# Genetically Encoded Multispectral Labeling of Proteins with Polyfluorophores on a DNA Backbone

Vijay Singh, Shenliang Wang†, Eric T. Kool*

Department of Chemistry, Stanford University, Stanford, California 94305-5080, United States

## General

Chemicals and solvents were purchased from Sigma Aldrich, Acros Chemicals and Fisher Scientific, and were used without further purification. Chemicals used for the solid-phase synthesis of oligodeoxyfluorosides (ODFs) such as spacers, solid supports and synthesizer reagents were acquired from Glen Research. Bioreagents were purchased from Invitrogen, Fermentas and BioRad. $^{1}$H NMR and $^{13}$C NMR were taken on Varian Inova 400-MHz NMR spectrometers, unless otherwise noted. Chemical shifts are reported in ppm downfield from tetramethylsilane (TMS). Data are reported as follows: chemical shift, multiplicity: singlet (s), doublet (d), triplet(t), multiplet (m), coupling constants J in Hz, and integration. High-resolution mass spectrometry (HRMS) was performed by the Stanford University mass spectrometry facility. High performance liquid chromatography (HPLC) was performed on a LC-CAD Shimadzu liquid chromatograph, equipped with a SPD-M10A VD diode array detector and a SCL 10A VP system controller. ODF masses were determined by the Stanford University Protein and Nucleic Acid facility using a Perspective Voyager-DE RP Biospectrometry MALDI-TOF Mass-Spectrometry instrument with 3-hydroxypicolinic acid/diammonium hydrogen citrate as matrix. Absorption spectra were measured on a Varian Cary 100 Bio UV-Vis spectrometer. Fluorescence emission spectra were obtained on a Jobin Yvon-Spex Fluorolog 3 spectrometer with 344 nm excitation, unless otherwise indicated. Emission lifetime measurements were carried out on a PTI Easy Life instrument using a 385-nm light emitting diode. HeLa cells (CCL-2) were purchased from ATCC. The plasmid encoding genetically engineered haloalkane dehalogenase enzyme was purchased from Promega. A plasmid encoding α-tubulin was purchased from Addgene (Addgene plasmid 30487)[1]. The plasmid for transmembrane protein pDisplay was generously provided by Dr. S. Gambhir (Stanford University). Cellular imaging was performed using Leica sp5 confocal microscopy available in Cell Science Imaging Facility at Stanford University. Dulbecco's modified essential medium (DMEM), fetal bovine serum (FBS) and transfection reagent Lipofectamine 2000 were purchased from Invitrogen; 8-well Lab-Tek RII chamber cover glass was from Thermo Fisher (Nunc 155409). Protein Gel analysis was performed on UV transilluminator (Ex 365 nm) and Typhoon 9410 Gel scanner (Ex 457 nm)

## 1. Synthesis of chlorolinker phosphoramidite (B8)

### tert-butyl (2-(2-hydroxyethoxy)ethyl)carbamate (A2)

To a solution of 2-(2-aminoethoxy)ethanol **A1** (1.0 g, 9.56 mmol) in anhydrous EtOH (20 mL) was added Boc$_2$O (2.18 g, 9.56 mmol) at 0 °C. After stirring at room temperature 2 h, the reaction mixture was evaporated. The product was then extracted with $CH_2Cl_2$ (3x20 mL). The combined organic layers were dried over $Na_2SO_4$ and evaporated under vacuum to obtain the product **A2** as colorless oil (1.93 g, 99%), which was used for the next step without further purification.

$^1$HNMR: (400MHz, $CDCl_3$, δ): 5.28 (br, 1H), 3.68-3.52 (m, 2H), 3.52-3.48 (m, 4H), 3.28-3.26 (m, 2H), 1.39 (s, 9H). $^{13}$CNMR: (100MHz, $CDCl_3$, δ): 156.32, 79.41, 72.37, 70.37, 61.62, 40.42, 28.48. HRMS (ESI, m/z): $[M+Na]^+$ calcd for $C_9H_{19}NNaO_4$, 228.1212; found 228.1207.

**tert-butyl (2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)carbamate (A3)**

To a solution of **A2** (2.0 g, 9.75 mmol) in THF (14 mL) and DMF (7 mL) at 0 °C was added NaH (354 mg, 95% in mineral oil, 11.7 mmol). After stirring at 0 °C for 30 min, 6-chloro-1-iodohexane (2 mL, 13.65 mmol) was added to the above solution. The reaction mixture was stirred overnight and quenched with saturated $NH_4Cl$. The mixture was extracted with EtOAc (3x30 mL), washed with $H_2O$ and brine. The combined organic layers were dried over $Na_2SO_4$ and concentrated. The crude product was purified by silica gel column chromatography using EtOAc/Hexane (20:80 to 30/70, v/v) to yield pure product **A3** (2.17 g, 69%) as colorless oil.

$^1$H NMR: (400MHz, $CDCl_3$, δ): 5.03 (br, 1H), 3.59-3.57 (m, 2H), 3.55-3.49 (m, 6H), 3.43 (t, $J$ = 6.6 Hz,, 2H), 3.31-3.27 (m, 2H), 1.79-1.72 (m, 2H), 1.62-1.55 (m, 2H), 1.47-1.33 (m, 13H); $^{13}$CNMR: (100MHz, $CDCl_3$, δ): 156.13, 79.28, 71.39, 70.36, 70.32, 70.14, 45.16, 40.45, 32.65, 29.54, 28.54, 26.80, 25.53; HRMS (ESI, m/z): $[M+Na]^+$ calcd for $C_{15}H_{30}ClNNaO_4$, 346.1761; found 346.1752.

**2-(2-((6-chlorohexyl)oxy)ethoxy)ethanamine (A4)**

To a solution of the above product **A3** (1.2 g, 3.09 mmol) in 10 mL anhydrous $CH_2Cl_2$ at 0 °C was added TFA (3.3 mL, 43 mmol) portion wise. Thereafter reaction mixture was warmed to rt and stirred for 2h. After completion of reaction as evident by TLC analysis, the solvent was removed and the residue was treated with anhydrous $K_2CO_3$ (5.9 g, 43 mmol) in MeOH (20 mL). The mixture was filtered and the filtrate was concentrated to obtain crude product, which was then purified by column chromatography. Impurities were removed by EtOAc/Hexane (60/40, v/v) and product **A4** (0.61 g, 81%, as colorless oil) was eluted with 2-5% methanol in $CH_2Cl_2$.

$^1$H NMR (400MHz, $CDCl_3$, δ): 4.73 (brs, 2H), 3.74-3.72 (m, 2H), 3.65.3.63 (m, 2H), 3.60.3.58 (m, 2H), 3.53 (t, $J$ = 6.6 Hz, 2H), 3.49-3.45 (m, 2H), 3.19-3.18 (m, 2H), 1.81-1.73 (m, 2H), 1.63-1.56 (m, 2H), 1.49-1.40 (m, 2H), 1.38-1.30 (m, 2H); $^{13}$CNMR: (125 MHz, $CDCl_3$, δ): 71.54, 70.51, 70.06, 69.99, 45.31, 39.93, 32.70, 29.46, 26.86, 25.51; HRMS (ESI, m/z): $[M+H]^+$ calcd for $C_{10}H_{23}ClNO_2$, 224.1417; found 224.1414.

**2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethoxy)ethanol (B2)**

To a solution of diethylene glycol **B1** (2.0 g, 20.9 mmol) in 30 mL anhydrous $CH_2Cl_2$ at rt was added $Et_3N$ (8 mL, 57.1 mmol). After stirring for 15 min, DMT-Cl ((3.2 g, 9.5 mmol) was added slowly over 30 min. Upon completion of the reaction (~6 h) the reaction was quenched by addition of 20 mL water. The organic layer was separated and dried over $Na_2SO_4$. After evaporation of the solvents, the product was purified by column chromatography (ethyl acetate/hexanes 10:90% to 30:70%, v/v) to yield the product **B2** as a colorless oil (2.6 g, 67%).

$^1$H NMR (400MHz, $CDCl_3$, δ): 7.47-7.45 (m, 2H), 7.34 (d, $J$ = 8.5 Hz, 4H), 7.31-7.26 (m, 2H), 7.23-7.19 (m, 1H), 6.83 (d, $J$ = 8.5 Hz, 4H), 3.79 (s, 6H), 3.77-3.62 (m, 2H), 3.69-3.67 (m, 2H), 3.64-3.61 (m, 2H), 3.27-3.25 (m, 2H); $^{13}$C NMR (100MHz, $CDCl_3$, δ): 158.57, 145.10, 136.36, 130.20, 128.34, 127.95, 126.89, 113.22, 86.25, 72.42, 70.83, 63.34, 62.05, 55.37. HRMS (ESI, m/z): [M+Na]$^+$ calcd for $C_{25}H_{28}NaO_5$, 431.1834; found 431.1816.

**Methyl 3-(2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethoxy)ethoxy)propanoate (B3)**

To a stirred solution of DMT-diethylene glycol **B2** (1.5 g, 3.6 mmol) and methyl acrylate (1.32 mL, 14.7 mmol) in 30 mL anhydrous THF was added catalytic amount of NaH (3x10 mg) at 0 $^o$C. The reaction was continued for 1 h at 0 $^o$C. After completion, the reaction was quenched by saturated solution of $NH_4Cl$ (2.0 mL). The organic layer was separated and dried over $Na_2SO_4$. After evaporation of the solvents, the product was purified by column chromatography (ethyl acetate/hexanes 10:90% to 20:80% having 1% $Et_3N$ in both cases) to yield the product **B3** as colorless oil (1.36 g, 75%). Due to trans-esterification, a small amount of acrylate ester of DMT-diethylene glycol was also isolated as a side product (see below).

$^1$H NMR (500MHz, $CDCl_3$, δ): 7.50-7.48 (m, 2H), 7.38-7.36 (m, 4H), 7.32-7.28 (m, 2H), 7.24-7.21 (m, 1H), 6.85-6.84 (m, 4H), 3.82-3.80 (m, 8H), 3.70-3.66 (m, 9H), 3.25 (t, , *J* = 4.2 Hz, 2H), 2.63 (t, , *J* = 5.2 Hz, 2H); $^{13}$C NMR (100MHz, $CDCl_3$, δ): 172.32, 158.61, 145.32, 136.58, 130.31, 128.45, 127.99, 126.89, 113.27, 86.16, 71.00, 70.92, 70.80, 66.92, 63.39, 55.46, 51.94, 35.15. HRMS (ESI, m/z): [M+Na]$^+$ calcd for $C_{29}H_{34}NaO_7$, 517.2202; found 517.2195.

**2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethoxy)ethyl acrylate (side product, trans-esterification product)**

$^1$H NMR (500MHz, $CDCl_3$, δ): 7.50-7.48 (m, 2H), 7.40-7.37 (m, 4H), 7.31-7.28 (m, 2H), 7.24-7.21 (m, 1H), 6.86-6.83 (m, 4H), 6.48, 6.44 (dd, $J_1$ = 1.4 Hz, $J_2$ = 17.4 Hz, 1H ), 6.21, 6.18 (dd, $J_1$ = 10.5 Hz, $J_2$ = 17.4 Hz, 1H ), 6.86, 6.84 (dd, $J_1$ = 1.4 Hz, $J_2$ = 11.3 Hz, 1H ), 4.39-3.37 (m, 2H), 3.82-3.80 (m, 8H), 3.71 (t, , *J* = 5.1 Hz, 2H), 3.26 (t, , *J* = 5.1 Hz, 2H); $^{13}$C NMR (125MHz, $CDCl_3$, δ): 166.49, 158.62, 145.29, 136.52, 130.31, 128.44, 128.02, 126.92, 113.29, 86.20, 71.02, 69.41, 64.13, 63.39, 55.46, HRMS (ESI, m/z): [M+Na]$^+$ calcd for $C_{28}H_{30}NaO_6$, 485.1940; found 485.1925.

**3-(2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethoxy)ethoxy)propanoic acid (B4)**

To a stirred solution of compound **B3** (1.1 g, 2.22 mmol) in MeOH/$H_2O$ (14 mL/4mL, v/v) was added LiOH (456mg, 11.1 mmol) at rt. After completion of reaction (~2 h) as evident by TLC analysis, side products were extracted with ether while reaction medium was basic. Thereafter the reaction mixture was neutralized (pH 6-7) with 1N $H_2SO_4$. Then product was extracted with EtOAc (3x50 mL). The organic layer was dried over $Na_2SO_4$ and evaporated to yield product **B4** (0.96 g, 93%) as colorless oil. The product was used for the next step without further purification.

$^1$H NMR(400MHz, $CDCl_3$, δ): 7.46-7.44 (m, 2H), 7.34-7.32 (m, 4H), 7.28-7.24 (m, 2H), 7.20-7.16 (m, 1H), 6.82-6.79 (m, 4H), 3.77-3.74 (m, 8H), 3.65-3.61(m, 6H), 3.21 (t, , *J* = 5.2 Hz, 2H), 2.55 (t, , *J* = 6.9 Hz, 2H); $^{13}$C NMR (100MHz, $CDCl_3$, δ): 176.78, 158.56, 145.29, 136.55, 130.27, 128.42, 127.96, 126.88, 113.25, 86.14, 70. 91, 70.83, 70.42, 63.36, 55.42, 36.85. HRMS (ESI, m/z): [M+Na]$^+$ calcd for $C_{28}H_{32}NaO_7$, 503.2046; found 503.2029.

**2,5-dioxopyrrolidin-1-yl 3-(2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethoxy)ethoxy) propanoate (B5)**

N-hydroxysuccinimide (0.79 g, 6.87 mmol) and compound **B4** (3.3 g, 6.87 mmol) were dissolved in anhydrous $CH_2Cl_2$ (30 mL) and cooled to $0^oC$. DCC (1.56 g, 7.56 mmol) was added to the resulting reaction mixture at 0 $^oC$. The reaction was then warmed to rt and stirred overnight. After completion of reaction as evident by TLC analysis, the reaction mixture was diluted with hexane and filtered. The filtrate was concentrated under vacuum to yield crude product (3.76g, 95%) as sticky white solid, which was utilized for the next step without further purification.

$^1$H NMR (500MHz, $CDCl_3$, $\delta$): 7.49-7.43 (m, 2H), 7.38-7.35 (m, 4H), 7.31-7.28 (m, 2H), 7.23-7.20 (m, 1H), 6.86-6.83 (m, 4H), 3.89 (t, *J* = 6.5 Hz, 2H), 3.81(s, 6H), 3.70-3.68 (m, , 6H), 3.25 (t, *J* = 5.2 Hz, 2H), 2.92 (t, *J* = 6.5 Hz, 2H), 2.82 (brs, 4H); $^{13}$C NMR (125MHz, $CDCl_3$, $\delta$): 169.21, 166.99, 158.60, 145.32, 136.59, 130.31, 128.45, 127.99, 126.90, 113.27, 86.16, 71.07, 71.03, 70.93, 66.06, 63.39, 55.47, 34.08, 25.81. HRMS (ESI, m/z): $[M+Na]^+$ calcd for $C_{32}H_{35}NNaO_9$, 600.2210; found 600.2206.

**3-(2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethoxy)ethoxy)-N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)propanamide (B6)**

The product **B5** (5.5 g, 9.53 mmol) and chloro linker **A4** (2.12 g, 9.53 mmol) were dissolved in anhydrous $CH_2Cl_2$ (25 mL). DIPEA (3.41 g, 19.0 mmol) was added to the resultant mixture at rt. The reaction mixture was stirred overnight. After completion of reaction, product was extracted with $CH_2CH_2$. The combined organic layers were dried over $Na_2SO_4$ and evaporated to dryness. The crude product was purified by column chromatography. Initially impurities were removed with EtOAC/Hexane (50:50 to 80:20 v/v +1% $Et_3N$) and finally product was eluted with 2-4% MeOH in $CH_2Cl_2$ (having 1% $Et_3N$ all the time) to yield product **B6** (5.35 g, 82%) as yellowish oil.

$^1$H NMR (500MHz, $CDCl_3$, $\delta$): 7.48-7.46 (m, 2H), 7.37-7.34 (m, 4H), 7.30-7.27 (m, 2H), 7.23-7.20 (m, 1H), 6.85-6.82 (m, 4H), 6.66 (brs, 1H), 3.80-3.76 (m, 8H), 3.70-3.65 (m, 6H), 3.60-3.51 (m, , 8H), 3.47-3.40 (m. 4H), 3.25 (t, , *J* = 5.2 Hz, 2H), 2.48 (t, , *J* = 5.95 Hz, 2H), 1.80-1.75 (m, 2H), 1.63-1.57 (m, 2H), 1.49-1.43 (m, 2H), 1.40-1.34 (m, 2H); $^{13}$C NMR (125MHz, $CDCl_3$, $\delta$): 171.66, 158.62, 145.25, 136.49, 130.29, 128.42, 128.01, 126.94, 113.28, 86.20, 70. 97, 70.68, 70.26, 67.55, 63.37, 55.46, 45.31, 37.25, 32.76, 29.69, 26.93, 25.66. HRMS (ESI, m/z): $[M+Na]^+$ calcd for $C_{38}H_{52}ClNNaO_8$, 708.3279; found 708.3276.

**N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-3-(2-(2-hydroxyethoxy)ethoxy)propanamide (B7)**

To the compound **B6** (4.3 g, 6.27 mmol ) in a flask was added 30 mL of a mixture of acetic acid and water (60:40, v/v) at room temperature. The reaction was continued for 2 h and after completion, product was extracted with EtOAc (30x6). The combined organic layers were dried over $Na_2SO_4$ and evaporated to dryness. The product was purified by column chromatography: impurities were removed with EtOAc/Hexane (50:50 ,v/v) and product was eluted with 5-10% methanol in $CH_2Cl_2$ to yield product (1.63 g, 68%) sticky white solid.

$^1$H NMR (400MHz, $CDCl_3$, $\delta$): 6.81 (brs, 1H), 3.74-3.69 (m, 4H), 3.62 (s, 4H), 3.58-3.49 (m, , 10H), 3.45-3.41 (m. 4H), 2.45 (t, , *J* = 5.9 Hz, 2H), 1.78-1.71 (m, 2H), 1.59-1.54 (m, 2H), 1.47-1.39 (m, 2H), 1.38-1.30 (m, 2H). $^{13}$C NMR (100MHz, $CDCl_3$, $\delta$): 171.69, 72.83, 71.46, 70.48, 70.43, 70.42, 70.26, 70.18, 67.42, 61.85, 45.26, 39.40, 37.22, 29.62, 26.88, 25.60. HRMS (ESI, m/z): $[M+Na]^+$ calcd for $C_{17}H_{34}ClNNaO_6$, 406.1972; found 406.1960.

**22-chloro-9-oxo-3,6,13,16-tetraoxa-10-azadocosyl (2-cyanoethyl) diisopropylphosphoramidite (B8)**

The compound **B7** (600 mg, 1.56 mmol) was dissolved in anhydrous $CH_2Cl_2$ (20 mL). The reaction mixture was degassed several times and purged with argon gas. The reaction was cooled with the help of ice bath and then diisopropylethylamine (1.39 mL, 7.8 mmol) was added. The reaction mixture was stirred for 10 min at 0 $^oC$, and thereafter 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (406 mg, 1.72 mmol) was added under argon atmosphere. The reaction was continued for an additional 45 min at 0 $^oC$. After completion of reaction as confirmed by TLC, solvent was evaporated and the residue was purified by silica gel column chromatography (Hex/EtOAc 10:90 + 1% TEA, under continuous stream of nitrogen) Evaporation of volatiles provided 900 mg of compound **B8** (98.5%) as colorless oil.

$^1H$ NMR (400MHz, $CDCl_3$, δ): 6.59 (brs, 1H), 3.92-3.70 (m, 6H), 3.68-3.66 (s, 4H), 3.65-3.60 (m, , 6H), 3.59-3.54 (m. 6H), 3.49-3.45 (m. 4H), 2.68-2.66 (m. 2H), 2.48 (t, , *J* = 4.8 Hz, 2H), 1.82-1.76 (m, 2H), 1.65-1.59 (m, 2H), 1.47-1.44 (m, 2H), 1.40-1.37 (m, 2H), 1.20 (t, , *J* = 5.0 Hz, 12H ). $^{31}P$ NMR (100MHz, $CDCl_3$, δ): 149.57. HRMS (ESI, m/z): $[M+Na]^+$ calcd for $C_{26}H_{51}ClN_3NaO_7P$, 606.3051; found 606.3044.

