BIOCHEMISTRY

 Volume 45, Number 32 August 15, 2006

*Articles*

# Novel Chromophores and Buried Charges Control Color in mFruits[†,‡]

Xiaokun Shu,[§] Nathan C. Shaner,[||] Corinne A. Yarbrough,[§] Roger Y. Tsien,[||] and S. James Remington*[,§]

*Institute of Molecular Biology and Department of Physics, University of Oregon, Eugene, Oregon 97403, and Departments of Pharmacology and Chemistry, Howard Hughes Medical Institute, University of California at San Diego, 9500 Gilman Drive, La Jolla, California 92093*

*Received April 20, 2006; Revised Manuscript Received June 8, 2006*

ABSTRACT: mFruits are second-generation monomeric red fluorescent proteins (mRFPs) that have improved brightness and photostability compared to the first-generation mRFP1. The emission and excitation maxima are distributed over the remarkably large ranges of about 550−650 and 540−590 nm, respectively; however, the variations in the spectra can be traced to a few key amino acids. Spectroscopic and atomic resolution crystallographic analyses of three representatives, mOrange, mStrawberry, and mCherry, reveal that different mechanisms operate to establish the excitation and emission maxima. Evidently, they all undergo the second oxidation step to produce an acylimine linkage in the polypeptide backbone. In comparison to the progenitor DsRed, direct covalent modification to this linkage (mOrange) and indirect modification of the chromophore environment (mStrawberry and mCherry) produce strong blue- and red-shifted variants. The blue shift of mOrange is induced by an unprecedented covalent modification of the protein backbone. The electron-density map indicates the formation of a third heterocycle, 2-hydroxy-dihydrooxazole, upon the reaction of Thr 66 O$\gamma$ with the polypeptide backbone, which in turn reduces the conjugation of the carbonyl at position 65 with the rest of the chromophore. In mStrawberry and mCherry, the movement of charged Lys 70 and protonation of Glu 215 are proposed to modify the chromophore electron-density distribution, inducing the red shift. pH-dependent spectral shifts of mCherry and mStrawberry appear to result from the titration of Glu 215, although, for mStrawberry, partial cyclization of Thr 66 may contribute at high pH.

Fluorescent proteins isolated from coral-reef organisms (*1*, *2*) have absorption and emission spectra that span most of the visible range, yet they can be clearly divided into two classes. The cyan- (*3*) and green-emitting (*4*) members of the family have chromophores that are chemically identical to that of green fluorescent protein (GFP).[1] However, as originally established for DsRed ($\lambda_{em}^{max}$ = 583 nm) (*5*, *6*), structural studies have confirmed (*7*−*10*) that red-shifted members of the family undergo an additional multistep reaction that includes one or more covalent modification to the polypeptide backbone, as outlined in Scheme 1. In

[†] This work was supported by grants from the National Science Foundation (MCB-0417290 to S.J.R.) and NIH (NS27177 and GM72033 to R.Y.T.).

[‡] The atomic coordinates and structure factors have been deposited in the Protein Data Bank (entry 2H5O for mOrange, 2H5P for mStrawberry at pH 9.5, 2H5Q for mCherry, 2H5R for mStrawberry at pH 10.5, and 2H8Q for DsRed/K83M).

* To whom correspondence should be addressed. Telephone: (541) 346-5190. Fax: (541) 346-5870. E-mail: jremington@uoxray.uoregon.edu.

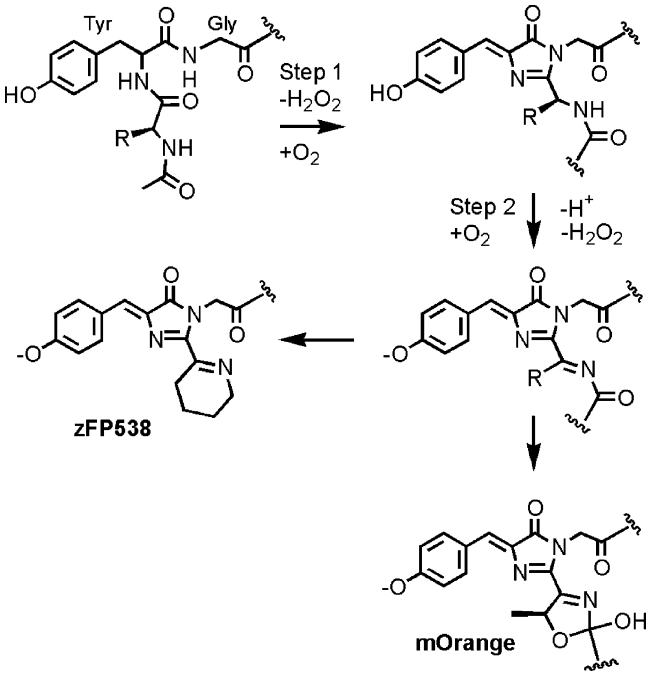
[§] University of Oregon.

[||] University of California at San Diego.

[1] Abbreviations: GFP, green fluorescent protein; mRFP, monomeric red fluorescent protein; rms, root mean square.

10.1021/bi060773l CCC: $33.50 
Published on Web 07/25/2006

Scheme 1




DsRed, the peptide bond immediately preceding the first amino acid in the chromophore triplet −X−Y−G− is oxidized to an acylimine, which delocalizes the chromophore electron density over the polypeptide backbone and increases both the excitation and emission maxima relative to GFP.

In two cases, additional covalent modifications have been proposed to follow the formation of the acylimine (Scheme 1). The yellow fluorescent protein zFP538 ($\lambda_{em}^{max}$ = 538 nm) undergoes a cyclization reaction in which the Nϵ of Lys 66 evidently attacks the acylimine α carbon to form a second heterocycle (*4*). The resulting chain cleavage produces a cyclic imine that is less effective than the acylimine in charge delocalization, resulting in the yellow emission. The hydrolysis of the acylimine linkage can also lead to polypeptide fragmentation as found in the "kindling fluorescent protein" (KFP) (*11*, *12*). The exact nature of the products of the cleavage reaction within native KFP has not been completely resolved (*13*). A model compound for the proposed chromogenic core of KFP [2-acetyl-4-(*p*-hydroxybenzylidene-1-methyl-5-imidazolone)] has been synthesized (*14*). In dimethylformamide solution, this compound has spectroscopic properties that closely match those of the native chromophore ($\lambda_{abs}^{max} \sim 570$ nm, in each case), arguing that it is an excellent model for the chromophore. As a substituent at the 2 position of the imidazolinone ring (Scheme 1), a carbonyl appears to be more effective in delocalizing the electron density than an imine (as found in zFP538), the intact acylimine linkage, or an olefin [as in model compounds for the DsRed chromophore (*15*)].

