5

MOLECULAR BIOLOGY AND MUTATION OF GREEN FLUORESCENT PROTEIN

David A. Zacharias

The Whitney Laboratory for Marine Bioscience, University of Florida
Department of Neuroscience, St. Auqustine, FL

Roger Y. Tsien

Department of Pharmacology, University of California, San Diego, La Jolla, CA

5.1 INTRODUCTION

Likely not since the discovery by McElroy (McElroy, 1947) of the involvement of ATP in the reaction catalyzed by firefly luciferase has there been as much interest in bioluminescence as currently exists in the scientific community. The interest was reignited, in large part, by the work of Chalfie et al. (1994), first showing the usefulness of the *Aequorea victoria* GFP. While *Aequorea* was first shown to fluoresce when irradiated with ultraviolet light in 1955 (Davenport and Nichol, 1955), it was not for another 40 years that it was shown for the first time that there exists a genetically encoded reporter molecule that is detectable in the absence of an enzymatic substrate or cofactor in a variety of cell types. The fluorescent properties of GFP make it especially useful in living cells and tissue. Recently, interest has surged again because there has been an exciting expansion in the discovery and characterization of homologous fluorescent proteins from distantly related sea creatures (discussed extensively in other chapters in this book).

The protein sequence, derived from the cDNA nucleotide sequence (Prasher et al., 1992), contains 238 amino acid residues and enabled determination of the chromophore structure (Shimomura, 1979; Cody et al., 1993). The *p*-hydroxybenzylideneimidazolinone chromophore is formed by the autocatalytic cyclization of Ser65, Tyr66, and Gly67 and dehydrogenation of the tyrosine. A mechanism for formation of the chromophore, which

Green Fluorescent Protein: Properties, Applications, and Protocols, Second Edition
Edited by Martin Chalfie and Steven R. Kain

is strongly fluorescent only in the intact protein (Ward and Bokman, 1982), has been proposed (Heim et al., 1994; Cubitt et al., 1995) and can be generally described as a three-step process in which cyclization is initiated by nucleophilic attack of the nitrogen atom of Glycine at position 67 on the carbonyl carbon of serine at position 65. This reaction creates the five-membered imidazolone ring. The carbonyl oxygen of S65 is then dehydrated, and finally the Cα–Cβ bond of Y66 is oxidized to conjugate the ring systems. While these steps must occur to arrive at the mature chromophore, the energetics and physical nature of the chromophore and its immediate environment that are necessary to drive the reaction have not been completely elucidated. Changing the Arg at position 96 to Ala (R96A) slows the cyclization reaction from minutes to months, allowing the isolation and crystallization of GFP in a state just prior to cyclization of the chromophore, providing a fixed image of a physical state of the protein that would be difficult to capture otherwise (Barondeau et al., 2003). This experimental result led the authors to propose a "conjugation-trapping mechanism" in which the thermodynamically unfavorable cyclization reaction of the chromophore is coupled to an electronic conjugation trapping reaction to yield the mature chromophore (Barondeau et al., 2003).

Nucleotide sequences derived from *Aequorea* indicate that at least five variants of GFP exist (Table 5.1). Four of the DNA sequences are encoded by cDNAs, while the fifth is encoded in the exons of a gfp gene. The variants differ generally by conservative amino acid replacements, suggesting they may have nearly identical physical properties. One of the genomic clones with GFP sequence contains three exons that can be matched to the

TABLE 5.1. Heterogeneous Amino Acid Residues Derived from the gfp Nucleotide Sequences

Residue Position	Locus1				
	AEVGFPA2	AEVGFP3	AVGFP14	AVGFP25	AEVGFPB6
14	Ile	Ile	Val	Ile	Ile
25	His	Gln	Gln	Gln	His
30	Ser	Ser	Ser	Arg	Ser
45	Lys	Lys	Asn	Lys	Lys
84	Phe	Phe	Phe	Leu	Phe
100	Phe	Tyr	Tyr	Tyr	Phe
108	Thr	Thr	Thr	Thr	Ser
141	Leu	Met	Met	Met	Met
154	Ala	Gly	Gly	Gly	Ala
157	Gln	Pro	Pro	Pro	Gln
172	Glu	Lys	Lys	Lys	Glu
209	Lys	Lys	Lys	Gln	Lys
212	Asn	Asn	Asn	His	Asn
213	Glu	Glu	Glu	Gly	Glu
219	Val	Ile	Ile	Val	Ile
226	Ala	Ala	Ala	Ser	Ala
228	Gly	Gly	Arg	Gly	Gly

1 As assigned by GenBank.
2 Derived from the gfp10 cDNA reported by Prasher et al. (1992), Accession No. M62653.
3 Derived from the gfp cDNA reported by Inouye & Tsuiji (1994), Accession No. L29345.
4 Derived from a cDNA submitted to GenBank by Watkins & Campbell, Accession No. X83959.
5 Derived from a cDNA submitted to GenBank by Watkins & Campbell, Accession No. X83960.
6 Derived from the exons of the gene encoded by gfp2 (Prasher et al., 1992), Accession No. M62654.

Source: With permission from the Annual Review of Biochemistry, volume 67 © 1998 by Annual Reviews www.annualreviews.org.

cDNA, while a fourth exon must exist to account for the 5′ end of the cDNA (Prasher et al., 1992). The tripeptide encoding the chromophore is located near the 3′ end of exon II (Prasher et al., 1992). A similarly large number of sequence variants exists in corals, and systematic analysis of the sequences of these variants from an evolutionary perspective is shedding considerable light on the genetic basis behind the diversity in coloration (Kelmanson and Matz, 2003; Chapter 3, this volume).

A flood of activity describing the use and optimization of the wild-type GFP and its modification followed a report (Chalfie et al., 1994) where it was shown that the chromophore forms when GFP is expressed from its cDNA in a prokaryote (*E. coli*) or a eukaryote (*C. elegans*); for reviews see Prasher (1995), Simon (1996), Tsien (1998), Zacharias et al. (2000), Matz et al. (2002), Zhang et al. (2002), Zimmer (2002), Miyawaki (2003), and other chapters in this volume. The earliest modifications of wild-type GFP were directed at improving the ability to express the protein in mammalian cell systems at 37°C. Other improvements that followed in rapid succession included simplification of the excitation spectrum, increasing the fluorescence intensity of the protein by increasing the extinction coefficient as well as increasing the speed and efficiency of folding, making different colors or spectral mutants, reducing sensitivity to pH and halides, and most recently reducing the tendency of the protein to dimerize. In this chapter we will describe, in general terms, the methods and strategies used to generate existing mutants of *A. victoria* GFP. We then present a classification of *A. victoria* GFPs based on the chemical structure and photophysical behavior of their chromophores and then catalog and describe many of the useful and interesting mutants that have been created and characterized to date.

5.2 MUTATIONAL STRATEGIES

Many strategies have been used to generate mutations in GFP; the degree of sophistication and power has risen considerably since the earliest methods were developed, partly because GFP itself is such a convenient reporter and testbed for the development of improved mutagenic strategies. We will present the strategies roughly in order of increasing degree of complexity: (1) random mutagenesis by chemical mutagens, (2) error-prone polymerase chain reaction (PCR) over the entire or a portion of the coding sequence, (3) deliberate site-directed mutations to test specific hypotheses (eventually aided by the crystal structure), (4) randomization (possibly with codon biases) of a predetermined, limited stretch of residues, (5) DNA "shuffling," (6) heteroduplex recombination, and (7) directed evolution to change or incorporate into fluorescent proteins (FPs) certain properties more rapidly than could be accomplished by rational design methodology.

1. *Random Mutagenesis by Chemical Mutagens.* The cDNA vector encoding GFP can be treated with mutagens such as hydroxylamine or nitrous acid before transfection into a host (Sikorski and Boeke, 1991). As far as mutagenic strategies for GFP are concerned, this method has been largely supplanted by protocols that rely solely on molecular biology.
2. *Error-Prone PCR.* cDNA can be amplified by PCR under conditions where the polymerase fidelity is compromised—for example, by replacing Mg^{2+} with Mn^{2+} in the reaction mixture and/or restricting the concentration of one nucleoside triphosphate at a time (Muhlrad et al., 1992). Error-prone PCR has the advantages of confining the mutations only to a desired region of the gene between the two

primers and in producing a greater variety of mutations at the nucleic acid level. Nevertheless, most mutations are typically changes of just one base in a relevant codon, so that amino acids for which codons differ in two or three positions from the starting codon are much less likely to be accessed. The importance of Y66, I167, T203, and E222 (Heim et al., 1994; Ehrig et al., 1995) were discovered this way, as well as some of the folding mutations such as Y145F, M153T, and V163A (Heim and Tsien, 1996), C70V (Zapata-Hommer and Griesbeck, 2003), and F46L (Nagai et al., 2002). This method is presumably random at the nucleic acid level and is a strategy that is still used commonly as a first step in some of the methods described below.

3. *Deliberate Site-Directed Mutations.* This method is generally used to mutagenize GFP to test specific hypotheses about certain amino acid replacements. This was the approach used to make monomeric versions of GFP (Zacharias et al., 2001). The coordinates from a crystal structure of dimeric GFP (Yang et al., 1996a) were used to predict which residues were responsible for the weak tendency to dimerize (Table 5.5; Fig. 5.4). The exchange of the aliphatic residues for positively charged residues was effective in each instance of reducing or eliminating the ability to dimerize. An earlier example of systematic single amino acid replacements was exchanging Thr203 for aromatic residues such as Tyr (Delagrave et al., 1995; Ormo et al., 1996). See Table 5.4 for a description of the resulting changes in the spectral characteristics imparted on the protein by these substitutions. Another example is testing the effect of adding one specific mutation known to impart a desirable property, such as efficient folding, to another mutant version of GFP to see if the benefits would be additive (Nagai et al., 2002).

4. *Randomization of a Predetermined Stretch of Amino Acids.* Delagrave et al. (1995) randomized positions 64–65 and 67–69 with the random codon NNK, where N is any base and K is G or T. They found (F64M, S65G, Q69L) as their favored mutant, code-named "RSGFP4" (Delagrave et al., 1995). Later, Cormack et al. (1996) randomized positions 55 through 74 with a 10% probability of substituting NNK for each wild-type codon. This produced three different mutants: (F64L, S65T), (S65A, V68L, S72A), and (S65G, S72A). All of these have excitation spectra shifted to 470–490 nm because of their replacement of Ser65, plus one or more mutations that improve folding efficiency. Subsequent crystal structures have provided some guidance by defining proximity to the chromophore in three-dimensional space rather than as a distance defined by the sequence of the amino acids. This knowledge has allowed the rational design of mutants with desired characteristics such as the introduction of simple metal-binding sites engineered into the staves of the beta barrel [e.g., Richmond et al. (2000)] (see Fig. 5.5A).

5. *DNA Shuffling.* This method (Crameri et al., 1996), reviewed in Giver and Arnold (1998) and in Minshull and Stemmer (1999), is one way to recombine an existing set of mutations spread throughout the gene of interest or to combine different mutations from each of two or more copies of the same gene. For example, it would be an ideal method to mix the individual (or sets of) beneficial mutations, such as those that improve folding, and produce composite proteins containing multiple mutations that might synergize with each other. In brief, a collection of cDNAs, each of which contains only one or a few mutations, is subjected to limited digestion with DNAse. The fragments are re-annealed in a PCR-like reaction with

nucleoside triphosphates but without primers. This process allows fragments with different mutations to re-sort with each other, repairs the breaks, and introduces additional point mutations. The resulting mixture is finally amplified by conventional PCR with primers, spliced into an expression vector, and screened. Obviously this cycle can be repeated as many times as desired. Using this technique, Crameri et al. (1996) [reviewed in Minshull and Stemmer (1999)] achieved a significant increase in brightness over wild-type GFP. Inspection of the mutations they generated (F99S, M153T, V163A) found that two of the three are ones found by other methods of random mutagenesis. So the individual mutants produced by DNA shuffling are not necessarily unique, but the potential for synergistically recombining them is a main attraction.

6. *Heteroduplex Recombination.* This is a labor-intensive twist on the polymerase-based DNA shuffling method. The general strategy consists of creating libraries of chimeric DNA sequences derived from homologous, nonidentical parental sequences (Volkov et al., 1999). cDNA heteroduplexes formed by denaturation and annealing are transformed into bacteria where the endogenous DNA repair system fix regions of nonidentity within the heteroduplexes, thereby creating a library of sequences comprised of elements of all parents. Among the potential advantages of this method is the possibility of simultaneously combining desired features from many individual, homologous cDNAs.
7. *Directed Evolution.* This methodology requires fairly loose definition. Realistically, the process can comprise one, several, or theoretically all of the mutagenic processes listed above. Generally, it is the repetitive mutagenesis of a cDNA encoding a protein that is subsequently expressed and exposed to some selective pressure or development criterion in an effort to capture a protein that has altered characteristics that are desirable. The combination of mutagenesis on many levels with selective pressure can yield remarkable changes in protein activity by incorporation of mutations that might never have been made rationally. Some examples where directed evolution has proven useful include the following: (a) the creation of allosteric GFP biosensors for beta-lactamase activity (Doi and Yanagawa, 1999), (b) to monitor folding efficiencies of proteins to which GFPs are fused (Waldo et al., 1999), and (c) perhaps the most dramatic example, the generation of a monomeric version the tetrameric red fluorescent protein [Campbell et al. (2002); see also Gibbs et al. (2001)] and subsequently a rainbow of color variatious from the monomer RFP (Shaner et al., 2004). Using fluorescent proteins as a reporter of the success of mutagenic strategies has probably decreased the development time and increased the efficiency of protocols designed to alter protein activity by mutagenesis.

5.3 SCREENING METHODS

1. Visual screening requires the least expensive equipment. The minimum is a source of excitation light, either 365 nm from a "black-light" illuminator or ~480 nm from a xenon lamp and interference filter or monochromator that can illuminate Petri dishes. Observation is through a long-pass filter. In the case of UV excitation, an external UV-blocking filter is advisable for health reasons. For blue excitation, one can tape pieces of yellow or orange gelatin filters (e.g., Kodak Wratten filters) over

disposable lab safety glasses. Alternatively, one can scan the dish with a fluorescence microscope with a low-power objective, typically 4×, using the appropriate filter cube to select excitation and emission wavelengths. The main advantage of visual screening is low capital cost; however, it is both laborious and insensitive to small changes in wavelength below the capacity of human color vision to recognize easily.

