

Green-fluorescent protein mutants with altered fluorescence excitation spectra

Torsten Ehrig^{a,*}, Dennis J. O’Kane^b, Franklyn G. Prendergast^a

^aDepartment of Pharmacology, Mayo Foundation, Rochester, Minnesota 55905, USA

^bDepartment of Laboratory Medicine and Pathology, Mayo Foundation, Rochester, Minnesota 55905, USA

Received 29 March 1995

Abstract Using random mutagenesis and visual selection of fluorescent clones, we have isolated a T203I and a E222G mutant of the *Aequorea* green-fluorescent protein. Each mutant has one of the two fluorescence excitation bands of the wild type deleted and retains the other without a wavelength shift. This finding is consistent with each excitation band corresponding to a distinct spectroscopic state of the chromophore. Both mutations are single amino acid exchanges which in the linear sequence are located remotely from the chromophore but in the folded protein may be situated in its vicinity. We conclude that the mutations influence the fluorescence properties by changing the interactions between the chromophore and its protein environment.

Key words: Luminescent protein; Fluorescence spectrometry; Mutagenesis (MeSH)

1. Introduction

The jellyfish *Aequorea victoria* and the sea pansy *Renilla reniformis* emit green light from a green-fluorescent protein (GFP), the in vitro emission characteristics of which match the in vivo bioluminescence emission [1–3]. Excitation of the GFPs is accomplished through energy transfer from the chemiluminescent (bioluminescent) proteins aequorin in *Aequorea* [4,5] and from a luciferase in *Renilla* [6]. The most intriguing feature of GFPs is that the fluorophore responsible for the green fluorescence is an integral part of the GFP peptide chain and not a prosthetic group. After limited protease digestion of *Aequorea* GFP, a covalently modified hexapeptide is recovered that has identical spectroscopic properties to that of the denatured, but otherwise intact GFP [7]. The essential features of the chromophore, as proposed originally by Shimomura [8] and extended by spectroscopic studies of synthetic model chromophores and two-dimensional NMR data [7] appears to be covalent modification of an internal tripeptide with the sequence SYG contained within the hexapeptide. Alignment with the amino acid sequence, deduced from the cDNA sequence [9], is consistent with a sequence FSYGVQ of the unmodified hexapeptide comprising amino acids 64–69 [7]. Since GFP expressed in organisms other than jellyfish such as *E. coli* [10,11], *Cenorhabditis elegans* [10] and *Drosophila* [12] still shows the characteristic green fluorescence, it has been concluded that the mechanism

of chromophore formation is either autocatalytic or employs ubiquitous cellular components.

Aequorea and *Renilla* GFPs most likely have a structurally identical chromophore. Characterization of a hexapeptide from *Renilla* GFP containing the chromophore [13] is consistent with the sequence FSYGDR of the unmodified hexapeptide, which contains the same internal sequence SYG as *Aequorea* GFP. The essential identity of the fluorescence emission spectra of the intact proteins and the identity of the absorbance spectra of the isolated hexapeptides containing the chromophore further support the structural identity of the chromophores [14]. The evolutionary relationship of the two GFPs has not been defined on a molecular level since only the amino acid sequence of *Aequorea* GFP, deduced from the cDNA sequence [9], but not of *Renilla* GFP is known.

There are indications that the fluorescence properties of the chromophore in *Aequorea* and *Renilla* GFP are influenced substantially by the surrounding protein matrix. In spite of the proposed structural identity of the chromophore in both proteins, the fluorescence excitation spectra of the folded proteins differ widely. The *Aequorea* GFP has two excitation maxima which were reported to be situated around 393 and 473 nm, while the *Renilla* GFP exhibits a maximum only at 498 nm [14]. This difference was interpreted as a consequence of different fluorophore environments in the folded proteins [14]. Furthermore, in the isolated hexapeptide [7] or in the denatured protein [14] the chromophore is not fluorescent and has different absorption spectra when compared to the native proteins, and synthetic model chromophores are not fluorescent [7]. However, the normal fluorescence is recovered when denatured *Aequorea* GFP is allowed to refold [15], which indicates that the spectral changes just described are not due to chemical modification of the ‘exposed’ chromophore.

Since the tertiary structures of GFPs have not been solved, we do not know which amino acid residues interact with the chromophore and might be responsible for the differences in excitation spectra. Accordingly, mutants with altered fluorescence properties have been isolated by random mutagenesis and visual selection [16,17]. Using a similar approach, we have identified *Aequorea* GFP mutants with amino acid exchanges which have profound effects on the fluorescence excitation spectrum.

