Supporting Information

Synthesis of an Artificial Cell Surface Receptor that Enables Oligohistidine Affinity Tags to Function as Metal-Dependent Cell-Penetrating Peptides

Siwarutt Boonyarattanakalin, Sonalee Athavankar, Qi Sun and Blake R. Peterson*

Department of Chemistry, The Pennsylvania State University, University Park, PA 16802

Supporting Figures

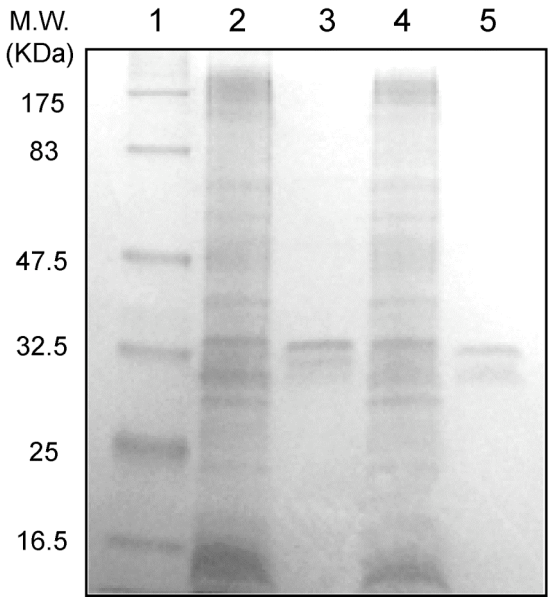


Figure S1. Analysis of oligohistidine-tagged AcGFP proteins by SDS PAGE. Lane 1: Protein molecular weight marker. Lane 2: Crude lysate from cells expressing AcGFP(His)$_{10}$. Lane 3: Purified AcGFP(His)$_{10}$. Lane 4: Crude lysate from cells expressing AcGFP(His)$_6$. Lane 5: Purified AcGFP(His)$_6$.

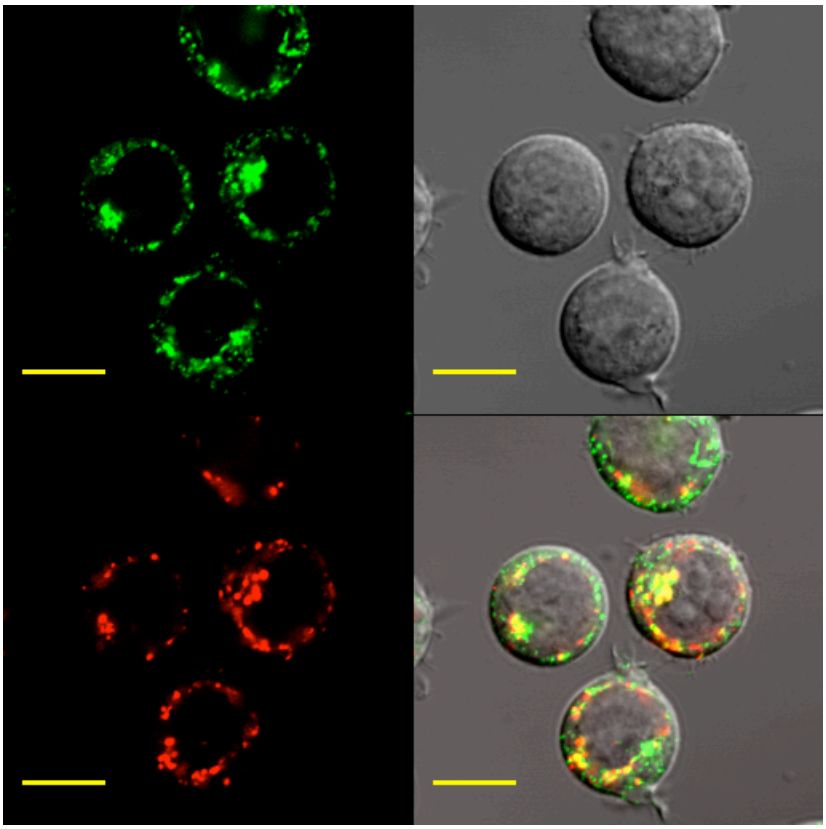


Figure S2. Delivery of oligohistidine-tagged AcGFP to endosomes and lysosomes. Jurkat lymphocytes were treated with synthetic receptor **1** (10 μM) at 37 °C for 1 h, the cells were washed with media, and $Ni(OAc)_2$ (100 μM), green fluorescent $AcGFP(His)_{10}$ (3.2 μM), and red fluorescent DiI-loaded LDL (50 μg / mL; this corresponds to 20 nM based on LDL[1] M.W. = 2.5 x 10^6 Da) were added at 37 °C for an additional 4 h. Cells were imaged by confocal laser scanning (left) and DIC microscopy (right). Green fluorescence is shown in the upper left panel. Red fluorescence is shown in the lower left panel. Colocalization of green and red fluorescence is shown in yellow overlaid on the DIC image in the lower right panel. Scale bar = 10 μm.

