A photoactivatable green-fluorescent protein from the phylum Ctenophora

Steven H. D. Haddock¹,*, Nadia Mastroianni²,³ and Lynne M. Christianson¹

¹Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA 95039, USA
²Axcam Spa, San Raffaele, Biomedical Science Park, via Olgetra 58, 20132 Milan, Italy
³Farmacia Posca, Via Cadorna 30, 22060 Carugo (CO), Italy

Genes for the family of green-fluorescent proteins (GFPs) have been found in more than 100 species of animals, with six or more copies producing a variety of colours. Thus far, however, these species have all been within three phyla: Cnidaria, Arthropoda and Chordata. We have discovered GFP-type fluorescent proteins in the phylum Ctenophora, the comb jellies. The ctenophore proteins share the xYG chromophore motif of all other characterized GFP-type proteins. These proteins exhibit the uncommon property of reversible photoactivation, in which fluorescent emission becomes brighter upon exposure to light, then gradually decays to a non-fluorescent state. In addition to providing potentially useful optical probes with novel properties, finding a fluorescent protein in one of the earliest diverging metazoans adds further support to the possibility that these genes are likely to occur throughout animals.

Keywords: comb jelly; GFP; bioluminescence; fluorescence; photoactivation

1. INTRODUCTION

The ctenophores are a phylum of gelatinous marine plankton found throughout the oceans, although most species are rarely collected owing to their fragility. These comb jellies are considered to be one of the earliest-diverging metazoan groups, along with sponges and cnidarians (Halanych 2004), and recent phylogenomic evidence suggests that they might even be the sister group to all other metazoans (Dunn et al. 2008). Ctenophores are well known for their bioluminescence (Haddock & Case 1999), which is produced by coleterazine bound to endogenous calcium-activated photoproteins. To date, however, they have not been known to contain any notably fluorescent compounds.

Fluorescence is a characteristic of many biological molecules, including chlorophyll, chitin and phycobiliprotein, and some of the most prominent naturally fluorescent molecules are those of the green-fluorescent protein (GFP) superfamily. These proteins were originally discovered and characterized in the bioluminescent hydrozoan Aequorea victoria (Prasher et al. 1992; Shimomura 2005), in which endogenous bioluminescence excites GFP fluorescence, altering the ultimate emission spectrum of the light. Since that original discovery, fluorescent proteins (FPs) have been cloned from other hydrozoans and many anthozoans (e.g. Labas et al. 2002), including luminescent octocorals (e.g. Renilla, Phlaosmar), non-luminescent anemones (e.g. Anemonea, Cerianthus) and soft and hard corals (many, including Corynactis, Montastrea, reviewed in Alieva et al. 2008). All of the known fluorescent-protein-bearing organisms were in the phylum Cnidaria until the surprising discovery (Shagin et al. 2004) of the genes for homologous FPs in non-luminous planktonic copepods (Crustacea, Arthropoda). Most recently, Dehey et al. (2007) found FPs in a third phylum by searching the genome of the lancelet chordate Branchiostoma floridae, which, like many of the anthozoans and copepods, is not bioluminescent. Given the evolutionary divergence between these three phyla—Chordata, Arthropoda, and Cnidaria—it might be expected that additional FPs would be found in additional metazoan taxa, as previously suggested (Shagin et al. 2004).

FPs, both natural and engineered, demonstrate a wide range of biophysical properties. They can have emission spectra from cyan at less than 470 nm (Chudakov et al. 2004) to red at greater than 600 nm peak emission (Matz et al. 1999; Wiedenmann et al. 2002; Schnitzler et al. 2008). They can also act as chromoproteins (Lukyanov et al. 2000), which absorb light to cause a visible colour without producing appreciable fluorescence. Although chromophore synthesis occurs autocatalytically within the folding polypeptide, maturation times for the chromophore, especially in the long-wave-emitting variants, can be very different. Maturation can also require a specific condition, such as oxygen, temperature or illumination by light of an appropriate wavelength (Lukyanov et al. 2005). Photoconversion processes have been found to trigger green-to-red maturation (Ando et al. 2002), which can be an important natural process (Leutenegger et al. 2007) and may even be an intrinsic property of most FPs (Bogdanov et al. 2009), although perhaps not under conditions found in nature. Less commonly, photoactivation can trigger transitions between non-fluorescent and fluorescent forms (Lukyanov et al. 2005).

