Supporting information

Cell-Specific Targeting by Heteromultivalent Ligands

Jatinder S. Josan, Heather L. Handl, Rajesh Sankaranarayanan, Liping Xu, Ronald M. Lynch, Josef Vagner, Eugene A. Mash, Victor J. Hruby and Robert J. Gillies

CONTENTS

EXPERIMENTAL SECTION

Abbreviations:

Abbreviations used for amino acids and designation of peptides follow the rules of the IUPAC-IUB Commission of Biochemical Nomenclature in *J. Biol. Chem.* **1972**, *247*, 977-983. The following additional abbreviations are used: AMBER, Asisted Model Building Energy Refinement; Boc, tert-butyloxycarbonyl; *t*Bu, tert-butyl; CCK-2R, Cholecystokinin Receptor subtype 2; CCK-6, Nle-Gly-Trp-Nle-Asp-Phe-NH_2; CCK-8, Asp-Tyr-Nle-Gly-Trp-Nle-Asp-Phe-NH_2; CD, Circular Dichroism; CDI, N,N'-carbonyldiimidazolide; CH_3CN, acetonitrile; Cy5, Cyanine 5 dye; δ-OR, delta-opioid receptor; DCM, dichloromethane; DIC, N,N'-diisopropylcarbodiimide; DIEA, diisopropylethylamine; DMEM, Dulbecco's Modified Eagle Medium; DMF, N,N'-dimethylformamide; DMSO, dimethylsulfoxide; DTPA, diethylenetriamine-N^1,N^2,N^3,N^4-pentaacetic acid; ESI-MS, Electrospray ionization - mass spectrometry; Et_2O, Diethyl ether; Fmoc, (9H-fluoren-9-ylmethoxy)carbonyl; FT-ICR, Fourier Transform - Ion Cyclotron Resonance; HBTU, 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluoro-phosphate; hMC4R, human melanocortin-4 receptor; HOBt, N-hydroxybenzotriazole; HOCt, 6-chloro-1H-hydroxybenzotriazole; htBVL, heterobivalent ligands; htMVL, heteromultivalent ligand; MALDI, Matrix Assisted Laser Desorption Ionization - Time of Flight; MD, Molecular Dynamics; MMFF, Merck Molecular Force Field; MSH, melanocyte-stimulating hormone; MSH-7, Ser-Nle-Glu-His-DPhe-Arg-Trp; NDP-α-MSH, Ac-Ser-Tyr-Ser-Nle-Glu-His-DPhe-Arg-Trp-Gly-Lys-Pro-Val-NH_2; Mtt, 4-methyltrityl; NHS, N-hydroxysuccinimide; Pbf, 2,2,4,6,7-pentamethyl-dihydrobenzofuran-5-sulfonyl; PEG, Polyethyleneglycol; PEGO, 19-amino-5-oxo-3,10,13,16-tetraoxa-6-azanonadecan-1-oic acid; RP-HPLC, reverse-phase high performance liquid chromatography; SPPS, solid-phase peptide synthesis; SD, stochastic dynamics; THF, Tetrahydrofuran; TFA, trifluoroacetic acid; TRF, time-resolved fluorescence; Trt, triphenylmethyl (trityl).

Materials:

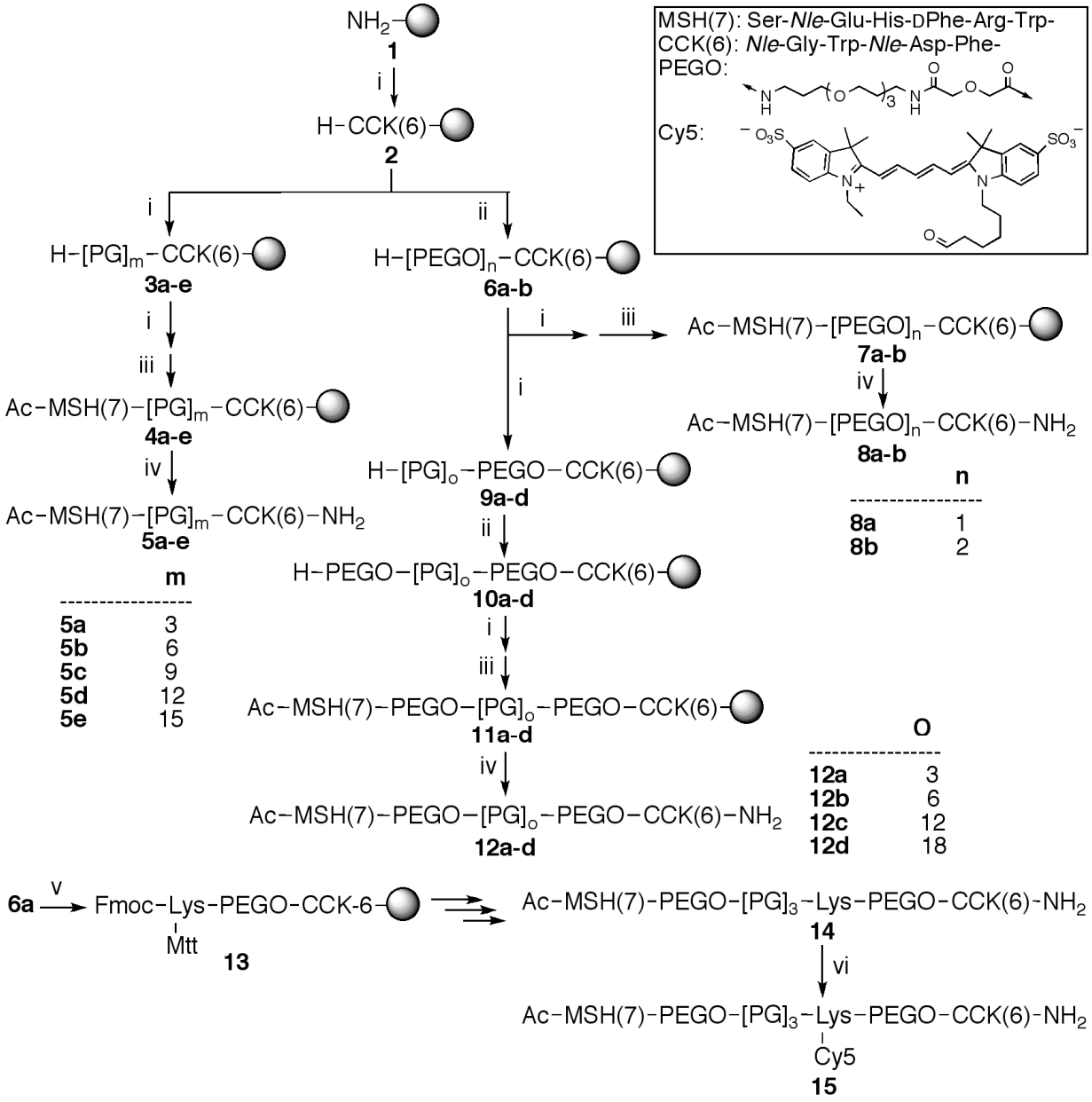
N^{α}-Fmoc protected amino acids, HBTU, and HOBt were purchased from SynPep (Dublin, CA) or from Novabiochem (San Diego, CA). Rink amide Tentagel S resin was acquired from Rapp Polymere (Tubingen, Germany). HOCt, DIC and DIEA were purchased from IRIS Biotech (Marktredwitz, Germany). The following side chain protecting groups were used for the amino acids: Arg(N^g-Pbf); Asp(O-*t*Bu); Glu(O-*t*Bu); His(N^{im}-Trt); Ser(*t*Bu), Trp(N^i-Boc); Lys(N^{ε}-Mtt). Cy5-NHS ester was acquired from Amersham Biosciences (Piscataway, NJ). Diglycolic anhydride and 4,7,10-trioxa-1,13-tridecandediamine were acquired from TCI America (Portland, OR). Peptide synthesis solvents, dry solvents, and solvents for HPLC were reagent grade, were acquired from VWR (West Chester, PA) or Sigma-Aldrich (Milwaukee, WI), and were used without further purification unless otherwise noted. All the peptides were manually assembled using 5 to 50 mL plastic syringe reactors equipped with a frit, and Domino manual synthesizer obtained from Torviq (Niles, MI). The C-18 Sep-Pak™ Vac RC cartdridges for solid phase extraction were purcahsed from Waters (Milford, MA).

