# Kindling Fluorescent Protein from *Anemonia sulcata*: Dark-State Structure at 1.38 Å Resolution[†,‡]

Michael L. Quillin,[§] David M. Anstrom,[§] Xiaokun Shu,[§] Shannon O'Leary,[§] Karen Kallio,[§] Dmitry M. Chudakov,[∥] and S. James Remington*[,§]

*Departments of Physics, Chemistry, and Biology and Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403, and Shemiakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Miklukho-Maklaya 16/10, 117997 Moscow, Russia*

*Received November 5, 2004; Revised Manuscript Received January 13, 2005*

ABSTRACT: When the nonfluorescent chromoprotein asFP595 from *Anemonia sulcata* is subjected to sufficiently intense illumination near the absorbance maximum ($\lambda_{abs}^{max}$ = 568 nm), it undergoes a remarkable transition, termed "kindling", to a long-lived fluorescent state ($\lambda_{em}^{max}$ = 595 nm). In the dark recovery phase, the kindled state relaxes thermally on a time scale of seconds or can instantly be reverted upon illumination at 450 nm. The kindling phenomenon is enhanced by the Ala143 → Gly point mutation, which slows the dark recovery time constant to 100 s at room temperature and increases the fluorescence quantum yield. To investigate the chemical nature of the chromophore and the possible role of chromophore isomerization in the kindling phenomenon, we determined the crystal structure of the "kindling fluorescent protein" asFP595-A143G (KFP) in the dark-adapted state at 1.38 Å resolution and 100 K. The chromophore, derived from the Met63-Tyr64-Gly65 tripeptide, closely resembles that of the nonfluorescent chromoprotein Rtms5 in that the configuration is trans about the methylene bridge and there is substantial distortion from planarity. Unlike in Rtms5, in the native protein the polypeptide backbone is cleaved between Cys62 and Met63. The size and shape of the chromophore pocket suggest that the cis isomer of the chromophore could also be accommodated. Within the pocket, partially disordered His197 displays two conformations, which may constitute a binary switch that stabilizes different chromophore configurations. The energy barrier for thermal relaxation was found by Arrhenius plot analysis to be ∼71 kJ/mol, somewhat higher than the value of ∼55 kJ/mol observed for cis−trans isomerization of a model chromophore in solution.

The use of green fluorescent protein (GFP) from *Aequorea victoria* and its red-shifted cousins as passive labels for protein localization and gene expression has revolutionized cell biology. The key properties are the autocatalytic formation of the chromophore, high molecular stability, and efficient, apparently steady-state fluorescence. However, when examined at extreme dilution, these proteins display complex photophysical behavior that belies such simple appearances. Aside from photobleaching, single-molecule techniques (among other methods) have revealed emission intensity variations, on−off blinking, and photoreversible switching between light and dark states at time scales ranging from microseconds to hours (*1*−*7*). This behavior has been attributed to diverse processes such as dynamic fluctuations in the protein matrix (*1*, *8*), transient changes in chromophore protonation states (*2*−*4*, *6*), and different thermally equilibrated or photoisomerizable ground states (*7*, *9*, *10*). For a recent review, see ref *11*.

The reversible transitions between fluorescent and dark states are of broad interest, as these properties may eventually find use in single-bit information storage in macromolecules or as active labels in biological systems. The photophysical basis for such behavior is at present unknown, but it has been observed in many fluorescent proteins. For example, Dickson et al. (*2*) reported that single molecules of yellow-emitting GFP mutants (S65G/S72A/T203F/Y) become semipermanently dark after emission of ∼$10^6$ photons. Emission could be restored by irradiation at 405 nm. They postulated a dark state corresponding to a neutral form of the chromophore that may be stabilized by an alternative hydrogen bond pattern. While the structural details for the transition have not been established, it is attractive to hypothesize that the light and dark states are related by cis−trans isomerization of the chromophore. In GFP, the chromophore configuration is cis and is found to be deeply embedded within the rigid environment of an 11-strand $\beta$-barrel. As this rigid shell is generally held to be responsible for the high quantum yield of fluorescence, it is difficult to envision how cis−trans isomerization might be accommodated.

Ground-state isomerization for model compound 4-hydroxybenzylidene-1,2-dimethylimidazolinone (HBDI) was investigated in solution by He et al. for different protonation states (*12*; Scheme 1).

[†] This work was supported by NSF Grant MCB-0417290 (S.J.R.), a NIH predoctoral traineeship (GM-00759) (D.M.A.), and a grant for Molecular and Cellular Biology from the Russian Academy of Sciences and NIH/NGMS Grant GM070358 (D.M.C.).

[‡] Atomic coordinates and structure factors have been deposited with the Protein Data Bank (entry 1XMZ).

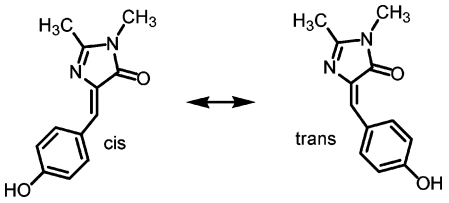
* To whom correspondence should be addressed: Institute of Molecular Biology, University of Oregon, Eugene, OR 97403. Phone: (541) 346-5190. Fax: (541) 346-5870. E-mail: jremington@uoxray.uoregon.edu.

[§] University of Oregon.

[∥] Russian Academy of Sciences.

10.1021/bi047644u CCC: $30.25 
Published on Web 03/19/2005

Scheme 1




These authors reported experimental free energy differences for cis−trans isomerization of 8.8 and 9.6 kJ/mol for the neutral and anionic forms of HBDI, respectively. The trans planar configuration is disfavored by steric clashes between the aromatic C$\delta$2-H and the imidazolinone oxygen. The activation energy barrier for isomerization was found to be 54.8 kJ/mol in each case, substantially lower than that predicted on theoretical grounds. The energy of visible photons is more than sufficient to effect these transitions. While the biological significance of light−dark transitions in GFP is unclear, light-driven isomerization of a closely related chromophore forms the basis for the activity of the blue light photoreceptor photoactive yellow protein (PYP) (briefly reviewed in ref *13*). Several groups have visualized intermediates in the PYP photocycle by time-resolved crystallography (e.g., see refs *14*−*16*), revealing complex adjustments on the part of the chromophore and surrounding side chains.

Among the GFP family proteins, there are both strongly fluorescent proteins [e.g., DsRed from *Discosoma* sp. (*17*) or eqFP611 from *Entacmaea quadricolor* (*18*)] and nonfluorescent proteins with a blue or purple appearance [e.g., asFP595 from *Anemonia sulcata* (*19*) or Rtms5 (*20*)]. The crystal structures of several of these proteins have been determined, revealing both cis [DsRed (*21*, *22*)] and trans [eqFP611 (*23*) and Rtms5 (*20*)] chromophore configurations. Therefore, cis−trans isomerization per se cannot be solely responsible for light−dark transitions in general. In the chromoprotein Rtms5, the chromophore not only has the trans configuration but also is significantly nonplanar as well (*24*), suggesting that some combination of these properties may be paramount in determining the emissive state.

asFP595 is a nonfluorescent member of a family of GFP-like proteins isolated from Anthozoa (see reviews in refs *17*, *25*, and *26*), in this case from the tentacles of the Mediterranean sea anemone *A. sulcata* (*19*). When this protein is subjected to sufficiently intense illumination near its absorbance maximum ($\lambda_{abs}^{max} = 568$ nm), it undergoes a remarkable transition, termed "kindling", to a long-lived fluorescent state ($\lambda_{em}^{max} = 595$ nm). This kindling property can be enhanced by the A143G point mutation (*27*), giving rise to a variant hereafter termed "kindling fluorescent protein" (KFP).[1] In the dark recovery phase, the brightly fluorescent kindled state of KFP relaxes thermally with a time constant of ~100 s at room temperature, or can instantly be reverted upon illumination at ~450 nm. In the kindled state, the absorbance spectrum shows an increase at 450 nm and a decrease at 568 nm, suggesting that in the light-activated state, a change in the chromophore environment favors the protonated neutral chromophore. Locations of mutations (including A143G) on a hypothetical model of KFP led to the suggestion that they could influence the ease of cis−trans isomerization (*28*).

The high-resolution crystallographic studies of the dark-adapted state of KFP reveal, as in the case of Rtms5 (*24*), a significantly nonplanar chromophore in the trans configuration. Additionally, the model supports the hypothesis that cis−trans isomerization is at least indirectly responsible for the kindling phenomenon. The crystal structure confirms the presence of a break in the polypeptide backbone within the native protein immediately adjacent to the chromophore, as previously proposed (*29*, *30*). The results are discussed in the context of recent studies on other members of this intriguing family of proteins. Taken together, these considerations lead to a proposal regarding the nature of the kindling transition and an explanation for the slow return to the nonfluorescent ground state.

## MATERIALS AND METHODS

*Protein Expression and Purification.* KFP was overexpressed in *Escherichia coli* strain JM109 DE-3 using the pQE-30 vector (Clontech), which also encodes an amino-terminal six-histidine tag. Cells were grown at 37 °C in a fermenter containing 4 L of LB with 100 mg/L ampicillin to an $A_{600}$ of 0.6. Protein expression was induced with 0.5 mM IPTG overnight at 20 °C. The cells were disrupted using a French press at 10K psi. Purification was achieved using Ni−NTA agarose (Qiagen) chromatography with a gradient from 100 to 300 mM imidazole. Purified KFP was dialyzed into 50 mM HEPES (pH 7.9), 0.3 M NaCl, and 1 mM $\beta$-mercaptoethanol, and then concentrated to an $A_{280}$ of 47.0 using centrifugal filters with a molecular mass cutoff of 30 kDa (Millipore).

*Arrhenius Plot Analysis.* The activation energy barrier for thermal relaxation of the light-induced fluorescent state was estimated by Arrhenius plot analysis, both in single crystals of KFP and in solution. The experimental setup consisted of a microspectrophotometer (4DX Systems, Uppsala, Sweden) using a weak incandescent white light source as a probe and a 2 mW 543.5 nm unfocused HeNe laser as an excitation source. For the solution studies, ~2 mm$^3$ samples of purified KFP ($A_{280} = 0.2$ for a path length of 1 mm) were placed into a 1 mm diameter capillary tube sealed with paratone oil. The temperature was measured at the sample using a type T thermocouple microprobe sealed into the tube and held at desired values using a dry air-stream cooler (FTS Systems, Stone Ridge, NY). The laser did not induce significant heating of the sample. After activation by the HeNe laser for 2 min, absorbance spectra (400−800 nm) were collected at 1−10 s intervals depending on temperature. The absorbance at 568 nm was extracted by Gaussian decomposition. Single-exponential curve fitting of the relaxation time course permitted the relaxation rate constant as a function of temperature to be determined.