Small changes in the protein matrix can perturb absorption and emission maxima by up to ∼20 nm. For example, the substitution T203Y in the yellow fluorescent mutant 10C increases $\lambda_{em}^{max}$ from 508 to 527 nm (*16*). Likewise, the K83M substitution in DsRed increases $\lambda_{em}^{max}$ from 583 to 602 nm (*5*). The excitation maxima are also red-shifted but by a smaller amount (5 and 8 nm for 10C and K83M, respectively), implying differential interactions of the protein matrix with the chromophore ground and excited states.

Here, we report the results of crystallographic investigations on the basis of spectral variation among a collection of monomeric proteins derived from a common precursor, monomeric red fluorescent protein 1 (mRFP1) (*17*). For a comparison, the crystal structure of the DsRed mutant K83M, the progenitor for mRFP1, is briefly described. The proteins, collectively named mFruits for their monomeric nature and variety of colors, were generated by a combination of random mutagenesis and directed evolution (*18*). The evolved proteins are thus very closely related in amino acid sequence. The excitation and emission maxima are grouped in the range of 540−590 and 550−610 nm, respectively. We chose three representative members: mCherry ($\lambda_{exc}^{max}$ = 587 nm, $\lambda_{em}^{max}$ = 610 nm), mStrawberry ($\lambda_{exc}^{max}$ = 574 nm, $\lambda_{em}^{max}$ = 596 nm), and mOrange ($\lambda_{exc}^{max}$ = 548 nm, $\lambda_{em}^{max}$ = 562 nm) for detailed study. An unprecedented mechanism was found to produce the striking blue shifts associated with mOrange and possibly with mStrawberry at high pH.

## MATERIALS AND METHODS

*Mutagenesis, Protein Expression, and Crystallization.* For an amino acid sequence alignment of relevant mFruits, see Figure S1 in the Supporting Information. The protein was expressed in *Escherichia coli* (Top10 or JM-109 for DsRed K83M variant) by use of the modified and original pBAD B His-tagged expression system (*18*). Mutagenesis was performed using the QuikChange method (Stratagene). The protein was purified by $Ni^{+2}$-affinity chromatography over Ni−NTA agarose (Qiagen, Chatsworth, CA) and then buffer-exchanged with PD-10 Sephadex columns (Amersham Pharmacia) into 50 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES) (pH 7.9). Crystals of mOrange were obtained by hanging drop vapor diffusion: 2 μL of the protein solution (50 mM HEPES at pH 7.9, $A_{280}$ = 20.3) with 2 μL of the well solution [200 mM $MgCl_2$, 100 mM Tris at pH 8.2, and 28% poly(ethylene glycol) (PEG) 1550] after 1 week. mStrawberry was crystallized in 1.1 M NaCitrate, 75 mM Glycine at pH 10.5, or 100 mM 3-(cyclohexylamino)-1-propanesulfonic acid (CAPS) at pH 11.5 after 1−2 months, and mCherry was crystallized in 100 mM NaOAc, 100 mM Tris at pH 8.5, and 30% PEG 4000 overnight. Crystals of DsRed/K83M were also obtained by hanging drop vapor diffusion: 2 μL of the protein solution (300 mM NaCl for DsRed/K83M, $A_{280}$ = 15.0) with 2 μL of the well solution (100 mM HEPES at pH 6.5 and 45% MPD) after 2 weeks. All crystals were fluorescent under UV light.

*Spectroscopy.* Fluorescence spectra were taken with a Perkin−Elmer LS-55 fluorometer. Protein samples were diluted to 44 μg/mL in 2 mL of 100 mM HEPES at pH 7.9 or 100 mM CAPS at pH 11.4. Absorbance spectra (400 μg/mL protein in the same buffers) were recorded with a HP 8453 UV−vis spectrophotometer.

*Data Collection and Structure Solution.* Diffraction data sets were collected from single flash-frozen crystals of mOrange and mStrawberry (100 K) using the ADSC-Q315 detector on beam line 8.2.2 at the Advanced Light Source (Berkeley, CA). Diffraction data for mCherry and DsRed/K83M were collected on Q315 and Q210 detectors on beam lines 14-BM-C and IMCA-CAT at the Advanced Photon Source (Argonne, IL), respectively. Data were reduced using

