Supporting Information for

"Structural Evidence of Photoisomerization Pathways in Fluorescent Proteins"

Jeffrey Chang[†], Matthew G. Romei[‡], and Steven G. Boxer*[,‡]

[†]Department of Physics, Stanford University, Stanford, California 94305, United States

[‡]Department of Chemistry, Stanford University, Stanford, California 94305, United States

*Correspondence to: sboxer@stanford.edu

Table of Contents

Section S1 Methods

Plasmid Construction

The rsEGFP2 gene in the pQE-31 plasmid was generously provided by Stefan Jakobs at the Max Planck Institute for Biophysical Chemistry[1] and inserted into the pET-15b vector (Novagen) using NEBuilder HiFi DNA assembly (New England Biolabs), retaining the start codon and polyhistidine affinity tag of the pQE-31 plasmid. The amber stop codon (TAG) was inserted at Y67 using the QuikChange Site-Directed Mutagenesis Kit (Agilent) according to the manufacturer's instructions. The protein and amino acid sequences are shown below. The chromophore-forming residues are underlined; the amber suppression site is **bolded**.

Protein Sequence

MRGSHHHHHHTDPMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPW
PTLVTTLA**Y**GVLCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDF
KEDGNILGHKLEYNYNSHNVYIMADKQKNGIKSNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLS
TQSKLSKDPNEKRDHMVLLEFVTAAGITLGMDELYK

DNA Sequence

ATGAGAGGATCTCACCATCACCATCACCATACGGATCCGATGGTGAGCAAGGGCGAGGAGCTGTTCACC
GGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGA
GGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCG
TGCCCTGGCCCACCCTCGTGACCACCCTGGCC**TAC**GGCGTGCTGTGCTTCAGCCGCTACCCCGACCACAT
GAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAA
GGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCG
AGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC
AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGTCTAACTTCAAGATCCGCCAC
AACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCC
CGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCG
CGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAA
GTAA

Protein Expression

Chemically competent BL21 *Escherichia coli* cells (Invitrogen) were co-transformed with (1) a pET-15b expression vector containing the Y67TAG rsEGFP2 gene and (2) a pEVOL vector containing the 3-chlorotyrosine (3-ClY) amber suppression machinery.[2] Single colonies of these cells were inoculated into 10 mL of LB Miller (25 g/L) containing 100 mg/L ampicillin (Sigma-

Aldrich, CAS 69-52-3) and 35 mg/L chloramphenicol (Sigma-Aldrich, CAS 56-75-7) and grown overnight to saturation at 37°C. The starter cultures were used to inoculate 1 L of 46.7 g/L Terrific Broth Modified (Fisher), 8 mL glycerol (Fisher, CAS 56-81-5), 0.1 mg/L 3-chlorotyrosine (Sigma-Aldrich, CAS 7423-93-0), 100 mg/L ampicillin, and 35 mg/L chloramphenicol. Cells were grown in baffled 3-L flasks at 37°C with 200 rpm shaking until an O.D. of 0.6-0.7 at 600 nm was reached. At this point, the flasks were placed on ice, and 0.5 g/L of L-arabinose (Sigma-Aldrich, CAS 5328-37-0) was added to the cell culture to induce expression of the amber suppression machinery. Once the flasks reached 20 °C, 0.24 g/L of isopropyl β-D-1-thiogalactopyranoside (Fisher, CAS 367-93-1) was added to induce rsEGFP2 expression. The cells were then incubated overnight (~17 hours) at 20 °C while shaking at 200 rpm.

Protein Purification

E. coli containing the protein of interest were separated from the media via centrifugation at 6500 rcf for 30 min. The cell pellet was resuspended in lysis buffer, an aqueous buffer at pH 8.0 containing 50 mM Tris-HCl (Fisher, CAS 1185-53-1) and 300 mM NaCl (Fisher, CAS 7647-14-5). The cells were then lysed by a high-pressure homogenizer (Avestin EmulsiFlex-C3). The lysate was centrifuged at 25000 rcf for 90 minutes, and the supernatant was centrifuged again at 25000 rcf for another 90 minutes to pellet any remaining cell debris. The resulting supernatant was passed through a 0.45 µm filter (Millex-HV Syringe Filter Unit, PVDF; Millipore) and then added to a column of Ni-NTA Agarose resin (QIAGEN) that was pre-equilibrated with lysis buffer. The column was rinsed with 3 column volumes of lysis buffer containing 20 mM imidazole (Aldrich, CAS 288-32-4) and then eluted with 3 column volumes of lysis buffer containing 200 mM imidazole. All the eluting fractions were inspected by eye, and those that were judged to contain rsEGFP2 were pooled together and exchanged into Buffer A, an aqueous buffer at pH 8.0 containing 10 mM NaCl and 50 mM Tris-HCl. The protein was then purified by anion-exchange chromatography (HiTrap 5 mL Q HP; GE Healthcare) with a gradient of Buffer A and Buffer B (1 M NaCl and 50 mM Tris-HCl pH 8.0 aqueous buffer).

