Published in final edited form as:

J Mol Biol. 2008 May 9; 378(4): 871–886. doi:10.1016/j.jmb.2008.02.069.

Crystal Structure and Raman studies of dsFP483, a Cyan Fluorescent Protein from *Discosoma striata*

Gabrielle D. Malo[¶], Meitian Wang[†], Di Wu[¶], Allison L. Stelling[§], Peter J. Tonge[§], and Rebekka M. Wachter^{¶,*}

[¶]Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287-1604

[§]Department of Chemistry, Stony Brook University, Stony Brook, NY 11794-3400

Abstract

To better understand the diverse mechanisms of spectral tuning operational in fluorescent proteins, we have determined the 2.1 Å X-ray structure of dsFP483 from the reef-building coral *Discosoma*. This protein is a member of the cyan color class of *Anthozoa* fluorescent proteins, and exhibits broad, double-humped excitation and absorbance bands, with a maximum at 437–400 nm and a shoulder at 453 nm. Although these features support a heterogeneous ground state for the protein-intrinsic chromophore, peak fluorescence occurs at 483 nm for all excitation wavelengths, suggesting a common emissive state. Optical properties are insensitive to changes in pH over the entire range of protein stability. The refined crystal structure of the biological tetramer (spacegroup C2) demonstrates that all protomers bear a cis-coplanar chromophore chemically identical to that in green fluorescent protein (avGFP). To test the roles of specific residues in color modulation, the optical properties of the H163Q and K70M variants were investigated. Although absorbance bands remain broad, peak excitation maxima are red-shifted to 455 and 460 nm, emitting cyan and green light respectively. To probe chromophore ground state features, Raman spectra were collected using 752 nm excitation. Surprisingly, the positions of key Raman bands of wild-type dsFP483 are most similar to those of the neutral avGFP chromophore, whereas the K70M spectra are more closely aligned with the anionic form. The Raman data provide further evidence of a mixed ground state with chromophore populations that are modulated by mutation. Possible internal protonation equilibria, structural heterogeneity in the binding sites, and excited state proton transfer mechanisms are discussed. Structural alignments of dsFP483 with the homologs DsRed, amFP486, and zFP538-K66M suggest that natural selection for cyan color is an exquisitely fine-tuned and highly cooperative process involving a network of electrostatic interactions that may vary substantially in composition and arrangement.

Keywords

Cyan fluorescent protein; green fluorescent protein; GFP; reef-building corals; Anthozoa; TLS refinement; color diversification; protein-based chromophores; Raman spectroscopy

*Correspondence should be addressed to R.M.W., Department of Chemistry and Biochemistry, Arizona State University, Tempe, AZ, 85287-1604, RWachter@asu.edu.

[†]Current address: Swiss Light Source, Paul Scherrer Institut, Villigen PSI, Switzerland 5232.

Accession numbers: Coordinates and structure factors for dsFP483 have been deposited in the Protein Data Bank (entry 3CGL). RCSB ID code is RCSB046751

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

Color diversification in reef-building corals has been shown to be the result of positive selection during natural evolution, providing the organisms with an adaptive advantage that has variously been ascribed to photoprotection of endosymbiotic algae, modulation of their photosynthetic activity, or interaction with regulatory photosensors intrinsic to zooxanthellae endosymbionts.¹ The colorful appearance of reef-building corals and related organisms originates in large part from the expression of a set of homologous fluorescent or colored proteins belonging to the family of GFP-like proteins.² Both cyan and red phenotypes appear to have evolved independently and repeatedly along different evolutionary branches of this group of proteins.³ Statistical sequence analyses have provided support for the idea that episodes of positive selection have occurred along all four major groups, clades A–D, within the *Anthozoa* GFP-family tree. It appears that Nature may have repeatedly adapted the GFP-fold to generate blue-shifted versions of GFP that emit in the cyan region of the spectrum, as is demonstrated by the evolution of a distinctly cyan color along different *Anthozoa* clades, as well as along different branches of the same clade. Based on “virtual evolution” experiments, these subgroups have been inferred to originate from common ancestral proteins exhibiting green color, based on the presence of a default GFP-like chromophore in “resurrected”, i.e. recreated ancestral proteins.⁴

In fluorescent proteins (FPs), the conjugated π -system of the covalently attached chromophore is formed by post-translational self-modification of three buried amino acid residues. The fluorescence emission color of members belonging to this protein family is primarily determined by the extent of the chromophore's π -conjugated system, the simplest form of which is that of a *p*-hydroxybenzylidene imidazolinone entity as found in avGFP. Based on covalent fluorophore modifications, FPs may be placed into two major categories, 5 those bearing chromophores chemically identical to avGFP, yielding cyan and green colors, and those that harbor additional oxidation reactions adjacent to the imidazolinone group, yielding chromophores with significantly red-shifted optical properties. In all naturally occurring FPs, the central residue of the tripeptide undergoing fluorophore-yielding modifications has been found to be a tyrosine, Tyr66 in the founding member avGFP derived from the *Aequorea victoria* jellyfish.^{2,5,6} Although the residue following the central tyrosine in the linear protein chain is an obligatory glycine (Gly67 in avGFP), the preceding residue (65 in avGFP) is known to vary substantially among naturally occurring FPs.⁷

Several cyan-fluorescent FPs isolated from reef-building corals have been characterized to date.^{7,8} Their emission spectra render them appropriate donor candidates for Förster resonance energy transfer (FRET) experiments, especially in combination with yellow-fluorescent acceptor GFPs.^{7,8} However, reef GFPs naturally exist as tightly bound tetramers,⁹ somewhat limiting their usefulness as reporter genes due to aggregation.¹⁰ Although the structures of several non-natural cyan FPs with a tryptophan-based chromophore have been described,^{11,12} the in vitro evolution of non-aggregating FPs that emit cyan light from a tyrosine-based chromophore has proven to be somewhat challenging. A monomeric teal-fluorescent FP ($\lambda_{em} = 492$ nm) has recently been engineered from cFP484, a cyan FP derived from a *Clavularia* coral species.¹³ An intermediate on the directed evolution pathway has been demonstrated to undergo reversible photochromism via cis-trans isomerization of the GFP-like chromophore.¹⁴

The cyan-fluorescent Anthozoa FPs exhibit a distinct spectral blue-shift compared to avGFPs, since their emission ranges from 483 to 486 nm, whereas the emission of green light peaks between 504 and 529 nm in GFP variants.⁶ Although the physical origin of this

shift remains unclear, it is apparent that the optical properties are tuned by the three-dimensional arrangement of protein groups around the chromophore.⁵ The series of non-covalent chromophore-protein contacts in the interior of the eleven-stranded β -barrel includes a complex network of electrostatic interactions that appears to be highly conserved among Anthozoa FPs of differing colors.⁸ Based on the recently published X-ray structures of amFP486 and the engineered mTFP0.7 derived from cFP484,^{8,14} tuning of fluorescence from the default green to the cyan state is accomplished by an electric charge quadrupole network involving the residues Arg72, Glu150, His199 and Glu217 (numbering for amFP486) (Figure 1). In amFP486, His199 has been proposed to play an essential role in absorbance blue-shifting, because its side chain π -stacks onto the chromophore's phenolic end, suggesting that the imidazolium cation completes the quadrupolar electrostatic interaction. The lack of pH sensitivity in this protein has been ascribed to an anionic chromophore stabilized by interaction with the positively charged His199 (consensus position #204 in Figure 1).⁸

However, sequence alignments indicate that in the cyan dsFP483 from *Discosoma striata*, the quadrupolar charge interaction observed in amFP486 is not conserved (Figure 1). Surprisingly, the amino acid position equivalent to the π -stacked histidine is occupied by a threonine in dsFP483 (Thr198, Figure 1). DsFP483 shares 55% of its primary amino acid sequence with its long-wavelength relative DsRed isolated from the same genus.¹⁵ In DsRed, the biological tetramer is composed of about 50% green- and 50% red-emitting molecules.¹⁶ Both proteins bear the internal residues Gln66-Tyr67-Gly68 from which the chromophore is constructed.^{2,17} In spite of the high sequence homology, revertants of DsRed produced by genetic engineering have been shown to emit green, not cyan light.¹⁸

To further elucidate the variety of structural and biochemical features that pertain to color tuning in FPs bearing a GFP-like chromophore, we present the high-resolution X-ray structure of the tetrameric dsFP483, in combination with a mutational analysis and Raman spectroscopic data. We find that the results are most consistent with a heterogeneous ground state of the chromophore that consists of a 400 nm absorbing species and a $\sim$ 453 nm absorbing species. Different mechanisms that may form the basis of mixed states or structural disorder are discussed.

