Targeted Protein Internalization and Degradation by ENDosome TArgeting Chimeras (ENDTACs)

Dhanusha A. Nalawansha[1], Stacey-Lynn Paiva[1], Diane N. Rafizadeh[2], Mariell Pettersson[1], Liena Qin[1] and Craig M. Crews[1,2,3]

[1]Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06511, United States
[2]Department of Chemistry, Yale University, New Haven, Connecticut 06511, United States
[3]Department of Pharmacology, Yale University, New Haven, Connecticut 06511, United States

Supporting Information

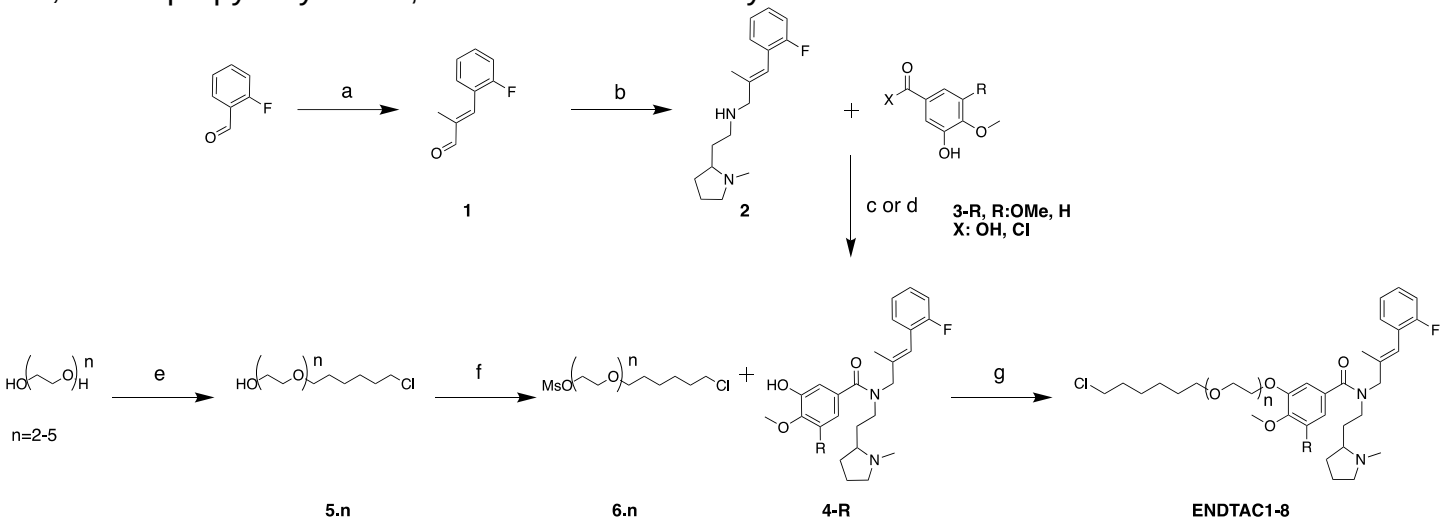
Materials and Methods

General Notes: Chemistry

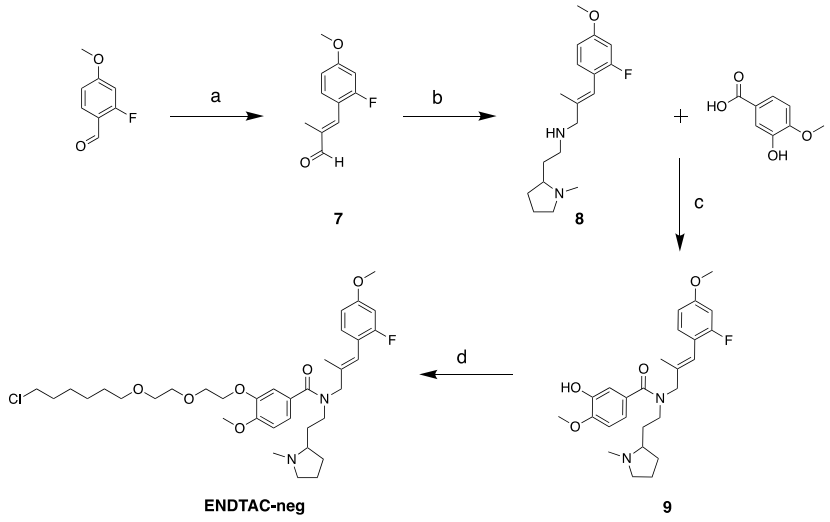
Unless otherwise indicated, common reagents or materials were obtained from commercial sources and used without further purification. Tetrahydrofuran (THF), Dimethylformamide (DMF) and Dichloromethane (DCM) were dried by a PureSolvTM solvent drying system. Flash column chromatography was performed using silica gel 60 (230-400 mesh). Analytical thin layer chromatography (TLC) was carried out on Merck silica gel plates with QF-254 indicator and visualized by UV or $KMnO_4$. Flash chromatography was performed using the Biotage Isolera One purification system. ^{1}H, ^{13}C, and ^{19}F NMR spectra were recorded on either Agilent DD2 500 (500 MHz ^{1}H; 125 MHz ^{13}C; 471 MHz ^{19}F) or Agilent DD2 600 (600 MHz ^{1}H; 151 MHz ^{13}C) or Agilent DD2 400 (400 MHz ^{1}H; 101 MHz ^{13}C, 376 MHz ^{19}F) or Varian 700 (700 MHz ^{1}H) spectrometer at room temperature (RT) unless otherwise indicated. Chemical shifts were reported in ppm relative to the residual DMSO-d_6 (δ 2.50 ppm ^{1}H; δ 39.52 ppm ^{13}C). NMR chemical shifts were expressed in ppm relative to internal solvent peaks, and coupling constants were measured in Hz. Low-resolution mass spectra (LRMS) were obtained using electrospray ionization (ESI) LCQ-Fleet mass spectrometer coupled to an Ultimate 3000 Ultra High Performance Liquid Chromatography (UHPLC) (C18 column) and Corona Veo RS. High-resolution mass spectra (HRMS) were obtained using a Waters Xevo Quad Time of Flight (Q-tof) mass spectrometer coupled to an UPLC system (C18 column). Preparative (prep) HPLC was carried out on 100 x 21.2 mm 110 Å C-18 column using gradient conditions (5 - 95% B, flow rate = 20.0 mL/min, 20 min) monitoring by UV for collection. The eluents used were: solvent A (H_2O with 0.1% trifluoroacetic acid (TFA)) and solvent B (CH_3CN with 0.1% TFA). **ENDTAC** purity was determined by LC-MS (λ= 254 nm) using gradient conditions (5 - 95% B_1, flow rate = 0.5 mL/min, 6 min). The eluents used here were: solvent A_1 (H_2O with 0.1% formic acid) and solvent B_1 (CH_3CN with 0.1% formic acid).

1-**9** were synthesized according to previously reported methods[1-2]. **ENDTAC-1-8** and **ENDTAC-neg** were synthesized according to **Scheme I** and **II**, respectively.

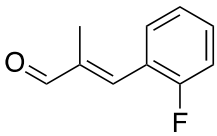
Supporting Scheme I. Synthesis of **ENDTAC1-8**. a) 1-propanal, KOH, EtOH/H_2O (95:5, v/v), 0 °C to RT, overnight; b) I. Na_2SO_4, DCM, RT, overnight, II. $NaBH_4$, MeOH, 0 °C to RT, 2 h; c) I. $COCl_2$, cat. DMF, DCM, RT, overnight, II. TEA, DCM, RT, 2-3 h; d) HATU, DIPEA, DMF, RT, overnight; e) I. NaH, THF, 0 °C, 30 min, II. 1-iodo-6-chlorohexane, 0 °C to RT, overnight; f) MsCl, Et_3N, DCM, RT, overnight; g) K_2CO_3, THF/DMF (1:1, v/v), 60 °C, overnight. TEA = Triethylamine; HATU = 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate; DIPEA = *N,N*-Diisopropylethylamine; MsCl = Methanesulfonyl chloride.



Supporting Scheme II. Synthesis of **ENDTAC-neg**. a) 1-propanal, KOH, EtOH/H_2O (95:5, v/v), 0 °C to RT, overnight; b) I. Na_2SO_4, DCM, RT, overnight, II. $NaBH_4$, MeOH, 0 °C to RT, 2 h; c) HATU, DIPEA, DMF, RT, overnight; d) **6.2**, K_2CO_3, THF/DMF (1:1, v/v), 60 °C, overnight.

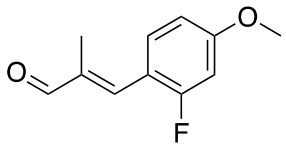


Characterization



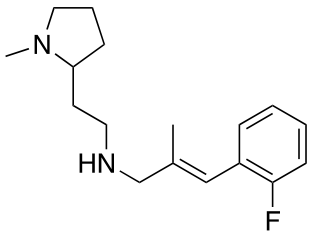
(E)-3-(2-Fluorophenyl)-2-methylacrylaldehyde (1)

2-Fluorobenzaldehyde (2.00 g, 16.1 mmol) was dissolved in a mixture of ethanol (EtOH)/H_2O (12.6 mL,110:1, v/v) and KOH (0.0904 g, 1.61 mmol). The mixture was cooled in an ice/ salt bath. 1-Propanal (1.03 g, 17.7 mmol) was added dropwise. The reaction mixture was stirred overnight allowing it to come to RT. Reaction was complete as analyzed by TLC and the solvent was removed *in vacuo*. The crude material was purified by flash chromatography using a gradient from 100% hexanes to 5% EtOAc in hexanes to provide the title compound as a pale yellow oil (1.70 g, 10.4 mmol, 64%).
^{1}H NMR (400 MHz, DMSO-d_6) δ 9.63 (s, 1H), 7.64 – 7.58 (m, 1H), 7.54 (s, 1H), 7.52 – 7.45 (m, 1H), 7.35 – 7.26 (m, 2H), 1.89 (s, 3H).
^{13}C NMR (101 MHz, DMSO-d_6) δ 195.5, 159.7 (d, *J* = 249.5 Hz), 140.9 (d, *J* = 4.3 Hz), 139.8, 131.7 (d, *J* = 8.7 Hz), 130.5 (d, *J* = 2.4 Hz), 124.6 (d, *J* = 3.5 Hz), 122.6 (d, *J* = 13.0 Hz), 115.8 (d, *J* = 21.6 Hz), 10.6 (d, *J* = 1.4 Hz).
^{19}F NMR (376 MHz, DMSO-d_6) δ -113.61 (dt, *J* = 11.9, 6.9 Hz).
LRMS (ESI): m/z [M+H]+ calc'd for $C_{10}H_{10}FO^+$ = 165.07; found = 164.95.



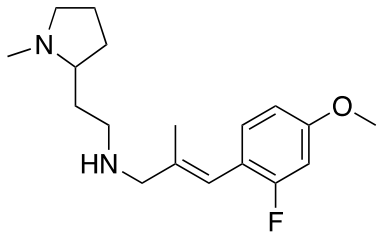
(*E*)-3-(2-Fluoro-4-methoxyphenyl)-2-methylacrylaldehyde (7)

2-Fluoro-5-methoxy-benzaldehyde (1.00 g, 6.49 mmol) was dissolved in a mixture of EtOH/H_2O (5.05 mL,110:1, v/v) and KOH (0.0728 g, 1.30 mmol). The mixture was cooled in an ice/salt bath. 1-Propanal (0.750 g, 13.0 mmol) was added dropwise. The reaction mixture was stirred overnight allowing it to come to RT. Reaction was complete as analyzed by analytical TLC and the solvent was removed *in vacuo*. The reaction mixture was diluted with EtOAc (100 mL) and washed with saturated $NaHCO_3$ (1 x 25 mL) followed by brine (3 x 25 mL each). The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude material was purified by flash chromatography using a gradient from 100% hexanes to 20% EtOAc in hexanes to provide the title compound as a pale yellow oil (0.850 g, 4.38 mmol, 67.5%).
^{1}H NMR (400 MHz, DMSO-d_6) δ 9.58 (s, 1H), 7.63 (t, *J* = 8.9 Hz, 1H), 7.48 (s, 1H), 6.99 (dd, *J* = 12.7, 2.5 Hz, 1H), 6.91 (dd, *J* = 8.8, 2.5 Hz, 1H), 3.83 (s, 3H), 1.92 (s, 3H).
^{13}C NMR (101 MHz, DMSO-d_6) δ 195.5, 162.1 (d, *J* = 11.7 Hz), 161.0 (d, *J* = 250.0 Hz), 140.8 (d, *J* = 4.3 Hz), 137.5, 131.3 (d, *J* = 4.2 Hz), 114.9 (d, *J* = 12.6 Hz), 111.0 (d, *J* = 2.8 Hz), 101.8 (d, *J* = 25.7 Hz), 55.9, 10.7.
^{19}F NMR (376 MHz, DMSO-d_6) δ -111.14 (dd, *J* = 12.6, 9.1 Hz).
LRMS (ESI): m/z [M+H]+ calc'd for $C_{11}H_{12}FO_2^+$ = 195.08; found = 195.01.



(*E*)-3-(2-Fluorophenyl)-2-methyl-N-(2-(1-methylpyrrolidin-2-yl)ethyl)prop-2-en-1-amine (2)

1 (0.910 g, 5.54 mmol), 2-(1-methylpyrrolidin-2-yl)ethanamine (0.711 g, 5.54 mmol), and Na_2SO_4 (4.52 g, 31.8 mmol) were stirred in dichloromethane (DCM, 18.0 mL) overnight at RT. Reaction was monitored by TLC for imine formation. Upon completion, reaction mixture was filtered, collected solid was washed with DCM, and the filtrate was concentrated *in vacuo*. The crude imine was solubilized in methanol (MeOH) (9 mL) and cooled to 0 °C over an ice bath. $NaBH_4$ (0.231 g, 6.1 mmol) was added to the imine and stirred for 1 h. Upon completion, the reaction was quenched with acetone (50 mL) for 10 min and concentrated *in vacuo*. Crude material was solubilized in EtOAc (100 mL) and washed with $NaHCO_3$ (3 x 25 mL washes) and brine (3 x 25 mL washes). The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude material was purified by flash chromatography using a gradient with a mixture of NH_4OH/MeOH/DCM (0.1:0.7:99.2 to 1:7:92, v/v/v) to afford a pale yellow oil (0.590 g, 2.13 mmol, 38.5%).
^{1}H NMR (400 MHz, DMSO-d_6) δ 7.34 – 7.23 (m, 2H), 7.19 – 7.12 (m, 2H), 6.38 (s, 1H), 3.22 (s, 2H), 2.94 – 2.84 (m, 1H), 2.58 – 2.40 (m, 3H), 2.18 (s, 3H), 1.99 (q, *J* = 8.9 Hz, 2H), 1.90 – 1.75 (m, 2H), 1.73 (s, 3H), 1.64 – 1.51 (m, 2H), 1.42 – 1.26 (m, 2H).
^{13}C NMR (101 MHz, DMSO-d_6) δ 159.43 (d, *J* = 244.5 Hz), 140.39, 130.58 (d, *J* = 3.7 Hz), 128.28 (d, *J* = 8.2 Hz), 125.21 (d, *J* = 14.6 Hz), 123.96 (d, *J* = 3.4 Hz), 116.72 (d, *J* = 2.7 Hz), 115.23 (d, *J* = 22.1 Hz), 64.09, 57.04, 56.69, 45.87, 40.19, 33.40, 30.25, 21.62, 16.44.
^{19}F NMR (376 MHz, DMSO-d_6) δ -115.48 – -115.59 (m).
LRMS (ESI): m/z [M+H]+ calc'd for $C_{17}H_{26}FN_2^+$= 277.21; found = 277.14.