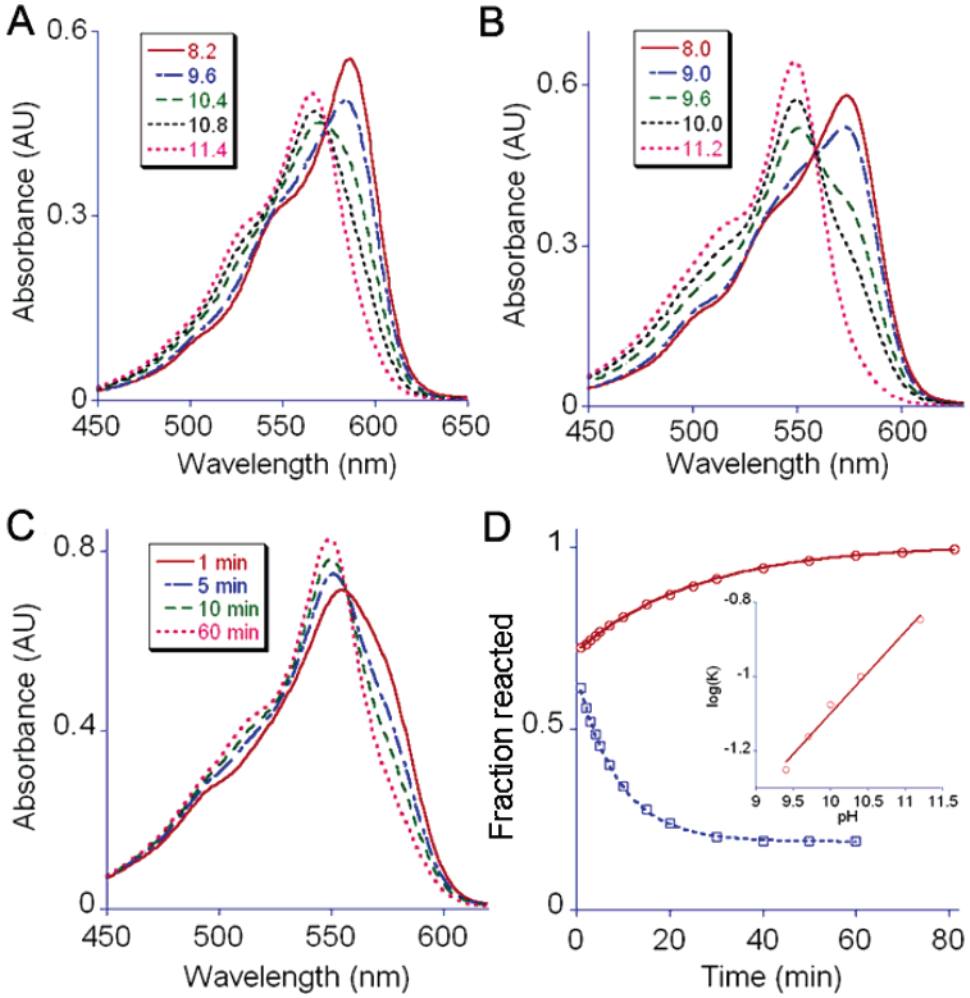



FIGURE 1: Time- and pH-dependent spectral shifts in mCherry and mStrawberry. Absorbance spectra of (A) mCherry and (B) mStrawberry at pH values shown in the inset. (C) Time-dependent absorbance spectra for mStrawberry upon the step increase from pH 7.9 to 10.5. (D) Time course showing reversible pH-dependent spectral shifts in mStrawberry. Blue squares indicate the fraction reacted upon the step increase in pH from 7.9 to 10.5, and red circles indicate the fraction reacted upon the step decrease in pH from 10.5 to 7.9. (Inset) Linear fit of log(*K*) versus final pH. See the Results for details of analysis.

HKL2000 (HKL Research). Molecular replacement was performed with EPMR (*19*) using the DsRed A chain (PDB entry 1G7K) as a search model, providing in all cases unambiguous identification of the space group and unique solutions to the rotation and translation problems. Rigid-body refinement was performed using TNT (*20*), and model building was conducted with O (*21*) in several stages of increasing resolution. The chromophore entries for the TNT geometry library were derived using AM4 as described (*7*). Shelx-97 (*22*) was used in the final stages of refinement for mFruits. Because of the ambiguity of the chemical linkage, residues 66 and 65 were unrestrained from the beginning of the refinement with Shelx-97.

The chromophore of mOrange was first modeled using a modified peptide bond (*7*) for the 65−66 linkage and the standard side chain for threonine at position 66. However, very large peaks appeared in ($F_o - F_c$) difference electron-density maps, indicating that this model was inappropriate. Geometric restraints were substantially loosened for these atoms, and upon refinement in Shelx-97 with Thr 66 and Phe 65 unrestrained, the resulting geometry resembled that of a five-membered ring (2-hydroxy-dihydrooxazole, Scheme 1) comprised of the carbonyl carbon of Phe 65 and the main-chain N, C$\alpha$, C$\beta$, and O$\gamma$ of Thr 66.

## RESULTS

*Absorbance and Emission Spectra.* Above pH 9, the absorbance and emission spectra of mCherry and mStrawberry are strongly pH-dependent, exhibiting blue shifts with increasing pH (parts A and B of Figure 1, respectively). Upon the increase in pH from 7.9 to 11.4, the absorbance maximum for mCherry is blue-shifted from 584 to 566 nm, while for mStrawberry, the shift is from 574 to 548 nm. Spectra at intermediate pH could be adequately decomposed using the endpoint spectra at pH 7.9 and 11.4, suggesting a two-state transition. The results could be fit to a single-site titration curve, with a calculated p$K_a$ of 10.29 ± 0.04 for mCherry and 9.75 ± 0.04 for mStrawberry. Similar behavior was noted for excitation and emission spectra, and in fact, the absorbance and excitation spectra are coincident. The emission maximum of mCherry blue shifts from 609 nm at pH 7.9 to 594 nm at pH 11.4, while the blue shift for mStrawberry emission is more substantial, from 596 to 565 nm.

Protein samples of mStrawberry were subjected to a step change in pH from 7.9 to 9.4 or higher, and the time dependence of the absorbance spectra (Figure 1C) suggested a very slow transition. Spectra at intermediate time points could be decomposed on the basis of the endpoint spectra collected

Table 1: Data Collection and Refinement Statistics

| | dsRED/K83M | mOrange | mStrawberry (pH 9.5) | mStrawberry (pH 10.5) | mCherry |
|---|---|---|---|---|---|
| data collection | | | | | |
| space group | $P6_522$ | $P2_12_12_1$ | $P2_1$ | $P2_1$ | $P2_1$ |
| cell dimensions | | | | | |
| *a*, *b*, *c* (Å) | 93.1,93.1,430.5 | 59.3, 67.4, 107.8 | 48.6, 43.4, 61.4 | 48.7, 44.2, 61.2 | 48.8, 42.9, 61.1 |
| $\beta$ (deg) | | | 112.3 | 106.0 | 112.31 |
| resolution (Å) | 30−2.25 (2.37−2.25)[a] | 50−1.08 (1.12−1.08) | 50−1.21 (1.25−1.21) | 50−1.60 (1.66−1.60) | 50−1.36 (1.41−1.36) |
| $R_{sym}$ or $R_{merge}$ (%) | 8.2 (45.0) | 5.5 (64.3) | 4.1 (29.3) | 7.6 (28.6) | 4.6 (32.8) |
| $I/\sigma(I)$ | 20.7 (2.4) | 38.4 (1.6) | 35.8 (2.6) | 19.7 (2.8) | 38.1 (5.1) |
| completeness (%) | 86.2 (84.3) | 99.5 (96.3) | 97.2 (88.2) | 93.3 (64.0) | 98.7 (93.8) |
| redundancy | NA[b] | 6.6 (4.2) | 3.3 (2.4) | 1.8 (1.5) | 3.5 (3.4) |
| refinement | | | | | |
| resolution (Å) | 5.0−2.25 | 10−1.08 | 10−1.21 | 10−1.60 | 10−1.36 |
| number of reflections | 61 886 | 170 025 | 70 282 | 30 663 | 47 050 |
| $R_{work}/R_{free}$ | 23.0/26.8 | 14.5/17.2 | 13.8/17.8 | 16.6/21.7[c] | 14.7/19.1 |
| number of atoms | | | | | |
| protein | 7044 | 3854 | 1826 | 1787 | 1753 |
| solvent | 580 | 550 | 247 | 283 | 225 |
| *B* factors (Å$^2$) | | | | | |
| protein | 45.57 | 17.50 | 19.42 | 22.50 | 15.63 |
| water | 57.01 | 31.50 | 33.25 | 37.60 | 24.61 |
| rms deviations | | | | | |
| bond lengths (Å) | 0.019 | 0.025 | 0.015 | 0.016 | 0.011 |
| bond angles (Å) | 2.9° | 0.040 | 0.034 | 2.3° | 0.029 |

