**Supporting Information for:**

# Fluorescence Multiplexing with Spectral Imaging and Combinatorics

Hadassa Y. Holzapfel[1,2], Alan D. Stern[1], Mehdi Bouhaddou[1], Caitlin M. Anglin[3], Danielle Putur[1], Sarah Comer[1] and Marc R. Birtwistle[1,3,#]

[1]Department of Pharmacological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

[2]Medical School for International Health, Faculty of Health Sciences, Ben-Gurion University of the Negev, Be'er Sheva, 84105, Israel

[3]Department of Chemical and Biomolecular Engineering, Clemson University, Clemson, SC 29634, USA

[#]To whom correspondence should be addressed mbirtwi@clemson.edu

## SUPPLEMENTAL METHODS

### *Computational Methods*

*Availability.* All MATLAB code used to perform simulations and analyze data, including the raw data, are included in the supplementary code zip file associated with this manuscript. The scripts "Fig2.m" and "Fig3.m" are the starting points for regenerating the analyses.

*Data Sources.* Data for fluorescent proteins (excitation spectra, emission spectra, brightness) were gathered from individual references (EBFP2[1], mTagBFP2[2], mT-Sapphire[3], mAmetrine[4], mCerulean3[5], mTFP1[6], LSSmOrange[7], EGFP[8], TagYFP[9], mPapaya1[10], mOrange2[11], mRuby2[12], TagRFP-T[11], mKate2[13], mCardinal[14], and iRFP[15]), the Nikon Imaging Center at UCSF (nic.ucsf.edu; fpvis.org), which has subsequently converted to FPbase (fpbase.org), and the Tsien lab website excel file (http://www.tsienlab.ucsd.edu/Documents.htm). All used spectra and other quantitative properties are also contained within the provided MATLAB code.

*Simulated FRET Efficiency.* FRET efficiency $E$ is related to the distance between donor and acceptor $r$, and a constant $R_0$ called the Förster radius as follows

$$E = \frac{1}{1+\left(r/R_0\right)^6}. \qquad (1)$$

The Förster radius depends on a dipole orientation factor, medium refractive index, donor quantum yield, multiple unit conversion factors and universal constants, and a spectral overlap integral $J$

$$J = \int_0^\infty f_D^{em} f_A^{ex} \lambda^4 d\lambda \qquad (2)$$

Where $f$ is a normalized spectrum (area 1) which depends on wavelength, $D$ denotes donor, $A$ denotes acceptor, *em* denotes emission, *ex* denotes excitation, and $\lambda$ is wavelength (in microns here

for numeric purposes). For the purposes of estimation in the simulation studies, we take all factors except the overlap integral as defined in Eq. 2 as roughly consistent across potential FRET pairs. The overlap integral is calculated using the function `trapz` in MATLAB. We aggregate the constants outside of the overlap integral and estimate its value based on data for mTFP1-mVenus[16] (73.4, see code). FRET efficiency is capped at 0.35, particularly for the red-shifted proteins which have higher estimated $E$ presumably due to the $\lambda^4$ term in the integral.

*Simulated Fluorescence Emission Spectra.* We calculate the emission intensity spectrum of a probe given a particular excitation wavelength as the sum of multiple emission processes. First, the relative excitation strength of an individual fluorescent protein (FP) is taken from the excitation spectra (with a maximum of 1). Thus, if the excitation wavelength is at the peak of the emission spectra, the excitation strength is 1, and the value along the curve otherwise. The excitation, regardless of the strength, leads to emission distributed across the entire emission spectrum. The overall emission spectra intensity is then multiplied by the relative probe levels (i.e. concentration) and the relative brightness of the FP (comparable across FPs). In the case when a two-FP probe is considered, direct excitation of the donor is calculated as above, but is reduced by the fraction of donor molecules estimated to be undergoing FRET. This fraction is calculated as the estimated FRET efficiency minus the fraction of acceptor that are estimated to be directly excited by the excitation light. We call this the adjusted FRET efficiency. The emission intensity of the acceptor is therefore then composed of two parts scaled by the adjusted FRET efficiency; the first due to FRET from the donor, and the second due to direct acceptor excitation. In the case of a three-FP probe the same logic is extended to the additional interactions to calculate the overall emission spectrum contributions. These calculations are contained within the Supplementary Code in the

