5276 *Biochemistry* **2009,** *48,* 5276–5283
DOI: 10.1021/bi801658p

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Crystal Structure of a New Cyan Fluorescent Protein and Its Hue-Shifted Variants[†,‡]

Akihiro Kikuchi,*[,§] Eiko Fukumura,[§,∥] Satoshi Karasawa,[⊥,#,△] Yoshitsugu Shiro,[§] and Atsushi Miyawaki[⊥]

[§]*Biometal Science Laboratory, RIKEN SPring-8 Center, 1-1-1, Kouto, Sayo, Hyogo 679-5148, Japan,* [∥]*Department of Life Science, Graduate School of Science, Himeji Institute of Technology/University of Hyogo, 3-2-1, Kouto, Kamigori, Ako, Hyogo 678-1297, Japan,* [⊥]*Laboratory for Cell Function and Dynamics, Advanced Technology Development Group, Brain Science Institute, RIKEN, 2-1, Hirosawa, Wako, Saitama 351-0198, Japan,* [#]*Amalgaam Co. Ltd., 2-9-3 Itabashi, Itabashi-ku, Tokyo 173-0004, Japan, and* [△]*Medical and Biological Laboratories Co. Ltd., 3-5-10 Marunouchi, Naka-ku, Nagoya, Aichi 460-0002, Japan*

Received August 31, 2008; Revised Manuscript Received March 12, 2009

ABSTRACT: Green fluorescent protein (GFP) based techniques are well established in molecular biology; however, the detailed mechanism for the fine-tuning of fluorescent colors remains unclear. Here, we report the cloning and crystal structure of a new cyan-emitting GFP-like protein, KCy. We also developed a mutant protein with a high folding efficiency (KCy-G4219: λ_{abs} = 453 nm; λ_{em} = 486 nm). X-ray diffraction analysis revealed that the KCy chromophore is formed from an internal Ser62-Tyr63-Gly64 tripeptide. The serine residue at the first position of the chromophore-forming tripeptide has a short polar chain (−OH) that forms a noncovalent interaction with the His38 imidazole at a distance of 2.96 Å. Substitution of His38 in KCy-G4219 with Gln (KCy-R1) or Leu residues resulted in a slight but significant red shift of the emission peak maximum from 486 to 492 or 496 nm, respectively. The crystal structure of KCy-R1 determined at a resolution of 1.58 Å showed that the noncovalent interaction between Ser62-OH and the substituted Gln38 occurred over a longer distance (3.07 Å) than that observed in the wild-type KCy. Such an interaction is absent in the Leu mutant, suggesting that this interaction is one of the key factors responsible for fine-tuning the emission peak maxima, which are affected by chromophore polarization. Moreover, the structural comparison suggests that an additional water molecule buried in the space between the Ala158 residue and the chromophore phenolate is also responsible for the chromophore polarization.

Green fluorescent protein from *Aequorea victoria* (avGFP)[1] is a ubiquitous tool in the fields of molecular and cellular biology for the visualization of gene expression, protein localization, and protein−protein interactions in living cells (*1*−*4*). The chromophore of avGFP assembles autocatalytically from the internal tripeptide Ser65-Tyr66-Gly67, resulting in the formation of a 4-(*p*-hydroxybenzylidene)-5-imidazolinone moiety (*5*, *6*). Since avGFP does not require external substrates or cofactors for the observation of its fluorescence, it has been possible to develop spectral variants by direct mutagenesis. The range of available colors has been further expanded with the discovery of GFP-like proteins mainly from the Anthozoa species (*7*−*9*).

Structural information from the avGFP variants and GFP-like proteins characterized to date has revealed an important correlation between the chromophore structure and fluorescence (*10*−*20*). Efficient fluorescence exclusively requires a coplanar *cis* conformation of the chromophore (*11*). Fluorescence is also dependent on the extension of the π-electron system of the chromophore; for example, an expanded conjugated π-electron system through the addition of an acylimine moiety results in a red shift, as found in DsRed (*12*).

More complicated factors are also thought to contribute to the tuning of the spectroscopic properties of the GFP-like protein family (*10*). It is well-known that excitation of the chromophore is accompanied by an intramolecular charge transfer (ICT) from the phenolate ring to the imidazolidinone ring that comprises the chromophore (*13*). Accordingly, the electronic polarization and dipole moment of the chromophore in the ground state are likely to be key factors for fine-tuning the spectroscopic properties of the proteins. In the case of cyan-emitting amFP486 from *Anemonia majano*, the additional water molecule buried in the space between the Ala165 residue and the chromophore phenolate is thought to play an important role in causing the blue shift relative to avGFP by increasing the polarization of the chromophore (*14*, *21*). Although some mechanisms of spectroscopic tuning are currently being elucidated, it remains a challenge to control such factors as chromophore polarization and to construct new *in vivo* marker proteins with desirable colors by rational mutagenesis. In order to rationally design desirable fluorescent proteins, it is imperative to obtain more three-

[†]This work was supported by grants from the Molecular Ensemble Program at RIKEN.

[‡]The atomic coordinates and structural factors have been deposited in the Protein Data Bank (entry 2ZO6 for wild-type KCy and 2ZO7 for KCy-R1).

*To whom correspondence should be addressed. Phone: +81-791-58-2817. Fax: +81-791-58-2818. E-mail: kikuchi@spring8.or.jp.

[1]Abbreviations: λ_{abs}, peak maxima of absorption spectrum; λ_{em}, peak maxima of emission spectrum; avGFP, green fluorescent protein from *Aequorea victoria*; ICT, intramolecular charge transfer; CFP, cyan-emitting fluorescent protein; BFP, blue-emitting fluorescent protein; KCy, Kusabira-Cyan.

 Published on Web 04/29/2009

dimensional (3D) structural information on GFP-like proteins, especially information relating to the structure and immediate environment of the chromophores. This is particularly important for blue-shifted GFP-like cyan-emitting fluorescent proteins (CFPs) and blue-emitting fluorescent proteins (BFPs), which currently have limited availability of 3D structures (*14*−*16*), although CFPs are not always required for *in vivo* imaging due to the photophysical properties of living cells.

Here, we report the molecular cloning and crystal structure characterization of a new cyan-emitting GFP-like protein, KCy, from the *Fungia* species. We also propose a strategy for fine-tuning the spectroscopic properties of this protein at the molecular level and introduce mutations to affect chromophore polarization in order to alter the emission peak maxima.

