SUPPLEMENTAL TABLES

Table S1. Saturation mutagenesis of E222

Using the "trick" method (*1*) , all of the 19 possible variants of E222 were obtained in a single round of screening 90 colonies. The list of variants classified as either bright (RF > 0.50), dim (RF ≤ 0.50) or dark (RF = 0.0). This experiment corroborates previously reported data and is presented here as a control

RF Class	**Amino acid at 222**
Bright	CDEGHIKNSTV
Dim	FLMQRY
Dark	APW

Table S2. Spectroscopic properties of selected mutants

None of the seven mutants that were purified and characterized showed dramatic shifts in their fluorescence excitation and emission peak wavelengths. However, changes in the width of their absorption and excitation spectra were observed. A full-width-half-height analysis of the spectral width did not reveal any significant differences as these were equal to or less than the excitation slit width used.

Protein	**Spectral width (FWHH; nm)**
Wild-type	58
F145M	57
F145W	58

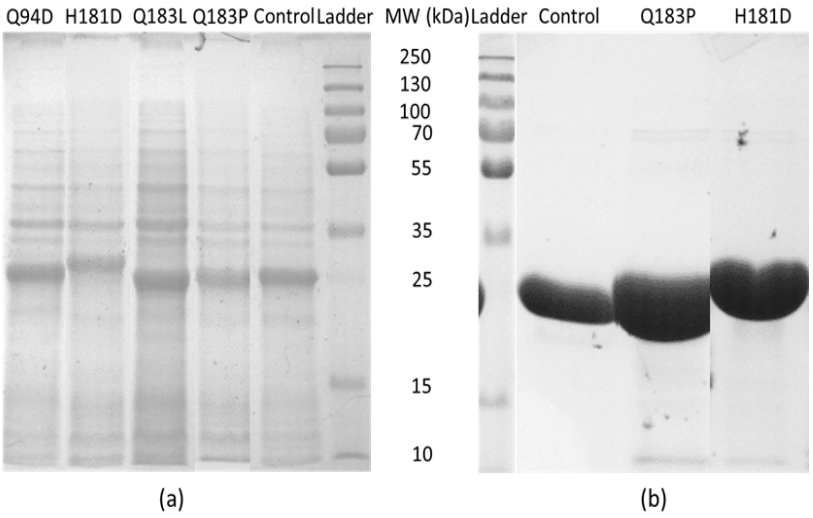
Table S3. Benchmark for expert sidechain modeling of point mutants. 12 out of 15 χ1 rotamer predictions were correct. (t=180°, m=-60°, p=+60°)

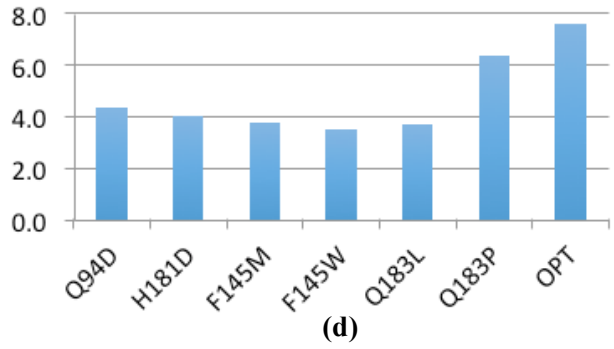
Target PDB code	**mutation**	***predicted χ1 rotamer***	***true χ1 rotamer***	***note***	***x=incorrect***
*Lysozyme point mutants, based on template **4w51**. All are minor sidechain differences, minimal backbone shifts.*					
118L	A130S	*t*	*t*		
119L	*A134S*	*m*	*m*		
120L	*A41S*	*m*	*m*		
122L	*A73S*	*t*	*t*		
123L	*A82S*	*t*	*m*		*x*
125L	*A98S*	*t*	*t*		
126L	*V149T*	*t*	*t*		
127L	*V75T*	*m*	*m*		
128L	*V87T*	*t*	*t*		
*DHFR point mutants, based on template **4dfr**. Major sidechain differences, some large backbone shifts.*					
1dhi	*D27S*	*t*	*t*		
4dra	*D27E*	*tmm*	*tmm*	*χ1,χ2,χ3 all correct*	
4qlq	*I14V*	*m*	*m*		
5cc9	*L28F*	*mm*	*m(0)*	*χ2 wrong, large backbone shifts*	
1dd2, 4qle, 4x5f, 4qlq, 5cc9	*N37D*	*m*	*(3m, 2t)*		*x*
all 9 dhfr mutants	*K154D*	*mt*	*(1 mt, 8 tt)*		*x*

