Evaluation of the Cytosolic Uptake of HaloTag Using a pH-Sensitive Dye

JoLynn B. Giancola,[§] Jonathan B. Grimm,[#] Joomyung V. Jun,[§] Yana D. Petri,[§]
Luke D. Lavis,[#] and Ronald T. Raines[§,*]

[§]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States
[#]Janelia Research Campus, Howard Hughes Medical Institute 19700 Helix Drive, Ashburn VA 20147

*rtraines@mit.edu

Table of Contents

I. Abbreviations Used

ACN	acetonitrile
CHCA	α-cyano-4-hydroxycinnamic acid
DIC	*N*,*N*′-diisopropylcarbodiimide
DIPEA	*N*,*N*′-diisopropylethylamine
DCM	dichloromethane
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
DTNB	5,5′-dithio-bis-(2-nitrobenzoic acid)
DPBS	Dulbecco's phosphate-buffered saline
Fmoc	fluorenylmethyloxycarbonyl
FPLC	fast protein liquid chromatography
HATU	(1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate)
HCTU	*O*-(1*H*-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HOBt	hydroxybenzotriazole
HTL	HaloTag ligand
IPTG	isopropyl β-D-1-thiogalactopyranoside
MALDI–TOF	matrix-assisted laser desorption/ionization time-of-flight
MFI	mean fluorescence intensity
MS	mass spectrometry
NTB	2-nitro-5-thiobenzoic acid
PBS	phosphate-buffered saline
PEG	polyethylene glycol
py	pyridine
Q–TOF	quadrupole time-of-flight
RPM	revolutions per minute
SA	sinapic acid
TEA	triethylamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIS	triisopropylsilane

II. General Methods

All reactions were performed in a reaction vial fitted with TFE-silicone septa under N_2(g) using standard Schlenk-line techniques. Reactions carried out at low temperature were cooled by cooling agents in a Dewar vessel (water-ice bath at 0 °C), whereas reactions performed above room temperature were heated on the IKA RCT basic plate. All reaction mixtures were stirred with a magnet and monitored by liquid chromatography–mass spectrometry (LC–MS) and by analytical thin layer chromatography (TLC). Purification was done with flash column

chromatography performed with silica gel or a Biotage Isolera One system unless indicated otherwise. It should be noted that all diazoamide compounds were purified by hand column chromatography to avoid any exposure to UV light from the UV detector in the Biotage system. Organic solutions were concentrated in vacuo with a Buchi rotary evaporator (model R-210).

Reagents and Solvents. Commercially available reagents and solvents were reagent grade or better, and were used directly without further purification. Reagents and solvents were from Sigma–Aldrich (St. Louis, MO) unless otherwise specified. Amino acids were from Chem-Impex International (Wood Dale, IL). Rink Amide ProTide resin (LL) and Oxyma Pure were from CEM Corporation (Matthews, NC). DIC and 4-methylpiperidine were from Oakwood Chemical (Tampa, FL). Anhydrous DMSO, TIS, and TFA were from Sigma–Aldrich. 2-Diazo-2-(*p*-methylphenyl)-*N*,*N*-dimethylacetamide (**1**) was synthesized by Chicago Discovery Solutions (Plainfeld, IL) as described previously.[1] Water was obtained from a Milli-Q IQ 7000 purification system and had a resistivity of 18.2×10^6 Ω cm.

Solvent Removal. The phrase "concentrated under reduced pressure" refers to the removal of solvents and other volatile materials with a rotary evaporator at water-aspirator pressure of <20 Torr and a water bath at ~25 °C. Residual solvents were removed from compounds by vacuum (<0.1 Torr) achieved by using a mechanical belt-drive oil pump.

Peptide Synthesis. All amino acids used in solid-phase peptide synthesis (SPPS) were of L stereochemistry and were protected at their N terminus with fluorenylmethoxycarbonyl (Fmoc) (Chem-Impex). Peptides were synthesized on Rink Amide ProTide Resin (LL) (0.1 mmol, 0.59 mmol/g, 1.0 equiv) with a Liberty Blue Automated Microwave Peptide Synthesizer from CEM Corporation (Matthews, NC) following the manufacturer's standard procedures. Standard solutions of Oxyma Pure (1.0 M in DMF), *N,N′*-diisopropylcarbodiimide (DIC) (0.5 M in DMF), 4-methylpiperidine (20% v/v in DMF), and Fmoc-protected amino acids (0.2 M in DMF) were used in coupling and deprotection steps. Peptides were purified with a 1260 Infinity II Preparative LC System from Agilent Technologies equipped with a XSelect Peptide CSH C18 OBD prep column (130 Å pore size, 5 µm particle size, 19 mm × 250 mm of width × length) from Waters Corporation (Milford, MA).

Conditions. Synthetic procedures to afford **2**-SSpy (page S6) were performed under a positive pressure of N_2(g) at ambient temperature (~22 °C) unless indicated otherwise. All other procedures were performed at ambient temperature (~22 °C) and pressure (~1.0 atm) unless indicated otherwise.

Nuclear Magnetic Resonance (NMR) Spectroscopy. ^{1}H and ^{13}C{^{1}H} NMR spectra were acquired with a Bruker Avance Neo 400 MHz or Bruker Avance Neo 500 MHz spectrometer at the Department of Chemistry Instrumentation Facility (DCIF) at the Massachusetts Institute of Technology or a Bruker AV400 or AV600 spectrometer at the Janelia Research Campus. Proton chemical shifts are reported in parts per million (ppm, δ scale) and are relative to residual protons in the deuterated solvent ($CDCl_3$: δ 7.26; D_2O: δ 4.79; CD_3CN: δ 1.94). Carbon chemical shifts are reported in parts per million (ppm, δ scale) and are relative to the carbon resonance of the solvent ($CDCl_3$: δ 77.2; CD_3CN: δ 1.3). $CDCl_3$ was supplied by Cambridge Isotope Laboratories (Tewksbury, MA). D_2O and CD_3CN were from Sigma–Aldrich. Multiplicities are abbreviated as:

s (singlet), br (broad), d (doublet), t (triplet), q (quartet), sept (septet), and m (multiplet). The $^{13}C\{^{1}H\}$ signal of the diazo carbon (C=N=N) is missing in most of the spectra, possibly due to a T1 relaxation effect.[2]

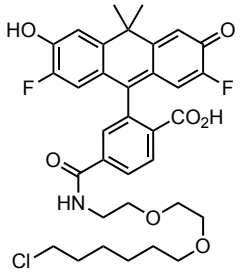
Mass Spectrometry. Mass spectra of small molecules were acquired on an LCT electrospray ionization (ESI) 1260 Infinity II instrument from Agilent Technologies (Santa Clara, CA) and an LC-MS column (Agilent Technologies, Poroshell 120, SB C18-reversed-phase, length 50 mm, internal diameter: 2.1 mm, particle size: 2.7 micron) with a gradient of 10−95% v/v MeCN (0.1% v/v formic acid) in water (0.1% v/v formic acid) over 10 min. To minimize the fragmentation of diazo moieties, the MSD parameters were set as follows: capillary voltage, 3000 V; drying gas temperature, 350 °C; gas flow, 13/min; fragmentor voltage, 30 V; nebulizer pressure, 35 psig; and cycle time, 0.83 s/cycle. HRMS of peptides and small molecules was performed with Agilent 6545 Q–TOF mass spectrometer coupled to an Agilent Infinity 1260 LC system (Q–TOF). The crude molecular mass of peptides, proteins, and protein conjugated was determined on either α-cyano-4-hydroxycinnamic acid (CHCA) matrix or sinapic acid matrix, respectively, by matrix-assisted laser desorption/ionization time-of-flight (MALDI–TOF) mass spectrometry with a microflex LRF instrument from Bruker (Billerica, MA). MALDI samples were all desalted using DOWEX 50WX4-400 strong cation exchange resin (CAS #11113-61-4) before spotting 1:1 v/v with the appropriate matrix. A more accurate assessment of the molecular mass of peptides, proteins, and protein conjugates was carried out using ESI mass spectrometry on a 6530C Accurate-Mass Q–TOF MS equipped with a PLRP-S column (1000 Å, 5-µm, 50 mm × 2.1 mm) from Agilent Technologies. A gradient of 5–95% v/v MeCN (0.1% v/v formic acid) in water (0.1% v/v formic acid) over 7 min was used unless otherwise indicated. Before Q–TOF LC-MS analysis, all samples were passed through a Spin-X Centrifugal Tube Filter (0.22-µm, cellulose acetate membrane) from R&D Systems (Minneapolis, MN).

Compound Purity. The purity of small molecules was judged to be ≥95%, as assessed by ^{1}H and ^{13}C NMR spectroscopy, mass spectrometry, and reversed-phase high performance liquid chromatography (HPLC) using an LCMS column and gradient of 10−95% v/v MeCN (0.1% v/v formic acid) in water (0.1% v/v formic acid) over 10 min unless indicated otherwise. The purity of peptides was accessed with a 1260 Infinity II Preparative LC System from Agilent Technologies equipped with an EC NUCLEOSIL 100-5 C18 analytical column (100 Å, 5 µm, 4.6 mm × 250 mm) from Macherey-Nagel (Düren, Germany) or 1200 Infinity System from Agilent Technologies equipped with a Microsorb-MV 100-5 C18 column (100 Å, 5 µm, 4.6 mm × 250 mm) from Varian (Palo Alto, CA).

Biological Reagents, Supplies, and Instrumentation. The Monarch PCR and DNA Cleanup Kit (5 µg) was from New England Biolabs (product #T1030S). Restriction enzymes were from New England Biolabs. The Gibson Assembly Master Mix was from New England Biolabs (product #E2611). The GeneJET Plasmid Miniprep Kit was from Thermo Fisher Scientific (Waltham, MA; product #K0502). Protein purification was performed with an ÄKTA Pure FPLC purification system from Cytiva (Marlborough, MA). The HiTrap Talon Crude 5-mL Column was from GE Healthcare (product #28-9538-09). TEV protease was from New England Biolabs (product #P8112S). Spectra Multicolor Broad Range Protein Ladder was from Thermo Fisher Scientific. Bio-Safe Coomassie G-250 Stain was from Bio-Rad Laboratories. Titer Plate Shaker was from Labline Instruments (Melrose Park, IL). Protein concentrations were determined with a

DS-11 UV–vis Spectrophotometer/Fluorimeter from DeNovix (Wilmington, DE). SDS–PAGE analyses were performed with Any kD Mini-PROTEAN TGX Precast Gels in a Mini-PROTEAN Tetra cell from Bio-Rad Laboratories (Hercules, CA). Gels were imaged with an Amersham Imager 600 from GE Healthcare Life Sciences (Marlborough, MA). Pierce Dye Removal Columns were from Thermo Fisher Scientific (product #22858). Amicon Ultra 0.5-mL 10K MWCO Centrifugal Filter Unit were from MilliporeSigma (Burlington, MA; product #UFC501024). Zeba Spin Desalting Columns, 0.5-mL 7K MWCO, were from Thermo Fisher Scientific (product #89882). Protein conjugate fluorescence for pK_a determination was assayed in 96-well half area, black flat bottom, non-binding plates from Corning (product #3993), and fluorescence readings were collected a Spark plate reader from Tecan (Männedorf, Switzerland). DMEM, high glucose, pyruvate for HeLa cells was from Thermo Fisher Scientific (product #11995065). Fetal Bovine Serum (FBS), Premium, US Sourced was from Corning (Corning, NY; product #45001-108). Penicillin-Streptomycin (10,000 U/mL) was from Thermo Fisher Scientific (product #15140122). Trypsin–EDTA (0.25% w/v) with phenol red was from Thermo Fisher Scientific (product #25200056). Cells were counted using a Countess II FL Automated Cell Counter (product #AMQAF1000) with Countess Cell Counting Chamber Slides from Thermo Fisher Scientific (product #C10283). mRNA was synthesized using the T7 ARCA mRNA Kit (with Tailing) kit from New England BioLabs (product #E2060S). Lipofectamine MessengerMAX (product #LMRNA001) was used for transient transfection of mammalian cells from Thermo Fisher Scientific. Differential interference contrast (DIC) and epifluorescent live cells images were acquired using an epifluorescent EVOS M7000 Imaging System (product #AMF7000) from Thermo Fisher Scientific. Protein conjugates were filtered with 0.22-µm Corning Costar Spin-X centrifuge tube filters from MilliporeSigma (product #CLS8160) to remove aggregates before cell treatment. IbiTreat (#1.5 polymer coverslip, tissue culture treated, sterilized) 8-well plates (product #80826) and 18-well plates (product #81816) from Ibidi (Fitchburg, WI) were used for live-cell imaging. DPBS with Ca^{2+}/Mg^{2+} (product #14040141) was from Gibco (Waltham, MA). DPBS without Ca^{2+}/Mg^{2+} (product #14190144) was from Gibco (Waltham, MA). FluoroBrite DMEM was from Thermo Fisher Scientific (product #A1896701). Life Technologies Attune NxT flow cytometer and SYTOX Blue Dead-Cell Indicator (product #S34857) from Thermo Fisher Scientific were used for flow cytometry.

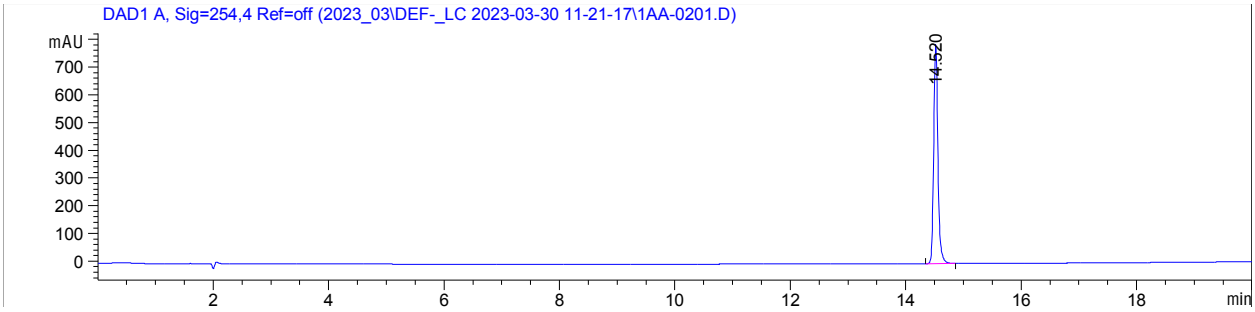
III. Synthesis of Virginia Orange-HaloTag Ligand (VO-HTL)

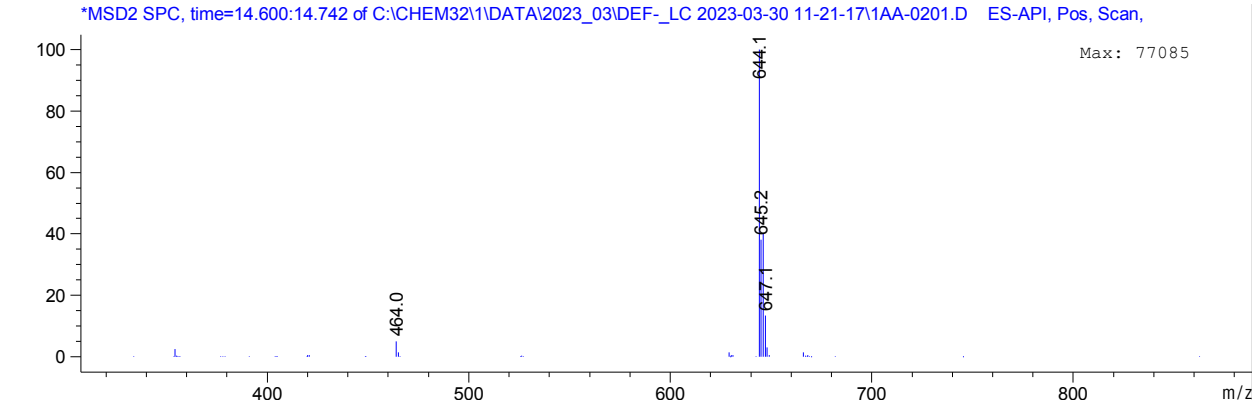


VO-HaloTag ligand. Ac_2VO-NHS[3] (8.0 mg, 12.9 µmol) and HaloTag(O2)amine (HTL-NH_2, TFA salt; 6.5 mg, 19.4 µmol, 1.5 equiv) were dissolved in DMF (1 mL). DIEA (6.7 µL, 38.7 µmol, 3 equiv) was added to the resulting solution. After stirring the reaction mixture for 2 h, MeOH (500 µL) and 1 M NaOH (50 µL) were added. The mixture was stirred for an additional 2 h, acidified with 1 M HCl (100 µL), diluted with water, and extracted with EtOAc (2×). The combined organic extracts were washed with brine, dried over anhydrous $MgSO_4$(s),

filtered, and concentrated *in vacuo*. Purification by silica gel chromatography (10–100% v/v EtOAc in CH_2Cl_2, linear gradient) afforded 7.5 mg (90%) of VO-HaloTag ligand.

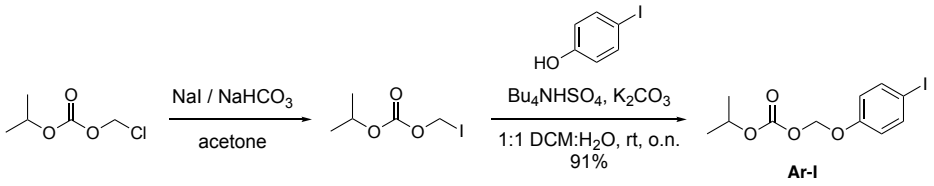
Physical state: off-white foam
^{1}H NMR ($CDCl_3$, 400 MHz, δ): 8.07 (d, J = 8.0 Hz, 1H), 7.97 (dd, J = 8.0, 1.4 Hz, 1H), 7.51 (d, J = 1.5 Hz, 1H), 7.18 (d, $^4J_{HF}$ = 8.6 Hz, 2H), 6.98 (t, J = 4.9 Hz, 1H), 6.32 (d, $^3J_{HF}$ = 11.3 Hz, 2H), 6.13 (s, 2H), 3.66–3.59 (m, 6H), 3.59–3.54 (m, 2H), 3.51 (t, J = 6.6 Hz, 2H), 3.42 (t, J = 6.7 Hz, 2H), 1.78–1.70 (m, 2H), 1.63 (s, 3H), 1.60 (s, 3H), 1.57–1.49 (m, 2H), 1.46–1.37 (m, 2H), 1.36–1.27 (m, 2H)
^{19}F{^{1}H} NMR ($CDCl_3$, 376 MHz, δ): −141.13 (dd, J_{FH} = 11.3, 8.6 Hz)
^{13}C{^{1}H} NMR ($CDCl_3$, 101 MHz, δ): 169.6 (C), 166.9 (C), 154.7 (C), 149.9 (d, $^1J_{CF}$ = 241.1 Hz, C), 145.2 (d, $^2J_{CF}$ = 13.7 Hz, C), 142.4 (d, $^4J_{CF}$ = 3.2 Hz, C), 141.0 (C), 128.7 (C), 128.4 (CH), 125.8 (CH), 123.4 (CH), 122.83 (d, $^3J_{CF}$ = 5.2 Hz, C), 115.5 (d, $^3J_{CF}$ = 2.6 Hz, CH), 114.4 (d, $^2J_{CF}$ = 19.1 Hz, CH), 86.4 (C), 71.4 (CH_2), 70.11 (CH_2), 70.06 (CH_2), 69.4 (CH_2), 45.1 (CH_2), 40.3 (CH_2), 37.8 (C), 35.1 (CH_3), 33.0 (CH_3), 32.6 (CH_2), 29.3 (CH_2), 26.7 (CH_2), 25.4 (CH_2)
Analytical HPLC: t_R = 14.5 min, >99% purity (10–95% v/v MeCN in H_2O, linear gradient with constant 0.1% v/v TFA; 20 min run; 1 mL/min flow; ESI; positive ion mode; λ = 254 nm)
HRMS (ESI): calcd for $C_{34}H_{37}ClF_2NO_7$ $[M + H]^+$ 644.2222, found 644.2224.