*Crystallization and Data Collection.* Crystals were grown by hanging-drop vapor diffusion using a well solution consisting of 26% PEG 1550, 0.14 M sodium citrate, and 0.1 M Tris (pH 9.5). Drops containing 3 $\mu$L of concentrated protein and 3 $\mu$L of well solution were equilibrated with 1 mL of well solution at room temperature. Crystals appeared

[1] Abbreviations: KFP, kindling fluorescent protein; BME, $\beta$-mercaptoethanol; rms, root-mean-square; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis.

in 3−5 days and continued to grow to a size of 0.3−0.5 mm in 3−4 weeks. For data collection, crystals were equilibrated briefly in well solution, passed through paratone (Chevron), and then flash-frozen in a nitrogen cold stream at 100 K. A preliminary 1.9 Å data set was collected on an RAXIS-IV detector (MSC) using a rotating-anode source. Diffraction data to 1.38 Å were measured subsequently on beamline 14 BM-C at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL) using a Quantum 4 detector (ADSC). Three hundred sixty images were collected with an oscillation range of 0.5°, a detector distance of 120 mm, and an exposure time of 3 s. The images were processed using the *HKL2000* suite (HKL Research).

*Structure Determination and Refinement.* The structure of KFP was determined by molecular replacement against the 1.9 Å data set using *EPMR* (*31*). The search model consisted of a polyalanine skeleton derived from the coordinates of the related protein DsRed (PDB entry 1G7K) (*21*). The preliminary model was refined against the 1.38 Å data set using torsion angle simulated annealing with *CNS* (*32*) at a resolution of 2.0 Å, followed by conjugate direction minimization with *TNT* (*33*) to a resolution of 1.6 Å. The final rounds of refinement were carried out with *SHELXL* (*34*) to permit inclusion of all data to the diffraction limit of 1.38 Å with 5% of the data excluded to permit analysis of $R_{free}$. This also facilitated refinement of anisotropic displacement parameters, modeling of disordered residues, and addition of riding hydrogen atoms. During the final refinement cycles, no restraints other than bond distances were applied to the chromophore. Between cycles of refinement, the model was rebuilt using *O* (*35*).

*Modeling of KFP Mutants.* To understand the effects of the A143S and S158V mutants on the chromophore environment, the structures of these mutants were modeled using the crystal structure of KFP. For the A143S mutant, the serine side chain at position 143 was modeled in the most probable conformation ($\chi_1 = 62°$), a conformation that is similar to that adopted by the side chain of Ser146 in DsRed ($\chi_1 = 74°$). Modeling the serine side chain in this position, however, introduced a close contact (1.3 Å) with the imidazole side chain of His197. To prevent this collision, His197 was modeled in the lower-occupancy conformation observed in the crystal structure. To model the structure of the S158V mutant, substitutions were made at two sites. First, to revert the structure of KFP (the A143G mutant of asFP595) to that of wild-type asFP595, C$\beta$ from the side chain of Ser143 (as modeled above) was used as the side chain of Ala143. Again, as in the case of the A143V mutant, a close contact with the side chain of His197 was avoided by modeling this side chain in the lower-occupancy conformation. Second, the side chain at position 158 was modeled as the most probable rotamer, for which $\chi_1 = 174°$. This conformation is similar to that of Ile161 in DsRed, with the C$\gamma$ atoms directed away from the chromophore.

*Analysis of Chromophore Planarity.* To analyze the distortions from planarity of the chromophores observed in the structures of GFP homologues, least-squares ("best") planes were calculated for the imidazolinone (including atoms C$\alpha$2, C2, N2, C1, N3, O2, C$\alpha$1, C$\alpha$3, and C$\beta$2) and *p*-hydroxybenzyl (including atoms C$\beta$2, C$\gamma$2, C$\delta$1, C$\delta$2, C$\epsilon$1, C$\epsilon$2, C$\zeta$, and O$\eta$) moieties using *GFPLANE* (M. L. Quillin, unpublished program). In addition to the determination of an overall dihedral angle, the relative orientation of the two planes was characterized further by calculating two orthogonal components. The first component, tilt, is defined as the sum of the angles between each plane and the line through the centers of mass of the atoms within each ring system. The second component, twist, is the dihedral angle formed by the planes containing the line through the centers of mass and perpendicular to the least-squares plane of each ring system.

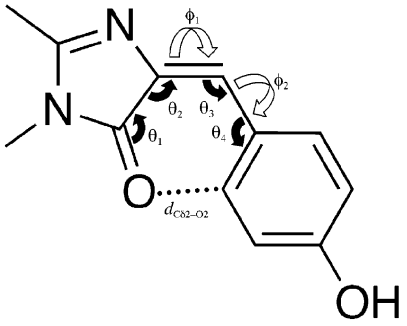



FIGURE 1: Schematic diagram of chromophore stereochemistry. The chromophore is shown in the trans configuration. Bond angles are shown as black arrows; torsion angles are shown as white arrows, and contact distances are shown as dotted lines. See Table 3 for a description of each angle.

*Modeling of Alternative Coplanar Chromophore Conformations.* We investigated the potential effects of the protein environment on coplanar chromophore configurations. An idealized starting model based on the crystallographic coordinates was obtained using the following protocol, as implemented in *GFPLANE*. First, the coordinates for each of the atoms that comprise the respective planes were projected onto the best plane, producing a flattened chromophore structure. The C$\beta$2 atom, which serves as a pivot for both planes, was projected onto the line formed by the intersection of the two planes. In proteins with very small angles between planes (GFP and DsRed), the original coordinates for the C$\beta$2 atom were retained to prevent large positional shifts resulting from uncertainty in the line of intersection. To obtain idealized cis conformations for each protein, the $\phi_1$ and $\phi_2$ torsion angles that describe the relative orientation of the planes (Figure 1) were fixed at 0°. For the trans conformation, the $\phi_1$ and $\phi_2$ angles were fixed at 180° and 0°, respectively.

*Calculation of Cavity Surfaces and Volumes.* Cavity calculations were performed with *MSP* (*36*) using default atomic radii and a probe radius of 1.4 Å. For display purposes, the output file containing polyhedral data was converted into a format that could be interpreted by *MOLSCRIPT* (*37*).

## RESULTS

*Arrhenius Plot Analysis.* The thermally activated decay of the fluorescent state was monitored by the increase in absorbance at 568 nm, following sample activation by 543.5 nm laser illumination. At temperatures above 15 °C, the decay could be accurately fit by a single exponential, but below that temperature, the decay was biphasic (data not shown). Although the decay at lower temperatures could be accurately modeled, the amplitudes and time constants of the two components differed by factors of only 2−3 and so were not well determined by the fitting process. Conse-

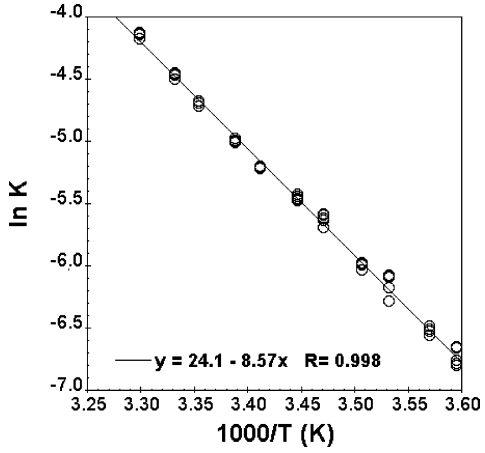



FIGURE 2: Arrhenius plot analysis for thermally driven relaxation from the activated fluorescent state of KFP to the nonfluorescent ground state. Open circles denote individual determinations of the relaxation rate constant at a given temperature (at least three per measurement).

quently, single-exponential fits were used to estimate the first-order rate constants as a function of temperature. The results are satisfactory and are displayed as an Arrhenius plot (Figure 2). From this analysis, the energy barrier for thermal relaxation from the fluorescent state to the nonfluorescent ground state was estimated to be 71 kJ/mol, which is somewhat higher than the values of ~55 kJ/mol reported for cis−trans isomerization of a model chromophore in solution (*12*). However, if cis−trans isomerization is occurring within the protein, the higher value is likely due to interactions with or rearrangements of the surrounding protein.

*Determination of the Structure of KFP*. Crystals of KFP grow in space group $C222_1$ with the following unit cell dimensions: $a = 76.52$ Å, $b = 125.46$ Å, and $c = 92.75$ Å. Although these crystals clearly diffract to higher resolution as shown by the low $R_{merge}$ and high signal-to-noise ratio in the highest-resolution shell (Table 1), the detector configuration limited the resolution to 1.38 Å. The resulting data set was very complete with a greater than 7-fold redundancy. A preliminary model of KFP was obtained by molecular replacement using the structure of the homologous protein DsRed [PDB entry 1G7K (*21*)] as the search model. Initial refinement of this model in *TNT* gave an *R*-factor of 20.8% to 1.6 Å resolution. Further refinement to 1.38 Å resolution using *SHELXL* gave values of 22.6 and 26.6% for *R* and $R_{free}$, respectively, before anisotropic *B*-factor refinement. Incorporation of anisotropic *B*-factors into the model gave *R* values of 15.1 and 20.3% for the working and test sets, respectively. Addition of riding hydrogen atoms gave an *R*-factor of 13.9% and an $R_{free}$ of 19.1%. Further refinement after combining working and test sets gave a final model with an *R*-factor of 14.0% with acceptable deviations from ideal stereochemistry (Table 1). Ramachandran analysis of main chain dihedral angles using PROCHECK (*38*) shows that, of the 372 non-proline, non-glycine residues observed in the structure, 339 (91.1%) are in most favored regions, 32 (8.6%) are in additional allowed regions, and only one (Asp151 of chain A) is in a disallowed region.

