



OPEN

Spectral and structural comparison between bright and dim green fluorescent proteins in *Amphioxus*

SUBJECT AREAS:

BIOLOGICAL
FLUORESCENCE

NANOCRYSTALLOGRAPHY

BIOMIMETICS

MOLECULAR EVOLUTION

Erin K. Bomati¹, Joy E. Haley², Joseph P. Noel³ & Dimitri D. Deheyn¹

¹Marine Biology Research Division, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92037 USA, ²Materials and Manufacturing Directorate, Air Force Research Laboratory, Wright Patterson Air Force Base, OH 45433, ³Howard Hughes Medical Institute, Salk Institute for Biological Studies, Jack H. Skirball Center for Chemical Biology and Proteomics, La Jolla, CA 92037.

Received
21 May 2013Accepted
23 May 2014Published
27 June 2014

Correspondence and requests for materials should be addressed to D.D.D. (ddeheyn@ucsd.edu)

The cephalochordate *Amphioxus* naturally co-expresses fluorescent proteins (FPs) with different brightness, which thus offers the rare opportunity to identify FP molecular feature/s that are associated with greater/lower intensity of fluorescence. Here, we describe the spectral and structural characteristics of green FP (bfloGFPa1) with perfect (100%) quantum efficiency yielding to unprecedentedly-high brightness, and compare them to those of co-expressed bfloGFPc1 showing extremely-dim brightness due to low (0.1%) quantum efficiency. This direct comparison of structure-function relationship indicated that in the bright bfloGFPa1, a Tyrosine (Tyr159) promotes a ring flipping of a Tryptophan (Trp157) that in turn allows a cis-trans transformation of a Proline (Pro55). Consequently, the FP chromophore is pushed up, which comes with a slight tilt and increased stability. FPs are continuously engineered for improved biochemical and/or photonic properties, and this study provides new insight to the challenge of establishing a clear mechanistic understanding between chromophore structural environment and brightness level.

First discovered in 1961 by Shimomura and colleagues in the cnidarian jellyfish *Aequorea victoria*¹, the green fluorescent protein (GFP) and its variants have become the cornerstone of fluorescent protein technologies, exponentially expanding the application of fluorescence spectroscopy and imaging in molecular, cellular and developmental biology, as well as the applied fields of biotechnology and bioengineering^{2,3}.

Critical steps in the wide-spread use of GFP included the elucidation of the three dimensional structure of the wild-type GFP⁴ and the S65T-GFP mutant that resulted in increased fluorescence, photostability and a red shifting the major excitation peak to 488 nm with the peak emission kept at 509 nm⁵. These initial atomic resolution views of GFP highlighted the unusual architecture and key functional groups of this protein family described as 'paint in a can'⁵. GFP folds into an 11-stranded β -barrel with a single helical segment threaded through the center of the barrel, and capped by several loops at either end of the barrel. A sharp turn places strain on three residues within the barrel-encapsulated stretch, which consequently drives cyclization via a dehydration/oxidation mechanism that requires only molecular oxygen⁶⁻⁸. The cyclized chromophore remains covalently attached to the polypeptide chain, and therefore, every cell expressing the GFP gene acquires fluorescence. This genetically encoded autonomy and self-assembly makes the GFP gene a powerful biological tag for the *in vivo* visualization of a wide array of cellular structures and processes^{2,9,10}.

Since the publication of the landmark crystal structures from the Remington and Phillips groups, respectively^{4,5}, more than 250 structures of engineered GFP variants from *Aequorea victoria* have been deposited into the Protein Data Bank (PDB), and are described in many publications (www.rcsb.org)¹¹. This extensive catalog of three-dimensional structures coupled with biochemical characterization has enabled both structure-guided engineering of the wild-type AvGFP and the ability to decipher the effect of random mutations and directed amino acid substitutions on the proteins' optical characteristics¹¹. These engineering efforts have produced GFPs with widely varying solubilities, oligomerization states, pH optima, halide and temperature sensitivities, and a diversity of fluorescence intensity, brightness, absorption and emission spectra (e.g.¹²⁻¹⁴). While these engineering efforts produced impressive improvements in a broad array of biophysical parameters, they were mostly associated with production of green fluorescence, while the production of GFP-like proteins fluorescing in the orange-range of the visible spectrum remaining more challenging¹⁵.



Some research efforts were also developed towards finding fluorescent proteins from other organisms than the *Aequorea* jellyfish, in order to identify whether Nature could have “engineered” GFPs with other structural backbones, or fluorescent proteins with different structure/s, and hopefully with different (and attractive) biophysical and photonic properties. As a result, an increasing number of GFP-like crystal structures from many different species of cnidarians initially, but more recently also from crustaceans, have been deposited in the PBD. In particular, a red fluorescing GFP-like protein, DsRED, was identified in the cnidarian stony coral *Discosoma*^{16–18}, and engineered to yield a set of proteins called mFruits, with colors ranging from honeydew to cherry hues¹⁹. Beyond their importance for expanding the color spectrum of genetically encoded imaging agents, this more phylogenetic approach revealed fascinating FP properties such as kindling (transition from non-fluorescent to fluorescent protein upon intense green light irradiation)²⁰, photo-conversion (changing both excitation and emission spectra upon irradiation with high energy light)^{21–23}, and photo-induced protein cleavage²⁴. Furthermore, while the β -can architecture is conserved for all GFPs examined to date, novel chromophore structures have been discovered in GFP-like proteins from other cnidarians species²⁵. It must be noted, however, that in most species only a few (and mostly one) distinct gene of GFP protein can be found per organism^{26,27}, thus limiting greatly the comparison of optical and biochemical performances from GFPs naturally co-expressed within the same organism.

Quantum yield (QY), also referred to as the quantum efficiency (QE), is the probability that an excitation of the electronic dipole of the chromophore leads to the emission of a photon instead of a heat dissipating transition as relaxation back to the ground state occurs²⁸. QE is a key-factor affecting the GFP “Brightness” defined as the product of the QE and the chromophore’s molar extinction coefficient, the latter representing the extent to which a chemical species absorbs light at a given wavelength. Despite extensive engineering efforts, QE improvements remain difficult for most GFP-like proteins. QEs associated with commercially available engineered-GFPs (possessing emission spectra maximums between 500–510 nm) range from 0.53 (TurboGFP) to 0.91 ($\text{mM} \cdot \text{cm}^{-1}$) (ZsGreen1)³. However, all commercial GFPs exhibit low molar extinction coefficients, resulting in only “modest GFP Brightness” [all expressed in ($\text{mM} \cdot \text{cm}^{-1}$): 37 for TurboGFP, 40 for ZsGreen1, as low as 26 for T-Sapphire, and as high as 41 for Azami Green³, thus improving by only a maximum of 121% the brightness of eGFP⁹.

The need for GFPs with improved brightness becomes apparent as researchers push the limits of signal detection. GFPs are widely used to address protein-protein interactions based on single molecule fluorescence detection, or fluorescence shifts from Förster Resonance Energy Transfer (FRET)^{29–31}. Using these techniques however is associated with a typical 40% loss of input light (photons), as is the case when using eGFP, which can clearly become a limiting

factor for optimal signal detection. Despite the wealth of three-dimensional information available, the specific structural factors modulating GFP QY remain a speculative area of biophotonic research. Ideally, comparing bright and dim GFPs with the same backbone would lead to constructive insights on how to direct engineering for spectral optimization of fluorescent proteins.