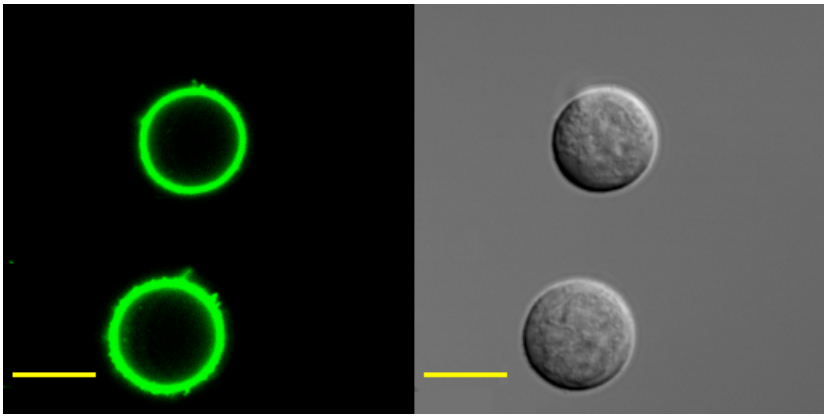
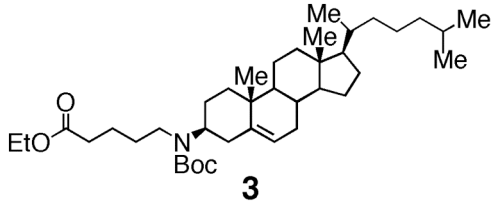


Figure S3. Inhibition of protein uptake at 4 °C. Jurkat lymphocytes were treated with synthetic receptor **1** (10 μM) at 37 °C for 1 h, washed to remove unincorporated receptors, the cells were cooled to 4 °C, and $AcGFP(His)_{10}$ (19 μM) and $Ni(OAc)_2$ (100 μM) were added at 4 °C for 4 h. Cells were imaged by confocal laser scanning (left) and DIC microscopy (right). Scale bar = 10 μm.

Experimental section

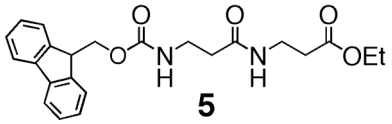
General. Chemical reagents were obtained from Acros, Aldrich, Alfa Aesar, or TCI America. Solvents were from EM Science. Media and antibiotics were purchased from Mediatech. DiI-loaded human low-density lipoprotein was from Invitrogen. Commercial grade reagents were used without further purification unless otherwise noted. Anhydrous solvents were obtained after passage through a drying column of a solvent purification system from GlassContour (Laguna Beach, CA). All reactions were performed under an atmosphere of dry argon or nitrogen. Reactions were monitored by analytical thin-layer chromatography on plates coated with 0.25 mm silica gel 60 F_{254} (EM Science). TLC plates were visualized by UV irradiation (254 nm) or stained with a solution of phosphomolybdic acid and sulfuric acid in ethanol (1:1:20). Flash column chromatography employed ICN SiliTech Silica Gel (32-63 μm). Purification by preparative reverse phase HPLC employed an Agilent 1100 preparative pump / gradient extension instrument equipped with a Hamilton PRP-1 (polystyrene-divinylbenzene) reverse phase column (7 μm particle size, 21.5 mm x 25 cm). The HPLC flow rate was increased from 10 mL/min (t = 0 min) to 20 mL/min (t = 2 min) and maintained at 20 mL/min for the remainder of the run unless otherwise noted. Melting points were measured with a Thomas Hoover capillary melting point apparatus and are uncorrected. Infrared spectra were obtained with a Perkin Elmer 1600 Series FTIR. NMR spectra were obtained with Bruker CDPX-300, DPX-300, AMX-360, or DRX-400 instruments with chemical shifts reported in parts per million (ppm, δ) referenced to either $CDCl_3$ (^{1}H 7.27 ppm; ^{13}C 77.23 ppm), DMSO-d_6 (^{1}H 2.50 ppm; ^{13}C 39.51 ppm), or $(CH_3)_4Si$. High-resolution mass spectra were obtained from the University of Texas at Austin and Penn State University Mass Spectrometry Facilities (ESI and CI). Peaks are reported as *m/z*.