SUPPLEMENTAL FIGURES

Figure S1: Protein expression

SDS-PAGE of (a) lysed cells that were subject to a 16 hr induction of protein expression and (b) proteins purified from only the supernatant. (c) Densitometry of cell lysate PAGE gels using ImageJ. PAGE gel lanes on top and densitometry trace (below each lane) are aligned to ladder (bottom frame). Numbers on the right are the integrated 27kD peak over the integrated 36kD peak. The baselines used for integration are drawn on the trace. (d) Relative expression levels of selected mutants compared to internal standard.





(d)

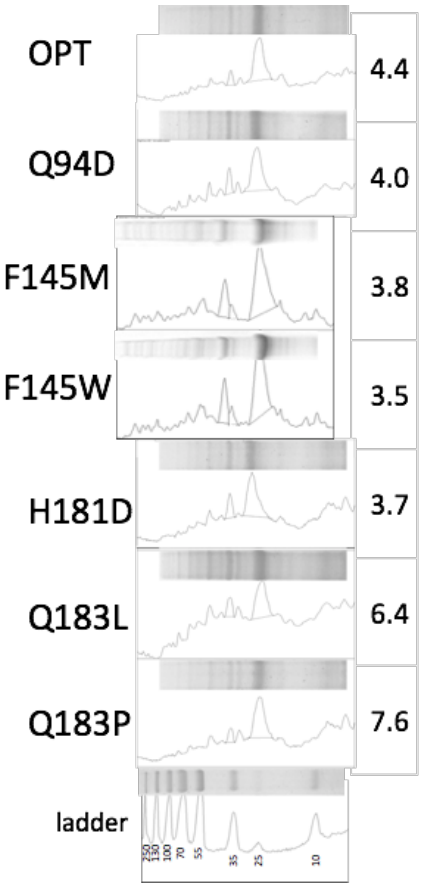


Figure S2: **Interpolating between the crystallized states**.
Interpolations (pink) between the precursor (orange), the cyclized intermediate (cyan), and the mature chromophore (green) were generated by averaging the chromophore coordinates between two adjacent states and minimizing the resulting structure within the context of the nearest full GFP state. Displayed in yellow are the positions considered for point mutation analysis. Image is cross-eyed stereo. (See also Video S1)