(*E*)-3-(2-Fluoro-4-methoxyphenyl)-2-methyl-N-(2-(1-methylpyrrolidin-2-yl)ethyl)prop-2-en-1-amine (8)

7 (0.500 g, 2.57 mmol), 2-(1-methylpyrrolidin-2-yl)ethanamine (0.330 g, 2.57 mmol), and Na_2SO_4 (2.10 g, 14.8 mmol) were stirred in (DCM, 12.0 mL) overnight at RT. Reaction was monitored by TLC for imine formation. Upon completion, reaction mixture was filtered, collected solid was washed with DCM, and the filtrate was concentrated *in vacuo*. The crude imine was solubilized in MeOH (6 mL) and cooled to 0 °C over an ice bath. $NaBH_4$ (0.107 g, 2.83 mmol) was added to the imine and stirred for 1 h. Upon completion, the reaction was quenched with acetone (20 mL) for 10 min and concentrated *in vacuo*. Crude material was solubilized in EtOAc (100 mL) and washed with $NaHCO_3$ (3 x 25 mL washes) and brine (3 x 25 mL washes). The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude material was purified by flash chromatography using a gradient with a mixture of DCM/MeOH/NH_4OH (99.2:0.7:0.1 to 92:7:1, v/v/v) to afford a pale yellow oil (0.340 g, 1.11 mmol, 43.1%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 7.22 (t, J = 8.8 Hz, 1H), 6.85 – 6.73 (m, 2H), 6.28 (s, 1H), 3.76 (s, 3H), 2.97 – 2.83 (m, 1H), 2.57 – 2.51 (m, 1H), 2.48 – 2.39 (m, 1H), 2.18 (s, 3H), 2.06 – 1.93 (m, 1H), 1.92 – 1.72 (m, 2H), 1.72 (s, 3H), 1.67 – 1.52 (m, 2H), 1.42 – 1.20 (m, 2H).

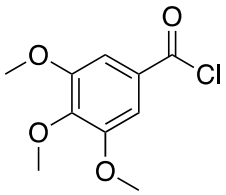
^{13}C NMR (101 MHz, DMSO-d_6) δ 160.0 (d, J = 244.4 Hz), 159.2, 138.8, 130.9 (d, J = 5.7 Hz), 117.2, 116.5 (d, J = 2.3 Hz), 109.9 (d, J = 2.9 Hz), 101.5, 101.2, 64.1, 57.2, 56.7, 55.5, 45.9, 33.4, 30.3, 21.6, 16.4.

^{19}F NMR (376 MHz, DMSO-d_6) δ -113.24 – -113.34 (m).

LRMS (ESI): m/z [M+H]+ calc'd for $C_{18}H_{28}FN_2O^+$= 307.22; found = 307.10.

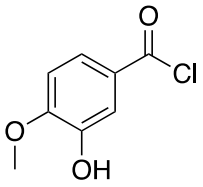
General acid chloride formation

DCM (0.1 M) was added to a reaction flask charged the benzoic acid and purged with argon. The reaction flask was then charged with DMF (~2-5 drops) followed by oxalyl chloride dropwise. Reaction was monitored by TLC (monitored for methyl ester formation in MeOH diluted sample) and/or LC-MS, where indicated. Upon completion, the reaction was concentrated *in vacuo*, azeotroped (3 x with DCM) and used immediately in this crude form.



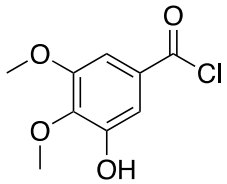
3,4,5-Trimethoxybenzoyl chloride

Using the general acid chloride formation procedure using 3,4,5-trimethoxybenzoic acid (2.00 g, 9.42 mmol) and oxalyl chloride (1.23 mL, 14.1 mmol), the crude product was isolated as an off-white solid.



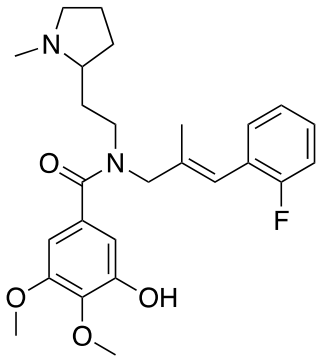
3,4-Dimethoxybenzoyl chloride

Using the general acid chloride formation procedure using 3,4,5-trimethoxybenzoic acid (2.00 g, 10.9 mmol) and oxalyl chloride (1.44 mL, 16.5 mmol), the crude product was isolated as a brown solid.



3-Hydroxy-4,5-dimethoxybenzoyl chloride (3, R = OMe, X = Cl)
Using the general acid chloride formation procedure using 3-hydroxy-4,5-dimethoxybenzoic acid (0.800 g, 4.04 mmol) and oxalyl chloride (0.528 mL, 6.06 mmol), the crude product was an off-white solid.
MS for methyl ester: calc. $[M+H]^+$ for $C_{10}H_{13}O_5^+$= 213.08; found = 212.93

Amide bond formation



(E)-N-(3-(2-Fluorophenyl)-2-methylallyl)-3-hydroxy-4,5-dimethoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide (4-OMe)

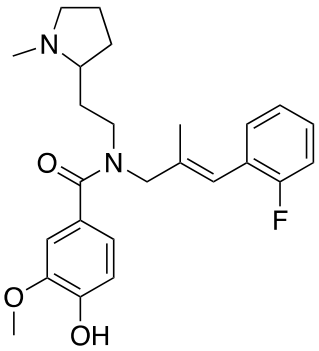
2 (0.270 g, 0.977 mmol) was mixed with triethylamine (TEA) (0.494 g, 0.681 mL, 4.88 mmol) in anhydrous DCM (5 mL), to which **3-OMe** (**X**=Cl) (0.38 g, 1.76 mmol) in anhydrous DCM (5 mL) was added dropwise. Reaction complete by TLC after stirring at RT for 2 h. The reaction was concentrated *in vacuo*. Crude material was solubilized in DCM (~200 mL) and washed with saturated $NaHCO_3$ (3 x 50 mL each), and brine (3 x 50 mL each). Organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude material was purified by flash chromatography using a gradient with a mixture of DCM/MeOH/NH_4OH (99.2:0.7:0.1 to 92:7:1, v/v/v) to afford a pale yellow oil (0.21 g, 0.46 mmol, 47.1%).

^{1}H NMR (500 MHz, DMSO-d_6, 100 °C) δ 8.98 (s, 1H), 7.36 – 7.26 (m, 2H), 7.23 – 7.12 (m, 2H), 6.52 (d, J = 12.3 Hz, 2H), 6.39 (s, 1H), 4.13 (s, 2H), 3.78 (s, 3H), 3.74 (s, 3H), 3.34 (t, J = 7.1 Hz, 2H), 2.99 (br, 1H), 2.93 – 2.85 (m, 1H), 2.17 (s, 3H), 2.11 – 1.96 (m, 2H), 1.92 – 1.82 (m, 1H), 1.82 – 1.75 (m, 1H), 1.71 (s, 3H), 1.63 – 1.50 (m, 2H), 1.38 – 1.24 (m, 1H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 170.5, 159.5 (d, J = 244.9 Hz), 153.0, 150.5, 137.0, 132.2, 130.6 (d, J = 3.4 Hz), 128.9 (d, J = 8.8 Hz), 124.5, 124.2 (d, J = 3.4 Hz), 118.8, 115.4 (d, J = 21.9 Hz), 107.8, 107.3, 101.7, 63.7, 56.6, 55.8, 50.7, 44.6, 41.7, 30.8, 29.9, 29.2, 21.6, 15.7.

^{19}F NMR (471 MHz, DMSO-d_6, 100 °C) δ -115.26.

LRMS (ESI): m/z [M+H]+ calc'd for $C_{26}H_{34}FN_2O_4^+$ = 457.25; found = 457.23.



(E)-N-(3-(2-Fluorophenyl)-2-methylallyl)-4-hydroxy-3-methoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide (4-H)

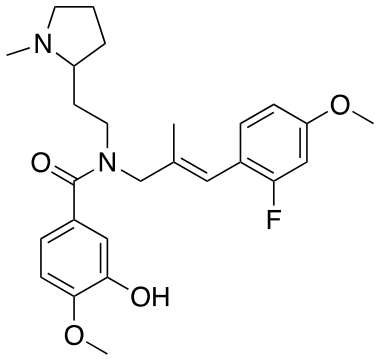
3-hydroxy-4-methoxy-benzoic acid (0.210 g, 0.760 mmol) was solubilized in DMF (2 mL) to which 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate(HATU) (0.166 g, 0.985 mmol) and N,N-Diisopropylethylamine (DIPEA) (0.295 g, 0.390 mL, 2.28 mmol) were added and allowed to stir at RT for 5 min. **2** (0.21 g, 0.76 mmol) in DMF (1 mL) was added and allowed to stir at RT over 72 h. Reaction was diluted in EtOAc (~100 mL) and washed with saturated $NaHCO_3$ (25 mL), and with brine (4 x 25 mL each). Organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude material was purified by flash chromatography using a gradient with a mixture of DCM/MeOH/NH_4OH (99.2:0.7:0.1 to 92:7:1, v/v/v) to afford a pale yellow oil (0.049 g, 0.11 mmol, 15.1%).

^{1}H NMR (500 MHz, DMSO-d_6, 100 °C) δ 8.75 (s, 1H), 7.38 – 7.25 (m, 2H), 7.22 – 7.12 (m, 2H), 6.96 (d, *J* = 8.2 Hz, 1H), 6.86 (d, *J* = 1.6 Hz, 1H), 6.82 (dd, *J* = 8.2, 1.8 Hz, 1H), 6.38 (s, 1H), 4.13 (s, 2H), 3.82 (s, 3H), 3.35 (t, *J* = 7.7 Hz, 2H), 2.95 – 2.85 (m, 1H), 2.17 (s, 3H), 2.10 – 1.94 (m, 2H), 1.92 – 1.81 (m, 1H), 1.82 – 1.73 (m, 1H), 1.69 (s, 3H), 1.64 – 1.47 (m, 3H), 1.32 (s, 1H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 170.7, 159.4 (d, *J* = 244.9 Hz), 148.5, 146.3, 137.0, 130.6 (d, *J* = 3.4 Hz), 129.2, 128.9 (d, *J* = 8.2 Hz), 124.2 (d, *J* = 3.4 Hz), 118.7, 117.6, 115.5, 115.3, 114.0, 111.7, 63.2, 56.6, 55.6, 44.6, 41.4, 30.9, 30.3, 29.7, 21.6, 15.7.

^{19}F NMR (471 MHz, DMSO-d_6, 100 °C) δ -115.2.

LRMS (ESI): m/z [M+H]+ calc'd for $C_{25}H_{32}FN_2O_3^+$ = 427.24; found = 427.25.



(E)-N-(3-(2-Fluoro-4-methoxyphenyl)-2-methylallyl)-3-hydroxy-4-methoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide (9)

3-Hydroxy-4-methoxy-benzoic acid (0.0543 g, 0.323 mmol) was solubilized in DMF (1.5 mL) to which HATU (0.223 g, 0.587 mmol) and DIPEA (0.144 g, 0.151 mL, 0.881 mmol) were added and allowed to stir at RT for 5 min. **8** (0.0900 g, 0.294 mmol) in DMF (1.5 mL) was added and allowed to stir at RT over 72 h. Reaction was diluted in EtOAc (~100 mL) and washed with saturated $NaHCO_3$ (25 mL), and brine (4 x 25 mL each). Organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude material was purified by flash chromatography using a gradient with a mixture of DCM/MeOH/NH_4OH (99.2:0.7:0.1 to 92:7:1, v/v/v) to afford a pale yellow oil (0.074 g, 0.16 mmol, 55.2%).

^{1}H NMR (500 MHz, DMSO-d_6, 100 °C) δ 8.75 (s, 1H), 7.23 (t, *J* = 8.5 Hz, 1H), 6.95 (d, *J* = 8.1 Hz, 1H), 6.86 – 6.76 (m, 4H), 6.30 (s, 1H), 4.11 (s, 2H), 3.82 (s, 3H), 3.80 (s, 3H), 3.33 (t, *J* = 7.6 Hz, 2H), 2.93 – 2.85 (m, 1H), 2.17 (s, 3H), 2.12 – 1.97 (m, 2H), 1.90 – 1.74 (m, 2H), 1.67 (s, 3H), 1.62 – 1.47 (m, 3H), 1.37 – 1.25 (m, 1H).

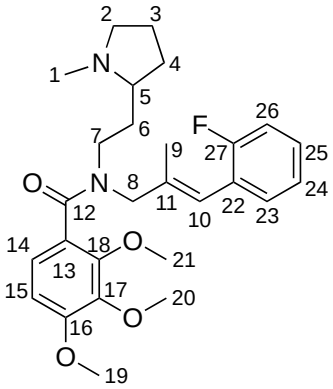
^{13}C NMR (151 MHz, DMSO-d_6) δ 170.7, 160.1 (d, *J* = 244.9 Hz), 159.5 (d, *J* = 11.0 Hz), 148.4, 146.2, 135.3, 130.9 (d, *J* = 5.4 Hz), 129.6 – 128.8 (m), 118.5, 117.6, 116.5 (d, *J* = 4.2 Hz), 115.1 – 113.0 (m), 111.6, 110.1 (d, *J* = 2.7 Hz), 101.5 (d, *J* = 25.9 Hz), 63.1, 56.5, 55.6, 50.8, 44.6, 41.5, 30.8, 29.9, 29.3, 21.6, 15.6.

^{19}F NMR (471 MHz, DMSO-d_6, 100 °C) δ -112.9.

LRMS (ESI): m/z [M+H]+ calc'd for $C_{26}H_{34}FN_2O_4^+$ = 457.25; found = 457.23.

Warhead synthesis

VUF11207 and **VUF11403** were synthesized as previously described to afford thick colourless oils [2].{



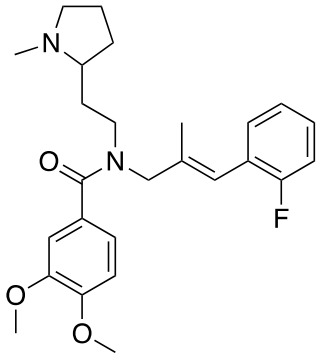
(E)-N-(3-(2-Fluorophenyl)-2-methylallyl)-3,4,5-trimethoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide (VUF11207)

^{1}H NMR (500 MHz, DMSO-d_6, 100 °C) δ 7.33 (q, J = 7.3 Hz, 2H), 7.20 (dd, J = 14.0, 6.4 Hz, 2H), 6.72 (s, 2H), 6.40 (s, 1H), 4.12 (s, 2H), 3.80 (s, 6H), 3.75 (s, 3H), 3.60 (br, 1H), 3.47 (br, 2H), 3.28 (br, 1H), 3.10 (br, 1H), 2.86 (s, 3H), 2.27 (br, 2H), 2.03 (br, 1H), 1.94 (br, 1H), 1.88 (br, 1H), 1.70 (s, 3H), 1.58 (br, 1H).

^{13}C NMR (151 MHz, DMSO-d_6) δ 170.7, 159.4 (d, J = 244.6 Hz), 152.7, 138.4, 137.0, 131.2, 130.6 (d, J = 2.9 Hz), 129.0, 124.2 (d, J = 3.3 Hz), 117.9, 117.2, 115.4 (d, J = 22.0 Hz), 115.3, 103.9, 66.3, 60.1, 55.9, 55.7, 55.2, 42.4, 40.1, 29.0, 27.8, 21.1, 15.9.

^{19}F NMR (471 MHz, DMSO-d_6, 100 °C) δ -74.37 (s) (TFA), -115.26 (br).

