DOI 10.1002/pro.2498

Title:

The structure of a far-red fluorescent protein, AQ143, shows evidence in support of reported red-shifting chromophore interactions.

Authors:

Wannier, Timothy M.[a] and Mayo, Stephen L.[a,1]

[a]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125

[1]To whom correspondence should be addressed:

Stephen L. Mayo

Division of Biology and Biological Engineering, MC 114-96

California Institute of Technology

1200 East California Blvd.

Pasadena, CA 91125

phone: 626-395-6408

e-mail: steve@mayo.caltech.edu

Running Title:

doi: 10.1002/pro.2498

Received: Apr 04, 2014; Revised: May 28, 2014; Accepted: May 29, 2014

Structure of AQ143, a far-red fluorescent protein

Total Number of Manuscript Pages:

Manuscript: 17 pages

Supplementary material: 2 pages + 8 figures + 1 table

Tables: 1

Figures: 2

Description of Supplementary Material:

We are providing eight supplementary figures with this article. They are listed below.

Figure S1 – Size-exclusion chromatography of AQ143 compared with mCherry

Figure S2 – Analytical ultracentrifugation of AQ143 and mCherry

Figure S3 – Excitation and emission spectra for the green chromophore of AQ143

Figure S4 – Excitation and emission spectra for the red chromophore of AQ143

Figure S5 – Absorbance spectrum of AQ143

Figure S6 – pH profile of AQ143

Figure S7 – Difference spectrum time course of AQ143 at alkaline pH

Figure S8 – Change in base-denatured and red chromophore absorbance over time

Table S1 – Analysis of the AB and AC interfaces of AQ143 as compared with other native tetrameric RFPs

We have also provided supplementary methods describing the dynamic difference method for determining the extinction coefficients for the red and green chromophore.

Abstract:

Engineering fluorescent proteins (FPs) to emit light at longer wavelengths is a significant focus in the development of the next generation of fluorescent biomarkers, as far-red light penetrates tissue with minimal absorption, allowing better imaging inside of biological hosts. Structure-guided design and directed evolution have led to the discovery of red FPs with significant bathochromic shifts to their emission. Here, we present the crystal structure of one of the most bathochromically shifted FPs reported to date, AQ143, a nine-point mutant of aeCP597, a chromoprotein from *Actinia equina*. The 2.19 Å resolution structure reveals several important chromophore interactions that contribute to the protein's far-red emission and shows dual occupancy of the green and red chromophores.

Keywords: Red Fluorescent Protein (RFP), Near-Infrared, Bathochromic Shift, Fluorescent Protein, Chromoprotein

Introduction

Fluorescent proteins (FPs) that emit light in the near-infrared (NIR) window (~650-900 nm) are in demand as biological imaging agents. The NIR window is a local minimum at which light penetrates tissue with minimal absorption from biological molecules such as melanin, hemoglobin, and water.[1] FPs natively do not emit light in the NIR; the longest maximum intensity emission wavelength (λ_{em}) reported to date for a native red FP (RFP) is 613 nm, found in NvFP-7R from *Nematostella vectensis.*[2] FPs with significant bathochromic shifts to λ_{em} have been produced with both rational design and directed evolution but these molecules tend to have low quantum yields, poor brightness and other characteristics that compromise their utility.[3-8] Many FP engineering strategies, including those that have induced bathochromic shifts in the λ_{em}, have relied on atomic-resolution structural data to guide intuition-based design, motivating continued efforts to obtain additional structural information for far-red FPs. AQ143, which was engineered from aeCP597, a chromoprotein from *Actinia equine,*[9] is one of only seven known FPs of the *Aequorea victoria* FP-like superfamily that exhibit a peak emission wavelength of at least 650 nm. The other five proteins (Neptune,[10] eqFP650,[7] TagRFP657,[11] mCardinal[12], eqFP670,[7] and TagRFP675[8]) are all variants of eqFP578, a native RFP from *Entacmaea quadricolor.*[13] There are known structures for five of these proteins (Neptune: 3IP2, eqFP650: 4EDO, mCardinal: 4OQW, eqFP670: 4EDS, TagRFP675: 4KGF), but as they are all derived from the same ancestral protein, there is limited sequence diversity among these structures. Here we report the 2.19 Å crystal structure of AQ143, which is derived from a more distantly related protein, aeCP597 (~60% sequence identity to eqFP578 and its variants). AQ143 has a novel chromophore

environment (defined as all internal-facing residues within 5 Å of the chromophore), which shares no more than 70% (16 of 23 positions) sequence identity with any other red fluorescent protein. Glu41 plays an important role in red-shifting AQ143's emission spectrum and is not seen in any other fluorescent protein. The reported structure also provides evidence in support of recently reported red-shifting chromophore interactions.[6,8,14]

Results and Discussion

The asymmetric unit contains eight protein molecules, which align with an all-atom *r.m.s.d*. of 0.27–0.74 Å with differences between molecules concentrated mostly in the loop regions and in the C- and N-terminal tails. The chromophore region also varies somewhat between molecules and shows weak electron density around the phenolate side chain, which could be attributed to mobility in the phenolate side chain and co-occupancy of two different chromophores -- green and red.

Oligomerization

AQ143 is a native tetramer, which is clear in the crystal packing. The asymmetric unit, however, contains eight monomers, or two such tetrameric assemblies with the C-terminal tail of each monomer involved in making inter-tetramer contacts. To verify the oligomerization state of AQ143, we ran both size exclusion chromatography (SEC) and analytical ultracentrifugation (AUC). SEC analysis indicates that AQ143 behaves as a tetramer, but that it has slight octomeric properties, while AUC confirms that the protein is predominantly tetrameric (Figs. S1 and S2). As oligomerization is an important

consideration in the engineering of RFPs, all of which are natively tetrameric,[15] we compared the AB and AC interfaces of AQ143 with those of four other native RFPs (Table S1) using the PISA server of the European Bioinformatics Institute[16] and report average buried surface area and average $\Delta^{i}G$ (the solvation free energy gain upon formation of the interface). The AC interface is known to be the tighter of the two interfaces,[17] which is consistent with the AC interface having more negative $\Delta^{i}Gs$. Interestingly, although AQ143 showed a similar amount of buried AB interface surface area, the $\Delta^{i}G$ for AQ143's AB interface is very high, indicating a large amount of hydrophobic residues at this interface. The AC interface is more difficult to compare as the amount of buried surface area varies widely, although this is in part due to the lack of crystallographic density at the C-terminal tails (which participate intimately in this interface) in many of these structures. DsRed, the first successfully monomerized RFP[17] shows the lowest $\Delta^{i}G$ for its AC interface, possibly indicating that future monomerization efforts of AQ143 may be more difficult.