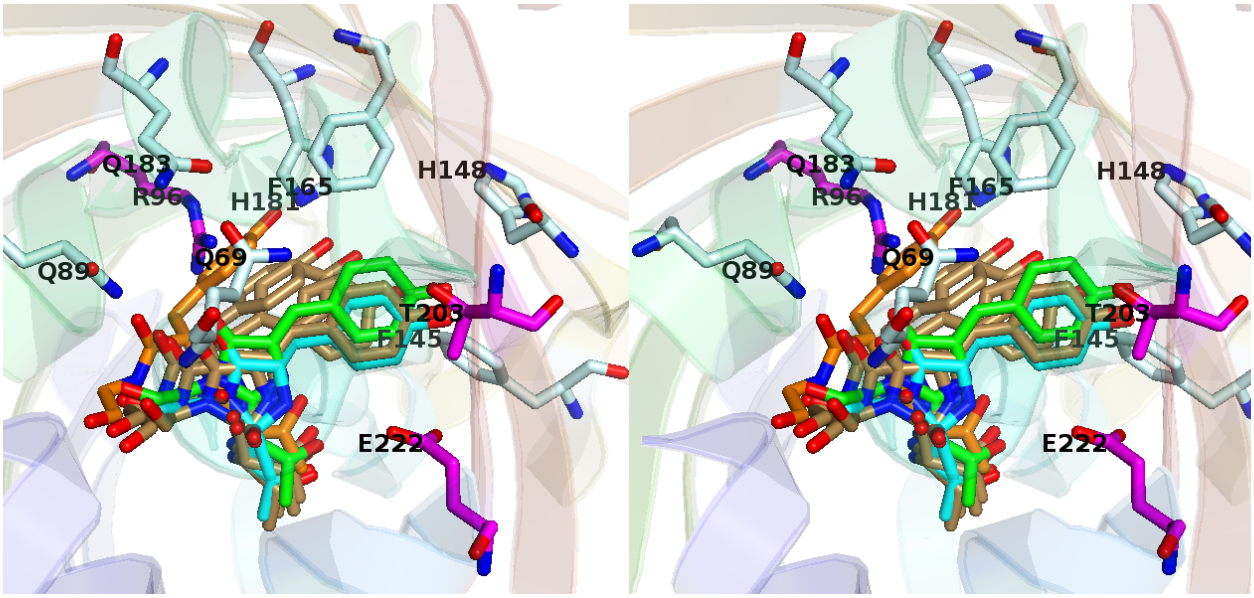


Figure S3: **Rosetta predictions of ΔΔG of mutation.**
One of the aims of this work was to test the ability of energy calculations to predict mutations that can impede chromophore maturation and function. This predictive ability could be used during protein design to avoid or minimize the loss of fluorescent function. If energy calculations correlated with GFP fluorescence, then fluorescence-preserving mutations could be predicted computationally. To test this hypothesis, rotamers were built using the Rosetta 3.5(*2*) PackRotamers mover function and the structure was energy minimized to convergence using the Minimization mover in for each of the 134 mutations. The *talaris2013* score function(*3*) was used for both protocols. ΔΔG is defined as the calculated free energy of the mutant relative to the WT, in Rosetta Energy units (REU), ΔΔG was plotted against the RF, but no correlation was found. Indeed, some very energetically agressive mutations were found to have high RF. Each point in the graph represents a variant, colored by position according to the legend.