2. Video imaging with a frame grabber and computerized false-color display of spectral features was dubbed digital imaging spectroscopy by Youvan et al. (1995). This method is similar to the excitation or emission ratioing now common in imaging physiological indicators inside cultured mammalian cells, except that the optics illuminate and observe a macroscopic rather than microscopic field of view. Digital imaging is more objective and sensitive than human vision to small changes in spectra, but the combined optical and computer setup is relatively expensive. Custom-designed instruments are becoming commonplace and being used successfully in labs (Delagrave et al., 1995; Sawano and Miyawaki, 2000; Griesbeck et al., 2001; Nagai et al., 2002; Neal Woodbury, personal communication) interested in developing fluorescent proteins for uses in cell biology. At least one platform is commercially available from Genetix, www.genetix.com.
3. Fluorescence-activated cell sorting is one of the most obvious high-throughput screening methods to search for useful mutants, especially when the sorting criterion is an unusual ratio of emissions at two wavelengths and no information on the spatial distribution of the protein is necessary. The earliest example of its use was to screen for random mutants that increased fluorescence intensity (Cormack et al., 1996). Two recent examples demonstrate the power of this screening technique, in combination with other innovative, directed-evolution techniques, to generate a broad range of spectral mutants of monomeric red fluorescent protein (Shaner et al., 2004; Wang et al., 2004).
4. High-throughput microscopy (HTM) is a methodology that is finding a home in large-scale drug discovery programs and, because of the increased abundance of commercially available, lower-priced and more powerful detection platforms, is becoming feasible for the academic researcher as well. Such systems provide throughput similar to that of FACS, but have the ability to quantify cellular morphology including the spatial distribution and concentration of fluorescently labeled molecules within the cells. While no group has as yet used the technology to discover useful mutants of GFP, such a task is certainly well within the capability of the technology. In many cases, GFP and GFP-fusion proteins fold with better efficiency when expressed in mammalian cells versus bacterial expression systems. The trapping of nascent proteins in inclusion bodies in bacterial expression systems could foreseeably prevent discovery of useful, but hard-to-fold, mutants of GFP. HTM would allow screening of very large mutagenic libraries of GFP expressed in mammalian cells. The ability to easily configure the excitation and emission characteristics in some HTM systems such as the EIDAQ100 (www.Q3DM.com) would allow users to screen for and select mutants in the exact format that the mutants would be imaged in later experiments. While existing protocols have not been adapted for the isolation of mutant cDNAs when using HTM to identify interesting mutants expressed by vertebrate cells in culture, there appear to be no technological barriers to doing so.

5.4 CLASSIFICATION OF SPECTRAL MUTANTS BY CHROMOPHORE TYPE

Classification of spectral mutants of *Aequorea* GFP by chromophore type has provided a logical and useful guide to users of the GFPs since its introduction in a comprehensive form in 1998 (Tsien, 1998). Here we update the compendium to include unique mutants described since 1998.

All known GFP variants may be divided into seven classes based on the distinctive component of their chromophores (Table 5.2): class 1, or wild type, is a mixture of neutral phenol and anionic phenolate; class 2 is a phenolate anion; class 3 is a neutral phenol; class 4 is a phenolate anion with stacked π-electron system; class 5 is an indole; class 6 is an imidazole; and class 7 is a phenyl. Each class has distinct wavelength distributions for their excitation and emission spectra (Table 5.1). The first four classes are derived from polypeptides with Tyr at Position 66, while classes 5–7 result from Trp, His, and Phe at position 66. The chromophore structures for each class are shown in Fig. 5.2 together with typical respective fluorescence spectra.

Class 1: Wild-Type Mixture of Neutral Phenol and Anionic Phenolate [Wild-Type Green Fluorescent Proteins (wtGFPs)]

Wild-type GFP has the most complex spectra of all the *Aequorea* GFPs. The major excitation peak is at 395 nm and is about three times larger than the secondary peak at 475 nm. In normal solution, excitation of the major peak gives rise to an emission maximum at 508 nm whereas exciting the secondary, minor peak results in an emission maximum at 503 nm (Heim et al., 1994). The fact that the emission maxima are dependent on the excitation maxima indicates that there are at least two chemically distinct populations of chromophore that do not equilibrate within the lifetime of the excited state. Though the nature of the transition from one state to the other is not completely understood, the simplest explanation is that the minor peak at 475 nm arises from a deprotonated, anionic chromophore while the major peak at 395 nm represents a protonated or neutral chromophore (Heim et al., 1994; Cubitt et al., 1995; Niwa et al., 1996; Bell et al., 2000). A third, transitional, intermediate state has been identified using time-resolved fluorescence spectroscopy and hole burning (Chattoraj et al., 1996; Creemers et al., 1999; Zimmer, 2002), but exactly what this state represents and what the exact mechanism by which this occurs is still under investigation.

Class 2: Phenolate Anion in Chromophore [Green Fluorescent Proteins (GFPs)]

GFPs of this class are perhaps the most commonly used due to their relative brightness, simple excitation, and emission spectra and the fact that the spectra match closely those of fluorescein. The prototypical GFP of this class incorporates the mutation S65T. This substitution results in a dramatic increase in the amplitude (five- to sixfold) and a red shift (from 470–475 nm to 489–490 nm) of the anionic peak while also dramatically suppressing the neutral phenol peak at 395 nm (Delagrave et al., 1995; Heim et al., 1995; Cheng et al., 1996). Formation of the mature chromophore in S65T (Heim et al., 1995) was about four times faster than in the wild type, and folding was fairly efficient at low temperatures as in the wild type at temperatures lower than 37°C. For this reason, a lot of effort was

TABLE 5.2. Spectral Characteristics of the Major Classes of GFP Variants

Mutation[a]	Common Name	λexc (ε)[b]	λem (QY)[c]	Reference[d]
Class 1, wild type				
-None or Q80R	Wild Type	395–397 (25–30) 470–475 (9.5–14)	504 (0.79)	(Patterson et al., 1997; Ward, 1997)
-F99S, M153T, V163A	Cycle 3	397 (30) 475 (6.5–8.5)	506 (0.79)	(Patterson et al., 1997; Ward, 1997)
Class 2, phenolate anion (green fluorescent proteins-GFPs)				
-S65T		489 (52–58)	509–511 (0.64)	(Patterson et al., 1997; Ward, 1997; Cubitt et al., 1999)
-F64L, S65T	EGFP	488 (55–57)	507–509 (0.60)	(Patterson et al., 1997; Ward, 1997; Cubitt et al., 1999)
-F64L, S65T, V163A	—	488 (42)	511 (0.58)	(Cubitt et al., 1999)
-S65T, S72A, N149K, M153T, I167T	Emerald	487 (57.5)	509 (0.68)	(Cubitt et al., 1999)
Class 3, neutral phenol				
-S202F, T203I	H9	399 (20)	511 (0.60)	(Cubitt et al., 1999)
-T203I, S72A, Y145F	H9-40/Sapphire	399 (29)	511 (0.64)	(Zapata-Hommer & Griesbeck, 2003)
-Q69M, C70V, V163A, S175G	T-Sapphire	399 (44)	511 (0.60)	(Zapata-Hommer & Griesbeck, 2003)
Class 4, phenolate anion with stacked p-electron system (yellow fluorescent proteins-YFPs)				
-S65G, S72A, T203F	—	512 (65.5)	522 (0.70)	(Cubitt et al., 1999)
-S65G, S72A, T203H	—	508 (48.5)	518 (0.78)	(Cubitt et al., 1999)
-S65G, S72A, T203Y	Topaz	514 (94.5)	527 (0.60)	(Cubitt et al., 1999)
-S65G, V68L, S72A, T203Y	10C	514 (83.4)	527 (0.61)	(Cubitt et al., 1999)
-10C + Q69K	10CQ69K	516 (62)	529 (0.71)	(Cubitt et al., 1999)
-10C + F64L, M153T, V163A S175G	Venus	515 (92.2)	528 (0.57)	(Nagai et al., 2002; Rekas et al., 2002)
-10C + V68L, Q69M	Citrine	516 (77)	529 (0.76)	(Griesbeck et al., 2001)

Class 5, indole chromophore (cyan fluorescent proteins-CFPs)				
-Y66W	—	436	485	(Heim et al., 1994)
-Y66W, N146I, M153T, V163A	W7	434 (23.9) 452 (shoulder)	476 (0.42) 505 (shoulder)	(Cubitt et al., 1999)
-F64L, S65T, Y66W N146I, M153T, V163A	W1B/ECFP	434 (23.9) 452 (shoulder)	476 (0.40) 505 (shoulder)	(Cubitt et al., 1999)
-S65A, Y66W, S72A N146I, M153T, V163A	W1C	435 (21.2)	495 (0.39)	(Cubitt et al., 1999)
Class 6, imidazole in chromophore (blue fluorescent proteins- BFPs)				
-Y66H	BFP	384 (21)	448 (0.24)	(Cubitt et al., 1999)
-Y66H, Y145F	P4-3	382 (22.3)	446 (0.30)	(Cubitt et al., 1999)
-F64L, Y66H, Y145F	EBFP	380–383 (26.3–31)	440–447 (0.17–0.26)	(Patterson et al., 1997; Cubitt et al., 1999)
Class 7, phenyl in chromophore				
Y66F		360	442	(Cubitt et al., 1995)

Source: Ann. Rev. Biochem.

[a] Substitutions from the primary sequence of GFP are given as the single-letter code for the amino acid being replaced, its numerical position in the sequence, and the single-letter code for the replacement. Note that many valuable mutants have been left out of this table for reasons of brevity and because quantitative spectral and brightness data were not available; therefore omission does not imply denigration. Phenotypically neutral substitutions such as Q80R, H231L have been omitted.

[b] λ_{exc} is the peak of the excitation spectrum in units of nanometers. E in parentheses is the absorbance extinction coefficient in units of $10^3\,M^{-1}\,cm^{-1}$. Estimates of extinction coefficients have tended to increase as expression and purification are optimized; obsolete older values have been omitted. Two numbers separated by a dash indicate a range of estimates from different authors working under slightly different conditions. Two numbers on separate lines indicate two distinct peaks in the excitation spectrum and have been designated as "shoulders" as they are also the minor of the two peaks.

[c] λ_{em} is the peak of the emission spectrum in units of nanometers. QY in parentheses is the fluorescence quantum yield, which is dimensionless. The best figure of merit for the overall brightness of properly matured GFPs is the product of e and QY. See footnote to b for explanation of values.

[d] References only for the quantitative spectral and brightness data. References to the origin and use of the mutants have been omitted for lack of space.

put into finding mutants that incorporated the benefits of S65T as well as having increased brightness at 37°C. Folding mutations that have worked well in combination with S65T include F64L (Cormack et al., 1996) and V163A (Kahana and Silver, 1996), V68L (Cormack et al., 1996), and I167T. Other mutations that improve folding used in combination with S65T have also proved useful. The photophysical mechanism by which replacement of Ser65 promotes chromophore ionization is reviewed in Tsien (1998) and Zimmer (2002).

Class 3: Neutral Phenol in Chromophore (GFP400 or Sapphire)

In the class 2 mutants, the protonated state of the chromophore is favored and the anionic state is suppressed. Conversely, suppression of this protonation can be accomplished by mutating Tyr203 to Ile, a change that basically eliminates the excitation peak at 475 nm, leaving only the lower wavelength peak at 399 nm, with the emission still occurring at 511 nm. This mutant produces the greatest separation of the excitation and emission maxima for any of the *Aequorea* GFPs. This mutant called, sapphire [Turbo sapphire, GFP400, or H9-40 when combined with other mutations that improve the folding (Zapata-Hommer and Griesbeck, 2003)], has the greatest Stokes shift (greater than 100 nm) of any GFP mutant. This protein has proved to be the only published, successful FRET donor when dsRED was used as a FRET acceptor (Mizuno et al., 2001). The complexity of the excitation spectra of dsRED and the fact that it has been, until recently (Campbell et al., 2002), an obligate tetramer composed of individual subunits with spectra very similar to GFP has meant that only a donor such as sapphire, with such a large Stoke's shift and blue-shifted excitation maximum, would do the job without unintentional cross-excitation of the immature, spectrally GFP-like, monomeric members of the dsRED tetramer. A full discussion of how the ionization state is influenced by the interior shape of the barrel of GFP and the sizes and shapes of the side chains of residue 203 can be found in Tsien (1998).

Class 4: Phenolate Anion with Stacked π-Electron System [Yellow Fluorescent Proteins (YFPs)]

This class of chromophore, the furthest red-shifted of all the *Aequorea* mutants, results from stacking an aromatic ring next to the phenolate anion of the chromophore. Residue 203 is properly positioned to provide an aromatic side chain. To promote ionization of the chromophore, position 65 must be Gly or Thr instead of Ser. All four aromatic residues at position 203 (His, Trp, Phe, and Tyr—in order of least to greatest red-shifting potential) increase the excitation and emission spectra up to 20 nm (Ormo et al., 1996). See Table 5.3. All of them cause a red shift in both the exciation and emission spectra by adding polarizability around the chromophore and by extending the π–π interaction, thereby reducing the excited-state energy level. These replacements were rationally designed, using information from the crystal structures of S65T (Ormo et al., 1996; Yang et al., 1996b). Given that all three of the nucleotides of the codon for Thr (ACA) must be replaced in order to swap-in any one of the aromatic residues, it is unlikely that these substitutions would have been discovered by random mutagenesis. The crystal structure of the T203Y mutant (Wachter et al., 1998) confirms that the aromatic side chain of Tyr stacks next to the chromophore. Mutations at residue Q69 like Q69K (Cubitt et al., 1999) and Q69M (Heikal et al., 2000; Griesbeck et al., 2001) cause an additional 1- to 2-nm red shift, making these the "reddest" mutants of *Aequorea* GFP yet discovered.

TABLE 5.3. Effects of Different Aromatic Amino Acids at Position 203 of GFP[a]

Mutations	λ_{exc}	ε	λ_{em}	Comments and References
T203H	512	19.4	524	(Ormo et al., 1996)
T203H, S65G, S72A	508		518	Cubitt (unpublished)
T203F, S65G, S72A	512		522	(Dickson et al., 1997)
T203Y, S65T	513	14.5	525	(Ormo et al., 1996)
T203Y, F64L, S65G, S72A	513	30.8	525	(Ormo et al., 1996)
T203Y, S65G, V68L, S72A	513	36.5	527	(Ormo et al., 1996) quantum yield 0.64
T203W, S65G, S72A	502	33	512	(Ormo et al., 1996)

[a] Abbreviations: λ_{exc}, excitation maximum in nanometers; e, extinction coefficient in $M^{-1} cm^{-1}$ at_λ_{exc}; λ_{em}, emission maximum in nanometers.

The mutation T203H appears to be the crucial alteration within photoactivatable GFPs (Patterson and Lippincott-Schwartz, 2002), discussed later under photoisomerization.

Class 5: Indole in Chromophore Derived from Y66W [Cyan Fluorescent Proteins (CFPs)]

An indole is formed in the chromophore when Tyr66 is replaced with a Trp (Y66W). The excitation and emission wavelengths are intermediate between the anionic phenolate and the neutral phenol-type chromphores, giving the protein a cyan appearance (see Table 5.2). The indole chromophore is bulky and requires a large number of additional mutations to restore the protein to reasonable brightness (Heim and Tsien, 1996). Many of these additional, beneficial mutations are the same as those required to restore brightness to the spectral mutants of other classes and are likely functioning in the same capacity.