Solid-Phase Peptide Synthesis

Peptides were synthesized on Tentagel Rink amide resin (initial loading: 0.2 mmol/g) using N^{α}-Fmoc protecting groups and a standard DIC/HOCt or HBTU/HOBt activation strategy. The resin was swollen in THF for an hour, washed with DMF, and Fmoc protecting group removed with 20% piperidine in DMF (2 min, washing, then 20 min). The resin was washed with DMF (3X), DCM (3X), 0.2 M HOBt in DMF (2X), and finally with DMF (2X) and the first amino acid coupled using pre-activated 0.3 M HOCt ester in DMF (3 eq. of N^{α}-Fmoc amino acid, 3 eq. of HOCt and 6 eq. of DIC). An on-resin test using Bromophenol Blue was used for qualitative and continuous monitoring of reaction progress.[1,2] To avoid deletion sequences and slower coupling rate in longer sequences, the double coupling was performed at all steps with 3 eq. of amino acid, 3 eq. of HBTU and 6 eq. of DIEA in DMF. A third coupling was performed with symmetric anhydride method (2 eq. of amino acid and 1 eq. of DIC in dichloromethane) wherever beads still tested Kaiser positive. Any unreacted NH_2 groups on the resin thereafter were capped using an excess of 50% acetic anhydride in pyridine for 5 min. When the coupling reaction was finished, the resin was washed with DMF, and the same procedure was repeated for the next amino acid until all amino acids were coupled. The PEGO spacers were introduced as described before.[1,3], Briefly, the N-terminal of the peptide on resin was coupled with the glycolic acid spacer using 10 eq. of diglycolic anhydride in DMF for 5 min. The resin was washed with DMF (3X), with the last washing with dry DMF, and the free carboxylic groups were activated using 10 eq. of carbonyldiimidazole in dry DMF for 30 min. The resin was washed with dry DMF (3X), and the PEG diamine coupled using 20 eq. of 4,7,10-trioxa-1,13-tridecanediamine in dry DMF for 30 min (vigorous vortexing for first 5 minutes).

Parallel Library Synthesis: The syntheses of htBVLs **5a** – **12d** consisting of MSH-7 and CCK-6 ligands connected in a head-to-tail fashion by PEGO and/or Pro-Gly linkers are depicted in Scheme S1. Starting with Tentagel Rink amide polystyrene resin (initial loading 0.2 mmol/g), the hexapeptide CCK-6 was constructed and the resin **2** split into two portions. To one portion, a PEGO linker was coupled on the N-terminus. The resin **6a** containing PEGO unit was again proportionally split for the synthesis of compounds **8a-b** and **12a-d**. At this stage, proline and glycine residues were added alternatively for resins **3a-e** and **9a-d**. After the final proline addition for these compounds, the N^{α}-Fmoc protecting group was removed. A second PEGO unit was coupled to the resins for **6b** and **10a-d**. The free amine terminals of all resins were coupled with Fmoc-Trp(Boc)-OH and syntheses continued to complete the MSH-7 sequence, the N^{α}-terminus deprotected, then acetylated to give **4a-e**, **7a-b**, and **11a-d**.

Frequently during the synthetic steps, a small amount of peptide was cleaved and analyzed by HPLC to monitor the



SCHEME S1. Synthetic route for heterobivalent ligands (htBVLs). The inset shows sequence of MSH-7 and CCK-6 ligands, structure of the PEGO linker (20 atoms), and the Cy5 dye with a hexanoyl linker. i. Fmoc/*t*Bu synthesis continued as follows: a) Fmoc-AA-OH (3eq), HOCt (3eq), and DIC (6eq) in DMF (10 mL / 1g of resin) for amino acid couplings; b) Piperidine/DMF (1:4) for Fmoc deprotection; ii. The PEGO linker was assembled as follows: a) Diglycolic anhydride (10eq) in DMF for 5 min, b) CDI (20eq) in anhyd. DMF for 30 min, c) 4,7,10-trioxa-1,13-tridecaneamine (20eq) in anhyd. DMF for 30 min; iii. Ac_2O/Pyridine (1:1) for acetylation (or N-capping); iv. TFA-scavengers cocktail (82.5% trifluoroacetic acid, 5% water, 5% triisopropylsilane, 5% thioanisole, 2.5% ethanedithiol); v. Fmoc-Lys(Mtt)-OH (3eq.), HOCt (3eq.), DIC (6eq.) in DMF; vi. Cy5-NHS ester (1.2 eq.) in DMSO. AA = amino acid.

synthesis and purity of the peptide. The resin was washed with DMF, DCM and THF and dried. A cleavage cocktail (10 mL per 1 g of resin) of TFA (82.5%), water (5%), triisopropylsilane (5%), thioanisole (5%), and ethanedithiol (2.5%) was injected into the resin and stirred for 3 h at room temperature. The crude peptides **5a-e**, **8a-b**, and **12a-d** were isolated from the resin by filtration, the filtrate was reduced to low volume by evaporation using a stream of nitrogen, and the peptides were precipitated in ice-cold diethyl ether, washed several times with ether, dried, dissolved in water and lyophilized to give off-white solid powders that were stored at -20°C until purified. The final compounds were purified by size-exclusion chromatography and RP-HPLC, and characterized by ESI-MS and/or MALDI-TOF and/or FT-ICR. The yields of the crude peptides were 50-80% based on the resin weight gain, and overall, the purified

yields for the syntheses were 5-30% based on the loading of the resin.

The Cy5 label was introduced in the bivalent ligand construct using lysine as a functional handle. After, the first PEGO linker incorporation in resin **6a**, an N^{α}-Fmoc-N^{ε}-Mtt protected lysine was incorporated into the sequence (**13**) and the peptide synthesis continued to the end as above, the peptide was then cleaved from the resin and purified using preparative HPLC to give **14**. For Cy5 labeling, the purified peptide **14** was dissolved in DMSO, 1.2 eq. of commercially available Cy5-NHS ester was added and the reaction was monitored using analytical HPLC at 280 nm. Finally, the labeled peptide was separated using size exclusion chromatography and lyophilized to yield a blue amorphous final product.

The lanthaligands (Eu-DTPA labeled NDP-α-MSH and CCK-8) for screening were synthesized using Fmoc/*t*Bu synthesis as described previously.[4,5] Briefly, the NDP-α-MSH and CCK-8 ligands were synthesized on Tentagel Rink amide resin. After the final amino acid addition, the N-terminal was deprotected and the DTPA chelator was attached to the N-terminus of the peptide using an ε-aminocaproic spacer. Using an *in situ* HOBt ester method, DTPA dianhydride was treated with HOBt for 20 min to obtain di-HOBt ester, which was coupled to the free ε-amino groups on resin. After cleavage, the peptides were purified by reverse-phase HPLC and chelate labeled with Eu(III) chloride in neutral pH buffer. The excess of metal salt was removed by solid-phase extraction (SPE). The purified products were characterized using high resolution ESI-MS and/or MALDI-TOF and/or FT-ICR.