EXPERIMENTAL PROCEDURES

cDNA Cloning, Gene Construction, Expression, and Mutagenesis. Total RNA isolation from *Fungia* and polymerase chain reaction (PCR) based amplification of cDNA were performed as described previously (*22*) using the degenerate primers 5′-GAAGGRTGYGTCAAYGGRCAY-3′ and 5′-ACVGGDCCATYDGVAAGAAARTT-3′. For bacterial expression of the protein, the cDNA was amplified using primers containing 5′ *Bam*HI and 3′ *Eco*RI sites; the amplified fragments were digested and cloned in-frame into the *Bam*HI/*Eco*RI site of the pRSET-B vector (Invitrogen, Carlsbad, CA), in which a 6 × His tag was introduced at the amino (N) terminus. The protein and its variants were expressed and purified as described previously (*22*).

Multiangle Light Scattering (MALS). MALS was measured with a multiangle light photometer (DAWN; Wyatt, Santa Barbara, CA) connected to a Shodex HPLC system (Showa Denko, Tokyo) on which size exclusion columns (tandem connection of KW-804 and KW-802.5; Showa Denko) were installed. The purified protein with N-terminal His tag was analyzed.

Crystallization and Data Collection. Purified wild-type KCy was concentrated by ultrafiltration to 10 mg mL^{-1} in 20 mM Tris-HCl (pH 7.9) and 0.2 M NaCl. Cyan fluorescent crystals were grown in sitting drops containing 1 μL of protein and 1 μL of mother liquid (20% (v/v) PEG 3350 and 0.2 M lithium chloride in 20 mM Tris-HCl (pH 8)) at room temperature for 7−10 days.

Removal of the His tag using enterokinase was performed prior to crystallization of KCy-R1. After purification by TARON and DEAE-Sepharose, the protein was concentrated to 12 mg mL^{-1} in 20 mM Tris−H_2SO_4 (pH 8), 5% glycerol, and 0.2 M Na_2SO_4. Crystals were obtained with 1 μL of protein added to 1 μL of mother liquid (35−40% PEG 8000, 0.2 M $(NH_4)_2SO_4$, and 0.1 M sodium cacodylate (pH 7.5)).

After soaking in a cryoprotectant, the crystals were flash-frozen at 100 K in a nitrogen stream before data collection. All diffraction data were collected at SPring-8 on the RIKEN beamline BL44B2. The data were merged and processed with the HKL2000 package (*23*). Both crystals KCy and KCy-R1 belonged to space group *C*2 with unit cell parameters as follows: a = 94.75 Å, b = 42.80 Å, c = 50.70 Å, and β = 96.28° for KCy; a = 96.69 Å, b = 45.02 Å, c = 51.16 Å, and β = 94.98° for KCy-R1. Both crystals had one protomer within the asymmetric unit.

Determination of Structure. The KCy structure was determined using the molecular replacement method with the CCP4 program suite Molrep (*24*). The structure of blue nonfluorescent pocilloporin (Protein Data Bank (PDB) entry 1MOU) was used as the initial phasing model with the deletion of the chromophore. Approximately 99% of the polypeptide chain was automatically traced into the electron density maps using ARP/wARP (*25*). The resulting chains were corrected and modified using TURBO-FRODO software (*26*). The model was refined using CNS version 1.1 (*27*). The electron density map permitted modeling of residues 2−218, which included the chromophore. The restraint parameters for the chromophore were generated by the PRODEG server (*28*). Water molecules were modeled into strongly different density peaks using CNS. Cycles of refinement, model rebuilding (including that of the chromophore), and the addition of solvent molecules culminated in a model of the KCy structure with an *R*-factor of 0.164 and a free *R*-factor of 0.181 at a resolution of 1.4 Å.

The KCy-R1 structure was determined using the molecular replacement method with the refined KCy model. The ARP/wARP program traced approximately 75% of the polypeptide chain into the electron density maps using the initial model. Refinement was carried out as outlined above. The final model of KCy-R1, which comprises residues 1−220 and two additional N-terminal residues derived from the vector, has an *R*-factor of 0.177 and a free *R*-factor of 0.211 to 1.58 Å resolution.

RESULTS AND DISCUSSION

KCy: A New CFP. Transformation of isolated *Fungia concinna* cDNA (cDNA) into *Escherichia coli* produced colonies with bright cyan fluorescence. The metal affinity chromatography-purified, His-tag-cleaved protein was named Kusabira-Cyan (KCy), since *Fungia* is known as *Kusabira-Ishi* in Japanese.

The sequence of KCy was aligned with those of CFPs and GFPs (Figure 1) and was found to have 56.7%, 51.9%, and 50.5% sequence identity with the sequences of amFP486 (*7*), mTFP1 (*15*), and MiCy (*22*), respectively. The alignment suggests that Ser62-Tyr63-Gly64 is a precursor of the fluorescent chromophore in KCy, in which the Tyr residue is conserved at the second position; a cyan-emitting avGFP variant was previously constructed by substitution of a Trp residue for Tyr at this position (*1*, *16*). The key residues responsible for chromophore formation (Arg91 and Glu208 in KCy numbering) (*7*, *12*) are highly conserved among the GFP-like proteins. Ala158 and His193 in KCy (red boxes in Figure 1) as well as Ala165 and His199 in amFP486 are conserved in both proteins; the latter two residues have been proposed to contribute to the blue shift and to efficient fluorescence in the emission spectra of amFP486 (*14*, *21*).

The spectral properties of recombinant KCy are compared with those of the other CFPs in Table 1. As shown in Figure 2a, KCy exhibits a major absorption maximum at 455 nm with an extinction coefficient of 38700 M^{-1} cm^{-1}. The shoulder around 400 nm corresponds to the neutral state of the chromophore (the phenol form). Thus, the spectrum clearly indicates that the KCy chromophore exists in an apparently ionized form in its dominant state. In the case of avGFP, excited-state proton transfer has been demonstrated experimentally. Since the excitation spectrum of KCy shows the neutral form of the chromophore, the neutral chromophore of KCy should change into an anion when excited. The details of excited-state proton transfer in KCy remain to be investigated. In the fluorescent spectrum, the emission peak maximum was observed at 488 nm (Figure 2b) with a fluorescence quantum yield (Φ) of 0.85. Both the absorption and emission spectra of KCy were pH-insensitive between pH 5 and pH 9, but the intensity of the fluorescence peak was reduced

