Supplemental Information

Excited State Proton Transfer in the Red Fluorescent Protein mKeima

J. Nathan Henderson[†], Maire F. Osborn[†], Nayden Koon[†],

Rinat Gepshtein[‡],

Dan Huppert[‡] & S. James Remington[†]

[†]*Institute of Molecular Biology and Department of Physics, University of Oregon, Eugene, Oregon 97403-1229, USA, and* [‡]*Raymond and Beverly Sackler Faculty of Exact Sciences, School of Chemistry, Tel Aviv University, Tel Aviv.*

Materials and Methods

Gene synthesis, mutagenesis, protein expression and purification

A synthetic gene coding for mKeima was produced, based on the deposited DNA sequence, by Blue Heron Bioltechnology in a pBAD expression vector containing a Tobacco Etch Virus (rTEV) protease-cleavable N-terminal His_6-tag. Mutagenesis was performed using the QuikChange method (Stratagene). Protein was expressed in *Escherichia coli* strain Top10 by growing cultures to an OD_{600} around 0.6 at 37 °C. The temperature was then lowered to 25 °C, L-arabinose was added at a concentration of 2g/L and growth was continued overnight. Cells were pelleted by centrifugation then resuspended in 50 mM HEPES pH 7.9 + 300 mM NaCl + 10% glycerol. After 5 minutes of sonication, cell lysate was centrifuged and the supernatant applied to a column of Ni-NTA agarose (Qiagen) for purification. The column was washed with 50 mM HEPES pH 7.9 + 300 mM NaCl + 20 mM imidazole and the protein eluted with 50 mM HEPES pH 7.9 + 300 mM NaCl + 100 mM imidazole. Eluate was dialyzed into 50 mM HEPES pH 7.9 + 300 mM NaCl. The His_6-tag was removed by digestion overnight at 16° C with addition of a 1 to 20 mg ratio of recombinant TEV protease to mKeima. Following

application to a column of Ni-NTA agarose (Qiagen) and elution with 50 mM HEPES pH 7.9 + 300 mM NaCl + 20 mM imidazole, the protein was concentrated using Centriprep centrifugal filter units (Millipore) with molecular weight cut off 10,000 Da to a volume of 2.5 mL and then buffer exchanged into 50 mM HEPES pH 7.9 + 300 mM NaCl + 2 mM dithiothreitol (DTT) prior to crystallization trials.

Maturation assays

Prior to induction of protein expression for maturation assays, cultures of *E. coli* Top cells were handled as detailed above for protein purification. Following induction with 2 g/L of L-arabinose, the culture temperature was lowered to 30 °C and growth continued for 4 hours. Cell pellets resuspended in Lysis buffer were then sonicated for 5 minutes on ice. Subsequent steps in protein purification were carried out as quickly as possible at 4 °C. Cell lysate was centrifuge and the supernatant applied to a 5 mL HisTrap HP affinity column (GE Healthcare). The column was washed with 5 column volumes (CVs) of 50 mM HEPES pH 7.9 + 300 mM NaCl + 50 mM imidazole, followed by 5 CVs of 50 mM HEPES pH 7.9 + 300 mM NaCl + 100 mM imidazole. Protein was eluted with 5 CVs of 50 mM HEPES pH 7.9 + 300 mM NaCl + 200 mM imidazole and then concentrated to a volume of 2.5 mL. After buffer exchange with a PD-10 desalting column into 50 mM HEPES pH 7.9 + 300 mM NaCl, the sample was diluted to an OD_{280} of 5.5 with the same buffer for a final volume between 3.5 and 5 mL.

The maturation time course was initiated by moving the sample to room temperature (~ 22 °C) where time points were measured at 30, 60 and 90 minutes and at 2, 3, 4, 8, 16 and 24 hours. At each time point a small fraction of sample was removed for

measurements of absorbance, fluorescence and mass spectroscopy as well as for SDS-PAGE. For the measurement of zero time point sample was taken just prior to the transition from 4 °C to room temperature. It should be noted that complete maturation takes weeks at 4 °C.

Data from gel electrophoresis and fluorescence excitation scans at intermediate time points are shown in Figures S3A, S3B, while the electron density maps for the partial models of the acylimine and hydrolyzed (presumed carboxamide) forms are shown in Figure S4. We conclude that the process of partial hydrolysis takes place concomitant with formation of the red-emitting chromophore. Several attempts were made to separate the two species using various types of column chromatography, but were ultimately unsuccessful. There is no evidence for formation of a green GFP-like chromophore at any intermediate stage or as a dead end product.