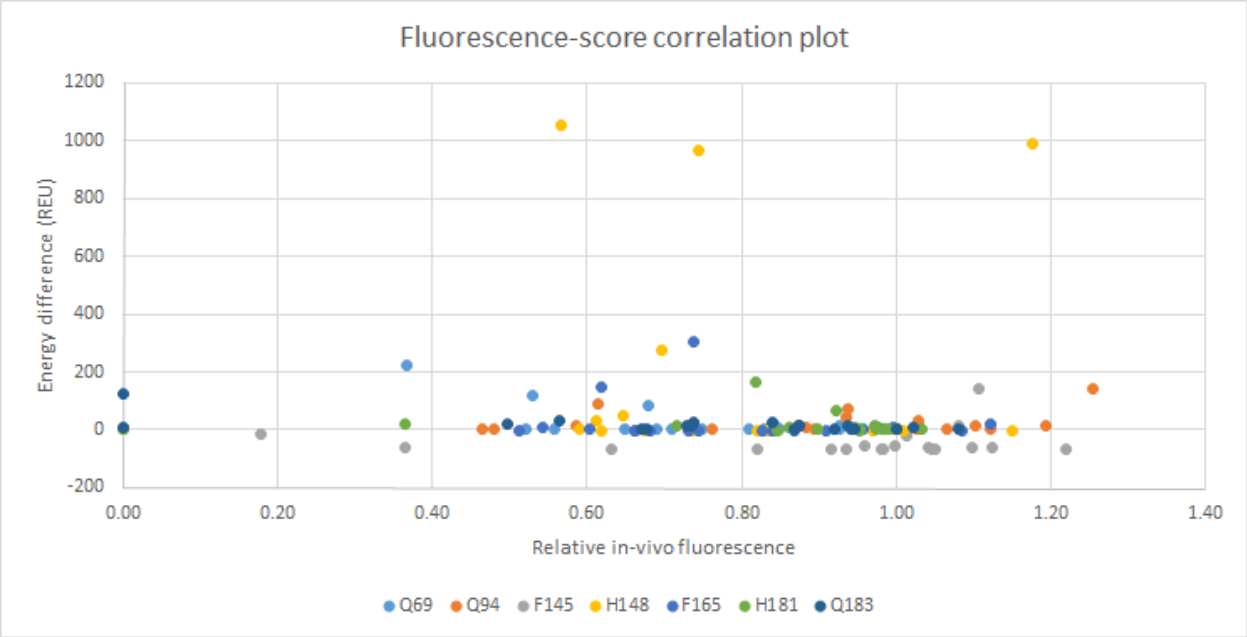


Figure S4: **Slow chromophore maturation in dark mutant Q94D matches slow phase of OPT-GFP.** Relative molar chromophore concentration (CM=A485/η for OPT-GFP and F145W (left axis), or A385/η for Q94D (right axis), where η = A280/ε280) was measured days, weeks and months past induction to characterize the rate of maturation. Standard error for triplicate measurements was 3%. The initial phase of maturation ($t_{1/2}$=34 minutes) (6), cannot explain the slow increase in chromophore concentration between 6 days and 120 days, with $t_{1/2}$ > 30 days. A single slow phase was used to fit the 6 to 120 day time range for Q94D simultaneously, giving $t_{1/2}$ > 100 days.