We recently identified a family of GFP proteins (bflGFP) in the cephalochordate, *Branchiostoma floridae*³², an invertebrate phylogenetically closest to vertebrates³³. In this particular species, an animal commonly called amphioxus or lancelet, the 16-member family of fluorescent proteins represents the largest set of GFPs yet discovered in a single species. These GFPs group into six clades, each clade possessing distinct fluorescence intensities, extinction coefficients, and absorption profiles, although always emitting light in the green color when collected from the field³⁴, and despite some red fluorescence reported in lancelets by other groups^{35,36}. Accessibility to such widely varying GFPs from a single organism represents a unique opportunity to investigate natural variation within a single species and the evolutionary consequences for properties of the encoded FPs. In a recent study, we described the spectroscopic properties associated with a representative member of each of four FP clades, showing that the fluorescence intensity is particularly high for some bflGFPs, while particularly low for others³⁴. In the present study, we focus our efforts on the structure-function analysis of a brightly fluorescent member, bflGFPa1, and a weakly fluorescent member, bflGFPc1. We present biochemical and spectral characteristics as well as three-dimensional structures derived from protein x-ray crystallography of both proteins, and discuss the structural differences in light of the chromophore environments and the resulting photonic properties of each GFP.

Results

Protein purification and crystallization. Brightly fluorescent bflGFPa1, weakly fluorescent bflGFPc1, and eGFP were expressed in *E. coli* with thrombin cleavable N-terminal hexahistidine tags, and purified via Ni^{2+} -NTA affinity and size exclusion chromatography (SEC). SEC confirmed the monomeric character of eGFP while both bflGFPa1 and bflGFPc1 behaved as dimers (Fig. S1). The *Amphioxus* GFPs’ peak spectral characteristics were (peak \pm FWHM) bflGFPa1 = 497 ± 45 nm and bflGFPc1 = 493 ± 55 nm for absorbance, and bflGFPa1 = 512 ± 62 nm and bflGFPc1 = 521 ± 58 nm for fluorescence emission. The excitation spectrum of bflGFPa1 peaked at 500 nm (FWHM = 38 nm) for emission fixed at 512 nm (Table 1), while the excitation spectrum for bflGFPc1 was below detection limits of the spectrophotometer.

bflGFPc1 and bflGFPa1 crystallized with 8 and 2 molecules in the asymmetric unit, respectively. In both cases, the common oligomeric units were dimers. The crystals were brightly colored (greenish for bflGFPa1 and yellow for bflGFPc1), and retained their color

Table 1 | Photonic properties and pKa values for eGFP, bflGFPa1 and bflGFPc1

	eGFP	bflGFPa1	bflGFPc1
Absorbance Maximum Peak (nm)	488	497	493
FWHM (nm)		45	55
Fluorescence Maximum Excitation Peak (nm)	488	500	n.d.
FWHM (nm)	35–45	38	
Fluorescence Maximum Emission Peak (nm)	508	512	521
FWHM (nm)	30–40	62	58
Extinction Coefficient ($\text{M}^{-1} \text{cm}^{-1}$) (per chain)	$56,600^{70}$	120,900	98,800
Quantum Yield (%)	60 ⁵	104 ± 5	0.15 ± 0.1
Brightness	34	120.9	0.148
pKa	5.65/5.9 ⁵	3.0	n.d.

n.d.: not determined.

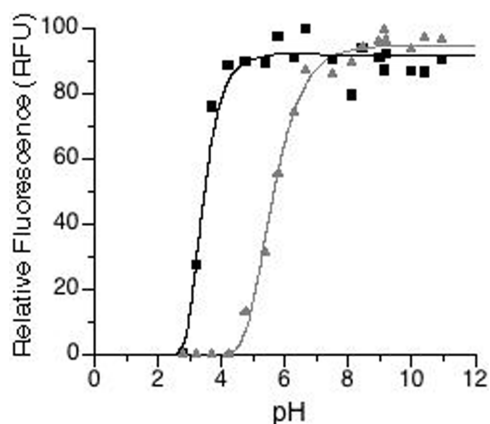


Figure 1 | Change in relative fluorescence intensity (RFU) with pH for eGFP (grey triangles) and bfloGFPa1 (black squares). Raw data (symbols) are fitted with a double-exponential (sigmoidal) dose-response model from which the pKa was calculated.

throughout the duration of the data collection at 100 K. bfloGFPa1 crystals diffracted x-rays to 1.35 Å resolution while bfloGFPc1 crystals yielded measurable diffraction to 1.9 Å resolution. Phasing of bfloGFPc1 (8 molecules per asymmetric unit) via molecular replacement afforded high quality electron density maps for monomers A–D while the electron density for monomers E–G were of a lower quality. In the case of bfloGFPa1 (2 molecules per asymmetric unit), electron density maps for both monomers were of a high quality. Model building and refinement yielded final GFP structures exhibiting electron density for all amino acid residues except residues 1 and 182–187 (bfloGFPa1 monomer A), residues 1 and 219 (bfloGFPa1 monomer B), and residues 1–2 (bfloGFPc1a) (Table S1).

Spectroscopic properties. Molar extinction coefficients measured for bfloGFPa1 and bfloGFPc1 were 120,900 M⁻¹ cm⁻¹ and 98,800 M⁻¹ cm⁻¹, respectively; both coefficients were significantly larger than the extinction coefficients of copGFP (70,000 M⁻¹ cm⁻¹)³⁷ and that the one of eGFP (reported between 55,000–57,000 M⁻¹ cm⁻¹¹⁵ and indeed measured here at 56,000 M⁻¹ cm⁻¹) (Table 1). bfloGFPs' extinction coefficients were significantly greater than those of commonly engineered GFPs^{9,29,38} and similar in magnitude – but not as high- as to the one of the cnidarian *Renilla reniformis* (sea pansy) GFP with a value of 133,000 M⁻¹ cm⁻¹^{39–41}.

The QE of eGFP calculated in our laboratory was 0.65, which was also comparable to the published value of 0.60⁴². In contrast, QEs were drastically different for *Amphioxus* GFPs compared to other GFPs; bfloGFPc1 has a very low QE of 0.0015, while bfloGFPa1 exhibited the maximum possible QE of 1.04 ± 0.05 (Table 1). These spectral properties remained unchanged in His-tagged bfloGFPs.