The identity of the protein was confirmed with electrospray ionization mass spectrometry (ESIMS) measured with LC-MS (Waters 2795 HPLC with ZQ single quadrupole MS in Stanford University Mass Spectrometry (SUMS) facility). As seen in Table S1, we observed a systematic discrepancy of about 145 Da between the observed and expected masses of the protein. This discrepancy was puzzling at first because the residues resolved in the crystal structure showed no deviation in the amino acid identities, and an MS/MS analysis also showed no mass discrepancies. However, the N-terminal tail is not completely resolved in the crystal structures, and after digestion with trypsin and LysC, the remaining N-terminal peptide was too short to resolve by MS/MS. Hence any possible modifications must be localized to the N-terminal residue. In agreement with an N-terminal modification, the N-terminal peptide formed from formic acid cleavage[3] of the Asp-Pro peptide bond had a mass of 146 Da greater than the

expected peptide (RGSHHHHHHTD). This mass difference can be attributed to an N-terminal methylated methionine. N-terminal methionine cleavage often occurs during recombinant protein expression in *E. coli*, but the native enzyme responsible for this cleavage, methionyl-aminopeptidase, is hindered by a bulky neighboring residue.[4] The arginine in the second position of the rsEGFP2 sequence likely prevents N-terminal methionine cleavage. Methylation at the N-terminus of proteins in *E. coli* has been observed previously.[5–7]

Table S1. Expected and observed mass for the proteins in this study.

Constructs	Expected mass[a] (Da)	Observed mass (Da)	Expected mass of N-terminal peptide after formic acid digestion[b] (Da)	Observed mass of N-terminal peptide after formic acid digestion (Da)
rsEGFP2	28358	28503	--	--
Cl-rsEGFP2	28392	28536 (+33 from wild-type)	1357	1503 (+146 from expected)

[a] Predicted from the primary sequence of the protein with N-terminal methionine removed,[4] and subtracting 20 Da for the dehydration and reduction of chromophore formation.

[b] Predicted from the primary sequence of the N-terminal peptide after cleavage of the Asp-Pro bond, with the N-terminal methionine removed.

Protein Crystallization and Cryoprotection

Y67(3-ClY) rsEGFP2, denoted Cl-rsEGFP2, was crystallized by hanging-drop vapor diffusion in a VDX plate with sealant (Hampton Research). The protein was first exchanged into an aqueous buffer at pH 7.5 of 50 mM HEPES (Sigma-Aldrich, CAS 7365-45-9) at a protein concentration of ~12 mg/mL. This protein solution was mixed 1:1 with the mother liquor (0.1 M HEPES pH 8.1, 1.76 M ammonium sulfate (Sigma-Aldrich, CAS 7783-20-2)) to form the initial 2 µL drop on a plastic cover slip, which was inverted over a well with 0.5 mL of mother liquor. Orthorhombic crystals of ~50-300 µm in size appeared after 3-7 days of incubation at room temperature in the dark.

Once the Cl-rsEGFP2 crystals were grown, they were looped from the mother liquor drop, placed briefly (~3-5 s) in solution of cryoprotectant (see below), optionally irradiated (see below), and immediately plunged into liquid nitrogen to trap the species formed by irradiation. The expanded Cl-rsEGFP2 crystals were soaked in a cryoprotectant containing 0.1 M HEPES pH 8.1, 1.7 M ammonium sulfate, and 1 M sucrose (Sigma-Aldrich, CAS 57-50-1), whereas the contracted crystals were soaked in a more dehydrating cryoprotectant containing 0.16 M HEPES pH 8.1, 2.72 M ammonium sulfate, and 1.6 M sucrose. The non-irradiated Cl-rsEGFP2 crystals were cryocooled immediately after cryoprotection, whereas the irradiated Cl-rsEGFP2 crystals were obtained by placing the loop containing the crystal in the path of a 27 mW 488 nm diode laser (85-BCD-030-115, Melles Griot) after incubation in cryoprotectant and prior to cryocooling.

We observed that the exposure time to the more dehydrating cryoprotectant was an important variable – too short an exposure did not allow the unit cell to contract sufficiently to

distort the chromophore pocket, whereas too long an exposure caused the diffraction quality to degrade. We also observed that ~3 s of irradiation with our 27 mW laser was sufficient to minimize the residual *cis* population visible in the electron density from subsequent x-ray diffraction analysis. Since the irradiated, contracted Cl-rsEGFP2 crystals are subjected both to osmotic shock from the dehydrating cryoprotectant and to mechanical stresses from photoisomerization, their diffraction quality degrades especially rapidly, but by limiting the irradiation time to ~1 second and moving quickly during the experimental procedure, we were able to preserve the crystal's high-quality diffraction. Subsequent modeling of the electron density showed that this brief irradiation photoswitches about 50% of the monomers in the crystal lattice to the *trans* state. The contracted irradiated crystal was modeled by a mixture of *cis* and *trans* states; the *trans* state is the one discussed throughout this paper.

