Electronic Supplementary Material (ESI) for Chemical Communications

Encoded peptide libraries and the discovery of new cell binding ligands

*Nina Svensen, Juan José Díaz-Mochón, and Mark Bradley**

School of Chemistry, The University of Edinburgh, Joseph Black Building, West Mains Road, Edinburgh, EH9 3JJ, UK.

Supplementary information

Materials and methods

Cell culture: HEK293T, HEK293T-CCR6 (courtesy of Dr. Dominic Compopiano, University of Edinburgh, UK), and D54 cells (courtesy of Dr. Darell D. Bigner, Duke University Medical Centre, California) were maintained in Dubelco's Modified Eagle Medium (DMEM, Sigma-Aldrich) complete medium [CM: 10% fetal bovine serum (FBS, Biosera), 2 mM L-glutamine, 100 mg/ml streptomycin, and 100 U/ml penicillin (Gibco)]. All cells were grown in T-75 or T-25 flasks (Nunc) in high humidity and 5% CO_2 at 37 ^{o}C. Cell lines were passaged by removing the old media from the flask, washing the surface with PBS (5 ml Oxoid), and detachment by standard trypsination: the media was removed and the cells washed with PBS. A minimal amount of 1 x trypsin (Fisher Scientific) in PBS was added and the cells WERE incubated at 37 ^{o}C until complete detachment (media was added to end trypsination).

Microarray hybridization: Custom DNA microarrays (Oxford Gene Technologies) complementary to the PNA sequences in Library-1 were hybridized with 5 µM PNA_Library-1 in 150 µl 1 x GenHyb (Genetix) buffer overnight lowering the temperature from 65 ^{o}C to 37 ^{o}C. The arrays were washed according to standard OGT protocols (10 min at 30 ^{o}C with Oxford Gene Buffer (100 mM NaCl, 40 mM citric acid, 0.7% (w/v) *N*-lauroylsarcosine sodium salt, 0.1 mM EGTD, pH7.5), rinsed with 3 x H_2O, 1 x Tris buffer pH8, and dried under a flow of N_2). Microarrays were scanned with a Tecan LS Reloaded microarray scanner with ArrayPro 32 Analyzer (Media Cybernetics Inc.) software and a FITC filter producing 4,000 x 12,000 pixel images. Analysis of the microarray images was performed using the software Bluefuse.

Cell growth on Library-1 microarrays: D54 and HEK293T-CCR6 cells were grown to 80% confluence in DMEM CM enriched with $CaCl_2$ (50 µM) and $MgCl_2$ (50 µM) and detached by trypsination (above). Cells were allowed to reattach for 8 h to allow regeneration of surface receptors.

The cells were then detached by gently tapping the flask and incubated (5 mL, 50,000 cells/mL) with the Library-1 hybridized microarray in DMEM enriched with $CaCl_2$ (50 μM) and $MgCl_2$ (50 μM) for 2 h at 37 °C. After which time the media was removed and fresh media was added and the incubation continued for 16 h. The microarray was washed PBS enriched with $CaCl_2$ (50 μM) and $MgCl_2$ (50 μM) for 30 s and attached cells were fixed (20 ml 4% paraformaldehyde in PBS for 15 min) and stained with phalloidin (200 μl, 10 nM in PBS for 15 min) and DAPI (20 ml, 10 μM in PBS for 15 min). The microarray was imaged using a Nikon Eclipse 50i microscope with Pathfinder (Imstar) software (20 x objective; FITC, DAPI, and TRITC filters) producing 5,000 x 13,000 pixel images. The DAPI-filtered image and the FITC-filtered image acquired prior to cellular incubation (above) were aligned in the following manner: the two FITC-filtered images acquired pre and post cellular incubation were superimposed and aligned using ImageViewer 1.0 software. Hereafter, the resulting realigned FITC-filtered image acquired prior to cellular incubation and the DAPI image were superimposed and analyzed using Bluefuse (BlueGenome) software (ArrayExpress database, accession number E-MEXP-3104). Note all microarray data complies with the Minimal Information About a Microarray Experiment (MIAME) guidelines.

Deconvolution of microarray data: Raw microarray data was obtained from Bluefuse, which allowed grid alignment and signal estimation, in an Excel format, with the top ~5% and the bottom ~5% of each of the replicate-sets being removed as outliers, (erroneous values caused by dust, scrapes etc[1,2]). The average fluorescence intensities were calculated for each of the remaining 14 replicates and over the non-complementary (negative control) features. The average intensities were corrected for the background by subtracting the average intensity of the non-complementary negative control features. Hereafter, the data was normalized for the PNA hybridization by calculation of the ratio between background corrected average intensities of the DAPI-label and the FAM-label.