## 2. Synthesis of monomer F

The synthesis of compounds **2, 3, 4** was reported in US2010072888. All spectral data agreed with the reference.[2]

**2,7-dibromo-9,9-dimethyl-fluorene (2)**

2,7-dibromo-fluorene (15 g), potassium hydroxide (10.3 g) and a catalytic amount of potassium iodide (0.77 g) were stirred in dimethylsulforxide (60 mL). Iodomethane (16.4 g) was added dropwise for 30 min. The reaction was stirred at room temperature for 24 hrs and then poured into 500 mL water. The product was extracted with dichloromethane and the combined organic layers were evaporated to dryness. The residue was subjected to a flash column chromatography ($CH_2Cl_2$) and the product was used directly in the following step. Yield 97%.

**7-bromo-9,9-dimethyl-fluorene-2-carbaldehyde (3)**

2,7-dibromo-9,9-dimethylfluorene (16 g) was stirred in THF (50 mL) at -78 °C. n-Butyllithium (18.2 mL, 2.5 M) was added via syringe. After 1 hr, DMF (3.5 mL) was added via syringe. The reaction was stirred for 2 h more and allowed to warm to room temperature. The reaction mixture was poured into HCl (50 mL, 2 M) and extracted with ethyl acetate. The combined organic layer was evaporated under reduced pressure and the product purified by column chromatography (10% $CH_2Cl_2$/hexanes). Yield 90%. HRMS: calculated $C_{16}H_{14}OBr$ [M+H] 301.0223, found 301.0219.

**7-(diphenylamino)-9,9-dimethyl-fluorene-2-carbaldehyde (4)**

7-bromo-9,9-dimethyl-fluorene-2-carbaldehyde (16 g), diphenylamine (13.5 g) and cesium carbonate (43.2 g) were dissolved/suspended in toluene (40 mL) and purged with argon. Palladium acetate (1.2 g) and tri-tert-butyl phosphine (2.14 g) were mixed in toluene (10 mL), purged with argon and stirred for 30 min. The resulting mixture was added into the previous solution via syringe. The reaction was heated under reflux for 24 h. After cooling, the solvent was evaporated under reduced pressure. The residue was partitioned between water and ethyl acetate. The combined organic layer was condensed and subjected to column chromatography (50% $CH_2Cl_2$ /hexanes to $CH_2Cl_2$). Yield 67%. HRMS: calculated $C_{28}H_{24}NO$ [M+H] 390.1852, found 390.1850.

**2-amino-5-bromobenzenethiol (5)**

2-Amino-6-bromobenzothiazole (5 g) was added to a solution of potassium hydroxide (25 g) in water (50 mL). The turbid mixture turned clear after short heating and the resulting solution was heated under reflux for 24 h. After cooling, the solution was neutralized by acetic acid and then extracted with toluene. The combined organic phase was condensed to dryness, subjected to column chromatography (20% ethyl acetate/hexane) to furnish the product as a light yellow solid. Yield 56%, ESI-MS: m/z 204, 206 [M+H]+. The product was used in the following steps without further purification.

**7-(6-bromobenzothiazol-2-yl)-9,9-dimethyl-N,N-diphenyl-fluoren-2-amine (6)**

2-Amino-5-bromobenzenethiol **5** (0.4 g), 7-(diphenylamino)-9,9-dimethyl-fluorene-2-carbaldehyde **4** (0.6 g), and a catalytic amount of p-toluenesulfonic acid was mixed in toluene (50 mL) and the mixture was heated under reflux for 2 days. After cooling, the solvent was evaporated and the product was purified by column chromatography (25% $CH_2Cl_2$ /hexanes to 60% $CH_2Cl_2$ /hexanes) to yield 0.85g yellow solid. Yield 80%.

$^1$HNMR (400 MHz, $CDCl_3$, $\delta$): 8.127-8.123 (d, 1H), 8.023-8.018 (d, 1H), 7.982-7.958 (dd, 1H), 7.919-7.896 (d, 1H), 7.699-7.678 (d, 1H), 7.610-7.565 (m, 2H), 7.278-7.250 (m, 4H), 7.201-7.196 (d, 1H), 7.163-7.133 (m, 4H), 7.070-7.028 (m, 3H), 1.482 (s, 6H). $^{13}$CNMR: (100 MHz, $CDCl_3$, $\delta$): 169.467, 156.143, 154.573, 153.368, 148.658, 147.976, 142.759, 136.813, 132.703, 131.277, 130.011, 129.556, 129.450, 127.554, 124.733, 124.688, 124.324, 124.233, 123.307, 123.088, 121.654, 121.548, 119.971, 118.667, 118.022, 47.315, 27.227. HRMS (ESI, m/z): $[M+H]^+$ calculated $C_{34}H_{26}N_2SBr$ 573.0995, found 573.0989.

**5-(2-(7-(diphenylamino)-9,9-dimethyl-fluoren-2-yl)benzothiazol-6-yl)-2-(hydroxymethyl)tetrahydrofuran-3-ol (7)**

Compound **6** (0.8 g, 1.4 mmol), 3’-O-(tert-butyldiphenylsilyl)-1,2-dehydro-2-deoxy-ribofuranose[3] (0.5 g, 1.4 mmol), N-methyldicyclohexylamine (0.45 mL, 2.1 mmol), bis(tri-tert-butylphosphine)palladium(0) (70 mg, 0.14 mmol) and tetrabutylammonium bromide (0.67 g, 2.1 mmol) were charged into a flame-dried round-bottom flask equipped with a septum. Anhydrous dioxane (30 mL), previously bubbled with argon for 1 h, was added via syringe. The mixture was bubbled with argon for 1 hr and heated at 90 °C for 36 h. After cooling, the mixture was concentrated to dryness and then partitioned between water and ethyl acetate. The organic layer was evaporated to dryness and then dissolved in THF (30 mL). After cooling to 0 $^o$C, acetic acid (0.5 mL) and tetrabutylammonium fluoride (1M, 2.1 mL) were sequentially added. The reaction mixture was stirred at 0 °C for 1 h. After concentration under vacuum, the residue was partitioned between water and ethyl acetate. The organic layer was evaporated to dryness and then dissolved in THF (20 mL) and acetonitrile (30 mL). After cooling to -10 °C, acetic acid (1 mL) and sodium triacetoxyborohydride (0.45 g, 2.1 mmol) were added. The reaction mixture was stirred at -10 °C for 1 h. Ammonium hydroxide (10 mL) was then added dropwise to quench the reaction. The mixture was concentrated and then partitioned between water and ethyl acetate.

After drying with sodium sulfate, the organic layer was evaporated under vacuum and the residue was purified by column chromatography ($CH_2Cl_2$ to 5% MeOH/$CH_2Cl_2$) yielding a yellow foam. Yield: 25%.
$^1$HNMR (400 MHz, $CDCl_3$, δ): 8.147 (s, 1H), 8.007-7.986 (d, 1H), 7.908-7.888 (m, 2H), 7.624-7.604 (d, 1H), 7.563-7.542 (d, 1H), 7.485-7.461 (d, 1H), 7.273-7.234 (m, 2H), 7.190 (s, 1H), 7.138-7.119 (m, 5H), 7.049-6.999 (m, 4H), 5.339-5.300 (m, 1H), 4.602 (m, 1H), 4.197 (m, 1H), 3.906-3.788 (m, 2H), 2.446-2.402 (m, 1H), 2.083-2.008 (m, 1H), 1.464 (s, 6H). HRMS (ESI, m/z): $[M+H]^+$ calculated $C_{39}H_{35}N_2O_3S$ 611.2363, found 611.2362.

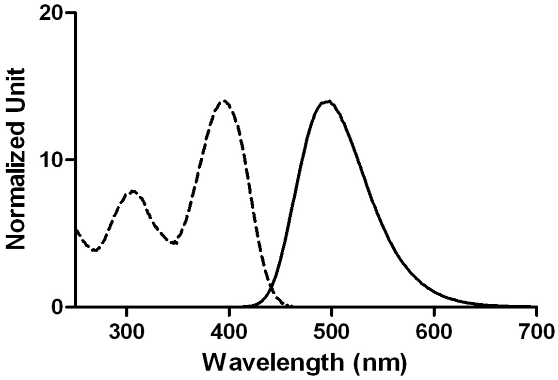


**Figure S1:** Absorption and fluorescence emission spectra of monomer **F** (Compound **7**) in methanol (Excitation 396; QY 0.37)

**(2R,3R,5R)-2-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-5-(2-(7-(diphenylamino)-9,9-dimethyl-9H-fluoren-2-yl)benzo[d]thiazol-6-yl)tetrahydrofuran-3-ol (8)**
The diol nucleoside **7** (500 mg, 0.82 mmol) was coevaporated with pyridine (25 mL) 3 times and then dried on vacuum pump for 4 h. DIPEA (3 mL, 17.2 mmol), 4,4'-dimethoxytrityl chloride (2 g, 5.9 mmol) and anhydrous pyridine (15 mL) were added sequentially. The reaction was stirred under argon in the dark and was monitored by TLC. Upon completion of reaction (~3 h), methanol (5 mL) was added to quench the remaining 4,4'-dimethoxytrityl chloride. After evaporation of the solvents, the product was purified by column chromatography (ethyl acetate/hexanes 1:1) to yield the product as a yellow solid (613 mg, 82%).
$^1$HNMR (500 MHz, $CDCl_3$, δ): 8.14-8.13 (d, *J* = 1.5 Hz, 1H), 8.02-8.00 (d, *J* = 8.4 Hz, 1H), 7.97 (m, 2H), 7.70-7.69 (d, *J* = 7.9 Hz, 1H), 7.61-7.60 (d, *J*= 8.2 Hz, 1H), 7.48 (m, 3H), 7.37(m, 4H), 7.29 (m, 6H), 7.24-7.22 (d, *J* = 7.2 Hz, 1H),7.20-7.19 (d, *J* = 2.0 Hz, 1H), 7.16-7.14 (d, *J* = 7.6 Hz, 4H), 7.05 (m, 3H), 6.84-6.83 (d, *J* = 8.8 Hz, 4H), 5.33 (m, 1H), 4.48 (m, 1H), 4.13 (m, 1H), 3.79 (s, 3H), 3.78 (s, 3H), 3.43-3.32 (m, 2H), 2.34 (m, 1H), 2.14 (m, 1H), 1.49 (s, 6H). 13C-NMR: (100 MHz, $CDCl_3$, δ): d =169.05, 158.75, 156.15, 154.52, 154.01, 148.53, 148.03, 145.06, 142.40, 139.39, 136.28, 136.24, 135.36, 132.91, 131.80, 130.38, 130.35, 129.57, 128.47, 128.16, 127.51, 127.12, 124.90, 124.70, 123.26, 123.17, 122.92, 121.62, 121.50, 119.95, 119.12, 118.16, 113.43, 86.72, 86.58, 80.20, 74.98, 66.14, 64.79, 55.50, 47.33, 44.48, 27.26, 15.55. HRMS (ESI, m/z): $[M+H]^+$ calculated $C_{60}H_{53}O_5N_2S$ 913.3670; found 913.3693.

**3. ODF-HaloTag Ligand Synthesis:** The ODF-HaloTag ligands were synthesized on Applied Biosystems 394 DNA/RNA synthesizer, using 3'-phosphate CPG column at 1 µmol scale with DMT-off method. Coupling of each monomer used standard 3' to 5' cyanoethyl phosphoramidite chemistry with extended coupling time (999 s). The oligomers were cleaved (from CPG resin) and deprotected by overnight incubation with 0.05M $K_2CO_3$ in methanol (Glen Research). The purification was carried out utilizing a Shimadzu Series HPLC with an Alltech C5 column with acetonitrile and TEAA buffer (100 mM, pH 7.2) as eluents. The identities of ODF-HaloTag ligands were confirmed by MALDI-TOF-MASS analysis.

## 4. Optical Properties of ODF-HaloTag Ligands

**a) Absorption and Fluorescence Analysis**: Absorption and fluorescence emission spectra were measured in PBS at 2.0 μM ODF-HaloTag ligand concentration. Absorption measurements were carried out on a Varian Cary 100 Bio UV Visible spectrophotometer. Fluorescence emission spectra were measured on a Jobin Yvon-Spex Fluorolog 3 spectrophotometer by exciting ODFs at 344 nm and collecting all the emission between 365 nm to 750 nm. The fluorescence emission spectra of protein-ODF conjugates were performed on FLEXstation II-384-fluorescent plate reader in the High-Throughput Bioscience Center (HTBC), Stanford University.

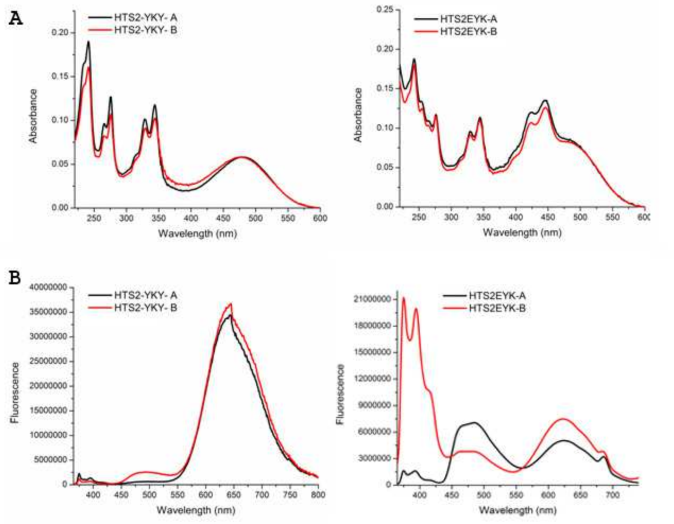


**Figure S2:** Photophysical properties of anomers of $htS_2EYK$ and $htS_2YKY$. A) Absorption spectra, and B) Fluorescence emission spectra (excitation 344 nm)

**b) Quantum yield measurements**. The quantum yields[4] of the ODFs dissolved in spectroscopic grade water were calculated by comparison of the UV-Vis absorbance readings (in triplicate) at $\lambda_{max}$ (OD) and the integrated area of the corrected emission spectrum (I) of the ODF solutions (X) to that of quinine sulfate (R) dissolved in 0.1 N aq. sulfuric acid solution ($\lambda_{ex} = 350$ nm, $\Phi_R = 0.577$), as a reference fluorophore, in accordance to the following equation:

$$\Phi_X = \Phi_R \times (I_X / I_R) \times (OD_R / OD_X) \times (n_X^2 / n_R^2)$$

The index of refraction (n) for each solvent was taken from the Handbook of Chemistry and Physics (91st ed.), **2010-2011**; Taylor and Francis Group, LLC. Pages 3-198, 8-75, and 10-251. Data was acquired on the above Varian Cary UV-Visible spectrophotometer and Jobin Yvon fluorimeter.

**Table S1: Mass spectral data and optical properties of ODF HaloTag ligands**

| # | ODF-HaloTag ligand | Calc. mass | Found mass | $\lambda_{max, abs}$ | Є, L.mol$^{-1}$ cm$^{-1}$ | $\lambda_{max, em}$ | $\Phi_f$[b] | brightness[b] Є • $\Phi_f$ | τ , ns |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 5'-HT-SS-YYYY-3' | 2343.54 | 2342.9 | 330,<br>346 | 96000[a] | 485 | 0.25[a] | 24000 | 6.7[a],<br>41.8[a] |
| 2 | 5'-HT-SS-EY-3' | 1634.80 | 1634.4 | 344,<br>447 | 47000[a] | 492 | 0.65[a] | 30600 | 5.1[a]<br>30.8[a] |
| 3 | 5'-HT-SS-EYK-3' | 2146.26 | 2146.37 | 344,<br>446,<br>498 | 61000[a] | 394,<br>482,<br>644 | 0.082 (a),<br>0.144 (b) | 5000(a),<br>8790(b) | 3.4[a],<br>34.6[a] |
| 4 | 5'-HT-SS-YKY-3' | 2096.20 | 2098.98 | 332,<br>344,<br>478 | 81000[a] | 488,<br>645 | 0.08[a] | 6500 | 1.7[a],<br>14.4[a] |
| 5 | 5'-HT-SS-YZY-3' | 2151.33 | 2151.76 | 329,<br>345,<br>464 | 65000[a] | 393,<br>672 | 0.001[a] | 65 | 2.1[a] |
| 6 | 5'-HT-SS-EYF-3' | 2307.53 | 2305.72 | 332,<br>348,<br>396 | 89000 | 540 | 0.028 | 2490 | 15.5 |
| 7 | 5'-HT-SS-FYF-3' | 2549.87 | 2550.20 | 316,<br>350,<br>386 | 62000 | 527 | 0.105 | 6510 | 3.3 |

[a] Earlier reported for the same dye lacking the haloalkyl group[5]

[b] Data are for free fluorophore; note that quantum yields and brightness increase for some of these dyes upon conjugation.

## 5. Vector Construction, Protein Expression and Purification

The gene encoding GST protein was amplified by PCR from the commercially available vector pGEX 6P3 (GE Healthcare.) The PCR fragment of GST gene was then inserted between the *Nco*I and *Bam*HI sites of vector encoding HaloTag protein pFN18A (Promega, G2751) by replacing the barnase site. The resulting fusion vector, encoding GST-HaloTag protein, was confirmed by agarose DNA gel and DNA sequencing analysis.

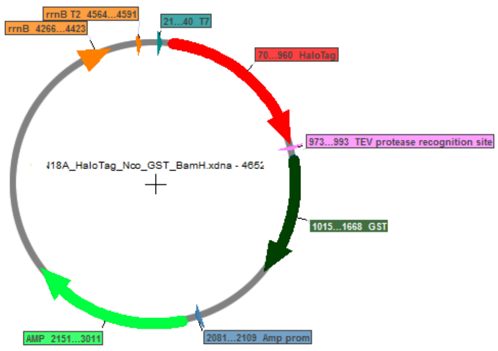


**Figure S3:** GST-Halotag fusion vector

GST-HaloTag fusion plasmid was transformed into KRX strain, an *E. coli* K12 derivative, (Promega, L3002) using a standard heat shock procedure (Sambrook J, Maniatis T, 1989, *Molecular Cloning, A Laboratory Manual*, Second Ed., Cold Spring Harbor Press, Cold Spring Harbor). A fresh colony of bacteria containing GST-HaloTag fusion plasmid was inoculated into 2 mL of LB media containing the ampicillin antibiotic (100 μg/mL) and 0.2–0.4% glucose. After 6–8 h incubation at 37 ºC the starter culture was diluted (1:100) in fresh LB media containing 0.05% glucose and 0.05% rhamnose and cells were grown at 25 ºC for another 17–20 h. The cells were then collected by centrifugation at 4,000 × *g* for 20 min at 4 ºC. Harvested cells were resuspended in 5 mL PBS. Cell lysis was performed through mechanical means with the help of a bead beater. The protease inhibitor (cOmplete, Mini, EDTA-free Protease Inhibitor Cocktail Tablets, Roche) was added as per manufacturer's protocol (1 tablet in 10 mL PBS) prior to cell lysis. The GST-HaloTag fusion protein was purified with the help of GST resin (GE Healthcare). The cleared cell lysate was added to the prewashed GST-resin, packed in a column and incubated for 20 min at rt. Following two 10 mL batch washes with PBS, the GST-HaloTag fusion protein was eluted with elution buffer containing 20 mM glutathione, as per the protocol suggested by Promega.

HaloTag protein was obtained by cleaving GST protein from GST-HaloTag fusion protein. The 10 µg fusion protein was incubated with 4 units of TEV protease (Promega, V6051) in a 1x ProTEV buffer for 30 min at 30 $^{o}$C. The cleaved GST protein was captured by GST-resin and the remaining TEV protease was removed by incubating the reaction mixture with His-resin, as per the protocol suggested by Promega. The GST-HaloTag fusion protein and HaloTag protein were further purified and concentrated using Amicon ultra centrifuge filter units (Millipore). Purified proteins were characterized by SDS-PAGE and ESI mass spectrometry.