HPLC Analysis, Size Exclusion Chromatography, Solid Phase Extraction and Peptide Conc. Determination:

The purity of final products was analyzed using Waters high-performance liquid chromatography (HPLC) apparatus and with a Vydac C18 reverse phase column (dia × length: 4.6 mm × 150 mm, pore size: 3 μm). Buffer A was with 0.1% TFA in water and buffer B was 0.1% TFA in acetonitrile. Peptides were analyzed using a linear gradient of buffer B under various gradient conditions at a flow rate of 0.3 - 1 mL/min and the separations monitored at 220 and 280 nm. Purification of compounds was achieved using a Waters 600 HPLC apparatus equipped with a Vydac C18 reverse phase column (22 × 250 mm, 5μm) with similar buffers and under optimized gradients and 3 - 10 mL/min flow rate, and monitored at 230 and 280 nm (see Table S1 for characterization data and Figure S1a for an example).

Size exclusion chromatography was performed on a borosilicate glass column (2.6 × 250 mm, Sigma, St. Louis, MO) filled with medium sized Sephadex G-25 or G-10. The compounds were eluted with an isocratic flow of 1.0 M aqueous acetic acid.

Solid-Phase Extraction (SPE) was employed where simple isolation of final compound was needed from excess salts and buffers for e.g., lanthaligand synthesis. For this purpose, C-18 Sep-Pak™ cartridges (100 mg or 500 mg) were used and pre-conditioned initially with 5 column volumes (5 times the volume of packed column bed) each of acetonitrile, methanol, and water, in that order. After loading the compound, the column was washed several times with water, and then gradually with 5, 10, 20, 30, 50, and 70% of aqueous acetonitrile to elute the peptide.

The peptide concentrations were determined by monitoring absorbance of peptides against 0.5 mM solution of Tryptophan (D or L) in DMSO at 280 nm. The peptides were initially dissolved in DMSO at approximately 1-5 mM concentration. Co-injections of peptide and Trp were made on analytical HPLC with a number of different volumes and peptide concentration calculated from area under the peaks (see Figure S1b for an example) using the formula given here:

$$\text{Peptide Conc.} = \frac{[\text{Abs. of Comp.}]}{[\text{Abs. of Trp}]} \times \frac{0.5}{\frac{\sum \varepsilon_{280}(\text{Trp} + \text{Tyr} + \text{Cys} + \text{Cy5} + \ldots)}{\varepsilon_{280}(\text{Trp})}} \times \frac{\text{Vol. of Trp}}{\text{Vol. of Comp.}}$$

where ε_{280} of compounds were determined by summation of tryptophans (ε_{280} = 5500), Cy5 dye (ε_{280} = 5800), and normalized to extinction coefficient of one tryptophan. Other amino acids in these peptides do not absorb significantly at this wavelength. For Cy5 dye, ε_{280} was determined in a similar manner by comparing the absorbances of 0.5 mM of both Trp and Cy5-NHS ester in DMSO at 280 nm wavelength.

Mass Spectrometry:

Mass spectra of positive ions were recorded either with a single stage reflectron MALDI-TOF mass spectrometer (Bruker Rexlex III, Bruker Daltonics, Billerica, MA; α-cyanocinnamic acid as a matrix) in reflectron mode or with a low resolution ESI mass spectrometer (Finnigan, Thermoquest LCQ ion trap instrument, Lake Forrest, CA) and/or using high resolution Fourier transform mass spectrometer (FT-ICR MS, Bruker Apex Qh, Bremen, Germany) equipped with an ESI source. For internal calibration, an appropriate mixture of standard peptides was used with an average resolution of *ca.* 10,000 on the Reflex III and 60,000 on the FT-ICR instrument (see Figure S2 for an example).

Molecular Modeling:

Conformational searches and molecular dynamics were performed with Macromodel version 9.1 implemented under Maestro 7.5 interface on a Linux workstation. The MacroModel implementations of Merck Molecular Force Field (MMFF), AMBER* and OPLS all-atom force fields were used.[6] AMBER* is a reparametrized AMBER force field containing a new set of torsional parameters that more closely reproduces *ab initio* calculations on the conformational preferences of simple peptides.[7] For solution phase calculations, the GB/SA continuum model for water was used. Amide bonds were required to be *trans* except in case of prolines whose imide bonds were intentionally sampled and accepted with either *cis* or *trans* geometry in the conformational searches. Conformational searches were performed with systematic Monte Carlo method of Goodman and Still.[8] For each search, 5000 starting structures were generated and minimized until the gradient was less than 0.05 kJ/mol/Å, using the truncated Newton-Raphson method implemented in MacroModel. Duplicate conformations and

those with energy greater than 50 kJ/mol above the global minimum were discarded.

All molecular dynamic simulations were performed at 310°K with Monte Carlo/Stochastic Dynamics (MC/SD) hybrid simulation algorithm[9] with either the AMBER* all-atom force field or the new OPLS-2005 force field in MacroModel 9.1. A time step of 1.5 fs was used for the stochastic dynamics (SD) part of the algorithm. The MC part of the algorithm used random torsional rotations between ± 60° and ± 180° that were applied to all rotatable bonds except the proline imide C-N bond where the random rotations between ± 90° and ± 180° were applied. No torsion rotations were applied to bonds in the pyrrolidine ring of proline as the barriers are low enough to permit adequate sampling from the SD part of the simulation. The total simulation time was up to 20,000 ps and samples were taken at 1 ps interval, yielding appropriate number of conformations for analysis. This analysis is represented in Figure S4.

For G-Protein Coupled Receptor size estimations, the PDB file of a recently described bovine rhodopsin protein in a trigonal crystal form (1GZM)[10] was loaded into PyMOL program (www.pymol.org). Several residues on the outer edges of the transmembrane domain were chosen and distances across the GPCR were measured. The "width" of the GPCR was taken as the average of 8 of these measured distances. Similarly, the homology modeled GPCR structures of hMC4R and CCK-2R were downloaded from www.gpcr.org and the distance of the binding pocket from the edge of the receptor and the depth of the binding pocket were determined. The residues forming the binding pocket were taken as available from the literature. This analysis is represented in Figure S3.

Circular Dichroism:

CD measurements were carried out using a JASCO J-720 spectrometer equipped with a water-jacketed cuvette holder. Data were collected with a quartz cuvette of 2 mm path length, from 300 to 185 nm, at 0.5 nm intervals, and with a scan rate of 1 nm per second. The spectra were recorded as an average of at least 3 scans and were corrected by subtracting the solvent/buffer base line contributions. No curve-smoothing procedure was applied. The solvent conditions included water, 95% TFE v/v in water, 95% nPrOH v/v in water, and PBS (phosphate buffered saline) buffer (pH 7.4). The concentration dependence was studied initially for Ac-$[PG]_6$-NH_2 peptide from 1 mM to 10 μM range, and was then kept constant at 100 μM for the rest of the studies. For melting temperature scan in water, a heating rate of 1°C/minute was used and the full CD spectra were collected at 5°C intervals from -10°C to 90°C. Data are presented in molar ellipticities ([θ], deg•cm^2•$dmol^{-1}$) (see Figure S5 for results).

Cloning, Transfection and Cell Culture:

HEK293 cells overexpressing the human melanocortin-4 receptor (hMC4R) were used to assess the binding at the hMC4R. The hMC4R vector was originally received from Dr. Ira Gantz, University of Michigan.[11] The coding region of the hMC4R gene was expressed in pcDNA3.1 (Invitrogen, V790-20). HEK293/hMC4R cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% FBS. Monovalent CCK binding was tested on HEK293 cells with stable expression of CCK-2 receptor. These cells were grown in DMEM supplemented with 10% FBS and were maintained under selection with 100 μg/mL Zeocin. Evaluation of bivalent ligand binding was completed on HEK293/hMC4R cells with transient expression of CCK-2R (referred to as HEK293/hMC4R/tCCK cells). For transient receptor expression, cells were plated at a density of 10,000 cells/ well in Wallac B&W Isoplate TC (Wallac/PerkinElmer, 1450-583) 96-well plates. The day after plating, cells were transfected with CCK-2R using FuGENE 6 Transfection Reagent (Roche, 1814-443). A 3:1 ratio of FuGENE to DNA was used, as per reagent protocol. Reagents (30 μL media, 0.15 μL FuGENE 6, and 0.05 μg CCK-2R DNA per well) were added to a sterile microtube and incubated at room temp for 15 min. The media used during the incubation was antibiotic free, serum free DMEM. After the 15 min incubation, the reaction mixture (30 μL) was added to the cells in their normal media. It was determined that 72 hours post-transfection was optimal for high surface expression of CCK-2R, thus all binding assays were performed 72-hours post-transfection.