RESULTS

Optical properties of wild-type dsFP483

The chromophore of the cyan fluorescent protein dsFP483 from the coral *Discosoma striata* is generated spontaneously from the residues Gln66, Tyr67 and Gly68, the same set of residues as in the closely related red-fluorescent protein DsRed, which exhibits 55% amino acid sequence identity (Figure 1).^{7,19,20} The dsFP483 chromophore matures slowly, with full protein maturation reached within six days after protein purification. Absorbance measurements do not provide evidence of any intermediates with optical density in the visible region of the spectrum. The absorbance spectrum of mature dsFP483 displays a flattened, double-humped band with a broad maximum at 440 nm (excitation maximum at 437 nm) and a significant shoulder around 453 nm, entailing a spectral band width of 3,278 cm^{-1} ($\sim$ 60 nm) at half-maximal absorbance (Table 1 and Figure 2). pH-titrations of the protein in solution yield no shift in absorbance maxima or band shape over the entire pH range of protein stability, pH 5–11. Base denaturation occurs around pH 12, and at pH 13, a single 448-nm band is observed, in line with a GFP-like chromophore anion attached to the denatured protein chain.²¹ Upon acid denaturation at pH 4, the absorbance band shifts to 380 nm, again essentially identical to acid-denatured avGFP with an attached chromophore in the neutral state.⁶ Thus, the optical properties of the denatured forms of the protein

support a chromophore that is chemically identical to the avGFP chromophore. However, the dsFP483 band shape (Figure 2) suggests the presence of multiple chromophore species.

The fluorescence excitation spectrum of dsFP483 mimics the absorbance spectrum exactly. Upon excitation at 415 nm (upslope), 440 nm/450 nm (maxima), or at 455 nm (downslope), the fluorescence emission spectrum contains a sharp peak at 483 nm with a small shoulder around 510 nm (Figure 2), suggesting a common emissive state for both excitation maxima. Emission scans taken in the range of pH 5 to 10 show no shift in maximum as a function of pH, and the fluorescence intensities are nearly constant over the entire pH range. In general, two different models may be invoked to explain the spectroscopic data. The chromophore ground state may entail two different populations that differ in ionization state, with the neutral state absorbing at 440 nm and the anionic state absorbing at 453 nm, whereas the emissive state may be anionic for both species. Alternatively, the ionization state of the chromophore may be fixed, but the chromophore may be surrounded by two different protein environments, giving rise to substantial band broadening due to conformational or charge heterogeneity. The second model would entail different main chain or side chain conformations that may be observable by X-ray crystallography.

Description of the dsFP483 crystal structure

To better understand the origin of the broad, double-humped absorbance spectrum and the structural parameters responsible for green-to-cyan spectral tuning in *Anthozoa* FPs, we solved the X-ray structure of dsFP483 to 2.1 Å resolution. The protein was expressed in *E. coli*, purified, and crystallized in the space group C2 as reported previously.¹⁵ The crystal structure was solved by molecular replacement using DsRed as a search model, and refined to $R_{\text{cryst}} = 18.5\%$ and $R_{\text{free}} = 21.7\%$ (Table 1). The asymmetric unit comprises six protein chains, each consisting of the GFP-like eleven-stranded β -barrel fold, organized into one tetramer (chains ABCD) and one dimer (chains EF). The Matthews coefficient was calculated to be $2.6 \text{ Å}^3 \text{ Da}^{-1}$, with a solvent content of 52.1%.

The molecular interfaces formed by the tetramer ABCD are consistent with the biological tetramer described for other *Anthozoa* FPs.^{8,19,20,22} Chains E and F are related to the adjacent chains E'F' by 2-fold crystallographic symmetry, forming a second biological tetramer judged to be structurally identical to ABCD, but considerably more disordered in the crystal. Therefore, the unit cell may be described as entailing two-dimensional layers of homo-tetramers, each layer composed of four-chain units that exhibit various degrees of disorder.¹⁵ The AB dimer is tightly packed in the crystal, forming a large number of crystal contacts with its biological partner CD, as well as with A'B', related by crystallographic symmetry. In addition, chain D forms a crystal contact to chain E, whereas chain F only interacts with its biological partners E and F'. This allows for the tetramer EFE'F' to pivot around an axis formed by its contacts to D and D'. Hence, F is more disordered than any of the other chains, a fact supported by an average B factor of 135 Å^2 prior to TLS refinement. Therefore, TLS refinement was utilized to model the anisotropic displacement of each protein chain as a group (Table 2),²³ producing a six-monomer model with average residual B factor of $\sim 16 \text{ Å}^2$ for all main chain atoms (Table 3). The refined TLS parameters (Table 2) illustrate that the most significant rigid body displacements occur for chains E and F, with F undergoing the largest movement and both chains entailing a dominant rotational component (L-matrix).²⁴ These observations suggest a larger degree of disorder or multiple positions for chains E and F in the crystal, reminiscent of the six-chain model described previously for the transcription factor GerE.²⁵ TLS refinement of dsFP483 yielded a reduced R-factor and more uniform residual (atomic) B-factors representing local isotropic motions. Based on the quality of the electron density map, the crystallographic model for the ABCD tetramer is judged to be more reliable than that of the EFE'F' tetramer. Consequently, the structural analysis provided below is based on the ABCD tetramer.

According to a PROCHECK analysis, 92.6% of residues in the refined model are found within the core region of the Ramachandran plot, and no residues are found in the “disallowed” or “generously allowed” regions.

A GFP-like chromophore surrounded by a DsRed-like protein environment

The electron density omit map of the dsFP483 chromophore demonstrates that its chemical nature is that of a planar benzylidene imidazolinone group in *cis*-configuration, as found in avGFP and many other GFP-like proteins (Figure 3).²⁷ The geometries of the chain A–D chromophores appear to be identical within model error (~ 0.2 Å), and exhibit average residual B-factors of 15.9 (A), 18.0 (B), 17.5 (C), and 19.0 Å² (D). As expected from the high degree of sequence homology to DsRed, the environment surrounding the chromophore π system is primarily polar, consisting of electrostatic and hydrogen bonding interactions involving the residues Arg95, Thr198, Glu148, Ser146, Glu144 (main-chain), His163, Lys70 and Glu216 in the first interaction shell (Figure 4). The relative spatial arrangement of protein groups around the chromophore is found to be nearly identical in all four chains, providing the same interaction network in chains B, C and D as shown for chain A (Figure 4). In the refined model, the largest variation in estimated hydrogen bonding distance occurs between the ϵ -amino group of Lys70 and the carboxylate of Glu216, where the separation ranges from 2.5 to 3.0 Å among the four chains. Since all NCS restraints were removed during the last stages of refinement, the four protein molecules constituting the biological tetramer may be considered to be structurally identical within model error (~ 0.2 Å), with the possible exception of the Lys70-Glu216 salt bridge. We conclude that the X-ray data do not provide clear evidence of alternate chromophore or protein conformations, although more subtle differences, such as positional variations of Lys70, may not be discernable at the present resolution.

Electrostatic interactions involving Lys70 and Glu216 near the chromophore’s methylene bridge—The ϵ -amino group of Lys70 is positioned “below” the plane of the chromophore (Figure 3), about 3.9 Å distant from Tyr67 C β and C γ , both part of the chromophore’s methylene bridge connecting the two ring systems (Figures 4 and 5). Lys70 occupies the same space as in DsRed (Figure 5), where the Lys70-Glu216 salt bridge is conserved and the charges are separated by 3.2 Å.¹⁹ However, in dsFP483, the B-factors of the side chain functional groups of Lys70 and Glu216 are elevated, clustering at 30 Å² for chains ABCD, whereas the average side chain B-factor for the tetramer as a whole lies at 19 Å². In addition, the electron density observed for Glu216 O ϵ_2 appears somewhat weak (Figure 3), likely due to partial side chain decarboxylation upon synchrotron radiation damage.^{28,29} Loss of electron density is also observed for the highly ordered Glu39 and Glu148 carboxyl groups, as well as for the Cys62 sulfur atom, providing additional support for radiation damage.

His163 hydrogen bonding to the chromophore—In the dsFP483 X-ray structure, the chromophore’s tyrosine-derived hydroxyl is found within hydrogen-bonding distance to the side chains of His163, Ser146 and an ordered water molecule, with donor-acceptor distances ranging from 2.5 to 2.8 Å (Figure 4). The DsRed Lys163 is replaced by His163 (Figures 1 and 5),¹⁹ and consequently, the DsRed salt bridge from the chromophore’s phenolic group to the cationic lysine is replaced by a hydrogen bond to His163 N δ_2 in dsFP483. The arrangement of polar groups around the His163 imidazole ring implies that this group remains primarily uncharged. N δ_1 is found within hydrogen bonding distance (2.9 Å) to the main chain nitrogen of Ala164, which may serve as hydrogen bond donor (but not acceptor), in an interaction that displays favorable donor-acceptor geometry (30° deviation from linearity). These features suggest that the lone electron pair of N δ_1 fulfills the role of H-bond acceptor, necessitating that N δ_1 remains deprotonated, and implying that N δ_2 must

bear a proton. N62 is found within hydrogen-bonding distance to the chromophore's hydroxyl group, which is shown in its protonated form in Panel a and in its anionic form in Panel b (Figure 4). The hydrogen bonding interactions of the imidazole ring are judged to be identical throughout the crystallographic model, since the refined distances of 2.8 and 2.9 Å to the chromophore and to Ala164 are invariant among all four chains A–D. This is not surprising, since the conserved Trp143 provides an edge-to-face aromatic-aromatic interaction with His163, locking the imidazole ring into place via steric restriction.

