Supplemental Information for Protein proximity observed using Fluorogen Activating Protein and Dye Activated by Proximal Anchoring (FAP-DAPA) System

M. Alexandra Carpenter [1,a], Yi Wang [1,b], Cheryl A. Telmer[b,c], Brigitte F. Schmidt[c], Zhipeng Yang[b], Marcel P. Bruchez *[,a,b,c]

Contents

Biochemical Methods: 3
Protein/Dye storage: 3
Bacterial Strains and Vectors: 3
Protein Expression in HEK293 cell culture: 3
Bacterial Protein purification: 3
Yeast cloning: 4
Yeast Protein purification: 4
Isothermal titration calorimetry (ITC): 4
Fluorescence correlation spectroscopy (FCS): 4
FAP protein specific activity: 4
Equilibrium K_D Measurements: 5
HaloTag/HexCl ligand Binding: 5
Rate of association: 5
Rate of dissociation: 6
Quantum Yield of fluorescence: 6
Chemical Inducer of Dimerization Experiments: 6
Cell culture: 7
Cell Permeability experiment: 7
HaloTag Cell surface labeling: 7
Cell contact experiment: 8
Cell contact time course: 8
Supplemental Video 1 8
Supplemental Figures: 9
Supplemental Table 1: Plasmids for bacterial expression 9
Supplemental Table 2: Plasmids for mammalian expression 9
Supplemental Figure 1: Sequences of dL5 and HaloTag** 11
Supplemental Figure 2: Validation of orthogonal dK' primers in a colony PCR test. 11
Supplemental Figure 3: SDS-PAGE gel of FRB, FKBP dK,HT fusions. 12

Supplemental Figure 4: HaloTag binding SDS-PAGE characterization 12
Supplemental Figure 5: Rapamycin induced dimerization receptor distance measurements ... 13
Supplemental Figure 6: Rapamycin Induced Dimerization Fusion Screening 14
Supplemental Figure 7: K_D for MHN ester. 15
Supplemental Figure 8: Rate of dissociation for MG derivatives when bound to dL5 or dK'** 15
Supplemental Figure 9: Rate of association versus concentration for FAP/ MG derivatives 16
Supplemental Figure 10: Quantum yield fluorescence versus concentration plots. 16
Supplemental Table 3: Summary of all mL5 (L91S E52X) properties shown in Figure 2. 17
Supplemental Table 4: Correlation of all mL5 (L91S E52X) properties 18
Supplemental Figure 11: NLS-dL5 cell permeability DIC+dye images with no-dye control** . 18
Supplemental Figure 12: Flow cytometry of dK-TM and dL5TM with MG Derivatives** 19
Supplemental Figure 13: HaloTag-TM-mCerulean3 labeling with Cy3$(SO_3)_2$-HexCl 19
Supplemental Table 5: Table of Image titles and ROI counts for dK-TM with HT-TM-mCer3 20
Supplemental Table 6: Table of Image titles and ROI counts for dL5-TM with HT-TM-mCer3** 20
Supplemental Figure 14: Time-lapse dK-TM and 10nM DAPA Supplemental Video1 stills 21
Supplemental Figure 15: Time-lapse of HT-TM-mCer3 and dK-TM with 10 nM DAPA 22
Supplemental Figure 16: "Cell-drop" experiment with HT-TM-mCer3 and dK-TM 23
Synthesis: 24
General Methods: 24
Fluorogenic Coumarin Alkyne Synthesis: 24
Schematic 1: Fluorogenic Coumarin Alkyne Synthesis 24
Supplemental Figure 17: ^{1}H-NMR of Alkyne-Coumarin-HexCl (8) 28
Supplemental Figure 18: ^{1}H-NMR MG-PEG_{3500}-HexCl (10) fractions 14 and 15 with assignments 30
Supplemental Figure 19: MALDI-TOF of MG-PEG_{3500}-HexCl (10): 31
Supplemental Figure 20: Identification of compound of 555m/z as 12 33
Supplemental Figure 21: Assigned ^{1}H-NMR of MG-HexCl (13) as reference compound 34
Supplemental Figure 22: Assigned ^{13}C-NMR of MG-HexCl (13) as reference compound 35
References: 36

Biochemical Methods:

Protein/Dye storage:

Proteins were stored at 4 °C in PBS or frozen at- 80 °C in 25% glycerol. Dyes were stored as dry aliquots in the dark and dissolved in ethanol at approximately 1 mM. DAPA dye was dissolved at ~100 μM. Concentration was determined using absorbance at 606 nm in acidic ethanol. Dyes were used at maximum 0.1% ethanol in any experiment.

Bacterial Strains and Vectors:

The *Escherichia coli* (*E. coli*) bacterial strain MACH1-T1 (Invitrogen) was used as the host for cloning and Rosetta-gami2(DE3) (Novagen) strain with a pET-21 vector was used for high level, inducible protein expression for purification (Wang et al. 2015). For expression of proteins in HEK293 cells the pcDNA3.1 vector (Invitrogen) was used.

The 3 FAP sequences, dL5** [1] and a variant of dL5**, called dK has a lysine at the E50 (Kabat numbering) in both monomers and a DNA sequence version of dK, called dK' are in Supplemental information. The dK variant was created by site directed mutagenesis of dL5**. The dK' version was synthesized as a gBLOCK (IDT) to maintain the amino acid sequence but create priming sequences so that a PCR screen would be selective for the dK' variant. Sequences of the dL5** variants are shown in the supplemental information.

The PDGFR transmembrane domain was amplified using PCR from pDISPLAY (Invitrogen) and inserted in the XhoI and SalI restriction sites of pBABE plasmid (Telmer et al. 2015). A version of pcDNA3.1 was created with the same MCS as the pBABE plasmid and the BamHI/SalI fragment was inserted.

.

Protein Expression in HEK293 cell culture:

For cell surface expression of the FAPs the pcDNA3.1 plasmids used in Telmer et al. 2015 were modified to include the murine Ig kappa-chain leader sequence and PDGFR transmembrane domain from pDISPLAY (Invitrogen). Table of plasmids used for HEK transformation in this study are in the supplemental information(Table S2).

Bacterial Protein purification:

Table of plasmids used to make protein for this study are listed in the supplemental information(Table S1).

Briefly, 500ml cultures were harvested by centrifugation, cells were lysed by sonication, cleared with a 20,000G centrifugation, bound to Ni-NTA agarose beads (Thermo Fisher) and then cleaved from the beads using HRV 3C protease. The protein released by the proteolytic digestion was collected as flow-through and was then purified on Superdex 75 Gel Filtration Column (GE Healthcare) by fast protein liquid chromatography (BioLogic DuoFlow, Biorad). Fractions were collected based on absorbance readout at 280nm(ε=36,900 $M^{-1}cm^{-1}$) and by dye binding with 1 μM MG-2p (TecanM1000 plate reader). Purity of selected fractions was evaluated by 12% SDS-PAGE Gel and protein was quantified using a DU730 UV/vis spectrophotometer based on the absorbance at 280 nm (Beckman Coulter Inc.).

Yeast cloning:
Site-directed mutagenesis from mL5 was performed using yeast 3-piece homologous recombination in pPNL9 vector. Digestion of plasmids was performed using CspCl restriction enzymes(NEB, Ipswich, MA). PCR was performed with primers with 30 base overlap, and mutagenesis primers flanked with 15 bases of mL5 sequence. Fragments for homologous recombination were transformed into competent yeast strain YVH10 using electroporation as described in the literature (2×10^8 cells/50µL of competent yeast, Bio-Rad Gene Pulser with 2mm cuvette, 0.54kV, 25µF)[2]. Electroporated cells recovered in YPD for 1 hour, were plated on SDCAA+Trp plates, and clones were confirmed by plasmid sequencing.

Yeast Protein purification:
YVH10 yeast cells expressing pPNL9-L5 mutants were grown in stages in SGRCAA+Trp medium at 30°C(5mL then 50mL overnight) then 500 mL at 20°C for 4 days. Supernatant was isolated via centrifugation and filtration, and then concentrated using by 10kDa MWCO filters (centricon Plus-70, EMD Millipore). Protein was dialyzed against PBS at 4°C overnight(0.5 M NaCl, 0.05% Tween 20), isolated on Ni-NTA beads, and eluted with 250 mM imidazole solution (10 mM Tris and 100 mM NaH_2PO_4, pH 8.0; 300 mM NaCl, 250 mM imidazole). Eluted protein was digested by His-tagged TEV protease and 6-His fragments were removed via second Ni-NTA bead incubation. Eluate from Ni-NTA beads was collected, concentrated, and further purified by FPLC, as before, using HiPrep Sephacryl S-100 column.

Isothermal titration calorimetry (ITC):
Microcal VP-ITC) was used to measure thermodynamics of FAP/Fluorogen complexation. A solution of 30 µM MG-2p in syringe was injected into a 4 µM monomeric protein or 2 µM dimeric protein in sample cell (Figure 2, Table S4).

Fluorescence correlation spectroscopy (FCS):
FCS was used to measure the molecular brightness of the mL5 mutant/MG complex on a LSM-510_META_MP_ConfoCor3 Confocal Carl Zeis microscope. No bleaching was applied prior the measurement; the measurement was set as 20s with 10 repeats. For complexation, 1 µM of protein was incubated with 5 nM of MG-2p in PBS+ buffer overnight prior measurement. Equilibrium binding experiments and kinetic experiments (association and dissociation) were performed as described below.

FAP protein specific activity:
Stock MG-2p concentrations were determined via extinction coefficients in acidic ethanol (A_{607}= 91600 $M^{-1}cm^{-1}$) and rough protein concentrations were determined via the absorbance at 280 nm. Protein specific activity was then determined via a saturation experiment utilizing constant MG-2p (100 nM for FKBP/FRB series, 500 or 1 µM for others) and varying concentrations of FAP.

Solutions were incubated for 1-12 hours at room temperature, vortexed, centrifuged, and then read on a TECAN M1000 plate reader (ex/em wavelength=636/664 nm, Bandwidth ex/em=10 nm). Active concentration was determined by solving for concentration of protein required to saturate the dye titrated.

Equilibrium K_D Measurements:

Stock dye concentrations were determined via extinction coefficients in acidic ethanol (A_{607}= 91 600 $M^{-1}cm^{-1}$) and rough protein concentrations were determined via the absorbance at 280 nm (ε_{208}=36 900 $M^{-1}cm^{-1}$). Protein specific activity was then determined via a saturation experiment utilizing constant MG-2p at 500 nM-1 μM and varying concentrations of FAP.

K_D was then tested using a range finding experiment prepared with serial dilutions of 1 μM FAP:10 μM dye. Data is multiplied by dilution factor then normalized. The concentration where the signal drops to half maximal is the approximate K_D and is an appropriate concentration at which to run subsequent K_D measurements.

K_D values are then determined using constant FAP and ranges of dyes prepared via a split dilution series. All dye protein complexes are incubated for 24 hours prior to reading in a Tecan M1000Pro plate reader measured at ex/em wavelength=636/664 nm (Bandwidth ex/em=10 nm, flashes=50, frequency=400 Hz, integration time=20 μs, settle time=0ms, bottom read). K_D data was fitted using the "One site total-accounting for ligand depletion" non-linear fitting parameter in GraphPad PRISM version 6.01 for Windows.

HaloTag/HexCl ligand Binding:

HaloTag proteins purified from RGII were quantified using Pierce BCA Protein Assay according to manufacturer protocol (1:8 dilution) and read in a TECAN M1000 plate reader. The proteins HT-FKBP or FRB-HT were incubated with HexCl ligands Cy3-HexCl or MG-PEG_{3500}-HexCl for 30 minutes at room temperature in PBS pH 7.4. Analysis of conjugation efficiency was performed with 12% SDS-PAGE gel separation. Samples containing 0.9 μg/well protein were prepared with SDS loading buffer (2%SDS, 1%BME, .1% bromophenol blue, 10% glycerol) and boiled for 5 minutes. SDS-PAGE gels were analyzed using ChemiDoc system fluorescence and colorimetric for dyes and coomassie staining respectively (Figure S1)

Rate of association:

Rate of association of FAP/dye complexes was measured on the TECAN M1000 plate reader. dL5** or dK protein stocks were added to a clear bottom polystyrene 96 well dish at 50 nM, 20 μL immediately prior to run. Dye at concentrations ranging from 11.5 nM to 200 nM was injected via injector carrier, well-wise at 200 μL/s. Fluorescence for each sample was read at λex/em= 636/664 nm in triplicate at suitable sampling interval until plateau was achieved. Resulting data was baseline subtracted with average of triplicate controls containing no protein but measured with same parameters including sampling interval and time. Data was fit using GraphPad PRISM 6.01 for windows "one-phase association" algorithm to give relative rates of association. Plots of

relative rates versus concentration of dye were fit with a linear regression in PRISM to give slopes equal to the concentration independent rate of association.

Rate of dissociation:
Quantitated dyes and FAPs were incubated overnight at room temperature in polypropylene tubes at 1 μM dye/dK or 100 nM dye/dL5** in PBS+0.1% PF127. These were then monitored for fluorescence upon 200-fold dilution with 2 μM MHN-ester PBS+0.1%PF127 solution to compete off the dye. The samples were run in quadruplicate on a TECAN M1000 plate reader (λex/em=636/664 nm, Gain=250, Bandwidth ex/em=10 nm, flashes=50, frequency=400 Hz, integration time=20 μs, settle time=0 ms, bottom read). The sampling intervals were: dK(interval, time: 15 s, 10 min then 5 min, 60 min) dL5(interval, time: 30 s, 5 min then 5 min, 300 min). Fitting was performed using GraphPad PRISM 6.01 for windows, non-linear fitting algorithm, "one-phase decay", to find the rate of dissociation , k=ln2/$t_{(1/2)}$,. The values are shown in Table 1 and the plots for selected data are show in Figure S8.

Quantum Yield of fluorescence:
The quantum yield of fluorescence was measured on the PTI fluorescence reader. Stock solutions of dye at 2 μM with 20 μM FAP (or buffer control) were incubated for 18 hours. The absorbances of these solutions were then matched and dilutions made to 1 μM, 750 nM, and 500 nM (with 10x protein or blank controls). Fluorescence was measured with excitation at 636 nm, emission at 652-780 nm and the integrated fluorescence was calculated for the average of 4 plots and blank subtracted. Plots of integrated fluorescence versus concentration for the standard (MG-2p/dL5, QY=0.20) and all unknowns were plotted and fit with a linear regression. The ratio of the slope in reference to the standard was used to calculate the unknown quantum yields according to the following equation:

$$\Phi_{unk} = \Phi_{std}\left(\frac{Slope_{unk}}{Slope_{std}}\right)$$

Chemical Inducer of Dimerization Experiments:
FKBP, FRB fusions of dK and HT were cloned and purified according to standard bacterial method (Supplemental Information) and quantified using protein specific activity(dK fusions) or BCA assays (HT fusions). Samples containing 100 nM dK fusion, 200 nM HT fusion, 250 nM rapamycin, and 10 nM DAPA were prepared in triplicate. DAPA was either pre-incubated with HT fusions or added at the same time as HT/dK and rapamycin solutions. Samples incubated for 12 hours at room temperature prior to measurement. Solutions were then vortexed, spun down, plated and read in a TECAN M1000 plate reader at λex/em=636/664 nm (BW ex/em=10 nM, gain=250 nm). The readings were baseline subtracted and then normalized to mean dK signal. Significance was determined via two-sided un-paired t-test and data was plotted in GraphPad PRISM. (Figure 3, S6). Crystal structure of FRB-HT and FKBP-dK interacting with rapamycin

and the cartoon of this structure were created using PyMol software (PDB ID: 4FAP (FKBP/rapa/FRB, 5UXZ (HaloTag), and 4K3G (FAP)).

Cell culture:

For cell surface expression of the FAPs the pcDNA3.1 plasmids used in Telmer et al. 2015 were modified to include the murine Ig kappa-chain leader sequence and PDGFR transmembrane domain from pDISPLAY (Invitrogen). Table of plasmids used for HEK transformation in this study are in the supplemental information(Table S2). For expression of proteins in HEK293 cells the pcDNA3.1 vector (Invitrogen) was used along with lipofectamine 2000 transfection reagent to generate stable cell lines from single colonies. Cells were cultured in DMEM (Thermo Fisher) supplemented with 10% fetal bovine serum (FisherBrand). Stable cell lines were passaged at maximum 90% confluency via gentle trituration and were maintained in selection pressure (G418) below passage number 20.

Cell Permeability experiment:

HEK 293 cells stably expressing dL5**-NLS (Addgene: 73205) or dL5**-TM (Addgene: 73206) were plated in a 8 well Ibidi dish and labeled with 250 uL 500 nM dye(DAPA, MG-nBu, or no dye control) in optiMEM, incubated for 15 minutes, and then imaged using a 40x oil objective on a Zeiss 880LSM. FAP was imaged using a 633 nm laser with powers of 0.4% and 6% for the surface and the nuclear FAP cells, respectively. Images were processed using FIJI image processing: lookup tables were matched for 633 nm channels and transmission channels were filtered with a FFT bandpass filter to filter large structures (shading, down to100 μM). **Flow cytometry:** Analysis of dL5**-TM and NLS-dL5** expressing HEK293 populations of cells was performed via flow cytometry. Sets of expressing cells or controls were suspended via trituration, concentrated via centrifugation at 1000 rpm, and labeled with 300 uL media containing 100 nM dye. Samples were run on the Accuri/Intellicyt High throughput Flow cytometer. Populations analyzed using FLoJo software. The cells were gated for live cells based on location in SSC/FSC(4300 cells (Range 2930-5898)), and then analyzed for fluorescence in the FL4 channel (640nm). Fraction of expressing cells was determined via analysis with MG-nBu control versus the no-dye negative control for each population (96% dL5**-TM, 55% NLS-dL5**). Median fluorescence of the expressing populations(75^{th} percentile for NLS) in the FL4 channel were obtained, normalized to MG-nBu fluorescence, and plotted using GraphPad PRISM. WT cell fluorescence data was normalized by scaling against the MG-nBu median or 75^{th} percentile sample for each population that it is compared to.