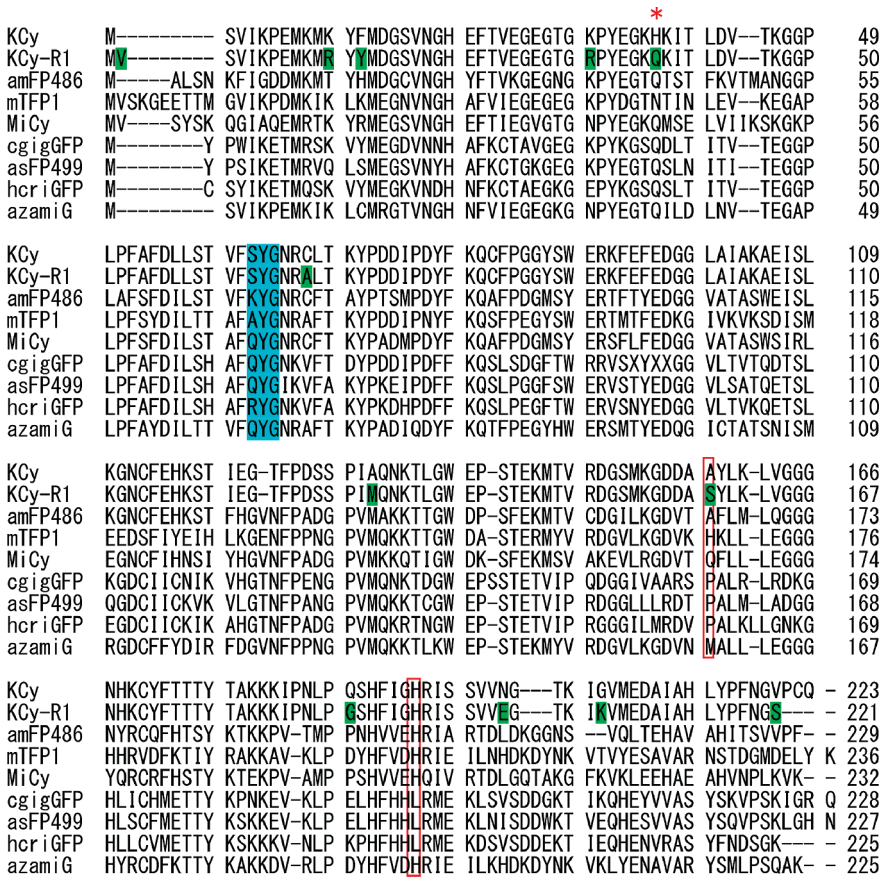

FIGURE 1: Amino acid sequence alignment of KCy, KCy-R1, cyan-emitting, and green-emitting GFP-like proteins. Sequences of green-emitting proteins are available from GenBank (cgigGFP cloned from *Calvatia gigantea*, GenBank accession no. AY037776; asFP499 cloned from *Anemonia sulcata*, GenBank accession no. AF322221; hcriGFP cloned from *Heteractis crispa*, GenBank accession no. AF420592; azamiG cloned from *Galaxea fascicularis*, GenBank accession no. AB107915). Chromophore-forming tripeptides are indicated by blue boxes. Substituted residues of KCy-R1 are indicated by green boxes. KCy residue corresponding to His38 is indicated by an asterisk. Positions of proposed key residues for cyan emission (Ala158 and His193 in KCy) are indicated by red boxes.

Table 1: Spectral Properties of CFPs

protein	amino acid no.	absorption/emission max (nm)	max extinction coeff (M^{-1} cm^{-1})	quantum yield	ref
KCy	223	455/488	38700	0.85	this work
KCy-G4219	221	453/486	21100	0.80	this work
KCy-G4219-38L	221	464/494			this work
KCy-R1	221	461/492	21200	0.75	this work
KCy-R1-158A	221	459/489			this work
KCy-R1-38H	221	459/489			this work
KCy-R1-38L	221	467/496	13000		this work
ECFP	238	435/478	28750	0.40	*21*
amFP486	229	458/486	40000	0.24	*7*
mTFP1	236	462/492	64000	0.85	*15*
MiCy	232	472/495	27250	0.90	*21*

below pH 5, probably due to protonation of the phenolate of the chromophore. These characteristics, namely, impressive brightness and stability against pH change, suggest that KCy has the potential to be a useful CFP and to act as a donor in fluorescence resonance energy transfer (FRET) techniques.

High-Folding Efficiency Mutant, KCy-G4219. We have rationally designed a variant that increased the folding efficiency of the protein. The scheme included the following: deletion of the C-terminal residues (221−223) that do not affect fluorescence (*vide infra*) in order to minimize the risk of aggregation; substitution of Ser for Val220 to increase solubility; introduction of four mutations (Lys11 → Arg, Phe13 → Tyr, Lys32 → Arg, and Gln187 → Lys) as well as two mutations based on a successful scheme used in Kusabira-Orange, which has 63.2% sequence identity with the KCy (Ala131 → Met, Asn200 → Glu) (*22*); and two additional mutations (Cys67 → Ala and Gly205 → Lys) to increase solubility. In addition, Val was inserted between Met1 and Ser2 due to the use of the Kozak sequence. It should be noted that "tetramer breaking" was not applied in this study because the dimerization appears to be an important feature of cyan-emitting KCy protein (*vide infra*). The resultant mutant was designated KCy-G4219. As shown in Figure 3, *E. coli* colonies expressing KCy-G4219 showed significantly brighter fluorescence than those expressing wild-type KCy after overnight

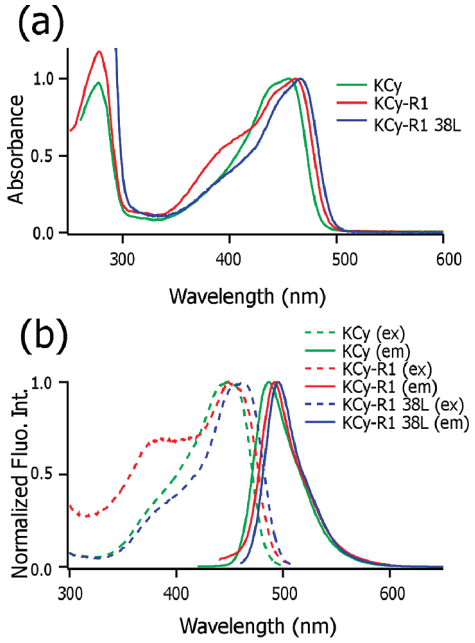

FIGURE 2: Spectroscopic properties of KCy (green), KCy-R1 (red), and KCy-R1-38L (blue). (a) Absorption spectra and (b) normalized excitation (broken line, abbreviated as ex) and emission spectra (solid line, abbreviated as em). Emission peak maxima are 488, 492, and 496 nm for KCy, KCy-R1, and KCy-R1-38L, respectively.