Receptor Number Determination:

The number of receptors present on the cell surface was determined through saturation binding analysis followed by correlation of the fluorescent signal achieved at B_{max}. Increasing amounts of Eu-labeled ligand were added to cells until saturation was achieved. For hMC4R, the B_{max} was determined to be 95, 244 ± 2480 AFU (Figure S6a). A standard curve relating fluorescent signal to the concentration of Eu-labeled ligand present in the well was produced (Figure S6b) and used to determine the concentration of Eu-labeled ligand present at saturation. In the case of hMC4R, a signal of 95, 244 AFU correlates with 380 fmol/well. Assuming that during the saturation study that each receptor is bound by a single Eu-labeled ligand, this correlates with 2.29×10^{11} receptors/well. The average number of cells/well was obtained via counting with a hemocytometer and this number (61,840 cells/well, average of 10 determinations) was then used to determine the number of hMC4R receptors per cell. The same process was followed for determining the number of CCK-2R per cell.

Lanthanide Based Binding Assays:

Lanthanide based competitive binding assays were conducted according to the method which has been previously described.[12] As an example, HEK293/hMC4R cells were plated in black and white 96-well isoplates (Wallac, 1450-584) at a density of 12,000 cells/well and were allowed to grow for 3 days. On the day of the experiment, media was aspirated from all wells. 50 μL of non-labeled ligand and 50 μL of Eu-labeled lanthaligand (final concentration of 10 nM for Eu-NDP-α-MSH) were added to each well. Ligands were diluted in binding media (DMEM, 1mM 1,10-Phenanthroline, 200 mg/L Bacitracin, 0.5 mg/L Leupeptin, 0.3% BSA) and samples were tested in quadruplicate, unless otherwise noted. Cells were incubated in the presence of ligands for 1 hr at 37°C. Following

the incubation, cells were washed 3× with 250 μL Wash Buffer (50 mM Tris-HCl, 0.2%BSA, 30 mM NaCl). Enhancement solution (Perkin Elmer; 1244-105) was added (100 μL/well) and the plate was incubated for at least 30 min at 37°C prior to reading. The plates were read on a Wallac VICTOR instrument using the standard Eu TRF measurement (340 nm excitation, 400 μsec delay, and emission collection for 400 μsec at 615 nm). Competition curves were analyzed with GraphPad Prism Software using the sigmoidal dose-response (variable slope) classical equation for non-linear regression analysis. See Figure 2 & S7 for representative binding curves.

Cell-Surface Labeling:

Cells were grown on #1 coverslips harbored in individual wells of six well plates. Individual slides were mounted in a chamber maintained at 37^{o}C on the stage of an inverted Olympus IX70 microscope equipped with a 40X 1.35 NA objective. For excitation of Cy5 fluorescence, white light emitted from a 150W Xe lamp was passed through a 10 nm band pass filter centered at 640 nm. The emitted light was selected using a band pass filter centered at 680 nm, and subsequently imaged onto a CCD camera (Photometrics CH-350; TEK-512 chip). Three control images were acquired at 5 min intervals prior to addition of the labeled htBVL to incubation medium. Immediately following addition of the ligand, images were acquired within 30 sec, 1 min and 3 min, then the media was then replaced with ligand-free medium prior to further image acquisition. Image analysis was performed on a SGI Indy-2 workstation using customized software. See Figure 5 for illustration of specific cell-surface labeling with htBVL **15**.

TABLE S1. HPLC & MS data of Ac-MSH(7)-NH_2, Ac-CCK(6)-NH_2 and heterobivalent ligands **5a – 15**.

Compound	Mass Calculated[a]	Mass Found	t_R (Purity %)	K'
Ac-MSH(7)-NH_2	1015.5035 $(M+1)^{1+}$	1015.4881[b] $(M+1)^{1+}$	17.0[e] (96)	9.2
Ac-CCK(6)-NH_2	791.4014 $(M+1)^{1+}$	791.1[b] $(M+1)^{1+}$	15.4[f] (96)	7.7
5a ($C_{107}H_{145}N_{27}O_{25}$)	1105.0452 $(M+2)^{2+}$	1105.2[b] $(M+2)^{2+}$	20.3[g] (99)	11.9
5b ($C_{128}H_{175}N_{33}O_{31}$)	1336.1566 $(M+2)^{2+}$	1336.3[b] $(M+2)^{2+}$	19.2[g] (100)	11.3
5c ($C_{149}H_{205}N_{39}O_{37}$)	1045.1786 $(M+3)^{3+}$	1045.1891[c] $(M+3)^{3+}$	18.3[g] (98)	10.8
5d ($C_{170}H_{235}N_{45}O_{43}$)	1199.2529 $(M+3)^{3+}$	1199.2559[c] $(M+3)^{3+}$	18.0[g] (99)	10.6
5e ($C_{191}H_{265}N_{51}O_{49}$)	4059.9942* $(M+1)^{1+}$	4060.5121*[d] $(M+1)^{1+}$	17.7[g] (87)	10.4
8a ($C_{100}H_{141}N_{23}O_{25}$)	1033.0235 $(M+2)^{2+}$	1032.9[b] $(M+2)^{2+}$	21.5[g] (100)	12.7
8b ($C_{114}H_{167}N_{25}O_{31}$)	1192.1129 $(M+2)^{2+}$	1192.0[b] $(M+2)^{2+}$	21.2[g] (100)	12.5
12a ($C_{135}H_{197}N_{31}O_{37}$)	949.1496 $(M+3)^{3+}$	949.2[b] $(M+3)^{3+}$	20.1[g] (100)	11.8
12b ($C_{156}H_{227}N_{37}O_{43}$)	827.6678 $(M+4)^{4+}$	827.6797[c] $(M+4)^{4+}$	26.1[h] (98)	11.4
12c ($C_{198}H_{287}N_{49}O_{55}$)	1058.7792 $(M+4)^{4+}$	1058.7916[c] $(M+4)^{4+}$	25.4[h] (98)	11.0
12d ($C_{240}H_{347}N_{61}O_{67}$)	5156.5621 $(M+1)^{1+}$	5156.5889[d] $(M+1)^{1+}$	24.7[h] (93)	10.7
15 ($C_{174}H_{246}N_{35}O_{45}S_2$)	3610.7478 $(M+1)^{1+}$	903.6916[c] $(M+4)^{4+}$	14.6[f] (96)	5.1

[a] Exact Mass calculated based on the most abundant isotope of the element; * This compound is shown with the exact mass of most abundant peak in the monoisotopic peak pattern; [b] ESI-MS; [c] FT-ICR MS; [d] MALDI-MS; [e] 10-40% B gradient in 30 min; [f] 10-90% B gradient in 30 min; [g] 20-60% B gradient in 50 min; [h] 10-40% B gradient in 50 min; HPLC eluents: Phase A is 0.1% TFA in water; Phase B is 0.1% TFA in acetonitrile; t_R is the retention time of compound peak in HPLC; (purity of final product in percentage is given in parenthesis); K' is retention time of compound peak/retention time of solvent peak.

FIGURE S1. HPLC profile of htBVL 8a. A) HPLC profile of crude compound **8a** at 220 and 280 nm, (B) Peptide concentration determination of compound **8a** using 0.5 mM D-Tryptophan standard (Retention time of 2.4 min) co-injected in analytical HPLC and monitored at 280 nm.