*Description of the Overall Structure of KFP*. As expected from the high degree of sequence homology (Supporting Information Figure S1), the overall fold of KFP closely resembles that of GFP and related proteins (Table 2). Even though the structure of DsRed proved to be sufficiently similar to serve as a search model for molecular replacement, KFP is most similar to the red fluorescent protein eqFP611, both in sequence and in structure. The tertiary structure of KFP consists of an 11-strand $\beta$-barrel with a central α-helix parallel to the barrel axis, commonly called a "$\beta$-can" fold (*39*). The two protomers in the asymmetric unit are related to the remaining protomers in the intact tetramer by a

Table 1: Crystallographic Statistics for KFP

| | |
|---|---|
| data collection | |
| space group | $C222_1$ |
| unit cell dimensions (Å) | $a = 76.52$, $b = 125.46$, $c = 92.75$ |
| no. of observations | |
| total | 653351 |
| unique | 89428 |
| resolution (Å) | 50.00−1.38 (1.43−1.38)[a] |
| completeness (%) | 97.5 (97.9)[a] |
| $I/\sigma(I)$ | 49.1 (8.4)[a] |
| $R_{merge}$ (%) | 4.4 (16.6)[a] |
| refinement | |
| no. of non-hydrogen atoms | |
| protein | 3728 |
| chromophore | 46 |
| $\beta$-mercaptoethanol | 28 |
| water | 463 |
| *R*-factor (%)[b] | |
| working | 13.9 (84670) |
| test | 19.1 (4745) |
| final | 14.0 (89415) |
| rms deviations from target values[c] | |
| bond lengths (Å) | 0.012 (3941) |
| bond angles (Å) | 0.031 (5276) |
| zero chiral volumes (Å$^3$) | 0.070 (583) |
| non-zero chiral volumes (Å$^3$) | 0.076 (558) |
| planarity (Å) | 0.424 (1258) |
| antibumping distances (Å) | 0.024 (209) |
| through-bond ADPs (Å$^2$) | 0.004 (9253) |
| through-space ADPs (Å$^2$) | 0.035 (24732) |
| isotropic ADPs for solvent (Å$^2$) | 0.090 (2778) |

[a] Numbers in parentheses correspond to values for the highest-resolution shell. [b] Numbers in parentheses are the numbers of reflections within each set. [c] Numbers in parentheses are the numbers of occurrences of each type of restraint.

Table 2: Structural and Sequence Similarity among GFP Homologues

Sequence Similarity (%)[a]

| | GFP | DsRed | Rtms5 | eqFP611 | KFP |
|---|---|---|---|---|---|
| GFP | − | 26.7/50.2 | 28.4/50.0 | 25.0/46.9 | 23.1/43.1 |
| DsRed | | − | 63.1/74.7 | 51.2/69.1 | 46.1/62.7 |
| Rtms5 | | | − | 28.4/50.0 | 46.8/64.7 |
| eqFP611 | | | | − | 58.9/71.9 |
| KFP | | | | | − |

Structural Similarity (Å)[b]

| | GFP | DsRed | Rtms5 | eqFP611 | KFP |
|---|---|---|---|---|---|
| GFP | −(225) | 1.33 (208) | 1.50 (211) | 1.50 (212) | 1.47 (209) |
| DsRed | | −(217) | 0.62 (217) | 0.82 (217) | 0.94 (214) |
| Rtms5 | | | −(218) | 0.91 (218) | 0.93 (214) |
| eqFP611 | | | | −(224) | 0.69 (217) |
| KFP | | | | | −(224) |

[a] Sequences were aligned using the Needleman−Wunsch algorithm as implemented in LSQMAN (*55*). Values before and after the slash are percent identity and percent similarity, respectively. [b] For each pair of structures, α-carbon atoms in chain A were structurally aligned using LSQMAN (*55*). The values correspond to rms deviations in α-carbon position (angstroms). The numbers in parentheses are the numbers of matched residues (total number of residues along the diagonal).

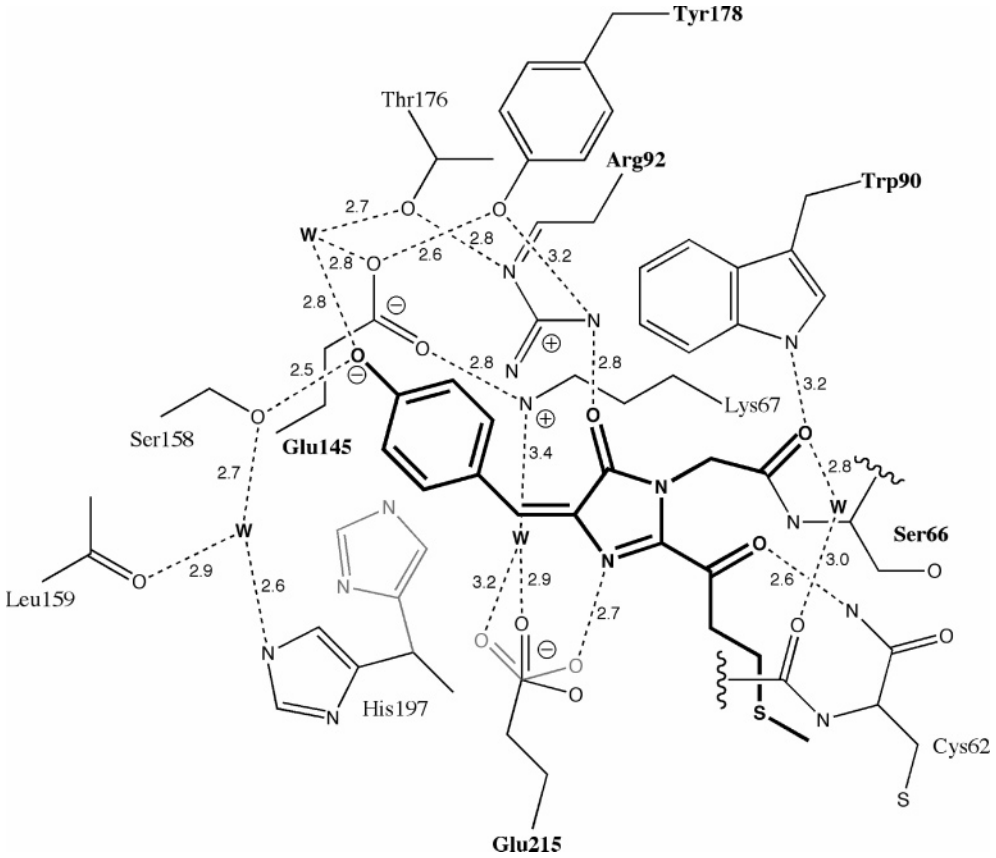



FIGURE 3: Schematic diagram of the structure and environment of the KFP chromophore. The chromophore as well as the names of conserved residues is shown in bold. Minor conformations of His197 and Glu215 are shown in gray. Water molecules are represented by the letter W. Hydrogen bonds are shown as dashed lines with numbers indicating the distance in angstroms.

crystallographic 2-fold axis parallel to the *y*-axis. Although both protomers are very similar (rmsd of 0.29 Å for residues 6−232), large differences can be seen near the amino terminus. In the first molecule (designated chain A), the amino terminus is disordered prior to residue 4, while in the second molecule (chain B), all amino-terminal residues in the native protein as well as four residues in the polyhistidine tag are clearly resolved. Residues 206 and 207 are completely disordered in both protomers. In contrast, the carboxy-terminal residues are well-ordered in both chains due to the formation of a salt bridge between the carboxyl terminus and Arg198 of an adjacent monomer. Seven $\beta$-mercapto-ethanol molecules and 463 water molecules were modeled in the structure. The BME molecules form disulfide bonds with the side chains of cysteines 114, 164, and 222 in each protomer in the asymmetric unit, as well as Cys210 of chain B.

*Alternate Conformations for Discretely Disordered Side Chains.* Electron density maps clearly indicated the presence of alternate conformations for 41 side chains [including 12 found in both chains (Met9, Thr13, Glu16, Glu47, Lys71, Thr95, Ile121, Thr144, Arg155, Lys181, His197, and Tyr211), 10 of which are only in chain A (Thr14, Lys25, Asn33, Glu36, Glu40, Lys42, Lys120, Lys188, Glu195, and Glu215) and seven of which are only in chain B (His22, Glu44, Glu97, Ser109, Lys118, Val150, and Asp223)].

Two of the disordered side chains, His197 and presumed catalytic Glu215, are highly conserved and are located near the chromophore (Figure 3). His197 is found in two discrete conformations (Figure 4). The predominant conformation exhibits side chain dihedral angles $\chi_1$ of 57° and 59° and $\chi_2$ of 88° and 87° in chains A and B, respectively, while the minor conformation has dihedral angles $\chi_1$ of −4° and −59° and $\chi_2$ of 57° and −106°, respectively, for these chains. The larger variability in dihedral angles for the minor conformation is presumably due to greater positional uncertainty of the lower-occupancy conformation. When the sum of the occupancies of these conformations is constrained to be unity, the relative occupancies are 0.86 and 0.14 in chain A and 0.61 and 0.39 in chain B. The unconstrained occupancies refined to values of 0.73 and 0.14 in chain A and 0.60 and 0.32 in chain B. In both molecules in the asymmetric unit, two water molecules form close contacts with the minor conformation and are presumably present only when the His197 side chain adopts the major conformation. Although it is likely that water molecules are similarly excluded by the major conformation, they might not be observed due to their low occupancy. The side chain of Glu215 is significantly disordered only in the structure of chain A, with relative proportions of 0.68 and 0.32 for the two modeled conformations (unconstrained occupancy values were 0.73 and 0.32, respectively). The predominant conformation differs from the lesser by 123° in $\chi_3$ (142° vs 19°), with changes of only 20° in $\chi_2$ (163° vs −177°) and 10° in $\chi_1$ (179° vs −171°). The conformation of Glu215 in chain B is nearly identical to that observed for the major conformation in chain A, with the following side chain dihedral angles: $\chi_1 = -174°$, $\chi_2 = 170°$, and $\chi_3 = 133°$. It is not clear why the side chain of Glu215 appears to be disordered only in one chain, but it may be related to the degree of disorder