Crystallization and X-ray diffraction data collection and structure solution

Primitive orthorhombic crystals of mKeima grew at room temperature in two weeks by hanging drop vapor diffusion using 2 μl of protein solution (OD_{280} = 32) mixed with 2 μl of well solution (26% PEG 3350 + 200 mM sodium chloride + 0.1 M Tris-HCl pH 7.0). Prior to data collection, a single crystal was equilibrated in 26% PEG 3350 + 200 mM sodium chloride + 0.1 M Tris-HCL + 15% ethylene glycol at room temperature for 20 minutes. Crystal were flash frozen in the cryostream at 100 K immediately prior to data collection.

Diffraction data were collected on an ADSC Q-210 CCD detector on beamline 5.0.2 at the Advanced Light Source (Berkeley, CA). Data were indexed and reduced using the HKL2000 suite (HKL Research). The structure was solved by molecular

replacement with EPMR[1] using the A chain of Rtms5 (PDB entry 1MOU) as a search model. Cycles of model building and crystallographic refinement were performed with the programs Coot[2] and Refmac 5[3]. It became apparent during refinement that the main chain is partially hydrolyzed at the acylimine linkage, as found in the "kindling fluorescent protein"[4]. Rather than attempt to model a complex chromophore mixture with partial occupancies, we adopted a different strategy. Two models were prepared, one containing an acylimine linkage while the other contained a carboxamide terminal group at Leu61 and a corresponding carbonyl at the alpha carbon of Gln62. The two models were independently refined against the data at 100% occupancy, resulting in very similar final R-factors. Difference Fourier maps were analyzed and found to be consistent with the presence of the two forms at approximately 50% occupancy each. Coordinates for only the acylimine model have been deposited in the Protein Data Bank, however coordinates for the presumed carboxamide form are available from the authors via email: remington@molbio.uoregon.edu

Time resolved spectroscopy

Time-resolved fluorescence was acquired using the time correlated single-photon counting (TCSPC) technique, the method of choice when sensitivity, large dynamic range, and low intensity illumination are important criteria in fluorescence decay measurements. For excitation, we used a cavity dumped Ti:sapphire femtosecond laser, Mira, Coherent, which provides short, 150 fs, pulses of variable repetition rate, operating at the SHG frequency, over the spectral range of 380-430 nm with the relatively low repetition rate of 500 kHz. The TCSPC detection system is based on a Hamamatsu

3809U, photomultiplier and Edinburgh instruments TCC 900 computer module for TCSPC. The overall instrumental response was about 35 ps (fwhm). Measurements were taken at 10 nm spectral width. The excitation pulse energy was reduced by neutral density filters to about 10 pJ. We checked the sample's absorption prior to and after time-resolved measurements. We could not find noticeable changes in the absorption spectra due to sample irradiation.

For the pump-probe experiments reported, we used an amplified femtosecond Ti:sapphire laser system. In brief, laser pulses (50 fs duration, centered near 800 nm with pulse energy of 600 μJ) at a 1 kHz repetition rate were generated by a Ti:sapphire-based oscillator (Coherent Mira seed) and amplified by a multipass Ti:sapphire amplifier (Odin Quantronix). Samples were excited by the second harmonic of the amplified laser (~400 nm). To obtain probe pulses, we generated a super continuum by focusing μJ of the 800 nm pulse onto a 2 mm thick sapphire window. The probe beam signal was measured by a combination of a chopper/lockin amplifier and computer averaging. Interference filters of 8 nm fwhm bandwidth at the proper wavelength were used in front of the probe beam detector, a silicon photodiode. Samples were placed in a rotating optical cell to avoid degradation.

References

1. Kissinger, C. R.; Gehlhaar, D. K.; Fogel, D. B. *Acta Crystallogr. D Biol. Crystallogr.* **1999**, 55, 484-491.
2. Emsley, P.; Cowtan, K. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, 60, 2126-2132.
3. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr. D Biol. Crystallogr.* **1997**, 53, 240-255.
4. Quillin, M. L.; Anstrom, D. M.; Shu, X.; O'Leary, S.; Kallio, K.; Lukyanov, K. A.; Remington, S. J., Biochemistry 2005, 44, 5774-5787.