Due to a very low QE, bfloGFPc1 is quite dim with a measured brightness of 0.148 (mM * cm)⁻¹. In contrast, bfloGFPa1 is exceedingly bright with a measured value of 121 (mM * cm)⁻¹ (Table 1). To the best of our knowledge, this value makes bfloGFPa1 the brightest natural GFP characterized to date, well above commercial GFPs considered to be very bright, including the engineered EmGFP with brightness of 39 (mM * cm)⁻¹⁴², and the recently reported native pmimGFPs from copepod with brightness of 70 (mM * cm)⁻¹⁴³, and the native VFP from coral with brightness of 107 (mM * cm)⁻¹³⁸.

pH-dependence of GFP fluorescent properties. The pK_a could not be determined for the bfloGFPc1 fluorescence due to the barely detectable fluorescence well below the detection limit for conventional spectroscopic instrumentation. As for bfloGFPa1, it exhibited a pK_a of 3.0 for its fluorescence, exhibiting high intensity for pH values ranging from 3.5 to 11 (Fig. 1). Notably, the high intensity for the acidic pH range is unusual compared to other GFPs such as eGFP. For instance, the pK_a for eGFP fluorescence calculated in our laboratory was 5.7, only slightly lower than the previously reported value of 6.0⁴² (Table 1, Fig. 1).

Re-oxidation and refolding. Re-oxidation and refolding rates for bfloGFPa1 were measured and compared to those for eGFP (Table 2). In both fast and slow phases, the rate of reoxidation of the bfloGFPa1 chromophore occurred two times faster than that of the eGFP chromophore. In the case of refolding, the fast phase occurred at approximately the same rate for both bfloGFPa1 and eGFP, while the slower phase occurred three times faster for bfloGFPa1 compared to eGFP (Table 2).

Three-dimensional architecture. Despite biochemical characteristics that were clearly different from one another (see above), we found that bfloGFPa1 and bfloGFPc1 shared essentially identical global tertiary and quaternary structures with a root mean square deviation (rmsd) for backbone atoms of 0.9 Å; for this reason, the results described in this section will use “bfloGFPs” to refer to both bfloGFPa1 and bfloGFPc1. bfloGFPs possess the classic 11-stranded β-barrel structure observed for all GFPs crystallized to date. In bfloGFPs, the barrel is capped by a combination of inter-strand loops at the bottom, and inter-strand loops and helices at the top. Following the central chromophore bearing helix (α1) there is a short helix at the top of the barrel (α2), common to all GFP structures, juxtaposed next to a third short helix (α3) that appears unique to *Amphioxus* and crustacean copepod GFP-structures (pdbid = 2G3O). The crustacean copepod GFP shares up to 35% amino acid sequence identity with bfloGFPs (compared to <19% between cnidarian GFPs and bfloGFPs), and is in fact also the closest GFP-containing evolutionary relative to *Amphioxus*³². GFPs from copepod and amphioxus bear high structural similarity; the rmsd of backbone atoms between bfloGFPa1 and copepod copGFP is 1.1 Å. The most obvious differences appear to be associated with

Table 2 | Kinetics parameters associated with refolding and reoxidation of eGFP and bfloGFPa1 when modeled with a two-phase exponential equation

	Refolding		Reoxidation	
	eGFP	bfloGFPa1 bfloGFPa1	eGFP	bfloGFPa1
Emission (nm)	508	516	508	516
K1 (s ⁻¹)	5.02 × 10 ⁻⁴	6.83 × 10 ⁻⁴	3.83 × 10 ⁻⁴	6.67 × 10 ⁻⁴
K1 _{std. error}	7.20 × 10 ⁻⁶	4.08 × 10 ⁻⁶	1.20 × 10 ⁻⁵	7.33 × 10 ⁻⁶
K2 (s ⁻¹)	7.63 × 10 ⁻³	2.18 × 10 ⁻²	5.07 × 10 ⁻³	1.16 × 10 ⁻²
K2 _{std. error}	2.01 × 10 ⁻⁴	2.69 × 10 ⁻³	7.15 × 10 ⁻⁴	1.72 × 10 ⁻³
T1 _(1/2) (sec)	1,380	1,014	1,812	1,040
T2 _(1/2) (sec)	90.89	31.82	136.70	59.93
r ²	0.9990	0.9992	0.9981	0.9984

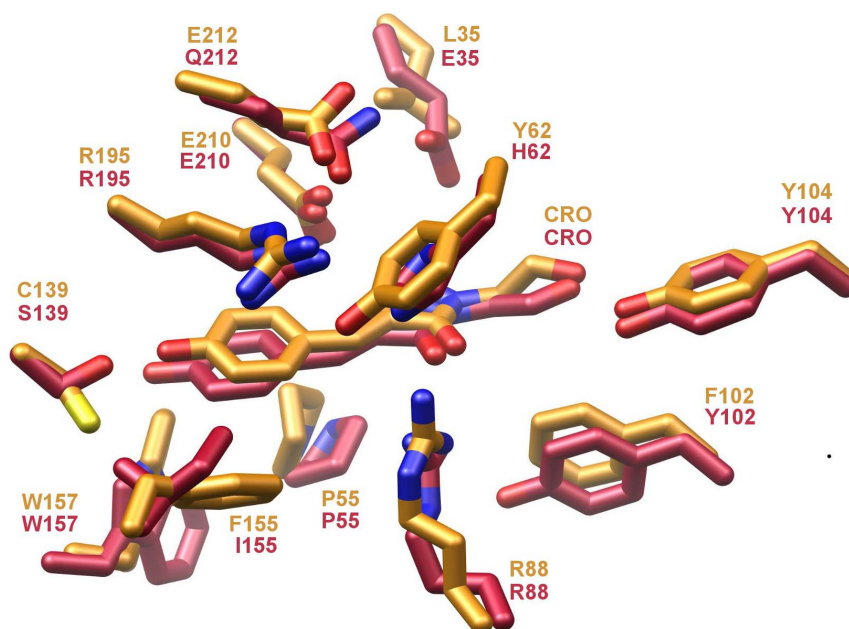


Figure 2 | Overlay of bflGFPa1 and bflGFPc1 chromophore sites. Side chains and chromophores (CRO) are shown in ball-and-stick representation with carbon atoms colored gold (bflGFPa1) and red (bflGFPc1). bflGFPc1 bears the Arg 195 and Glu 210 conserved across all GFPs as well as the Tyr 104 and Arg 88 involved in making H-bonds with the chromophore. As for the Phe 155, Phe 102, and Tyr 62, they were substituted in bflGFPc1 by Ile, Tyr, and His, respectively. The three other residues common to bflGFPa1 and bflGFPc1 were therefore unique to amphioxus GFPs only, while found divergent in copGFP. These residues are Trp 157, Pro 55 and Leu 208, and could play a critical role in the unique biochemical characteristics/differences of the amphioxus GFPs as presented earlier (see also³²). In particular, Trp 157 (Arg in copGFP) and Pro 55 (His in copGFP) form the base of the chromophore binding site where the phenolic ring of the chromophore sits, thus promoting Van der Waals contacts, especially with Pro 55.

the conformation of the loop and helical regions capping the top of the GFP β -barrel (Fig. S2).

