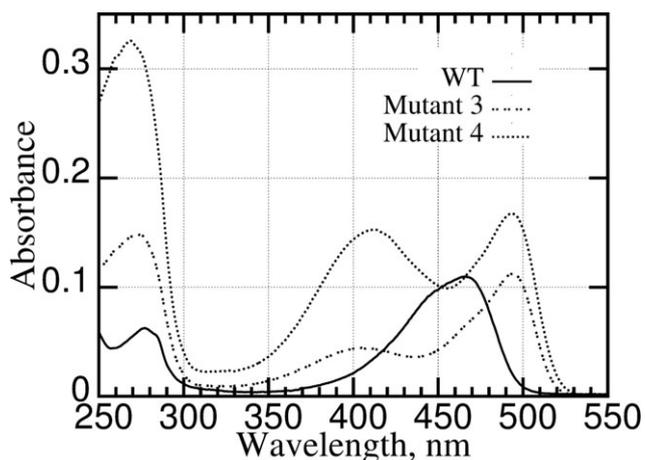


Figure 4. Absorption spectra of purified *Rhacostoma* fluorescent proteins. Purified proteins were diluted into 10 mM Tris-HCl pH 8.0 and the absorption spectra were collected on a Cary 100 spectrophotometer. WT is the wild-type *Rhacostoma* fluorescent protein. Mutant 3 is the T186N/V204Y/M221T protein. Mutant 4 is the A68S/T186N/V204Y/M221T protein.

a monoexponential decay model, *i.e.* $k = 1$, Eq. (4), for excitation at either 444 nm or 467 nm, and for emission recorded between 480 and 540 nm. A typical fit is shown in Fig. 9.

Mutagenesis

In order to generate a red-shifted mutant of the *Rhacostoma* FP, we initiated mutagenesis using a site-directed approach. Based on mutagenesis performed on the wt *Aequorea* GFP (21), we mutated valine 204 to tyrosine (V204Y, mutant 1). *Escherichia coli* expressing the mutant showed little or no fluorescence (Fig. 10).

Assuming additional mutations were needed to physically accommodate the V204Y, we initiated two rounds of random mutagenesis using the V204Y mutant as a template. Mutations were screened using a DR-46B transilluminator. The first round of random mutagenesis yielded a sequence with a second mutation (V204Y/M221T, mutant 2). The expressed mutant 2 protein was red shifted (Table 1), but the expression was low or the maturation was slow (Fig. 10). The second round of random mutagenesis introduced a third mutant (T186N/V204Y/M221T, mutant 3, Fig. 2), resulting in an additional red shift in the excitation spectrum (Table 1, Fig. 11) and better expression or maturation in *E. coli* (Fig. 10).

The absorption spectrum had peaks at 409 and 494 nm (Fig. 4). A fourth round of random mutation was targeted at the chromophore region alanine 68. Mutation was achieved by performing PCR amplification with a degenerate primer with all three nucleotides of alanine 68 codon varied. The mutant identified (A68S/T186N/V204Y/M221T, mutant 4, Fig. 2) resulted in a protein with better expression or maturation, but the protein had a lower extinction coefficient (Table 1, Fig. 10). Both mutant 3 and mutant 4 have a prominent absorption peak, or shoulder, in the 400–410 nm range. The lower extinction coefficient of mutant 4 at 493 nm was likely the result of the decrease in the ratio of peak absorbances (493 nm band to 400–415 nm region, Fig. 4). We did notice that the 400–415 nm peak could decrease upon long-term storage (70 days or longer stored at 4°C), especially for mutant 3. The 495/400 nm ratio of mutant 3 also displayed some concentration dependence. For concentrations between 0.01 and 0.04 OD at 495 nm (approximately 0.25–1 μ M, based on the extinction coefficient in Table 1), the 495/400 nm ratio was 1.6 ± 0.3 ($n = 9$, \pm standard deviation) for material less than 70 days old. For concentrations between 0.11 and 0.17 (about 2.6–4 μ M), the ratio was 2.2 ± 0.3 ($n = 4$). The concentration used for the extinction coefficient measurements was between 0.11 and 0.15 OD (about 2.6–3.6 μ M) at 495 nm. It was possible to drive the ratio higher at higher concentrations, but the effect was not reproducible. Mutant 4 seemed more stable, varying from a 493/405 nm ratio of 1.15 for low-OD samples to 1.19 for high-OD samples less than 70 days old.

