A FLASH of insight into cellular chemistry: genetically encoded labels for protein visualization *in vivo*

Kevin J Luebke

Genetically encoded fluorescent labels, such as green fluorescent protein, make it possible to visualize a protein's natural distribution and environment in living cells. A new approach to protein labeling in living cells has been devised in which a small, membrane-permeable ligand binds with high affinity and specificity to a short peptide motif that can be incorporated into the protein of interest; the ligand becomes brightly fluorescent after binding to the peptide.

Address: Center for Biomedical Inventions, Department of Internal Medicine, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Boulevard, Dallas, TX 75235-8573, USA.

E-mail: Luebke@ryburn.swmed.edu

Chemistry & Biology December 1998, 5:R317–R322
http://biomednet.com/elecref/10745521005R0317

When you look at a living cell under a light microscope, only the most morphologically profound cellular events are visible. But underlying the events of major morphological reorganization is a complex, unseen bustle of chemical events. Cell biologists have often taken advantage of the sensitivity and versatility of fluorescent probes to explore the chemical details of the intracellular environment, using such probes as tags and positional markers to measure distance and intracellular transport, and as sensors of intracellular concentrations of small molecules and ions [1].

A fluorescent label attached to a specific protein allows that protein's distribution and environment to be visualized. Chemical labeling of a purified protein *in vitro* does not lend itself to visualization of the unperturbed intracellular milieu, however. Introduction of fluorescently labeled antibodies raised against a protein of interest is feasible only in special cases, and antibody binding can interfere with the function of the protein. In contrast, intrinsically fluorescent proteins have found wide applicability as genetically encodable fluorescent labels. By encoding a label into the primary sequence of a protein of interest at the level of its DNA, that protein can be specifically observed in the context of the cell itself. Disruptive post-translational manipulation of the protein is avoided.

The archetype of genetically encoded fluorescent labels is the green fluorescent protein (GFP) from the jellyfish *Aequorea victoria*, a 238-residue protein in which three consecutive residues, serine, tyrosine and glycine, spontaneously react to form a fluorogenic species (Figure 1) [2–5]. GFP is well suited to the role of a genetically encoded label, because the fluorophore forms upon expression of the protein without requirement for any additional factors. As a result, it can be expressed in fluorescent form in a wide variety of non-native cell types and fused to the amino or carboxyl termini of other proteins [2].

Griffin, Adams and Tsien [6] have recently developed a new approach to fluorescently labeling a protein of interest in living cells. They devised a peptide–small-molecule pair that affords specific, covalent labeling of proteins containing the short peptide. This methodology overcomes some of the limitations of using a fluorescent protein as an encoded *in situ* label and makes new experiments possible. The work by Griffin *et al.* [6] is reviewed here in the context of some of the applications of fluorescent protein-labeling methods, including some recent innovations.