HRMS (Q-tof): m/z $[M+H]^+$ calc'd for $C_{27}H_{36}FN_2O_4^+$= 471.2659; found = 471.2660.



(E)-N-(3-(2-Fluorophenyl)-2-methylallyl)-3,4-dimethoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide (VUF11403)

^{1}H NMR (500 MHz, DMSO-d_6, 100 °C) δ 7.33 (q, J = 6.8, 6.2 Hz, 2H), 7.24 – 7.14 (m, 2H), 7.03 (s, 3H), 6.39 (s, 1H), 4.13 (s, 2H), 3.82 (s, 3H), 3.78 (d, J = 1.3 Hz, 3H), 3.59 (s, 1H), 3.49 (d, J = 7.8 Hz, 2H), 3.28 (s, 1H), 3.16 – 3.05 (m, 1H), 2.86 (s, 3H), 2.26 (s, 2H), 2.03 (s, 1H), 1.94 (s, 1H), 1.86 (d, J = 9.9 Hz, 1H), 1.68 (s, 3H), 1.65 (s, 1H).

^{13}C NMR (151 MHz, DMSO-d_6) δ 170.9, 159.4 (d, J = 244.7 Hz), 149.8, 148.2, 136.9, 130.6, 129.0, 124.2, 119.6, 118.9, 117.9, 116.9, 115.4 (d, J = 21.9 Hz), 115.0, 111.2, 110.2, 66.3, 55.6, 55.5, 55.2, 42.2, 40.0, 29.0, 27.8, 21.0, 15.8.

^{19}F NMR (471 MHz, DMSO-d_6, 100 °C) δ -74.56 (s) (TFA), -115.23 (s).

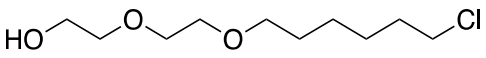
HRMS (Q-tof): m/z $[M+H]^+$ calc'd for $C_{26}H_{34}FN_2O_3^+$= 441.2553; found = 441.2562.

Chloroalkane linker synthesis

6n were synthesized according to previous reported methods[1].

General chloroalkane alkylation

The select ethylene glycol (5 equiv) was added dropwise to a mixture of NaH in 1:1 DMF:THF (0.2 M) over an ice bath under inert atmosphere. After 30-40 min, 1-chloro-6-iodohexane (1 equiv) was added, and the mixture was stirred overnight, allowing it to reach RT (RT). The mixture was then quenched with water, diluted with 1 M HCl (~ 200 mL total aqueous volume), and extracted with DCM (3 x 50 mL each). The organic layers were combined, dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude material was purified via flash chromatography using a gradient from 25 to 100% EtOAc in hexanes to give the monoalkylated product as a colorless oil.



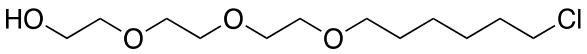
2-(2-((6-Chlorohexyl)oxy)ethoxy)ethan-1-ol (5.2)

Using the general chloroalkane mono-alkylation procedure with diethylene glycol (3.51 g, 3.13 mL, 33.1 mmol), NaH (0.633 g, 16.5 mmol), and 1-chloro-6-iodohexane (1.63 g, 1 mL, 6.61 mmol) **5.2** was afforded as a colorless oil (0.470 g, 2.09 mmol, 31.6%).

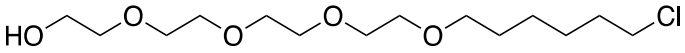
^{1}H NMR (400 MHz, DMSO-d_6) δ 3.62 (t, J = 6.6 Hz, 2H), 3.54 – 3.44 (m, 6H), 3.43 – 3.34 (m, 4H), 1.71 (p, J = 6.7 Hz, 2H), 1.49 (p, J = 6.7 Hz, 2H), 1.42 – 1.24 (m, 4H).

^{13}C NMR (100 MHz, DMSO-d_6) δ 72.3, 70.2, 69.8, 69.5, 60.2, 45.3, 32.0, 29.1, 26.1, 24.9.

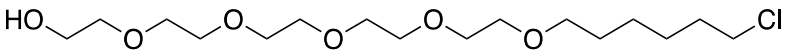
LRMS (ESI): m/z [M+H]+ calc'd for $C_{10}H_{22}ClO_3^+$ = 225.13; found = 225.04.



2-(2-(2-((6-Chlorohexyl)oxy)ethoxy)ethoxy)ethan-1-ol (5.3)
Using the general chloroalkane mono-alkylation procedure with triethylene glycol (4.96 g, 4.51 mL, 33.1 mmol), NaH (0.633 g, 16.5 mmol), and 1-chloro-6-iodohexane (1.63 g, 1 mL, 6.61 mmol) **5.3** was afforded as a colorless oil (0.944 g, 3.51 mmol, 53.1%).
^{1}H NMR (400 MHz, DMSO-d_6) δ 3.62 (t, J = 6.6 Hz, 2H), 3.54 – 3.44 (m, 10H), 3.43 – 3.34 (m, 4H), 1.70 (p, J = 6.7 Hz, 2H), 1.49 (p, J = 6.7 Hz, 2H), 1.43 – 1.23 (m, 4H).
^{13}C NMR (100 MHz, DMSO-d_6) δ 72.3, 70.2, 69.8, 69.8, 69.8, 69.5, 60.2, 45.3, 32.0, 29.0, 26.1, 24.9.
LRMS (ESI): m/z [M+H]+ calc'd for $C_{12}H_{26}ClO_4^+$ = 269.15; found = 269.09.

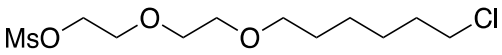


18-Chloro-3,6,9,12-tetraoxaoctadecan-1-ol (5.4)
Using the general chloroalkane mono-alkylation procedure with tetraethylene glycol (6.42 g, 5.71 mL, 33.1 mmol), NaH (0.633 g, 16.5 mmol), and 1-chloro-6-iodohexane (1.63 g, 1 mL, 6.61 mmol) **5.4** was afforded as a colorless oil (0.77 g, 2.4 mmol, 37%).
^{1}H NMR (400 MHz, DMSO-d_6) δ 3.61 (t, J = 6.6 Hz, 2H), 3.55 – 3.43 (m, 12H), 3.43 – 3.32 (m, 4H), 2.89 (s, 1H), 2.73 (s, 1H), 1.70 (p, J = 6.7 Hz, 2H), 1.48 (p, J = 6.7 Hz, 2H), 1.43 – 1.24 (m, 4H).
^{13}C NMR (100 MHz, DMSO-d_6) δ 72.3, 70.2, 69.8, 69.8, 69.8, 69.5, 60.2, 45.3, 32.0, 29.0, 26.1, 24.9.
Note: 12/14 C nuclei observed, equivalent carbons in Polythethylene Glycol (PEG) chain
LRMS (ESI): m/z [M+H]+ calc'd for $C_{14}H_{30}ClO_5^+$ = 313.18; found = 313.17.

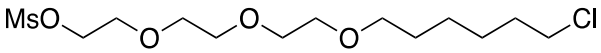


21-Chloro-3,6,9,12,15-pentaoxahenicosan-1-ol (5.5)
Using the general chloroalkane mono-alkylation procedure with pentaethylene glycol (7.88 g, 7 mL, 33.1 mmol), NaH (0.633 g, 16.5 mmol), and 1-chloro-6-iodohexane (1 mL, 6.61 mmol) **5.5** was afforded as a colorless oil (0.89 g, 2.5 mmol, 37.7%).
^{1}H NMR (400 MHz, DMSO-d_6) δ 3.62 (t, J = 6.6 Hz, 2H), 3.49 (d, J = 15.1 Hz, 18H), 3.43 – 3.32 (m, 4H), 1.70 (p, J = 6.7 Hz, 2H), 1.49 (p, J = 6.7 Hz, 2H), 1.43 – 1.25 (m, 4H).
^{13}C NMR (101 MHz, DMSO-d_6) δ 70.2, 69.6, 67.9, 45.5, 45.3, 36.7, 32.0, 28.9, 26.1, 24.8, 8.5.
Note: 11/16 C nuclei observed, equivalent carbons in PEG chain
LRMS (ESI): m/z [M+H]+ calc'd for $C_{16}H_{34}ClO_6^+$ = 357.20; found = 357.17.

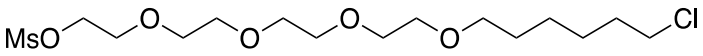
General mesylation of chloroalkanes (6n)
The select chloroalkane (1 equiv) was solubilized in DCM (0.25 M) at RT to which TEA (3 equiv) and methanesulfonyl chloride (MsCl) (1.5 equiv) were added. The reaction was stirred at RT overnight. Completion of the reactions were monitored via TLC. Upon completion, the reaction mixture was diluted with 10% (w/v) citric acid in water (100 mL total volume) and extracted with DCM (3 x 25 mL each). The combined organic layers were dried over magnesium sulfate, filtered and concentrated *in vacuo* to afford the desired mesylate as a pale yellow oil, which was used without further purification.



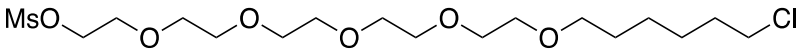
2-(2-((6-Chlorohexyl)oxy)ethoxy)ethyl methanesulfonate (6.2)
Using the general mesylation protocol with **5.2** (0.470 g, 2.09 mmol), MsCl (0.359 g, 0.243 mL, 3.14 mmol), and TEA (0.635 g, 0.875 mL, 6.27 mmol), **6.2** was afforded as a colorless oil (0.61 g, 2 mmol, 97%). ^{1}H NMR (400 MHz, DMSO-d_6) δ 4.33 – 4.27 (m, 2H), 3.70 – 3.65 (m, 2H), 3.62 (t, J = 6.6 Hz, 2H), 3.58 – 3.53 (m, 2H), 3.51 – 3.45 (m, 2H), 3.38 (t, J = 6.5 Hz, 2H), 3.17 (s, 3H), 1.71 (p, J = 6.7 Hz, 2H), 1.49 (p, J = 6.7 Hz, 2H), 1.43 – 1.25 (m, 4H).
^{13}C NMR (100 MHz, DMSO-d_6) δ 70.2, 69.7, 69.7, 69.4, 68.3, 45.3, 36.8, 32.0, 29.0, 26.1, 24.9.
LRMS (ESI): m/z [M+H]+ calc'd for $C_{11}H_{24}ClO_5S^+$ = 303.10; found = 303.05.



2-(2-(2-((6-Chlorohexyl)oxy)ethoxy)ethoxy)ethyl methanesulfonate (6-3)
Using the general mesylation protocol with **5.3** (0.944 g, 3.51 mmol), MsCl (0.603 g, 0.408 mL, 5.27 mmol), and TEA (1.07 g, 1.47 mL, 10.5 mmol), **6.3** was afforded as a colorless oil (1.22 g, 3.52 mmol, quantitative yield).
^{1}H NMR (400 MHz, DMSO-d_6) δ 4.33 – 4.27 (m, 2H), 3.69 – 3.64 (m, 2H), 3.62 (t, J = 6.6 Hz, 2H), 3.58 – 3.48 (m, 4H), 3.46 (s, 2H), 3.37 (t, J = 6.5 Hz, 2H), 3.17 (s, 3H), 3.10 – 3.01 (m, 2H), 1.70 (p, J = 6.7 Hz, 2H), 1.49 (p, J = 6.7 Hz, 2H), 1.43 – 1.24 (m, 4H).
^{13}C NMR (100 MHz, DMSO-d_6) δ 70.2, 69.8, 69.7, 69.7, 69.6, 69.5, 68.3, 45.4, 36.8, 32.0, 29.0, 26.1, 24.9.
LRMS (ESI): m/z [M+H]+ calc'd for $C_{13}H_{28}ClO_6S^+$ = 347.13; found = 347.04.



18-Chloro-3,6,9,12-tetraoxaoctadecyl methanesulfonate (6.4)
Using the general mesylation protocol with **5.4** (0.766 g, 2.5 mmol), MsCl (0.421 g, 0.284 mL, 3.67 mmol), and TEA (0.735 g, 1.02 mL, 7.35 mmol), **6.4** was afforded as a colorless oil (0.957 g, 2.5 mmol, quantitative yield).
^{1}H NMR (400 MHz, DMSO-d_6) δ 4.33 – 4.27 (m, 2H), 3.70 – 3.65 (m, 2H), 3.62 (t, J = 6.6 Hz, 2H), 3.58 – 3.52 (m, 4H), 3.51 – 3.43 (m, 6H), 3.37 (t, J = 6.5 Hz, 2H), 3.18 (s, 3H), 3.12 – 3.03 (m, 2H), 1.70 (p, J = 6.7 Hz, 2H), 1.49 (p, J = 6.7 Hz, 2H), 1.43 – 1.25 (m, 4H).
^{13}C NMR (101 MHz, DMSO-d_6) δ 70.2, 69.8, 69.7, 69.7, 69.7, 69.5, 68.3, 45.5, 45.4, 36.8, 32.0, 29.0, 26.1, 24.9.
Note: 14/15 C nuclei observed, equivalent carbons in PEG chain
LRMS (ESI): m/z [M+H]+ calc'd for $C_{15}H_{32}ClO_7S^+$ = 391.16; found = 391.10.

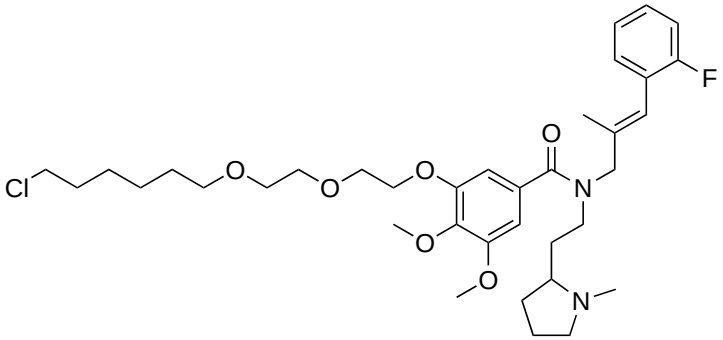


21-Chloro-3,6,9,12,15-pentaoxahenicosyl methanesulfonate (6.5)
Using the general mesylation protocol with **6.5** (0.890 g, 2.49 mmol), MsCl (0.429 g, 0.290 mL, 3.74 mmol), and TEA (0.757 g, 1.04 mL, 7.48 mmol), **6.5** was afforded as a colorless oil (1.08 g, 2.48 mmol, 99.6%).
^{1}H NMR (400 MHz, DMSO-d_6) δ 4.35 – 4.28 (m, 2H), 3.70 – 3.65 (m, 2H), 3.62 (t, J = 6.6 Hz, 2H), 3.58 – 3.51 (m, 8H), 3.51 – 3.43 (m, 6H), 3.40 – 3.33 (m, 2H), 3.17 (s, 3H), 3.13 – 3.00 (m, 2H), 1.70 (p, J = 6.7 Hz, 2H), 1.49 (p, J = 6.7 Hz, 2H), 1.43 – 1.23 (m, 4H).
^{13}C NMR (101 MHz, DMSO-d_6) δ 70.2, 69.8, 69.8, 69.8, 69.7, 69.7, 69.7, 69.5, 68.3, 45.4, 36.8, 32.0, 29.0, 26.1, 24.9.
Note: 15/17 C nuclei observed, equivalent carbons in PEG chain
LRMS (ESI): m/z [M+H]+ calc'd for $C_{17}H_{36}ClO_8S^+$ = 435.18; found = 435.10.