The hydrogen bonding network involving Ser146, Thr198, and Glu148—The charge state of the conserved residue Glu148 cannot be assigned solely based on its hydrogen bonding pattern, and therefore remains indeterminate in dsFP483. Its carboxyl group may be anionic (Figure 4, Panel a) as it is positioned 5.3 and 6.5 Å distant from the positive charges of Lys70 and Arg95, in line with the interpretation previously provided for DsRed.19 As in DsRed, Glu148 is connected to Lys70 via a hydrogen-bonded network that involves ordered solvent molecules and a side-chain hydroxyl, Thr198 in dsFP483 and Ser197 in DsRed. At a distance of 3.4 Å, W₂₀ is the water molecule in closest proximity to the ε-amino group of Lys70 (Figure 4), whereas in DsRed, the equivalent distance is 3.1 Å. 22 Unlike DsRed, the dsFP483 chromophore is connected to Glu148 by hydrogen bonding interactions that involve Ser146 and two solvent molecules (Figure 4). This arrangement suggests that only one or the other group bears a proton, providing a proton relay path for charge equilibration between the chromophore's phenolic hydroxyl and the Glu148 carboxyl group (see Discussion).

In combination, the four residues Glu148, Lys70, Arg95 and Glu216 may account for a total of two positive and two negative charges in the vicinity of the chromophore, provided that His163 and Glu214 remain neutral (Figure 4, Panel a). The carboxyl of Glu214 is inferred to be protonated since it serves as a hydrogen bond donor to the amide oxygen of the Gln66 side chain. The assignment of side chain oxygen and nitrogen atoms of residues Gln66 and Asn42 (which are hydrogen bonded) is based on placement of the Asn42 side chain nitrogen atom, which serves as hydrogen bond donor to Glu216 (Figure 4).

Hydrophobic contacts—While the dsFP483 charge network is primarily located near one face of the chromophore plane, as previously observed for other *Anthozoa* FPs such as DsRed and amFP486,8,19 the opposite face of the dsFP483 chromophore is exposed to several hydrophobic protein groups belonging to residues Ile161, Leu220 and Pro63.

Molecular interfaces of the tetramer

The dsFP483 homo-tetramer may be thought of as a dimer of dimers, entailing two types of molecular interfaces with characteristics very similar to those described previously for the DsRed tetramer and other *Anthozoa* FPs.19,20 The contacts between chains A and B (and C and D) consist primarily of hydrophobic interactions, but also include a few hydrogen bonds and one intermolecular disulfide bridge involving CysA106 and CysB106. In DsRed, the equivalent positions are occupied by threonine residues that are hydrogen-bonded to each other. In dsFP483, the cysteines are modeled with 50% of the side chains oxidized and cross-linked to the equivalent residue in the adjacent subunit, and 50% bearing free thiols. The A–C (and B–D) interface of dsFP483 is more hydrophilic in nature, consisting mainly of salt bridge and hydrogen bonding interactions, as described for DsRed. The π-stacking interaction of HisA162 with HisC162 is conserved in dsFP483, as is the clasp-like interaction of the C-terminal tails of the A and C chains, where the side chains of His222 and Phe224 partition into indentations on the adjacent subunit. Surprisingly, the A–C interface of the dsFP483 tetramer also involves a sodium ion, coordinated to the carboxylate groups of AspA176 and AspC176, as well as four water molecules. The E–F interface of the

more disordered tetramer EFE'F' in the crystal is equivalent to the A–B interface, and the F–F' interface is equivalent to the A–C interface.

Optical features of active site mutants

Based on the dsFP483 X-ray structure in combination with sequence alignments (Figure 1) and previously published works,^{1,8} Lys70 and His163 appeared likely candidates for residues that play important roles in color tuning. Therefore, we generated the single-site substitution mutants K70M, H163Q, H163K, and H163I of dsFP483. The K70M variant was highly soluble and brightly colored with a greenish-yellow hue. In contrast, only very small amounts of the variants H163Q, H163K and H163I were expressed in soluble form, and most of the protein was lost due to misfolding. Absorbance spectra collected on affinity-purified H163Q protein gave evidence of trace amounts of a GFP-like chromophore, whereas the spectra of H163K and H163I were entirely devoid of absorbance in the visible region.

Spectral properties, fluorescence quantum yield, and apparent Stokes shift for the mutants H163Q and K70M in comparison to wild-type dsFP483 and amFP486 are summarized in Table 1. In the variants, the double-humped band shape is maintained but red-shifted, with the excitation maximum now occurring at 455 nm in H163Q, and 460 nm in K70M (Figure 2). The fluorescence emission of K70M is green and peaks at 500 nm, whereas the emission of H163Q remains in the cyan region of the spectrum with a peak at 486 nm. This effect is due to an increased effective Stokes shift in K70M ($1,739\text{ cm}^{-1}$) compared to H163Q ($1,402\text{ cm}^{-1}$), whereas the largest apparent Stokes shift is observed for wild-type dsFP483 ($2,179\text{ cm}^{-1}$) in this series. However, all values, including that for amFP486 ($1,646\text{ cm}^{-1}$),⁸ are intermediate to the Stokes shifts observed for avGFP ($3,450\text{ cm}^{-1}$ for the neutral, and $1,210\text{ cm}^{-1}$ for the anionic form).^{6,30}

Overall, the optical features of H163Q are very similar to those of its parent molecule, wild-type dsFP483. As the engineered glutamine maintains the hydrogen bonding potential of the replaced imidazole group, the H163Q spectra provide additional support for a neutral His163 in the wild-type protein (Figure 4). The red-shifted features and reduced Stokes shift of the H163 and K70M mutants support the presence of a chromophore population that bears a negative charge in the ground state, suggesting that the two overlapping absorbance humps of the parent protein arise from a heterogeneous ground state, possibly (but not necessarily) with distinct charge states. These states appear to be highly coupled in the excited state, such that excitation of either yields a common emissive state that yields cyan color. The high quantum yield of 0.78 determined for wild-type dsFP483 is essentially identical to that of amFP486, and is consistent with an anionic emissive state.⁸ In the H163Q variant, the quantum yield remains high (0.63), in contrast to K70M (0.05), where fluorescence appears to be quenched by the engineered methionine sulfur atom.

Raman spectroscopy

Wild-type dsFP48—To better characterize ground state features such as protonation equilibria of the dsFP483 chromophore, steady-state Raman spectra were collected at pH 7.9 utilizing excitation at 752 nm as described previously.³¹ Under these conditions, avGFP Raman spectra are dominated by pre-resonance-enhanced contributions from the chromophore, and the chromophore's vibrational bands have been assigned previously via isotopic labeling methods.³² Figure 6 shows the Raman spectrum of dsFP483 together with the Raman spectra of avGFP-S65T at pH 5 and pH 8 published previously.³¹ The chromophore in avGFP-S65T is sensitive to pH, titrating with a pK_a of 6.0. At pH 5, the chromophore is largely in the neutral form, while at pH 8 only the anionic form is present. The Raman spectra of avGFP-S65T reveal two principal features that are sensitive to

chromophore ionization: a band at 1640–1620 cm^{-1} , assigned to a delocalized mode involving stretching motions of the exocyclic C=C bond ($\nu_{\text{C}=\text{C}}$), and a band at 1560–1540 cm^{-1} assigned to a mode involving motions of the imidazolinone group. For S65T, the C=C mode shifts from 1643 to 1619 cm^{-1} upon ionization, while the imidazolinone mode at 1558 cm^{-1} for neutral S65T is found at 1540 cm^{-1} for the anion (Figure 7). The shoulder at ~1544 cm^{-1} in the pH 5 spectrum of S65T indicates the presence of a small amount of anionic chromophore.

In the Raman spectrum of dsFP483, the C=C and imidazolinone modes are found at 1634 and 1555 cm^{-1} , respectively (Figures 6 and 7). Initial comparison of the data with those for S65T suggests that the chromophore of dsFP483 is in the neutral state. This conclusion is supported in particular by the similarity in band positions for the imidazolinone mode in dsFP483 (1555 cm^{-1}) and neutral avGFP-S65T (1558 cm^{-1}). In addition, the position of the C=C mode for dsFP483 is closer to that for neutral avGFP-S65T than to anionic avGFP-S65T, again consistent with a neutral chromophore in the cyan protein. However, we note that the C=C mode for dsFP483 is red shifted 9 cm^{-1} compared to the frequency of the corresponding band in neutral S65T. This is significant since the position of this mode for other neutral forms of the chromophore is largely insensitive to the environment. For example, neutral 4'-hydroxybenzylidene-2,3-dimethylimidazolinone (HBDI) in H_2O ,³² and the neutral forms of S65T and S65T/H148D, have C=C bands at 1641, 1643 and 1642 cm^{-1} , respectively, while the corresponding values of λ_{max} for these chromophores are 368, 394 and 415 nm, respectively. In contrast, the C=C mode for anionic chromophores is much more sensitive to the environment, and a linear correlation has been shown to exist between $\nu_{\text{C}=\text{C}}$ and λ_{max} for anionic HBDI in various solvents and for different anionic fluorescent proteins.³¹ Indeed, the λ_{max} (437 nm) and $\nu_{\text{C}=\text{C}}$ values for dsFP483 place this protein at the blue end of the linear correlation with spectral parameters similar to anionic HBDI in a solvent such as 2-propanol. However, the spectral parameters for dsFP483 are considerably blue-shifted compared to other anionic FPs. Overall, the data appear more consistent with a red-shifted neutral chromophore than a blue-shifted anion, although significant ground state heterogeneity is indicated by the intermediate band positions (see Discussion).