HaloTag Cell surface labeling:

A mixture of HaloTag-TM-mCerulean3 (Addgene: 145767) and WT HEK293 cells were labeled with 500 nM DAPA or MG-Btau in DMEM+10%FBS for 5 minutes. The dye was removed, and dish washed 6 times with DMEM+10%FBS over 2 hours. The media was then exchanged with imaging media (OptiMem) containing 1 μM dL5** to activate bound DAPA, MG moiety. Cells were imaged using a Zeiss 880 LSM at 633 and 405 nm with laser powers of 55% and 25.7%

respectively. Images were processed with FIJI image processing software: all lookup tables were matched and the DIC image was filtered using a FFT bandpass filter to remove large objects (shadows, above 100 μm).

Cell contact experiment:

HT-TM-mCerulean and dK-TM or dL5**-TM cells were plated on a poly-L-lysine coated dish. The cells were labeled with optiMEM containing 0.5-100nM DAPA and imaged within 30 minutes without rinsing cells. All imaging was performed at 37°C with 5% CO_2 supply on a Zeis880 LSM with 405 nm diode laser (25.7% laser power, emission 408 - 486 nm) and varying laser powers for the 633 nm HeNe laser channel (emission 641 - 695 nm, dK 0.5 nM = 55%, dK 1 nM = 25%, dK 10 nM= 25%, dK 100 nM and dL5** 0.5-100 nM = 0.4%). Images were created using FIJI where DIC images were filtered with a 1 nM bandpass filter to lower periodic noise and remove shadows (below1 μm, above 100 μm). **Quantitation**: In FIJI, circular ROIs of area 6.37 pixels were drawn and used to measure brightness in 633 nm channel for contact sites and non-contact sites. Contact sites were measured more densely 1-3 ROIs per contact site. The non-contact sites were measured at 1/4- density of contact sites, randomly along the edge of every FAP expressing cell (determined from FAP channel/DIC channel overlay). Data was normalized by the mean of the fluorescence in the contact zone and compared for the dL5** and dK populations. The ratio of signals was also calculated. Ratio=(mean ROI signal in contact zone)/(mean ROI signal in non-contact zone for that concentration).

Cell contact time course:

HT-TM-mCerulean and dK-TM cells were plated on a poly-L-lysine coated dish. The cells were labeled with optiMEM containing 10 nM DAPA and imaged within 15 minutes without rinsing cells. All imaging was performed at 37°C with 5% CO_2 supply on a Zeis880 LSM with 405 nm diode laser (25.7% laser power, emission 409 - 486 nm) and 633 nm HeNe laser (6.0% laser power, 641 - 695 nm). Images were taken at intervals of 150 seconds. Movies were created using FIJI and show 6 frames per second. 1 field of view with three zoomed regions is available as SupplementalVideo 1 and stills located in Supplemental Figure 14. An additional set of images from this experiment taken prior to or during the panels for the video are included in Supplemental Figure 15.

Supplemental Video 1

SupplementalVideo 1 of the cell contact time-course with 10nM DAPA was created using FIJI and Windows Video Editor and is available online.

Supplemental Figures:

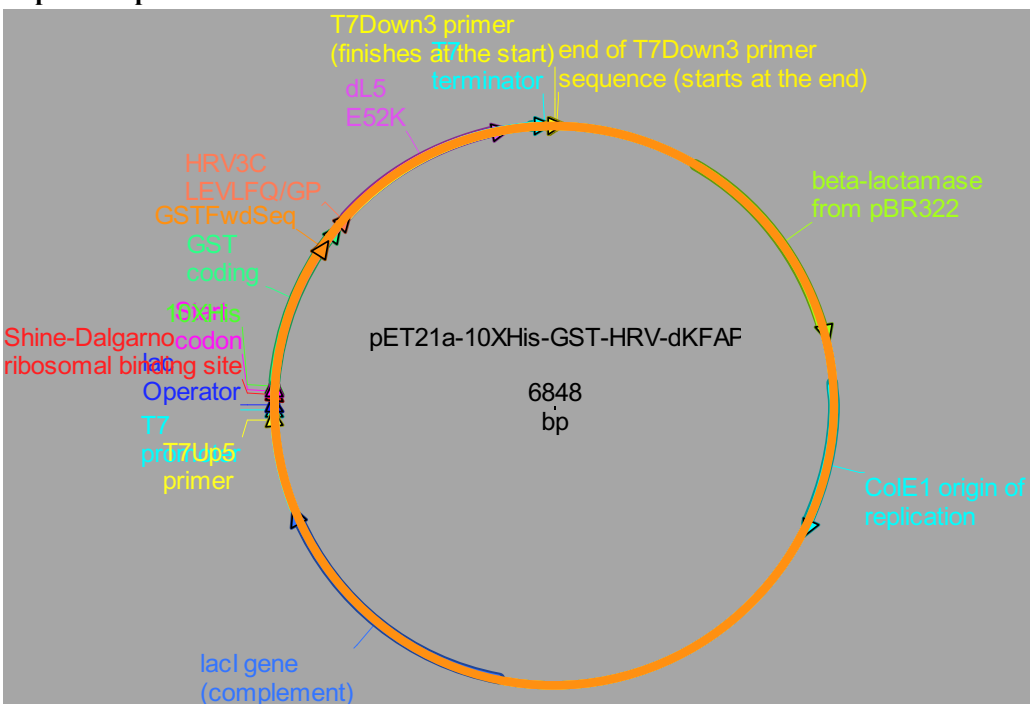
Vector	Fusion1	Fusion 2	Fusion 3	Addgene
pET21	GST	dL5**		73214
pET21	GST	dK’		145765
pET21	GST	dK	FKBP	
pET21	GST	FKBP	dK	
pET21	GST	dK	FRB	
pET21	GST	FRB	dK	
pET21	GST	HaloTag	FKBP	
pET21	GST	FKBP	HaloTag	
pET21	GST	HaloTag	FRB	
pET21	GST	FRB	HaloTag	

Supplemental Table 1: Plasmids for bacterial expression

Vector	Targeting sequence	Fusion 1	(Fusion 2)	(Fusion 3)	Addgene
pcDNA3.1	kappa	myc-dL5	TM		73206
pcDNA3.1	kappa	myc-dK	TM		
pcDNA3.1	kappa	myc-dK’	TM		145766
pcDNA3.1	kappa	myc-HaloTag	TM	mCer3	145767
pcDNA3.1	NLS	Myc-dL5	mCer3		73205

Supplemental Table 2: Plasmids for mammalian expression

A.pET21 plasmid with dL5-E52K:



B. dL5:**

CAGGCCGTCGTTACCCAAGAACCTAGTGTTACCGTTAGCCCAGGTGGTACTGTTATACTTACTTGTGGAAGT
GGTACGGGTGCCGTCACATCTGGTCATTATGCAAATTGGTTTCAACAAAAACCAGGACAAGCTCCAAGAGCT
TTGATTTTTGATACTGATAAGAAGTATTCTTGGACCCCAGGTAGATTTTCTGGATCTTTGCTGGGAGCAAAGG
CAGCTTTGACAATATCAGATGCTCAGCCTGAGGACGAAGCCGAGTATTACTGTTCTCTTAGCGACGTGGATG
GCTACTTGTTTGGCGGTGGAACACAACTGACGGTTCTGTCCGGTGGTGGCGGCTCTGGTGGCGGTGGCAGCG
GCGGTGGTGGTTCCGGAGGCGGCGGTTCTCAGGCTGTGGTGACTCAGGAGCCGTCAGTGACTGTGTCCCCAG
GAGGGACAGTCATTCTCACTTGTGGCTCCGGCACTGGAGCTGTCACCAGTGGTCATTATGCCAACTGGTTCC
AGCAGAAGCCTGGCCAAGCCCCCAGGGCACTTATATTTGACACCGACAAGAAGTATTCCTGGACCCCTGGCC
GATTCTCAGGCTCCCTCCTTGGGGCCAAGGCTGCCCTGACCATCTCGGATGCGCAGCCTGAAGATGANGCTG
AGTATTACTGTTCGCTCTCCGACGTTGACGGTTATCTGTTCGGAGGAGGCACCCAGCTGACCGTCCTCTCC

C. dL5(E52K), dK:

CAGGCCGTCGTTACCCAAGAACCTAGTGTTACCGTTAGCCCAGGTGGTACTGTTATACTTACTTGTGGAAGT
GGTACGGGTGCCGTCACATCTGGTCATTATGCAAATTGGTTTCAACAAAAACCAGGACAAGCTCCAAGAGCT
TTGATTTTTAAAACTGATAAGAAGTATTCTTGGACCCCAGGTAGATTTTCTGGATCTTTGCTGGGAGCAAAG
GCAGCTTTGACAATATCAGATGCTCAGCCTGAGGACGAAGCCGAGTATTACTGTTCTCTTAGCGACGTGGAT
GGCTACTTGTTTGGCGGTGGAACACAACTGACGGTTCTGTCCGGTGGTGGCGGCTCTGGTGGCGGTGGCAGC
GGCGGTGGTGGTTCCGGAGGCGGCGGTTCTCAGGCTGTGGTGACTCAGGAGCCGTCAGTGACTGTGTCCCCA
GGAGGGACAGTCATTCTCACTTGTGGCTCCGGCACTGGAGCTGTCACCAGTGGTCATTATGCCAACTGGTTC
CAGCAGAAGCCTGGCCAAGCCCCCAGGGCACTTATATTTAAAACCGACAAGAAGTATTCCTGGACCCCTGGC
CGATTCTCAGGCTCCCTCCTTGGGGCCAAGGCTGCCCTGACCATCTCGGATGCGCAGCCTGAAGATGANGCT
GAGTATTACTGTTCGCTCTCCGACGTTGACGGTTATCTGTTCGGAGGAGGCACCCAGCTGACCGTCCTCTCC

D. dL5(E52K) codon optimized, dK':

CAaGCaGTaGTaACaCAgGAgCCatcaGTTACCGTTAGCCCAGGTGGTACTGTTATACTTACTTGTGGAAGTGGTA
CGGGTGCCGTCACATCTGGTCATTATGCAAATTGGTTTCAACAgAAACCAGGACAAGCTCCAAGAGCTTTGA
TcTTTaaaACTGATAAGAAGTATTCTTGGACCCCAGGTAGATTcTCTGGATCTTTGCTGGGAGCAAAGGCAGCT
TTGACAATATCAGATGCTCAGCCTGAGGACGAAGCCGAGTATTACTGTTCTCTTAGCGACGTGGATGGCTAC
TTGTTTGGtGGaGGAACACAACTGACGGTTCTGTCCGGaGGTGGCGGCTCTGGTGGCGGTGGCAGCGGCGGTG
GTGGTTCCGGAGGCGGCGGTTCTCAGGCTGTGGTGACTCAGGAGCCGagcGTGACTGTGTCCCCAGGAGGGA
CAGTCATTCTCACTTGTGGCTCCGGCACTGGAGCTGTCACCAGTGGTCATTATGCCAACTGGTTCCAGCAGA
AGCCTGGCCAAGCaCCCAGGGCACTTATATTTaagACCGACAAGAAGTATTCCTGGACtCCTGGCCGATTCTCA

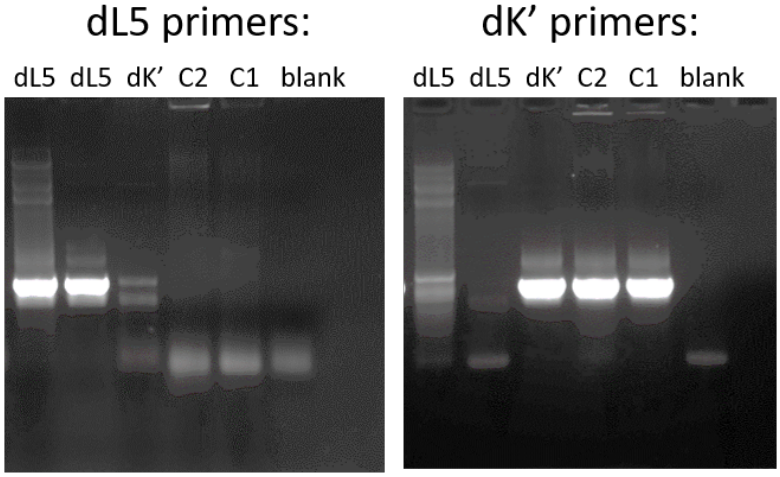
GGCTCCCTCCTTGGaGCCAAGGCTGCCCTGACCATCTCGGATGCGCAGCCTGAAGATGAGGCTGAGTATTAC
TGTTCGCTCTCCGACGTTGACGGTTATCTGTTCGGcGGtGGtACaCAatTaACtGTatTgagC

E. HaloTag7

ATGGCAGAAATCGGTACTGGCTTTCCATTCGACCCCCATTATGTGGAAGTCCTGGGCGAGCGCATGCACTACGTCGATGTTGGTCC
GCGCGATGGCACCCCTGTGCTGTTCCTGCACGGTAACCCGACCTCCTCCTACGTGTGGCGCAACATCATCCCGCATGTTGCACCGA
CCCATCGCTGCATTGCTCCAGACCTGATCGGTATGGGCAAATCCGACAAACCAGACCTGGGTTATTTCTTCGACGACCACGTCCGC
TTCATGGATGCCTTCATCGAAGCCCTGGGTCTGGAAGAGGTCGTCCTGGTCATTCACGACTGGGGCTCCGCTCTGGGTTTCCACTG
GGCCAAGCGCAATCCAGAGCGCGTCAAAGGTATTGCATTTATGGAGTTCATCCGCCCTATCCCGACCTGGGACGAATGGCCAGAA
TTTGCCCGCGAGACCTTCCAGGCCTTCCGCACCACCGACGTCGGCCGCAAGCTGATCATCGATCAGAACGTTTTTATCGAGGGTAC
GCTGCCGATGGGTGTCGTCCGCCCGCTGACTGAAGTCGAGATGGACCATTACCGCGAGCCGTTCCTGAATCCTGTTGACCGCGAG
CCACTGTGGCGCTTCCCAAACGAGCTGCCAATCGCCGGTGAGCCAGCGAACATCGTCGCGCTGGTCGAAGAATACATGGACTGGC
TGCACCAGTCCCCTGTCCCGAAGCTGCTGTTCTGGGGCACCCCAGGCGTTCTGATCCCACCGGCCGAAGCCGCTCGCCTGGCCAA
AAGCCTGCCTAACTGCAAGGCTGTGGACATCGGCCCGGGTCTGAATCTGCTGCAAGAAGACAACCCGGACCTGATCGGCAGCGA
GATCGCGCGCTGGCTGTCGACGCTCGAGATTTCCGGC

Supplemental Figure 1: Sequences of dL5 and HaloTag**

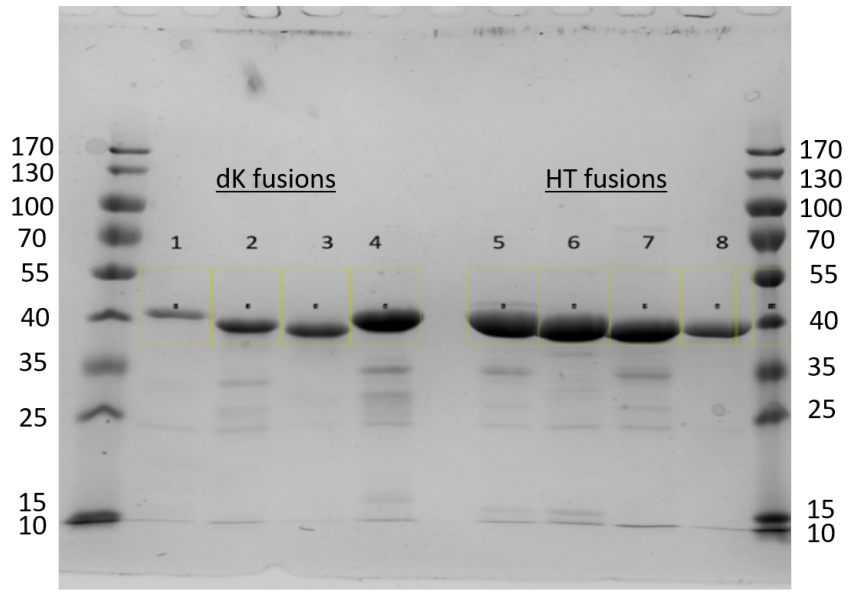
A. Map of standard pET21 protein purification vector. dK' is inserted and labeled as dL5E52K. **B.** Sequence of dL5**, **C**. Sequence of dL5E52K with the change highlighted in red, **D**. Sequence of codon optimized dK'. The lowercase letters are location of sequence changes. Additionally, the green underlined region was altered to create orthogonal dK' and dL5** primers. **E.** sequence of HaloTag7 (Promega).



