Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The 1.6 Å resolution structure of a FRET-optimized Cerulean fluorescent protein

Jennifer L. Watkins,[a] Hanseong Kim,[a] Michele L. Markwardt,[b] Liqing Chen,[a] Raimund Fromme,[a] Mark A. Rizzo[b] and Rebekka M. Wachter[a]*

[a]Department of Chemistry and Biochemistry, Arizona State University, Tempe, AZ 85287-1604, USA, and [b]Department of Physiology, University of Maryland School of Medicine, Baltimore, MD 21201-1559, USA

Correspondence e-mail: rwachter@asu.edu

Received 8 August 2012
Accepted 16 January 2013

PDB Reference: mCerulean3, 4en1

Genetically encoded cyan fluorescent proteins (CFPs) bearing a tryptophan-derived chromophore are commonly used as energy-donor probes in Förster resonance energy transfer (FRET) experiments useful in live cell-imaging applications. In recent years, significant effort has been expended on eliminating the structural and excited-state heterogeneity of these proteins, which has been linked to undesirable photophysical properties. Recently, mCerulean3, a descendant of enhanced CFP, was introduced as an optimized FRET donor protein with a superior quantum yield of 0.87. Here, the 1.6 Å resolution X-ray structure of mCerulean3 is reported. The chromophore is shown to adopt a planar *trans* configuration at low pH values, indicating that the acid-induced isomerization of Cerulean has been eliminated. β-Strand 7 appears to be well ordered in a single conformation, indicating a loss of conformational heterogeneity in the vicinity of the chromophore. Although the side chains of Ile146 and Leu167 appear to exist in two rotamer states, they are found to be well packed against the indole group of the chromophore. The Ser65 reversion mutation allows improved side-chain packing of Leu220. A structural comparison with mTurquoise2 is presented and additional engineering strategies are discussed.

1. Introduction

A large variety of GFP-like proteins have been developed for imaging applications in the live cell environment (Shaner *et al.*, 2007; Chudakov *et al.*, 2010). Based on the multitude of available colours, Förster resonance energy transfer (FRET) has become a popular method to monitor macromolecular proximity on a subcellular level (Vogel *et al.*, 2006; Piston & Kremers, 2007). In FRET, one measures the nonradiative transfer of excited-state energy between a donor and an acceptor fluorophore with overlapping emission and absorption spectra, respectively (Miyawaki, 2011; Piston & Kremers, 2007; Vogel *et al.*, 2006). The most common fluorescent protein pair used in FRET experiments involves variants of enhanced cyan fluorescent protein (ECFP) and yellow fluorescent protein (YFP) (Vogel *et al.*, 2006; Piston & Kremers, 2007), both of which are engineered descendants of green fluorescent protein (GFP). The chromophore of GFP consists of a *para*-hydroxybenzylidene imidazolinone group formed autocatalytically from a Ser-Tyr-Gly tripeptide (Tsien, 1998). Although a variety of naturally occurring cyan fluorescent proteins have been characterized (Malo *et al.*, 2008), the central tyrosine residue has been replaced by a tryptophan in the ECFP line of engineered GFPs. ECFP emits light maximally at 475 nm, while YFP absorbs light in a similar region of the spectrum (λ_{max} = 515 nm). Although the two proteins

Table 1
Residue substitutions in ECFP-derived proteins compared with avGFP.

Roman type, avGFP (*Aquorea victoria* GFP) wild-type residue; bold type, residues engineered to improve the photophysical properties (the list of substitutions is not comprehensive).

	ECFP†	Cerulean‡	mCerulean2§	mCerulean3§	mTurquoise§	mTurquoise-GL¶	mTurquoise2¶
Quantum yield††	0.36	0.48	0.60	0.87	0.84	0.82	0.93
Residue							
64	**L**	**L**	**L**	**L**	**L**	**L**	**L**
65	**T**	**T**	**T**	S	S	S	S
66	**W**	**W**	**W**	**W**	**W**	**W**	**W**
72	S	**A**	**A**	**A**	**A**	**A**	**A**
145	Y	**A**	**A**	**A**	Y	Y	Y
146	**I**	**I**	**I**	**I**	**I**	**I**	**F**
147	S	S	**H**	**H**	S	S	S
148	H	**D**	**G**	**G**	**D**	**G**	**D**
153	**T**	**T**	**T**	**T**	**T**	**T**	**T**
163	**A**	**A**	**A**	**A**	**A**	**A**	**A**
166	K	K	**G**	**G**	K	K	K
167	I	I	**L**	**L**	I	I	I
168	R	R	**N**	**N**	R	R	R
169	H	H	**C**	**C**	H	H	H
175	S	S	S	S	**G**	**G**	**G**
224	V	V	V	V	V	**L**	V