DsFP483-K70M—The imidazolinone ring mode of the K70M variant is shifted from 1555 cm^{-1} (wild-type) to 1547 cm^{-1} , rendering the ground state of K70M more similar to the anion than the neutral form when compared to GFP-S65T (1558=neutral, 1540=anion). In addition, the methylene bridge C=C mode is shifted from 1634 cm^{-1} (wild-type) to about 1610 cm^{-1} (K70M), again more similar to the anionic form of the chromophore. In comparison, the GFP-S65T anionic mode is observed at 1619 cm^{-1} , and that of the YFP-H148Q anion is observed at 1612 cm^{-1} .³¹ The significant shift observed for the C=C mode provides support for the presence of an anionic population in K70M, as this mode is most sensitive to the environment when the chromophore bears a negative charge.³¹ This interpretation does not preclude a coexisting neutral population, as the red form tends to dominate Raman spectra due to preresonance, and the blue form may not be easily observable if the population is mixed.

DISCUSSION

DsFP483 structural features and optical properties

In the cyan-emitting coral protein dsFP483, ground state heterogeneity is indicated for its 4-hydroxybenzylidene imidazolinone chromophore, as the excitation band appears to be much broader than the emission band (3,278 cm^{-1} vs. 2,017 cm^{-1} full width at half-maximum, see Figure 2). However, in all four chains comprising the tetramer, the chromophore geometry is that of the *cis*-coplanar avGFP chromophore. Therefore, the origin of the heterogeneity may

lie in coexisting protonation equilibria of the chromophore itself, or in a variation in the shape or electrostatics of the chromophore binding pocket.

In a large number of avGFP variants, the phenolic end of the fluorophore titrates with pH, exhibiting a large variation in pK_a values depending on the particular protein environment.³³ However, in wild-type avGFP, populations of protonated and deprotonated chromophores coexist at a constant ratio, as determined by the internal electrostatic and hydrogen bonding arrangement, that is independent from external pH conditions.⁶ Similarly, *Anthozoa* FPs exhibit a general lack of pH sensitivity of spectral properties, and appear to largely bear chromophores with fixed protonation states. For dsFP483, the absorption and emission bands peak at 437–440 nm and at 483 nm respectively, red-shifted from those of the neutral avGFP chromophore ($\lambda_{ex} = 397$ nm and $\lambda_{em} = 460$ nm), and blue-shifted from those of the anionic avGFP chromophore ($\lambda_{ex} = 475$ nm and $\lambda_{em} = 504$ nm).^{6,30} According to the dsFP483 X-ray structure presented here, the chromophore's phenolic hydroxyl is involved in three hydrogen bonding interactions, to Ser146, to a solvent molecule, and to the imidazole ring of His163. These interactions allow for complete satisfaction of the hydrogen bonding potential irrespective of protonation state (Figure 4 a,b). The chromophore serves as hydrogen bond acceptor to N82 of His163, which is inferred to bear a proton, as N81 must remain deprotonated to accept a hydrogen bond from Ala164. However, a neutral interaction partner does not necessarily preclude stabilization of the chromophore anion within the protein fold,³⁴ and the dsFP483 X-ray structure is consistent with both neutral and anionic forms of the chromophore. For the most part, structural variations in the immediate chromophore environment appear to be lacking, however, the B-factors of the Lys70 ϵ -amino group and the carboxylate oxygen atoms of Glu216 are somewhat elevated. The increased disorder or positional heterogeneity of these groups is reflected in a variation of their charge separation among the four protein chains of the biological tetramer. The Lys70-Glu216 salt bridge ranges from 2.5 to 3.0 Å among the ABCD protomers of the refined model, indicative of increased positional error or actual conformational differences among the four chains. At the current resolution limit, these two possibilities cannot be distinguished by the X-ray data. However, electrostatic modulations in the chromophore surroundings could form the basis for the significant band broadening observed in absorbance spectra. In addition, ground state equilibria involving the distribution of protons and other charges in the tetrameric assembly may be established by means of solvent channels connecting the chromophores of the AC and BD protomers. For example, in DsRed, a solvent path of hydrogen-bonded waters and protein groups has been shown to connect two chromophores of the tetramer through the hydrophilic A–C interface.²⁰ In dsFP483, this interface involves two aspartic acid carboxyl groups in position 176 of each chain, as well as a sodium ion and four water molecules. In combination, these groups may allow for proton relay and charge equilibration of the respective chromophores across the interface.

Ground state chromophore spectroscopic features

The spectroscopic results obtained for dsFP483 and its mutants remain ambiguous in terms of chromophore charge state. Quantum yield data (Table 1) and the lack of pH sensitivity argue in favor of a chromophore anion. However, Raman spectra collected on dsFP483 at the same pH as the crystal growth conditions (pH 7.9) are more consistent with a red-shifted neutral chromophore than a blue-shifted anion (Figures 6 and 7). According to this interpretation, the observed decrease in wave number of the highly delocalized C=C mode would be associated with a red-shift in the absorption maximum, reflecting a reduction in the energy gap between ground and excited states. We speculate that the shift from 1643 cm^{-1} (neutral GFP-S65T) to 1634 cm^{-1} (dsFP483) may result from increased delocalization of electron density throughout the chromophore's ring system, as stabilized by the positive charge of the Lys70 ϵ -amino group positioned in close proximity to the methylene bridge.

Increased charge delocalization has previously been associated with red-shifted optical properties in the yellow fluorescent GFP variants.³¹

In the K70M variant of dsFP483, the positive charge near the methylene bridge (Figure 3) is replaced by a hydrophobic group, eliminating the salt bridge to Glu216. Likely, structural adjustments yield a modified distribution of charges around the chromophore's phenolic end, shifting the chromophore population towards a state that is more similar to the avGFP anion. Although the detailed effects of the K70M substitution are currently unknown, the Raman spectra indicate that the wild-type dsFP483 chromophore is rather distinct from that found in K70M. The Raman features of wild-type are more similar to the neutral avGFP chromophore, whereas those of K70M are reminiscent of the avGFP anion. In spite of this resemblance, a more detailed description of chromophore subpopulations in dsFP483 is not currently possible. Preliminary spectral deconvolution experiments on the broad, highly delocalized C=N Raman mode in the 1555 cm⁻¹ region further suggest multiple ground state species that may correspond to the double peak of the absorption spectrum. A chemical and photophysical description of the nature of these species will need to await a more detailed spectroscopic characterization utilizing steady-state and time-resolved methods.

Can the fluorescence emission properties be rationalized by excited-state proton transfer (ESPT)?

The possibility must be considered that the cyan color originates from an emissive state comprising the chromophore anion, produced during excited-state dynamic processes entailing the loss of a proton. Ultrafast proton transfer from the chromophore's phenolic hydroxyl to a nearby protein group was originally described for avGFP,³⁰ but has also been demonstrated for various mutant GFPs,^{35,36,37,38} and other wild-type FPs such as Kaede.³⁹ In comparison, the observed dsFP483 Stokes shift of 2,179 cm⁻¹ appears somewhat small to support ESPT, however, a hydrogen-bond network connecting the chromophore's phenolic hydroxyl to the carboxylate of Glu148 may provide a facile pathway for this type of process (Figure 4). Indeed, a reciprocal internal protonation equilibrium may exist, in which either the chromophore or the Glu148 carboxyl bears a proton but not both, reminiscent of wild-type avGFP.⁴⁰ Intriguingly, a chromophore sub-population with excitation at 435 nm has recently been characterized in GFP-S65T/H148D, where a proton is thought to be shared between the chromophore's phenolic oxygen and the Asp carboxyl group.^{35,36,41} However, such an unusually short hydrogen bond is not apparent in the dsFP483 X-ray structure.

Structural comparison with the *Anthozoa* red fluorescent protein DsRed

When comparing the chromophore environment of dsFP483 with that of the closely related DsRed protein,²² it is apparent that the only noteworthy difference concerns residue 163 (consensus # 169, Figure 1), a histidine in the cyan and a lysine in the red protein. All other hydrogen bonding and charge-charge interactions originally described for DsRed are closely mimicked in dsFP483, including hydrogen bonds with the side chain hydroxyl of Thr198 (DsRed Ser197, consensus #204, Figure 1), several ordered solvent molecules, the carboxylate of Glu148 (consensus #152) and the hydroxyl of Ser146 (consensus #150), as well as a salt bridge between Lys70 (consensus #72) and Glu216 (consensus #223).

Interestingly, 40 to 50% of the chains in the DsRed tetramer bear a GFP-like chromophore, 16,42 which emits green light around 500 nm.⁴³ The green chromophore population of DsRed is anionic based on its absorbance at 485 nm, in line with the salt bridge formed by the chromophore phenolate and the ϵ -amino group of Lys163. Strikingly, the emission from the anionic chromophore embedded in a protein cavity that otherwise mimics dsFP483 yields green, not cyan color, suggesting that the color change from green to cyan is

controlled by substitution of cationic Lys163 with neutral His163 in dsFP483. Therefore, the replacement of a positively charged residue with a polar one appears to be coupled to cyan light emission in a DsRed-like protein environment.

Structural Comparison with the *Anthozoa* cyan-fluorescent amFP486

DsFP483 exhibits optical properties very similar to those observed in amFP486, another cyan fluorescent protein from *Anthozoa* with 40% sequence identity to dsFP483. AmFP486 exhibits a double-humped absorbance band that peaks at 454 nm, and lacks pH-dependent spectral changes in a manner reminiscent of dsFP483.⁸ A structural alignment of the dsFP483 and amFP486 chromophores gives a rms deviation of 0.12 Å, hence the cis-coplanar geometries of the π -conjugated system are judged to be identical within error.