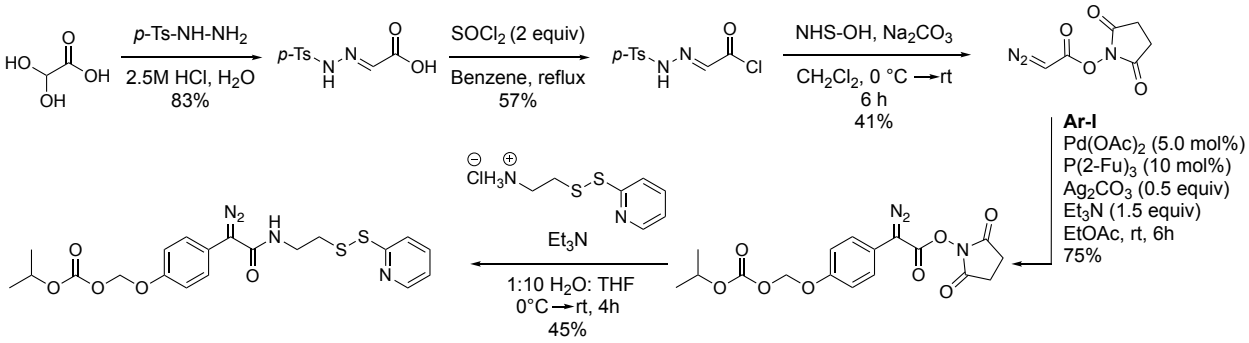


IV. Synthesis of Diazo Compound 2-SSpy

Synthetic Route to Aryl Iodide (Ar-I)

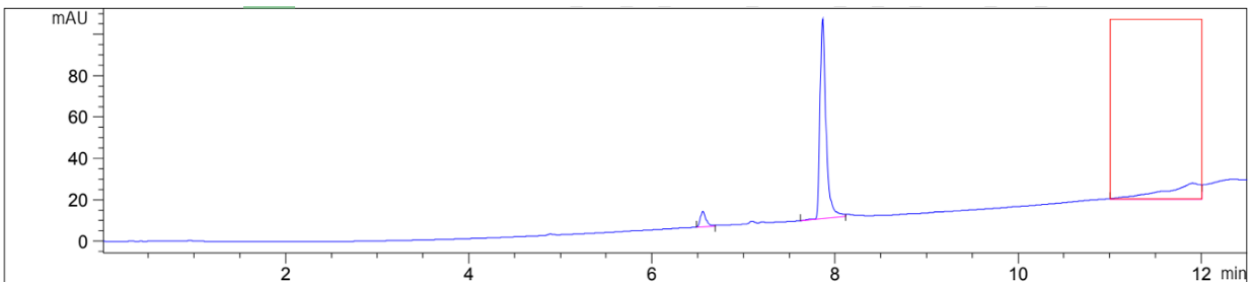


Synthetic Route to Diazo Compound 2-SSpy



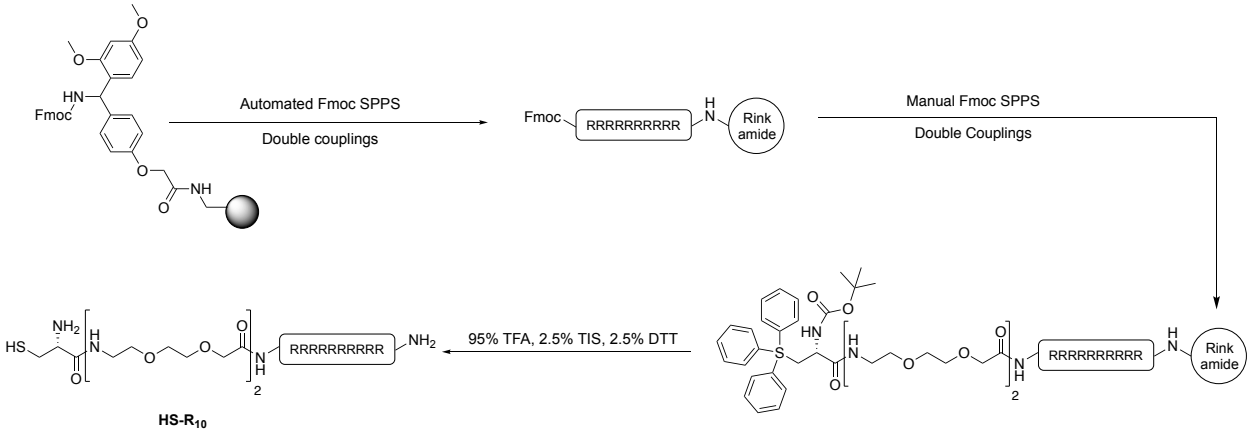
(4-(1-Diazo-2-oxo-2-((2-(pyridin-2-yldisulfaneyl)ethyl)amino)ethyl)phenoxy)methyl isopropyl Carbonate (2-SSpy). 2-SSpy was synthesized as described previously[4] to yield the title compound (54 mg, 0.12 mmol, 45% yield). Safety considerations regarding this diazo compound were likewise described previously.[4]

Physical State: Orange amorphous solid
TLC: $R_f = 0.34$ (1:1 hexanes/EtOAc)
^{1}H NMR (500 MHz, CD_3CN, δ): 8.22–8.09 (m, 1H), 7.73–7.63 (m, 2H), 7.46 (d, $J = 8.9$ Hz, 2H), 7.17–7.10 (m, $J = 8.9$ Hz, 3H), 6.89 (t, $J = 6.0$ Hz, 1H), 5.77 (s, 2H), 4.88 (s, 1H), 3.55 (s, 2H), 2.96 (s, 2H), 1.27 (d, $J = 6.3$ Hz, 6H). ^{1}H NMR matched the literature.[4]
HRMS (ESI–TOF): Calc'd for $C_{20}H_{23}N_2O_5S_2$ $[M - N_2 + H]^+$, 435.1043; found, 435.1047
LC-MS: product peak at ~8 min
10−95% v/v MeCN in water containing formic acid (0.1% v/v) over 10 min (UV trace at 250 nm)



V. Peptide Synthesis and Characterization

HS-R10



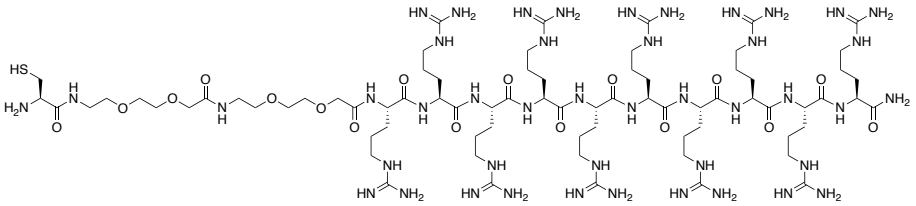
Synthesis. A linear R10 peptide with the sequence Fmoc-(Arg(Pbf))$_{10}$–NH_2 was synthesized with double-couplings of amino acid monomers. After the synthesis, the resin was transferred to a 24-mL polypropylene luer-lock syringe equipped with a filter frit, where further elaboration of the peptide was performed by hand according to traditional Fmoc-based methods. Fmoc-8-amino-3,6-dioxaoctanoic acid (4 equiv) was double coupled using HATU (4 equiv) and DIPEA (8 equiv) in 3 mL of DMF. Boc-Cys(Trt) (4 equiv) was performed using HCTU (4 equiv) and DIPEA (8 equiv) in 3 mL of DMF. Fmoc protecting groups were cleaved by treating the resin with a solution of 20% v/v methyl-piperidine in DMF (2 × 5 min each time). The resin was washed after deprotection steps and in between amino acid couplings using DMF (5×), DCM (5×), and DMF (5×).

Cleavage and Precipitation. Peptides were first washed with DMF (5×) followed by DCM (15×) and were cleaved from the resin for 5 h (7 mL mixture of 95% TFA, 2.5% TIS, 2.5% DTT). The resin was washed with an additional 4 mL of cleavage cocktail, and the pooled cleavage eluate was blown under a stream of N_2(g) to evaporate the cleavage cocktail. When the peptide had concentrated to a thick red oil, the peptide was precipitated in 50 mL of ice-cold diethyl ether. The peptide was pelleted by centrifugation for 10 min at 1500 RPM at 4 °C. The ether supernatant was decanted, and the crude peptide was stored at −70 °C until performing reversed-phase chromatographic purification.

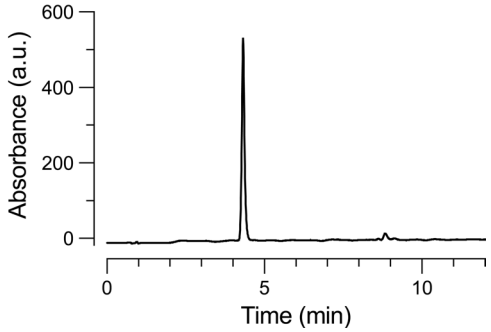
Purification. Reversed-phase liquid chromatography was performed by using a VP 250/21 Nucleosil 100-5 C18 column from Macherey–Nagel (Bethlehem, PA) and a 1260 Infinity II instrument from Agilent Technologies (Santa Clara, CA). The crude peptide was reconstituted in a minimal amount of ACN, passed through a 0.22-µm PFTE filter, and separated using a gradient of 5–95% v/v ACN in H_2O containing TFA (0.1% v/v). A second round of purification was performed from 5–20% v/v ACN in H_2O containing TFA (0.1% v/v) to remove additional impurities. Fraction purity was assessed by MALDI–TOF MS in positive mode using a microflex LRF instrument from Bruker (Billerica, MA) and a CHCA matrix. Pure fractions were pooled and lyophilized using a FreeZone benchtop instrument from Labconco (Kansas City, MO) to yield the peptides as a fluffy white TFA salt (39 mg, 12.1 µmol, 12.1% yield). The molar mass of the purified material was confirmed by MALDI–TOF MS (molar mass = 1972.38 Da, molar mass (TFA_{11}-salt) = 3226.64 Da). Final purity was assessed by RP-HPLC using an EC 250/4.6

Nucleosil 100-5 C18 column from Macherey–Nagel and a 1260 Infinity II instrument from Agilent Technologies.

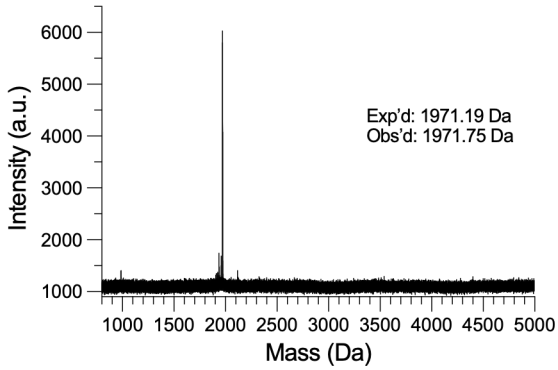
HS-R10 Peptide



Analytical RP-HPLC



MALDI–TOF MS



Analytical RP-HPLC of HS-R10

Purity was assessed with analytical RP-HPLC using a linear gradient from 5–25% B over 12 min at 2 mL/min (λ = 210 nm). Product was quantified as being >97% pure.

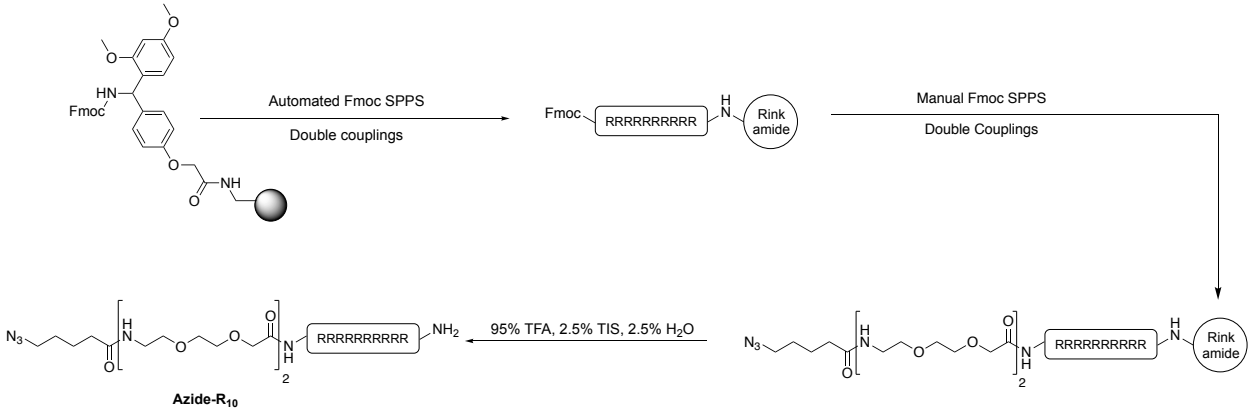
LC/MS Analysis of HS-R10

0–15% v/v B over 4 min

Exp'd *m*/*z*: $[M + 3H]^{3+}$: 658.3, $[M + 4H]^{4+}$: 494.0, $[M + 5H]^{5+}$: 395.4, $[M + 6H]^{6+}$: 329.7

Obs'd *m*/*z*: $[M + 3H]^{3+}$: 658.2, $[M + 4H]^{4+}$: 493.9, $[M + 5H]^{5+}$: 395.4, $[M + 6H]^{6+}$: 329.7

Azide-R10

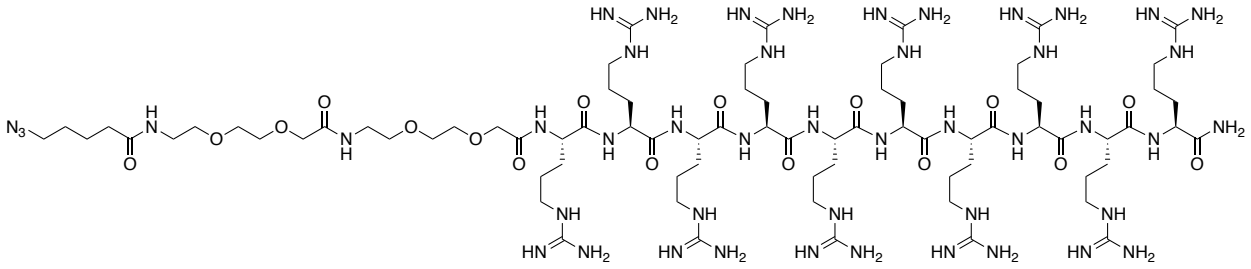


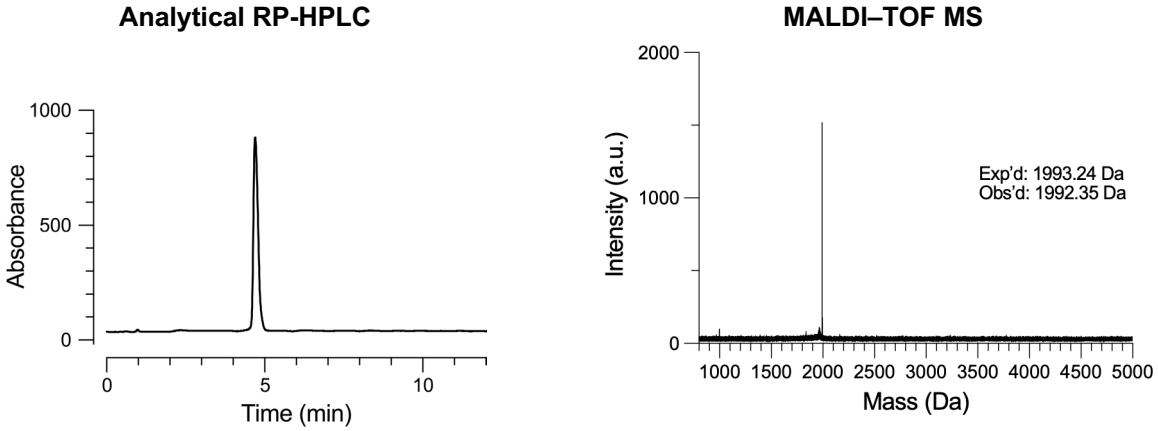
Synthesis. After the synthesis of the Fmoc-(Arg(Pbf))$_{10}$-NH$_2$ peptide as described above, the resin was transferred to a 24-mL polypropylene luer-lock syringe equipped with a filter frit, where further elaboration of the peptide was performed by hand according to traditional Fmoc-based methods. Fmoc-8-amino-3,6-dioxaoctanoic acid (4 equiv) was double coupled using HATU (4 equiv) and DIPEA (8 equiv) in 3 mL of DMF. A final Fmoc deprotection was performed, and subsequent double coupling using 5-azidopentanoic acid (5 equiv) was performed using HATU (5 equiv) and DIPEA (10 equiv) in 2 mL of DMF for 1 h at room temperature. Fmoc protecting groups were cleaved through treatment of the resin with a solution of 20% v/v *N*-methylpiperidine in DMF (2 × 5 min each time). The resin was washed after deprotection steps and in between amino acid couplings using DMF (5×), DCM (5×), and DMF (5×).