Peptide synthesis: Peptides were synthesized on an automated CEM-peptide synthesizer using NH_2-Rink-polystyrene resin (10.6 mg; 1.07 mmol/g) according to previously reported procedures.[3,4] Amino acid couplings: amino acids [Phe, Ile, Pro, *d*-Pro, Tyr, Lys, Ser, Arg, Asp, Glu; 5.5 equiv.] were coupled using HTBU/HOBt (5 equiv.) and DIPEA in DMF (10 equiv., 500 μl) under microwave irradiation for 20 min at 60 °C. Fmoc-deprotection: 20% piperidine in DMF (2 mL) was added to the resin and the reaction was stirred for 2 x 10 min. Hereafter, the *N*-termini were coupled with Fmoc-6-aminohexanoic acid followed by Fmoc cleavage and capping with 5(6)-carboxyfluorescein (5.5 equiv.), PyBOP (5 equiv.), and *N*-ethylmorpholine (NEM) in DMF (10 equiv., 500 μl), 3 h at room temperature. After synthesis the resins were washed twice with 20% piperidine in DMF (2 mL) for 10 min to cleave unwanted fluorescein dimers[3]. The peptide-PNA conjugates were cleaved from the solid support by treatment with 5% TIS in TFA (500 μl) for 2 x 2 h and the products were precipitated with cold diethyl

ether (x 2) and collected by centrifugation. The peptides were purified by HPLC (Column: Phenomenex Luna, C18, 15 cm x 1.00 cm, 5 µm.; λ = 254 nm; Buffer A: H_2O with 0.1% formic acid; Buffer B: MeCN with 0.1% formic acid. eluting with 95% A to 95% B over 3 min; 100% B for 1 min) and analyzed by ES-MS (positive mode): FAM-Ahx-FQpIYIIIK-NH_2: 7.35 mg (41%). RP-HPLC: 4.195 min. (46.4% crude purity). MALDI-TOF^+/MS: m/z (%) 1577.5 $[M+H]^+$ (100). FAM-Ahx-FQKKSRIK-NH_2: 13.5 mg (72%). RP-HPLC: 2.818 min. (65.0% crude purity). MALDI-TOF^+/MS: m/z (%) 1652 $[M+H]^+$ (100). FAM-Ahx-FQIPYIIIK-NH_2: 6.53 mg (36%). RP-HPLC: 4.190 min. (55.2% crude purity). MALDI-TOF^+/MS: m/z (%)1576.6 $[M+H]^+$ (100). FAM-Ahx-FQPKYSFIK-NH_2: 14.4 mg (79%). RP-HPLC: 3.457 min. (79.5% crude purity). MALDI-TOF^+/MS: m/z (%) 1600.4 $[M+H]^+$ (100). FAM-Ahx-RGD-NH_2: 6.65 mg (74%). RP-HPLC: 3.030 min. (95.4% crude purity). MALDI-TOF^+/MS: m/z (%) 789.2 $[M+H]^+$ (100).

Microarray printing: FAM-Ahx-FQpIYIIIK-NH_2 (200 nM), FAM-Ahx-FQKKSRIK-NH_2 (200 nM), FAM-Ahx-FQIPYIIIK-NH_2 (200 nM), and FAM-Ahx-FQPKYSFIK-NH_2 (200 nM) in sodium phosphate (50 mM, pH 8.5) were contact printed on Codelink slides (GE Healthcare) in 45% humidity using 4 solid pins, 1 stamp per ink, 3 stamps per spot, and 10 ms stamp- and inking time in 9 spots/feature (squares) in a 3 x 3 grid in 3 subarrays using a Genetix QArray Mini Microarrayer. The peptides were immobilized by incubating the microarrays in a box with filter paper soaked with NaCl (3 M) overnight. Unreacted sites were blocked with 2 M perfluorooctyl propylamine and DIPEA (2 M) in acetonitrile for 30 min. The microarrays were washed with acetonitrile (2 x 20 mL) for 30 s, H_2O (2 x 20 mL) for 30 s, 0.1% SDS in 4 x SSC (20 mL) at 50 °C for 30 min, H_2O (2 x 20 mL) for 30 s, and Tris pH 8 (20 mL, 10 mM) for 30 s. The array was imaged with a BioTEc LaVision BioAnalyzer using a FITC filter.