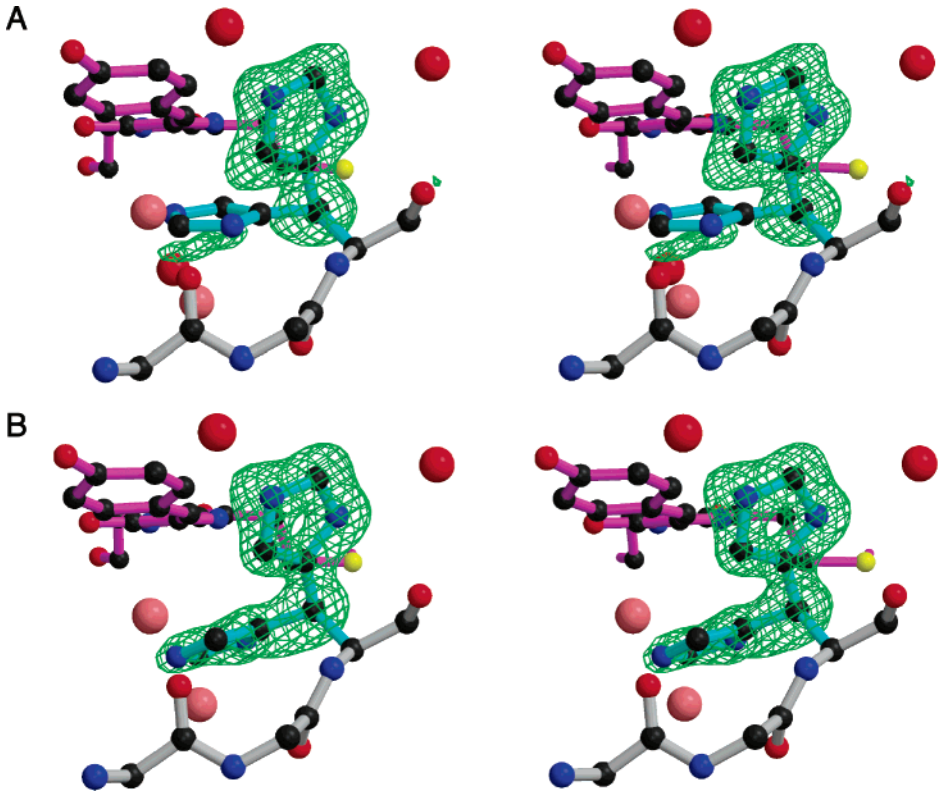



FIGURE 4: Stereo images of the electron density surrounding the alternate conformations of His197 in both molecules in the asymmetric unit, designated chains A and B. The omit map calculated after removal of atoms of the histidine side chain is contoured at 2.5$\sigma$. The side chain of His197 is colored cyan, while the chromophore is colored magenta. Water molecules that are mutually exclusive from either conformation are colored pale red. This figure was created using BOBSCRIPT (*57*) and RASTER3D (*58*).

present at position 197. Because of their proximity to the chromophore, the conformational heterogeneity observed at His197 and Glu215 may be central to the light-driven changes in chromophore structure proposed herein (see the Discussion).

*Structure and Environment of the KFP Chromophore.* As might be expected from its deeply buried location, the chromophore in KFP makes extensive contacts with neighboring residues (Figure 3). Its trans, nonplanar configuration is extremely well defined by the electron density map (Figure 5A). Most of the 45 contacts (defined as atoms within 3.5 Å of the chromophore) are with the *p*-hydroxybenzyl and imidazolinone ring systems (16 and 17 contacts, respectively), with the remaining contacts involving chromophore atoms derived from either the methionine (eight contacts) or glycine (five contacts) residues (Supporting Information Table S1). Because of the many polar substituents on the chromophore, it forms the nucleus of an extended network of hydrogen bonds within the core of KFP. The phenolic oxygen of the *p*-hydroxybenzyl moiety hydrogen bonds to the side chain hydroxyl of Ser158 and to a nearby water molecule. This water molecule interacts in turn with the side chains of Thr176 and Glu145. Ser158 also accepts a hydrogen bond from a water molecule that is hydrogen bonded to the main chain carbonyl of Leu159 and the major conformation of the His197 side chain. One of the carboxyl oxygens from the side chain of Glu215 interacts through a water molecule with Lys67, while the other forms a hydrogen bond with a nitrogen atom in the imidazolinone ring. Another water-mediated interaction joins the main chain carbonyl oxygens of Ser61 and Gly65. The side chains of aromatic residues Tyr178 and Trp90 make hydrogen bonds to the side chains of Arg92 and Glu145, and to the carbonyl oxygen of Gly65, respectively. Despite the relative abundance of positively (Arg92 and Lys67) and negatively (Glu145 and Glu215) charged residues, there is only one salt bridge near the chromophore, involving Lys67 and Glu145. The second positively charged residue, Arg92, interacts with the imidazolinone oxygen and the hydroxyl oxygens of Thr174 and Tyr178, all of which are uncharged. Finally, the chromophore carbonyl and the amide of Cys62 that are the remnants of the hydrolyzed acylimine form a hydrogen bond in the mature protein.

*Main Chain Fragmentation in KFP.* One of the chief differences between the chromophore structure as observed in KFP and in related proteins is the presence of a main chain break immediately preceding the chromophore. Samples of KFP used for crystallization exhibited greater than 90% cleavage as determined by SDS−PAGE (data not shown), confirming previous reports of chain fragmentation in KFP (*29*, *30*). Furthermore, absorbance spectra of purified KFP show no evidence for partial maturation, such as a green-emitting GFP-like chromophore. In agreement with these results, electron density maps revealed no significant density for the main chain between Cys62 and Met63, the first amino acid in the chromophore triplet (Figure 5B). We assume that the fragmentation can be attributed to hydrolysis of an acylimine moiety that forms during chromophore maturation, as in other reef GFPs [e.g., DsRed (*40*)] upon harsh treatment such as boiling. In modeling the chromophore structure, it was assumed that hydrolysis generates a carboxamide group on Cys62 and a carbonyl on the Cα1 atom of Met63 of the chromophore (see Figure 3), although even at 1.38 Å resolution, nitrogen and oxygen atoms cannot be distin-

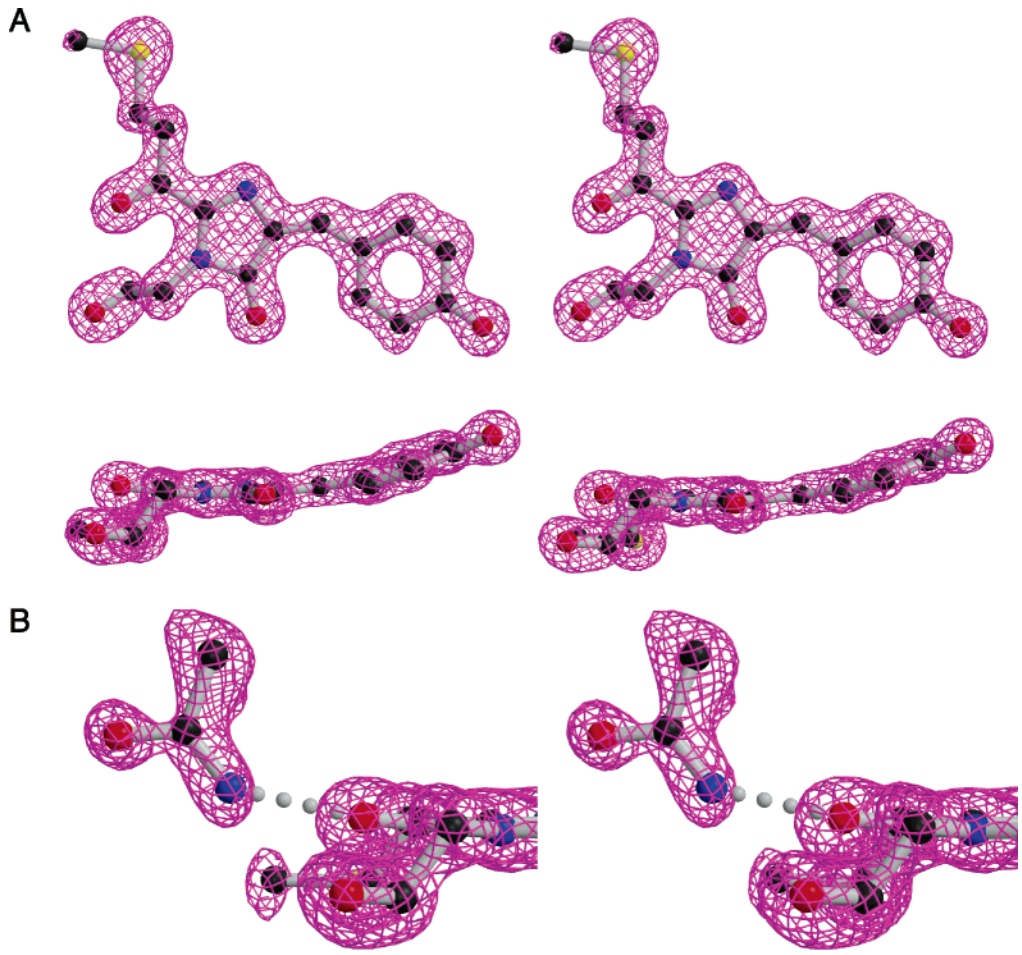



FIGURE 5: Stereo images of the omit electron density map surrounding the KFP chromophore. (A) Views perpendicular (top) and parallel (bottom) to the approximate plane of the chromophore. (B) Close-up of the chain break (dotted line). The omit map calculated after removal of chromophore and main chain atoms near the chain break is contoured at 4$\sigma$ in panel A and at 3$\sigma$ in panel B. This figure was created using BOBSCRIPT (*57*) and RASTER3D (*58*).

guished unambiguously. The presumptive amide nitrogen does appear to form a hydrogen bond with the O1 atom of the chromophore. Similarly, the carbonyl oxygen appears to form a hydrogen bond with the side chain hydroxyl of Thr13.

*Chromophore Conformation.* In the ideal case, the double bond present between the imidazolinone and *p*-hydroxybenzyl rings (Figure 1) requires that the chromophore adopt a conformation that is either cis or trans. GFP homologues with cis chromophores include GFP and DsRed, while those with trans isomers include Rtms5, eqFP611, and KFP. In the trans conformation, there is an internal collision between the O2 and C$\delta$2-H atoms of the chromophore (Figure 1) that results in distortions of the chromophore geometry from ideal bond lengths and angles. One such distortion is an increase in the bond angles that sequentially connect these atoms (Table 3A). The most significant changes occur between the C2−C$\alpha$2−C$\beta$2 ($\theta_2$) and C$\alpha$2−C$\beta$2−C$\gamma$2 ($\theta_3$) angles, with average increases of 12° and 4°, respectively. These changes, combined with smaller increases in the remaining angles, give rise to an overall increase of 20° in the bond angles joining the O2 and C$\delta$2 atoms in the trans compared to the cis conformation. When comparing cis and trans chromophores, one might expect to find similar distortions in the torsion angles that relate the O2 and C$\delta$2 atomic positions; however, no clear trends in these angles were observed (Table 3A). In the cis chromophore found in DsRed (*21*), both torsion angles are small, reflecting the coplanarity of the ring systems. In contrast, the angles observed in the cis chromophore present in GFP (*41*) are much larger despite the absence of an internal steric clash. Similarly, the torsion angles observed among chromophores in the trans conformation exhibit considerable variability. For eqFP611 and KFP, values of $\chi_1$ are approximately 10° out of plane, in contrast to the 40° deviation seen in Rtms5. Values for $\chi_2$ are around 5° in eqFP611 with 2- and 3-fold greater values observed in Rtms5 and KFP, respectively. Given the large variability in torsion angles among structures with either geometric isomer, the close contact present in the trans planar conformation does not appear to affect these angles in a consistent manner.