Synthetic procedures and compound characterization data



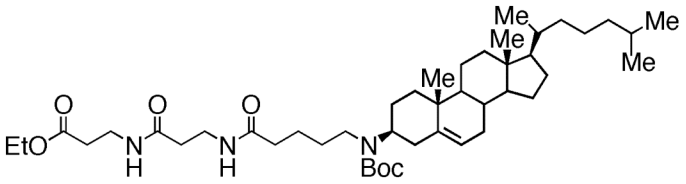
3

Ethyl 5-{(tert-butoxycarbonyl)[(3β)-cholest-5-en-3-yl]amino}pentanoate (3). To DMF (10 mL) was added 3β-amino-5-cholestene (**2**, 386 mg, 1.0 mmol),[2] ethyl 5-bromovalerate (174 μL, 1.1 mmol) and K_2CO_3 (276 mg, 2.0 mmol). The solution was heated to 60 °C and stirred for 24 h. The reaction was cooled to 22 °C and DMF was removed *in vacuo*. To the resulting solid residue was added CH_2Cl_2 (10 mL), insoluble salts were removed by filtration, and solids were washed with additional CH_2Cl_2 (5 mL). To this solution containing the crude secondary amine product was added $(Boc)_2O$ (327 mg, 1.5 mmol) and DIEA (0.5 mL, 3.0 mmol). The reaction was stirred for 3 h at 22 °C and concentrated *in vacuo*. Flash column chromatography (hexanes / ethyl acetate, 10:1) afforded **3** (417 mg, 68%) as a white solid, mp 78-79 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.30 (d, 1H), 4.10 (q, *J* = 7.1 Hz, 2H), 3.73 (br s, 1H), 3.09 (br s, 2H), 2.30 (t, *J* = 7.1 Hz, 2H), 2.01-0.83 (m, 56H), 0.66 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.6, 155.5, 141.3, 121.4, 79.3, 60.4, 56.9, 56.3 (× 2), 50.3, 42.6 (× 2), 39.9, 39.7, 38.6, 37.1, 36.8, 36.3, 35.9, 34.2, 32.0 (× 2), 28.7 (Boc Me_3), 28.4, 28.1 (× 2), 26.9, 24.4, 24.0, 23.0, 22.7, 22.6, 21.2, 19.6, 18.9, 14.4, 12.0; IR (film) ν max 2935, 2867, 1737, 1692, 1466, 1409, 1365, 1247, 1173 cm^{-1}; CI *m/z* 614.5146 (MH^+, $C_{39}H_{68}NO_4$ requires 614.5148).