Green and red chromophores

Many engineered far-red FPs exhibit slow or incomplete maturation to the red chromophore,[4,11,18] and it has recently been shown that maturation to the green and red chromophores in DsRed-type FPs occurs via a branched pathway (i.e., the two forms of the chromophore are separate endpoints in chromophore maturation; the green is not an intermediate in the maturation to the red chromophore as had been previously proposed).[19] AQ143 is a DsRed-type FP with a chromophore composed of a methionine/tyrosine/glycine triad (MYG) that matures to both a green and a red

chromophore (Fig. 1), as evidenced by its absorbance, excitation, and emission spectra (Figs. S3, S4, and S5). To calculate the percentage of chromophores that mature to the green and to the red, we determined the extinction coefficients of the two species by the dynamic difference method. In this procedure, AQ143 was pH-adjusted to alkaline conditions, in which the green and red chromophores denature at different rates and their respective contributions to the 450 nm alkali-denatured absorbance peak can be determined (Figs. S6, S7, S8 and Supplementary Methods).[5] We calculated the extinction coefficient to be 58,000 ± 11,000 $M^{-1}cm^{-1}$ for the red chromophore and 47,000 ± 5,000 $M^{-1}cm^{-1}$ for the green chromophore. From these data, we estimated the percentages in the fully mature protein to be 33 ± 6% for the red and 67 ± 6% for the green chromophore. Measurements of the protein in the crystal condition suggested that this fraction did not change upon crystallization. Corroborating the spectroscopic evidence, we observed that the refined electron density map of AQ143 shows a mixture of chromophores containing both the oxidized N-acylimine (red) and the unoxodized N-acylamine (green) at the N-terminal residue of the chromophoric triad. The estimated occupancy of the red and green chromophores averaged across all eight monomers in the asymmetric unit is 24 ± 9% and 76%, respectively. Thus the spectroscopic calculations of chromophore occupancy in the crystal condition are consistent with the crystallographic refinement.

Cis vs. trans phenolate

The phenolate side chain of the chromophore (the phenolate group) in DsRed-type FPs and related chromoproteins can occupy either a *cis* or a *trans* conformation, indicating

its proximity to the N1 nitrogen of the imidazolinone ring of the chromophore. For many RFPs, a *trans* to *cis* isomerization of this phenolate moiety, which is sometimes pH-inducible,[20,21] has been implicated in fluorescence. In non-fluorescent chromoproteins, for instance, the chromophore is found in the *trans* conformation, and mutations to these chromoproteins that stabilize the *cis* conformation have created FPs such as HcRed and AQ143. In engineering AQ143 from the chromoprotein aeCP597, Cys143Ser was reported to be responsible for inducing weak fluorescence,[9] as the mutation to serine stabilizes the *cis* chromophore by providing a hydrogen bond to the hydroxyl oxygen of the phenolate side chain. In the referenced work, fluorescence was improved by removing a serine hydrogen bond to the hydroxyl of the *trans* phenolate with a Ser158Ala mutation, further stabilizing the *cis* over the *trans* chromophore. By inducing fluorescence in an otherwise non-fluorescent chromoprotein, these mutations seem to imply that the *cis* chromophore represents the fluorescent moiety in AQ143.

Indeed, the refined structure shows good electron density for all parts of the chromophore with the exception of the phenolate side chain, which we modeled in the *trans* configuration. However, the difference map shows that the modeled phenolate is not a perfect fit, as the electron density is not sufficient to describe a chromophore that is solely found in the modeled *trans* configuration, while residual density appears in the position we expect that the *cis* phenolate would occupy. The refined electron density is such that we expect there is a co-occupancy in the crystal of two or more chromophore orientations and also possibly that the phenolate is mobile in one or both of these chromophore species. This would be consistent with a *cis-trans* isomerization of the chromophore upon fluorescence excitation, as has been seen in other FPs,[5,22,23]. The

lack of clear electron density for the phenolate moiety implies that the fluorescence-inducing mutations in AQ143 may have had their predicted effects, namely in destabilizing the native *trans* chromophore, and allowing for the phenolate to occupy the *cis* conformation. Given the ambiguity associated with the chromophore orientation and the lack of clear density for the *cis* conformation, we elected to model-build the *cis* phenolate post-refinement (Fig. 1). The modeled position of the *cis* phenolate accommodates a hydrogen bond between the hydroxyl of the fluorescence-inducing Cys143Ser mutation and the phenolate oxygen, supporting the hypothesis that this interaction is linked to the induction of fluorescence in AQ143 (Fig. 2). A second water-mediated hydrogen bond to the phenolate oxygen appears to further stabilize the *cis* conformation.

Interestingly, neither the *trans* nor the modeled *cis* conformations of the chromophore are coplanar with the imidazoline ring. This non-coplanarity is relatively uncommon in FPs and has been proposed to be responsible for low quantum yields.[10] AQ143 indeed has a very low quantum yield (0.04)[9], and improving the coplanarity of the two chromophore rings may represent an opportunity to further improve its fluorescence.

Mechanisms of bathochromic shift

AQ143 exhibits a number of red-shifting chromophore interactions that have been well documented in the literature.[8] A network of direct and water-mediated hydrogen bonds has been proposed to lower the energy of the photoexcited state of the chromophore's conjugated π-electron system, resulting in bathochromic shifts to $\lambda_{em.}$[10,24] In AQ143, three hydrogen bonds to the chromophore are good candidates to provide such

stabilization including two hydrogen bonds to the acylimine oxygen, as well as one to the phenolate oxygen (Fig. 2).

Glu41 and a water molecule coordinated by Gln106 and the chromophore's C-terminal acyl oxygen both form hydrogen bonds to the chromophore N-acylimine (Fig. 2). To our knowledge, the only other FPs known to have two hydrogen bonds to the acylimine oxygen are CjBlue, the furthest red-shifted chromoprotein, and TagRFP675, the furthest red-shifted FP, although in TagRFP675, the hydrogen bond donor at the position equivalent to Glu41 in AQ143 is a glutamine.[8,14] mPlum, the furthest red-shifted monomeric FP,[3] has a similar hydrogen bonding interaction between Glu16 and the chromophore N-acylimine, but is lacking a coordinated water molecule to provide the second hydrogen bond. The importance of hydrogen bonds to the N-acylimine was shown in mPlum variants, in which Glu16 is mutated to other residues including proline and glutamine, causing significant hypsochromic shifts to λ_{em}.[18,25,26]

Additionally, flexibility in the hydrogen bonding network to the phenolate oxygen of the chromophore, particularly via water-mediated hydrogen bonds, has been proposed to be responsible for extended stokes' shifts and significant bathochromic shifts to fluorescence emission.[8] The modeled *cis* chromophore, which we believe to be the fluorescent moiety, can accommodate two hydrogen bonds to the phenolate oxygen from the hydroxyl of Ser143 and a structural water molecule (Fig. 2). The *trans* chromophore, despite the mutation away from Ser158, makes a hydrogen bond contact with a structural water molecule stabilized by Glu145 and Thr176, although the effect of this interaction is less clear as the *trans* chromophore is not thought to be fluorescent.