† Heim & Tsien (1996). ‡ Rizzo *et al.* (2004). § Markwardt *et al.* (2011). ¶ Goedhart *et al.* (2010). †† Quantum yield measurements, as reported from the references cited, were generally performed at ambient temperature under solution conditions.

fulfill the requirement for significant spectral overlap, ECFP exhibits only dim fluorescence, reducing the maximally obtainable signal-to-noise ratio. Additionally, the fluorescence lifetime of ECFP is multi-exponential, indicating the presence of multiple excited-state decay pathways that complicate data analysis in lifetime imaging applications (Rizzo *et al.*, 2004). In 2003, the ECFP crystal structure was reported to contain two conformations of β-strand 7 (residues 144–151), with substantial positional heterogeneity for His148 and Tyr145 (Bae *et al.*, 2003). In one conformation, the Tyr145 side chain was reported to point into the β-barrel, while the side chain of His148 was found to be solvent-exposed. In the second conformation, the side-chain positions were found to be reversed (Bae *et al.*, 2003).

Improvements in interior packing interactions are generally thought to decrease the dynamic motions of buried protein groups, so that collisional quenching of a buried fluorophore is reduced. The effort to improve ECFP for more effective use in FRET experiments began by introducing site-directed mutations to eliminate structural heterogeneity. In subsequent work, this approach was expanded to include semi-random mutagenesis in combination with colony screening. By using the ECFP X-ray model as a basis for structure-guided engineering, several cyan variants with improved photophysical properties were developed. One line of ECFP descendants was termed Cerulean (Rizzo *et al.*, 2004; Markwardt *et al.*, 2011), whereas another line was termed Super Cyan Fluorescent Protein (SCFP), which evolved into the Turquoise set of fluorescent proteins (Goedhart *et al.*, 2010, 2012). Initially, both lines included the H148G substitution, which was retained in the Cerulean line but was ultimately replaced by an aspartic acid in mTurquoise2, the most highly evolved variant along this line (Table 1). In addition, both lines involved the reversion substitution T65S. Ser65 represents a wild-type residue in *Aequorea victora* GFP and was retained in some of the early versions of cyan fluorescent proteins (Cubitt *et al.*, 1999). Recently, the Cerulean line culminated in mCerulean3, with a quantum yield of 0.87 (Markwardt *et al.*, 2011), and the SCFP/Turquoise line in mTurquoise2, with a quantum yield of 0.93 (Goedhart *et al.*, 2012). Key residue substitutions of selected variants along each line are summarized in Table 1 together with their reported quantum yields.

Along the Cerulean line, the H148D substitution was introduced into ECFP to stabilize a single conformation of β-strand 7 (Rizzo *et al.*, 2004). In combination with S72A and Y145A (Table 1), this variant, which was termed Cerulean, proved to be 2.5 times brighter than ECFP and was shown to be useful for lifetime imaging in FRET applications (Rizzo *et al.*, 2004). Crystallization of Cerulean revealed the existence of two different chromophore conformations that depended on the pH of the crystallization buffer. At pH 5, the chromophore appeared to exist in an unusual *cis* conformation in which the indole group was rotated around the Trp66 $C^{\beta}-C^{\gamma}$ bond to generate a close contact to the imidazolinone ring (Fig. 1; Malo *et al.*, 2007). This conformation provided a structural basis for the observed fourfold loss of fluorescence under acidic conditions. Chromophore isomerization was proposed to occur hand-in-hand with movement of the Asp148 side chain from the inside of the barrel at low pH to the outside at high pH (Malo *et al.*, 2007). Subsequently, a pH 8.0 Cerulean X-ray model was reported to bear a chromophore in the *trans* or *Z*,*Z*′ conformation (Fig. 1), confirming acid-induced isomerization (Lelimousin *et al.*, 2009).

To augment the shape complementarity of the protein pocket to the nearby tryptophan-derived chromophore, bulky and mobile residues were replaced, so that radiative excited-state relaxation would be favoured over thermal relaxation. Random pairwise mutations were introduced into the seventh and eighth strands of the β-barrel, leading to mCerulean2 with the substitutions S147H/D148G/K166G/I167L/R168N/H169C. This variant was shown to exhibit a 25% increase in quantum yield and was reported to be 1.3 times as bright as Cerulean (Table 1; Markwardt *et al.*, 2011). However, the six mutations in mCerulean2 seemed to cause an increase in photoswitching behavior, thus decreasing the photostability (Markwardt *et al.*, 2011). To relieve the conformational strain that may reduce chromophore rigidity, the T65S wild-type reversion substitution was incorporated, yielding the variant mCerulean3. The

S65T substitution was originally introduced into wild-type GFP, where it improved fluorescence by modulating the chromophore pK_a value (Ormö *et al.*, 1996; Yang *et al.*, 1996; Brejc *et al.*, 1997). More recently, reversion to wild-type Ser65 has proven to be useful in some blue variants (Mena *et al.*, 2006; Kremers *et al.*, 2007). In mCerulean3, the T65S reversion dramatically improved the brightness and the quantum yield, while eliminating photoswitching behavior almost entirely (Markwardt *et al.*, 2011). Here, we present the X-ray structure of mCerulean3 and provide a structural rationale for its superior optical characteristics that are useful in FRET applications (Markwardt *et al.*, 2011).

