Supporting Information

Intracellular Photocatalytic-Proximity Labeling for Profiling Protein–Protein Interactions in Microenvironments

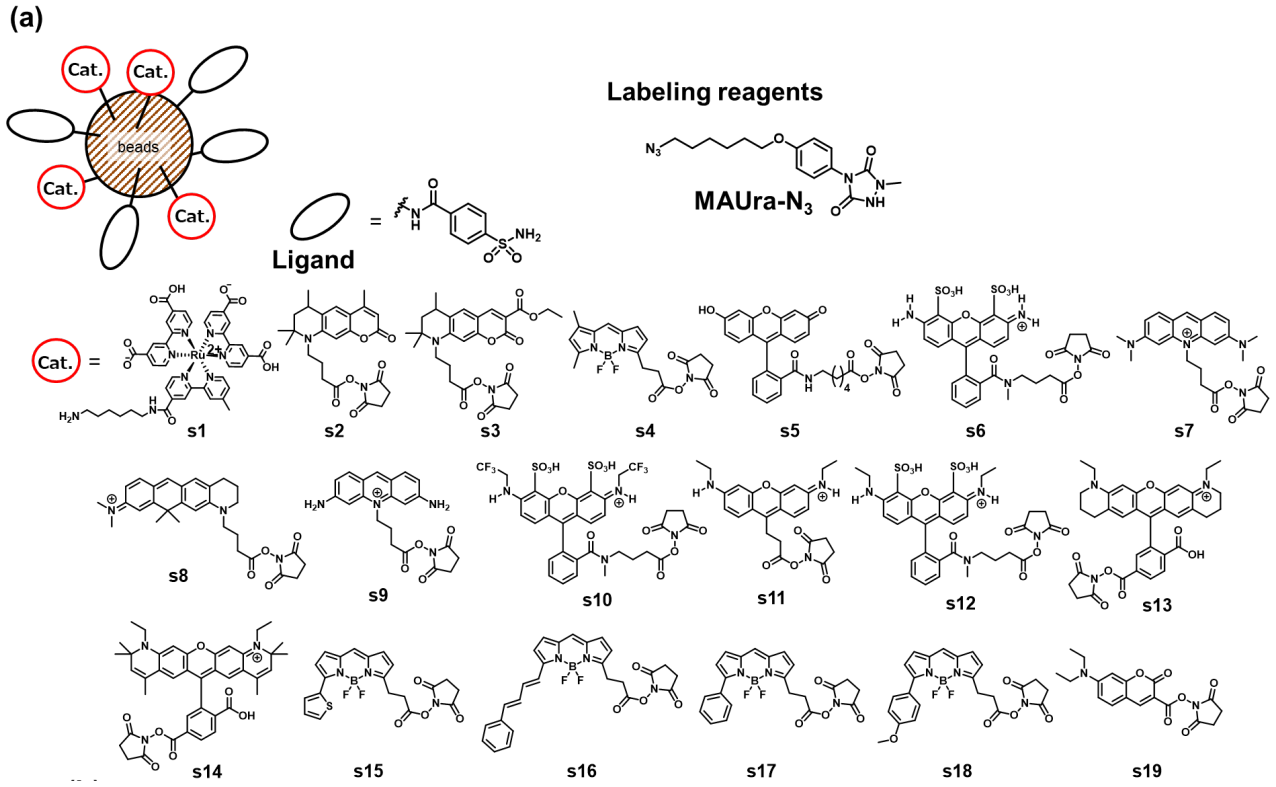
Michihiko Tsushima,[1] Shinichi Sato,*[,2] Kazuki Miura,[1,3] Tatsuya Niwa,[1,4] Hideki Taguchi,[1,4] Hiroyuki Nakamura*[,1,3]

1) School of Life Science and Engineering, Tokyo Institute of Technology, Yokohama, 226-8501, Japan
2) Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, Sendai-shi, Miyagi, 980-8577, Japan.
3) Laboratory for Chemistry and Life Science, Institute of Innovative Research, Tokyo Institute of Technology, Yokohama, 226-8503, Japan
4) Cell Biology Centre, Institute of Innovative Research, Tokyo Institute of Technology, Yokohama, 226-8503, Japan

List of contents

1. Supporting Data

We conducted a proximity labeling evaluation using catalyst-immobilized affinity beads. We chose carbonic anhydrase II (CAII) and benzenesulfonamide as the model protein and its ligand, respectively, for catalyst screening using the proximity labeling evaluation system. The photocatalysts (**s1**–**s19**) and ligand were immobilized on magnetic beads and the labeling was performed in HeLa cell lysate containing CAII using the reagent azide-conjugated MAUra (MAUra-N_3) (Fig. S1a). Among the photoredox catalysts, the coumarin derivatives (**s3**, **s19**), BODIPY derivatives (**s4**, **s17**), acridine derivatives (**s7**, **s9**), fluorescein (**s5**), and rhodamine derivatives (**s13**, **s14**) acted as photocatalysts for protein labeling (Fig. S1b). Proximity labeling on affinity beads immobilized with **s4**, **s13**, **s14**, **s17**, and **s19** proceeded with a high target protein selectivity similar to that of the previously reported ruthenium photocatalyst **s1.** In contrast, ligand-independent proximity labeling was observed on affinity beads immobilized with **s7** and **s9**. A pulldown assay of CAII using photocatalyst-immobilized affinity beads showed that the acridine structure induced ligand-independent protein interactions, which is why proximity labeling on beads immobilized with **s7** and **s**9 resulted in a low target protein selectivity



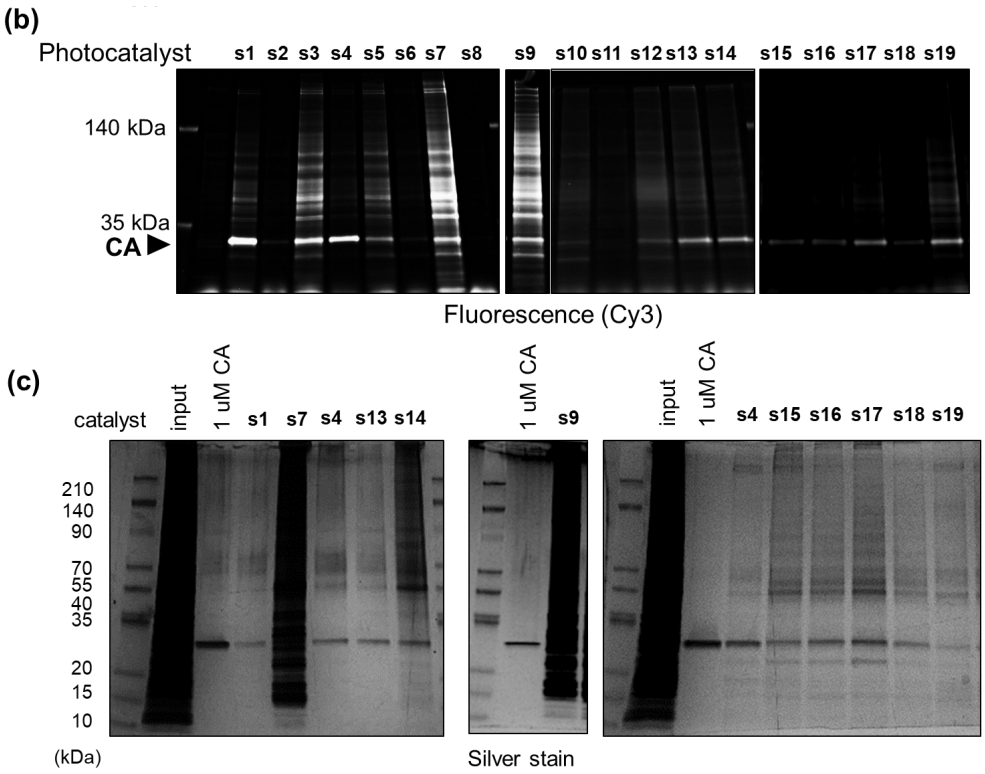


Figure S1. Screening of photocatalyst on catalyst-functionalized affinity beads. (a) Structures of ligand, labeling reagent and photocatalysts. (b) Fluorescence images labeled proteins with MAUra-N_3 under photoirradiation (**s1**-**s12**, **s15**-**s19**: 455 nm, **s13**, **s14**: 540 nm) followed by strain-promoted azide-alkyne cycloaddition (SPAAC) reaction with DBCO-Cy3. (c) Carbonic anhydrase (CA) pull down assay using photocatalyst-functionalized affinity beads. (Input: 1 μM CA in HeLa cell lysate (1.0 mg/mL proteins))

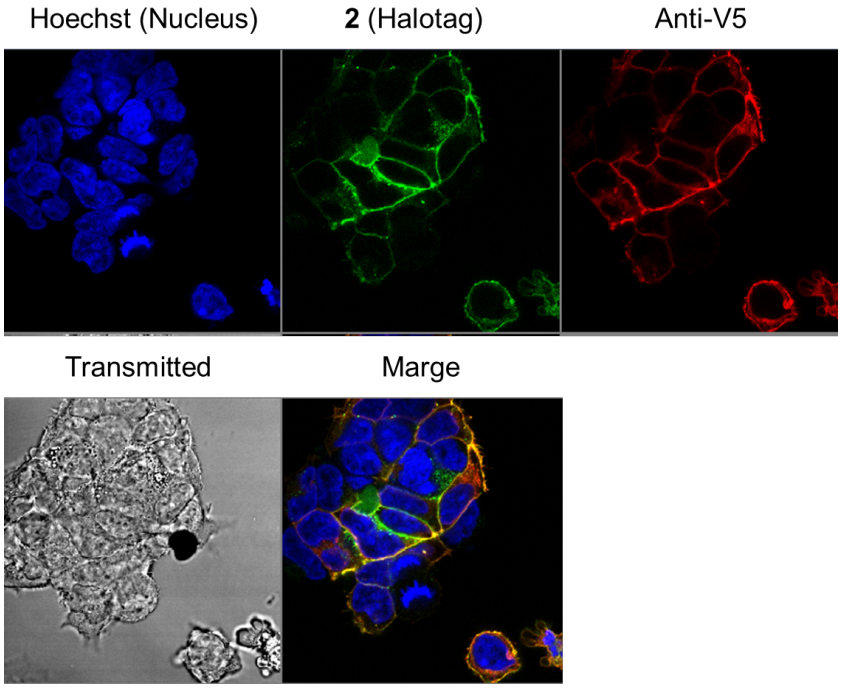


Figure S2. Localization of acriflavine **2** and V5-HaloTag-TM in V5-HaloTag-TM expressed HEK293FT cells.

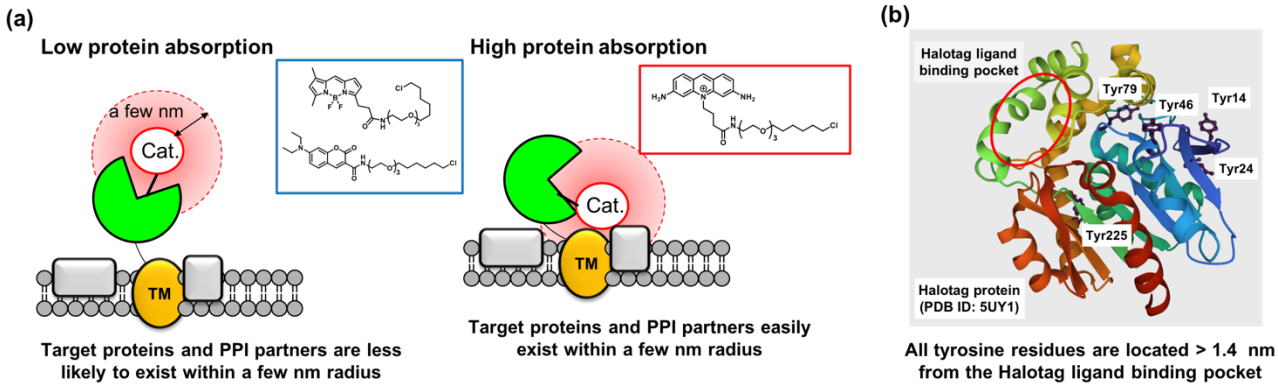


Figure S3. Image of catalysts (BODIPY, coumarin and acriflavine) on HaloTag-TM. (a) Catalysts on HaloTag hardly approach to protein-protein interaction (PPI) partners, but an acriflavine catalyst approaches PPI partners due to its protein absorptive property. (b) Calculation of distance from Halotag ligand binding pocket to each tyrosine residues (PDB: 5UY1, purple: tyrosine residues). The distance from Halotag ligand binding pocket to Try14, Try24, Try46, Try79, and Try225 residues were calculated 2.8, 2.7, 1.8, 21.2, and 29.1 nm, respectively. The distances were calculated on PyMOL software.

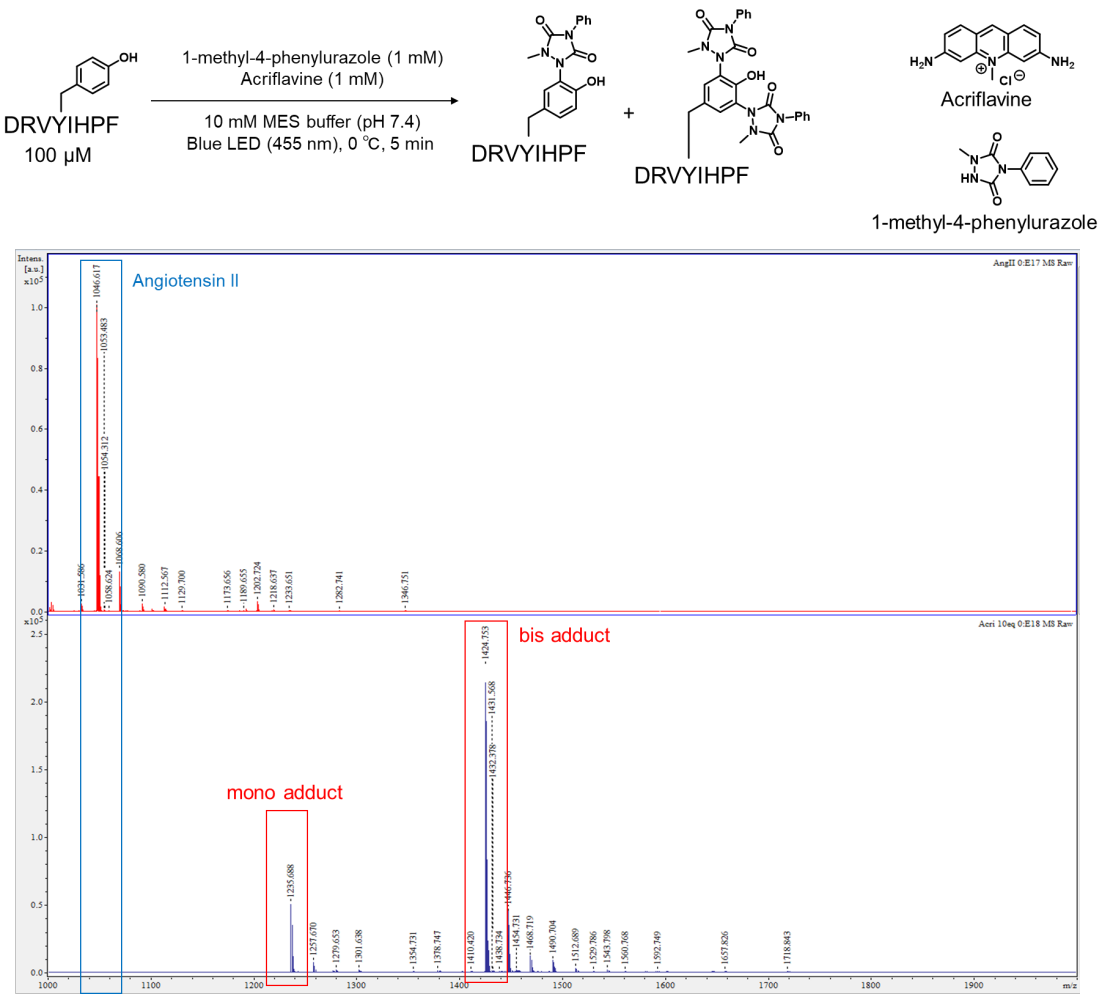


Figure S4. Angiotensin II (DRVYIHPF) labeling using acriflavine and 1-methyl-4-phenylurazole as photocatalyst and labeling reagent. Angiotensin II (100 μM) was labeled by 1-methyl-4-phenylurazole (1 mM) using acriflavine (1 mM) in 10 mM MES buffer (pH 7.4) under blue LED irradiation (455 nm) at 0 °C for 5 min. MALDI TOF-MS spectra were measured using a Bruker Daltonics UltrafleXtreme.