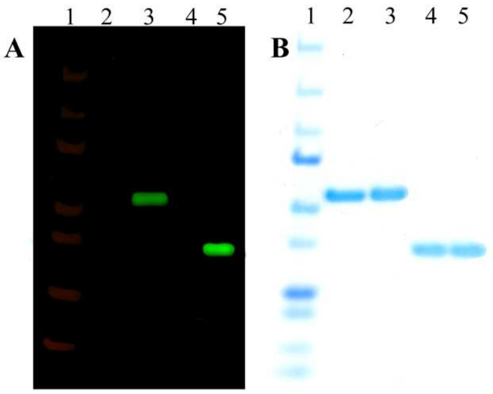


**Figure S4** SDS PAGE analysis of purified proteins: A) fluorescence scan (Ex 457 nm, Emission filters: 520 BP 40 Cy2 and 670 BP 30 Cy5); B) Coomassie blue staining. Lane 1: marker, lane 2: GST-HaloTag protein, lane 3: GST-HaloTag protein labeled with htS$_2$EY, lane 4: HaloTag protein, lane 4: HaloTag protein labeled with htS$_2$EY

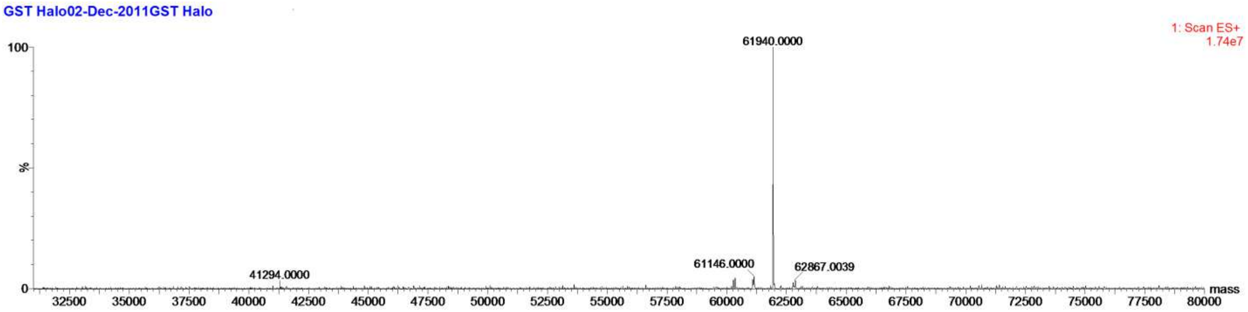


**Figure S5.** ESI mass spectrum of HaloTag-GST fusion protein.

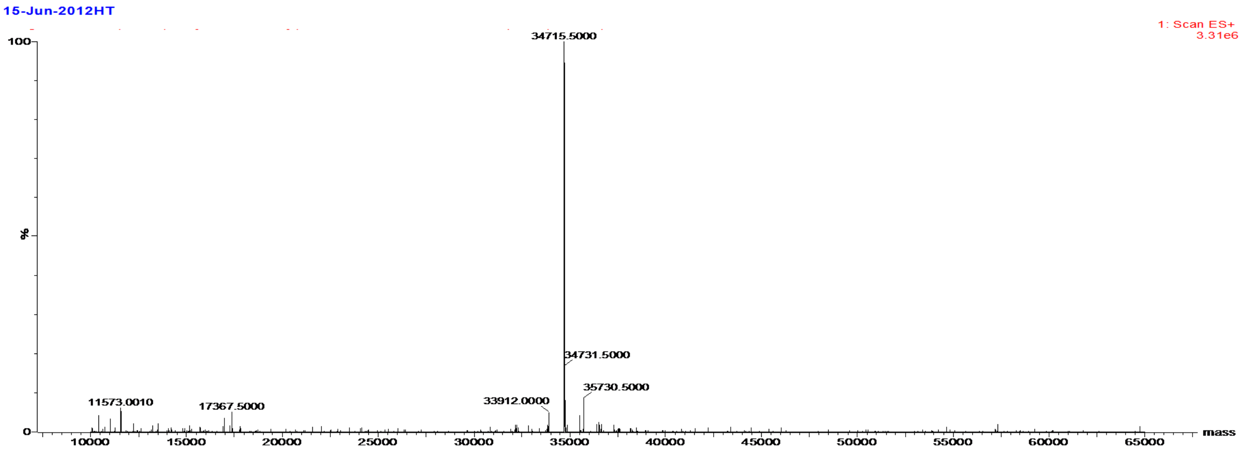


**Figure S6.** ESI mass spectrum of HaloTag protein.

## 6. In Vitro Protein Labeling Experiments

**a) Concentration-dependent experiment.** 1.0 µg GST-HaloTag fusion protein in 8.0 µL PBS buffer (protein concentration ~2.0 µM) was incubated with increasing amounts of ODF-HaloTag ligand (0.5, 1, 2, 4, 8, 16, 32 µM) at 37 °C. After 30 min, the reaction mixture was quenched with 2x SDS gel loading dye and heated to 95 $^{o}$C for 5 min. The labeled protein was then resolved by SDS-PAGE (Mini-Protean 4-15% Tris-HCl gels), and in-gel fluorescence analysis was carried out by Typhoon gel scanner (Ex 457).

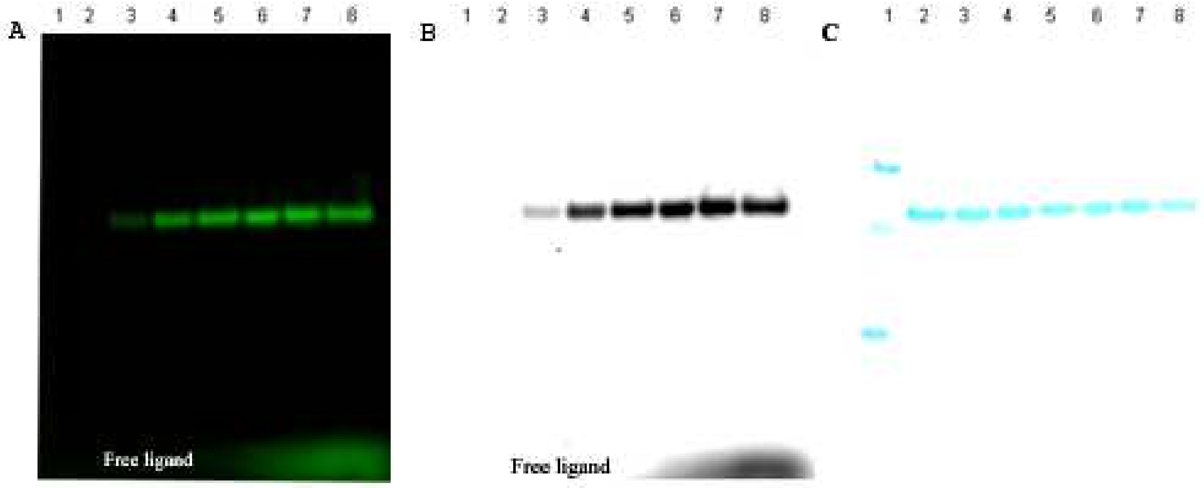


**Figure S7**. ODF-HaloTag ligand concentration-dependent experiment: GST-HaloTag protein 1.0 µg (~2 µM) in 8 µL PBS, 37 $^{\circ}$C, 30 min. A) fluorescence image (Ex 457 nm, Emission filters: 520 BP 40 Cy2 and 670 BP 30 Cy5), B) fluorescence image (Ex 457 nm, Emission filter 390 BP 100 Phosphor), and C) Coomassie blue staining: lane 1: marker, lane 2: 0.5 µM, lane 3: 1 µM, lane 4: 2 µM, lane 5: 4 µM, lane 6: 8 µM, lane 7: 16 µM, lane 8: 32 µM htS$_2$EY.

**b) Time-dependent Experiment**

7.0 µg GST-HaloTag fusion protein in 56.0 µL PBS buffer (protein concentration ~2µM) was incubated with 4.0 µM ODF-HaloTag ligand at 37 °C. At selected time points (5, 15, 30, 45, 60, 90 min), aliquots (8 µL) were removed from the reaction mixture and quenched with 2x SDS gel loading dye and heated to 95 $^{o}$C for 5 min. The samples were then resolved by SDS-PAGE and imaged as in **6a** above.

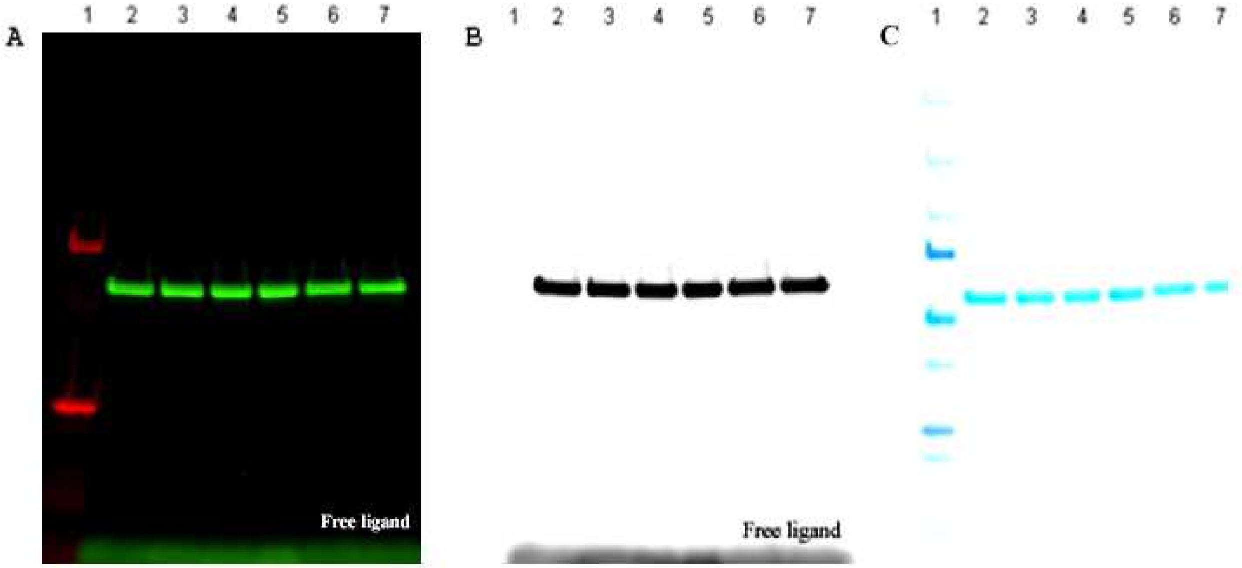


**Figure S8.** Time-dependent experiments: GST-HaloTag protein 1.0 µg (~2 µM) in 8.0 µL PBS, 4.0 µM $htS_2EY$, 37 °C. A) fluorescence image (Ex 457 nm, Emission filters: 520 BP 40 Cy2 and 670 BP 30 Cy5), B) fluorescence image (Ex 457 nm, Emission filter: 390 BP 100 Phosphor), and C) Coomassie blue staining: lane 1: marker, lane 2: 5 min, lane 3: 15 min, lane 4: 30 min, lane 5: 45 min, lane 6: 60 min, lane 7: 90 min.

### c) Multispectral protein labeling

1.0 µg of GST-HaloTag fusion protein and 0.5 µg of Halotag protein in 8.0 µL PBS (protein concentration ~2.0 µM) were separately incubated with 4.0 µM ODF-HaloTag ligands ($htS_2Y_4$, $htS_2EY$, $htS_2YKY$, $htS_2EYKa$, $htS_2EYKb$, $htS_2EYF$, $htS_2FYF$, $htS_2YZY$) at 37°C. After 30 min, reaction mixture was quenched with SDS gel loading dye and heated to 95 °C for 5 min. The samples were resolved by SDS-PAGE and imaged as in **6a** above.

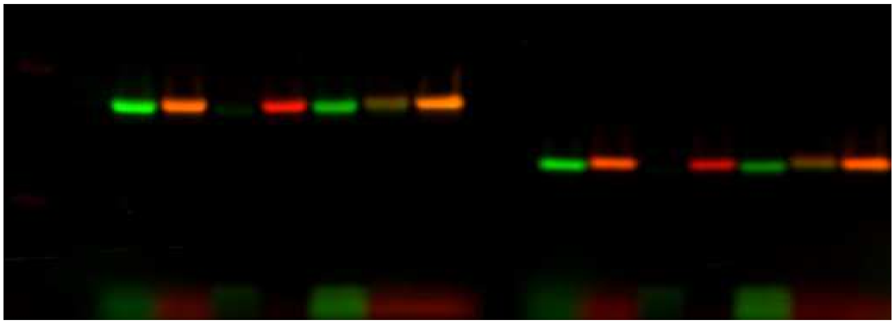


**Figure S9**: Multispectral labeling of GST-HaloTag fusion protein (lane 2-9) and Halotag protein (lane 10-17) with ODFs (Ex 457 nm, Emission filters: 520 BP 40 Cy2 and 670 BP 30 Cy5): lane 1: marker, lane 2 & 10: $htS_2Y_4$, lane 3 & 11: $htS_2EY$, lane 4 & 12: $htS_2YKY$, lane 5 & 13: $htS_2FYF$, lane 6 & 14: $htS_2YZY$, lane 7 & 15: $htS_2EYF$, lane 8 & 16: $htS_2EYKa$, lane 9 & 17: $htS_2EYKb$. The image was taken on a Typhoon Gel scanner (Ex 457 nm).

## 7. Protein-ODF conjugates:

In a 96-well plate, 1.0 µM of each ODF-HaloTag ligand (ht$S_2Y_4$, ht$S_2$EY, ht$S_2$YKYb, ht$S_2$EYKa, ht$S_2$EYKb, ht$S_2$EYF, ht$S_2$FYF, ht$S_2$YZY) was incubated in presence of protein (HaloTag protein 2.5 µM) and in absence of protein, in 50 µL PBS at 37 $^{o}$C for 30 min. The fluorescence emission spectra of ODFs and protein-ODF conjugates were recorded using FLEXstation II-384-fluorescent plate reader (Ex 344 nm, Em >365-750nm).

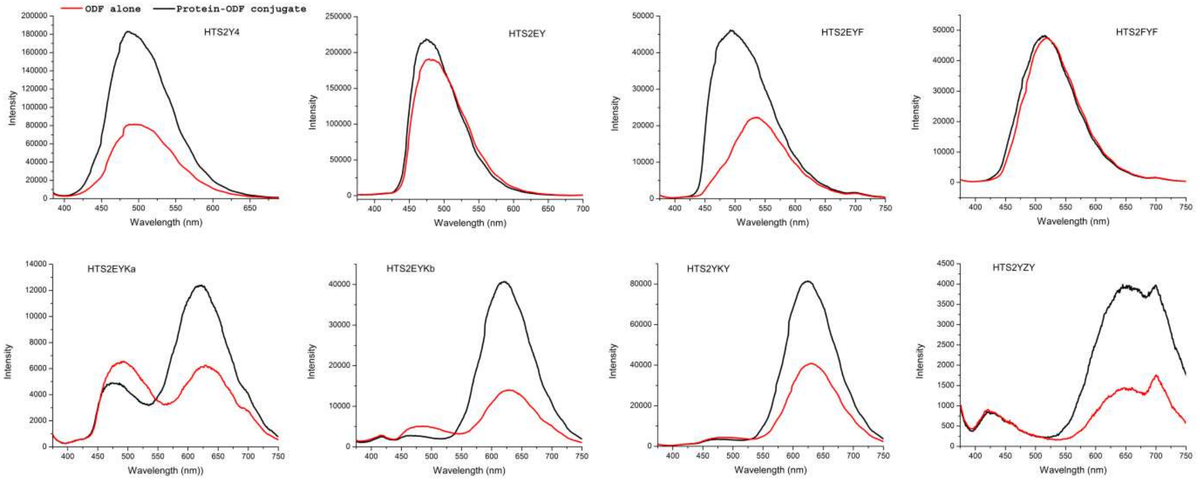


**Figure S10:** Fluorescence emission spectra of ODFs and protein-ODF conjugates (red line ODF alone and black line protein-ODF conjugates). In order to make ODF-protein conjugates 1.0 µM ODF was treated with 2.5 µM ODF-Halotag ligand in 50 µL PBS at 37 $^{o}$C for 30 min.

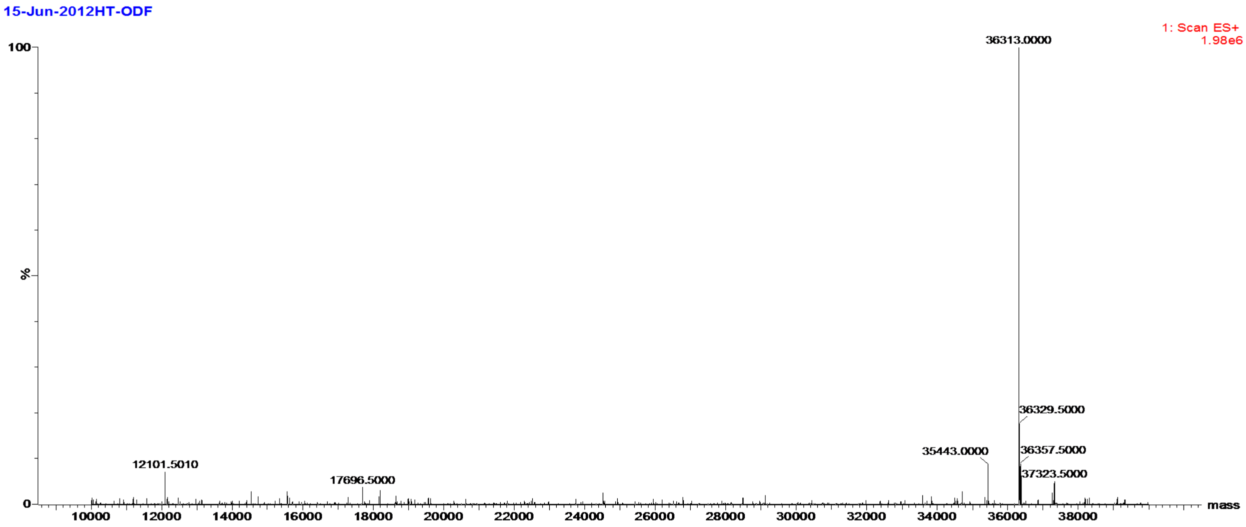


**Figure S11**: ESI mass spectrum of a representative protein-ODF conjugate (HaloTag protein-ht$S_2$EY conjugate). Calculated mass: 36312.9 (found 36313.0)

## 8. Cellular Protein Labeling Experiments

### a) Halotag Fusion Vector Construction

The vector encoding α-tubulin-HaloTag fusion protein was constructed by inserting α-tubulin gene between the NcoI and BamHI sites of commercially available HaloTag plasmid pFN21A (Promega, G2821), a mammalian expression vector. The plasmid encoding cell surface-HaloTag fusion protein was obtained from Dr. S. Gambhir (Stanford University), which was constructed by inserting HaloTag protein gene into the pDisplay vector (Invitrogen). The resulting plasmids were then tranformed into Top-10 bacterial cells using a standard heat shock method. The transfected cells were propagated in LB media at 37 $^{o}$C. Thereafter plasmids were isolated from bacterial cells using a plasmid isolation kit (Fermentas) and charectorized by agarose DNA gel and DNA sequencing analysis

### b) Human Cell Culture, Transfection and Protein Expression

HeLa cells were cultured in DMEM w/ glutamine (Gibco #11995) with 10% v/v fetal bovine serum and 1% v/v Pen/Strep. All cells were maintained under 5% $CO_2$ at 37 °C. For live cell imaging, cells were plated in Lab-Tek 8-well chambered coverglass slides (Nunc 155409) 24 h before transfection. Thereafter HeLa cells were transfected with HaloTag fusion plasmids (2.0 μg DNA per well) using Lipofectamine 2000 transfection reagent (Invitrogen) following the manufacturer's protocol. After transfection, cells were incubated in growth media for 48 h. The protein labeling was performed by incubating HeLa cells in 200 μL growth media DMEM (without FBS and Pen/Strep) containing 5.0 μM ODF HaloTag ligand for 15 min at 37 °C. After labeling, staining media was removed and cells were washed two times with PBS, and a final washing was performed by incubating cells in phenol-free media (DMEM) at 37 $^{0}$C for 30 min. Before imaging, the media was replaced with fresh phenol-free media.