2. Materials and methods

2.1. Crystallization

The mCerulean3 protein was expressed and purified as described previously (Markwardt *et al.*, 2011). The protein stock consisted of 20 m*M* HEPES pH 7.9 with 20 m*M* NaCl and about 12 mg ml^{-1} protein. Crystallization screens were carried out using the hanging-drop vapor-diffusion method with drops consisting of 2 µl protein solution and 1 µl mother liquor. Diffraction-quality crystals were obtained from wells with mother liquor consisting of 0.08 *M* sodium acetate trihydrate pH 4.7, 0.16 *M* ammonium sulfate, 9%(*w*/*v*) PEG 4000, 19%(*v*/*v*) glycerol.

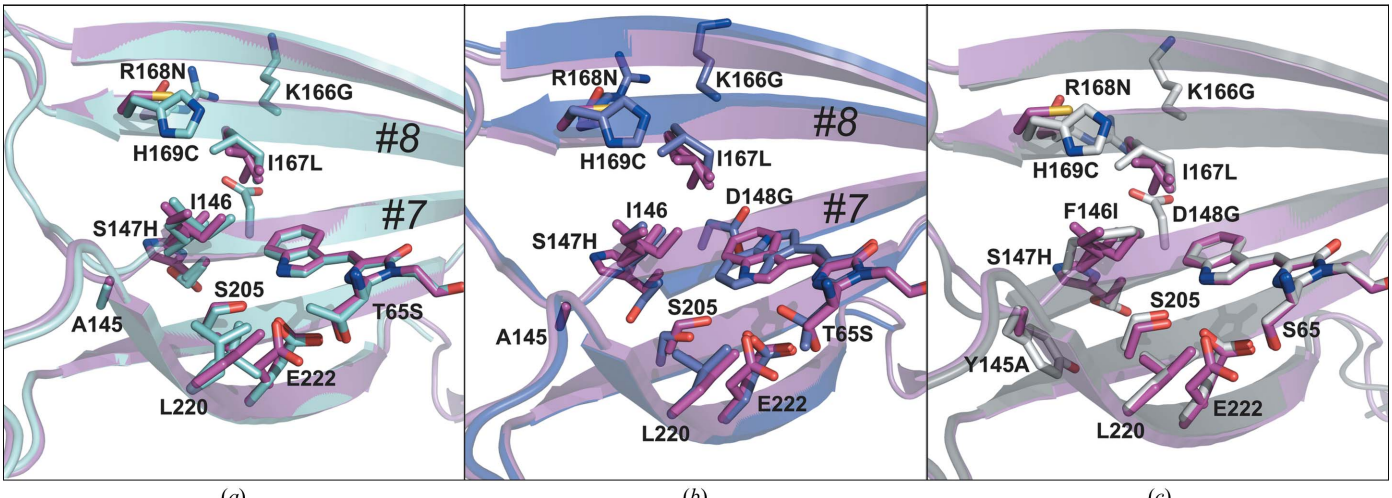
2.2. Data collection and processing

The mounted crystal was flash-cooled in liquid nitrogen before exposure to X-rays. Additional cryoprotectant was not needed because the mother liquor contained 19% glycerol. Diffraction data were collected at 100 K on beamline 8.2.1 at the Advanced Light Source (ALS; Berkeley, California, USA), which is equipped with an ADSC 315 detector (3072 × 3072 pixels). Data were indexed and scaled using *HKL*-2000 (Otwinowski & Minor, 1997).

2.3. Structure determination and refinement

Data processing was carried out with *HKL*-2000 (Otwinowski & Minor, 1997) and 5% of the reflections were flagged for R_{free} calculations. *Phaser* v.2.3.0 (McCoy *et al.*, 2007) from the *CCP*4 program suite v.6.2.0 (Winn *et al.*, 2011) was employed for molecular replacement using the pH 8 structure of Cerulean (PDB entry 2wso; Lelimousin *et al.*, 2009) as a search model. This model was selected based on its high sequence identity to mCerulean3 and its high resolution of 1.37 Å. Iterative rounds of refinement and model adjustment were performed in the resolution range 30–1.62 Å using *REFMAC*5 v.5.6.0117 (Murshudov *et al.*, 2011) and *Coot* v.0.6.2 (Emsley & Cowtan, 2004). No NCS restraints were utilized to refine the two protein chains in the asymmetric unit. Isotropic *B*-factor refinement was carried out on individual atoms using *REFMAC* and the Wilson *B* factor was calculated using *CTRUNCATE* v.1.0.12 (Winn *et al.*, 2011).