Supplemental Figure 2: Validation of orthogonal dK' primers in a colony PCR test.

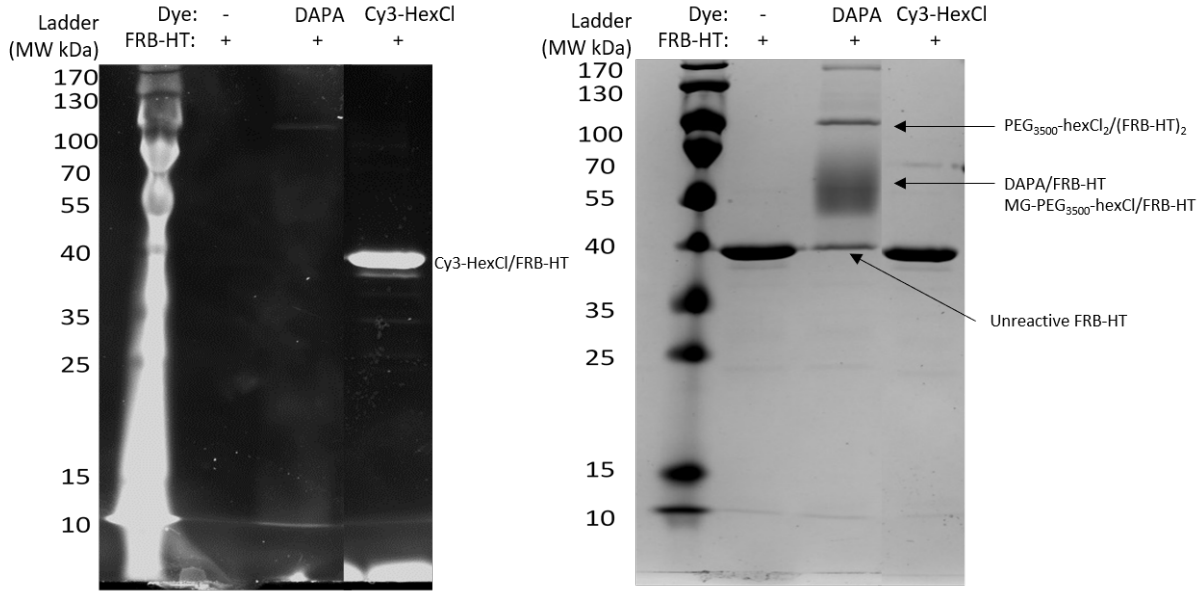
PCR using dL5** primers and dK' primers were done on dL5** at two concentrations and dK' positive plasmids. The method was also tested on colony PCR of MachT1 colonies likely containing dK' plasmids. All positive controls show bright bands and the negative controls show weak bands, multiple bands, and/ or bands of incorrect size. (colony PCR: where a colony is suspended in 10uL water and 1uL added to the PCR mixture directly instead of using isolated DNA).

Validation of the orthogonality of the dK' primers was demonstrated when applied to dL5** or dK' and colonies. The two primers were tested for replication in PCR with Taq polymerase. dL5** positive controls for show high signal at two concentrations. dK' positive control and the colonies C1,C2 do not show bright bands in the dL5** primer test. This is reversed for ability to copy plasmids using dK' primers; the dL5** positive controls show only faint smears while the dK' positive control and colonies show bright bands at the same molecular weight as the dL5** bands. These were quantitated to be of about 750 base pairs in length in subsequent tests. Therefore, we can conclude that the dL5** primers are specific to dL5** systems and dK' primers are specific to dK' containing systems.

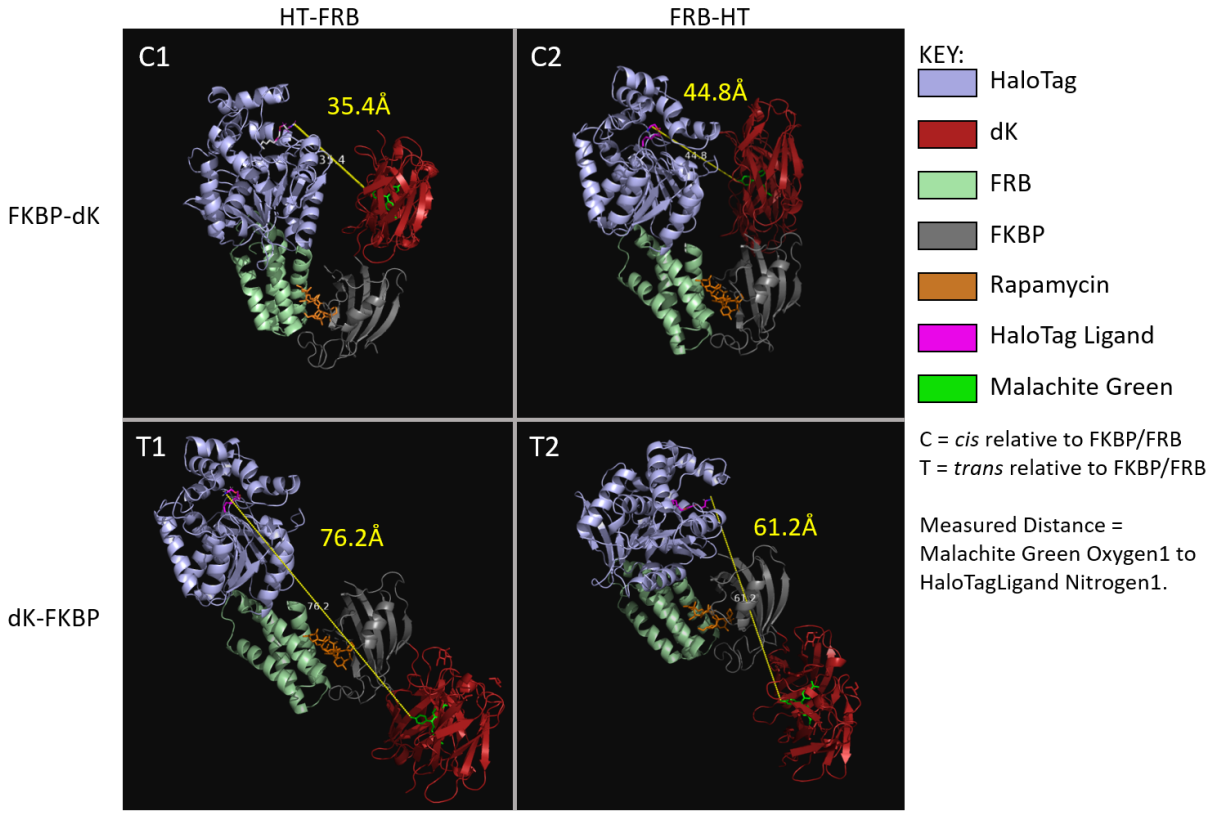


Label	protein
1	dK-FKBP
2	FKBP-dK
3	dK-FRB
4	FRB-dK
5	HT-FKBP
6	FKBP-HT
7	HT-FRB
8	FRB-HT

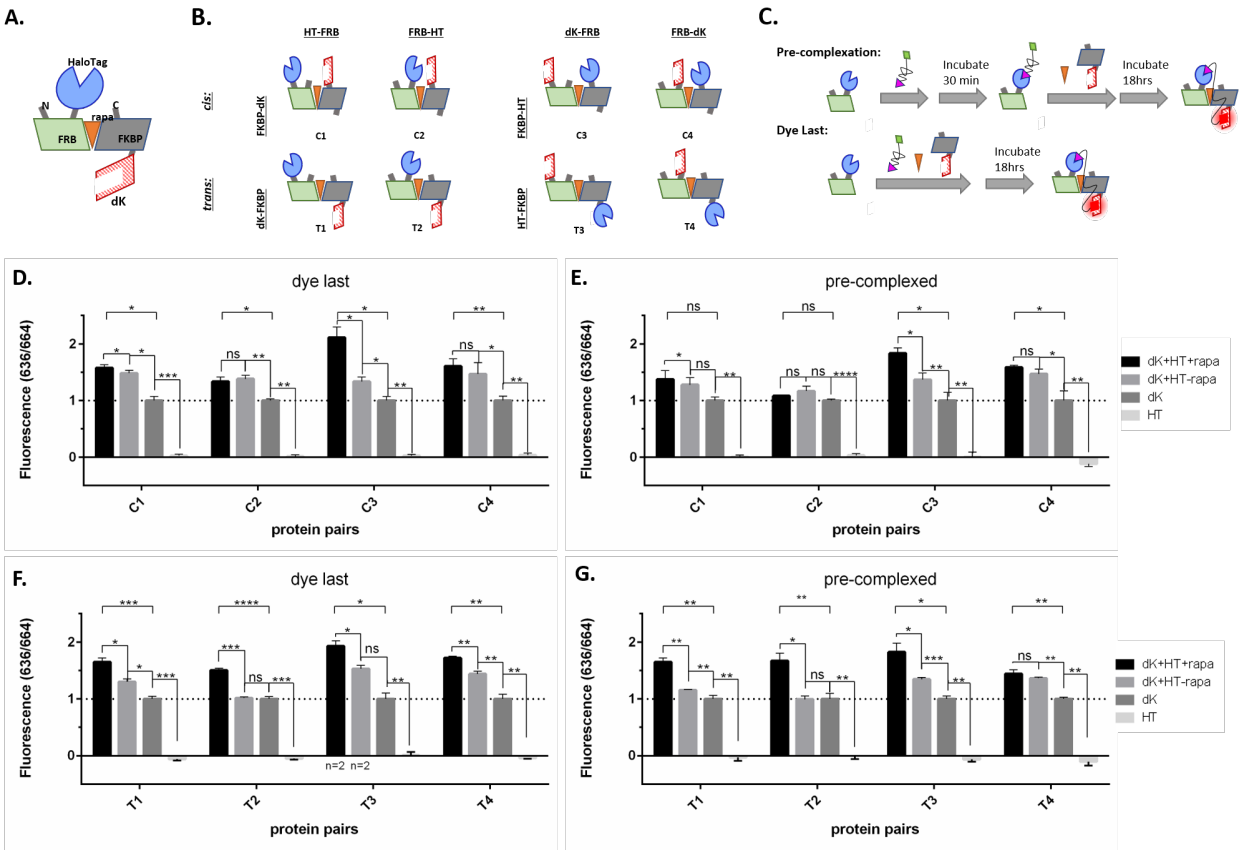
Supplemental Figure 3: SDS-PAGE gel of FRB, FKBP dK,HT fusions.
To confirm purity of the purified proteins, SDS-PAGE gels were run. The proteins (0.9 µg/well) were then run on a 12% agarose SDS-PAGE gel with 0.4% BME. Samples were boiled for 5 minutes prior to loading to denature protein.



Supplemental Figure 4: HaloTag binding SDS-PAGE characterization
FRB-HT in PBS at 1 µM was incubated without dye, with 5 µM DAPA(MG-80p), or with 5 µM Cy3-HexCl dye for 30 minutes at room temperature. The proteins (0.9 µg/well) were then run on a 12% agarose SDS-PAGE gel with 0.4% BME with BioRad Pre-stained Ladders. Fluorescence (Left) shows bright Cy3-HexCl/FRB-HT labeling and expected lack of fluorescence in other lanes. Coomassie staining (Right) confirms no molecular weight shift for Cy3-HexCl conjugate with respect to the control. Conversely, a large molecular weight shift is seen for MG-PEG_{3500} -HexCl conjugated to FRB-HT (55 kDa). Small amounts of unreactive FRB-HT are seen in this lane (45 kDa).

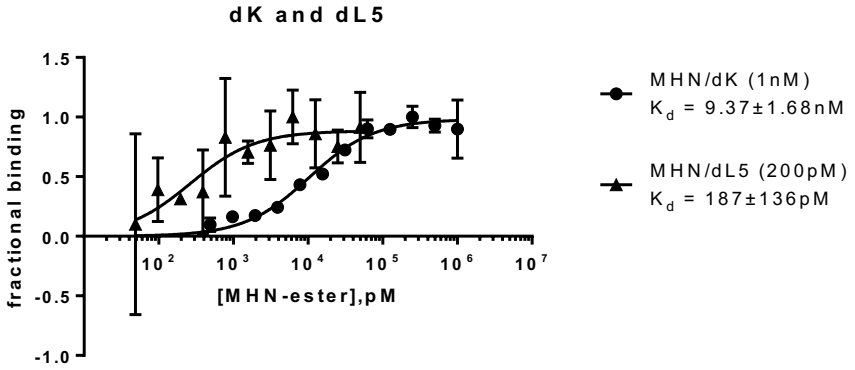


Supplemental Figure 5: Rapamycin induced dimerization receptor distance measurements
Prepared protein orientations were estimated from solved crystal structures(PDBID: 4FAP, 5UXZ, and 4K3G). Fusions illustrated using The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. The orientation of the fusions *cis* or *trans* relative to the plane of the FKBP/rapamycin/FRB trimer dictates inter-receptor distances; *cis*-pairs have loosely half the inter-receptor distance compared to *trans*-pairs. It can also be noted here that the C2 orientation allows very little space for the tail of ligands to either FAP or HaloTag receptor. This lack of space for DAPA reagent to bind could explain reductions in proximity induced FAP activation for C2 and C4 analogue under any labeling conditions (Supplemental Figure 6).

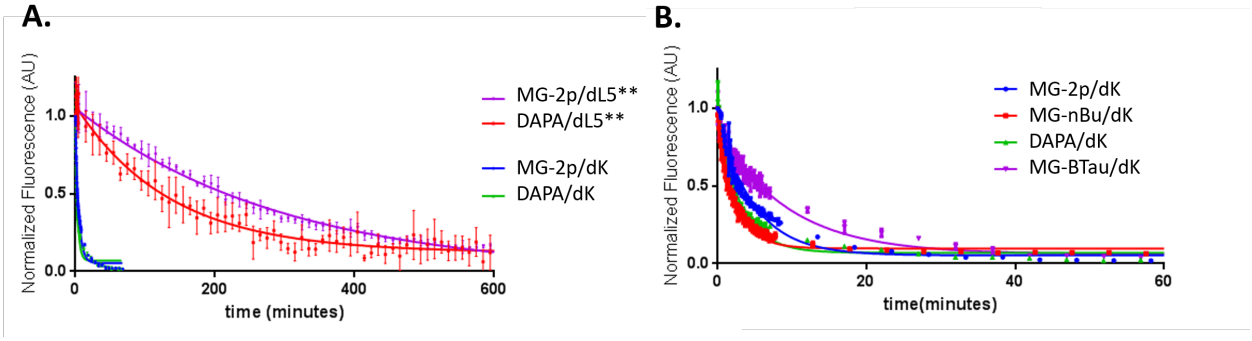


Supplemental Figure 6: Rapamycin Induced Dimerization Fusion Screening

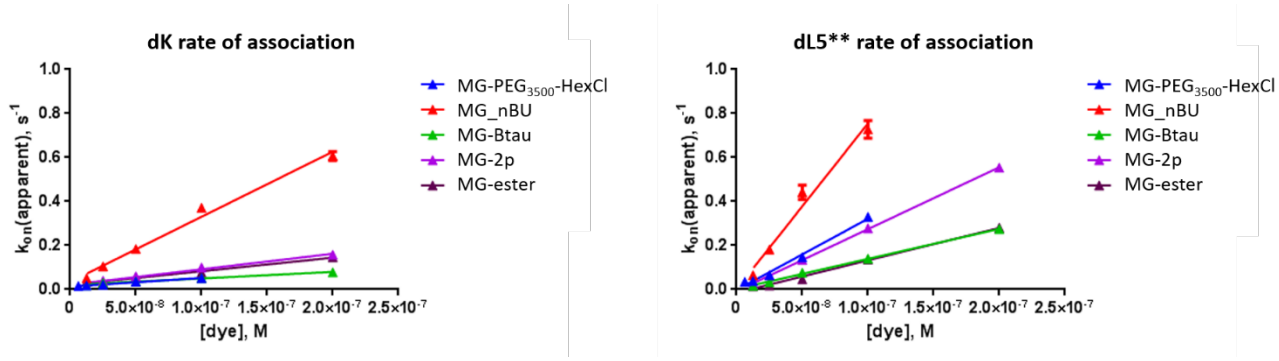
A. Cartoon representation of FRB/rapamycin/FKBP fusions showing locations of C and N-termini. **B.** Cartoon of all fusions of FRB/FKBP fusions with dK and HaloTag (HT). All orientations were sub-divided into *cis*- or *trans*- fusions based on location of HT and FAP fusions relative to the FRB/rapa/FKBP trimer. **C.** Schematic of order of reagent addition for the two experimental conditions. Pre-complexation is expected to provide superior results but, adding dye last is more representative of labeling conditions required in tissue culture/ *in vivo*. **D-G.** Fluorescence of rapamycin induced dimerization of FKBP, FRB fusions of dK and HT were measured at 100 nM dK fusion, 200 nM HT fusion, 250 nM rapamycin, and 10 nM DAPA. Addition of rapamycin to solutions containing FRB/FKBP fusions show significant increase in FAP-DAPA signal for most pairs when compared to both proteins in absence of rapamycin or dK fusions alone (unpaired, two sided, t-test). ($p \leq 0.0001$=****, $0.0001 < p \leq 0.001$=***, $0.0001 < p \leq 0.01$=**, $p \leq 0.05$=*).