MATLAB files `CalcIntensity1F.m`, `CalcIntensity2F.m`, and `CalcIntensity3F.m`. Noise is added to spectra by sampling from a standard normal distribution that is multiplied by 10 (arbitrary units) using the MATLAB function `randn`. This denotes a constant signal to noise of 100 at the maximum probe concentration possible of 1000 (arbitrary units). Thus, this modeled noise has much greater effects on low abundance probes that constitute the majority of the simulated cases.

*Unmixing.* Linear unmixing is performed similarly on simulated and experimental data. First, each emission spectrum is normalized such that the area under the curve is 1 (i.e. divide by sum). For references, the sum used for such normalization is stored for adjustment of probe level estimates later. Then, the MATLAB function `lsqlin` is used to estimate probe levels given a normalized reference matrix and normalized sample vector. Lower bounds of zero and upper bounds of one are used for the levels of each probe, and we constrain the sum of all probe levels to be equal to one, since when the unmixing problem is formulated like this, the probe levels are fractional abundances in the mixture (once corrected for the above noted sums). The function `Unmix.m` in the Supplementary Code contains these calculations. In this case of experimental data, first the averages across triplicates are taken, and then both the reference and mixture data are blank and background (PBS) subtracted prior to analysis.

***Experimental Methods***

*Fluorescent Protein Cloning*

Fluorescent protein DNA templates were obtained (LSSmOrange[7]: Addgene #37135; EBFP2[1]: Addgene #14893; ECFP: Gift from Susana Neves; mTFP1 and mVenus, as previously[16]), and sequences amplified by PCR for recombination into DONR221 (Invitrogen) via BP clonase (ThermoFisher Cat: 11789020) following the manufacturer's protocol. For single probes, a stop

codon was included. For dual probes, the first fluorescent protein was amplified without a stop codon, followed at the 3' end with AgeI-linker-HindIII (linker: GCC GGA GGT GGG GGC CTA GGA). The second fluorescent protein was then added via restriction digest and ligation cloning using the AgeI and HindIII sites. This resulted in an amino acid sequence of TGAGGGGLG between FPs in a tandem probe. Constructs were then recombined using LR Clonase (ThermoFisher Scientific, Cat: 11791100) into pQLinkHD[17] for protein expression with an N-terminal His-tag (Addgene, Cat: 13668). Constructs were verified by Sanger sequencing.

*Protein Expression*

*E. coli* (BL-21 strain) containing the plasmid-of-interest were taken from a glycerol stock using a sterile pipette tip and added to 5 mL LB (Alfa Aesar, Cat: H26760-36) containing ampicillin (100 ug/mL, Sigma, Cat: A9518-25G) in a 14 mL culture tube (Falcon, Cat: 352063). The cultures were incubated overnight (not more than 16 hours) in an orbital shaker at 37°C and 180 RPM. The following day the culture was added to 100 mL of sterile expression media [1L solution: 20 g bactotryptone (VWR, Cat: 90000-282), 15 g yeast extract (VWR, Cat: 97063-370), 8 g NaCl (Sigma, Cat: S7653-1KG), 4 g $NaH_2PO_4$ (Amresco, Cat: 0571-500g), 2 g $KH_2PO_4$ (VWR, Cat: 97062-350) pH to 7.5 with NaOH], 5 mL 40% (40 g / 100 mL) sterile glucose (Alfa Aesar, Cat: A16828) and 100 ug/mL ampicillin in a 1-2L Erlenmeyer flask and incubated in an orbital shaker at 180 RPM and 37ºC until OD600 was ~1 (roughly 2-3 hours). Expression was then induced with 1 mM IPTG (VWR, Cat: TCI0328-005G) and incubation continued. For single proteins, cells were cultured overnight at 37ºC; for tandem proteins, cells were cultured overnight at 25ºC. The next day, cells were pelleted in two 50mL sterile conical tubes (Falcon, Cat: 14-432-22) at 5,000xg for 20 minutes. The pellets were either stored at -20°C until further use or immediately processed for protein purification (see below).