### c) Cellular Imaging

The cellular imaging was performed by using Leica sp5 confocal microscope with a PL APO 63x oil objective. During imaging HeLa cells were in phenol-free DMEM. The ODFs were excited at 405 nm with an argon laser source.

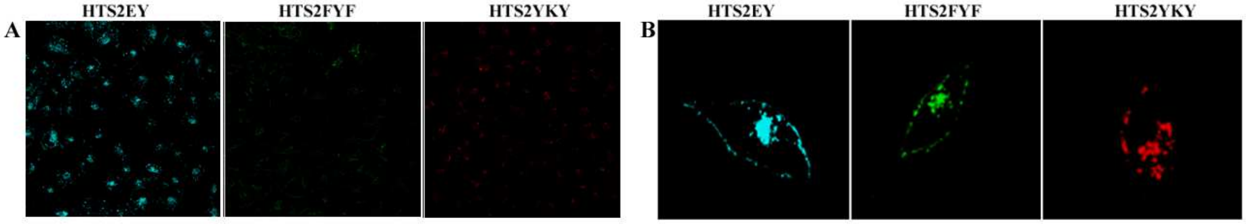


**Figure S12**. Control HeLa cells labeling with ODF-HaloTag ligands. A) Control for cell surface protein labeling; B) control for cytoplasmic protein labeling.

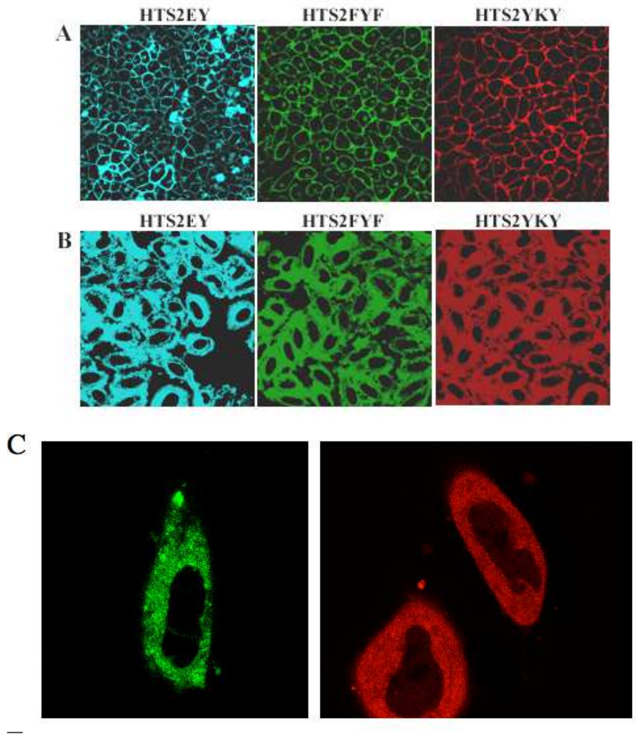


**Figure S13**. A) Cell surface protein (PDGFR-TM) labeling, and B) cytoplasm protein (α-tubulin) labeling (at same scale) with ODF-HaloTag ligands (sequences shown). (C) closeup of α-tubulin-labeled cells; (left) HTS2FYF dye; (right) commercial TMR-halotag dye for comparison.

## Dual Protein Labeling

HeLa cells were cultured in DMEM (Gibco #11995) with 10% v/v fetal bovine serum and 1% v/v Pen/Strep. All cells were maintained under 5% $CO_2$ at 37 °C. The cells were plated in Lab-Tek 8-well chambered coverglass slides (Nunc 155409) 24 h before transfection. The cells were first transfected with the vector encoding cytoplasmic protein (α-tubulin-HaloTag fusion vector, 2.0 μg plasmid per well) using Lipofectamine 2000 transfection reagent. The cells were incubated for 48 h after transfection. Thereafter, expressed cytoplasmic protein was labeled with cyan color ODF dye (htS$_2$EY). This was achieved by incubating cells in 200 μL culture media (DMEM without FBS and Pen/Strep) containing 5.0 μM ODF dye for 60 min. After washing, cells were again transfected with a vector encoding cell surface fusion protein (pDisplay-HaloTag fusion vector, 2.0 μg plasmid per well). After 48 h of incubation, cell surface protein was labeled with red color ODF dye (htS$_2$YKY) for 15 min. Thereafter, cells were washed two times with PBS and final washing was performed by incubating cells in phenol-free media (DMEM) at 37 $^{o}$C for 30 min. Before imaging, the medium was replaced with fresh phenol-free medium. The dual protein labeling was imaged by Leica sp5 confocal microscope with 405 nm excitation and using cyan and red emission filters.

## Additional Figures

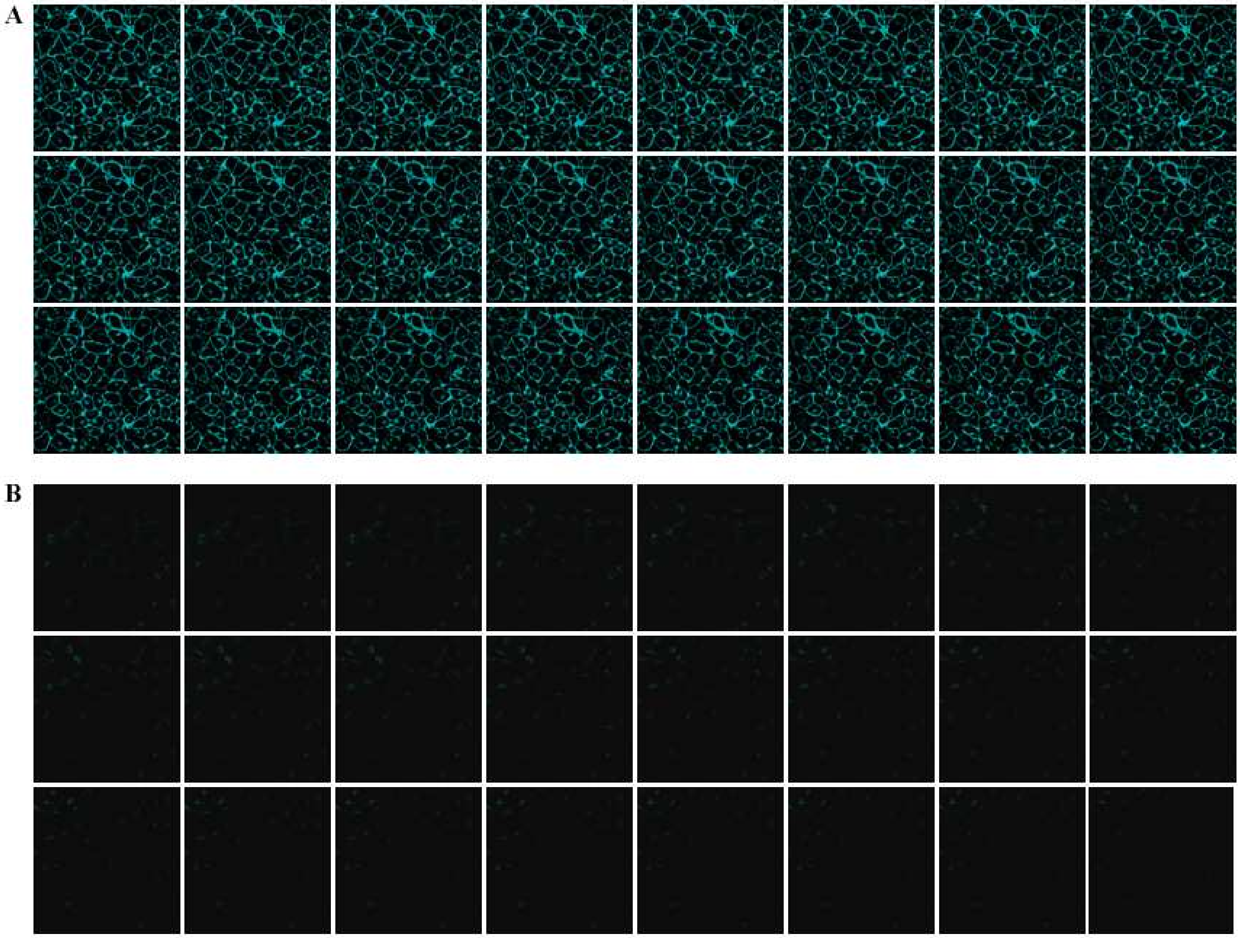


**Figure S15.** Z stacks from laser confocal microscopy. A) HeLa cells expressing cell surface protein (pDisplay) labeled with $htS_2EY$, and B) control HeLa cells labeled with $htS_2EY$.

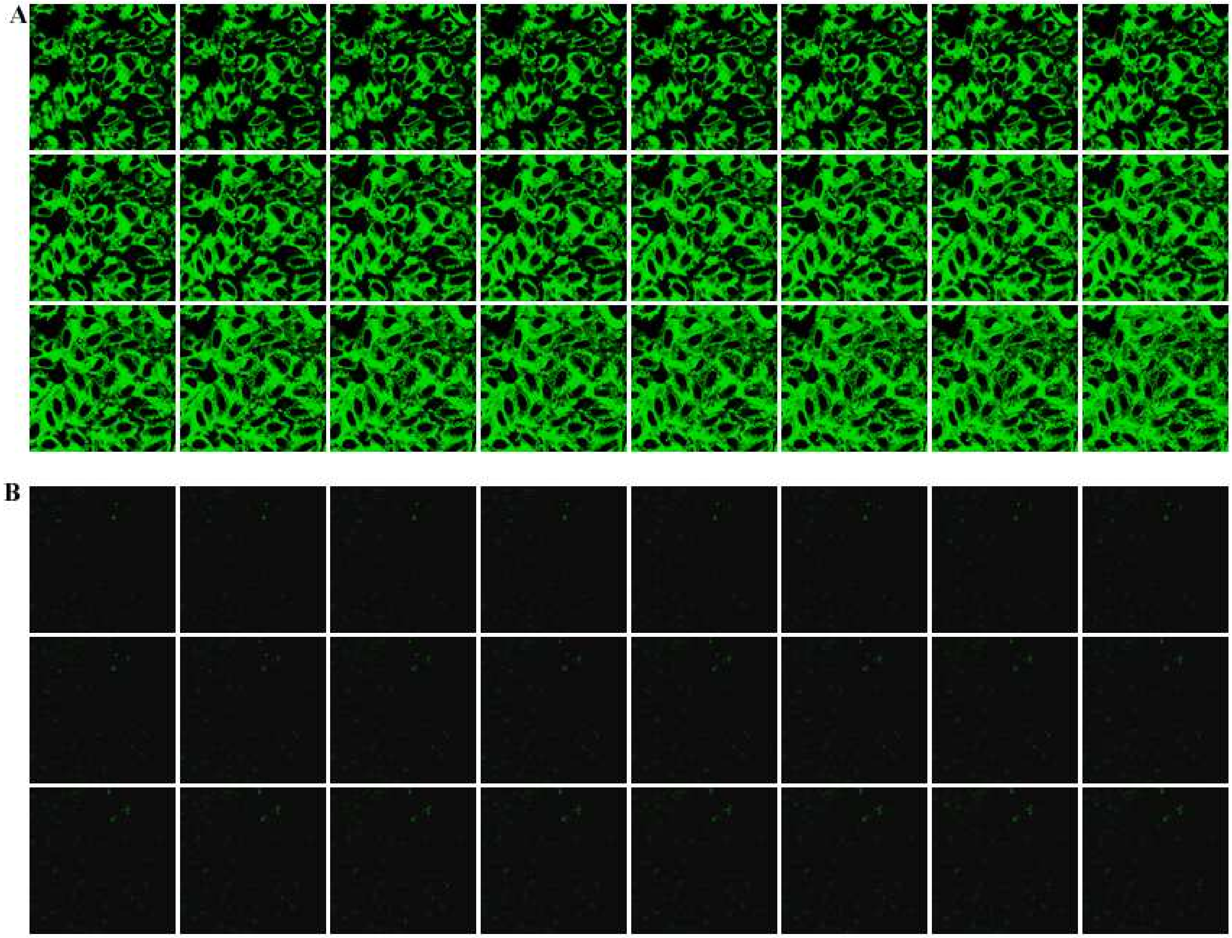


**Figure S16.** Z stacks from laser confocal microscopy. A) HeLa cells expressing cytoplasmic α-tubulin protein labeled with htS$_2$FYF, and B) control HeLa cells labeled with htS$_2$FYF.

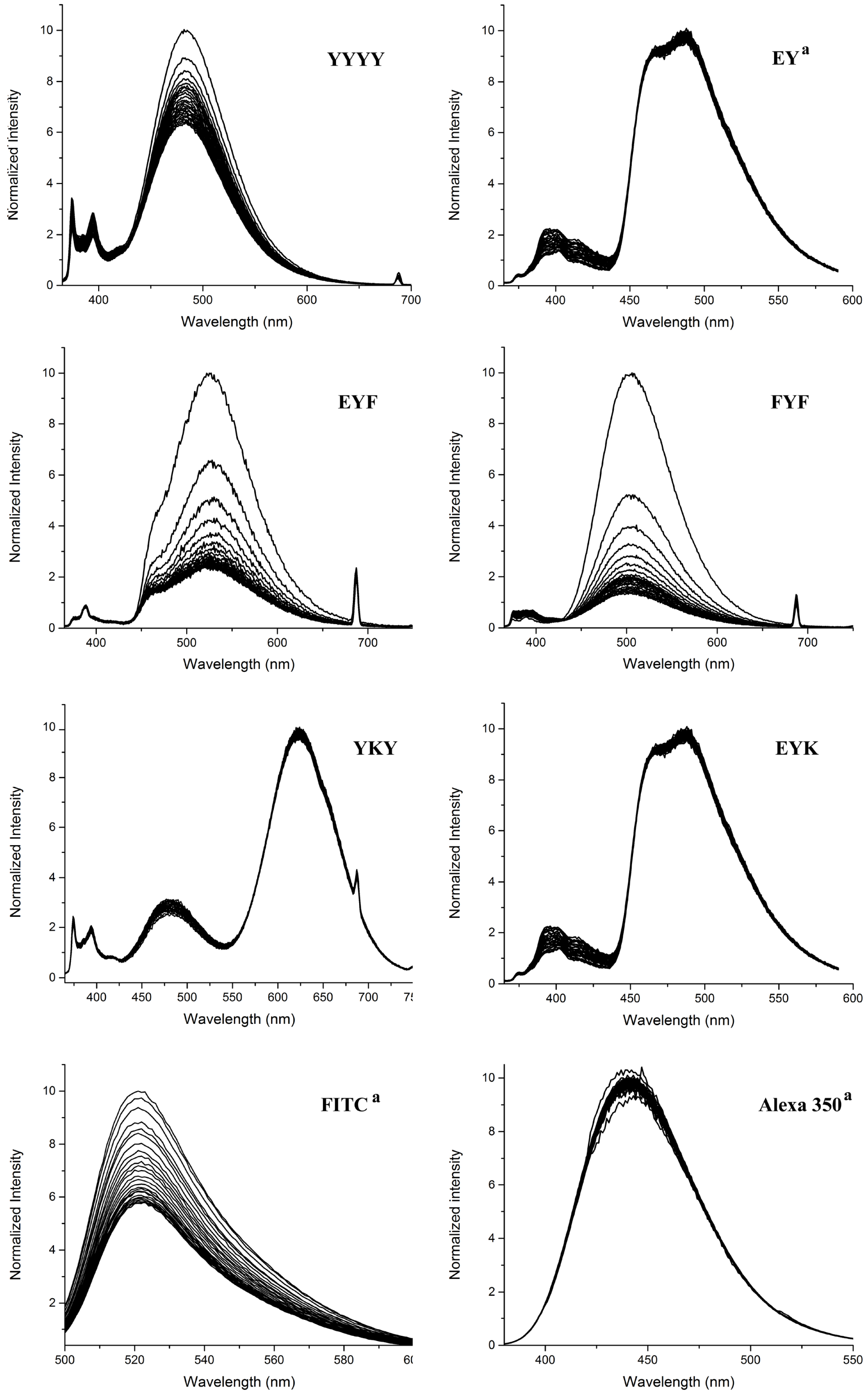


**Figure S17. Testing stability of ODF dyes against photobleaching in PBS (25 °C).** The fluorescence emission spectra were recorded every 5 min for 2.5 h, exposing the sample (1 μM dye) to excitation light for ca. 60 sec for each recording. ODF sequences are as indicated. Standard organic dyes FITC and Alexa 350 were included for comparison. FITC was excited at 495 nm, all others excited at 344 nm. No antioxidants or oxygen scavengers were included. [a]Data taken from ref. 5c (no haloalkyl group).

---

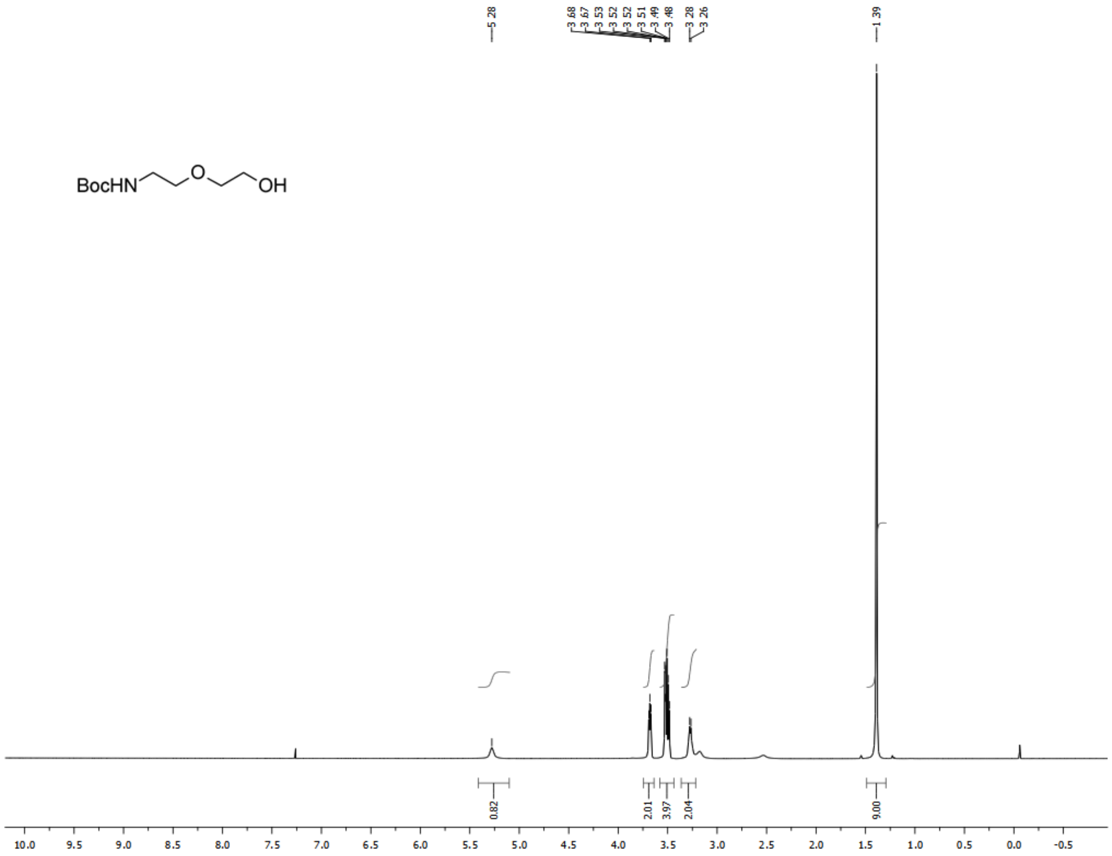
[1] Y.S. Gao, C.C. Hubbert, T.P. Yao. The microtubule-associated histone deacetylase 6 (HDAC6) regulates epidermal growth factor receptor (EGFR) endocytic trafficking and degradation, *J Biol Chem*. **2010**; 285, 11219-26.

