Development and Applications of Enhanced Green Fluorescent Protein Mutants

ABSTRACT

The introduction of several mutations resulted in the generation of improved mutants of the green fluorescent protein (GFP). A strong green (GFPsg25) and blue (BFPsg50) fluorescent protein, gave 50-fold–100-fold brighter fluorescence compared to wild-type GFP and BFP (Tyr66His), respectively, upon expression in mammalian cells. GFPsg25 and BFPsg50 have different excitation and emission maxima. This allows their use as an efficient dual-color tagging system and their independent detection in living cells.

INTRODUCTION

The green fluorescent protein (GFP) of Aequorea victoria (21) has been used extensively as a live marker and a unique tool to monitor dynamic processes in living cells and whole organisms (5,8). This has been possible because GFP, in contrast to other bioluminescent molecules, operates independently of cofactors, is active in a variety of organisms (20) and can be detected rapidly and easily (18). The various applications of GFP range from the study of protein localization and trafficking (13,17,28) to marking cell lines during development (32) and to the expression by recombinant viruses (14,15). However, there is still a need for brighter GFP mutants and mutants with different spectral properties to expand the potential applications. For this purpose, we generated enhanced GFP mutants and present their use in several applications including double-color fluorescence-activated cell sorting (FACS) and dual-color labeling in living cells. We also used these new mutants in an assay for gene expression and in retroviral vectors for single-copy gene expression of GFP-tagged proteins.

MATERIALS AND METHODS

Generation of Enhanced Green and Blue Fluorescent Protein Mutants

The GFP sequences were derived from plasmid TU#58 (2). Plasmid pBSGFP was created by inserting the GFP coding region from pCMV-GFP (28), digested with BssHII and BamHI and subsequently treated with Klenow, into the EcoRV-digested pBluescript® SK(+) vector (Stratagene, La Jolla, CA, USA). pBSGFP was used to mutate the GFP coding sequence by single-stranded (ss)DNA site-directed mutagenesis using appropriate oligonucleotides as described (24). pCMV-GFPsg11 was constructed by replacing the wild-type GFP gene of pCMV-GFP with the corresponding mutagenized GFP genes from pBSGFP by NheI/BamHI subcloning. In plasmid pCMV-GFPsg11, Ser65 was changed to Cys, resulting in pCMV-GFPsg25. To generate pCMV-GFPsg12, the AvrII/PmlI fragment of pCMV-GFP was exchanged by the corresponding DNA sequence from pCMV-GFPsg11. Similarly, the blue fluorescent protein (BFP)-expressing plasmids pCMV-BFPsg42 and pCMV-BFPsg50 were engineered by introducing the desired changes by site-directed mutagenesis into plasmid pCMV-GFPsg12.

Spectroscopy

Fluorescence spectra were measured on a Model L550B Spectrofluorimeter (Perkin-Elmer, Norwalk, CT, USA) as described (19).

Quantitation of Cellular Fluorescence

Relative fluorescence of the GFP/BFP mutants was obtained by comparing the fluorescence of cytoplasmic extracts prepared from 293 cells.
transiently transected in duplicate or triplicate with different amounts of GFP expression plasmids. Similar results were obtained in independent experiments including luciferase activity as an internal transfection control. The fluorescent signal of living cells in monolayers or cellular extracts was quantitated using a Cytofluo® II Fluorescence Plate Reader (PerSeptive Biosystems, Framingham, MA, USA) equipped with a 360/40- or 485/20-nm excitation filter and a 460/40- or 530/30-nm emission filter, for the detection of BFPsg50 or GFPsg25, respectively. For the measurements of Table 2, equal amounts of plasmid DNAs expressing the wild-type or mutant proteins were transfected into 293 cells, and the corresponding cellular fluorescence for each protein was quantitated in cell extracts using Cytofluo II. The following filter combinations were used: wild-type GFP, sg12, 360/40-530/30 nm; sg11, sg25, 485/20-530/30 nm; BFP (Y66H = P4 in Table 2), sg42, sg49, sg50 360/40–460/40 nm.

Construction of GFP/BFP Fusion Proteins

Fusions to the N terminus of GFP/BFP were constructed by inserting the coding regions of the specific genes into the unique NheI site in the plasmid pCMV-GFPsg25 or pCMV-BFPGsg50. To generate fusions to the C terminus of GFP, we used pF25GFP-Hyg (K. Horie and G.N. Pavlakis, unpublished), which expresses a GFP-hygromycin fusion. The coding region for Hyg in pF25GFP-Hyg was replaced by the gene of interest, using the unique NarI and XbaI sites flanking the hygromycin gene. In the resulting plasmids, the expression of the hybrid gene is under the control of the bovine growth hormone polyadenylation signal.