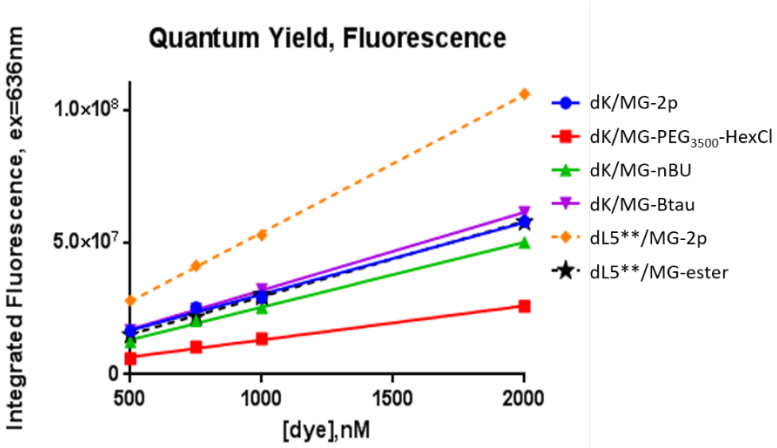
Supplemental Figure 7: K_D for MHN ester.
K_D values for FAP/ dye complexes was determined using dL5** at 200 pM and dK' at 1 nM (based on range finding experiments that suggested K_Ds at 200 pM and 10 nM respectively). Both were prepared in quadruplicate and incubated for 24 hours at room temperature in polypropylene tubes prior to vortexing and transferring to a plate reader. NOTE: dL5** equilibrium binding determined K_Ds are thought to be a factor of 10 over K_Ds determined by ratios of k_{off}/k_{on} due to ligand depletion. MHN ester was used as a competitor in some rate of dissociation experiments since it is 'dark' at λex/em=630/664 nm.



Supplemental Figure 8: Rate of dissociation for MG derivatives when bound to dL5 or dK'**
Quantitated dyes and FAPs were incubated overnight at room temperature in polypropylene tubes at 1 µM dye/dK (**A., B.**) or 100nM dye/dL5** (**A.**) in PBS+0.1% PF127. These were then monitored for fluorescence upon 200 fold dilution with 2 µM MHN-ester PBS+0.1%PF127 solution to compete off the dye. The samples were run in quadruplicate on a TECAN M1000 plate reader (λex/em=636/664 nm, Gain=250, Bandwidth ex/em=10 nm, flashes=50, frequency=400 Hz, integration time=20 µs, settle time=0 ms, bottom read). The sampling intervals were: dK(interval, time: 15s,10min then 5min,60min) dL5(interval, time: 30s, 5min then 5min,300min). Fitting was performed using GraphPad PRISM 6.01 for windows, non-linear fitting algorithm, "one-phase decay", to find the rate of dissociation , k=ln2/$t_{(1/2)}$,. The values are shown in Table 1.



Supplemental Figure 9: Rate of association versus concentration for FAP/ MG derivatives
Rate of association of FAP/dye complexes was measured on the TECAN M1000 plate reader. dL5** or dK protein stocks were added to a clear bottom polystyrene 96 well dish at 50 nM, 20 uL immediately prior to run. Dye at suitable concentrations to saturate the protein in greater than 20 s (11.5 nM to 200 nM) was injected via injector carrier, well-wise at 200 uL/s. Fluorescence for each sample was read at λex/em= 636/664 nm in triplicate at suitable sampling interval until plateau was achieved. (interval= 200 ms – 1s, flashes=10, flash frequency=100Hz, Integration time=20 µs, gain=150, BWex/em=10 nm/10 nm, bottom read). Resulting data was baseline subtracted with average of triplicate controls containing no protein but measured with same parameters including sampling interval and time. Data was fit using GraphPad PRISM 6.01 for windows "one-phase association" algorithm to give relative rates of association. Plots of relative rates versus concentration of dye were fit with a linear regression in PRISM to give slopes equal to the concentration independent rate of association. Errors are reported as error in linear fit of data (Table 1).



Supplemental Figure 10: Quantum yield fluorescence versus concentration plots.
The quantum yield of fluorescence was measured on the PTI fluorescence reader. Stock solutions of dye with 10 fold excess FAP were incubated for 18 hours and measured after dilution to various concentrations. A standard curve of fluorescence versus concentration for unknowns and a standard were plotted. The ratio of the slope in reference to the standard (MG-2p/dL5** pair with known quantum yield of fluorescence of 0.2) was used to calculate the unknown quantum yield.

	FCS	Equilibrium Measurements		Kinetic Measurements					ITC		
Residue (E52**X**)	Brightness CPM(kHz)	K_d (M)	Bmax (AU)	on-rate (s-1M-1)	t1/2 (s)	off-rate (s)	t1/2 off (s)	K_d (M)	K_d (M)	ΔH (kCal/mol)	side chain weight (g/mol)
A	10.1	3.27E-08	5.75E+04	8.00E+02	1.00E-03	1.23E-04	5.63E+03	1.54E-07	4.29E-09	-18.5	15.03
D	10.0	7.96E-09	6.13E+04	1.49E+04	4.65E-05	1.60E-04	4.34E+03	1.07E-08	9.84E-09	-19.5	59.04
E	13.0	7.63E-07	4.61E+04	***2.14E+03***	***3.23E-04***	1.63E-03	4.24E+02		2.51E-07	-20.4	73.06
F	11.8	4.92E-07	1.83E+03	2.14E+04	3.24E-05	2.47E-03	2.81E+02	1.15E-07	3.09E-07	-28.6	91.12
G	10.5	6.71E-07	1.93E+04	1.71E+03	4.06E-04	1.01E-03	6.87E+02	5.92E-07	1.14E-07	-22.8	1.00
H	11.0	6.32E-07	2.41E+04	3.02E+03	2.29E-04	4.49E-04	1.55E+03	1.49E-07	7.53E-08	-19.7	81.09
K	6.8	1.18E-05	4.48E+03	1.30E+04	5.32E-05	***1.53E-01***	***4.53E+00***		1.55E-06	-30.9	72.12
L	11.9	2.72E-08	2.85E+04	2.09E+04	3.32E-05	1.45E-02	4.77E+01	6.96E-07	6.74E-07	-27.6	57.11
M	13.5	2.38E-06	5.51E+04	1.32E+04	5.24E-05	2.16E-03	3.22E+02	1.63E-07	1.27E-07	-24.3	75.14
N	11.3	4.52E-07	7.66E+04	8.68E+03	7.99E-05	***3.93E-03***	***1.77E+02***		1.35E-08	-20.2	58.05
Q	12.5	2.11E-07	4.53E+03	5.89E+03	1.18E-04	1.58E-03	4.38E+02	2.69E-07	2.36E-07	-23.2	72.08
R	11.6	9.41E-06	3.50E+04	7.91E+03	8.76E-05	3.24E-03	2.14E+02	4.09E-07	6.25E-07	-21.7	100.13
S	13.1	1.51E-08	5.49E+04	1.53E+04	4.53E-05	1.94E-04	3.58E+03	1.27E-08	1.03E-08	-19.9	31.02
T	14.4	2.91E-08	2.29E+04	1.40E+04	4.96E-05	1.47E-04	4.72E+03	1.05E-08	7.74E-09	-20.7	45.05
V	9.0	2.38E-07	5.65E+03	2.71E+02	2.56E-03	7.69E-04	9.01E+02	2.84E-06	2.52E-08	-19.7	43.08
W	8.8	7.37E-08	5.01E+03	***5.55E+03***	***1.25E-04***	4.09E-04	1.69E+03		2.16E-07	-28.5	130.16
Y	12.2	2.59E-07	7.93E+03	***4.43E+03***	***1.57E-04***	1.15E-03	6.04E+02		3.68E-07	-31.3	107.12

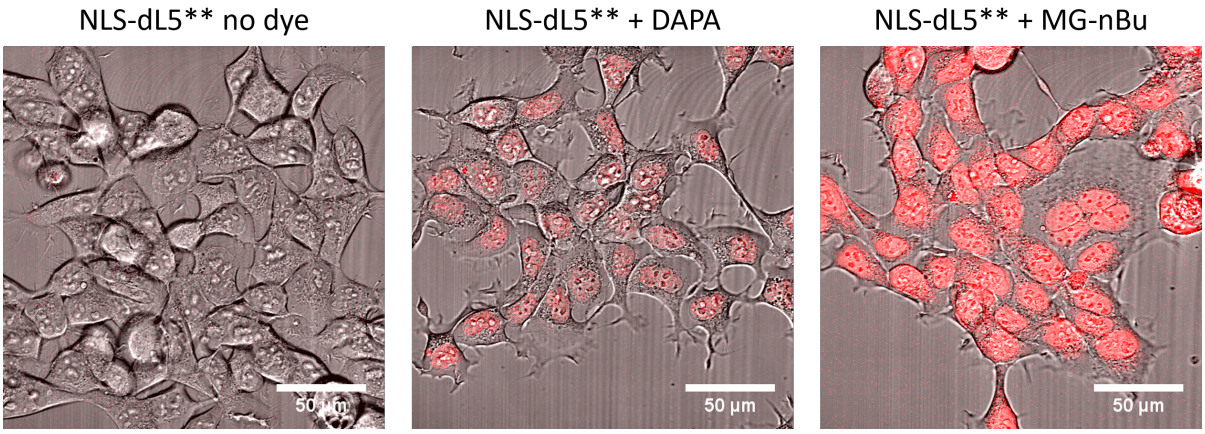
Supplemental Table 3: Summary of all mL5 (L91S E52X) properties shown in Figure 2.

The properties of mutants of mL5(L91S)(E52X) were determined in purified proteins secreted from yeast as described above. The molecular brightness of mL5(E52X) with MG-2p (1 μM protein, 5nM MG-2p) was measured using fluorescence correlation spectroscopy (FCS) as described above. Equilibrium measurements of K_D and complex brightness (Saturation point, B_{max}) were performed as described above. Kinetic measurements on purified mL5(E52X) were performed as described above. Half-lives are calculated from rates using the first order kinetics equation $t_{1/2} = \ln(2)/k$. ***Bolded data*****:** Some mutant data was not acquired for on-rate (W,W,Y) or off-rate (K,N). These holes were filled using the relationship between K_D, on and off rate (K_D(equilibrium measurement)=k_{off}/k_{on}). K_D values for the kinetic data were also calculated with the equation K_D(kinetic)=k_{off}/k_{on}. Isothermal titration calorimetry (ITC) was used to measure K_D and binding enthalpy of mL5(E52X) as described above.

	FCS	Equilibrium Measurements			Kinetic Measurements					ITC		side chain weight
	Brightness (kHz)	K_d (M)	log K_d (M)	Bmax (AU)	on-rate (s-1M-1)	t1/2 on (s)	off-rate (s)	t1/2 off (s)	K_d (M)	K_d (nM)	ΔH (kCal/mol)	(g/mol)
Brightness (kHz)	1											
Equilibrium K_d	-0.405	1										
log(Equilib. K_d)	-0.207	0.745	1									
Equilibrium Bmax	0.281	-0.152	-0.235	1								
on-rate	0.238	0.093	-0.191	0.068	1							
t1/2	-0.344	-0.166	-0.052	-0.163	-0.536	1						
off-rate	-0.586	0.760	0.464	-0.267	0.194	-0.126	1					
t1/2 off (s)	0.095	-0.349	-0.709	0.357	-0.023	0.087	-0.248	1				
Kinetics $K_d = k_{off}/k_{on}$	-0.557	-0.037	0.133	-0.432	-0.409	0.893	0.094	-0.322	1			
ITC K_d	-0.452	0.820	0.555	-0.394	0.272	-0.247	0.877	-0.477	-0.003	1		
ΔH (kCal/mol)	0.240	-0.304	-0.294	0.615	-0.336	0.337	-0.475	0.514	0.096	-0.661	1	
side chain weight	-0.051	0.232	0.316	-0.357	0.110	-0.339	0.056	-0.426	-0.159	0.298	-0.559	1
side chain weight (-Glycine)	-0.123	0.228	0.440	-0.484	-0.035	-0.372	0.024	-0.564	-0.154	0.286	-0.628	1

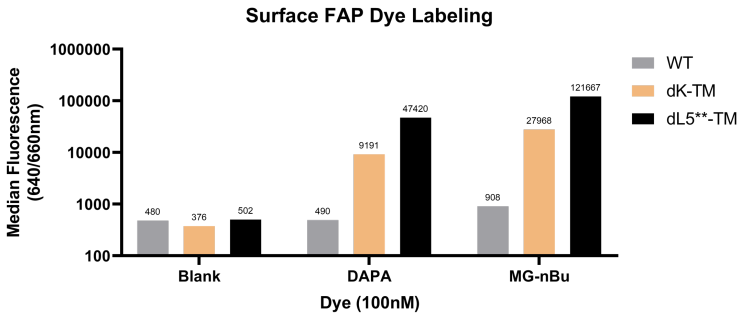
Supplemental Table 4: Correlation of all mL5 (L91S E52X) properties

Data listed in Table S3(N=17, Degrees of freedom=15), or data from Table S3 excluding Glycine (N=16, Degrees of freedom=14)were used to calculate Pearson's correlational coefficients. Kinetics highlighted for ease of viewing. Critical values of R (r(N-2) = abs(critical r value), resulting one sided p value): (r(15) = 0.327, p = 0.1), (r(15) = 0.412, p = 0.05), (r(15) = 0.558, p = 0.01), (r(14) = 0.338, p = 0.1), (r(14) = 0.426, p = 0.05), (r(14) = 0.574, p = 0.01).

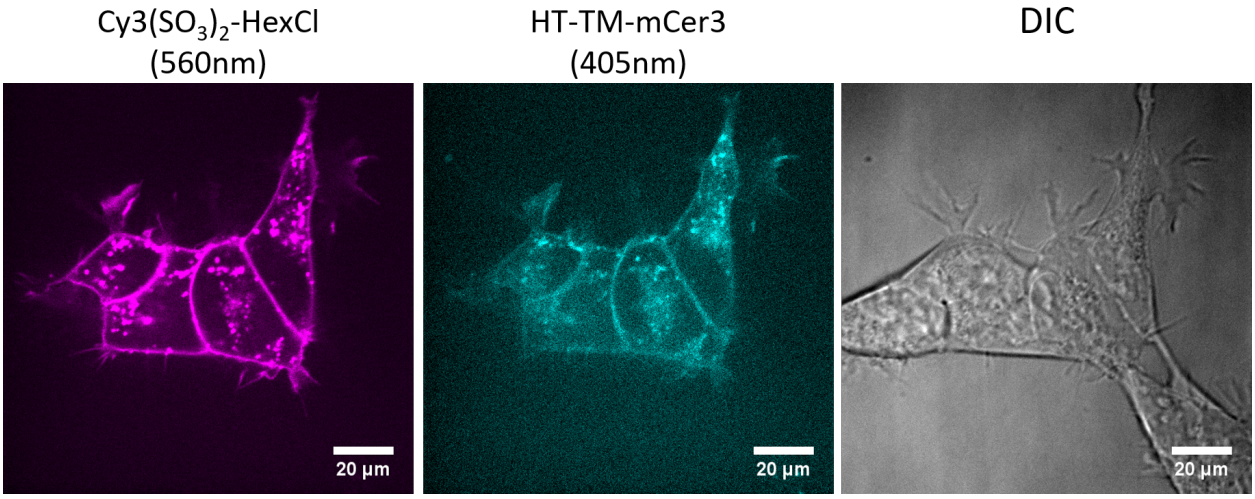


Supplemental Figure 11: NLS-dL5 cell permeability DIC+dye images with no-dye control**

Microscopy Analysis: HEK 293 cells stably expressing nuclear FAP (NLS-dL5) were labeled with 500 nM MG-PEG_{3500}-HexCl or MG-nBu, a cell permeant dye.[3] Dyes were incubated for 15 minutes, and then imaged. Lookup tables were matched and laser power at 633 nm were 6%. Scale bars are 50 µm. The no-dye control shows lack of red-fluorescence background. The MG-PEG_{3500}- HexCl shows very faint labeling in comparison to the MG-nBu. The low cell permeability of MG-PEG_{3500}- HexCl is expected since MG-2p, MG-11p-biotin, and other MG-PEG derivatives have been cell impermeant[4].



Supplemental Figure 12: Flow cytometry of dK-TM and dL5TM with MG Derivatives**
Flow cytometry analysis: HEK293 cells expressing dK-TM or dL5**-TM were labeled with 100 nM dye. The blank contains no dye. The median fluorescence of the expressing population was measured in the 640/660nm laser channel, median values appear above the log-scale bars. The dK-TM signals are about 25% of the dL5**-TM signals, potentially due to the combination of dK-TM expression level, lower quantum yield of fluorescence, and loosened K_d relative to dL5**-TM cells and construct properties.