Finally, many red-shifted fluorescent proteins have been described that exhibit π-stacking interactions with the phenolate group of the chromophore.[6,27,28] Histidine and tyrosine have both been reported at positions analogous to His197 in AQ143 with histidine present in eqFP578, RFP639, and mRuby[13,29,30], and tyrosine present in mRojoA, TagRFP657, and mGrape3[6,10,11]. In engineering mRojoA, a tyrosine π-stacking interaction with the *cis* phenolate was explicitly designed into the protein which resulted in a 7nm red-shift[6]. In AQ143, His197 appears to form a π-stacking interaction with the *trans* phenolate (Fig. 2), which we presume to be the non-fluorescent entity. Interestingly, in mRuby and eqFP578, the histidine also π-stacks with the *trans* phenolate, whereas in the further red-shifted RFP639, the π-stack occurs with the *cis* phenolate. This implies that there may be room to further stabilize the photo-excited state of the *cis* phenolate of AQ143 and red-shift its emission by optimizing the π-stacking interaction with the *cis* chromophore.

Conclusion

AQ143 is one of the furthest red-emitting FPs of the GFP family, and the structure reported in this study helps elucidate some of the features underlying its far-red emission. A recently reported FP, TagRFP675, shares many of the same chromophore interactions responsible for AQ143's bathochromic shift.[8]

Materials and Methods

Protein expression and purification

A synthetic gene construct encoding an N-terminal poly-histidine tagged AQ143

(GenBank KF479351) was assembled *in vitro,* expressed in *Escherichia coli* BL21(DE3) cells, purified, and crystallized. Cultures were grown at 37°C to an optical density of □0.6 in LB, induced, then allowed to express protein at 20°C for 24 hours. Protein was purified via His-tag affinity chromatography, run over a size exclusion column to remove trace contaminants and move the protein into storage buffer (1 × PBS pH 7.4), and finally concentrated to 18 mg/ml.

Crystallization, data collection, and structure determination

Rectangular plate crystals grew in 7 days by the sitting-drop vapor diffusion method in 100 mM Tris pH 7.0 with 50 mM lithium sulfate and 20% w/v PEG 3350. Crystals were flash frozen in 2-Methyl-2,4-pentanediol (MPD) and shipped to beamline 12-2 at the Stanford Synchrotron Radiation Lightsource, where a 2.19 Å data set was collected. Phases were obtained through molecular replacement using the crystal structure of the FP asFP595 (PDB ID 1A50).

Following molecular replacement, model building and refinement were run with COOT and PHENIX.[31,32] NCS restraints were applied to early refinement steps and removed at the final stages of refinement. TLS parameters were used throughout. The chromophore was initially left out of the refinement and added at a later stage when clear density became evident for it. First the chromophore was added without the phenolate side chain, as little density appeared for this group. Subsequently, as density became clearer, a *trans* chromophore was added. The final modeled chromophore has a *trans* phenolate ring, an imidazoline heterocyclic ring, and dual occupancy of a green N-acylamine and a red N-acylimine. Coordinates were deposited in the Protein Data Bank

with the code 4OHS. Data collection and refinement statistics are listed in Table 1.

Modeling the cis chromophore post refinement

We modeled the *cis* chromophore after refining the structure because there was poor density for this conformation. There was, however, residual density in the region we expected the *cis* chromophore to be. We introduced the alternate conformation in COOT, fit it to the residual density, and ran the model through several rounds of PHENIX refinement, which resulted in the modeled positions shown in figures 1 and 2 in turquoise.

Supplementary Material

We are providing five supplementary figures with this article. They are listed below.

Figure S1 – Size-exclusion chromatography of AQ143 compared with mCherry

Figure S2 – Analytical ultracentrifugation of AQ143 and mCherry

Figure S3 – Excitation and emission spectra for the green chromophore of AQ143

Figure S4 – Excitation and emission spectra for the red chromophore of AQ143

Figure S5 – Absorbance spectrum of AQ143

Figure S6 – pH profile of AQ143

Figure S7 – Difference spectrum time course of AQ143 at alkaline pH

Figure S8 – Change in base-denatured and red chromophore absorbance over time

Table S1 – Analysis of the AB and AC interfaces of AQ143 as compared with other native tetrameric RFPs

We have also provided supplementary methods describing the dynamic difference method for determining the extinction coefficients for the red and green chromophore.

Acknowledgments

The authors are grateful for the use of beamline 12-2 at the Stanford Synchrotron Radiation Lightsource (SSRL) in Menlo Park, CA, operated by Stanford University and supported by the Department of Energy and the National Institutes of Health. They additionally are thankful to Jens Kaiser and Pavle Nikolovski at the California Institute of Technology for helpful discussions. Finally they thank the Gordon and Betty Moore Foundation, the Beckman Institute, and the Sanofi-Aventis Bioengineering Research Program for support of the Molecular Observatory at the California Institute of Technology.