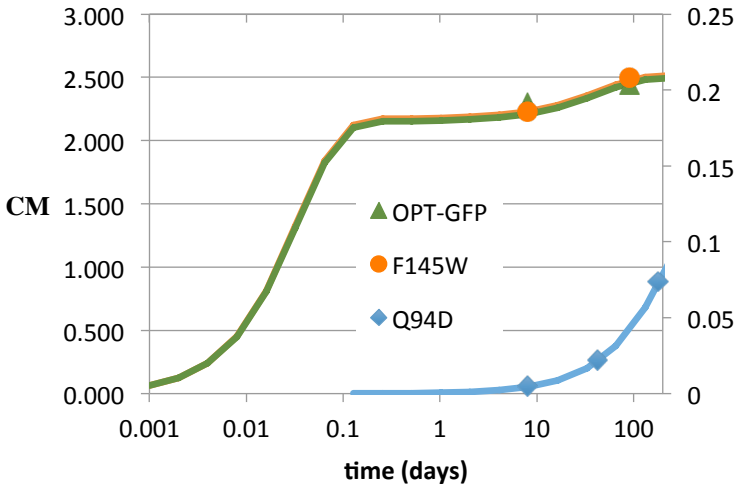
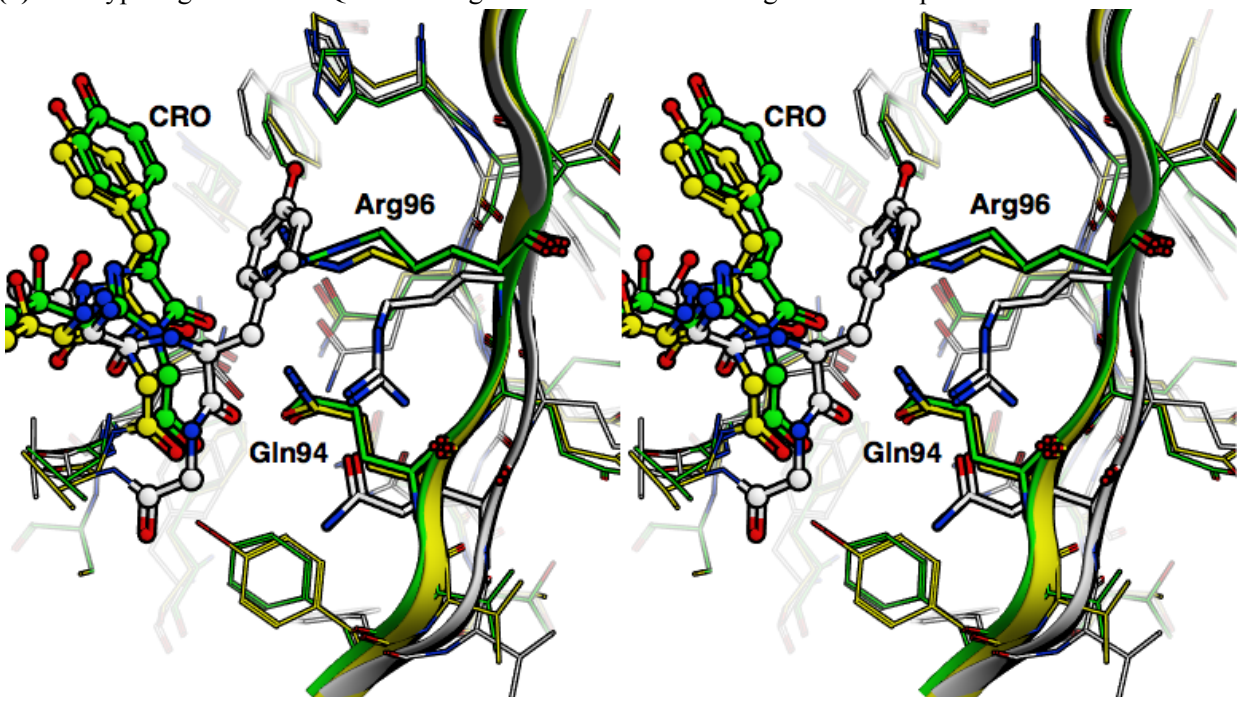
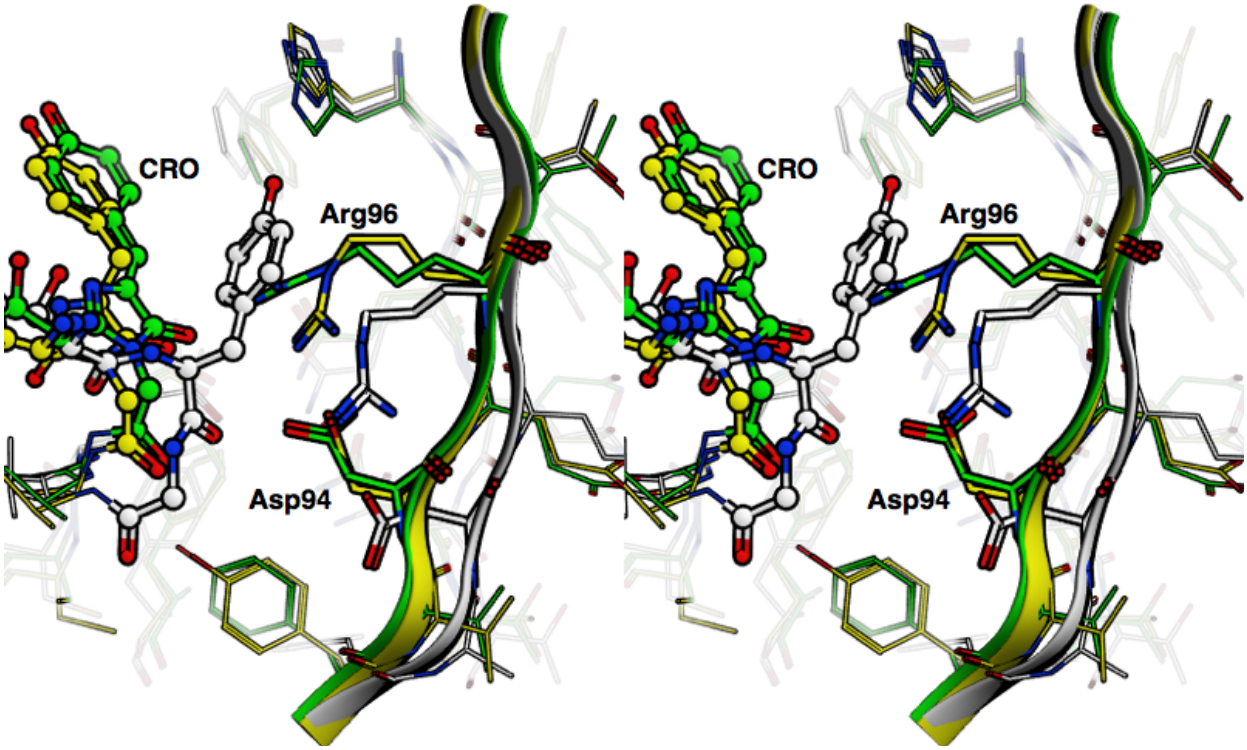