* Author for correspondence (haddock@mbari.org).

2. MATERIAL AND METHODS

(a) Cloning and expression

Total RNA was isolated from frozen Haecella beehleri tissue using the RNeasy Kit (Qiagen). cDNA was generated from total RNA with the Superscript III First Strand Synthesis Kit for RT-PCR (Invitrogen) following the manufacturer’s instructions.

Ctenophore cDNA synthesis was performed using the SuperScript System (Invitrogen) following the manufacturer’s instructions. Because there was a small amount of total RNA as starting material (approx. 100 ng), long-distance PCR was used to enhance yield (Barnes 1994; Cheng et al. 1994). Primer sequences were as follows: 5′-aagggcttctggcagtgctggttccttcgag-3′ and 5′-aagggcggccccctcaggggtgag-3′, with SaI and a NdeI restriction sites for subsequent cloning into the chosen expression vector. Amplified cDNAs were digested, selected by fractionation with column chromatography and ligated into pSPORT 1/SalI–NdeI vector (Invitrogen).

Escherichia coli BL21-Gold (DE3) cells (Stratagene) were transformed with the ligation mixture by electroporation. Transformants were plated on LB agar plates containing 100 μg ml−1 ampicillin and grown at 37°C. After overnight colony growth, induction was obtained by adding 1 mM IPTG and incubating for 3 h at RT. About 1.5 × 10⁶ recombinant colonies were screened visually using a fluorescence microscope and a GFP filter cube for excitation/emission. One fluorescent clone was identified, picked, plated on a new LB agar plate and subjected to a second screening. Single positive colonies were isolated and the purified plasmids were sequenced using an ABI PRISM 3100 Gene Analyser (Applied Biosystems).

(b) Isolation of FP genes

FP genes were PCR-amplified from cDNA with primers 5′-atggagtttgagcctgaattcttcaataaa-3′ and 3′-cgccctgcggcggccgac-5′ using Easy-A Polymerase (Stratagene) with the following cycling parameters: one cycle of 95°C for 2 min; 35 cycles of 95°C for 40 s, 50°C for 30 s and 72°C for 1 min; one cycle of 72°C for 7 min.

A single band of approximately 800 bp was excised from a 1% agarose gel and purified using Ultrafree DNA Gel Extraction Kit (Millipore), cloned into the pCR2.1 vector from the TOPO-TA Cloning Kit (Invitrogen) and used to transform TOP10 cells (Invitrogen). Clones were sequenced using the ABI Big Dye Terminator V3.1 sequencing kit and an ABI 3100 DNA Genetic Analyzer. Sequences were analysed using CodonCode Aligner.

(c) Cloning full-length sequence with RACE

Using the cDNA prepared from a second specimen, the 5′ and 3′ ends were determined by RACE using the SMART RACE cDNA Amplification Kit (Clontech). RACE-ready cDNA was prepared following the manufacturer’s instructions. The 5′ end was generated using primer 5′-gtttaaagtgggatggcgggagtcgtc-3′ and the 3′ end was generated using primer 5′-gcttctgctgcctagacgacactg-3′. Reactions were run on a 1% agarose gel and bands were excised and cloned into the TOPO- TA vector pCR2.1.

(d) Phylogenetic analysis

Phylogenetic trees were generated using maximum-likelihood methods using both programs phyml (Guindon & Gascuel 2003) andraxml (Stamatakis et al. 2005) under a WAG amino acid substitution model (Whelan & Goldman 2001) selected based on results from the program modelgenerator (Keane et al. 2006). The alignment was generated using muscle (Edgar 2004), with minor adjustments when default gap insertion was inconsistent between taxa. Trees were examined and visualized using FigTree 1.2 (Rambaut 2007). The variant sites in a sequence alignment were determined and visualized using a custom gapmap.py script, available at http://www.mbari.org/~haddock/scripts/. Sequences HEGFP1 and HbGFP6 have been deposited in GenBank under accession numbers GQ903692 and GQ903693.