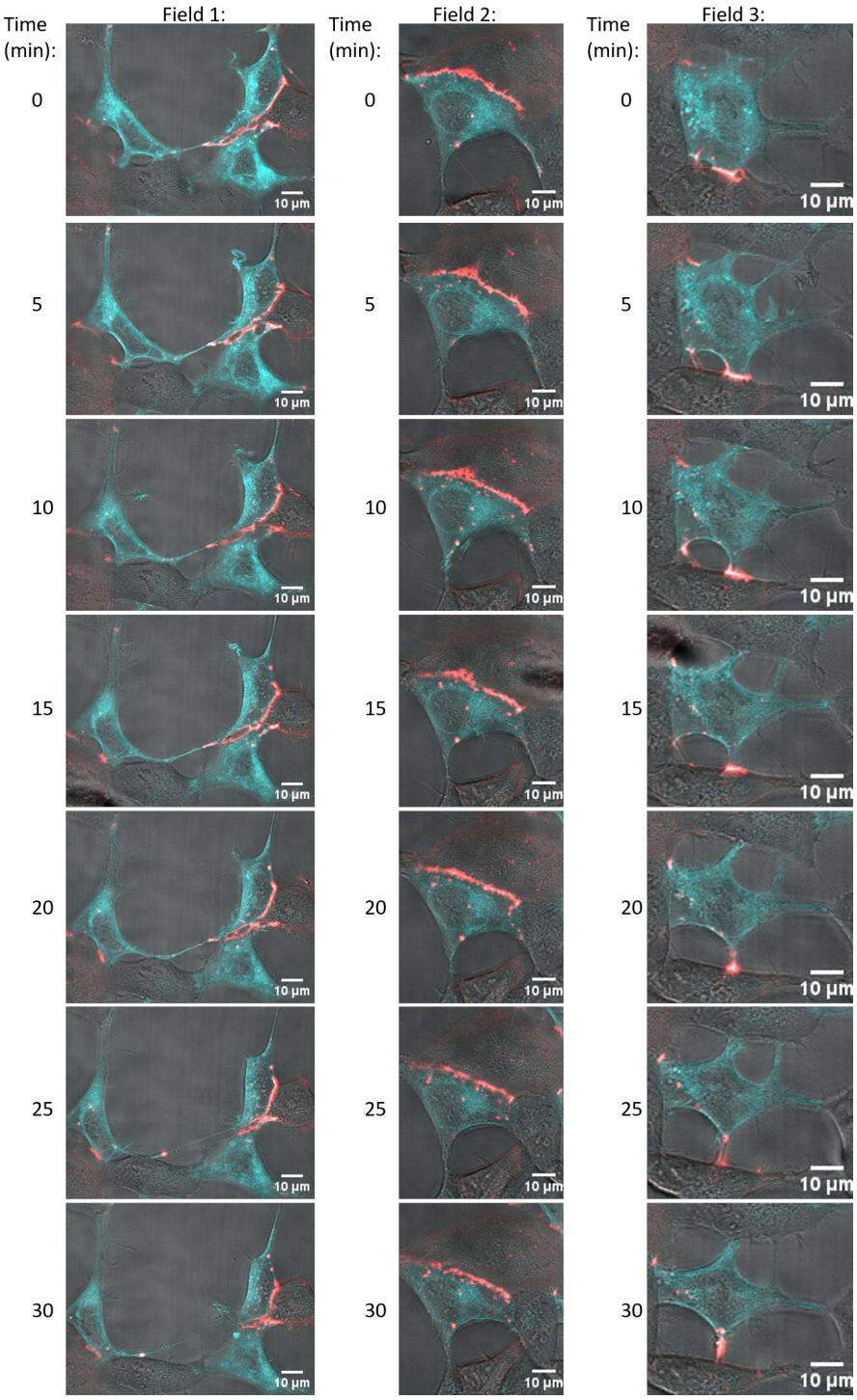
Supplemental Figure 13: HaloTag-TM-mCerulean3 labeling with $Cy3(SO_3)_2$-HexCl
HEK293 cells expressing HT-TM-mCer3 were labeled with 500 nM $Cy3(SO_3)_2$-HexCl for 5 minutes. The unbound dye was then removed via three washes of DMEM+10% FBS followed by a final media exchange to OptiMem. Cells were imaged two hours after media exchange using an Andor spinning disc microcope, 60x oil objective. Cy3(magenta) was imaged using 560 nm laser at 9% power and mCer3 (cyan) was imaged at 405 nm at 37% laser power.

Image Titles dK		N (contact ROI)	N (non-contact ROI)	N (images w/ contact sites)	N (images)
dK_.5nM_55_1b-1		5	7		
dk_.5_55_2-1		1	7		
dK_.5_55_3-1		4	7		
	subtotal	10	21	3	3
dk_1nM_25_2-1		8	10		
dK_1nM_25_3-1		4	8		
	subtotal	12	18	2	2
dK_10nM_25_1-1		5	9		
dK_10nM_25_2_z-9		13	12		
dK_10nM_25_4-1		9	11		
	subtotal	27	32	3	3
dK_100nM_.4_1-1		5	9		
dK_100nM_.4_2-1-NOCONT		0	15		
dK_100nM_.4_3-1		1	9		
	subtotal	6	33	2	3

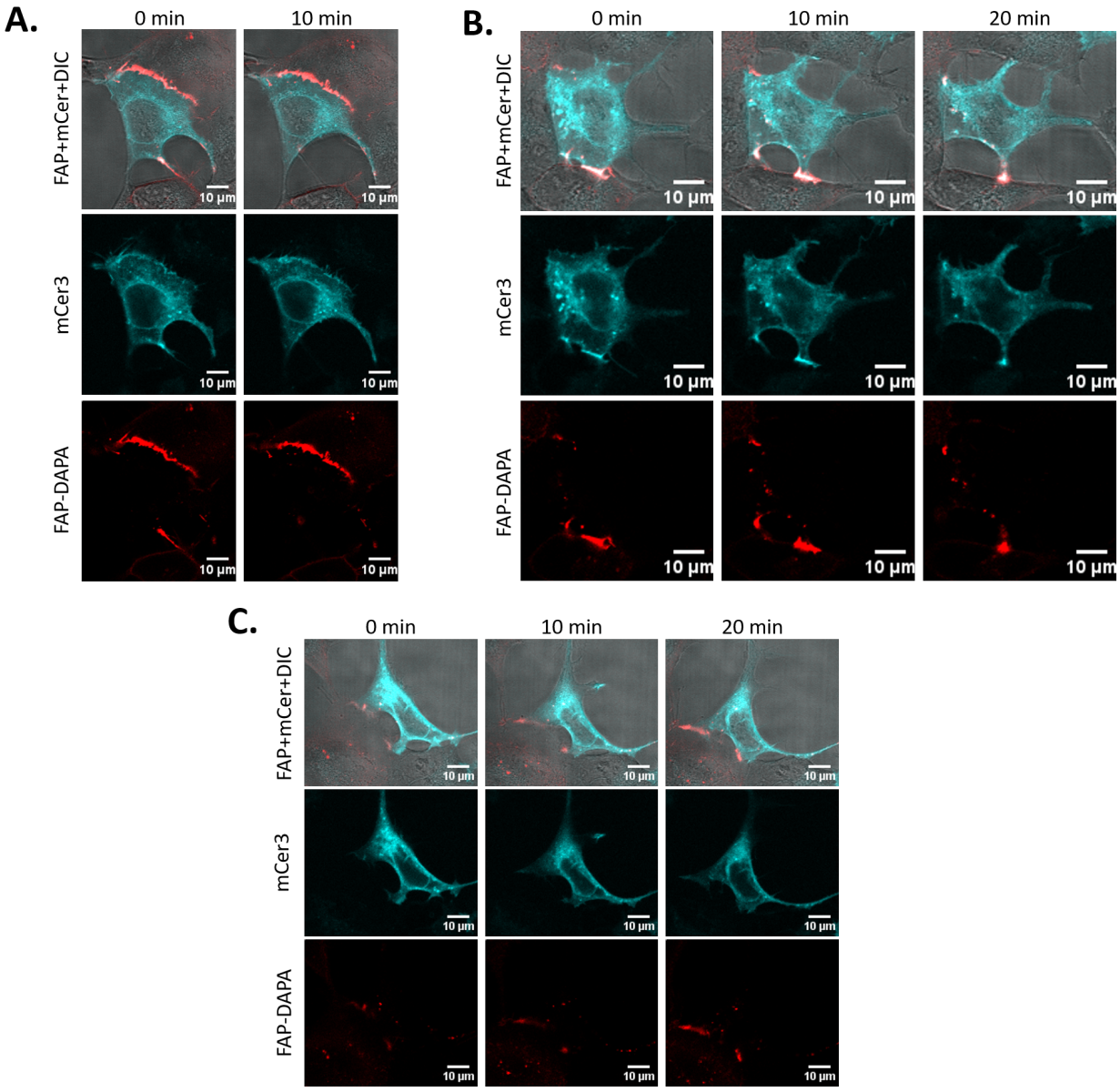
Supplemental Table 5: Table of Image titles and ROI counts for dK-TM with HT-TM-mCer3

Image Titles dL5**		N (contact ROI)	N (non-contact ROI)	N (images w/ contact sites)	N(images)
dL5_.5nM_.4_1-1		6	7		
dL5_.5nM_.4_1-2			11		
dL5_.5nM_.4_2-1		7	8		
dL5_.5nM_.4-1-4		0	13		
	subtotal	13	39	3	4
dL5_1nM_.4_2-2		3	9		
dL5_1nM_.4_2-1		4	8		
dL5_1nM_.4_1-1		4	10		
	subtotal	11	27	3	3
dL5_10_.4_3-2		8	12		
dL5_10_.4_3-1		8	11		
dL5_10nM_0.4_2b-2		6	12		
dL5_10nM_0.4_2b-1		11	11		
dL5_10nM+.4_1-1		3	10		
	subtotal	36	56	5	5
dl5_100nM_0.4_3-1		5	13		
dL5_100nM_0.4_2-1		4	13		
dL5_100nM_0.4_1-1		3	10		
dL5_100nM_0.4_1-4		3	13		
	subtotal	15	49	4	4

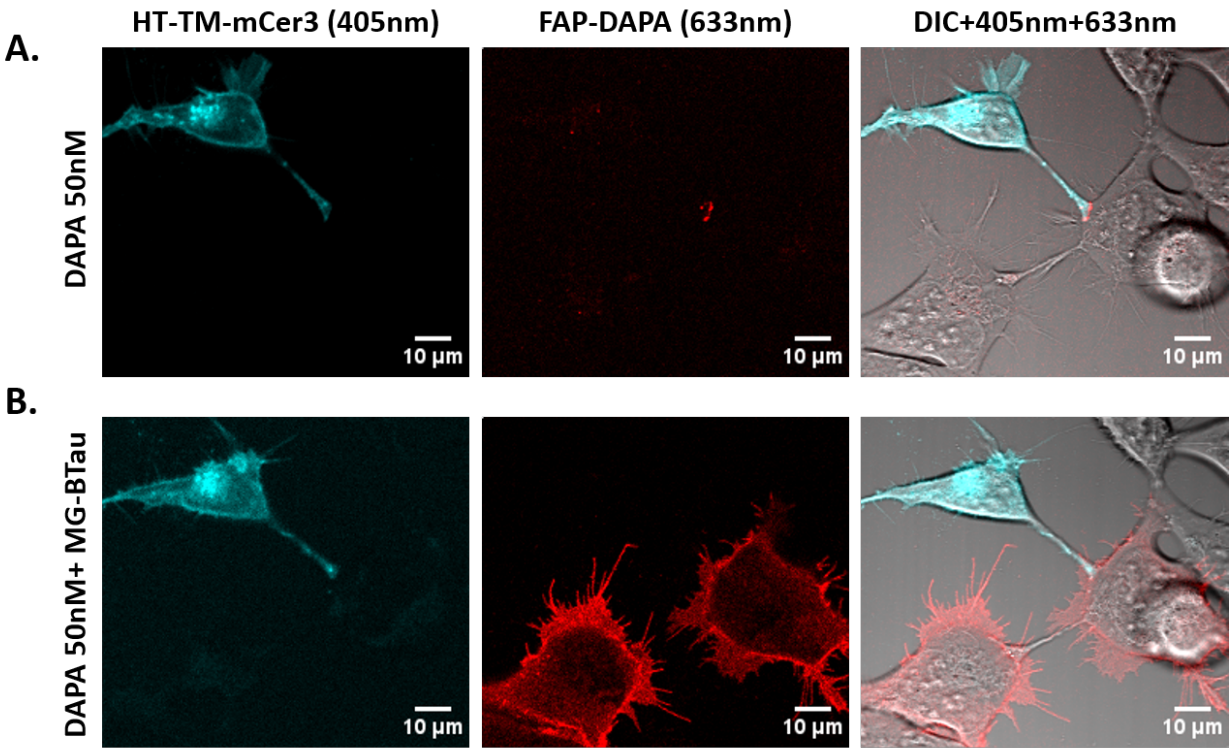
Supplemental Table 6: Table of Image titles and ROI counts for dL5-TM with HT-TM-mCer3**



Supplemental Figure 14: Time-lapse dK-TM and 10nM DAPA Supplemental Video1 stills HT-TM-mCerulean and dK-TM cells were plated on a poly-L-lysine coated dish. The cells were labeled with optiMEM containing 10 nM DAPA and imaging was initiated within 15 minutes of labeling without rinsing cells. All imaging was performed at 37°C with 5% CO_2 supply on a Zeis880 LSM with 405 nm diode laser (25.7% laser power, emission 409 - 486 nm) and 633 nm HeNe laser (6.0% laser power, 641 - 695 nm).



Supplemental Figure 15: Time-lapse of HT-TM-mCer3 and dK-TM with 10 nM DAPA
HT-TM-mCerulean and dK-TM expressing HEK 293 cells were plated on a poly-L-lysine coated dish. The cells were labeled with optiMEM containing 10 nM DAPA and imaged without rinsing cells. All imaging was performed at 37°C with 5% CO_2 supply on a Zeiss880 LSM with 405 nm diode laser (25.7% laser power, emission 409 - 486 nm) and 633 nm HeNe laser (6.0% laser power, 641 - 695 nm). Time 0 is consistent between B and C but not A. **A.** Within 10 minutes a site of cell contact is pruned showing loss of FAP-DAPA signal at the contact site. **B**. A broad site of cellular juxtaposition is reduced to a narrow point of contact upon contraction of central HT-TM-mCer3 – expressing cell. **C.** A site of contact between a HT-TM-mCer3 cell and a dK-TM expressing cell is broadened over 20 minutes.



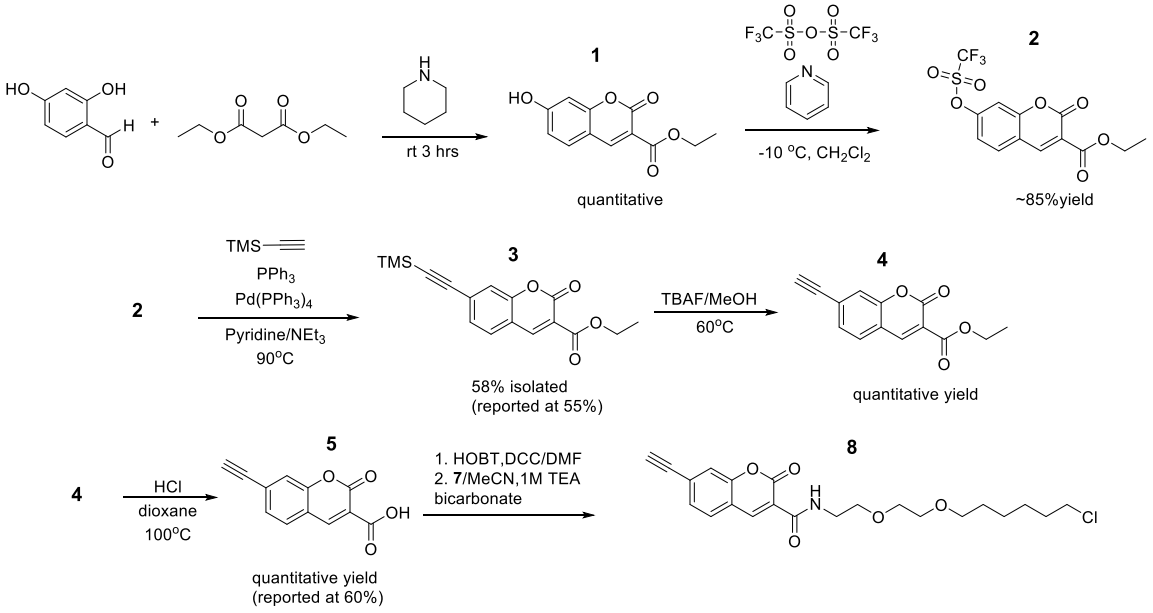
Supplemental Figure 16: "Cell-drop" experiment with HT-TM-mCer3 and dK-TM
HT-TM-mCerulean3 cells were plated into 8 well Ibidi dishes coated with poly-L-lysine. Cells were grown overnight and then labeled with 50nM DAPA for 5 minutes in DMEM+10%FBS to saturate all HT receptors with DAPA reagent. Cells were gently washed 6 times with DMEM + 10% FBS over a period of 1 hour. At this point, a test well was confirmed for HT/DAPA labeling by addition of purified dL5** at 1µM(not shown). Cells expressing dK-TM, suspended via gentle trituration, were then seeded into the HT-TM-mCer3/DAPA dishes and allowed to attach for 6 hours prior to imaging. Imaging was performed on a Zeiss 880 LSM at excitation = 633nm(55% laser power) and 405nm (25% laser power), **(A).** After imaging DAPA labeled cells, saturating MG-BTau was added to confirm the identity of the FAP expressing cells, **(B)**. The results show that DAPA dye persists on the surface of labeled HT-TM-mCer3 cells for many hours. Additionally, this panel demonstrates the utility of MG-BTau in conjunction with the dK-TM cells to confirm identity of FAP expressing cells once contacts are observed.