Superimposition of dsFP483 and amFP486 structures indicates that the dsFP483 residues 69 to 72 are shifted by approximately 0.75 Å, placing the dsFP483 Lys70 ϵ -amino group 1.2 Å distant from the amFP486 His199 imidazole ring in the overlay (Figure 5). The positive charge of Lys70 is found within van der Waals distance (3.8 to 4.0 Å) to C $_{\beta}$ and C $_{\gamma}$ of the chromophore-forming Tyr67, both bridging atoms that connect the two ring systems. Therefore, a resonance form may be dominant that places increased charge density onto the methylene bridge, leading to increased delocalization over the entire chromophore skeleton. In contrast, the His199 imidazolium group in amFP486 interacts with the chromophore's phenolic end via π stacking, and the guanidinium group of Arg72 (equivalent to dsFP483 Lys70) is rotated away from the chromophore's methylene bridge.

The side chain of amFP486 His199, structurally equivalent to dsFP483 Thr198 (consensus #204) has been inferred to carry a positive charge, based on a direct interaction with the carboxyl groups of Glu217 and Glu250 (Figure 5, cyan color). A quadrupole network of alternating positive and negative charges, consisting of Arg72, Glu150, His199 and Glu217 (consensus #72, 152, 204, and 223) has been proposed to play an important role in the production of cyan color in this protein.⁸ Elimination of a negative charge via the E150Q substitution provided support for this notion, based on spectral shifting to green emission in this mutant. A similar spectral red-shift was also observed for the H199T mutant, however, the quantum yield was reduced dramatically due to a non-planar chromophore geometry.⁸ Regardless, the equivalent quadrupole electrostatic network appears to be disrupted in dsFP483, as the two positive charges Arg72 and His199 (amFP486) have been replaced by one, Lys70 (dsFP483). In the structural overlay, the Lys70 ϵ -amino group is positioned in between the two positive charges of amFP486. Therefore, both electrostatic arrangements are consistent with cyan emission, although the origin of spectral tuning remains elusive.

Structural comparison with the *Anthozoa* green fluorescent protein zFP538-K66M

A π -stacked histidine, His202 (consensus #204), has previously been described in the green variant zFP538-K66M, derived from a *Zoanthus* FP (Figure 5, green model).²² In this protein, the imidazolium cation of His202 has been implicated in stabilizing the chromophore anion, and partakes in a quadrupole arrangement of charges involving the residues Glu221, His202, Glu150, and Arg70, an arrangement essentially identical to that found in amFP486 (Figure 5, cyan model). Based on the fact that zFP538-K66M emits green light, it appears unlikely that either the positive charge of the π -stacked histidine (consensus #204) or the quadrupole network of electrostatic interactions is sufficient to effect cyan emission, although these are likely contributing factors. On the other hand, consensus position #169 (His 163 in dsFP483) clearly plays an important role in color modulation. This position is occupied by an alanine in amFP486, and by methionine in the green mutant of zFP538. In amFP486, the Ala165 side chain does not contact the chromophore directly, instead, a water molecule is found to provide hydrogen bonding interactions with the

chromophore.⁸ Consensus position #169 was identified previously as a determinant of cyan color by sequence analysis, and was found to be occupied by either His or Ala in the proteins dsFP483, cFP484 and amFP486.⁴⁴ In contrast, the equivalent position is taken up by an arginine in the cyan protein mcFP485, further demonstrating that no one residue appears to be the sole color determinant in cyan FPs.

Sequence comparisons and evolutionary relationships

A NCBI data base search produced 14 hits of wild-type protein sequences for cyan fluorescent proteins belonging to the family of GFP-like proteins. A subset of these is shown in the CLUSTALW sequence alignment in Figure 1. The cyan FPs with known X-ray structures have arisen from different clades of the phylogenetic tree of Anthozoa FPs: dsFP483 belongs to clade B,⁴⁵ amFP486 to clade C, and cFP484 to clade D.³ Currently, available structures for cFP484 are those of highly engineered teal-fluorescent mutants.^{13,14} Several studies have demonstrated significance of consensus position #169 in the positive selection for cyan color (Figure 1), a position that corresponds to dsFP483 His163 (clade B), and amFP486 Ala164 (clade A). To-date, extensive evolutionary studies have been performed only within clade D,¹ where position #169 scored highest as a determinate of cyan color in ‘virtual evolution’ experiments.¹ Convincing evidence has been provided that this residue changed from a methionine in a green common ancestor to either an arginine (mcFP485) or a histidine (cFP484) in extant cyan proteins. Forward (M167R) and reverse (R167M) mutagenesis performed on the *Montastrea cavernosa* green ancestor and cyan extant protein mcFP485, respectively, confirmed that this position plays an essential role in the distinction between cyan and green emission in this clade, in line with a previously published sequence analysis.⁴⁴

Conclusions

The structural comparison of fluorescent proteins presented above allows us to delineate likely features essential in color tuning to cyan emission wavelengths, provided that the analysis is carried out in a pair-wise manner. The identified features appear to depend on the particular pair under consideration, be it dsFP483 in relation to the green form of DsRed, the green form of zFP538, or in relation to the cyan proteins amFP486 or cFP484. Therefore, with respect to specific residue types or placement of charges in the vicinity of the chromophore, the available data do not allow for the description of one overriding feature that yields cyan color in all naturally occurring fluorescent proteins bearing this characteristic. Clearly, structural alignment of crystallographic models is not sufficient to fully understand the subtleties of non-covalent interactions that lead to the fine-tuning of fluorescence emission over a very narrow wavelength range spanning 483–486 nm. It appears that adaptive evolution has yielded several different avenues to achieve precise spectral modulation along different *Anthozoa* lineages that have diverged from a common green ancestor.¹ For this reason, the mechanism of tuning to cyan color of a default green fluorophore may differ from one cyan FP to another, just as red color has evolved several times independently within the family of GFP-like proteins.^{3,17} The molecular features appear to be complex, and the underlying photophysical mechanisms are difficult to extract from structural alignments or steady-state spectroscopy.

Materials and Methods

Absorbance, fluorescence, and quantum yield measurements

Room temperature absorbance spectra were collected on dsFP483 solutions ranging from pH 4.0 to 14.0. 50 μ l aliquots of 1.2 mg/ml protein solution (50 mM HEPES, 300 mM NaCl, 10 mM EDTA) were diluted 10-fold into appropriate buffers, mixed gently, and allowed to equilibrate. pH 4.0 to 11.0 buffers consisted of 50 mM citrate, acetate, PIPES, HEPES,

CHES or CAPS, 100 mM NaCl and 1mM EDTA. Buffers at pH 12, 13 and 14 were prepared by adjusting NaOH concentrations to 10, 100, and 400 mM. Absorbance scans were collected on a Shimadzu UV-2401 spectrophotometer, and fluorescence excitation and emission spectra were collected a Jobin Yvon Fluoromax-3[®] fluorimeter. Fluorescence emission scans ($\lambda_{\text{ex}} = 450 \text{ nm}$) were collected on protein solutions ranging from pH 4.0 to 10.0, utilizing 1.0 nm increments, an integration time of 1.0 s, and a slit width of 1.0 nm for both excitation and emission. Excitation spectra of dsFP483 were collected at pH 7.0 with λ_{em} set to 483 nm.

The fluorescence quantum yield of wild-type dsFP483 (0.015 mg/mL) was determined using fluorescein as a standard (1 μM in 0.1 M NaOH), and integrating the fluorescence emission intensity upon excitation at 456 nm. The quantum yield Φ_{FI} for the mutants dsFP483-H163Q and dsFP483-K70M was determined with wild-type dsFP483 (~0.03 mg/mL) as a standard, using the same excitation wavelength. Buffer for all protein preparations was 50 mM HEPES pH 7.9, 300 mM NaCl, 1mM EDTA. The data were fit to the following equation $\Phi_{\text{FI}} = \Phi_{\text{R}} \cdot (I/I_{\text{R}}) \cdot (\text{OD}_{\text{R}}/\text{OD})$, where Φ is the quantum yield, I the integrated intensity, and OD the optical density at 456 nm.⁴⁶ The subscript R stands for the reference fluorophore with known quantum yield.

Preparation of dsFP483 wild-type and mutant protein

Bacterial expression and Ni-affinity purification of N-terminally 6His-tagged wild-type dsFP483 protein has been described previously.¹⁵ Single-site mutants were introduced with a PCR-based method using the QuikChange site-directed mutagenesis kit (Stratagene), following the manufacturer's instructions. The dsFP483 mutants containing the single residue substitutions K70M, H163Q, H163K, H163I were each expressed in the *E. coli* strain M15 using the plasmid pQE30 (Qiagen). Protein expression was induced by addition of 1 mM IPTG to 4 L cultures ($\text{OD}_{600} = 0.6$) at 25 °C, and allowed to proceed for seven hours or overnight. The soluble fraction was purified over a Ni-NTA affinity column (Qiagen) following the manufacturer's instructions, and the 200 mM imidazole eluent was dialyzed immediately against buffer containing 50 mM HEPES pH 7.9, 300 mM NaCl, 1 mM EDTA.