CFPs usually have two distinct peaks in their excitation spectra, two peaks in their emission spectra, and two lifetimes in their excited state decay. This microheterogeneity has been attributed to two distinct conformations visible in the crystal structure, associated with differing solvent exposures of Y145 and H148 ((Hyun Bae et al., 2003)). Therefore these two residues have been varied. The best improvement is the mutant S72A/Y145A/H148D, which has been dubbed "Cerulean" (Rizzo et al., 2004). Cerulean is reported to be 2.5-fold brighter than ECFP, due to improvements of about 1.5-fold in extinction coefficient and 1.7-fold in quantum yield. Cerulean is slightly more photostable than ECFP and shows single exponential kinetics for the decay of its excited state. Curiously, Cerulean retains the double-humped excitation and emission spectra of ECFP, suggesting that the molecular basis for this spectral shape remains somewhat mysterious.

Class 6: Imidazole in Chromophore Derived from Y66H [Blue Fluorescent Proteins (BFPs)]

Exchanging Tyr66 for His puts an imidazole in the chromophore (Heim et al., 1994) and results in blue-shifting the excitation and emission wavelengths yet a bit more than replacement by the indole in class 5. As such, this class has been dubbed blue florescent proteins (BFP). Crystal structures (Palm et al., 1997; Wachter et al., 1997; Palm and Wlodawer, 1999) have been solved. As has been the case in the other classes, the brightness of these proteins is improved by additional mutations such as F64L (Patterson et al., 1997) and

Y145F (Heim and Tsien, 1996; Yang et al., 1998). Optimizing the codon usage to improve expression in mammalian cells was also beneficial (Yang et al., 1996b). However, even with these improvements, BFP remains among the least-used of the spectral mutants because it is easily photobleached and is dim [due to a very low fluorescence quantum yield (Rizzuto et al., 1996; Palm and Wlodawer, 1999; Kummer et al., 2002)], and the excitation wavelengths necessary to cause fluorescence are relatively more phototoxic and induce more tissue or cellular autofluorescence.

Class 7: Phenyl in Chromophore Derived from Y66F

Replacement of Tyr66 with Phe (Y66F) results in the lowest known excitation (360 nm) and emission (442 nm) maxima (Cubitt et al., 1995). While this substitution illustrates nicely that any aromatic residue at position 66 can form a chromophore, it has been of little practical use.

The described classification above (and Table 5.2) provides information concerning the fundamental chemical and photophysical nature of each fluorophore type. Ultimately it is direct manipulation, of these fluorophores or their environment, that is the goal of most mutagenic experiments. The following sections include descriptions of how certain amino acid changes in *A. victoria* GFP affect fluorophore behavior and ultimately how, why, and when it glows.

5.5 MUTANTS THAT COVALENTLY ALTER THE FLUOROPHORE: Y66FHW

The most fundamental mutations that do not destroy fluorescence entirely are those that alter the covalent core structure of the chromophore. Such mutations are inevitably at position 66, because the crucial π-electron conjugated framework is derived from Tyr66 plus the invariant carbonyl carbon and amino nitrogen of residues 65 and 67, respectively. All the aromatic amino acids allowable by the genetic code—namely, phenylalanine, histidine, and trytophan—have been tested and found to generate fluorescent proteins of very different wavelengths (Heim et al. 1994; Heim and Tsien, 1996). The structures of the chromophores expected to be formed from these amino acids and the observed excitation and emission wavelengths are compared with wild-type GFP in Fig. 5.2. The rank order of shortest to longest excitation wavelengths is as expected from the electronic properties of the side chains. The benzene ring of phenylalanine has no electron donor group to conjugate to the electron-withdrawing carbonyl at the other end of the chromophore, so Phe gives the shortest wavelengths. The imidazole of His has a moderately electron-donating HN< group, but this is weakened by the electron-withdrawing =N— on the same ring. Next after imidazole is the nonionized phenol of tyrosine, which is responsible for the 395-nm excitation maximum of wild-type GFP. Yet more electron-rich is the >NH in the indole of tryptophan, which also has the largest conjugated system. The strongest electron-donating group is the phenolate anion of tyrosine, which is responsible for the minor excitation maximum of wild-type GFP at 470 nm. The chromophore with an ionized phenolate has about twice the extinction coefficient as that with a neutral phenol (Chattoraj et al., 1996). Therefore wild-type GFP, whose 395-nm peak is typically about three times as high as its 470-nm peak, is approximately a 6:1 mixture of neutral and anionic chromophores. This ratio has recently been confirmed by X-ray crystallography of wild-type GFP, in which two internal isomers corresponding to neutral and anionic chromophores

in an 85:15 ratio can be resolved (Brejc et al., 1997). Five unnatural tyrosine-analog substitutions (Wang et al., 2003) have also been substituted at position 66 on the background of the Cycle 3 GFP variant (Crameri et al., 1996) as a showcase for incorporation of unnatural amino acids. Similarly, the spectral properties of these mutant GFPs, including the absorbance and fluorescence maxima and quantum yields, correlate with the structural and electron-donating ability of the substituents on the amino acids (Wang et al., 2003).

Similar to most classes of GFP chromophore variants, practical applications of position-66 mutants have required additional amino acid substitutions to increase folding efficiency and fluorescence intensity (Heim and Tsien, 1996; Mitra et al., 1996; Siemering et al., 1996). The visual appearance of bacteria expressing first-generation improved versions of Y66H (blue) and Y66W (cyan) are shown in Fig. 5.3; normalized excitation and emission spectra are in Fig. 5.1. Based on crystallographic information obtained from the mutant Y66H (Palm et al., 1997; Wachter et al., 1997; Palm and Wlodawer, 1999), it was suggested that an increased flexibility around the chromophore led to a reduced quantum yield rather than poor folding efficiency as is often the case among the various classes. However, picosecond time-resolved fluorescence measurements from Y66F and

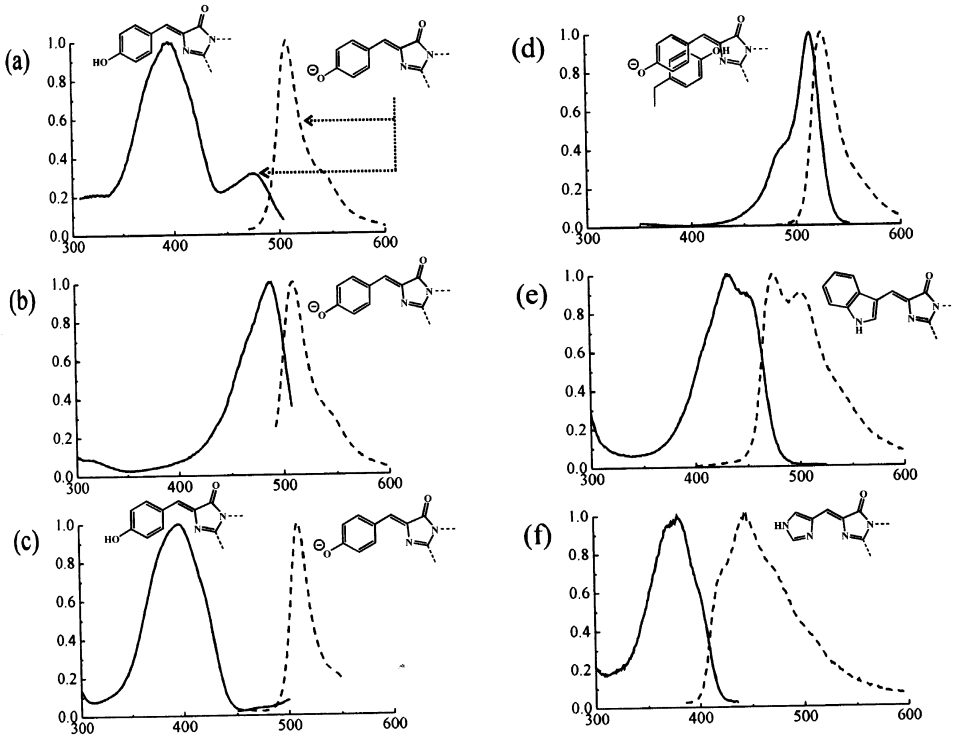

Figure 5.1. Fluorescence excitation and emission spectra for typical members of the major classes of GFP mutants together with the chromophore structures believed to be responsible for the spectra. The spectra have been normalized to a maximum amplitude of 1. When only one structure is represented in the figure, both excitation and emission spectra arise from the same state of chromophore protonation. The actual GFP proteins depicted are (a) wild type, (b) emerald, (c) H9-40, (d) topaz, (e) W1B, and (f) P4-3. The detailed substitutions within each of these variants are listed in Table 5.1.

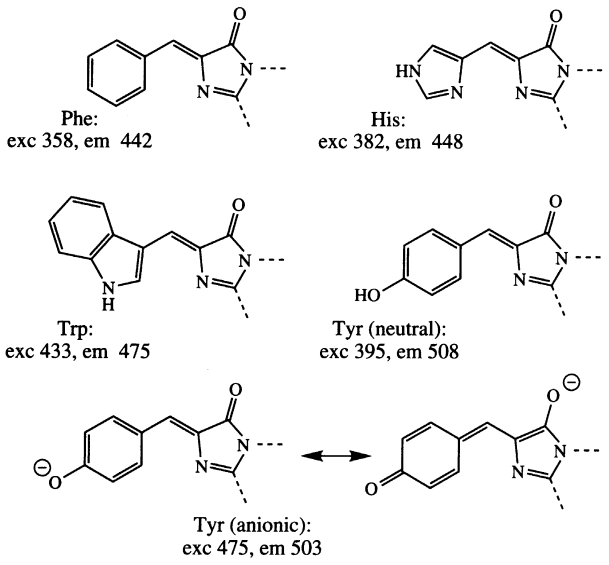

Figure 5.2. Structures and fluorescence wavelengths of fluorophores resulting from different aromatic amino acids at position 66 of GFP. Each structure is labeled by the amino acid occupying position 66 before chromophore formation. Excitation (exc) and emission (em) peak wavelengths are given in nanometers. The native amino acid, Tyr, gives two peaks depending on whether the phenolic hydroxyl is neutral or ionized. For the latter, two of the possible resonance structures are drawn. Spectral data from Heim and Tsien (1996) and Cubitt et al. (1995).

Y66H indicate that while the free volume of the chromophore, cannot be discounted, the primary channel for loss of excitation energy is dependent more on hydrogen bond interactions between the chromophore and its immediate protein environment (Kummer et al., 2002). Heterocycle formation begins even with a nonaromatic residue such as Gly at 66 (Barondeau et al., 2003), although no useful fluorescence properties have been reported for such mutations.

5.6 MUTANTS WITH AN ALTERED RATIO BETWEEN THE TWO WT EXCITATION PEAKS: I167T, T203I, T203C, E222G, S65GACTVL, Q69L; CIRCULAR PERMUTATIONS, DURABLE PHOTOISOMERIZATION

The next most fundamental mutations are those that alter the ionization of Tyr66 and change the ratio between the excitation peaks due to the neutral (~395 nm) and anionic (~470 nm) species. Mutation of Thr203 to Ile (T203I) suppresses the anionic peak (Heim et al., 1994; Ehrig et al., 1995; Zapata-Hommer and Griesbeck, 2003). This result is readily understandable from crystal structures because the hydroxyl of Thr203 donates a hydrogen bond to the chromophore phenolate (Ormo et al., 1996; Yang et al., 1996a; van Thor et al., 2002). In the neutral form of the chromophore, the hydroxyl of Thr203 rotates away from the protonated phenol (Brejc et al., 1997; van Thor et al., 2002). Ile cannot form a hydrogen bond, so the anionic phenolate would be destabilized relative to the neutral phenol. However, once the neutral chromophore absorbs light and reaches the excited state, the phenol ionizes (Chattoraj et al., 1996; Brejc et al., 1997; van Thor et al., 2002) despite

the relatively poor solvation, so the subsequent emission still peaks at 511 nm. The most recent mutant published in this class is called Turbo or "T-Sapphire" (Zapata-Hommer and Griesbeck, 2003). The mutations Q69M/C70V/V163A/S175G were added on the background of H9-40 (see Table 5.2), resulting in a protein that remained insensitive to most cellular pHs (pK_a 4.9), became fluorescent much faster, and recovered from denaturation more quickly and to a greater extent. Of these, V163A and S175G are common folding mutations and Q69M was known to suppress pH and halide sensitivity of YFPs (Griesbeck et al., 2001), C70V was a serendipitous, PCR-induced mutation that rescued the otherwise nonfluorescent H9-40 Q69M/V163A/S175G. Replacement of T203 with Cys (on the background S65T/H148G-"deGFP1" or S65T/C48S/H148C-"deGFP4") results in ratiometric sensors in which the absorbance/excitation and the emission maxima vary in a pH-dependent way between the protonated (~400 nm) and the anionic (~508 nm) state of the chromophore (Hanson et al., 2002; McAnaney et al., 2002). Interestingly, these deGFPs appear to support excited-state proton transfer through a mechanism novel for a GFP that includes Ser147 and two water molecules. It permits rapid proton transfer between the chromophore hydroxyl and the bulk solvent at high pH. This network rearranges, and Ser147 is removed from contact with the chromophore, thereby eliminating the proton relay at low pH (Hanson et al., 2002).

Many researchers have been interested in shifting the excitation to longer rather than shorter wavelengths, because ultraviolet excitation is potentially injurious to cells, excites more cellular autofluorescence, and generally requires more expensive optics and detection instrumentation. Shifts toward longer wavelengths require favoring the anionic peak. The first such mutants were I167T and I167V, which gave anionic excitation peaks about twice the amplitude of the neutral peaks (Heim et al., 1994). Given the ratio of extinction coefficients, these mutants probably contain about a 1 : 1 molar ratio of the two species. The crystal structure of GFP (Ormo et al., 1996; Yang et al., 1996a) allowed one to rationalize why Thr at 167 gives a moderately, but not overwhelmingly, higher ratio of anion to neutral than does wild-type Ile167. Thr should point its hydroxyl toward the phenolate without coming close enough to form a direct hydrogen bond. The dipole moment of the hydroxyl should therefore favor the phenolate. This rationalization, however, does not explain why Val at 167, which lacks the hydroxyl, produces much the same effect as Thr. Perhaps the smaller steric bulk of either Val or Thr relative to Ile is more important.