Cleavage and Precipitation. Peptides were first washed with DMF (5×) followed by DCM (15×) and cleaved from the resin for 3 h (6 mL mixture of 95% v/v TFA, 2.5% v/v TIS, 2.5% v/v water). The resin was washed with an additional 4 mL of cleavage cocktail, and the pooled cleavage eluate was blown under a stream of N_2(g) to evaporate the cleavage cocktail. When the peptide had concentrated to a thick brown oil, the peptide was precipitated in 50 mL of ice-cold diethyl ether. The peptide was pelleted by centrifugation for 10 min at 1500 RPM at 4 °C. The ether supernatant was decanted, and the crude peptide was stored at −70 °C until performing reversed-phase chromatographic purification.

Purification. Reversed-phase liquid chromatography was performed using a VP 250/21 Nucleosil 100-5 C18 column from Macherey–Nagel (Bethlehem, PA) and a 1260 Infinity II instrument from Agilent Technologies (Santa Clara, CA). The crude peptide was reconstituted in a minimal amount of ACN, passed through a 0.22 μm PFTE filter, and separated using a gradient of 5–35% ACN in H_2O containing TFA (0.1% v/v). Fraction purity was assessed by MALDI–TOF MS in positive mode using a microflex LRF instrument from Bruker (Billerica, MA) and a CHCA matrix. Pure fractions were pooled and lyophilized using a FreeZone benchtop instrument from Labconco (Kansas City, MO) to provide the peptide as a fluffy white TFA salt (17.8 mg, 5.48 μmol, 5.5% yield). The molar mass of the purified material was confirmed by MALDI–TOF MS (molar mass = 1994.37 Da, molar mass (TFA$_{11}$-salt) = 3248.63 Da). Final purity was assessed by RP-HPLC using an EC 250/4.6 Nucleosil 100-5 C18 column from Macherey–Nagel and a 1260 Infinity II instrument from Agilent Technologies.

Azide-R10 Peptide





Analytical RP-HPLC of Azide-R10

Purity was assessed with analytical RP-HPLC using a linear gradient from 5–25% B over 12 min at 2 mL/min (λ = 210 nm). Product was quantified as being >98% pure.

LC/MS Analysis of Azide-R10

Exp'd *m*/*z*: $[M + 3H]^{3+}$, 665.4; $[M + 4H]^{4+}$, 499.3; $[M + 5H]^{5+}$, 399.6; $[M + 6H]^{6+}$, 333.2
Obs'd *m*/*z* $[M + 3H]^{3+}$, 655.5; $[M + 4H]^{4+}$, 499.5; $[M + 5H]^{5+}$, 399.8; $[M + 6H]^{6+}$, 333.1

L17E

Synthesis. A peptide with the sequence:

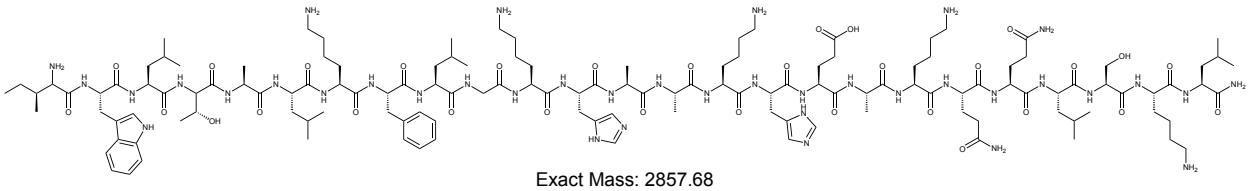
H-Ile-Trp(Boc)-Leu-Thr(tBu)-Ala-Leu-Lys(Boc)-Phe-Leu-Gly-Lys(Boc)-
His(Boc)-Ala-Ala-Lys(Boc)-His(Boc)-Glu(OtBu)-Ala-Lys(Boc)-Gln(Trt)-
Gln(Trt)-Leu-Ser(OtBu)-Lys(Boc)-Leu-NH_2

was synthesized with single-coupling of amino acid monomers.

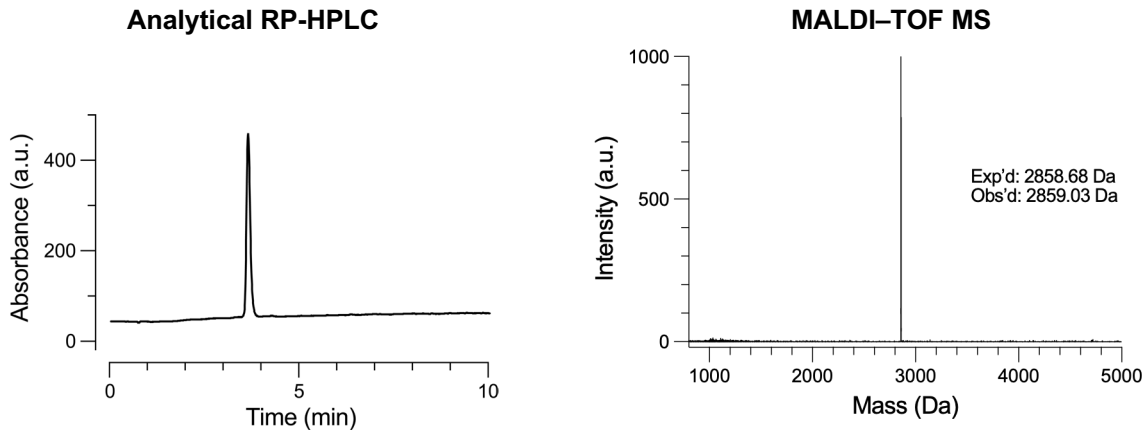
Cleavage and Precipitation. After the synthesis, the peptide on resin was transferred to a 24-mL polypropylene syringe equipped with a filter frit. After washing and drying the resin, the L17E peptide was deprotected and cleaved from solid support for 3 h in 8 mL of the cleavage cocktail (95% v/v trifluoroacetic acid (TFA), 5% v/v 1,2-ethanedithiol (EDT)). The resin then was washed with 4 mL of the cleavage cocktail. The flow throughs were combined and evaporated under a stream of N_2(g). The peptide was precipitated in 50 mL of diethyl ether twice, redissolved in 10 mL of 1:1 MeCN/water, and lyophilized overnight.

Purification. The peptide was purified via preparative reversed-phase HPLC (15–45% v/v MeCN (0.1% v/v TFA) in H_2O (0.1% v/v TFA) over 40 min) to yield a trifluoroacetate salt of L17E (27.2 mg, 7.7% yield, expected molecular weight of the L17E·TFA_6 salt = 3542.60026).

L17E Peptide



Exact Mass: 2857.68
Molecular Weight: 2859.47



Analytical RP-HPLC of L17E

Purity was assessed with analytical RP-HPLC using a linear gradient from 10–70% v/v acetonitrile in water containing TFA (0.1% v/v) over 10-min at 2 mL/min (λ = 218 nm). Product was quantified as being >99% pure by analytical RP-HPLC.

LC/MS Analysis of L17E

Exp'd *m*/*z* $[M + 4H]^{4+}$, 715.4; $[M + 5H]^{5+}$, 572.5; $[M + 6H]^{6+}$, 477.3; $[M + 6H]^{7+}$, 409.2
Obs'd *m*/*z* $[M + 4H]^{4+}$, 715.7; $[M + 5H]^{5+}$, 572.8; $[M + 6H]^{6+}$, 477.5; $[M + 6H]^{7+}$, 409.4

VI. Recombinant DNA

A commercially available, codon-optimized HaloTag vector for bacterial expression equipped with an N-terminal His-tag and a C-terminal cleavage site (Promega product #G7971) was purchased. The vector was linearized with EcoRI-HF and StyI-HF restriction enzymes (New England Biolabs) for 12 h at 37 °C followed by 65 °C heat inactivation for 20 min. A Monarch PCR and DNA Cleanup Kit (5 μg) was used according to manufacturer's instructions to isolate the linearized vector.

A gBlock was obtained from IDT with the following sequence for use in Gibson cloning:
cactatagggtattttaactttactaaggagaattcatcatgaaagcagaaatcggtactggctttccattcgacccccattatgtggaagtcctg
ggcgagcgcatgcactacgtcgatgttggtccgcgcgatggcacccctgtgctgttcctgcacggtaacccgacctcctcctacgtgtggc
gcaacatcatcccgcatgttgcaccgacccatcgctgcattgctccagacctgatcggtatgggcaaatccgacaaaccagacctgggttat
ttcttcgacgaccacgtccgcttcatggatgccttcatcgaagccctgggtctggaagaggtcgtcctggtcattcacgactggggctccgct
ctgggtttccactgggccaagcgcaatccagagcgcgtcaaaggtattgcatttatggagttcatccgccctatcccgacctgggacgaatg
gccagaatttgcccgcgagaccttccaggccttccgcaccaccgacgtcggccgcaagctgatcatcgatcagaacgtttttatcgagggta
cgctgccgatgggtgtcgtccgcccgctgactgaagtcgagatggaccattaccgcgagccgttcctgaatcctgttgaccgcgagccact
gtggcgcttcccaaacgagctgccaatcgccggtgagccagcgaacatcgtcgcgctggtcgaagaatacatggactggctgcaccagtc
ccctgtcccgaagctgctgttctggggcaccccaggcgttctgatcccaccggccgaagccgctcgcctggccaaaagcctgcctaactg
caaggctgtggacatcggcccgggtctgaatctgctgcaagaagacaacccggacctgatcggcagcgagatcgcgcgctggctgtcga
cgctcgagatttccggcgagccaaccactgaggatctgtactttcagagcTCCTCCGGGGTCGATTTGGGGcatcatcac
catcaccactaactagcataaccccttggggcctctaaacgggtcttgaggggt

The gene fragment was used directly in a Gibson Assembly by combining the Gibson Assembly Master Mix (New England Biolabs) and the linearized vector according to

manufacturer's instructions. After Gibson Assembly, the plasmid was transformed into KCM chemically competent DH10B cells and plated on ampicillin selective media. Colonies were isolated and overnight cultures were setup with ampicillin selective media. Cultures with bacterial growth were pelleted and miniprepped with the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) according to the manufacturer's recommendations. Samples were submitted for sequencing using QuintaraBio Basic Sequencing Services. DNA sequence alignment between sequencing results and the expected template were performed with SnapGene software (Dotmatics).

VII. Recombinant Protein Production

Transformation and Overexpression. Using a modified expression protocol that was previously reported,[5] HaloTag was expressed and purified. BL21-DE3 electrocompetent *E. coli* were transformed with the cloned HaloTag construct and grown overnight at 37 °C on 2% w/v agar plates of Luria–Bertani (Miller) medium supplemented with glucose (10% w/v) and ampicillin (100 μg/mL). Using a single, isolated colony, a 150-mL overnight starter culture of 10% w/v glucose and 100 μg/mL ampicillin in Luria–Bertani (Miller) growth medium was inoculated and shaken at 200 RPM. Four 1-L cultures of Luria–Bertani (Miller) growth medium supplemented with glucose (2% w/v), $MgSO_4$ (2 mM), and ampicillin (100 μg/mL) were inoculated with the starter culture to a starting OD_{600} = 0.05. Flasks were incubated in a floor shaker at 37 °C with shaking at 200 RPM until induction. Protein expression was induced at OD_{600} = 0.59–0.72 by the addition of IPTG to 0.25 mM after cooling at 4 °C for 45 min. Flasks were incubated at 18 °C overnight with shaking at 200 RPM. Cells were pelleted via centrifugation at 5000*g* for 20 min and stored at −70 °C until the time of purification. SDS–PAGE gel electrophoresis confirmed overexpression.

Lysis and Purification. Collected cells were suspended in 50 mM sodium phosphate buffer, pH 7.4, containing NaCl (300 mM) at 10 mL/g of pellet and lysed with a cell disrupter (Constant Biosystems). Lysate was collected on ice and clarified by centrifugation at 30,000*g* for 1 h at 4 °C, and the supernatant was subsequently passed through a 0.45-μm PES filter. Clarified cell lysate was loaded onto an ÄKTA Pure FPLC purification system, equipped with a HiTrap Talon Crude 5 mL Column (GE Healthcare) after equilibration with 50 mM sodium phosphate buffer, pH 7.4, containing NaCl (300 mM). The column was washed with 5 column volumes (CV) of 50 mM sodium phosphate buffer, pH 7.4, containing NaCl (300 mM) and imidazole (5 mM). The target protein was eluted using a linear gradient of 0–150 mM imidazole over 15 CV. Fractions containing the target protein were identified by UV absorbance, pooled, and subsequently treated with 10,000 units/mL of TEV protease (New England Biolabs) to remove the C-terminal His-tag. The digestion was monitored via Q–TOF MS and was complete after overnight incubation at 30 °C using standard buffer conditions. The tag-less target protein was loaded onto a HiTrap Talon Crude 5-mL column equilibrated with 50 mM sodium phosphate buffer, pH 7.4, containing NaCl (300 mM). HaloTag was collected in the flowthrough, while the TEV Protease and His-tag were retained on the column. The eluate was exchanged into a 25 mM TRIS, 150 mM NaCl, and 5 mM DTT buffer system at pH 7.5 using 15 mL, 10 kDa MWCO MilliporeSigma Amicon Ultra-15 Centrifugal Filter Unit. The target protein was purified further by gel-filtration chromatography using a High Resolution Superdex 75 pg 26/600 FPLC column equilibrated in a 25 mM Tris–HCl, pH 7.5, containing NaCl (150 mM) and DTT (5 mM). An

SDS–PAGE gel and Q–TOF MS confirmed the purity of the material, and the protein was flash-frozen in liquid nitrogen and stored at −70 °C until the time of use.

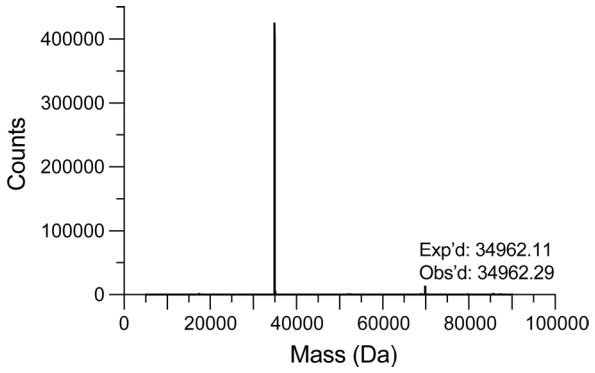


Figure S1. Deconvoluted Q–TOF MS spectrum of HaloTag.

VIII. Chemical Reactivity of HaloTag with a Ligand

The chemical reactivity of HaloTag with a ligand was validated by dye-labeling of the active site with JaneliaFluor dyes JF_{585}-HTL or VO-HTL. Janelia Fluor Dye stock solutions were reconstituted at 1 mM in DMSO. HaloTag was incubated with JF_{585}-HTL or VO-HTL (1.1 equiv) in a 25 mM Tris–HCl buffer, pH 7.5, containing NaCl (150 mM) and DTT (5 mM) at room temperature on a shaker overnight. Reaction progress was checked the following day by Q–TOF MS and was complete.

When using the HaloTag conjugate in subsequent experiments, excess dye was removed using Pierce Dye Removal Columns according to the manufacturer's recommendations. After active site labeling, the HaloTag conjugate was exchanged into the reaction buffer of the subsequent step using either an Amicon Ultra 0.5-mL 10K MWCO Centrifugal Filter Unit or a 0.5-mL 7K MWCO Zeba Spin Desalting Column according to manufacturer's recommendations. HaloTag conjugate concentrations were assessed by absorbance values of the JaneliaFluor dye via a DeNovix DS-11 Spectrophotometer/Fluorometer.

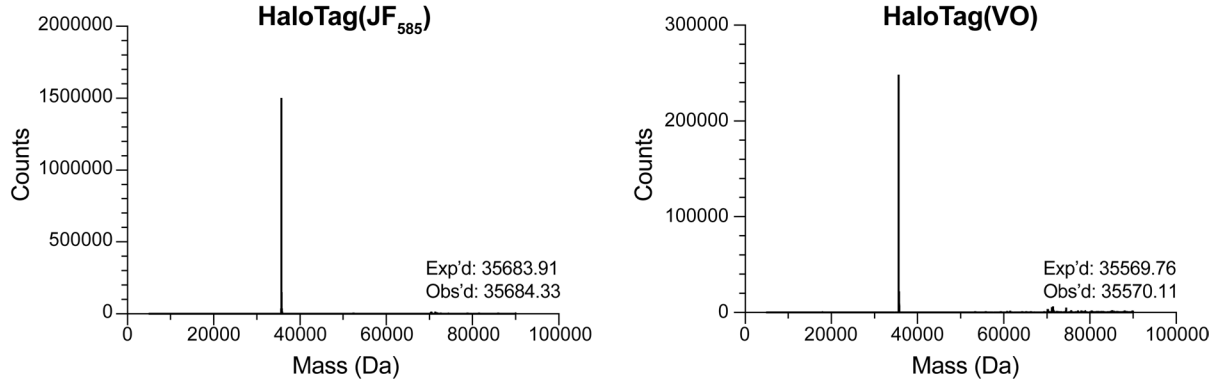


Figure S2. Representative deconvoluted Q–TOF mass spectra of HaloTag(JF_{585}) (left) and HaloTag(VO) (right). The observed masses indicate that labeling has gone to completion.

IX. Determination of Fluorophore pK_a Values of HaloTag(JF_{585}) and HaloTag(VO)

To determine the pK_a of JF_{585} and VO bound to the HaloTag active site, the HaloTag(JF_{585}) and HaloTag(VO) constructs were subjected to a pH titration in buffers composed of 150 mM NaCl and 10 mM buffer (*vide infra*). This experimental design was based on a previously designed assay used to determine the pK_a of VO.[6] The buffers for this experiment were citrate (pH 4.0–6.5), phosphate (pH 6.75–7.5), tris (pH 8–9), and carbonate (pH 9.5–10). A stock solution of both protein–dye conjugates was prepared at 1.25 μM in 150 mM NaCl. A 96-well half area, black flat bottom, non-binding plate (Corning) was used to assay the fluorescence. A working volume of 100 μL was used in each well, and protein-dye conjugates were diluted to a final concentration of 62.5 μM. Fluorescence readings were collected for HaloTag(JF_{585}) at $\lambda_{ex}/\lambda_{em}$ = 585 nm/609 nm and for HaloTag(VO) at $\lambda_{ex}/\lambda_{em}$ = 555 nm/581 nm on a Tecan Spark plate reader (Männedorf, Switzerland). A bandwidth of 5 nm was used to minimize fluorescence crosstalk between the excitation and emission. Gain values were optimized for the experiment using samples that were of the highest fluorescence intensity, and these optimal settings were applied to all subsequently collected data. Each value was collected in triplicate, and normalized to the average background fluorescence signal from buffer alone. Data were plotted using GraphPad Prism software and fitted to a normalized sigmoidal dose response curve with a variable slope to ascertain the Hill coefficient and pK_a of each protein-fluorophore construct.