2. Materials and methods

2.1. Expression of GFP in *E. coli*

Random mutagenesis and visual screening employed the GFP expression plasmid Tu#65 propagated in the *E. coli* host DH5 α (Gibco BRL Life Technologies, Gaithersburg, MD). Tu#65 had been constructed by Chalfie and collaborators by cloning the GFP coding

*Corresponding author. Present address: Dermatologische Klinik der Philipps-Universität Marburg, Deutschhausstr. 9, 35033 Marburg, Germany. Fax: (49) 6421 28 29 02.

Abbreviations: GFP, Green-fluorescent protein; IPTG, Isopropyl- β -D-thiogalactopyranoside

sequence into the Bluescript II KS(+) vector (Stratagene, LaJolla, CA), allowing expression of GFP under the control of the IPTG-inducible *lac* promoter as a fusion protein with 23 additional amino acids derived from the 5'-region of the β -galactosidase gene and from the plasmid polylinker [10]. The fluorescence properties of the expressed wild-type GFP fusion protein are identical to those of authentic *Aequorea* GFP.

2.2. Random mutagenesis and visual screening

Random mutagenesis of the GFP cDNA was achieved using the *E. coli* XL1-Red strain (Stratagene), which is deficient in DNA polymerase proofreading activity and DNA repair mechanisms and introduces random mutations into a plasmid during cell division. After propagating Tu#65 in *E. coli* XL1-Red for 2 days, the randomly mutated plasmid DNA was isolated (Wizard Midiprep, Promega, Madison, WI) and transformed into the nonmutagenic expression host DH5 α . Bacteria were subsequently plated on 0.45 μ M pore size nitrocellulose membranes (Schleicher and Schuell, Keene, N.H.), providing a suitable visual background. The membranes were mounted on LB agar containing 50 μ M ampicillin and 1 mM IPTG (Sigma, St. Louis, MO). The colonies were screened under illumination with monochromatic light of alternating 390 and 470 nm wavelengths which was generated by passing light from a 150 W Xenon lamp (Xenon Corporation, Woburn, MA) through interference filters (CVI Laser Corporation, Albuquerque, N.M.). The fluorescence emission of the bacterial colonies was observed through a Schott KV 500 filter (500 nm cutoff wavelength). Plasmid DNA was isolated from both clones as above (Wizard midiprep), and the insert was sequenced using the dideoxy chain termination method [18] on an automated 370A DNA Sequencer (Applied Biosystems, Foster City, CA). The mutated plasmids were designated pGFP_{T203I} and pGFP_{E222G}, respectively.

2.3. Protein purification

The *Aequorea* GFPs were isolated from *E. coli* DH5 α harboring the wild-type GFP expression plasmid Tu#65 or the mutated expression plasmids pGFP_{T203I} and pGFP_{E222G}. Bacteria were grown in 1 L LB media containing 50 μ M ampicillin and 1 mM IPTG, harvested by centrifugation, and lysed by freezing (-60°C) and thawing of the cell pellet. Debris was removed by centrifugation at 14,000 \times g, and the supernatant was loaded on a G-75 Sephadex column (Pharmacia, Piscataway, NJ) equilibrated with 10 mM phosphate buffer, pH 7.0. The column had been calibrated with molecular weight markers ribonuclease A, chymotrypsinogen A and ovalbumin of 13.7, 25 and 43 kDa molecular weight, respectively (Pharmacia). Fractions containing GFP were identified by fluorescence emission at 510 nm when excited by the appropriate wavelength. All three *Aequorea* GFPs eluted from the gel filtration column with apparent molecular weights between 28 and 33 kDa, in good agreement with the calculated molecular weight of the expressed GFP fusion proteins of 30.2 kDa. The peak fraction of each preparation was used directly for the fluorescence spectral measurements. The Renilla GFP was isolated from *Renilla reniformis* as reported previously [3].

2.4. Fluorescence spectroscopy

Fluorescence spectra were recorded on a SPEX Fluorolog fluorimeter (SPEX Industries, Edison, NJ) with excitation and emission band-passes set at 2 nm.

2.5. Reagents

If not indicated otherwise, reagents were obtained from Sigma (St. Louis, MO).

3. Results

Approximately 200,000 clones from a library of the *Aequorea* GFP cDNA carrying randomly distributed mutations were visually screened for altered fluorescence properties using alternating exciting wavelengths of 390 and 470 nm. Whereas the wild type clone exhibited visible green fluorescence with both excitation wavelengths, one clone was identified which was fluorescent only when illuminated at 470 nm. DNA sequencing of the coding region revealed a single base mutation changing