An effective way of suppressing the neutral species was found to be mutation of Ser65 to any of a variety of small uncharged amino acids including Gly, Ala, Cys, Thr, or Leu (Heim et al., 1995). Such mutations cause essentially complete ionization of the chromophore even in the ground state and simplify the excitation spectrum to a single peak between 470 and 490 nm, with an amplitude about six times higher than that of the 475-nm subsidiary peak of the wild-type protein. Thus at such blue excitation wavelengths, the mutants are about six times brighter per molecule than wild type. The improved brightness of such Ser65 mutants when excited in this blue region is understandable, because the wild-type protein was only about one-sixth ionized. The exact positions of the excitation and emission maxima depend on which amino acid replaces Ser65, as shown in Table 5.4. Ala gives the shortest wavelengths, Thr or Gly the longest. Very polar or bulky residues such as Arg, Asn, Asp, Phe, or Trp at position 65 do not seem to be tolerated (Heim et al., 1995).

The mutant with Thr at 65, or S65T, was chosen by Heim et al. (1995) for further analysis and use because it combined the longest peak wavelengths (489 nm excitation, 511 nm emission) with the most conservative substitution. The visual appearance of bacteria expressing S65T is included in Fig. 5.3, and its excitation and emission spectra are

TABLE 5.4. Effects of Different Amino Acids at Position 65 of GFP[a]

Amino Acid	λ_{exc}	ε	λ_{em}	Q	Comments and References
Ser (wt)	395 475	21,000 7,150	508 503	0.77	At room temperature (Heim et al., 1995); the 395- and 475-nm peaks probably represent excitation of the neutral and anionic fluorophore respectively
Ser (wt)	404 471		504 482		At 77K in 1 : 1 glycerol-water to slow down proton transfers (Chattoraj et al., 1996)
Ala	471		503		(Heim et al., 1995)
Ala	481		507		Measured in combination with V68L, S72A (Cormack et al., 1996)
Cys	479	47,400	507		(Heim et al., 1995)
Leu	484		510		R. Heim (unpublished data)
Thr	489	39,200	511		(Heim et al., 1995)
Thr		52,900			R. Heim (unpublished data)
Thr	488		507		Measured in combination with F64L (Cormack et al., 1996)
Gly	490		505		Measured in combination with F64M, Q69L (Delagrave et al., 1995)
Gly	501		511		Measured in combination with S72A (Cormack et al., 1996)
Arg, Asn, Asp Phe, Trp					Weak or no fluorescence (Heim et al., 1995)

[a] Abbreviations: λ_{exc}, excitation maximum in nanometers; ε, extinction coefficient in $M^{-1}\,cm^{-1}$ at λ_{exc}; λ_{em}, emission maximum in nanometers; Q, fluorescence quantum yield. Discrepancies of 2–3 nm between different laboratories in estimating λ_{exc} and λ_{em} are probably not significant.

shown in Fig. 5.1. S65T also proved to have at least two other advantages over wild type: (1) a fourfold increase in the rate of the final oxygen-dependent step in the fluorophore formation and (2) a considerable increase in photostability (Heim et al., 1995). The increased photostability of S65T is manifest in two ways. Wild-type GFP, when irradiated at wavelengths short enough to excite the neutral species, tends to photoisomerize toward the anionic species (Cubitt et al., 1995; Chattoraj et al., 1996). Therefore, excitation of the 395-nm peak rapidly diminishes its amplitude and boosts that of the 470-nm peak (Chalfie et al., 1994). This conversion involves transfer of a proton from the neutral chromophore to Glu222 (Brejc et al., 1997; van Thor et al., 2002) and can partially reverse upon standing in the dark over the time scale of hours to days (Chattoraj et al., 1996). Since the fluorophore of S65T is already entirely anionic, there is no way for photoisomerization to push the ionization any further. In addition to the wild-type photoisomerization, both wild-type and mutant GFPs irreversibly photobleach. The reciprocal interaction between Glu222 and the fluorophore predicts that if Glu222 were replaced by a nonionizable group, the fluorophore would be permanently anionic. Indeed, Ehrig et al. (1995) had found E222G by random mutagenesis and visual screening and had shown that its only excitation peak is at 481 nm, consistent with full ionization of the fluorophore.

Delagrave et al. (1995) combinatorially mutagenized positions 64–69 and screened colonies by imaging spectroscopy. They found six mutants, which they described as "red-shifted" because their excitation spectra peaked near 490 nm, while the position of the emission maxima were largely unchanged. These mutations are now interpretable as

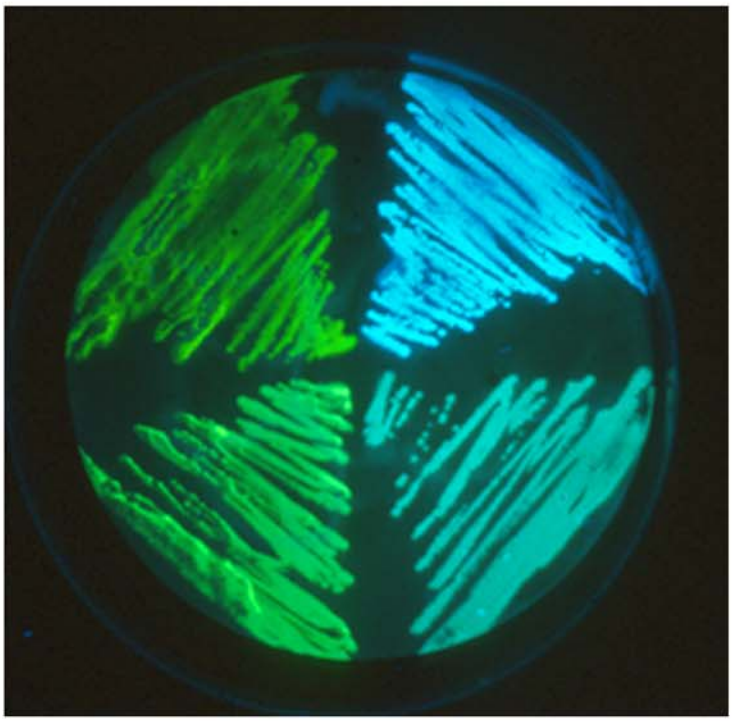

Figure 5.3. Visual appearance of *E. coli* expressing four differently colored mutants of GFP. **Clockwise from upper right**: Blue mutant P4-3 (= Y66H, Y145F) (Heim and Tsien, 1996); cyan mutant W7 (Y66W, N146I, M153T, V163A, N212K) (Heim and Tsien, 1996); green mutant S65T (Heim et al., 1995); yellow mutant 10C (= S65G, V68L, S72A, T203Y) (Ormo et al., 1996). In each of these lists of mutants, the mutation most responsible for the special alterations is underlined, while the other substitutions improve folding or brightness. The bacteria were streaked onto nitrocellulose, illuminated with a Spectraline B-100 mercury lamp (Spectronics Corp., Westbury, NY) emitting mainly at 365 nm, and photographed with Ektachrome 400 slide film through a low-fluorescence 400 nm and a 455-nm colored glass long-pass filter in series. The relative brightness of the bacteria in this image is not a good guide to the true brightness of the GFP mutants. Expression levels are not normalized, and the 365 nm excites the blue and cyan mutants much more efficiently than the green and yellow mutants, but the blue emission is significantly filtered by the 455-nm filter required to block violet haze. See color insert.

emphasizing the anionic fluorophore at the expense of the neutral form. Many biologists have misunderstood the term "red-shifted" to imply that the fluorescence is actually red, but in fact the emission spectra are not significantly changed from the normal green emission of wild-type protein, peaking at 505–510 nm. Five out of six of these mutants had Gly, Ala, Cys, or Leu at position 65, which are undoubtedly the key substitutions responsible for the shift in excitation spectra.

Photoisomerization of GFP was discovered just after the cDNA for GFP was cloned (Chalfie et al., 1994; Cubitt et al., 1995). Exposure to light at ~400 nm causes photoisomerization in wild-type GFP that involves a shift in the chromophore population from the protonated form to the unprotonated. van Thor et al. (2002) have provided evidence for a two-step mechanism to explain this conversion. The first step is decarboxylation of Glu222 followed by structural rearrangements of Thr203 and His148. The most dramatic example

of photoconversion of *Aequorea* GFP being useful in cell biology was described by Patterson and Lippincott-Schwartz (2002). The mutation T203H (on a wild-type GFP background) reduces the absorbance/excitation peak at 490 nm to an even greater extent that in the T203I mutant. Illumination of this mutant called photoactivatable GFP (PA-GFP) with laser light at 413 nm caused a dramatic shift in the excitation spectrum from the peak at 400 nm to the peak at 490 nm. The optical contrast between the photoactivated and nonphotoactivated PA-GFP was up to 100-fold when excited with laser light at 488 nm (60-fold in cells) and was stable at 37°C for at least 1 week. This mutant is potentially very valuable for tracking the location and fate of a certain subpopulation of activated proteins (Patterson and Lippincott-Schwartz, 2002). An even more dramatic example of photoactivation occurs in the coral protein Kaede (Ando et al., 2002), in which UV illumination changes the emission from green to red due to a photochemically driven beta-elimination between the GFP-like chromophore and the imidazole within an immediately adjacent histidine (Mizuno et al., 2001).

Circular permutants (cp) of EYFP and EFYP with calmodulin inserted at position 145 (Camgaroo) have two absorbance peaks (396–404 and 490–496 nm) corresponding to the two major excitation peaks of wild-type GFP (Baird et al., 1999). In both cases, and unlike GFPs, excitation of the most blue peak (~400 nm) does not result in fluorescent emission. Insertion of calmodulin at position 145 of YFP modulates, in a Ca^{2+}-dependent manner, the relative magnitudes of these two absorption maxima; Ca^{2+} increases the protonated (~400 nm) peak while dramatically diminishing the nonprotonated (~496 nm) peak, while increasing Ca^{2+} at the same pH reverses the relative absorption profile and increases the fluorescence seven- to eightfold when excited at ~496 nm. The conformational change in calmodulin presumably torques the staves of the beta barrel in such a way that in the Ca^{2+}-bound state, the chromophore is more susceptible to quenching by acid (Fig. 5.2). These will be discussed in further detail later.

The only other mutant with an increased ratio of anionic to neutral excitation peaks that cannot be explained by the above substitutions of S65 and E222 is "RSGFP1" of Delagrave et al. (1995), which is F64G, V68L, Q69L. Mutations of F64 and V68 are known to affect folding efficiency (see below) but not wavelengths, so Q69L is most likely responsible for the increased ionization of the fluorophore. The crystal structure (Ormo et al., 1996) shows that Q69 anchors a cluster of water molecules that also participate in solvating the Glu222 carboxylate. Replacement by Leu would disrupt this hydrogen-bonding network and destabilize the carboxylate anion more than the neutral carboxyl of Glu222, thereby indirectly promoting fluorophore ionization.

5.7 MUTANTS THAT MORE SUBTLY MODIFY THE ENERGY LEVELS

A few mutants shift both the excitation and emission wavelength distributions. Obviously the major effort has been directed toward obtaining mutants with longer rather than shorter wavelengths. It is this class of mutants that can most accurately be termed "red-shifted." One of the best understood mutants is T203Y, which was designed rationally when the crystal structure of GFP (Ormo et al., 1996) revealed the close proximity of Thr203 to the fluorophore. Replacement of that aliphatic residue by aromatic residues next to the fluorophore was intended to increase the local polarizability of the chromophore. Starting from S65 mutants that were already fully ionized, T203Y increased the excitation and emission maxima by 24 and 16 nm, respectively, to 513 and 527 nm, respectively (Ormo et al., 1996). These wavelengths, which are the longest so far published for an *Aequorea* GFP

mutant, permit detection through at least some standard filter sets for rhodamines—for example 510- to 560-nm excitation, 565-nm dichroic, 572- to 647-nm emission. Although pure 527 nm is itself still green, the long tail of emission at yet longer wavelengths makes the emission yellowish to the eye and distinguishable from S65T in side-by-side comparisons. Figure 5.3 shows such a comparison captured on color film. The hue difference is somewhat more impressive in real life than in this picture. T203F and T203H were almost as effective (Table 5.3), whereas T203W produced less of a shift, perhaps because its steric bulk was excessive or because dipole moments of the new substituents are playing contributory roles. Another mutant of even smaller and less easily explained effect is M153A, which increases the excitation and emission wavelengths of S65T by 15 and 3 nm, respectively (Heim and Tsien, 1996).

5.8 MUTANTS THAT IMPROVE FOLDING AT 37°C

A large class of mutations improves the percentage yield of GFP molecules that fold correctly and become fluorescent without seeming to affect the spectral properties of those properly matured proteins. The yields of fluorescent protein expressed from the original jellyfish gene fall steeply as the temperature increases above about 15–20°C, which is not surprising given the low temperature of Puget Sound, where *Avictoria victoria* lives. Also, high-level expression in *E. coli* tends to give extensive deposition of nonfluorescent protein in inclusion bodies, in which the chromophore has not even formed (Siemering et al., 1996). Both of these problems can be greatly ameliorated by suitable amino acid substitutions, which have been found mostly by random mutagenesis and visual or flow-cytometric screening for brighter bacteria at temperatures up to 37°C. Many have been arrived at independently by several groups (Cormack et al., 1996; Crameri et al., 1996; Heim and Tsien, 1996; Kahana and Silver, 1996; Siemering et al., 1996; Kimata et al., 1997; Nagai et al., 2001), which may imply that existing mutational strategies are approaching saturation.

In our opinion, all new GFP constructs should incorporate some subset of mutations that improve maturation because they do no harm and can often produce large increases in brightness. The only exception would be experiments in which one desires to shut off the formation of newly fluorescent GFP molecules by raising the temperature (Kaether and Gerdes, 1995; Lim et al., 1995). Note that even wild-type GFP is fairly heat-stable once properly folded and matured (Lim et al., 1995), becoming denatured only above 65°C (Ward and Bokman, 1982); only during the folding process is it highly temperature-sensitive (Siemering et al., 1996).

Although folding mutations are quite valuable, one should not expect indefinite further improvements in GFP brightness. The ultimate brightness of any fluorophore is limited at the molecular level by the product of extinction coefficient and fluorescence quantum efficiency. Folding mutations do not significantly improve the brightness of GFP molecules that are correctly folded (Siemering et al., 1996), and unfortunately few laboratories characterizing GFP mutants have reported extinction coefficients or quantum yields along with their studies. The fact that the brightness of bacteria expressing very high levels of GFP can be raised by a factor of 20 to 50 (Cormack et al., 1996; Crameri et al., 1996; Siemering et al., 1996) is more indicative of the wretched folding efficiency of the wild-type protein at high temperature and concentrations; at lower temperatures or expression levels—that is, when one wants to be able to detect as few GFP molecules as possible—the improvement due to folding mutations is not as impressive. Furthermore, the improve-

ment factors due to individual mutations do not multiply: Two mutations that separately improve folding efficiency by factors of x and y generally give much less than an improvement of xy when combined (Zapata-Hommer and Griesbeck, 2003). Folding efficiency can asymptotically approach but can never exceed 100%.