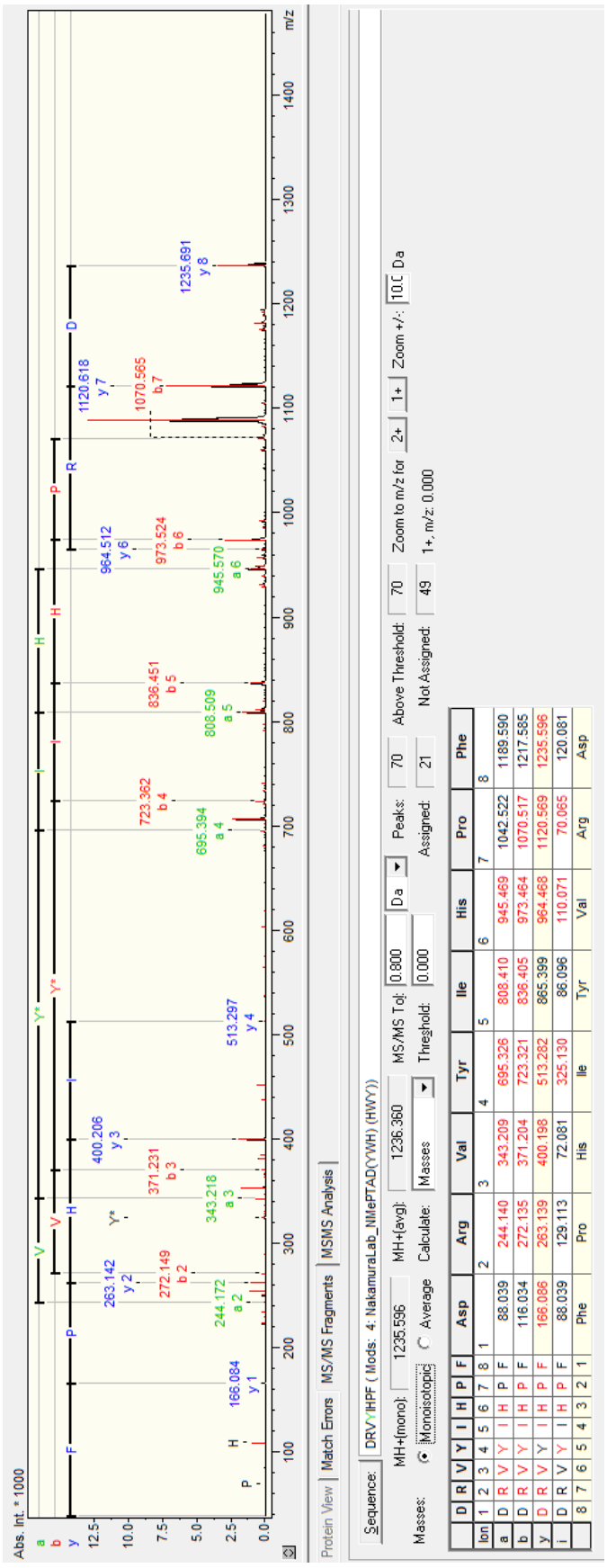


Figure S5. MS/MS analysis of mono-labeled angiotensin II (DRVYIHPF). Angiotensin II (100 μM) was labeled by 1-methyl-4-phenylurazole (1 mM) using acriflavine (1 mM) in 10 mM MES buffer (pH 7.4) under blue LED irradiation (455 nm) at 0 °C. Y4 of Angiotensin II was selectively labeled with 1-methyl-4-phenylurazole. MS/MS spectra were measured using a Bruker Daltonics UltrafleXtreme.

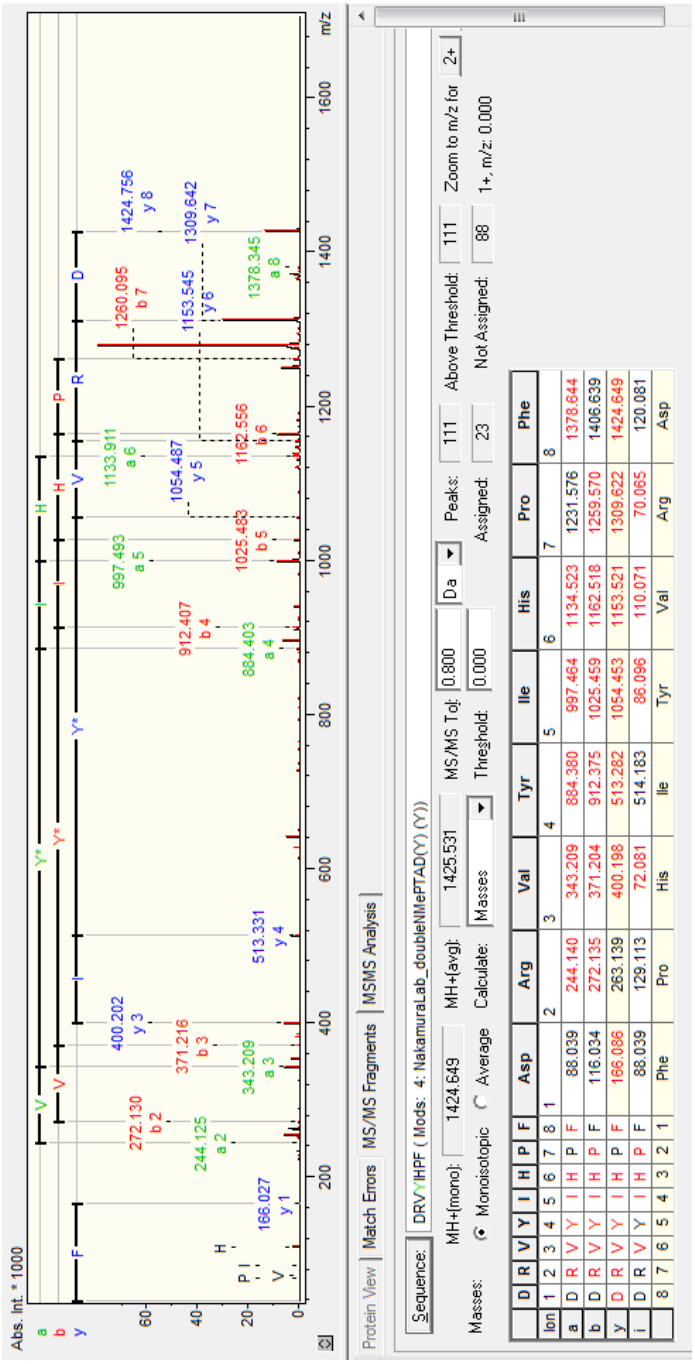


Figure S6. MS/MS analysis of bis-labeled angiotensin II (DRVYIHPF). Y4 of Angiotensin II was selectively labeled with 1-methyl-4-phenylurazole.

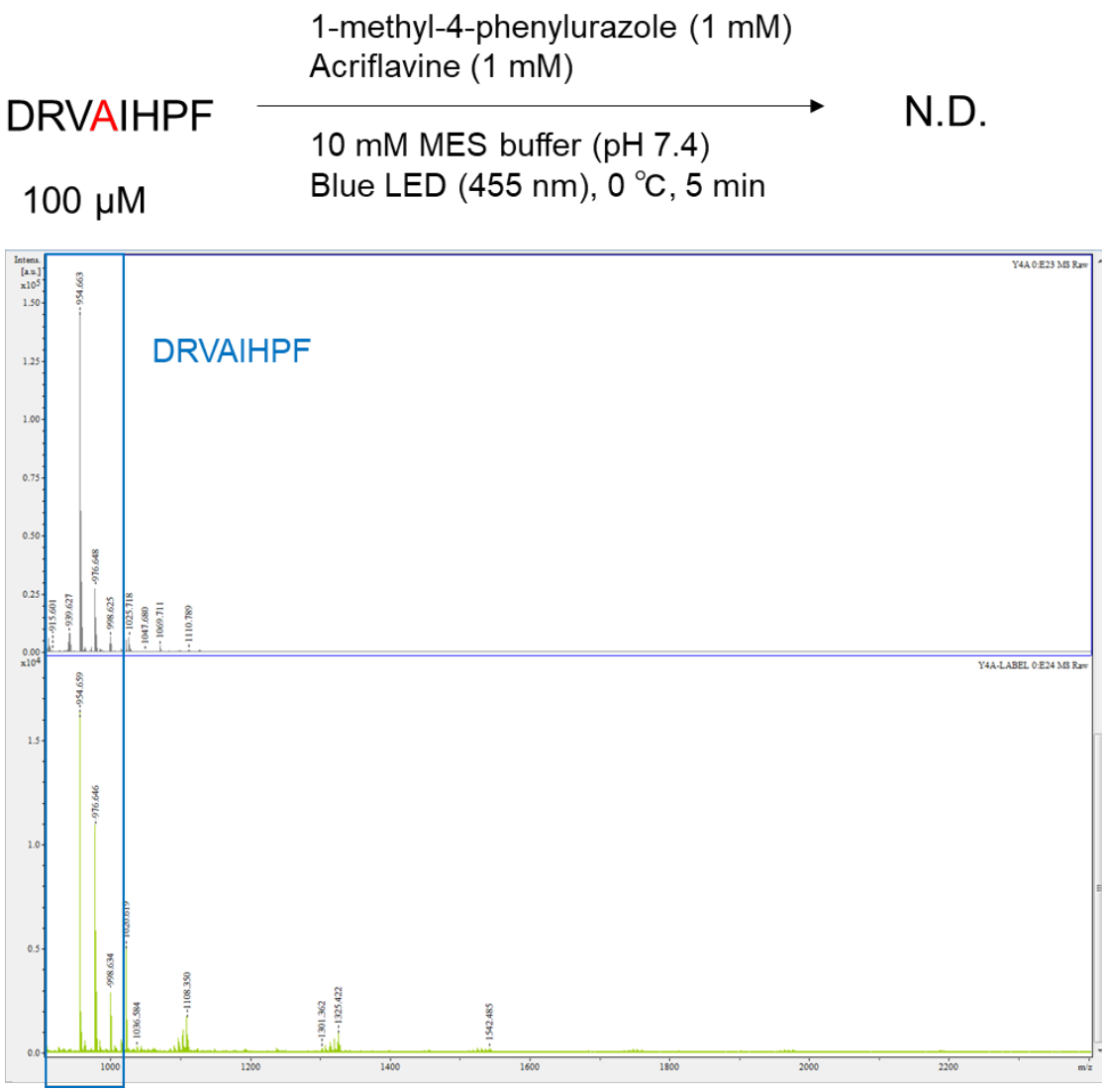


Figure S7. Labeling of Y4A substituted angiotensin II (DRVAIHPF) using acriflavine and 1-methyl-4-phenylurazole as photocatalyst and labeling reagent. Peptide (100 μM) was labeled by 1-methyl-4-phenylurazole (1 mM) using acriflavine (1 mM) in 10 mM MES buffer (pH 7.4) under blue LED irradiation (455 nm) at 0 °C for 5 min. MALDI TOF-MS spectra were measured using a Bruker Daltonics UltrafleXtreme.

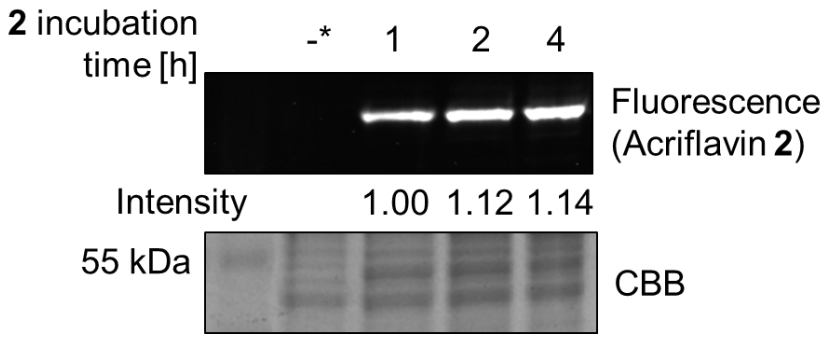


Figure S8. Optimization of acriflavine **2** incubation time. This result suggested that 2 h incubation is enough for catalyst **2** binding to HaloTag-H2B. *Without transfection of HaloTag-H2B.