The overall fluorescence lifetimes of mutant 3 and mutant 4 are similar to the wild-type proteins, although both mutants have a shorter lifetime component (Table 2). Mutant 3 shows a small variation across emission wavelength, while mutant 4 has a lifetime independent of emission wavelength.

DISCUSSION

Sequence comparisons

A PCR-based approach was used to identify a novel FP from *R. atlantica*—the first protein from this genus ever to be sequenced. The most closely related FP known in the literature is the yellow FP from *Phialidium* (11). *Phialidium* is a member of the Campanulariidae family, even though *Rhacostoma* is

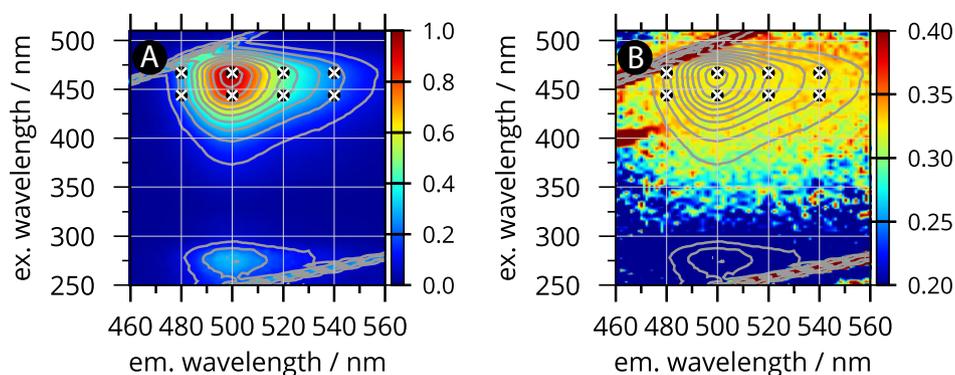


Figure 5. (A) Excitation–emission matrix (EEM) recorded for *Rhacostoma* FP in Tris-EDTA (pH 7.6) at 21°C. The circles (white) indicate where fluorescence lifetimes have been recorded (see Table 2). The color scale is normalized to 1 at the peak emission intensity centered around (500, 460) nm. (B) Anisotropy map corresponding to data presented in panel A. Note that the contour lines are shown for the intensity data, and not for the anisotropy values. Also, the anisotropy values are color coded from 0.2 to 0.4.

Table 1. Fluorescence properties of wild-type and mutant proteins.

<i>Rhacostoma</i> fluorescent protein	Longest wavelength excitation peak*, nm	Emission peak, nm	Quantum yield†	Extinction coefficient‡ M ⁻¹ cm ⁻¹
Wild type	470	498	0.77 ± 0.05 <i>n</i> = 8	58 200 ± 2170 <i>n</i> = 9
Mutant 1 V204Y	ND	ND	ND	ND
Mutant 2 V204Y/M221T	484	508	ND	ND
Mutant 3 T186N/V204Y/M221T	490	512	0.58 ± 0.05 <i>n</i> = 4	41 800 ± 4000 <i>n</i> = 3
Mutant 4 A68S/T186N/V204Y/M221T	492	514	0.60 ± 0.05 <i>n</i> = 3	26 770 ± 2190 <i>n</i> = 2

*Fluorescence peaks were extracted from the EEM data in Figs. 5 and 11, except for mutant 2. For mutant 2, fluorescence spectra were collected on a Gilford Fluoro IV spectrofluorometer. Data for wild-type, mutant 3 and mutant 4 were collected using purified protein. Data for mutant 2 were generated from a partially purified protein. †Quantum yields were determined as described in Materials and Methods using a Gilford Fluoro IV spectrofluorometer. ‡Extinction coefficients were determined as described in Materials and Methods for the longest wavelength absorbance peak, at an optical density between 0.1 and 0.2.