References

1. Tromberg B, Shah N, Lanning R, Cerussi A, Espinoza J, Pham T, Svaasand L, Butler J (2000) Non-invasive in vivo characterization of breast tumors using photon migration spectroscopy. Neoplasia 2:26-40.
2. Ikmi A, Gibson M (2010) Identification and in vivo characterization of NvFP-7R, a developmentally regulated red fluorescent protein of Nematostella vectensis. PloS one 5:[PAGE].
3. Wang L, Jackson W, Steinbach P, Tsien R (2004) Evolution of new nonantibody proteins via iterative somatic hypermutation. Proc Natl Acad Sci USA 101:16745-16749.
4. Shcherbo D, Merzlyak E, Chepurnykh T, Fradkov A, Ermakova G, Solovieva E, Lukyanov K, Bogdanova E, Zaraisky A, Lukyanov S, Chudakov D (2007) Bright far-red fluorescent protein for whole-body imaging. Nature Methods 4:741-746.
5. Kredel S, Nienhaus K, Oswald F, Wolff M, Ivanchenko S, Cymer F, Jeromin A, Michel F, Spindler K-D, Heilker R, Nienhaus G, Wiedenmann J (2008) Optimized and far-red-emitting variants of fluorescent protein eqFP611. Chem Biol 15:224-233.
6. Chica R, Moore M, Allen B, Mayo S (2010) Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries. Proc Natl Acad Sci USA 107:20257-20262.
7. Shcherbo D, Shemiakina I, Ryabova A, Luker K, Schmidt B, Souslova E, Gorodnicheva T, Strukova L, Shidlovskiy K, Britanova O, Zaraisky A, Lukyanov K, Loschenov V, Luker G, Chudakov D (2010) Near-infrared fluorescent proteins. Nature Methods 7:827-829.
8. Piatkevich K, Malashkevich V, Morozova K, Nemkovich N, Almo S, Verkhusha V (2013) Extended Stokes shift in fluorescent proteins: chromophore-protein interactions in a near-infrared TagRFP675 variant. Scientific Rep 3:1847.
9. Shkrob M, Yanushevich Y, Chudakov D, Gurskaya N, Labas Y, Poponov S, Mudrik N, Lukyanov S, Lukyanov K (2005) Far-red fluorescent proteins evolved from a blue chromoprotein from Actinia equina. Biochem J 392:649-654.
10. Lin M, McKeown M, Ng H-L, Aguilera T, Shaner N, Campbell R, Adams S, Gross L, Ma W, Alber T, Tsien R (2009) Autofluorescent proteins with excitation in the optical window for intravital imaging in mammals. Chem Biol 16:1169-1179.
11. Morozova K, Piatkevich K, Gould T, Zhang J, Bewersdorf J, Verkhusha V (2010) Far-red fluorescent protein excitable with red lasers for flow cytometry and superresolution STED nanoscopy. Biophys J 99:5.
12. Jun C, Russell DH, Stéphane YC, Pengpeng L, Emilio G-G, John SB, Niloufar JA, Amy JL, Paula JC, Michelle AB, Michael WD, Ho-Leung N, Garcia KC, Christopher HC, Kang S, Helen MB, Michael ZL (2014) Non-invasive intravital imaging of cellular differentiation with a bright red-excitable fluorescent protein. Nature Methods [VOL:PAGE #S].
13. Merzlyak E, Goedhart J, Shcherbo D, Bulina M, Shcheglov A, Fradkov A, Gaintzeva A, Lukyanov K, Lukyanov S, Gadella T, Chudakov D (2007) Bright monomeric red fluorescent protein with an extended fluorescence lifetime. Nature Methods 4:555-557.

14. Chan MCY, Karasawa S, Mizuno H, Bosanac I, Ho D, Prive GG, Miyawaki A, Ikura M (2006) Structural characterization of a blue chromoprotein and its yellow mutant from the sea anemone Cnidopus japonicus. J Biol Chem 281:[PAGES].
15. Alieva NO, Konzen KA, Field SF, Meleshkevitch EA, Hunt ME, Beltran-Ramirez V, Miller DJ, Wiedenmann J, Salih A, Matz MV (2007) Diversity and evolution of coral fluorescent proteins. PloS one 3:[PAGES].
16. Krissinel, E. a. H., K. (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372:774-797.
17. Robert EC, Oded T, Amy EP, Paul AS, Geoffrey SB, David AZ, Roger YT (2002) A monomeric red fluorescent protein. Proc Natl Acad Sci USA [VOL:PAGES].
18. Moore M, Oteng-Pabi S, Pandelieva A, Mayo S, Chica R (2012) Recovery of red fluorescent protein chromophore maturation deficiency through rational design. PloS one 7:[PAGES].
19. Strack R, Strongin D, Mets L, Glick B, Keenan R (2010) Chromophore formation in DsRed occurs by a branched pathway. J Am Chem Soc 132:8496-8505.
20. Mudalige K, Habuchi S, Goodwin P, Pai R, De Schryver F, Cotlet M (2010) Photophysics of the red chromophore of HcRed: evidence for cis-trans isomerization and protonation-state changes. J Phys Chem B 114:4678-4685.
21. Pletnev S, Shcherbo D, Chudakov D, Pletneva N, Merzlyak E, Wlodawer A, Dauter Z, Pletnev V (2008) A crystallographic study of bright far-red fluorescent protein mKate reveals pH-induced cis-trans isomerization of the chromophore. J Biol Chem 283:28980-28987.
22. Adam V, Lelimousin M, Boehme S, Desfonds G, Nienhaus K, Field M, Wiedenmann J, McSweeney S, Nienhaus G, Bourgeois D (2008) Structural characterization of IrisFP, an optical highlighter undergoing multiple photo-induced transformations. Proc Natl Acad Sci USA 105:18343-18348.
23. Loos D, Habuchi S, Flors C, Hotta J-I, Wiedenmann J, Nienhaus G, Hofkens J (2006) Photoconversion in the red fluorescent protein from the sea anemone Entacmaea quadricolor: is cis-trans isomerization involved? J Am Chem Soc 128:6270-6271.
24. Wall M, Socolich M, Ranganathan R (2000) The structural basis for red fluorescence in the tetrameric GFP homolog DsRed. Nature Struct Biol 7:1133-1138.
25. Abbyad P, Childs W, Shi X, Boxer S (2007) Dynamic Stokes shift in green fluorescent protein variants. Proc Natl Acad Sci USA 104:20189-20194.
26. Shu X, Wang L, Colip L, Kallio K, Remington SJ (2009) Unique interactions between the chromophore and glutamate 16 lead to far-red emission in a red fluorescent protein. Protein Sci 18:460-466.
27. Strack RL, Hein B, Bhattacharyya D, Hell SW, Keenan RJ, Glick BS (2009) A rapidly maturing far-red derivative of DsRed-Express2 for whole-cell labeling. Biochemistry 48:8279-8281.
28. Suto K, Masuda H, Takenaka Y, Tsuji FI, Mizuno H (2009) Structural basis for red-shifted emission of a GFP-like protein from the marine copepod Chiridius poppei. Genes to Cells 14:727-737.