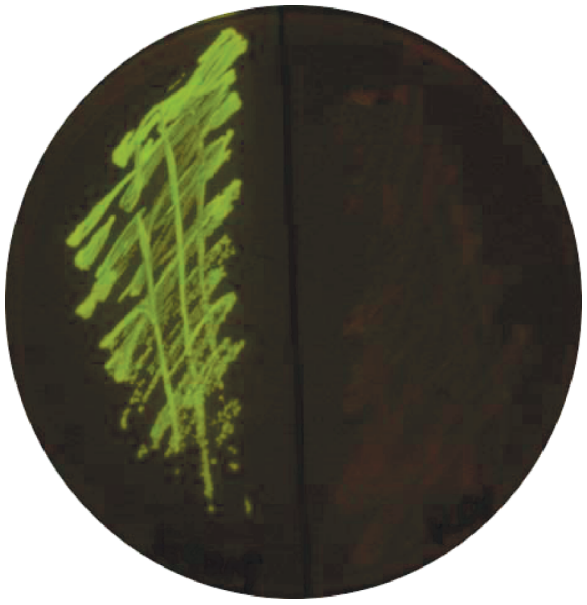

FIGURE 3: *E. coli* colonies expressing KCy (right) and KCy-G4219 (left) after overnight incubation at 37 °C, examined under an ultraviolet (UV) illuminator.

incubation at 37 °C. Therefore, it appeared that the maturation of mutated KCy, namely, KCy-G4219, was faster than that of the wild-type KCy. The variant exhibited spectral properties similar to those of the wild-type KCy (Table 1 and see Supporting Information Figure S1), although it showed a decrease in the extinction coefficient to 21100 M^{-1} cm^{-1} and a slight decrease in the fluorescence quantum yield (0.80). Nevertheless, with regard to brightness in terms of the multiplication of the molar extinction coefficient and quantum yield (Table 1), the fluorescence of KCy-G4219 was clearly brighter than those of the other CFPs such as amFP486 (*7*). Thus, KCy-G4219 appears to be a promising candidate for use as a donor in FRET measurements.

Table 2: Crystallographic Statistics for KCy Proteins

	KCy	KCy-R1
PDB ID	2ZO6	2ZO7
	Data Collection	
space group	*C*2	*C*2
unit cell		
a/*b*/*c* (Å)	94.75/42.80/50.70	96.69/45.02/51.16
β (deg)	96.28	94.98
beamline	SPring-8 BL44B2	SPring-8 BL44B2
wavelength (Å)	1.0	1.0
overall resolution (Å)		
(outer shell)	50−1.40 (1.45−1.40)	50−1.58 (1.64−1.58)
measured reflections	127366	104654
unique reflections	36604	29798
completeness (%)	91.6 (62.5)	98.0 (91.7)
$\langle I/\sigma(I)\rangle$	28.7 (2.8)	34.7 (5.3)
R_{merge} (%)	3.6 (21.8)	6.0 (25.7)
	Refinement	
resolution range (Å)	20−1.4	20−1.58
reflections used	36588	29784
R_{cryst} (%)	16.4	17.7
R_{free} (%)	18.1	21.1
water atom	315	319
rmsd bond length (Å)	0.013	0.009
rmsd bond angles (deg)	1.7	1.6
average *B*-factor (Å)		
protein	16.7	21.0
water	30.9	34.5
Ramachandran plot		
most favorable (%)	92.7	93.1
additionally allowed (%)	7.3	6.9

Overall KCy Structure. We determined the crystal structure of the wild-type KCy at a resolution of 1.4 Å by the molecular replacement technique. The final crystallographic statistics are given in Table 2. The structure of the KCy protomer was similar to the previously determined structures of avGFPs and GFP-like proteins, which are characterized by an 11-stranded β-barrel fold with a central α-helix holding the chromophore (*1*, *5*, *12*). Comparison of the overall structure of KCy with that of amFP486 (PDB entry 2A46) and mTFP1 (PDB entry 2HQK) yielded a root-mean-square (rms) deviation of 1.0 Å for 192 equivalent Cα atoms and 0.83 Å for 190 equivalent Cα atoms, respectively, which is consistent with the high sequence identity among the three CFPs (*14*, *15*).

Although one protomer of KCy is present in an asymmetric unit, a dimer structure can be generated by crystallographic symmetry (Figure 4a). The dimer interface is mediated by hydrogen bonds and hydrophobic interactions at the surface. In addition, the carboxy (C) terminus (residues 209−217) of one protomer interacts with the C-terminus of the other protomer (Figure 4b), and the distances between each chain are longer than 3.5 Å in the crystal structure. For example, the distance between the nitrogen atom in the His213 side chain and the oxygen atom in the Tyr215 side chain was estimated to be 4.35 Å. Despite such long-range weak interactions, the interaction via the C-termini appears to be crucial for the fluorescent properties of KCy, since deletion of the C-terminal residues, namely, Pro216, Phe217, and Asn218, converts KCy into a nonfluorescent protein (data not shown). Deletion of residues from Gly219 to Gln223 did not affect the fluorescence properties, and thus the C-terminal residues (221−223) were deleted in the construction of KCy-G4219 (*vide supra*).

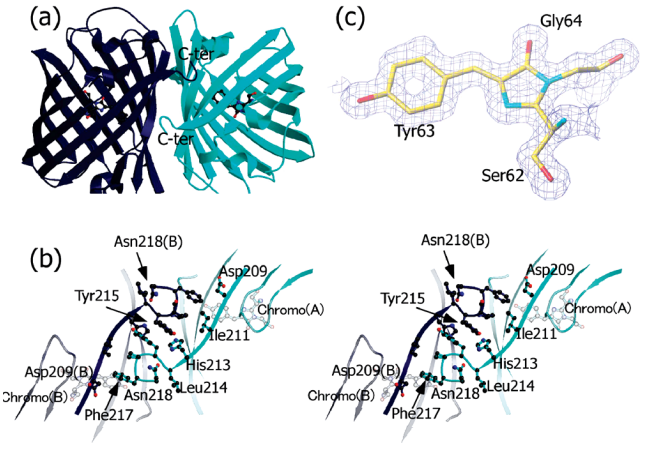

FIGURE 4: Crystal structure of the wild-type KCy. (a) Overall dimeric structure of KCy generated by crystallographic symmetry. Protomers are represented in different colors, and chromophores are shown as ball-and-stick representations. (b) Stereoview of observed dimer interface near the C-termini. (c) Final $2F_o - F_c$ electron density (contoured at 1.2σ) superimposed onto the KCy chromophore formed from the internal Ser-Tyr-Gly tripeptide. The chromophore is represented as a stick model in which carbon, oxygen, and nitrogen are shown as yellow, red, and blue, respectively.