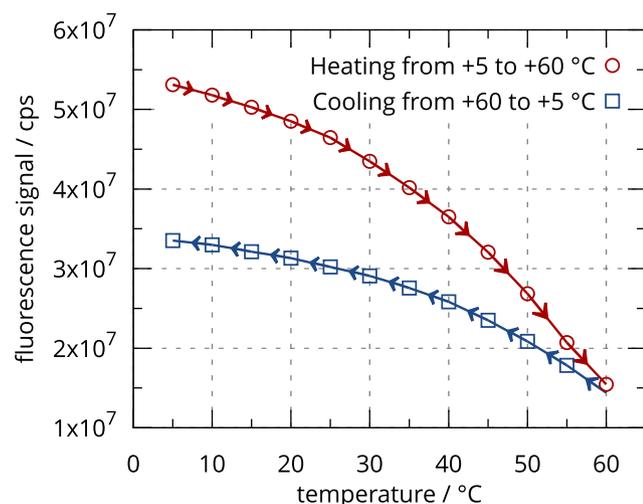


Figure 6. Stability of *Rhacostoma* FP in Tris-EDTA (pH 7.6) during heat and cool cycle. The excitation wavelength was 460 nm. The fluorescence signal is the integrated emission band centered at 500 nm. Note: First the sample is heated from 5 to 60°C, then the sample is cooled back to 5°C again.

placed in the family Aequoreidae (9,10). The amino acids that form the *Rhacostoma* FP chromophore are Ala-Tyr-Gly, a fairly rare combination, but also seen in the teal FP (22).

The *Rhacostoma* FP is readily purified from *E. coli*

The *Rhacostoma* protein expresses well in *E. coli* and like all other FPs we have studied, this protein can be highly purified by the use of TPP. Stage I of TPP releases soluble FP from whole, unlysed *E. coli* cells, while leaving the cell walls intact. Within the congealed disk, removed at the end of Stage I, no disrupted cells are seen by phase contrast microscopy (Dr. Diane Davis, personal communication). It is assumed that tertiary butanol, used in conjunction with an equal volume of 1.6 M aqueous ammonium sulfate, dissolves the membrane lipids upon 30 s of shaking, allowing soluble proteins to exit through large pores in the cell wall. Most of the other macromolecules (DNA and proteins) aggregate into large clumps, becoming entombed within the cells. This congealed material is easily removed by low-speed centrifugation (20 min at 2500 *g*). Final purification involves chromatography on a Phenyl Sepharose FF column, occasionally followed by a polishing step on Sepharose CL-6B, if needed.

Rhacostoma FP mass and size

The SDS gel of the purified protein (Fig. 3) shows a band consistent with the molecular weight calculated from the predicted amino acid sequence (26.5 kDa). The high fluorescence anisotropy observed for the *Rhacostoma* FP has also been seen in the *Aequorea* protein and is consistent with a rigid chromophore

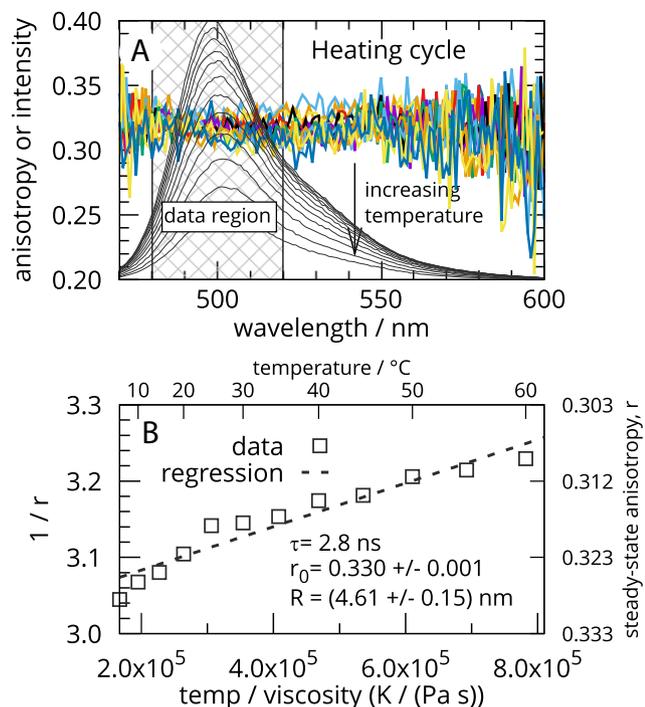


Figure 7. (A) Steady-state anisotropy and sum curves recorded for *Rhacostoma* FP dissolved in Tris-EDTA (pH 7.6) from 5 to 60°C, *i.e.* this is the heat cycle. The dashed vertical lines indicate the region where the average value of the steady-state anisotropies were calculated and used for the Perrin analysis. Note that the intensity of the sum curve decreases with temperature. (B) Perrin graph for anisotropies obtained from panel A.