X-Ray Data Collection and Structure Refinement

X-ray diffraction data were collected at 100 K at the Stanford Synchrotron Radiation Lightsource on beamlines 12-2 for expanded *cis*, 14-1 for expanded *trans*, and 9-2 for contracted *cis* and *trans*. The data were processed (indexing, integration, scaling, and merging) with X-ray Detector Software (XDS)[8] using the autoxds script.[9] Molecular replacement was performed with phenix.phaser[10] using the wild-type rsEGFP2 structures (Protein Data Bank entries: 5dtx and 5dty[11]) as search models for the *cis* and *trans* states, respectively. Cycles of model building and refinement were done using Coot[12] and phenix.refine. During the initial rounds of refinement, the occupancy of the chromophore was set to 0 to avoid biasing its isomerization state or substituent orientation. Once the $2mF_{obs}$ - DF_{calc} electron density of the chromophore became unambiguous, the chromophore was added to the atomic model, and the structures were refined until the R-free score converged. Data collection and refinement statistics are listed in Table S2.

Table S2. Crystallographic data collection and refinement statistics.

	Expanded *cis*	Expanded *trans*	Contracted *cis*	Contracted *trans*
PDB ID	6PFR	6PFT	6PFS	6PFU
Wavelength	0.97946	1.305053	0.97946	0.97946
Resolution range	39.37 - 1.51 (1.564 - 1.51)	34.28 - 1.45 (1.502 - 1.45)	25.67 - 1.759 (1.822 - 1.759)	38.86 - 1.619 (1.677 - 1.619)
Space group	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21
Unit cell volume (Å^3)	214140	219540	198950	199660
a (Å)	50.699	51.014	51.422	51.320
b (Å)	62.489	62.716	59.246	59.479
c (Å)	67.591	68.619	65.302	65.408
α	90	90	90	90
β	90	90	90	90

γ	90	90	90	90
Total reflections	64376 (6726)	77790 (7299)	40521 (3886)	51967 (5060)
Unique reflections	32194 (3364)	38992 (3674)	20330 (1968)	25996 (2533)
Multiplicity	2.0 (2.0)	2.0 (2.0)	2.0 (2.0)	2.0 (2.0)
Completeness (%)	93.52 (99.61)	98.16 (93.61)	99.60 (98.35)	99.38 (98.37)
Mean I/sigma(I)	19.50 (1.86)	21.46 (5.52)	8.52 (1.68)	14.19 (1.67)
Wilson B-factor	20.60	17.03	21.74	20.13
R-merge	0.01152 (0.3537)	0.008496 (0.08949)	0.03846 (0.4464)	0.02333 (0.4409)
R-meas	0.0163 (0.5002)	0.01201 (0.1266)	0.05439 (0.6313)	0.03299 (0.6235)
R-pim	0.01152 (0.3537)	0.008496 (0.08949)	0.03846 (0.4464)	0.02333 (0.4409)
CC1/2	1 (0.854)	1 (0.982)	0.999 (0.703)	1 (0.703)
CC*	1 (0.96)	1 (0.995)	1 (0.909)	1 (0.909)
Reflections used in refinement	32167 (3358)	38991 (3674)	20325 (1968)	25986 (2533)
Reflections used for R-free	1610 (161)	1562 (147)	1998 (194)	1998 (195)
R-work	0.1876 (0.3330)	0.1503 (0.1999)	0.1734 (0.3029)	0.1789 (0.3012)
R-free	0.2116 (0.3565)	0.1757 (0.2449)	0.2105 (0.3464)	0.2076 (0.3597)
CC(work)	0.959 (0.863)	0.964 (0.931)	0.968 (0.855)	0.962 (0.817)
CC(free)	0.960 (0.861)	0.965 (0.905)	0.965 (0.830)	0.960 (0.776)
Number of non-hydrogen atoms	2153	2316	2095	2073
macromolecules	1915	1981	1887	1854
ligands	21	26	46	52
solvent	217	309	162	167
Protein residues	240	241	232	229
RMS(bonds)	0.008	0.021	0.007	0.008
RMS(angles)	1.18	1.90	1.21	1.25
Ramachandran favored (%)	97.87	98.31	96.92	98.21
Ramachandran allowed (%)	2.13	1.69	2.64	1.34
Ramachandran outliers (%)	0.00	0.00	0.44	0.45
Rotamer outliers (%)	1.42	0.92	0.00	0.00
Clashscore	3.92	3.80	1.84	3.22
Average B-factor	32.57	24.04	28.82	25.87
macromolecules	32.10	22.60	27.88	25.16
ligands	23.33	19.92	42.28	20.42
solvent	37.65	33.61	36.00	35.48
Number of TLS groups	9	12	10	2

Section S2 Minor populations in the irradiated crystal structures

In the main text we discussed the *trans* structural models in the context of the irradiated crystals. For the contracted irradiated crystal, the electron density is well explained by a 50-50 mixture of *cis* and *trans* chromophores. For the expanded irradiated crystal, when we model the electron density with just a *trans* structural model, the difference density map (mF_{obs} - DF_{calc}) displays small peaks suggestive of other minor populations of the chromophore. These peaks likely show the position of the chlorine atom in minor populations because Cl is electron-rich and scatters x-rays most strongly.

One peak of the difference map is found at the position of the chlorine atom in the *cis* chromophore. This peak, along with another peak on the *cis* state rotamer of Tyr146, suggest that a fraction of the monomers in the crystal lattice remain in the *cis* state and were not photoswitched. (The residual *cis* population was also detected in a previous serial femtosecond crystallography study,[13] where the irradiated crystal of wild-type rsEGFP2 was modeled as a mixture of 90% *trans* and 10% *cis* states.) Note that the rest of the *cis* chromophore (the non-chlorine atoms) is undetectable in the difference density map.