ENDTAC synthesis

General phenol alkylation

The desired phenol (1 equiv) and mesylate (1.5-2.5 equiv) were solubilized in DMF (0.015 M) to which K_2CO_3 (2.5-3 equiv) was added. The reaction mixture was stirred at 70 °C overnight. Upon completion, the reaction mixture was diluted in EtOAc (100 mL), washed once with saturated $NaHCO_3$ (~25 mL) and with brine (~25 mL each). The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude material was purified by prepHPLC as thick colorless oils containing TFA as observed by ^{19}F NMR.



ENDTAC-1

(E)-3-(2-(2-((6-Chlorohexyl)oxy)ethoxy)ethoxy)-N-(3-(2-fluorophenyl)-2-methylallyl)-4,5-dimethoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide

Using the general phenol alkylation protocol with **4-OMe** (0.0275 g, 0.0602 mmol), **6.2** (0.0365 g, 0.120 mmol), and K_2CO_3 (0.0250 g, 0.181 mmol) in DMF (4 mL), **ENDTAC-1** was afforded as a thick colorless oil (0.00320 mg, 0.00482 mmol, 8.01%). **ENDTAC-1** was determined as 98.3% pure UV (Retention time (R_t)= 5.5 min).

^1H NMR (700 MHz, DMSO-d_6, 100 °C) δ 7.33 (t, *J* = 7.8 Hz, 2H), 7.18 (dt, *J* = 21.7, 8.4 Hz, 2H), 6.73 (dd, *J* = 6.3, 1.9 Hz, 2H), 6.39 (s, 1H), 4.14 – 4.11 (m, 4H), 3.80 (s, 3H), 3.77 (s, 3H), 3.76 (t, *J* = 4.8 Hz, 2H), 3.62 – 3.56 (m, 4H), 3.53 – 3.48 (m, 3H), 3.47 (br, 2H), 3.40 (t, *J* = 6.3 Hz, 2H), 2.85 (s, 3H), 2.26 (br, 2H), 2.02 (br, 1H), 1.94 (br, 1H), 1.92 – 1.86 (m, 1H), 1.73 (q, *J* = 6.9 Hz, 2H), 1.70 (s, 3H), 1.67 (br, 1H) 1.50 (p, *J* = 6.8 Hz, 2H), 1.44 – 1.38 (m, 2H), 1.33 (dt, *J* = 16.0, 7.9 Hz, 2H).

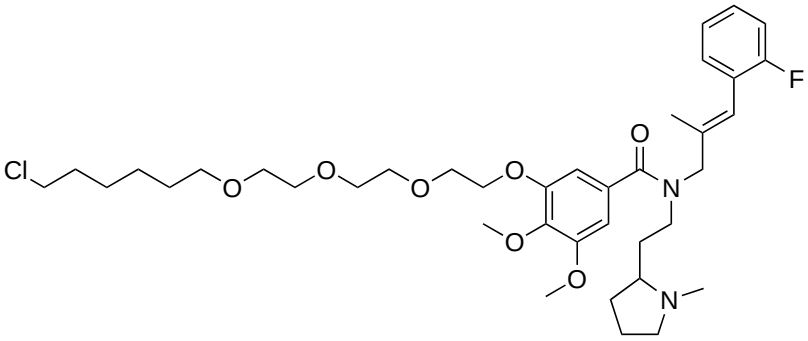
Note: 50/52 H nuclei observed, however 5 (1H) and 2 (1H), as numbered in **VUF11207**, are overlapping with the DMSO-d_6 water signal, but are observed in COSY (data not shown here).

^{13}C NMR (151 MHz, DMSO-d_6) δ 170.7, 159.5 (d, *J* = 245.0 Hz), 151.9, 138.8, 137.0, 131.1, 130.6, 129.0, 124.2, 117.0, 115.4 (d, *J* = 21.8 Hz), 105.2, 104.1, 70.2, 70.0, 69.8, 69.5, 68.9, 66.3, 60.1, 55.9, 55.3, 45.4, 42.3, 40.1, 32.0, 29.1, 29.0, 27.8, 26.1, 25.0, 21.1, 15.9.

Note: 33/36 C nuclei observed, equivalent carbons in PEG chain

^{19}F NMR (471 MHz, DMSO-d_6, 100 °C) δ -74.44 (s) (TFA), -115.16 – -115.24 (m).

HRMS (Q-tof): m/z [M+H]+ calc'd for $C_{36}H_{53}ClFN_2O_6^+$ = 663.3576; found = 663.3578.



ENDTAC-2

(E)-3-(2-(2-(2-((6-Chlorohexyl)oxy)ethoxy)ethoxy)ethoxy)-N-(3-(2-fluorophenyl)-2-methylallyl)-4,5-dimethoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide

Using the general phenol alkylation protocol with **4-OMe** (0.0275 g, 0.0602 mmol), **6.3** (0.0418 g, 0.120 mmol), and K_2CO_3 (0.0250 g, 0.181 mmol) in DMF (4 mL), **ENDTAC-2** was afforded as a thick colorless oil (0.0078 mg, 0.0110 mmol, 18.3%). **ENDTAC-2** was determined as 97.9% pure UV (R_t= 5.43 min).

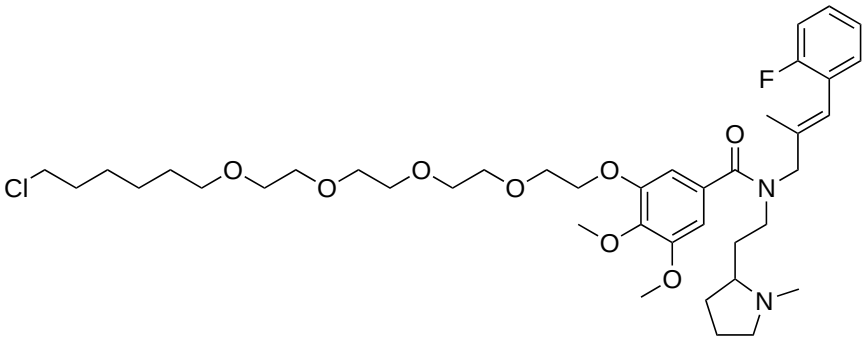
^{1}H NMR (500 MHz, DMSO-d_6, 100 °C) δ 7.36 – 7.29 (m, 2H), 7.23 – 7.15 (m, 2H), 6.73 (q, *J* = 1.8 Hz, 2H), 6.39 (s, 1H), 4.12 (dd, *J* = 5.8, 3.8 Hz, 4H), 3.80 (s, 3H), 3.77 (s, 3H), 3.76 (d, *J* = 4.8 Hz, 2H), 3.63 – 3.58 (m, 4H), 3.40 (t, *J* = 6.5 Hz, 2H), 3.28 (br, 1H), 3.09 (br, 1H), 2.86 (s, 3H), 2.26 (br, 2H), 2.02 (br, 1H), 1.94 (br, 1H), 1.92 – 1.82 (m, 1H), 1.74 (q, *J* = 7.1 Hz, 2H), 1.70 (s, 3H), 1.66 (br, 1H), 1.51 (p, *J* = 6.8 Hz, 2H), 1.41 (dq, *J* = 8.8, 6.8 Hz, 2H), 1.38 – 1.30 (m, 2H).

Note: 47/56 H nuclei observed, however, 2´ (1H), 7 (2H) nuclei, as numbered in **VUF11207**, and 6H nuclei from the PEG chain overlap with the DMSO-d_6 water signal.

^{13}C NMR (151 MHz, DMSO-d_6) δ 170.7, 159.4 (d, *J* = 244.7 Hz), 152.8, 151.9, 138.8, 136.9, 131.1, 130.6, 129.0, 124.2, 118.0, 117.0, 115.4 (d, *J* = 22.2 Hz), 115.0, 105.2, 104.1, 70.2, 70.0, 69.8, 69.8, 69.5, 68.9, 68.3, 66.3, 60.1, 55.9, 55.2, 45.4, 42.3, 40.1, 32.0, 29.0, 29.0, 27.8, 26.1, 24.9, 21.1, 15.9.

^{19}F NMR (471 MHz, DMSO-d_6, 100 °C) δ -74.51 (s) (TFA), -115.16 – -115.25 (m).

HRMS (Q-tof): m/z [M+H]+ calc'd for $C_{38}H_{57}ClFN_2O_7^+$ = 707.3838; found = 707.3854.



ENDTAC-3

(E)-3-((18-Chloro-3,6,9,12-tetraoxaoctadecyl)oxy)-N-(3-(2-fluorophenyl)-2-methylallyl)-4,5-dimethoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide

Using the general phenol alkylation protocol with **4-OMe** (0.0275 g, 0.0602 mmol), **6.4** (0.0471 g, 0.120 mmol), and K_2CO_3 (0.0250 g, 0.181 mmol) in DMF (4 mL), **ENDTAC-3** was afforded as a thick colorless oil (0.0092 mg, 0.0122 mmol, 20.3%). **ENDTAC-3** was determined as 98.3% pure UV (R_t= 5.33 min).

^{1}H NMR (500 MHz, DMSO-d_6, 100 °C) δ 7.36 – 7.28 (m, 2H), 7.23 – 7.13 (m, 2H), 6.73 (q, *J* = 1.8 Hz, 2H), 6.39 (s, 1H), 4.14 – 4.10 (m, 4H), 3.80 (s, 3H), 3.77 (s, 3H), 3.6 – 3.74 (m, 2H), 3.64 – 3.57 (m, 4H), 3.60 – 3.53 (m, 2H), 3.53 (d, *J* = 8.2 Hz, 6H), 3.50 – 3.47 (m, 3H), 3.47 (br, 2H), 3.40 (t, *J* = 6.5 Hz, 2H), 3.12 (s, 1H), 2.86 (s, 3H), 2.26 (br, 2H), 2.02 (br, 1H), 1.94 (br, 1H), 1.93 – 1.82 (m, 1H), 1.78 – 1.70 (m, 2H), 1.70 (s, 3H), 1.66 (br, 1H), 1.51 (p, *J* = 6.7 Hz, 2H), 1.42 (dq, *J* = 8.9, 6.8 Hz, 2H), 1.39 – 1.29 (m, 2H).

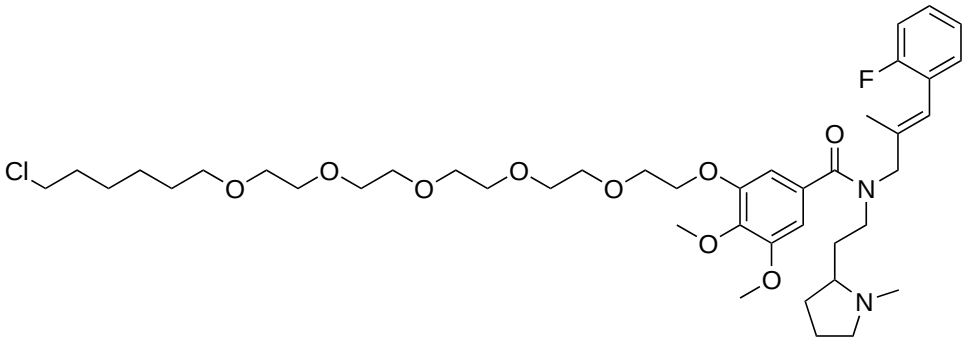
Note: 59/60 H nuclei observed, 5 (1H) nuclei, as numbered in **VUF11207**, is overlapping with the DMSO-d_6 water signal.

^{13}C NMR (151 MHz, DMSO-d_6) δ 170.7, 159.6 (d, *J* = 276.6 Hz), 151.9, 138.8, 137.0, 131.1, 130.6, 129.0, 124.2, 118.0, 117.0, 115.4 (d, *J* = 21.7 Hz), 115.1, 105.2, 104.1, 70.2, 70.0, 69.8, 69.5, 68.9, 68.3, 66.3, 60.1, 55.9, 55.8, 55.3, 45.4, 42.3, 40.1, 32.0, 29.1, 29.0, 27.8, 26.1, 25.0, 21.1, 15.9.

Note: 37/40 C nuclei observed, equivalent carbons in PEG chain

^{19}F NMR (471 MHz, DMSO-d_6, 100 °C) δ -74.53 (s) (TFA), -115.16 – -115.27 (m).

HRMS (Q-tof): m/z [M+H]+ calc'd for $C_{40}H_{61}ClFN_2O_8^+$ = 751.4100; found = 751.4117.



ENDTAC-4

(E)-3-((21-Chloro-3,6,9,12,15-pentaoxahenicosyl)oxy)-N-(3-(2-fluorophenyl)-2-methylallyl)-4,5-dimethoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide

Using the general phenol alkylation protocol with **4-OMe** (0.0275 g, 0.0602 mmol), **6.5** (0.0524 g, 0.120 mmol), and K_2CO_3 (0.0250 g, 0.181 mmol) in DMF (4 mL), **ENDTAC-4** was afforded as a thick colorless oil (0.0040 mg, 0.00503 mmol, 8.35%). **ENDTAC-4** was determined as 99.0% pure UV (R_t= 5.44 min). ^{1}H NMR (500 MHz, DMSO-d_6, 100 °C) δ 7.36 – 7.30 (m, 2H), 7.23 – 7.14 (m, 2H), 6.73 (q, *J* = 1.8 Hz, 2H), 6.39 (s, 1H), 4.12 (dd, *J* = 5.9, 3.7 Hz, 4H), 3.80 (s, 3H), 3.77 (s, 3H), 3.75 (d, *J* = 4.9 Hz, 2H), 3.64 – 3.58 (m, 2H), 3.59 – 3.52 (m, 14H), 3.51 – 3.47 (m, 3H), 3.46 (br, 2H), 3.40 (t, *J* = 6.5 Hz, 2H), 2.85 (s, 3H), 2.26 (br, 2H), 2.09 – 1.98 (m, 1H), 1.95 (br, 1H), 1.92 – 1.82 (m, 1H), 1.74 (p, *J* = 6.9 Hz, 2H), 1.70 (s, 3H), 1.66 (br, 1H), 1.51 (p, *J* = 6.7 Hz, 2H), 1.42 (dq, *J* = 8.7, 6.9 Hz, 2H), 1.38 – 1.30 (m, 2H).

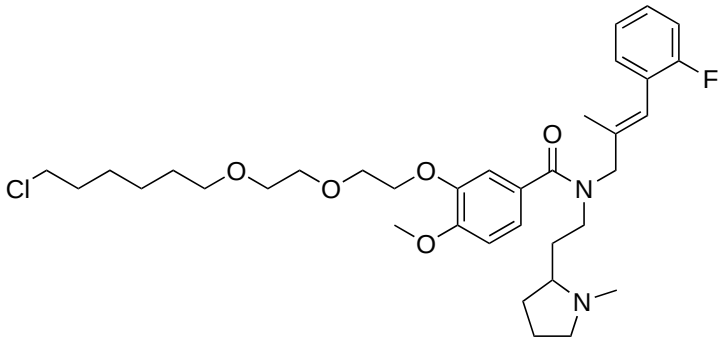
Note: 62/64 H nuclei observed, however 5 (1H) and 2 (1H) nuclei, as numbered in **VUF11207**, are overlapping with the DMSO-d_6 water signal.

^{13}C NMR (151 MHz, DMSO-d_6) δ 159.5 (d, *J* = 244.3 Hz), 151.9, 138.8, 137.0, 130.6, 129.0, 124.2, 118.0, 115.4 (d, *J* = 21.9 Hz), 115.3, 105.2, 104.1, 70.2, 70.0, 69.8, 69.8, 69.5, 68.9, 68.3, 66.3, 60.1, 55.9, 55.8, 55.2, 45.4, 42.3, 40.1, 32.0, 29.1, 29.0, 27.8, 26.1, 25.0, 21.1, 15.9.

Note: 35/42 C nuclei observed, equivalent carbons in PEG chain, C=O and two ipso-carbons not observed.