The chromophore was modeled based on the CRF chromophore geometry of the Cerulean model 2wso, which describes the tryptophan-derived chromophore (Lelimousin *et al.*, 2009). The CRF geometry file was modified to reflect that position 65 has been reverted from a threonine to a serine residue in mCerulean3. Two protein chains were modeled in the asymmetric unit, with chain *A* comprising residues 1–238 and chain *B* residues 3–230 of the 238 residues of the mature protein, which was crystallized in the presence of an N-terminally attached 6×His tag. After refinement of the protein chains and internal chromophores, solvent molecules were modeled into 3σ positive difference density peaks where suitable hydrogen-bonding partners were available. In total,



(*a*) (*b*) (*c*)

Figure 1
C$^{\alpha}$ superposition of the structure of mCerulean3 (magenta) onto (*a*) the high-pH X-ray structure of Cerulean (cyan; PDB entry 2wso; Lelimousin *et al.*, 2009), (*b*) the low-pH X-ray structure of Cerulean (blue; PDB entry 2q57; Malo *et al.*, 2007) and (*c*) the X-ray structure of mTurquoise2 (grey; PDB entry 3ztf; Goedhart *et al.*, 2012). Side chains are shown for the residues discussed in the text, with labels indicating substitutions found in mCerulean3 compared with Cerulean (*a*, *b*) and mTurquoise2 (*c*). In all cases, the amino acid following the position number indicates the mCerulean3 residue. β-Strands 7 and 8 are labeled #7 and #8.

Table 2
X-ray data-collection and refinement statistics.

Values in parentheses are for the highest resolution bin.

Data collection	
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 79.41, b = 88.73, c = 94.74
Wavelength (Å)	0.97950
Resolution limits (Å)	50.0–1.62 (1.65–1.62)
Total reflections	250772
Unique reflections	84496
Multiplicity	3.0 (3.0)
Completeness (%)†	98.7 (99.7)
$\langle I/\sigma(I)\rangle$	14.5 (2.04)
$R_{\rm merge}$‡ (%)	7.1 (47.7)
Wilson B factor (Å)	18.4
Refinement	
Resolution range (Å)	29.983–1.620
No. of reflections	79895
$R_{\rm cryst}$§ (%)	17.77
$R_{\rm free}$§ (%)	20.55
R.m.s. deviation	
Bond lengths (Å)	0.0176
Bond angles (°)	2.38
Average B factor, protein atoms (Å^2)	20.27
B-factor deviation (Å^2)	
C^α atoms (isotropic)	18.68
Side chain (isotropic)	20.35
Average B factor, solvent atoms (Å^2)	29.92
No. of protein atoms	3745
No. of waters	372
No. of other solvent atoms (see §2)	90
Total No. of atoms	4207
No. of protein chains per asymmetric unit	2
Ramachandran analysis (%)	
Favored	91.7
Allowed	8.3
Disallowed	0
PDB code	4en1

† Completeness is the ratio of the number of observed reflections with $I > 0$ to the theoretically possible number of reflections. ‡ $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the average of individual measurements of $I(hkl)$. § $R_{\rm cryst}$ and $R_{\rm free} = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}||/\sum_{hkl}|F_{\rm obs}|$ for reflections in the working and test sets (5% of all data), respectively.

372 water molecules, one sulfate ion, ten glycerol molecules, three diethyleneglycol molecules and one acetate molecule were modeled. Figures were created using *PyMOL* and data-collection and refinement statistics are shown in Table 2.

3. Results and discussion

We have determined the crystal structure of mCerulean3 and refined the model to 1.62 Å resolution. The protein crystallized in space group $P2_12_12_1$, with two protein chains in the asymmetric unit. Refinement gave an R factor of 17.8% and an $R_{\rm free}$ of 20.5%, with good geometry (Table 2). The overall fold is the standard 11-stranded β-barrel typical of GFP-like proteins. Although monomeric in solution (Markwardt *et al.*, 2011), the two protein chains form a molecular interface in the crystal involving β-strands 9, 10 and 11, such that the long axes of the β-barrels are aligned at approximately right angles with respect to each other. The two chains were refined in the absence of noncrystallographic symmetry (NCS) restraints and therefore represent independent views of the structure of mCerulean3. A structural superposition of the C^α atoms of each chain gave an r.m.s.d. of 0.17 Å, indicating that the two models are nearly identical. The coordinates for the X-ray model have been deposited in the Protein Data Bank as entry 4en1.