Synthesis:

General Methods:

No unexpected or unusually high safety hazards were encountered. NMR spectra were obtained on a Bruker Avance 500 MHz Instrument. The electrospray ionization mass spectrometry (ESI-MS) experiments were run on a ThermoFisher Scientific ESI/APCI LCQ Ion Trap MS. MALDI-TOF run on Applied Biosystems Voyager DE-STR MALDI-TOF instrument. (NSF grant #CHE-9808188). NH_2PEG_{3500}alkyne was purchased from Jen Kem. ESI and MALDI-TOF experiments were performed in the Center for Molecular Analysis at Carnegie Mellon University. MALDI TOF supported by NSF grant number CHE-9808188. Thin-Layer chromatography (TLC) was performed using silica gel on aluminum plates. NMR assignments were made with assistance of MestreNova auto-assignment tool.

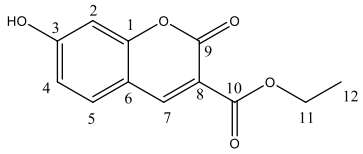
Fluorogenic Coumarin Alkyne Synthesis:



Schematic 1: Fluorogenic Coumarin Alkyne Synthesis

Ethyl 7-hydroxy-2-oxo-2H-chromene-3-carboxylate "hydroxycoumarin ethyl ester" (1):

As described by Wright *et al.*,[5] Dihydroxybenzaldehyde (9.75 g, 70.6 mmol) and diethylmalonate (24 mL, 160 mmol) were mixed and piperidine (12 mL, 121mmol) was added dropwise under constant stirring. The mixture dissolved to give a purple solution with a yellow sheen. After 3 hours the reaction was quenched via acidification with 3M HCl (50 mL). The solids that formed, were filtered and washed with 50 mL ice cold water and dried overnight to yield 17.2 g of **1** as a yellow flaky solid (quantitative). **1** was used without further purification. TLC: silica gel, 30% ethyl acetate/dichloromethane Rf (**1**)=0.3. ESI-MS *m/z*: $[M]^-$ Calcd for $C_{12}H_9O_5^-$, 233.05; Found 233.2 m/z. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.67 (d, *J* = 0.7 Hz, 1H, *7*), 7.75 (d, *J* = 8.6 Hz, 1H, *5*), 6.84 (dd, *J* = 8.6, 2.3 Hz, 1H, *4*), 6.73 (dd, *J* = 2.2, 0.6 Hz, 1H, *2*), 4.26 (q, *J* = 7.1 Hz, 2H, *11*), 3.35(br s, 4H, *H2O*,*OH*)1.30 (t, *J* = 7.1 Hz, 3H, *12*). ^{13}C NMR (500 MHz, DMSO-d_6) 164.5(*3*), 163.4(*10*), 157.6(*1*), 156.9(*9*), 149.9(*7*), 132.6(*5*), 114.5(*4*), 112.6(*8*), 110.9(*6*), 102.3(*2*), 61.3(*12*), 40(*dmso*), 14.6(*11*).

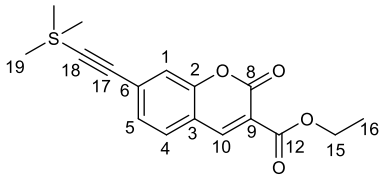


Ethyl 2-oxo-7-(((trifluoromethyl)sulfonyl)oxy)-2H-chromene-3-carboxylate "coumarin triflate ethyl ester" (2) :

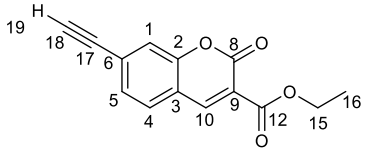
Based on protocol from Wright *et al.,* [5] **1** (9.51g, 27.8 mmol), anhydrous dichloromethane (50 mL), and freshly distilled triethylamine (4 mL,31.2 mmol) were combined in a three neck round bottom flask. The coumarin dissolved upon addition of triethylamine to yield a bright yellow solution which was cooled to 0°C. An ice cold solution of triflic anhydride (4.8 mL, 28.5 mmol) in dichloromethane (30 mL) was added to the coumarin solution via canula over a period of about 30 minutes. The solution was titrated with two more portions of triflic anhydride (2x 4.8 mL, 28.5 mmol) and triethylamine (2x 4 mL, 31.2 mmol) until the solution was no longer yellow. Upon verification of completed reaction by TLC (silica gel/ dichloromethane/ethyl acetate 7:3), the reaction was filtered and the filtrate washed with 3x 30 mL water and 20 mL brine. The organic layer was dried over sodium sulfate and concentrated to yield 11.2g **2** as light brown solids (75 % yield). To achieve a higher yield, the organic layer can be concentrated to solids and pure **2** was isolated after recrystallization in hot ethyl acetate. TLC: Hexanes/Ethyl acetate Rf(7)=0.9 ESI $[M+H]^+$ calcd. 367.0m/z (367.0), $[M+Na]^+$ calcd 388.99 (388.9), (429.5), $[2M+Na]^+$ calcd 754.99 (754.5). ^{1}H NMR (300 MHz, DMSO-d_6) δ 8.81 (s, 1H), 8.14 (d, J = 8.7 Hz, 1H), 7.81 (d, J = 2.4 Hz, 1H), 7.58 (dd, J = 8.7, 2.5 Hz, 1H), 4.32 (q, J = 7.1 Hz, 2H), 4.15 – 3.97 (m, 1H), 3.17 (d, J = 5.2 Hz, 1H), 1.32 (t, J = 7.1 Hz, 3H), 1.27 – 1.13 (m, 1H).

Ethyl 2-oxo-7-(trimethylsilyl)ethynyl)-2H-chromene-3-carboxylate "TMS-acetylene-coumarin-ethyl ester" (3):

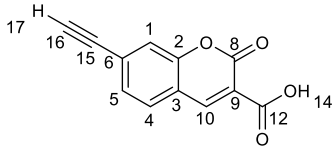
In a modified protocol, [6,7] **2** (7.0g, 19mmol), triphenyl phosphine (1.08g, 4.00 mmol), and palladium tetrakis-triphenylphosphine (1.2 mg, 1 mmol, 0.05 mol%) were added to a three neck round bottom flask fitted with a condenser under argon. Freshly distilled triethylamine (10 mL) and anhydrous pyridine (20 mL) were added and heated up. The mixture formed a brown solution with yellow sheen when it reached 90°C. TMS-acetylene (5.0 mL, 1.9 eq) was added dropwise to the solution in 4 portions over a period of about 1.5 hrs. The mixture was concentrated in vacuo advancing further the reaction. The concentrate was purified by column chromatography on a 3x40cm silica gel column applying a slow gradient of hexane 5-20% /ethyl acetate allowing the separation of **3** from the TMS-acetylene-pyridine salts, **2**, and byproducts. Alternatively: a crude chromatography to isolate high Rf fractions followed by their slow evaporation in vacuo can lead to pure **3** as transparent light-yellow crystals (3.5g, 58% isolated yield). TLC: (20% $CHCl_3$/ hexane or 20% ethyl acetate/hexane) RF(**3**)=0.6, Pyridine-TMS-acetylene salt: RF=0.9. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.42 (d, J = 0.7 Hz, 1H,*10*), 7.51 (d, J = 8.0 Hz, 1H, *4*), 7.35 – 7.28 (m, 2H, 1, *5*), 4.34 (q, J = 7.1 Hz, 2H, *15*), 1.34 (t, J = 7.1 Hz, 3H, *16*), 0.23 (s, 9H, *19*). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 162.71 (*12*), 156.22 (*8*), 154.70 (*2*), 147.62 (*10*), 129.31 (*6*), 129.17 (*4*), 128.26 (*5*), 119.47 (*3*), 118.23 (*1*), 117.76 (*9*), 103.00 (*17*), 100.44 (*18*), 61.90 (*15*), 14.18 (*16, 19*).



Ethyl 7-ethynyl-2-oxo-2H-chromene-3-carboxylate (4): [6,7] Compound **3** (260 mg, 0.82 mmol) was dissolved in methanol (30 mL) and brought to 60ºC. Under argon, tert-butylaluminum fluoride (3 mL, 1M in THF, 3 mmol) was added and stirred for 15 minutes. The reaction mixture was cooled to room temperature and quenched via addition of cold water until sample precipitated (~20mL). The white solids were filtered. To achieve quantitative yield the mother liquor volume was reduced in vacuo and re-filtered. Drying yielded a white sticky solid (0.24g, quantitative). NMR was obtained in warm DMSO. TLC: 50% hexane/chloroform , Rf **(4)**=0.45, Rf(**3**)=0.50 ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.46 (s, 1H, *10*), 7.56 (d, *J* = 8.4 Hz, 1H, *5*), 7.37 (d, *J* = 5.9 Hz, 2H, *1, 4*), 4.38 (q, *J* = 7.1 Hz, 2H, *15*), 3.37 (s, 1H, *19*), 1.38 (t, *J* = 7.2 Hz, 3H, *16*).
^{13}C NMR (126 MHz, Chloroform-*d*) δ 162.71(*12*), 156.17 (*8*), 154.66 (*2*), 147.60 (*10*), 129.42 (*6*), 128.37 (*5*), 128.15 (*4*), 119.93 (*1*), 118.60 (*3*), 118.08 (*9*), 81.87 (*18*), 77.36 (*17*), 62.04 (*15*), 14.19 (*16*).



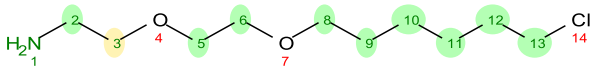
7-Ethynyl-2-oxo-2H-chromene-3-carboxylic acid (5): Compound **4** (1.08g) was suspended in dioxane (25 mL) and hydrochloric acid (20 mL, 6 M). The mixture dissolved upon heating to 92ºC and was refluxed for 6 hours. The progression of the reaction was monitored by TLC (silica gel, 30% ethyl acetate/hexane) showing a drop in Rf from 0.6 to ~0, but NMR was required to confirm 100% conversion. The reaction mixture was quenched with water and extracted into 3 x 30 mL ethyl acetate. Residue **5** was isolated upon concentration of the organic phase and was used without further purification. TLC: RF(**5**)=0.0 in EtOAc. ^{1}H NMR (300 MHz, DMSO-d_6) δ 13.29 (s, 1H, *14*), 8.73 (s, 1H, *10*), 7.90 (d, *J* = 8.0 Hz, 1H, *4*), 7.54 (d, *J* = 1.4 Hz, 1H, *1*), 7.47 (dd, *J* = 8.0, 1.5 Hz, 1H, *5*), 4.60 (s, 1H, *17*).



tert-butyl (2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)carbamate, "BOC-hexCl" (6):
Compound **6** was generously provided by Brigitte Schmidt and Bruce Feldman who synthesized the reagent according to the literature. [8]

2-(2-((6-chlorohexyl)oxy)ethoxy)ethan-1-amine "hexCl amine" (7)

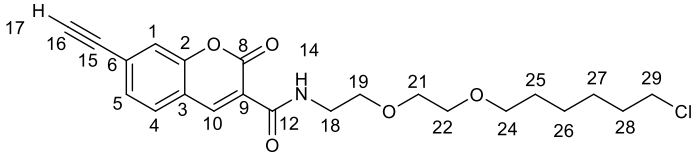
Compound **6** (53 mg, 0.16 mmol) was deprotected in anhydrous dichloromethane (5 mL) with excess TFA (500 µL). After stirring for 30 minutes the mixture was concentrated and the TFA removed in vacuo. The paste was re-dissolved in acetonitrile and brought to pH 7 using Tetraethyl ammonium bicarbonate (1M, 800 µL) and DIPEA(400 µL) to yield **7** which was used immediately. ^{1}H NMR (300 MHz, Methanol-d_4) δ 3.63 (ddt, J = 7.2, 5.5, 2.2 Hz, 5H, *6, 13*), 3.52 (dt, J = 10.8, 6.6 Hz, 4H, *5, 8*), 3.36 (s, 3H, *3*), 2.99 (t, J = 5.1 Hz, 2H, *2*), 1.77 (dq, J = 8.1, 6.5 Hz, 2H, *12*), 1.71 – 1.30 (m, 4H, *9, 10, 11*), 0.91 (dt, J = 16.7, 7.3 Hz, 0H, salts).

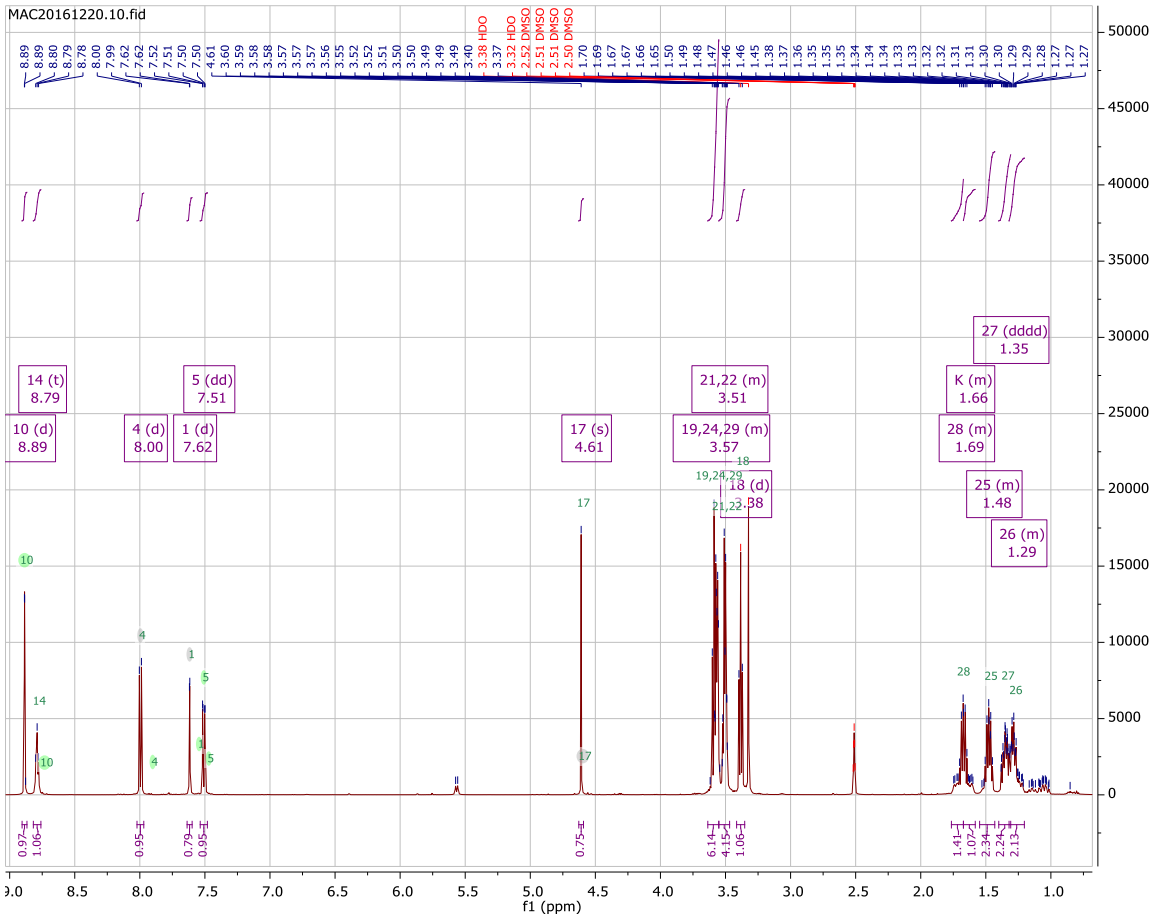


N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-7-ethynyl-2-oxo-2H-chromene-3-carboxamide "alkyne coumarin hexCl" (8):

5 (0.20 g, 0.90 mmol, 1 eq) was dissolved in 2 mL DMF. HOBT(0.15 g, 1.1 mmol, 1.2 eq) and DCC (0.38 g, 1.8 mmol, 2 eq) were added to give a suspension. After 1 hour the urea had crashed out to form a brown paste.

The NHS ester was added to the solution of freshly de-boc protected **7**, portion-wise, along with additional acetonitrile. Additional DIPEA was added to maintain pH 7 to 8 until the reaction was complete (monitored by TLC). Upon workup, the solution was filtered, concentrated to an oil, diluted with ethyl acetate (150 mL) and extracted with water (3 x 50 mL), HCl (2 x 40 mL), water(1 x 30 mL) and brine(1 x 30 mL). The organic layer was then dried over sodium sulfate and concentrated to give an orange-yellow paste (395 mg). The crude residue was purified on silica (hexane/ ethyl acetate) to yield a flakey yellow solid (300 mg, 77% yield). TLC silica gel, ethyl acetate (Rf=0.8, yellow spot, UV active, faint ninhydrin stain with heat). ESI: $[M]^+$ calcd. 419.15; found 420.1m/z. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.89 (d, J = 0.7 Hz, 1H, *10*), 8.79 (t, J = 5.5 Hz, 1H, *14*), 8.00 (d, J = 8.1 Hz, 1H, *4*), 7.62 (d, J = 1.4 Hz, 1H, *1*), 7.51 (dd, J = 8.0, 1.5 Hz, 1H, *5*), 4.61 (s, 1H, *17*), 3.64 – 3.55 (m, 6H, *19, 24, 29*), 3.55 – 3.47 (m, 4H, 21, *22*), 3.38 (d, J = 13.1 Hz, 1H, *18*), 1.76 –1.58 (m, 2H, *28*), 1.55 – 1.43 (m, 2H, *25*), 1.35 (dddd, J = 12.0, 9.1, 5.9, 1.5 Hz, 2H, *27*), 1.32 – 1.20 (m, 2H, *26*). ^{13}C NMR (126 MHz, DMSO-d_6) δ 161.29 (*12*), 160.57 (*8*), 154.06 (*2*), 147.24 (*10*), 130.97 (*4*), 128.66 (*5*), 127.26 (*6*), 119.59 (*3*), 119.44 (*9*), 119.39 (*1*), 85.62 (*15*), 82.69 (*16*), 70.68 (*24*), 70.16 (*22*), 69.94 (*21*), 69.19 (*19*), 47.96 (*29*), 39.58 (*18*), 33.80 (*28*), 29.53 (*26*), 26.58 (*27*), 24.91 (*25*).