Figure S5: Models of mutations that significantly reduced fluorescent function, shown in walleye stereo, at three stages of GFP chromophore maturation. White bonds: precursor state (model based on PDB:1QXT), Yellow bonds: cyclized-only intermediate (model based on PDB: 2QYQ). Green bonds: fully mature state (PDB: 2B3P). (a) Wildtype region around Q94 showing rotameric motion of Arg96 between precursor and mature state. (b) Q94D mutant, region around Arg96 showing preferred non-native Arg rotamer in cyclized-only state. (c) Wildtype region around F145. (d) F145W region showing different preferred totamers in models of different stages of chromophore maturation. The model generation process was described in Materials and Methods in the main document.

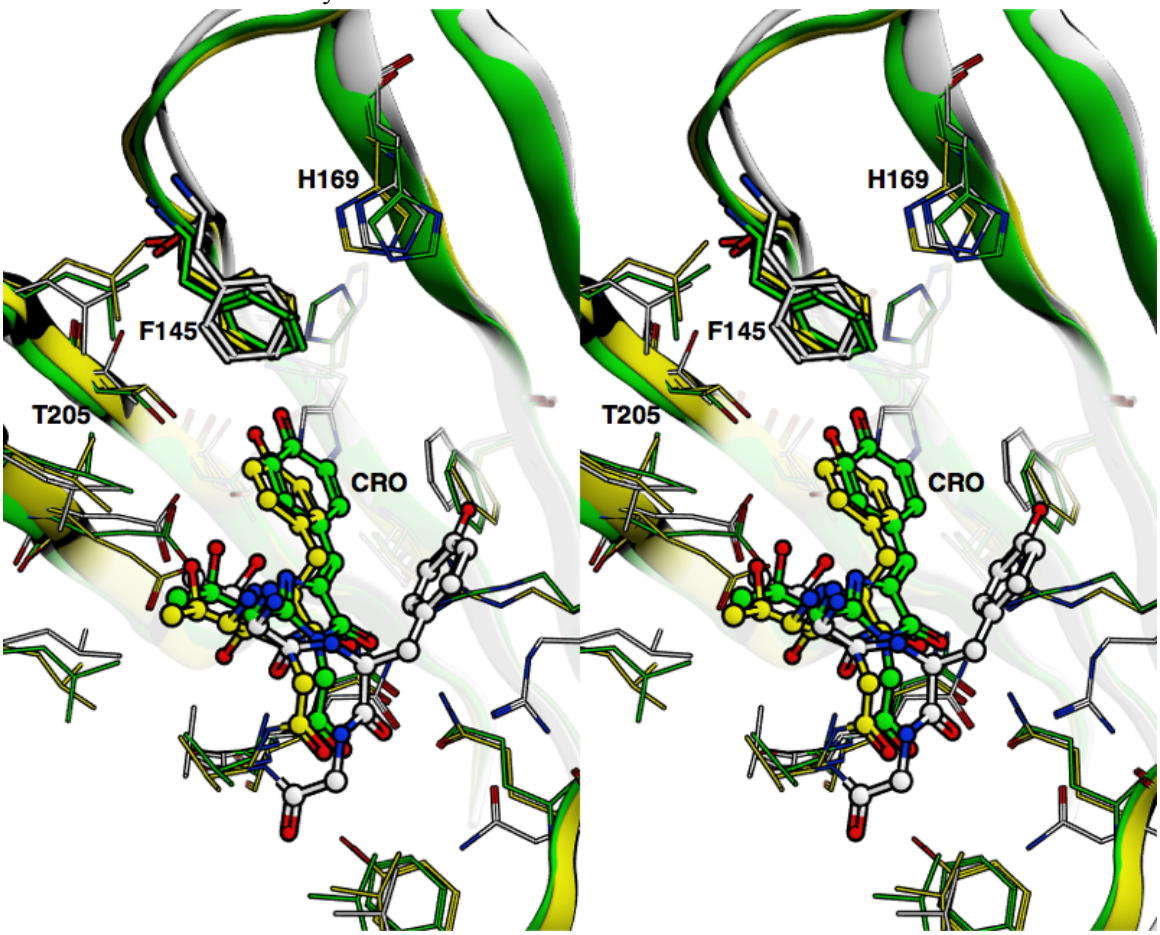
(a) Wildtype region around Q94 showing rotameric motion of Arg96 between precursor and mature state.