*Protein Purification*

Bacterial cell pellets were resuspended in 15 mL of 1X Wash Buffer [5X Wash Buffer: 250 mM $NaH_2PO_4$ (Amresco, Cat: 0571-500g), 1.5M NaCl (Sigma, Cat: S7653-1KG), 100 mM imidazole (Fisher, Cat: 03196-500), pH adjusted to 8.0 with NaOH]. PMSF (VWR, Cat: 10187-508) was added fresh to each resuspended cell pellet at 1 mM final concentration. The protein solutions were probe sonicated at max power for six cycles of 30 seconds on ice. Subsequently, lysates were spun down at 15,000g for 15 minutes at 4$^{o}$C and the cleared supernatant was collected and kept on ice. A Kimble Flex-Column (VWR, Cat: KT420401-2520) was filled with 25 mL HisPur NiNTA Resin (Thermo Scientific, Cat: 88222), then washed and equilibrated with 15 mL 1X Wash Buffer. The protein supernatant was carefully added to the resin (saving ~500 uL for SDS-PAGE or other analyses later) and allowed to bind to the column for 10-15 minutes. The supernatant was drained from the column and the resin was washed four times with 15 mL 1X Wash Buffer. After washing, 5 mL of Elution Buffer (50 mM $NaH_2PO_4$ (Amresco, Cat: 0571-500g), 300 mM NaCl (Sigma, Cat: S7653-1KG), 250 mM imidazole (Fisher, Cat: 03196-500), pH 8.0 with NaOH) were added to the column and incubated for 5 minutes. Elution was repeated 10X and each elution was collected. 10 mL PBS (Sigma, Cat: P4417-100TAB) was added to each elution, and then the mixture was added to Amicon Ultra-15 10kDa filter tubes (EMD Millipore, Cat: UFC901024). The tubes were spun at 4000xg for 20 minutes and the flow through was discarded. The concentrate was then washed with 15 mL PBS another 4 times using the same Amicon filter tube. The final concentrate (~200 uL) was kept at 4 ºC for further use.

Expression and purification was verified by SDS-PAGE and Coomassie staining of the original lysate supernatant along with all washes and elutions. 50 µL β-mercaptoethanol (Sigma, Cat: M3148-100ML) was added to 950 µL 4X Laemmli buffer (Bio-Rad, Cat: 161-0737) in a

microcentrifuge tube. 15 µL of each sample was added to 5 uL of the Laemmli stock in separate 1.5mL microcentrifuge tubes and heated for 5 minutes at 95°C. Mini-Protean TGX Precast Gels (Bio-Rad, Cat: 4569033) with a 15-well comb were used in the BioRad gel electrophoresis cassette (Bio-Rad, Cat: 1658004) with 1L Running Buffer (100 mL 10X Tris/Glycine/SDS Buffer (Bio-Rad, Cat: 1610732) and 900 mL ddH20). Chameleon Precision Plus Protein Dual color standards protein ladder (5 µL; Bio-Rad, Cat:1610374) was used in one lane for each gel, and 15 µL of each sample was added to individual wells. The gels were run at 125V for ~30 minutes until the dye front was close to the bottom. Gels were stained with Coomassie solution [1.2g Coomassie Blue (Thermo Scientific, Cat: 20278) added to 300 mL Methanol (VWR, Cat: BDH1135-4LP) and 60 mL Acetic Acid (Fisher, Cat: A38SI-212)] by microwaving for 30 seconds and then letting the gel stand at room temperature for 10 minutes, and repeating this once. Gels were de-stained by covering the gel with destaining solution (400 mL methanol (VWR, Cat: BDH1135-4LP) added to 100 ml acetic acid (Fisher, Cat: A38SI-212) and 500 mL of ddH20), microwaving for 45 seconds while covered with saran wrap, adding a few crumbled KimWipes (Fisher, Cat: 06-666A) on top of the gel and allowing the mixture to sit for 10 minutes, discarding the solution, and repeating until the gel was no longer blue. The gel was left in ddH20 overnight and scanned on a LI-COR Odyssey Infrared Scanner at 169 µm resolution and 0.5 mm focus offset for 700-channel fluorescence for visualization. We verified that expressed proteins were the expected molecular weight and the predominant band after purification.