Another peak of the difference map is found near carbons 4 and 5 (Figure 2) of the *trans* structural model. This peak could be explained by the chlorine atom of a *trans anti* chromophore, suggesting that a small fraction of the Cl-rsEGFP2 molecules in the expanded irradiated crystal may undergo a hula-twist and end up in the *trans anti* state. Oddly, this peak is located over the C4-C5 bond, rather than a C-Cl bond length's distance away from C5, so the ring geometry of this putative *trans anti* minor conformer would have to be slightly distorted compared to the predominant *trans syn* conformer. Since the non-chlorine atoms of the chromophore are undetectable in the difference density map, it is not possible to build a reliable atomic model of this minor conformer.

Section S3 Discussion of the timescale of ring-flipping between *anti* and *syn* conformers

In reaching our conclusion on the photoisomerization pathway, we assumed that the *trans* state structure observed in the cryocooled crystal reflects the configuration of the chromophore immediately after photoisomerization. That is, we assumed that the chlorine substituent is locked in a particular orientation (either *syn* or *anti*). The *trans syn* and *trans anti* conformers could interconvert upon a 180-degree torsion of the φ-bond, which we will call a ring-flip. For the assumption to hold, the kinetic barrier to ring-flipping must be sufficiently high to prevent any interconversion during the experiment. In our procedure, once the crystal is photoswitched, it spends a few seconds at room temperature, and then it is stored for weeks to months at liquid nitrogen temperatures. We assume that the ring-flip occurs too slowly at cryogenic temperatures to be relevant. Thus, if ring-flipping occurs at all in our experiment, it must happen during the few seconds between photoisomerization and cryocooling. It is important to note that the observed fractional occupancies of *trans syn* and *trans anti* in the expanded and contracted lattices, respectively, are predominantly one conformer rather than a mixture. If the two conformers were freely exchanging, this observation could only be explained by an energetic bias of many kT for one conformer over the other between the two lattices, which seems unlikely.

In the following we consider whether the φ-bond rotation barrier is high enough to prevent thermal ring-flipping on the second timescale at room temperature. There are few experiments (short of full serial crystallography) which can probe this process directly, so we present a qualitative argument based on what is known about the rate of related processes. The relevant chemical process is a 180-degree ring-flip of the 3-chlorophenolate moiety about the φ-bond, at room temperature, in the confined environment of a protein cavity. A similar process is the ring-flipping of buried tyrosine residues in the interiors of proteins. NMR studies of chemical exchange have shown that this process occurs with typical rates of ~10-1000 s^{-1},[14–16] and compared to the Cl-rsEGFP2 chromophore ring-flip, the ring-flip of a buried tyrosine involves a similar steric size and occurs in a similar cavity environment. The key difference for the Cl-rsEGFP2 chromophore is that the rotating φ-bond lies in a conjugated system and has significant double bond character. This effect hinders rotation about the φ-bond and hence is expected to slow down the ring-flipping process compared to buried tyrosines.

The partial double bond character of the φ-bond can be estimated using a resonance picture. The Cl-rsEGFP2 chromophore has an additional resonance form where the locations of the single and double bonds are swapped; in this case the φ-bond is a double bond and the τ-bond is a single-bond. The amount of double bond character of the φ-bond is determined by the size of the contribution from the secondary resonance form. If there were no contribution, then the φ-bond and the τ-bond would be "pure" C-C single and double bonds, respectively, and their bond lengths would differ by roughly 0.2 Å;[17] if there were equal weightings of the two resonance forms (as in the allyl anion), then the φ-bond and the τ-bond would be roughly

equal in length. The measured difference in bond lengths between the φ-bond and the τ-bond is 0.10 Å in small-molecule crystallography of the model GFP chromophore[18] and 0.06 Å in protein crystallography of GFP at subatomic resolution.[19] (Both these bond lengths are of the protonated *cis* state which is more experimentally accessible; while high resolution structural data is not available for the protonated *trans* state because of its transient nature, it is reasonable to expect similar bond length differences in the *trans* state.) These crystallographic bond lengths indicate that the φ-bond has double bond character.

Since double bonds rotate less freely than single bonds, the partial double bond nature of the φ-bond is expected to slow down ring-flipping in the Cl-rsEGFP2 compared to the analogous process in typical buried tyrosine residues. One way to estimate the rate of φ-bond rotation is to start with the ~10-1000 s^{-1} of buried tyrosines and then account for how much the double bond character increases the rotation barrier. Each kcal/mol increase in barrier height causes the rate to decrease by a factor of exp(1/RT) =5.4 at room temperature. A typical C=C double bond has a rotation barrier on the order of 50-65 kcal/mol,[20] but since the φ-bond is only partially a double bond, its barrier is likely only a fraction of this. Even a modest increase of 5 kcal/mol in the barrier height would slow down the rate 5000-fold to a range of ~0.002-0.2 s^{-1}, which would lock the chlorine substituent in place for the duration of our experiment. While this analysis suggests that ring-flipping is likely not occurring faster than the rate of freezing following irradiation, a definitive result could be obtained using time-resolved crystallographic approaches.