Cell growth on "hit"-peptide microarrays: Cell growth on the FAM-Ahx-FQpIYIIIK-NH_2/FAM-Ahx-FQKKSRIK-NH_2 displaying microarrays was carried out as described above with D54 (3 mL, 100,000 cells/mL) in DMEM enriched with $CaCl_2$ (50 µM) and $MgCl_2$ (50 µM). In addition, D54 (3 mL, 20,000 cells/mL) were pre-incubated with FAM-Ahx-RGD-NH_2 (100 µM) in DMEM enriched $CaCl_2$ (50 µM) and $MgCl_2$ (50 µM) for 20 min before incubation on the microarray. Likewise, HEK293T-CCR6 or HEK293T cells (3 mL, 100,000 cells/mL) were incubated on FAM-Ahx-FQIPYIIIK-NH_2/FAM-Ahx-FQPKYSFIK-NH_2 displaying microarrays in DMEM enriched with $CaCl_2$ (50 µM) and $MgCl_2$ (50 µM). In addition, HEK293T-CCR6 cells (3 mL, 100,000 cells/mL) were pre-incubated with PerCP-Anti human CCR6 (5 µg/mL, R&D systems) in DMEM enriched $CaCl_2$ (50 µM) and $MgCl_2$ (50 µM) for 20 min before incubation on the microarray. Cells were fixed, stained and analyzed as described above a BioTEc LaVision BioAnalyzer using a FITC filter and high-resolution images were aquired

using a Nikon Eclipse 50i microscope with a 40x objective and FITC, DAPI, and TRITC filters and Pathfinder software (Imstar).

Cytotoxicity assays: D54 and HEK293T-CCR6 cells were grown to ~50% confluence (to avoid cells reaching complete confluence during the 48 h experiment) in 96-well plates and incubated with FAM-labeled peptide (100 μM) in appropriate growth media for 24 h at 37 °C and the media was removed and fresh media was added and the incubation continued for 24 h with the last row was used as a blank (no cells seeded). Cells were washed with PBS and 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT, 50 μL, 1mg/mL, Sigma-Aldrich) in Iscove's Modified Dulbecco's Media (IMDM) was added and incubated for 4 h at 37 °C. Hereafter, MTT solubilizing solution [60 μL, 10% Triton-X 100 (Sigma-Aldrich), 0.1 N HCl (Fisher Scientific) in isopropanol] was added and the well-plate was shaken till complete dissolution of the formazan crystals. Absorbance was measured at 570 nm in a Bio-Rad Benchmark microplate reader. Untreated cells were considered to be 100% viable.