*Spectral Scanning Measurements*

Fluorescence spectra were measured on a Shimadzu RF-5301PC spectrofluorometer. Fluorescent protein probes were diluted with ice-cold PBS (Sigma, Cat: P4417-100TAB) such that peak fluorescence emission intensity was ~1000 AU. Each sample was excited at 10 nm increments

from 380 nm to 480 nm. A 2D emission spectral scan was taken of each sample with an emission start 20 nm after the excitation channel center and an emission end at 700 nm. The instrument parameters were set to high sensitivity with an excitation slit width of 5 nm and an emission slit width of 10 nm. Scanning parameters were set to very fast scanning speed, auto response time and 1 nm emission wavelength intervals. Each sample was made in triplicate and individually measured at each excitation wavelength. Data was exported to Excel for further analysis in MATLAB.

**References**


(1) Ai, H.; Shaner, N. C.; Cheng, Z.; Tsien, R. Y.; Campbell, R. E. Exploration of New Chromophore Structures Leads to the Identification of Improved Blue Fluorescent Proteins. *Biochemistry.* **2007**, *46* (20), 5904–5910.

(2) Subach, O. M.; Cranfill, P. J.; Davidson, M. W.; Verkhusha, V. V. An Enhanced Monomeric Blue Fluorescent Protein with the High Chemical Stability of the Chromophore. *PLoS One.* **2011**, *6* (12), e28674.

(3) Zapata-Hommer, O.; Griesbeck, O. Efficiently Folding and Circularly Permuted Variants of the Sapphire Mutant of GFP. *BMC Biotechnol.* **2003**, *3*, 5.

(4) Ai, H.; Hazelwood, K. L.; Davidson, M. W.; Campbell, R. E. Fluorescent Protein FRET Pairs for Ratiometric Imaging of Dual Biosensors. *Nat. Methods.* **2008**, *5* (5), 401–403.

(5) Markwardt, M. L.; Kremers, G.-J.; Kraft, C. A.; Ray, K.; Cranfill, P. J. C.; Wilson, K. A.; Day, R. N.; Wachter, R. M.; Davidson, M. W.; Rizzo, M. A. An Improved Cerulean Fluorescent Protein with Enhanced Brightness and Reduced Reversible Photoswitching. *PLoS One.* **2011**, *6* (3), e17896.

(6) Ai, H.; Henderson, J. N.; Remington, S. J.; Campbell, R. E. Directed Evolution of a Monomeric, Bright and Photostable Version of Clavularia Cyan Fluorescent Protein: Structural Characterization and Applications in Fluorescence Imaging. *Biochem. J.* **2006**, *400* (3), 531–540.

(7) Shcherbakova, D. M.; Hink, M. A.; Joosen, L.; Gadella, T. W. J.; Verkhusha, V. V. An Orange Fluorescent Protein with a Large Stokes Shift for Single-Excitation Multicolor FCCS and FRET Imaging. *J. Am. Chem. Soc.* **2012**, *134* (18), 7913–7923.