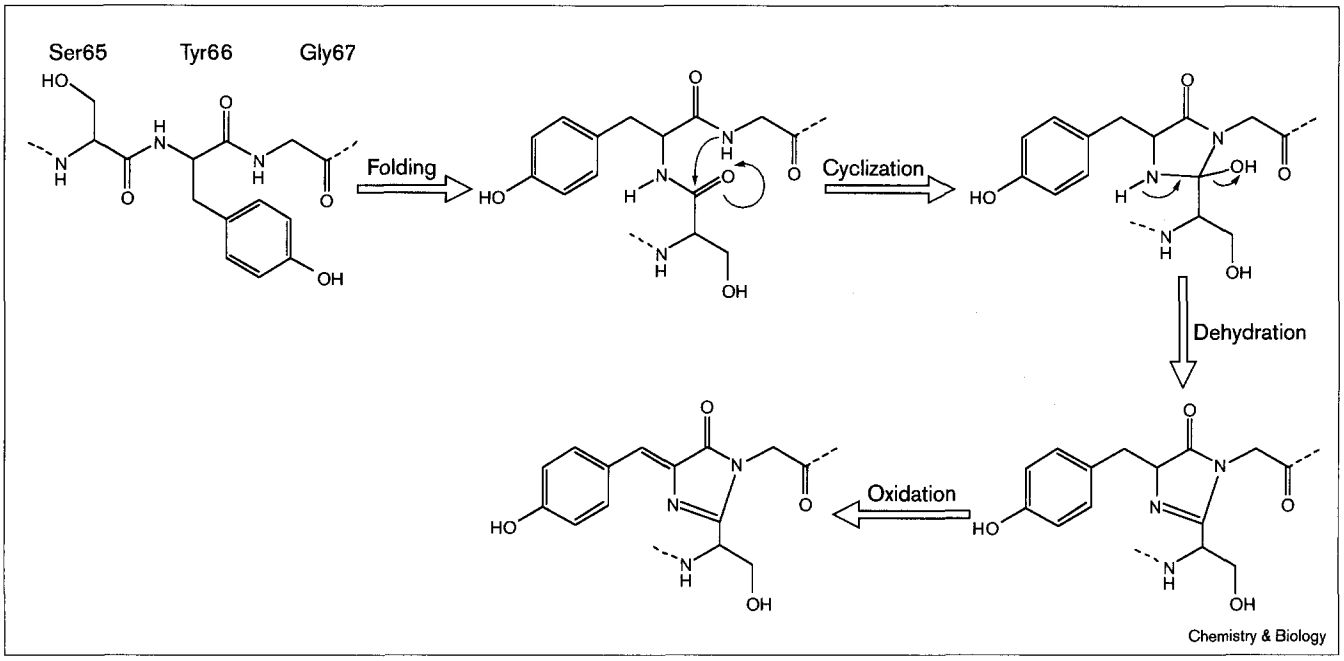
Fluorescent proteins as reporters of gene expression and protein localization

Among the first applications of GFP as an *in situ* protein label was the visualization of gene expression. Chalfie and coworkers [7] showed GFP to be a reporter of gene expression in *Escherichia coli* and *Caenorhabditis elegans*. In *C. elegans*, control of GFP expression by the promoter for a β-tubulin gene that is expressed selectively in certain touch-sensitive cells allowed selective fluorescence visualization of those cells in the larval nematode. In addition to monitoring gene expression, GFP can be employed to study protein localization. For many proteins, fusion of GFP at the amino or carboxyl terminus preserves both the fluorescence of the GFP and the native function and cellular localization of the attached protein. For that reason, GFP fusion proteins have been used as markers to follow cellular protein traffic [8–10] and the dynamic behavior of specifically labeled organelles [11,12].

Fluorescent proteins for monitoring protein–protein interactions *in vivo*

The interaction of two fluorescently labeled proteins can be monitored using fluorescence resonance energy transfer (FRET) [13]. FRET occurs between two fluorophores incorporated into two proteins that are associated with each other and diminishes precipitously upon their dissociation. Taking advantage of this phenomenon with GFP requires a FRET partner with appropriate excitation and emission spectra; GFP mutants with spectral maxima shifted from those of the wild-type protein provide such partners [14]. Mutation of

Figure 1



Formation of the fluorophore of green fluorescent protein. Folding of the protein promotes cyclization, which is followed by dehydration of the ring and oxidation of the tyrosine. For more details please see the text.

the tyrosine constituent of the fluorophore in GFP to histidine or tryptophan results in shifts of the excitation and emission maxima to shorter wavelengths; further mutagenesis can recover some of the brightness lost in these mutants [14]. A variety of other GFP mutants with altered spectral properties have been found, including a yellow mutant [15] and green mutants with enhanced photostability and brightness [14,16].

FRET between green and blue mutants of GFP has been used elegantly by Mahajan *et al.* [17] to demonstrate an interaction between two proteins, Bax and Bcl-2, in mammalian cells. Bax and Bcl-2 are both involved in apoptosis, a process in which the cell kills itself by executing a controlled program of biochemical events. Bax mediates a series of events that lead to cell death, and several experiments have suggested that apoptosis is regulated by direct interaction between Bax and Bcl-2. Fusions of GFP with Bax and BFP (blue fluorescent protein) with Bcl-2 were co-expressed in mammalian cell lines. These fusions were shown in control experiments to have biological activities similar to those of the wild-type proteins. Not only was FRET observed between the two fluorescent proteins, consistent with a direct interaction between Bax and Bcl-2, but the fluorescence from FRET was localized to the mitochondria. In addition to demonstrating the interaction of Bax and Bcl-2, the cellular location of the interaction was visualized directly.

