

# Supporting Information for Fluorescence from Multiple Chromophore Hydrogen-Bonding States in the Far-Red Protein TagRFP675

Patrick E. Konold<sup>1,2</sup>, Eunjin Yoon<sup>3</sup>, Junghwa Lee<sup>3</sup>, Samantha Allen,<sup>1,2</sup> Prem P. Chapagain<sup>4</sup>, Bernard S. Gerstman<sup>4</sup>, Chola K. Regmi<sup>4,5</sup>, Kiryl D. Piatkevich<sup>6</sup>, Vladislav V. Verkhusha,<sup>7</sup> Taiha Joo<sup>3</sup> and Ralph Jimenez<sup>1,2,\*</sup>

<sup>1</sup> JILA, University of Colorado and NIST, Boulder, CO 80309

<sup>2</sup> Department of Chemistry & Biochemistry, University of Colorado, Boulder, CO 80309

<sup>3</sup> Department of Chemistry, Pohang University of Science and Technology (POSTECH), Pohang, South Korea 790-784

<sup>4</sup> Department of Physics, Florida International University, Miami, FL 33199

<sup>5</sup> Department of Physics, Virginia Tech, Blacksburg, VA 24061

<sup>6</sup> Massachusetts Institute of Technology Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139

<sup>7</sup> Department of Anatomy & Structural Biology, Albert Einstein College of Medicine, Bronx, NY 10461

\* Correspondence should be addressed to R.J. (rjimenez@jila.colorado.edu) or T.J. (thoo@postech.ac.kr)

## Contents:

1. Experimental and computer simulation methods
2. Absorbance/emission spectra of TagRFP675 mutants
3. Contour plots of SRTG data for TagRFP675 mutants
4. SRTG SE band peak shifts and exponential fit parameters
5. DAS from global analyses of SRTG data and table of DAS time constants
6. Linear-log plot of TRF peak positions and fits and table of fit parameters
7. Excitation-dependent emission spectra of TagRFP675 and mKate/M41Q

## Methods

### *SRTG Spectroscopy*

Transient grating measurements were performed as previously described with the JILA MONSTR nonlinear optical platform.<sup>1-2</sup> Upon irradiation with 3 excitation pulses (20 fs,  $\sim 10$  nJ/beam, 50  $\mu\text{m}$  spot size, 20 kHz repetition rate), the spatially isolated four-wave mixing signal is emitted in the prescribed  $\mathbf{k}_s = -\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3$  phase matching direction. The excitation wavelength was chosen to maximize contrast between Ground State Bleach (GSB) and Stimulated Emission (SE) components, which varied slightly for each mutant. Measured excitation spectra are provided in Figure S1. The sample was refreshed with a spinning cell to mitigate photobleaching and accumulation of laser-induced photoproducts. The sample concentration ( $\sim 50$   $\mu\text{M}$ ; 0.15 peak optical density) and optical path length (0.5 mm) were selected to avoid protein aggregation and reabsorption of the nonlinear signal. The spectrally-resolved homodyne signal was collected as a function of the  $T$  time delay ( $\mathbf{k}_1$  and  $\mathbf{k}_2$  temporally overlapped) from a range of 0 to 1.3 ns, spaced logarithmically, with a background spectrum collected at -500 fs delay to eliminate contributions from excitation scatter and spontaneous emission. The baseline-subtracted spectra were fit to a sum of two Gaussians using a nonlinear least squares fitting algorithm. The SE peak center position and amplitude was extracted at each time delay and fit to multiexponential decays.

### *Femtosecond TRF*

Output from a home-built cavity dumped optical parametric oscillator (OPO) employing a magnesium-doped periodically poled lithium niobate (MgO:PPLN) crystal as the gain medium was frequency-doubled in a 2 mm-thick lithium triborate (LBO) crystal to generate the pump pulse.<sup>3</sup> The pulse spectrum was 310  $\text{cm}^{-1}$  FWHM and centered at 16670  $\text{cm}^{-1}$  (600 nm) for the TRF of TagRFP675 and at 17540  $\text{cm}^{-1}$  (570 nm) for mKate/M41Q. The excitation wavelength was tuned near the absorption maximum to wavelengths where the laser system gave nearly identical bandwidth excitation pulses (310  $\text{cm}^{-1}$  FWHM) for the two measurements. The excitation wavelength for mKate/M41Q would 270  $\text{cm}^{-1}$  blue-shifted from the absorption peak. For a room-temperature measurement, this difference in excitation conditions is unlikely to result in measurable differences on ps/ns timescales. The remaining infrared pulse energy was used for the gate pulse. Femtosecond time