^{19}F NMR (471 MHz, DMSO-d_6, 100 °C) δ -74.37 (s) TFA, -115.17 – -115.25 (m).

HRMS (Q-tof): m/z [M+H]+ calc'd for $C_{42}H_{65}ClFN_2O_9^+$= 795.4363.; found = 795.4370.



ENDTAC-5

(E)-3-(2-(2-((6-Chlorohexyl)oxy)ethoxy)ethoxy)-N-(3-(2-fluorophenyl)-2-methylallyl)-4-methoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide

Using the general phenol alkylation protocol with **4-H** (0.0120 g, 0.0281 mmol), **6.2** (0.0170 g, 0.0561 mmol), and K_2CO_3 (0.0117 g, 0.0844 mmol) in DMF (2 mL), **ENDTAC-5** was afforded as a thick colorless oil (0.0050 mg, 0.00790 mmol, 28.1%). **ENDTAC-5** was determined as 99.9% pure UV (R_t= 5.53 min).

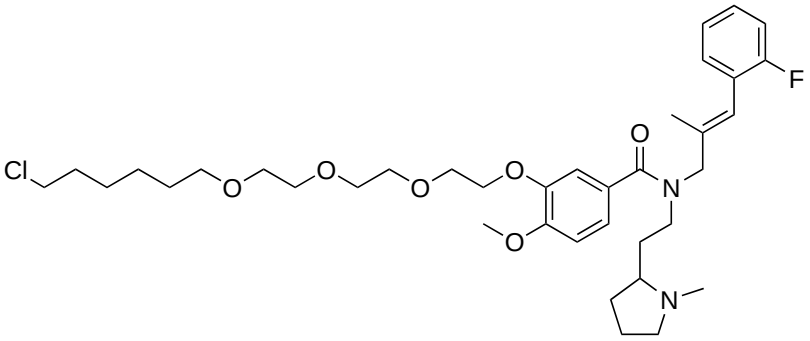
^{1}H NMR (700 MHz, DMSO-*d*$_6$, 100 °C) δ 7.32 (q, *J* = 6.6, 5.7 Hz, 2H), 7.22 – 7.14 (m, 2H), 7.07 – 7.01 (m, 3H), 6.39 (s, 1H), 4.13 (br, 2H), 4.10 (t, *J* = 5.0 Hz, 2H), 3.83 (s, 3H), 3.75 (t, *J* = 5.0 Hz, 2H), 3.59 (q, *J* = 7.0, 5.8 Hz, 4H), 3.50 (t, *J* = 5.1 Hz, 3H), 3.49 (br, 2H), 3.41 (t, *J* = 6.5 Hz, 2H), 3.09 (br, 1H), 2.86 (s, 3H), 2.26 (br, 2H), 2.03 (br, 1H), 1.95 (br, 1H), 1.92 – 1.84 (m, 1H), 1.73 (p, *J* = 6.9 Hz, 2H), 1.68 (s, 4H), 1.51 (p, *J* = 6.8 Hz, 2H), 1.41 (p, *J* = 7.4 Hz, 2H), 1.34 (p, *J* = 7.5 Hz, 2H).

Note: 49/50 H nuclei observed, however 5 (1H) nuclei, as numbered in **VUF11207**, is overlapping with the DMSO-d$_6$ water signal.

^{13}C NMR (151 MHz, DMSO-*d*$_6$) δ 170.8, 159.4 (d, *J* = 244.7 Hz), 150.0, 147.4, 136.9, 130.6, 129.0, 124.2, 119.8, 117.9, 117.1, 115.4 (d, *J* = 22.1 Hz), 115.1, 111.5, 111.5, 70.2, 69.9, 69.5, 68.9, 67.9, 66.2, 55.6, 55.5, 55.2, 45.4, 42.1, 40.1, 32.0, 29.1, 29.0, 27.8, 26.1, 24.9, 21.0, 15.8.

^{19}F NMR (471 MHz, DMSO-*d*$_6$, 100 °C) δ -74.44 (s) (TFA), -115.13 – -115.21 (m).

HRMS (Q-tof): m/z [M+H]+ calc'd for $C_{35}H_{51}ClFN_2O_5^+$ = 633.3471; found = 633.3488.



ENDTAC-6

(E)-3-(2-(2-(2-((6-Chlorohexyl)oxy)ethoxy)ethoxy)ethoxy)-N-(3-(2-fluorophenyl)-2-methylallyl)-4-methoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide

Using the general phenol alkylation protocol with **4-H** (0.0120 g, 0.0281 mmol), **6.3** (0.0250 g, 0.0721 mmol), and K_2CO_3 (0.0117 g, 0.0844 mmol) in DMF (2 mL), **ENDTAC-6** was afforded as a thick colorless oil (0.0074 mg, 0.0109 mmol, 38.8%). **ENDTAC-6** was determined as above 99.9% pure UV (R_t= 5.53 min).

^{1}H NMR (500 MHz, DMSO-*d*$_6$, 100 °C) δ 7.32 (tt, *J* = 6.8, 2.9 Hz, 2H), 7.23 – 7.14 (m, 2H), 7.04 (t, *J* = 2.8 Hz, 3H), 6.39 (s, 1H), 4.13 (s, 2H), 4.12 – 4.07 (m, 2H), 3.83 (d, *J* = 2.5 Hz, 3H), 3.75 (dq, *J* = 5.3, 3.0 Hz, 3H), 3.60 (dd, *J* = 5.9, 3.0 Hz, 5H), 3.55 (ddt, *J* = 12.5, 5.0, 3.2 Hz, 4H), 3.48 (dt, *J* = 6.9, 3.6 Hz, 4H), 3.40 (td, *J* = 6.6, 2.9 Hz, 2H), 3.10 (br, 1H), 2.86 (s, 3H), 2.25 (br, 2H), 2.02 (br, 1H), 1.94 (br, 1H), 1.90 – 1.83 (m, 1H), 1.73 (td, *J* = 7.2, 2.5 Hz, 2H), 1.68 (s, 3H), 1.65 (br, 1H), 1.51 (pd, *J* = 6.8, 2.7 Hz, 2H), 1.41 (t, *J* = 7.4 Hz, 2H), 1.34 (dp, *J* = 8.6, 3.1, 2.6 Hz, 2H).

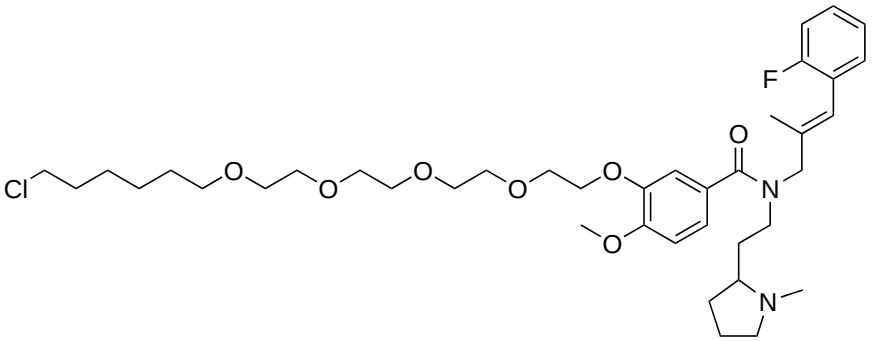
Note: 53/54 H nuclei observed, however 5 (1H) nuclei, as numbered in **VUF11207**, is overlapping with the DMSO-*d*$_6$ water signal.

^{13}C NMR (151 MHz, DMSO-d$_6$) δ 159.4 (d, *J* = 244.6 Hz), 150.0, 147.4, 136.7, 130.6, 129.0, 124.2, 119.9, 118.0, 117.2, 115.5, 115.3, 115.2, 111.5 (d, *J* = 8.6 Hz), 70.2, 69.9, 69.9 – 69.7 (m), 69.5, 68.9, 67.9, 66.3, 55.6, 55.5, 55.2, 45.4, 41.9, 40.1, 32.0, 29.1, 28.9, 27.7, 26.1, 24.9, 21.0, 15.8.

Note: 35/37 C nuclei observed, equivalent carbons in PEG chain, and C=O carbon not observed.

^{19}F NMR (471 MHz, DMSO-*d*$_6$) δ -74.40, -115.13 – -115.21 (m).

HRMS (Q-tof): m/z [M+H]+ calc'd for $C_{37}H_{55}ClFN_2O_6^+$ = 677.3733; found = 677.3745.



ENDTAC-7

(E)-3-((18-Chloro-3,6,9,12-tetraoxaoctadecyl)oxy)-N-(3-(2-fluorophenyl)-2-methylallyl)-4-methoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide

Using the general phenol alkylation protocol with **4-H** (0.0120 g, 0.0281 mmol), **6.4** (0.022 g, 0.0563 mmol), and K_2CO_3 (0.0117 g, 0.0844 mmol) in DMF (2 mL), **ENDTAC-7** was afforded as a thick colorless oil (0.0056 mg, 0.00776 mmol, 27.6%). **ENDTAC-7** was determined as 99% pure UV (R_t= 5.42 min).

^{1}H NMR (500 MHz, DMSO-*d*$_6$, 100 °C) δ 7.33 (tq, *J* = 7.3, 4.3, 3.4 Hz, 2H), 7.18 (dt, *J* = 14.3, 8.6 Hz, 2H), 7.04 (q, *J* = 2.7, 2.3 Hz, 3H), 6.39 (s, 1H), 4.13 (s, 2H), 4.11 – 4.07 (m, 2H), 3.83 (d, *J* = 2.2 Hz, 3H), 3.75 (td, *J* = 5.2, 2.3 Hz, 2H), 3.65 – 3.57 (m, 4H), 3.57 – 3.51 (m, 5H), 3.48 (dt, *J* = 6.2, 3.3 Hz, 2H), 3.28 (br, 1H), 3.09 (br, 1H), 2.85 (s, 3H), 2.25 (br, 2H), 2.02 (br, 1H), 1.94 (br, 1H), 1.91 – 1.81 (m, 1H), 1.77 – 1.70 (m, 2H), 1.68 (s, 3H), 1.65 (br, 1H), 1.56 – 1.48 (m, 2H), 1.41 (q, *J* = 7.6 Hz, 2H), 1.38 – 1.30 (m, 2H).

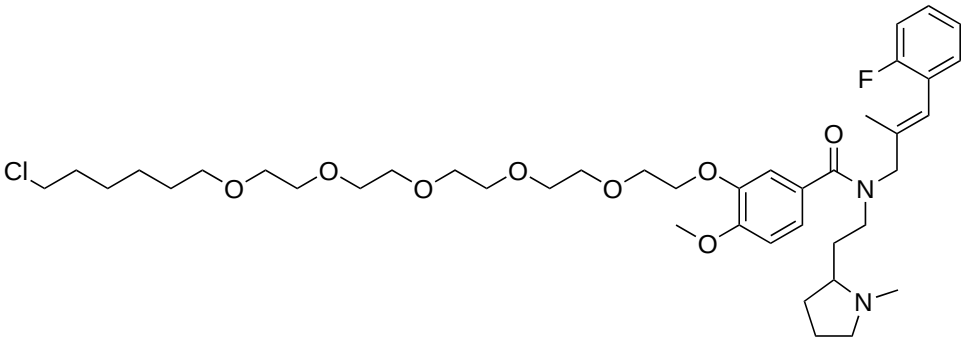
Note: 50/58 H nuclei observed, however 2' (1H) and 7 (2H), as numbered in **VUF11207**, and 6H nuclei from the PEG chain are overlapping with the DMSO-*d*$_6$ water signal.

^{13}C NMR (151 MHz, DMSO-d$_6$) δ 159.4 (d, *J* = 244.7 Hz), 150.1, 147.4, 136.9, 130.6, 129.0, 124.2, 119.9 (d, *J* = 8.7 Hz), 118.1, 115.4 (d, *J* = 22.1 Hz), 115.2, 111.5, 111.5, 70.2, 69.9, 69.8, 69.5, 68.9, 67.9, 66.3, 55.8, 55.6, 55.2, 45.4, 42.1, 40.1, 32.0, 29.1, 28.8, 27.8, 26.1, 24.9, 21.0, 15.8.

Note: 34/39 C nuclei observed, equivalent carbons in PEG chain and C=O carbon not observed.

^{19}F NMR (471 MHz, DMSO-*d*$_6$, 100 °C) δ -74.39, -115.12 – -115.23 (m).

HRMS (Q-tof): m/z [M+H]+ calc'd for $C_{39}H_{59}ClFN_2O_7^+$ = 721.3995; found = 721.3997.



ENDTAC-8

(E)-3-((21-Chloro-3,6,9,12,15-pentaoxahenicosyl)oxy)-N-(3-(2-fluorophenyl)-2-methylallyl)-4-methoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide

Using the general phenol alkylation protocol with **4-H** (0.0120 g, 0.0281 mmol), **6.5** (0.0306 g, 0.0703 mmol), and K_2CO_3 (0.0117 g, 0.0844 mmol) in DMF (2 mL), **ENDTAC-8** was afforded as a thick colorless oil (0.0071 mg, 0.0093 mmol, 33.0%). **ENDTAC-8** was determined as 98.9% pure UV (R_t= 5.41 min).

^{1}H NMR (500 MHz, DMSO-d_6, 100 °C) δ 7.36 – 7.29 (m, 2H), 7.24 – 7.14 (m, 2H), 7.04 (dd, *J* = 3.7, 1.9 Hz, 3H), 6.39 (d, *J* = 2.1 Hz, 1H), 4.13 (s, 2H), 4.10 (dd, *J* = 5.6, 4.2 Hz, 2H), 3.83 (br, 3H), 3.75 (dd, *J* = 5.6, 4.2 Hz, 2H), 3.63 – 3.58 (m, 4H), 3.54 (d, *J* = 2.2 Hz, 9H), 3.41 (t, *J* = 6.5 Hz, 2H), 3.28 (br, 1H), 3.09 (br, 1H), 2.86 (s, 3H), 2.26 (br, 2H), 2.02 (br, 1H), 1.94 (s, 1H), 1.91 – 1.82 (m, 1H), 1.77 – 1.70 (m, 2H), 1.68 (s, 3H), 1.65 (s, 1H), 1.51 (p, *J* = 6.8 Hz, 2H), 1.42 (dq, *J* = 8.5, 6.8 Hz, 2H), 1.38 – 1.31 (m, 2H).

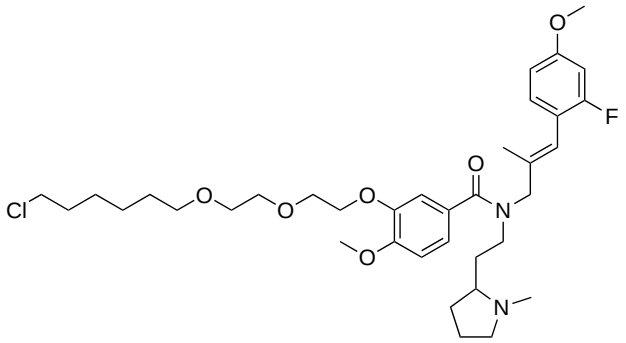
Note: 54/62 H nuclei observed, however 2' (1H) and 7 (2H), as numbered in **VUF11207**, and 6H nuclei from the PEG chain are overlapping with the DMSO-d_6 water signal.

^{13}C NMR (151 MHz, DMSO-d_6) δ 171.0, 159.4 (d, *J* = 245.0 Hz), 150.1, 147.5, 136.9, 130.6, 129.0, 128.1, 124.2, 119.8, 118.0, 115.4 (d, *J* = 21.6 Hz), 115.1, 111.5, 111.5, 70.2, 69.9, 69.8, 69.5, 68.9, 67.9, 66.3, 55.6, 55.4, 55.2, 45.4, 42.1, 40.1, 32.0, 29.1, 28.9, 27.8, 26.1, 24.9, 21.0, 15.8.