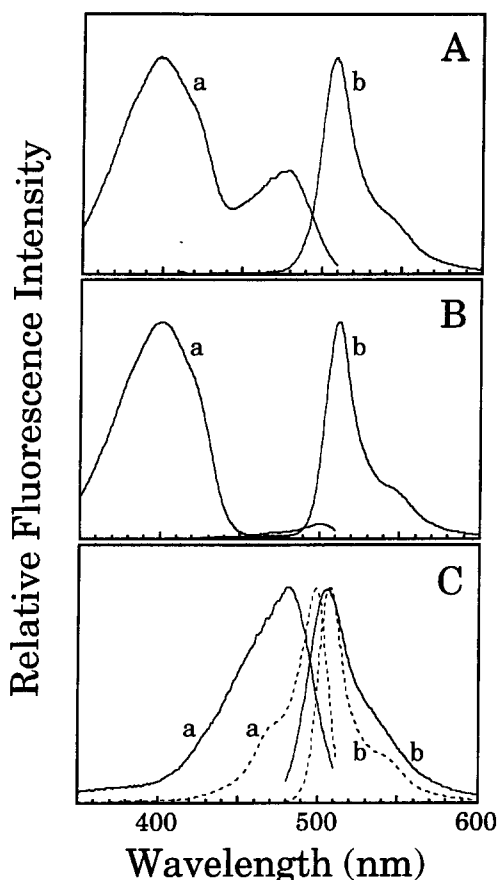


Fig. 1. Fluorescence excitation (a) and emission (b) spectra of soluble cell extracts containing green-fluorescent proteins. A, *Aequorea* wild-type GFP; B, *Aequorea* GFP_{T203I} mutant; C, *Aequorea* GFP_{E222G} mutant (solid lines) and *Renilla* GFP (broken lines). The excitation spectra were recorded at a fixed emission wavelength of 520 nm. For the emission spectra, the fixed exciting wavelength was 400 nm for the *Aequorea* wild-type and for GFP_{T203I}, and 470 nm for GFP_{E222G} and *Renilla* GFP. The spectra are normalized to equal peak heights. The buffer was 10 mM sodium phosphate, pH 7.0, and the temperature was 20°C.

the codon of E222 from GAG to GGG which results in a glycine at this position. In addition, one clone was isolated which exhibited the inverse behavior in that it was fluorescent only when excited at 390 nm. The DNA sequence of this clone showed a single base mutation which changes the codon of T203 from ACA to ATA. This exchange results in an isoleucine and thus turns out to be identical to a mutation T203I which had been reported previously [16].

The fluorescence spectra of protein extracts confirm the profound dissimilarity between the mutated and the wild-type proteins (Fig. 1, Table 1). The positions of the emission maxima of both mutants do not differ substantially from those of the wild-type, but the excitation spectra show marked changes which, comparing the two mutants, can be viewed as reciprocal events: whereas GFP_{T203I} has the 480 nm peak deleted and retains the 400 nm peak, GFP_{E222G} has the 400 nm peak deleted and retains the 480 nm peak. In both cases, the maximum intensity wavelength of the retained band is not changed compared to the wild type.

Table 1
Fluorescence excitation and emission maxima of green-fluorescent proteins

Protein	Excitation maximum (nm)	Emission maximum (nm)
<i>Renilla</i> GFP	499	507
<i>Aequorea</i> GFPs		
Wild-type	398 480	509
GFP _{T203I}	400	512
GFP _{E222G}	481	506
RSGFP4	490	505

Data for the RSGFP4 mutant are from reference [17].

4. Discussion

In the present report as well as in previous publications [16,17] *Aequorea* GFP mutants with altered fluorescence properties are described which have been isolated by random mutagenesis and selection of phenotypes of interest. These mutations can be tentatively divided into two groups. One group has in common that not only changes in intensities, but also in the wavelength positions of fluorescence excitation bands are found. These mutations concern amino acids within the FSYGVQ hexapeptide and comprise the single-amino acid mutations Y66H and Y66W [16], as well as the triple mutant named RSGFP4 which has three substitutions within the FSYGVQ hexapeptide, resulting in the sequence MGYGVL [17]. In contrast, four *Aequorea* GFP mutations can be grouped together because the relative intensities of the excitation bands are changed, in some cases to the extreme of completely deleting one band, *without* alteration of the wavelength positions of those bands still present. These mutations comprise T203I and S202F, which have the 480 nm peak deleted ([16] and Fig. 1), I167T which has the intensity ratio of the excitation bands changed in favor of the 480 nm peak [16], and the mutation E222G described here which has the 400 nm peak deleted (Fig. 1). Interestingly, in these four cases the mutated residues are located in the C-terminal part of the peptide chain and remote from the chromophore in the linear amino acid sequence, suggesting that in the folded protein the C-terminal part interacts with the chromophore and influences its fluorescence properties. The non-conservative nature of the observed amino acid exchanges suggests that polarity and electrical charge of the environment of the chromophore are important determinants of its fluorescence.