Section S4 Supplementary Figures

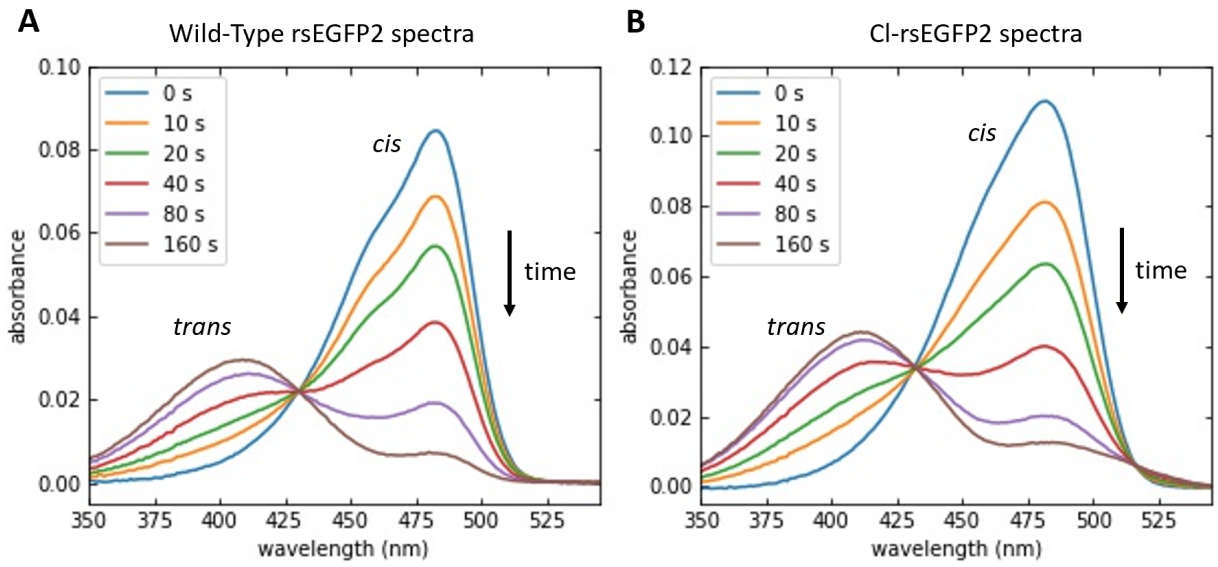

Figure S1. Cl-rsEGFP2 maintains the photoswitching property of wild-type rsEGFP2. Absorption spectra of (A) wild-type rsEGFP2 and (B) Cl-rsEGFP2 upon 488 nm laser irradiation. The emerging peaks around 410 nm correspond to the *trans* states.

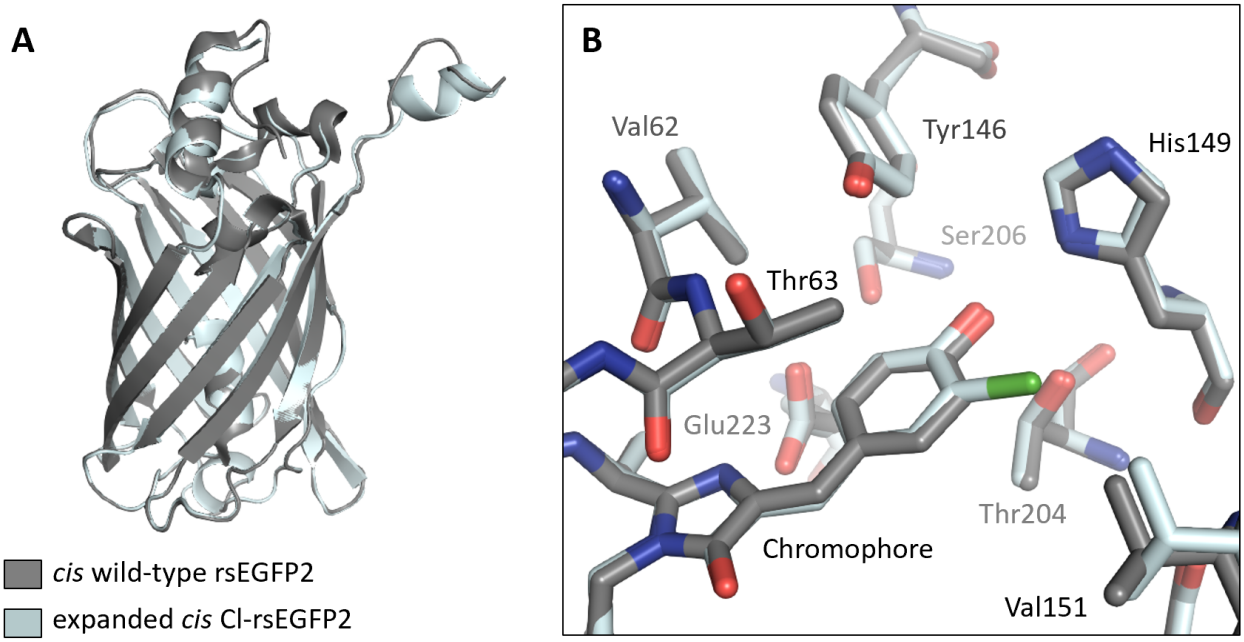

Figure S2. Comparison of Cl-rsEGFP2 to wild-type rsEGFP2 *cis* structure. (A) The structure of *cis* Cl-rsEGFP2 (light blue) in the expanded unit cell overlays onto the previously reported wild-type *cis* rsEGFP2 structure (gray; PDB ID: 5DTX[11]) with an α-carbon root-mean-square-deviation of 0.17Å. The wild-type *cis* rsEGFP2 has lattice constants $a = 51.0$ Å, $b = 62.9$ Å, $c = 70.8$ Å, and the expanded *cis* Cl-rsEGFP2 has lattice constants $a = 50.7$ Å, $b = 62.5$ Å, $c = 67.6$ Å. The proteins are shown in cartoon representation. (B) Inset of chromophore pocket with atoms shown in stick representation.