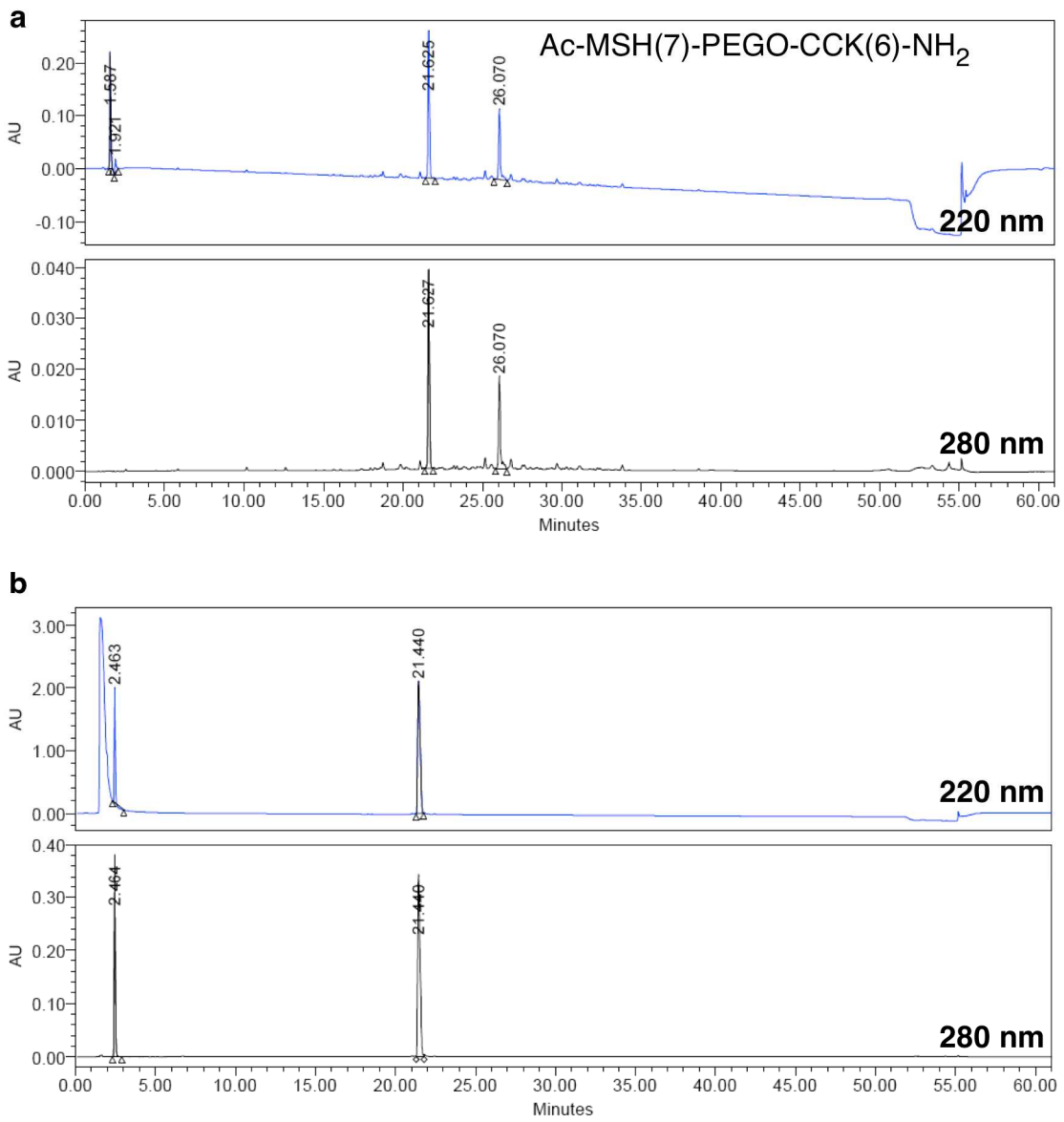


FIGURE S2. ESI-MS of compound **8a**

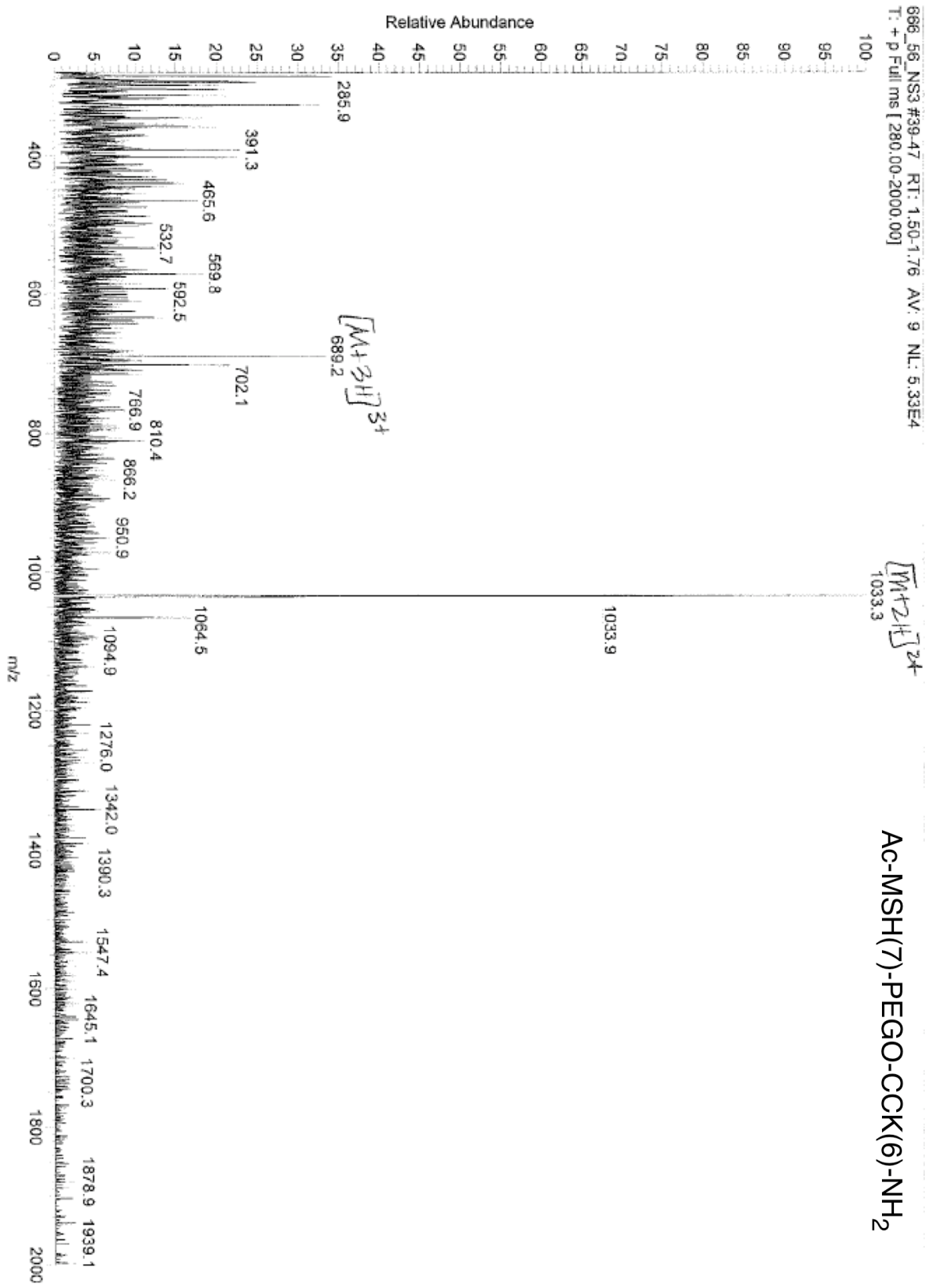


FIGURE S3. GPCR modeling a) PDB file 1GZM loaded in PyMOL software and shows the two GPCR molecules per each asymmetric crystal unit. The end-to-end distance along the longer elliptical axis is ~ 70 □.[10] b) The distance between two adjacent TM4 helices is shown in the picture. Note that this distance is not between the binding pockets as the two GPCRs are oriented head-to-tail, and merely suggests an approximate distance that would be between two off-center binding pockets arranged head-to-head. c) The homology modeled GPCR structure of hMC4R with off-center binding pocket (green surface) and measured distances. d) GPCR dimer orientations. The two receptors can pack in a (*top*) head-to-head, (*middle*) head-to-tail or (*bottom*) tail-to-tail or any number of intermediate orientations. The distance span between the two binding pockets could be up to 50 □ long in a tightly packed dimer. The possibility of domain swapping or involvement of lipid rafts could shorten or lengthen this distance span.