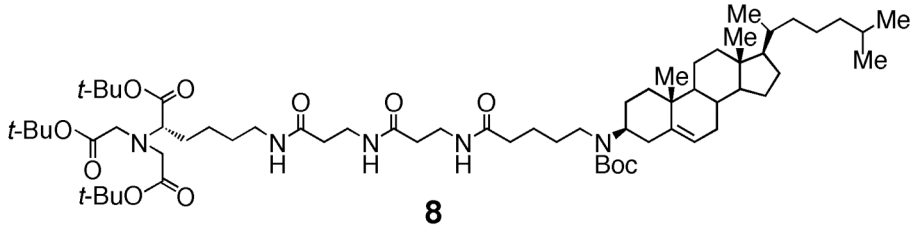
5

***N*-Fmoc-β-alanine-β-alanine ethyl ester (5).** *N*-Fmoc-β-alanine (311 mg, 1.0 mmol) in anhydrous CH_2Cl_2 (20 mL) under N_2 was cooled to 4 °C. HOBt (168 mg, 1.1 mmol) and EDC (230 mg, 1.2 mmol) were added and the solution was stirred for 30 min at 4 °C. β-Alanine ethyl ester hydrochloride (169 mg, 1.1 mmol) in anhydrous CH_2Cl_2 (10 mL) and DIEA (180 μL, 1.1 mmol) were added. The reaction was warmed to 22 °C and stirred for 16 h. The reaction solution was diluted with CH_2Cl_2 (30 mL) and washed with aqueous HCl (5%, 30 mL), followed by aqueous NaOH (0.1 M, 30 mL), and deionized H_2O (30 mL). The organic layer was dried over anhydrous Na_2SO_4 and concentrated *in vacuo*. Flash column chromatography (hexanes / ethyl acetate, 1:2) afforded **5** (394 mg, 96%) as a white solid, mp 158-159 °C; ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.74 (d, *J* = 7.5 Hz, 2H), 7.58 (d, *J* = 7.5 Hz, 2H), 7.38 (m, *J* = 7.5 Hz, 2H), 7.29 (m, *J* = 7.5 Hz, 2H), 6.53 (br m, 1H), 5.77 (br m, 1H), 4.34 (d, *J* = 7.1 Hz, 2H), 4.18 (t, *J* = 7.1 Hz, 1H), 4.11 (q, *J* = 7.1 Hz, 2H), 3.50 (m, 4H), 2.52 (t, *J* = 6.0 Hz, 2H), 2.40 (t, *J* = 5.7 Hz, 2H), 1.23 (t, *J* = 7.1 Hz, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 172.6, 171.6, 156.7, 144.0, 141.3, 127.8, 127.1, 125.2, 120.1, 66.8, 60.9, 47.3, 37.2, 36.0, 35.0, 34.1, 14.3; IR (film) ν max 3319, 3069, 2978, 2955, 2884, 1732, 1689, 1634, 1538, 1446, 1270, 1190 cm^{-1}; CI 411.1922 m/z (MH^+,$C_{23}H_{27}N_2O_5$ requires 411.1920).



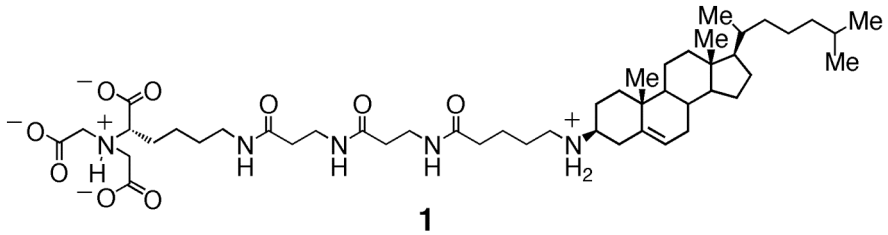
7

Ethyl 5-[(3β)-cholest-5-en-3-yl]-2,2-dimethyl-4,10,14-trioxo-3-oxa-5,11,15-triazaoctadecan-18-oate (7). Aqueous LiOH (10 mL, 0.5 M) was added dropwise to a solution of **3** (315 mg, 0.51 mmol) in a mixture of MeOH (15 mL) and THF (10 mL). The solution was stirred for 4 h at 22 °C and the organic solvents were removed *in vacuo*. The remaining aqueous solution was acidified with aqueous HCl (10%) and the resulting carboxylic acid precipitated as a white solid. This solid was collected by vacuum filtration, washed with cold water, and dried *in vacuo*. The dried solid was dissolved in anhydrous CH_2Cl_2 (20 ml) under dry N_2 and cooled to 4 °C. HOBt (92 mg, 0.6 mmol) and EDC (125 mg, 0.65 mmol) were added and the solution was stirred at 4 °C for 30 min. To this solution was added β-alanine-β-alanine ethyl ester (**6,** 141 mg, 0.75 mmol, prepared by treatment of **5** (310 mg) in DMF (8 mL) with piperidine (2 mL) for 10 min, followed by removal of solvent under high vacuum) in anhydrous CH_2Cl_2 (15 mL) over 15 min. The reaction was allowed to warmed to 22 °C and stirred for an additional 16 h. The solution was diluted with CH_2Cl_2 (50 mL) and washed with aqueous NaOH (0.1 M, 40 mL) followed by deionized H_2O (40 mL). The organic layer was dried over anhydrous Na_2SO_4 and concentrated *in vacuo*. Flash column chromatography (CH_2Cl_2 / MeOH, 30:1) afforded **7** (330 mg, 86%) as colorless oil; ^{1}H NMR (300 MHz, $CDCl_3$) δ 6.62 (br m, 2H), 5.29 (d, *J* = 4.9 Hz 1H), 4.11 (q, *J* = 7.1 Hz, 2H), 3.48 (m, 4H), 3.33 (br, 1H), 3.07 (br s, 2H), 2.49 (t, *J* = 6.0 Hz, 2H), 2.35 (t, *J* = 5.4 Hz, 2H), 2.16 (br, 2H), 1.99-0.75 (m, 56H), 0.64 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 173.1, 172.6, 171.8, 155.6, 141.5, 121.4, 79.4, 60.9, 56.8, 56.3 (× 2), 50.2, 42.4 (× 2), 39.9, 39.6 (× 2), 38.5 (× 2), 37.2, 36.8, 36.3, 35.9, 35.6, 35.2, 34.2, 32.0 (× 2), 28.7 (Boc Me_3), 28.4, 28.1 (× 2), 26.9, 24.4, 23.9, 23.2, 22.9, 22.7, 21.2, 19.6, 18.8, 14.3, 12.0; IR (film) ν max 3293, 3082, 2868, 1737, 1693, 1644, 1549, 1466, 1412, 1366, 1251, 1175, 1156 cm^{-1}; CI *m/z* 756.5891 (MH^+, $C_{45}H_{78}N_3O_6$ requires 756.5891).