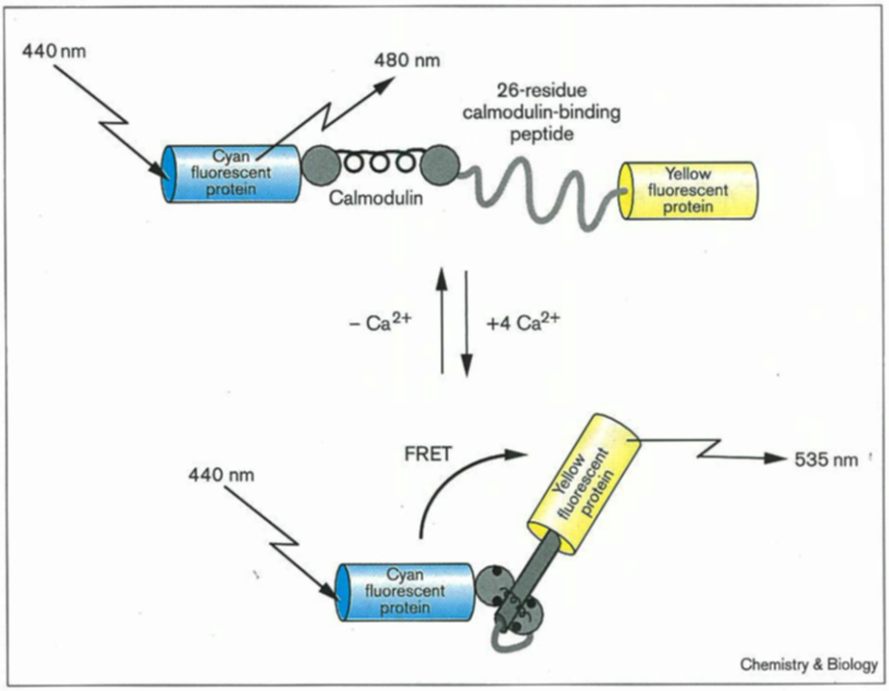
Sensing ligand-dependent conformational changes of proteins to monitor intracellular concentrations of Ca^{2+}

Just as FRET between proteins can be used to observe their interactions, it can also be used to observe conformational changes within a protein that alter the distance and orientation between two fluorophores. If the conformational change depends upon binding of a specific ligand, the fluorescence signal resulting from FRET can be used as an indicator for that ligand. Many cellular events are regulated by small-molecule effectors, concentrations of which can vary locally within the cell and change rapidly in time; measurement of these events therefore requires high spatial and temporal resolution, and indicators based on conformational changes within proteins have the potential to provide that resolution.

This type of approach has been used to create genetically encoded fluorescent indicators of Ca^{2+} concentration [18,19]. Calcium ions regulate a variety of cellular processes, mediated primarily by the calcium-binding protein calmodulin. Miyawaki *et al.* [18] fused a 26-residue calmodulin-binding peptide from myosin light-chain kinase to the carboxyl terminus of calmodulin and a different-colored variant of GFP to each end of the resulting fusion (Figure 2). When Ca^{2+} binds to the fusion protein's calmodulin domain, and calmodulin-binding domains, the protein is transformed from an extended structure to a more compact globular structure in which the two fluorescent protein domains are closer

Figure 2

Design of fluorescence indicators for Ca^{2+} using fluorescent proteins and calmodulin. Variants of GFP are appended to the termini of a fusion of calmodulin with a 26-residue calmodulin-binding peptide. Upon association with Ca^{2+}, calmodulin wraps around the 26-residue peptide, bringing the two fluorescent proteins closer to each other and enhancing FRET between them.



to each other, and FRET between them is enhanced. This indicator was sensitive to changes in Ca^{2+} from $<10^{-7}$ to $>10^{-4}$ M *in vitro*, and its response could be tuned by mutagenesis of the calmodulin components. When expressed in cultured human epithelial cells, this indicator had a response to cytosolic Ca^{2+} similar to its Ca^{2+} response *in vitro*. One benefit of a genetically encoded Ca^{2+} sensor is its intracellular targetability. Intracellular Ca^{2+} concentrations have previously been measured using synthetic fluorescent chelators, which are difficult to target to specific intracellular locations [18,20–22]. Miyawaki *et al.* [18] successfully targeted Ca^{2+}-responsive fluorescent proteins to the nucleus or the endoplasmic reticulum by adding a nuclear localization signal or an endoplasmic reticulum retention sequence, respectively.