Figure S1a.
Synthetic spectra for emission from the A* (neutral protonated, blue curve) and I* (anionic green curve) forms of the mKeima chromophore in D_2O. The observed spectrum is shown in black while the summed synthetic spectra are shown in red.

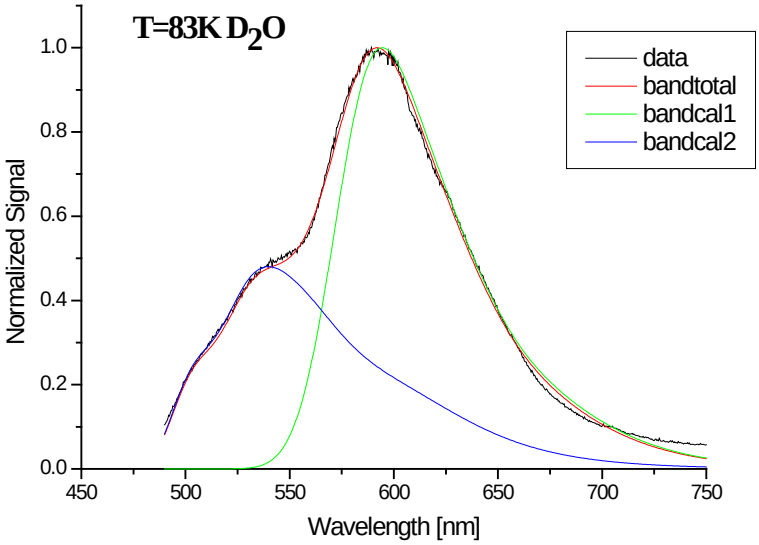


Figure S1b.
Synthetic spectra for emission from the A* (neutral protonated, red curve) and I* (anionic, cyan curve) forms of the mKeima chromophore in H_2O. The observed spectrum is shown in black while the summed synthetic spectra are shown in blue.