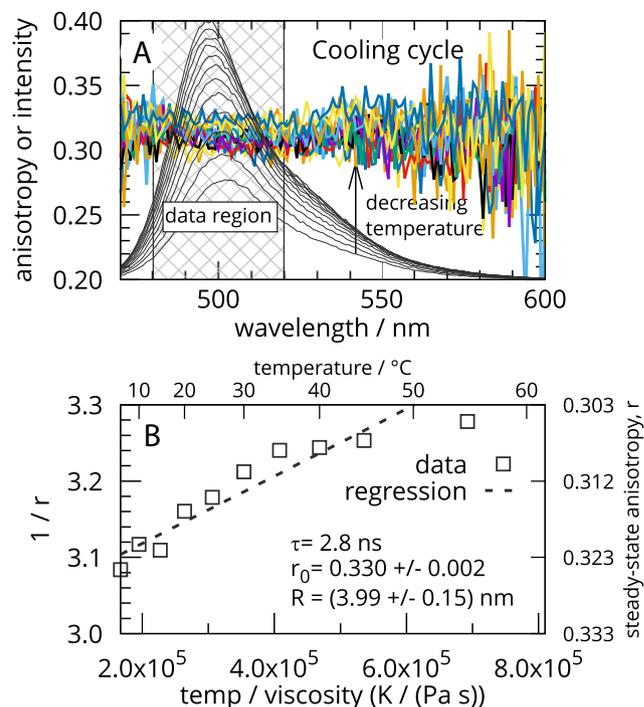


Figure 8. (A) Steady-state anisotropy and sum curves recorded for *Rhacostoma* FP dissolved in Tris-EDTA (pH 7.6) from 60 to 5°C, *i.e.* this is the cool cycle. The dashed vertical lines indicate the region where the average value of the steady-state anisotropies were calculated and used for the Perrin analysis. Note that the intensity of the sum curve decreases with temperature. (B) Perrin graph for anisotropies obtained from panel A.

(23,24), although the limiting anisotropy calculated for the *Rhacostoma* FP ($r_0 = 0.330$) is slightly lower than that calculated for the *Aequorea* FP ($r_0 = 0.39$) (24). From the Perrin graphs (Figs. 7 and 8), the radius of the *Rhacostoma* FP was calculated to be 4.0 or 4.6 nm, depending on the heating cycle. The radius of the *Aequorea* FP calculated from time-resolved anisotropy decay was 2.3 nm (24). The larger diameter of the *Rhacostoma* FP may be indicative of dimerization. A comparison of the mass determined from size exclusion chromatography (46 kDa, Fig. 3B) and the predicted molecular weight of 26.5 kDa would support the premise that the *Rhacostoma* FP is a dimer in solution.

The wild-type *Rhacostoma* FP as a potential FRET donor

The wild-type *Rhacostoma* FP has an absorption peak at 466 nm, with a shoulder at about 440 nm. The extinction coefficient is $58\,200\text{ M}^{-1}\text{ cm}^{-1}$ and the quantum yield is 0.77. The brightness (product of extinction coefficient and quantum yield, $45\text{ nm}^{-1}\text{ cm}^{-1}$) compares favorably to EGFP ($34\text{ nm}^{-1}\text{ cm}^{-1}$) (25). The emission peak at 498 nm (Table 1, Fig. 5) makes this protein a potential FRET donor to proteins such as YFP (13) and Citrine (26). A commonly used FRET donor to the YFP is the cyan FP (CFP) derived from the *Aequorea* GFP (27). The CFP has a tryptophan replacing the tyrosine in the chromophore, resulting in a wide emission spectrum, about 60 nm at half height. In contrast, the emission spectrum of *Rhacostoma* FP is less than 40 nm at half height. A narrow emission spectrum could be helpful in FRET assays, as a broad emission from the

Table 2. Fluorescence lifetimes recorded for *Rhacostoma* FP and mutants dissolved in Tris-EDTA (pH 7.6) at 21°C.