Chromophore environment. While globally the architectures of bflGFPa1 and copGFP displayed a high degree of structural similarity, the local environments of the chromophores in both cases were highly divergent, with only 7 out of 19 chromophore-contacting residues strictly conserved (Fig. S3). Arg 195 and Glu 210, conserved across all GFPs described to date, were also present in bflGFPa1, as were Tyr 104 and Arg 88, that latter of which form hydrogen bonds with the chromophore's peptidic and imidazolinone carbonyls, respectively. Phe 155, Phe 102, and Tyr 62, characteristic of copepod GFP, were also observed in amphioxus GFPs. These three aromatic residues contributed to the well-packed hydrophobic surface encapsulating the bflGFPa1 chromophore. However, significant variation in chromophore contacting residues includes Cys 139 (GFPa1) to Thr 138 (copGFP), Val 197 (GFPa1) to Cys 197 (copGFP), Trp 157 (GFPa1) to Arg 156 (copGFP), Ser 141 (GFPa1) to Glu 140 (copGFP), and Pro 55 (GFPa1) to His 54 (copGFP). Some of these residues appear critical for the overall energetic stability of the protein since substitution of one of them using site-specific mutagenesis lead to protein precipitation, especially upon exposure to blue excitation light (see Mutagenesis section in Discussion).

When bflGFPa1 is compared to bflGFPc1, the degree of conservation between chromophore sites was 37% with only 7 out of 19 residues retained. However, only 4 out of 7 residues for each set were common to all three proteins (Fig. 2). Interestingly, we observed that Pro 55 (His in copGFP), conserved in both bflGFPc1 and bflGFPa1, was part of a cis peptide bond in bflGFPa1 and a trans peptide bond in bflGFPc1 suggestive that these two conformationally distinct states may play a critical role in modulating the pronounced differences in brightness between these two otherwise similar *Amphioxus* GFPs.

Likewise, the indole ring of Trp 157, abutting the phenolic oxygen moiety of the chromophore, flips its relative orientation for bflGFPa1 compared to bflGFPc1. In bflGFPc1, the Ne1 nitrogen of the Trp 157 forms a hydrogen bond with the phenolic oxygen while in bflGFPa1 the ring flip negates a similar interaction (Fig. 3). These distinct rotamers for Trp 157 may correlate with what appears to be a well-stabilized edge-to-face interaction of the indole moiety with Tyr 159 in bflGFPa1. In contrast, bflGFPc1 bears a Cys at position 159 (Fig. 3). Finally, the chromophore of bflGFPa1 is tilted 10° relative to the bflGFPc1 chromophore. In total, the changes just described result in significant differences in chromophore cavity volumes between bflGFPa1 and bflGFPc1, the latter

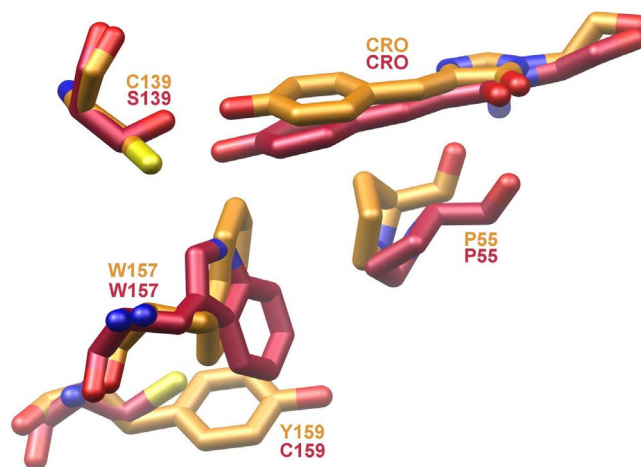


Figure 3 | Ball-and-stick representation of key residues in bflGFPa1 (gold) and bflGFPc1 (red) that appear determinants of chromophore energetic stability.



ition between fluorescent and non-fluorescent states is associated with the combination of both tilting and twisting of the two chromophoric rings relative to one another⁵⁰. Such process takes places without affecting the planarity *per se*, indicating that high fluorescence QE is not necessarily associated with planar chromophoric rings⁵⁰. This could explain the fact that the bflGFPa1 has maximum QE while also have 6.8° coplanar difference between the chromophoric rings. The maximum QE in bflGFPa1 seems to be related instead to a combination of other factors related to the slight 10° tilt of the chromophore when pushed into its pocket (no twisting was detected), and/or to the vibrational freedom of the rings, as shown by analysis of the B-factor (see section below). Comparison to other maximum QE is limited: the Verde fluorescent protein (VFP) from the coral *Cyphastrea microphthalmia* is the only other native GFP reported so far with 100% QE³⁸, although with lesser brightness than bflGFPa1 because of lower extinction coefficient, being 107,000 M⁻¹ cm⁻¹ compare to 120,900 M⁻¹ cm⁻¹ for bflGFPa1. Crystallography data from VFP are currently not available to address similarities or differences with structural features of bflGFPa1.

The crystallographic atomic displacement factor (or B-factor) is a partial measure of this rigidity, using as a proxy the level of flexibility induced by thermal energy. The B-factor is therefore resolution dependent, and when comparing structures resolved at the same crystallographic resolution, the lower the B-factors of a particular groups of atoms, the more rigid the associated structure. We screened the PDB for GFP structures refined to a similar resolution as bflGFPa1 and found 15 of them (Table S2) also studied elsewhere⁵¹. We then analyzed the average temperature factor of the atoms comprising the chromophores of these GFPs and found that the B-factor ranges between 9.5 Å² (2CD1) and 21.1 Å² (2DUG) (Table S2). All these GFPs are reported with QE similar to the original GFP template and thus well lower than 100%, thus reinforcing the observation that the lower the B-factor, the greater the stiffness, and the greater the QE. Indeed, in comparison, the 100% QE bflGFPa1 bears a relatively low average B-factor of 8.7 Å² for the atoms making up its chromophore, suggesting that the chromophore is very rigidly restrained in the chromophore pocket in bflGFPa1. Such trend was in fact observed specifically at, or around, the chromophore *per se* since analyses of the B-factor around residue 60 were about 40 Å² for bflGFPc1 (*viz.* clearly indicative of a high flexibility region), while being <10 Å² for bflGFPa1. The low B-factor for bflGFPa1 clearly indicates that the chromophore and its surrounding have limited structural flexibility (hence high stiffness), which is known widely in the literature to be associated with greater QE (e.g.,^{29,46,52}).

If the B-factor clearly plays an important role driving the QE in GFPs, it is also likely not the only one, and several factors might have to act in synergy for reaching maximal QE. This might be indicated for example by the Glycine-containing mutant 3CB9, which corresponds to a redox sensitive GFP with insertion of an extra amino acid, in this case arginine⁵¹. In this mutant, the QE is reported low despite however particularly low B-factor of 6.2 Å², which is indicative of high stiffness of the chromophore and thus high QE. This suggests that there is another factor that affects QE, which could be the chromophore pocket. Indeed, we believe that the conformational rigidity of the bflGFPa1 chromophore may be in part the result of the smaller volume of its chromophore pocket relative to its very dim cousin bflGFPc1. Nevertheless, the volume of the bflGFPa1 chromophore pocket is approximately 65% smaller than the same pocket calculated for the static x-ray crystal structure of eGFP, and roughly equivalent to the one found in an evolutionarily related GFP, namely copGFP, yet both of these latter GFPs are 5–10 times dimmer than bflGFPa1. This observation indicates that volume of the chromophore pocket alone cannot be the only factor modulating QE and ultimately brightness. A restricted pocket size would likely contribute to maintaining the chromophore in a strained, high-energy con-

formation, thus contributing to an increased efficiency for photon transfer and fluorescence, while also increasing protection of the chromophore from quenching agents^{53,54}.