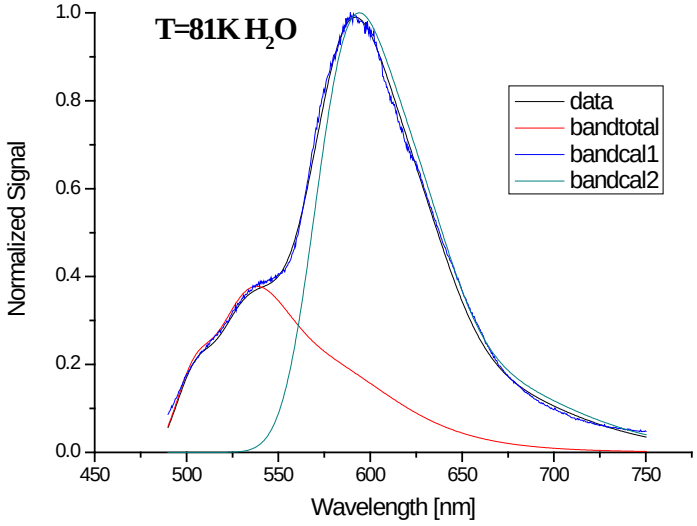


Figure S2A

mKeima Pump Probe Fit Results for 610 nm Emission

Table of Fit			
Amplitude1	0.14	τ1	0.01
Amplitude2	0.19	τ2	4
		τ3	450

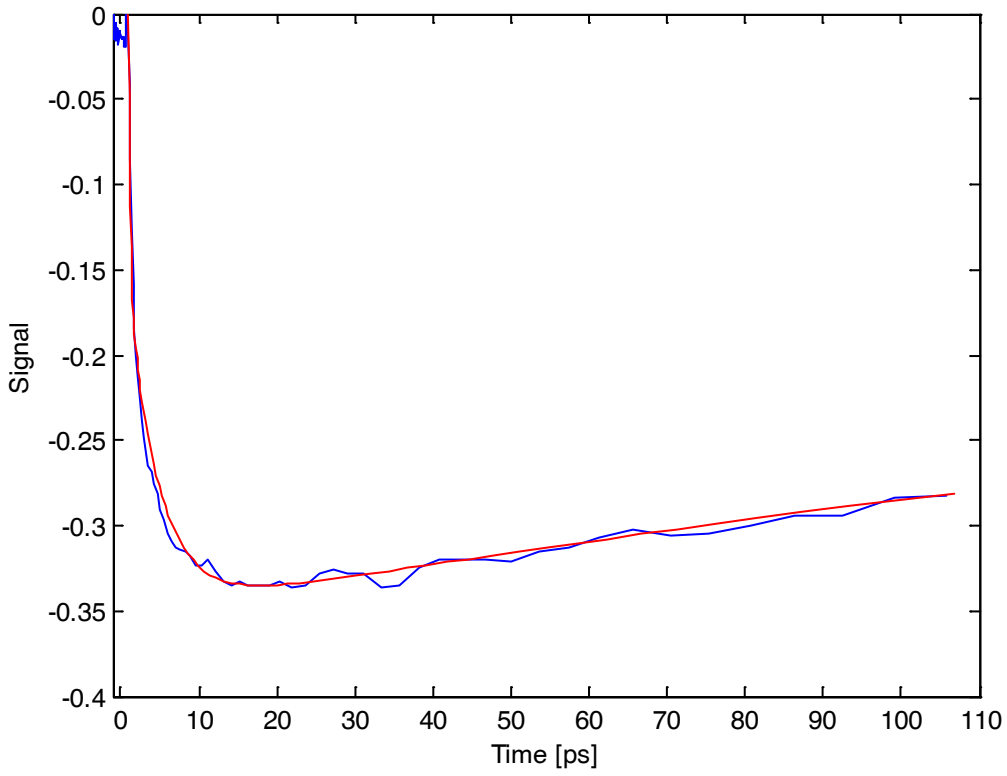
$Y = \exp(-t/tau3)^{x} (Amplitude1^{x}(1-\exp(-t/tau1)) + Amplitude2^{x}(1-\exp(-t/tau2)))$