resolution was achieved by sum frequency generation (SFG) between the fluorescence and the gate pulse in 300  $\mu\text{m}$ -thick  $\beta$ -barium borate (BBO) crystal. The SFG signal was spectrally filtered by a double monochromator (DH10, Horiba Jobin Yvon) and detected by a photomultiplier (H6180-01, Hamamatsu) and a photon counter (SR400, Stanford Research). TRF spectra (TRFS) were measured directly without the spectral reconstruction by simultaneously controlling the phase matching angle of the BBO crystal, the time delay of the gate pulse, and the detected SFG wavelength. The instrument response function of the TRF setup was  $<200$  fs (FWHM). A 200  $\mu\text{m}$ -thick flow cell (Starna Inc.) containing the protein samples ( $\sim 200$   $\mu\text{M}$ ) was shaken to avoid any photodamage during the measurements. To avoid the scattering of the pump pulse and the second harmonic of the gate pulse, the blue edge of the fluorescence spectra was eliminated.

#### *Picosecond TRF and TRF Anisotropy*

The OPO used for the femtosecond TRF experiments was employed for Time-Correlated Single Photon Counting (TCSPC) to record picosecond-nanosecond TRFS. Details of the TCSPC setup have been presented elsewhere<sup>4</sup>. The fluorescence signal was collected and dispersed by a monochromator (SP-300i, Acton Research) and detected with an avalanche photodiode (ID-100-50-STD, ID Quantique). The instrument response function of the TCSPC system was 50 ps (FWHM). TCSPC time traces were obtained at every 5 nm step size to reconstruct the picosecond TRFS. Magic angle ( $54.7^\circ$ ) detection was used to avoid the effect of polarization on timescales of the picosecond TRF measurement.

#### *MD Simulations*

Time series trajectories were obtained with explicit solvent, all-atom simulations using the NAMD molecular dynamics package with the CHARMM27 force field<sup>5-6</sup>. The initial X-ray crystallographic structures were obtained from the Protein Data Bank (TagRFP675 PDB 4KGF and mKate PDB 3BXB). We used CHARMM to make the amino acid substitutions in the TagRFP675 mutants F62A, Q106M, and F62A/Q106M as well as for the mKate/M41Q mutant. Force field parameters for the mature chromophore were adopted from the anionic GFP chromophore developed by Reuter *et al.* and from CHARMM27 parameters for

acylimine nitrogen<sup>7</sup>. The anionic form of the chromophore in the ground state was used. The VMD package was used to setup the system for simulations<sup>8</sup>. The initial structure of TagRFP675 with crystallographic water molecules was solvated using VMD. For all mutants, the box cutoff was set to 10 Å. For TagRFP675, this resulted in a simulation box of dimensions of 66 x 72 x 77 Å<sup>3</sup>, with similar dimensions for the other mutants. Each solvated system for TagRFP675 and its variants was electrically neutralized by adding 1 Na<sup>+</sup> ion randomly in the bulk water. The final system contained about 35,000 atoms. The particle mesh Ewald method was used to treat long-range interactions with a 12 Å nonbonded cutoff<sup>9</sup>. Energy minimization was performed using the conjugate gradient and line search algorithm. The system was then heated for 90 ps with a linear gradient of 20K/6ps from 20 to 300 K. At 300 K, the system was equilibrated for 910 ps with a 2 fs integration time step in the NVT (constant number, volume, and temperature) ensemble. Langevin dynamics was used to maintain the temperature at 300 K. The production run was 50 ns using NVT dynamics with 2 fs time steps.