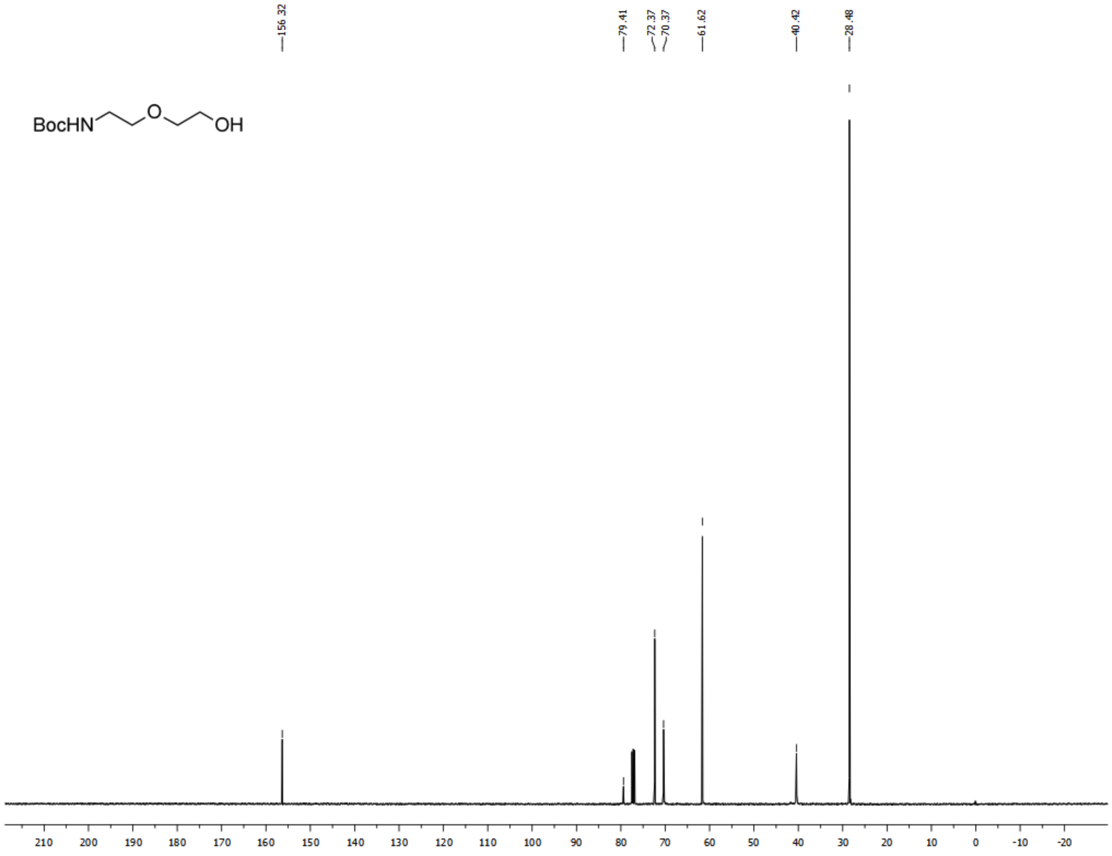
[2] a) K. Y. Gil, et al, Novel organic electroluminescent compounds and organic electroluminescent device using the same. US2010072888. b) K.D. Belfield, K.J. Schafer, W. Mourad, B.A. Reinhardt *J.Org.Chem*.**2000**, *65*, 4475–4481. c) R. Kannan**,** G. He**,** L. Yuan, F. Xu**,** P. N. Prasad**,** A. G. Dombroskie**,** B. A. Reinhardt *Chem. Mater.*, **2001**, *13*, 1896–1904.

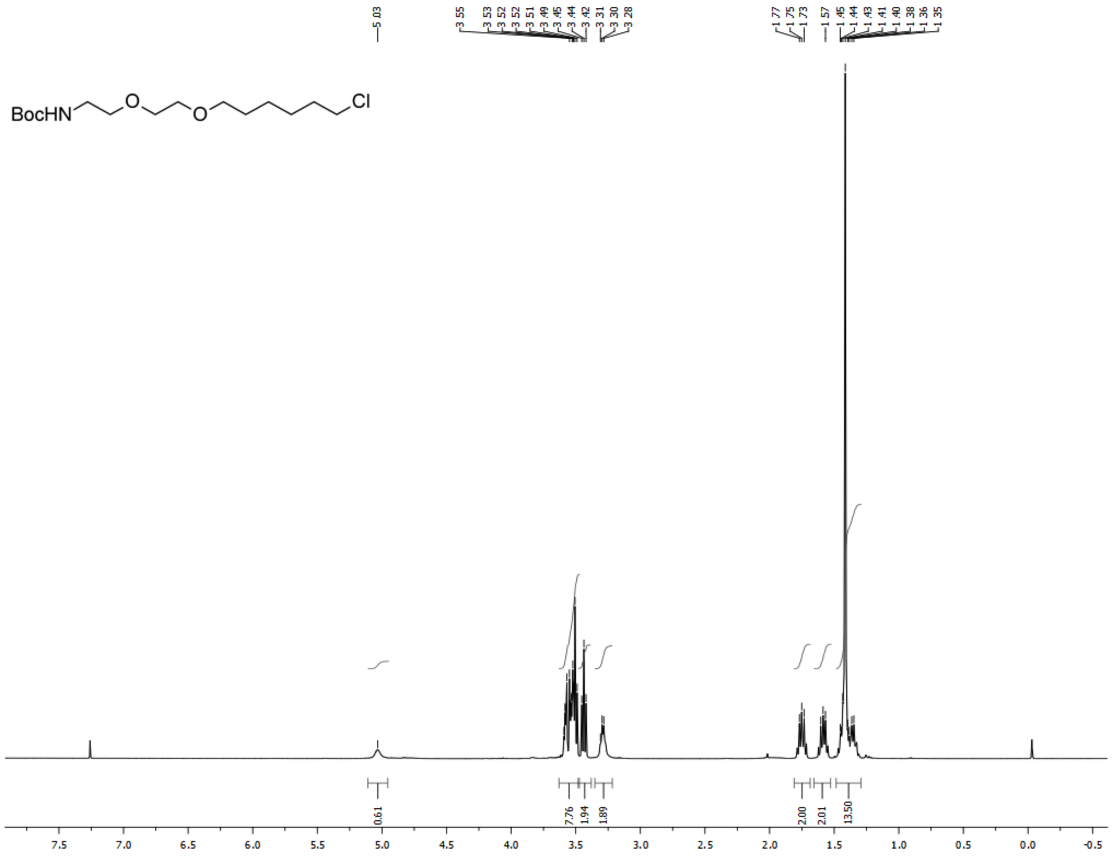
[3] M. A. Cameron, S. B. Cush, R. P. Hammer, Facile Preparation of Protected Furanoid Glycals from Thymidine. *J. Org. Chem.* **1997**, *62*, 9065-9069

[4] J.R Lakowicz,. (Editor). Principles of Fluorescence Spectroscopy (3rd ed.), Boston, MA, USA: Springer US, 2006, pgs. 54-55

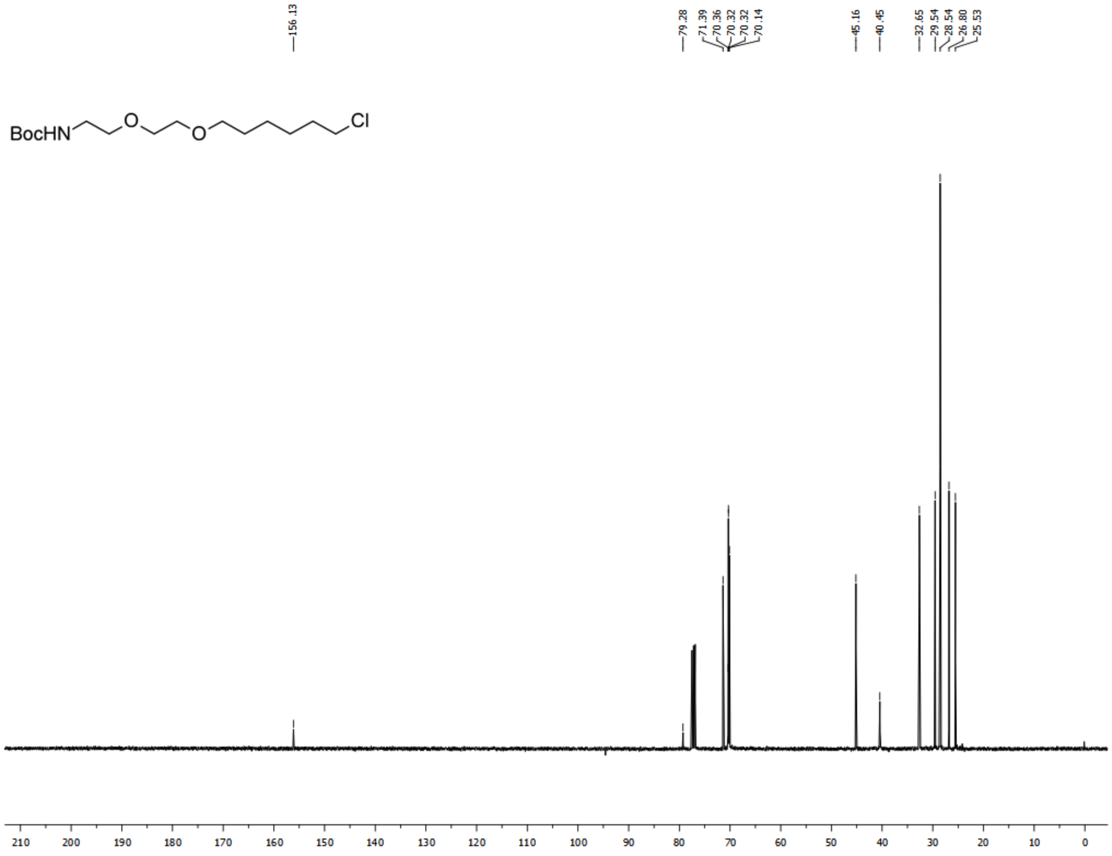
[5] a) Y. N. Teo, J. N. Wilson, E. T. Kool, *J. Am. Chem. Soc.*, **2009**, *131*, 3923-3933 b) J. Guo, S. Wang, N. Dai, Y. N. Teo, E. T. Kool, *PNAS*, **2011**, *108,* 3493-3498. c) S. Wang, J. Guo, T. Ono, E. T. Kool, *Angew. Chem. Int. Ed.*, **2012,** *51*, 7176-7180.


BocHN O OH
5.28
3.68
3.67
3.53
3.52
3.52
3.51
3.49
3.48
3.28
3.26
1.39
0.82
2.01
3.97
2.04
9.00
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5

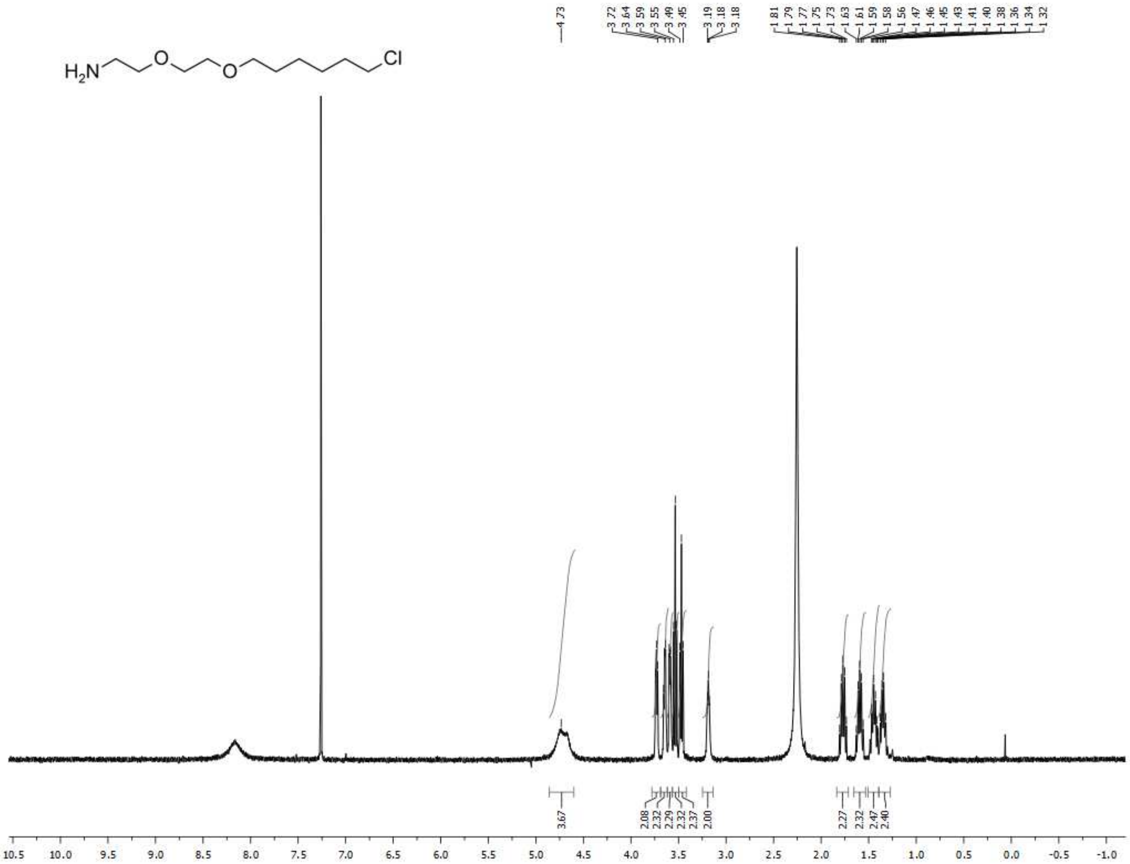

BocHN O OH
156.32
79.41
72.37
70.37
61.62
40.42
28.48
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10 -20



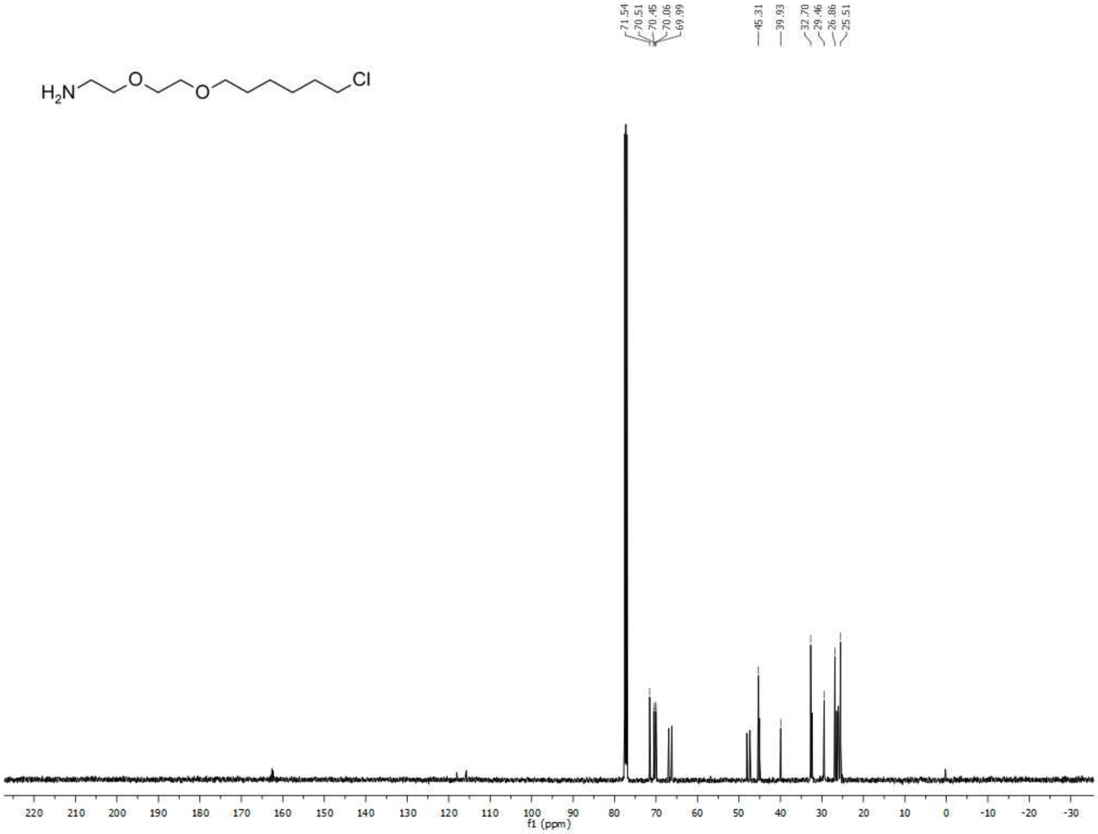
BocHN
O
O
Cl
5.03
3.55
3.53
3.52
3.52
3.51
3.49
3.45
3.44
3.42
3.31
3.30
3.28
1.77
1.75
1.73
1.57
1.45
1.44
1.43
1.41
1.40
1.38
1.36
1.35
0.61
7.76
1.94
1.89
2.00
2.01
13.50
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5



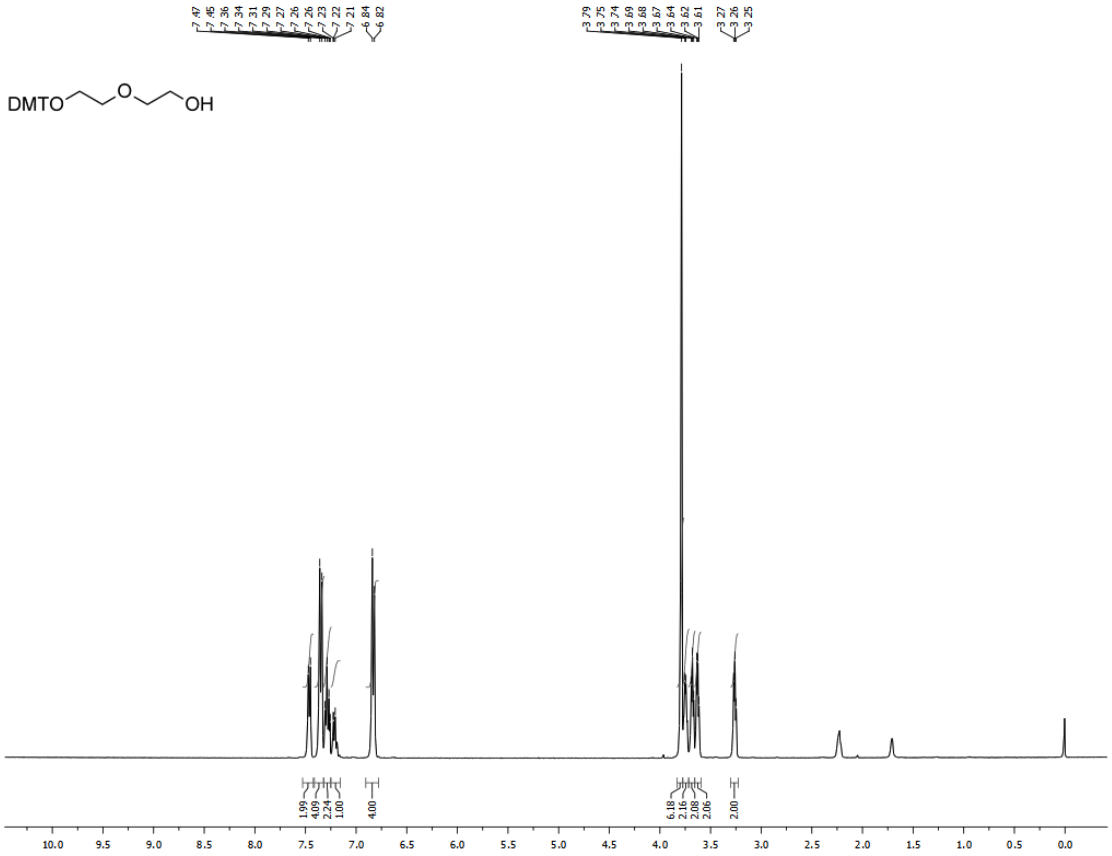
BocHN
O
O
Cl
156.13
79.28
71.39
70.36
70.32
70.32
70.14
45.16
40.45
32.65
29.54
28.54
26.80
25.53
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