X. Electrostatic Surface of HaloTag and Its Coulombic Repulsion from the Mammalian Cell Surface

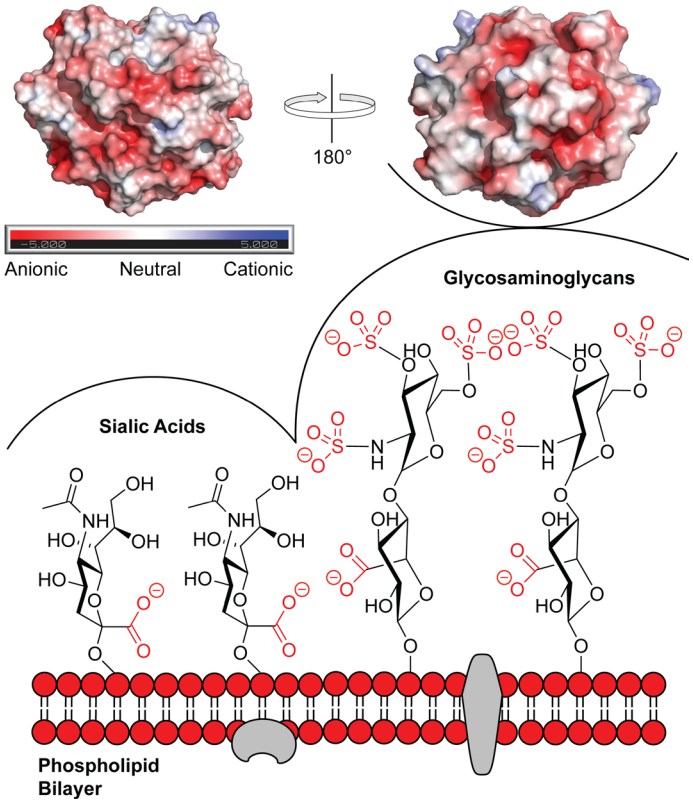


Figure S3. (Top) Depiction of the electrostatic potential map of HaloTag and the mammalian cell surface. Space-filling model, rotated about a 180° axis, demonstrating the anionic character of the HaloTag surface. The image was created with the APBS Electrostatics Plugin of PyMOL software (Schrödinger) and PDB entry 5y2y. (Bottom) Representation of the glycocalyx, decorated with sialic acids and glycosaminoglycans, and the Coulombic barrier between cellular uptake of HaloTag.

XI. Electrostatic Surface of HaloTag

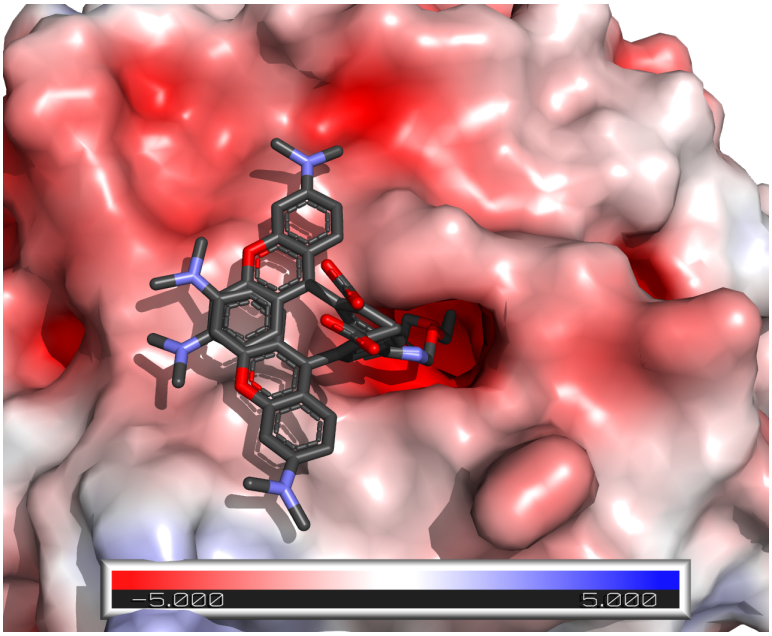
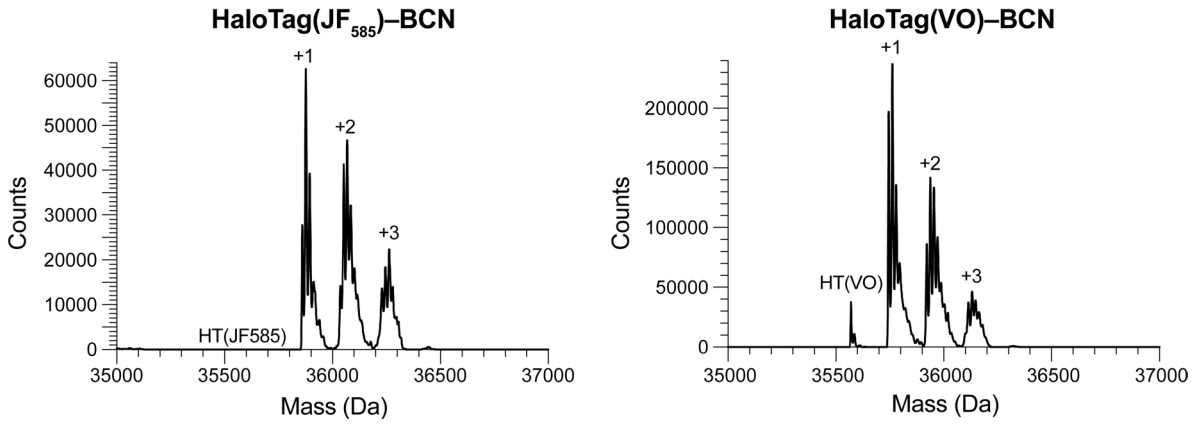


Figure S4. Electrostatic model of HaloTag near a covalently bound tetramethylrhodamine ligand. The fluorophore moiety is in two distinct conformations. The image was created with the APBS Electrostatics Plugin of PyMOL software (Schrödinger) and PDB entry 6u32.[7] Red, anionic surface; gray, neutral surface; blue, cationic surface. The environment around the bound dye is predicted to be a highly anionic.

XII. Protein Modification

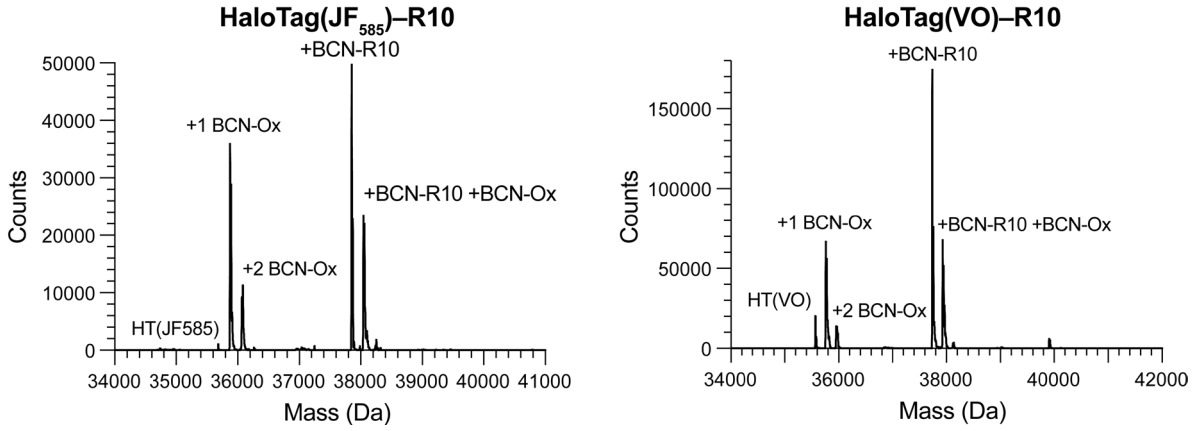
HaloTag–R10

To conjugate the azide-containing peptide to the HaloTag–dye protein conjugate, lysine residues were reacted with click reagent cross linker NHS–BCN ((1*R*,8*S*,9*s*)-bicyclo[6.1.0]non-4-yn-9-ylmethyl *N*-succinimidyl carbonate) from Sigma–Aldrich (product #744867). HaloTag–dye conjugate was exchanged into DPBS with Ca^{2+}/Mg^{2+} at 40 to 50 μM. A 50 mM stock of the NHS–BCN reagent was prepared in DMSO, and 25 equiv of the reagent was added to the labeling reaction. The reaction mixture was incubated overnight under agitation. Reaction progress was monitored by Q–TOF MS. After the overnight incubation, the reaction was complete, and excess NHS–BCN reagent was removed using a 0.5-mL 7K MWCO Zeba Spin Desalting Column according to the manufacturer's recommendations. To prepare the HaloTag(JF_{585}/VO)–R10 conjugate, the BCN-equipped material was subsequently incubated with the azide–R10 peptide. A 10 mM stock of the R10–azide peptide was prepared in water. The solution azide–R10 peptide (25 equiv) was neutralized by the addition of 100 mM sodium phosphate buffer, pH 8.0, to a final peptide concentration of 450 mM, and the resulting solution was added to the BCN-equipped protein. The reaction mixture was incubated for 24 h under agitation, and reaction progress was monitored by Q–TOF MS. After this incubation, a distribution of 0–2 R10 labels were present on the protein, and excess azide–R10 was removed with a 0.5-mL 7K MWCO Zeba Spin Desalting Column according to the manufacturer's recommendations and exchanged into DPBS with Ca^{2+}/Mg^{2+}.



+177.22 Da/label		HaloTag–dye	1 label	2 labels	3 labels
HaloTag(JF585)–BCN	Exp'd mass	35683.91	35861.13	36038.35	36215.57
	Obs'd mass	35684.33	35860.95	36036.61	36213.31
HaloTag(VO)–BCN	Exp'd mass	35569.76	35746.98	35924.2	36101.42
	Obs'd mass	35570.11	35748.55	35923.08	36099.10

Figure S5. Representative deconvoluted Q–TOF mass spectra of HaloTag(JF_{585})–BCN and HaloTag(VO)–BCN.

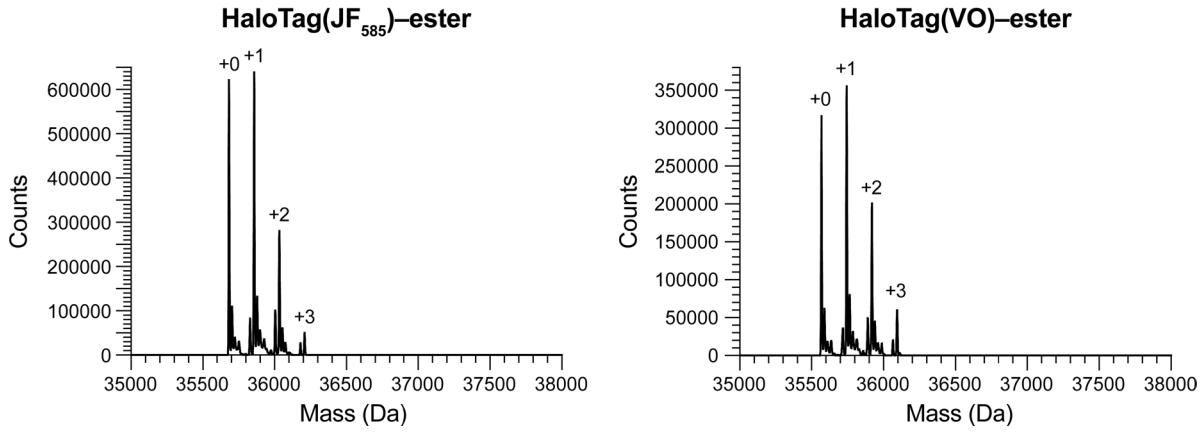


+2171.59 Da/label		1BCN-R10	2 labels	3 labels
HaloTag(JF_{585})–R10	Exp'd mass	37855.50	40027.09	42198.68
	Obs'd mass	37855.26	—	—
HaloTag(VO)–R10	Exp'd mass	37741.35	39912.94	42084.53
	Obs'd mass	37740.78	—	—

Figure S6. Representative Q–TOF mass spectra of HaloTag(JF_{585})–R10 and HaloTag(VO)–R10. We observe that several BCN moieties remain unreacted with the azide-containing peptide due to hydrolysis or oxidation, as had been observed previously.[4]

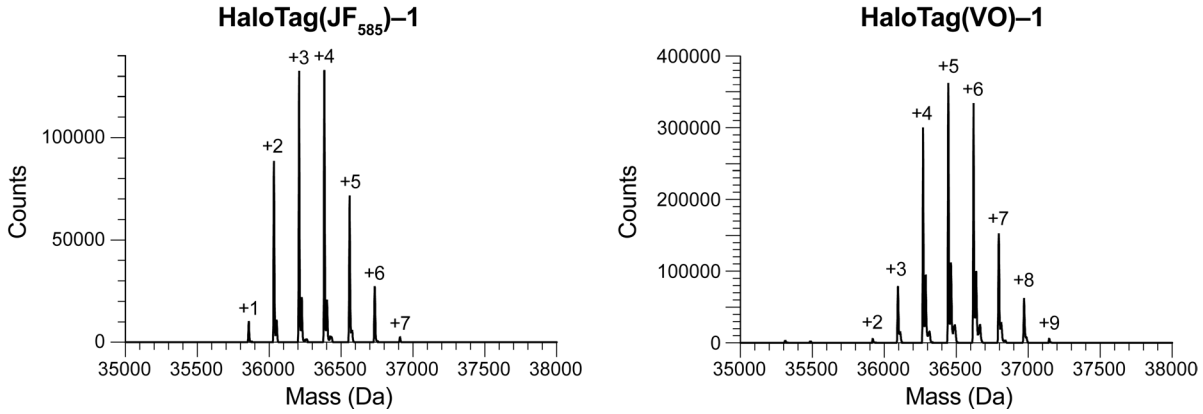
HaloTag–1

HaloTag(JF_{585}) or HaloTag(VO) was exchanged into a 10 mM Bis-Tris buffer, pH 6.5, containing NaCl (500 mM). A stock solution of diazo compound **1** was prepared (161 mM in ACN), and 25 equiv (low ester condition, Figure S7) or 100 equiv (high ester condition, Figure S8) were diluted with ACN such that the final ACN concentration in the reaction mixture was 10% v/v. After the addition of the diluted diazo compound, the reaction mixture was incubated overnight under agitation. Reaction progress was monitored by Q–TOF MS. The protein conjugate was exchanged into DPBS with Ca^{2+}/Mg^{2+} with an Amicon Ultra 0.5-mL 10K MWCO Centrifugal Filter Unit.



+175.23 Da/label		Unlabeled	1 label	2 labels	3 labels
HaloTag(JF_{585})–**1**	Exp'd mass	35683.91	35859.14	36034.37	36209.60
	Obs'd mass	35684.23	35859.28	36034.46	36209.58
HaloTag(VO)–**1**	Exp'd mass	35569.76	35744.99	35920.22	36095.45
	Obs'd mass	35570.16	35745.24	35920.32	36095.52

Figure S7. Representative deconvoluted Q–TOF mass spectra of low-labeled HaloTag(JF_{585})–**1** and HaloTag(VO)–**1**.

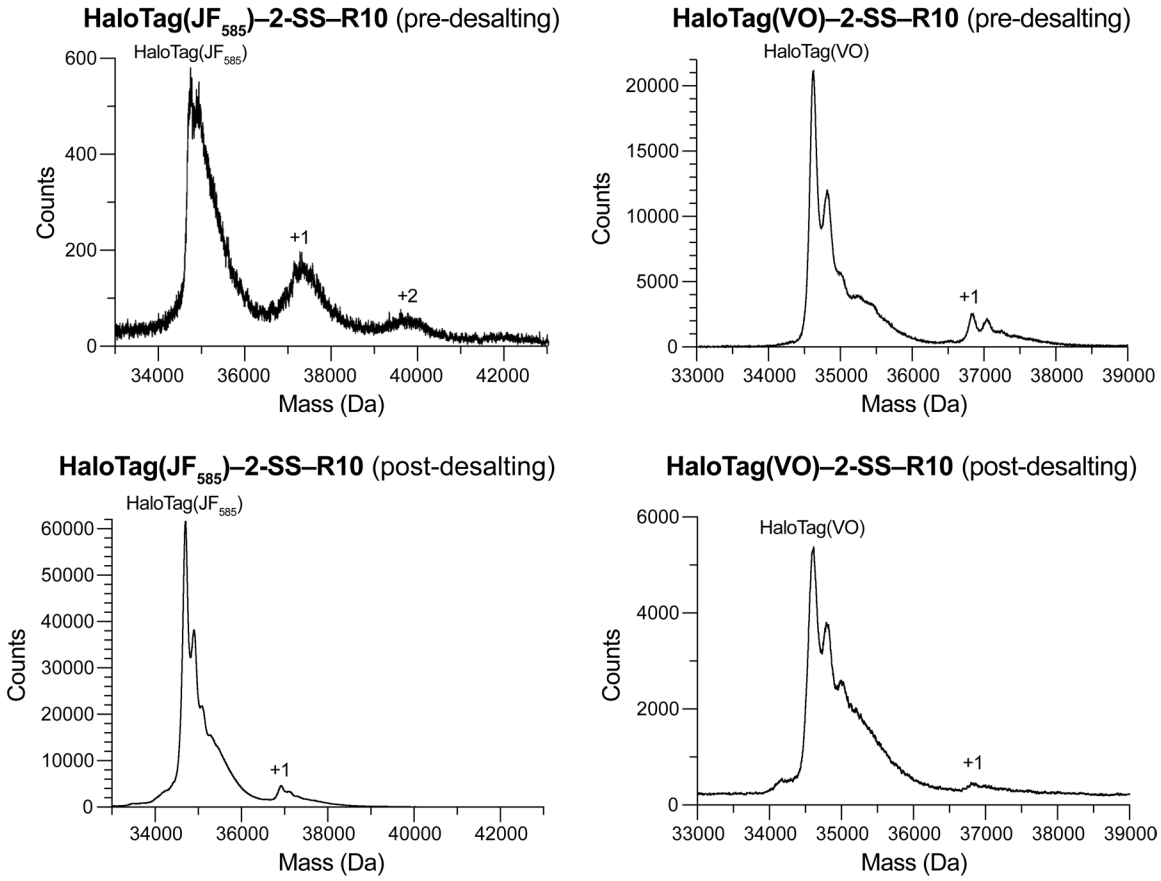


+175.23 Da/label		1 label	2 labels	3 labels	4 labels	5 labels	6 labels	7 labels	8 labels	9 labels
HaloTag (JF_{585})–**1**	Exp'd mass	35859.14	36034.37	36209.6	36384.83	36560.06	36735.29	36910.52	—	—
	Obs'd mass	35859.31	36034.60	36209.72	36384.96	36560.78	36735.73	36911.33	—	—
HaloTag (VO)–**1**	Exp'd mass	—	35920.22	36095.45	36270.68	36445.91	36621.14	36796.37	36971.6	37146.83
	Obs'd mass	—	35921.31	36096.36	36270.91	36446.12	36621.55	36797.17	36972.85	37147.90

Figure S8. Representative deconvoluted Q–TOF mass spectra of high-labeled HaloTag(JF_{585})–**1** and HaloTag(VO)–**1**.