#### *Protein Mutagenesis, Expression, and Purification*

FPs were expressed with the pBAD vector containing the corresponding genes under control of the arabinose promoter. Point mutations were introduced with QuikChange mutagenesis using custom designed primers (IDT). Expression procedures differed slightly for the two primary lineages studied. For mKate and mutants, plasmids were transformed into Top10 *E. coli*, grown to OD<sub>600</sub>=0.6 at 37 °C, followed by induction with 0.02% arabinose for 24 hours at 30 °C. Plasmids of TagRFP675 and mutants were transformed into the LMG194 *E. coli* strain, grown to OD<sub>600</sub>=0.6 at 37 °C, followed by induction with 0.02% arabinose for 48 hours at 18 °C. Samples were purified via 6X-His tag/Ni-NTA chromatography and then buffer-exchanged into 15 mM pH 8.0 TRIS buffer, 100 mM KCl. Emission spectra of each mutant were collected with a diode array spectrometer, with DPSS laser excitation at 18797 cm<sup>-1</sup> (532 nm).

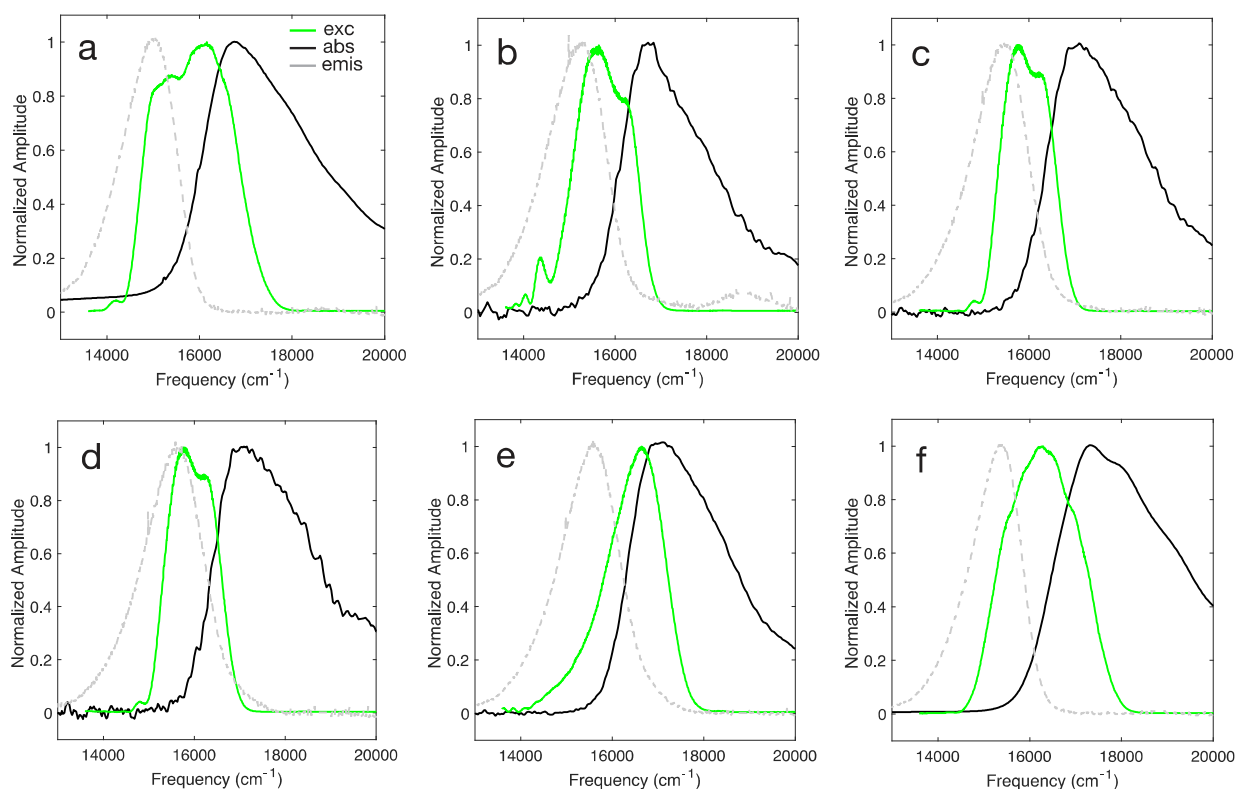
#### *Global fitting method*

Transients at each wavelength were nonlinear least-squares fitted to a sum of exponential functions convoluted with an instrument response function, which was assumed to be a