Pump-Probe of mKeima

Pump at 400nm ; Det at 610nm

data=blue curve, fit=red curve

Rg07gd1008



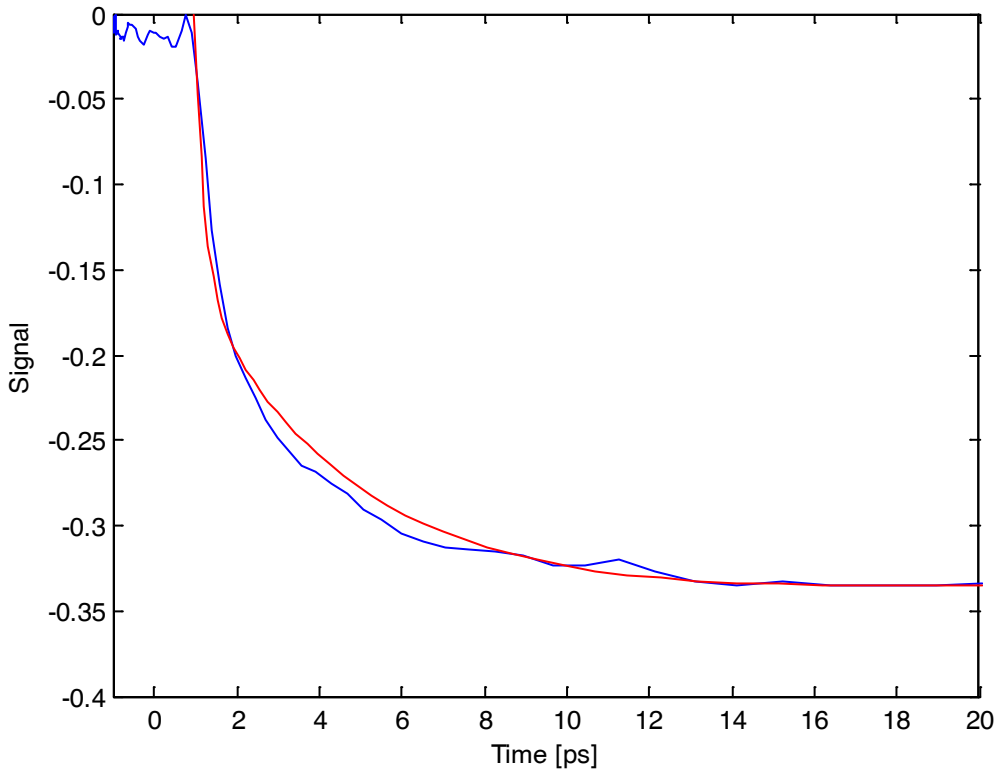
07/10/08

Figure S2B

Pump-Probe of mKeima, expanded time axis

Pump at 400nm ; Det at 610nm

Rg07gd1008



07/10/08

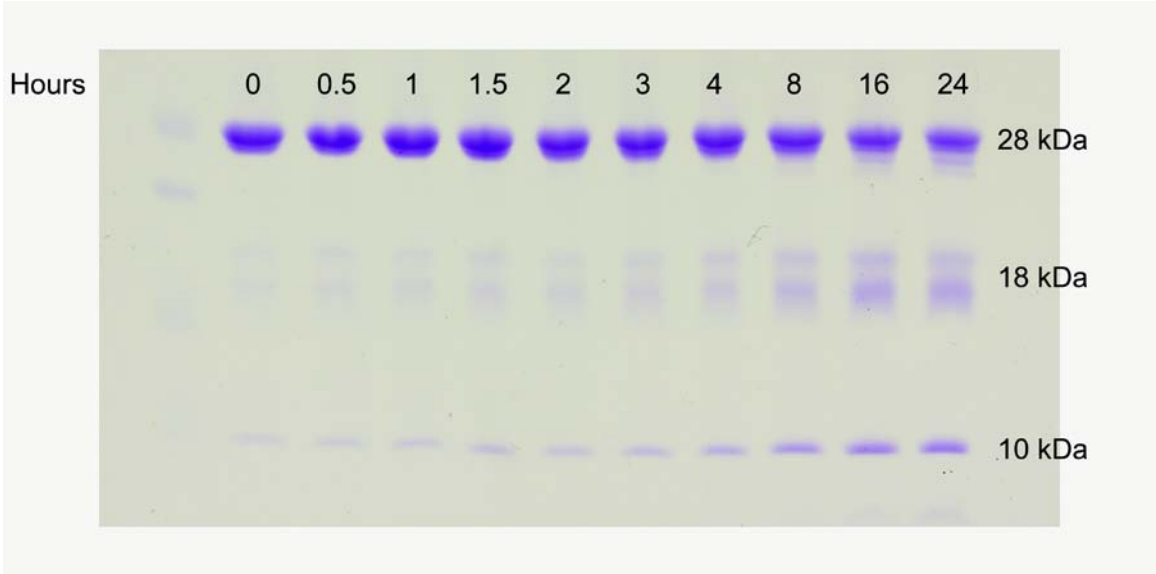


Figure S3A. Gel electrophoresis showing maturation of mKeima over the course of 24 hours. The given molecular weights correspond to the full length (28 kDa, including the His tag), C-terminal (18 kDa) and N-terminal (10 kDa, including the His tag) mKeima fragments and are corroborated by mass spectroscopy (data not shown). Note: a single peak is seen in the mass spectra for the C-terminal fragment, suggesting that multiple bands on the SDS gel of intermediate molecular weight are due to a gel artifact. The maturation process continues past 24 hours (see Methods, above). After 1 month, approximately the time between protein purification and crystal flash freezing for x-ray diffraction studies, the ratio between full length and cleaved protein had reached about 1:1 as determined by SDS PAGE of protein solutions and crystals (data not shown).