[a] The values in parentheses indicate statistics for the highest resolution shell. [b] NA = not available. [c] $R_{free}$ is estimated by simulated annealing.

at pH 7.9 and 11.4. The fraction $X(t)$, where $X$ is the fraction of the pH 7.9 baseline (blue squares in Figure 1D), could be fit to a single exponential (blue dashed line) and a rate constant extracted. The log of rate constants was linearly dependent upon the final pH (red line in the inset of Figure 1D); however, a linear fit of the rate constants to the final pH is also satisfactory. The extracted rate constants range from 0.056 min$^{-1}$ at pH 9.4 to 0.142 min$^{-1}$ at pH 11.2. The spectral shifts were fully reversible upon reduction of pH.

*Atomic Models and Chromophore Environment.* Data collection and refinement statistics for the atomic models are presented in Table 1. The crystal structure of DsRed/K83M was determined at moderate resolution, yet it contained sufficient information to serve as a useful reference point with which to compare the high-resolution models of the mFruits. The asymmetric unit for the crystal of DsRed/K83M contains two half-tetramers; however, crystal symmetry operators generate tetramers that are essentially identical to that described for wild-type DsRed (*7*). The mutation Lys83 → Met results in minor side-chain rearrangements in the immediate vicinity of the chromophore. The red shifts in absorption and emission can be rationalized in terms of these rearrangements as described below.

In contrast to DsRed, the crystals of mFruits provided no evidence for oligomer formation. No electron density is apparent for residues 1−5 of mOrange and mCherry or for residue 1 of mStrawberry, indicating disorder. Density at the C terminus is also weak for mOrange 222−236, mCherry 228−236, and mStrawberry 231−236. The electron density for loop residues 168−170 of chain A in mOrange suggests multiple backbone conformations. PROCHECK (*23*) reveals that there are no residues in either the disallowed or "generously allowed" regions of the Ramachandran diagram. Root-mean-square (rms) deviations of 0.35, 0.40, and 0.46 Å are observed for mCherry, mStrawberry, and mOrange upon superposition of α carbons with the A subunit of the parent protein DsRed, respectively. The models of mCherry, mStrawberry, and mOrange are thus remarkably similar to that of DsRed, which is somewhat surprising considering the dramatic change in quaternary structure: from tetrameric to monomeric.

In mOrange, maturation is incomplete, which complicates the crystallographic analysis. Upon excitation at 470 nm, an emission peak from a presumably GFP-like chromophore at 510 nm is readily detectable (data not shown) and the electron-density maps are consistent with partial occupancy of a GFP-like chromophore. Both the structural results and spectra suggest that the percentage of immature mOrange in the samples is less than 20%. In mCherry and mStrawberry, maturation to the acylimine is essentially complete as Met 66 Cα and Thr 66 Cα clearly adopt sp$^2$ geometry. The peptide bond between residues 65 and 66 is in the cis conformation and oxidized to an acylimine as in DsRed (*7*).

Schematic diagrams (parts A−C of Figure 2) of chromophore environments found in mCherry, mStrawberry, and mOrange show that the environments are, in each case, substantially more hydrophobic than in DsRed [compare to Figure 3C of Yarbrough et al. (*7*)] because of the replacement of the internal polar side chains Lys 83 and Lys 163 with bulky nonpolar groups. The substitutions K163Q (mCherry) or K163M (mStrawberry and mOrange) lead to a loss of hydrogen-bonding opportunities for the phenolate oxygen of the chromophore and presumably lead to a net shift of electron density toward the imidazolinone moiety. Common to all of these structures is the mutation of Lys 83 → hydrophobic (K83L in mCherry and mStrawberry; K83F in mOrange). In all cases, the primary effect of this substitution is that the terminal Nϵ of conserved Lys 70 is observed to move away (by about 2.7 Å) from its position over the methylene bridge of the chromophore in DsRed (parts A and B of Figure 3). Its new position is stabilized through the interaction with Glu 148. The 2.25 Å resolution crystal structure of the DsRed variant K83M shows similar movement of Lys 70 (Figure 3A).

All fluorescent proteins contain the internal conserved side chain Glu 215, which in mFruits adopts different conforma-

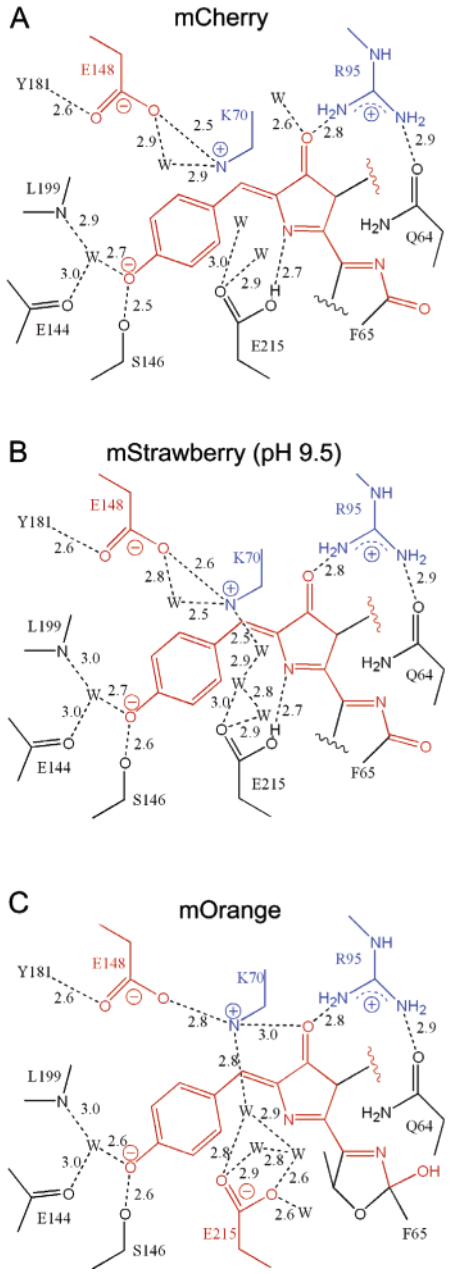



FIGURE 2: Schematic diagram of chromophore environment of mFruits. (A) mCherry, (B) mStrawberry at pH 9.5, and (C) mOrange. Hydrogen bonds are shown in dashed lines, labeled with approximate lengths in angstroms.