Sample	λ_{ex} (nm)	λ_{em} (nm)	f_1 (%)	τ_1 (ns)	f_2 (%)	τ_2 (ns)	τ_{avg} (ns)	χ^2
<i>Rhacostoma</i> FP	444	480	100	2.84*			2.84	1.08
	444	500	100	2.89			2.89	1.05
	444	520	100	2.90			2.90	1.07
	444	540	100	2.94			2.94	1.02
<i>Rhacostoma</i> FP	467	480	100	2.88*			2.88	1.12
	467	500	100	2.87			2.87	1.08
	467	520	100	2.91			2.90	1.03
Mutant 3	444	500	6.7	0.75	93.3	2.92	2.77	1.07
	444	520	5.9	1.56	94.1	3.07	2.98	1.04
	444	540	7.3	1.55	92.7	3.09	2.98	0.99
Mutant 4	444	560	4.7	1.58	95.3	3.08	3.01	1.06
	444	540	8.4	1.11	91.6	3.25	3.07	1.16
	444	560	9.6	1.14	90.4	3.34	3.05	1.19

*The data were better fitted to a biexponential decay model reflecting a contamination of scattered excitation light. The dominating decay component is tabulated.

donor chromophore can be problematic. The homogeneous lifetime decay seen for the *Rhacostoma* FP may be desirable for fluorescence-lifetime imaging microscopy (FLIM)-based energy transfer studies. Much effort has been placed on improving the CFP for energy transfer studies, part of which was the development of homogeneous lifetime decays (7,8), although the shapes of the emission spectra were essentially unchanged. Another potential energy donor to the YFP is the teal FP (TFP) (22),

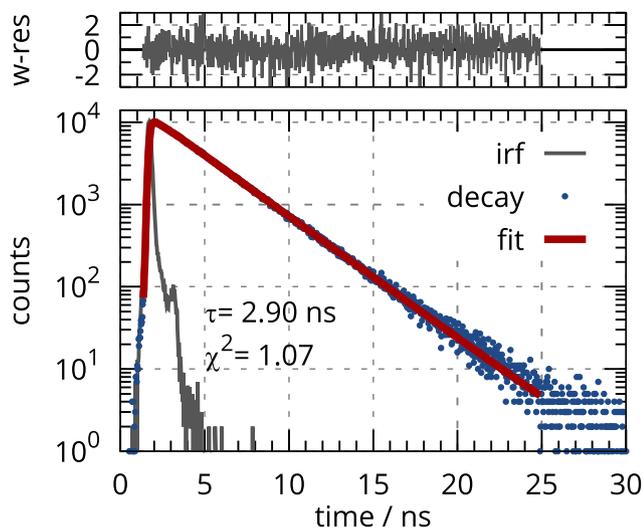


Figure 9. Time-resolved decay recorded for *Rhacostoma* FP in Tris-EDTA (TE) 1× solution at pH 7.6, excited at 444 nm and with emission collected at 520 nm. The data fits well to a monoexponential decay function. The instrumental response function (irf) was recorded from a scattering solution.

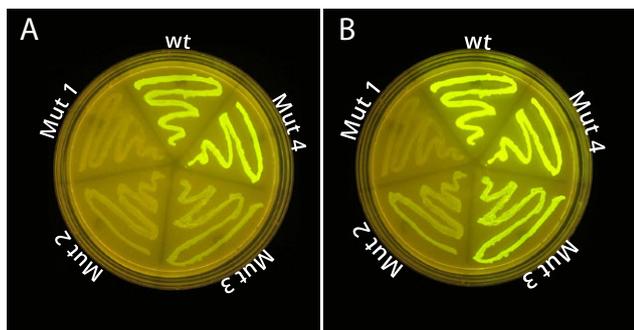


Figure 10. Mutant expression. *Escherichia coli* (NEB10 cells) expressing wild-type (WT) or mutant fluorescent proteins were streaked on an LB agar plate with 100 $\mu\text{g mL}^{-1}$ carbenicillin and 0.1% L-arabinose. The plate was allowed to grow overnight at 37°C and then room temperature for 8 h and photographed on a Clare Chemical DR-46B transilluminator (A). The plate was then incubated at 4°C for 48 h and photographed again (B). Mut1 is V204Y, Mut2 is V204Y/M221T, Mut3 is T186N/V204Y/M221T, and Mut4 is A68S/T186N/V204Y/M221T.