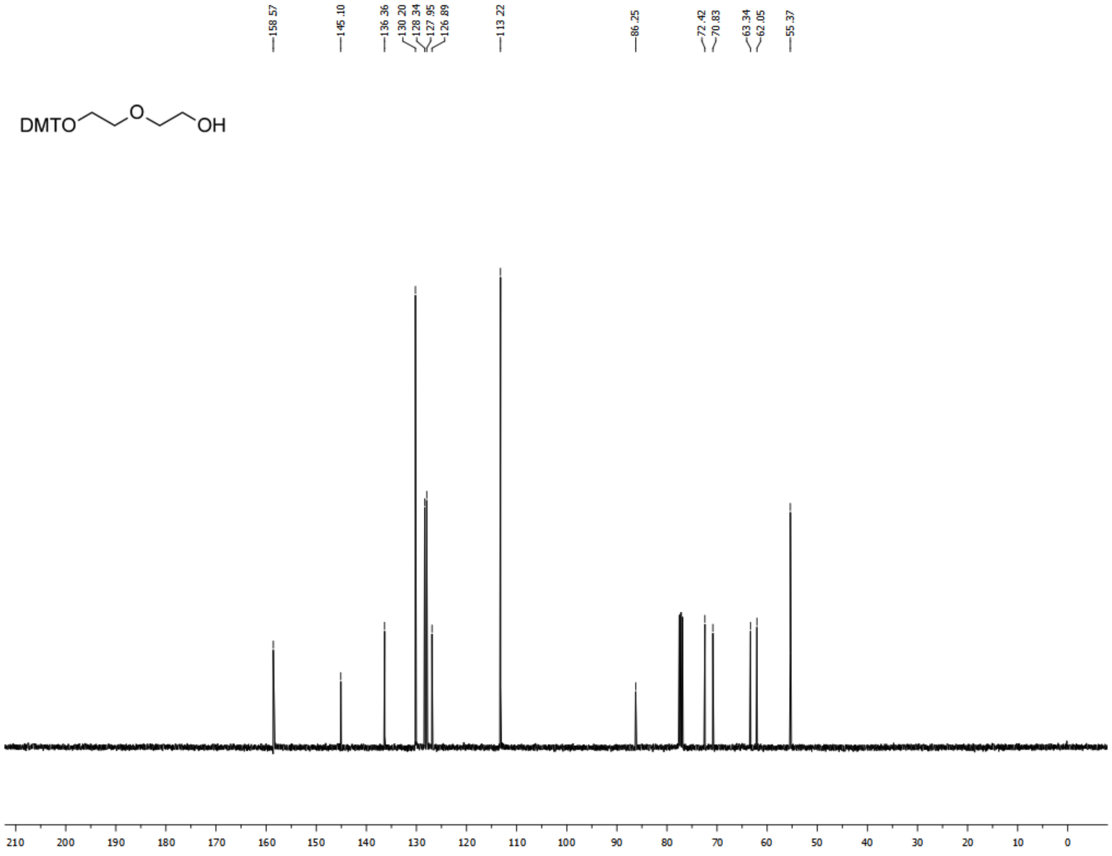


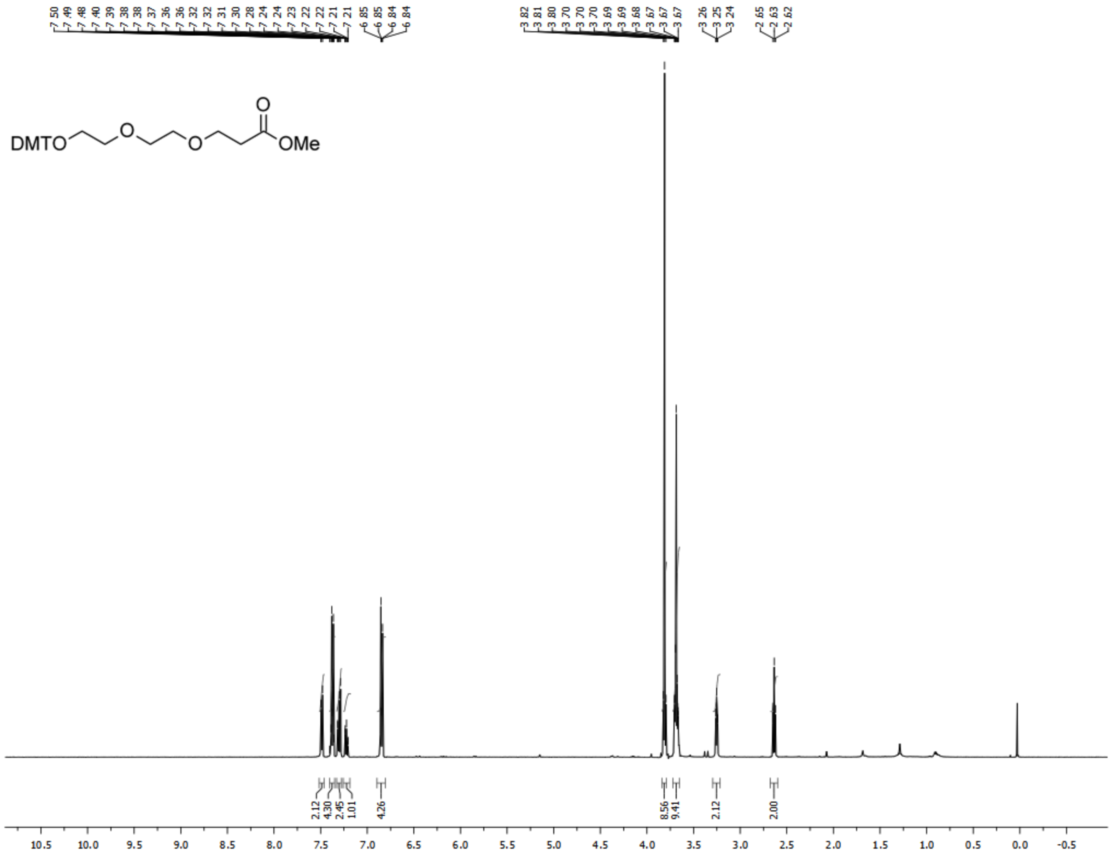
H2N O O Cl
4.73
3.72
3.64
3.59
3.55
3.49
3.45
3.19
3.18
3.18
1.81
1.79
1.77
1.75
1.73
1.63
1.61
1.59
1.58
1.56
1.47
1.46
1.45
1.43
1.41
1.40
1.38
1.36
1.34
1.32
3.67
2.08
2.32
2.29
2.32
2.37
2.00
2.27
2.32
2.47
2.40
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0

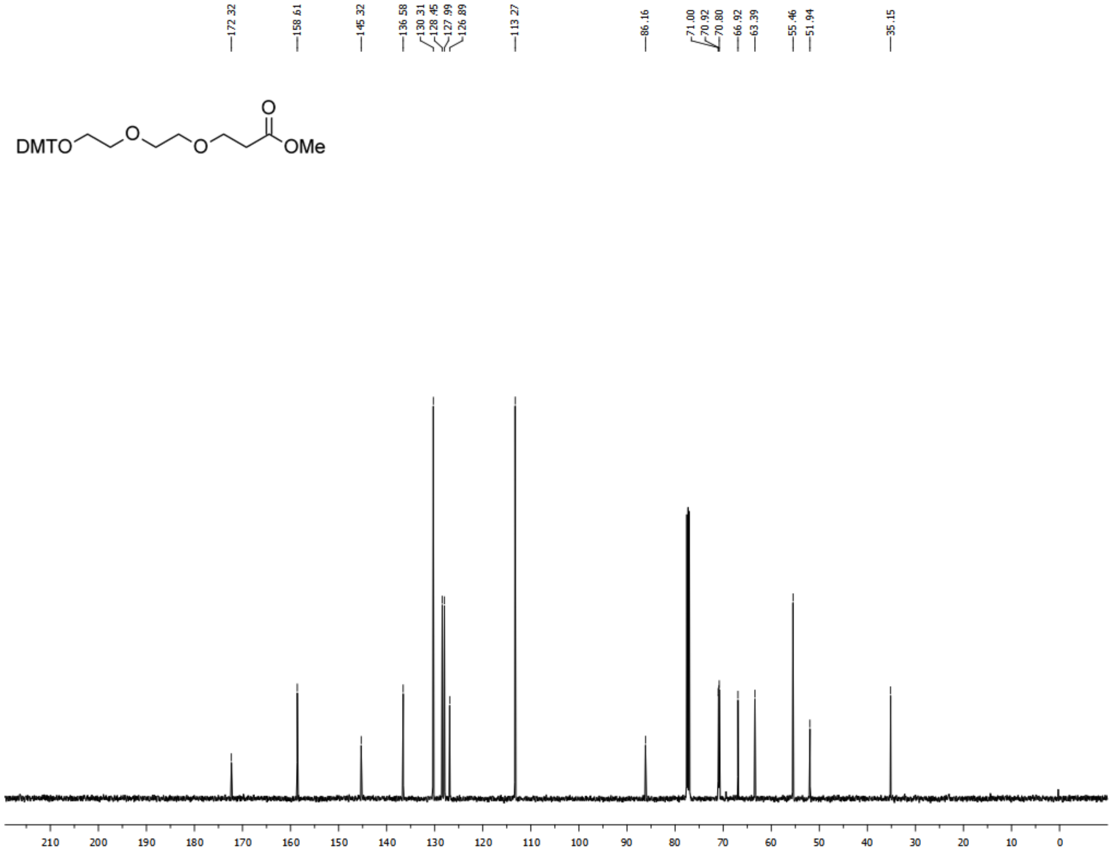


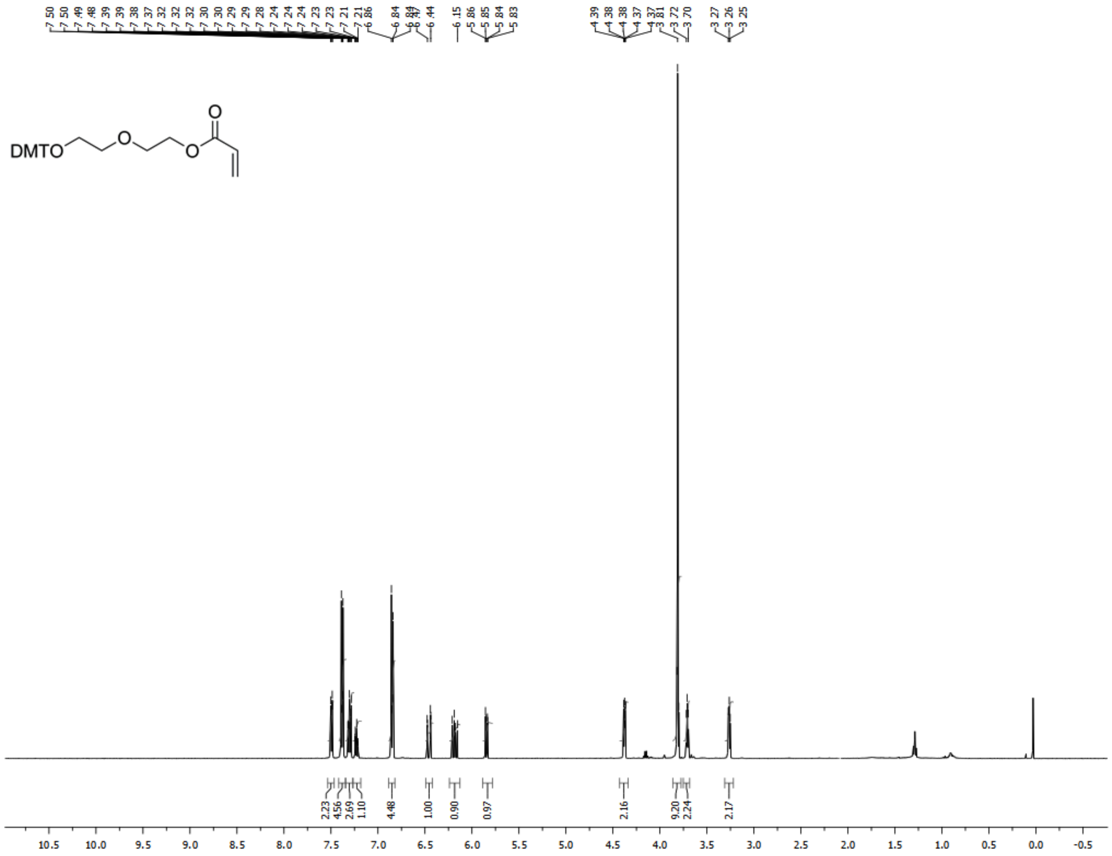
71.54
70.51
70.45
70.06
69.99
45.31
39.93
32.70
29.46
26.86
25.51
H2N O O Cl
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
f1 (ppm)

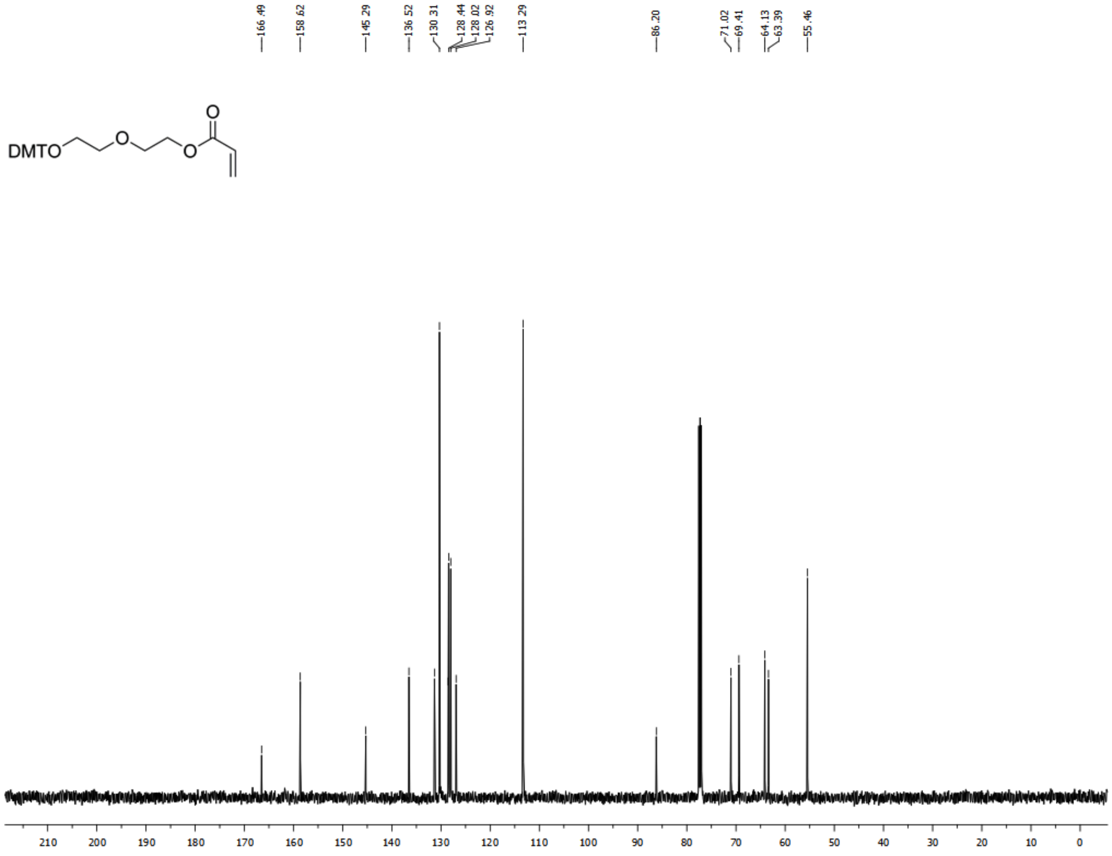

7.47
7.45
7.36
7.34
7.31
7.29
7.27
7.26
7.26
7.23
7.22
7.21
6.84
6.82
3.79
3.75
3.74
3.69
3.68
3.67
3.64
3.62
3.61
3.27
3.26
3.25
DMTO
O
OH
1.99
4.09
2.24
1.00
4.00
6.18
2.16
2.08
2.06
2.00
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

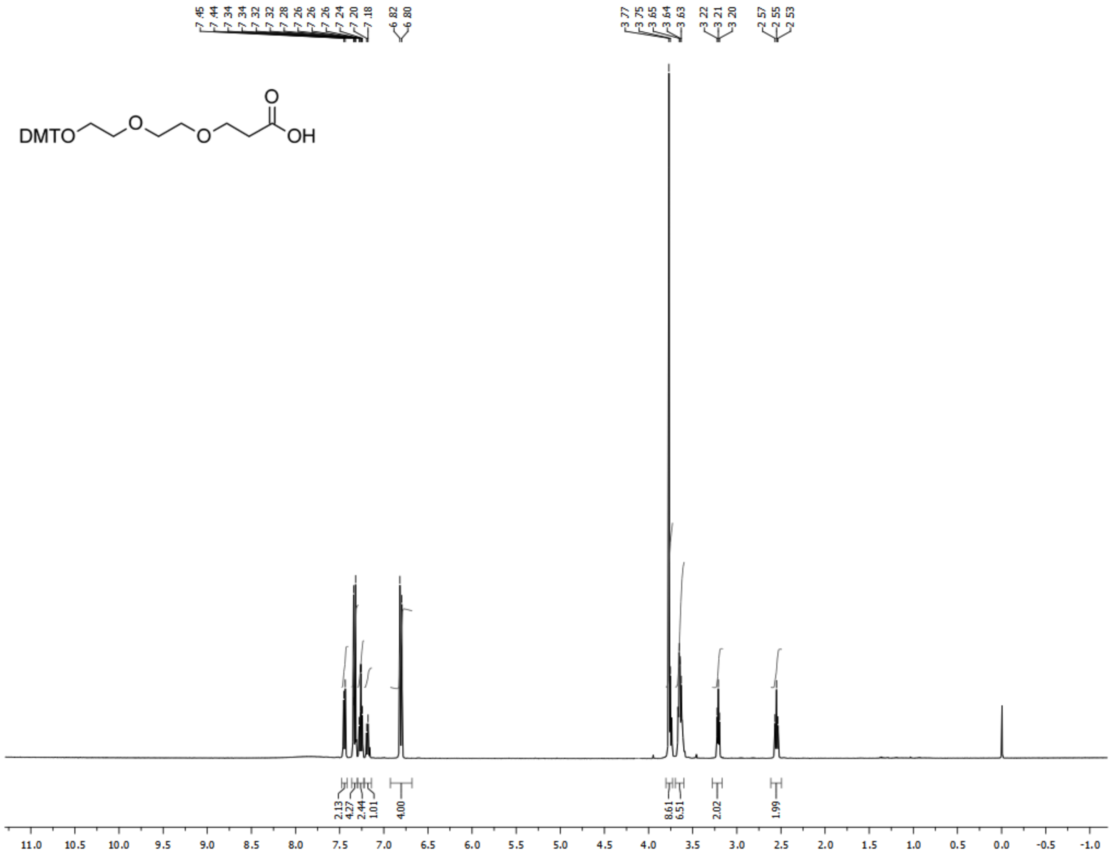

158.57
145.10
136.36
130.20
128.34
127.95
126.89
113.22
86.25
72.42
70.83
63.34
62.05
55.37
DMTO
O
OH
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