Through comparisons of GFP structures possessing widely varying spectral properties but isolated from the same organism, we were able to clearly define other structural features not obvious from sequence alignments alone that likely contribute to the modulation of QE and brightness in these naturally evolved FPs. In the *Amphioxus* GFP structures elucidated and described here, it is clear that the chromophore packs against a conserved Pro, Pro 55, that assumes distinct backbone conformational states in each GFP; in the intensely bright GFP, bflGFPa1, the peptide bond, Ile 54 – Pro 55, adopts a *cis* orientation while the same peptide, Ile 54 – Pro 55 adopts a *trans* orientation in the very dim bflGFPc1. These divergent backbone conformations shift the positions of the encapsulated chromophore helical segment, resulting in the formation of a single hydrogen bond between the imidazolinone nitrogen and the backbone carbonyl oxygen of the Pro 55 in bflGFPc1. Such an alteration of the hydrogen bonding patterns of the β-barrel with the chromophore would expectedly lead to change(s) in the chromophores' emission efficiencies and the brightness of the resultant fluorescence¹².

An additional set of changes in the chromophore environments when comparing bflGFPa1 and bflGFPc1 centers around the rotameric state of the chromophore contacting Trp 157. In bflGFPa1, this rotation disrupts a hydrogen bond between the indole NH of Trp 157 and the hydroxyl moiety of Tyr 159 that occurs in bflGFPc1. This structural distortion results in the phenyl portion of Trp 157 abutting the phenolic end of the chromophore, thus increasing the hydrophobicity of the bflGFPa1 chromophore environment compared to that of bflGFPc1. This rotation appears to be stabilized by the edge-to-face interaction between the aromatic moieties of Trp 157 and Tyr 159. Without an actual experimentally determined three-dimensional structure, this type of second tier interaction can be challenging to identify unequivocally.

Currently, genetic engineering of GFPs is focusing more and more on improving the QEs but remains an ongoing challenge compared to modifying the biochemical and/or biophotonic properties of FPs^{3,13,55}. Therefore, bflGFPa1 may represent an alternative evolutionarily optimized GFP with which to pursue the directed evolution of other desirable parameters, including oligomeric state, greater tolerance to acidic pH, better resistance to bleaching, faster folding and chromophore maturation, large extinction coefficient, and broad excitation and emission spectra. Given its perfect quantum efficiency, its broad fluorescence pKa range, and its relatively fast folding and chromophore maturation rates, *Amphioxus* bflGFPa1 appears to be an ideal candidate for future applications and engineering efforts. This has been materialized already by the use of bflGFPa1 in bioassays where increased brightness and biochemical stability were key in providing signal not otherwise available using conventional FPs^{56,57}. In addition, a GFP from *Branchiostoma lanceolatum*, the amphioxus species from Europe, is now commercially available (under the name of lanGFP) and showing about 4x-increased brightness compared to eGFP⁵⁸, thus comparable to bflGFPa1. This strongly suggests that amphioxus GFPs indeed hold attractive promises for future use in an extended range of applications than currently available.

We used conventional techniques of mutagenesis of critical residues in a first attempt to demonstrate the role of key residues from the chromophore pocket that are associated with high versus low brightness or quantum efficiency. We performed site directed amino acid substitutions on bflGFPa1 based on differences with bflGFPc1 and copGFP. bflGFP mutants were generated using the QuikChange (Stratagene, San Diego, CA) PCR-based method. Mutant enzymes were expressed and purified as described in this study for wild-type bflGFP. In addition to wtGFP controls, we



completed five different single mutations chosen to provide further understanding of the relationship amongst the diverse GFPs within amphioxus³⁴. We then attempted to perform spectroscopic and biochemical comparative analysis on the five mutants using the exact same proteomic and spectroscopic protocols as used for bfoGFPa1. The mutants were P55H and C139S (P55 and C139 are both unique to GFPa1), F155L and Y62H (F155 and Y62 are common to GFPa1 and copGFP and contribute to chromophore encapsulation), and R195A (common to all GFPs and responsible for hydrogen bonding in the chromophore cavity). These mutations thus involve residues with significant role in providing energetic stability to the chromophore (Fig. 4).

The five bfoGFPs mutants were successfully expressed and exhibiting strong fluorescence, except for C139S, F155L and R195A that showed relatively lower fluorescence intensity, yet always with a spectrum similar to that of bfoGFPa1. The absorbance spectrum was also measurable for each mutant, and remained similar amongst all bfoGFPs, yet showing relatively greater values for F155L. Complete and detailed biochemical and spectroscopic characterization of the mutants could however not be performed because all mutants were unstable, constantly precipitating out of solution, which was otherwise not observed for the wtGFP control, or when maintained in darkness. Consequently, determining accurate protein concentration necessary for calculation of the extinction coefficient, for performing biochemical experiments, and for interpretation of the spectroscopic measurements from the bfoGFPa1 mutants was not possible using the current protocols. However, these data suggests that the five mutations we performed in the chromophore pocket do not seem to qualitatively affect the spectral characteristics of bfoGFPs, while clearly indicating that each of the substituted residues (proposed to each have a critical function in the chromophore pocket) plays a key role in contributing to energetic stability of the whole protein.

Amphioxus contains 16 bfoGFPs organized in six clades³⁴, each showing various substitutions of one (or more) of the residues that we experimentally mutated in this study. The co-occurrence of many different GFPs in one single organism therefore indicates that the unbalance in energy due to substitution of the residues considered here must be compensated in the natural system by additional substitutions in other areas of the protein, in order to provide protein stability. This is clearly a possible scenario considering that protein sequence identity amongst clades varies from 49 to 65%³⁴, and that FPs in general appear to have distinct regions (the relatively conserved central chromophore region versus the N- and C- terminal variable regions) with divergent evolutions and different molecular functions⁵⁹. At this stage, genetic engineering of bfoGFPa1 appears attractive, yet requiring the screening of tandem substitutions and/or alternative protocols in order to preserve stability of the protein in solution^{12,13}. Performing mutations of residues other than the five ones presented here would be key in future engineering studies, especially considering that mutations leading to other colors of fluorescence are different from the ones we tested⁹, thus giving the prospect to preserve maximal quantum efficiency for engineered bfoGFPa1 with different emission spectra.

Methods

Cloning, protein expression, and purification. *Amphioxus* bfoGFPa1 and bfoGFPc1 were cloned and purified as previously reported³⁴. eGFP and YFP were kindly provided by Air Force Research Laboratories (WPAFB, Ohio), sub-cloned into pET24b, and purified as described for *Amphioxus* GFPs³⁴. Spectral characteristics (absorbance and fluorescence) of all these fluorescent proteins were measured using a spectrophotometer SpectraMax M2 (Molecular Devices, Sunnyvale, CA) with complete scanning of the spectral range (comprise between 400–800 nm) but expressed here as peak and full width at half maximum (FWHM) of each spectrum.