The existing crystal structures do not clearly explain how most of the folding mutations exert their favorable effects. Many of the mutations may only make a difference while the protein is still relatively disordered and in an unfavorable environment, whereas the crystals are typically grown from mature protein produced under conditions chosen to maximize successful expression. In a few cases, such as F99S and M153T, for which the side chains face outward, the mutations reduce patches of surface hydrophobicity, possibly inhibiting aggregation and enhancing brightness. However, these mutations did not alter the overall speed of fluorescence development at 37°C compared to wild type (Crameri et al., 1996). Mutations V163A and S175G together actually slow the final aerobic development of fluorescence (Siemering et al., 1996), even though they greatly improve the yield of properly matured protein. The folding and unfolding reactions of Cycle3 (F99S, M153T, V163A) and wild-type GFP were compared using fluorescence and circular dichroism spectroscopy and measurement of hydrogen exchange reactions using Fourier transform infrared spectroscopy (Fukuda et al., 2000). The results illustrate that while folding and unfolding were relatively slow processes, they were basically the same for both proteins and that the decreased exterior surface hydrophobicity of Cycle3 contributed more to the general increase in brightness of Cycle3 than did the rates of folding. Similarly, Miyawaki and colleagues (Nagai et al., 2002) identified a novel mutation (F46L) that decreased the time required to form mature, fluorescent YFP. When they looked closely at the rates of the folding and the oxidation of the chromophore, the mutation F46L slowed the rate of folding slightly, compared to YFP without F46L, but increased the rate of oxidation of the chromophore so that the resulting maturation time was reduced (Nagai et al., 2002). Griesbeck et al. (2001) reported a series of mutations Q69M/C70V/V163A/S175G that improved the folding of GFP400 or Sapphire. The fact that Q69M has some beneficial effects on YFP had been reported (Griesbeck et al., 2001) previously, but C70V, which was the fortunate result of a PCR error, turned out to be crucial because Sapphire with Q69M/V163A/S175G was nonfluorescent. F46L, F99S, and M153T were found not to improve the overall efficiency of folding or brightness of Sapphire. An understanding at the molecular level of how other folding mutations work, especially those affecting buried side chains, may require detailed comparison of the resulting final crystal structures with the baseline structures already on hand, as well as continued investigations of the folding intermediates and dynamics.

5.9 MUTATIONS THAT MODULATE AGGREGATION

Currently, *Aequorea* GFP is the only fluorescent protein known that is not an obligate homo-oligomer in its natural state; the GFP from *Renilla* is an obligate dimer (Ward and Cormier, 1979; Ward, 1998) and RFP (dsRED from coral) is an obligate tetramer (Baird et al., 2000) (as are all other characterized fluorescent proteins form coral species see other chapters in this volume). Quite some time before a rigorous determination of the homoaffinity was made, it was known that even *Aequorea* GFP dimerized to some degree in solution (Yang et al., 1996b). GFP crystallizes as either a dimer (Yang et al., 1996a; Palm et al., 1997; Battistutta et al., 2000) or a monomer (Ormo et al., 1996; Brejc et al., 1997; Wachter et al., 1997; Wachter et al., 1998; Wachter et al., 2000). In the dimeric

crystal structure, the unit cell consists of two monomers associated in a slightly-twisted, head-to-tail fashion via many hydrophilic contacts as well as several hydrophobic contacts. The very different solvent conditions used by each group are sufficient to explain the differing results (Ward, 1998). Residing within a large hydrophobic patch, residues A206, L221 and F223 (Fig. 5.4) are sufficient to cause formation of the dimer at relatively low concentrations in solution and in living cells. However, changing these residues singly or in combination to positively charged residues such as A206K, L221K and F223R (Fig. 5.4; Table 5.5), effectively eliminated the interaction of the monomers (Zacharias et al., 2001); the resulting monomeric GFPs have been termed mGFPs. To determine the strength

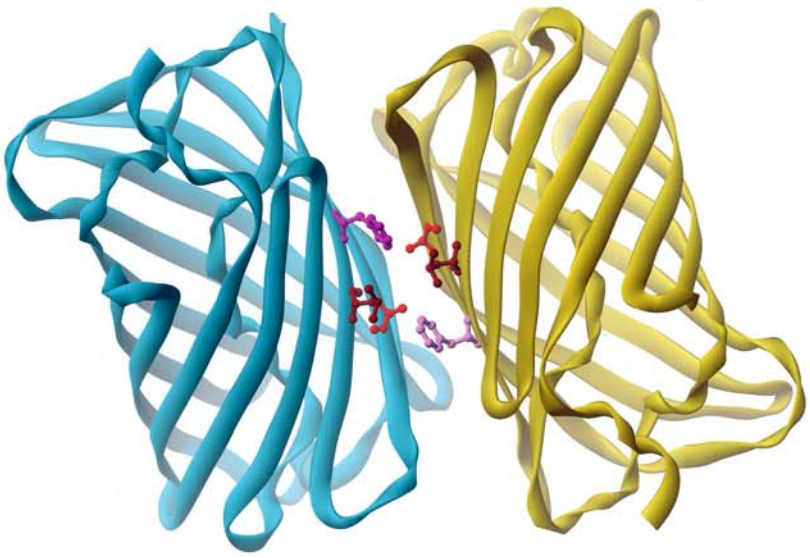

Figure 5.4. The crystal structure of dimeric GFP (1GFL) (Yang et al., 1996a). The residues A206 (red), L221 (orange), and F223 (lavender) are shown as ball-and-stick representations. Replacing any of these residues with the positively charged residues lysine or arginine effectively monomerizes the protein. See color insert.

TABLE 5.5. Mutations that Eliminate Dimerization of GFPs

Mutation	Quantum Yield (Ward, 1998)	Extinction Coefficient[1]	K_d[2]
wild type	0.67	67	0.11
L221K	0.67	64	9.7
F223R	0.53	65	4.8
L221K, F223R	0.68	59	2.4
A206K	0.62	79	74[3]

Fluorescence and dissociation of wild-type YFP versus monomeric YFP. [1] Extinction coefficient × 1000 in $M^{-1} cm^{-1}$ at # λ_{exc}. [2] The dissociation constant K_d (mM) as measured by sedimentation equilibrium analytical ultracentrifugation (Zacharias et al., 2002). [3] Due to the extreme monomeric nature of this protein it was difficult to determine an accurate dissociation constant for a hypothetical dimer.

of the interaction in solution, Zacharias et al. (2001) used sedimentation equilibrium analytical ultracentrifugation to characterize the affinity of GFPs with the wild-type interface as well as the mGFPs. In significant contrast to X-ray crystallography the experimental conditions used in the analytical ultracentrifugation experiments approximates cellular physiological conditions and were able to provide definitive information (McRorie and Voelker, 1993) about the affinity of the complex. The dimer dissociation constants for wtGFP and several of the mutants are compiled in Table 5.5. Other mutations are also thought to affect the state of GFP aggregation. F99S and M153T first described in relation to aggregation by Stemmer and colleagues (Crameri et al., 1996) reduce obvious patches of surface hydrophobicity and could inhibit aggregation, but no dissociation constant has yet been determined for Cycle3. Indeed, the triple mutant (F99S, M153T, V163A) (Crameri et al., 1996) has a diffusion coefficient inside mammalian cells one order of magnitude higher than that of wild-type GFP, implying a corresponding reduction in binding to other macromolecules (Yokoe and Meyer, 1996). Because V163 points into the interior of the protein (Ormo et al., 1996) and because F99S and M153 face outwards, the latter two are most likely the culprits in wild-type GFP. However, the triple mutation did not alter the overall speed of fluorescence development at 37°C compared to wild-type GFP (Crameri et al., 1996). Mutations V163A and S175G together actually slow the final aerobic development of fluorescence (Siemering et al., 1996), even though they greatly improve the yield of properly matured protein.

The crystal structure of cyclized Cycle3 GFP was determined (Hofmann et al., 2002). The authors found a crystallographic dimer interface different than those previously reported and concluded that various polar and nonpolar patches on the surface of GFP could serve to dimerize the proteins in ways previously undescribed. However, it is unlikely that the dimer interface described in this work exists under physiological conditions with great affinity in non-Cycle3 versions of GFP because it is clear from sedimentation equilibrium analytical ultracentrifugation experiments on GFP containing A206K (and other mutations) (Zacharias et al., 2001) that there is virtually no remaining affinity when the more commonly observed dimer interface is altered. Further biophysical characterization of Cycle3 and mutants derived from it is clearly warranted.

The issue of GFP oligomerization is significant for several reasons. Most of the potential for trouble arises when GFP or its spectral mutants are fused to other proteins to track protein localization or expression or to measure interactions by fluorescence resonance energy transfer (FRET) (most often CFP and YFP). If GFP dimerizes in the context of being part of a fusion protein, it could also foreseeably dimerize the protein to which it is fused. The situation could become even "stickier" if the host protein is itself an oligomer. When measuring the interactions of molecules by FRET, the fluorophores used to report the interactions must not themselves in any way influence—or worse yet, create—the interactions being measured. Obviously, if the fluorophores have affinity for each other, then doubt is cast on the accuracy of any measurement made to the presumed interaction of the host proteins. The problems associated with GFP dimerization are most troublesome when measuring intermolecular FRET in a two-dimensional space such as a membrane (Fung and Stryer, 1978; Wolber and Hudson, 1979; Dewey and Hammes, 1980; Snyder and Freire, 1982; Dewey and Datta, 1989; Yguerabide, 1994; Zimet et al., 1995). In this situation, we found that wtGFPs were very likely to dimerize even when expressed at very low surface densities (Zacharias et al., 2001). Since the monomerizing mutations alter nothing but the homoaffinity of GFP, we recommend including them (preferably A206K) in all GFP expression constructs where dimerization is not desirable.

Another interesting phenomenon associated with dimerization of type I (wild-type) GFPs is that when they are dimerized, there is a suppression of the ~500-nm peak in the absorbtion/excitation spectrum and a concomitant but relatively smaller increase in the ~400-nm peak (Morise et al., 1974). In GFPs, a dimer-induced shift of the excitation spectrum ~100 nm toward the blue (neutral form of the chromophore) will result in off-peak excitation (unless one changes to appropriate filters) and much lower fluorescence output. In YPF, excitation near the 400-nm peak yields a nonfluorescent species. Additionally, for FRET-related experiments a shift in the absorbance of YFP (acceptor) toward the ~400-nm peak will result in decreased spectral overlap with the emission spectrum of a CFP (donor) and therefore reduce the efficiency of FRET. This particular phenomenon was exploited to make non-FRET-based sensors for molecular proximity (De Angelis et al., 1998).

It has been a point of curiosity as to why GFP and FPs in general should oligomerize. It has been proposed, in the case of nonbioluminescent corals (and likely the anemones; also anthozoans), that FPs are acting as a protective barrier between the harmful UVA irradiation from the sun and the resident symbiotic organisms that produce energy by photosynthesis (Salih et al.,2000). Corals living closer to the surface of the water are exposed to greater amounts of UVA irradiation from the sun, and it is in these Salih and co-workers found that these corals had the greatest abundance and diversity of FPs. They went on to show that the FP complexes dissipate excess energy at wavelengths of low photosynthetic activity and also serve to reflect visible and infrared light. The complexes are able to do this is by capturing light at the bluer, more damaging end of the spectrum and to shuttle it by FRET to red-shifted (FRET acceptors), oligomeric, partner proteins, finally "spilling" the energy in a range that is not compatible with photosynthesis. In this view, oligomerization between FPs with different chromophores (Gross et al., 2000; Salih et al., 2000; Cotlet et al., 2001; Dove, 2001; Garcia-Parajo et al., 2001) serves to maximize FRET (Zacharias, 2002). However, this type of explanation is less appropriate for bioluminescent sea pansies and jellyfish. In their native species, these FPs serve as acceptors for luciferases, and there are no known spectral differences among FPs expressed in the same organism, so oligomerization does not promote intersubunit FRET. A speculative alternative comes from the observations that resistance to photobleaching of DsRed mutants decreased as the stoichiometry was reduced from tetrameric to dimeric to monomeric (Campbell et al., 2002), and that the most monomeric FP (*Aequorea*) comes from 49°N latitude, obligate dimeric FPs (*Renilla*) from 31–33°N latitude, and obligate tetramers from corals in tropical waters. Perhaps oligomerization protects the FPs themselves from photobleaching by decreasing surface-to-volume ratio, hindering access of oxygen to the chromophores. The higher the intensity of solar irradiation, the more selection pressure for photostability. Water temperature also is inversely correlated with latitude, and thermostability and photostability requirements run in parallel. Further characterization of more related FPs from geographically disparate sources should help to clarify the reasons for these interesting distinctions (Zacharias, 2002).

5.10 TRUNCATIONS OF GFP

A frequently asked question about GFP is whether it can be significantly reduced in size. In the two most systematic investigations of this possibility, Dopf and Horiagon (1996) and Kim and Kaang (1998) produced a family of genetic truncations from the N- and C-

termini. The N-terminal methionine could be replaced by a polyhistidine tag (Dopf and Horiagon, 1996), but deletion of residues 2–8 prevented fluorescence or chromophore development. The C-terminus was slightly more tolerant, in that 6 (Dopf and Horiagon, 1996–9) (Kim and Kaang, 1998) but not 13 (Dopf and Horiagon, 1996; Kim and Kaang, 1998) residues could be eliminated. These narrow limits are in good agreement with the crystal structure (Ormo et al., 1996), in which Met1 and residues 230–238 are too disordered to be located accurately. There are no major regions of internal sequence that appear dispensable. The core domain of GFP, residues 2–229, seems to be a monolithic entity required to shield the fluorophore from solvent. A few local loops might conceivably be shortened somewhat, but the slight net reduction in size would hardly be worth the effort. Fortunately, it seems possible to target full-length GFP to essentially any compartment in the cell, so its size has not yet severely restricted its versatility. A report (Wiedenmann et al., 2000) of a coral chromoprotein with only 148 amino acids proved to be an artifact of incorrect DNA sequencing combined with cleavage of the peptide backbone during chromophore formation (Martynov et al., 2001; Wiedenmann et al., 2002).