Structure determination of dsFP483

Crystal growth, data collection and molecular replacement solution has been described previously.¹⁵ Briefly, dsFP483 protein bearing an N-terminal 6His-tag was expressed in *E. coli* and purified by nickel affinity chromatography, and crystals were grown at 4°C from 100 mM HEPES pH 7.9, 300 mM NaCl, 200 mM calcium acetate, and 16% v/v PEG4000. Diffraction data were collected at the Advanced Light Source Beamline 5.0.3, and indexing, integration and scaling were carried out with DENZO/SCALEPACK.⁴⁷ Data were processed to 2.1 Å resolution in spacegroup C2 ($a = 111.2 \text{ Å}$, $b = 78.27 \text{ Å}$, $c = 188.6 \text{ Å}$, $\alpha = \gamma = 90$, $\beta = 91.35$) and yielded $R_{\text{sym}} = 7.0\%$.¹⁵ MOLREP molecular replacement searches, 48 utilizing DsRed (PDB code 1G7K) as a search model,¹⁹ revealed a tetrameric β -barrel protein with pseudo-222 point group symmetry, and resulted in the localization of a total of five protein chains ABCDE in the asymmetric unit, though their packing suggested the presence of a sixth chain F.¹⁵ Crystal twinning tests based on reflection intensity statistics provided no evidence of hemihedral crystal twinning.^{15,49}

Crystallographic model building and initial refinement

All crystallographic refinement procedures were carried out using REFMAC5 within the CCP4 suite of programs,^{50,51} and models containing the correct amino acid sequence were built using the program XtalView.⁵² Several rounds of model building followed by positional and B-factor refinement to 2.1 Å yielded an initial model comprising the five

protein chains ABCDE without solvent, and resulted in $R_{\text{cryst}} = 0.273$ and $R_{\text{free}} = 0.302$. At this stage, the electron density of the ABCD tetramer was well defined, whereas the density for chain E remained poor. The average atomic B factor was $22 \text{ \AA}^2$ for chains AB, $38 \text{ \AA}^2$ for chains CD, and $64 \text{ \AA}^2$ for chain E. Careful inspection of the $|F_o| - |F_c|$ difference map revealed faint electron density for the sixth monomer F, thought to pair with monomer E to form half of a tetramer adjacent to a crystallographic 2-fold axis. Therefore, a second, more disordered layer of tetramers appears to cross the crystal lattice, constructed from chains E and F via 2-fold crystallographic symmetry.¹⁵ Tetramer assembly was judged to be identical for the more ordered (ABCD) and the less ordered (EFE'F') layers in the crystal, hence a model for chain F was generated by superimposition of the dimer AB onto the monomer E, matching the α -carbons of chain A with those of chain E by a least-squares fitting algorithm. Chain F was subsequently modeled from the superimposed coordinates of chain B. However, extensive rigid body and positional refinement tests with inclusion of all six chains did not lower the crystallographic R factor, and modeling of any atoms of chain F was therefore postponed (see below). Solvent molecules were built into the five-monomer model, reducing R_{cryst} and R_{free} to 0.244 and 0.292 respectively, though the quality of the electron density maps for chains E and F did not improve. To further exclude the possibility that pseudo-merohedral crystal twinning may be the underlying cause of the poor density (the β angle in the C2 cell is close to 90°),⁵³ additional refinement tests were carried out that included utilization of a crystal twin law within the SHELX97 crystallographic software package.⁵⁴ However, this approach neither resolved the disordered chain F nor reduced the crystallographic R factor.

Refinement strategy with inclusion of TLS and NCS parameters

Given the high quality of the diffraction data, the unsatisfactory value of R_{free} suggested that thermal vibrations in the crystal lattice may not be modeled adequately by use of atomic B factors alone. Although the resolution limit of $2.1 \text{ \AA}$ does not allow for the refinement of anisotropic displacement parameters for individual atoms, TLS (translation, libration, screw-rotation) refinement provides a method to model large-scale anisotropic molecular motions.^{23,24} In this approach, a set of twenty TLS parameters describes the mean square translation (T), libration (L), and correlation between translation and libration (S) for each rigid body. Accordingly, we implemented a TLS refinement strategy in which each protein chain was defined as one rigid body TLS group.²⁵ Solvent molecules were omitted from the TLS group definitions, and all atomic B factors were set to 50 at the beginning of TLS refinement. In the first step, TLS parameters were refined in REFMAC5 while keeping atomic positions constant,⁵⁰ and in the second step, the converged TLS parameters were kept constant, while refining atomic coordinates and isotropic B factors. These residual atomic B factors model deviations from the rigid-body description, i.e. local atomic displacements after large-scale rigid body movements have been subtracted. After twenty cycles of TLS refinement, followed by ten cycles of restrained maximum likelihood refinement in REFMAC5, R_{cryst} and R_{free} decreased to 0.234 and 0.259 respectively, validating this approach. After inclusion of TLS parameters in refinement, the average residual B factor for each of the five protein chains converged to a similar value.

Additional TLS refinement included tight non-crystallographic symmetry (NCS) positional restraints of chain E to chain A, implemented due to the poor quality of the chain E map. To further improve the model, chains A–E and solvent molecules were manually adjusted based on visual inspection of $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ maps, followed by intermittent refinement cycles that included TLS refinement (initial B factor set to 30) and NCS positional refinement. Two sodium ions were modeled into positive difference density. Cys106 of chains A and B and Cys106 of chains C and D were modeled as intermolecular disulfide linkages in two conformations, each with an occupancy of 0.5.

When R_{cryst} and R_{free} had dropped to 19.3% and 22.7% respectively, several continuous stretches of 2σ positive difference density were visible in the $|F_o|-|F_c|$ map of chain F, near the E–F interface and along the central helix bearing the chromophore. A model for monomer F was generated by superimposition as described above, and TLS refinement was carried out with six chains in the model, positionally restraining E and F to chain A by NCS. Regions of chain F that were not visible in 1σ $2|F_o|-|F_c|$ density were deleted from the PDB file. Atoms with a residual B factor larger than twenty, 3σ negative difference density, or an absence of 1σ $2|F_o|-|F_c|$ density, were given an occupancy of zero or 0.5, depending on the characteristics.

A final round of TLS refinement in the absence of any NCS restraints yielded $R_{\text{cryst}} = 18.5\%$ and $R_{\text{free}} = 21.7\%$ (Table 2). The refined TLS parameters are listed in Table 3. The final crystallographic model consists of the following: chain A (residues 1–226), B (residues 7–226), C (residues 6–228), D (residues 7–228), E (residues 7–226), F (residues 20–23, 25–27, 49–51, 53–63, 96–101, 103–108, 124–139, 160–164, 174–178), 567 water molecules, and two sodium ions (Table 2). About half of the amino acid residues present in the chain F model contain side chains.

Raman Spectroscopy

Upon purification, freshly prepared wild-type dsFP483 protein and its K70M variant were incubated at room temperature for 6 days to allow for complete protein maturation. The samples were then buffer-exchanged into 20 mM HEPES pH 7.9, 300 mM NaCl, and the concentration was adjusted to 2.4 mg/ml (92 μM). Steady state Raman spectra were acquired as described previously,³¹ using instrumental settings providing 670 mW of 752 nm excitation. 60 μL of protein was placed in a 2 mm by 2 mm fluorimeter cell and data were collected for 400 s (2 s accumulation $\times$ 200 scans). Subsequently, the sample was replaced by buffer and a reference spectrum was acquired. The Raman spectra of the proteins were obtained by spectral subtraction, and the data were wavenumber-calibrated using the spectrum of cyclohexanone. Peak positions are accurate to $\pm 2\text{ cm}^{-1}$.

Acknowledgments

This work was supported by a grant from the National Science Foundation (NSF grant MCB-0615938) to RMW, and a grant from the National Institutes of Health (GM63121) to PJT. Crystallographic data were collected at the Advanced Light Source Beamline 5.0.3, Lawrence Berkeley National Laboratory, which is supported by the US Department of Energy.

ABBREVIATIONS

HBDI	4'-hydroxybenzylidene-2,3-dimethylimidazolinone
avGFP	<i>Aequoria victoria</i> green fluorescent protein
FP	fluorescent protein
ESPT	excited state proton transfer

References

1. Field SF, Bulina MY, Kelmanson IV, Bielawski JP, Matz MV. Adaptive evolution of multicolored fluorescent proteins in reef-building corals. *J Mol Evol.* 2006; 62:332–339. [PubMed: 16474984]
2. Wachter RM. The family of GFP-like proteins: Structure, function, photophysics and biosensor applications. *Photochem Photobiol.* 2006; 82:339–344. [PubMed: 16223340]
3. Shagin DA, Barsova EV, Yanushevich YG, Fradkov AF, Lukyanov KA, Labas YA, Semenova TN, Ugalde JA, Meyers A, Nunez JM, Widder EA, Lukyanov SA, Matz MV. GFP-like proteins as