Cytometry and Sorting of Living Cells

Cytometry of 293 cells transiently transfected with plasmids expressing GFPsg25 and BFPsg50, or of transduced NIH3T3 cells, was performed on a two-laser FACStarPlus™ platform (Becton Dickinson Immunocytometry, San Jose, CA, USA). An argon ion laser at a wavelength of 488 nm was used to excite GFP (run at 200 mW with a 500-nm, long-pass emission filter), whereas a multiline UV laser at 364-nm wavelength was used to excite BFP. Emitted GFP fluorescence was detected with a 510–540-nm band-pass filter, or a 420-nm cut-off filter was used to detect BFP emission. Normaski images were made using a 543-nm green laser and appropriate polarized lenses.

Construction of GFP-Neo Retroviral Expression Vectors

To express a GFPsg25-Neo fusion, the neo gene was cloned into pCMV-GFPsg25, and the CMV promoter was subsequently replaced by the pgk promoter (1). The DNA fragment containing the GFPsg25-Neo fusion under the control of the pgk promoter was then subcloned into the XhoI site of pGen-(27) resulting in pGen-PGKgfo25RO.

Production of Pseudotyped Retroviral Particles

Production of retroviral stocks and infections and viral titer determination were carried out as described by Soneoka et al. (26). Briefly, 293T cells were transfected by the calcium phosphate method using 6 µg pHCMV-G (31). The combination of these vectors allowed the generation of infectious retrovirus transducing the expression cassette for GFPsg25-neo. pHIT60 expressed Gag-Pol polyproteins of murine leukemia virus (MLV), while pHCMV-G expressed the vesicular stomatitis virus (VSV) G protein, which is incorporated on the surface of the retroviral particles and allows infection of mammalian cells. Two days after transfection, the supernatant was

<table>
<thead>
<tr>
<th>Table 1. GFP and BFP Mutants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>wtGFP</td>
</tr>
<tr>
<td>sg12</td>
</tr>
<tr>
<td>sg11</td>
</tr>
<tr>
<td>sg25</td>
</tr>
<tr>
<td>BFP</td>
</tr>
<tr>
<td>sg42</td>
</tr>
<tr>
<td>sg49</td>
</tr>
<tr>
<td>sg50</td>
</tr>
</tbody>
</table>

The numbering corresponds to amino acids of the wild-type GFP.
harvested, filtered and titrated on NIH3T3 cells. To generate cell lines containing only one retroviral integration event, NIH3T3 cells were infected at a multiplicity of infection (MOI) of 0.001 and put under G418 selection (500 µg/mL) for 21 days. Independent clones were selected and analyzed by Southern blotting and FACS.

RESULTS AND DISCUSSION

Generation and Characterization of Enhanced GFP and BFP Protein Mutants

Starting from the wild-type gfp sequence (2), we have introduced several mutations by ssDNA site-directed mutagenesis. During mutagenesis, we discovered a mutation in which the phenylalanine at position 64 was changed into a leucine (Phe64Leu), resulting in a GFP protein with increased fluorescent signal. This mutation, which was independently reported by Cormak et al. (3), caused higher yields of soluble GFP when expressed in E. coli and is also suspected to increase folding efficiency (19). The already high quantum yield of wild-type GFP was not further improved (Table 2). To optimize the beneficial properties of GFP, we combined Phe64Leu with several mutations generated by us and others (11). Being interested in applications in mammalian cells, we compared the fluorescence intensities and spectral properties of the mutants directly in human cells after transient expression. This approach takes into account the efficiency of translation, protein folding, chromophore formation and stability inside mammalian cells under physiological conditions. Because Phe64Leu did not change the excitation maxima, we introduced the Ile167Thr mutation, which caused an excitation shift to a longer wavelength (11). The combination of Phe64Leu, Ser65Cys and Ile167Thr resulted in a synergistic effect on the fluorescence intensity and still preserved the shift of the excitation maximum to 470 nm.