Note: 36/41 C nuclei observed, equivalent carbons in PEG chain.

^{19}F NMR (471 MHz, DMSO-d_6, 100 °C) δ -74.48 (d, *J* = 2.5 Hz), -115.13 – -115.22 (m).

HRMS (Q-tof): m/z [M+H]+ calc'd for $C_{41}H_{63}ClFN_2O_8^+$ = 765.4257; found = 765.4257.



ENDTAC-neg
(*E*)-3-(2-(2-((6-Chlorohexyl)oxy)ethoxy)ethoxy)-N-(3-(2-fluoro-4-methoxyphenyl)-2-methylallyl)-4-methoxy-N-(2-(1-methylpyrrolidin-2-yl)ethyl)benzamide

Using the general phenol alkylation protocol with **9** (0.0375 g, 0.0821 mmol), **6.2** (0.0373 g, 0.123 mmol), and K_2CO_3 (0.0284 g, 0.205 mmol) in DMF (2 mL), **ENDTAC-neg** was afforded as a thick colorless oil (0.0032 mg, 0.00482 mmol, 5.87 %). **ENDTAC-neg** was determined as 99.4% pure UV (R_t= 5.46 min).

^{1}H NMR (700 MHz, DMSO-*d*$_6$, 100 °C) δ 7.23 (t, *J* = 8.8 Hz, 1H), 7.03 (d, *J* = 4.4 Hz, 3H), 6.82 – 6.75 (m, 2H), 6.30 (s, 1H), 4.10 (br, 4H), 3.83 (s, 3H), 3.80 (s, 3H), 3.75 (t, *J* = 4.9 Hz, 2H), 3.61 – 3.56 (m, 4H), 3.52 – 3.49 (m, 3H), 3.47 (br, 2H), 3.41 (t, *J* = 6.5 Hz, 2H), 3.27 (br, 1H), 2.85 (s, 3H), 2.24 (br, 2H), 2.02 (br, 1H), 1.94 (br, 1H), 1.87 (br, 1H), 1.73 (p, *J* = 6.9 Hz, 2H), 1.66 (s, 4H), 1.51 (p, *J* = 6.9 Hz, 2H), 1.41 (q, *J* = 7.5 Hz, 2H), 1.34 (q, *J* = 7.5 Hz, 2H).

Note: 51/52 H nuclei observed, however 2 (1H), as numbered in **VUF11207**, is overlapping with the DMSO-d$_6$ water signal.

^{13}C NMR (151 MHz, DMSO-*d*$_6$) δ 171.0, 160.9, 159.7, 159.3, 158.9 (d, *J* = 220.5 Hz), 150.0, 147.3, 135.3, 131.0, 120.0, 117.7, 116.4, 111.4 (d, *J* = 18.7 Hz), 110.2, 101.6, 101.4, 70.2, 70.0, 69.5, 68.9, 67.9, 66.3, 55.7, 55.6, 55.2, 45.5, 42.1, 40.1, 32.1, 29.1, 29.0, 27.8, 26.2, 25.0, 21.1, 15.9.

^{19}F NMR (471 MHz, DMSO-d$_6$, 100 °C) δ -74.00 (d, *J* = 3.6 Hz) (TFA), -112.83 – -112.90 (m).

HRMS (Q-tof): m/z [M+H]+ calc'd for $C_{36}H_{53}ClFN_2O_6^+$ = 663.3576; found = 663.3591.

Methods: Biology

Antibodies and reagents

Antibodies against HA (CST-3724) was purchased from Cell Signaling Technology, Alexa Fluor 488 conjugated tubulin antibody (16-232) was purchased from Millipore. EEA1 antibody (14-9114-82), Hoechst 33342 stain (H3570), Lysotracker™ Red DND-99 (L7528), Super Signal West Femto substrate (34095), Zeba spin desalting coulmns (7K) (89882) and Alexa Fluor 488 transferrin conjugate (T13342) from Thermo fisher scientific. CXCR7 antibody was purchased from Abcam (ab138509). Bright Glo™ luciferase substrate (E2610) and Nano-Glo® Luciferase Assay system (N1130) was obtained from Promega. Tissue culture treated white 96 well plates (CLS3610) and poly lysine solution (P8920) were purchased from Sigma. µ-slide 8 well chambered coverslips (80826) were obtained from Ibidi. CXCR7 Tango plasmid (66265) was purchased from Addgene. Bafilomycin A1 from Alfa Aesar.

Cell culture

HTLA (a HEK293 cell line stably expressing a tTA-dependent luciferase reporter and a β-arrestin2-TEV fusion gene) cells were generously provided by Gilad Barnea (Brown University, RI). MCF7 cells were cultured in RPMI-1640 medium supplemented with 10% FBS, Penicillin (100 U/mL) and Streptomycin (100 µg/mL), and MIA PaCa-2 cells were cultured in DMEM+10% FBS, 2.5% Horse serum, Penicillin (100 U/mL) and Streptomycin (100 µg/mL). HTLA cells were cultured in DMEM supplemented with 10% FBS, Penicillin (100 U/mL) and Streptomycin (100 µg/mL), Hygromycin (200 µg/mL), Puromycin (5 µg/mL) and G418 (0.5 mg/mL). 293T or MDA-MB-231 cells were cultured in DMEM medium supplemented with 10% FBS, Penicillin (100 U/mL) and Streptomycin (100 µg/mL). All cells were grown in a humidified atmosphere at 37 °C in 5% CO_2.

Cloning

HA-eGFP-Halotag7 was cloned into a pET28a bacterial expression plasmid. Nanoluc sequence was PCR amplified from pNL1.1 plasmid (Promega) and was cloned in between HA and Halotag7 (excluding eGFP) in the pET28a plasmid using USER cloning (NEB). Prior to protein expression, all constructs were fully sequence-verified by Sanger sequencing. Full sequence information for HA-eGFP-HaloTag7 and HA-Nanoluc-HaloTag7 constructs is provided below.

Protein Expression and purification

Large-scale expression and purification of HA-eGFP-HaloTag7 was performed in BL21-CodonPlus (DE3)-RIPL *E. coli* cells. A 10 mL starter culture was grown overnight at 37 °C before dilution into 2 L (1:200) LB broth. The culture was allowed to grow at 37 °C until OD_{600} reaches 0.7-0.9. Then the culture was induced with 0.5 mM IPTG and moved to 25 °C to express protein for 4 h. After 4 h, the cells were harvested by centrifugation and washed once with cold PBS. Cells were then resuspended in cold lysis buffer (150 mM NaCl, 50 mM Tris-HCl (pH = 7.5), 5% glycerol, 10 mM imidazole supplemented with Roche protease inhibitor cocktail (EDTA-free) and lysed by sonication. Cell lysate was clarified by ultracentrifugation at 16000 rpm for 45 min in a Sorvall SS34 rotor. Then, the cleared lysate was incubated with 5 mL of Ni-NTA agarose for 2 h at 4 °C. After capture of His- tagged HT7 onto Ni-NTA resin (Qiagen), the resin was washed with 10 column volumes of lysis buffer supplemented with increasing concentrations of imidazole (20 mM and 35 mM imidazole). Following washes, protein was eluted using lysis buffer supplemented with 150 mM imidazole. Eluted fractions were collected and injected to MonoQ5/50GL anion exchange chromatography using an Äktaexplorer FPLC system (GE Healthcare Life Sciences). Protein was eluted by running a gradient over 40 CV (from 100% Buffer A [50 mM Tris-HCl [pH = 7.5], 50 mM NaCl, 5% glyercol] to 50 % Buffer B (50 mM Tris-HCl [pH = 7.5], 1 M NaCl, 5% glycerol). Fractions were analyzed by SDS-PAGE and purified fractions were collected,

concentrated (~1 mL) and then subjected to HiLoad 16/60 Superdex 200 column using buffer containing 150 mM NaCl, 50 mM Tris-HCl (pH = 7.5), 5% glycerol. Protein fractions were analyzed for purity by Coomassie staining, concentrated, flash frozen, aliquoted and stored at -80 °C.

Purification of Nanoluc-HaloTag7 was conducted in a similar fashion as that described above for eGFP-HaloTag7 with several exceptions. MonoQ anion exchange chromatography was performed using the same buffers and gradient as that used for eGFP-HaloTag7 (see above). Size exclusion chromatography was performed similarly in a different equilibration buffer. Size exclusion buffer consisted of 25 mM Tris-HCl (pH = 7.5), 100 mM NaCl. Protein fractions were analyzed for purity by Coomassie staining, concentrated and stored at -20 °C with 50% Glycerol to avoid freeze-thawing between multiple experiments.

Reaction of ENDTAC 1-8 and ENDTAC-neg with HT7

The eGFP-HT7 protein (40 µM) was reacted with ENDTACs in a 1:2 molar ratio in PBS for 1 h at RT. Then, the reaction mixture was subjected to Zeba 0.5 mL spin columns (7K MWCO) to remove excess ENDTAC and used the reacted protein for subsequent experiments. Similarly, Nanoluc-HT7 (50 µM) was reacted with ENDTACs in a 1:2 molar ratio and subjected to Zeba 0.5 mL spin columns (7K MWCO) to remove excess ENDTAC and used the reacted protein for subsequent experiments. To confirm the completion of reaction of HT7 with ENDTAC 1-4, LC-MS analysis was performed. 1 µM of protein (filtered) was subjected to LC-MS analysis and unreacted HT7 was used as a reference standard. Protein mass spectrometry was acquired on Waters Acquity^{TM} Ultra Performance LC with C4 column, PDA λR detector and Waters Zspray^{TM} electrospray ionization (ESI) mass detector. MaxEnt software was used for protein mass deconvolution.

Tango GPCR assay

HTLA cells were cultured in DMEM supplemented with 10% FBS, Penicillin (100 U/mL) and Streptomycin (100 µg/mL), Hygromycin (200 µg/mL), Puromycin (5 µg/mL) and G418 (0.5 mg/mL). For transfection of CXCR7 Tango plasmid, cells were plated at 1.5X10^6 cells in a 6 well dish (day1). The following day (day 2), cells were transfected using the PEI reagent. Next day (day 3) transfected cells were plated at 15000 cells per well in 100 µL medium into a poly-lysine coated white, clear bottomed 96 well cell culture plates. Non-transfected cells were included in one lane to use as a negative control. On day 4, cells were serum starved for 1 h, prior to adding the increasing concentrations of ENDTACs. One lane of transfected cells was used as a DMSO control. After 24 h incubation, medium containing ENDTACs were removed and 20 µL of Bright Glo substrate (Promega) was added to each well. The plates were incubated at RT for 15 min and luminescence was measured using Wallac Victor 2 Plate Reader (Perkin Elmer). Relative luminescence units were exported into an Excel sheet and data were analyzed using GraphPad prism.

For the pulse chase Tango assay to detect saturation kinetics of HT7 reacted with ENDTAC, cells were prepared as described above for a dose response experiment. On day 4, cells were serum starved for 1 h, prior to adding DMSO, VUF-11207, ENDTAC-1 and ENDTAC-1:eGFP-HT7 (4 µM) to cells. Cells were incubated with compounds for 1 h, 2 h, 4 h, 6 h, 8 h and 24 h. At each time point, compounds were removed, washed with serum free media twice and maintained in serum free media for up to 24 h to get the luminescence readings as described above.

Analysis of internalization by Confocal Microscopy

For microscopy experiments, MCF7 cells were seeded on 8 well chambered poly-D-lysine coated coverslips (Ibidi) and grown until 50-60% confluent. Next day, cells were incubated with serum free RPMI media supplemented with 1% BSA for 1 h at 37 °C and then kept at 4 °C for 30 min. eGFP-HT7 protein pre-reacted with either ENDTAC-1 or ENDTAC-neg (10 µM) was added to cells in RPMI with BSA and incubated at 37 °C for 4 h. After 4 h, cells were cooled on ice, washed twice with complete RPMI media and PBS before fixing with 4% PFA for 15 min at RT. To analyze the cellular uptake of HT7, cells were washed 2X with PBS after fixing and stained with Hoechst stain (3 µg/mL) for 10 min at RT. Cells were washed with PBS and analyzed by Zeiss confocal microscope. For transferrin uptake experiments, 25 µg/mL TFN488 was added to serum free medium and incubated for 30 min before fixing. To track the lysosomal accumulation, LysoTracker Red DND99 (300 nM) was added to cells during the last 45 min of 4 h incubation time and fixed. For immunofluorescence experiments with EEA1, cells were blocked for 1 h at RT in PBS, 3% BSA and 0.2% Triton X-100. Then, the cells were incubated with primary EEA1 antibody (1.5 µg/mL) in PBS, 1% BSA and 0.2% Triton-X-100 overnight at 4 °C. Next day, after washing 3X with PBS, secondary antibody conjugated to Alexa Fluor 546 (1:500 dilution) in PBS, 1% BSA and 0.2% Triton-X-100 for 1 h at RT. Images were acquired on a Zeiss LSM 880. Z-stacks (1-µm slices) spanning the entire volume of the cells were recorded with oil-immersion 40× and 63× lenses. Images were processed using ImageJ.

Analysis of HT7 uptake by Nanoluc-based plate assay

For Nanoluc luciferase assays, MCF7 cells were plated at 12000 cells per well in 100 µL medium into a poly-lysine coated white, clear bottomed 96 well cell culture plates. Following day, cells were blocked with RPMI containing 1% BSA for 1 h, cooled at 4 °C for 30 min prior to starting the uptake experiment. Nanoluc-HT7 pre-reacted with ENDTAC-neg or ENDTAC-1 was added to cells (final 1 µM) and incubated for 2 h at 37 °C. After the pulse, cells were washed twice with complete RPMI media and incubated at 37 °C for different time points. At the end of each time point, cells were washed once with PBS, acid washed to reduce non-specific binding followed by another PBS wash. Then, RPMI media was added to each well followed by the addition of reconstituted NanoGlo® substrate (1:1 media to NanoGlo® substrate ratio), incubated 3 min at RT and luminescence was measured using Wallac Plate Reader (Perkin Elmer). Reactions were performed in six replicates, and the data were normalized to inactive compound control (ENDTAC-neg:Nanoluc-HT7) and presented as the Fold uptake.