(e) Quantum yield

Quantum yield of fluorescence was determined for photo-activated HbGFP ratiometrically, by comparison with fluorescein in 0.1 N NaOH as a standard, and assuming a QY of 0.925 (Magde et al. 2002). Absorbance was monitored at 490 nm for both fluorescein and HbGFP.

(f) Fluorescence excitation and decay

Because exposure to excitatory light could also cause photo-activation during the measurement of decay rates, fluorescence was sampled every 5 min by a short excitation pulse from a xenon strobe, following initial induction with a 460 nm light source. Fluorescent measurements were taken with a Tecan Infinite 200, equipped with a cuvette holder and plate reader.

3. RESULTS AND DISCUSSION

(a) Organism collection and observation

During surveys in the eastern Pacific using blue-water diving techniques (Haddock & Heine 2005), we collected the ctenophore H. beehleri, which has a distinct greenish tint under white light (figure 1a). Fluorescence microscopy showed the presence of bright fluorescent granules throughout the outer epithelium of the ctenophore (figure 1b). The organism is also bioluminescent, and in vivo bioluminescence was at notably longer wavelengths than most other ctenophore emission spectra (Haddock & Case 1999), although bioluminescence spectra did not show the narrow bandwidth characteristic of bioluminescent organisms containing GFP-family FPs (Haddock & Case 1999, Shagin et al. 2004). Individual ctenophores were kept alive in the dark at the temperatures at which they were collected, typically 14–17°C,
before RNA extraction. This species of ctenophore eats other gelatinous plankton (mainly small hydrozoan medusae; Haddock 2007), and its transparent body wall made it possible to confirm that they contained no noticeable gut contents before they were frozen in liquid nitrogen for subsequent processing.

(b) Genetic screening

An expression library constructed from total H. beehleri cDNA in E. coli was initially screened for fluorescence using a microscope and blue-light (470 nm) illumination. From roughly $1.5 \times 10^4$ colonies, one bright fluorescent clone was picked and its plasmid insert was sequenced. To confirm that the sequence was from the ctenophore and not from dietary contamination, primers were designed based on this sequence and were then used to screen cDNA from another specimen collected and prepared several months after the original. These primers recovered 27 unique PCR products, and the inferred amino acid sequences of these products had high similarity to previously characterized FPs. There were 19 variable amino acid sites in a total of 262, and the sequences were at least 97.5 percent identical to the original amino acid sequence recovered from the cDNA library screening. A single PCR product (HbGFP6) was only 68.3 percent identical to the rest of the sequences, although it was clearly more similar to the other H. beehleri proteins than to any other known FPs because it contained the unique insertions that have been found only in these ctenophore proteins. Despite having the same chromophore sequence as the other H. beehleri FPs, this dissimilar protein was not found to be fluorescent. For five of these sequences (called HbGFP1–5), full-length genes were generated by SMART-RACE, expressed and found to be also fluorescent. One of these proteins, HbGFP1, was selected for more extensive characterization.

(c) Gene sequence and gene trees

The majority of H. beehleri proteins were 262 amino acids in length—longer than the mean of 227.5 for 138 other known full-length FPs (electronic supplementary material, fig. S1) and longer than other representative FPs from several phyla (excerpted in table 1). Based on the amino acid alignment and on homology modelling of the tertiary structure (Bystroff et al. 2000; Kelley & Sternberg 2009) using X-ray diffraction structures of other known FPs, the additional length of the HbGFPs comes as a result of three unique insertions four to nine amino acids long. These insertions occur in extended loops near the ends of the strands that form the β barrel (electronic supplementary material, fig. S2). The chromophore of the ctenophore FP was AYG in all cases—a relatively infrequent chromophore previously identified in four anthozoan genera (Fungia, Eusmilia, Astrangia and Condylactis) out of 70 species with FPs currently listed in GenBank.