The previously reported blue variant of GFP (BFP, Tyr66His) (12) has not been very useful in practical applications because the emitted signal is very weak. To improve its properties, we introduced several mutations suspected to increase the signal intensity or folding, chromophore maturation and stability of the protein and tested the mutants in human cells. The beneficial effect of Phe64Leu in combination with Val163Ala on protein solubility and improved folding (4,25) seemed to account predominantly for the increased blue fluorescence because the quantum yields of the BFP mutants are similar to BFP (Tyr66His) (12) (Table 2). Our strategy led to a series of GFP and BFP mutants with enhanced fluorescent signals (Table 1). Although the introduced mutations increase the adsorption coefficients, the brighter fluorescence is most likely caused by a higher percentage of active GFP molecules in a given population due to improved chromophore maturation and stability under physiological conditions. GFPsg25 and BFPsg50 were 50–100 times brighter than wild-type GFP or BFP, respectively (Table 2). In addition, the mutants GFPsg25 and BFPsg50 have different excitation (474 vs. 386 nm) and emission maxima (509 vs. 450 nm) with no overlap (Figure 1), which allows the independent detection of BFP and GFP as a dual-color tagging system.

Multiparameter Cell Sorting Based on Two Different-Colored Autofluorescent Proteins

We tested the suitability of the dual-color system for rapid and specific identification of tagged cell populations. Figure 2 illustrates the FACS analysis of cells transfected with plasmids expressing GFPsg25, BFPsg50 or both. Using an argon or a UV laser for the detection of GFP or BFP, respectively, a bright green and a blue cell population were easily detectable (Figure 2, A and B). Subsequently, the GFPsg25- or BFPsg50-transfected cells were mixed (Figure 2D) and sorted by gating on the green or blue signal. This resulted in the successful re-separation of two homogeneous green and blue cell populations (Figure 2, E and F). The co-transfection of the two plasmids resulted in cells with a combination of the green and the blue signals (Figure 2C). GFPsg25 and BFPsg50 are better tags for flow cytometry than wild-type GFP and BFP used previously (23). Ropp et al. (23) performed analysis of cells expressing wild-type GFP and BFP but not cell sorting based on the two colors. The wavelength used (407 nm) is not optimal for wild-type BFP excitation, and

Figure 1. Excitation and emission spectra of GFPsg25 and BFPsg50.
Research Reports

Sorting of a mixture of GFP- and BFP-expressing cells needs to be based solely on different emission signals. The GFP and BFP variants described in this report not only have distinct emission maxima (Figure 1) facilitating the separation of different tagged cell populations. In addition, the strong fluorescence of GFPsg25/BFPsg50 resulted in a high percentage of bright green and blue cells. The development of the strong BFPsg50 allows the routine use of multiparameter flow cytometry to assess differential gene expression and the simultaneous tagging of different cell populations.

Combining the power of flow cytometry single cell sorting with the improved properties of the GFP mutants resulted in an efficient approach to generate stable cell lines expressing GFP-tagged proteins, without the time-consuming selection for drug resistance first (29). Because the expression of the trans gene correlates to the intensity of the GFP signal, the FACS approach allows to focus immediately on specific cell populations; i.e., on high, medium

### Table 2. Properties of GFP Mutants

<table>
<thead>
<tr>
<th>Protein</th>
<th>Excitation Maximum</th>
<th>Emission Maximum</th>
<th>Increased Green Fluorescence$^a$</th>
<th>Increased Blue Fluorescence$^a$</th>
<th>E$^c$ (cm/M)</th>
<th>Quantum Yield$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>wtGFP</td>
<td>398 nm</td>
<td>509 nm</td>
<td></td>
<td></td>
<td></td>
<td>0.86</td>
</tr>
<tr>
<td>S65T$^d$</td>
<td>490 nm</td>
<td>511 nm</td>
<td>6$^{d}$</td>
<td></td>
<td>39 200$^d$</td>
<td>0.66$^d$</td>
</tr>
<tr>
<td>F64L, S65T$^e$</td>
<td>488 nm</td>
<td>507 nm</td>
<td></td>
<td>35$^e$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S65T, V163A$^f$</td>
<td>395 nm/473 nm</td>
<td>507 nm</td>
<td></td>
<td>33$^f$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BFP (P4) Y66H$^d$</td>
<td>383 nm</td>
<td>447 nm</td>
<td></td>
<td></td>
<td>13 500$^d$</td>
<td>0.21$^d$</td>
</tr>
<tr>
<td>P4-3 Y66H, Y145F$^d$</td>
<td>381 nm</td>
<td>445 nm</td>
<td></td>
<td></td>
<td>14 000$^d$</td>
<td>0.38$^d$</td>
</tr>
<tr>
<td>sg12</td>
<td>398 nm</td>
<td>509 nm</td>
<td>9$\times$–12$\times$</td>
<td></td>
<td>18 300$^d$</td>
<td>0.85</td>
</tr>
<tr>
<td>sg11</td>
<td>471 nm</td>
<td>508 nm</td>
<td>19$\times$–38$\times$</td>
<td></td>
<td>12 700$^d$</td>
<td>0.84</td>
</tr>
<tr>
<td>sg25</td>
<td>474 nm</td>
<td>509 nm</td>
<td>50$\times$–100$\times$</td>
<td></td>
<td>26 200$^d$</td>
<td>0.76</td>
</tr>
<tr>
<td>sg42</td>
<td>386 nm</td>
<td>450 nm</td>
<td></td>
<td>27$^g$</td>
<td>15 600$^d$</td>
<td>0.24</td>
</tr>
<tr>
<td>sg49</td>
<td>386 nm</td>
<td>450 nm</td>
<td></td>
<td>37$^g$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sg50</td>
<td>386 nm</td>
<td>450 nm</td>
<td></td>
<td>63$^g$</td>
<td>16 500$^d$</td>
<td>0.25</td>
</tr>
</tbody>
</table>