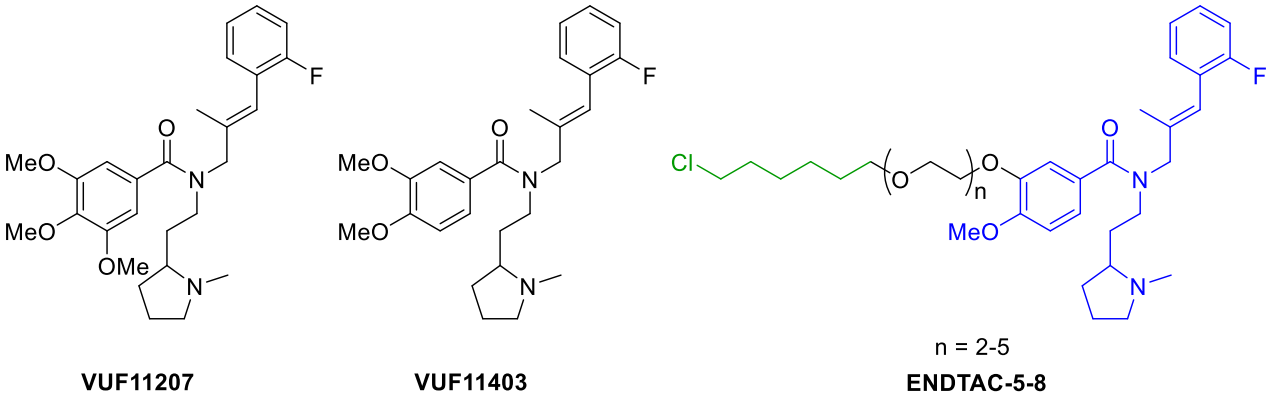
Analysis of HT7 uptake and degradation by Western blotting

To probe the cellular uptake of HT7 using Western blotting, MCF7 cells or MIA PaCa-2 cells were chosen as they both express endogenous CXCR7 receptor. MCF7 or MIA PaCa-2 cells were plated at $1.4X10^6$ cells per well in a 6 well dish and allowed to grow up to 70-80% confluency. Next day, cells were pre-blocked with complete media supplemented with 1% BSA for 1 h at 37 °C prior to the addition of pre-reacted ENDTAC-1:eGFP-HT7 or ENDTAC-neg:eGFP-HT7 (500 nM). After 4 h incubation, cells were washed twice with complete media and PBS. Cells were acid washed (0.2M glycine, 0.15M NaCl, pH 2.5) to remove surface-bound protein, followed by another PBS wash. Next, the cells were harvested by trypsinization. Cell were lysed and protein concentration was determined by BCA assay. Proteins were separated by SDS-PAGE, transferred to PVDF membrane, and blocked with 5% (w/v) non-fat dry milk in TBST for 1 h at RT. The membrane was incubated with primary antibody against HA-tag (CST 3724, 1:1000) overnight at 4 °C. Following three washes, secondary anti-rabbit HRP antibody was incubated for 1 h at RT. The membrane was developed using Super Signal West Femto substrate (ThermoFisher). For the competition experiments with warhead VUF11207, MCF7 cells were incubated with pre-reacted DMSO:eGFP-HT7, ENDTAC-neg:eGFP-HT7 or ENDTAC-1:eGFP-HT7 (250 nM) for 4 h in the absence or presence of 100X excess VUF11207. After 4 h incubation, cells were

washed and harvested by trypsinization. To compare the effect of CXCR7 expression on eGFP-HT7 uptake, 293T cells (non-transfected or CXCR7-RFP transfected) were incubated with pre-reacted DMSO:eGFP-HT7, ENDTAC-neg:eGFP-HT7 or ENDTAC-1:eGFP-HT7 (250 nM) for 4 h before harvesting. MDA-MB-231 or MCF7 cells were incubated with pre-reacted ENDTAC-neg:eGFP-HT7 or ENDTAC-1:eGFP-HT7 (250 nM) for 4 h before harvesting.

For the pulse chase experiment, the experiment was performed and harvested as described above with some exceptions. After 4 h incubation, cells were washed twice with complete medium and incubated for another 3 h at 37 °C in the absence or presence of lysosome inhibitor-Bafilomycin A1 (100 nM) before harvesting.

Safety Statement: no unexpected or unusually high safety hazards were encountered.



VUF11207 **VUF11403** **ENDTAC-5-8**

Figure S1: Structures of the two active agonists **- VUF11207, VUF11403** and **ENDTAC-5-8.**

Protein sequence of HA-eGFP-HT7 (575 aa)

MGSSHHHHHHSSGLVPRGSHMASYPYDVPDYAGVSKGEELFTGVVPILVELDGDVNGHKFSVSGE
GEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIF
FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIR
HNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDE
LYKARDTNSAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYVWRNIIPHVAPTH
RCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHDWGSALGFHWAKRNPERVKGI
AFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRKLIIDQNVFIEGTLPMGVVRPLTEVEMDHYREPFL
NPVDREPLWRFPNELPIAGEPANIVALVEEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKSLPN
CKAVDIGPGLNLLQEDNPDLIGSEIARWLSTLESRGPV*

Protein sequence of HA-nluc-HT7 (507 aa)

MGSSHHHHHHSSGLVPRGSHMASYPYDVPDYAGVFTLEDFVGDWRQTAGYNLDQVLEQGGVSSL
FQNLGVSVTPIQRIVLSGENGLKIDIHVIIPYEGLSGDQMGQIEKIFKVVYPVDDHHFKVILHYGTLVIDG
VTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSLLFRVTINGVTGWRLCERILAAR
DTNSAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYVWRNIIPHVAPTHRCIAP
DLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHDWGSALGFHWAKRNPERVKGIAFMEFI
RPIPTWDEWPEFARETFQAFRTTDVGRKLIIDQNVFIEGTLPMGVVRPLTEVEMDHYREPFLNPVDR

EPLWRFPNELPIAGEPANIVALVEEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDI
GPGLNLLQEDNPDLIGSEIARWLSTLESRGPV*

Figure S2: Protein sequence of eGFP-HT7 (top) and Nanoluc HT7 (bottom).

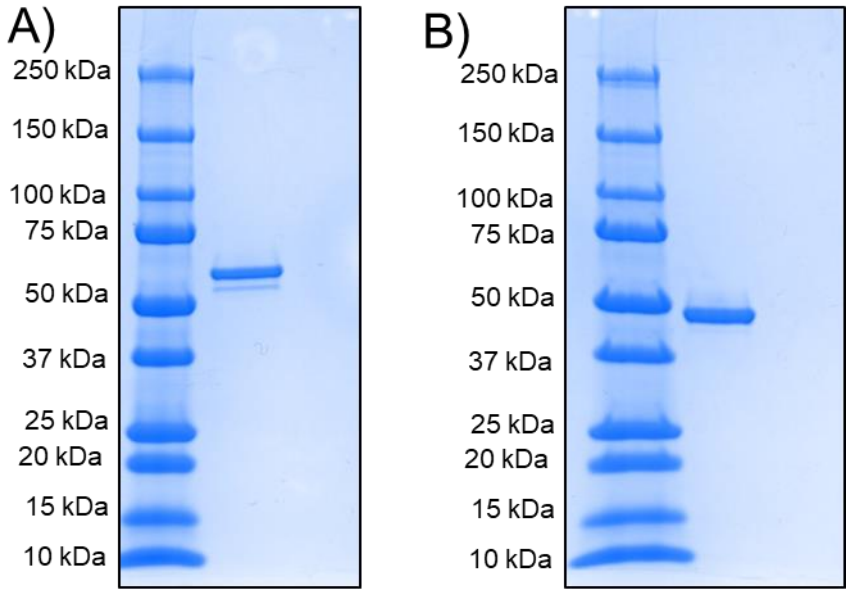


Figure S3: SDS-PAGE analysis of A) purified HA-eGFP-HT7 (62 kDa) and B) HA-Nanoluc-HT7 (50 kDa).

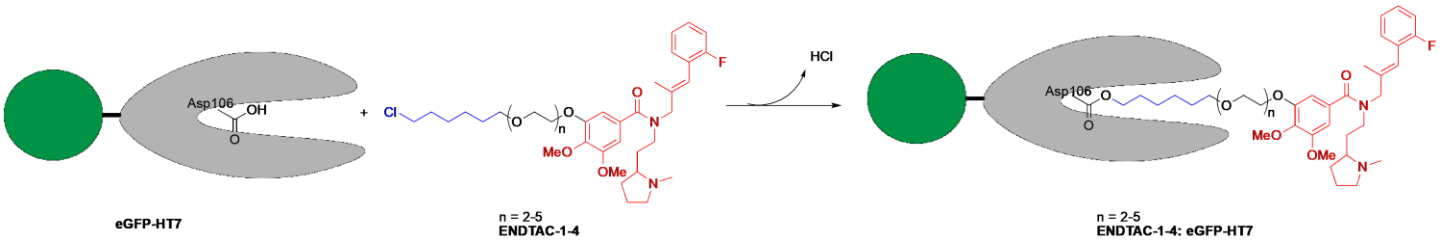
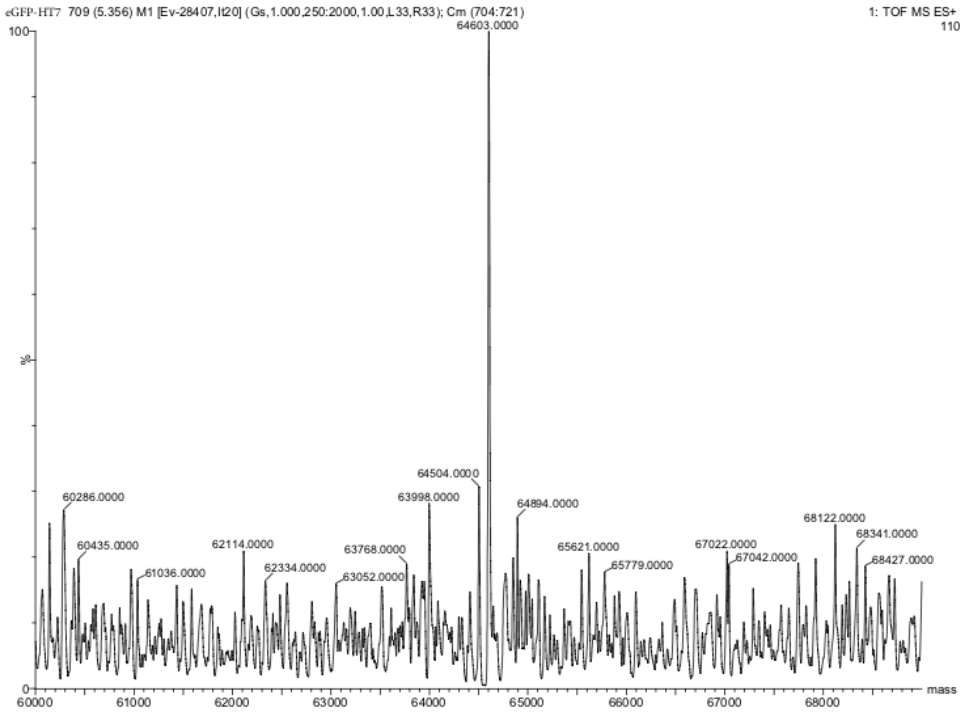


Figure S4- Reaction between **ENDTAC-1-4** with eGFP-HT7

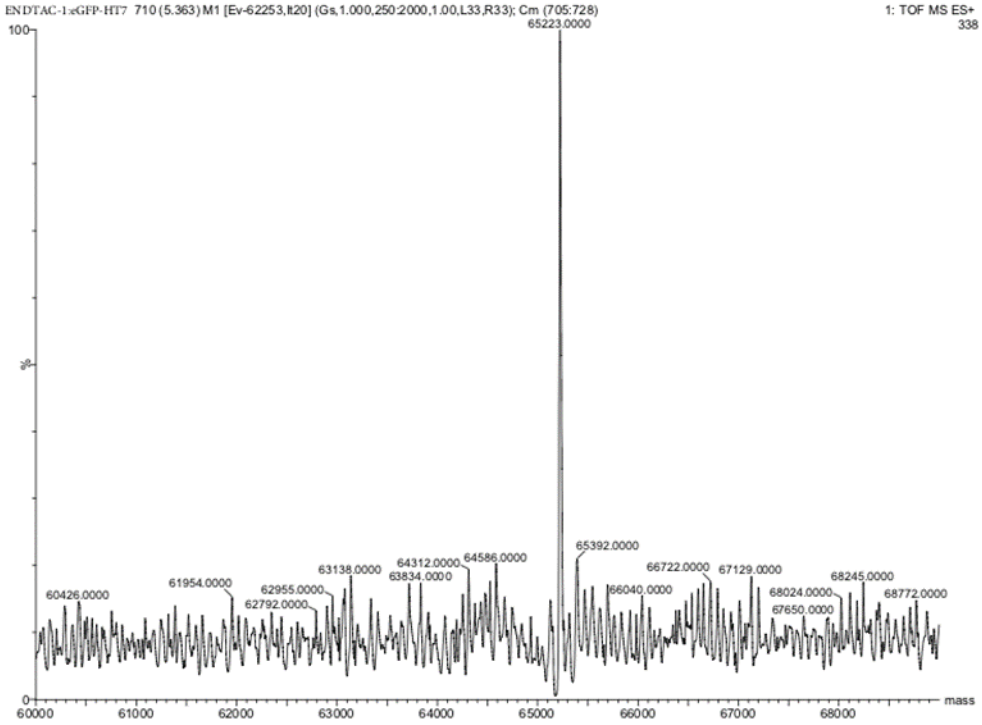
Table S1- Expected and observed masses of LC-MS analysis for **ENDTAC-1-4** with eGFP-HT7

Entry	Compound	ENDTAC mass	Expected mass	Observed mass
1	HA-eGFP-HT7	-	64748	64603
2	**ENDTAC-1**: eGFP-HT7	662.35	65229	65223
3	**ENDTAC-2**: eGFP-HT7	706.38	65273	65267
4	**ENDTAC-3**: eGFP-HT7	750.40	62317	62313
5	**ENDTAC-4**: eGFP-HT7	794.43	65361	65356

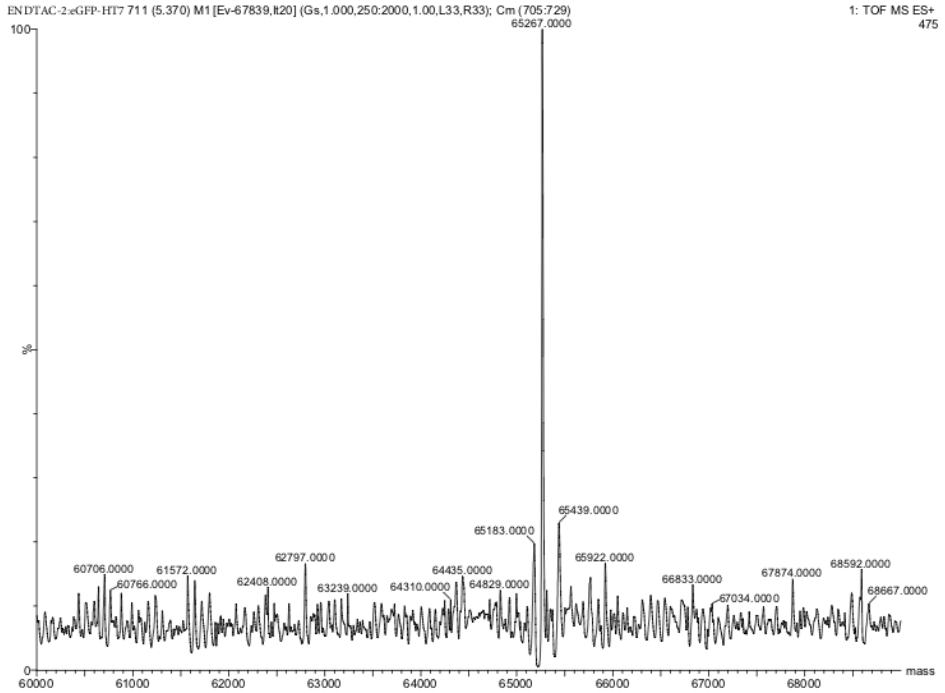
A)


eGFP-HT7 709 (5.356) M1 [Ev-28407,It20] (Gs,1.000,250:2000,1.00,L33,R33); Cm (704:721)
1: TOF MS ES+
110
100
%
0
64603.0000
64504.0000
63998.0000
64894.0000
60286.0000
60435.0000
62114.0000
63768.0000
65621.0000
67022.0000
68122.0000
68341.0000
61036.0000
62334.0000
63052.0000
65779.0000
67042.0000
68427.0000
60000
61000
62000
63000
64000
65000
66000
67000
68000
mass