*Chromophore Planarity.* A simple measure of chromophore planarity can be obtained by determining the dihedral angles between the least-squares planes of the imidazolinone and *p*-hydroxybenzyl rings. In proteins with cis chromophores, the ring systems that comprise the chromophore are in general more coplanar than in proteins with trans chromophores (Figure 6 and Table 3B). In the DsRed tetramer (*21*), for instance, all four chromophores have dihedral angles (and consequently, tilt and twist angles) of less than 1°. Similarly, in GFP (*41*) the dihedral angle between the ring systems is nearly 4°, due almost entirely to a tilting of the planes. In contrast, the dihedral angles of chromophores in the trans conformation are much less coplanar (Table 3B). The first prominent example of nonplanarity was seen in Rtms5 (*24*), which has an overall

Table 3: Chromophore Geometry in KFP and Other GFP Homologues

(A) Angles

| structure | chain | bond angles (deg)[a] | | | | | | torsion angles (deg)[b] | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\theta 1$ | $\theta 2$ | $\theta 3$ | $\theta 4$ | $\Sigma(\theta)$ | $d_{C\delta 2-O2}$ | $\chi_1$ | $\chi_2$ |
| Proteins with Cis Chromophores | | | | | | | | | |
| GFP | – | 125 | 125 | 127 | 117 | 493 | – | −13 | 18 |
| DsRed | a | 131 | 120 | 134 | 124 | 509 | – | 4 | 0 |
| | b | 132 | 119 | 137 | 123 | 511 | – | −2 | 1 |
| | c | 131 | 119 | 136 | 124 | 510 | – | −1 | 1 |
| | d | 132 | 120 | 136 | 124 | 511 | – | −1 | 2 |
| Proteins with Trans Chromophores | | | | | | | | | |
| Rtms5 | – | 130 | 130 | 139 | 121 | 520 | 3.3 | −137 | −11 |
| eqFP611 | a | 130 | 133 | 138 | 128 | 529 | 3.2 | −173 | −4 |
| | b | 130 | 133 | 138 | 128 | 529 | 3.2 | −173 | −5 |
| KFP | a | 136 | 133 | 141 | 116 | 526 | 3.1 | −171 | −17 |
| | b | 138 | 132 | 134 | 125 | 529 | 3.1 | −172 | −15 |

(B) Planes

| structure | chain | dihedral angle (deg)[c] | twist | tilt[d] |
|---|---|---|---|---|
| Proteins with Cis Chromophores | | | | |
| GFP | – | 3.7 | 0.0 | −3.7 |
| DsRed | a | 0.6 | −0.5 | −0.4 |
| | b | 0.3 | −0.3 | 0.2 |
| | c | 0.1 | −0.1 | 0.1 |
| | d | 0.7 | −0.7 | 0.0 |
| Proteins with Trans Chromophores | | | | |
| Rtms5 | – | 35.1 | −30.7 | −17.4 |
| eqFP611 | a | 4.4 | −2.5 | −3.6 |
| | b | 4.3 | −1.7 | −3.9 |
| KFP | a | 19.6 | −5.1 | −18.9 |
| | b | 20.2 | −6.1 | −19.2 |

[a] The atoms that form each bond angle are as follows: $\theta_1$, O2−C2−Cα2; $\theta_2$, C2−Cα2−C$\beta$2; $\theta_3$, Cα2−C$\beta$2−C$\gamma$2; $\theta_4$, C$\beta$2−C$\gamma$2−C$\delta$2. $\Sigma(\theta)$ is the sum of all bond angles. See Figure 1. [b] The atoms that form each torsion angle are as follows: $\chi_1$, C2−Cα2−C$\beta$2−C$\gamma$2; $\chi_2$, Cα2−C$\beta$2−C$\gamma$2−C$\delta$2. See Figure 1. [c] The overall dihedral angle is the angle between the planes of the imidazolinone (defined by atoms Cα2, C2, N2, C1, N3, O2, Cα1, Cα3, and C$\beta$2) and *p*-hydroxybenzyl (defined by atoms C$\beta$2, C$\gamma$2, C$\delta$1, C$\delta$2, C$\epsilon$1, C$\epsilon$2, C$\zeta$, and O$\eta$) moieties. [d] These angles are orthogonal components of the overall dihedral angle as described in the text.

dihedral angle of 35° and correspondingly large tilt and twist components. Likewise, the KFP chromophore has large dihedral angles, with values of 20° in both protomers in the asymmetric unit. While the magnitude and direction of the tilt component are similar in both KFP and Rtms5, the magnitude of the twist component is much larger in Rtms5 than in KFP (31° vs 4°, respectively). Because a large tilt component is coupled with a small twist component in KFP, the chromophore in this structure appears to be bowed along its long axis (Figure 5A). Interestingly, the structure of fluorescent eqFP611 (*23*) provides an important exception to the trend in dihedral angles seen in trans chromophores. The overall dihedral angle in this protein is 4°, a value much closer to those seen among the cis homologues. eqFP611 is the only known example of a fluorescent protein with a trans chromophore conformation, leading us to suggest that it is the degree of planarity rather than the geometric isomer that determines the fluorescent properties.

*Volume of the Chromophore Cavity.* In all GFP homologues whose crystal structures have been determined thus far, the *p*-hydroxybenzyl ring of the chromophore is enclosed in a relatively large, partially hydrated cavity. To determine the volume available for possible cis−trans isomerization, cavity volumes were calculated after omission of the *p*-hydroxybenzyl group (Table 4). The resulting cavities have volumes ranging from 120 to 180 Å$^3$ (for comparison, a water molecule has a volume of 21 Å$^3$ and the *p*-hydroxybenzyl ring has a volume of 110 Å$^3$). In KFP, the disordered

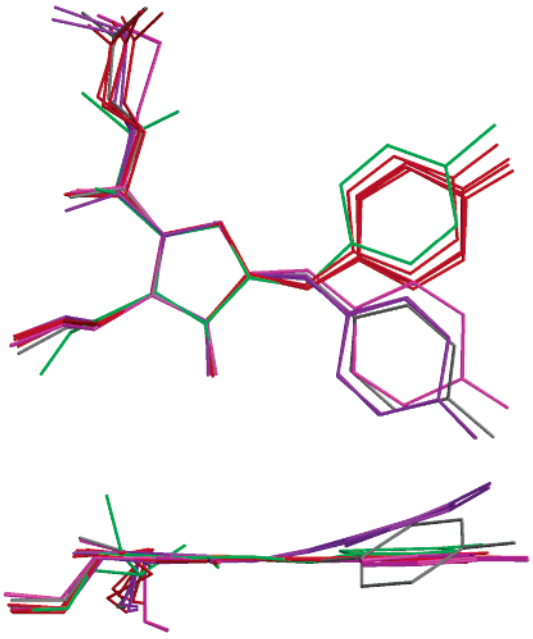


FIGURE 6: Overlay of chromophore structures found in GFP homologues. The chromophores included in the comparison are from GFP (green; PDB entry 1EMA), DsRed (red; PDB entry 1G7K), eqFP611 (magenta; PDB entry 1UIS), Rtms5 (gray; PDB entry 1MOU), and KFP (purple). The structures were superposed using nine atoms within and adjacent to the imidazolinone ring. This figure was created using MOLSCRIPT (*37*) and RASTER3D (*58*).

Table 4: Volumes of Cavities Enclosing the *p*-Hydroxybenzyl Ring in GFP Homologues

| protein | cavity volume (Å$^3$) |
|---|---|
| DsRed | 151 |
| Rmts5 | 140 |
| Rtms5 H146S mutant | 153 |
| eqFP611 | 169 |
| KFP (197a, 215a)[a] | 118 |
| KFP (197a, 215b)[a] | 118 |
| KFP (197b, 215a)[a] | 156 |
| KFP (197b, 215b)[a] | 153 |
| KFP A148S mutant (modeled) | 164 |
| KFP S158V mutant (modeled) | 178 |

[a] The a and b designations refer to the major and minor conformations of the disordered side chains, respectively.

conformers of His197 and Glu215 form part of the cavity wall. Because each side chain was modeled with two conformations, there are four possible combinations of cavity volumes. The observed volumes (Table 4) suggest that variability in the conformation of Glu215 does not substantially affect the volume of the cavity in KFP. However, the conformation chosen for the side chain of His197 has a dramatic effect on the volume accessible to the chromophore. With His197 in the predominant conformation, the volume of the cavity is approximately 120 Å$^3$ while the minor conformation gives a volume of 155 Å$^3$ (Figure 7). Although slight rearrangements of the chromophore and surrounding side chains might be necessary, it appears that the cis form of the chromophore could be readily accommodated within this cavity.

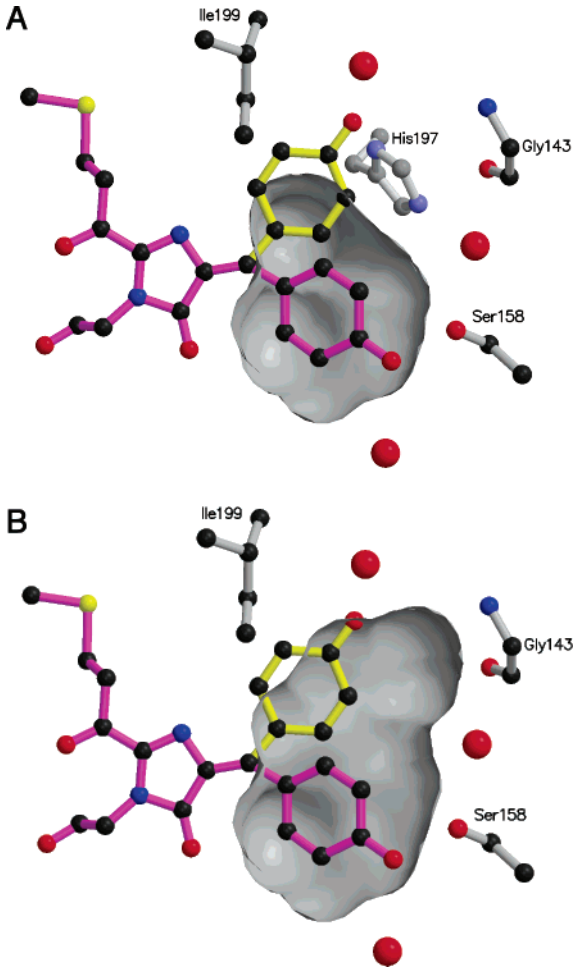



FIGURE 7: Cavity occupied by the *p*-hydroxybenzyl moiety in KFP. Cavities were calculated with the disordered side chain of His197 (shown in pale colors) in either the major (A) or minor (B) conformation. In Figures 7−9, the observed conformations of the chromophore are depicted in magenta, while the canonical cis (worst-case model) and trans isomers are colored yellow and cyan, respectively. Figures 7−9 were created using MSP (*36*), MOLSCRIPT (*37*), and RASTER3D (*58*).