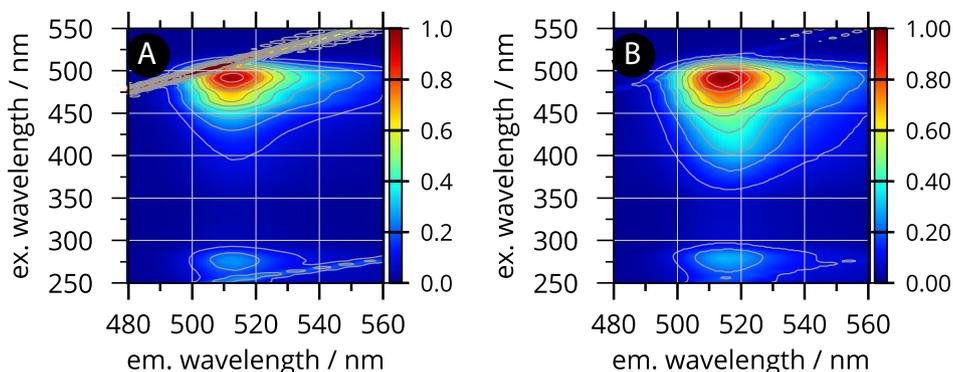


Figure 11. Contour emission graph recorded for mutant 3 and mutant 4. (A) Mutant 3 in Tris-EDTA pH 7.6 at 20°C. (B) Mutant 4 in Tris-EDTA pH 7.6 at 20°C.

having excitation and emission peaks close to the *Rhacostoma* FP peaks. Although the TFP and *Rhacostoma* FP share the same Ala-Tyr-Gly sequence that forms the chromophore, the TFP, derived from the distantly related *Clavularia* soft coral protein, has only 26% amino acid identity with *Rhacostoma* (based on alignment using the ClustalX program and comparing to gi110589865). Thus, the *Rhacostoma* FP provides a new source for development and modification of proteins for use in this color range.

Red-shifted mutants of *R. atlantica*

We attempted to generate a red-shifted mutant of *R. atlantica* by replacing the valine at position 204 with a tyrosine. In the *A. victoria* protein, introduction of a tyrosine in the homologous position contributes to forming a yellow FP, the spectral shift probably aided by stacking of the tyrosine with the chromophore (21). In the *Phialidium* FP, the naturally occurring residue at that homologous position is tyrosine (Fig. 2), and that protein has a yellow emission spectrum as well (peak at 537 nm) (11). Since the *Phialidium* protein was the closest related FP to *Rhacostoma*, making this substitution of valine to tyrosine seemed like a reasonable starting point for mutagenesis. However, introduction of this single mutant resulted in a protein with very low fluorescence, very poor expression or slow maturation (Fig. 10). It is likely that additional mutations would be needed to accommodate the introduction of the tyrosine. One option would be to design additional site-directed mutants based on available sequence data from *Phialidium* and the *Aequorea* yellow fluorescent mutants. An alternative approach is to take advantage of the nonfluorescent mutant as a template for random mutagenesis. Using the nonfluorescent mutant as the DNA template, additional colored mutants would be easy to detect on the DR-46B transilluminator light box. One round of random mutagenesis using the V204Y template yielded the double mutant V204Y/M221T.

The residues homologous to methionine 221 in *Phialidium* and *A. victoria* FPs are leucines. Random mutation of an EGFP with a disrupted chromophore (EGFP_{pevo}) also resulted in a substitution of leucine 220 that restored fluorescence (28). In the *Phialidium* structure (29), leucine 220 is pointed in toward the chromophore. It is possible that replacing leucine with the smaller residue, threonine, allowed some accommodation for the introduction of the tyrosine in the V204Y mutant. It is also

possible that the substitution allowed for a more optimal orientation of the critical Glu222.

Although fluorescence was detected, the V204Y/M221T mutant (mutant 2) was still very dim. Either it had a low intrinsic brightness or was a poorly expressed or was slow to mature in the *E. coli*. Another round of random mutagenesis resulted in the triple mutant T186N/V204Y/M221T (mutant 3). The absorption max was shifted to 494 nm, but there was also the appearance of a large absorption peak at about 410 nm (Fig. 4). The ~410 nm peak is not present in the excitation spectrum (Fig. 11), indicating it may be a nonfluorescent form of the chromophore, presumably the A form, or nonionic form of the chromophore (13,21). The fluorescence of the triple mutant was improved over the double mutant, but the yield of FP, upon expression in *E. coli*, appeared much lower than that of the wild-type, due to either poor expression or slow chromophore formation (Fig. 10). It is interesting to note that the homologous amino acid in the FPs from several other Leptomedusae at position 186 is an asparagine (Fig. 2). The random mutagenesis produced a FP with the more conserved asparagine in that position. The homologous residue in the *Phialidium* FP (N185) is within 3 Å of glutamine 183, an appropriate space for a hydrogen bond. Glutamine 183 is, in turn, within 3 Å of R94, which interacts with the chromophore (29). Thus, in *Phialidium*, it seems that asparagine 185 is indirectly participating in orienting the chromophore through an extended network. These residues are conserved in the *Rhacostoma* FP (Fig. 2).