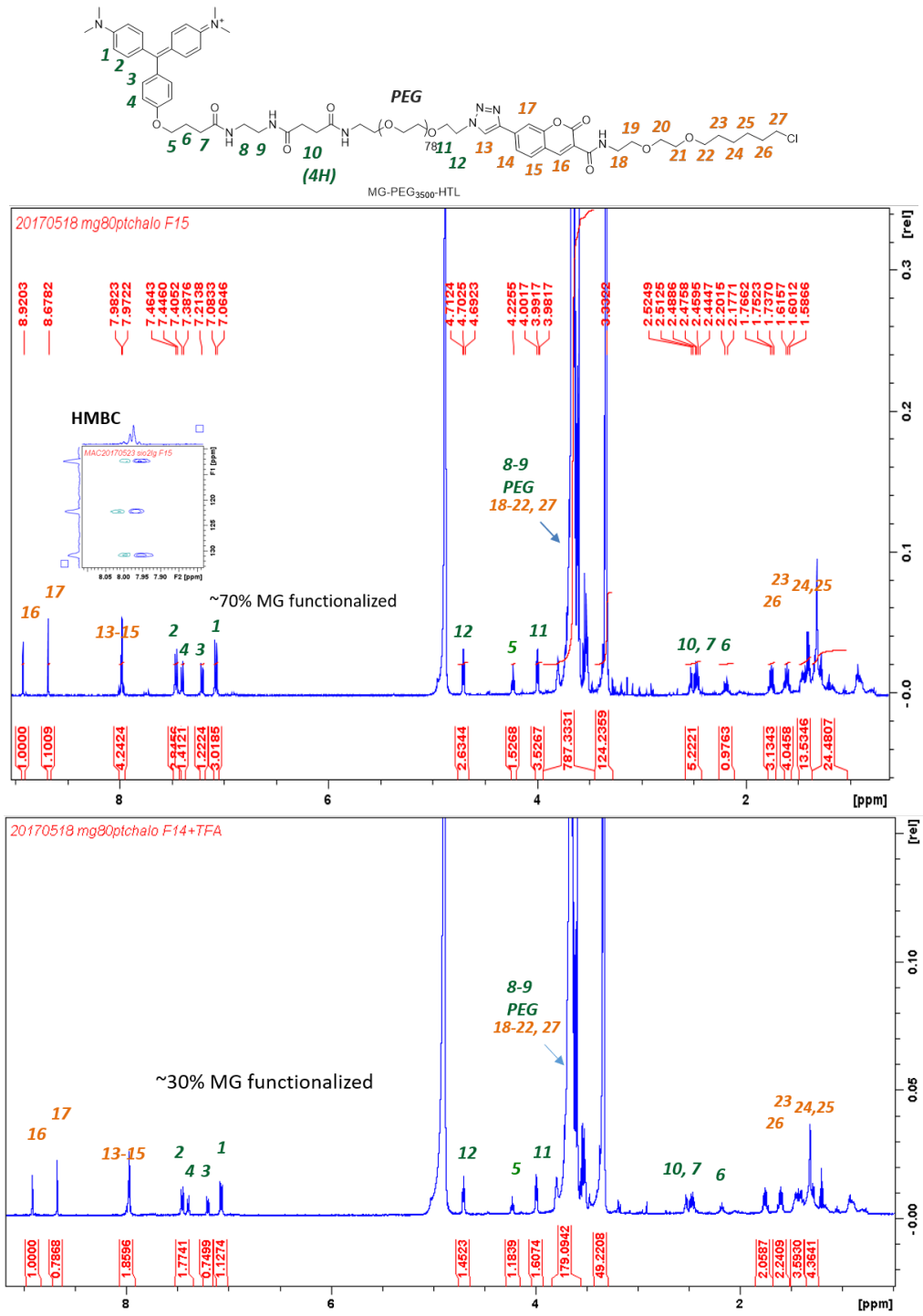


Supplemental Figure 17: ^{1}H-NMR of Alkyne-Coumarin-HexCl (8)
Peaks assigned to 28 and "K" should not be split into two integrals. Peaks at 1.69 and 1.66ppm are in one signal assigned to 28.

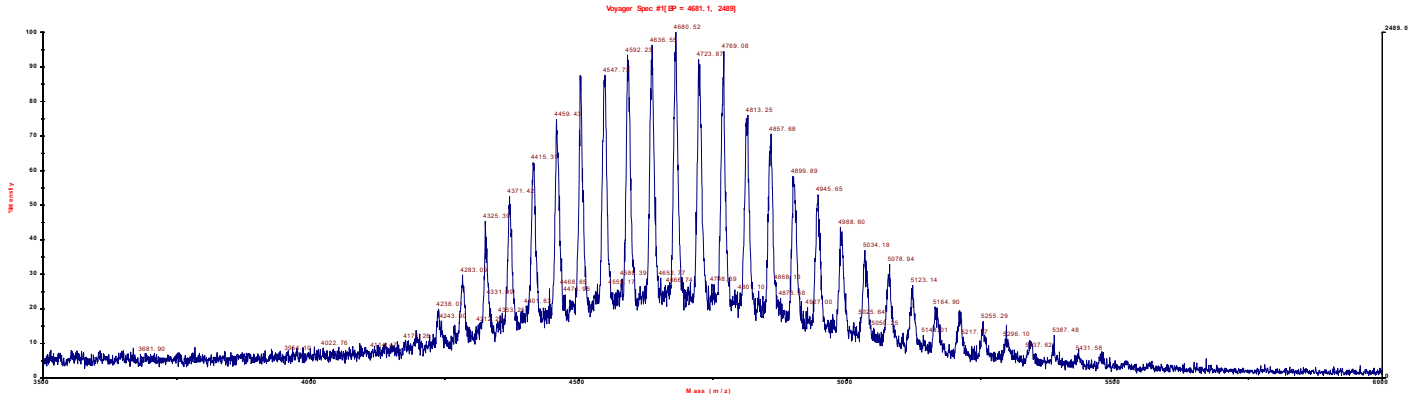
NH_2-PEG_{3500}- HexCl (9): Compound **7** (20 mg, 48 µmol, 1.15eq) and NH_2-PEG_{3500}-alkyne (145 mg, 42.1 µmol, 1eq) were dissolved in anhydrous DMF (1.5mL) and purged with argon for 5 minutes. PMDTA (10 µL, 48 µmol,1.15 eq) , Cu(II)acetate (2.5mg, 14 µmol) were added neat under positive pressure of argon to yield a blue solution which became red upon addition of solid sodium ascorbate (2.8 mg, 14 µmol). After 1 hour the reaction was completed by TLC (silicagel, 15% methanol in chloroform with visualization with UV, iodine stain, and ninhydrin (Rf=0.8, ninhydrin negative, HexCl-PEG_{3500}-HexCl or X-PEG_{3500}-HexCl, where X is unreacted end group) (Rf 0.7= ninhydrin positive, NH_2-PEG_{3500}-HexCl). The reaction was quenched by addition of water and was isolated by extraction into chloroform (3 x 30 mL) followed by washes of the organic layer with 1M HCl (2 x 15 mL), 1M EDTA (1 x15 mL), and brine (1 x 15 mL). The organic layer was dried in vacuo to yield crude **8** (115 mg, 70% yield) as a yellow residue with a light blue fluorescent sheen. The product was used without further purification.

MG-PEG_{3500}-HexCl (10): MG-ethylenediamine-succinate (MG-EDA-succinate)[9,10] (29mg, 0.5 mmol) in oxidized form was dissolved in dry dichloromethane (15 mL) under argon. TSTU (21mg, 0.70 mmol) was added neat and allowed to stir overnight. **9** (115 mg, 0.03 mmol) was dissolved in anhydrous DCM (5

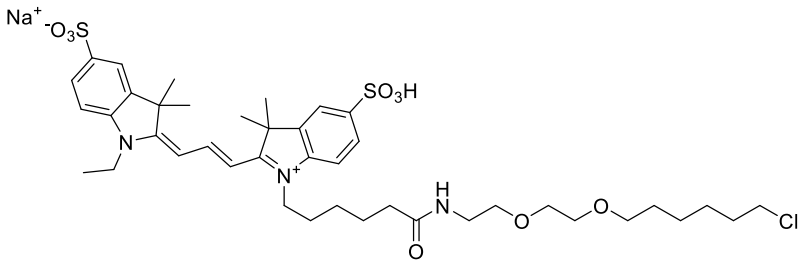
mL) along with DIPEA (50 μL, 0.28 mmol) to give an orange solution. MG-EDA-succinate-NHS (0.1 mmol, 50mM in DCM) was added to the amine in two portions along with more DIPEA (50uL) to give a light green solution. After several hours, the sample was isolated via precipitation into ether followed by size exclusion chromatography on a P2 sephadex column in 10-15% acetonitrile/water. Fractions were confirmed to contain PEG with I_2 vapor staining. Isolated polymers were then separated into mono-HexCl, bis-HexCl, and bis-MG functionalized polymer derivatives via a slow silica gel column starting at 5% methanol/chloroform and increasing through 7% 10% and 15% methanol/chloroform. Fractions were characterized using TLC in 5% methanol/chloroform followed by 10% methanol/chloroform. (Rf(0.8)= **HexCl-PEG$_{3500}$-HexCl** or X-PEG$_{3500}$-HexCl, Rf(0.7)= **10** (green spot only), Rf(0.6) = MG-PEG$_{3500}$-MG. Fractions of isolated samples were acidified with TFA and characterized by NMR in $CDCl_3$.
Yields: F14 3.2 mg, F15, 2.1mg. F16: 5.1mg Fraction 15: ^{1}H NMR (500 MHz, Methanol-d_4)
δ 8.92 (s, 1H), 8.68 (s, 1H), 8.05 – 7.91 (m, 3H), 7.46 (d, J = 9.0 Hz, 1H), 7.44 – 7.30 (m, 1H), 7.30 – 7.16 (m, 1H), 7.13 – 7.01 (m, 1H), 4.70 (t, J = 5.0 Hz, 1H), 4.23 (t, J = 6.3 Hz, 1H), 3.99 (t, J = 5.0 Hz, 2H), 3.82 – 3.78 (m, 1H), 3.83 – 3.58 (m, 185H), 3.74 – 3.57 (m, 190H), 3.53 (dt, J = 9.4, 6.6 Hz, 3H), 3.39 – 3.22 (m, 30H), 2.53 (td, J = 6.2, 1.8 Hz, 1H), 2.50 – 2.38 (m, 1H), 1.75 (p, J = 6.8 Hz, 2H), 1.60 (p, J = 6.8 Hz, 2H), 1.51 – 1.37 (m, 4H), 1.37 – 1.22 (m, 7H), 0.98 – 0.82 (m, 3H). See Supplemental Figure 18 for assignments.



Supplemental Figure 18: ^{1}H-NMR MG-PEG$_{3500}$-HexCl (10) fractions 14 and 15 with assignments



Supplemental Figure 19: MALDI-TOF of MG-PEG$_{3500}$-HexCl (10):
MG-PEG$_{3500}$-HexCl column isolate fraction 14 was co-crystalized with matrix, α-cyano-4-hydroxycinnamic acid (10 mg/mL) from methanol/water 1:1. Mn=4634.0, Mw=4673.3.

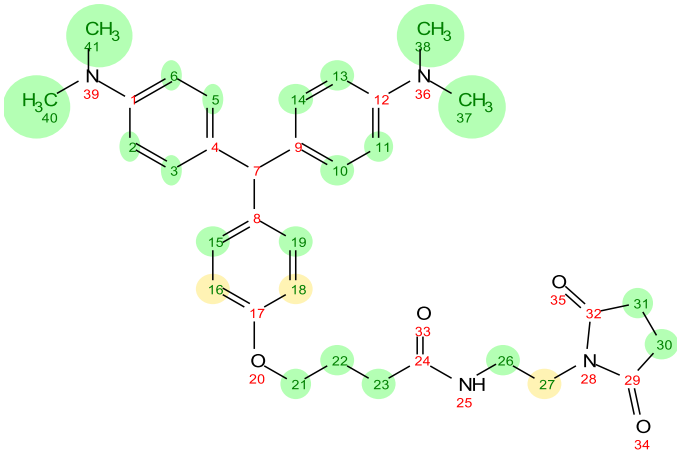


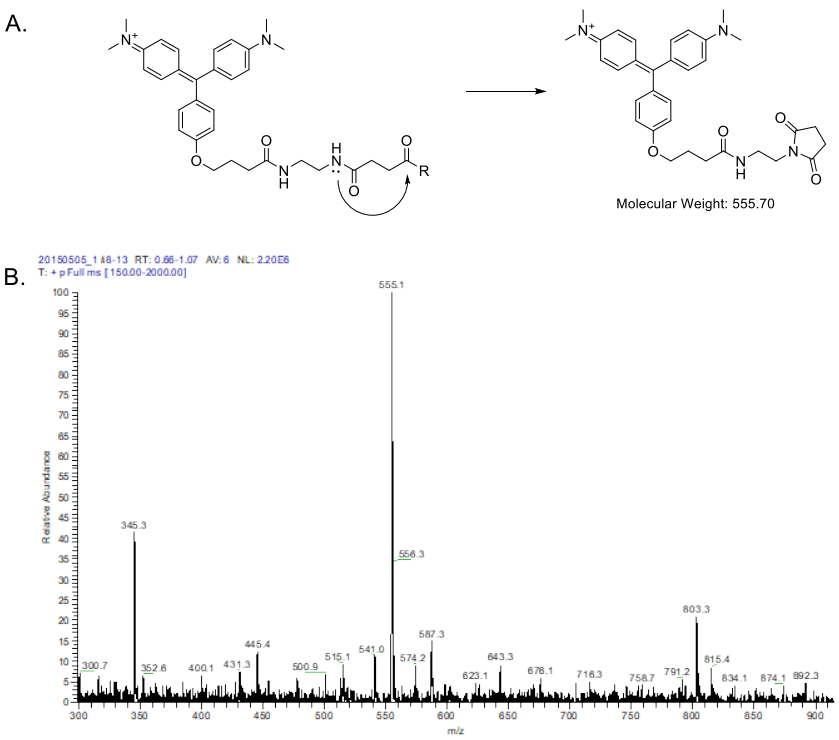
Bis sulfonate Cy3-HexCl (11): The sodium salt of Cy3.29 (bis-sulfonated Cy3, 66 mg, 0.1 mmol) was dissolved in 1.5 mL dry DMF to which was added TSTU (40 mg, 0.132 mmol) and DIEA (40 µL, 0.42 mmol) and allowed to stir for 1 hour. The conversion to the Cy3.29-NHS was tested via precipitation in ether followed by dissolution in 30% MeCN/water 0.1% TFA and reverse phase TLC in 15% ethanol/water (NHS, rf=0.2 vs rf=0.6). **6** (32.3 mg, 0.1 mmol) was deprotected in TFA/MeCN 1:1 100 µL, stirred overnight and isolated in vacuo. The remaining TFA was neutralized with 100 uL of 1M NaH_2CO_3, dried, and used immediately. **7** was dissolved in 0.5mL dry DMF. The Cy3.29-NHS mixture was added dropwise to the amine. After 100 µL addition the reaction was monitored via precipitation into ether followed by dissolution in 1M $NaHCO_3$ and reverse phase TLC in 15% Ethanol/water (rf=.01). After addition of all Cy3.29-NHS, the solution was stirred overnight followed by isolation via precipitation into 50 mL dry ethyl ether and column chromatography on a RP18 column. The product was eluted with 20% ethanol/water to yield 45 mg of product (52% yield). ESI: m/z 880.3[Prod+2Na]$^+$, 858.3[Prod+Na]$^+$, 451.6 [Prod+4Na]$^{2+}$, 441.1 [Prod+2Na]$^{2+}$. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.43 (t, *J* = 13.4 Hz, 1H), 8.02 (d, *J* = 1.5 Hz, 1H), 7.90 (dd, *J* = 8.2, 1.6 Hz, 1H), 7.43 (dd, *J* = 7.8, 6.7 Hz, 2H), 7.30 (t, *J* = 7.5 Hz, 1H), 7.23 – 7.16 (m, 1H), 7.13 (d, *J* = 8.0 Hz, 2H), 6.57 (d, *J* = 13.5 Hz, 1H), 6.50 (d, *J* = 13.3 Hz, 1H), 4.23 (q, *J* = 7.2 Hz, 2H), 4.08 (t, *J* = 7.7 Hz, 2H), 3.74 (hept, *J* = 6.7 Hz, 1H), 3.61 (dd, *J* = 5.9, 3.3 Hz, 2H), 3.58 – 3.50 (m, 6H), 3.43 (dt, *J* = 12.0, 6.3 Hz, 4H), 3.20 (dq, *J* = 29.5, 7.4 Hz, 2H), 2.68 (s, 8H), 2.28 (t, *J* = 7.4 Hz, 2H), 2.08 (s, 2H), 1.86 – 1.78 (m, 3H), 1.75 (d, *J* = 4.7 Hz, 13H), 1.71 (q, *J* = 7.5 Hz, 2H), 1.59 (h, *J* = 7.7, 7.0 Hz, 2H), 1.55 – 1.30 (m, 15H).^{13}C NMR (126 MHz, $CDCl_3$) δ 174.70, 174.17, 174.10, 173.81, 172.75, 150.75, 143.64, 142.61, 141.51, 140.78, 140.21, 128.96, 127.37, 125.64, 122.31, 120.89, 110.86, 110.47, 103.48, 103.19, 77.29, 77.04, 76.78, 71.16,

70.10, 69.95, 69.51, 54.11, 49.34, 49.31, 46.17, 45.10, 44.53, 42.38, 39.66, 39.15, 35.97, 32.49, 29.37, 28.01, 27.02, 26.64, 26.21, 25.48, 25.34, 25.23, 20.87, 12.54, 12.10, 8.66, -0.04.