- ubiquitous Metazoan superfamily: evolution of functional features and structural complexity. *Mol Biol Evol.* 2004; 21:841–850. [PubMed: 14963095]
4. Ugalde JA, Chang BSW, Matz MV. Evolution of coral pigments recreated. *Science.* 2004; 305:1433. [PubMed: 15353795]
 5. Remington SJ. Fluorescent proteins: maturation, photochemistry and photophysics. *Curr Opin Struct Biol.* 2006; 16:1–8. [PubMed: 16439109]
 6. Tsien RY. The Green Fluorescent Protein. *Ann Rev Biochem.* 1998; 67:509–544. [PubMed: 9759496]
 7. Matz MV, Fradkov AF, Labas YA, Savitsky AP, Zaraisky AG, Markelov ML, Lukyanov SA. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nature Biotechnol.* 1999; 17:969–973. [PubMed: 10504696]
 8. Henderson JN, Remington SJ. Crystal structures and mutational analysis of amFP486, a cyan fluorescent protein from *Anemonia majano*. *Proc Natl Acad Sci USA.* 2005; 102:12712–12717. [PubMed: 16120682]
 9. Baird GS, Zacharias DA, Tsien RY. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proc Natl Acad Sci USA.* 2000; 97:11984–11989. [PubMed: 11050229]
 10. Verkhusha V, Lukyanov KA. The molecular properties and applications of Anthozoa fluorescent proteins and chromoproteins. *Nat Biotechnol.* 2004; 22:289–296. [PubMed: 14990950]
 11. Bae JH, Rubini M, Jung G, Wiegand G, Seifert MHJ, Azim MK, Kim J-S, Zumbusch A, Holak TA, Moroder L, Huber R, Budisa N. Expansion of the genetic code enables design of a novel "Gold" class of green fluorescent proteins. *J Mol Biol.* 2003; 328:1071–1081. [PubMed: 12729742]
 12. Malo GD, Pouwels LJ, Wang M, Weichsel A, Montfort WR, Rizzo MA, Piston DW, Wachter RM. X-ray structure of Cerulean GFP: A tryptophan-based chromophore useful for fluorescence lifetime imaging. *Biochemistry.* 2007; 46:9865–9873. [PubMed: 17685554]
 13. Ai H-W, Henderson JN, Remington SJ, Campbell RE. Directed evolution of a monomeric, bright, and photostable version of *Clavularia* cyan fluorescent protein: structural characterization and applications in fluorescent imaging. *Biochem J.* 2006; 400:531–540. [PubMed: 16859491]
 14. Henderson JN, Ai H-w, Campbell RE, Remington SJ. Structural basis for reversible photobleaching of a green fluorescent protein homologue. *Proc Natl Acad Sci USA.* 2007; 104:6672–6677. [PubMed: 17420458]
 15. Wang M, Patel HN, Wachter RM. X-ray diffraction analysis and molecular-replacement solution of the cyan fluorescent protein dsFP483. *Acta Cryst.* 2005; F61:922–924.
 16. Gross LA, Baird GS, Hoffman RC, Baldrige KK, Tsien RY. The structure of the chromophore within DsRed, a red fluorescent protein from coral. *Proc Natl Acad Sci USA.* 2000; 97:11990–11995. [PubMed: 11050230]
 17. Wachter RM. Chromogenic cross-link formation in green fluorescent protein. *Acc Chem Res.* 2007; 40:120–127. [PubMed: 17309193]
 18. Campbell RE, Tour O, Palmer AE, Steinbach PA, Baird GS, Zacharias DA, Tsien RY. A monomeric red fluorescent protein. *Proc Natl Acad Sci USA.* 2002; 99:7877–7882. [PubMed: 12060735]
 19. Yarbrough D, Wachter RM, Kallio K, Matz MV, Remington SJ. Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-Å resolution. *Proc Natl Acad Sci USA.* 2001; 98:462–467. [PubMed: 11209050]
 20. Wall MA, Socolich M, Ranganathan R. The structural basis for red fluorescence in the tetrameric GFP homolog DsRed. *Nature Structural Biology.* 2000; 7:1133–1138.
 21. Sniegowski JA, Phail ME, Wachter RM. Maturation efficiency, trypsin sensitivity, and optical properties of Arg96, Glu222, and Gly67 variants of green fluorescent protein. *Biochem Biophys Res Comm.* 2005; 332:657–663. [PubMed: 15894286]
 22. Remington SJ, Wachter RM, Yarbrough DK, Branchaud BP, Anderson DC, Kallio K, Lukyanov KA. zFP538, a yellow fluorescent protein from *Zoanthus*, contains a novel three-ring chromophore. *Biochemistry.* 2005; 44:202–212. [PubMed: 15628861]

23. Winn MD, Isupov MN, Murshudov GN. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Cryst.* 2001; D57:122–133.
24. Winn MD, Murshudov GN, Papiz MZ. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol.* 2003; 374:300–321. [PubMed: 14696379]
25. Ducros VM, Lewis RJ, Verma CS, Dodson EJ, Leonard G, Turkenburg JP, Murshudov GN, Wilkinson AJ, Brannigan JA. Crystal structure of GerE, the ultimate transcriptional regulator of spore formation in *Bacillus subtilis*. *J Mol Biol.* 2001; 306:759–771. [PubMed: 11243786]
26. Laskowski RA, Moss DS, Thornton JM. Main-chain bond lengths and bond angles in protein structures. *J Mol Biol.* 1993; 231:1049–1067. [PubMed: 8515464]
27. Ormo M, Cubitt AB, Kallio K, Gross LA, Tsien RY, Remington SJ. Crystal structure of the *Aequorea victoria* Green Fluorescent Protein. *Science.* 1996; 273:1392–1395. [PubMed: 8703075]
28. Burmeister WP. Structural changes in a cryo-cooled protein crystal owing to radiation damage. *Acta Cryst.* 2000; D56:328–341.
29. Banumathi S, Zwart PH, Ramagopal U, Dauter M, Dauter Z. Structural effects of radiation damage and its potential for phasing. *Acta Cryst.* 2004; D60:1085–1093.
30. Chattoraj M, King BA, Bublitz GU, Boxer SG. Ultra-fast excited state dynamics in Green Fluorescent Protein: Multiple states and proton transfer. *Proc Natl Acad Sci USA.* 1996; 93:8362–8367. [PubMed: 8710876]
31. Bell AF, He X, Wachter RM, Tonge PJ. Probing the ground state structure of the green fluorescent protein chromophore using Raman spectroscopy. *Biochemistry.* 2000; 39:4423–4431. [PubMed: 10757992]
32. He X, Bell AF, Tonge PJ. Isotopic labeling and normal-mode analysis of a model green fluorescent protein chromophore. *J Phys Chem B.* 2002; 106:6056–6066.
33. Elsliger MA, Wachter RM, Hanson GT, Kallio K, Remington SJ. Structural and spectral response of green fluorescent protein variants to changes in pH. *Biochemistry.* 1999; 38:5296–5301. [PubMed: 10220315]
34. Wachter RM, Elsliger MA, Kallio K, Hanson GT, Remington SJ. Structural basis of spectral shifts in the yellow-emission variants of Green Fluorescent Protein. *Structure.* 1998; 6:1267–1277. [PubMed: 9782051]
35. Shi X, Abbyad P, Shu X, Kallio K, Kanchanawong P, Childs W, Remington SJ, Boxer SG. Ultrafast excited-state dynamics in the green fluorescent protein variant S65T/H148D. 2. Unusual photophysical properties. *Biochemistry.* 2007; 46:12014–12025. [PubMed: 17918960]
36. Shu X, Kallio K, Shi X, Abbyad P, Kanchanawong P, Childs W, Boxer SG, Remington SJ. Ultrafast excited-state dynamics in the green fluorescent protein variant S65T/H148D. 1. Mutagenesis and structural studies. *Biochemistry.* 2007; 46:12005–12013. [PubMed: 17918959]
37. Leiderman P, Genosar L, Huppert D, Shu X, Remington SJ, Solntsev KM, Tolbert LM. Ultrafast excited-state dynamics in the green fluorescent protein variant S65T/H148D. 3. Short- and long-time dynamics of the excited-state proton transfer. *Biochemistry.* 2007; 46:12026–12036. [PubMed: 17918961]
38. Shu X, Leiderman P, Gepshtein R, Smith NR, Kallio K, Huppert D, Remington SJ. An alternative excited-state proton transfer pathway in green fluorescent protein variant S205V. *Protein Sci.* 2007; 16:2703–2710. [PubMed: 17965188]
39. Hosoi H, Mizuno H, Miyawaki A, Tahara T. Competition between energy and proton transfer in ultrafast excited-state dynamics of an oligomeric fluorescent protein Kaede. *J Phys Chem B.* 2006; 110:22853–22860. [PubMed: 17092037]
40. Brejc K, Sixma TK, Kitts PA, Kain SR, Tsien RY, Ormö M, Remington SJ. Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* Green Fluorescent Protein. *Proc Natl Acad Sci USA.* 1997; 94:2306–2311. [PubMed: 9122190]
41. Stoner-Ma D, Jaye AA, Ronayne KL, Nappa J, Meech SR, Tonge PJ. An alternate proton acceptor for excited-state proton transfer in green fluorescent protein: Rewiring GFP. *J Am Chem Soc.* 2008 in press.
42. Garcia-Parajo MF, Koopman M, van Dijk EMHP, Subramaniam V, van Hulst NF. The nature of fluorescence emission in the red fluorescent protein DsRed, revealed by single-molecule detection. *Proc Natl Acad Sci USA.* 2001; 98:14392–14397. [PubMed: 11724943]