(8) Cormack, B. P.; Valdivia, R. H.; Falkow, S. FACS-Optimized Mutants of the Green Fluorescent Protein (GFP). *Gene.* **1996**, *173* (1 Spec No), 33–38.

(9) Xia, N.-S.; Luo, W.-X.; Zhang, J.; Xie, X.-Y.; Yang, H.-J.; Li, S.-W.; Chen, M.; Ng, M.-H. Bioluminescence of Aequorea Macrodactyla, a Common Jellyfish Species in the East China Sea. *Mar. Biotechnol. (NY).* **2002**, *4* (2), 155–162.

(10) Hoi, H.; Howe, E. S.; Ding, Y.; Zhang, W.; Baird, M. A.; Sell, B. R.; Allen, J. R.; Davidson, M. W.; Campbell, R. E. An Engineered Monomeric Zoanthus Sp. Yellow Fluorescent Protein. *Chem. Biol.* **2013**, *20* (10), 1296–1304.

(11) Shaner, N. C.; Lin, M. Z.; McKeown, M. R.; Steinbach, P. A.; Hazelwood, K. L.; Davidson, M. W.; Tsien, R. Y. Improving the Photostability of Bright Monomeric Orange and Red Fluorescent Proteins. *Nat. Methods.* **2008**, *5* (6), 545–551.
(12) Lam, A. J.; St-Pierre, F.; Gong, Y.; Marshall, J. D.; Cranfill, P. J.; Baird, M. A.; McKeown, M. R.; Wiedenmann, J.; Davidson, M. W.; Schnitzer, M. J.; et al. Improving FRET Dynamic Range with Bright Green and Red Fluorescent Proteins. *Nat. Methods.* **2012**, *9* (10), 1005–1012.
(13) Shcherbo, D.; Murphy, C. S.; Ermakova, G. V; Solovieva, E. A.; Chepurnykh, T. V; Shcheglov, A. S.; Verkhusha, V. V; Pletnev, V. Z.; Hazelwood, K. L.; Roche, P. M.; et al. Far-Red Fluorescent Tags for Protein Imaging in Living Tissues. *Biochem. J.* **2009**, *418* (3), 567–574.
(14) Chu, J.; Haynes, R. D.; Corbel, S. Y.; Li, P.; González-González, E.; Burg, J. S.; Ataie, N. J.; Lam, A. J.; Cranfill, P. J.; Baird, M. A.; et al. Non-Invasive Intravital Imaging of Cellular Differentiation with a Bright Red-Excitable Fluorescent Protein. *Nat. Methods.* **2014**, *11* (5), 572–578.
(15) Shcherbakova, D. M.; Verkhusha, V. V. Near-Infrared Fluorescent Proteins for Multicolor in Vivo Imaging. *Nat. Methods.* **2013**, *10* (8), 751–754.
(16) Birtwistle, M. R.; von Kriegsheim, A.; Kida, K.; Schwarz, J. P.; Anderson, K. I.; Kolch, W. Linear Approaches to Intramolecular Förster Resonance Energy Transfer Probe Measurements for Quantitative Modeling. *PLoS One.* **2011**, *6* (11), e27823.
(17) Scheich, C.; Kümmel, D.; Soumailakakis, D.; Heinemann, U.; Büssow, K. Vectors for Co-Expression of an Unrestricted Number of Proteins. *Nucleic Acids Res.* **2007**, *35* (6), e43.