8

Di-tert-butyl 2-(2-tert-butoxy-2-oxoethyl)-22-[(3β)-cholest-5-en-3-yl]-25,25-dimethyl-9,13,17,23-tetraoxo-24-oxa-2,8,12,16,22-pentaazahexacosane-1,3-dicarboxylate (8). To a solution of **7** (83 mg, 0.11 mmol) in a mixture of MeOH (6 mL) and THF (4 mL) was slowly added aqueous LiOH (2 mL, 0.5 M). The solution was stirred for 4 h at 22 °C. Solvents were removed *in vacuo*, the solid residue was resuspended in aqueous HCl (10%, 10 mL), and the resulting carboxylic acid was extracted with ethyl acetate (2 × 20 mL). The organic layer was dried over anhydrous Na_2SO_4 and solvents were removed *in vacuo*. The solid residue was dissolved in of anhydrous CH_2Cl_2 (10 mL) under dry N_2 and cooled to 4 °C. HOBt (22 mg, 0.14 mmol) and EDC (29 mg, 0.15 mmol) were added, the solution was stirred for 30 min at 4 °C, and N^{α}, N^{α}-bis[(tert-butoxycarbonyl)methyl]-L-lysine tert-butyl ester (75 mg, 0.17 mmol) was added. This amine was prepared from N^{α},N^{α}-*Bis[(tert*-butoxycarbonyl)methyl]-*N*-ε-benzyloxycarbonyl-L-lysine tert-butyl ester (the Cbz precursor) as previously reported,[3] but the Cbz precursor (3.4 g) in EtOH (50 mL) was deprotected by catalytic hydrogenation (Pd(C), 10%) using 300 psi H_2 for 12 h to afford 2.51 g of amine (97%). The reaction was warmed to 22 °C and stirred for 16 h. The solution was diluted with CH_2Cl_2 (20 mL) and washed with aqueous NaOH (0.1 M, 15 mL) followed by deionized H_2O (15 mL). The organic layer was dried over anhydrous Na_2SO_4 and concentrated *in vacuo*. Flash column chromatography (CH_2Cl_2 / MeOH, 30:1) afforded **8** (102 mg, 85%) as colorless oil; ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.86 (br s, 1H), 6.65 (br m, 2H), 5.31 (d, *J* = 4.9 Hz 1H), 3.50 (m, 4H), 3.42 (s, 2H), 3.43 (s, 2H), 3.32 (m, 2H), 3.28-3.16 (m, 2H), 3.08 (br s, 2H), 2.42 (m, 2H), 2.35 (t, *J* = 5.9 Hz, 2H), 2.17 (br, 2H), 2.01-0.84 (m, 86H), 0.66 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.4, 172.2, 171.9, 170.7 (× 3), 155.7, 141.8, 121.4, 81.7, 81.3 (× 2), 79.4, 64.9, 56.9, 56.3 (× 2), 54.2 (× 2), 50.3, 42.5 (× 2), 39.9, 39.7 (× 2), 39.2, 38.6 (× 2), 37.3, 36.8, 36.3, 36.0, 35.9, 35.6, 35.1, 32.0 (× 2), 29.4, 28.7 (Boc Me_3), 28.5, 28.4 (*t* -Bu Me_3), 28.3 (*t*-Bu Me_3 × 2), 28.2 (× 2), 27.8, 26.9, 24.4, 24.0, 23.4, 23.0, 22.7, 22.6, 21.2, 19.7, 18.9, 12.0; IR (film) ν max 3288, 3082, 2934, 2868, 1744, 1687, 1642, 1550, 1458, 1367, 1250, 1222, 1148 cm^{-1} ; CI *m/z* 1140.8524 (MH^+, $C_{65}H_{114}N_5O_{11}$ requires 1140.8515).