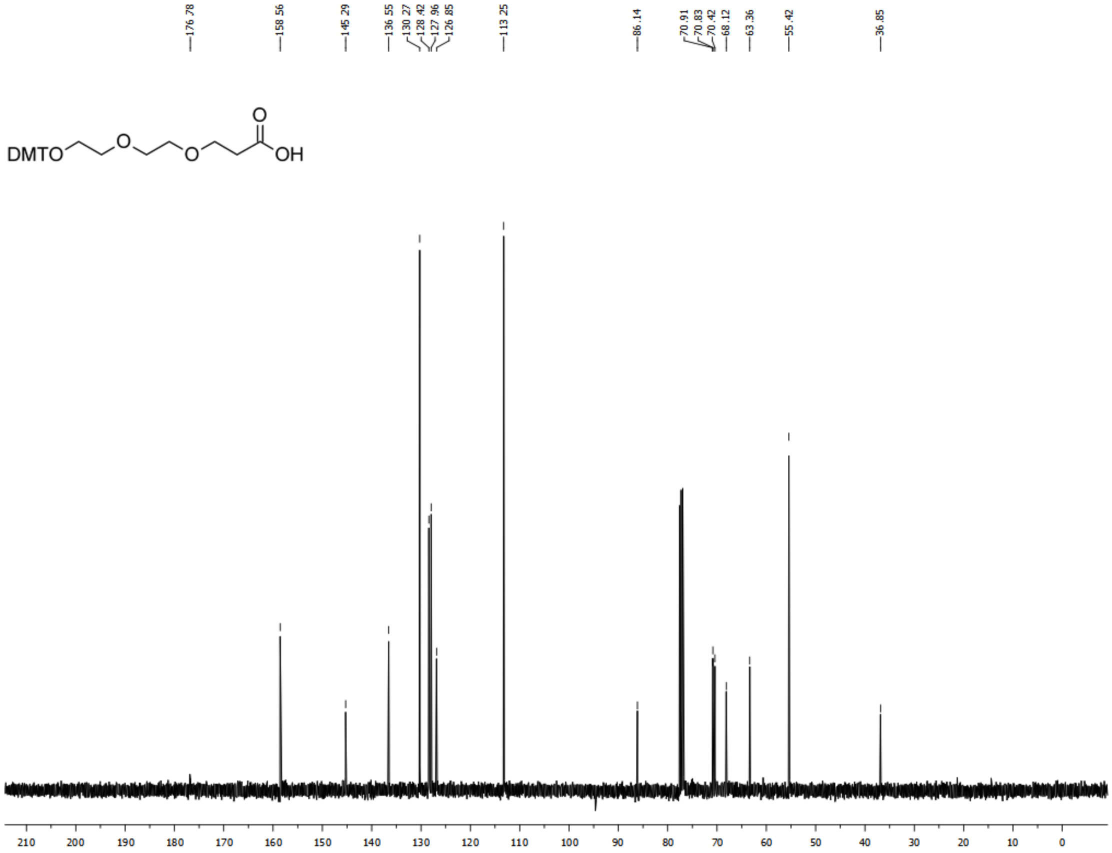

7.50
7.49
7.48
7.40
7.39
7.38
7.38
7.37
7.36
7.36
7.32
7.32
7.31
7.30
7.28
7.24
7.24
7.23
7.22
7.22
7.21
7.21
6.85
6.85
6.84
6.84
3.82
3.81
3.80
3.70
3.70
3.70
3.69
3.69
3.68
3.67
3.67
3.67
3.26
3.25
3.24
2.65
2.63
2.62
DMTO
O
O
O
OMe
2.12
4.30
2.45
1.01
4.26
8.56
9.41
2.12
2.00
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5

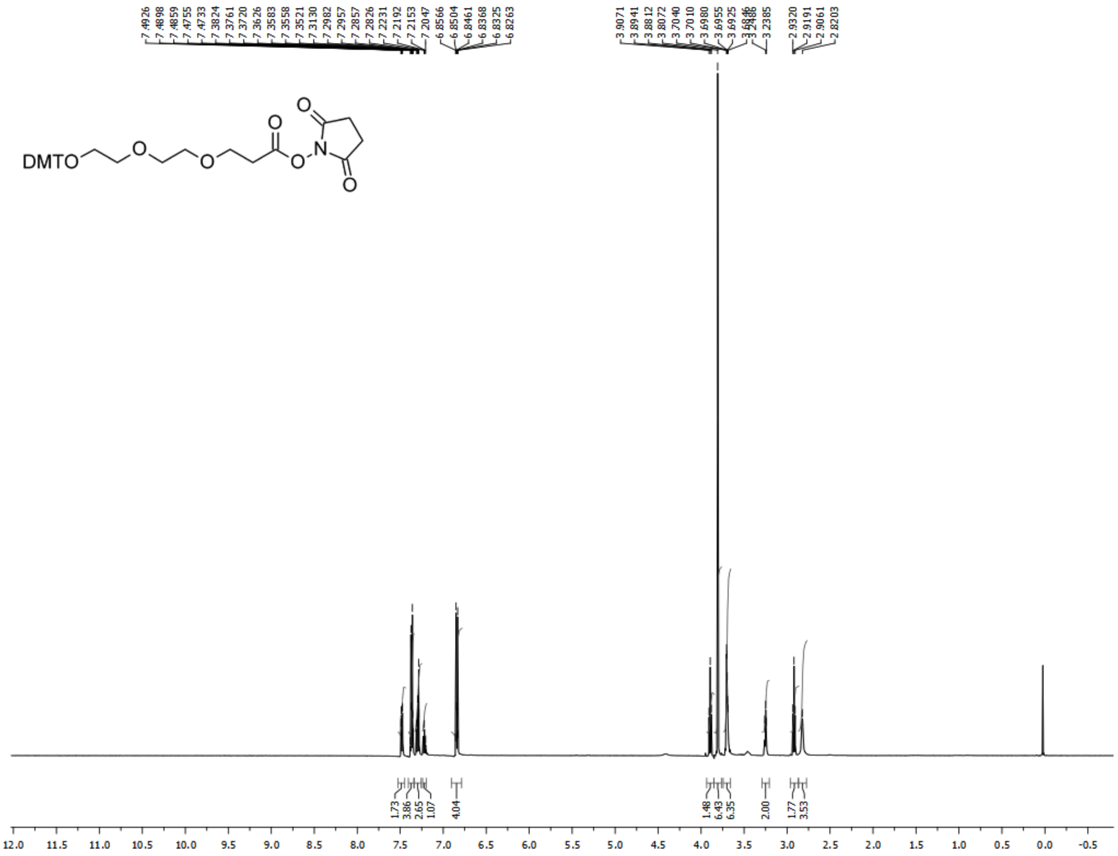

172.32
158.61
145.32
136.58
130.31
128.45
127.99
126.89
113.27
86.16
71.00
70.92
70.80
66.92
63.39
55.46
51.94
35.15
DMTO
O
O
O
OMe
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

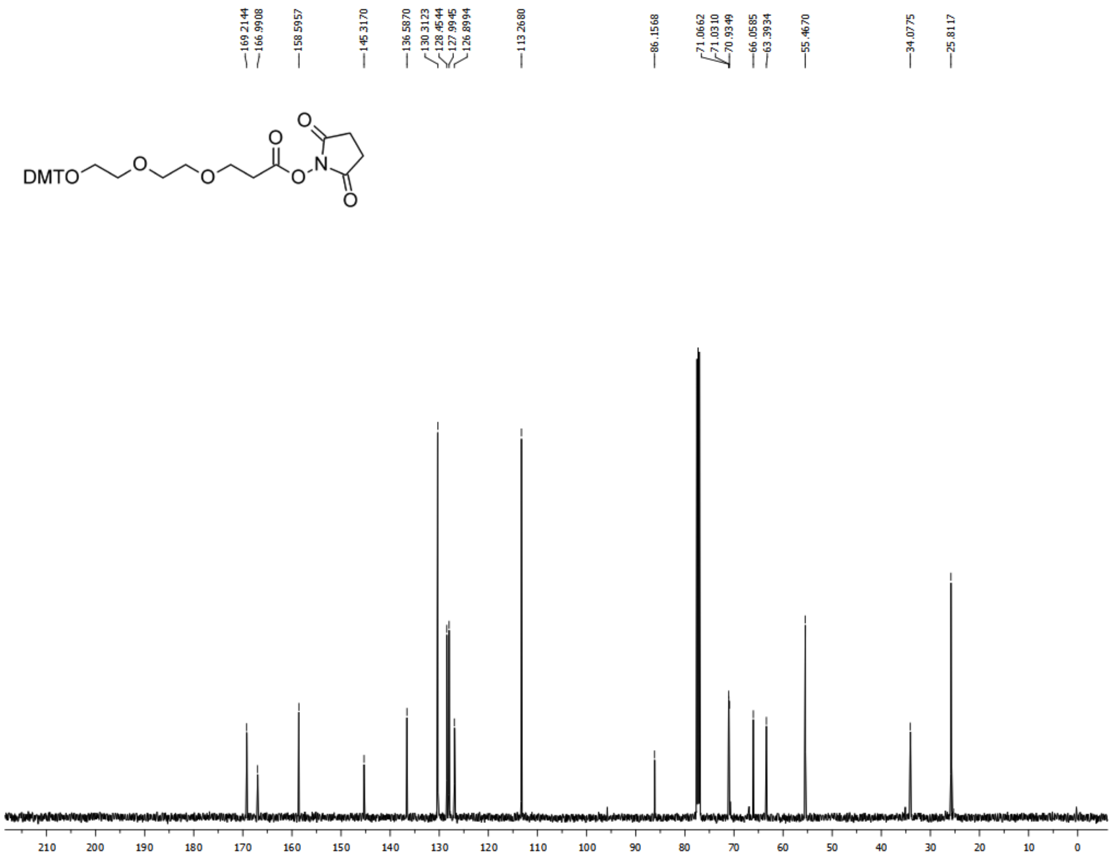

7.50
7.50
7.49
7.48
7.39
7.39
7.38
7.37
7.32
7.32
7.32
7.30
7.30
7.29
7.29
7.28
7.24
7.24
7.24
7.23
7.23
7.21
7.21
6.86
6.84
6.84
6.47
6.44
6.15
5.86
5.85
5.84
5.83
4.39
4.38
4.38
4.37
4.37
3.81
3.72
3.70
3.27
3.26
3.25
DMTO
O
O
O
2.23
4.56
2.69
1.10
4.48
1.00
0.90
0.97
2.16
9.20
2.24
2.17
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5

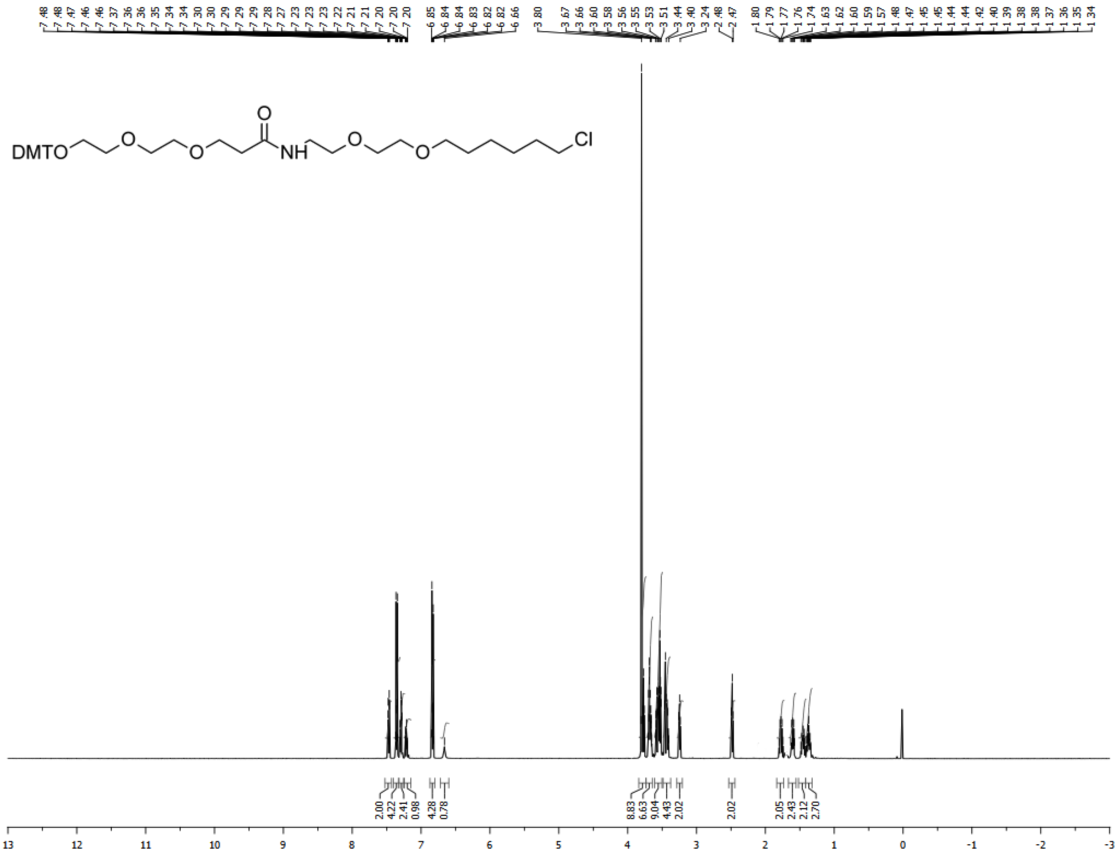

166.49
158.62
145.29
136.52
130.31
128.44
128.02
126.92
113.29
86.20
71.02
69.41
64.13
63.39
55.46
DMTO
O
O
O
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

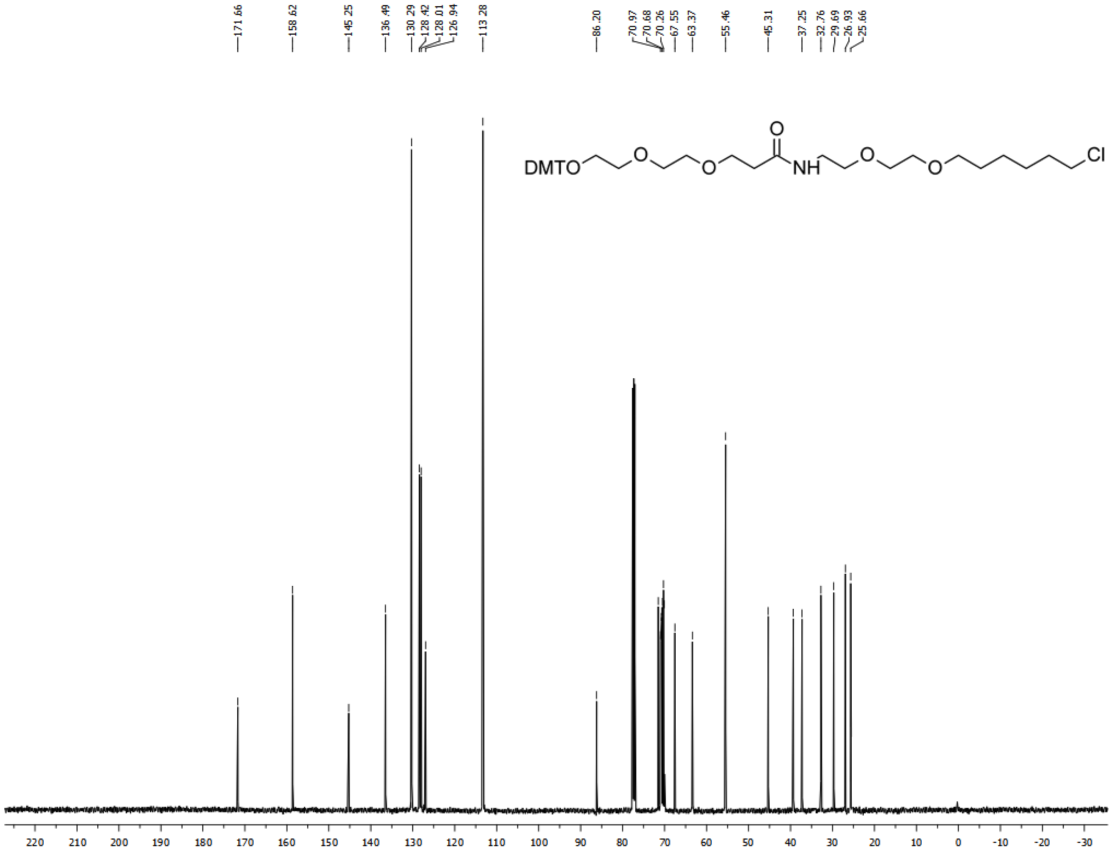

7.45
7.44
7.34
7.34
7.32
7.32
7.28
7.26
7.26
7.26
7.24
7.20
7.18
6.82
6.80
3.77
3.75
3.65
3.64
3.63
3.22
3.21
3.20
2.57
2.55
2.53
DMTO
O
O
O
OH
2.13
4.27
2.44
1.01
4.00
8.61
6.51
2.02
1.99
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0

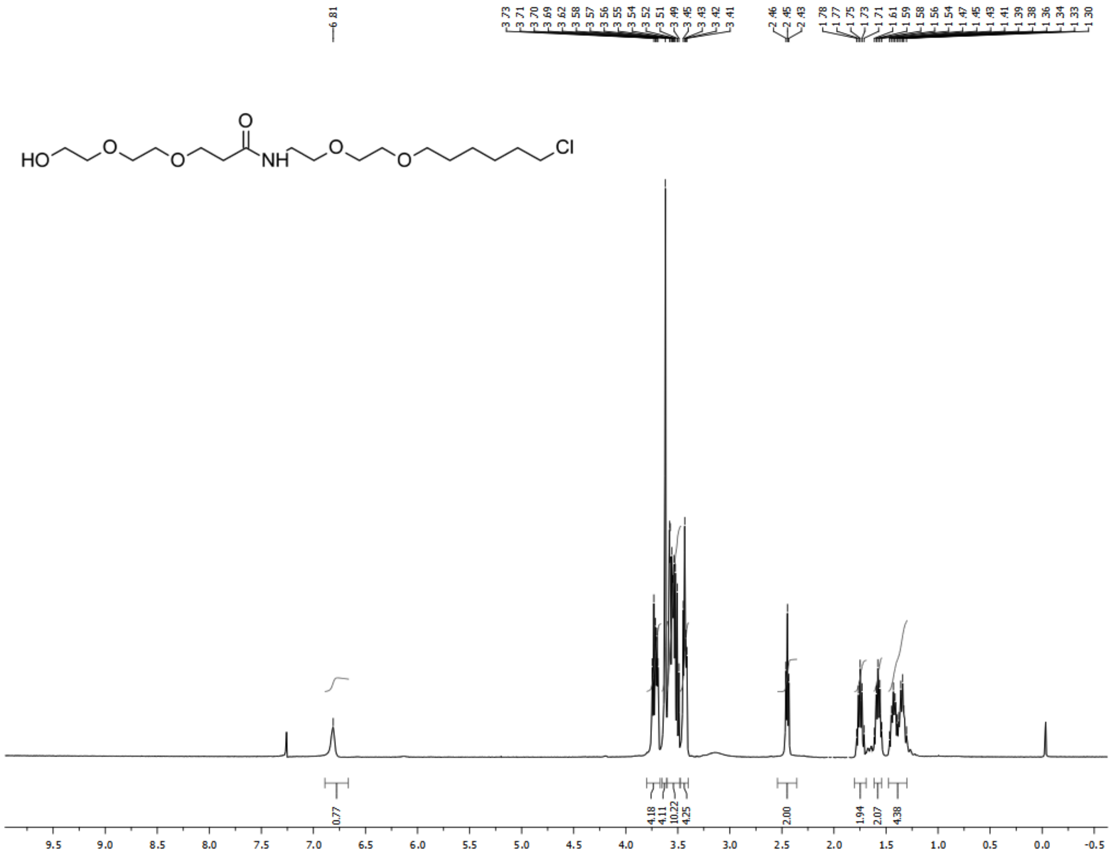

176.78
158.56
145.29
136.55
130.27
128.42
127.96
126.85
113.25
86.14
70.91
70.83
70.42
68.12
63.36
55.42
36.85
DMTO
O
O
O
OH
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