MG- pyrrolidine-2,5 dione (555m/z) (12)

A MG-pyrrolidine derivative was isolated as a standard degradation product and perpetual contaminant. MG-EDA-pyrrolidine-2,5 dione is seen in almost every reaction of MG. This compound has an rf ~0.8 in 20% Methanol/chloroform which makes it difficult to separate from other functionalized MG derivatives (MG-HexCl, MG2p-boc, MG-2p-BriB, MG-EDA-succinate). This compound was isolated via NMR to confirm the identity. The two pyrrolidine methylene peaks appear as a singlet integrating to 4H which also confused previous attempts to characterize the product because it seemed like no alkyl groups were present. The evidence for this compound was cited in a classic polymer functionalization paper as a potential side product of pegylation of proteins and likely forms through mechanism shown [11] ESI: calcd 555.7, found 555.1. ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.43 (d, J = 8.9 Hz, 3H, *3, 5, 10, 14*), 7.37 (d, J = 8.5 Hz, 2H, *15, 19*), 7.21 – 7.16 (m, 3H, *16, 18*), 7.05 (d, J = 8.9 Hz, 3H, 2, 6, 11, 13), 4.19 (t, J = 6.4 Hz, 2H, *21*), 3.70 – 3.56 (m, 3H, *26*), 3.40 (t, J = 5.6 Hz, 2H, *27*), 3.32 (s, 15H, *37, 38, 40, 41*), 2.91 (s, 2H), 2.67 (s, 3H, *30, 31*), 2.37 (t, J = 7.4 Hz, 2H, *23*), 2.13 (p, J = 6.9 Hz, 2H, *22*).

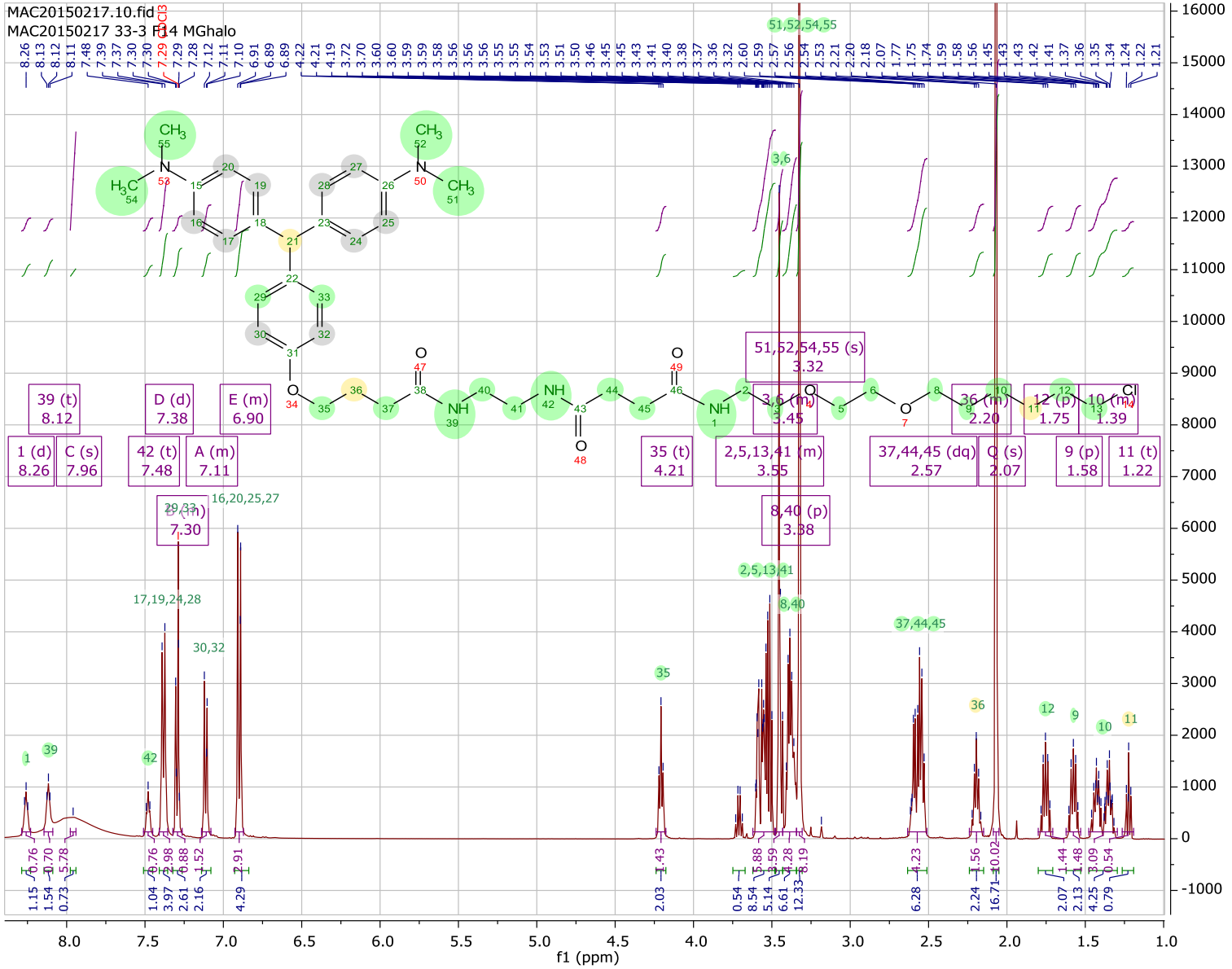




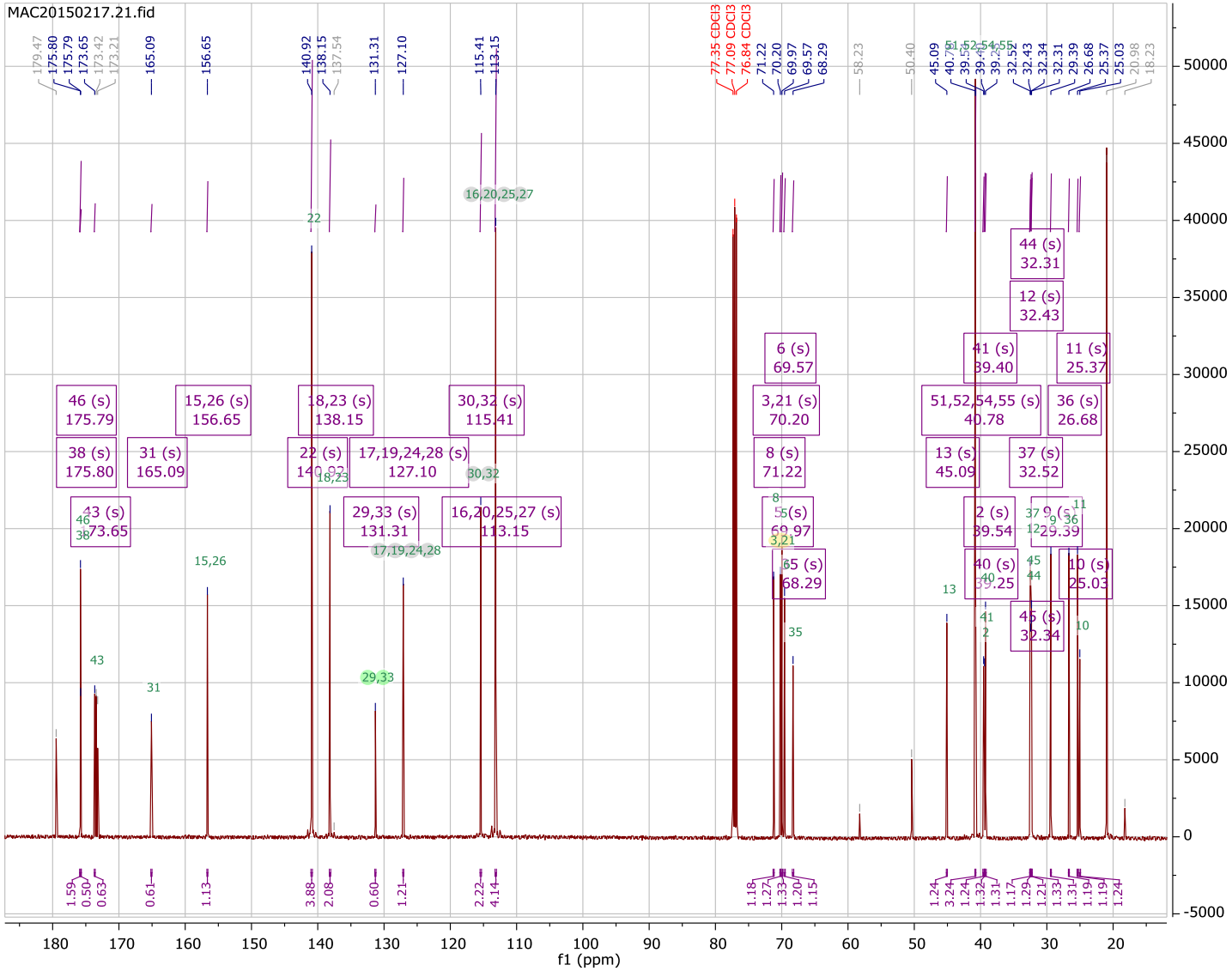
Supplemental Figure 20: Identification of compound of 555m/z as 12

A. Formation of the pyrolodinedione compound occurs through nucleophilic substitution reaction. **B.** ESI of crude MG-2p-derivative(803 m/z) with MG-pyrolidine-2,5-dione contaminant (555.1 m/z).

MG-HexCl (13): Reduced MG-EDA-succinate (0.20 g, 360 mmol, 1 eq) was dissolved in DMF (600 µL). TSTU (0.13 g, 430 mmol, 1.2 eq) was added and stirred for one hour. Di-isopropyl ethylamine(75 µL) and freshly de-protected **7** (92 mg,0.415 mmol,1.1 eq) were added. After 1.5 hours TLC in 20% MeOH/$CHCl_3$ indicated complete conversion (rf 0.8). Workup was done via dilution with saturated $NaHCO_3$ followed by extraction into ethyl acetate (2 x 15mL). Chromatography on silica gel using 5-20% MeOH/$CHCl_3$ yielded 250 mg MG-HexCl as a mixture of the oxidized and reduced forms. Chloranil was dissolved in boiling acetonitrile and chloroform (1:1) and added to the MG-HexCl (250 mg,0.32 mmol) in acetonitrile (20 mL) at 50 ºC. The mixture was brought to 80 ºC and stirred for 1 hour to yield a dark green solution. A flash column using first acetonitrile followed by 0% to 20% Methanol in Chloroform yielded 274 mg dark green oxidized MG-hexCl, 61% yield. ESI: calcd: 778.4, found: 778m/z. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.26 (d, *J* = 5.0 Hz, 1H, 1), 8.12 (t, *J* = 5.0 Hz, 1H, 39), 7.96 (s, 6H), 7.48 (t, *J* = 5.7 Hz, 1H, 42), 7.38 (d, *J* = 8.9 Hz, 3H), 7.32 – 7.26 (m, 1H), 7.14 – 7.08 (m, 2H), 6.93 – 6.87 (m, 3H), 4.21 (t, *J* = 6.6 Hz, 1H, 35), 3.62 – 3.48 (m, 6H, 2, 5, 13, 41), 3.48 – 3.43 (m, 4H, 3, 6), 3.38 (p, *J* = 6.7, 6.1 Hz, 4H, 8, 40), 3.32 (s, 8H, 51, 52, 54, 55), 2.57 (dq, *J* = 20.3, 6.8 Hz, 4H, 37, 44, 45), 2.24 – 2.15 (m, 2H, 36), 2.07 (s, 10H), 1.75 (p, *J* = 6.9 Hz, 1H, 12), 1.58 (p, *J* = 6.9 Hz, 1H, 9), 1.48 – 1.30 (m, 3H, 10), 1.22 (t, *J* = 7.0 Hz, 1H, 11). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 175.80 , 38, 175.79 , 46, 173.65 , 43, 165.09 , 31, 156.65 , 15, 26, 140.92 , 22, 138.15 , 18, 23, 131.31 , 29, 33, 127.10 , 17, 19, 24, 28, 115.41 , 30, 32, 113.15 , 16, 20, 25, 27, 71.22 , 8, 70.20 , 3, 21, 69.97 , 5, 69.57 , 6, 68.29 , 35, 45.09 , 13, 40.78 , 51, 52, 54, 55, 39.54 , 2, 39.40 , 41, 39.25 , 40, 32.52 , 37, 32.43 , 12, 32.34 , 45, 32.31 , 44, 29.39 , 9, 26.68 , 36, 25.37 , 11, 25.03 , 10. See Supplemental Figure 21 and 22 for NMR and assignments.



Supplemental Figure 21: Assigned ^{1}H-NMR of MG-HexCl (13) as reference compound



Supplemental Figure 22: Assigned ^{13}C-NMR of MG-HexCl (13) as reference compound

References:

(1) Yan, Q.; Schwartz, S. L.; Maji, S.; Huang, F.; Szent-Gyorgyi, C.; Lidke, D. S.; Lidke, K. A.; Bruchez, M. P. Localization Microscopy Using Noncovalent Fluorogen Activation by Genetically Encoded Fluorogen-Activating Proteins. *ChemPhysChem* **2014**, *15* (4), 687–695.

(2) Colby, D. W.; Kellogg, B. A.; Graff, C. P.; Yeung, Y. A.; Swers, J. S.; Wittrup, K. D. Engineering Antibody Affinity by Yeast Surface Display. *Methods Enzymol.* **2004**, *388* (2000), 348–358.

(3) Perkins, L. A.; Fisher, G. W.; Naganbabu, M.; Schmidt, B. F.; Mun, F.; Bruchez, M. P. High-Content Surface and Total Expression SiRNA Kinase Library Screen with VX-809 Treatment Reveals Kinase Targets That Enhance F508del-CFTR Rescue. *Mol. Pharm.* **2018**, *15* (3), 759–767.

(4) Szent-gyorgyi, C.; Schmidt, B. a; Creeger, Y.; Fisher, G. W.; Zakel, K. L.; Adler, S.; Fitzpatrick, J. a J.; Woolford, C. a; Yan, Q.; Vasilev, V.; et al. Supp: Fluorogen Activating Proteins : Technology for Imaging and Assaying Cell Surface Proteins . Methods and Supplementary Figures. *Nat. Biotechnol.* **2008**.

(5) Wright, A. T.; Song, J. D.; Cravatt, B. F. Supporting- A Suite of Activity-Based Probes for Human Cytochrome P450 Enzymes. *J. Am. Chem. Soc.* **2009**, *131* (30), 10692–10700.

(6) Zhou, Z.; Fahrni, C. J. A Fluorogenic Probe for the Copper (I)-Catalyzed Azide-Alkyne Ligation Reaction. *J. Am. Chem. Soc.* **2004**, *126* (I), 8862–8863.

(7) Wright, A. T.; Song, J. D.; Cravatt, B. F. A Suite of Activity-Based Probes for Human Cytochrome P450 Enzymes. *J. Am. Chem. Soc.* **2009**, *131* (30), 10692–10700.

(8) Zhang, Y.; So, M. K.; Loening, A. M.; Yao, H.; Gambhir, S. S.; Rao, J. Supporting-HaloTag Protein-Mediated Site-Specific Conjugation of Bioluminescent Proteins to Quantum Dots. *Angew. Chemie - Int. Ed.* **2006**, *45* (30), 4936–4940.

(9) Szent-Gyorgyi, C.; Schmidt, B. A. F.; Schmidt, B. A. F.; Creeger, Y.; Fisher, G. W.; Zakel, K. L.; Adler, S.; Fitzpatrick, J. A. J.; Woolford, C. A.; Yan, Q.; et al. Fluorogen-Activating Single-Chain Antibodies for Imaging Cell Surface Proteins. *Nat. Biotechnol.* **2008**, *26* (2), 235–240.

(10) Ackerman, D. S.; Vasilev, K. V; Schmidt, B. F.; Cohen, L. B.; Jarvik, J. W. Tethered Fluorogen Assay to Visualize Membrane Apposition in Living Cells. *Bioconjug. Chem.* **2017**, *28* (5), 1356–1362.

(11) Zalipsky, S. Functionalized Poly(Ethylene Glycols) for Preparation of Biologically Relevant Conjugates. *Bioconjug. Chem.* **1995**, *6* (2), 150–165.