1

4-Carboxy-3-(carboxymethyl)-22-[(3β)-cholest-5-en-3-ylamino]-10,14,18-trioxo-3,9,13,17-tetraazadocosan-1-oic acid (1). TFA (1.5 mL) was added dropwise to a solution of **5** (22 mg, 0.019 mmol) in CH_2Cl_2 (5 mL). The solution was stirred for 16 h and the solvent was removed *in vacuo*. Purification by preparative reverse-phase HPLC (gradient: 49.9% MeCN, 50% H_2O, and 0.1% TFA to 99.9% MeCN, 0% H_2O, and 0.1% TFA over 30 min; retention time = 16.3 min (215 nm)) afforded **1** (15.6 mg, 94%) as a white solid, mp 101-103 °C; ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.34 (br s, 1H), 3.48 (m, 4H), 3.35 (br m, 5H), 3.10 (br s, 2H), 2.85 (br s, 3H), 2.31 (br m, 7H), 2.16 (br m, 2H), 1.94-0.75 (m, 50H), 0.65 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 174.7 (× 3), 172.2, 172.0 (× 2), 137.6, 123.8, 57.9, 56.4, 55.9, 55.0, 54.9 (× 2), 49.7, 44.1, 42.0, 40.0 (× 2), 39.4, 39.3 (× 2), 38.6, 36.8, 36.5, 36.0, 35.8, 35.6, 35.4, 34.9, 34.7, 31.6, 31.5, 29.2, 28.2, 28.0, 27.8 (× 2), 24.9, 24.0, 23.6, 23.1, 22.5, 22.3, 20.7, 18.8, 18.4, 11.6; IR (film) ν max 3700-2500 (br), 3295, 3084, 2938, 2868, 1732, 1686, 1634, 1556, 1468, 1441, 1382, 1202, 1143 cm^{-1}; ESI+ *m/z* 872.6093 (MH^+, $C_{48}H_{82}N_5O_9$ requires 872.6107).

Biological assays and protocols

General. Standard protocols were employed for microbiological techniques and plasmid construction. Oligonucleotide synthesis and sequencing of all new plasmid constructs was by the Pennsylvania State University Huck Institute Nucleic Acid Facility. Genes were constructed using the Polymerase Chain Reaction (PCR) with either *Pfu* polymerase (Stratagene) or Platinum *Taq* polymerase (Gene Choice). *E. coli* strain DH5-α (Clontech / BD biosciences) was used for subcloning, and proteins were expressed in *E. coli* strain BL21(DE3)pLysS (Novagen).

Cell culture. Jurkat lymphocytes (human acute T-cell leukemia, ATCC #TIB-152) were maintained in Roswell Park Memorial Institute (RPMI) 1640 media supplemented with Fetal Bovine Serum (FBS, 10%), penicillin (100 units / mL), and streptomycin (100 μg / mL). RPMI media used for cell culture and wash steps contained antibiotics and FBS unless otherwise noted.

Microscopy. A Zeiss LSM 5 Pascal confocal laser-scanning microscope fitted with a Plan Apochromat objective (63 X) was employed. Alexa Fluor 488 was excited with a 488 nm Argon ion laser and emitted photons were collected through 505 nm LP filter. Excitation of DiI-loaded LDL employed a 543 nm HeNe laser and a 560 nm LP filter.

Flow cytometry. Analyses were performed with a Beckman-Coulter XL-MCL bench-top flow cytometer. Forward-scatter (FS) and side-scatter (SSC) dot plots afforded cellular physical properties of size and granularity that allowed gating of live cells. After gating, 10,000 cells were counted. For studies of uptake of AcGFP His tag fusion proteins, AcGFP was excited at 488 nm with a 15 mW air-cooled argon-ion laser, the emission was split with a 550 nm dichroic and filtered through a 510 nm long pass filter and 530/30-nm band pass filter. The PMT voltage for this instrument was set to 724 for detection of AcGFP. Calibration with Sphero Rainbow Calibration particles (Spherotech) bearing 330,000 molecules of fluorescein/particle provided a fluorescence of 16.7 at this voltage.