3.1. Elimination of conformational heterogeneity along β-strand 7

Overall, the main-chain structure of mCerulean3 was found to be similar to that of Cerulean (C^α r.m.s.d. of 0.22 Å). However, the structure indicated that the extensive modification of strands 7 (S147H/D148G) and 8 (K166G/I167L/R168N/H169C) has provided the desired result of stabilizing strand 7 in a single conformation. In the Cerulean parent, strand 7 only adopted a single conformation at low pH (Malo *et al.*, 2007), whereas two conformational states were observed at positions 143–146 at more elevated pH values. As these residues reside near the indole group of the chromophore, their positional disorder was proposed to be the underlying cause of the excited-state heterogeneity (Lelimousin *et al.*, 2009). In contrast, the structure of mCerulean3 exhibits well defined and continuous electron density for all main-chain atoms of β-strand 7 (residues 144–151) and this density is well accounted for by modeling a single backbone trace. Therefore, the redesign of packing interactions near the chromophore (Fig. 1) has resulted in a significant reduction of dynamic motion in the vicinity of the chromophore, as further supported by a B-factor analysis. The average B factor for the strand 7 C^α atoms was calculated to be 18.3 Å^2, which is close to the average B factor of 18.7 Å^2 for all C^α atoms in the model, suggesting that the β-barrel is evenly rigid and lacks additional thermal motion along strand 7. A structural superposition of β-strands 7 (residues 143–153) and 8 (residues 161–170) of chain A onto the equivalent residues of chain B provides a C^α r.m.s.d. of 0.21 Å, suggesting that the C^α coordinate error of strands 7 and 8 is similar to that calculated for the entire protein chain (r.m.s.d. of 0.17 Å).

Compared with the high-pH Cerulean conformation A (Lelimousin *et al.*, 2009), residues 144–146 of mCerulean3 are displaced towards the outside, with a maximum shift of 0.8 Å for residue 145. Their new position is almost perfectly intermediate to Cerulean conformations A and B, supporting reduced dynamic motion without significant structural perturbation. Residues 167–169 of strand 8 are also displaced towards the outside, with a maximum shift of 0.8 Å for residue 167. Therefore, the structure of mCerulean3 demonstrates that the use of pairwise mutagenesis of residues facing each other has been highly successful in improving the side-chain packing near the chromophore. It appears that the single conformation of strand 7 is primarily responsible for the increased brightness of mCerulean3 compared with the Cerulean parent protein.

3.2. Chromophore '*trans*' isomer, planarity and packing interactions

Although mCerulean3 was crystallized at low pH, the chromophore exists as the Z,Z' or *trans* isomer with respect to rotation around the β-methylene bridge that connects the

imidazolinone ring to the indole group (Fig. 2). This is the same isomer as was observed in the Cerulean structure at pH 8 (Fig. 1; Lelimousin *et al.*, 2009). To test the crystallographic model, the planarity restraints of the chromophore π-system were released and the model was subjected to additional rounds of refinement. The resulting chromophore structure demonstrated no significant deviation from the constrained model, indicating that the planar *trans* geometry of the refined structure is a result of the experimental data rather than of model bias (Fig. 2). The elimination of pH-dependent chromophore isomerization is further supported by the lack of any pH-dependent spectral changes in mCerulean3 (Markwardt *et al.*, 2011). Upon acidification, the Cerulean parent protein was originally reported to undergo blue-shifting by 10–15 nm (Malo *et al.*, 2007). Absorbance spectra collected from single crystals in 20% glycerol at pH 5 and 100 K were shown to be essentially identical to solution spectra collected in the absence of glycerol at the same pH but at room temperature (Malo *et al.*, 2007). The insensitivity of the Cerulean spectra to glycerol was recently verified independently (von Stetten *et al.*, 2012) and spectral changes owing to the crystalline state or owing to glycerol are not expected for mCerulean3 either. In the X-ray structure of mCerulean3 there is no indication of a trapped organic group within the chromophore cavity. The removal of the titratable aspartic acid side chain *via* the D148G mutation has effectively eliminated the pH-dependent conformational switch originally observed in Cerulean (Malo *et al.*, 2007). Acid-induced chromophore isomerization is absent in mCerulean3 because steric clashes with a protonated Asp148 have been removed.

The improvements in molecular packing around the chromophore of mCerulean3 are partly owing to the I167L substitution. In and of itself, the I167L substitution has the effect of moving one methyl group closer to the chromophore, increasing the available bulk near the chromophore so that isomerization is hindered. The hydrophobic side chains of Ile146 (strand 7) and I167L (strand 8) are found to be in van der Waals contact with each other and are also in van der Waals contact with the six-membered ring of the indole group of the chromophore. Both side chains are modeled with dual occupancy, as the electron density is clearly consistent with multiple conformational states (Fig. 1). For Ile146, the distances of closest approach were found to be 4.2 and 4.0 Å for the two conformational states, whereas for I167L the distances of closest approach were 3.8 and 3.3 Å. Therefore, residues 146 and 167 provide a substantial number of hydrophobic contacts with the chromophore indole face opposite the side facing Glu222, Thr203 and Ser205 (Fig. 1).