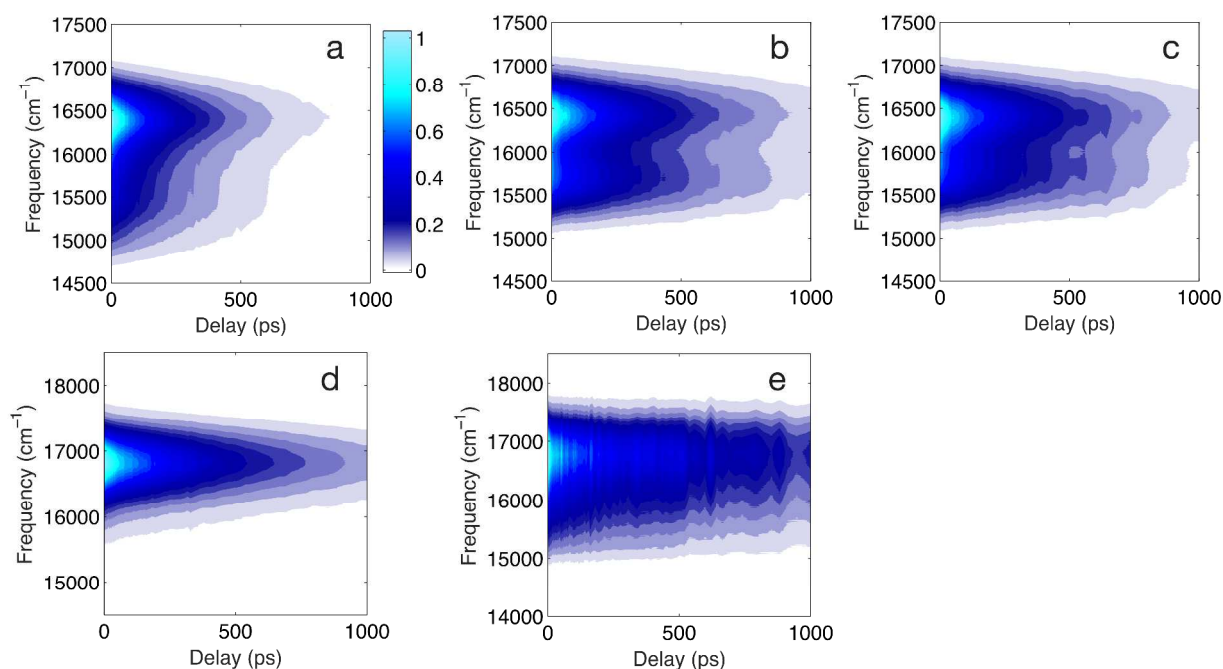
Gaussian with a standard deviation of 70 fs determined by the cross correlation between the gate and scattered pump pulse. Amplitudes  $A(\lambda)$  are assumed to be wavelength-dependent with a common set of time constants that applies across the spectra (Equation S1):