29. Kredel S, Oswald F, Nienhaus K, Deuschle K, Röcker C, Wolff M, Heilker R, Nienhaus GU, Wiedenmann J (2008) mRuby, a bright monomeric red fluorescent protein for labeling of subcellular structures. PloS one 4:[PAGES].
30. Nienhaus K, Nar H, Heilker R, Wiedenmann J, Nienhaus GU (2008) Trans-cis isomerization is responsible for the red-shifted fluorescence in variants of the red fluorescent protein eqFP611. J Am Chem Soc 130:12578-12579.
31. Adams P, Afonine P, Bunkóczi G, Chen V, Davis I, Echols N, Headd J, Hung L-W, Kapral G, Grosse-Kunstleve R, McCoy A, Moriarty N, Oeffner R, Read R, Richardson D, Richardson J, Terwilliger T, Zwart P (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Cryst D66:213-221.
32. Emsley P, Lohkamp B, Scott W, Cowtan K (2010) Features and development of Coot. Acta Cryst D66:486-501.

Tables:

Table I. X-ray data reduction and crystallographic refinement statistics

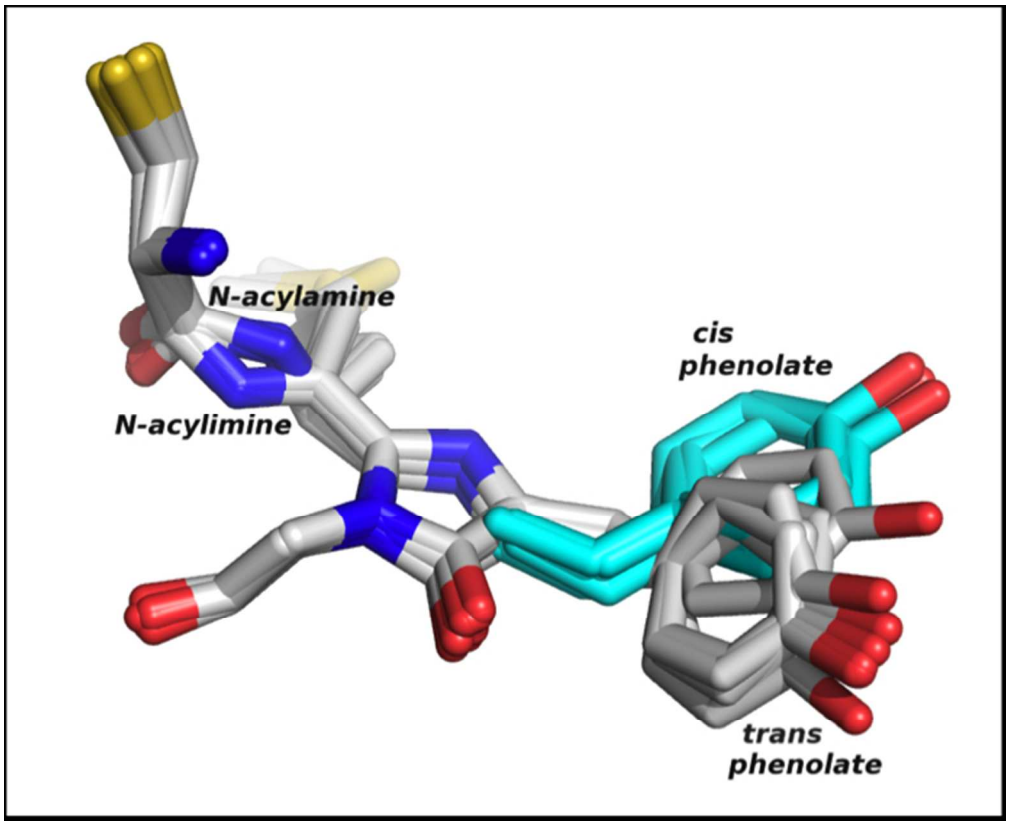
(A) X-ray data reduction statistics	
Space group	P1
Unit cell dimensions (*a, b, c*)	51.0 Å, 68.1 Å, 132.8Å
Resolution	39.1 Å – 2.19 Å
(last shell)	2.31 Å – 2.19 Å
Total measurements (last shell)	281,018 (30,290)
Number of unique reflections (last shell)	72,946 (8,028)
Wavelength	
R-merge (last shell)	0.072 (0.749)
I/σ(I) (last shell)	11.9 (1.7)
Completeness (last shell)	0.861 (0.648)
Multiplicity (last shell)	3.9 (3.8)
(B) Crystallographic refinement statistics	
Resolution	131.1 Å - 2.19 Å
(last shell)	2.22 Å – 2.19 Å
No. of reflections (working set)	69,234
No. of reflections (test set)	3,647
R-factor (last shell)	0.190 (0.315)
R-free (last shell)	0.221 (0.338)
No. of amino acid residues	1,770
No. of atoms	14,508

No. of solvent molecules	355
Average *B*-factor	
Protein	62.5 Å^2
Solvent	49.6 Å^2
R.m.s.d. from ideal geometry	
Bond lengths	0.006 Å
Bond angles	0.987°

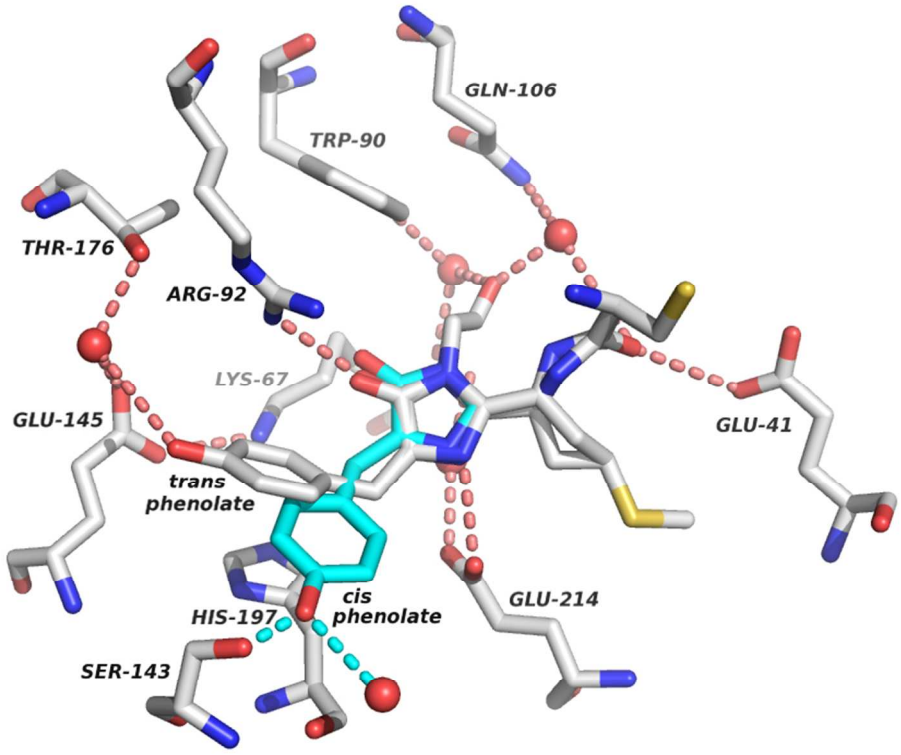
Figure Legends:

Figure 1. Alignment of the chromophores and C-terminal cysteine from each of the eight monomers in the asymmetric unit. The modeled *cis* phenolate is shown in turquoise. The N-acylamine and N-acylimine are present in the green and red chromophores respectively.