Visualizing exocytosis and synaptic transmission with pH-sensitive fluorescent proteins

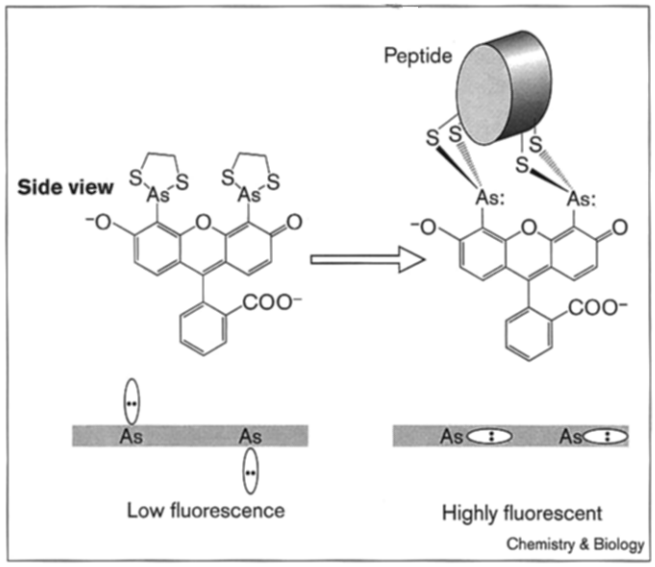
The genetic accessibility and targetability of fluorescent proteins makes them valuable as indicators of the intracellular environment. To create targetable, intracellular pH sensors, Miesenbock, De Angelis and Rothman [23] have screened for pH-sensitive GFP mutants, and have identified two classes: the ecliptic class of mutants loses fluorescence as pH is lowered from 7.5 to 5.5; and the ratiometric mutants undergo a rapid and reversible change in their excitation ratio at different excitation maxima in response to changes in pH. The response of a ratiometric indicator was calibrated by targeting the indicator to the surface of human epithelial cells and imaging the cells in buffers of known pH. With that calibration, the pH inside cellular compartments was measured.

Miesenbock *et al.* [23] used the pH-indicator fluorescent proteins to observe the fusion of vesicles with the plasma membrane that occurs during exocytosis. The pH inside secretory vesicles is acidic, but as the vesicle membrane fuses with the plasma membrane and the contents of the vesicle are spilled into the extracellular fluid, the pH environment of the inner surface of the vesicle membrane equilibrates with that of the extracellular fluid, at about 7.4. An optical sensor of pH attached to the inner surface of the vesicle membrane therefore has the potential to report individual exocytotic events. Miesenbock *et al.* [23] targeted a ratiometric indicator protein to the inner walls of vesicles in hippocampal neurons forming an array of synapses in culture. By monitoring the change in the emission ratio, they were able to observe synaptic transmission induced in these neurons. Vesicle proteins are recycled by endocytosis into new vesicles, and the acidification of the environment of the vesicle proteins was also observed as a change in the emission ratio. Ecliptic indicators targeted to the inner surface of vesicles eliminate background from resting vesicles, because with long-wavelength excitation they are not fluorescent under the acidic conditions inside the vesicle.

Covalent labeling of recombinant proteins in living cells using a small molecule

Their great utility notwithstanding, GFP and its variants have limitations as *in situ* protein labels and chemical

Figure 3



Association of a designed peptide–small-molecule pair. Four cysteines are incorporated into the sequence of an α-helical peptide at positions that present the four thiol groups on one face of the helix. The ligand (FLASH) contains two arsenic centers that are spatially disposed toward simultaneous interaction of each with a thiol pair of the peptide. A molecule of 1,2-ethanedithiol (EDT) chelates each arsenic center in the free ligand, preventing association with cellular thiols. Association of the ligand with the peptide displaces EDT.

sensors [6]. For one, control of the spectral properties of the fluorophore is limited to what can be achieved by a limited set of mutations to the protein; most mutations in GFP result in a loss of fluorescence with little other change in its absorbance or emission spectra [2]. Another drawback of GFP is that it is a relatively large label and, as such, could influence the functional behavior of the protein to which it is conjugated. In addition, GFP can only be attached at a terminus of the protein of interest. The limitation to fluorescence visualization is another drawback. It is easy to imagine systems in which other types of reporters would be valuable. For example, visualizing processes in whole multicellular organisms (beyond optically transparent organisms such as nematodes) could benefit from the availability of magnetic resonance reporters that are cell-type-specific and intracellularly targetable.