$^a$Increase in green fluorescence is in comparison to wild-type GFP, and increase in blue fluorescence is in comparison to BFP(Y66H). In each case, the optimal filter combination was used as described in Materials and Methods.

$^b$Quantum yields are as reported by Palm et al. (19) except where noted otherwise.

$^c$Absorption coefficients at excitation maximum as reported by Palm et al. (19) except where noted otherwise.

$^d$Values as reported by Heim et al. (10)

$^e$Values as reported by Cormack et al. (3). Emission measured for excitation at 488 nm for both wild-type GFP and F64L, S65T.

$^f$Values as reported by Siemering et al. (25). Emission measured for excitation at 488 nm.

Figure 2. FACS analysis and sorting of living cells transfected with BFP/GFP. 293 cells were transfected with 2 µg of either pCMV-GFPsg25 (A) or pCMV-BFPsg50 (B) alone or in combination (C) and analyzed 24 h later by FACS. An argon and a UV laser were used for the detection of green and blue cell populations, respectively. Co-transfection resulted in cells emitting a combination of both signals. The GFP- and BFP-transfected cells were mixed together (D) and sorted by gating on the green or blue signal. The successful re-separation into homogeneous green (E) and blue (F) cell populations is demonstrated by analyzing the sorted cells.
or low expressors to avoid possible artifacts of overexpression. In addition, GFP helps to continuously monitor the expression of the trans gene to detect possible changes in expression levels caused, for example, by promoter inactivation.

**GFP/BFP as Protein Tags**

Many fusion proteins with GFP and BFP have been studied by us and others. In general, GFP fusions assume the localization and functional properties of the fusion partner. The two-color system based on the mutant GFP and BFP allows the study of differential gene expression and protein-protein interactions in the same cell. As an example of dual tagging, Figure 3 illustrates the localization of the GFPsg25- and BFPsg50-tagged HIV-1 proteins Gag and Rev. Upon coexpression in human cells, Rev-BFPsg50 is localized primarily in the nucleolus. The presence of the functional Rev-BFPsg50 protein is prerequisite for high levels of Gag-GFPsg25 protein production, which accumulates in the cytoplasm. Using an argon (Figure 3A) or UV (Figure 3B) laser for the detection of GFPsg25 or BFPsg50, respectively, we were able to visualize independently the different tagged proteins in the same cell. This system can be used to screen for compounds that specifically affect the function of the HIV-1 Rev protein and present an example of a cell-based, functional dual-color fluorescence assay. Rizzuto et al. (22) described dual labeling of subcellular structures using GFP(Ser65Thr) and BFP(Tyr66His, Tyr145Phe). They reported that the BFP signal was weak and strongly affected by bleaching. In addition, the overlap in excitation maximum of the GFP and BFP pair used made an independent detection of the labeled organelles in the same cell difficult. BFPsg50, on the other hand, was readily detectable, not affected significantly by bleaching and could be detected independently from GFPsg25.