$$I(\lambda, t) = \sum_j A(\lambda) e^{-t/\tau_j} \quad (\text{S1})$$

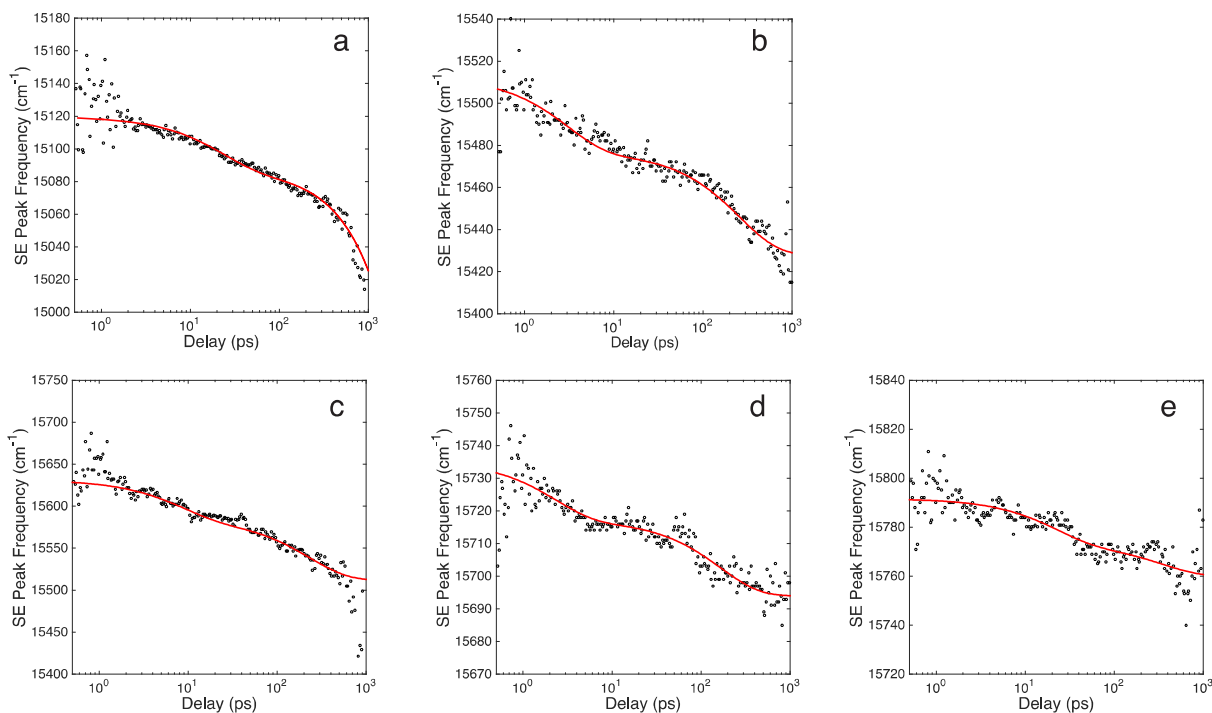
Initial values for the amplitudes were all set to the amplitude at each wavelength at time zero.



**Figure S1.** Steady state absorption and emission spectra with SRTG excitation spectra overlaid for (a) TagRFP675, (b) TagRFP675/F62A, (c) TagRFP675/Q106M, (d) TagRFP675/F62A/Q106M (e) TagRFP675/Q41M (f) mKate/M41Q.



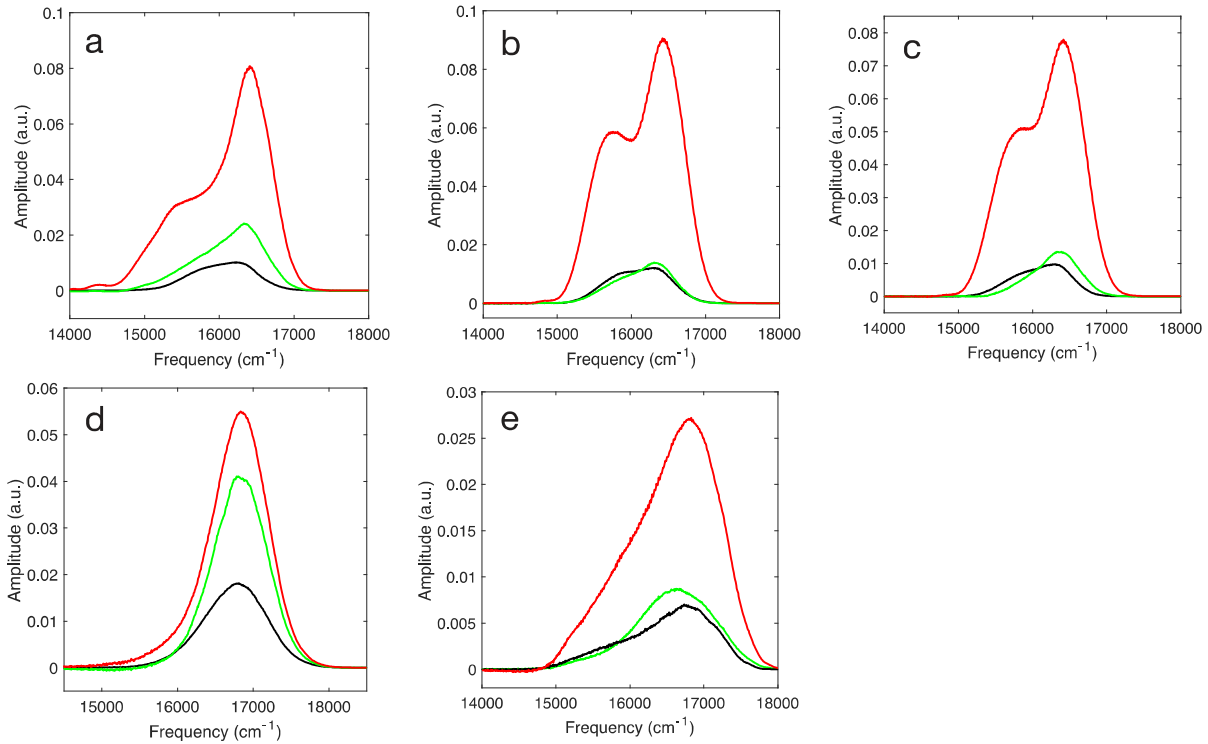
**Figure S2.** SRTG spectra of (a) TagRFP675/F62A, (b) TagRFP675/Q106M, (c) TagRFP675/F62A/Q106M, (d) TagRFP675/Q41M, (e) mKate