HaloTag–2-SS–R10

2-SSpy (100 equiv, 100 mM stock in ACN) and HS-R10 (100 equiv, 10 mM stock in water) were premixed for 1 min. After pre-formation of a mixed disulfide between the **2**-SSpy and the HS-R10, the reaction mixture was added to HaloTag(JF_{585}) or HaloTag(VO) (1 equiv, 40 μM) in 50 mM Bis-Tris buffer, pH 6.5, containing NaCl (500 mM). The volume of acetonitrile added from the premixing reaction was 10% of the total reaction volume. The reaction mixture was incubated with agitation for 1.5–2 h, and reaction progress was monitored by MALDI–TOF MS. Excess **2**-SSpy and HS-R10 were removed from the reaction mixture with a 0.5-mL 7K MWCO Zeba Spin Desalting Column according to the manufacturer's recommendations and exchanged into 50 mM sodium acetate buffer, pH 4.5.



HaloTag(JF_{585})–2-SS–R10

R = R10	Observed Mass	Δ (Observed mass − Peak 1 mass)	1971 Da/label
Peak 1	34946	0	0
Peak 2	37318	2372	1.2
Peak 3	39596	4650	2.4

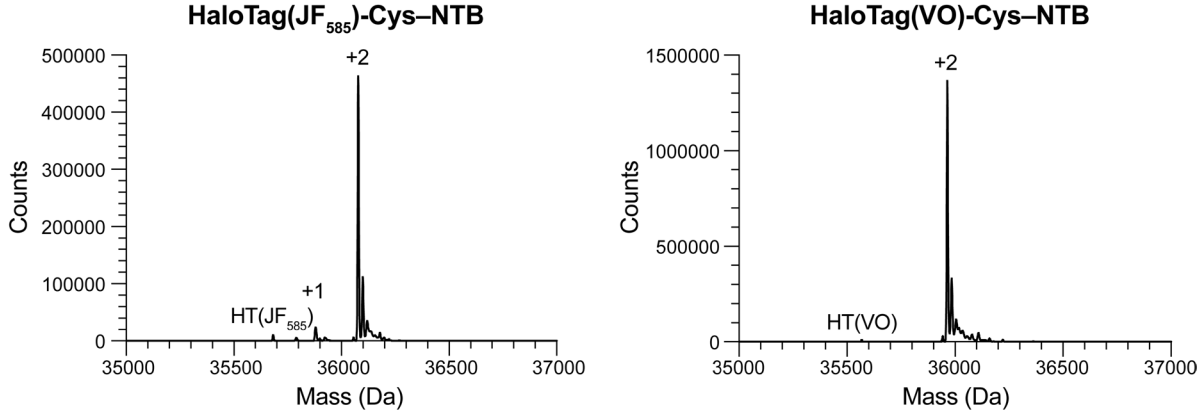
HaloTag(VO)–2-SS–R10

R = R10	Observed Mass	Δ (Observed mass − Peak 1 mass)	1971 Da/label
Peak 1	34634	0	0
Peak 2	36872	2238	1.1

Figure S9. Representative MALDI–TOF spectra of HaloTag(JF_{585})–**2**-SS–R10 and HaloTag(VO)–**2**-SS–R10 reactions.

HaloTag-Cys–R10

To conjugate the cysteine-containing peptide to the HaloTag(JF_{585}) protein conjugate, cysteine residues were first activated with Ellman's reagent (5,5′-dithiobis(2-nitrobenzoic acid)). HaloTag(JF_{585}) or HaloTag (VO) was exchanged into DPBS with Ca^{2+}/Mg^{2+} at 110 μM. A 50 mM stock of Ellman's reagent was prepared in a 9:1 solution of 0.1 M sodium phosphate buffer, pH 8.0, and DMSO, and the reagent was added to the labeling reaction at a final concentration of 1.12 mM (10 equiv). The reaction mixture was incubated for 2 h at room temperature under agitation. Within 10 s, the reaction progress could be assessed visually by the appearance of yellow color. Reaction progress was monitored by Q–TOF MS. After 2 h, the reaction was complete, and excess Ellman's reagent was removed using a 0.5mL 7K MWCO Zeba Spin Desalting Column according to product recommendations. To prepare the HaloTag(JF_{585})–R10 or the HaloTag(VO)–R10 conjugate, the NTB-activated material was subsequently incubated with the HS-R10 peptide. A 10 mM stock of the HS-R10 peptide was prepared in water. The HS-R10 peptide (20 equiv) was basified by the addition of 3.0 M Tris base to a final concentration of 450 mM, and the resulting solution was added to the NTB-activated protein. The reaction mixture was incubated for 10–20 min under agitation, and reaction progress was monitored by MALDI–TOF MS. After incubation, the hydrolysis of the NTB labels was observed, rather than installation of the HS-R10 peptide on the protein surface.



+197.17 Da/label		HaloTag–dye	1 label	2 labels
HaloTag(JF_{585})-Cys–NTB	Exp'd mass	35863.91	35881.08	36078.25
	Obs'd mass	35684.33	35881.22	36078.56
HaloTag(VO)-Cys–NTB	Exp'd mass	35569.76	35766.93	35964.10
	Obs'd mass	35570.11	—	35964.20

Figure S10. Representative deconvoluted Q–TOF mass spectra of HaloTag(JF_{585})-Cys–NTB and HaloTag(VO)-Cys–NTB.

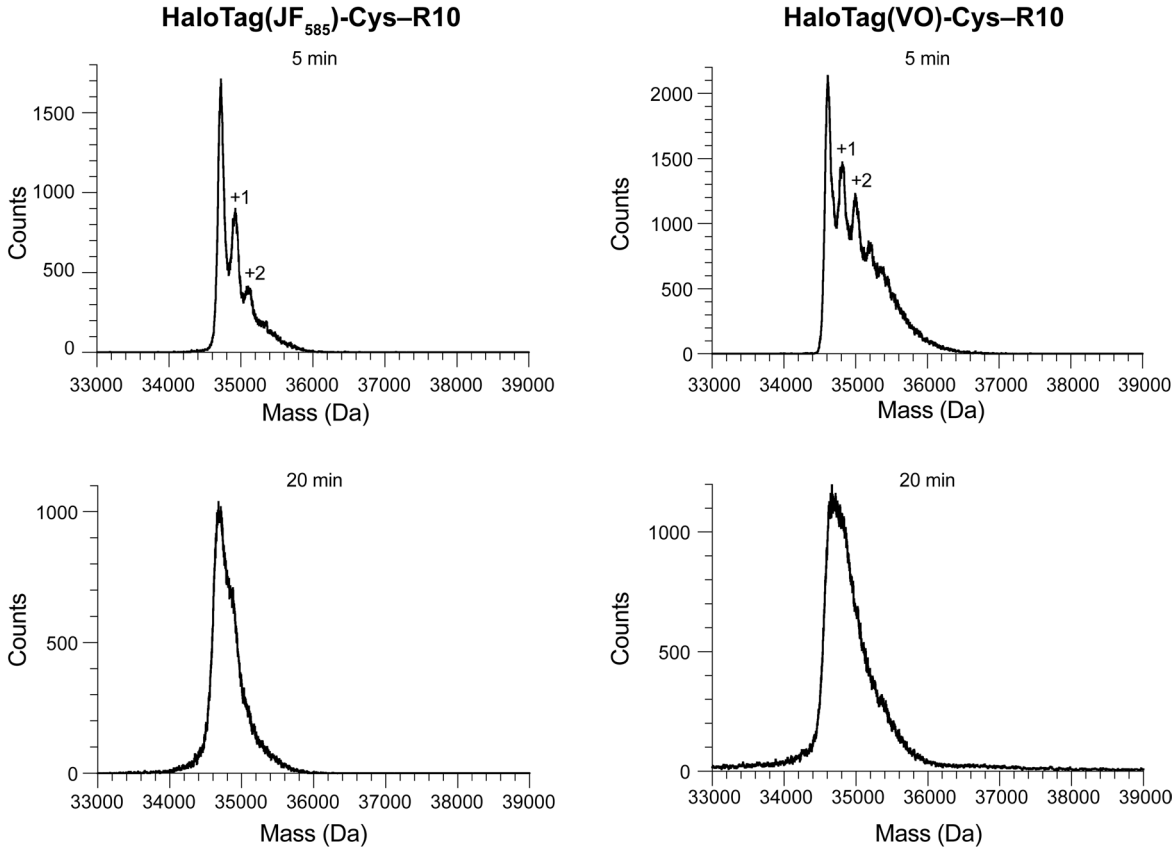


Figure S11. Representative MALDI–TOF spectra of HaloTag(JF_{585})-Cys–R10 and HaloTag(VO)-Cys–R10 reactions after 5 and 20 min. Data indicate the lability of the NTB label, labeled as +1 and +2 in each conjugate at 5 min.

XIII. Mammalian Cell Culture

HeLa cells (ATCC product #CCL-2) were cultured according to ATCC guidelines. HeLa cells were cultured in DMEM supplemented with FBS (10% v/v), penicillin (100 units/mL), and streptomycin (100 μg/mL). Cells were cultured in an incubator maintained at 37 °C and humidified to 5% v/v CO_2. Cells tested negative for mycoplasma.

XIV. Transient Transfection

HeLa cells were transiently transfected with a cytosolically-localized variant of HaloTag to confirm that both JF_{585} and VO fluoresce in the cytosol. mRNA was synthesized using HiScribe T7 ARCA mRNA Kit (with Tailing) (New England BioLabs) according to manufacturer recommendations with a HaloTag DNA template (derived from Promega product #G6591), linearized downstream of the HaloTag gene and upstream of the poly(A) tail.

HeLa cells were seeded in sterile 18-well IbiTreat dish (Ibidi) and were confluent at the time of transfection. Cells transfected with 0.15 μL/well of Lipofectamine MessengerMAX (Thermo

Fisher Scientific) and 100 ng/well mRNA according to the manufacturer's recommendations. After transfection, cells were incubated for 16–24 h to allow for adequate protein expression. Transfected cells were washed with DPBS without Ca^{2+}/Mg^{2+} (2 × 100 μL) and subsequently treated with a 1 μM concentration of JF_{585}-HTL or VO-HTL incubated overnight at 37 °C in a humidified incubator at 5% v/v CO_2. A sample of un-transfected cells was treated with either JF_{585}-HTL or VO-HTL as a negative control for both imaging and flow cytometry experiments. After incubation with the HTL-equipped fluorophore, cells were washed with DPBS without Ca^{2+}/Mg^{2+} (2 × 100 μL) and 100 μL of full DMEM was added to cells for a minimum of 15 min to allow for unbound dye to wash out. After the washout period, cells were washed with DPBS without Ca^{2+}/Mg^{2+} (2 × 100 μL) and Fluorobrite DMEM (100 μL) was used for imaging. Cells were imaged using an Evos M7000 Epifluorescent Microscope from Thermo Fisher Scientific.

After imaging experiments, cells were washed with DPBS without Ca^{2+}/Mg^{2+} (2 × 100 μL) and 100 μL of full DMEM was added to cells for a minimum of 30 min to allow cells to recover prior to flow cytometry analysis. To prepare cells for flow cytometry, cells were washed with DPBS without Ca^{2+}/Mg^{2+} (2 × 200 μL) and lifted from the plate with 50 μL of 0.25% v/v trypsin–EDTA. Trypsin was quenched by the addition of 50 μL of full medium, and cells were strained through netted caps into flow tubes (Falcon product #352054). Cells were pelleted by centrifugation for 5 min at 1200 RPM at 4 °C, and pellets were resuspended in 1 mL of ice-cold DPBS with Ca^{2+}/Mg^{2+} supplemented with bovine serum albumin (0.1% w/v). Each sample was stained with SYTOX Blue Dead-Cell Indicator (1 μL of a 1.0 mM stock) for at least 10 min on ice protected from light. The fluorescence intensity of at least 10,000 events was measured by flow cytometry with an Attune NxT Flow Cytometer (405 nm, 488 nm, 561 nm, and 640 nm lasers, Thermo Fisher Scientific). Control cells treated with SYTOX Blue, followed by untransfected cells treated with JF_{585}-HTL or VO-HTL, were analyzed first to set gates and laser intensities. For transfected cells treated with JF_{585}-HTL, the 561-nm laser was used for excitation and the 620/15 filter was used to detect fluorescence. For transfected cells treated with VO-HTL, the 561-nm laser was used for excitation and the 585/16 filter was used to detect fluorescence. Events were collected using standardized laser intensity values. Data were analyzed using the FlowJo software package (FlowJo LLC). The median fluorescence intensity (MFI) of live, single cells is reported.

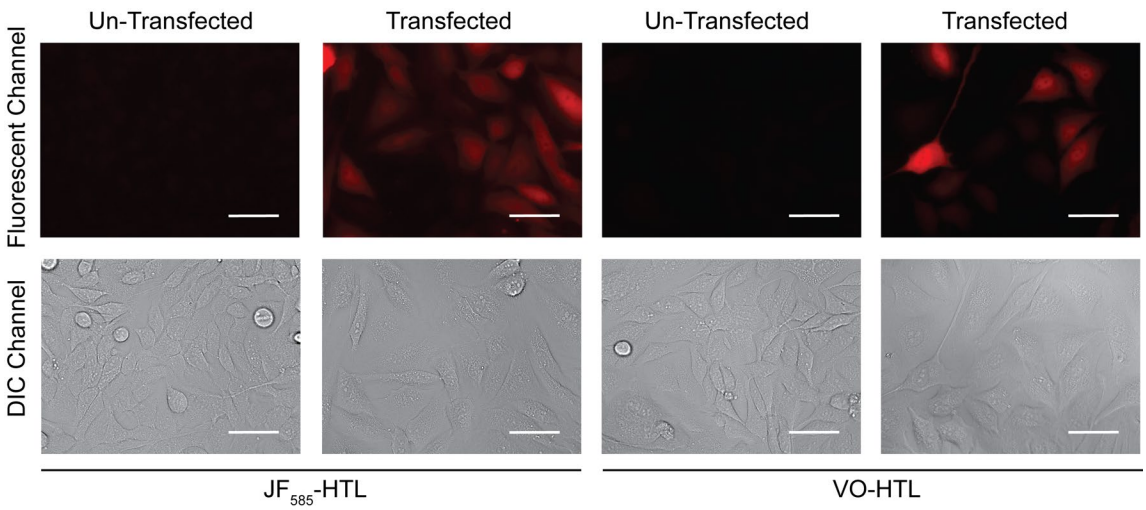


Figure S12. Fluorescence microscopy images of HeLa cells transfected to express cytosolic HaloTag and pulse-chased with JF_{585}-HTL or VO-HTL, and untransfected cells treated with JF_{585}-HTL and VO-HTL (negative controls) at 37 °C. JF_{585}: λ_{ex} = 585/29 nm, λ_{em}= 628/32 nm (Texas Red channel). VO: λ_{ex} = 542/20 nm, λ_{em} = 593/40 nm (RFP channel). Scale bars, 50 μm. All images were acquired with identical laser settings.

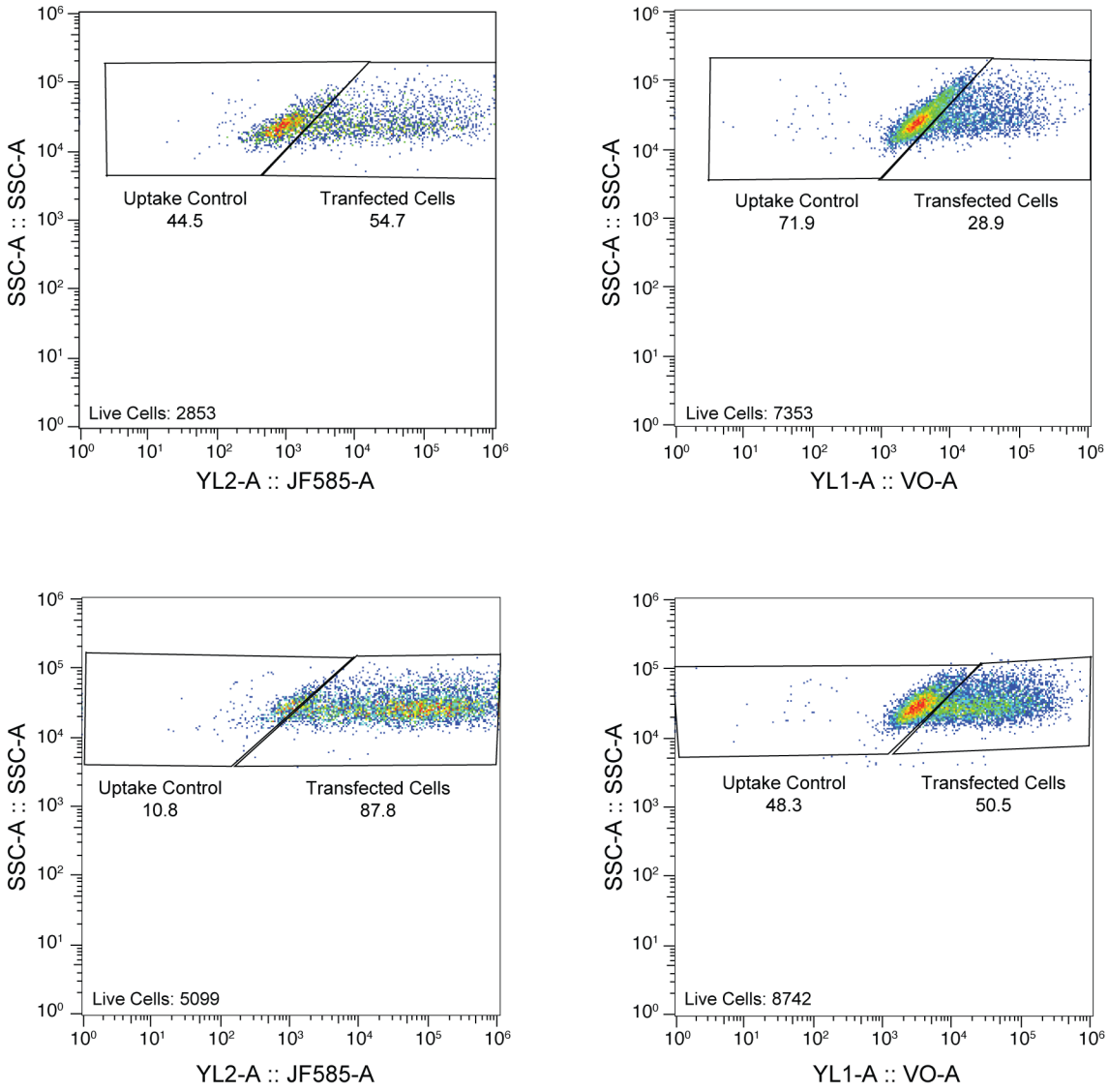


Figure S13. Flow cytometry gating to quantitate the cytosolic localization of HaloTag(JF_{585}) (left panels) and HaloTag(VO) (right panels) into transiently transfected HeLa cells. Results are shown for individual samples from a representative experiment performed in biological duplicate.