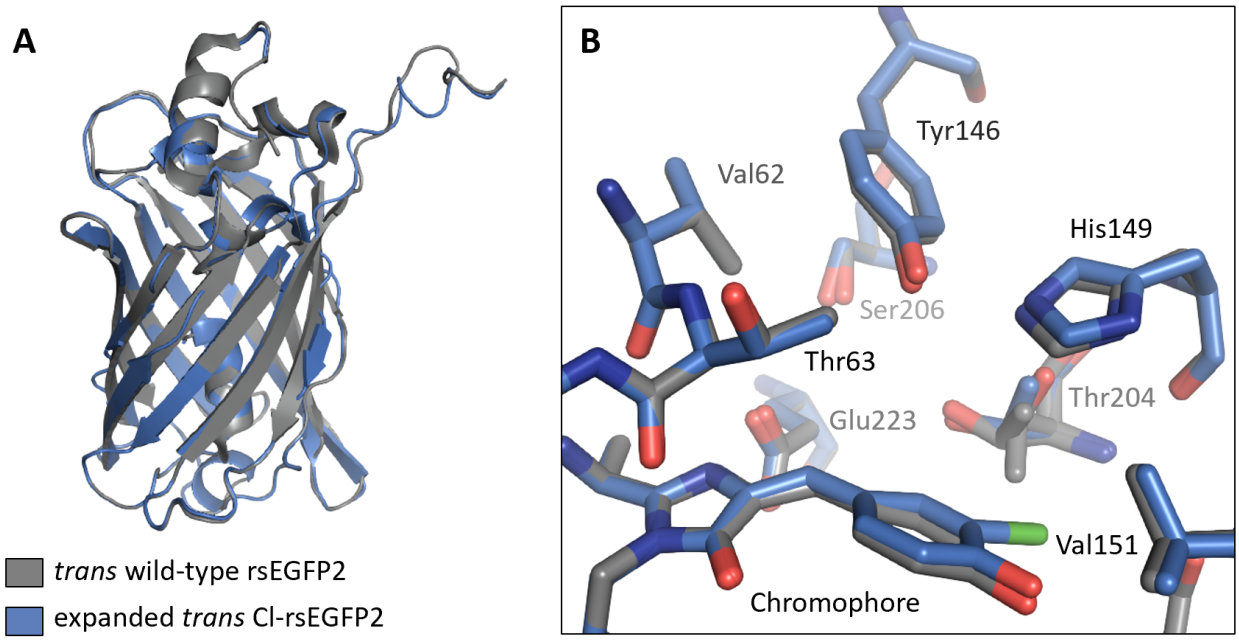

Figure S3. Comparison of Cl-rsEGFP2 to wild-type rsEGFP2 *trans* structure. (A) The structure of *trans* Cl-rsEGFP2 (blue) in the expanded unit cell overlays onto the previously reported wild-type *trans* rsEGFP2 structure (gray; PDB ID: 5DTY[11]) with an α-carbon root-mean-square-deviation of 0.11Å. The wild-type *trans* rsEGFP2 has lattice constants $a = 51.2$ Å, $b = 62.9$ Å, $c = 70.6$ Å, and the expanded *trans* Cl-rsEGFP2 has lattice constants $a = 51.0$ Å, $b = 62.7$ Å, $c = 68.6$ Å. The proteins are shown in cartoon representation. (B) Inset of the chromophore pocket with atoms shown in stick representation.