The mutant A68S/T186N/V204Y/M221T (mutant 4) showed similar fluorescence spectra to the triple mutant 3 (Fig. 11), except the absorbance ratio of 494/~410 was lower and the extinction coefficient at 494 nm was lower (Table 1, Fig. 4). Mutant 4 did, however, appear to have a higher apparent expression in *E. coli*, close to that of the wild type. The higher expression may be due to a better/faster chromophore formation. Alanine 68 in the *Rhacostoma* FP is homologous to serine 65 in *A. victoria*, making up part of the chromophore. The wild-type *A. victoria* GFP (with a serine in position 65) displays a high ratio of the A form (nonionic chromophore) to the B form (ionic chromophore). Both forms are fluorescent in the *Aequorea* GFP. The presence of both forms is explained by the ability of Ser65 to stabilize Glu222 as an anion, leading to an inhibition of the B form in *A. victoria* GFP (13,21). In *Rhacostoma* wild-type FP, there is an unresolved peak in the absorption or excitation spectra that could correspond to the A band. An A-like band is apparent in the mutant 3 absorption spectrum, but not the excitation spectrum (Figs. 4 and 5), implying that the chromophore form absorbing in the 400–415 nm region is nonfluorescent. Both the wild-type *Rhacostoma* and mutant 3 have an alanine in position 68 (corresponding to serine 65 in *Aequorea*). The introduction of the A68S mutant (mutant 4) seems to increase the amount of the A-like form in *Rhacostoma*, but the A-like form is still apparent in the T186N/V204Y/M221T mutant, which has the alanine in chromophore region. This implies that the alanine in the chromophore region can, in the context of the triple mutant, lead to the formation of an A-like band.

CONCLUSIONS

The *Rhacostoma* FP has a novel amino acid sequence, showing the most variation in the C-terminal region when compared to *A. victoria* GFP and other GFP-like proteins from the

Leptomedusae order. Random mutagenesis has highlighted some areas of potential interest in the C-terminal region, such as T186, which in other FPs is a highly conserved asparagine. The *Rhacostoma* sequence, especially in the C-terminal region, supplies additional diversity to the collection of FPs. The availability of additional sequences for mutagenesis and structural studies will increase our understanding of FPs and perhaps lead to the discovery of molecules with novel and beneficial properties.

Acknowledgements—We would like to acknowledge Bratati Ganguly for assistance in preparing a *Rhacostoma* pGEM plasmid library and Karla Guerrero for assistance with SDS gels.