These crystallographic observations are consistent with the findings of the MALS analysis, in which the absolute molecular mass of KCy (56 kDa) was 2.18 times larger than that deduced from the primary structure of the protein (25.7 kDa) (Supporting Information Figure S2). Thus, it appears that the protein is present as a dimer in solution as well as in the crystal form. On the basis of the crystal structure, the accessible surface areas for the KCy monomer and dimer are estimated to be 9960 and 17830 Å^2, respectively, and the buried surface area is calculated to be 1045 Å^2 per monomer. This value is smaller than that of DsRed (1360 Å^2 at the AC interface) (*12*), indicating that the dimer interaction is weaker, which is in agreement with the presence of long-range weak interactions at the dimer interface of the C-terminus. It is likely that weak dimerization is important for the formation of the chromophore and its environment in KCy. Although most of the GFP-like proteins characterized thus far form obligate tetramers, the formation of dimers has been confirmed for members of the GFP-like protein family, such as MiCy (*Acropora*) (*22*) and HcRed (*Heteractis*) (*19*, *29*), as well as for GFP from the sea pansy (*Renilla*) (*30*). This dimer/tetramer formation might reflect different functions and/or evolution of the protein families.

Structure of the KCy Chromophore. The electron density map of KCy was of sufficient quality to enable direct interpretation of the chromophore structure and its environment (Figure 4c). This showed that the chromophore is formed by the autocatalytic cyclization of the Ser62-Tyr63-Gly64 tripeptide, as expected from the sequence alignment (Figure 1). The chromophore possesses a 4-(*p*-hydroxybenzylidene)-5-imidazolinone group with a coplanar *cis* conformation, and the peptide bond between Phe61 and Ser62 has normal sp^3 geometry. Interestingly, the structural study revealed that the chromophore is formed from the same tripeptide as that of the wild-type avGFP. It should be noted that almost all of the CFPs characterized so far possess a residue with a longer side chain such as Lys in amFP486 at the first position of the chromophore-forming tripeptide. Thus, structural differences in noncovalent interactions between the chromophore and its surroundings, including not only hydrogen bonds but also others such as van der Waals' forces, appear to contribute to differences in chromophore polarization, resulting in the blue shift of KCy (cyan-emitting) relative to avGFP (green-emitting). In particular, the absence of a hydrogen bond between the Ser62-OH in the chromophore and the well-conserved Glu208 in KCy is distinct from avGFP, where the Ser65-OH in the chromophore directly forms a hydrogen bond with well-conserved Glu222.

KCy Chromophore Environment and Structural Basis for the Blue Shift. The chromophore environments of KCy and amFP486 are illustrated in panels a and c of Figure 5, respectively.

We observed two important similarities between the environments of KCy and amFP486, which would affect the blue shift of the two CFPs relative to avGFP. The first similarity was the presence of His193 in KCy, which lies in direct contact with the phenolate moiety of the chromophore, resulting in the formation of a $\pi-\pi$ stacking interaction (*31*); this interaction is absent in avGFP since the His residue is replaced with Thr203. It is noteworthy that, in KCy, His193 forms an intricate charge network in the chromophore cavity: the two imidazole nitrogen atoms interact with carboxylates from Glu143 (2.75 Å) and Glu208 (2.65 Å), and the interaction extends to Arg66 and a buried water molecule (W1 in Figure 5a) to achieve a cyclic interaction, indicating that the imidazole of His193 could be positively charged through the charge network. A similar "quadrupole" network using the stacking His residue was observed not only in amFP486 (Figure 5c) but also in mTFP1, implying that these structural characteristics would be related to the cyan-emitting properties of the GFP-like proteins. Here, it can be proposed that the quadrupole network including the $\pi-\pi$ stacking interaction between the phenolate moiety of the chromophore and the imidazole ring of His193 serves to stabilize the anionic phenolate form of the chromophore. As a result, the ICT in the excited state requires higher energy and thereby gives rise to the blue shift. This proposal is supported by our mutagenesis studies using colony-based fluorescence imaging, which showed that His193 was always conserved in all KCy mutants that emitted bright fluorescence (data not shown). In the case of amFP486, the substitution of Thr for His199, corresponding to His193 in KCy, resulted in a dramatic lowering of the quantum yield, and its crystal structure revealed a *cis* and *trans* chromophore conformation in a 1:1 ratio (*14*). It should be noted that similar quadupole networks have been observed in the crystal structures of not only CFPs but also a green-emitting protein showing photochromism, namely, Dronpa (*20*), as well as a yellow- and an orange-emitting GFP-like proteins (*17*, *18*). The quadrupole network may cause these proteins to have blue-shifted peaks in their spectra relative to the chromophores themselves.

The second important similarity between KCy and amFP486 is the hydrogen-bonding interaction between the phenolate oxygen atom and nearby residues. As shown in Figure 5a, Ser141 and two water molecules (W2 and W3) contribute to the stabilization of the anionic phenolate form in the KCy chromophore through hydrogen bonding. In comparison with avGFP, a presence of a water molecule, W3 in Figure 5a, in the cavity between Ala158 and the chromophore in KCy is particularly important for the polarization of the chromophore, which leads to the blue shift. This finding is consistent with the observations in amFP486 (*14*, *21*), although the interactions are not conserved in mTFP1 (*15*). In the case of mTFP1, the phenolate oxygen atom makes contact with the imidazole ring of