Griffin, Adams, and Tsien [6] sought a method for specific labeling of proteins in living cells that overcomes these limitations. Their strategy was to devise a chemically complementary pair comprising a small peptide module that could be genetically incorporated into proteins to be labeled and a small ligand that binds to that module with sufficient affinity and specificity that the peptide is selectively labeled in the context of a mammalian cell. The ligand had to be cell-permeable and it had to be derivatizable with various reporter groups, such as spectroscopic probes.

As a basis of molecular recognition between a peptide and small ligand, Griffin *et al.* [6] sought a high affinity chemical interaction that could be formed rapidly with functional groups present in genetically encodable peptides. They chose the formation of covalent bonds between trivalent arsenic centers and pairs of thiols. The thiols from a pair of cysteine residues appropriately placed in the peptide can chelate the arsenic center of an organic ligand. The interaction had to be specific for thiols in the target peptide over other pairs of cellular thiols — specificity is required for visualization of the protein of interest and also to prevent perturbation of cellular processes by nonspecific binding. To create a specific interaction, Griffin *et al.* [6] relied on the cooperativity of two pairs of thiols from the target peptide simultaneously chelating two arsenic centers in a single ligand. They designed a peptide that was expected to fold into an α helix in which four cysteines, at positions i, $i+1$, $i+4$, and $i+5$, present their thiol groups on one face (Figure 3). They reasoned that an organic molecule containing two trivalent arsenic centers spaced appropriately could bind to this peptide through interaction of each of the arsenic centers with a thiol pair. They anticipated that the cooperativity of this interaction would make it highly favored over interactions with single thiols or individual pairs of thiols. Chelation of the arsenic centers with ethanedithiol can be used to prevent the ligand from binding to endogenous cellular thiols; the ethanedithiol is displaced by formation of the thermodynamically favored complex of the ligand with the tetracysteine peptide.

One of 14 bi-arsenical ligands tested, a fluorescein derivative, bound to the tetracysteine peptide in the presence of a small excess of ethanedithiol. Serendipitously, this compound, termed FLASH (fluorescein arsenical helix binder) by Griffin *et al.* [6], fluoresces brightly when bound to the peptide, but is more than four orders of magnitude less fluorescent (practically nonfluorescent) when bound to ethanedithiol. Conjugation of the arsenic lone-pair electrons with the fluorescein orbitals is thought to allow quenching of the excited state by vibrational de-activation or photo-induced electron transfer. In the complex with the peptide, the arsenic lone pairs are expected to be held out of conjugation with the fluorescein orbitals.

To explore the use of the FLASH–peptide pair in cells, Griffin *et al.* [6] expressed a fusion of the designed peptide and a cyan mutant of GFP in human epithelial cells. In the absence of FLASH, the cells expressing the fusion were identified by their bright fluorescence at the emission maximum of the fluorescent protein. Upon treatment of the cells with 1 μM FLASH bis-ethanedithiol (accompanied by 10 μM ethanedithiol), FRET was observed between the protein fluorophore and FLASH in the cells that were expressing the fusion, confirming the cell-permeability of FLASH. Complete binding required approximately one hour. This slow equilibration limits the time

scale at which observations of events such as gene expression can be made. It is noteworthy, however, that GFP acquires fluorescence after expression on approximately the same time scale [3,24].