5.11 GFP BIOSENSORS

The crystal structure of GFP gives us a picture of a protein that appears to be rather monolithic, seemingly without much tolerance for gross rearrangements, insertions, or deletions. However, during some semirandom mutagenesis experiments, Baird et al. (1999) made the serendipitous discovery that a six-residue peptide (FKTRHN) had been inserted into CFP at position 145 without destroying its fluorescence (e.g., Fig. 5.5B). This suggested that the two halves of CFP had the capability to fold autonomously and that circular permutations generated by fusing the original N- and C-termini and breaking the main chain backbone at other locations might be tolerated (Fig. 5.5C). A systematic procedure (Graf and Schachman, 1996) to find other potential break points uncovered 10 sites within CFP that could serve as new N- and C-termini while retaining fluorescent proteins (Fig. 5.6). Another report (Topell et al., 1999) that followed closely on the heels of the report by Baird and colleagues found 20 possible breakpoints, but among them only five retained significant fluorescence. When new termini are generated at position 145, a minimum of 3 residues must form the linker between the wild-type termini to reestablish fluorescence, and linkers of four to seven residues resulted in restoration of fluorescence to the same level as EGFP (Akemann et al., 2001). Aside from their sheer novelty, such circular permutations and insertions have proved to be useful in generating single-fluorophore, GFP-based sensors for Ca^{2+} (Nagai et al., 2001) and for changing the orientation of a fluorophores to try to improve the a FRET signal (Baird et al., 1999). GFPs have also been completely circularized by split intein technology (Hofmann et al., 2002).

Some of the same breakpoints identified in circularly permuted GFP can also be locations to serve as acceptor sites into which other proteins can be fused (Fig. 5.5B). The first such insertion of a whole, functional protein was *Xenopus* calmodulin into position 145 of YFP (Baird et al., 1999), resulting in a calcium sensor named Camgaroo which has proven useful in measuring calcium inside of mitochondria and the mushroom bodies in the brain of *Drosophila* (Yu et al., 2003). Other groups combined insertions with circular permutations of EGFP (Nakai et al., 2001) or EYFP (V68L/Q69K) (Nagai et al., 2001) and created a multimember suite of calcium sensors each with a different fluorometric readout for calcium. One sensor for Zn^{2+}, with an apparent K_d of 400 μM made by inserting the zinc finger motif from zif268 into YFP at position 145, gave a maximum response

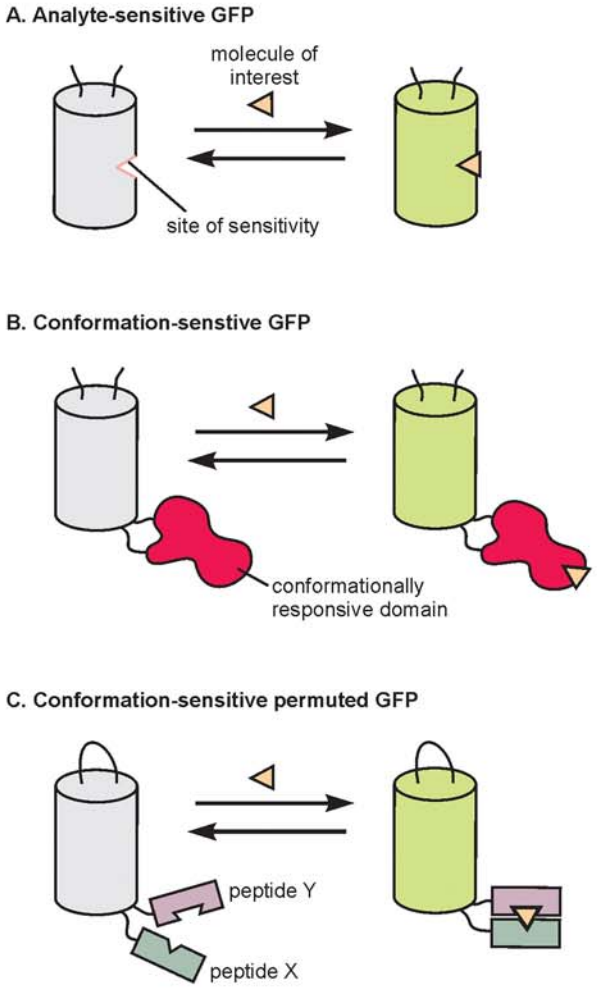

Figure 5.5. GFP biosensors. **(A)** GFP can be engineered to be directly sensitive to a small molecule of interest. **(B)** Insertion of a conformationally dynamic domain into GFP can result in a chimera in which the fluorescence properties of GFP are modulated by a change in conformation of the domain. **(C)** Similarly, proteins or peptides with dyanmic, associative properties can be fused to the N and C termini of circularly permuted GFPs, thereby reporting on the changes in the association in response to a stimulus. See color insert.

of about 1.7-fold change in fluorescence intensity (Baird et al., 1999); another group engineered Zn^{2+}-binding sites directly into the chromophore of BFP (Richmond et al., 2000). The resulting protein had about a twofold change in fluorescence intensity with a K_d for Zn^{2+} of 50 μM. The slow on-rate for Zn^{2+} for the second probe ($T_{1/2}$ of >4 hours) limits its usefulness in cell biological experiments. Similarly, attempts have been made (Richmond

Figure 5.6. A backbone representation of the three-dimensional structure of GFP (1EMG) (Elsliger et al., 1998). The residues where circular permutations are permitted while retainin fluorescence are color highlighted. E142, hot pink; Y143, gray; Y145, dark blue; H148, fuchsia; D155, yellow; H169, red; E172, light blue; D173, orange; A227, Cyan; I229, light purple. These residues represent sites where the main chain can be interrupted. In most cases, resumption of GFP sequence can occur one to four residues following the initial interruption. See color insert.

et al., 2000) to rationally design sites, on the external surface of the barrel (Fig. 5.5A) of YFP (10C) that could bind to metals. Mutants on the 10C background (S147H/Q204H and S202D or F223E) were generated that successfully quenched YFP fluorescence when exposed to various divalent cations, but nothing was reported of the selectivity concerning the ability of these mutants to discriminate among the cations. It would be ideal if robust, selective sensors of such simplicity could be generated for cellular analytes.

Similar to the case where Zn^{2+} sites were engineered directly into the chromophore, individual GFPs with only small modifications can be made into sensors for pH and halides (Fig. 5.5A).

5.12 pH

High pH, 11–12, causes a relative redistribution of the two absorbance/excitation maxima of wtGPF toward the longer wavelength (470 nm) peak while low pH ($pK_a \sim 5.5$) causes a quenching of fluorescence (Bokman and Ward, 1981; Ward et al., 1982). A fluorophore with sensitivity to variations in pH is often viewed as a lemon, but in the case of GFPs, the lemon has been made into some very informative lemonade. While it is good to have a fluorescent probe that emits stably over a broad range of pHs, it was also apparent that the pH-dependent change in fluorescence behavior could be exploited to measure pH in

TABLE 5.6. pk_a Values of Gfps Used as ph Indicators

FP	pK_a	Mutations	Reference
GFP	4.8	T203I	(Kneen et al., 1998)
YFP	5.7	Q69M	(Griesbeck et al., 2001)
GFP	5.9	S65	(Kneen et al., 1998)
GFP	6.0	F64L, S65T	(Kneen et al., 1998)
GFP	6.1	Y66H	(Kneen et al., 1998)
GFP	6.15	F64L, S65T, H231L	(Llopis et al., 1998)
CFP	6.4	K26R, F64L, S65T, Y66W, N146I, M153T, V163A, N164H, H231L	(Llopis et al., 1998)
YFP	7.1	S65G, S72A, T203Y, H231L	(Llopis et al., 1998)
GFP	7.3	S65T/C48S/H148C/T203C	(Hanson et al., 2002)
GFP	8.0	S65T, H148G, T203C	(Hanson et al., 2002)
GFP	8.0	H148G	(Maysuyama et al., 2000)

FP indicates the spectral mutant upon which the indicated mutations were incorporated. pKa is the pH value at which 50% of the molecules are fluorescent.

subcellular domains (Llopis et al., 1998). pK_a values range for GFPs vary widely covering the range of most cellular pHs (Table 5.6). pHluorins comprise a set of pH-sensitive GFPs used to monitor exocytosis (Miesenbock et al., 1998; Sankaranarayanan et al., 2000). These sensors are targeted to the luminal side of secretory vesicles where the pH is below the pK_a of pHluorin (which causes them to be nonfluorescent). Upon fusion, exocytosis, and exposure of the membrane-associated pHluorin to the extracellular milieu held at a desired physiological pH, the pHluorin becomes fluorescent (ecliptic pHluorin) or shifts its excitation maximum from 395 nm to 475 nm (ratiometric pHluorin). The broad range of pK_a values for the fluorophores is generated by the diversity mutations in and around the fluorophore. Many of the physical reasons for the various pH-sensitive behaviors are summarized in a theoretical study (Scharnagl et al., 1999) that incorporated a broad range of existing physical data from many experimental sources. In YFP (class 4 chromophores) the mutation Q69M (named "Citrine") retains virtually identical excitation and emission spectra but lowers the pK_a of the chromophore to 5.7, renders it insensitive to chloride, increases the photostability over previous versions of YFP by about twofold, and improves expression at 37°C in cells (Griesbeck et al., 2001). In the crystal structure of Citrine, the Met at position 69 is well-ordered, tightly packed into the cavity, and unlikely to be able to undergo the same sort of conformational change that is seen with the apo- and iodide-bound forms of EYFP.

5.13 HALIDES

Wachter and Remington (1999) first reported halide sensitivity of a YFP (S65G/V68L/S72A/T302Y and H148Q or H148G). Shortly thereafter, they and Verkman's group made more detailed biophysical characterization of the nature of the Cl^- sensitivity of YFP (T203Y/S65G/V68L/S72A/H148Q) and made the first steps toward developing it as the first genetically encoded Cl^- sensor (Jayaraman et al., 2000). Almost simultaneously, Kuner and Augustine (2000) discovered the Cl^- sensitivity of YFP while using CFP (K26R/F64L/S65T/Y66W/N146I/M153T/V163A/N164H/H231L) and YFP (S65G/S72A/K79R/T203Y/H231L) to study protein–protein interactions by FRET. They fused CFP and

YFP with a short intervening peptide linker (ENLYFQG) to make a protein that, in the resting state, had a high degree of FRET, and they called it "Clomeleon." When Cl^- bound to YFP, quenching its fluorophore and rendering it a nonfunctional FRET acceptor, CFP subsequently became dequenched or brighter. This ratiometric sensor for Cl^- was able to measure small changes in Cl^- at physiological concentrations. The crystal structure of YFP H148Q has been solved (Wachter et al., 2000), and the selecitive halide-binding site was found in a small, amphiphilic, buried cavity adjacent to R96. The halide ion was found to be hydrogen-bonded to the phenol group of tyrosine at position 203, illustrating why this mutation is critical to the formation of a halide-sensitive GFP (Wachter et al., 2000).

Conversely, YFPs have been rendered virtually insensitive to the effects of halide binding by the mutation Q69M (Griesbeck et al., 2001) or F64L/M153T/V163A/S175G ("Venus") (Nagai et al., 2002). In the first case, the bulkier side chain of Met fills the cavity into which chloride ion could reside with Q69, and it is conformationally stable. Even if a conformational change in the thioether side chain of Met were permitted in this space, it is unlikely to contribute to halide binding because it is incapable of hydrogen bonding in the same manner as the carboxamide nitrogen of a Gln side chain (Griesbeck et al., 2001). In the second case, F64L induced large conformational changes in the molecule, leading to the removal of halide sensitivity by preventing ion access to the binding site (Rekas et al., 2002). The "Venus" variant of GFP is also very insensitive to changes in pH.

5.14 INSERTION OF GFP INTO OTHER PROTEINS

Insertion of GFPs into other proteins is another important twist on this theme. An early example was the insertion of GFP [GFPΔC; Chalfie et al. (1994)] into a nonconducting mutant of the Shaker K^+ channel (Siegel and Isacoff, 1997) and subsequent, improved generations thereof (Guerrero et al., 2002). This first version of the fusion was able to monitor changes in membrane voltage with a maximal fractional fluorescence change of 5.1%. Similarly, wtGFP was inserted into an intracellular loop of a reversibly nonconducting form of the rat mu I skeletal muscle voltage-gated sodium channel (Ataka and Pieribone, 2002). The resulting protein called SPARC (sodium channel protein-based activity reporting construct) can report depolarizing pulses as short as 2 ms and does not inactivate during prolonged depolarizations, but the size of its optical response is very small.

5.15 TANDEM CONCATENATIONS OF TWO GFPS

The availability of GFP mutants of different colors, UV-excited blue emitters and blue-excited green emitters, enables fluorescence resonance energy transfer (FRET) from one to the other. FRET is strongly dependent on the angular orientation and distance of the fluorophores from one another, falling off steeply as the distance exceeds the Förster distance R0 at which FRET is 50% efficient (Tsien et al., 1993; Lakowicz, 1999). For the blue emitter P4-3, containing the point mutations Y66H and Y145F, donating energy either to S65T or S65C, R0 is calculated to be about 40 Å (Heim and Tsien, 1996), assuming that the mutual orientation of the fluorophores is random or freely tumbling. The larger the R0, the better; this is because GFP is a cylindrical structure of about 12-Å radius and 42-Å length (Ormo et al., 1996), and so much of R0 is used up simply within the two GFPs. A systematic study determined the Förster distances between all homo and hetero

pairings of BFP, CFP, GFP, YFP, and DsRed [Patterson et al., 2000; see also Wu (1994)] The maximum R0 measured for any pair was 56.4 Å between EGFP and EYFP. The Förster distance between CFP and YFP is 49.2 Å. This aspect, combined with greater distance between the peaks of excitation and a favorable overlap integral (J), makes CFP and YFP the pair used most commonly in FRET studies.

Sensors that track the activity of proteases were one of the earliest applications for which FRET between concatenated GFPs was exploited (Xu et al., 1998 Heim, 1999; Harpur et al., 2001; Luo et al., 2001; Tawa et al., 2001) and reviewed (Jones et al., 2000). Sensors for caspase 3 have been used successfully as reporters in high-throughput drug discovery programs (Tawa et al., 2001). The first examples of these constructs included the concatenation of the genes encoding S65C or S65T (GFPs) and P4-3 (BFP) with an intervening 25-residue linker connecting the two GFP-derived domains (Heim and Tsien, 1996). Likewise, BFP5 and RSGFP4 have been fused with a 20-residue linker sensitive to factor Xa (Mitra et al., 1996). In either case, before protease cleavage, UV excitation gives rise to some blue emission but also substantial green emission due to FRET from the blue- to the green-emitting domain. After protease cleavage to separate the two domains, FRET is abolished, the blue emission is increased, and the green emission is nearly abolished. For the S65C:P4-3 construct, the ratio between blue and green emission intensities increased by a factor of 4.6 upon cleavage (Heim and Tsien, 1996), while the RSGFP4::BFP5 fusion showed about a 1.9-fold increase in emission ratio (Mitra et al., 1996). In the case of the S65C/P4-3 construct, the large change in ratio between blue and green emissions was shown to result from separation of the two fluorophores rather than from an effect on either one separately, because control experiments with the two separate proteins showed no spectral sensitivity to protease under matching conditions.