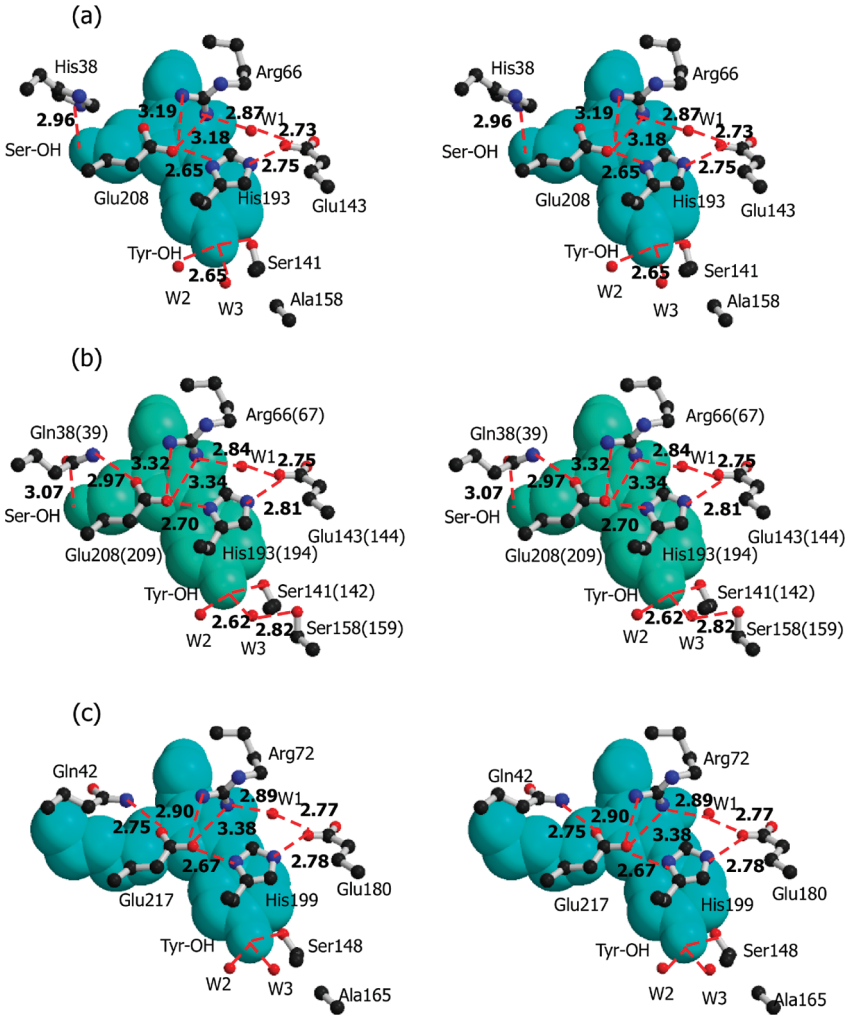

FIGURE 5: Stereoview of the immediate chromophore environment of (a) KCy, (b) KCy-R1, and (c) amFP486 (PDB entry 2A46). Noncovalent interactions between the chromophore and nearby residues are shown as a red-dotted line with the distance in angstroms. KCy-R1 sequence numbers in the PDB file (PDB ID 2ZO7) are shown in parentheses.

His163, which exactly corresponds to Ala158 in KCy (red box in Figure 1). With regard to position 158, it should be noted that this position is occupied by a hydrophobic Pro or Met residue in green-emitting proteins, as shown in the alignment in Figure 1. However, the additional water molecule in the cavity seems to not always be required for cyan emission, since bulkier residues His and Gln are located at this position in mTFP1 and MiCy (Figure 1), respectively. Therefore, it may be interesting to determine the crystal structure of MiCy to further explore the relationship between 3D structure and cyan-emitting color.

Hence, we concluded that the combination of the structural factors mentioned above contributes to the stabilization of negative charge localization, i.e., polarization, in the phenolate moiety. The resulting chromophore polarization leads to a greater energy requirement for ICT in the excited state than that of avGFP, which contributes to the blue shift of the fluorescent peak.

Fine-Tuning of the KCy Emission Peak Maxima. In the first residue forming the chromophore, the direction of the side chain of KCy, namely, the hydroxyl group (−OH), differs from that of avGFP. While Ser65 of avGFP interacts with the well-conserved Glu222 residue (*5*, *13*), Ser62-OH of KCy forms a noncovalent interaction only with His38, which corresponds to Leu42 of avGFP, with a distance of 2.96 Å (Figure 5a). To test whether the noncovalent interaction between Ser62-OH and His38 affects the spectroscopic properties, we introduced two mutations into KCy-G4129:[2] His38 → Gln and Ala158 →Ser. The variant, named KCy-R1, showed a major absorption maximum at 461 nm (ε = 21200 M^{-1} cm^{-1}) with a shoulder at approximately 400 nm (Figure 2a) and a slight but significant red shift of the fluorescent maximum (λ_{em} = 492 nm) in the emission spectra compared with KCy-G4219 (λ_{em} = 486 nm; Figure 2b).

To examine the mutational effect on the structure, we determined the crystal structure of KCy-R1 (Table 2) and observed that the overall fold was identical to that of the wild-type KCy (rms deviation of 0.28 Å for all Cα atoms; data not shown). The immediate environment of the KCy-R1 chromophore (Figure 5b) was then compared with those of KCy (Figure 5a) and amFP486 (Figure 5c).

Position 38. We first assessed the structure of the mutated His38 → Gln region. On the basis of the structural comparison between KCy and KCy-R1, we found two subtle structural differences among them. First, the Ser62-OH establishes a contact with the replaced Gln residue of KCy-R1, which is

[2]Sequence numbers of KCy-G4129 and KCy-R1 are increased by one compared to that of wild-type KCy due to the insertion of the Val residue between Met1 and Ser2. Thus, Ser63 of KCy-G4219 and KCy-R1 corresponds to Ser62 of the wild-type KCy and is referred to as Ser62. In this paper, amino acid sequence numbers are based on that of the wild type for simplicity.

similar to KCy. Although no significant difference was found for the directions of the side chains of Ser62 as shown in Figure 5a,b, the distance between Ser62-OH and the Gln residue in KCy-R1 increased to 3.07 Å, while that between the His38 residue and Ser62-OH in KCy was 2.96 Å. This difference suggests that the noncovalent interaction between Ser62 and position 38 in KCy-R1 would be weaker than that in KCy. Thus, it would alter the electronic polarization and/or dipole moment of the chromophore. Brejc and colleagues proposed that the direction and strength of the dipole moment of the chromophore Ser-OH affect its spectroscopic properties (*13*). They compared the absorption spectra of avGFP (λ_{abs} = 475 nm) and its Ser65Thr mutant (λ_{abs} = 489 nm) and explained the 14 nm red shift as a "dipole moment effect". In avGFP, the dipole moment of Ser65-OH is roughly parallel to the long axis of the chromophore and points in the direction of the phenolate, which will electrostatically increase the energy required for ICT. In the Ser65Thr mutant, however, the dipole moment is roughly perpendicular to the long axis of the chromophore, which has little electrostatic interaction. This proposal provides evidence that subtle structural differences in the chromophore and its surrounding are coupled to the electronic polarization and/or dipole moment of the chromophore. Consequently, the structural difference between Ser62 and position 38 could be a key factor that accounts for the red shift.