tions that seem to be related to its protonation state. In mCherry and mStrawberry, Glu 215 appears to be protonated at pH < 10, while in mOrange, Glu 215 is probably deprotonated. As evidence for the protonation state and in contrast to DsRed, in mCherry, Glu 215 forms a hydrogen bond (2.7 Å) with the imidazolinone ring nitrogen (N2, Figure 3B). In mStrawberry at pH 9.5, the electron-density map suggests disorder of Glu 215 with two conformations: in conformation A, it is hydrogen-bonded to the N2 (2.7 Å) as in mCherry, while in conformation B, it is further away from N2 (3.7 Å), as in mOrange (data not shown). In mStrawberry at pH 10.5, Glu 215 has only one conformation (4.0 Å from N2) and appears to be hydrated by several water molecules. We take this observation as evidence that the chromophore N2 is uncharged, as is commonly assumed.

*Unusual Covalent Chromophore Modifications.* The ($F_o - F_c$) omit map of mOrange, contoured at the 8$\sigma$ level (green in Figure 4A), reveals unprecedented covalent modifications to the protein backbone. The side-chain O$\gamma$ of Thr 66 forms a covalent bond with the carbonyl carbon of Phe 65 to form 2-hydroxy-dihydrooxazole (Scheme 1 and Figure 4A). However, after automated refinement, the apparent C−O bond length is unusually long, ~1.55 Å averaged over chains A and B. We interpret this apparently long bond to be a consequence of partial maturation and the presence in the crystal of a statistical mixture of reacted and unreacted chromophore species (see the Discussion).

In mStrawberry at pH 9.5, Thr 66 O$\gamma$ may also react with the carbonyl carbon of Phe 65 although to a lower extent. The distance between these atoms refined to be 2.8 Å, which is approximately the sum of van der Waals radii. However, at pH 10.5, this distance is refined to be 2.5 Å, which, as in the case of mOrange, we interpret to be a superposition of partially occupied structures containing either a normal covalent C−O bond (1.43 Å) or a nonbonded contact. The ($2F_o - F_c$) electron-density map (Figure 4B) can be adequately modeled using these two conformations by setting the occupancy as 80% for the model containing the covalent bond and 20% for the noncovalent interaction. The *B* values were restrained to be 15 Å$^2$, similar to those of nearby atoms. After 10 cycles of TNT refinement, the distance for the nonbonded C−O contact, which is not restrained, refined to be 3.1 Å.

*Nonplanar Chromophore Geometry.* The chromophores of all mFruits show significant deviations from planarity, which is adequately modeled by allowing the two torsion angles associated with the linkage between the five- and six-membered rings to deviate from 180° (*12*) (Figure 4C). The twist and tilt angles describing the distortion are significantly larger than those of the fluorescent proteins of the highest quantum yields such as DsRed and GFP (*12*) (Table 2) but less than those of the nonfluorescent chromoproteins KFP (*12*) and Rtms5 (*10*).

## DISCUSSION

In fluorescent proteins, changes within the chromophore cavity have been observed to cause shifts in excitation and emission maxima by up to 20 nm. Baird et al. (*5*) drew attention to the importance of the K83M mutation in DsRed, which increases the emission (excitation) maximum from 583 (558) to 602 (564) nm. Structural studies of DsRed/K83M at 2.25 Å resolution suggested that the primary effect of the mutation was a shift in the position of the charged Lys 70 N$\epsilon$ away from the chromophore. Emission shifts of similar magnitude were observed in GFP resulting from the substitution T203Y. The introduced aromatic side chain stacks against the chromophore, producing the popular yellow fluorescent protein YFP (*16*). In each case, the substitutions are presumed to alter the electron-density distribution of the chromophore in such a way as to increase both emission and excitation wavelengths, although to different extents.

On the other hand, excitation and emission maxima exhibited by mFruits span a remarkably large range (540−590 nm for excitation and 550−650 nm for emission), suggesting that more than one effect, including additional covalent modifications to the chromophore, may be operative.

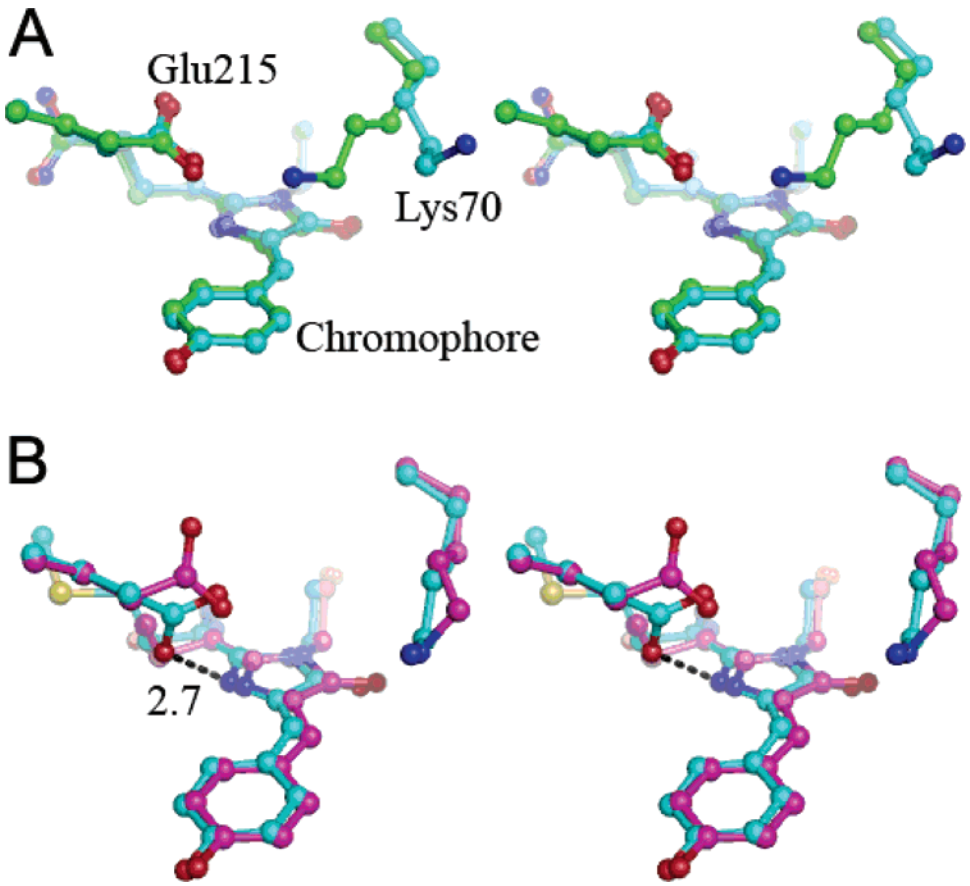



FIGURE 3: Chromophore environment detailing the configuration of conserved residues Glu 215 and Lys 70. (A) DsRed/K83M (cyan) is compared with DsRed (green). (B) Disposition of Glu 215 and Lys 70 in mCherry (cyan) and mOrange (magenta). A presumed hydrogen bond is indicated by the dashed line, with the indicated length in angstroms.