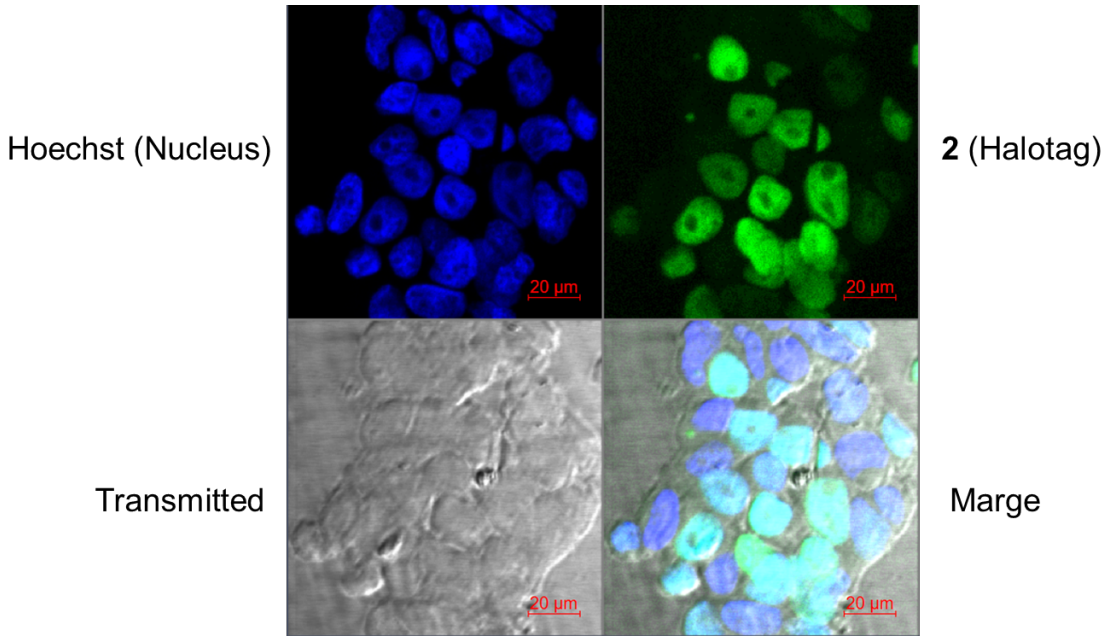


Figure S9. Localization of **2** in HaloTag-H2B expressed HEK293FT cells.

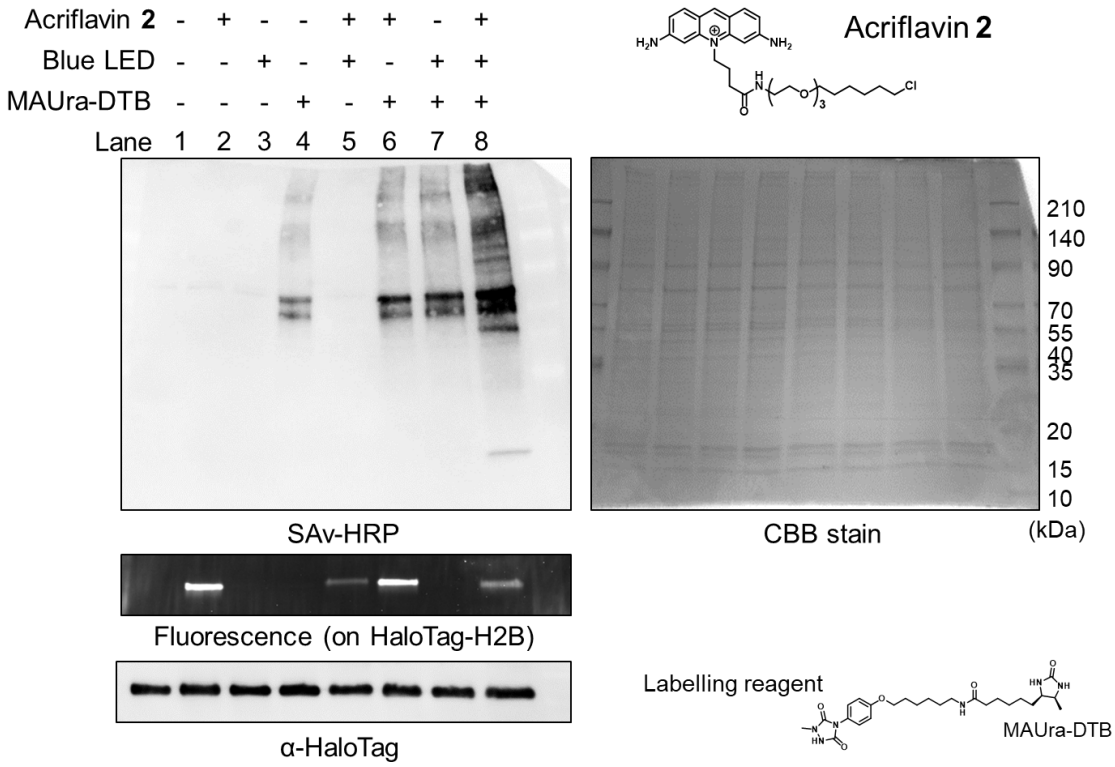


Figure S10. Labeled and unlabeled proteins using acriflavine **2** and MAUra-DTB systems under various conditions in HaloTag-H2B expressed HEK293FT cells. Only adding MAUra-DTB to the cells, some labeled proteins were detected (lane 4), however, HaloTag-H2B self-labeling and labeling of endogenous histones can be only detected under photo-irradiation with MAUra-DTB and acriflavine **2**.

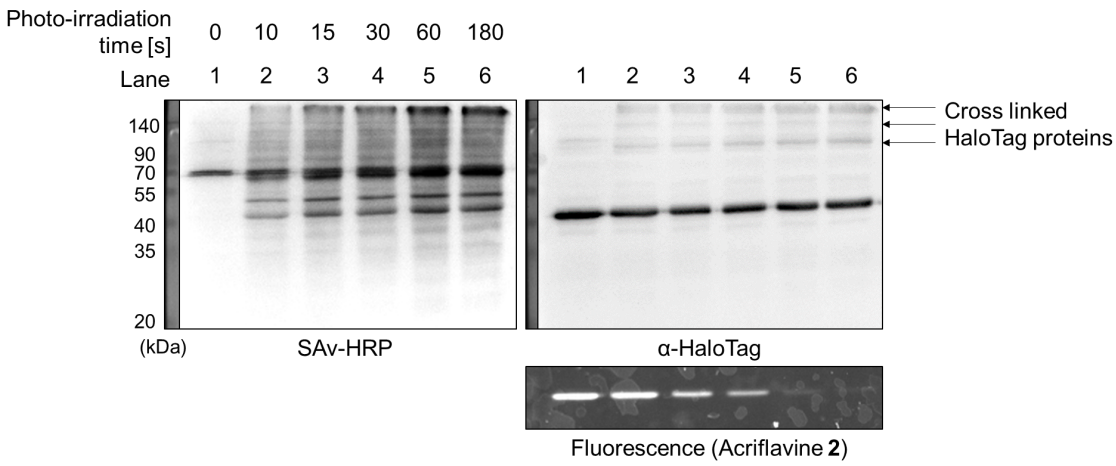


Figure S11. Light irradiation time-dependency in the photocatalyst-proximity labeling. The labeling was completed within 60 s via photobleaching of **2**. Not only labeling reaction but also cross-linking reaction via tyrosine-tyrosine radical coupling proceeded.

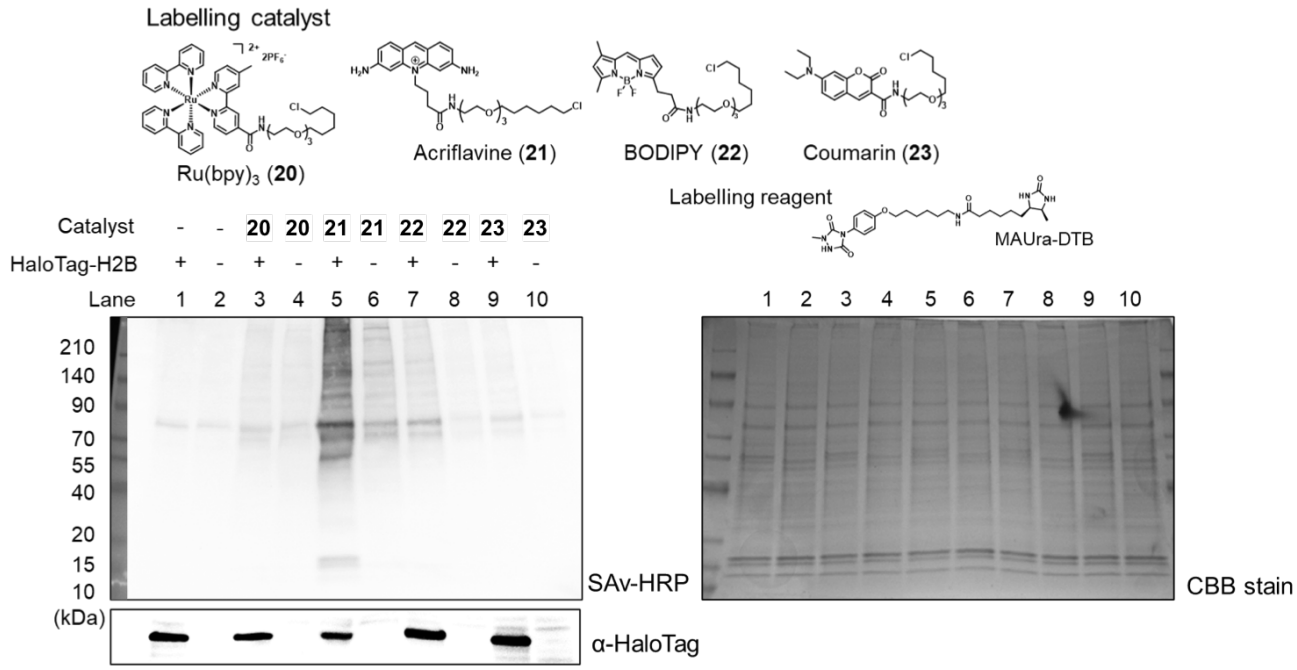


Figure S12. Catalyst-proximity labeling with MAUra-DTB using various photocatalysts (**20**-**23**) in HaloTag-H2B expressed HEK293FT cells. Labeling of HaloTag-H2B (ca. 51 kDa) and histone monomers (15 kDa) was observed only when acriflavine **21** was used as photocatalyst.

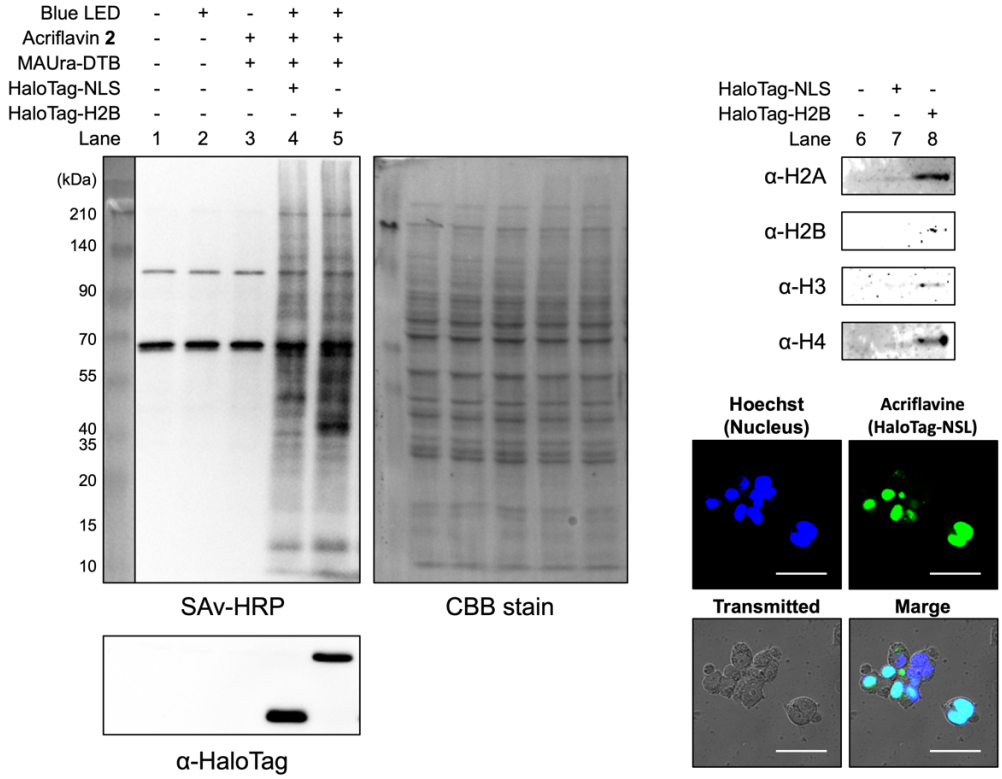


Figure S13. Catalyst-proximity labeling with MAUra-DTB using acriflavine **2** in HaloTag-NSL or HaloTag-H2B expressed HEK293FT cells. Labeled proteins and HaloTag protein (left), the total proteins (center), enriched labeled histone proteins (upper right), and HaloTag-NSL protein localization in cells (right lower) are shown. Scare bars; 50 μm.

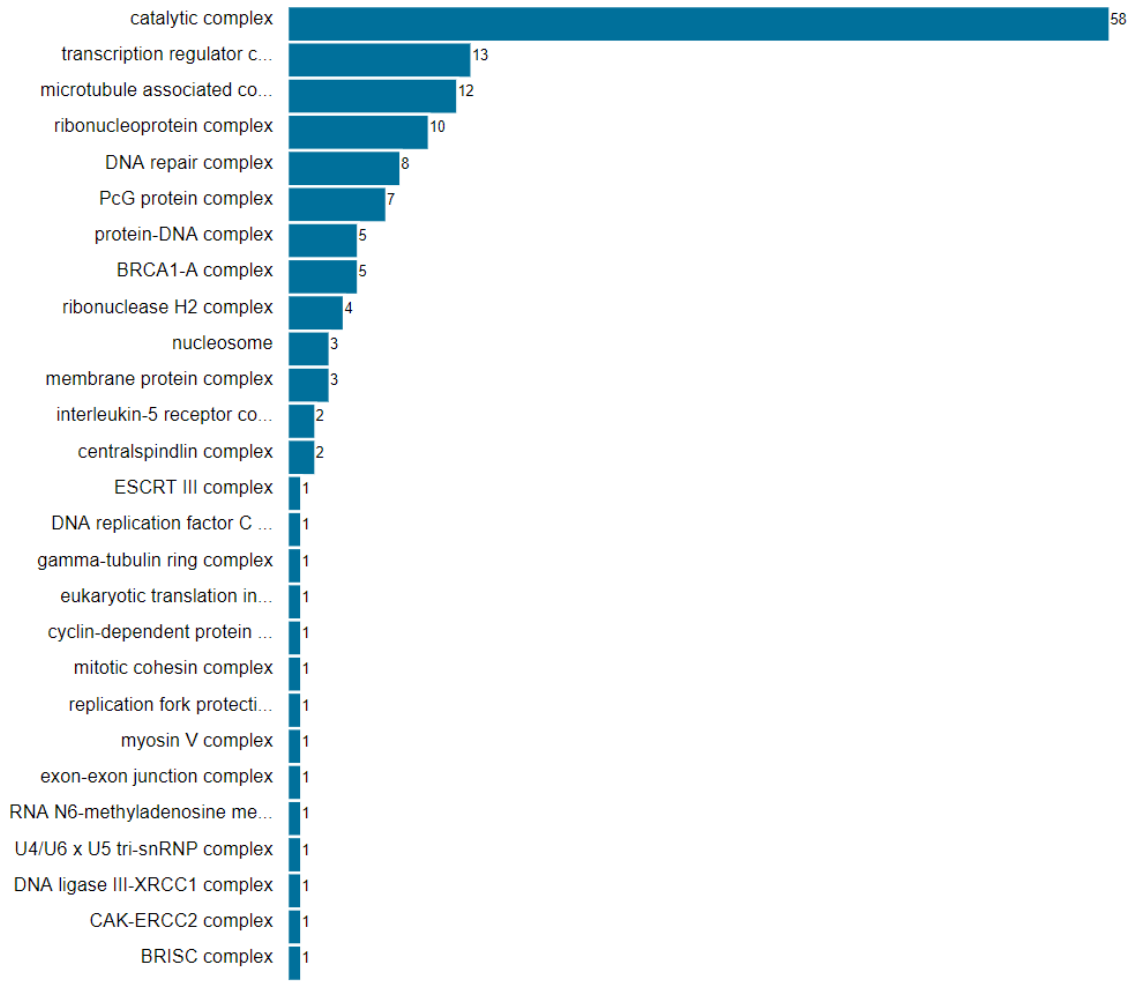


Figure S14. Gene ontology (GO)-term enrichment analysis of significantly labeled proteins in BioID. The labeled proteins list was obtained from J.-P. Lambert and co-workers' report[1] and proteins list were re-analyzed using Uniprot.