Another concatenated, intramolecular FRET-based reporter that has been used broadly is the calcium sensor CaMeleon (Miyawaki et al., 1997). In this sensor, CFP and YFP flank calmodulin and a Ca^{2+}-calmodulin-binding peptide, M13 from myosin light-chain kinase. Wheng CaMeleon encounters a change in Ca^{2+}, M13 and calmodulin respond either by associating in increased [Ca^{2+}] or dissociating in decreased [Ca^{2+}]. The conformational change that occurs causes a change in the FRET efficiency (which is low at lower [Ca^{2+}] and high at higher [Ca^{2+}]) largely via a change in kappa squared or the intermolecular angle of fluorophore orientation (Atsushi Miyawaki, personal communication). This probe has been especially useful in reporting calcium changes in subcellular domains like the nucleus and endoplasmic reticulum (Arnaudeau et al., 2002; Demaurex and Frieden, 2003; Malli et al., 2003; Palme et al., 2004) and caveolae (Isshiki et al., 2002), as well as in *Drosophila* (Reiff et al., 2002; Liu et al., 2003) and *C. elegans* (Kerr et al., 2000). The most recent versions of CaMeleon incorporate YFPs (Citrine) that are insensitive to pH and Halides (Griesbeck et al., 2001). Most recently, conformationally sensitive concatenations of GFP have been used to track the activity of kinases (Ting et al., 2001; Zhang et al., 2001b; Violin et al., 2003), elegantly showing links between activity of these kinases and other signal transduction pathways.

5.16 SILENT AND LOSS-OF-FUNCTION MUTATIONS

During any random or semirandom mutagenesis screen, the great majority of colonies typically are either indistinguishable from the starting phenotype or of significantly reduced brightness. In principle, these mutants could be sequenced to provide a list of neutral or deleterious substitutions, but such a list would be laborious to collect and of negligible

interest to those wishing to improve GFP or obtain novel properties. Perhaps this list deserves compilation, because it might increase understanding of how GFP folds and builds its chromophore. A few mutations are known to be neutral, such as the ubiquitous Q80R, which may have arisen from a PCR error in the initially distributed cDNA clone (Chalfie et al., 1994). The substitutions within the natural isoforms are presumably all permissive for fluorescence, although other properties may well be altered. As mentioned earlier, mutations of S65 to bulky or highly polar residues, mutations of Y66 to any nonaromatic amino acid, and mutations of G67 to anything else are probably not tolerated.

5.17 NUCLEIC ACID CHANGES THAT DO NOT CHANGE THE PREDICTED AMINO ACID SEQUENCE—THAT IS, OPTIMIZATION OF CODON USAGE AND ELIMINATION OF CRYPTIC SPLICE SITES

GFP expression levels can often be increased by redesigning the nucleic acid sequence in ways that should have no significant effect on the final protein sequence. For example, the codon usage in the jellyfish gene is not optimal for mammalian cells, so the gene has been resynthesized with mammalian-preferred codons (Crameri et al., 1996; Levy et al., 1996; Zolotukhin et al., 1996). Translation in eukaryotes can be optimized by inclusion of an optimal translation-initiation sequence (Kozak, 1989). This redesign sometimes involves inserting a new codon that begins with G immediately after the start (AUG) codon. This introduces an extra amino acid such as Ala or Val, which in some articles adds one to the numbering of all amino acids from 2 upwards (Crameri et al., 1996), whereas we prefer to call it 1a to preserve wild-type numbering. Fortunately, the N-terminus is tolerant of such additions. For ease of comparison of mutants, this chapter numbers residues according to their position in the original gfp gene. In plant cells, mRNA derived from the original gfp gene undergoes undesired splicing, which can be eliminated by codon changes (Haseloff and Amos, 1995; see also Chapter 12). GFP cDNA coding sequences have also been altered to reflect the codon bias, and thereby increase the level of expression, of a wide variety of organisms such as Chlamydomonas (Franklin et al., 2002), yeast (Gerami-Nejad et al., 2001), paramecium (Hauser et al., 2000), and sugar beets (Zhang et al., 2001a).

5.18 ODDS AND ENDS

Perhaps one of the most interesting and persistent questions concerning the existence of GFPs is that of their functional role in the animal: Why should they glow? What advantage is afforded to the creatures who harbor such a protein? The recent discoveries of GFP-like proteins from nonbioluminescent Anthozoan organisms indicates that the proteins primary function cannot be linked exclusively to bioluminescence. Similarly, discoveries of chromoproteins in these same animals indicate that the proteins function may not even necessarily be tied to fluorescence. Konstantin Lukyanov's group (Gurskaya et al., 2003) has cloned a colorless, nonfluorescent GFP (acGFPL) from *Aequorea coerulescens*. They showed convincingly that the protein was not an artifact of cloning and that fluorescence could be imparted by a reintroducing the invariant G222 which existed naturally in acGFPL as E222. In the living organism, this protein cannot serve as an acceptor for the bioluminescence energy of aequorin, suggesting that this protein may have some completely unique role in the jellyfish. When one considers that the major absorbtion of wild-

type *Aequorea victoria* GFP is at 396 nm with only a minor peak at 475 nm (the primary emission band of aequorin), it seems plausible that the primary role of normally fluorescing *A. victoria* GFP is not necessarily as an acceptor for aequorin. The absorbance maximum of wild-type *A. victoria* GFP can be shifted to the 475-nm peak by a single-point mutation at position 65, a seemingly easy evolutionary maneuver that would optimize the transfer of energy between the two proteins. Perhaps future screens will produce the optimal GFP from *Aqueorea*.

It seems that the structure of GFP has persisted throughout the evolutionary tree and has possibly ended up as an invariant component called Nidogen of basement membranes in organisms up to humans (Hopf et al., 2001). Despite having only a 10% sequence identity with *Aequorea* GFP, all of the general structural components similar to GFP exist with remarkable similarity. The Nidogen residues that are equivalent to the chromophore-forming SYT residues in GFP are IGG. Several other residues within GFP that are known to be critical for forming fluorescence are also represented in Nidogen by residues that would eliminate the opportunity to form a functional fluorophore.

These final two examples suggest that the striking beta-barrel structure of GFP has served several important functions in the many organisms known to express such a structure. We may simply count ourselves fortunate that at some point in the past the evolution chanced upon a version that glows.

REFERENCES

Akemann, W., Raj, C. D., and Knopfel, T. (2001). Functional characterization of permuted enhanced green fluorescent proteins comprising varying linker peptides. *Photochem. Photobiol.* 74: 356–363.

Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and Miyawaki, A. (2002). An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc. Natl. Acad. Sci. USA* 99:12651–12656.

Arnaudeau, S., Frieden, M., Nakamura, K., Castelbou, C., Michalak, M., and Demaurex, N. (2002). Calreticulin differentially modulates calcium uptake and release in the endoplasmic reticulum and mitochondria. *J. Biol. Chem.* 277:46696–46705.

Ataka, K., and Pieribone, V. A. (2002). A genetically targetable fluorescent probe of channel gating with rapid kinetics. *Biophys. J.* 82:509–516.

Baird, G. S., Zacharias, D. A., and Tsien, R. Y. (1999). Circular permutation and receptor insertion within green fluorescent proteins. *Proc. Natl. Acad. Sci. USA* 96:11241–11246.

Baird, G. S., Zacharias, D. A., and Tsien, R. Y. (2000). Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proc. Natl. Acad. Sci. USA* 97: 11984–11989.

Barondeau, D. P., Putnam, C. D., Kassmann, C. J., Tainer, J. A., and Getzoff, E. D. (2003). Mechanism and energetics of green fluorescent protein chromophore synthesis revealed by trapped intermediate structures. *Proc. Natl. Acad. Sci. USA.* 100:12111–12116.

Battistutta, R., Negro, A., and Zanotti, G. (2000). Crystal structure and refolding properties of the mutant F99S/M153T/V163A of the green fluorescent protein. *Proteins* 41:429–437.

Bell, A. F., He, X., Wachter, R. M., and Tonge, P. J. (2000). Probing the ground state structure of the green fluorescent protein chromophore using Raman spectroscopy. *Biochemistry* 39: 4423–4431.

Bokman, S. H., and Ward, W. W. (1981). Renaturation of *Aequorea* green-fluorescent protein. *Biochem. Biophys. Res. Commun.* 101:1372–1380.

Brejc, K., Sixma, T. K., Kitts, P. A., Kain, S. R., Tsien, R. Y., Ormo, M., and Remington, S. J. (1997). Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. *Proc. Natl. Acad. Sci. USA* 94:2306–2311.

Campbell, R. E., Tour, O., Palmer, A. E., Steinbach, P. A., Baird, G. S., Zacharias, D. A., and Tsien, R. Y. (2002). A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. USA* 99:7877–7882.

Campbell, R. E., Tour, O., Palmer, A. E., Steinbach, P. A., Baird, G. S., Zacharias, D. A., and Tsien, R. Y. (2002). A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. USA* 99:7877–7882.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science* 263:802–805.

Chattoraj, M., King, B. A., Bublitz, G. U., and Boxer, S. G. (1996). Ultra-fast excited state dynamics in green fluorescent protein: Multiple states and proton transfer. *Proc. Natl. Acad. Sci. USA* 93:8362–8367.

Cheng, L., Fu, J., Tsukamoto, A., and Hawley, R. G. (1996). Use of green fluorescent protein variants to monitor gene transfer and expression in mammalian cells. *Nat. Biotechnol.* 14:606–609.

Cody, C. W., Prasher, D. C., Westler, W. M., Prendergast, F. G., and Ward, W. W. (1993). Chemical structure of the hexapeptide chromophore of the *Aequorea* green-fluorescent protein. *Biochemistry* 32:1212–1218.

Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* 173:33–38.

Cotlet, M., Hofkens, J., Habuchi, S., Dirix, G., Van Guyse, M., Michiels, J., Vanderleyden, J., and De Schryver, F. C. (2001). Identification of different emitting species in the red fluorescent protein DsRed by means of ensemble and single-molecule spectroscopy. *Proc. Natl. Acad. Sci. USA* 27:27.

Crameri, A., Whitehorn, E. A., Tate, E., and Stemmer, W. P. (1996). Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat. Biotechnol.* 14:315–319.

Creemers, T. M., Lock, A. J., Subramaniam, V., Jovin, T. M., and Volker, S. (1999). Three photoconvertible forms of green fluorescent protein identified by spectral hole-burning. *Nat. Struct. Biol.* 6:557–560.

Cubitt, A. B., Heim, R., Adams, S. R., Boyd, A. E., Gross, L. A., and Tsien, R. Y. (1995). Understanding, improving and using green fluorescent proteins. *Trends Biochem. Sci.* 20:448–455.

Cubitt, A. B., Woollenweber, L. A., and Heim, R. (1999). Understanding structure-function relationships in the Aequorea victoria green fluorescent protein. *Methods Cell Biol.* 58:19–30.

Davenport, D., and Nichol, J. A. C. (1955). *Proc. R. Soc. London, Ser. B* 144:399–411.

De Angelis, D. A., Miesenbock, G., Zemelman, B. V., and Rothman, J. E. (1998). PRIM: proximity imaging of green fluorescent protein-tagged polypeptides. *Proc. Natl. Acad. Sci. USA* 95:12312–12316.

Delagrave, S., Hawtin, R. E., Silva, C. M., Yang, M. M., and Youvan, D. C. (1995). Red-shifted excitation mutants of the green fluorescent protein. *Biotechnology (NY)* 13:151–154.

Demaurex, N., and Frieden, M. (2003). Measurements of the free luminal ER $Ca^{(2+)}$ concentration with targeted "cameleon" fluorescent proteins. *Cell Calcium* 34:109–119.

Dewey, T. G., and Hammes, G. G. (1980). Calculation on fluorescence resonance energy transfer on surfaces. *Biophys. J.* 32:1023–1035.

Dewey, T. G., and Datta, M. M. (1989). Determination of the fractal dimension of membrane protein aggregates using fluorescence energy transfer. *Biophys. J.* 56:415–420.

Doi, N., and Yanagawa, H. (1999). Design of generic biosensors based on green fluorescent proteins with allosteric sites by directed evolution. *FEBS Lett.* 453:305–307.

Dopf, J., and Horiagon, T. M. (1996). Deletion mapping of the *Aequorea victoria* green fluorescent protein. *Gene* 173:39–44.

Dove, S. G., Hoegh-Guldberg, O., and Ranganathan, S. (2001). Major colour patterns of reef-building corals are due to a family of GFP-like proteins. *Coral Reefs* 19:197–204.

Ehrig, T., O'Kane, D. J., and Prendergast, F. G. (1995). Green-fluorescent protein mutants with altered fluorescence excitation spectra. *FEBS Lett.* 367:163–166.

Fukuda, H., Arai, M., and Kuwajima, K. (2000). Folding of green fluorescent protein and the cycle3 mutant. *Biochemistry* 39:12025–12032.

Fung, B. K., and Stryer, L. (1978). Surface density determination in membranes by fluorescence energy transfer. *Biochemistry* 17:5241–5248.

Garcia-Parajo, M. F., Koopman, M., van Dijk, E. M., Subramaniam, V., and van Hulst, N. F. (2001). The nature of fluorescence emission in the red fluorescent protein DsRed, revealed by single-molecule detection. *Proc. Natl. Acad. Sci. USA* 27:27.

Gerami-Nejad, M., Berman, J., and Gale, C. A. (2001). Cassettes for PCR-mediated construction of green, yellow, and cyan fluorescent protein fusions in *Candida albicans*. *Yeast* 18:859–864.

Gibbs, M. D., Nevalainen, K. M., and Bergquist, P. L. (2001). Degenerate oligonucleotide gene shuffling (DOGS): A method for enhancing the frequency of recombination with family shuffling. *Gene* 271:13–20.

Giver, L., and Arnold, F. H. (1998). Combinatorial protein design by in vitro recombination. *Curr. Opin. Chem. Biol.* 2:335–338.

Graf, R., and Schachman, H. K. (1996). Random circular permutation of genes and expressed polypeptide chains: Application of the method to the catalytic chains of aspartate transcarbamoylase. *Proc. Natl. Acad. Sci. USA* 93:11591–11596.

Griesbeck, O., Baird, G. S., Campbell, R. E., Zacharias, D. A., and Tsien, R. Y. (2001). Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *J. Biol. Chem.* 276:29188–29194.

Gross, L. A., Baird, G. S., Hoffman, R. C., Baldridge, K. K., and Tsien, R. Y. (2000). The structure of the chromophore within DsRed, a red fluorescent protein from coral. *Proc. Natl. Acad. Sci. USA* 97:11990–11995.

Guerrero, G., Siegel, M. S., Roska, B., Loots, E., and Isacoff, E. Y. (2002). Tuning Flash: Redesign of the dynamics, voltage range, and color of the genetically encoded optical sensor of membrane potential. *Biophys. J.* 83:3607–3618.