REFERENCES

- Prasher, D., V. Eckenrode, W. Ward, F. Prendergast and M. Cormier (1992) Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* **111**, 229–233.
- Chalfie, M., Y. Tu, G. Euskirchen, W. Ward and D. Prasher (1994) Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805.
- Heim, R. and R. Tsien (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Biol.* **6**, 178–182.
- Shaner, N., G. Patterson and M. Davidson (2007) Advances in fluorescent protein technology. *J. Cell Sci.* **120**, 4247–4260.
- Shaner, N., P. Steinbach and R. Tsien (2005) A guide to choosing fluorescent proteins. *Nat. Methods* **2**, 905–909.
- Mishin, A. S., F. V. Subach, I. V. Yampolsky, W. King, K. A. Lukyanov and V. V. Verkhusha (2008) The first mutant of the *Aequorea victoria* green fluorescent protein that forms a red chromophore. *Biochemistry* **47**, 4666–4673.
- Goedhart, J., D. von Stetten, M. Noirclerc-Savoie, M. Lelimosin, L. Joosen, M. Hink, L. van Weeren, T. Gadella and A. Royant (2012) Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. *Nat. Commun.* **3**, 751.
- Fredj, A., H. Pasquier, I. Demachy, G. Jonasson, B. Levy, V. Derrien, Y. Bousmah, G. Manoussaris, F. Wien, J. Ridard, M. Erard and F. Merola (2012) The single T65S mutation generates brighter cyan fluorescent proteins with increased photostability and pH insensitivity. *PLoS ONE* **7**, e49149.
- Agassiz, L. (1851) [a new naked-eyed medusa... *Rhacostoma atlanticum*]. *Proc. Boston Soc. Nat. Hist.* **3**, 342–343.
- Kramp, P. (1961) Order leptomedusae. *J. Mar. Biol. Assoc. U.K.* **40**, 132–212.
- Shagin, D., E. Barsova, Y. Yanushevich, A. Fradkov, K. Lukyanov, Y. Labas, T. Semenova, J. Ugalde, A. Meyers, J. Nunez, E. Widder, S. Lukyanov and M. Matz (2004) GFP-like proteins as ubiquitous metazoan superfamily: evolution of functional features and structural complexity. *Mol. Biol. Evol.* **21**, 841–850.
- Ward, W., C. W. Cody, R. C. Hart and M. Cormier (1980) Spectrophotometric identity of the energy transfer chromophores in Renilla and Aequorea green-fluorescent proteins. *Photochem. Photobiol.* **31**, 611–615.
- Tsien, R. (1998) The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509–544.
- Altschul, S., T. Madden, A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. Lipman (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.
- Larkin, M., G. Blackshields, N. Brown, R. Chenna, P. McGettigan, H. McWilliam, F. Valentin, I. Wallace, A. Wilm, R. Lopez, J. Thompson, T. Gibson and D. Higgins (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948.
- Ho, S., H. Hunt, R. Horton, J. Pullen and L. Pease (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51–59.
- Dennison, C. and R. Lovrien (1997) Three phase partitioning: concentration and purification of proteins. *Protein Expr. Purif.* **11**, 149–161.

18. Ward, W. W. (2009) Three-phase partitioning for protein purification. *Innovations Pharm. Technol.* **2009**, 28–34.
19. Ward, W. W. (1981) Properties of the Coelenterate green-fluorescent proteins. In *Bioluminescence and Chemiluminescence: Basic Chemistry and Analytical Application*, (Edited by M. DeLuca and W. McElroy), pp. 235–242. Academic Press, New York, NY.
20. Lakowicz, J. (1983) *Principles of Fluorescence Spectroscopy*. Plenum Press, New York, NY.
21. Ormö, M., A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien and S. J. Remington (1996) Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392–1395.
22. Ai, H., J. Henderson, S. Remington and R. Campbell (2006) Directed evolution of a monomeric, bright and photostable version of *Clavularia* cyan fluorescent protein: structural characterization and applications in fluorescence imaging. *Biochem. J.* **400**, 531–540.
23. Swaminathan, R., C. P. Hoang and A. Verkman (1997) Photobleaching recovery and anisotropy decay of green fluorescent protein gfp-s65t in solution and cells: cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion. *Biophys. J.* **72**, 1900–1907.
24. Suhling, K., D. M. Davis and D. Phillips (2002) The influence of solvent viscosity on the fluorescence decay and time-resolved anisotropy of green fluorescent protein. *J. Fluoresc.* **12**, 91–95.
25. Heim, R., A. Cubitt and R. Tsien (1995) Improved green fluorescence. *Nature* **373**, 663–664.
26. Griesbeck, O., G. Baird, R. Campbell, D. Zacharias and R. Tsien (2001) Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *J. Biol. Chem.* **276**, 29188–29194.
27. Day, R. and M. Davidson (2012) Fluorescent proteins for FRET microscopy: monitoring protein interactions in living cells. *BioEssays* **34**, 341–350.
28. Wang, C., D. Yang and M. Wabl (2004) Directed molecular evolution by somatic hypermutation. *Protein Eng. Des. Sel.* **17**, 659–664.
29. Pletneva, N., V. Pletnev, E. Souslova, D. Chudakov, S. Lukyanov, V. Martynov, S. Arhipova, I. Artemyev, A. Wlodawer, Z. Dauter and S. Pletnev (2013) Yellow fluorescent protein phiYFPv (*Phialidium*): structure and structure-based mutagenesis. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1005–1012.
30. Beitz, E. (2000) TEXshade: shading and labeling of multiple sequence alignments using LATEX2 epsilon. *Bioinformatics* **16**, 135–139.