An amino acid alignment was constructed from 51 FP sequences from all known groups of GFP-bearing metazoans. Conserved residues between these FPs included the sYG motif of the chromophore that is found in all known GFPS, as well as many glycine residues (electronic supplementary material, fig. S2), several of which were conserved across all of the proteins used in the full alignment. Molecular trees were generated from this alignment using maximum-likelihood methods (figure 2), and the same relationships between major clades were recovered by distance-based and parsimony methods. There was a strong bootstrap support for all major clade relationships within the tree, although moderate support for the placement of Anthozoa relative to other groups. Cnidarian FP sequences did not form a monophyletic group, as hydrozoan FPs clustered with the ctenophores, although the distances between the two groups were greater than between crustacean and chordate FPs. Because of interplay between evolutionary history, selection on protein optical properties and the presence of paralogues, it is not unusual for FPs to form clusters that are incongruous with taxonomy (e.g. in corallimorpharians; Schnitzler et al. 2008). Although ctenophores are widely accepted to be basal to cnidarians (e.g. Halanych 2004; Minelli 2009), and perhaps sister to all other metazoans (Dunn et al. 2008), given the long distances between taxa and the difficulty in rooting the tree, it is not possible to ascribe an ancestral state to the ctenophore FPs. To the FP alignment, we added the G2F fragment of nidogen, previously suggested as a GFP homologue (Hopf et al. 2001; Shagin et al. 2004), extracted from searching the Nematostella and Trichoplax genomes. Outgroup analysis tentatively placed the root

<table>
<thead>
<tr>
<th>name</th>
<th>taxon</th>
<th>length</th>
<th>charge</th>
<th>mol wt</th>
<th>accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haeckelia GFP</td>
<td>Ctenophora</td>
<td>262</td>
<td>-9</td>
<td>30.3</td>
<td>GQ903692</td>
</tr>
<tr>
<td>Aequorea GFP</td>
<td>Hydromedusa</td>
<td>238</td>
<td>-5</td>
<td>26.9</td>
<td>M62653</td>
</tr>
<tr>
<td>Anthomedusa GFP</td>
<td>Hydromedusa</td>
<td>225</td>
<td>4</td>
<td>25.0</td>
<td>AY485334</td>
</tr>
<tr>
<td>Discosoma RFP</td>
<td>Corallimorph, Hexacoral</td>
<td>225</td>
<td>1</td>
<td>25.8</td>
<td>AF168419</td>
</tr>
<tr>
<td>Corynactis YFP</td>
<td>Corallimorph, Hexacoral</td>
<td>221</td>
<td>0</td>
<td>25.0</td>
<td>AY823227</td>
</tr>
<tr>
<td>Renilla GFP</td>
<td>sea pansy, Octocoral</td>
<td>222</td>
<td>-4</td>
<td>24.8</td>
<td>AF372525</td>
</tr>
<tr>
<td>Cerianthus GFP</td>
<td>tube anemone, Anthozoa</td>
<td>222</td>
<td>-1</td>
<td>25.1</td>
<td>AF545830</td>
</tr>
<tr>
<td>Pontella GFP</td>
<td>Copepod, Crustacea</td>
<td>221</td>
<td>-2</td>
<td>24.9</td>
<td>AY268074</td>
</tr>
<tr>
<td>Branchiostoma GFP</td>
<td>lancelet, Chordata</td>
<td>223</td>
<td>-2</td>
<td>23.7</td>
<td>EF157660</td>
</tr>
</tbody>
</table>

Table 1. Measured and predicted properties of H. beehleri GFP compared with representative fluorescent proteins from the three other known FP-bearing phyla. Species were chosen as the first-discovered or most commonly encountered representative of that taxonomic group. Length is in amino acids, charge is the hypothetical charge based on the primary sequence. Molecular weights in kilodaltons. Calculated values were determined using the protcalc.pl script, available at http://www.mbari.org/~haddock/scripts/.
between the (Hydrozoa + Ctenophora) and (Anthozoa + Bilateria) clades.