Assays of uptake of AcGFP - oligohistidine fusion proteins.
Preload conditions: To Jurkat lymphocytes (7 x 10^5) in RPMI media (0.5 mL) was added synthetic receptor **1** (typical final [Receptor **1**] = 10 μM) in DMSO (final [DMSO] = 1%). These cells were incubated at 37 °C for 1 h to load the receptor into cellular plasma membranes. The cells were washed with RPMI media (0.5 mL) to remove unincorporated receptor / DMSO and resuspended in fresh media (87.6 μL). To these cells was added a solution containing aqueous $Ni(OAc)_2$ (4 μL, 2.5 mM, typical final [$Ni(OAc)_2$] = 100 μM) and oligohistidine tagged AcGFP (8.4 μL of 1.19 mg/mL, typical final [AcGFP] = 3.2 μM). Cells were maintained at 37 °C for 4 h to promote synthetic receptor-mediated endocytosis. Prior to analysis, treated cells were washed with disodium NTA in media (400 μM, 0.5 mL) and further incubated in media containing disodium NTA (400 μM, 0.5 mL) for 30 min to compete away any non-internalized protein. Cells were centrifuged, the supernatant was discarded, and the cells were resuspended in media (0.5 mL) for analysis by confocal microscopy and flow cytometry. For the competitive inhibition of uptake experiment shown in Figure 3, Panel A (labeled: + NTA (0.4 mM)), disodium NTA (400 μM) was included in the media during the 4 h incubation with AcGFP and $Ni(OAc)_2$. In Figure 3, Panels A-D, the concentration of oligohistidine-tagged AcGFP protein was 3.2 μM. In Figure 3, Panels A, C, and D, the metal acetate concentration was 100 μM. In Figure 3, Panels B-D, the concentration of synthetic receptor **1** was 10 μM.

Premix conditions (Figure 3, Panel A, last bar on the right): Jurkat lymphocytes (7 x 10^5) in RPMI media (100 μL total volume) were incubated 37 °C with a preequilibrated (30 min) solution of synthetic receptor **1** (10 μM), $Ni(OAc)_2$ (100 μM), AcGFP$(His)_{10}$ (3.2 μM), and DMSO (1%) for 4 h. Prior to analysis, treated cells were washed with disodium NTA in media (400 μM, 0.5 mL) followed by an additional 30 min incubation with media containing disodium NTA (400 μM, 0.5 mL) to compete away any non-internalized protein. Cells were centrifuged, the supernatant was discarded, and the cells were resuspended in media (0.5 mL) for analysis by flow cytometry.

Analysis of cytotoxicity of protein delivery. Jurkat lymphocytes were treated under the preload protein uptake assay conditions ([Receptor **1**] = 10 μM (preloaded for 1 h); [AcGFP] = 3.2 μM; [$Ni(OAc)_2$] = 100 μM (protein and metal acetate added for an additional 4 h). These cells were washed with disodium NTA, washed with fresh media, and incubated for an additional 48 h at 37 °C. The dead-cell stain propidium iodide (10 μg / mL) was added to cells prior to analysis, and viability was quantified by flow cytometry forward and side-scatter dot plots. No significant effects on viability or cellular morphology were observed under these conditions.

Quantification of the number of receptors on the surface of Jurkat lymphocytes.
Jurkat lymphocytes (7 x 10^5) in RPMI media (0.5 mL) were treated with synthetic receptor **1** (10 μM final conc.) in DMSO (1% final DMSO conc.). These receptor-treated cells were incubated at 37 °C for 1 h, washed with ice-cold RPMI media (0.5 mL) to remove unincorporated receptor / DMSO, and resuspended in ice-cold media (100 μL) containing $Ni(OAc)_2$ (100 μM) and excess AcGFP$(His)_{10}$ (final concentration = 19 μM). Cells were maintained at 4 °C for 1 h to saturate the NTA sites on the cell surface. Prior to analysis by confocal microscopy and flow cytometry, treated cells were washed ice-cold media (0.5 mL). As shown in Figure S2, confocal microscopy revealed that AcGFP$(His)_{10}$ was exclusively localized on the cell surface, and analysis by flow cytometry revealed a cellular fluorescence of 377.1. Construction of a fluorescence calibration curve with Sphero Rainbow Calibration particles (Spherotech) enabled calculation of 39,300,000 molecules of equivalent fluorescein (MEFL) per cell (based on analysis of 680,000 cells). From the fluorescence quantum yield of fluorescein of 0.93 and the fluorescence quantum yield of 0.82 for AcGFP, the average number of receptors per cell was calculated to be ~ 45,000,000.

Plasmid construction.
The gene encoding AcGFP was amplified by PCR using primers 5'-EcoRI-XhoI-AcGFP (5'-AGTCGAATTCGGTCTCGAGATGGTGAGCAAGGGC-3'), 3'-SalINoStop-AcGFP (5'-GACTGTCGACCTTGTACAGCTCATC-3'), and pAcGFP1(BD biosciences) as the template to append flanking *Eco*RI and *Sal*I restriction sites. This PCR product was digested with *Eco*RI and *Sal*I and inserted into *Eco*RI / *Xho*I-digested vectors pSA4 and pSLH2, derivatives of vector pLM[4] that add to the C-terminus of this gene decahistidine (pSA4) or hexahistidine (pSLH2) peptides followed by a stop codon (and *Sal*I site). The resulting AcGFP$(His)_{10}$ and AcGFP$(His)_6$ genes were digested with *Eco*RI / *Sal*I and ligated to the *Eco*RI / *Sal*I-digested vector pSM1, a derivative of pLM[4] modified by S. Martin to add a hemagglutinin (HA) epitope tag (flanked by in-frame restriction sites: *Mfe*I-MASYPYDVPDYASP-*Eco*RI) to the N-terminus of these proteins.