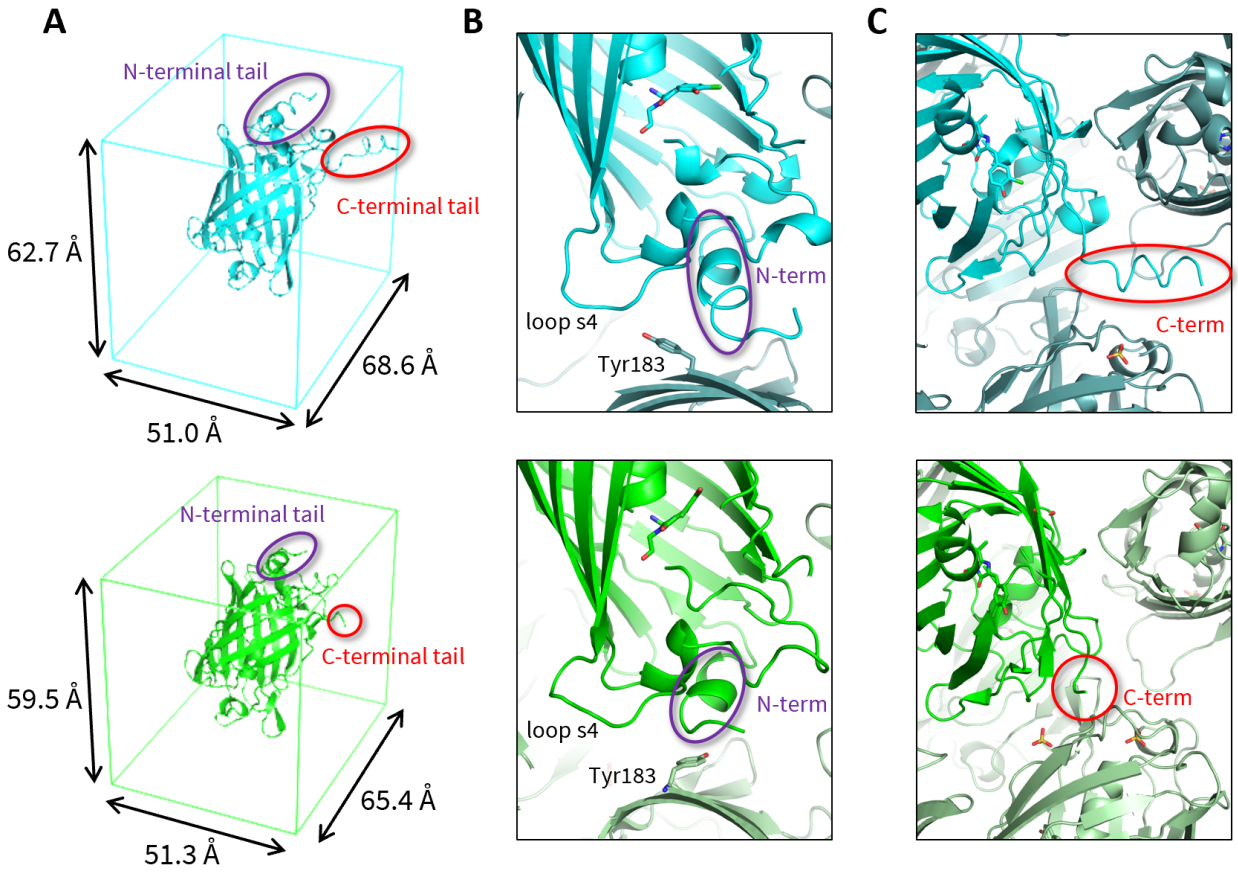

Figure S4. Comparison between expanded and contracted *trans* crystals. Cl-rsEGFP2 maintains the same overall β-barrel fold in the expanded (above) and contracted (below) crystals, but the packing interactions between neighboring units affect the structure of tails, loops, and certain outward-facing residues. (A) A Cl-rsEGFP2 monomer is shown in the orthorhombic unit cell with lattice constants labeled (Supp. Table 2). The N-terminal and C-terminal tails are circled in purple and red, respectively. (B) In the contracted unit cell (below), the N-terminal helix (circled in purple) is pressed closer to the β-barrel. The tighter packing also distorts the loop following strand 4 (loop s4) and forces Tyr183 to adopt a different rotamer. (C) The C-terminal loop (circled in red) is resolved in the expanded crystal (above) but unresolved in the contracted crystal (below). The protein and its neighboring crystallographic monomers are shown in cartoon representation, and the chromophore, sulfates, and Tyr183 are shown in stick representation.

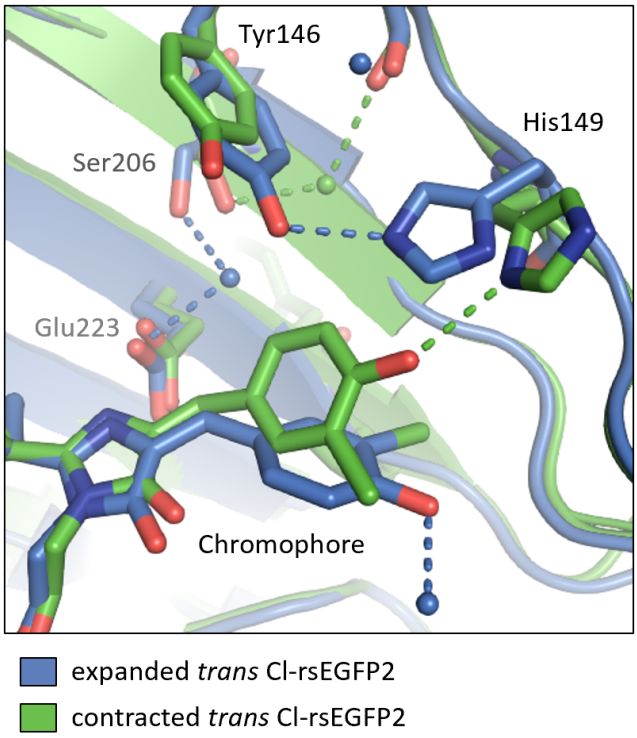

Figure S5. Effect of unit cell contraction on the *trans* Cl-rsEGFP2 chromophore pocket. Shown above are the chromophore pockets of the expanded (blue) and contracted (green) *trans* Cl-rsEGFP2 crystal structures, with water molecules shown as spheres and hydrogen bonds (≤3.1Å) shown as dotted lines. A 7% reduction in unit cell size (Supp. Table 2) results in the internal structural changes shown above.