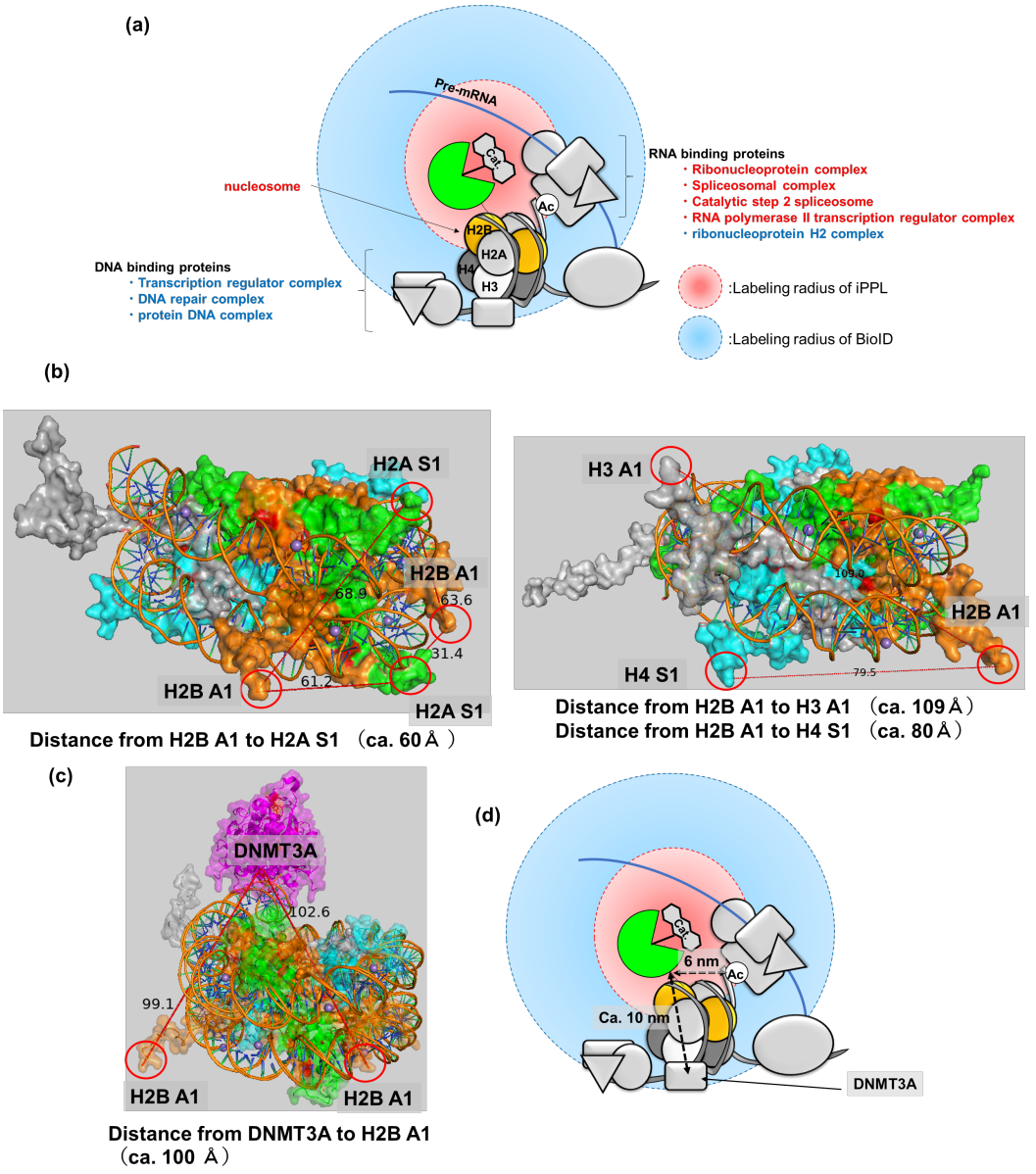


Figure S15. Estimation of labeling radius. (a) Labeled proteins list labeled by iPPL (red) or BioID (blue). DNA binding proteins were labeled by only BioID method. (b) Calculation of distance from H2B tail to H2A tail (PDB: 1kx5, green: H2A, orange: H2B, gray: H3, cyan: H4, red: tyrosine residues). Distance from H2B tail to H2A tail is ca. 6 nm, so RNA binding proteins which can bind to H2A tail (e.g. NONO, SFPQ, these are labeled proteins by iPPL, see table S1) locate within approximately 6 nm radius. (c) Calculation of distance from DNA methyltransferase 3A (DNMT3A) to H2B tail (PDB: 6pa7 (merge with 1kx5), magenta: DNMT3A, green: H2A, orange: H2B, gray: H3, cyan: H4, red: tyrosine residues). The distances were calculated on PyMOL software. (d) The image of location of nucleosome and other components. Distance from H2B tail to DNMT3A which labeled by BioID is ca. 10 nm and DNMT3A are not labeled by iPPL method. These estimations imply iPPL can label proteins located within 6 nm radius.

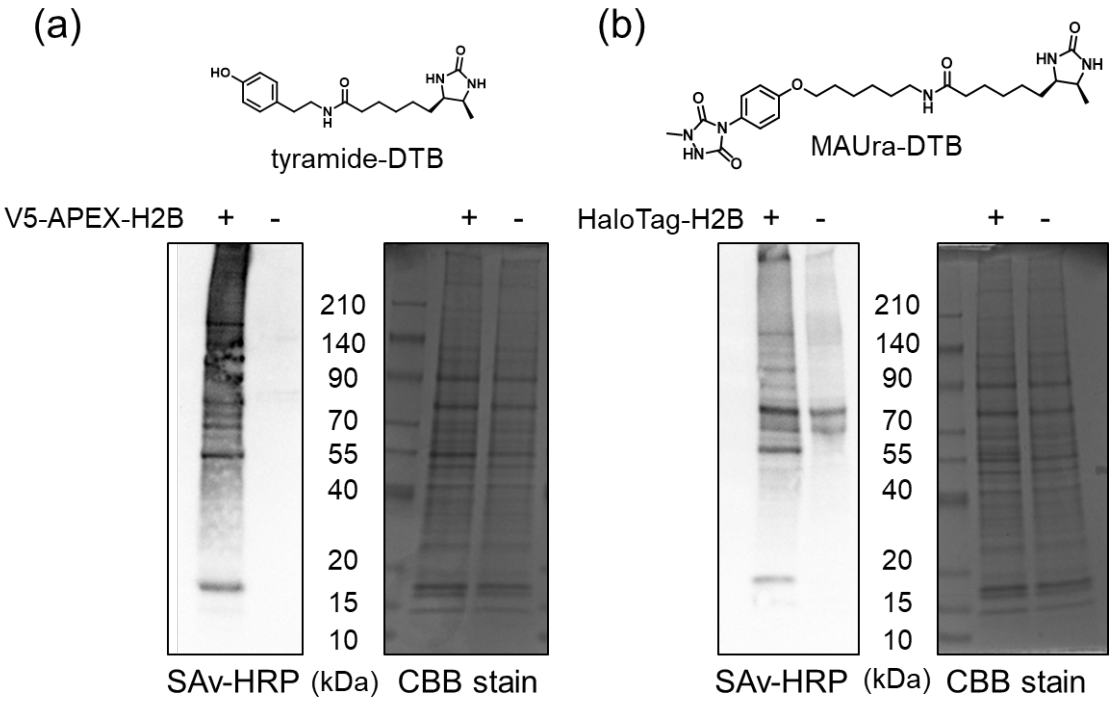


Figure S16. Comparison of APEX method and our photocatalyst-proximity labeling system. (a) histone proximity labeling in V5-APEX-H2B expressed HEK293FT cells using tyramide-DTB as a labeling reagent. (b) histone proximity labeling in HaloTag-H2B expressed HEK293FT cells using MAUra-DTB as a labeling reagent.

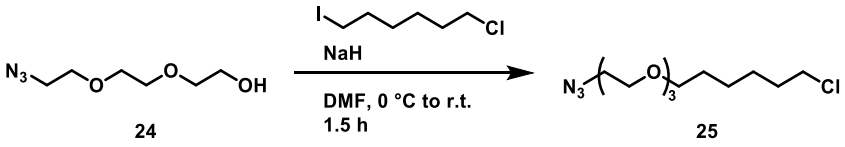
Table S1. List of significant enriched proteins (LED+/ LED- > 2, $p < 0.05$) in our system.

Accession	Description	$\log_2$(hv+/hv-)	$-\log_{10}p$	Cellular Component
Q9P258	Protein RCC2 OS=Homo sapiens OX=9606 GN=RCC2 PE=1 SV=2	> 6.64	3.94	chromosome;cytoplasm;cytoskeleton;cytosol;membrane;nucleus
P40926	Malate dehydrogenase, mitochondrial OS=Homo sapiens OX=9606 GN=MDH2 PE=1 SV=3	> 6.64	2.58	membrane;mitochondrion;nucleus;organelle lumen
P18669	Phosphoglycerate mutase 1 OS=Homo sapiens OX=9606 GN=PGAM1 PE=1 SV=2	6.37	3.06	cytoplasm;cytosol;extracellular;membrane;organelle lumen
P07737	Profilin-1 OS=Homo sapiens OX=9606 GN=PFN1 PE=1 SV=2	5.82	2.82	cytoplasm;cytoskeleton;cytosol;membrane;nucleus
P23246	Splicing factor, proline- and glutamine-rich OS=Homo sapiens OX=9606 GN=SFPQ PE=1 SV=2	4.84	2.77	chromosome;cytoplasm;nucleus
P04406	Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens OX=9606 GN=GAPDH PE=1 SV=3	4.60	2.12	cytoplasm;cytoskeleton;cytosol;membrane;nucleus
P09651	Heterogeneous nuclear ribonucleoprotein A1 OS=Homo sapiens OX=9606 GN=HNRNPA1 PE=1 SV=5	4.42	2.63	cytoplasm;membrane;nucleus;spliceosomal complex
Q71UI9	Histone H2A.V OS=Homo sapiens OX=9606 GN=H2AFV PE=1 SV=3	3.29	2.43	chromosome;nucleus
Q14676	Mediator of DNA damage checkpoint protein 1 OS=Homo sapiens OX=9606 GN=MDC1 PE=1 SV=3	3.21	2.84	chromosome;nucleus
Q15233	Non-POU domain-containing octamer-binding protein OS=Homo sapiens OX=9606 GN=NONO PE=1 SV=4	3.18	2.25	membrane;nucleus
P11142	Heat shock cognate 71 kDa protein OS=Homo sapiens OX=9606 GN=HSPA8 PE=1 SV=1	2.97	2.31	cytoplasm;cytosol;endosome;extracellular;membrane;nucleus;organelle lumen;spliceosomal complex
P19338	Nucleolin OS=Homo sapiens OX=9606 GN=NCL PE=1 SV=3	2.97	2.40	cytoplasm;membrane;nucleus
Q14683	Structural maintenance of chromosomes protein 1A OS=Homo sapiens OX=9606 GN=SMC1A PE=1 SV=2	2.81	2.59	chromosome;cytosol;nucleus
Q9UQE7	Structural maintenance of chromosomes protein 3 OS=Homo sapiens OX=9606 GN=SMC3 PE=1 SV=2	2.78	2.21	chromosome;cytosol;nucleus
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1 OS=Homo sapiens OX=9606 GN=HNRNPA2B1 PE=1 SV=2	2.75	2.43	cytoplasm;extracellular;membrane;nucleus;spliceosomal complex
Q00839	Heterogeneous nuclear ribonucleoprotein U OS=Homo sapiens OX=9606 GN=HNRNPU PE=1 SV=6	2.70	2.43	Cell surface; chromosome; cytoplasm;membrane; nucleus; spliceosomal complex
P62805	Histone H4 OS=Homo sapiens OX=9606 GN=H4C1 PE=1 SV=2	2.60	2.37	chromosome;extracellular;membrane;nucleus
Q14980	Nuclear mitotic apparatus protein 1 OS=Homo sapiens OX=9606 GN=NUMA1 PE=1 SV=2	2.56	2.79	chromosome;cytoplasm;cytoskeleton;cytosol;membrane;nucleus
Q92841	Probable ATP-dependent RNA helicase DDX17 OS=Homo sapiens OX=9606 GN=DDX17 PE=1 SV=2	2.42	2.35	cytoplasm;cytosol;membrane;nucleus
Q96A08	Histone H2B type 1-A OS=Homo sapiens OX=9606 GN=H2BC1 PE=1 SV=3	2.17	2.07	chromosome;nucleus
Q7L7L0	Histone H2A type 3 OS=Homo sapiens OX=9606 GN=HIST3H2A PE=1 SV=3	2.16	2.84	chromosome;nucleus
P0DMV9	Heat shock 70 kDa protein 1B OS=Homo sapiens OX=9606 GN=HSPA1B PE=1 SV=1	2.05	2.31	cytoplasm;cytoskeleton;cytosol;endoplasmic reticulum;extracellular;mitochondrion;nucleus
Q99878	Histone H2A type 1-J OS=Homo sapiens OX=9606 GN=H2AC14 PE=1 SV=3	1.92	2.23	chromosome;nucleus
Q8N257	Histone H2B type 3-B OS=Homo sapiens OX=9606 GN=HIST3H2BB PE=1 SV=3	1.84	2.24	chromosome;cytosol;nucleus
P17844	Probable ATP-dependent RNA helicase DDX5 OS=Homo sapiens OX=9606 GN=DDX5 PE=1 SV=1	1.71	2.16	cytoplasm;membrane;nucleus;spliceosomal complex
P84243	Histone H3.3 OS=Homo sapiens OX=9606 GN=H3-3A PE=1 SV=2	1.66	3.34	chromosome;extracellular;nucleus
Q05639	Elongation factor 1-alpha 2 OS=Homo sapiens OX=9606 GN=EEF1A2 PE=1 SV=1	1.64	2.35	cytoplasm;nucleus
P14618	Pyruvate kinase PKM OS=Homo sapiens OX=9606 GN=PKM PE=1 SV=4	1.55	2.23	cytoplasm;cytosol;extracellular;mitochondrion;nucleus;organelle lumen
P68363	Tubulin alpha-1B chain OS=Homo sapiens OX=9606 GN=TUBA1B PE=1 SV=1	1.37	2.14	cytoplasm;cytoskeleton;membrane
Q5QNW6	Histone H2B type 2-F OS=Homo sapiens OX=9606 GN=HIST2H2BF PE=1 SV=3	1.37	2.35	chromosome;cytosol;nucleus
P0DPK5	Putative histone H3.X OS=Homo sapiens OX=9606 PE=5 SV=1	1.33	2.59	chromosome;nucleus
P17066	Heat shock 70 kDa protein 6 OS=Homo sapiens OX=9606 GN=HSPA6 PE=1 SV=2	1.20	2.27	cytoplasm;cytosol;extracellular;organelle lumen
P07437	Tubulin beta chain OS=Homo sapiens OX=9606 GN=TUBB PE=1 SV=2	1.19	1.74	cytoplasm;cytoskeleton;extracellular;membrane;nucleus;organelle lumen