Protein concentration and extinction coefficient calculation. Protein concentrations were calculated using the extinction coefficient of the chromophore after denaturation in 0.1 N NaOH (44,000 M⁻¹ cm⁻¹ at 446 nm)³⁹. Absorbances of bfoGFPa1, bfoGFPc1 and eGFP were measured with the spectrophotometer and

extinction coefficients calculated according to the Beer Lambert law, $A = \epsilon \cdot l \cdot c$, where “A” is absorption of a given wavelength of light, “ ϵ ” is the molar extinction coefficient of a certain species, “l” is the path-length, and “c” is the concentration of that given sample.

Quantum efficiency calculation. Fluorescence quantum efficiency (QE) of bfoGFPa1 and bfoGFPc1 was measured from six independent replicate expression batches, with and without His-tag, using the method of relative actinometry. Emission of the GFPs was then expressed relative to a known standard with the same absorption at the excitation wavelength of 450 nm, keeping all instrumental conditions identical⁶⁰. Here, fluorescein in 0.1 N NaOH was used as a standard since its QY is 0.90 at 25 °C⁶⁰, while eGFP was used to inter-calibrate our experimental setup and measurements in comparison to published values for eGFP. All samples were maintained in a dilute state adjusted to an absorbance of 0.11 at 450 nm for accurate comparisons. The area under the emission curve extending from 460 to 800 nm was then integrated and the following formula used for calculating the quantum yield:

$$\Phi_{unk} = \Phi_{ref} \frac{A_{ref} \eta_{unk}^2 \Delta A_{unk}}{A_{unk} \eta_{ref}^2 \Delta A_{ref}}$$

where “F” is the fluorescence quantum yield, “A” is the absorbance of the unknown and standard at the exciting wavelength, “h” is the refractive index of the solvent, and “DA” is the area of the 460–800 nm emission. The refractive index values were $h = 1.3576$ for the 0.1 N NaOH standard, and $h = 1.5442$ for the 50 mM Tris-HCl (pH 8), 400 mM NaCl sample buffer, which is the value for a solution in which NaCl is the predominant analyte⁶¹.

GFP refolding and reoxidation. Refolding and re-oxidation experiments were performed as previously described⁶². Briefly, pure GFP in 50 mM Tris-HCl, pH 8, 400 mM NaCl, 1 mM DTT was diluted to 2 mg/ml in 8 M Urea, 1 mM DTT for refolding experiments, and 8 M Urea, 1 mM DTT, 5 mM Dithionite for re-oxidation experiments. The samples were heated to 95 °C for 5 min, cooled to room temperature (~23 °C), and diluted 1 : 100 in renaturation buffer (35 mM KCl, 2 mM MgCl₂, 50 mM Tris-HCl, pH 8, and 1 mM DTT). The SpectraMax M2 plate reader equipped with a 495 nm excitation filter was used to follow refolding and re-oxidation, using as a proxy the fluorescence intensity of the GFPs. Samples were excited at 475 nm and fluorescence emission recorded at 508 nm (eGFP) or 516 nm (bfoGFPa1) every second for 3,000 sec (50 min). Both refolding and re-oxidation curves were modeled employing a two-phase exponential equation using Prism 4.00 (GraphPad). The equation was as follow: $Y = P + C_F \cdot \exp(-K_F \cdot X) + C_S \cdot \exp(-K_S \cdot X)$, where “P” is the Plateau value reached at infinite time for “X”, “C_F” and “C_S” are the two time constants for the Fast and Slow half-life, respectively, and “K_F” and “K_S” are the rate constants for the Fast and Slow half-life, respectively. The Fast and Slow half-lives were then computed as $\ln(2)/K_F$ and $\ln(2)/K_S$, respectively. Parameters of the equations were estimated by iteration from the raw values of fluorescence, following the nonlinear estimation method of Quasi-Newton, while the resulting best fit model is indicated by the R corrected for non-linear systems⁶³.

pH titration. *Amphioxus* GFP fluorescence at various pH values was evaluated using citric acid - sodium citrate (pH 3–5), sodium phosphate (pH 5–6), Tris-HCl - Tris-Base (pH 6–9), and glycine - NaOH (pH 9–12). Each sample consisted of concentrated GFP in a weakly buffered solution (~5 mM) diluted approximately 200-fold (to 50 nM final concentration) into 50 mM buffer at a given pH. The SpectraMax M2 plate reader with a 495 nm cut-off filter was used to record fluorescence emission at 508 (eGFP) and 516 (bfoGFPa1) upon excitation at 475 nm. Absorbance spectra of each sample were subsequently recorded and used as internal controls for calibration of protein concentrations.

Crystallization. bfoGFPc1 crystals were grown overnight by vapor diffusion at 4 °C in 2.0 μL drops, consisting of 1 μL crystallization reservoir [28% (w/v) polyethylene glycol 8,000, 1 M NaCl, 100 mM HEPES-Na+ (pH 7.5)], and 1 μL protein solution [419 μM bfoGFPc1]. Diffraction data were collected at the Berkeley Advanced Light Source (ALS) synchrotron beamline 8.2.2 on a Quantum Q315 CCD detector. bfoGFPc1 crystallized in spacegroup C2, $a = 158.76 \text{ \AA}$, $b = 130.46 \text{ \AA}$, $c = 106.33 \text{ \AA}$, $a = g = 90^\circ$, $b = 128.39^\circ$ with eight monomers in the asymmetric unit. Data were indexed, integrated, and scaled to 1.95 Å with HKL2000⁶⁴. bfoGFPa1 crystals were grown overnight by vapor diffusion at 4 °C in 2.0 μL drops, consisting of 1 μL crystallization reservoir [30% (w/v) polyethylene glycol 4,000, 3% (v/v) isopropanol, 100 mM citric acid - sodium citrate (pH 5.5)], and 1 μL protein solution [336 μM bfoGFPa1]. Diffraction data were collected at the ALS synchrotron beamline 8.2.1 on a Quantum Q315 CCD detector. GFPa1 crystallized in spacegroup C222(1), $a = 59.25 \text{ \AA}$, $b = 125.57 \text{ \AA}$, $c = 106.46 \text{ \AA}$, $a = b = g = 90^\circ$ with two monomers in the asymmetric unit. Data were indexed, integrated, and scaled to 1.35 Å with HKL2000⁶⁴. Structure elucidation process is described in supplementary information.

Structure elucidation. Phase determination of bfoGFPc1 was accomplished via molecular replacement (MR) using the program Phaser⁶⁵. The MR search model was a mixed model⁶⁶ based on the structure of copepod GFP (2G3O). MR phases were used for manual model building of the bfoGFPc1 tertiary structure in Coot⁶⁷. Iterative stages of building and refinement were carried out using Coot and CNS⁶⁸, respectively. Refinement was completed imposing restrained non-crystallographic symmetry between the 8 GFP monomers. The final structure was evaluated with



PROCHECK⁶⁹. The bfoGFPc1 structure had 92.6% and 7.4% of residues in the most favored and allowed regions of the Ramachandran plot, respectively. The final structural coordinates and structure factors were deposited to the Protein Data Bank under PDB ID 3GJV.