Protein overexpression.
E. coli strain BL21(DE3)pLysS was transformed with bacterial expression vectors pSM1-AcGFP(His)$_6$ or pSM1-AcGFP(His)$_{10}$. These transformed cells were grown to saturation in Luria Broth medium (LB, Difco/BD biosciences) containing ampicillin (100 μg / mL) and chloramphenicol (20 μg / mL). Saturated cultures (8 mL) were inoculated into LB (100 mL) containing ampicillin (100 μg / mL) and chloramphenicol (20 μg / mL) and incubated with shaking at 37 °C to an optical density of 0.6 (600 nm). Protein expression was induced by addition of IPTG (Invitrogen, final concentration = 1 mM). The cultures were shaken at 30 °C for 5 h and cells harvested by centrifugation (4400 rpm, 10 min). The cell pellets were resuspended in Bacterial Protein Extraction Reagent (BPER, Pierce, 3 mL). The cells were shaken at 30 °C for 30 min, centrifuged (10,000 rpm, 5 min) and the supernatant was applied to packed Talon resin (7.5 mL, Clontech / BD biosciences) prewashed with phosphate buffered saline (DPBS, pH 8). The resin was washed with cold DPBS (5 x 5 mL), and the protein was eluted from the resin by washing with DPBS containing imidazole (300 mM, 2 x 2 mL). The eluate was concentrated by centrifugation against a protein-impermeable membrane (Millipore centricon concentrator, 10K) and washed at 4 °C with cold DPBS (5 x 1.5 mL) to remove the imidazole. As shown in Figure S1, proteins were analyzed by SDS PAGE on a 15% polyacrylamide gel (Cambrex) and detected by staining with coomassie dye. The concentration of the protein was quantified using the Coomassie Plus protein assay reagent (Pierce). The purified proteins were stored at final concentrations of: 1.19 mg / mL (AcGFP(His)$_{10}$, 1000 μL) and 1.09 mg / mL, (AcGFP(His)$_6$, 750 μL).

Amino acid sequences of AcGFP proteins expressed in *E. coli*. In addition to the C-terminal oligohistidine sequence, these proteins also include an N-terminal hemagglutinin (HA) epitope tag (sequence: YPYDVPDYA).

AcGFP(His)$_6$

MRGSGTELQLMASYPYDVPDYASPEFGLEMVSKGAELFTGIVPILIELNGDVNGHKFSVSGEGEGDATYGK
LTLKFICTTGKLPVPWPTLVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYIQERTIFFEDDGNYKSRAEV
KFEGDTLVNRIELTGTDFKEDGNILGNKMEYNYNAHNVYIMTDKAKNGIKVNFKIRHNIEDGSVQLADHY
QQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMIYFGFVTAAAITHGMDELYKVEGTHHHHHH

AcGFP(His)$_{10}$

MRGSGTELQLMASYPYDVPDYASPEFGLEMVSKGAELFTGIVPILIELNGDVNGHKFSVSGEGEGDATYGK
LTLKFICTTGKLPVPWPTLVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYIQERTIFFEDDGNYKSRAEV
KFEGDTLVNRIELTGTDFKEDGNILGNKMEYNYNAHNVYIMTDKAKNGIKVNFKIRHNIEDGSVQLADHY
QQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMIYFGFVTAAAITHGMDELYKVEGTHHHHHHHHHH

References for the supporting information

(1) Orlova, E. V.; Sherman, M. B.; Chiu, W.; Mowri, H.; Smith, L. C.; Gotto, A. M., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 8420-8425.
(2) Boonyarattanakalin, S.; Martin, S. E.; Dykstra, S. A.; Peterson, B. R. *J. Am. Chem. Soc.* **2004**, *126*, 16379-16386.
(3) Dorn, I. T.; Neumaier, K. R.; Tampe, R. *J. Am. Chem. Soc.* **1998**, *120*, 2753-2763.
(4) Sodeoka, M.; Larson, C. J.; Chen, L.; Leclair, K. P.; Verdine, G. L. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1089-1094.