**d** Spectroscopy and optical properties

Cloned proteins were given an N-terminal hexa-histidine tag, expressed in *E. coli* and purified using Ni–NTA affinity chromatography. Subsequent optical characterization used purified recombinant protein from HbGFP1, with the properties measured in the fluorescent activated state ($\S$3e). Fluorescence emission peaked at 512 nm, with the width at half maximum of 27 nm and the ‘typical’ shape of a GFP-type emission, including a red-shifted shoulder at about 542 nm (figure 3a). The absorbance spectrum during peak fluorescence had a maximum at 495 nm. The quantum yield of the HbGFP was 0.83 with a calculated extinction coefficient of 39 420 m$^{-1}$1. The charge was among the most negative of natural FPs, partly in relation to the length of the protein, along with a larger proportion (approx. 15%) of acidic residues near the carboxy terminus (electronic supplementary material, fig. S1).

Based on size-exclusion HPLC and non-denaturing gel electrophoresis (figure 4) relative to monomeric and tetrameric DsRed, HbGFP appears to occur as a dimer, with a tendency to form higher-order aggregates at high concentrations.

**e** Photoactivation

During the screening and characterization of clones, we noticed that the fluorescence brightened during examination. Further experimentation showed that the expressed proteins were strongly photoactivated and did not fluoresce until exposed to blue or violet light. In the non-fluorescent state, the 495 nm absorbance peak was absent, and a 389 nm absorbance peak appeared in its place (figure 3a). This absorbance peak is nearly identical to that found in the protonated (non-fluorescent)
4. CONCLUSIONS

While bioluminescence is known to occur across a vast diversity of taxa, this marks only the fourth phylum known to possess FPs. The large phylogenetic distances between FP-bearing taxa, however, suggest that these proteins may ultimately be found in many Metazoan lineages.

REFERENCES


Retracted

function. FPs in bioluminescent animals are generally believed to shift the wavelength of luminescent light emission to a longer wavelength that may be more ecologically and behaviourally appropriate (Haddock & Case 1999). However, in the darkness required for bioluminescence to be visible, the unactivated FP would not be functional unless there were enough dim blue ambient light to maintain photoactivation. If the light of bioluminescence itself were responsible for the activation, then the overall emission of *H. beehleri* would change from blue to green during the course of a luminescent display. Ctenophore photoproteins used to make bioluminescence are also known to be inactivated by light (Ward & Seliger 1976), so it may even be possible that these FPs are serving a photoprotective function. It is possible that proteins only exhibit photoactivation in *in vitro*, and not within the live organism. *Haecelata beehleri* is uncommon enough that we have not been able to thoroughly test its natural fluorescent properties since isolating the protein.

Figure 4. Native PAGE gel of purified recombinant FPs. *(a)* White light image of coomassie-blue-stained standards and novel FPs. *(b)* A blue-illuminated image of the same gel, showing fluorescence emission of native proteins. The native ladder (first column) includes a fluorescent marker that is visible in *(b).* Lanes 1–5 in both images: native protein ladder, monomeric DsRed, tetrameric DsRed, HbGFP sample 1, HbGFP sample 2. The two *H. beehleri* samples were from clone HbGFP1 used in this study, but they were expressed and prepared at different times.

form of the ‘Dronpa’ coral FP (Ando et al. 2004; Habuchi et al. 2005) and in the *Aequorea*-based photoactivatable FP PA-GFP (Patterson & Lippincott-Schwartz 2002). Interestingly, the Dronpa FP loses fluorescence under blue illumination, whereas the HbGFP gains fluorescence in these conditions. In both proteins, during the interconversion between the fluorescent and non-fluorescent states, the absorbance spectra show an isosbestic point at 430 nm (figure 3a). Unlike many photoswitchable FPs, when in the ‘off’ state, HbGFP show almost no trace of the longer-wavelength, fluorescence-associated absorbance spectrum (arrowhead in figure 3a). The fluorescence emission of HbGFP increases at least 130–160-fold during the course of blue-light activation (figure 3b), with an initial doubling time of approximately 15 s, dependent on stimulus intensity.