Phase determination of bfoGFPa1 was accomplished via molecular replacement (MR) using the program Phaser. The MR search model was the refined structure of bfoGFPc1. MR phases were used for manual model building of the bfoGFPa1 tertiary structure in Coot⁶⁷. Iterative stages of building and refinement were carried out using Coot and CNS⁶⁸, respectively. The final structure was evaluated with PROCHECK⁶⁹. The bfoGFPa1 structure had 91.5% and 8.5% of residues in the most favored and allowed regions of the Ramachandran plot, respectively. The final structural coordinates and structure factors were deposited to the Protein Data Bank under PDB ID 3GIH.

- Shimomura, O., Johnson, F. H. & Saiga, Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusa, *Aequorea*. *J. Cell. Comp. Physiol.* **59**, 223–239 (1962).
- Mocz, G. Fluorescent proteins and their use in marine biosciences, biotechnology, and proteomics. *Mar. Biotechnol.* **9**, 305–328 (2007).
- Shaner, N. C., Patterson, G. H. & Davidson, M. W. Advances in fluorescent protein technology. *J. Cell Sci.* **120**, 4247–4260 (2007).
- Ormo, M. *et al.* Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392–1395 (1996).
- Yang, F., Moss, L. G. & Phillips, G. N. The molecular structure of green fluorescent protein. *Nat. Biotechnol.* **14**, 1246–1251 (1996).
- Heim, R., Cubitt, A. B. & Tsien, R. Y. Improved Green Fluorescence. *Nature* **373**, 663–664 (1995).
- Heim, R., Prasher, D. C. & Tsien, R. Y. Wavelength mutations and post-translational autooxidation of green fluorescent protein. *P. Natl. Acad. Sci. USA* **91**, 12501–12504 (1994).
- Rosenow, M. A., Patel, H. N. & Wachter, R. M. Oxidative chemistry in the GFP active site leads to covalent cross-linking of a modified leucine side chain with a histidine imidazole: Implications for the mechanism of chromophore formation. *Biochemistry* **44**, 8303–8311 (2005).
- Stepanenko, O. V., Verkhusha, V. V., Kuznetsova, I. M., Uversky, V. N. & Turoverov, K. K. Fluorescent proteins as biomarkers and biosensors: Throwing color lights on molecular and cellular processes. *Curr. Protein Pept. Sci.* **9**, 338–369 (2008).
- Chudakov, D. M., Lukyanov, S. & Lukyanov, K. A. Fluorescent proteins as a toolkit for in vivo imaging. *Trends Biotechnol.* **23**, 605–613 (2005).
- Ong, W. J. H. *et al.* Function and structure of GFP-like proteins in the protein data bank. *Mol. Biosyst.* **7**, 984–992 (2011).
- Barondeau, D. P., Kassmann, C. J., Tainer, J. A. & Getzoff, E. D. Understanding GFP chromophore biosynthesis: Controlling backbone cyclization and modifying post-translational chemistry. *Biochemistry* **44**, 1960–1970 (2005).
- Nienhaus, G. U. & Wiedenmann, J. Structure, dynamics and optical properties of fluorescent proteins: Perspectives for marker development. *ChemPhysChem* **10**, 1369–1379 (2009).
- Tsien, R. Y. The green fluorescent protein. *Ann. Rev. Biochem.* **67**, 509–544 (1998).
- Mishin, A. S. *et al.* The first mutant of the *Aequorea victoria* green fluorescent protein that forms a red chromophore. *Biochemistry* **47**, 4666–4673 (2008).
- Baird, G. S., Zacharias, D. A. & Tsien, R. Y. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proc. Natl. Acad. Sci. USA* **97**, 11984–11989 (2000).
- Gross, L. A., Baird, G. S., Hoffman, R. C., Baldrige, K. K. & Tsien, R. Y. The structure of the chromophore within DsRed, a red fluorescent protein from coral. *Proc. Natl. Acad. Sci. USA* **97**, 11990–11995 (2000).
- Wall, M. A., Socolich, M. & Ranganathan, R. The structural basis for red fluorescence in the tetrameric GFP homolog DsRed. *Nat. Struct. Biol.* **7**, 1133–1138 (2000).
- Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp red fluorescent protein. *Nat. Biotechnol.* **22**, 1567–1572 (2004).
- Quillin, M. L. *et al.* Kindling fluorescent protein from *Anemonia sulcata*: Dark-state structure at 1.38 angstrom resolution. *Biochemistry* **44**, 5774–5787 (2005).
- Andresen, M. *et al.* Structure and mechanism of the reversible photoswitch of a fluorescent protein. *Proc. Natl. Acad. Sci. USA* **102**, 13070–13074 (2005).
- Mizuno, H. *et al.* Photo-induced peptide cleavage in the green-to-red conversion of a fluorescent protein. *Mol. Cell* **12**, 1051–1058 (2003).
- Wiedenmann, J. *et al.* EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. *Proc. Natl. Acad. Sci. USA* **101**, 15905–15910 (2004).
- Nienhaus, K., Nienhaus, G. U., Wiedenmann, J. & Nar, H. Structural basis for photo-induced protein cleavage and green-to-red conversion of fluorescent protein EosFP. *Proc. Natl. Acad. Sci. USA* **102**, 9156–9159 (2005).
- Remington, S. J. *et al.* zFP538, a yellow-fluorescent protein from *Zoanthus*, contains a novel three-ring chromophore. *Biochemistry* **44**, 202–212 (2005).
- Matz, M. V. *et al.* Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat. Biotechnol.* **17**, 969–973 (1999).
- Kelmanson, I. V. & Matz, M. V. Molecular basis and evolutionary origins of color diversity in great star coral *Montastraea cavernosa* (Scleractinia: Faviida). *Mol. Biol. Evol.* **20**, 1125–1133 (2003).
- Zhang, J. The colorful journey of green fluorescent protein. *ACS Chem. Biol.* **4**, 85–88 (2009).
- Heim, R. & Tsien, R. Y. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Biol.* **6**, 178–182 (1996).
- Nguyen, A. W. & Daugherty, P. S. Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nat. Biotechnol.* **23**, 355–360 (2005).
- Grohmann, D., Werner, F. & Tinnefeld, P. Making connections - strategies for single molecule fluorescence biophysics. *Curr. Opin. Chem. Biol.* **17**, 691–698, doi:10.1016/J.Cbpa.2013.05.020 (2013).
- Deheyn, D. D. *et al.* Endogenous green fluorescent protein (GFP) in amphioxus. *Biol. Bull.* **213**, 95–100 (2007).
- Holland, L. Z. *et al.* The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res.* **18**, 1100–1111 (2008).
- Bomati, E. K., Manning, G. & Deheyn, D. D. Amphioxus encodes the largest known family of green fluorescent proteins, which have diversified into distinct functional classes. *BMC Evol. Biol.* **9**, 1–11, doi:10.1186/1471-2148-9-77 (2009).
- Baumann, D. *et al.* A family of GFP-like proteins with different spectral properties in lancelet *Branchiostoma floridae*. *Biol. Direct* **3**, 1–12, doi:10.1186/1745-6150-3-28 (2008).
- Pletnev, V. Z. *et al.* Structure of the red fluorescent protein from a lancelet (*Branchiostoma lanceolatum*): a novel GYG chromophore covalently bound to a nearby tyrosine. *Acta Crystallogr. D* **69**, 1850–1860, doi:10.1107/S0907444913015424 (2013).
- Shagin, D. A. *et al.* GFP-like proteins as ubiquitous metazoan superfamily: Evolution of functional features and structural complexity. *Mol. Biol. Evol.* **21**, 841–850 (2004).
- Ilgan, R. P. *et al.* A new bright green-emitting fluorescent protein - engineered monomeric and dimeric forms. *FEBS J.* **277**, 1967–1978 (2010).
- Ward, W. W. in *Bioluminescence and chemiluminescence: Basic chemistry and analytical applications* (eds DeLuca, M. & McElroy, W. D.) pp. 235–242 (Academic Press, Inc., 1981).
- Ward, W. W., Cody, C. W., Hart, R. C. & Cormier, M. J. Spectrophotometric identity of the energy-transfer chromophores in *Renilla* and *Aequorea* green fluorescent proteins. *Photochem. Photobiol.* **31**, 611–615 (1980).
- Ward, W. W. & Cormier, M. J. Energy-transfer protein in coelenterate bioluminescence - Characterization of the *Renilla* green fluorescent protein. *J. Biol. Chem.* **254**, 781–788 (1979).
- Shaner, N. C., Steinbach, P. A. & Tsien, R. Y. A guide to choosing fluorescent proteins. *Nat. Methods* **2**, 905–909 (2005).
- Hunt, M. E., Scherrer, M. P., Ferrari, F. D. & Matz, M. V. Very bright green fluorescent proteins from the Pontellid Copepod *Pontella mimocerami*. *PLoS One* **5**, e11517 doi:10.1371/journal.pone.0011517 (2010).
- Laurie, A. T. R. & Jackson, R. M. Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics* **21**, 1908–1916 (2005).
- Rhodes, G. *Crystallography made crystal clear: a guide for users of macromolecular models*. 3rd edn, (Elsevier/Academic Press, 2006).
- Krasnenko, V., Tkaczyk, A. H., Tkaczyk, E. R., Farkas, D. & Mauring, K. Vibrations-determined properties of green fluorescent protein. *Biopolymers* **78**, 140–146 (2005).
- Voityuk, A. A., Michel-Beyerle, M. E. & Rosch, N. Structure and rotation barriers for ground and excited states of the isolated chromophore of the green fluorescent protein. *Chem. Phys. Lett.* **296**, 269–276 (1998).
- Rick, S. W. & Cachau, R. E. The nonplanarity of the peptide group: Molecular dynamics simulations with a polarizable two-state model for the peptide bond. *J. Chem. Phys.* **112**, 5230–5241 (2000).
- Nijegorodov, N., Luhanga, P. V. C., Nkoma, J. S. & Winkoun, D. P. The influence of planarity, rigidity and internal heavy atom upon fluorescence parameters and the intersystem crossing rate constant in molecules with the biphenyl basis. *Spectrochim. Acta A* **64**, 1–5 (2006).
- Brakemann, T. *et al.* Molecular Basis of the Light-driven Switching of the Photochromic Fluorescent Protein Padron. *J. Biol. Chem.* **285**, 14603–14609 (2010).
- Lohman, J. R. & Remington, S. J. Development of a family of redox-sensitive green fluorescent protein indicators for use in relatively oxidizing subcellular environments. *Biochemistry* **47**, 8678–8688, doi:10.1021/Bi800498g (2008).
- Hanson, G. T. *et al.* Green fluorescent protein variants as ratiometric dual emission pH sensors. 1. Structural characterization and preliminary application. *Biochemistry* **41**, 15477–15488, doi:10.1021/Bi026609p (2002).
- Agmon, N. Proton pathways in green fluorescence protein. *Biophys. J.* **88**, 2452–2461 (2005).
- Phillips, G. N. Structure and dynamics of green fluorescent protein. *Curr. Opin. Struct. Biol.* **7**, 821–827 (1997).
- Wachter, R. M. The family of GFP-like proteins: Structure, function, photophysics and biosensor applications. Introduction and perspective. *Photochem. Photobiol.* **82**, 339–344 (2006).
- Harper, J. C. *et al.* Orthogonal Cell-Based Biosensing: Fluorescent, Electrochemical, and Colorimetric Detection with Silica-Immobilized Cellular