43. Cotlet M, Hofkens J, Habuchi S, Dirix G, van Guyse M, Michiels J, Vanderleyden J, De Schryver FC. Identification of different emitting species in the red fluorescent protein DsRed by means of ensemble and single-molecule spectroscopy. *Proc Natl Acad Sci USA*. 2001; 98:14398–14403. [PubMed: 11724946]
44. Gurskaya NG, Savitsky AP, Yanushevich YG, Lukyanov SA, Lukyanov KA. 2001Color transitions in coral's fluorescent proteins by site-directed mutagenesis. *BMC Biochemistry*. : 2.<http://biomedcentral.com/1472-2091/2/6>
45. Labas YA, Gurskaya NG, Yanushevich YG, Fradkov AF, Lukyanov KA, Lukyanov SA, Matz MV. Diversity and evolution of the green fluorescent protein family. *Proc Natl Acad Sci USA*. 2002; 99:4256–4261. [PubMed: 11929996]
46. Lakowicz, JR. *Principles of Fluorescence Spectroscopy*. 2. Kluwer Academic/Plenum; New York: 1999.
47. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. 1997; 276:307–326.
48. Vagin A, Teplyakov A. An automated program for molecular replacement. *J Appl Cryst*. 1997; 30:1022–1025.
49. Yeates TO. Detecting and overcoming crystal twinning. *Methods Enzymol*. 1997; 276:344–358. [PubMed: 9048378]
50. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst*. 1997; D53:240–255.
51. Collaborative Computational Project 4 N. The CCP4 Suite: Programs for protein crystallography. *Acta Cryst*. 1994; D50:760–763.
52. McRee DE. XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J Struct Biol*. 1999; 125:156–165. [PubMed: 1022271]
53. Rudolph MG, Wingren C, Crowley MP, Chien Y-h, Wilson IA. Combined pseudo-merohedral twinning, non-crystallographic symmetry and pseudo-translation in a monoclinic crystal form of the gamma-delta T-cell ligand T10. *Acta Cryst*. 2004; D60:656–664.
54. Herbst-Irmer R, Sheldrick GM. Refinement of twinned structures with SHELEXL97. *Acta Cryst*. 1998; B54:443–449.
55. DeLano WL. 2002The PyMOL molecular graphics system. World Wide Web. <http://www.pymol.org>

Figure 1. CLUSTALW alignment of the primary amino acid sequences of selected cyan, red, and green fluorescent proteins. Amino acid sequence numbers are provided based on the alignment, whereas the residue numbers for individual clones vary depending on loop insertions (see text). Underlined = chromophore-forming residues. Positions implicated in formation of a quadrupole network of charges near the chromophore are denoted with a diamond (blue = positively charged; red = negatively charged; green = charge unknown; black = uncharged). Position 169, which is close to the chromophore hydroxyl, is shown in italics. GenBank Accession numbers: AY181556.1 (mcFP485), AF168421.1 (amFP486), AF168420.1 (dsFP483), AF168419.2 (DsRed), AF168423.1 (zFP538), and AF168424.1 (cFP484).

Figure 2. Normalized excitation and fluorescence emission spectra of dsFP483 and its single-site mutants, collected at room temperature. Protein solutions contained 50 mM HEPES pH 7.9, 300 mM NaCl, and 10 mM EDTA.

Figure 3.

Stereoview of the $|F_{\text{obs}}| - |F_{\text{calc}}|$ electron density omit map calculated to 2.1 Å resolution and contoured at 3.0 σ . The chromophore and select atoms of surrounding residues (as shown) were omitted from the model, and 10 cycles of maximum likelihood refinement were carried out before map calculation. The refined chromophore model is superimposed onto the omit map. The figure was generated using PyMol.55

Figure 4. Schematic representation of the dsFP483 chromophore and environment for Chain A. Panel a represents a schematic of the neutral chromophore, and Panel b that of the anionic chromophore. Dashed lines indicate likely hydrogen bonding interactions based on favorable geometry, with the estimated donor-acceptor distances indicated in Å. Ordered solvent molecules are indicated by W. Acidic residues are shown in red and basic residues in blue. Proposed proton positions (green) are shown for those sites deemed relevant for the assignment of charges (see text).

Figure 5. Structural comparison of the chromophore surroundings of dsFP483 (dark blue), amFP486 (cyan), zFP538-K66M (green), and DsRed (red). Structural superimposition of all chromophore atoms that are part of the π -system was carried out utilizing a least-squares algorithm available in the program Xfit.⁵¹ This figure was generated using PyMol.⁵⁵

Figure 6.

Raman spectra of dsFP483 (Cyan) and dsFP483-K70M (K70M Cyan) in aqueous solution buffered at pH 7.9, and at a protein concentration of 92 μ M. The solvent background has been subtracted. For comparison, spectra of GFP-S65T at pH 8.0 (chromophore anion) and pH 5.0 (neutral chromophore) are also shown.³¹

Figure 7. Expanded view of the Raman spectra shown in Figure 6 between 1550 and 1750 cm^{-1} (A, GFP-S65T anion; B, GFP-S65T neutral; C, dsFP483-K70M; D, dsFP483 wild-type).

Table 1
Optical properties of dsFP483 and variants. For comparison, data for amFP486^a are also shown.

Protein	Excitation $\lambda_{\text{max}}(1)$ (nm)	Excitation $\lambda_{\text{max}}(2)$ (nm)	Emission λ_{max} (nm)	Quantum Yield Φ_F	Effective Stokes Shift ^b (cm ⁻¹)
dsFP483, wild-type	437 (major)	~ 453 (minor)	483	0.78	2,179
dsFP483-H163Q	~ 443 (minor)	455 (major)	486	0.63	1,402
dsFP483-K70M	~ 445 (minor)	460 (major)	500	0.05	1,739
amFP486 ^a	~ 440 (minor)	450 (major)	486	0.71	1,646

^aHenderson et al. (ref. 8).
^bCalculated utilizing the overall excitation maximum.

Table 2**Final refinement parameters**

No. protein atoms/molecule	MolA: 1817 (residues 1 to 226)	
	MolB: 1757 (residues 7 to 226)	
	MolC: 1799 (residues 6 to 228)	
	MolD: 1764 (residues 7 to 228)	
	MolE: 1613 (residues 7 to 226)	
	MolF: 357 (residues 20–23, 25–27, 49–51, 53–63, 96–101, 103–108, 124–139, 160–164, 174–178)	
No. protein molecules in asymmetric unit	6	
No. solvent water molecules	567	
No. modeled ions	2 sodium ions	
Resolution used in refinement (Å)	188.98–2.09	
R _{cryst} (%) ^a	18.5	
R _{free} (%) ^a	21.7	
rms deviation bond lengths (Å)	0.011	
rms deviation bond angles (°)	1.520	
rms deviation chiral volumes (Å ³)	0.093	
Residual average main-chain B value (Å ²)	Mol A–F:	
	16.8, 18.1, 17.0, 18.0, 18.6, 24.0	
	Average 18.8 (not including HOHs)	
Residual average side-chain B value (Å ²)	Mol A–F:	
	18.7, 19.0, 18.8, 19.3, 18.7, 25.6	
	Average 20.0 (not including HOHs)	
rms deviation Ca (Å)	A–B 0.22	A–C 0.22
	A–D 0.23	A–E 0.27

^aR_{cryst} and R_{free} = $\sum_{\mathbf{h}} (||F(\mathbf{h})_{\text{obs}}| - |F(\mathbf{h})_{\text{calc}}||) / |F(\mathbf{h})_{\text{obs}}|$ for reflections in the working and test sets (5% of all data).

Table 3

TLS Parameters

Chain A	T	-0.2285	-0.2851	-0.2199	-0.0857	0.0039	0.0300		
	L	2.0284	1.7776	1.7584	0.7422	0.4345	0.2637		
	S	0.3108	-0.1118	0.2167	-0.0128	0.0824	-0.0859	0.1063	-0.0558
Chain B	T	0.0273	0.2659	-0.2019	-0.4305	-0.0101	0.0045		
	L	2.4021	1.3039	2.6036	0.6062	-0.2916	0.0805		
	S	0.8846	-0.4152	0.8505	-0.1145	0.0265	-0.4013	0.1420	-0.0954
Chain C	T	-0.2213	-0.3215	-0.1663	-0.0682	-0.0202	0.0373		
	L	2.6920	1.7388	1.8018	1.2512	0.3036	0.3327		
	S	0.4013	-0.3150	0.2042	0.3141	0.1401	-0.0742	-0.1951	0.0437
Chain D	T	-0.0223	0.2681	-0.1632	-0.4237	0.0206	0.0398		
	L	2.1205	1.3536	2.2270	0.7300	0.4406	0.3259		
	S	0.9105	-0.5308	0.8043	0.1203	-0.0564	-0.3519	-0.1649	0.4048
Chain E	T	0.2635	0.6416	-0.1448	-0.2918	-0.0027	-0.0096		
	L	3.7461	2.3572	8.4972	0.2685	2.1586	-0.2861		
	S	-0.9225	0.0786	-0.5353	0.2312	0.0170	0.3935	-0.6268	-0.3118
Chain F	T	0.5377	0.4814	0.5033	-0.1427	-0.0029	-0.0262		
	L	6.6629	13.9125	18.2200	4.0426	2.5396	-7.6264		
	S	-0.6879	-0.2304	-0.2065	1.2579	-0.4037	-0.0173	-0.5766	-0.0646

The T, L, and S values are given for each rigid body group, and correspond to the T-matrix (T11, T22, T33, T12, T13, T23), the L-matrix (L11, L22, L33, L12, L13, L23), and the S-matrix (S11, S12, S13, S21, S22, S23, S31, S32, S33). The large values for L in chains E and F indicate large rotational motion for the corresponding chains, resulting in poor electron density for these chains.