3.3. Structural comparison with mTurquoise2

Along the Turquoise line of cyan fluorescent proteins, mTurquoise2, a variant that exhibits a remarkable quantum yield of 0.93, was recently isolated (Goedhart *et al.*, 2012). The differences between mCerulean3 and mTurquoise2 consist of a total of nine residues in strands 7 and 8 (Table 1). A number of Cerulean residues that were mutated in mCerulean3 were retained in mTurquoise2 (residues 147, 148, 166, 167, 168 and 169), whereas residues 145, 146 and 175 were retained in mCerulean3 but were modified to Tyr, Phe and Gly, respectively, in mTurquoise2 (Fig. 1*c*). The high quantum yield of mTurquoise2 was primarily attributed to optimized van der Waals interactions of the I146F substitution with the indole ring of the chromophore (Goedhart *et al.*, 2012). This residue was replaced based on its high degree of positional disorder in a number of variants such as ECFP, high-pH Cerulean and SCFP3A. Although the phenylalanine aromatic group appears to occupy a similar volume element to Ile146 in mCerulean3, Phe146 may only pack well in combination with Ile167, not with the Leu167 present in mCerulean3 (Fig. 1*c*). Since Ile146 also occupies multiple positions in mCerulean3, targeting this position for further mutagenesis in combination with its contact residues may be a fruitful strategy for making additional improvements to the quantum yield.

3.4. Effect of the reversion of residue 65 to wild-type GFP (T65S)

mCerulean3 differs from mCerulean2 by one substitution: the reversion of position 65 to the wild-type serine residue (T65S). As this reversion is also present in mTurquoise2, the set of residues Ser205, Leu220, Glu222 and Ser65 located near the 'lower' face of the chromophore remain identical in mCerulean3 and mTurquoise2 (Fig. 1*c*). The removal of the threonine-derived methyl group from mCerulean2 was predicted to allow the chromophore to fit better into the central cavity and was shown to result in a 25% increase in brightness (Markwardt *et al.*, 2011). Unfortunately, diffraction-quality crystals of mCerulean2 have not yet been obtained, despite extensive efforts. Regardless, the effects of the T65S substitution can be examined reasonably well by comparing the structure of mCerulean3 with that of its Cerulean parent protein (Figs. 1*a* and 1*b*). Unlike Cerulean, the side chain of Glu222 in

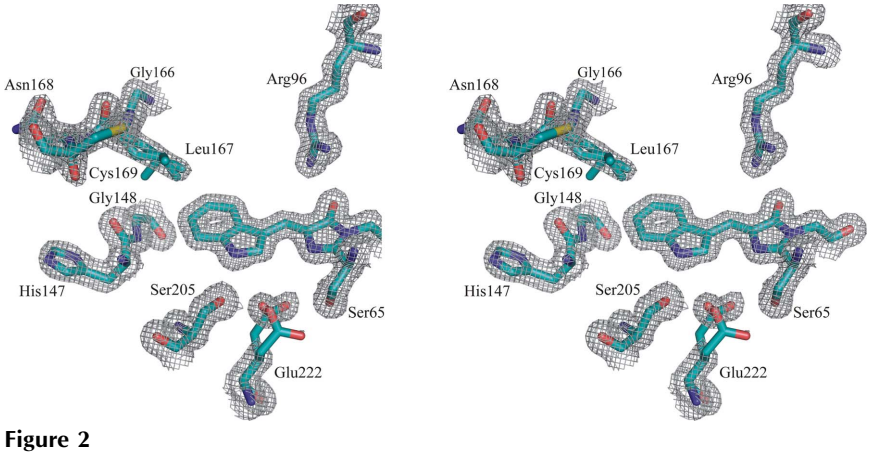


Figure 2
Cross-eyed stereoview of the mCerulean3 chromophore and its surrounding residues. The $2F_o - F_c$ electron density is shown at 2.2σ.