## DISCUSSION

*Chromophore Structure.* As in all chromoproteins in the GFP family, KFP undergoes an autocatalytic cyclization reaction involving three residues (Met-Tyr-Gly) in the central α-helix followed by oxidation to form the mature chromophore. In red-emitting proteins such as DsRed, a second oxidation step follows the cyclization/oxidation reaction to generate an acylimine bond between the α-carbon and nitrogen of the first residue of the chromophore tripeptide (*40*, *42*). We assume that this reaction occurs in KFP as well, but in this case, the acylimine bond is subsequently cleaved, producing a break in the main chain immediately preceding the chromophore (*29*, *30*; see below). Zagranichny et al. have studied chromopeptides isolated from KFP and, on the basis of NMR and spectroscopic studies, have proposed that the cleavage reaction results in an N-unsubstituted ketimine instead of a carbonyl oxygen on the chromophore (*30*). However, no direct evidence for the existence of the ketimine in either the intact protein or the isolated peptides was presented. We chose to model the chromophore as resulting from hydrolysis of an acylimine linkage, which would most likely result in a carboxamide terminus on Cys62 and a carbonyl oxygen on the chromophore (Figure 3). In the following paper (*43*), the complete synthesis of 2-acetyl-4-(*p*-hydroxybenzylidene)-1-methyl-5-imidazolone, the proposed chromogenic core of KFP, is presented. As described therein, the absorbance and fluorescence emission spectra of this compound in dimethylformamide solution closely match those of KFP, which argues that this compound accurately represents the chromophore state within intact KFP. As noted in that paper, the free chromophore in solution is approximately one order of magnitude more fluorescent than the bound state within the protein. This suggests that the protein shell induces a conformation that favors nonradiative energy dissipation.

*Structural Basis for Fluorescent Behavior.* There are three distinctive features of the KFP chromophore that may give rise to the observed lack of fluorescence and the unique kindling behavior. First, the chromophore in KFP adopts a trans conformation about the bond between the imidazolinone and *p*-hydroxybenzyl moieties. Second, the chromophore is detached at one end from the polypeptide backbone as a result of the main chain cleavage. Third, the chromophore configuration is substantially distorted from planarity. Drawing from the literature, one can arrive at some tentative conclusions about the relative importance of these observations and their possible consequences.

*(1) Stabilization of the Trans Isomer.* On the basis of small-molecule crystal structures (*44*, *45*), quantum mechanical calculations (*46*), and recent NMR results (*12*), the preferred conformation of the chromophore is thought to be planar with the *p*-hydroxybenzyl group in the cis conformation. This preference for the cis over the trans isomer can be attributed

to a steric collision between the O2 and C$\delta$2-H atoms in the trans form (Figure 1). Although in solution the cis conformation predominates, the energetic difference between geometric isomers of the chromophore is small [3.3−9.6 kJ/mol (*12*)], so interactions with the local protein environment could lead to preferential stabilization of either conformation. Among native GFP homologues, the trans isomer has been observed in only Rtms5 (*24*) and eqFP611 (*23*) and, as described here, KFP. As pointed out previously, both cis and trans configurations can give rise to strongly fluorescent proteins.

We investigated the structural basis for the selective stabilization of the trans isomer in KFP by modeling the KFP chromophore in the cis configuration. In the worst-case (and very unlikely) scenario, this is accomplished by a 180° change in the torsion angle associated with the bridging double bond, holding the five-membered ring fixed. Of course, cis−trans isomerization could result from a more complex motion [such as a "hula twist" involving torsional motions of the methylene bridge (*46*)]. This would conserve the approximate positions of both the five- and six-membered rings and would be much more readily accommodated than the following worst-case scenario suggests. Nevertheless, the exercise is instructive. The simple torsional rotation described above results in large movements of the phenol moiety; however, it is surprising how few unfavorable steric clashes are observed (Figure 7A; Supporting Information Table S1). The worst violations involve the major conformation of the His197 side chain. Modeling this side chain in the minor conformation (see above) eliminates these clashes (Figure 7B) and in addition creates a favorable stacking interaction between imidazole and the *p*-hydroxybenzyl ring. Precisely such a stacking interaction is found in zFP538 (*42*) and eqFP611 (*23*), as well as in the engineered YFP (*47*). In Rtms5, the *p*-hydroxybenzyl ring stacks against the guanidinium moiety of Arg197, which replaces His197 in KFP.

Other steric collisions involving the chromophore in the cis conformation include the terminal side chain atoms of Ile199 and Glu215. Because, as in the case of His197, the side chain of Glu215 exhibits multiple conformations, this residue must be relatively unconstrained by the protein environment and, consequently, might be able to reorient to minimize the steric clash. By comparison, the side chain of Ile199 is not as flexible and may be an important determinant of chromophore selectivity in KFP (see below).

A second mechanism through which one isomer might be selectively stabilized is through the gain or loss of hydrogen bonds. This possibility is especially intriguing since the energy differences between cis and trans forms in solution are on the order of the energy of a single hydrogen bond (*12*). The modeled cis conformation results in the loss of two hydrogen bonds to the chromophore hydroxyl, one to the side chain of Ser158 and the other to a bound water molecule. The large number of water molecules within the chromophore cavity makes the prediction of hydrogen bond networks uncertain at best. In conclusion, although it is difficult to clearly identify the packing or hydrogen bond interactions that stabilize the observed trans configuration in the ground state, the exercise makes it clear that cis−trans isomerization could be readily accommodated within the chromophore cavity.

*(2) Main Chain Break.* A remarkable feature present in KFP that might contribute to its fluorescent properties is the main chain break adjacent to the chromophore. This break would clearly confer additional conformational flexibility to the chromophore. By comparison with other red-shifted fluorescent proteins, we assume that the break is the result of hydrolysis of an acylimine bond that is formed between the main chain amide nitrogen and the $\alpha$-carbon of Met63 during chromophore maturation. To date, such fragmentation has been observed in only one other native fluorescent protein structure, that of zFP538 (*42*, *48*). In zFP538, the main chain cleavage apparently results from an autocatalytic cyclization reaction in which the $\epsilon$-amino group of Lys66 attacks an acylimine formed from the backbone between residues 65 and 66, forming a three-ring chromophore. A similar break is induced in DsRed (*40*) and gtCP (*49*) upon boiling (highlighting the reactivity of the acylimine linkage) and in Kaede upon exposure to ultraviolet light (*50*). In the cases of KFP and zFP538, the fragmentation reaction must be autocatalytic. Although the catalytic residues that are involved in acylimine hydrolysis in KFP are not known, differences in the amino acid sequence between KFP and other fluorescent homologues near the site of bond cleavage point to groups that may be responsible. Candidates for catalytic groups include Thr13 (Val16 in GFP), which is only found in chromoproteins from *A. sulcata*, and Cys62 (Phe64 in GFP) and Ser66 (Val68 in GFP), which are commonly observed among other members of the class Anthozoa (*51*). Which, if any, of these residues are important for the presumed hydrolysis reaction will not be known until further mutagenesis studies have been performed. Given that both zFP538 and Kaede are strongly fluorescent proteins, it seems unlikely that the main chain break in KFP is centrally important in determining the fluorescent properties.

*(3) Chromophore Planarity.* A most remarkable feature of KFP is the pronounced distortion of the chromophore from planarity (Figure 5A). Similar "bowing" has been observed in the carotenoid ligand bound within $\beta$-crustacyanin and has been proposed to be one source of the bathochromic shift seen in the visible spectrum of this protein (*52*, *53*). To determine which of the surrounding residues are responsible for the lack of planarity, contact distances were analyzed for a model structure containing a hypothetical planar chromophore. It was again surprising that serious clashes were formed between the chromophore and the side chains of only two residues: His197 (∼2.6 Å, minor conformation) and Glu145 (∼2.0 Å). Both sets of interactions could presumably be relieved by minor adjustments in side chain conformations or in the protein matrix. A similar analysis was conducted on Rtms5, but similarly, this effort failed to reveal steric collisions that could convincingly be argued to prevent the chromophore from adopting a planar conformation. Although new contacts are formed with the side chains of Glu148 and Arg197 upon flattening, none of these contacts is closer than 3 Å.

These observations suggest that the energy required to distort the chromophore from planarity is small. Nevertheless, significant nonplanarity is observed only in the chromophores within the nonfluorescent proteins Rtms5 and KFP. For this reason, it seems very likely that the nonplanar configuration is largely responsible for the nonfluorescent state. Furthermore, the large volume of the chromophore cavity, evident

from the packing study discussed in subsection 1 above, may permit vibrational degrees of freedom not accessible to a more tightly packed chromophore. Nonradiative pathways for energy dissipation may be accelerated in an active site that accommodates such distortions, thereby accounting for the lack of fluorescence. In solution, there is no reason to suppose that such large distortions would be found, and indeed, in solution the quantum yield for fluorescence of the KFP chromophore is 1 order of magnitude higher than of the native protein (*43*).

*Proposed Conformational Switch for Fluorescence.* The kindling phenomenon cannot with confidence be attributed to any of the three unusual characteristics of KFP discussed above (trans conformation, chain break, or chromophore nonplanarity). Nevertheless, an attractive scenario draws upon all three features and includes a conformational switch, the side chain configuration of His197. It is clear that in the trans form present in the dark state, the KFP chromophore is nonplanar and nonfluorescent, so that nonradiative pathways of energy dissipation predominate. However, absorption of a photon of appropriate energy will induce an excited-state configuration that facilitates trans−cis isomerization of the chromophore (*28*). As described above, the most serious clash discovered by the modeling exercise results from an interaction of the chromophore cis-phenol with the major conformation of the His197 side chain. If the isomerization were to be accompanied by rearrangement of His197 to the minor conformation, the resulting stacking interaction may act to stabilize a planar, fluorescent configuration of the chromophore (either cis or trans). Assuming that the metastable state is cis, thermally driven recovery to the more stable trans nonplanar isomer should be slow with a high energy barrier, because it would require the concomitant rearrangement of His197. The break in the polypeptide backbone presumably increases the flexibility of the chromophore and may lower the activation energy for the process. Given such a combination of unusual features, one might expect kindling to be a rare phenomenon.

Structural information from observed and modeled chromophore environments in GFP homologues, as well as data from mutational analyses (*28*), provides evidence in support of this hypothesis. In Figure 8, the results of modeling KFP mutants A143S and S158V are illustrated. In Figure 9, the chromophore cavities within DsRed, eqFP611, and Rtms5 are compared. From the structural studies, the conformation of the chromophore appears to be largely determined by the nature of the amino acid side chains at positions 143, 158, and 199 (KFP numbering) as well as the conformation of His197. With the correlation of the residues at these positions with the observed orientation of the chromophore and the corresponding fluorescent properties, a simple scheme that predicts chromophore conformation can be developed.