Figure 2. Chromophore contacts in AQ143. Residues that directly interact with the chromophore or help to coordinate structural waters (red spheres) are shown along with the immediate hydrogen-bonding network. A representative chromophore was chosen (chain E) to illustrate the contacts. Hydrogen bonds (dotted lines) are shown for interactions with the chromophore. The modeled *cis* conformation is shown in turquoise, along with two putative hydrogen bonds to its hydroxyl group. Two hydrogen bonds to the acylimine oxygen from Glu41 and a coordinated water can be seen in the right of the figure.



117x95mm (150 x 150 DPI)



323x255mm (72 x 72 DPI)

Supplementary Figure Legends

Figure S1. SEC traces of mCherry and AQ143. 100 µL of a 60 µ*M* aliquot of each protein was run through a Superdex 75 column. mCherry is a monomer, whereas AQ143 appears to be a tetramer by AUC (Fig. S2), but is an octomeric assembly in crystal packing and demonstrates a clear octomeric peak by SEC.

Figure S2. Sedimentation velocity analysis of mCherry and AQ143. mCherry, a monomer, sediments at its molecular weight, which is 27 kD. AQ143 sediments at 108 kD, very near its tetrameric weight.

Figure S3. Excitation and emission spectra of the green chromophore of AQ1143. Emission was measured with an excitation wavelength of 465 nm, while excitation was measured at an emission wavelength of 560 nm. These spectra were taken in a Photon Technology International fluorometer.

Figure S4. Excitation and emission spectra of the red chromophore of AQ1143. Emission was measured with an excitation wavelength of 550 nm, while excitation was measured at an emission wavelength of 660 nm. These spectra were taken in a Photon Technology International fluorometer.

Figure S5. Absorbance spectrum of AQ143 taken on a Tecan Safire2 platereader.

Figure S6. pH profile of the absorbance of 30 µ*M* AQ143 between pH 5.0 and 12.0. The protein undergoes two distinct transitions. The first transition is between an acid-denatured absorbance profile at which there is little absorbance by either chromophore to the growth of both the green and red absorption peaks (510 nm and 588 nm, respectively). The second transition is a rapid base denaturation of the protein in which both the red and green

chromophores are converted into a yellow, 450 nm-absorbing species. The extinction coefficient grows as the pH is increased, topping out near pH 10.0.

Figure S7. Time-dependent absorbance of AQ143 diluted 1:100 into alkaline buffer (100mM Na_2HPO_4; 150mM NaCl; pH 10.5). Time points are measured 5 minutes apart. Early time points are denoted with darker lines while late time points are denoted with lighter lines.

Figure S8. The loss in absorbance of the 588 nm (red chromophore) peak correlates linearly to an increase in the 470 nm (base-denatured chromophore) peak.

Table S1. AQ143 is compared to four other native RFPs (all tetramers). The surface area and solvation free energy of the AB and AC interfaces were evaluated with the PISA server from the European Bioinformatics Institute (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html).

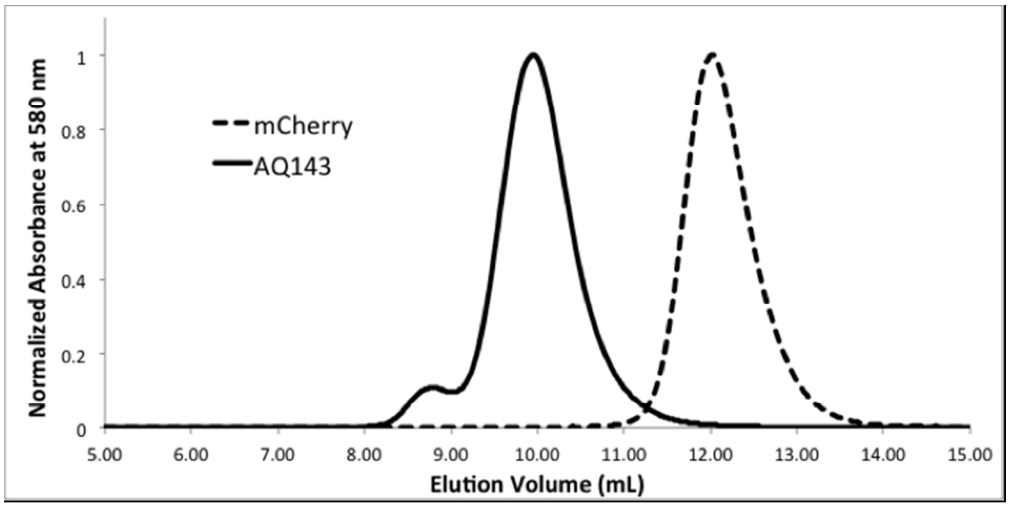
Table S1

Protein	PDB ID	AB Interface: Average buried surface area (Å^2)	AB Interface: Δ^iG (kcal/mol)*	AC Interface: Average buried surface area (Å^2)	AC Interface: Δ^iG (kcal/mol)*
AQ143	4OHS	979.6	-11.2	1203.9	-13.8
KFP†	2A50	977.7	-6.4	1691.4	-19.2
DsRed	1ZGO	993.9	-8.6	1326.8	-7.2
eqFP578	3PIB	1043.1	-2.2	1564.6	-20.8
eqFP611†	1UIS	988.1	-0.5	1185.6	-17.3