**Supplemental Figure 1, Holzapfel et al.**

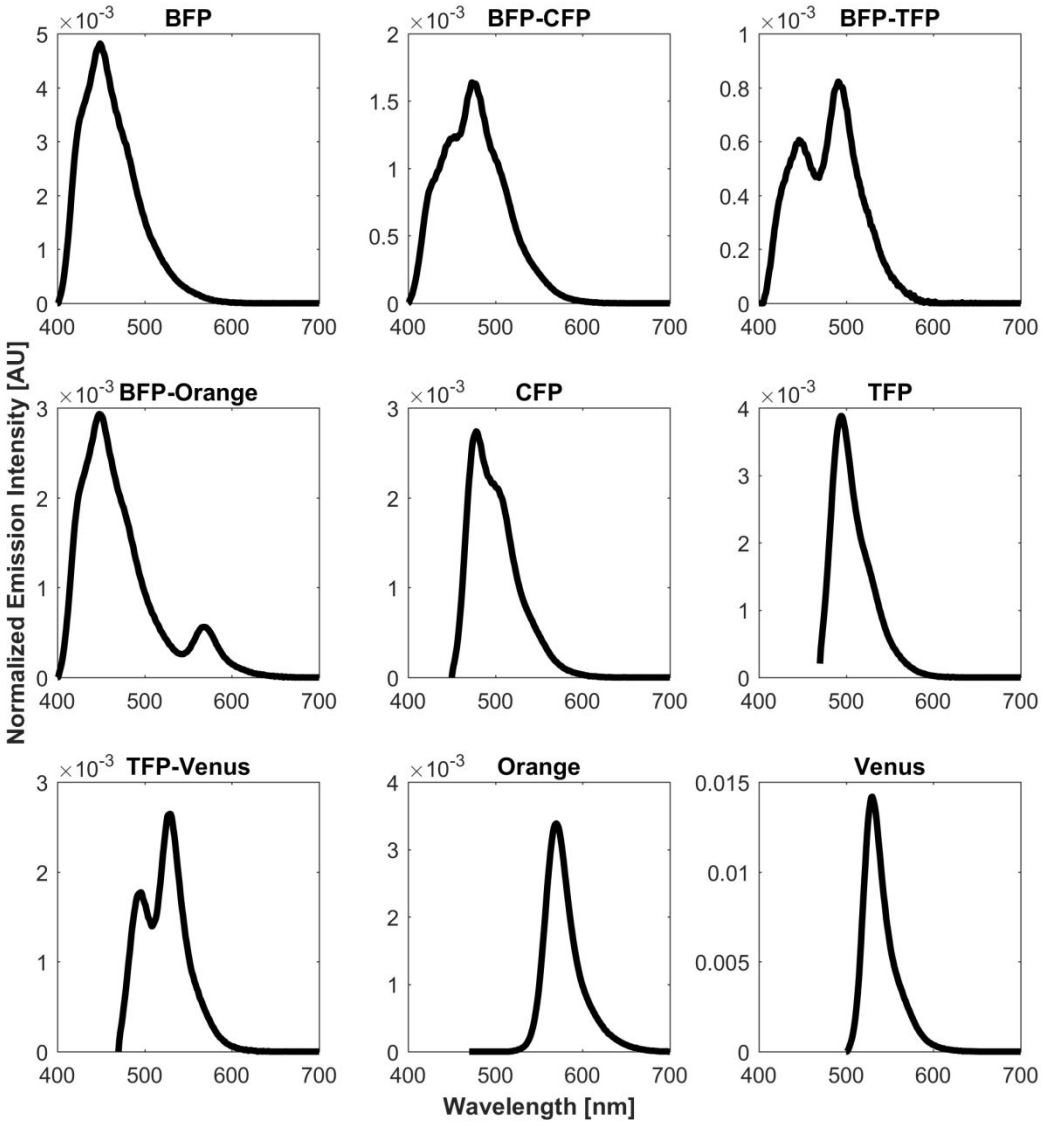


**Figure S1. Normalized Emission Spectra of Individual Probes.** Emission intensity was background subtracted and then normalized to unit integral. A representative spectra is shown from a suitable excitation channel. BFP, BFP-CFP, BFP-TFP, BFP-Orange: 380 nm; CFP: 430 nm; TFP, TFP-Venus, Orange: 450 nm; Venus: 480 nm