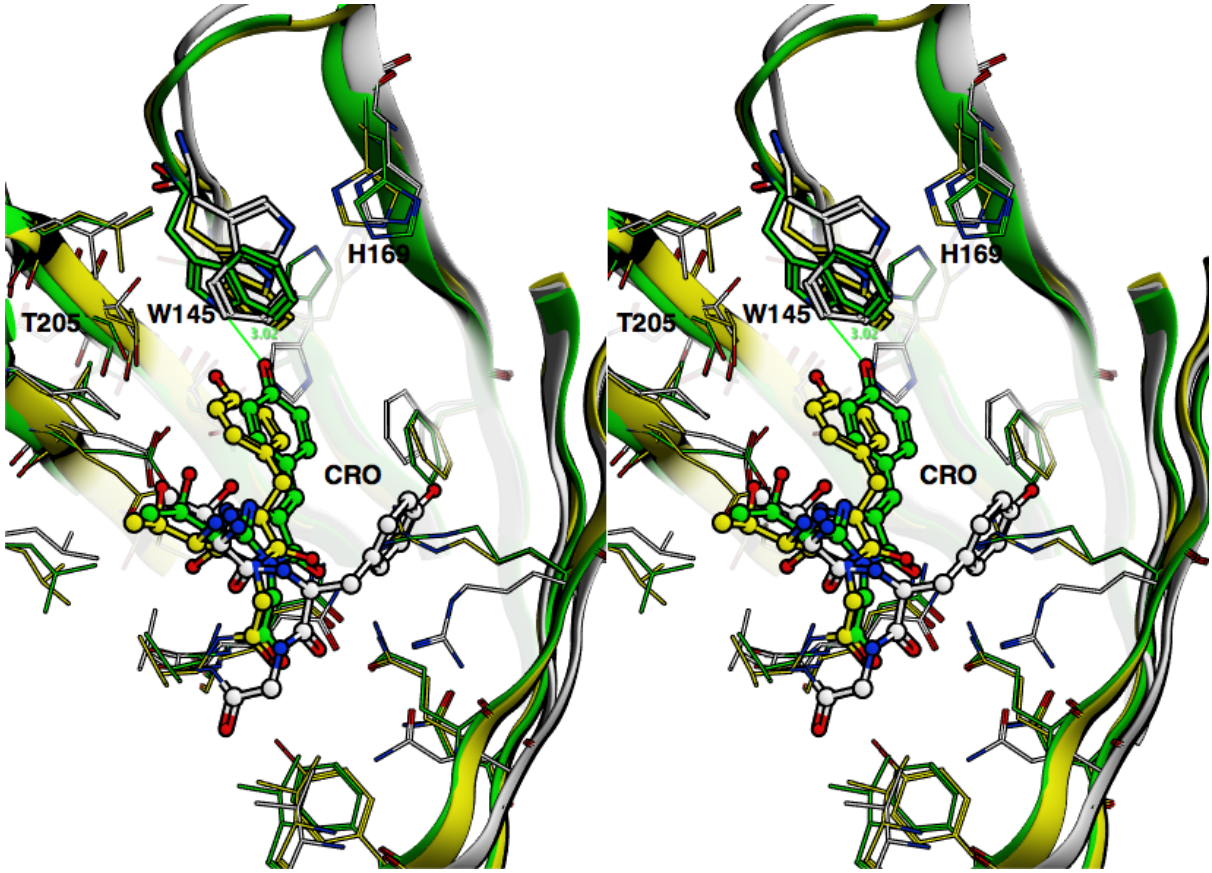
(b) Q94D mutant, region around Arg96 showing preferred non-native Arg rotamer in cyclized-only state. White bonds: precursor state. Yellow bonds: cyclized-only intermediate. Green bonds: fully mature state.