*Position 143.* The residue at position 143 is highly variable among GFP homologues (Supporting Information Figure S1). The A143S mutation renders asFP595 strongly and permanently fluorescent, while the A143G (KFP) and A143T mutations result in weakly fluorescent proteins that can be further kindled (*28*, *54*). In DsRed, the corresponding Ser146 hydroxyl forms a hydrogen bond with the chromophore hydroxyl group, stabilizing the observed cis form of the chromophore (Figure 9A). In both asFP595 and Rtms5, serine residues modeled at this position (positions 143 and 146, respectively) can act to stabilize the cis isomer through a similar mechanism (Figures 8A and 9D), explaining the observed increases in fluorescence in both proteins. Presumably, the A143G mutation enlarges the cavity, favoring the cis configuration of the chromophore. In the ground state, the major conformation of His197 occupies the void. We assume that His197 must rearrange to the minor conformation to accommodate the cis chromophore. Interestingly, the residues occupying positions analogous to Gly143 in KFP are His146 in Rtms5 and Asp143 in eqFP611. In Rtms5, the imidazole of His146 (Figure 9C) adopts a conformation that is parallel to, but displaced from, the plane of the chromophore and would not appear to favor either cis or trans forms. Finally, Asn143 in eqFP611 should be able to hydrogen bond equally well with both isomers, again leading to no net stabilization of either configuration.

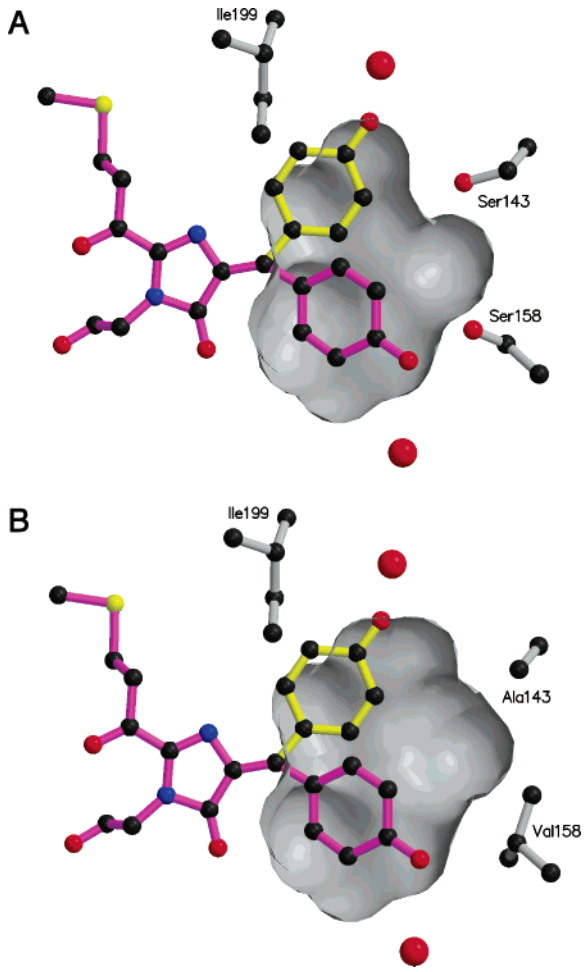



FIGURE 8: Cavities occupied by the *p*-hydroxybenzyl moeity in modeled chromophore conformations in the permanently fluorescent A143S (A) and S158V (B) mutants of asFP595. See the legend of Figure 7 for additional details.

*Position 158.* In contrast to positions 143 and 199, the amino acid at position 158 interacts exclusively with the trans isomer of the chromophore, providing it with significant influence in determining the favored configuration. When the residue at this position is capable of hydrogen bonding to the chromophore hydroxyl, as in asFP595 and eqFP611 (Ser158 in both cases; Figures 7 and 9B), and in Rtms5 (Asn161; Figure 9C and D), the trans isomer is stabilized. In contrast, the incorporation of large apolar residues at this position leads to a net destabilization of the trans conforma-

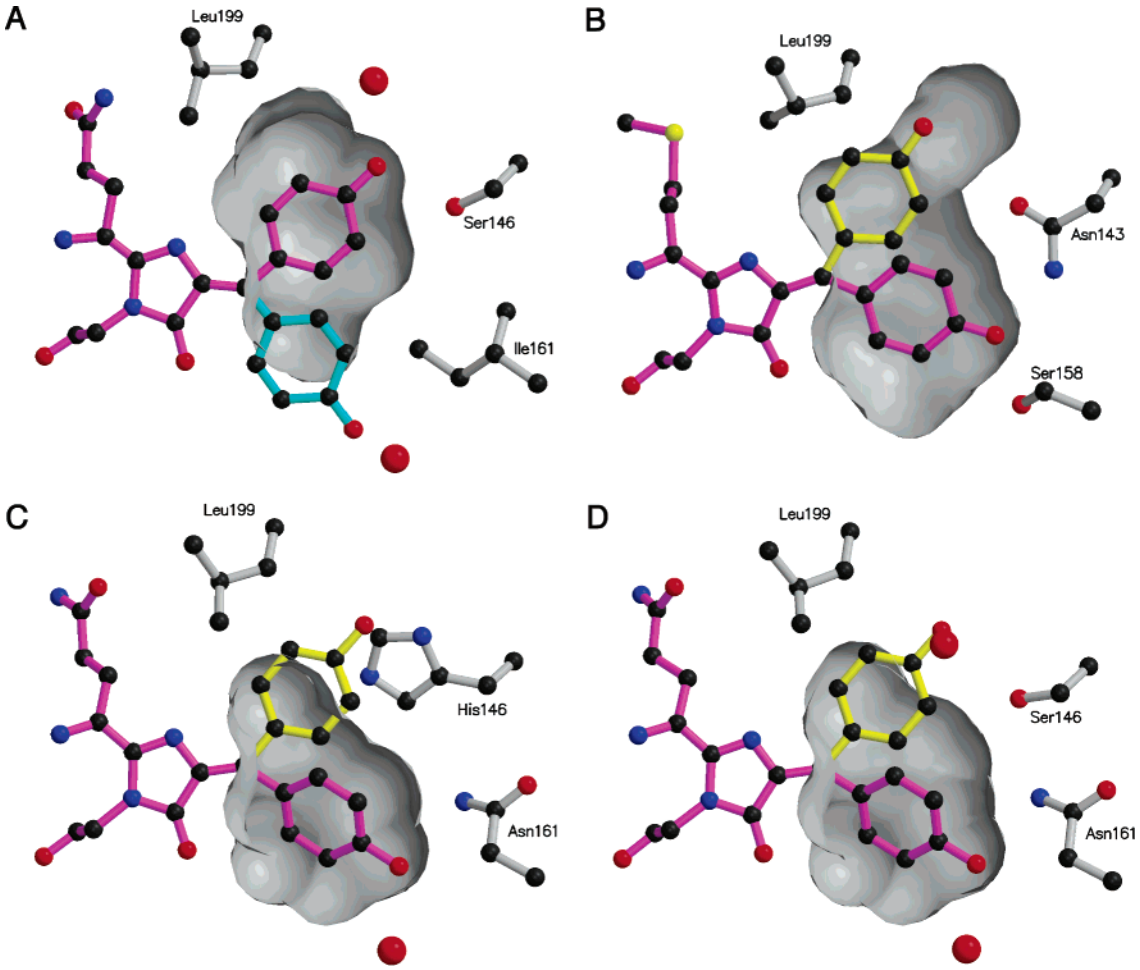



FIGURE 9: Cavities occupied by the *p*-hydroxybenzyl moiety in DsRed (A), eqFP611 (B), Rtms5 (C), and Rtms5 H146S (D). See the legend of Figure 7 for additional details.

tion due to repulsive steric interactions. This is clearly the case in DsRed with an isoleucine at this position (Figure 9A) as well as in GFP with a corresponding phenylalanine (not shown). While the structure of the S158V mutant of asFP595 is not known, modeling the smaller apolar residue valine at this position favors the cis isomer due to the loss of a hydrogen bond without the introduction of a steric collision (Figure 8B). Accordingly, the S158V single mutant (as well as all known double mutants containing this mutation) is constitutively fluorescent (*28*, *54*).

*Position 199.* While the choice of amino acid at positions 143 and 158 can give rise to both positive and negative interactions, the interactions of the chromophore with the amino acid at position 199 appear to be exclusively repulsive. In all structures of GFP-related proteins, this amino acid is a branched chain aliphatic residue. In KFP, this residue is an isoleucine, while in the remaining proteins, it is a leucine. When leucine occupies this position, the aliphatic side chain is directed away from the chromophore (Figure 9). In KFP, however, the isoleucine at this position is too bulky to adopt the same orientation (Figure 7). In this case, the C$\delta$ atom of Ile199 assumes a position that would result in a steric clash with the chromophore in the cis conformation and is presumably the basis for stabilization of the trans isomer in KFP.

There are many potential applications for a photoswitchable fluorescent label in biotechnology (*27*). In addition, the mechanism underlying the transition between fluorescent and nonfluorescent states is of fundamental interest to spectroscopists. The crystal structure of the dark state of KFP described in this report gives clear indications for the structural basis for the kindling phenomenon and points the way for future investigations. Site-directed mutagenesis combined with crystal structure analysis of the light-activated state of KFP, currently underway, should provide very valuable insight into the structural basis for this most unusual phenomenon.

## NOTE ADDDED IN PROOF

Recently a 2.1-Å resolution structure of KFP in a different crystal form appeared (PDB entry 1XQM) (*59*). Aside from minor differences in the conformations of surface residues, the models are very similar (rmsd 0.46 Å for residues 4-232). An important exception is that His197 was modeled in what we have identified as the minor conformation. The significance of this is unclear as the deposited model reveals a void where the alternate conformation of His197 would be located. We could not assess the possibility that His197 is partially disordered in this structure as the observed structure factors were not deposited.

## ACKNOWLEDGMENT

We thank Tad Spencer and Kristin Anderson at the University of Oregon and Spencer Anderson and Vukica

Srajer at BioCARS, Argonne National Laboratory, for help with crystallization and data collection. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract W-31-109-Eng-38. Use of BioCARS Sector 14 was supported by the National Institutes of Health, National Center for Research Resources, under Grant RR07707.