**Figure S3.** SRTG stimulated emission band peaks vs. time determined by fitting each band to a Gaussian function for (a) TagRFP675 and (b) mKate/M41Q, (c) TagRFP675/F62A, (d) TagRFP675/Q106M, (e) TagRFP675/F62A/Q106M. Exponential fits are overlaid in red. Fit parameters given in Table S2.

**Table S1.** Parameters for the multi-exponential fits to SRTG SE band peak position to the functional form  $I(t) = a_1 \exp(-t/\tau_1) + a_2 \exp(-t/\tau_2)$

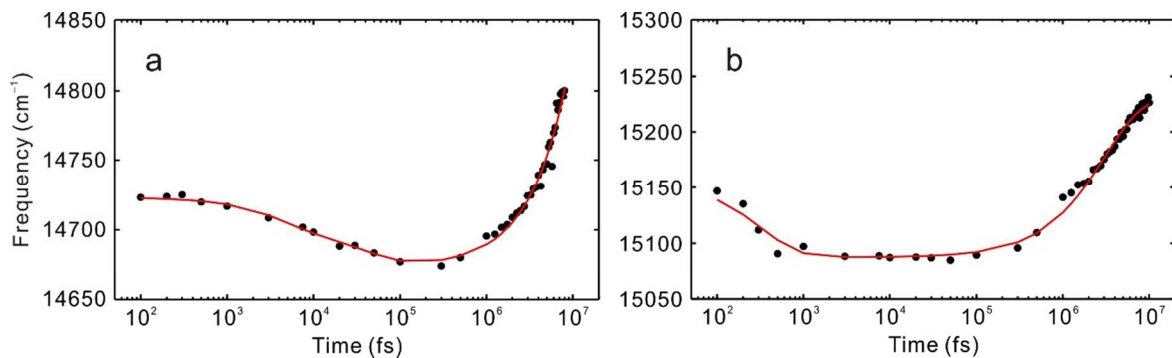
Mutant	$a_1$ (cm <sup>-1</sup> )	$\tau_1$ (ps)	$a_2$ (cm <sup>-1</sup> )	$\tau_2$ (ps)
mKate/M41Q	30	5	52	355
TagRFP675	33	17	44	397
TagRFP675/F62A	41	7	68	159
TagRFP675/Q106M	12	4	24	219
TagRFP675/F62A/Q106M	18	24	17	537



**Figure S4.** DAS for SRTG of (a) TagRFP675/F62A, (b) TagRFP675/Q106M, (c) TagRFP675/F62A/Q106M, (d) TagRFP675/Q41M, (e) mKate. Time constants for each component are given in Table S2.

**Table S2.** Time constants for DAS in Figure S4 and Figure 2 tabulated according to color. Red – slowest ( $\tau_1$ ), green - next fastest ( $\tau_2$ ), black -fastest ( $\tau_3$ ).