Gurskaya, N. G., Fradkov, A. F., Pounkova, N. I., Staroverov, D. B., Bulina, M. E., Yanushevich, Y. G., Labas, Y. A., Lukyanov, S., and Lukyanov, K. A. (2003). A colourless green fluorescent protein homologue from the non-fluorescent hydromedusa *Aequorea coerulescens* and its fluorescent mutants. *Biochem. J.* 373:403–408.

Hanson, G. T., McAnaney, T. B., Park, E. S., Rendell, M. E., Yarbrough, D. K., Chu, S., Xi, L., Boxer, S. G., Montrose, M. H., and Remington, S. J. (2002). Green fluorescent protein variants as ratiometric dual emission pH sensors. 1. Structural characterization and preliminary application. *Biochemistry* 41:15477–15488.

Harpur, A. G., Wouters, F. S., and Bastiaens, P. I. (2001). Imaging FRET between spectrally similar GFP molecules in single cells. *Nat. Biotechnol.* 19:167–169.

Haseloff, J., and Amos, B. (1995). GFP in plants. *Trends Genet.* 11:328–329.

Hauser, K., Haynes, W. J., Kung, C., Plattner, H., and Kissmehl, R. (2000). Expression of the green fluorescent protein in *Paramecium tetraurelia*. *Eur. J. Cell. Biol.* 79:144–149.

Heikal, A. A., Hess, S. T., Baird, G. S., Tsien, R. Y., and Webb, W. W. (2000). Molecular spectroscopy and dynamics of intrinsically fluorescent proteins: Coral red (dsRed) and yellow (Citrine). *Proc. Natl. Acad. Sci. USA* 97:11996–12001.

Heim, R. (1999). Green fluorescent protein forms for energy transfer. *Methods Enzymol.* 302:408–423.

Heim, R., and Tsien, R. Y. (1996). Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Biol.* 6:178–182.

Heim, R., Prasher, D. C., and Tsien, R. Y. (1994). Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci. USA* 91:12501–12504.

Heim, R., Cubitt, A. B., and Tsien, R. Y. (1995). Improved green fluorescence. *Nature* 373:663–664.

Hofmann, A., Iwai, H., Hess, S., Pluckthun, A., and Wlodawer, A. (2002). Structure of cyclized green fluorescent protein. *Acta Crystallogr. D Biol. Crystallogr.* 58:1400–1406.

Hopf, M., Gohring, W., Ries, A., Timpl, R., and Hohenester, E. (2001). Crystal structure and mutational analysis of a perlecan-binding fragment of nidogen-1. *Nat. Struct. Biol.* 8:634–640.

Hyun Bae, J., Rubini, M., Jung, G., Wiegand, G., Seifert, M. H., Azim, M. K., Kim, J. S., Zumbusch, A., Holak, T. A., Moroder, L., Huber, R., and Budisa, N. (2003). Expansion of the genetic code enables design of a novel "gold" class of green fluorescent proteins. *J. Mol. Biol.* 328:1071–1081.

Isshiki, M., Ying, Y. S., Fujita, T., and Anderson, R. G. (2002). A molecular sensor detects signal transduction from caveolae in living cells. *J. Biol. Chem.* 277:43389–43398.

Jayaraman, S., Haggie, P., Wachter, R. M., Remington, S. J., and Verkman, A. S. (2000). Mechanism and cellular applications of a green fluorescent protein-based halide sensor. *J. Biol. Chem.* 275:6047–6050.

Jones, J., Heim, R., Hare, E., Stack, J., and Pollok, B. A. (2000). Development and application of a GFP-FRET intracellular caspase assay for drug screening. *J. Biomol. Screen* 5:307–318.

Kaether, C., and Gerdes, H. H. (1995). Visualization of protein transport along the secretory pathway using green fluorescent protein. *FEBS Lett.* 369:267–271.

Kahana, J. A., and Silver, P. A. (1996). Use of the *A. Victoria* green fluorescent protein to study protein dynamics *in vivo*. In *Current Protocols in Molecular Biology*, Aosobel, F. M., Brent, R., Kingston, R. E., Moore, D. E., Seidman, J. G., Smith, J. A., and Struhl, K., Eds., (New York, NY, John Wiley & Sons, Inc.), pp. 9.7.22–29.27.27.

Kelmanson, I. V., and Matz, M. V. (2003). Molecular basis and evolutionary origins of color diversity in great star coral *Montastraea cavernosa* (Scleractinia: Faviida). *Mol. Biol. Evol.* 20:1125–1133.

Kerr, R., Lev-Ram, V., Baird, G., Vincent, P., Tsien, R. Y., and Schafer, W. R. (2000). Optical imaging of calcium transients in neurons and pharyngeal muscle of *C. elegans*. *Neuron* 26:583–594.

Kim, H. K., and Kaang, B. K. (1998). Truncated green fluorescent protein mutants and their expression in *Aplysia* neurons. *Brain. Res. Bull.* 47:35–41.

Kimata, Y., Iwaki, M., Lim, C. R., and Kohno, K. (1997). A novel mutation which enhances the fluorescence of green fluorescent protein at high temperatures. *Biochem. Biophys. Res. Commun.* 232:69–73.

Kozak, M. (1989). The scanning model for translation: An update. *J. Cell Biol.* 108:229–241.

Kummer, A. D., Wiehler, J., Schuttrigkeit, T. A., Berger, B. W., Steipe, B., and Michel-Beyerle, M. E. (2002). Picosecond time-resolved fluorescence from blue-emitting chromophore variants Y66F and Y66H of the green fluorescent protein. *Chembiochem* 3:659–663.

Kuner, T.. and Augustine G. J. (2000). A genetically encoded ratiometric indicator for chloride: capturing chloride transients in cultured hippocampal neurons. *Neuron* 27:447–459.

Lakowicz, J. (1999). *Principles of Fluorescence Spectroscopy*, 2nd ed., Plenum, New York.

Levy, J. P., Muldoon, R. R., Zolotukhin, S., and Link, C. J., Jr. (1996). Retroviral transfer and expression of a humanized, red-shifted green fluorescent protein gene into human tumor cells. *Nat. Biotechnol.* 14:610–614.

Lim, C. R., Kimata, Y., Oka, M., Nomaguchi, K., and Kohno, K. (1995). Thermosensitivity of green fluorescent protein fluorescence utilized to reveal novel nuclear-like compartments in a mutant nucleoporin NSP1. *J. Biochem. (Tokyo)* 118:13–17.

Liu, L., Yermolaieva, O., Johnson, W. A., Abboud, F. M., and Welsh, M. J. (2003). Identification and function of thermosensory neurons in *Drosophila* larvae. *Nat. Neurosci.* 6:267–273.

Llopis, J., McCaffery, J. M., Miyawaki, A., Farquhar, M. G., and Tsien, R. Y. (1998). Measurement of cytosolic, mitochondrial, and Golgi pH in single living cells with green fluorescent proteins. *Proc. Natl. Acad. Sci. USA* 95:6803–6808.

Luo, K. Q., Yu, V. C., Pu, Y., and Chang, D. C. (2001). Application of the fluorescence resonance energy transfer method for studying the dynamics of caspase-3 activation during UV-induced apoptosis in living HeLa cells. *Biochem. Biophys. Res. Commun.* 283:1054–1060.

Malli, R., Frieden, M., Osibow, K., and Graier, W. F. (2003). Mitochondria efficiently buffer subplasmalemmal Ca^{2+} elevation during agonist stimulation. *J. Biol. Chem.* 278:10807–10815.

Martynov, V. I., Savitsky, A. P., Martynova, N. Y., Savitsky, P. A., Lukyanov, K. A., and Lukyanov, S. A. (2001). Alternative cyclization in GFP-like proteins family. The formation and structure of the chromophore of a purple chromoprotein from Anemonia sulcata. *J. Biol. Chem.* 276:21012–21016.

Matz, M. V., Lukyanov, K. A., and Lukyanov, S. A. (2002). Family of the green fluorescent protein: Journey to the end of the rainbow. *Bioessays* 24:953–959.

McAnaney, T. B., Park, E. S., Hanson, G. T., Remington, S. J., and Boxer, S. G. (2002). Green fluorescent protein variants as ratiometric dual emission pH sensors. 2. Excited-state dynamics. *Biochemistry* 41:15489–15494.

McElroy, W. D. (1947). The energy source for biolominescence in an isolated system, *Proc. Natl. Acad. Sci. USA* 33:342–345.

McRorie, D., and Voelker, P. (1993). *Self-Associating Systems in the Analytical Ultracentrifuge*, Beckman Instruments, Fullerton, CA.

Miesenbock, G., De Angelis, D. A., and Rothman, J. E. (1998). Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature* 394:192–195.

Minshull, J., and Stemmer, W. P. (1999). Protein evolution by molecular breeding. *Curr. Opin. Chem. Biol.* 3:284–290.

Mitra, R. D., Silva, C. M., and Youvan, D. C. (1996). Fluorescence resonance energy transfer between blue-emitting and red-shifted excitation derivatives of the green fluorescent protein. *Gene* 173:13–17.

Miyawaki, A., Llopis, J., Heim, R., McCaffery, J. M., Adams, J. A., Ikura, M., and Tsien, R. Y. (1997). Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin. *Nature* 388:882–887.

Miyawaki, A. (2003). Visualization of the spatial and temporal dynamics of intracellular signaling. *Dev. Cell.* 4:295–305.

Mizuno, H., Sawano, A., Eli, P., Hama, H., and Miyawaki, A. (2001). Red fluorescent protein from *Discosoma* as a fusion tag and a partner for fluorescence resonance energy transfer. *Biochemistry* 40:2502–2510.

Morise, H., Shimomura, O., Johnson, F. H., and Winant, J. (1974). Intermolecular energy transfer in the bioluminescent system of *Aequorea*. *Biochemistry* 13:2656–2662.

Muhlrad, D., Hunter, R., and Parker, R. (1992). A rapid method for localized mutagenesis of yeast genes. *Yeast* 8:79–82.

Nagai, T., Sawano, A., Park, E. S., and Miyawaki, A. (2001). Circularly permuted green fluorescent proteins engineered to sense Ca^{2+}. *Proc. Natl. Acad. Sci. USA* 98:3197–3202.

Nagai, T., Ibata, K., Park, E. S., Kubota, M., Mikoshiba, K., and Miyawaki, A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* 20:87–90.

Nakai, J., Ohkura, M., and Imoto, K. (2001). A high signal-to-noise $Ca^{(2+)}$ probe composed of a single green fluorescent protein. *Nat. Biotechnol.* 19:137–141.

Niwa, H., Inouye, S., Hirano, T., Matsuno, T., Kojima, S., Kubota, M., Ohashi, M., and Tsuji, F. I. (1996). Chemical nature of the light emitter of the *Aequorea* green fluorescent protein. *Proc. Natl. Acad. Sci. USA* 93:13617–13622.

Ormo, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., and Remington, S. J. (1996). Crystal structure of the Aequorea victoria green fluorescent protein. *Science* 273:1392–1395.

Palm, G., and Wlodawer, A. (1999). Spectral variants of green fluorescent protein. *Methods Enzymol.* 302:378–394.

Palm, G. J., Zdanov, A., Gaitanaris, G. A., Stauber, R., Pavlakis, G. N., and Wlodawer, A. (1997). The structural basis for spectral variations in green fluorescent protein. *Nat. Struct. Biol.* 4:361–365.

Palmer, A. E., Jin, C., Reed, J. C., and Tsien, R. Y. (2004). Bcl-2-mediated alterations in endoplasmic reticulum Ca^{2+} analyzed with an improved genetically encoded fluorescent sensor. *Proc. Natl. Acad. Sci. USA* 101:17404–17409.

Patterson, G. H., and Lippincott-Schwartz, J. (2002). A photoactivatable GFP for selective photolabeling of proteins and cells. *Science* 297:1873–1877.

Patterson, G. H., Knobel, S. M., Sharif, W. D., Kain, S. R., and Piston, D. W. (1997). Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophys. J.* 73:2782–2790.

Patterson, G. H., Piston, D. W., and Barisas, B. G. (2000). Forster distances between green fluorescent protein pairs. *Anal. Biochem.* 284:438–440.

Prasher, D. (1995). Using GFP to see the light. *Trends Genet.* 11:320–232.

Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G., and Cormier, M. J. (1992). Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* 111:229–233.

Reiff, D. F., Thiel, P. R., and Schuster, C. M. (2002). Differential regulation of active zone density during long-term strengthening of *Drosophila* neuromuscular junctions. *J. Neurosci.* 22:9399–9409.

Rekas, A., Alattia, J. R., Nagai, T., Miyawaki, A., and Ikura, M. (2002). Crystal structure of Venus, a yellow fluorescent protein with improved maturation and reduced environmental sensitivity. *J. Biol. Chem.* 4:4.

Richmond, T. A., Takahashi, T. T., Shimkhada, R., and Bernsdorf, J. (2000). Engineered metal binding sites on green fluorescence protein. *Biochem. Biophys. Res. Commun.* 268:462–465.

Rizzo, M. A., Springer, G. H., Granada, B., and Piston, D. W. (2004). An improved cyan fluorescent protein variant useful for FRET. *Nat. Biotechnol.* 22:445–449.

Rizzuto, R., Brini, M., De Giorgi, F., Rossi, R., Heim, R., Tsien, R. Y., and Pozzan, T. (1996). Double labelling of subcellular structures with organelle-targeted GFP mutants *in vivo*. *Curr. Biol.* 6:183–188.

Salih, A., Larkum, A., Cox, G., Kuhl, M., and Hoegh-Guldberg, O. (2000). Fluorescent pigments in corals are photoprotective. *Nature* 408:850–853.

Sankaranarayanan, S., De Angelis, D., Rothman, J. E., and Ryan, T. A. (2000). The use of pHluorins for optical measurements of presynaptic activity. *Biophys. J.* 79:2199–2208.

Sawano, A., and Miyawaki, A. (2000). Directed evolution of green fluorescent protein by a new versatile PCR strategy for site-directed and semi-random mutagenesis. *Nucleic Acids Res.* 28:E78.

Scharnagl, C., Raupp-Kossmann, R., and Fischer, S. F. (1999). Molecular basis for pH sensitivity and proton transfer in green fluorescent protein: Protonation and conformational substates from electrostatic calculations. *Biophys. J.* 77:1839–1857.

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E., and Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat. Biotechnol.* 22:1567–1572.

Shimomura, O. (1979). Structure of the chromophore of *Aequorea* green fluorescent protein. *FEBS Lett.* 104:220–222.

Siegel, M. S., and Isacoff, E. Y. (1997). A genetically encoded optical probe of membrane voltage. *Neuron* 19:735–741.

Siemering, K. R., Golbik, R., Sever, R., and Haseloff, J. (1996). Mutations that suppress the thermosensitivity of green fluorescent protein. *Curr. Biol.* 6:1653–1663.