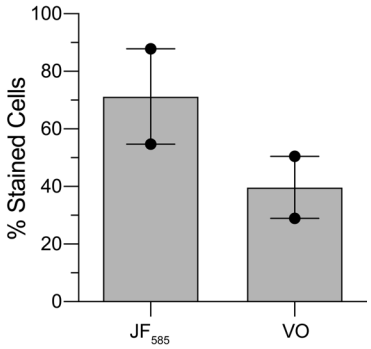


Figure S14. Bar graph of flow cytometry data on the cytosolic localization of HaloTag in transiently transfected HeLa cells. Data for each fluorophore are from two replicate samples from one representative experiment. In cell culture medium, JF_{585} is neutral but VO is largely a dianion and thus less permeable.

XV. Epifluorescence Microscopy

Cells were seeded to be 90% confluent at the time of the experiment. Specifically, cells were seeded at 36,000 cells/well if performing the experiment 24 h later; cells were seeded at 18,000 cells/well if performing the experiment 48 h later; cells were seeded at 9,000 cells/well if performing the experiment 72 h later. In each case=, cells were seeded into a sterile 18-well IbiTreat dish (Ibidi). Prior to treatment, cells were washed with DPBS without Ca^{2+}/Mg^{2+} (2 × 100 µL), and serum-free medium was added to each well. All protein conjugates were sterile-filtered with a 0.22-µm Corning Costar Spin-X centrifuge tube filter, dosed into each well according to the conditions below (Table S1), and incubated at 37 °C in a humidified incubator at 5% v/v CO_2. The volume of DPBS in each treatment condition did not exceed 25% of the medium volume, and the final volume of each well after the addition of protein was equal to 100 µL. After incubation with the protein conjugate, cells were washed with DPBS without Ca^{2+}/Mg^{2+} (2 × 100 µL) and Fluorobrite DMEM (100 µL) was used for imaging. Cells were protected from light at room temperature until imaged. Epifluorescent imaging was performed using an Evos M7000 Epifluorescent Microscope from Thermo Fisher Scientific. For HaloTag(JF_{585}) conjugates, the Texas Red light cube (λ_{ex} = 585/29 nm, λ_{em}= 628/32 nm) was used for excitation. For HaloTag(VO) conjugates, the RFP light cube (λ_{ex} = 542/20 nm, λ_{em} = 593/40 nm) was used for excitation. Images were collected using standardized laser intensity values. Images were analyzed using the open-source Fiji distribution of ImageJ, adjusting for brightness and contrast, and processing was identically applied to all fluorescence images collected in a session.

Sample Preparation and Treatment Conditions: The concentrations of each conjugate were determined via an absorbance reading of the corresponding fluorophore (JF_{585}, 585 nm; VO, 555 nm) or 205 nm using a DeNovix DS-11 Spectrophotometer/Fluorimeter. For JF_{585}, the source information[8] was used to derive a molar absorptivity coefficient of 120,000 M^{-1} cm^{-1}. For VO, the reported value of 90,900 M^{-1} cm^{-1} was used.[6]

When constructs were exchanged into neutral buffer, absorbance measurements of the corresponding fluorophore were used to calculate concentration. When constructs were exchanged into acidic buffer, we observed a slight dependence of the absorbance of HaloTag(VO) conjugates on buffer acidity. To ensure consistency in acidic conditions, the absorbance at 205 nm was used to calculate concentrations of both HaloTag(JF_{585})–**2**-SS–R10 and HaloTag(VO)–**2**-SS–R10 conjugates. Molar absorptivity coefficients for HaloTag(JF_{585}) and HaloTag(VO) were derived at 205 nm using a known concentration of each conjugate. Molar absorptivity coefficients at 205 nm for HaloTag(JF_{585}) and HaloTag(VO) were calculated to be 2,991,569 M^{-1} cm^{-1} and 3,868,854 M^{-1} cm^{-1}, respectively. These constants were used to calculate concentrations of HaloTag(JF_{585})–**2**-SS–R10 and HaloTag(VO)–**2**-SS–R10 in 50 mM sodium acetate buffer, pH 4.5.

Table S1. Stability and Treatment Conditions for HaloTag Conjugates

Construct	Figure No.	Shelf Life	Vehicle	Treatment Concentration (μM)	Treatment Time	Recovery Time
HaloTag–R10	S15	2 weeks	DPBS with Ca^{2+}/Mg^{2+}	2.5	1.5 h	1 h
HaloTag–**1** (high)	S16	2–3 days	DPBS with Ca^{2+}/Mg^{2+}	10	2 h	30 min
HaloTag–**1** (low)	S19	2–3 days	DPBS with Ca^{2+}/Mg^{2+}	5 or 15	2 h	30 min
HaloTag–**2**-SS–R10	S17	1 day	50 mM NaOAc, pH 4.5	1.0–1.5	1.5 h	1 h
HaloTag + L17E	S18	N/A	DPBS with Ca^{2+}/Mg^{2+}	HaloTag: 5 or 15	5–10 min	1 h
			H_2O	L17E: 40		

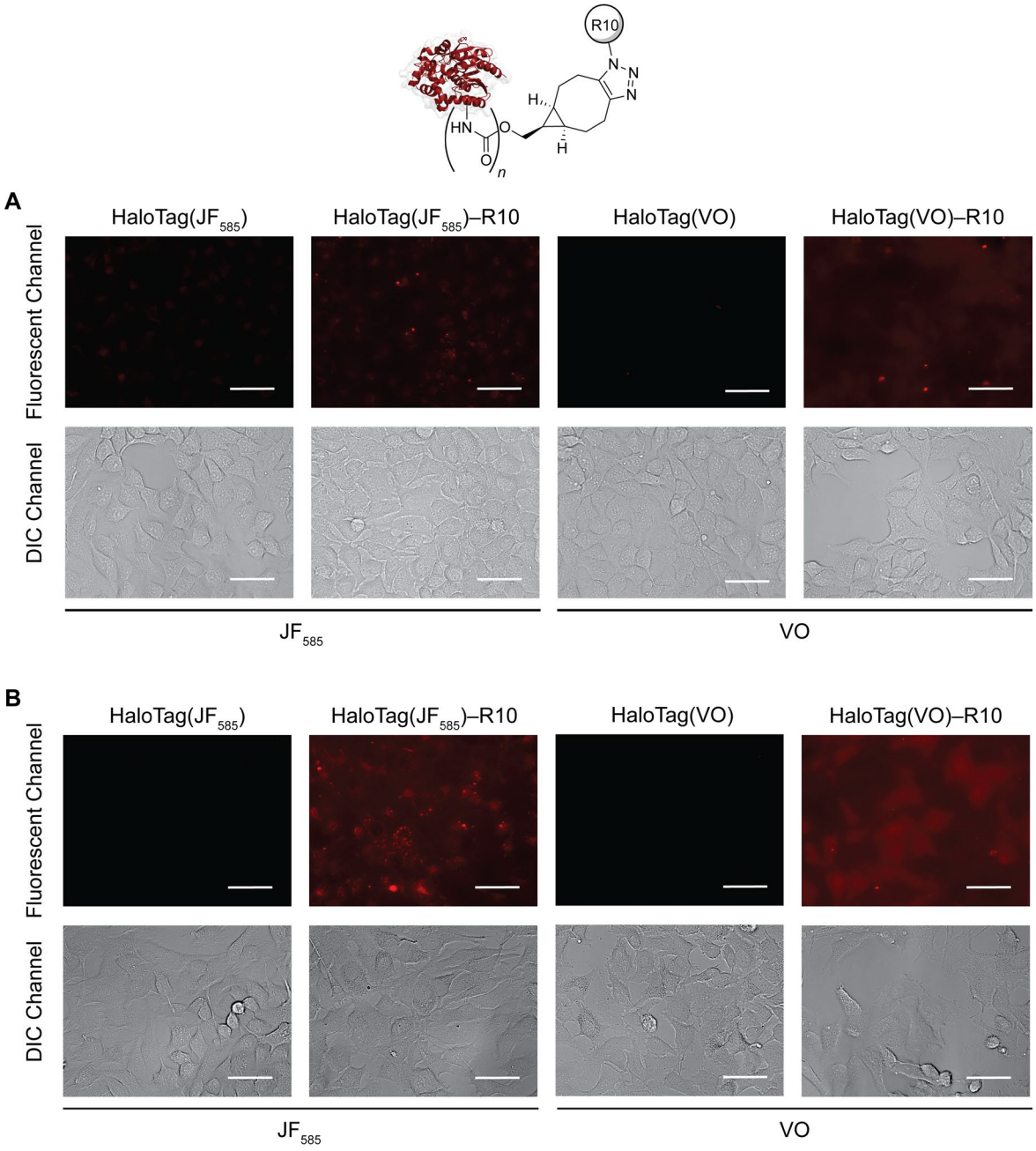


Figure S15. Fluorescence microscopy images showing the effect of conjugation to R10 on the uptake of HaloTag by HeLa cells after a 1.5-h treatment period followed by a 1-h rest period at 37 °C. (A) 2.5 μM. (B) 5 μM. JF_{585}: λ_{ex} = 585/29 nm, λ_{em}= 628/32 nm (Texas Red channel). VO: λ_{ex} = 542/20 nm, λ_{em} = 593/40 nm (RFP channel). Scale bars, 50 μm. All images were acquired with identical laser settings.

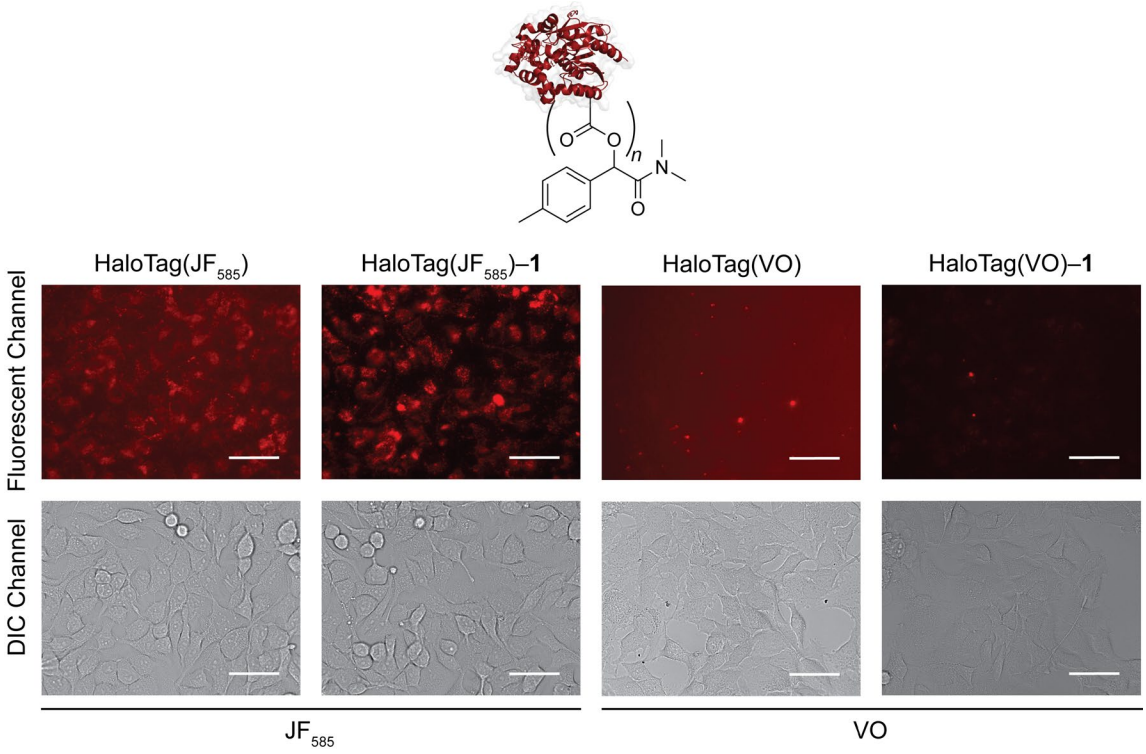


Figure S16. Fluorescence microscopy images showing the effect of esterification with **1** on the cellular uptake of HaloTag (10 μM) by HeLa cells after a 2-h treatment period followed by a 30-min recovery at 37 °C. JF_{585}: λ_{ex} = 585/29 nm, λ_{em}= 628/32 nm (Texas Red channel). VO: λ_{ex} = 542/20 nm, λ_{em} = 593/40 nm (RFP channel). Scale bars: 50 μm. All images were acquired with identical laser settings.

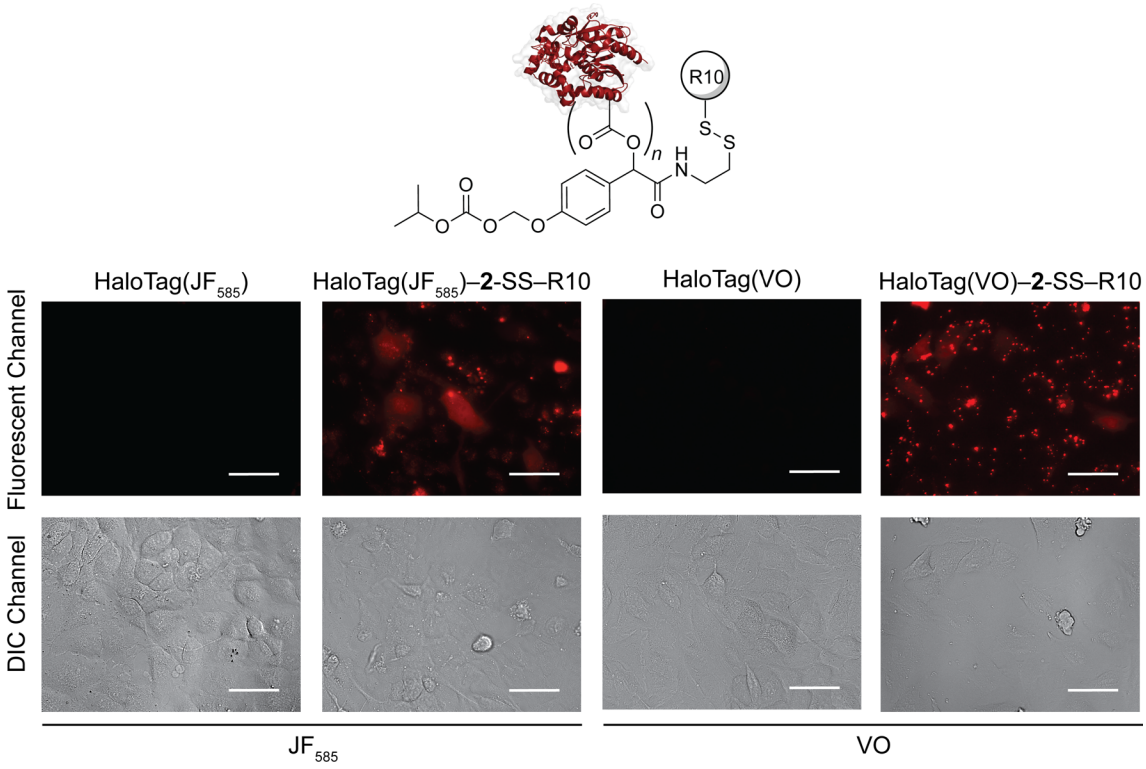


Figure S17. Fluorescence microscopy images showing the effect of esterification with **2**-SS–R10 on the uptake of HaloTag (1.25-1.5 μM) by HeLa cells after a 1.5-h treatment period followed by a 1-h recovery at 37 °C. JF_{585}: λ_{ex} = 585/29 nm, λ_{em}= 628/32 nm (Texas Red channel). VO: λ_{ex} = 542/20 nm, λ_{em} = 593/40 nm (RFP channel). Scale bars, 50 μm. All images were acquired with identical laser settings.

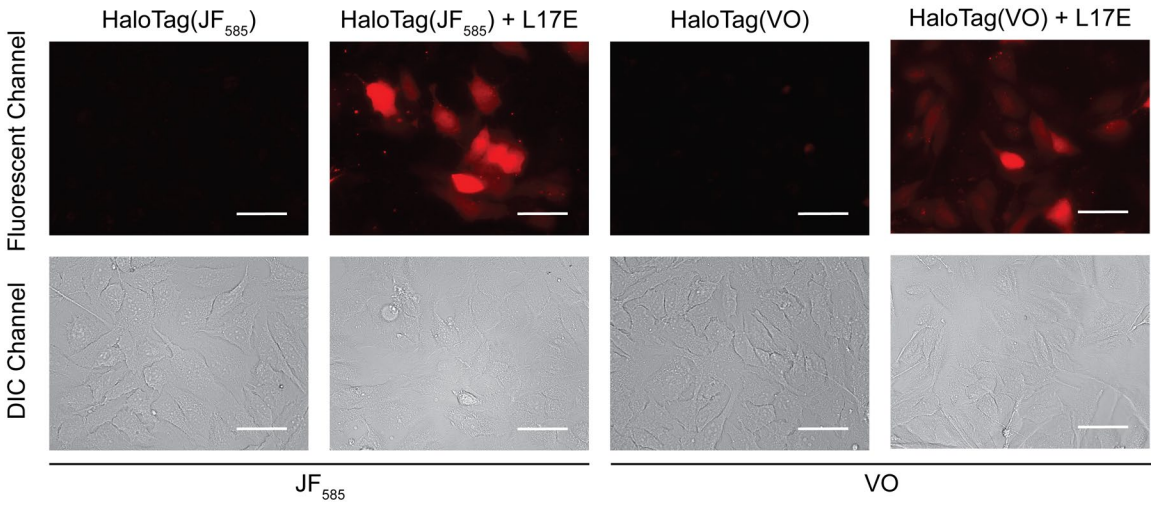


Figure S18. Fluorescence microscopy images showing the effect of unconjugated L17E (40 μM) on the uptake of HaloTag (15 μM) by HeLa cells after a 5–10 min treatment period followed by a 1-h recovery at 37 °C. JF_{585}: λ_{ex} = 585/29 nm, λ_{em}= 628/32 nm (Texas Red channel). VO: λ_{ex} = 542/20 nm, λ_{em} = 593/40 nm (RFP channel). Scale bars, 50 μm. All images were acquired with identical laser settings. Scale bars: 50 μm. All images were acquired with identical laser settings.