In mFruits, the crystal structures and spectral properties unequivocally establish that both environmental effects and covalent chromophore modifications are required to explain the observed excitation and emission spectra.

*Structural Basis for Red Shifts of mCherry/mStrawberry.* Excitation and emission in mCherry and mStrawberry are substantially red-shifted with respect to the progenitor DsRed. Two linked changes in the chromophore environment appear to be responsible. First, all mFruits contain a substitution of the form K83 → large nonpolar side chain. mCherry and mStrawberry both contain the K83L substitution, and the emission maxima 609 and 596 nm are close to that (602 nm) of DsRed/K83M. A comparison of the location of Lys 70 in both structures shows that they all reflect the ~2.7 Å shift in position observed for DsRed/K83M (Figure 3A). Thus, it is reasonable to assume that this positional shift could be due to the loss of an electrostatic interaction with residue 83 resulting from the K83L mutation and that this accounts for some portion of the red shift. Attempts to test this proposition by reversion to the original Lys 83 failed, because the substitution appears to be required for proper folding of mFruits (data not shown). In both structures, Lys 70 interacts with Glu 148, which, in the absence of the former AC dimer interface found in the parent tetramer, may be required for proper folding (*17*).

A second effect may also be linked to protonation of Glu 215. In contrast to DsRed, the structure of mCherry reveals Glu 215 to be hydrogen-bonded to the imidazolinone ring nitrogen, suggesting that Glu 215 is protonated. The structure of mStrawberry at pH 9.5, which is near the transitional p$K_a$ ~ 9.8, suggests that one of the two conformations of Glu 215 is similar to that in mCherry. The proposition that protonation of Glu 215 may be linked to the red shifts is supported by the following analysis of pH-dependent spectral shifts in both proteins.

*Novel Five-Membered Ring and Blue Shifts in mOrange.* The observation of a novel chromophore modification, cyclization of Thr 66 in mOrange to form an oxazole ring, explains the large blue shift in emission relative to the progenitor, DsRed. This modification reduces the order of the main chain C═O from two to one, thus eliminating the conjugation between the carbonyl and the chromophore. This observation provides strong experimental evidence that, in red fluorescent proteins, the carbonyl of residue 65 is indeed conjugated with the chromophore despite the fact that it is almost perpendicular to the chromophore plane, a result consistent with previous theoretical calculations (*6*). A reduction in the extent of charge delocalization over the polypeptide backbone reduces the effective size of the chromophore, thus reducing excitation and emission wavelengths. This proposal is consistent with previous explanations for the excitation and emission spectrum of yellow zFP538, in which a similar tricyclic chromophore, formed by cyclization of Lys 66, was observed (*4*).

However, in the refined model of mOrange, the carbonyl carbon−O$\gamma$ bond appears to be unusually long (1.55 Å). We considered the possibility that this bond is a partial bond, as originally proposed in the classic work by Burgi and Dunitz