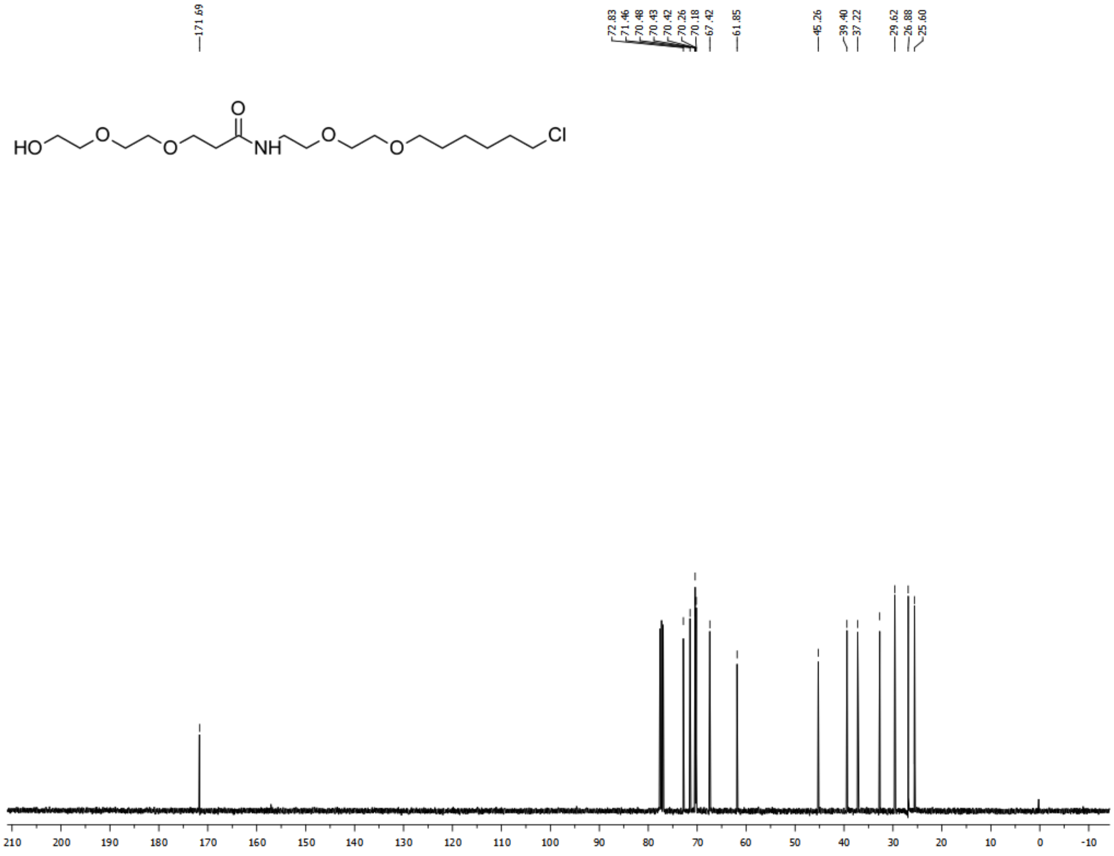

7.4926
7.4898
7.4859
7.4755
7.4733
7.3824
7.3761
7.3720
7.3626
7.3583
7.3558
7.3521
7.3130
7.2982
7.2957
7.2857
7.2826
7.2231
7.2192
7.2153
7.2047
6.8566
6.8504
6.8461
6.8368
6.8325
6.8263
3.9071
3.8941
3.8812
3.8072
3.7040
3.7010
3.6980
3.6955
3.6925
3.2385
2.9320
2.9191
2.9061
2.8203
DMTO
O
O
O
O
N
O
O
1.73
3.86
2.65
1.07
4.04
1.48
6.43
6.35
2.00
1.77
3.53
12.0
11.5
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5

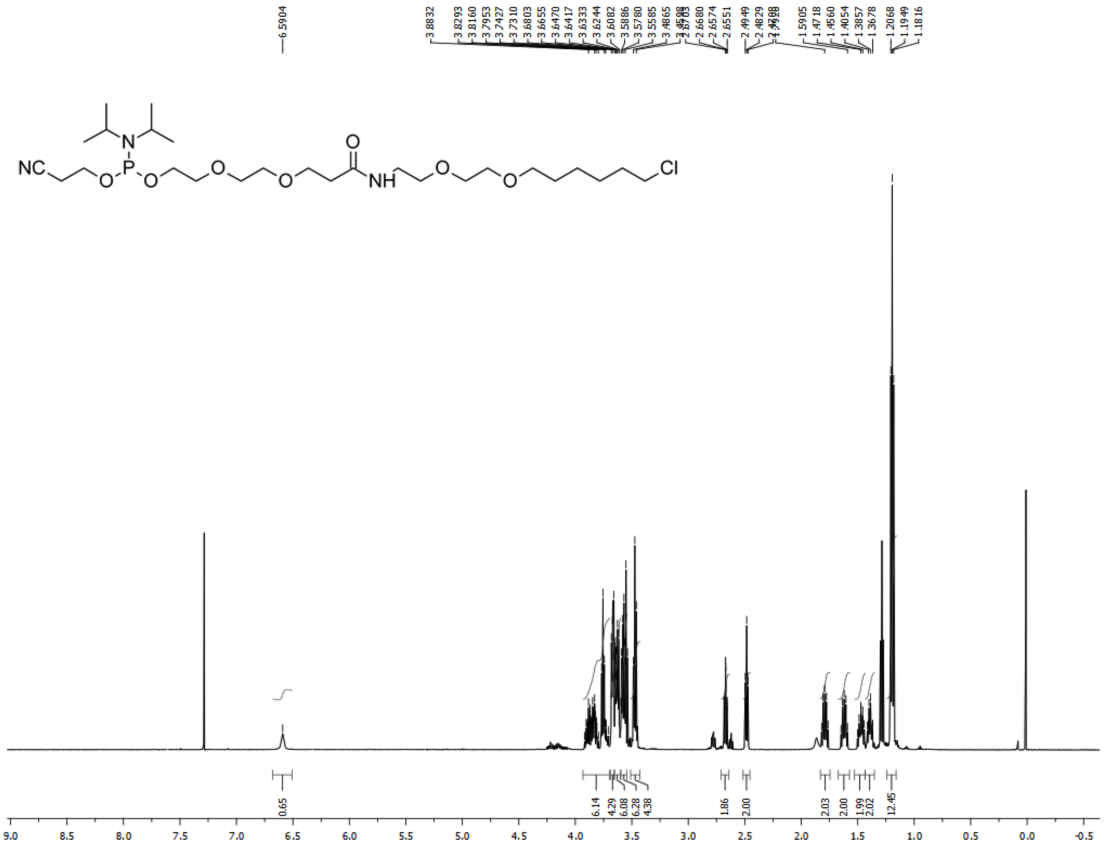

169.2144
166.9908
158.5957
145.3170
136.5870
130.3123
128.4544
127.9945
126.8994
113.2680
86.1568
71.0662
71.0310
70.9349
66.0585
63.3934
55.4670
34.0775
25.8117
DMTO
O
O
O
O
N
O
O
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

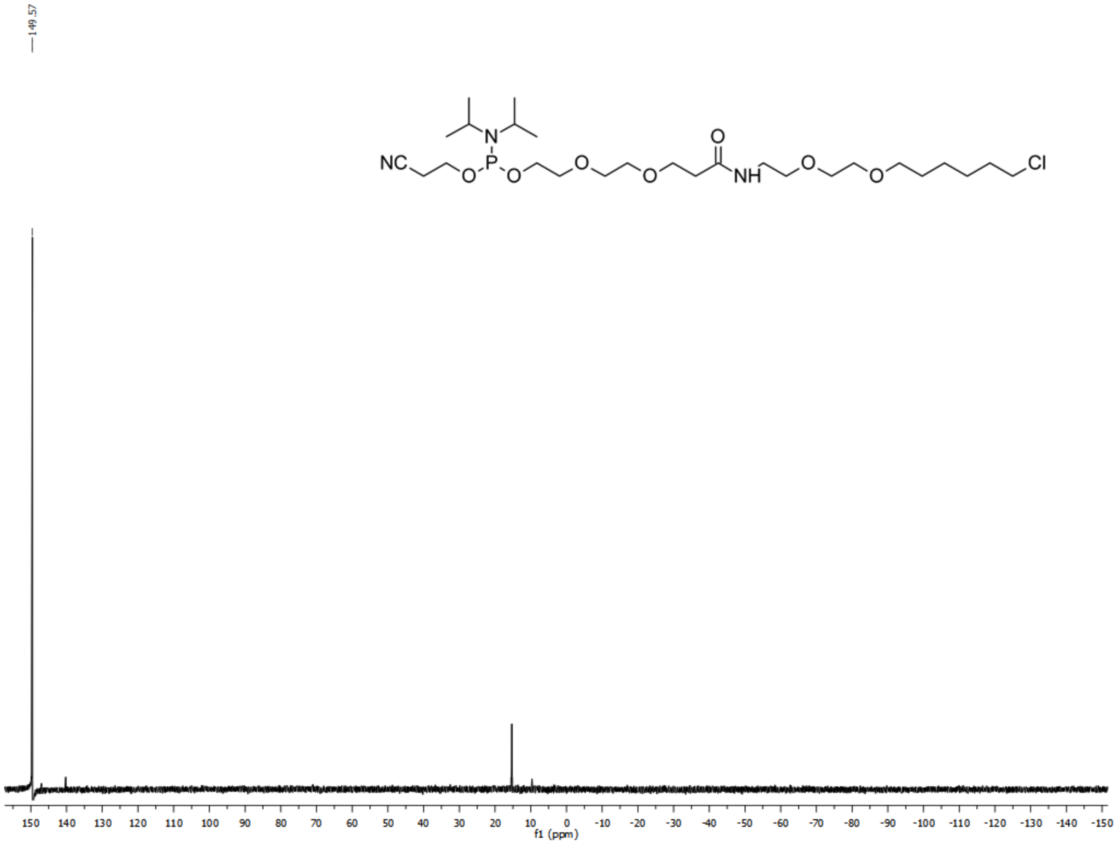

7.48
7.48
7.47
7.46
7.46
7.37
7.36
7.36
7.35
7.34
7.34
7.30
7.30
7.29
7.29
7.29
7.28
7.27
7.23
7.23
7.23
7.22
7.21
7.21
7.20
7.20
7.20
6.85
6.84
6.84
6.83
6.82
6.82
6.66
3.80
3.67
3.66
3.60
3.58
3.56
3.55
3.53
3.51
3.44
3.40
3.24
2.48
2.47
1.80
1.79
1.77
1.76
1.74
1.63
1.62
1.60
1.59
1.57
1.48
1.47
1.45
1.45
1.44
1.44
1.42
1.40
1.39
1.38
1.38
1.37
1.36
1.35
1.34
DMTO
O
O
O
NH
O
O
Cl
2.00
4.22
2.41
0.98
4.28
0.78
8.83
6.63
9.04
4.43
2.02
2.02
2.05
2.43
2.12
2.70
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
-2
-3

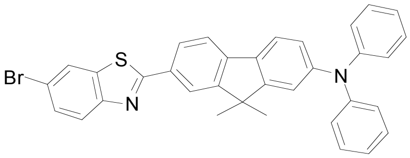

171.66
158.62
145.25
136.49
130.29
128.42
128.01
126.94
113.28
86.20
70.97
70.68
70.26
67.55
63.37
55.46
45.31
37.25
32.76
29.69
26.93
25.66
DMTO
O
O
O
NH
O
O
Cl
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30


6.81
3.73 3.71 3.70 3.69 3.62 3.58 3.57 3.56 3.55 3.54 3.52 3.51 3.49 3.45 3.43 3.42 3.41
2.46 2.45 2.43
1.78 1.77 1.75 1.73 1.71 1.61 1.59 1.58 1.56 1.54 1.47 1.45 1.43 1.41 1.39 1.38 1.36 1.34 1.33 1.30
HO
O
O
O
NH
O
O
Cl
0.77
4.18
4.11
10.22
4.25
2.00
1.94
2.07
4.38
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5

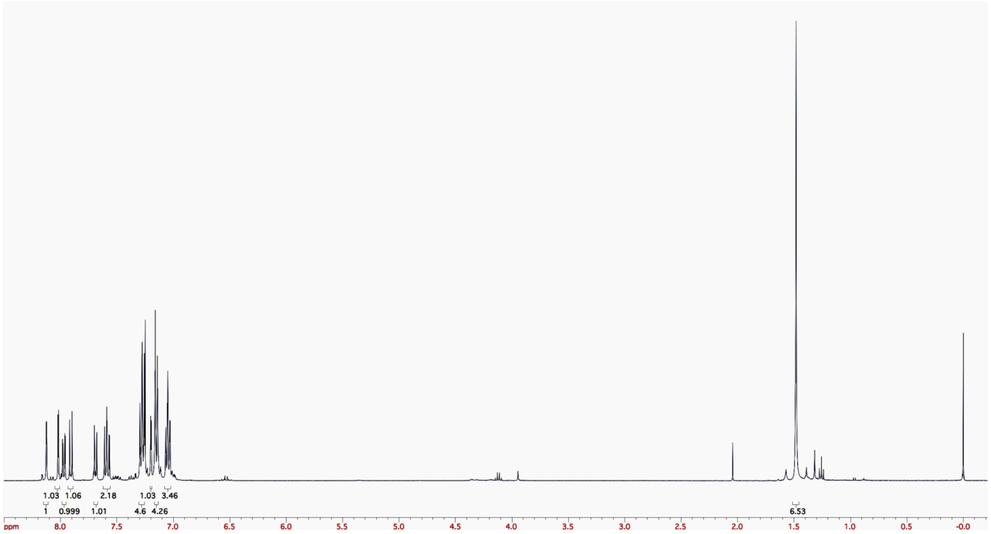

171.69
72.83 71.46 70.48 70.43 70.42 70.26 70.18 67.42
61.85
45.26
39.40 37.22
29.62 26.88 25.60
HO
O
O
O
NH
O
O
Cl
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10


6.5904
3.8832
3.8293
3.8160
3.7953
3.7427
3.7310
3.6803
3.6655
3.6470
3.6417
3.6333
3.6244
3.6082
3.5886
3.5780
3.5585
3.4865
2.6680
2.6574
2.6551
2.4949
2.4829
1.5905
1.4718
1.4560
1.4054
1.3857
1.3678
1.2068
1.1949
1.1816
NC
O
P
N
O
O
O
O
NH
O
O
Cl
0.65
6.14
4.29
6.08
6.28
4.38
1.86
2.00
2.03
2.00
1.99
2.02
12.45
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5

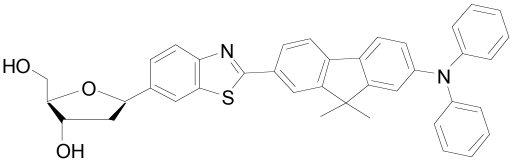

149.57
NC
O
P
N
O
O
O
O
NH
O
O
Cl
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
-110
-120
-130
-140
-150
f1 (ppm)



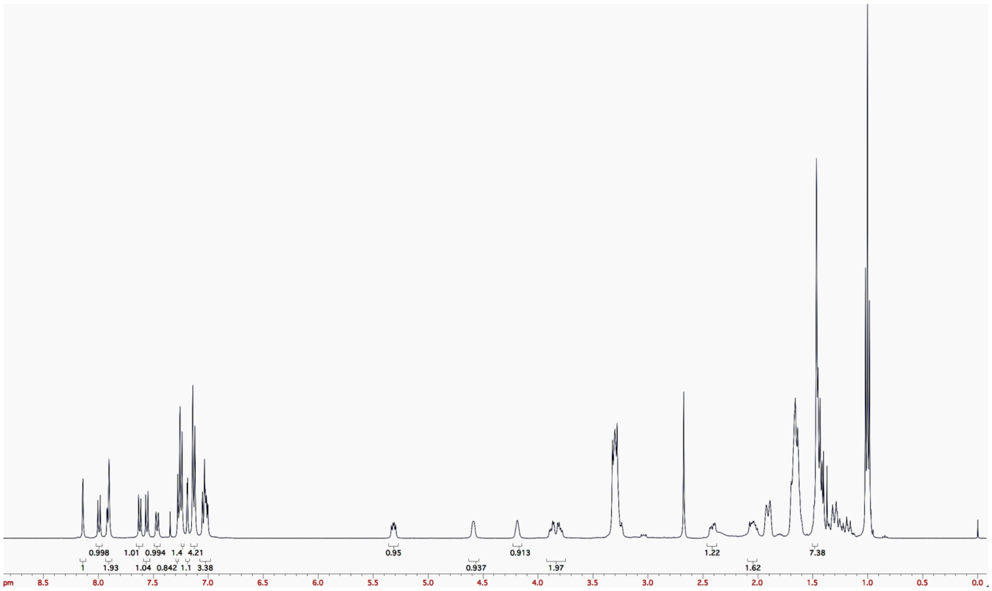
Br
S
N
N



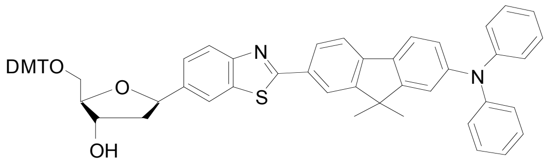
1.03
1.06
2.18
1.03
3.46
1
0.999
1.01
4.6
4.26
6.53
ppm
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
-0.0

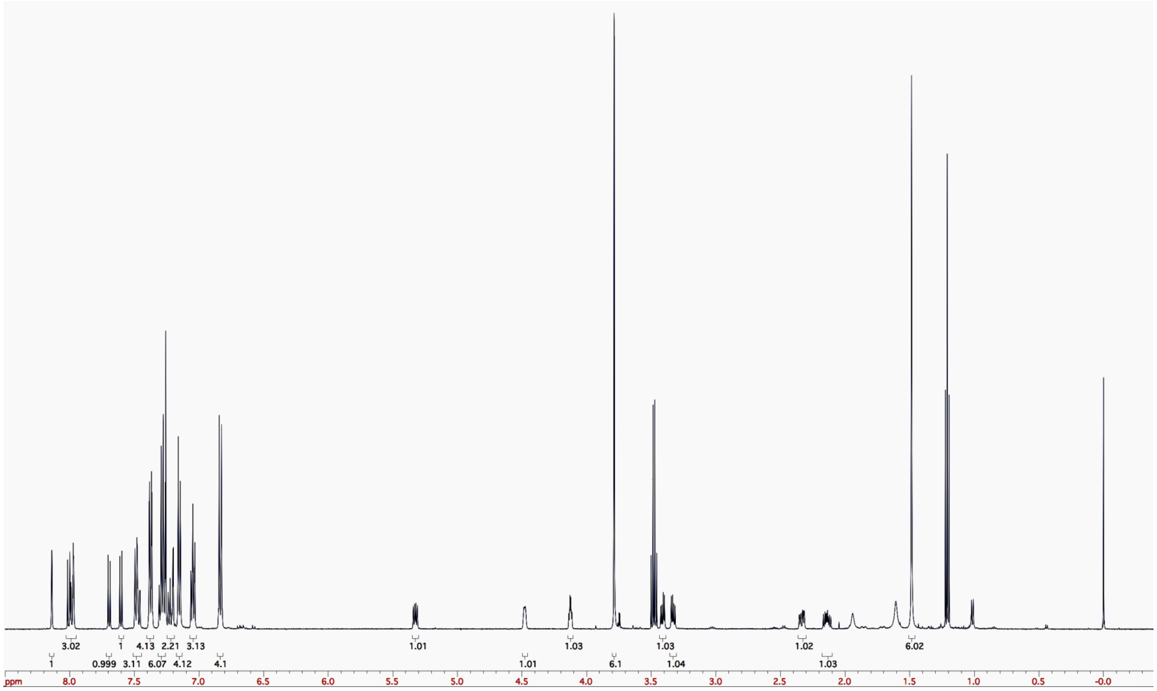


HO
O
OH
N
S
N



0.998
1.01
0.994
1.4
4.21
1
1.93
1.04
0.842
1.1
3.38
0.95
0.937
0.913
1.97
1.22
1.62
7.38
pm
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0


DMTO
O
OH
N
S
N


3.02
1
4.13
2.21
3.13
1
0.999
3.11
6.07
4.12
4.1
1.01
1.01
1.03
6.1
1.03
1.04
1.02
1.03
6.02
ppm
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
-0.0