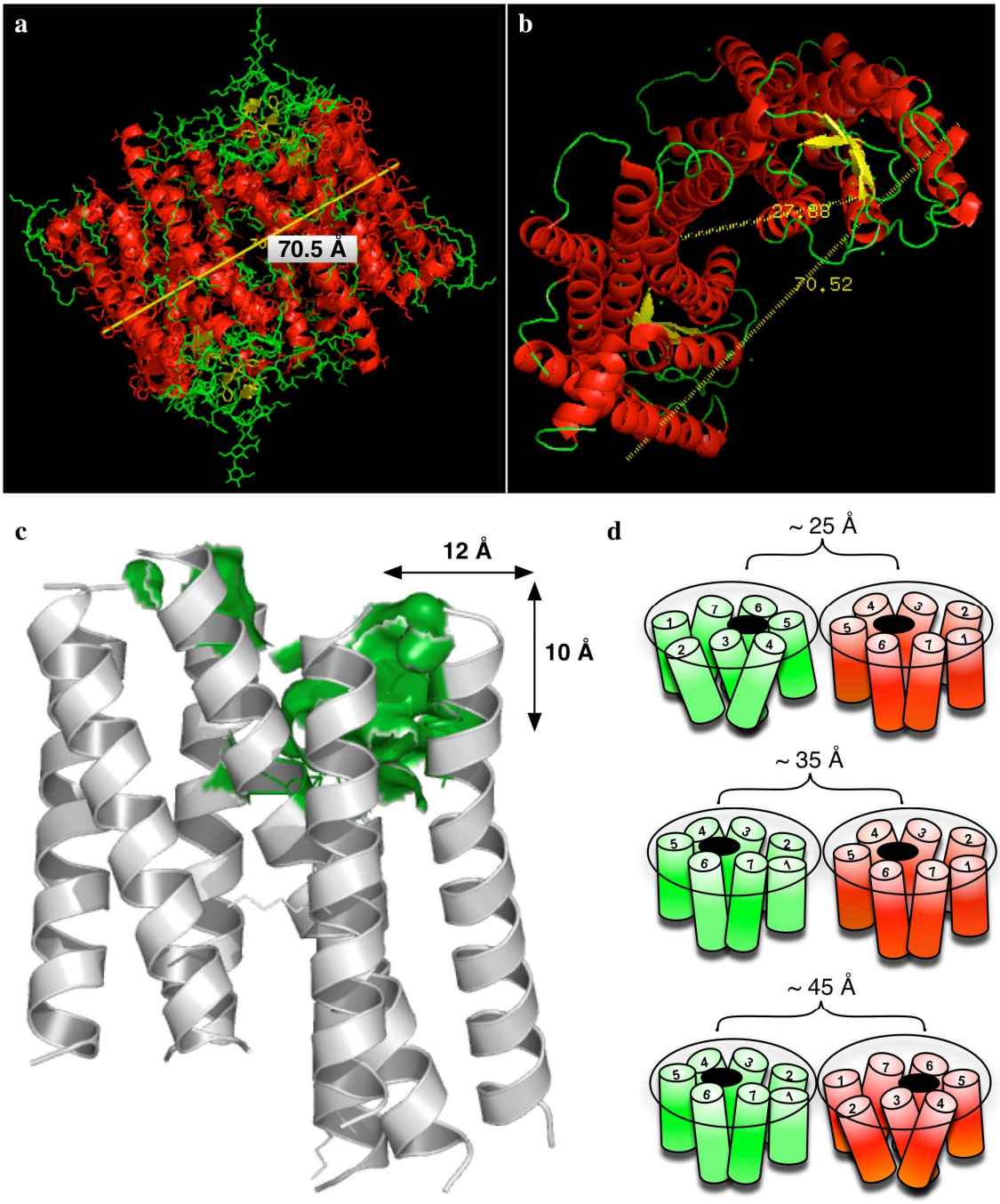


FIGURE S4. Molecular modeling. Computational studies were carried out on Ac-PEGO-NH_2, Ac-$[PG]_n$-NH_2, where n = 3, 6, or 9, htBVL **12b**, and were used to estimate the linker length in each htBVL given in Table 1. Representative cases of linkers are described here. a,b) One of the low energy conformers of modeled Ac-PEGO-NH_2 linker with a bent conformation. The MD simulations indicated that this linker could extend up to 18 Å length (distance between atoms labeled 1 and 50); c) One of the low energy conformers of Ac-$[PG]_9$-NH_2 linker highlighting the helical secondary structure. The modeling studies indicated the presence of PolyProline Type II (PPII) structure in this proline-rich sequence (also see CD studies in Figure S4). d,e) MD simulations carried on the htBVL **12b** depicting the excursion of linker region $[PG]_6$ and PEGO-$[PG]_6$-PEGO in the sampled MD simulation of 20,000 ps. The PEGO-$[PG]_6$-PEGO linker could extend up to 53 Å in this simulation. The $[PG]_6$ part of the linker exhibited some random conformations that could extend up to nearly 30 Å but most conformations retained a certain degree of helicity with linker length less than 20 Å. More detailed studies on $[PG]_n$ linker will be published elsewhere.