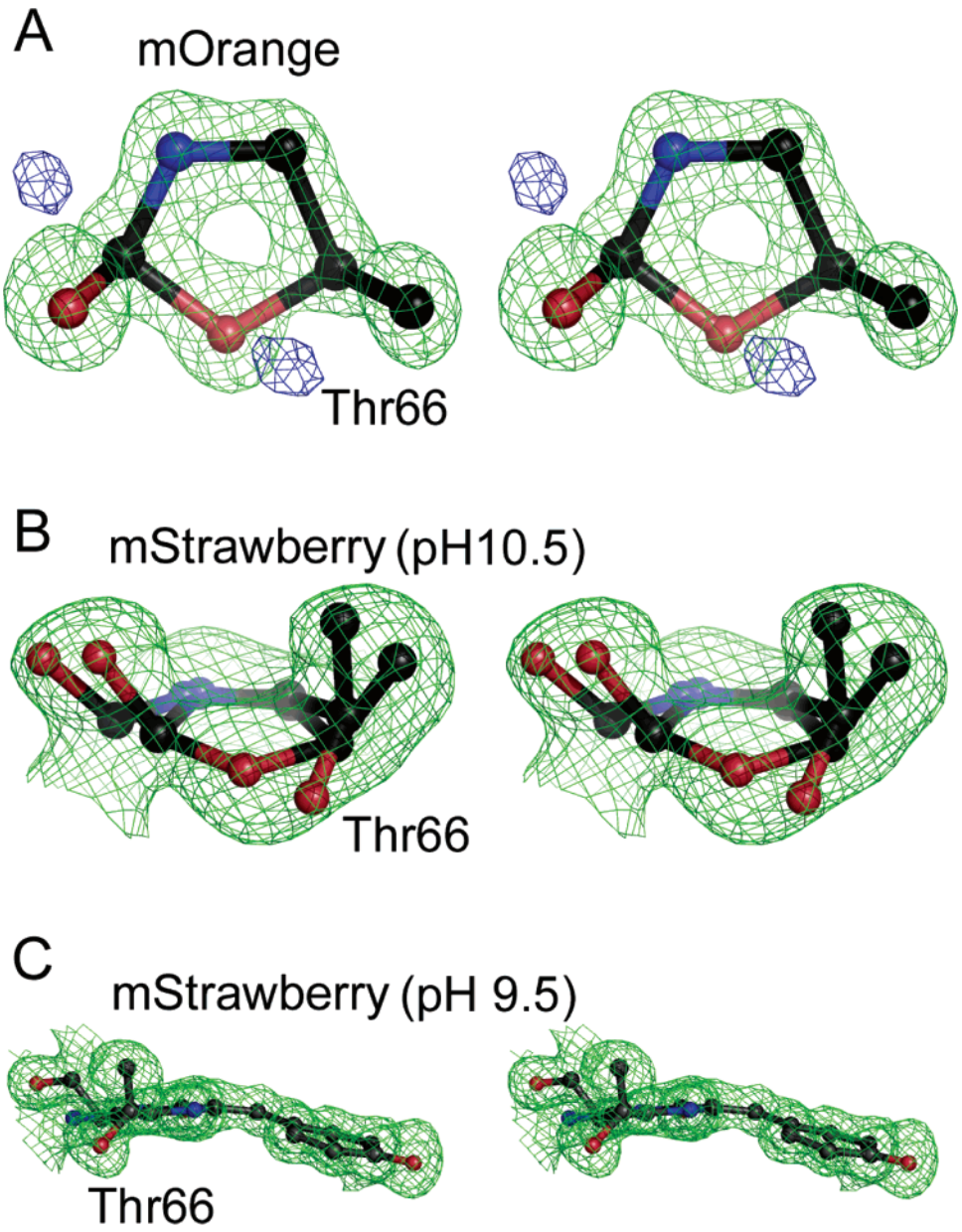



FIGURE 4: Electron-density maps in stereo showing chromophore details in mOrange and mStrawberry. (A) Omit (green contours, 8$\sigma$) and final ($F_o - F_c$) difference electron-density maps (blue contours, 4$\sigma$) in the vicinity of Thr 66 in mOrange. See the text for discussion. (B and C) ($2F_o - F_c$) electron-density map of mStrawberry at pH 10.5, with the contour level at 1$\sigma$. The final refined models of the cyclized and open forms of Thr 66 are superimposed. (C) Chromophore nonplanarity revealed for mStrawberry at pH 9.5.

Table 2: Summary of the Quantum Yield and Deviations from Chromophore Planarity in mFruits[a]

| model | QY[b] | twist[c] | tilt[c] |
|---|---|---|---|
| wt-GFP | 0.79 | 0.0 | 3.7 |
| DsRed | 0.79 | 0.9 | 0.2 |
| mRFP | 0.25 | NA[d] | NA[d] |
| mOrange | 0.69 | 1.5 | 10.3 |
| mStrawberry (pH 9.5) | 0.31 | 5.9 | 17.9 |
| mStrawberry (pH 10.5) | 0.36 | 0.4 | 4.6 |
| mCherry | 0.22 | 11.3 | 13.7 |

[a] For the definition of twist and tilt angles, see ref *12*. [b] Quantum yield data are from ref *18*. [c] Angles are in degrees. [d] NA = not available.

(*24*, *25*). These authors discussed several small-molecule crystal structures in which N···C=O interactions are observed within the range of 1.5−2.0 Å. Such interactions were characterized as partial bonds, arrested in an intermediate stage of nucleophilic N−C addition by ring strain or other geometric effects. The compounds methadone, clivorine, and protopine are often cited as examples of this principle (*26*). It has been proposed that similar O···C=O interactions might be observed; however, no evidence for O···C contacts in the range of 1.5−2.5 Å has been reported, which these authors attributed to the reduced nucleophilicity of oxygen compared to nitrogen.

A search of the Cambridge Structural Database [CSD (*27*)] revealed that compounds containing oxazole rings are rare. Among the dozen or so examples discovered, none were found to contain an abnormal C−O bond. Indeed, no examples of an O···C=O approach in the range of 1.5−2.5 Å were found among the ~335 000 entries in the CSD, substantiating the 1974 study (*24*).

In a test refinement with no geometric restraints on the atoms of Phe 65 and the chromophore, a ($F_o - F_c$) difference

map (Figure 4A) exhibits positive peaks exceeding 4$\sigma$ near the carbonyl of Phe 65 and side chain of Thr 66. This indicates that no single atom model can satisfactorily account for the electron density. This result is consistent with statistical contamination of mature mOrange by a small fraction of immature mOrange containing a GFP-like chromophore as recently described for DsRed/K70M (*28*). Model-building experiments suggest that the electron density can be explained by such a mixture; however, automated refinement of mixed models failed to converge, most likely because of the low percentage of the immature form. Nevertheless, both the electron-density maps and the spectroscopic results are consistent with a superposition of structures containing a normal covalent bond (of length approximately 1.4 Å) and a nonbonded C−O contact.

We therefore concluded that the 1.55 Å C−O bond length observed in this study is an artifact of incomplete maturation.

*Structural Basis for pH-Dependent Spectral Shifts of mCherry/mStrawberry.* The absorbance and emission spectra of both mCherry and mStrawberry exhibit pH-dependent blue shifts as the pH is increased. In the case of mStrawberry, the crystal structures at pH 9.5 and 10.5 show the movement of Glu 215, suggesting that it is the titrating group. In the case of mCherry, blue shifts in excitation and emission maxima make the spectra of the protein at pH 11.4 similar to those of DsRed/K83M. Again, a structural comparison provides evidence that, in mCherry, Glu 215 is the titrating group. Because electrostatic effects are long-range and additive, the change in charge distribution near the chromophore because of deprotonation of Glu 215 may act to reduce the effect of the red shifts that we propose are associated with the movement of Lys 70.

This analysis is complicated by the structural evidence that, in mStrawberry at pH 10.5, Thr 66 may slowly react to form an oxazole ring as found in mOrange. Upon a step increase in pH, the observed rate constants for the spectral transition are highly correlated with the final pH, consistent with deprotonation of Thr 66 O$\gamma$ being a prerequisite for the attack on the carbonyl carbon of residue 65.

*Quantum Yield and Chromophore Planarity.* Naturally occurring fluorescent proteins such as DsRed and GFP have quantum yields approaching 1.0, and the five- and six-membered rings in their chromophores are found to be highly coplanar. However, in the mFruits described in this study, the chromophores are all bent or otherwise distorted from coplanarity (Table 2). In mFruits, the distortion may be adequately modeled by permitting the two torsional angles defined by the methylene linkage connecting the five- and six-membered rings to deviate from 180° (see ref *12* for a definition). Similar but larger distortions have been invoked to explain the lack of fluorescence in the chromoproteins KFP (*12*) and Rtsm5 (*10*). We suspect that this may be a general phenomenon and propose that the reduction in the quantum yield of mFruits compared to the progenitor DsRed is associated with the increased chromophore noncoplanarity. Several trends support this argument. At pH 10.5, mStrawberry has a higher quantum yield (0.36) and correspondingly better coplanarity than at pH 9.5. Likewise, the higher quantum yield of mOrange is consistent with a chromophore that is more planar than those found in mStrawberry and mCherry at pH $< 10$.