One indication of the specificity of FLASH for its cognate peptide relative to the other components of a mammalian cell can be seen from the background fluorescence in cells not expressing a tetracysteine-peptide-labeled protein in the presence of FLASH and ethanedithiol. Dim fluorescence, attributable to FLASH staining, is observed associated with mitochondria. This background was reduced by increasing the concentration of ethanedithiol, suggesting that it is due to relatively weak association of the ligand with abundant binding sites. The specificity of the desired labeling interaction could be increased, therefore, by a modest increase in the affinity of the ligand for the target peptide, allowing the association to occur at higher ethanedithiol concentrations. The authors suggest that this improvement might be accomplished by a combinatorial optimization of the noncysteine residues in the receptor. Even without such optimization, competing binding sites that induce fluorescence in the ligand are sufficiently rare in mammalian cells that fluorescence detection of the target peptide over background when expressed in transfected cells is permitted. Another indication of the specificity of FLASH for its cognate peptide is its lack of toxicity to the cell. The toxic effects of arsenic compounds are largely attributable to binding to cellular thiols, but cells treated with 1 μM FLASH in the presence of 10 μM ethanedithiol remained viable for at least four hours.

The constellation of four cysteines that form the ligand-binding site was also engineered into an existing α helix within a protein. Griffin *et al.* [6] introduced the four-cysteine array, cysteine–cysteine–X–X–cysteine–cysteine (where X is any amino acid), into the amino-terminal α helix of *Xenopus* calmodulin, and expressed it in human epithelial cells. The cytosol and nuclei of these cells became brightly fluorescent in the presence of 1 μM FLASH and 10 μM ethanedithiol. Although it might not be possible to incorporate the FLASH-binding tetracysteine motif arbitrarily into α helices without disrupting either protein structure or function or the affinity for the label, judicious selection of incorporation sites based on structural knowledge might sometimes obviate the need to engineer new secondary structural elements into proteins of interest.

The biarsenical–peptide pair invented by Griffin, Adams, and Tsien [6] provides the ability to target a small-molecule probe to a single, genetically specified (i.e. labeled) protein *in vivo*. It can be used in many of the same types of experiment as GFP. Experiments that rely on FRET will, however, require an energy-transfer partner for the label. This requirement could be satisfied using a GFP variant, but one of the benefits of the peptide–biarsenical pair is that it is potentially less disruptive to protein structure and function than fusion with GFP. This benefit is partially lost if fusion to a GFP variant is still required. Thus, an important challenge in the development of this methodology is the invention of a second peptide–small-molecule pair that does not cross react with the biarsenical–helix pair.

One of the primary benefits of this *in situ* protein labeling system is its potential flexibility for use with different probes attached to the FLASH ligand. The fortuitous enhancement of fluorescence upon binding of FLASH to the peptide minimizes interference of free ligand when detecting that fluorescence signal, but enhancement of signal with binding is not necessarily required for observation of the complex with a different probe. The FLASH–peptide complex, once formed, dissociates imperceptibly in the absence of excess vicinal dithiol like ethanedithiol, so background signal could be diminished by removing free ligand without diminishing signal from the peptide-bound ligand. A higher standard of specificity, against all modes of binding to nontargeted cellular components, not just those that result in enhanced fluorescence signal, is required in that case. In principle, though, any probe can be conjugated to the biarsenical ligand, as long as the modification preserves cell permeability and binding specificity of the ligand. Whereas fluorescent proteins have been engineered to create sensitivity to pH or Ca^{2+}, a fluorophore with its own sensitivity to pH or small molecules or ions could be delivered to a specified protein, attached to the biarsenical helix-binding ligand.

Genetically encoded protein labels provide the spatial and temporal resolution to observe intracellular molecular processes in real time. In essence, they allow the chemistry inside a living cell to be observed much as the morphological behavior of a cell can be watched. Fluorescent proteins have been used as labels to monitor gene expression, protein localization, protein–protein interactions and the intracellular environment. These experiments hint at the types of observation that might be made using specific peptide–small-molecule pairs such as the tetracysteine–helix–bi-arsenical pair created by Griffin, Adams, and Tsien [6], but they do not define the limits. The inherent versatility of this new approach to protein labeling promises to illuminate new avenues of discovery.

Acknowledgements

I thank Thomas Kodadek and Kathlynn Brown for helpful discussions and critical reading of the manuscript.