In the isolated GFP chromophore, two spectroscopic states have been observed which correspond to the two absorbance maxima at 380 and 445 nm. The relative intensities of these bands can be shifted by varying the pH, and the observation of an isosbestic point indicates that not more than two spectroscopic states are involved [7]. That in the folded protein the chromophore exhibits two similar states is strongly supported by the existence of the mutations described above which change only the relative intensities of the excitation bands, but not the position of the bands on the wavelength scale. These mutations suggest that in the folded protein the chromophore assumes two different spectroscopic states with fluorescence excitation maxima of 400 and 480 nm, the relative intensities of which

depend on interactions between the chromophore and the C-terminal part of the protein.

The *Aequorea* GFP has two fluorescence excitation maxima at 400 and 480 nm, whereas the *Renilla* GFP has one excitation maximum at 499 nm ([14] and Fig. 1). Two *Aequorea* GFP mutants, the triple mutant RSGFP4 [14] (see above) and the GFP_{E222G} mutant, are now known which have the 400 nm excitation peak deleted and thus have their excitation spectrum shifted towards that of *Renilla* GFP (Table 1). To determine whether similar or analogous substitutions are responsible for the spectroscopic differences between the *Renilla* and *Aequorea* GFPs will require elucidation of the *Renilla* GFP amino acid sequence and alignment with the *Aequorea* GFP sequence.

The ability to express fluorescent GFP in a variety of organisms makes it an extremely useful fluorescence marker. For example, when GFP was expressed under the control of a *mec-7* gene promoter, the characteristic fluorescence of GFP appeared *in vivo* in those cells that normally express *mec-7*, demonstrating that GFP is a useful reporter for promoter activation [10]. Furthermore, GFP can be used as a fluorescent label to study the cellular distribution of a protein of interest *in vivo*. As an example, a fusion construct of the *Exu* protein and GFP was expressed in *Drosophila* oocytes, and the spatial distribution and temporal appearance of the fluorescent Exu-GFP conformed to that of the endogenous *Exu* protein [12]. The introduction of mutants which change the fluorescence properties of GFP will enhance the usefulness of this protein. Mutants with independent fluorescence excitation are available through the pair T203I [16] and RSGFP4 [17], as well as through the pair described here, and will be useful for labelling of two cellular proteins or following the activity of two promoters simultaneously. The profound effects of mutations within the C-terminal region of the GFP protein should encourage further attempts to change the fluorescence properties of GFP by mutations not only within the chromophore, but also within the C-terminal region, and possibly also other parts of the peptide chain.

Acknowledgements: We thank M. Chalfie, Columbia University, NY, and D.C. Prasher, Institute for Marine Biology, Woods Hole, MA) for providing the wild-type GFP cDNA clone Tu#65. This work was supported by grants Gm-46300 (D.J.O.) and Gm-34847 (F.G.P.) from the National Institutes of Health, and by a Thompson-Mayo Fellowship (T.E.).

References

- [1] Morise, H., Shimomura, O., Johnson, F.H. and Winant, J. (1974) *Biochemistry* 13, 2656–2662.
- [2] Ward, W.W. and Cormier, M.J. (1976) *J. Phys. Chem.* 80, 2289–2291.
- [3] Ward, W.W. and Cormier, M.J.J. (1979) *J. Biol. Chem.* 254, 781–788.
- [4] Shimomura, O. and Johnson, F.H. (1969) *Biochemistry* 8, 3991–3997.
- [5] Shimomura, O. and Johnson, F.H. (1970) *Nature* 227, 1356–1357.
- [6] Charbonneau, H. and Cormier, M.J. (1979) *J. Biol. Chem.* 254, 769–780.
- [7] Cody, C.W., Prasher, D.C., Westler, W.M., Prendergast, F.G. and Ward, W.W. (1993) *Biochemistry* 32, 1212–1218.
- [8] Shimomura, O. (1979) *FEBS Lett.* 104, 220–222.
- [9] Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G. and Cormier, M.J. (1992) *Gene* 111, 229–233.
- [10] Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. and Prasher, D.C. (1994) *Science* 263, 802–805.

- [11] Inouye, S. and Tsuji, F.I. (1994) *FEBS Lett.* 341, 277–280.
- [12] Wang, X. and Hazelrigg, T. (1994) *Nature* 369, 400–403.
- [13] SanPietro, R.M., Prendergast, F.G. and Ward, W.W. (1993) *Photochem. Photobiol.* 57, 63S.
- [14] Ward, W.W., Cody, C.W., Hart, R.C. and Cormier, M.J. (1980) *Photochem. Photobiol.* 31, 611–615.
- [15] Ward, W.W. and Bokman, S.H. (1982) *Biochemistry* 21, 4535–4540.
- [16] Heim, R., Prasher, D.C. and Tsien, R.Y. (1994) *Proc. Natl. Acad. Sci. USA* 91, 12501–12504.
- [17] Delagrave, S., Hawtin, R.E., Silva, C.M., Yang, M.M. and Youvan, D.C. (1995) *Biotechnology* 13, 151–154.
- [18] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.