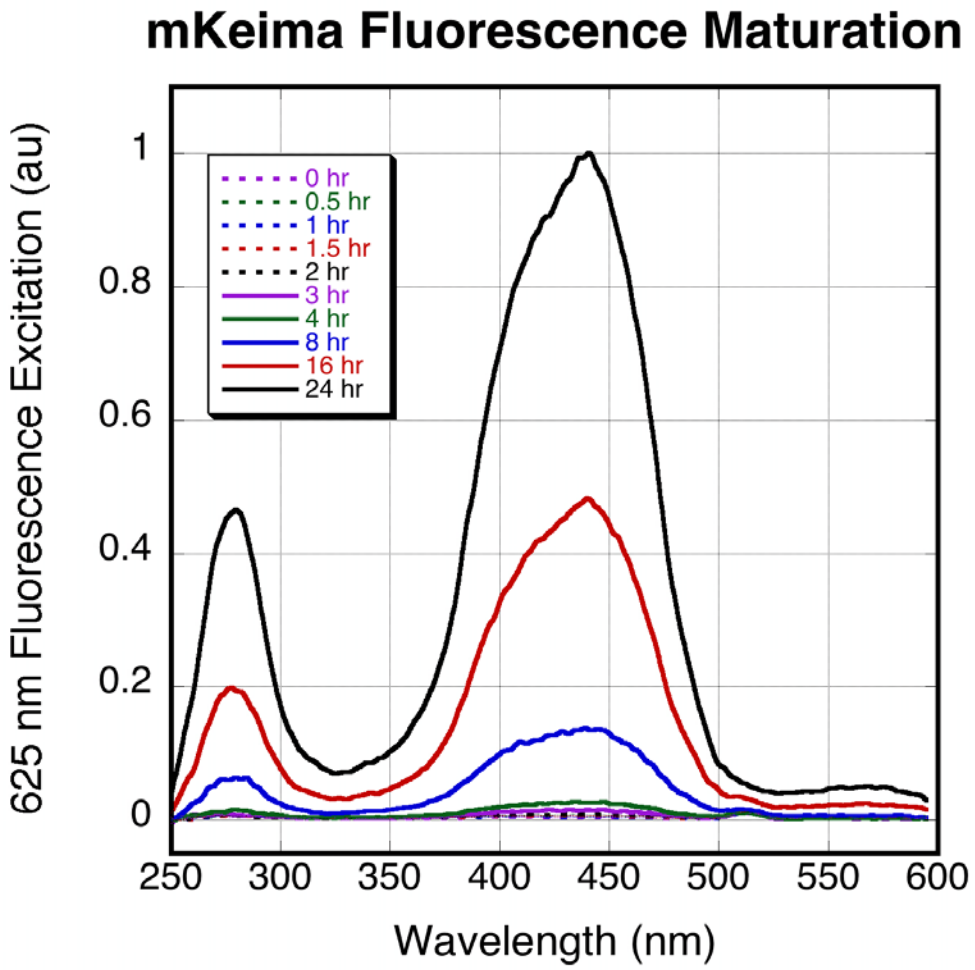


Figure S3B. Excitation spectra monitoring emission at 625 nm showing maturation of mKeima at room temperature over the course of 24 hours. Refer to the legend of Figure S3A and the text in the Materials and Methods section above for experimental details.