2. Experimental section

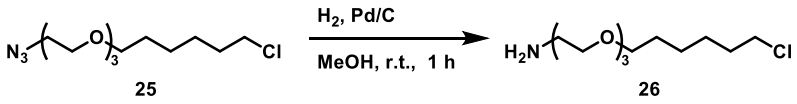
General. NMR spectra were recorded on a Bruker biospin AVANCE III (500 MHz for ^{1}H, 125 MHz for ^{13}C) instrument in the indicated solvent. Chemical shifts are reported in units parts per million (ppm) relative to the signal (0.00 ppm) for internal tetramethylsilane for solutions in $CDCl_3$ (7.26 ppm for ^{1}H, 77.0 ppm for ^{13}C) or CD_3CN (1.94 ppm for ^{1}H, 118.26 ppm for ^{13}C). Multiplicities are reported using the following abbreviations: s; singlet, d; doublet, dd; doublet of doublets, t; triplet, q; quartet, m; multiplet, br; broad, *J*; coupling constants in Hertz. IR spectra were recorded on a JASCO FT/IR-4100 spectrometer. Only the strongest and/or structurally important peaks were reported as the IR data given in cm^{-1}. High-resolution mass spectra (HRMS) were recorded on a Bruker ESI-TOF-MS (micrOTOF II). Analytical thin layer chromatography (TLC) was performed on a glass plate of silica gel 60 GF254 (Merck). Silica gel (Fuji Silysia, CHROMATOREX PSQ 60B, 50-200 μm) was used for column chromatography. Gel permeation chromatography (GPC) for purification was performed on Japan Analytical Industry Model LC-9225 NEXT (recycling preparative HPLC) and a Japan Analytical Industry Model UV-600 NEXT ultra violet detector with a polystyrene gel column (JAIGEL-1H, 20 mm x 600 mm), using chloroform as solvent (3.5 mL/min). Preparative high-performance liquid chromatography (HPLC) was performed with LC-forte/R (YMC) using a C18 reverse phase column (Kanto, Mightysil RP-18 250 × 20 mm, 5 μm). MALDI TOF-MS spectra were measured using a Bruker Daltonics UltrafleXtreme. All chemicals and reagents for biological experiments were obtained from commercial sources and used without further purification.

Synthesis of compounds. Bis(2,2'-bipyridine)[4'-methyl-(2,2'-bipyridine)-4-carboxylicacid] ruthenium(II)bis(hexafluorophosphate) ($Ru(bpy)_2(mcbpy)(PF_6)_2$ (**27**)[3], MAURa-DTB[4], 2-(2-(2-azidoethoxy)ethoxy)ethan-1-ol (**24**)[5] and BODIPY FL (**28**)[6] were synthesized according to previously reported procedure.

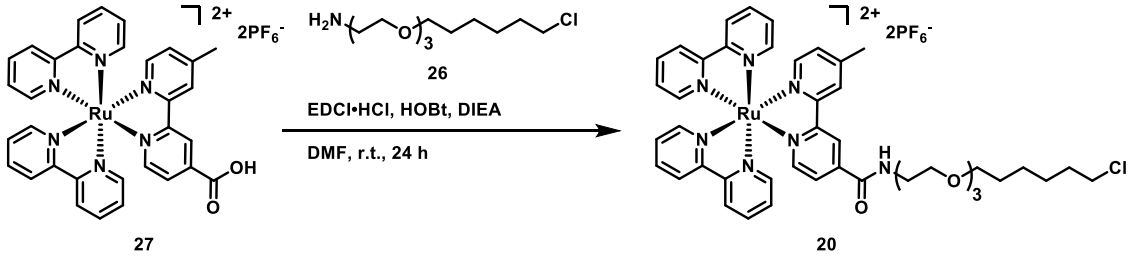


1-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)-6-chlorohexane (25, azide containing HaloTag ligand). To a solution of **27** (232.3 mg, 1.33 mmol) in 1.0 mL of DMF was added NaH (60%, dispersion in paraffin liquid, 63.6 mg, 1.59 mmol) at 0 °C. After stirring at 0 °C for 30 min, to a solution of 1-chrolo-6-iodohexane (490.6 mg, 1.99 mmol) in 1.0 mL of DMF was added the reaction mixture at 0 °C. After stirring at room temperature for 1 h, reaction was quenched by addition of aqueous HCl solution (1 M) and AcOEt, extracted with AcOEt, dried over Na_2SO_4, filtered and concentrated *in vacuo*. The residue was purified by silica gel chromatography with Hexane : AcOEt = 3 : 1 to give **25** as a colorless oil (128.4 mg, 0.585 mmol, 44%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 3.69-3.65 (m, 8H), 3.60-3.58 (m, 2H), 3.53 (t, *J* = 6.7 Hz, 2 H), 3.46 (t, *J* = 6.5 Hz, 4H), 3.91 (t, *J* = 6.0 Hz, 2H), 1.78 (q, *J* = 6.8 Hz, 2H), 1.63-1.57 (m, 2H), 1.48-1.42 (m, 2H), 1.40-1.34 (m, 2H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 71.3, 70.9, 70.9, 70.8, 70.2, 70.2, 50.8, 45.2, 32.7, 29.6, 26.8, 25.6. FT-IR (neat) 2936, 2864, 2104, 1446, 1349, 1301, 1287, 1252, 1122 cm^{-1}; HRMS (ESI, Positive): *m/z* calced. for $C_{12}H_{24}ClN_3O_3$ $[M+Na]^+$: 316.1398, found 316.1397.

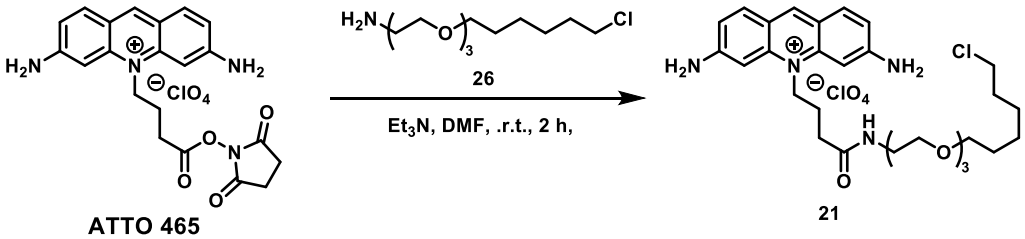


2-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethoxy)ethan-1-amine (26, amine containing HaloTag ligand). To a solution of **25** (54.0 mg, 0.183 mmol) in 1.0 mL of MeOH was added 2.7 mg of Pd/C (50% moisture content). After stirring for 1 h under H_2 at room temperature, Pd/C was removed by Celite filtration and the filtrate was concentrated under the reduced pressure to give **26** as a colorless oil. The colorless oil was applied next step directly without further purification.



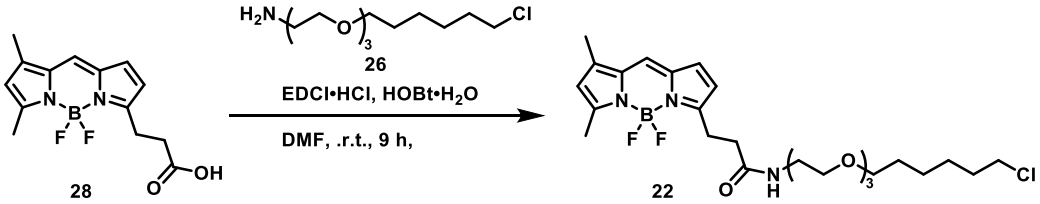
Synthesis of HaloTag ligand-conjugated Ru(bpy)$_3$ complex (20). To a solution of **27** (11.0 mg, 0.0120 mmol), HOBt・H_2O (2.8 mg, 0.0180 mmol) and **26** (3.9 mg, 0.0144 mmol) in 1.0 mL of DMF was added EDCI・HCl (2.8 mg, 0.0144 mmol) at room temperature. After stirring at room temperature for 24 h, reaction was quenched by addition of water (20 mL) and CH_2Cl_2 (10 mL), extracted with CH_2Cl_2, dried over Na_2SO_4, filtered and concentrated *in vacuo*. The residue was purified by HPLC (20-100% MeCN/H_2O) to give **20** as a red oil (3.9 mg, 0.0033 mmol, 28%).

^{1}H NMR (500 MHz, CD_3CN) δ 9.08 (s, 1H), 8.66 (s, 1H), 8.50 (d, J = 8.1 Hz, 4 H), 8.05 (td, J = 7.8, 9.1 Hz, 4H), 7.83 (d, J = 5.9 Hz, 1H), 7.75-7.70 (m, 5H), 7.54 (d, J = 5.8 Hz, 1H), 7.40-7.36 (m, 4H), 7.26 (d, J = 5.7 Hz, 1 H), 3.65-3.63 (m, 2H), 3.59-3.52 (m, 10H), 3.45-3.43 (m, 2H), 3.31 (t, J = 6.6 Hz, 2H), 2.55 (s, 3H), 1.69 (q, J = 6.8 Hz, 2H), 1.76 (q, J = 7.0 Hz, 2H), 1.45-1.41 (m, 2H), 1.38-1.32 (m, 2H), 1.29-1.22 (m, 2H); ^{13}C NMR (125 MHz, CD_3CN δ 164.2, 158.8, 158.0, 157.9, 157.9, 157.8, 157.1, 153.2, 152.7, 152.6, 152.5, 151.8, 151.7, 143.5, 140.6, 138.8, 138.8, 138.7, 129.6, 128.6, 128.5, 126.7, 125.8, 125.2, 125.2, 122.6, 71.5, 71.0, 70.8, 70.7, 69.8, 46.1, 40.8, 33.2, 30.1, 27.2, 26.1, 21.1. FT-IR (neat) 3413, 3111, 3080, 2930, 2864, 1677, 1542, 1465, 1446, 1201, 1132 cm^{-1}; HRMS (ESI, Positive): *m/z* calced. for $C_{44}H_{50}ClN_7O_4Ru$ $[M]^{2+}$: 438.6323, found 438.6326.



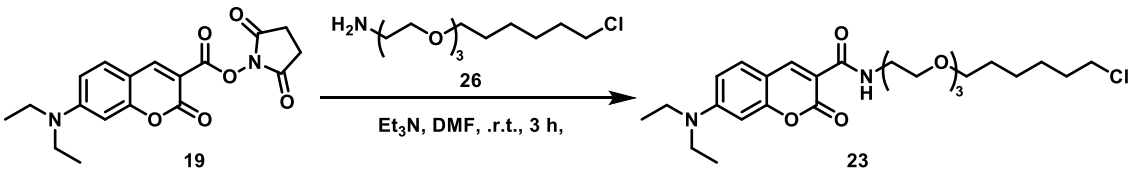
Synthesis of HaloTag ligand conjugated acriflavine (21). To a solution of ATTO 465 (5.0 mg, 0.010 mmol) in 1.0 mL of DMF were added **26** (from a 100 mM stock solution in DMF, 0.010 mmol) and Et_3N (from a 100 mM stock solution in DMF) at room temperature. After stirring at room temperature for 2 h, DMF was removed under reduced pressure. The residue was purified by HPLC (10-80% MeCN/H_2O) to give **21** as a yellow oil (5.7 mg, 0.0088 mmol, 88%).

^{1}H NMR (500 MHz, CD_3OD) δ 8.61 (s, 1H), 8.50 (brs, 1H), 7.83 (d, *J* = 9.4 Hz, 2 H), 7.04-7.01 (m, 4H), 4.57-4.54 (m, 2H), 3.65-3.62 (m, 6H), 3.58-3.56 (m, 2H), 3.53-3.47 (m, 6H), 3.39 (t, *J* = 6.6 Hz, 2H), 2.60 (t, *J* = 6.4 Hz, 2H), 2.25-2.20 (m, 2H), 1.71 (q, *J* = 6.8 Hz, 2H), 1.52 (q, *J* = 7.0 Hz, 2H), 1.44-1.38 (m, 2H), 1.35-1.30 (m, 2H); ^{13}C NMR (125 MHz, CD_3OD) δ 174.7, 159.3, 144.9, 144.7, 135.0, 118.7, 118.1, 94.9, 72.1, 71.6, 71.5, 71.2, 71.1, 70.5, 45.7, 40.5, 33.7, 32.6, 30.5, 27.7, 26.5, 22.6; FT-IR (neat) 3353, 3207, 2931, 2863, 1651, 1599, 1496, 1386, 1326, 1192 cm^{-1}; HRMS (ESI, Positive): *m/z* calced. for $C_{29}H_{42}ClN_4O_4$ $[M]^+$: 545.2889, found 545.2889.



Synthesis of HaloTag ligand conjugated BODIPY (22). To a solution of **28** (11.2 mg, 0.0383 mmol) in 1.0 mL of DMF were added HOBt-H_2O (8.8 mg, 0.046 mmol), EDCI-HCl (8.8 mg, 0.046 mmol) and **26** (from a 100 mM stock solution in DMF, 0.0421 mmol) at room temperature. After stirring at room temperature for 9 h, the reaction was quenched by addition of water (20 mL) and AcOEt (20 mL). Organic layer was washed with cold water (20 mL x 2), saturated aqueous NH_4Cl solution (20 mL x 2), saturated aqueous $NaHCO_3$ solution (20 mL x 2) and brine (10 mL x 1). The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo*. The residue was purified by GPC to give **22** as an orange amorphous (5.7 mg, 0.0105 mmol, 27%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.08 (s, 1H), 6.88 (d, *J* = 4.0 Hz, 1 H), 6.30-6.29 (m, 2H), 6.11 (s, 1H) 3.62-3.55 (m, 8H), 3.53-3.50 (m, 4H), 3.45-3.42 (m, 4H), 3.29 (t, *J* = 7.6 Hz, 2H), 2.63 (t, *J* = 7.6 Hz, 2H), 2.56 (s, 3H), 2.25 (s, 3H), 1.75 (q, *J* = 6.8 Hz, 2H), 1.57 (q, *J* = 6.8 Hz, 2H), 1.46-1.40 (m, 2H), 1.37-1.32 (m, 2H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 171.9, 160.1, 158.0, 143.8, 135.1, 133.5, 128.4, 123.8, 120.4,117.6, 71.3, 70.7, 70.6, 70.4, 70.2, 70.0, 45.1, 39.4, 36.0, 32.6, 29.8, 29.5, 26.8, 25.5, 24.9, 15.0, 11.4; FT-IR (neat) 3314, 2930, 2864, 1659, 1606, 1529, 1487, 1443, 1252, 1174, 1135 cm^{-1}; HRMS (ESI, Positive): *m/z* calced. for $C_{26}H_{39}BClF_2N_3O_4$ $[M+Na]^+$: 564.2582, found 564.2586.