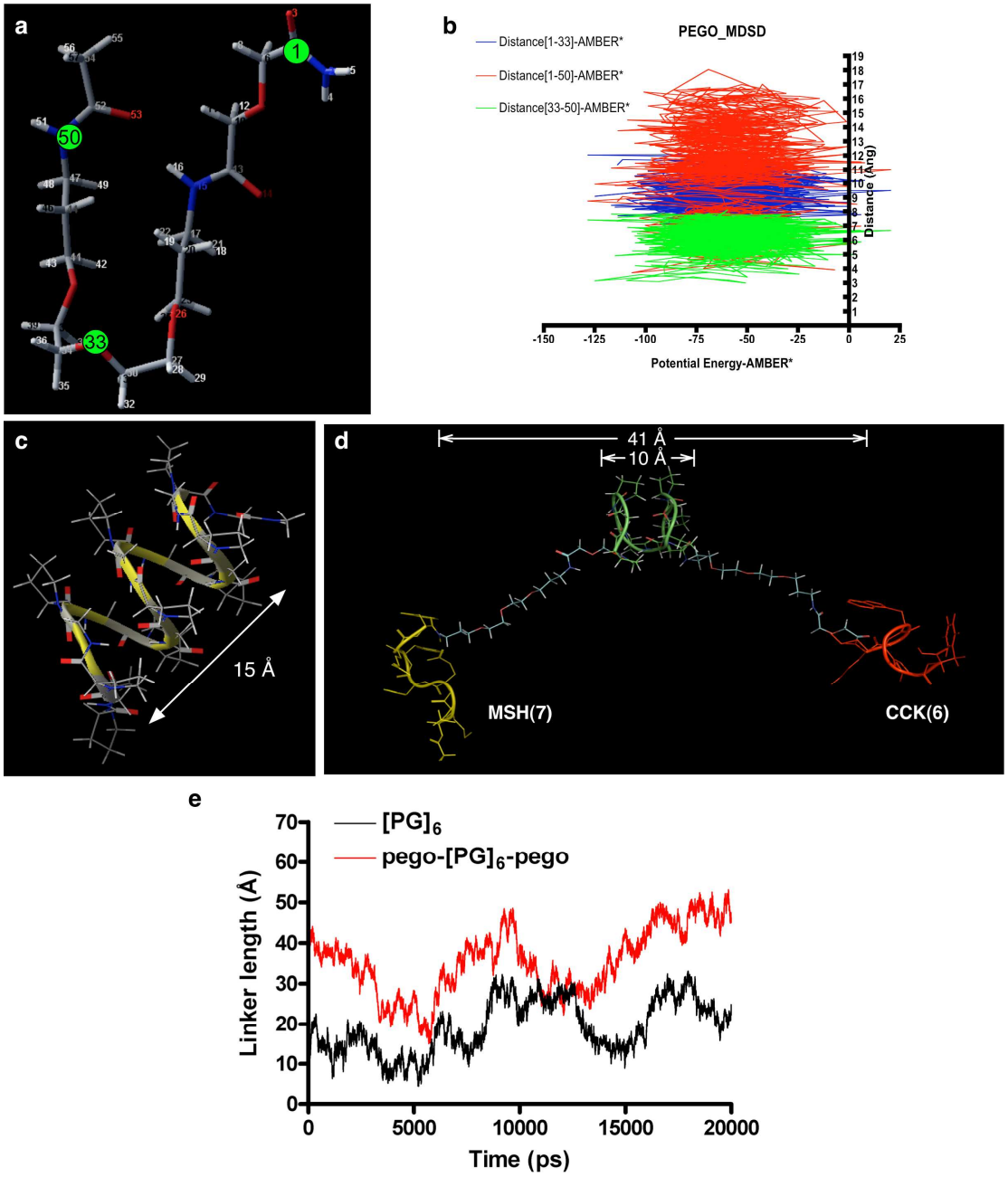


FIGURE S5. CD study on poly(Pro-Gly) linker. a) CD spectrum of Ac-$[PG]_6$-NH_2 in different solvent at 100 μM concentration and 25 °C temperature. The peptide was synthesized on solid-phase using procedures described earlier in the supporting text. This linker shows negative absorption band at 205 nm, which is characteristic of polyproline type II (PPII) secondary structure. However, the weak the postive band usually centered at 225 nm in PPII is absent in this linker, and may reflect the presence of some other structural element. b) Temperature scan of Ac-$[PG]_6$-NH_2 linker in water at 100 μM concentration, highlighting the trace similar to other polyproline sequences and the presence of an isodichroic point at 213 nm that suggests the presence of two equilibrating structures. Further deconvolution analysis of these linkers with different peptide lengths revealed a mixture of PPII (~ 40% content), β-turn, and random coils, reinforcing our design consideration of a semi-rigid linker. More experimental details on poly(Pro-Gly) linkers will be published elsewhere.

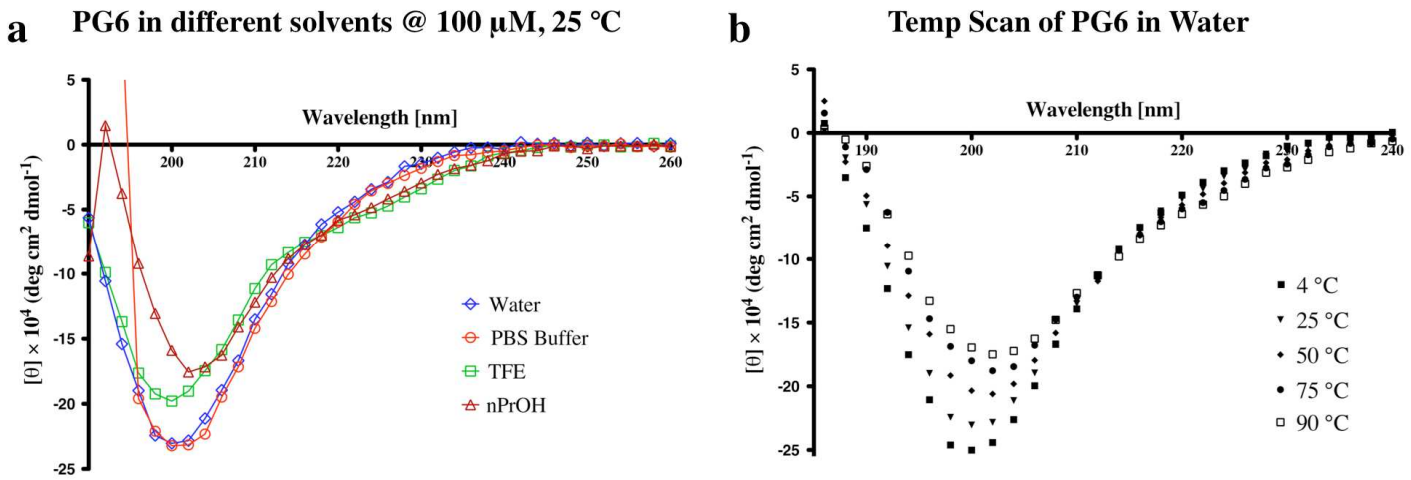


FIGURE S6. Receptor number determination. a) Saturation binding analysis of Eu-NDP-α-MSH binding to HEK293/hMC4R/tCCK cells where (■), (o) and (▲) indicate total, specific and non-specific binding, respectively. Increasing amounts of Eu-NDP-α-MSH were added to cells and incubated for 1 hr at 37ºC. Non-specific binding was determined in the presence of 100 μM NDP-α-MSH. From these data, the K_d = 1.30 ± 0.14 nM, and B_{max} = 95,244 ± 2480 AFU. The data were fit using GraphPad Prism software using the non-linear regression, one site-binding equation, with a R^2 value of 0.93. Each data point indicates the average of 4 samples, with error bars indicating the standard error mean. b) Standard curve relating [Eu-NDP-α-MSH] to fluorescent signal. Increasing amounts of Eu-NDP-α-MSH were added to wells of a 96-well plate. Each data point indicates the average of 4 replicates with error bars indicating the standard error mean. The standard curve allowed for the determination of the amount of ligand present at the B_{max} value obtained from the saturation binding curve. For binding to the hMC4R, the B_{max} was determined to be 95,244 ± 2480 AFU which correlates to 380 fmol/well.