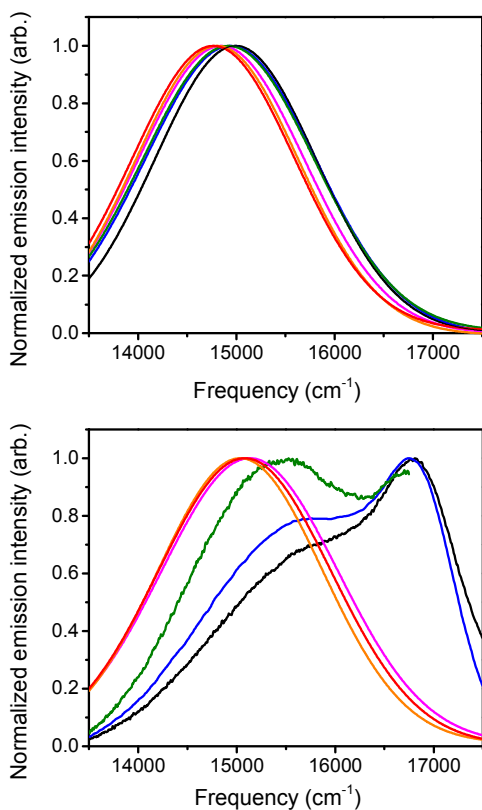
	$\tau_1$ (ps)	$\tau_2$ (ps)	$\tau_3$ (ps)
TagRFP675	434	134	19
TagRFP675/F62A	585	190	14
TagRFP675/Q106M	794	88	8
TagRFP675/F62A/Q106M	818	112	19
TagRFP675/Q41M	1224	520	88
mKate	1916	142	24
mKate/M41Q	1200	457	9



**Figure S5.** Transient emission peak positions (first moments) extracted from femtosecond and picosecond TRF spectroscopy for (a) TagRFP675 (b) mKate/M41Q.

**Table S3.** Fit parameters for the time evolution of the first moment of TRF measurements. Fit function is  $\nu(t) = a_1 \exp(-t/\tau_1) + a_2 \exp(-t/\tau_2) + \dots$

Mutant	$a_1$ (cm <sup>-1</sup> )	$\tau_1$ (ps)	$a_2$ (cm <sup>-1</sup> )	$\tau_2$ (ps)	$a_3$ (cm <sup>-1</sup> )	$\tau_3$ (ns)	$a_4$ (cm <sup>-1</sup> )	$\tau_4$ (ns)
TagRFP675	23	4.5	26	44	1330	16	-3770	23
mKate/M41Q	69	0.33	2	15	-128	5.4		



**Figure S6.** 7. Excitation-dependent emission spectra of (top) TagRFP675 and (bottom) mKate/M41Q. For both plots, the colors correspond to excitation at red-15310 cm<sup>-1</sup>, orange-15850 cm<sup>-1</sup>, magenta-16420 cm<sup>-1</sup>, green- 17040 cm<sup>-1</sup>, blue- 18050 cm<sup>-1</sup>, and black- 19190 cm<sup>-1</sup>.



## References

1. Bristow, A. D.; Karaiskaj, D.; Dai, X.; Zhang, T.; Carlsson, C.; Hagen, K. R.; Jimenez, R.; Cundiff, S. T. A Versatile Ultrastable Platform for Optical Multidimensional Fourier-Transform Spectroscopy. *Review of Scientific Instruments* **2009**, *80* (7).
2. Konold, P.; Regmi, C. K.; Chapagain, P. P.; Gerstman, B. S.; Jimenez, R. Hydrogen Bond Flexibility Correlates with Stokes Shift in mPlum Variants. *The Journal of Physical Chemistry B* **2014**, *118* (11), 2940-8.
3. Min, C.-K.; Joo, T. Near-Infrared Cavity-Dumped Femtosecond Optical Parametric Oscillator. *Opt. Lett.* **2005**, *30* (14), 1855-1857.
4. Manoj, P.; Min, C.-K.; Aravindakumar, C. T.; Joo, T. Ultrafast Charge Transfer Dynamics in 2-Aminopurine Modified Double Helical DNA. *Chemical Physics* **2008**, *352* (1-3), 333-338.
5. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *Journal of Computational Chemistry* **2005**, *26* (16), 1781-1802.
6. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM - a Program for Macromolecular Energy, Minimization and Dynamics Calculations. *Journal of Computational Chemistry* **1983**, *4* (2), 187-217.
7. Reuter, N.; Lin, H.; Thiel, W. Green Fluorescent Proteins: Empirical Force Field for the Neutral and Deprotonated Forms of the Chromophore: Molecular Dynamics Simulations of the Wild Type and S65T Mutant. *Journal of Physical Chemistry B* **2002**, *106* (24), 6310-6321.
8. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *Journal of Molecular Graphics & Modelling* **1996**, *14* (1), 33-38.
9. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *Journal of Chemical Physics* **1995**, *103* (19), 8577-8593.