(c) Wildtype region around F145. White bonds: precursor state. Yellow bonds: cyclized-only intermediate. Green bonds: fully mature state.



(d) F145W region showing different preferred totamers in models of different stages of chromophore maturation. White bonds: precursor state. Yellow bonds: cyclized-only intermediate. Green bonds: fully mature state.



SUPPLEMENTAL VIDEO

Video S1. Chromophore maturation mechanism

Chromophore maturation pathway modeled based on pre-cyclized (PDB: 2AWJ), cyclized pre-oxidation (PDB: 2QRF) and native (PDB: 2B3P) structures showing the chromophore microenvironment and the transient appearance of molecular oxygen which leaves as hydrogen peroxide.
All required molecular editing was performed using the Molecular Operating Environment (MOE), version 2013.08 (Chemical Computing Group Inc).Interpolations of atomic coordinates in the movie were generated using Chimera (version 1.10)(*4*) . All visual rendering was performed using PyMOL (version 1.8.0.0)(*5*) .

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Image acquistion parameters for in vivo relative fluorescence, as reported in Table 1

Petri plates were imaged 24 h after induction using a DarkReader blue-light source and orange filter. The following equipment and setttings were used to obtain the relative fluorescence data.

Image Size: 3872 pixels x 2592 pixels. Image Quality: Compressed RAW (12-bit)
Camera Info -- Device: Nikon D3000. Lens: VR 18-55mm f/3.5-5.6G. Focal Length: 55mm. Focus Mode: Manual. AF-Area Mode: Single. VR: ON
Exposure -- Aperture: f/5.6. Shutter Speed: 1/5s. Exposure Mode: Shutter Priority. Exposure Comp.: 0EV. Metering: Matrix. ISO Sensitivity: ISO 400
Image Settings -- White Balance: Incandescent, 0, 0. Color Space: sRGB. High ISO NR: OFF. Long Exposure NR: OFF. Active D-Lighting: OFF.
Picture Control -- Picture Control: [VI] VIVID. Base: [VI] VIVID.Quick Adjust: 0.Sharpening: 4. Contrast: 0.Brightness: 0.Saturation: 0.Hue: 0.

SUPPLEMENTAL REFERENCES

(1) Kille, S., Acevedo-Rocha, C. G., Parra, L. P., Zhang, Z.-G., Opperman, D. J., Reetz, M. T., and Acevedo, J. P. (2013) Reducing Codon Redundancy and Screening Effort of Combinatorial Protein Libraries Created by Saturation Mutagenesis. *ACS Synth. Biol. 2*, 83–92.
(2) Leaver-Fay, A., Tyka, M., Lewis, S. M., Lange, O. F., Thompson, J., Jacak, R., Kaufman, K., Renfrew, P. D., Smith, C. A., Sheffler, W., Davis, I. W., Cooper, S., Treuille, A., Mandell, D. J., Richter, F., Ban, Y. E., Fleishman, S. J., Corn, J. E., Kim, D. E., Lyskov, S., Berrondo, M., Mentzer, S., Popović, Z., Havranek, J. J., Karanicolas, J., Das, R., Meiler, J., Kortemme, T., Gray, J. J., Kuhlman, B., Baker, D., and Bradley, P. (2011) ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzym. 487*, 545–574.
(3) O'Meara, M. J., Leaver-Fay, A., Tyka, M. D., Stein, A., Houlihan, K., DiMaio, F., Bradley, P., Kortemme, T., Baker, D., Snoeyink, J., and Kuhlman, B. (2015) Combined Covalent-Electrostatic Model of Hydrogen Bonding Improves Structure Prediction with Rosetta. *J. Chem. Theory Comput. 11*, 609–622.
(4) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and

Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605–12.
(5) Schrödinger, LLC. (2015) The {PyMOL} Molecular Graphics System, Version~1.8.
(6) Zhang, L., Patel, H. N., Lappe, J. W., & Wachter, R. M. (2006). Reaction progress of chromophore biogenesis in green fluorescent protein. Journal of the American Chemical Society, 128(14), 4766-4772.