Synthesis of HaloTag ligand conjugated coumarin (23). To a solution of **19** (8.4 mg, 0.023 mmol) in 1.0 mL of DMF were added **26** (from a 100 mM stock solution in DMF, 0.046 mmol) and Et_3N (from a 100 mM stock solution in DMF, 0.046 mmol) at room temperature. After stirring at room temperature for 3 h, the reaction was quenched by addition of water (20 mL) and AcOEt (20 mL). Organic layer was washed with cold water (20 mL x 2), saturated aqueous NH_4Cl solution (20 mL x 2), saturated aqueous $NaHCO_3$ solution (20 mL x 2) and brine (10 mL x 1). The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo*. The residue was purified by silica gel chromatography with AcOEt to give **23** as a yellow oil (11.9 mg, 0.023 mmol, 99%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 9.01 (brs, 1H), 8.69 (s, 1H), 7.42 (d, *J* = 9.0 Hz, 1 H), 6.64 (dd, *J* = 2.4, 9.0 Hz, 1 H) 6.49 (d, *J* = 2.3 Hz, 1H), 3.70-3.64 (m, 10H), 3.60-3.58 (m, 2H), 3.52 (t, *J* = 6.8 Hz, 2H), 3.47-3.43 (m,

6H), 1.77 (q, J = 6.8 Hz, 2H), 1.59 (q, J = 7.0 Hz, 2H), 1.47-1.41 (m, 2H), 1.39-1.33 (m, 2H), 1.24 (t, J = 7.0 Hz, 6H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 163.3, 162.6, 157.7, 152.6, 148.1, 131.2, 110.4, 109.9, 108.4, 96.6, 71.3, 70.8, 70.7, 70.6, 70.2, 69.9, 45.1, 39.5, 32.6, 29.5, 26.8, 25.5, 12.5; FT-IR (neat) 3341, 2969, 2933, 2866, 1703, 1618, 1584, 1531, 1515, 1418, 1352, 1230, 1189, 1135 cm^{-1} ; HRMS (ESI, Positive): *m/z* calced. for $C_{26}H_{39}ClN_2O_6$ $[M+Na]^+$: 533.2389, found 533.2388.

Photocatalyst screening. Preparation of photocatalyst **1**-**19** functionalized affinity beads and target protein labeling was operated according to our previous report[7]. Photocatalyst **2**-**19** was purchased from commercial sources. ATTO390-NHS (**2**, ATTO-TEC); ATTO425-NHS (**3**, ATTO-TEC); ATTO488-NHS (**6**, ATTO-TEC); ATTO495-NHS (7, ATTO-TEC); ATTO610-NHS (**8**, ATTO-TEC); ATTO465-NHS (**9**, ATTO-TEC); ATTO514-NHS (**10**, ATTO-TEC); ATTO520-NHS (**11**, ATTO-TEC); ATTO532-NHS (**12**, ATTO-TEC); ATTO565-NHS (**13**, ATTO-TEC); ATTO590-NHS (**14**, ATTO-TEC), BDP FL-NHS (**4**, Lumiprobe), BDP 558/568-NHS (**15**, Lumiprobe), BDP 581/591 (**16**, Lumiprobe), BDP RG6-NHS (**17**, Lumiprobe), BDP TMR-NHS (**18**, Lumiprobe), 7-(diethylamino)coumarin-3-carboxylicacid *N*-succinidimyl ester (**19**, invitrogen); 5-SFX (**5**, invitrogen).

Cell culture. HEK293FT cells were incubated with the RPMI-1640 medium, supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (Gibco; Thermo Fisher Scientific, Waltham, MA). Cells were incubated in a cell incubator with 5% CO_2 at 37 °C.

Plasmid preparation.

pFN21A/Halo-H2BA. The plasmid was purchased from Promega.

pcDNA3/V5-APEX2-H2B. Nuclear export signal (NES) sequence was removed from pcDNA3/V5-APEX2-NES (kindly provided from Prof. Hyun-Woo Rhee, Seoul National University) is amplified by PCR using primers (TAATAGCTCGAGCATGCATCTAGAGGGC and GGCATCAGCAAACCCAAGCTCG). H2B sequence was obtained from pFN21A/Halo-H2BA using primers (GAGCCAACCACTGAGGATCTGTAC and CATGGTGGCTTTGCTAGCCCTATAG). The obtained H2BA sequence and pcDNA3/V5-APEX2 fragment were ligated with In-Fusion HD cloning kit (Takara Bio Inc.).

pDisplay/CMV-V5-HaloTag-myc-TM. The vector, pDisplay/CMV-V5-APEX2-myc-TM (kindly provided from Prof. Hyun-Woo Rhee, Seoul National University) was amplified except APEX2 gene by PCR (primers: GAACAAAAACTCATCTCAGAAGAGG and GGTGCTGTCCAGGCCCAG) and HaloTag gene amplified by PCR (primers: GGCCTGGACAGCACCATGGCAGAAATCGGTACTG and GATGAGTTTTTGTTCCTCCATGGCGATCGCGTT) from pFN21/HaloTag-HOMX1 (addgene, FHC09676) was ligated into it with In-Fusion HD cloning kit.

pCI-HaloTag-NSL. HaloTag gene was amplified from pFN21A/Halo-H2BA vector by PCR. To introduce a C-terminal NSL (nuclear localization signal), we performed PCR with the NSL (PKKKRKV) codon primers. The obtained PCR product including HaloTag-NSL codon was subcloned into EcoRI/MluI sites of the pCI-neo vector (Promega Corporation, Madison, WI).

Transfect cells. Transfection of HaloTag-POI was operated according to protocol of Lipofectamine^{TM} 3000 reagent (invitrogen).

Protein labeling in cells. HaloTag-POI transfected 293FT cells were incubated on 24 well plate (2×10^5 cells/well) in CO_2 incubator for 24 h. These cells were treated with ligand conjugated photocatalyst (**20**-**23**, from a 10 mM stock solution in DMF; final 1 μM in medium) and incubated in CO_2 incubator for 2 h. HaloTag-POI transfected 293FT cells were transferred to centrifuging tube, discard supernatant, add PBS and centrifuge (100 × g, 5 min, 4 °C) to remove albumin derived from medium and unbound photocatalyst. After discarding the supernatant, to the cells added 500 μL of 500 μM labeling reagent solution in PBS. The cells were incubated at room temperature for 30 min, and operated photo-irradiation on ice for 1 min. After the labeling reaction, reaction was quenched by addition of 1 x SDS-PAGE sample buffer (50 mM Tris–HCl pH 6.8, 125 mM 2-mercaptoethanol, 2% sodium dodecyl sulfate (SDS), 0.025% bromophenol blue (BPB), 10% glycerol) and heated at 95 °C for 5 min. Proteins were separated by SDS-PAGE using 4-20% acrylamide gels (Biorad), transferred to polyvinylidene difluoride (PVDF) membrane (GE Healthcare), blocked with Immuno Block (DS Pharma), treated with horseradish peroxidase (HRP)-conjugated streptavidin (streptavidin-HRP, Sigma-Aldrich), and a blot was treated with Immobilon® Forte Western HRP Substrate (Millipore) and chemical luminescence was detected with a iBright CL 1500 | FL 1500 Imaging Systems (invitrogen). For checking of HaloTag, V5-Tag or Myc-Tag linked protein expression, tagged proteins were visualized by western blotting analysis (details are described-above) using anti-HaloTag pAb (rabbit IgG, promega), anti-V5-Tag pAb (rabbit IgG, proteintech) or MYC-tag rabbit pAb (rabbit IgG, proteintech).

Immunocytochemical study (ICC) of Acriflavine-HaloTag-H2B. Cells were seeded at a density of 8×10^5 cells/mL with media 0.1 mL on the collagen-coated cover glasses in 6 well plate dishes, and incubated for 2 h in CO_2 incubater. After incubation, to the cells was added 1.0 mL of additional medium and incubated for 12 h in CO_2 incubater. To the cells was added acriflavine **21** (from a 10 mM stock solution in DMF, final 1 μM) and incubated for 2 h in CO_2 incubater. The cells were gently washed once in the dish with 1 mL of PBS buffer, fixed with 4% paraformaldehyde (PFA) for 15 min, the cells were washed with PBS buffer once. Nuclei were counterstained with Hoechst33452. Fluorescence signals were observed using a confocal laser microscope (LMS780 spectral confocal system, Zeiss).

Protein purification from photo-irradiated cells. After photo-irradiated (protein-labeled) cells (3.0×10^6 cells) was added 1.0 mL of RIPA buffer (50 mM tris-HCl (pH 8.0), 150 mM NaCl, 0.5% sodium deoxycholate (SDC), 0.1% SDS, 1% NP-40, supplemented 1 mM DTT, 1000 U/mL Micrococcal nuclease (Takara-bio), 1 mM $CaCl_2$ and EDTA, EGTA-free protease inhibitor cocktail (Sigma-Aldrich). Then, the samples were incubated at 4 °C for 15 min, followed by incubation at 37 °C for 10 min. Precipitation was removed by centrifuging (16000 × g, 4 °C, 10 min) to obtain protein extraction. Dynabeads (MyOne StreptAvidin C1, 0.2 mg/sample) was washed with ultrapure water three times and added to protein extraction. After stirring for 4 h at 4 °C, beads were collected by magnetic separation and washed with RIPA buffer (1 mL x 1), wash buffer I (100 mM tris (pH 9.0), 12 mM sodium deoxy cholate (SDC), 12 mM sodium *N*-lauroyl sarcosinate (SLS), 1 mL x 3) wash buffer II (20 mM Tris buffer (pH 7.5), 150 mM NaCl, 1 M Urea, 1 mL x 3), wash buffer III (20 mM

tris buffer (pH 7.5), 1 M NaCl, 1 mL x 3) and 20 mM Tris buffer (pH 8.0, 1 mL x 2). Then, labeled protein binding beads was applied next step (Western-blotting or trypsin digestion for LC-MS/MS).

Western-blotting analysis of labeled histone. Desthiobiotin labeled proteins were eluted from beads (prepared in section "Protein purification from photo-irradiated cells") by adding 1 x SDS-PAGE sample buffer supplemented 1 mM biotin (from a 100 mM stock solution in DMSO) and heating at 95 °C for 5 min. The protein mixture was separated by SDS-PAGE, transferred to polyvinylidene difluoride (PVDF) membrane (GE Healthcare), blocked with Immuno Block (DS Pharma), treated with anti-H2A (rabbit, Active Motif), anti-H2B (rabbit, Active Motif), anti-H3 (rabbit, Active Motif) or anti-H4 (rabbit, Active Motif) and anti-rabbit-HRP (Santa Cruz Biotechnology), a blot was treated with Immobilon® Forte Western HRP Substrate (Millipore) and chemical luminescence was detected with a iBright CL 1500 | FL 1500 Imaging Systems (invitrogen).

Preparation of digested peptides for nano LC-MS/MS analysis. Desthiobiotin labeled proteins were enriched according to section "Protein purification from photo-irradiated cells". The beads were suspended in 100 μL of 20 mM Tris buffer (pH 8.0) and to the suspension was added trypsin (Promega, final 100 ng/sample). After incubation at 37 °C overnight, to the solution was added 10% TFA solution in ultrapure water (final 0.1% TFA). The solution was desalted using C18 pipette tips (Nikkyo Technos Co., Ltd.). Desalted solution was concentrated under reduced pressure, dissolved in 5% MeCN/0.1% TFA and applied to LC-MS/MS analysis.

Nano LC-MS/MS analysis. LC-MS/MS analysis was performed by nanoLC-ESI-MS system composed of a quadrupole-orbitrap hybrid mass spectrometer (Q-Exactive; Thermo Fisher Scientific) equipped with a nanospray ion source and a nano HPLC system (Easy-nLC 1000; Thermo Fisher Scientific). The trap column used for the nano HPLC was a 2 cm × 75 μm capillary column packed with 3 μm C18-silica particles (Thermo Fisher Scientific) and the separation column was a 12.5 cm × 75 μm capillary column packed with 3μm C18-silica particles (Nikkyo Technos Co., Ltd.). The flow rate of the nano HPLC was 300 nL/min. The separation was conducted using a 10−40% linear acetonitrile gradient at 70 min in the presence of 0.1% formic acid. The LC-MS/MS data were acquired in data-dependent acquisition mode controlled by Xcalibur 4.0 (Thermo Fisher Scientific). The settings of data-dependent acquisition were as follows: the resolution was 70,000 for a full MS scan and 17,500 for MS2 scan; the AGC target was 3.0E6 for a full MS scan and 5.0E5 for MS2 scan; the maximum IT was 60 ms for both a full MS scan and MS2 scan; the scan range was 310−1,500 m/z for a full MS scan and no fixation of first mass for MS2 scan; the top 10 signals were selected for MS2 scan per one full MS scan; only +2 ~ 4 charged MS1 signals was taken for MS2 scan; and dynamic exclusion was set as 15 sec. The measurement was done three times for each sample.