Substitution of threonine 203 with histidine was required to introduce photoactivatability in PA-GFP, an engineered version of *Aequorea* GFP (Patterson & Lippincott-Schwartz 2002), and the corresponding position is also occupied with a histidine in the photoswitchable cerianthid GFP cerFP505 (Vogt et al. 2008). In HbGFP, however, that position is occupied by the original threonine, indicating that this histidine substitution is not mandatory for photoactivation. To our knowledge, the AYG chromophore of the HbGFP has also not been found in other photoactivatable FPs.

After blue-light activation, fluorescence of the protein decayed exponentially (dashed line in figure 3c), following the relationship

\[ F(t) = F_0 e^{-0.022174t} \]  

(3.1)

with \( t \) in minutes (\( r = 0.996 \)). This resulted in a fluorescence half-life of approximately 32 min.

The presence of a photoactivatable FP in a bioluminescent organism raises questions as to its natural

REFERENCES


Haddock, S. H. D. & Heine, J. N. 2005 *Scientific blue-water diving*, p. 49. La Jolla, CA: California Sea Grant.


Rambaut, A. 2007 FigTree, a graphical viewer of phylogenetic trees. See http://tree.bio.ed.ac.uk/software/figtreec.


Downloaded from http://rspb.royalsocietypublishing.org/ on July 25, 2017
A photoactivatable green-fluorescent protein from the phylum Ctenophora

Steven H. D. Haddock, Nadia Mastroianni and Lynne M. Christianson

In the years following the publication of ‘A photoactivatable green-fluorescent protein from the phylum Ctenophora’ [1], another research group [2] sequenced hydrozoan (Cnidaria) fluorescent proteins that were very similar to those we cloned from cDNA prepared from ctenophore specimens. We therefore now believe that the green-fluorescent proteins reported in our study are not from ctenophores, and were due to incorporation of cnidarian prey into the ctenophore tissues.

Although our genes were cloned from two independent samples of mRNA taken years apart, the similarity of our sequences to those obtained from a siphonophore (Cnidaria, Hydrozoa) indicate that this mRNA was carried over from ingested material. This ctenophore species is known to prey on cnidarians, but we did not expect that mRNA would persist for as long as it did, and we regret our erroneous conclusion.

What is remarkable about this situation is that the ctenophore, in addition to maintaining full-length mRNA for the fluorescent protein, also incorporated the proteins into its tissue to the extent that its bioluminescence emission was green-shifted. While the exact species of origin for these proteins is unknown, the photoactivatable properties of this fluorescent protein are as originally described, and it remains an interesting target for future research and application.

References

Retraction

In the years following the publication of ‘A photoactivatable green-fluorescent protein from the phylum Ctenophora’ [1], another research group [2] sequenced hydrozoan (Cnidaria) fluorescent proteins that were very similar to those we cloned from cDNA prepared from ctenophore specimens. We therefore now believe that the green-fluorescent proteins reported in our study are not from ctenophores, and were due to incorporation of cnidarian prey into the ctenophore tissues.

Although our genes were cloned from two independent samples of mRNA taken years apart, the similarity of our sequences to those obtained from a siphonophore (Cnidaria, Hydrozoa) indicate that this mRNA was carried over from ingested material. This ctenophore species is known to prey on cnidarians, but we did not expect that mRNA would persist for as long as it did, and we regret our erroneous conclusion.

What is remarkable about this situation is that the ctenophore, in addition to maintaining full-length mRNA for the fluorescent protein, also incorporated the proteins into its tissue to the extent that its bioluminescence emission was green-shifted. While the exact species of origin for these proteins is unknown, the photoactivatable properties of this fluorescent protein are as originally described, and it remains an interesting target for future research and application.

References


Steven H. D. Haddock
Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA 95039, USA
e-mail: haddock@mbari.org

Nadia Mastroianni
Axxam SpA, San Raffaele, Biomedical Science Park, via Olgettina 58, 20132 Milan, Italy
e-mail: Na_mastro@yahoo.it

Lynne M. Christianson
Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA 95039, USA
e-mail: lynne@mbari.org

© 2015 The Author(s) Published by the Royal Society. All rights reserved.