Second, the substituted Glu at position 38 forms a hydrogen bond with Glu208 at a distance of 2.97 Å in KCy-R1, while His38 in KCy is far away from Glu208 at a distance of 3.62 Å even if we assume the imidazole ring to be flipped by 180°. This difference would substantially affect the charge on the quadrupole network, resulting in the red shift. The 12 nm red shift observed in the mTFP1 Thr73Ala variant was interpreted by such mechanism. Namely, the Thr73 is hydrogen-bonded to Arg70, which is in the quadrupole netwok. However, such hydrogen bonding should be absent in the Thr73Ala variant, although its crystal structure was not available (*32*).

Both of the structural differences mentioned above are possible factors for the fine-tuning of the KCy protein. To gain insight into the role of position 38, we introduced a Gln38 → X mutation into KCy-R1, where X was Asp, Phe, Asn, Ser, Thr, Met, or Leu, and investigated the spectroscopic properties of the proteins. The variants could be classified into three classes: (1) In the case of the Asp and Phe variants, the proteins were colorless and nonfluorescent (data not shown), indicating that residues with anionic character and bulky hydrophobic side chains were unfit for position 38. Such residues seemed to prevent the autocatalytic formation of the chromophore and/or to destabilize the protein. (2) The Asn, Ser, Thr, and Met variants demonstrated fluorescence with each having similar properties in their fluorescence spectra; i.e., these variants had a peak maximum at around 493 nm (shown in Supporting Information Figure S3). This second class of variants, with the exception of Met, had polar residues at position 38. Thus, the residues were capable of forming electrostatic interactions between Ser62 and/or Glu208 as found in the KCy-R1 (38Gln) structure. However, because of their shorter side chains relative to Gln, the interaction would be weaker than in KCy-R1. The fact that the ratio between the neutral (400 nm) and anionic (around 460 nm) chromophores varied with this mutation in the absorption and excitation spectra might reflect this weaker interaction. Thus, position 38, in any proposed mechanism, should influence the electronic polarization and/or dipole moment of the chromophore in the ground state. Further high-resolution structural studies will be required to gain a clearer understanding of the interactions in these variants since not only distance but also geometry are important for discussion of noncovalent interactions. (3) The Leu variant (KCy-R1-38L), in which the substituted Leu is unable to form noncovalent interactions such as hydrogen bonds with neither the Ser62 nor Glu208 residues and a packing energy of such environment may also increase, had emitting ability and exhibited a red-shifted peak at 496 nm in the emission spectrum (Figure 2b). However, the value of the molar extinction coefficient of KCy-R1-38L was lower than that of KCy-R1 (Table 1). Furthermore, we also introduced a His38 → Leu mutation into KCy-G4219. In this case, Ala158 was unaltered. This mutation also resulted in a red shift of the fluorescent peak maximum (λ_{em} = 494 nm) compared with that of KCy-G4219 (KCy-G4219-38L in Table 1 and in Supporting Information Figure S1). Consequentially, we can suggest that position 38 of KCy should have influence on the polarization and/or charge of the chromophore, which are detectable in absorption and fluorescence spectra.

Position 158. As mentioned earlier, Ala158 in KCy permits a water molecule (W3) to interact with the phenolate of the chromophore at a distance of 2.65 Å (Figure 5a). We expected that the mutation of Ala158 to Ser would result in the obstruction of the cavity where the water molecule is positioned, resulting in a red shift in the spectrum. Contrary to our expectations, the water molecule W3 remained in its original position in the crystal structure of KCy-R1 (Figure 5b). Due to its structure, the side chain of the substituted Ser158 is unable to fully occupy the cavity. Nevertheless, an additional hydrogen bond was formed between Ser158 and the buried water, W3 at the distance of 2.82 Å in KCy-R1 (Figure 5b). This additional interaction should weaken the hydrogen bond between W3 and the chromophore phenolate. Accordingly, we predicted that such an environment would potentially decrease the energy required for ICT in KCy-R1, relative to KCy, and would result in a red shift of the emission spectrum. To test this hypothesis, Ala was restored at position 158 in KCy-R1. The resulting variant, in which Glu was still at position 38, showed an emission peak maximum at 489 nm (Table 1 and Supporting Information Figure S4), which is almost identical with KCy (λ_{em} = 488 nm). Therefore, we suggest that position 158 is likely responsible for fine-tuning the emission maximum because it has capability to change chromophore polarization through the buried water molecule. This is perfectly consistent with previous studies on amFP486 (*14*, *21*).

Meanwhile, the substitution of Gln at position 38 for the His in KCy-R1, where Ser158 was unchanged, also altered the emission spectrum. The variant, KCy-R1-38H, demonstrated a peak maximum at 489 nm in its emission spectrum (Table 1 and Supporting Information Figure S3). Although, as mentioned above, Ser158 should decrease the chromophore polarization in KCy-R1-38H, the substituted His38 again increases it as much as a Ser158 → Ala mutation. Therefore, we conclude that position 158 has a potent effect on the fine-tuning of CFPs as well as position 38 (*vide supra*).

The results of this study clearly indicate that noncovalent interactions including hydrogen-bonding and electrostatic interactions between the chromophore and the surrounding residues, particularly His38 and Ala158 in the case of KCy, are key factors in the fine-tuning of KCy. In the absence of high-resolution crystal structures of all variants, we cannot rule out the effect of an interaction between position 38 and Glu208. However, since KCy is the first example of a CFP with a short polar side chain (Ser62-OH) in the chromophore, we propose that the noncova-

lent interaction between Ser62 and position 38 plays a vital role in controlling the polarization of the chromophore in KCy. This finding suggests the possibility of fine-tuning other CFPs through the use of double mutations at the corresponding positions. In addition, the interaction between the phenolate oxygen atom of the chromophore and the buried water molecule (W3 in Figure 5a,b) in the cavity formed by position 158 should affect the chromophore polarization. Accordingly, the results of this study indicate that chromophore polarization should be coupled with the residues not only at position 38 but also at position 158 in CFPs. In order to gain a more quantitative understanding of the role of residues in the polarization of the chromophore, theoretical calculations using the SAC-CI method (*33*) are under way.