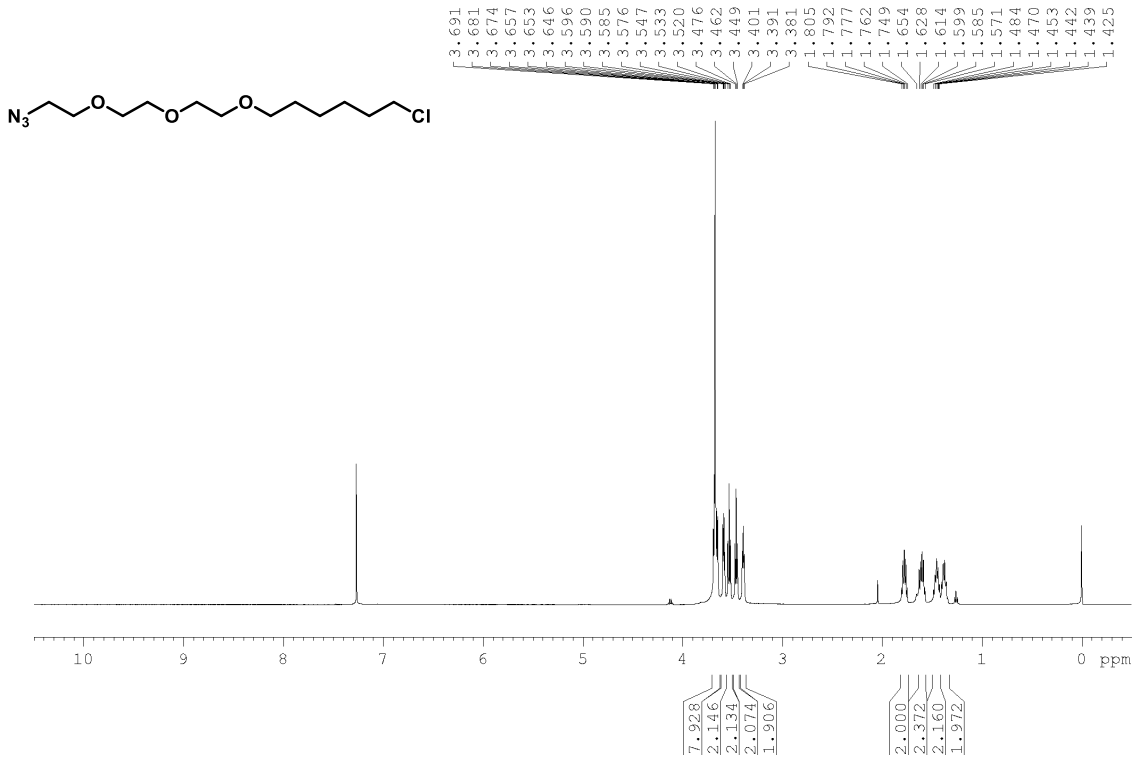
The identification and label-free quantification of the proteins were conducted by using the Proteome Discoverer 2.4 software embedded with the Sequest algorithm (Thermo Fisher Scientific). The list of the human proteins was obtained from the UniProt database (taxonomy 9606; downloaded on Feb. 29, 2016). The Percolator algorithm was used for the PSM validation. The settings of label-free quantification were as follows: "Normalization Mode" was set as "None", "Protein Abundance Calculation" was set as "Summed Abundance",

“Protein Ratio Calculation” was set as “Protein Abundance Based”, “Maximum Allowed Fold Change” was set as “100”, “Imputation Mode” was set as “None”, and “Hypothesis Test” was set as “ANOVA (Individual Proteins)”. Only the proteins whose FDR annotation was “High (< 1% FDR)” and Master annotation was “IsMasterProtein” were extracted and used for the calculation of the foldchange. The proteins whose detected peptide number (Number of Peptides) was only one was omitted.

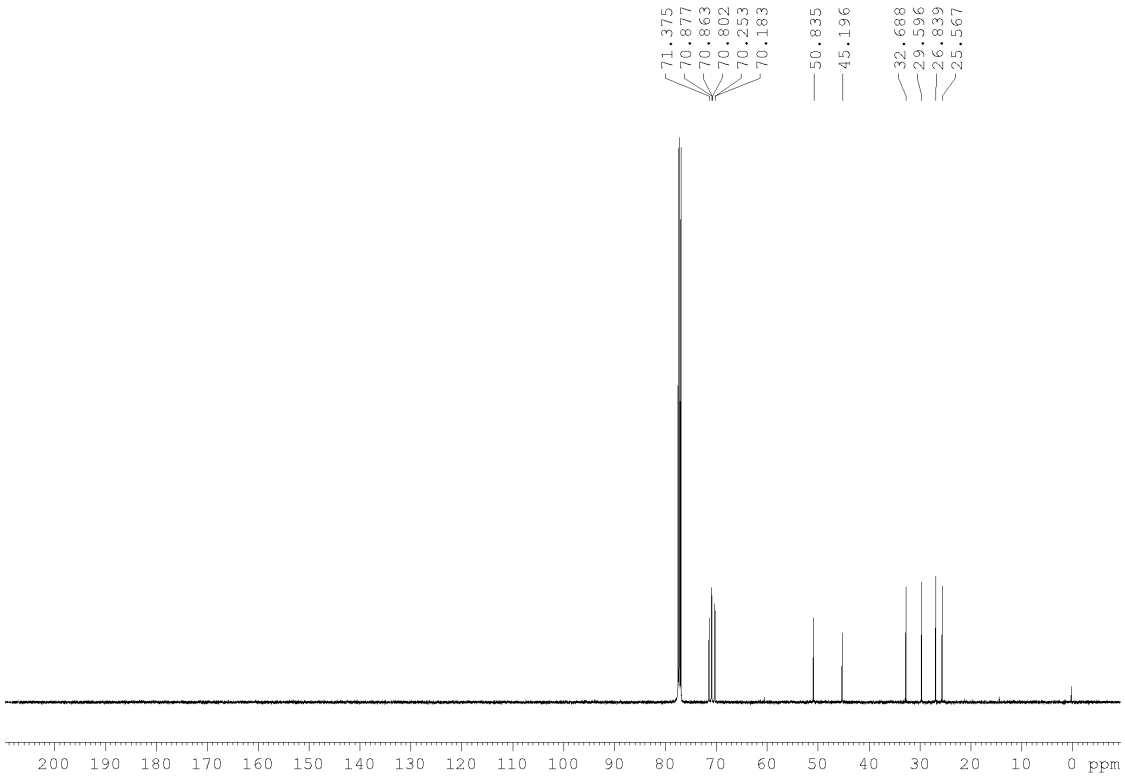
3. ^{1}H and ^{13}C NMR spectra of compounds

1-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)-6-chlorohexane (28).

^{1}H NMR (500 MHz, $CDCl_3$)

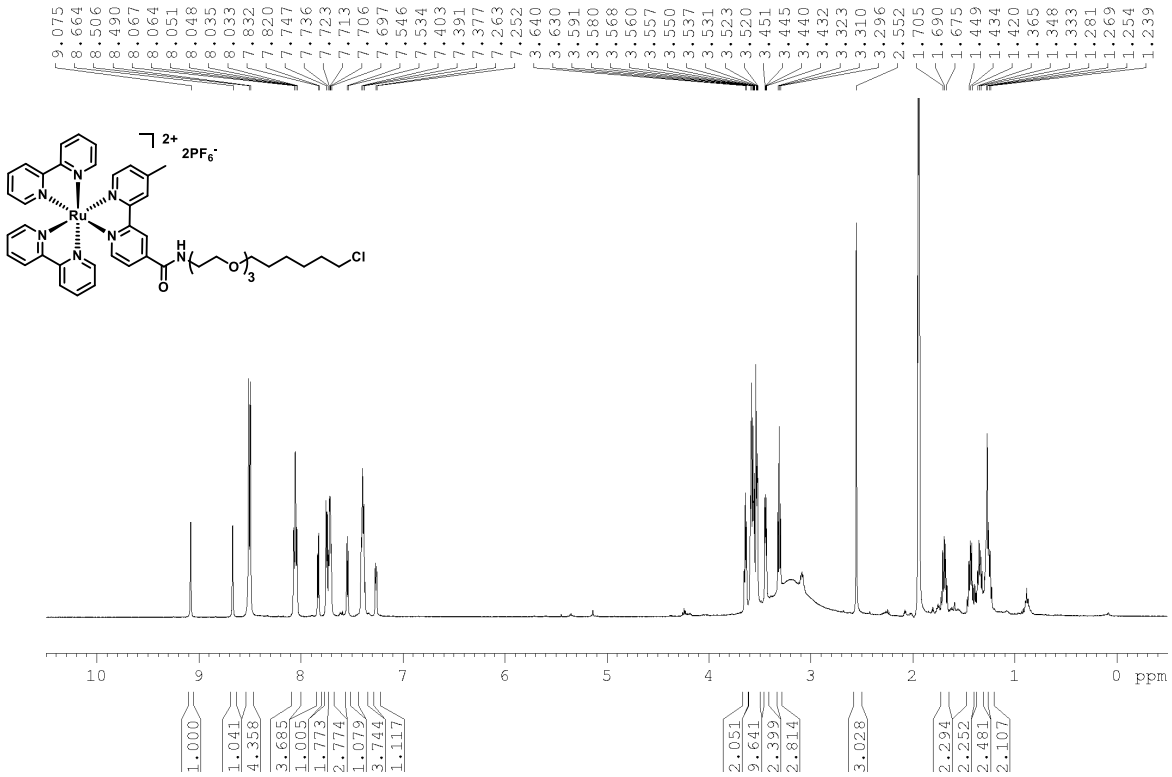


^{13}C NMR (125 MHz, $CDCl_3$)

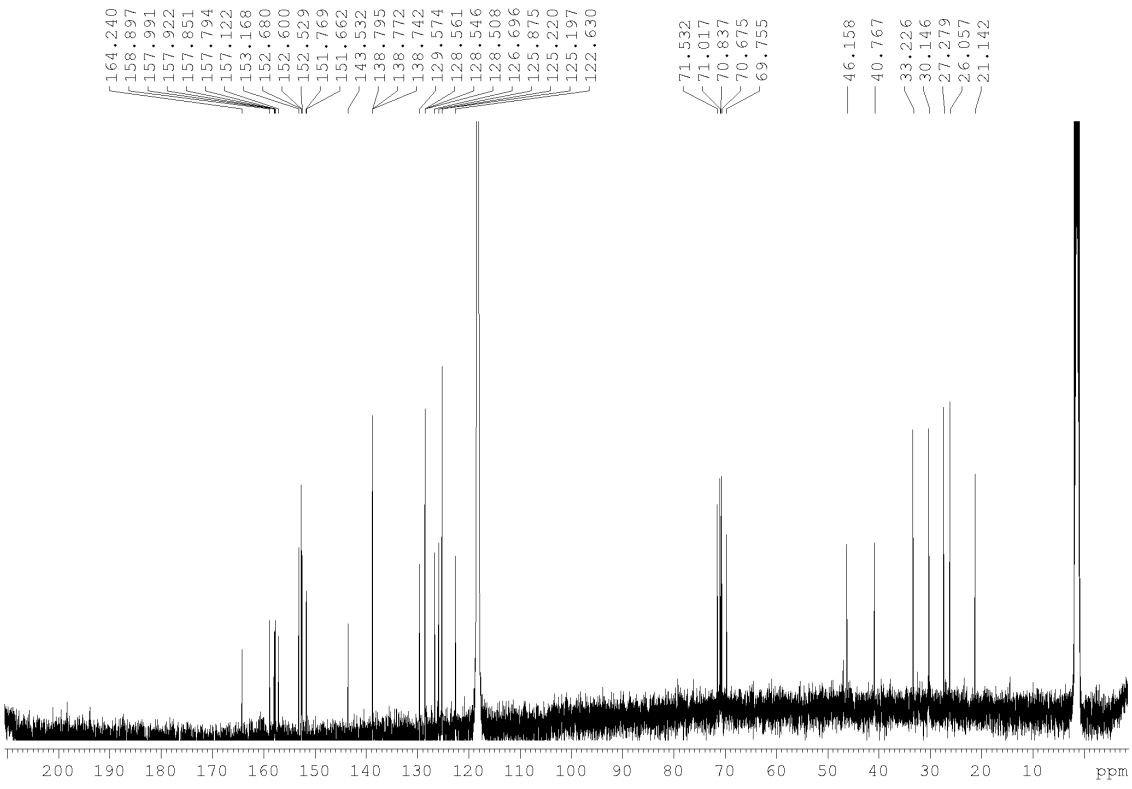


HaloTag ligand conjugated Ru(bpy)$_3$ complex (20)

^{1}H NMR (500 MHz, CD$_3$CN)

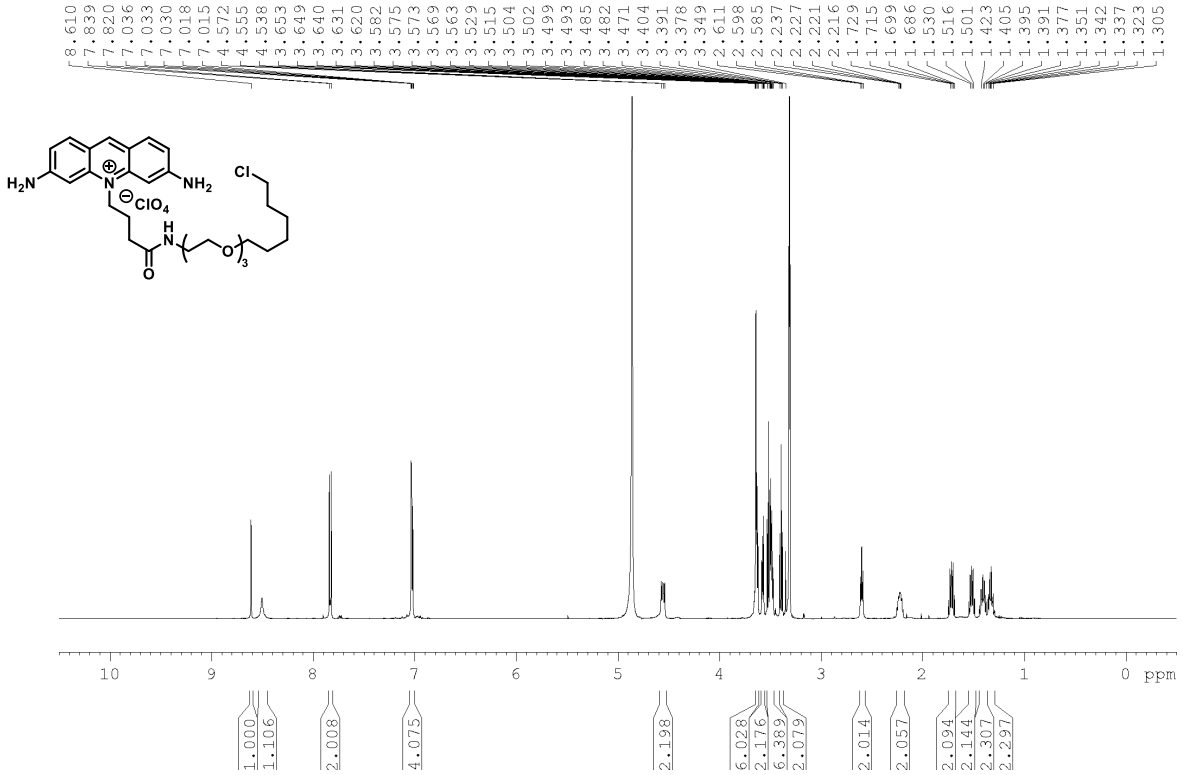


^{13}C NMR (125 MHz, CD$_3$CN)