mCerulean3 is found to adopt two rotamer conformations (Fig. 3). The major conformation represents about 60% of the side-chain population and its position is essentially identical to that in Cerulean and mTurquoise. In this conformation, one of the Glu222 carboxyl O atoms forms a hydrogen bond to the Ser65 hydroxyl and the other forms a hydrogen bond to the Ser205 hydroxyl. The interaction with Ser205 involves a short hydrogen bond that refines to 2.4 Å, suggesting tight side-chain packing. In turn, Ser205 is hydrogen-bonded to the indole N atom of the chromophore, which serves as a donor with nearly ideal hydrogen-bonding geometry (2.8 Å). The second, minor, conformation of Glu222 exhibits an occupancy of about 40% (Fig. 3). In this rotamer, the carboxylate is rotated by 70° around the $C^{\alpha}-C^{\beta}$ bond. Although the hydrogen-bonding network of the major conformation is maintained in the minor conformation, the Ser205–Glu222 interaction refines to a normal hydrogen-bonding distance of 2.8 Å, whereas the hydrogen bond to Ser65 now appears to be very short (2.5 Å). Although this feature could be interpreted in terms of a particularly strong interaction with the Ser65 hydroxyl, the sandwiching of the Glu222 carboxyl group between Ser65 and Ser202 suggests that the hydrogen-bonded contacts may be correlated in crystallographic refinement, resulting in one short and one normal hydrogen bond in each of the two conformations. Although this notion is consistent with residual molecular crowding near the heterocycle of the chromophore, the removal of a methyl group by means of the T65S substitution appears to relieve some steric restrictions around Glu222, so that additional rotamer positions become populated.

The T65S substitution also appears to affect Leu220. In the Cerulean parent, the Leu220 side chain appears to adopt two rotamer conformations both in van der Waals contact with the methyl group of Thr65 (Fig. 1*a*). However, in mCerulean3, the Leu220 side chain adopts only one conformation and is found in van der Waals contact with both C^{β} of Ser65 and C^{β} of Ser205, similar to the geometry reported in mTurquoise2 (Fig. 1*c*; Goedhart *et al.*, 2012). Therefore, it appears that the removal of a methyl group from position 65 allows better packing of Leu220, while reducing steric crowding of Glu222 and possibly introducing an alternate short hydrogen bond to the chromophore. In combination, the variety of interactions affected by T65S suggests that there may be room for additional improvements *via* synergistic optimization of residues arranged around Ser65 while conserving the catalytically essential Glu222 (Sniegowski *et al.*, 2005).

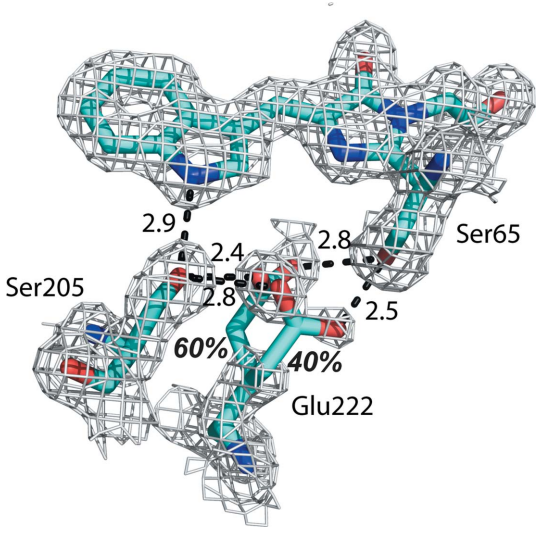


Figure 3
Model of the two alternate conformations of Glu222 in mCerulean3. The $2F_o - F_c$ electron density is shown at 1.5σ. Conformation *A* is dominant, with 60% occupancy. Inferred hydrogen bonds are shown by dashed lines, with distances shown in Å.

4. Summary and conclusions

The high fluorescence efficiency and the outstanding photostability of mCerulean3 renders this cyan variant a superior probe for use as a donor fluorophore in quantitative FRET experiments (Markwardt *et al.*, 2011). The X-ray structure of mCerulean3 presented here is consistent with a *trans* chromophore that is well packed in the interior of the protein. The set of substitutions introduced into mCerulean3 seems to have eliminated unfavorable steric interactions, reducing the likelihood of fluorescence quenching by *cis–trans* isomerization and hula-twist motions, so that a planar chromophore structure associated with a high quantum yield is favored (Ong *et al.*, 2011).

Furthermore, the structural model suggests that additional improvements in interior side-chain packing may be possible to push the quantum yield even higher. Residues with multiple side-chain conformational states may be targeted for synergistic mutagenesis to stabilize conformations with optimized steric interactions and with a large number of van der Waals interactions with the chromophore in order to provide a protein cavity with a high degree of shape complementary to the tryptophan-derived chromophore. The strategic engineering of planar chromophore structures by the elimination of dynamic motion associated with conformational exchange may be considered as a more general approach for generating brighter fluorescent proteins.

mCerulean3 exhibits a quantum yield of 0.87 (Markwardt *et al.*, 2011), which is close to the quantum yield of 0.93 recently reported for mTurquoise2 (Goedhart *et al.*, 2012). Although these variants differ at nine residue positions near the chromophore, both proteins exceed the quantum yield of the first generation of cyan fluorescent proteins by about 2.5-fold (Table 1; Cubitt *et al.*, 1995). Therefore, the work presented here demonstrates that different combinations of correlated residue substitutions can accomplish the same end goal of optimized photophysical properties in fluorescent proteins. The structural comparisons carried out here provide support for the idea that genetic engineering of protein-sequence space can provide several parallel solutions to the same problem.