SUPPORTING INFORMATION AVAILABLE

Spectra comparisons of KCy, KCy-G4219, and KCy-G4219-38L (Figure S1); MALS data of wild-type KCy (Figure S2); absorption, excitation, and emission spectra of the position 38 variants of KCy-R1 (Figure S3); emission spectra of the KCy-R1 variants (Figure S4). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Tsien, R. Y. (1998) The green fluorescent protein. *Annu. Rev. Biochem.67*, 509–544.
2. Zimmer, M. (2002) Green fluorescent protein (GFP): Application, structure, and related photophysical behavior. *Chem. Rev.102*, 759–781.
3. Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005) A guide to choosing fluorescent proteins. *Nat. Methods2*, 905–909.
4. Zhang, J., Campbell, R. E., Ting, A. Y., and Tsien, R. Y. (2002) Creating new fluorescent probes for cell biology. *Nat. Rev. Mol. Cell Biol.3*, 906–918.
5. Ormo, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., and Remington, S. J. (1996) Crystal sStructure of the *Aequorea victoria* green fluorescent protein. *Science273*, 1392–1395.
6. Yang, F., Moss, L. G., and Phillips, G. N.Jr. (1996) The molecular structure of green fluorescent protein. *Nat. Biotechnol.14*, 1246–1251.
7. Matz, M. V., Fradkov, A. 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.
8. 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.
9. Patterson, G. H. (2004) A new harvest of fluorescent proteins. *Nat. Biotechnol.22*, 1524–1525.
10. Remington, S. J. (2006) Fluorescent proteins: Maturation, photochemistry and photophysics. *Curr. Opin Struct. Biol.16*, 714–721.
11. Shu, X., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y., and Remington, S. J. (2006) Novel chromophores and buried charges control color in mfruits. *Biochemistry45*, 9639–9647.
12. 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.A98*, 462–467.
13. Brejc, K., Sixma, T. K., Kitts, P. A., Kain, S. R., Tsien, R. Y., Ormo, M., and Remington, S. J. (1997) Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A.94*, 2306–2311.
14. 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.A102*, 12712–12717.
15. Ai, H-w., Henderson, J. N., Remington, S. J., and Campbell, R. E. (2006) Directed evolution of a monomeric, bright and photostable version of Clavelaria cyan fluorescent protein: Structual characterization and applications in fluorescence imaging. *Biochem. J.400*, 531–540.
16. Bae, J. H., Rubini, M., Jung, G., Wiegand, G., Seifert, M. H. J., Azim, M. K., Kim, J.-S., Zumbusch, A., Holak, T. A., Moroder, L., Huber, R., and Budisa, N. (2003) Expansion of the genetic code enables design of a novel "gold" class of green fluorescent proteins. *J. Mol. Biol.328*, 1071–1081.
17. 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. *Biochemistry44*, 202–222.
18. Kikuchi, A., Fukumura, E., Karasawa, S., Mizuno, H., Miyawaki, A., and Shiro, Y. (2008) Structural characterization of a thiazoline-containing chromophore in an orange fluorescent protein, monomeric kusabira-orange. *Biochemistry* (in press).
19. Wilmann, P. G., Petersen, J., Pettikiriarachchi, A., Buckle, A. M., Smith, S. C., Olsen, S., Perugini, M. A., Devenish, R. J., Prescott, M., and Rossjohn, J. (2005) The 2.1 Å crystal structure of the far-red-fluorescent protein HcRed: Inherent conformational flexibility of the chromophore. *J. Mol. Biol.349*, 223–237.
20. Mizuno, H., Mal, T. K., Walchli, M., Kikuchi, A., Fukano, T., Ando, R., Jeyakanthan, J., Taka, J., Shiro, Y., Ikura, M., and Miyawaki, A. (2008) Light-dependent regulation of structural flexibility in a photochromic fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A.105*, 9227–9232.
21. Gurskaya, N. G., Savitsky, A. P., Yanushevich, Y. G., Lukyanov, S. A., and Lukyanov, K. A. (2001) Color transitions in coral's fluorescent proteins by site-directed mutagenesis. *BMC Biochem.2*, 6.
22. Karasawa, S., Araki, T., Nagai, T., Mizuno, H., and Miyawaki, A. (2004) Cyan-emitting and orange-emitting fluorescent proteins as a donor/acceptor pair for fluorescence resonance energy transfer. *Biochem. J.381*, 307–312.
23. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.276*, 307–326.
24. Vagin, A. A., and Isupov, M. N. (2001) Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. *Acta Crystallogr.D57*, 1451–1456.
25. Lamzin, V. S., Perrakis, A., and Wilson, K. S. (2001) The ARP/WARP suite for automated construction and refinement of protein models, in *International Tables for Crystallography. Vol. F: Crystallography of biological macromolecules* (Rossmann, M. G., and Arnold, E., Eds.) pp 720−722, Kluwer Academic Publishers, Dordrecht, The Netherlands.
26. Roussel, A., and Cambillau, C. (1989) TURBO-FRODE, *Silicon Graphics Geometry Partner Directory*, pp 77−78, Silicon Graphics, Mountain View, CA.
27. 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 System: A new software suite for macromolecular structure determination. *Acta Crystallogr.D54*, 905–921.
28. Schuttelkopf, A. W., and van Aalten, D. M. (2004) PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr.D60*, 1355–1363.
29. Gurskaya, N. G., Fradkov, A. F., Terskikh, A., Matz, M. V., Labas, Y. A., Martynov, V. I., Yanushevich, Y. G., Lukyanov, K. A., and Lukyanov, S. A. (2001) GFP-like chromoproteins as a source of far-red fluorescent proteins. *FEBS Lett.507*, 16–20.
30. Ward, W. W., and Cormier, M. J. (1979) An energy transfer protein in coelenterate bioluminescence. Characterization of the *Renilla* green-fluorescent protein. *J. Biol. Chem.254*, 781–788.
31. McGaughey, G. B., Gagné, M., and Rappé, A. K. (1998) π-Stacking interactions alive and well in proteins. *J. Biol. Chem.273*, 15458–15463.
32. Ai, H-w., Olenych, S. G., Wong, P., Davidson, M. W., and Campbell, R. E. (2008) Hue-shifted monomeric variants of *Clavularia* cyan fluorescent protein: Identification of the molecular determinants of color and applications in fluorescence imaging. *BMC Biol.6*, 13.
33. Hasegawa, J., Fujimoto, K., Swerts, B., Miyahara, T., and Nakatsuji, H. (2007) Excited states of GFP chromophore and active site studied by the SAC-CI method: Effect of protein-environment and mutations. *J. Comput. Chem.28*, 2443–2452.