References

1. Tsien, R.Y. (1989). Fluorescent probes of cell signaling. *Ann. Rev. Neurosci.* **12**, 227-253.
2. Cubitt, A.B., Heim, R., Adams, S.R., Boyd, A.E., Gross, L.A. & Tsien, R.Y. (1995). Understanding, improving, and using green fluorescent proteins. *Trends Biochem. Sci.* **20**, 448-455.
3. Reid, B.G. & Flynn, G. (1997). Chromophore formation in green fluorescent protein. *Biochemistry* **36**, 6786-6791.

4. Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G. & Cormier, M.J. (1992). Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* **111**, 229-233.
5. Cody, C.W., Prasher, D.C., Westler, W.M., Prendergast, F.G. & Ward, W.W. (1993). Chemical structure of the hexapeptide chromophore of the *Aequorea* green fluorescent protein. *Biochemistry* **32**, 1212-1218.
6. Griffin, B.A., Adams, S.R. & Tsien, R.Y. (1998). Specific covalent labeling of recombinant protein molecules inside live cells. *Science* **281**, 269-272.
7. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. & Prasher, D. (1994). Green fluorescent protein as a marker for gene expression. *Science* **263**, 802-805.
8. Partikian, A., Olveczky, B., Swaminathan, R., Li, Y. & Verkman, A.S. (1998). Rapid diffusion of green fluorescent protein in the mitochondrial matrix. *J. Cell Biol.* **140**, 821-829.
9. Wang, S. & Hazelrigg, T. (1994). Implications for bcd mRNA localization from spatial distribution of exu protein in *Drosophila* oogenesis. *Nature* **369**, 400-403.
10. Lippincott-Schwartz, J., Cole, N. & Presley, J. (1998). Unravelling Golgi membrane traffic with green fluorescent protein chimeras. *Trends Cell Biol.* **8**, 16-20.
11. Rizzuto, R., Brini, M., Pizzo, P., Murgia, M. & Pozzan, T. (1995). Chimeric green fluorescent protein as a tool for visualizing subcellular organelles in living cells. *Curr. Biol.* **5**, 635-642.
12. Rizzuto, R., *et al.*, & Pozzan, T. (1996). Double labelling of subcellular structures with organelle-targeted GFP mutants *in vivo*. *Curr. Biol.* **6**, 183-188.
13. Mitra, D., Silva, C.M. & 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
14. Heim, R. & Tsien, R.Y. (1996). Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Biol.* **6**, 178-182.
15. Ormo, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y. & Remington, S.J. (1996). Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392-1395.
16. Heim, R., Cubitt, A.B. & Tsien, R.Y. (1995). Improved green fluorescence. *Nature* **373**, 663-664.
17. Mahajan, N.P., Linder, K., Berry, G., Gordon, G.W., Heim, R. & Herman, B. (1998). Bcl-2 and Bax interactions in mitochondria probed with green fluorescent protein and fluorescence resonance energy transfer. *Nat. Biotechnol.* **16**, 547-552.
18. Miyawaki, A., *et al.*, & Tsien R.Y. (1997). Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin. *Nature* **388**, 882-887.
19. Romoser, V.A., Hinkle, P.M. & Pereschini, A. (1997). Detection in living cells of Ca^{2+}-dependent changes in the fluorescence emission of an indicator composed of two green fluorescent protein variants linked by a calmodulin-binding sequence. *J. Biol. Chem.* **272**, 13270-13274.
20. Tse, F.W., Tse, A. & Hill, B. (1994). Cyclic Ca^{2+} changes in intracellular stores of gonadotropes during gonadotropin-releasing hormone-stimulated Ca^{2+} oscillations. *Proc. Natl Acad. Sci. USA* **91**, 9750-9754.
21. Hofer, A.M. & Schulz, I. (1996). Quantification of intraluminal free [Ca] in the agonist-sensitive internal calcium store using compartmentalized fluorescent indicators: some considerations. *Cell Calcium* **20**, 235-242.
22. Golovina, V.A. & Blaustein, M.P. (1997). Spatially and functionally distinct Ca^{2+} stores in sarcoplasmic and endoplasmic reticulum. *Science* **275**, 1643-1648.
23. Miesenbock, G., De Angelis, D.A. & Rothman, J.E. (1998). Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature* **394**, 192-195.
24. Heim, R., Prasher, D.C. & Tsien, R.Y. (1994). Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl Acad. Sci. USA* **91**, 12501-12504.