This work was supported by a grant from the National Science Foundation (NSF MCB-0615938) to RMW and a grant from the National Institutes of Health (NIH/NIDDK

1R01DK077140) to MAR. X-ray diffraction data collection was carried out on Advanced Light Source (ALS) beamline 8.2.1, which is supported by the Department of Energy.

References

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. & Budisa, N. (2003). *J. Mol. Biol.* **328**, 1071–1081.

Brejc, K., Sixma, T. K., Kitts, P. A., Kain, S. R., Tsien, R. Y., Ormö, M. & Remington, S. J. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 2306–2311.

Chudakov, D. M., Matz, M. V., Lukyanov, S. & Lukyanov, K. A. (2010). *Physiol. Rev.* **90**, 1103–1163.

Cubitt, A. B., Heim, R., Adams, S. R., Boyd, A. E., Gross, L. A. & Tsien, R. Y. (1995). *Trends Biochem. Sci.* **20**, 448–455.

Cubitt, A. B., Woollenweber, L. A. & Heim, R. (1999). *Methods Cell Biol.* **58**, 19–30.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Goedhart, J., van Weeren, L., Hink, M. A., Vischer, N. O., Jalink, K. & Gadella, T. W. (2010). *Nature Methods*, **7**, 137–139.

Goedhart, J., von Stetten, D., Noirclerc-Savoye, M., Lelimousin, M., Joosen, L., Hink, M. A., van Weeren, L., Gadella, T. W. & Royant, A. (2012). *Nature Commun.* **3**, 751.

Heim, R. & Tsien, R. Y. (1996). *Curr. Biol.* **6**, 178–182.

Kremers, G.-J., Goedhart, J., van den Heuvel, D. J., Gerritsen, H. C. & Gadella, T. W. (2007). *Biochemistry*, **46**, 3775–3783.

Lelimousin, M., Noirclerc-Savoye, M., Lazareno-Saez, C., Paetzold, B., Le Vot, S., Chazal, R., Macheboeuf, P., Field, M. J., Bourgeois, D. & Royant, A. (2009). *Biochemistry*, **48**, 10038–10046.

Malo, G. D., Pouwels, L. J., Wang, M., Weichsel, A., Montfort, W. R., Rizzo, M. A., Piston, D. W. & Wachter, R. M. (2007). *Biochemistry*, **46**, 9865–9873.

Malo, G. D., Wang, M., Wu, D., Stelling, A. L., Tonge, P. J. & Wachter, R. M. (2008). *J. Mol. Biol.* **378**, 869–884.

Markwardt, M. L., Kremers, G.-J., Kraft, C. A., Ray, K., Cranfill, P. J. C., Wilson, K. A., Day, R. N., Wachter, R. M., Davidson, M. W. & Rizzo, M. A. (2011). *PLos One*, **6**, e17896.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Mena, M. A., Treynor, T. P., Mayo, S. L. & Daugherty, P. S. (2006). *Nature Biotechnol.* **24**, 1569–1571.

Miyawaki, A. (2011). *Annu. Rev. Biochem.* **80**, 357–373.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Ong, W. J., Alvarez, S., Leroux, I. E., Shahid, R. S., Samma, A. A., Peshkepija, P., Morgan, A. L., Mulcahy, S. & Zimmer, M. (2011). *Mol. Biosyst.* **7**, 984–992.

Ormö, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y. & Remington, S. J. (1996). *Science*, **273**, 1392–1395.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Piston, D. W. & Kremers, G.-J. (2007). *Trends Biochem. Sci.* **32**, 407–414.

Rizzo, M. A., Springer, G. H., Granada, B. & Piston, D. W. (2004). *Nature Biotechnol.* **22**, 445–449.

Shaner, N. C., Patterson, G. H. & Davidson, M. W. (2007). *J. Cell Sci.* **120**, 4247–4260.

Sniegowski, J. A., Lappe, J. W., Patel, H. N., Huffman, H. A. & Wachter, R. M. (2005). *J. Biol. Chem.* **280**, 26248–26255.

Stetten, D. von, Batot, G. O., Noirclerc-Savoye, M. & Royant, A. (2012). *Acta Cryst.* D**68**, 1578–1583.

Tsien, R. Y. (1998). *Annu. Rev. Biochem.* **67**, 509–544.

Vogel, S. S., Thaler, C. & Koushik, S. V. (2006). *Sci. STKE*, **2006**, re2.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Yang, F., Moss, L. G. & Phillips, G. N. Jr (1996). *Nature Biotechnol.* **14**, 1246–1251.