XVI. Imaging and Flow Cytometry

Cells were seeded to be 90% confluent at the time of the experiment. Specifically, cells were seeded at 36,000 cells/well if performing the experiment 24 h later; cells were seeded at 18,000 cells/well if performing the experiment 48 h later; cells were seeded at 9,000 cells/well if performing the experiment 72 h later. In each case ,cells were seeded in a sterile 18-well IbiTreat dish (Ibidi). Prior to treatment, cells were washed with DPBS without Ca^{2+}/Mg^{2+} (2 × 200 μL), and serum-free medium was added to each well. Proteins were dosed into each well and incubated at 37 °C in a humidified incubator at 5% v/v CO_2. The volume of DPBS in each treatment condition did not exceed 25% of the medium volume, and the final volume of each well after the addition of protein was equal to 100 μL. After incubation with the protein conjugate, cells were washed with DPBS without Ca^{2+}/Mg^{2+} (2 × 100 μL) and Fluorobrite DMEM (100 μL) was used for imaging. Cells were kept protected from light at room temperature until imaging was performed. Epifluorescent imaging was performed using an Evos M7000 Epifluorescent Microscope as described above. After imaging had concluded, cells were washed with DPBS without Ca^{2+}/Mg^{2+} (2 × 100 μL) lifted from the plate with 50 μL of 0.25% v/v trypsin–EDTA. Trypsin was quenched by the addition of 50 μL of full medium. Cells were strained into flow tubes and pelleted via centrifugation for 5 min at 1200 RPM at 4 °C. Cells were resuspended in 1 mL of ice-cold DPBS with Ca^{2+}/Mg^{2+} supplemented with bovine serum albumin (0.1% w/v). Each sample was stained with SYTOX Blue Dead-Cell Indicator (1 μL of a 1.0 mM stock) for ≥5 min on ice protected from light. Cells were kept on ice and protected from light until the time of analysis. The fluorescence intensity of at least 10,000 events was measured by flow cytometry with an Attune NxT Flow Cytometer from Thermo Fisher Scientific (405 nm, 488 nm, 561 nm, and 640 nm lasers). Control cells treated with SYTOX Blue Dead-Cell Indicator (1 μL of a 1.0 mM stock), followed by cells treated with unmodified HaloTag(JF_{585}) or

HaloTag(VO) protein, were analyzed first to set gates and laser intensities. For the JF_{585}-labelled HaloTag conjugates, the 561-nm laser was used for excitation and the 620/15 filter was used to detect fluorescence. For the VO-labeled HaloTag conjugates, the 561-nm laser was used for excitation and the 585/16 filter was used to detect fluorescence. Events were collected using standardized laser intensity values. A template was established for laser intensity values, and all experiments were performed using these laser powers to enable comparison between data sets. Data were analyzed using the FlowJo software package (FlowJo LLC). The median fluorescence intensity (MFI) of live, single cells is reported.

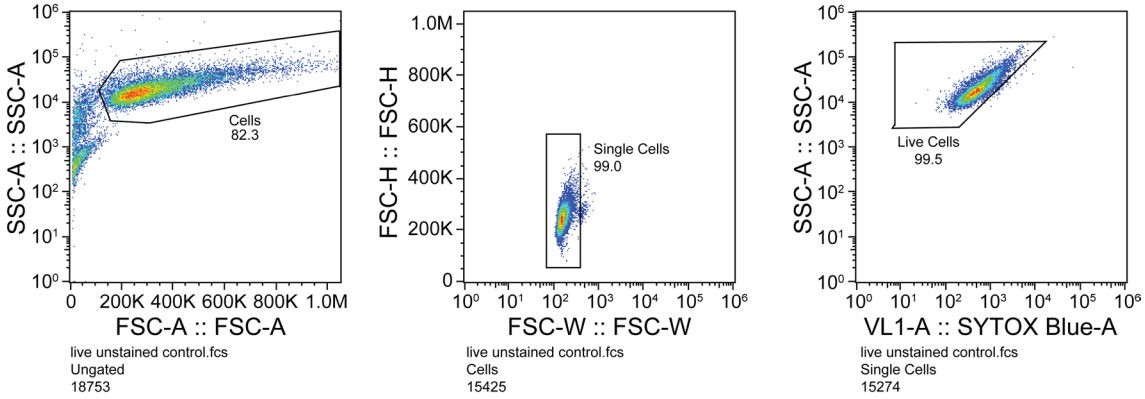


Figure S19. Representative flow cytometry gating strategy to identify live cells in a sample.

Minimum number of ester labels needed for cellular uptake. Uptake experiments were performed with HeLa cells treated with esterified HaloTag(JF_{585}) or HaloTag(VO) containing 0–3 ester labels. Imaging was done with fluorescence microscopy, and uptake was quantitated with flow cytometry.

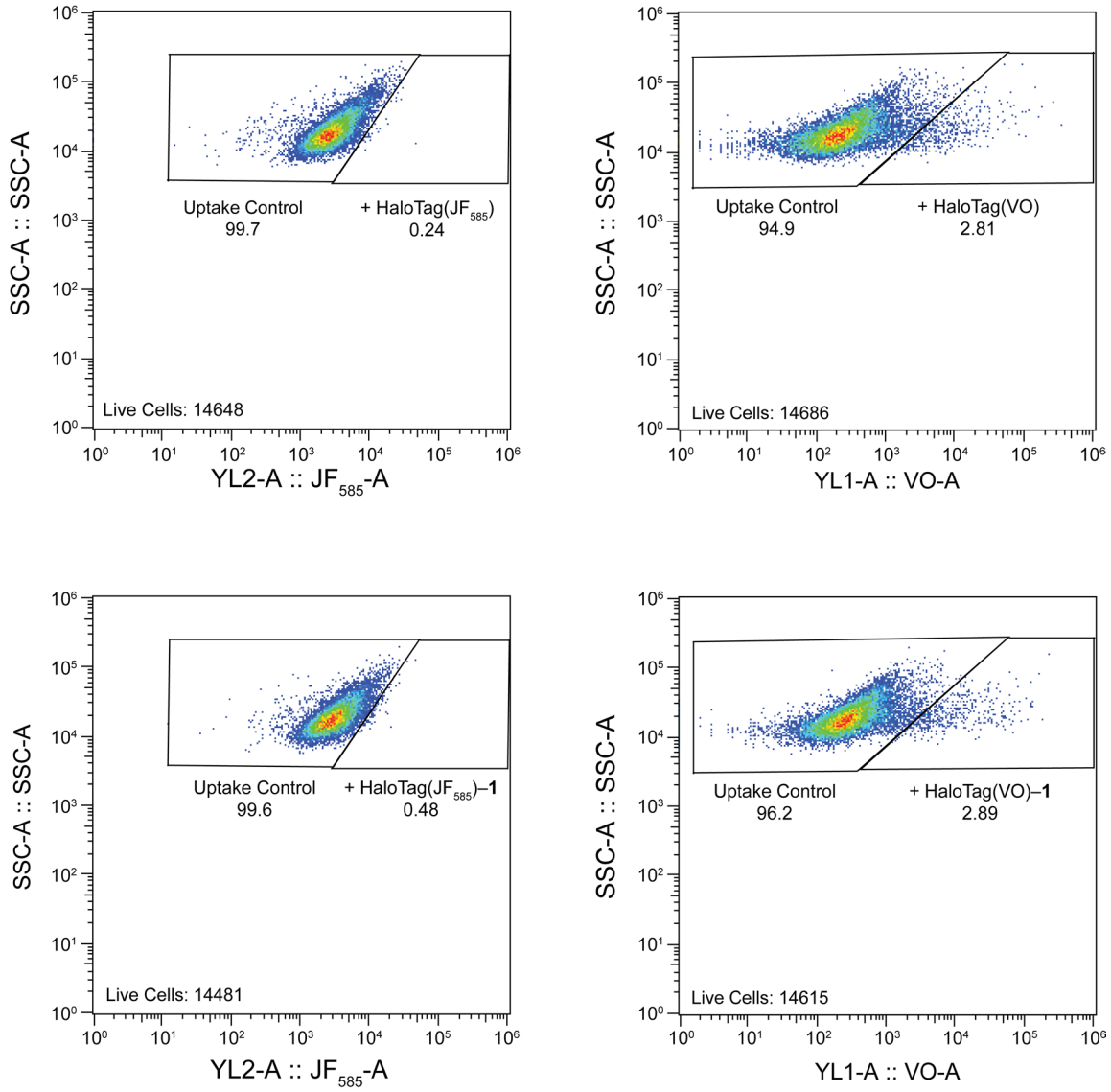


Figure S20. Flow cytometry gating of HeLa cells treated with 5 μM HaloTag(JF_{585}) (top left), HaloTag(JF_{585})–**1** (bottom left), HaloTag(VO) (top right), or HaloTag(VO)–**1** (bottom right), each containing 0–3 ester labels. There is little difference in relative fluorescence between either population relative to the respective control, indicating that a low degree of ester labeling does not appear to effect cellular uptake.

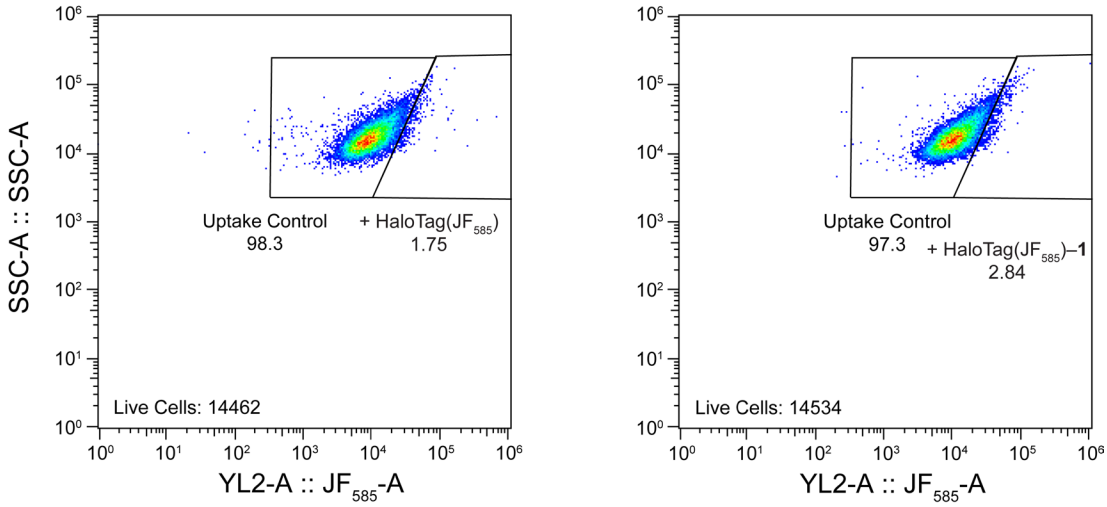


Figure S21. Flow cytometry gating of HeLa cells treated with 15 μM HaloTag(JF_{585}) (left) or HaloTag(JF_{585})–**1** (right), each containing 0–3 ester labels. There is no difference in relative fluorescence between either population, indicating that a lower degree of ester labeling does not appear to substantially increase cellular uptake even upon treatment at a higher concentration.

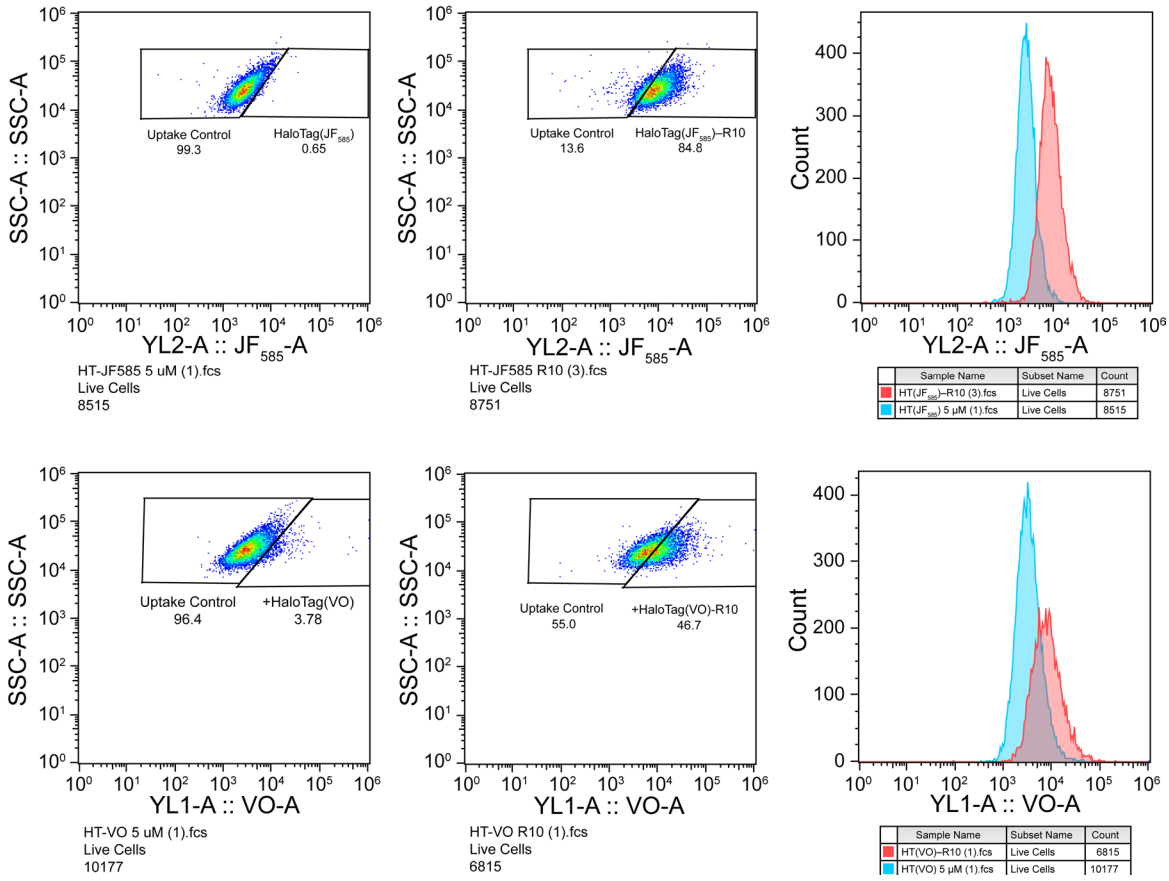


Figure S22. Flow cytometry gating of HeLa cells treated with 5 µM HaloTag(JF_{585}) (top left), HaloTag(JF_{585})–R10 (top middle), HaloTag(VO) (bottom left), or HaloTag(VO)–R10 (bottom middle). These representative plots are shown as histograms overlaying the relative fluorescence of HaloTag(JF_{585}) versus HaloTag(JF_{585})–R10 (top right) and HaloTag(VO) versus HaloTag(VO)–R10 (bottom right). The histograms demonstrate that there is a shift in relative fluorescence of the bulk population when cells are treated with HaloTag(JF_{585})–R10 and with HaloTag(VO)–R10. We note, though, that cell viability and conjugate uptake was variable in cells treated with 5 µM HaloTag(dye)–R10. Due to this variability, flow cytometry for HaloTag(dye)–R10 was subsequently performed at 2.5 µM as those cells were consistently viable (see Figure S23).

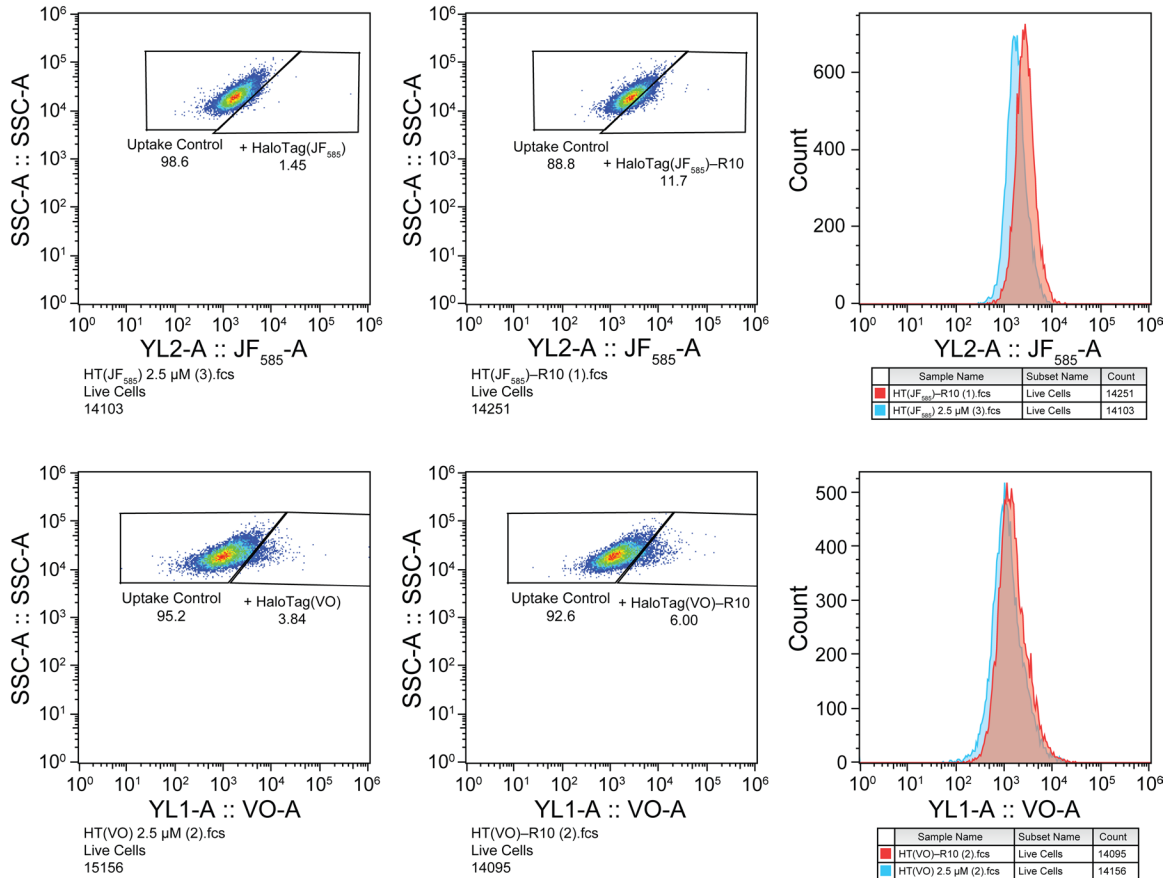


Figure S23. Flow cytometry gating of HeLa cells treated with 2.5 µM HaloTag(JF_{585}) (top left), HaloTag(JF_{585})–R10 (top middle), HaloTag(VO) (bottom left), or HaloTag(VO)–R10 (bottom middle). These representative plots are shown as histograms overlaying the relative fluorescence of HaloTag(JF_{585}) versus HaloTag(JF_{585})–R10 (top right) and HaloTag(VO) versus HaloTag(VO)–R10 (bottom right). The histograms demonstrate that there is a modest shift in relative fluorescence of the bulk population when cells are treated with HaloTag(JF_{585})–R10, and this fluorescent shift is abrogated in cells treated with HaloTag(VO)–R10. The MFI of 2.5 µM HaloTag(dye)–R10 is reported in the main text, and subsequent delivery efficiency comparisons were drawn based on these data.