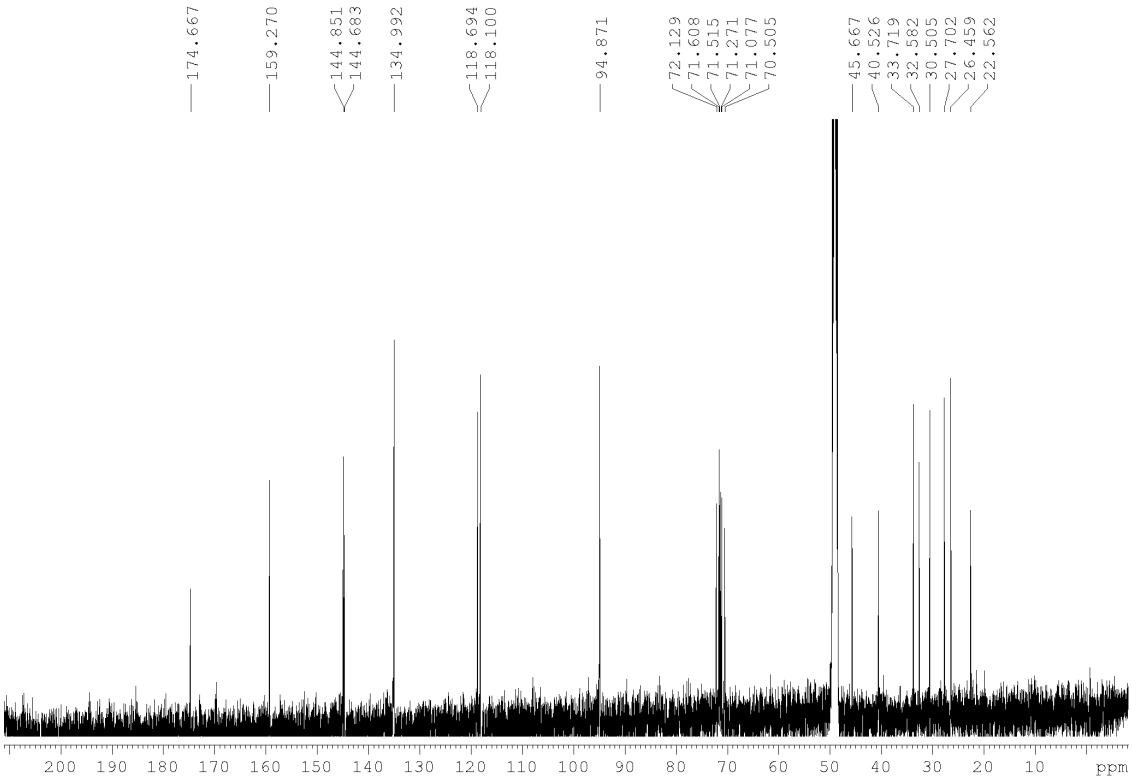


HaloTag ligand conjugated acriflavine (21)

^{1}H NMR (500 MHz, CD_3OD)

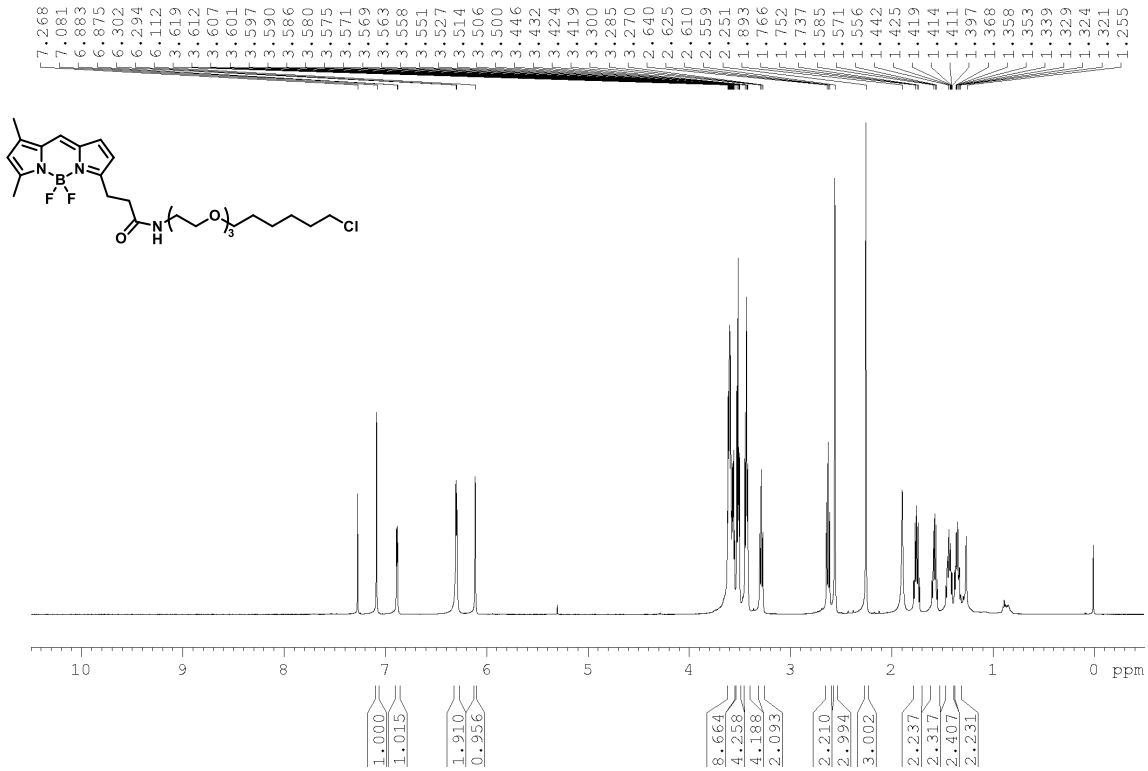


^{13}C NMR (125 MHz, CD_3OD)

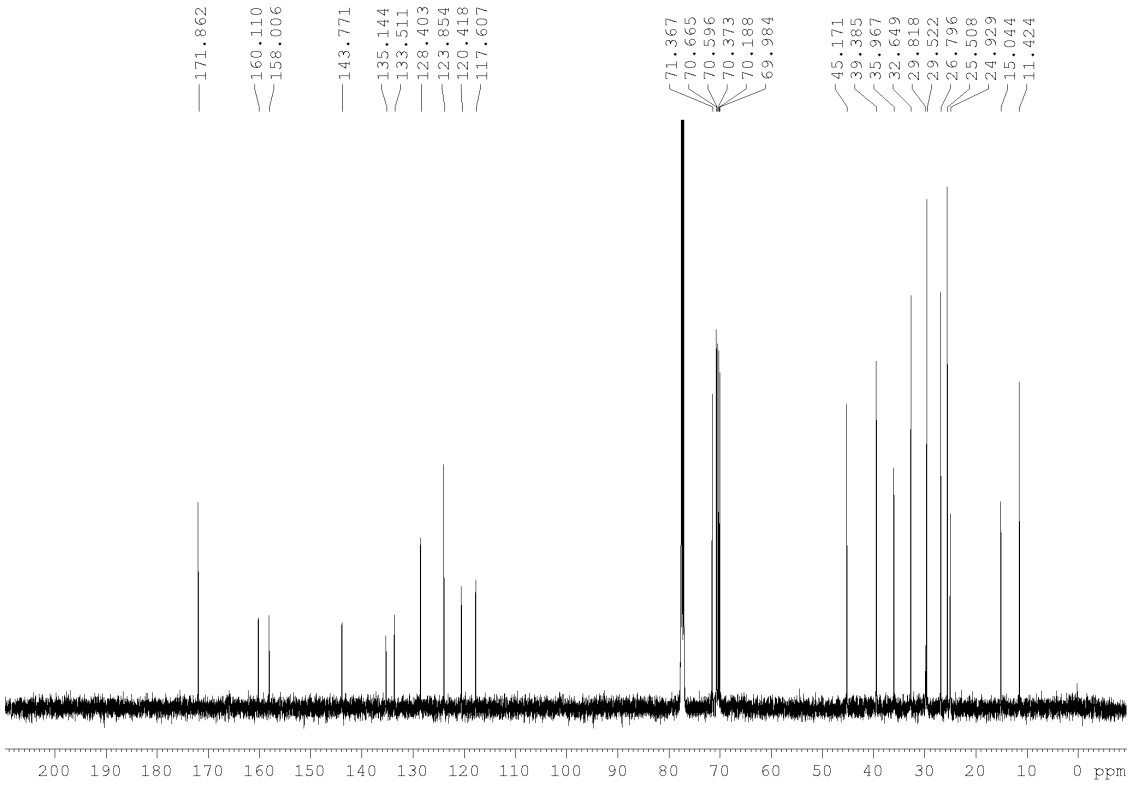


HaloTag ligand conjugated BODIPY (22)

^{1}H NMR (500 MHz, $CDCl_3$)

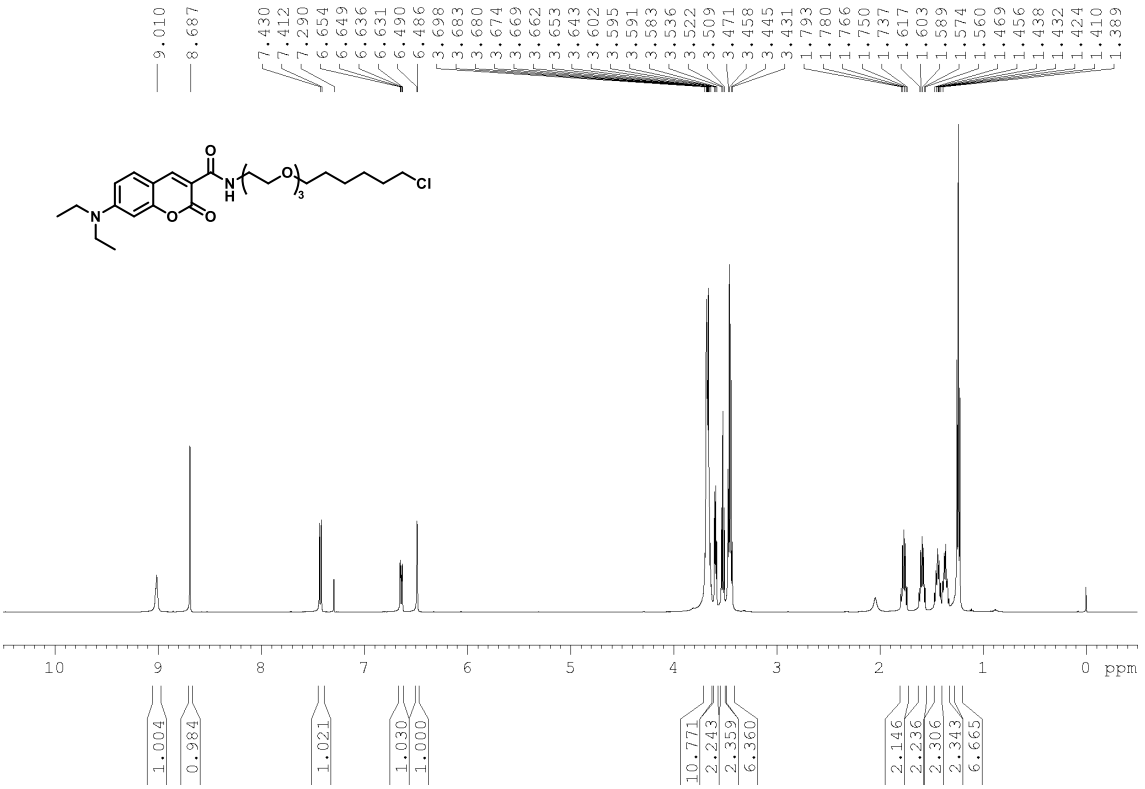


^{13}C NMR (125 MHz, $CDCl_3$)

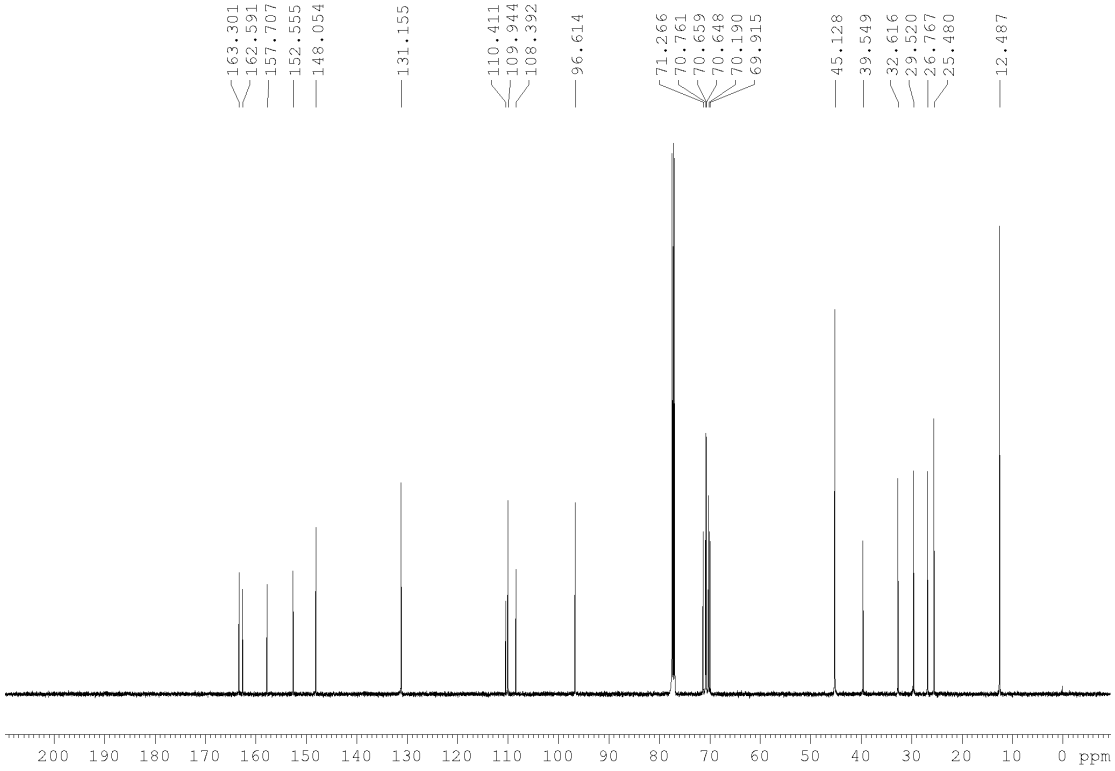


HaloTag ligand conjugated coumarin (23)

^{1}H NMR (500 MHz, $CDCl_3$)



^{13}C NMR (125 MHz, $CDCl_3$)



4. Reference

1. J.-P. Lambert, M. Tucholska, C. Go, J.D.R. Knight, A.-C. Gingras, *J. Proteomics,* **2015**, *118*, 81-94.
2. Y-. E. Kim, C. Park, K. E. Kim, K. K. Kim, *Biochem. Biophys. Res. Commun.,* **2018,** *499,* 30-36.
3. S. Sato and H. Nakamura, *Angew. Chem. Int. Ed.* **2013,** *52,* 8681-8684.
4. S. Sato, K. Hatano, M. Tsushima, H. Nakamura, *Chem. Commun.* **2018,** *54,* 5871-5874.
5. J. A. Walker, J. J. Bohn, F. Ledesma, M. R. Sorkin, S. R. Kabaria, D. N. Thornlow, and C. A. Alabi, *Bioconjugate Chem.* **2019,** *30,* 2452.
6. A. V. Saura, M. I. Burguete, F. Galindo and S. V. Luis, *Org. Biomol. Chem*., **2017**, *15,* 3013-3024.
7. M. Tsushima, S. Sato, K. Nakane, H. Nakamura, *Current Protocols*, **2020**, *101*, 1.