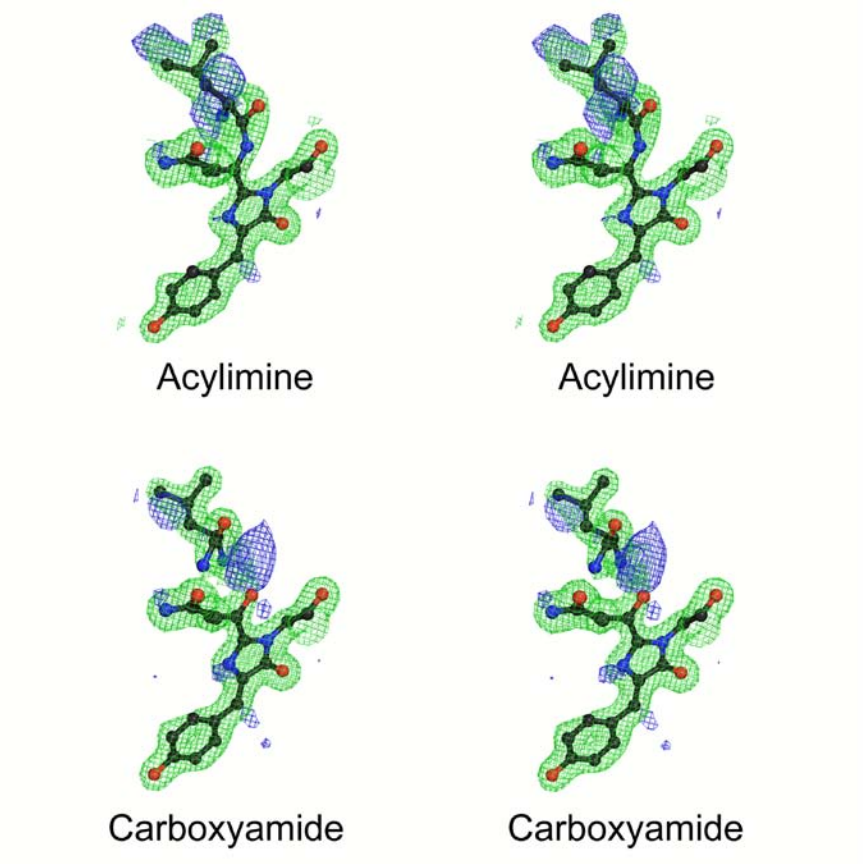


Figure S4. Stereo pairs of the two mKeima chromophore models, acylimine (above) and carboxyamide (below), superimposed on corresponding portions of the electron density maps contoured at 1 σ for $2F_o$-F_c map (green) and 2.5 σ for the F_o-F_c map (blue). Following refinement of each model against the data at 100% occupancy, the major residual difference density features spatially correspond with and are accounted for by the alternate model.

Figure S5. a) Sequence alignment of mKeima, Rtms5 and DsRed, with individual sequence numbering. b) Overlay of chromophore environment in Rtms5 (green) and mKeima (cyan). Residues are labeled according to (mKeima/Rtms5) numbering schemes. Although the chromophores in DsRed, intact mKeima and Rtms5 are chemically identical, of especial interest is the unusual placement of a positive charge (Lys71 in DsRed, also shown in orange; Arg193 in mKeima and Arg197 in Rtms5) that is required for chromophore maturation.

Fig S6a

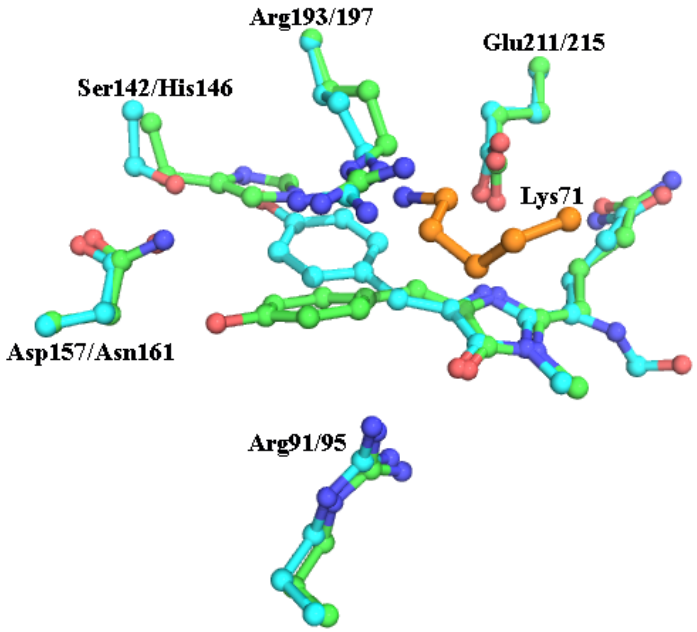


Fig. S6B

Table S1: Crystallographic Statistics	
Data collection	
Space group	$P2_12_12_1$
Cell dimensions (Å)	66.96, 75.69, 85.67
Total Reflections	1,520,095
Unique Reflections	101,363
Resolution[a] (Å)	50-1.63 (1.67-1.63)
R_{merge}[a,b]	0.054 (0.468)
$I/\sigma(I)$[a]	18.8 (2.0)
Completeness[a] (%)	96.7 (92.2)
Redundancy[a]	3.2 (2.7)
Refinement	
Resolution (Å)	31.19 (1.63)
R_{work}[c]	0.193
R_{free}	0.235
R_{factor} (all data)	0.195
Number of molecules[d]	2
Number of protein atoms[d]	3408
Number of solvent atoms[d]	306
Average B factors (Å^3)	16.4
Protein atoms	15.7
Solvent atoms	24.0
rmsd bond lengths (Å)	0.025
rmsd bond angles (deg)	2.41

[a] The values in parentheses indicate statistics for the highest resolution shell. [b] $R_{merge} = \Sigma|I - \langle I \rangle|/\Sigma\langle I \rangle$, where I is the observed intensity and $\langle I \rangle$ is the average intensity from multiple observations of symmetry-related reflections. [c] $R = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are the observed and calculated structure amplitudes, respectively. [d] Per asymmetric unit.