Section S5 References

(1) Grotjohann, T.; Testa, I.; Reuss, M.; Brakemann, T.; Eggeling, C.; Hell, S. W.; Jakobs, S. rsEGFP2 enables fast RESOLFT nanoscopy of living cells. *eLife* **2012**, *1*, e00248.

(2) Liu, X.; Jiang, L.; Li, J.; Wang, L.; Yu, Y.; Zhou, Q.; Lv, X.; Gong, W.; Lu, Y.; Wang, J. Significant expansion of fluorescent protein sensing ability through the genetic incorporation of superior photo-induced electron-transfer quenchers. *J. Am. Chem. Soc.* **2014**, *136* (38), 13094–13097.

(3) Crimmins, D. L.; Mische, S. M.; Denslow, N. D. Chemical cleavage of proteins in solution. *Curr. Protoc. Protein Sci.* **2005**, *41* (1), 11.4.1-11.4.11.

(4) Hirel, P. H.; Schmitter, M. J.; Dessen, P.; Fayat, G.; Blanquet, S. Extent of N-terminal methionine excision from Escherichia coli proteins is governed by the side-chain length of the penultimate amino acid. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86* (21), 8247–8251.

(5) Apostol, I.; Aitken, J.; Levine, J.; Lippincott, J.; Davidson, J. S.; Abbott-Brown, D. Recombinant protein sequences can trigger methylation of N-terminal amino acids in Escherichia coli. *Protein Sci.* **1995**, *4* (12), 2616–2618.

(6) Varland, S.; Osberg, C.; Arnesen, T. N-terminal modifications of cellular proteins: The enzymes involved, their substrate specificities and biological effects. *Proteomics* **2015**, *15* (14), 2385–2401.

(7) Stock, A.; Clarke, S.; Clarke, C.; Stock, J. N-terminal methylation of proteins: structure, function and specificity. *FEBS Lett.* **1987**, *220* (1), 8–14.

(8) Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66* (Pt 2), 133–144.

(9) A Quick XDS Tutorial for SSRL http://smb.slac.stanford.edu/facilities/software/xds/#autoxds_script (accessed Jun 19, 2019).

(10) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66* (2), 213–221.

(11) El Khatib, M.; Martins, A.; Bourgeois, D.; Colletier, J.-P.; Adam, V. Rational design of ultrastable and reversibly photoswitchable fluorescent proteins for super-resolution imaging of the bacterial periplasm. *Sci. Rep.* **2016**, *6* (1), 18459.

(12) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60* (12), 2126–2132.

(13) Coquelle, N.; Sliwa, M.; Woodhouse, J.; Schirò, G.; Adam, V.; Aquila, A.; Barends, T. R. M.; Boutet, S.; Byrdin, M.; Carbajo, S.; De la Mora, E.; Doak, R. B.; Feliks, M.; Fieschi, F.; Foucar, L.; Guillon, V.; Hilpert, M.; Hunter, M. S.; Jakobs, S.; Koglin, J. E.; Kovacsova, G.; Lane, T. J.; Lévy, B.; Liang, M.; Nass, K.; Ridard, J.; Robinson, J. S.; Roome, C. M.; Ruckebusch, C.; Seaberg, M.; Thepaut, M.; Cammarata, M.; Demachy, I.; Field, M.; Shoeman, R. L.; Bourgeois, D.; Colletier, J.-P.; Schlichting, I.; Weik, M. Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography. *Nat. Chem.* **2018**, *10* (1), 31–37.

(14) Wagner, G. Characterization of the distribution of internal motions in the basic pancreatic trypsin inhibitor using a large number of internal NMR probes. *Q. Rev. Biophys.* **1983**, *16* (1), 1–57.
(15) Li, H.; Yamada, H.; Akasaka, K. Effect of pressure on the tertiary structure and dynamics of folded basic pancreatic trypsin inhibitor. *Biophys. J.* **1999**, *77* (5), 2801–2812.
(16) Weininger, U.; Modig, K.; Akke, M. Ring flips revisited: 13C relaxation dispersion measurements of aromatic side chain dynamics and activation barriers in basic bancreatic trypsin inhibitor. *Biochemistry* **2014**, *53* (28), 4519–4525.
(17) Berg, J. M.; Stryer, L. *Biochemistry*; W.H. Freeman, 2002.
(18) Dong, J.; Solntsev, K. M.; Tolbert, L. M. Activation and tuning of green fluorescent protein chromophore emission by alkyl substituent-mediated crystal packing. *J. Am. Chem. Soc.* **2009**, *131* (2), 662–670.
(19) Takaba, K.; Tai, Y.; Eki, H.; Dao, H.-A.; Hanazono, Y.; Hasegawa, K.; Miki, K.; Takeda, K. Subatomic resolution X-ray structures of green fluorescent protein. *IUCrJ* **2019**, *6* (3), 387–400.
(20) Barrows, S. E.; Eberlein, T. H. Understanding rotation about a C=C double bond. *J. Chem. Educ.* **2005**, *82* (9), 1329–1333.