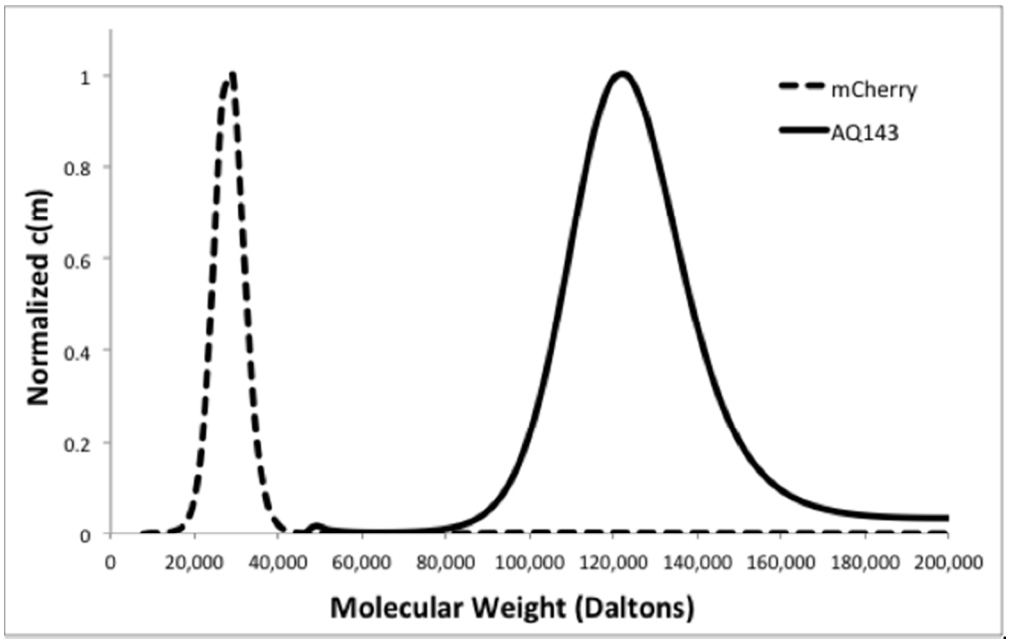
* Δ^iG indicates the solvation free energy gain upon formation of the interface.
† Only two monomers were present in the asymmetric unit, so symmetry mates were generated to visualize the tetramer.

Supplementary Methods

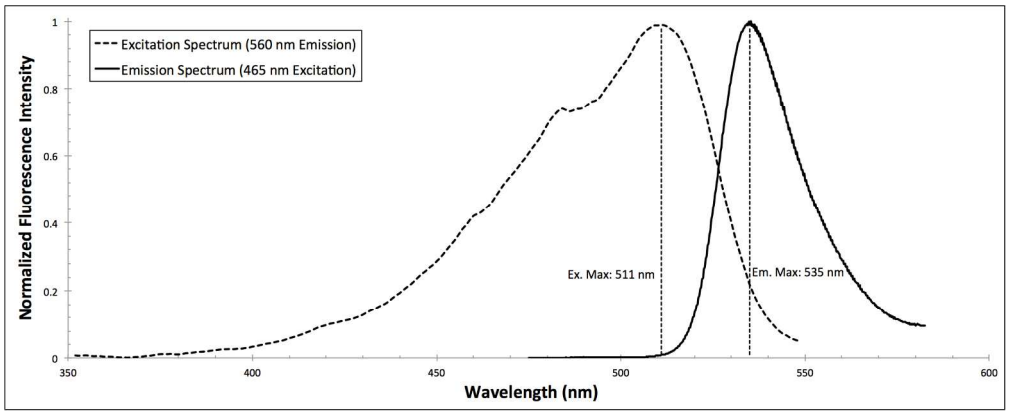
To determine the respective concentrations of green and red chromophore, the dynamic difference method was used as explained in Kredel et al. (2008).[5] AQ143 in 1x PBS was diluted 1:100 into alkaline phosphate buffer (100mM Na_2HPO_4; 150mM NaCl; pH 10.5). Absorption was then immediately measured over two hours as the protein denatured. The loss of red chromophore absorption at 590 nm occurs much faster than loss of green chromophore absorption at about 506 nm (Fig S7). Absorbtion at 588 nm (red chromophore) is replaced by absorption at 470 nm (base-denatured chromophore). The difference over time between these two peaks is linear (Fig S8). The slope of the best-fit line was used to estimate the ratio between the absorbance of the two chromophore species, and this in turn was used to calculate the extinction coefficient of the red chromophore. The denatured chromophore was assumed to have an extinction coefficient of 43,000 $M^{-1}cm^{-1}$. Then, after total denaturation of AQ143 at pH 12+, the total contribution of the red chromophore to the protein's absorbance is known, and the concentration of the green chromophore can therefore be calculated as the difference between the total chromophore concentration and the red chromophore concentration.



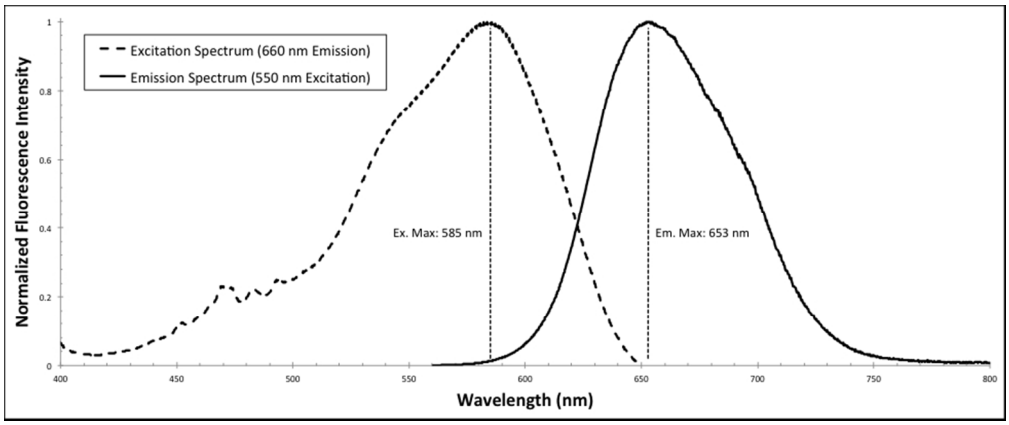
128x64mm (150 x 150 DPI)