In conclusion, these studies confirm that, in fluorescent proteins, simple rearrangements in the chromophore cavity, in this case, the position of charged groups, can lead to significant but limited spectral perturbations. Larger perturbations, such as the blue shift of mOrange relative to DsRed, will most likely require covalent modifications to the chromophore itself. Evidently, the reactivity of the acylimine linkage can lead to novel covalent rearrangements; however, the structural features that control both the formation of this linkage and its subsequent reactivity remain to be determined.

## ACKNOWLEDGMENT

We thank Prof. Kim Baldridge of the University of Zurich for helpful discussions and Karen Kallio, Alaine Garrett, and Leslie Colip for help in protein crystallization. N. C. S. was a Howard Hughes Predoctoral Fellow.

## SUPPORTING INFORMATION AVAILABLE

Figure S1, amino acid sequence alignment of DsRed, mCherry, mStrawberry, and mOrange. This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES

1. Matz, M. V., Arkady, F. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L., and Lukyanov, S. A. (1999) Fluorescent proteins from nonbioluminescent *Anthozoa* species, *Nat. Biotechnol. 17*, 969−973.
2. Verkhusha, V. V., and Lukyanov, K. A. (2004) The molecular properties and applications of *Anthozoa* fluorescent proteins and chromoproteins, *Nat. Biotechnol. 22*, 289−196.
3. Henderson, J. N., and Remington, S. J. (2005) Crystal structures and mutational analysis of amFP486, a cyan fluorescent protein from *Anemonia majano*, *Proc. Natl. Acad. Sci. U.S.A. 102*, 12712−12717.
4. Remington, S. J., Wachter, R. M., Yarbrough, D. K., Branchaud, B., Anderson, D. C., Kallio, K., and Lukyanov, K. A. (2005) zFP538, a yellow fluorescent protein from *Zoanthus*, contains a novel three-ring chromophore, *Biochemistry 44*, 202−212.
5. 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. U.S.A. 97*, 11984−11989.
6. 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. U.S.A. 87*, 11990−11995.
7. Yarbrough, D., Wachter, R. M., Kallio, K., Matz, M. V., and Remington, S. J. (2001) Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0 Å resolution, *Proc. Natl. Acad. Sci. U.S.A. 98*, 462−467.
8. Wall, M. A., Socolich, M., and Ranganathan, R. (2000) The structural basis for red fluorescence in the tetrameric GFP homolog DsRed, *Nat. Struct. Biol. 7*, 1133−1138.
9. Petersen, J., Wilmann, P. G., Beddoe, T., Oakley, A. J., Devenish, R. J., Prescott, M., and Rossjohn, J. (2003) The 2.0 Å crystal structure of eqFP611, a far-red fluorescent protein from the sea anemone *Entacmaea quadricolor*, *J. Biol. Chem. 278*, 44626−44631.
10. Prescott, M., Ling, M., Beddoe, T., Oakley, A. J., Hoegh-Guldberg, O., Devenish, R. J., and Rossjohn, J. (2003) The 2.2 Å crystal structure of a pocilloporin pigment reveals a nonplanar chromophore conformation, *Structure 11*, 275−284.
11. Wilmann, P. G., Petersen, J., Dvenish, R. J., Prescott, M., and Rossjohn, J. (2005) A polypeptide fragmentation within the chromophore revealed in the 2.1 Å crystal structure of a nonfluorescent chromoprotein from *Anemonia sulcata*, *J. Biol. Chem. 280*, 2401−2404.
12. Quillin, M. L., Anstrom, D. M., Shu, X., O'Leary, S., Kallio, K., Lukyanov, K. A., and Remington, S. J. (2005) The kindling

fluorescent protein from *Anemonia sulcata*: Dark state structure at 1.38 Å resolution, *Biochemistry 44*, 5774−5787.
13. Zagranichny, V. E., Rudenko, N. V., Gorokhovatsky, A. Y., Zakharov, M. V., Balashova, T. A., and Arseniev, A. S. (2004) Traditional GFP-type cyclization and unexpected fragmentation site in a purple chromoprotein from *Anemonia sulcata*, asFP595, *Biochemistry 43*, 13598−13603.
14. Yampolsky, I. V., Remington, S. J., Martynov, V. I., Potapov, V. K., Lukyanov, S., and Lukyanov, K. A. (2005) Synthesis and properties of the chromophore of asFP595 chromoprotein from *Anemonia sulcata*, *Biochemistry 44*, 5788−5793.
15. He, X., Bell, A. F., and Tonge, P. J. (2002) Synthesis and spectroscopic studies of model red fluorescent protein chromophores, *Org. Lett. 9*, 1523−1526.
16. 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.
17. 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. U.S.A. 99*, 7877−7882.
18. 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.
19. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) Rapid automated molecular replacement by evolutionary search, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 55*, 484−491.
20. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) An efficient general-purpose least-squares refinement program for macromolecular structures, *Acta Crystallogr., Sect. A: Found. Crystallogr. 43*, 489−503.
21. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr., Sect. A: Found. Crystallogr. 47*, 110−119.
22. Sheldrick, G. M., and Schneider, T. R. (1997) SHELXL: High-resolution refinement, *Methods Enzymol. 277*, 319−343.
23. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
24. Burgi, H. B., Dunitz, J. D., Lehn, J. M., and Wipff, G. (1974) Stereochemistry of reaction paths at carbonyl centres, *Tetrahedron 30*, 1563−1572.
25. Dunitz, J. D. (1979) *X-ray Analysis and the Structure of Organic Molecules*, Cornell University Press, Ithaca, NY.
26. Cieplak, A. S. (1994) Organic addition and elimination reactions; transformation paths of carbonyl derivatives, in *Structure Correlation* (Burgi, H. B., and Dunitz, J. D., Eds.) pp 205−302, VCH, Weinheim, Germany.
27. Allen, F. H. (2002) The Cambridge Structural Database: A quarter of a million crystal structures and rising, *Acta Crystallogr., Sect. B: Struct. Sci. 58*, 380−388.
28. Tubbs, J. L., Tainer, J. A., and Getzoff, E. D. (2005) Crystallographic structures of *Discosoma* red fluorescent protein with immature and mature chromophores: Linking peptide bond trans−cis isomerization and acylimine formation in chromophore maturation, *Biochemistry 44*, 9833−9840.

BI060773L