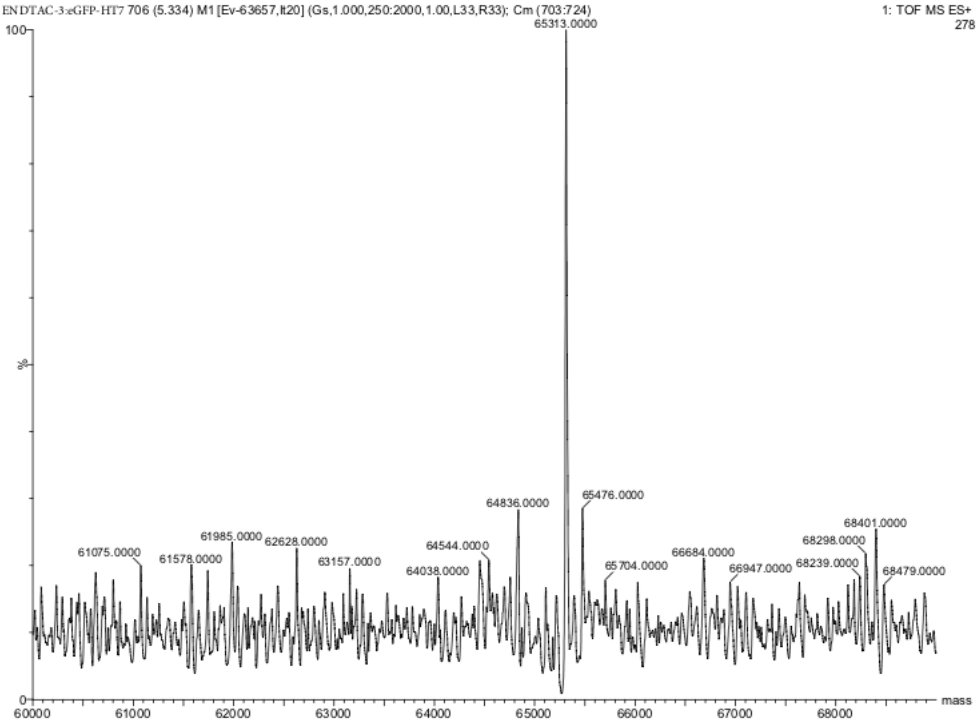
B)


ENDTAC-1-eGFP-HT7 710 (5.363) M1 [Ev-62253,It20] (Gs,1.000,250:2000,1.00,L33,R33); Cm (705:728)
1: TOF MS ES+
338
100
%
0
65223.0000
65392.0000
63138.0000
64312.0000
64586.0000
63834.0000
66722.0000
67129.0000
61954.0000
62955.0000
62792.0000
66040.0000
68245.0000
68024.0000
68772.0000
60426.0000
67650.0000
60000
61000
62000
63000
64000
65000
66000
67000
68000
mass

C)


ENDTAC-2-eGFP-HT7 711 (5.370) M1 [Ev-67839,lt20] (Gs,1.000,250:2000,1.00,L33,R33); Cm (705:729)
1: TOF MS ES+
475
65267.0000
65439.0000
65183.0000
65922.0000
62797.0000
60706.0000
60766.0000
61572.0000
62408.0000
63239.0000
64310.0000
64435.0000
64829.0000
66833.0000
67034.0000
67874.0000
68592.0000
68667.0000
100
%
0
60000
61000
62000
63000
64000
65000
66000
67000
68000
mass

D)


ENDTAC-3-eGFP-HT7 706 (5.334) M1 [Ev-63657,lt20] (Gs,1.000,250:2000,1.00,L33,R33); Cm (703:724)
1: TOF MS ES+
278
65313.0000
65476.0000
64836.0000
68401.0000
68298.0000
61985.0000
62628.0000
61075.0000
61578.0000
63157.0000
64544.0000
64038.0000
65704.0000
66684.0000
66947.0000
68239.0000
68479.0000
100
%
0
60000
61000
62000
63000
64000
65000
66000
67000
68000
mass

E)

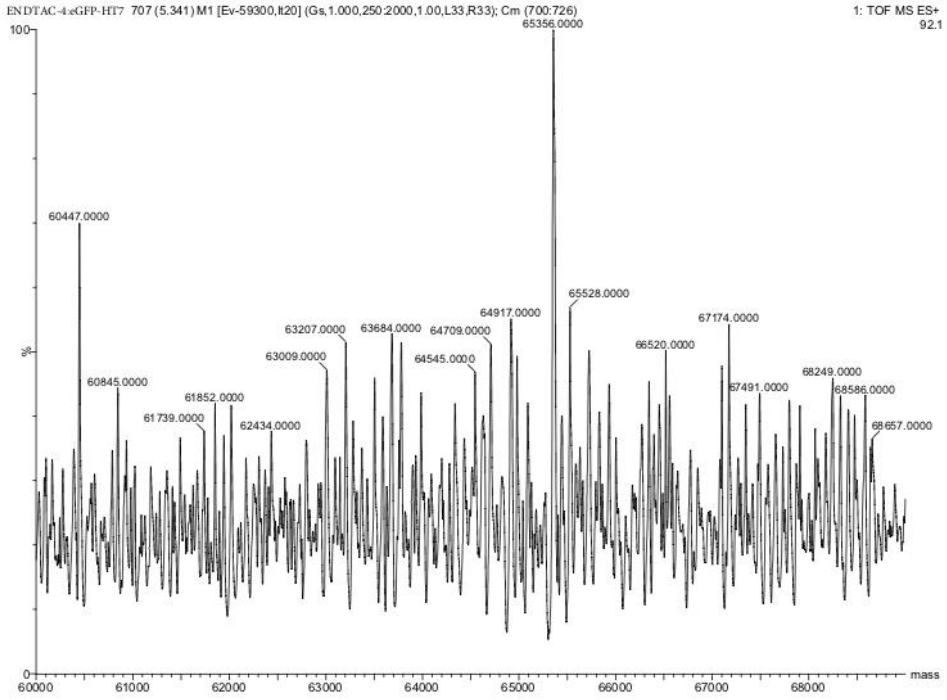


Figure S5: MS trace of eGFP-HT7 and eGFP-HT7 covalently linked to ENDTAC1-4, respectively. A) eGFP-HT7 only B) **ENDTAC-1**: eGFP-HT7 C) **ENDTAC-2**: eGFP-HT7 D) **ENDTAC-3**: eGFP-HT7 and E) **ENDTAC-4**: eGFP-HT7.

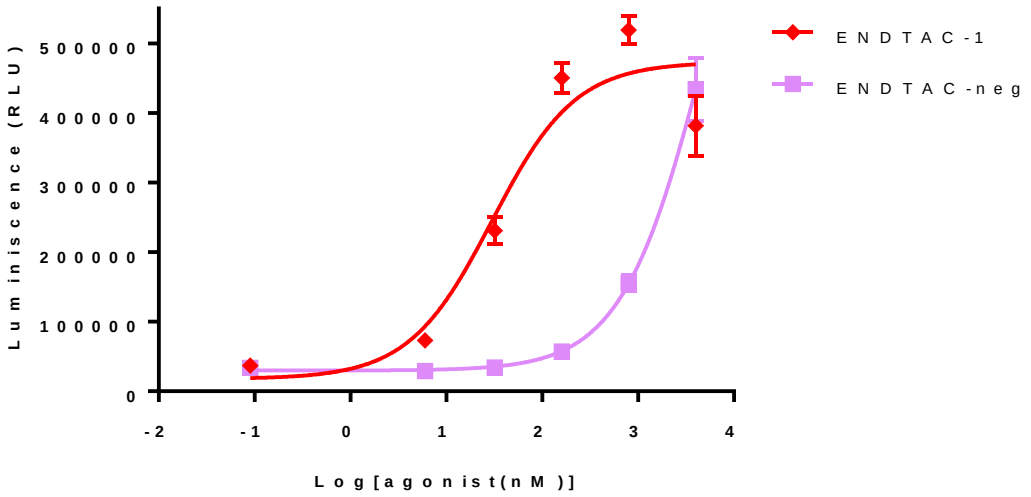


Figure S6: Characterization of agonist activity of **ENDTAC-1** and **ENDTAC-neg** using Tango assay (n=3). The data represent mean±SEM.

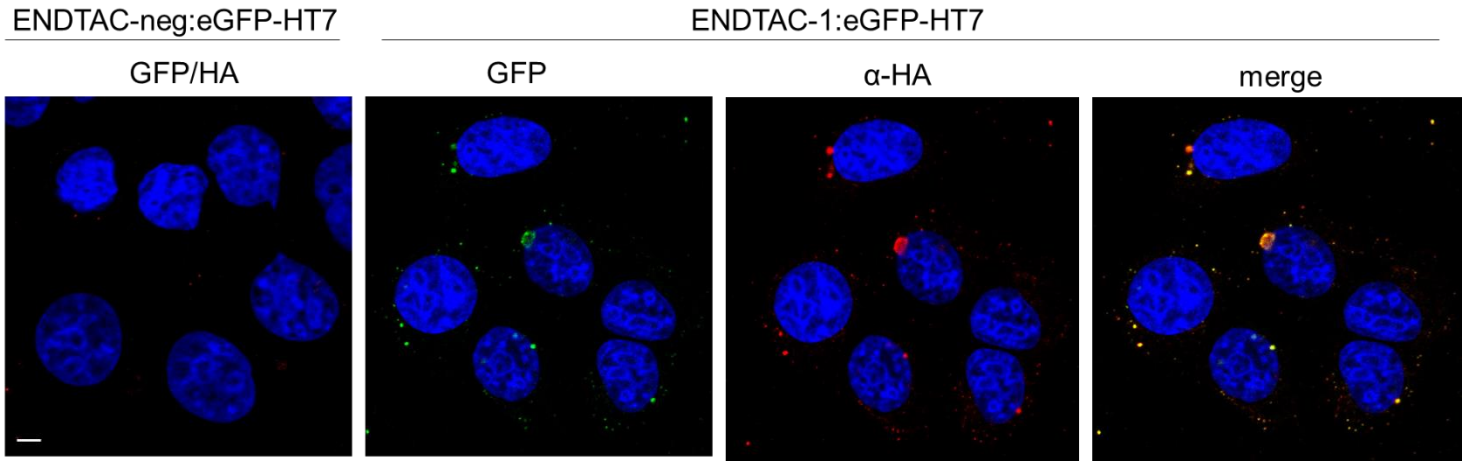


Figure S7: Internalization of eGFP-HT7 in the presence of ENDTAC-1. A) Confocal microscopy analysis of internalized eGFP-HT7 (n=3) (10 µM) (green puncta) or immunostaining with anti-HA antibody (red), nuclei are stained with Hoechst stain (blue). Scale bar: 5 µm.

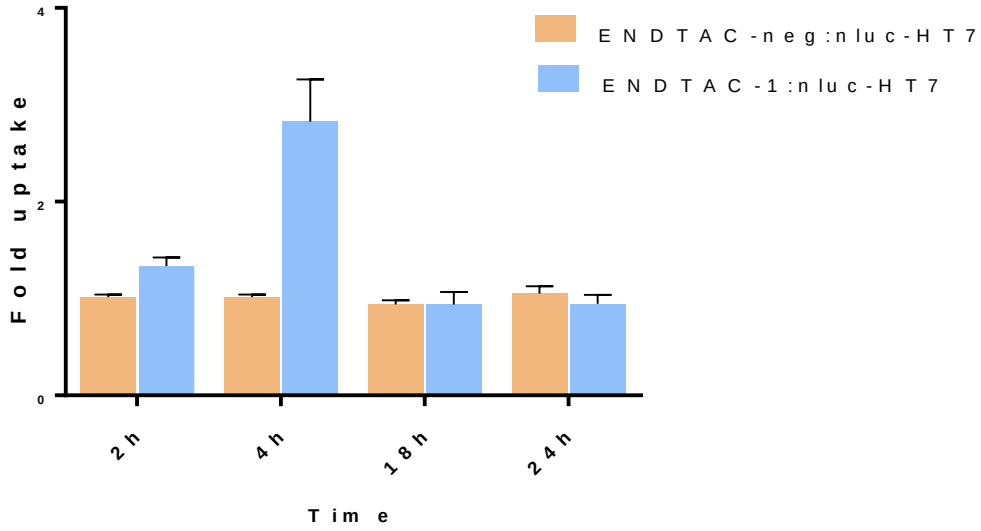


Figure S8: **Cellular uptake of Nanoluc-HT7 (nluc-HT7) in HTLA cells analyzed by luciferase assay system.** The relative luminescence units of **ENDTAC-1**:Nanoluc-HT7 was normalized to **ENDTAC-neg**:Nanoluc-HT7 and presented as the fold uptake using GraphPad Prism 5 (n=4). The data represent mean±SEM.

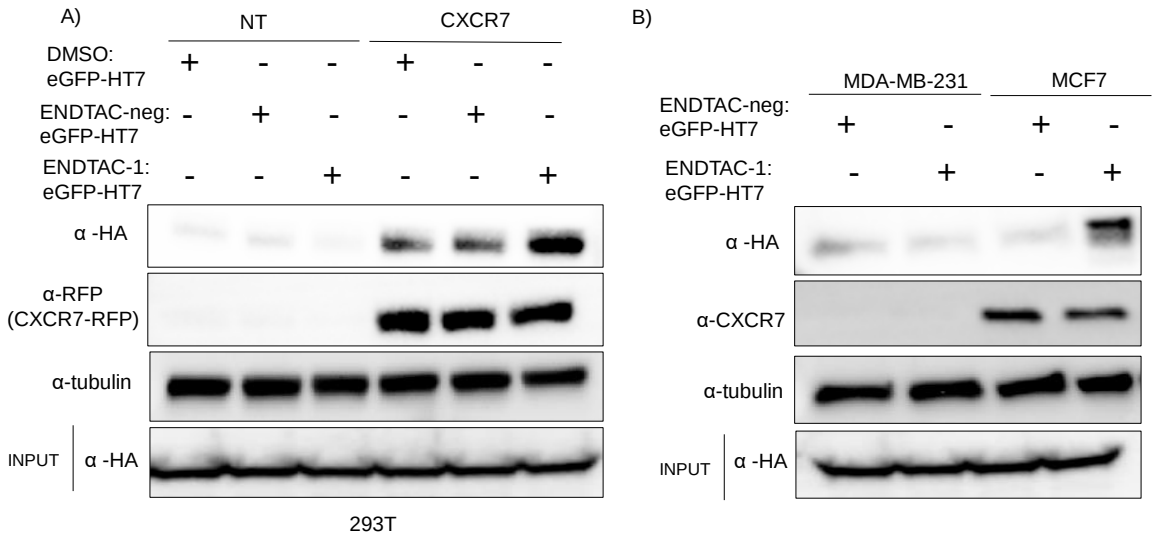


Figure S9- Expression of CXCR7 regulates cellular uptake of eGFP-HT7. A) B) Cellular uptake of eGFP-HT7 in MDA-MB-231 (CXCR7 negative) and MCF7 (CXCR7 positive) cells was analyzed by immunoblotting after incubating for 4 h with **ENDTAC-neg**:eGFP-HT7 or **ENDTAC-1**:eGFP-HT7 (250 nM).

References

1. Buckley, D. L.; Raina, K.; Darricarrere, N.; Hines, J.; Gustafson, J. L.; Smith, I. E.; Miah, A. H.; Harling, J. D.; Crews, C. M., HaloPROTACS: Use of Small Molecule PROTACs to Induce Degradation of HaloTag Fusion Proteins. *ACS Chemical Biology* **2015,** *10* (8), 1831-1837.

2. Wijtmans, M.; Maussang, D.; Sirci, F.; Scholten, D. J.; Canals, M.; Mujic-Delic, A.; Chong, M.; Chatalic, K. L.; Custers, H.; Janssen, E.; de Graaf, C.; Smit, M. J.; de Esch, I. J.; Leurs, R., Synthesis, modeling and functional activity of substituted styrene-amides as small-molecule CXCR7 agonists. *European journal of medicinal chemistry* **2012,** *51*, 184-92.