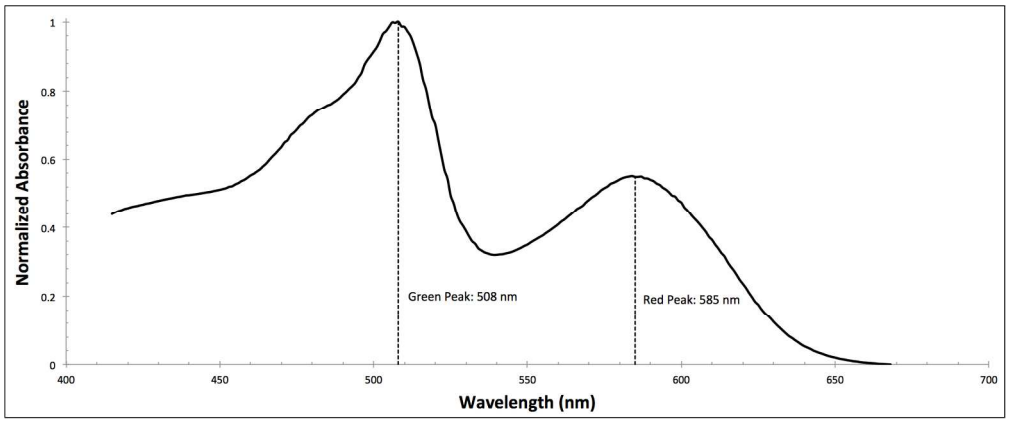
111x70mm (150 x 150 DPI)



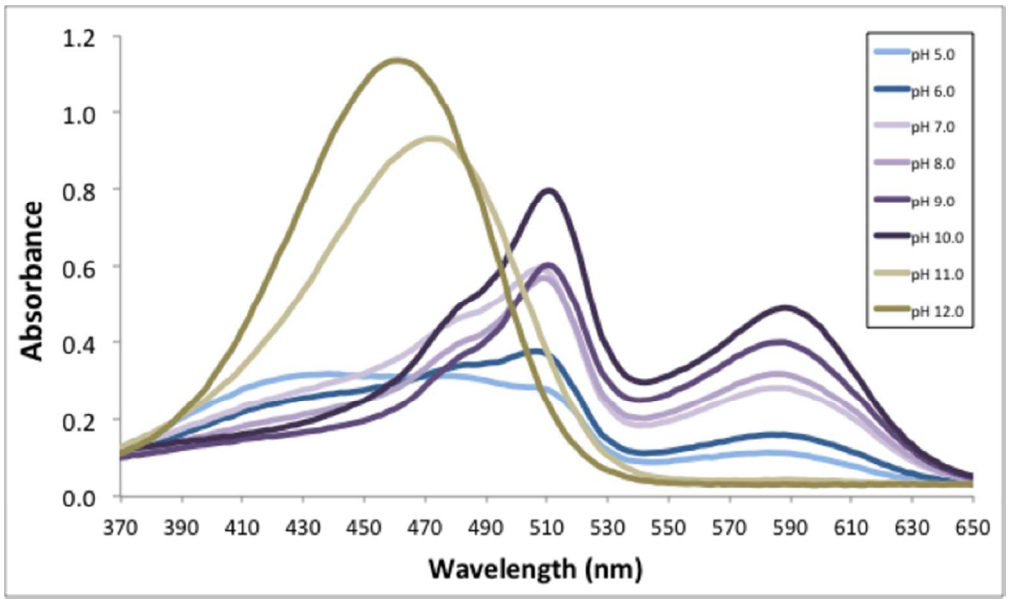
381x154mm (150 x 150 DPI)



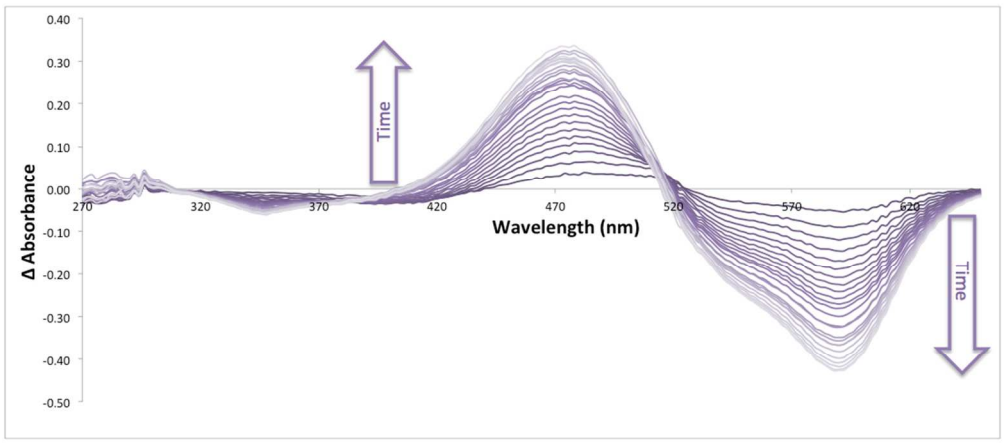
165x68mm (150 x 150 DPI)



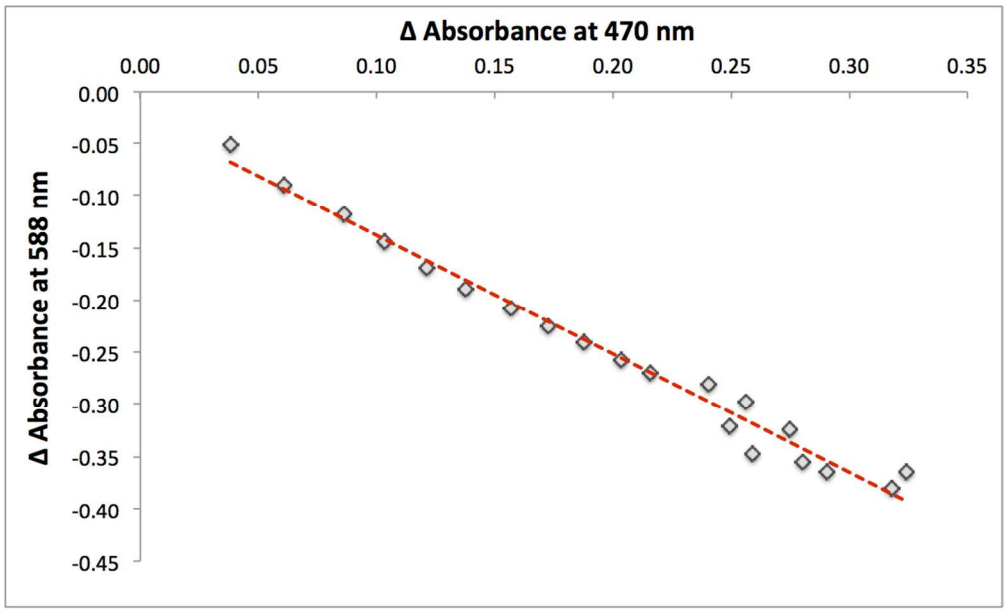
372x154mm (150 x 150 DPI)



130x77mm (150 x 150 DPI)



388x170mm (72 x 72 DPI)



217x131mm (150 x 150 DPI)