## SUPPORTING INFORMATION AVAILABLE

Structure-based sequence alignment of KFP homologues (Figure S1) and contact distances for observed and modeled chromophore conformations in KFP (Table S1). This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES

1. Pierce, D. W., Hom-Booher, N., and Vale, R. D. (1997) Imaging individual green fluorescent proteins, *Nature 388*, 338.
2. Dickson, R. M., Cubitt, A. B., Tsien, R. Y., and Moerner, W. E. (1997) On/off blinking and switching behaviour of single molecules of green fluorescent protein, *Nature 388*, 356−358.
3. Haupts, U., Maiti, S., Schwille, P., and Webb, W. W. (1998) Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy, *Proc. Natl. Acad. Sci. U.S.A. 95*, 13573−13578.
4. Schwille, P., Kummer, S., Heikal, A. A., Moerner, W. E., and Webb, W. W. (1999) Fluorescence correlation spectroscopy reveals fast optical excitation-driven intramolecular dynamics of yellow fluorescent proteins, *Proc. Natl. Acad. Sci. U.S.A. 97*, 151−156.
5. Lounis, B., Deich, J., Rosell, F. I., Boxer, S. G., and Moerner, W. E. (2001) Photophysics of DsRed, a red fluorescent protein, for the ensemble to the single-molecule level, *J. Phys. Chem. B 105*, 5048−5054.
6. Heikal, A. A., Hess, S. T., and Webb, W. W. (2001) Multiphoton molecular spectroscopy and excited-state dynamics of enhanced green fluorescent protein (EGFP): Acid−base specificity, *Chem. Phys. 274*, 37−55.
7. Nifosi, R., Ferrari, A., Arcangeli, C., Tozzini, V., Pellegrini, V., and Beltram, F. (2003) Photoreversible Dark State in a Tristable Green Fluorescent Protein Variant, *J. Phys. Chem. B 107*, 1679−1684.
8. Pierce, D. W., and Vale, R. D. (1999) Single-molecule fluorescence detection of green fluorescent protein and application to single-protein dynamics, *Methods Cell Biol. 58*, 49−73.
9. Creemers, T. M. H., 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.
10. Cinelli, R. A. G., Ferrari, A., Pellegrini, V., Signorelli, A., Tyagi, M., Giacca, M., and Beltram, F. (2001) Green fluorescent proteins as optically controllable elements in bioelectronics, *Appl. Phys. Lett. 79*, 3353−3355.
11. Moerner, W. E. (2002) Single-molecule optical spectroscopy of autofluorescent proteins, *J. Chem. Phys. 117*, 10925−10937.
12. He, X., Bell, A. F., and Tonge, P. J. (2003) Ground-state isomerization of a model green fluorescent protein chromophore, *FEBS Lett. 549*, 35−38.
13. Schlichting, I., and Berendzen, J. (1997) Out of the blue: The photocycle of the photoactive yellow protein, *Structure 5*, 735−739.
14. Genick, U. K., Borgstahl, G. E. O., Kingman, N., Ren, Z., Pradervanc, C., Burke, P. M., Srajer, V., Teng, T.-Y., Schildkamp, W., McRee, D. E., Moffat, K., and Getzoff, E. D. (1997) Structure of a protein photocycle intermediate by millisecond time-resolved crystallography, *Science 275*, 1471−1475.
15. Perman, B., Srajer, V., Ren, Z., Teng, T., Pradervand, C., Ursby, T., Bourgeois, D., Schotte, F., Wulff, M., Kort, R., Hellingwerf, K., and Moffat, K. (1998) Energy transduction on the nanosecond time scale: Early structural events in a xanthopsin photocycle, *Science 279*, 1946−1950.
16. Anderson, S., Srajer, V., and Moffat, K. (2004) Structural heterogeneity of cryotrapped intermediates in the bacterial blue light photoreceptor, photoactive yellow protein, *Photochem. Photobiol. 80*, 7−14.
17. 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.
18. Wiedenmann, J., Schenk, A., Rocker, C., Girod, A., Spindler, K. D., and Nienhaus, G. U. (2002) A far-red fluorescent protein with fast maturation and reduced oligomerization tendency from *Entacmaea quadricolor* (Anthozoa, Actinaria), *Proc. Natl. Acad. Sci. U.S.A. 99*, 11646−11651.
19. Lukyanov, K. A., Fradkov, A. F., Gurskaya, N. G., Matz, M. V., Labas, Y. A., Savitsky, A. P., Markelov, M. L., Zaraisky, A. G., Zhao, X., Fang, Y., Tan, W., and Lukyanov, S. A. (2000) Natural animal coloration can be determined by a nonfluorescent green fluorescent protein homolog, *J. Biol. Chem. 275*, 25879−25882.
20. Beddoe, T., Ling, M., Dove, S., Hoegh-Guldberg, O., Devenish, R. J., Prescott, M., and Rossjohn, J. (2003) The production, purification and crystallization of a pocilloporin pigment from a reef-forming coral, *Acta Crystallogr. D59*, 597−599.
21. 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.
22. 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.
23. 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.
24. 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.
25. 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.
26. Verkhusha, V. V., and Lukyanov, K. A. (2004) The molecular properties and applications of Anthozoa fluorescent proteins and chromoproteins, *Nat. Biotechnol. 22*, 289−196.
27. Chudakov, D. M., Belousov, V. V., Zaraisky, A. G., Novoselov, V. V., Staroverov, D. B., Zorov, D. B., Lukyanov, S., and Lukyanov, K. A. (2003) Kindling fluorescent proteins for precise in vivo photolabeling, *Nat. Biotechnol. 21*, 191−194.
28. Chudakov, D. M., Foefanov, A. V., Mudrik, N. N., Lukyanov, S., and Lukyanov, K. A. (2003) Chromophore environment provides clue to "kindling fluorescent protein" riddle, *J. Biol. Chem. 278*, 7215−7219.
29. 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.
30. 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.
31. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) Rapid automated molecular replacement by evolutionary search, *Acta Crystallogr. D55*, 484−491.
32. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography and NMR Systems: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
33. 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. A43*, 489−503.
34. Sheldrick, G. M., and Schneider, T. R. (1997) SHELXL: High-resolution refinement, *Methods Enzymol. 277*, 319−343.
35. 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. A47*, 110−119.

36. Connolly, M. L. (1993) The molecular surface package, *J. Mol. Graphics 11*, 139−141.
37. Kraulis, P. J. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−950.
38. 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.
39. Yang, F., Moss, L. G., and Phillips, G. N. J. (1996) The molecular structure of green fluorescent protein, *Nat. Biotechnol. 14*, 1246−1251.
40. 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.
41. 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.
42. 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.
43. 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.
44. Drew, M. G. B., Mok, F. K., Ang, K. P., and Tan, S. F. (1987) Structure of the Z isomer of 5-[(4-methoxyphenyl)methylene]-imidazolidine-2,4-dione, *Acta Crystallogr. A47*, 392−400.
45. Kurimoto, M., Subramony, P., Gurney, R. W., Lovell, S., Chmielewksi, J., and Kahr, B. (1999) Kinetic stabilization of biopolymers in single-crystal hosts: Green fluorescent protein in α-lactose monohydrate, *J. Am. Chem. Soc. 121*, 6952−6953.
46. Weber, W., Helms, V., McCammon, J. A., and Langhoff, P. W. (1999) Shedding light on the dark and weakly fluorescent states of green fluorescent proteins, *Proc. Natl. Acad. Sci. U.S.A. 96*, 6177−6182.
47. Wachter, R. M., Elsliger, M. A., Kallio, K., Hanson, G. T., and Remington, S. J. (1998) Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein, *Structure 6*, 1267−1277.
48. Zagranichny, V. E., Rudenko, N. V., Gorokhovatsky, A. Y., Zakharov, M. V., Shenkarev, Z. O., Balashova, T. A., and Arseniev, A. S. (2004) zFP538, a yellow fluorescent protein from coral, belongs to the DsRed subfamily of GFP-like proteins but possesses the unexpected site of fragmentation, *Biochemistry 43*, 4764−4772.
49. Martynov, V. L., Maksimov, B. I., Matynova, N. Y., Pakhomov, A. A., Gurskaya, N. G., and Lukyanov, S. A. (2003) A purple-blue chromoprotein from *Goniopora tenuidens* belongs to the DsRed subfamily of GFP-like proteins, *J. Biol. Chem. 278*, 46288−46292.
50. Mizuno, H., Mal, T. K., Tong, K. I., Ando, R., Furuta, T., Ikura, M., and Miyawaki, A. (2003) Photo-induced peptide cleavage in the green-to-red conversion of a fluorescent protein, *Mol. Cell 12*, 1051−1058.
51. Shagin, D. A., Barsova, E. V., Yanushevich, Y. G., Fradkov, A. F., Lukyanov, K. A., Labas, Y. A., Semenova, T. N., Ugalde, J. A., Meyers, A., Nunez, J. M., Widder, E. A., Lukyanov, S. A., and Matz, M. V. (2004) GFP-like proteins as ubiquitous metazoan superfamily: Evolution of functional features and structural complexity, *Mol. Biol. Evol. 21*, 841−850.
52. Cianci, M., Rizkallah, P. J., Olczak, A., Raftery, J., Chayen, N. E., Zagalsky, P. F., and Helliwell, J. R. (2002) The molecular basis of the coloration mechanism in lobster shell: β-Crustacyanin at 3.2-Å resolution, *Proc. Natl. Acad. Sci. U.S.A. 99*, 9795−9800.
53. Zagalsky, P. F. (2003) β-Crustacyanin, the blue-purple carotenoprotein of lobster carapace: Consideration of the bathochromic shift of the protein-bound astaxanthin, *Acta Crystallogr. D59*, 1528−1531.
54. Bulina, M. E., Chudakov, D. M., Mudrik, N. N., and Lukyanov, K. A. (2002) Interconversion of Anthozoa GFP-like fluorescent and nonfluorescent proteins by mutagenesis, *BMC Biochem. 3*, 7.
55. Kleywegt, G. J., and Jones, T. A. (1997) Detecting folding motifs and similarities in protein structures, *Methods Enzymol. 277*, 525−545.
56. Barton, G. J. (1993) ALSCRIPT a tool to format multiple sequence alignments, *Protein Eng. 6*, 37−40.
57. Esnouf, R. M. (1999) Further additions to MolScript version 1.4, including reading and contouring of electron-density maps, *Acta Crystallogr. D55*, 938−940.
58. Merritt, E. A., and Bacon, D. J. (1997) Raster3D: Photorealistic molecular graphics, *Methods Enzymol. 277*, 505−524.
59. Wilmann, P. G., Petersen, J., Devenish, R. J., Prescott, M., and Rossjohn, J. (2005) Variations on the GFP chromophore: 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.

BI047644U