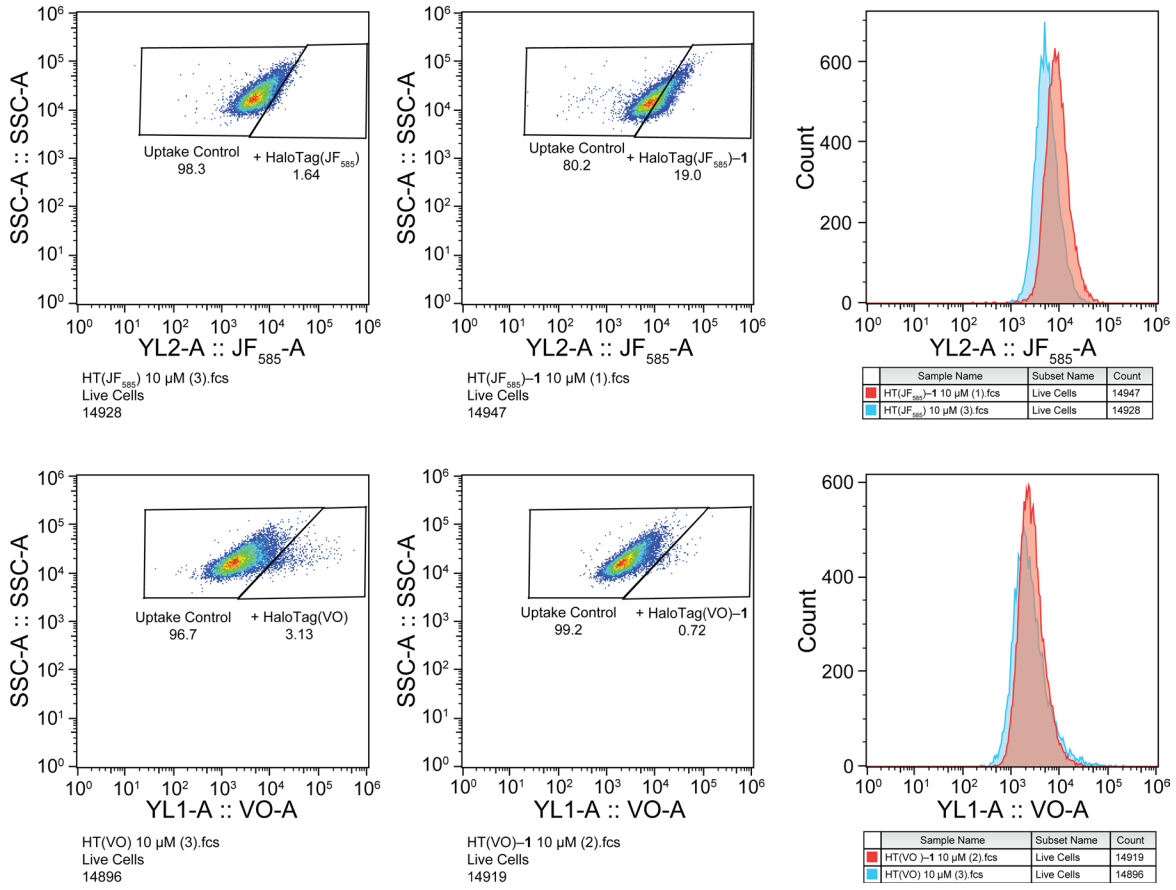


Figure S24. Flow cytometry gating of HeLa cells treated with 10 µM HaloTag(JF$_{585}$) (top left), HaloTag(JF$_{585}$)–**1** (top middle), HaloTag(VO) (bottom left), or HaloTag(VO)–**1** (bottom middle) each containing 2–8 ester labels. There is little difference in relative fluorescence between either population relative to the respective control, indicating that a low degree of ester labeling does not appear to effect cellular uptake. These representative plots are shown as histograms overlaying the relative fluorescence of HaloTag(JF$_{585}$) versus HaloTag(JF$_{585}$)–**1** (top right) and HaloTag(VO) versus HaloTag(VO)–**1** (bottom right). The histograms demonstrate that there is a modest shift in relative fluorescence of the bulk population when cells are treated with HaloTag(JF$_{585}$)–**1**, but that there is no observed shift in cells treated with HaloTag(VO)–**1**.

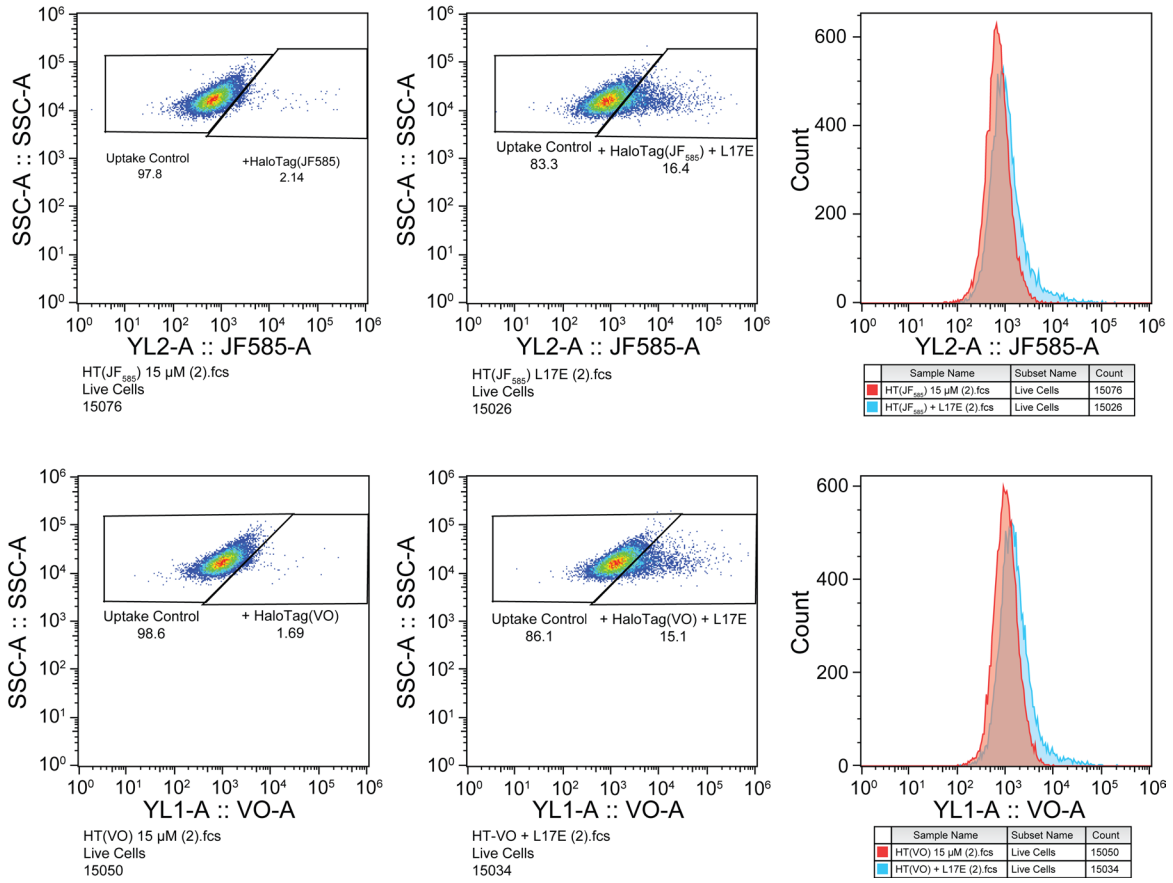
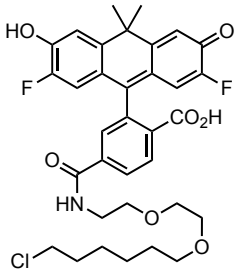


Figure S25. Flow cytometry gating of HeLa cells treated with 15 µM HaloTag(JF_{585}) (top left), HaloTag(JF_{585}) + 40 µM L17E (top middle), HaloTag(VO) (bottom left), or HaloTag(VO) + 40 µM L17E (bottom middle). In each case in which L17E was co-treated, we observed a highly fluorescent subset of cells, corresponding to cells exhibiting robust cytosolic uptake via microscopy. These representative plots are shown as histograms overlaying the relative fluorescence of HaloTag(JF_{585}) versus HaloTag(JF_{585}) + L17E (top right) and HaloTag(VO) versus HaloTag(VO) + L17E (bottom right). The histograms demonstrate that there is little difference in relative fluorescence of the bulk population relative to the respective control and highly fluorescent populations appear as a tailing population in samples treated with L17E.

XVII. NMR Spectra

VO-HTL

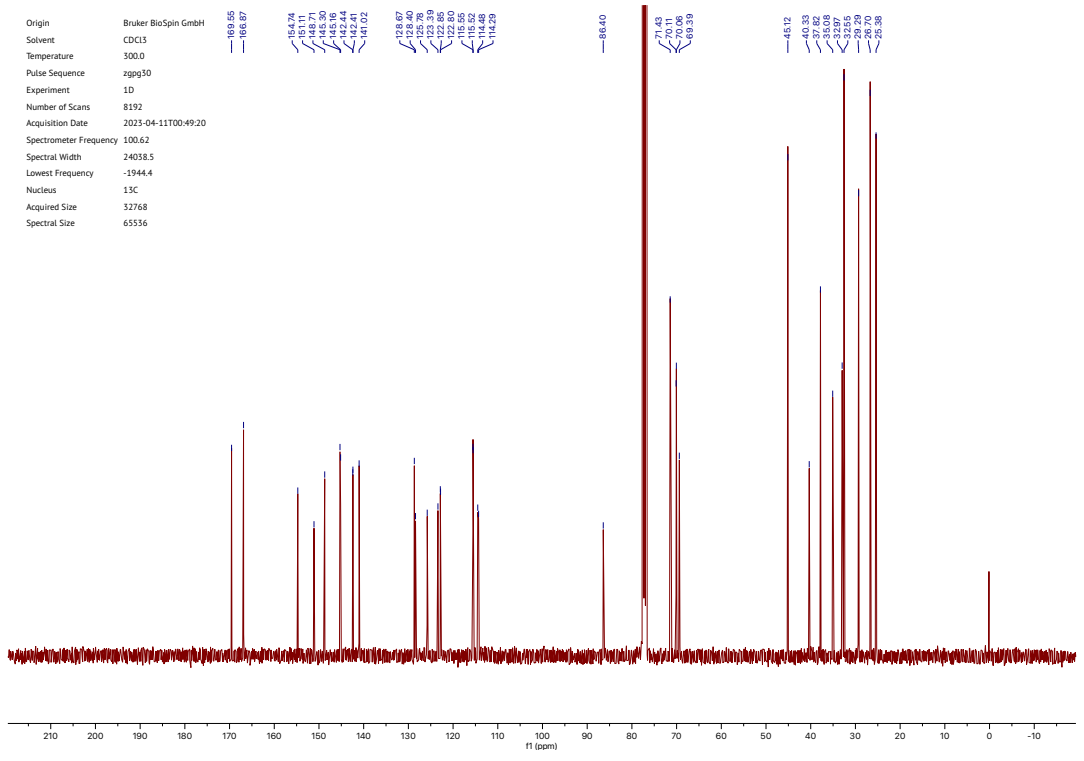
^{1}H NMR (400 MHz) in $CDCl_3$



Origin	Bruker BioSpin GmbH
Solvent	CDCl3
Temperature	300.0
Pulse Sequence	zg30
Experiment	1D
Number of Scans	16
Acquisition Date	2023-04-10T14:06:02
Spectrometer Frequency	400.13
Spectral Width	8012.8
Lowest Frequency	-1545.3
Nucleus	1H
Acquired Size	32768
Spectral Size	65536

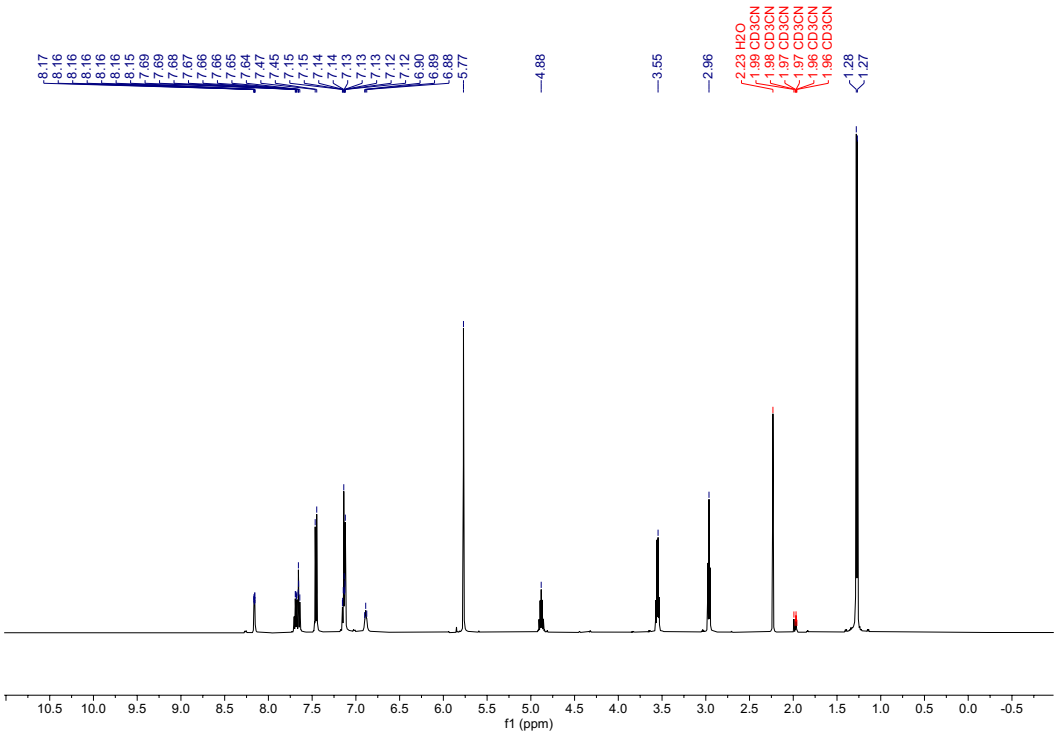
^{13}C{^{1}H} NMR (100 MHz) in $CDCl_3$

Origin	Bruker BioSpin GmbH
Solvent	CDCl3
Temperature	300.0
Pulse Sequence	zgpg30
Experiment	1D
Number of Scans	8192
Acquisition Date	2023-04-11T00:49:20
Spectrometer Frequency	100.62
Spectral Width	24038.5
Lowest Frequency	-1944.4
Nucleus	13C
Acquired Size	32768
Spectral Size	65536



2-SSpy

^{1}H NMR (500 MHz) in CD_3CN



XVIII. References

(1) Mix, K. A.; Raines, R. T. Optimized diazo scaffold for protein esterification. *Org. Lett.* **2015**, *17*, 2358–2361.

(2) Jun, J. V.; Raines, R. T. Two-step synthesis of α-aryl-α-diazoamides as modular bioreversible labels. *Org. Lett.* **2021**, *23*, 3110–3114.

(3) Martineau, M.; Somasundaram, A.; Grimm, J. B.; Gruber, T. D.; Choquet, D.; Taraska, J. W.; Lavis, L. D.; Perrais, D. Semisynthetic fluorescent pH sensors for imaging exocytosis and endocytosis. *Nat. Commun.* **2017**, *8*, 1412.

(4) Jun, J. V.; Petri, Y. D.; Erickson, L. W.; Raines, R. T. Modular diazo compound for the bioreversible late-stage modification of proteins. *J. Am. Chem. Soc.* **2023**, *145*, 6615–6621.

(5) Kang, M. G.; Lee, H.; Kim, B. H.; Dunbayev, Y.; Seo, J. K.; Lee, C.; Rhee, H. W. Structure-guided synthesis of a protein-based fluorescent sensor for alkyl halides. *Chem. Commun.* **2017**, *53*, 9226–9229.

(6) Grimm, J. B.; Gruber, T. D.; Ortiz, G.; Brown, T. A.; Lavis, L. D. Virginia Orange: A versatile, red-shifted fluorescein scaffold for single- and dual-input fluorogenic probes. *Bioconjugate Chem.* **2016**, *27*, 474–480.

(7) Deo, C.; Abdelfattah, A. S.; Bhargava, H. K.; Berro, A. J.; Falco, N.; Farrants, H.; Moeyaert, B.; Chupanova, M.; Lavis, L. D.; Schreiter, E. R. The HaloTag as a general scaffold for far-red tunable chemigenetic indicators. *Nat. Chem. Biol.* **2021**, *17*, 718–723.

(8) Grimm, J. B.; Muthusamy, A. K.; Liang, Y.; Brown, T. A.; Lemon, W. C.; Patel, R.; Lu, R.; Macklin, J. J.; Keller, P. J.; Ji, N.; Lavis, L. D. A general method to fine-tune fluorophores for live-cell and in vivo imaging. *Nat. Methods* **2017**, *14*, 987–994.