- Communities Integrated with an ITO-Glass/Plastic Laminate Cartridge. *Small* **8**, 2743–2751, doi:10.1002/Sml.201200343 (2012).
57. Eby, D. M. *et al.* Bacterial Sunscreen: Layer-by-Layer Deposition of UV-Absorbing Polymers on Whole-Cell Biosensors. *Langmuir* **28**, 10521–10527 (2012).
 58. Shaner, N. C. *et al.* A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nat. Methods* **10**, 407–409, doi:10.1038/Nmeth.2413 (2013).
 59. Gruber, D. F. *et al.* Novel Internal Regions of Fluorescent Proteins Undergo Divergent Evolutionary Patterns. *Mol. Biol. Evol.* **26**, 2841–2848 (2009).
 60. Demas, J. N. & Crosby, G. a. Measurement of photoluminescence quantum yields - Review. *J. Phys. Chem.* **75**, 991–1024, doi:10.1021/j100678a001 (1971).
 61. Lide, D. R. in *CRC Handbook of Chemistry and Physics* (CRC Press, Boca Raton, FL, 1990).
 62. Reid, B. G. & Flynn, G. C. Chromophore formation in green fluorescent protein. *Biochemistry* **36**, 6786–6791 (1997).
 63. Zar, J. H. *Biostatistical analysis*. (Prentice-Hall, 1996).
 64. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Macromol. Crystallogr. A* **276**, 307–326 (1997).
 65. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
 66. Lunin, V. Y. *et al.* Phase improvement in protein crystallography using a mixed electron-density model. *Acta Crystallogr. A* **41**, 166–171 (1985).
 67. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
 68. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
 69. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486 (1996).
 70. Patterson, G., Day, R. N. & Piston, D. Fluorescent protein spectra. *J. Cell Sci.* **114**, 837–838 (2001).

Acknowledgments

We thank Dr. Michael Latz (SIO) for providing access to additional laboratory space and facilities, and Dr. Evelien DeMeulenaere (KUL, Belgium) for comments on the manuscript. Research supported by the Air Force Office of Scientific Research, grant # AFOSR FA9550-07-1-0027 (to D.D.D.).

Author contributions

D.D.D. designed research; E.K.B. and J.E.H. performed research; J.P.N. provided support for crystallization and x-ray data collection; E.K.B. and D.D.D. wrote the paper; all authors read and edited the manuscript. The authors declare no conflict of interest.

Additional information

Crystal structures deposition Structures of bflGFPa1 and bflGFPc1 were deposited at the RCSB Protein data bank under “Crystal structure of amphioxus green fluorescent protein, GFPa1, and GFPc1”, with assigned RCSB ID code rcsb061402 and rcsb061411, respectively.

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Bomati, E.K., Haley, J.E., Noel, J.P. & Deheyn, D.D. Spectral and structural comparison between bright and dim green fluorescent proteins in *Amphioxus*. *Sci. Rep.* **4**, 5469; DOI:10.1038/srep05469 (2014).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>