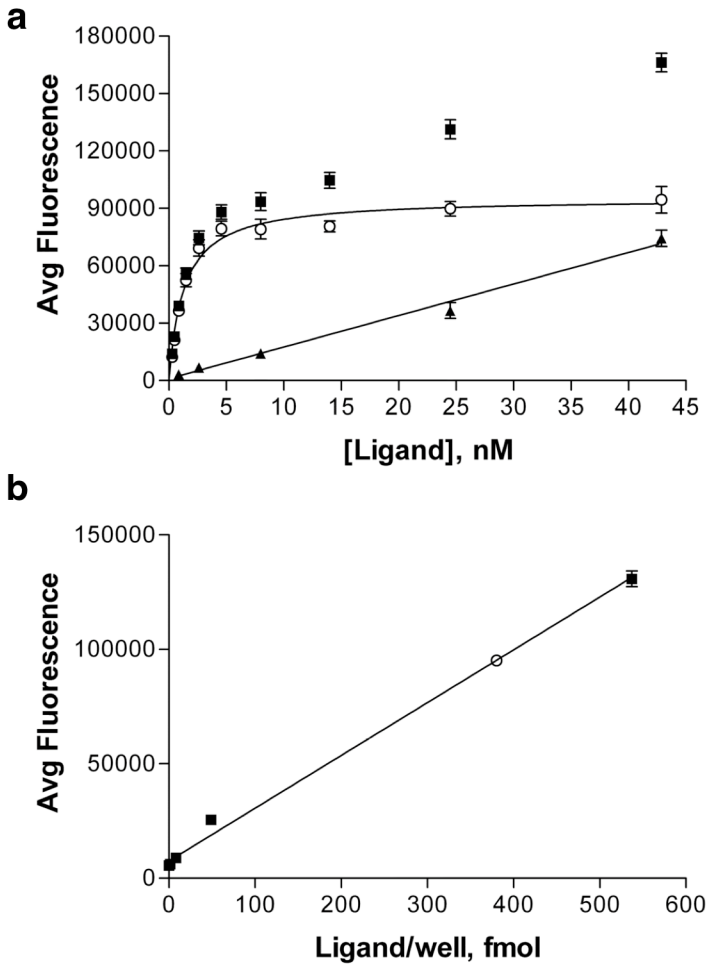
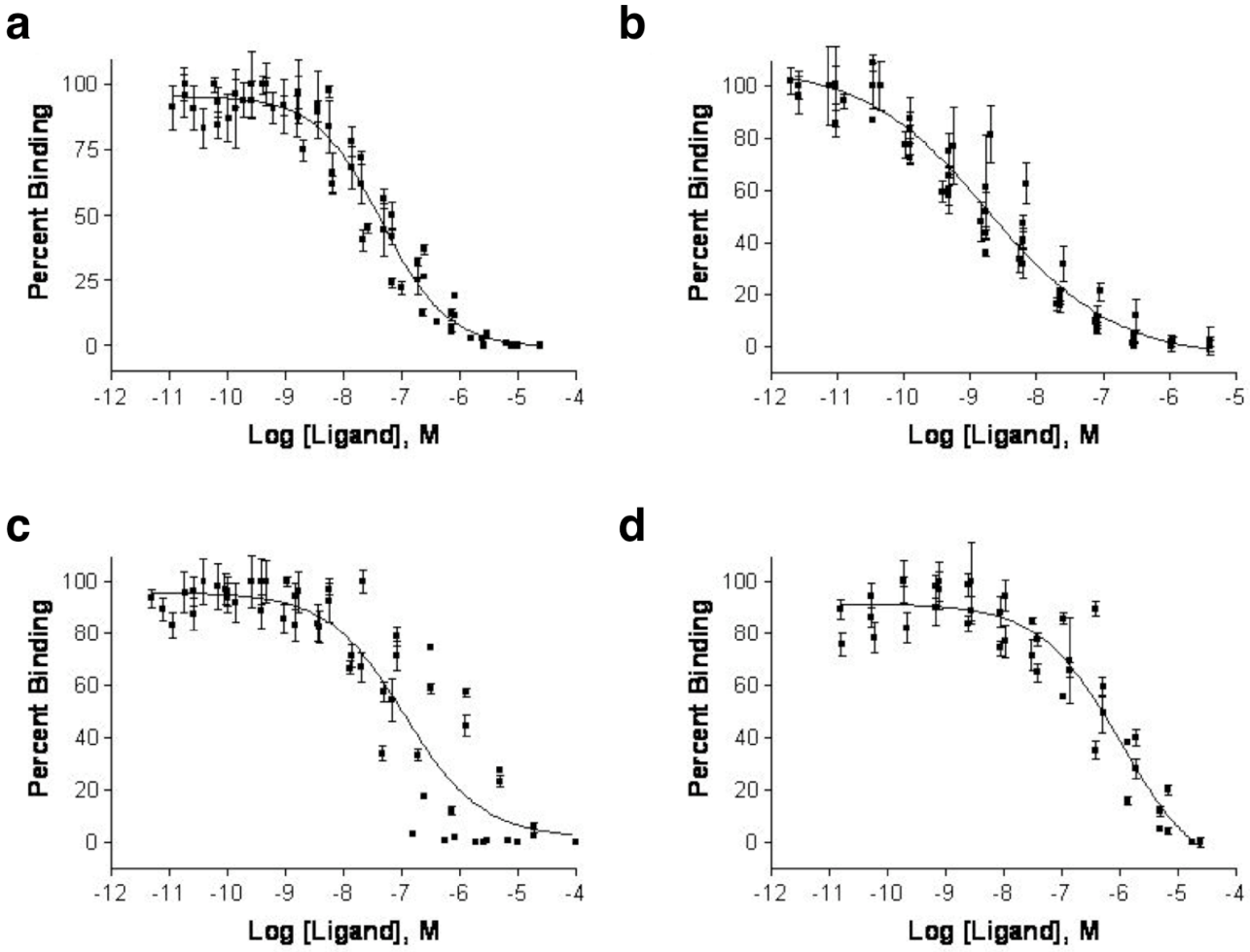


FIGURE S7. Representative curves from the competitive binding assay of heterobivalent ligands. The ligands were evaluated for their monovalent and bivalent binding by competing them against Eu-labeled NDP-α-MSH and CCK-8 lanthaligands. Single plot IC_{50} values were determined where data from all n measurements were pooled first and non-linear regression analysis performed. a) Binding of Ligand **5b** competed with 0.1 nM Eu-CCK8 in HEK293/CCK cells, with an IC_{50} of 46 nM (R^2 = 0.90). b) Binding of Ligand **5b** competed with 0.1 nM Eu-CCK8 in HEK293/MC4R/CCK cells, with an IC_{50} of 2.3 nM (R^2 = 0.89). c) Binding of Ligand **5b** competed with 0.1 nM Eu-NDP-α-MSH in HEK293/MC4R cells, with an IC_{50} of 110 nM (R^2 = 0.81) d) Binding of Ligand **5b** competed with 0.1 nM Eu-NDP-α-MSH in HEK293/MC4R/CCK cells, with an IC_{50} of 1000 nM (R^2 = 0.86).



REFERENCES:

[1] V. Krchnák, J. Vágner, M. Lebl, *Int. J. Pept. Prot. Res.* **1988**, *32*, 415.
[2] J. S. Josan, J. Vagner, H. L. Handl, R. Sankaranarayanan, R. J. Gillies, V. J. Hruby, *Int. J. Pep. Res. Ther.* **2008**, *14*, 293.
[3] J. Vagner, H. L. Handl, R. J. Gillies, V. J. Hruby, *Bioorg. Med. Chem. Lett.* **2004**, *2004*, 211.
[4] H. L. Handl, J. Vagner, H. I. Yamamura, V. J. Hruby, R. J. Gillies, *Anal. Biochem.* **2005**, *343*, 299.
[5] J. Reubi, B. Waser, J. C. Schaer, U. Laederach, J. Erion, A. Srinivasan, M. A. Schmidt, J. E. Bugaj, *Eu. J. Nucl. Med.* **1998**, *25*, 481.
[6] F. Mohmadi, N. G. J. Richards, W. C. Guida, R. Lizkamp, M. Lipton, C. Caulfield, G. Chang, T. Hendickson, W. C. Still, *J. Comput. Chem.* **1990**, *11*, 440.
[7] D. Q. McDonald, W. C. Still, *Tet. Lett.* **1992**, *33*, 7743.
[8] J. M. Goodman, W. C. Still, *J. Comput. Chem.* **1991**, *12*, 1110.
[9] F. Guarnieri, W. C. Still, *J. Comput. Chem.* **1994**, *15*, 1302.
[10] J. Li, P. C. Edwards, M. Burghammer, C. Villa, G. F. X. Schertler, *J. Mol. Biol.* **2004**, *343*, 1409.
[11] I. Gantz, H. Miwa, Y. Konda, Y. Shimoto, T. Tashiro, S. J. Watson, J. Delvalle, T. Yamada, *J. Biol. Chem.* **1993**, *268*, 15174.
[12] H. L. Handl, J. Vagner, H. I. Yamamura, V. J. Hruby, R. J. Gillies, *Anal. Biochem.* **2004**, *330*, 242.