Tagging several proteins revealed that the fluorescent signal varied from very bright fusions to constructs that gave only dim signals. The weak fluorescence could be caused by a protein conformation unfavorable for GFP function and/or by factors influencing the expression level of the hybrid proteins (e.g., translation efficiency, protein maturation, degradation). In such cases or in situations in which the levels of protein expression are low (e.g., stable cell lines), the availability of the strong fluorescent mutants described in this report is essential for fusion protein detection and study. This is illustrated in Figure 4 by comparing stable cell lines expressing the cholecystokinin receptor (CCKAR) (29) tagged with GFPsg25 or...
with the commonly used GFPS65T (10). Using isogenic expression constructs, CCKAR-GFPS65T was detectable by transient transfections in Cos-1 cells (29). However, it was almost undetectable in stable cell lines (Figure 4A). Therefore, tagging CCKAR with GFPsg25 provided the signal intensity necessary for studying receptor trafficking and localization (Figure 4C). Using a ligand-binding assay, it was estimated that $1.75 \times 10^5$ GFP-tagged receptor molecules were present on the cell surface, providing sufficient signal intensity to be detected easily by confocal laser scanning microscopy.

GFP combines the properties of commonly used protein tags with the advantage of easy detection in living cells and extracts. Using purified bacterial-expressed GFP, we generated polyclonal rabbit anti-GFP sera that allowed us to also analyze GFP fusion proteins by indirect immunofluorescence, immunoblotting and immunoprecipitation (33).

Detection of Single-Copy Gene Expression Using Retroviral Vectors

The enhanced fluorescence of the GFP mutants allowed the detection of low-copy GFP gene expression not only by itself (16) but also as a fusion with heterologous proteins like aminoglycosyl phosphotransferase (Neo). By creating retroviral vectors, we could detect single-copy gene expression of GFPsg25 alone or as a fusion with Neo, even when driven by a weak promoter such as pgk (1), by FACS analysis (Figure 5) and microscopic observation. Southern blot analysis verified that the GFP signal was expressed from a single-copy gene (data not shown). The positive cells were G418-resistant, demonstrating the biological activity of the tagged Neo gene. The retroviral vector containing the GFPsg25-Neo fusion facilitates the selection and identification of transduced cell clones by FACS in the absence or presence of drug selection.

Fluorescent Proteins as Reporter Genes

Because the bright GFP mutants are easy to detect, they represent ideal reagents for the development of new and convenient functional assays. Transient transfection of plasmid DNA is a widely used system to study transcriptional regulation, protein localization or signal transduction. Quantitative assays for transfection efficiencies and gene expression using various markers have been established (6,9). Using GFPsg25, we set up a simple assay for gene transfer by quantitating GFP fluorescence not only in cell extracts but also in living cell monolayers applying a Cytofluor II plate reader. In contrast to other markers like luciferase or chloramphenicol acetyltransferase (CAT), the assay is easy and fast. Sequential measurements in living cells can be performed, and the protein is detectable as early as 6 h post-transfection. Figure 6 shows an example for optimizing transfection conditions by GFP quantitation in living cells. The fluorescent signal increased with
increasing amounts of transfected plasmid DNA until an optimum was reached. The plate reader was calibrated using purified GFPsg25 or BFPsg50, and the signal/protein ratio increased linearly at protein concentrations ranging from 1 to 100 µg/mL (data not shown). The simultaneous measurement of 96 different cell extract samples can be performed in less than 5 min.

CONCLUDING REMARKS

The variety of applications already established or presented in our study justifies the recent popularity of GFP as a powerful tool in molecular biology. Because the protein structures of the wild-type (30) and mutant GFPS/BFPs presented above have recently been solved (19), the rational design of brighter mutants or mutants with different spectral properties can be undertaken. The spectral characteristics of GFPsg25 and BFPsg50 make them an interesting pair of fluorescent probes for energy transfer experiments (FRET). In this context, it will be important to design mutants with diminished or enhanced GFP-GFP interactions and to develop functional assays based on FRET. Combining the autofluorescent proteins with synthetic fluorescent dyes and probes will result in a multicolor approach to set up new standard assays in molecular biology or drug testing. The development of the various GFP and BFP mutants together with the knowledge of the GFP structure promise certainly a “bright” future.

ACKNOWLEDGMENTS

We thank B.K. Felber for suggestions and discussions. Research sponsored by the National Cancer Institute, DHHS, under contract with ABL. The contents of this publication do not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organizations imply endorsement by the U.S. Government.

REFERENCES


Adress correspondence to:
George N. Pavlakis
ABL-Basic Research Program
Bldg. 535, Rm. 210
NCI-FRDC
PO Box B
Frederick, MD 21702-1201, USA
Internet: pavlakis@ncifcrf.gov

Received 17 March 1997; accepted 16 October 1997.

Vol. 24, No. 3 (1998)