Supplementary figures

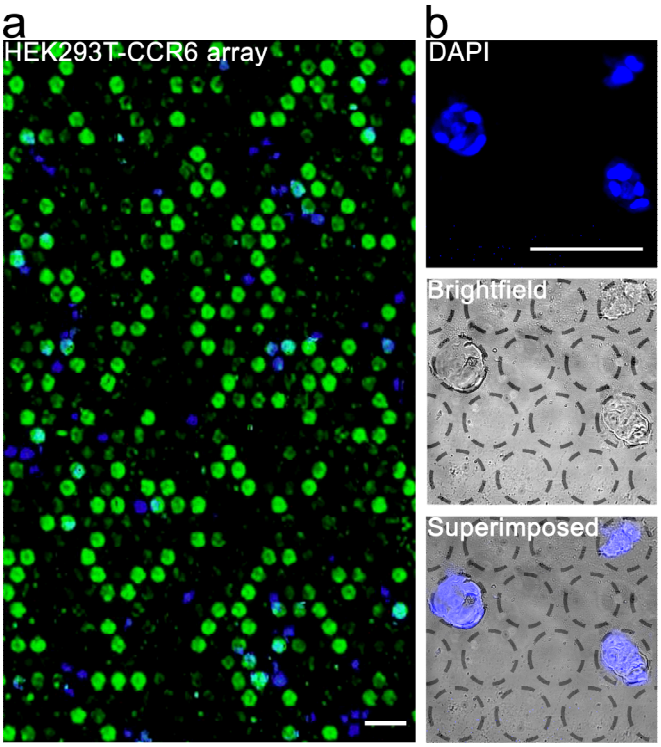

Figure S1: Microarray analysis of cell binding onto PNA-encoded peptides. DNA microarrays were designed with 4 sub-arrays of 44,000 features each with 4 replicates of each oligonucleotide complementary to each member of the 10,000 member library as well as 4,000 non-coding negative controls. HEK293T-CCR6 cells incubated on Library-1 hybridized microarrays in serum free media (2 h) followed by incubation in complete media (16 h) and fixation (4% paraformaldehyde) and nuclei staining with DAPI (blue). (**a**) FAM and DAPI fluorescent imaging and (**b**) Brightfield and DAPI fluorescent imaging of the HEK293T-CCR6 cell array. Scale bars = 20 μm.

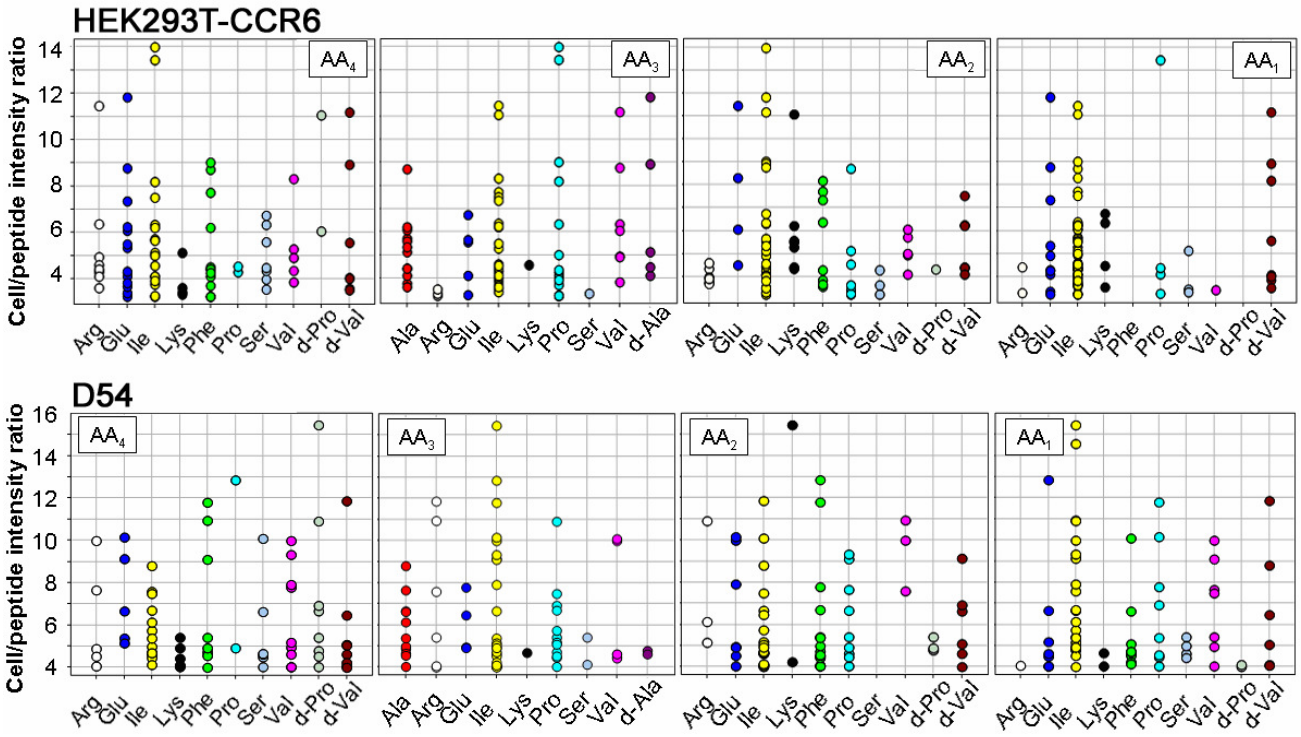

Figure S2: Scatter plots of the top 100 relative microarray intensities derived from the binding of D54 and HEK293T-CCR6 cells on Library-1 hybridized microarrays versus the variable amino acids for positions $AA_{1\text{-}4}$. Cellular preference for specific peptides was elucidated by combining the most preferred amino acids for each of the variable positions (i.e. the amino acids with the highest number of high-intensity dots/peptides). Consensus sequences of the least preferred peptides were obtained by combining the amino acids with the lowest intensity peptides.

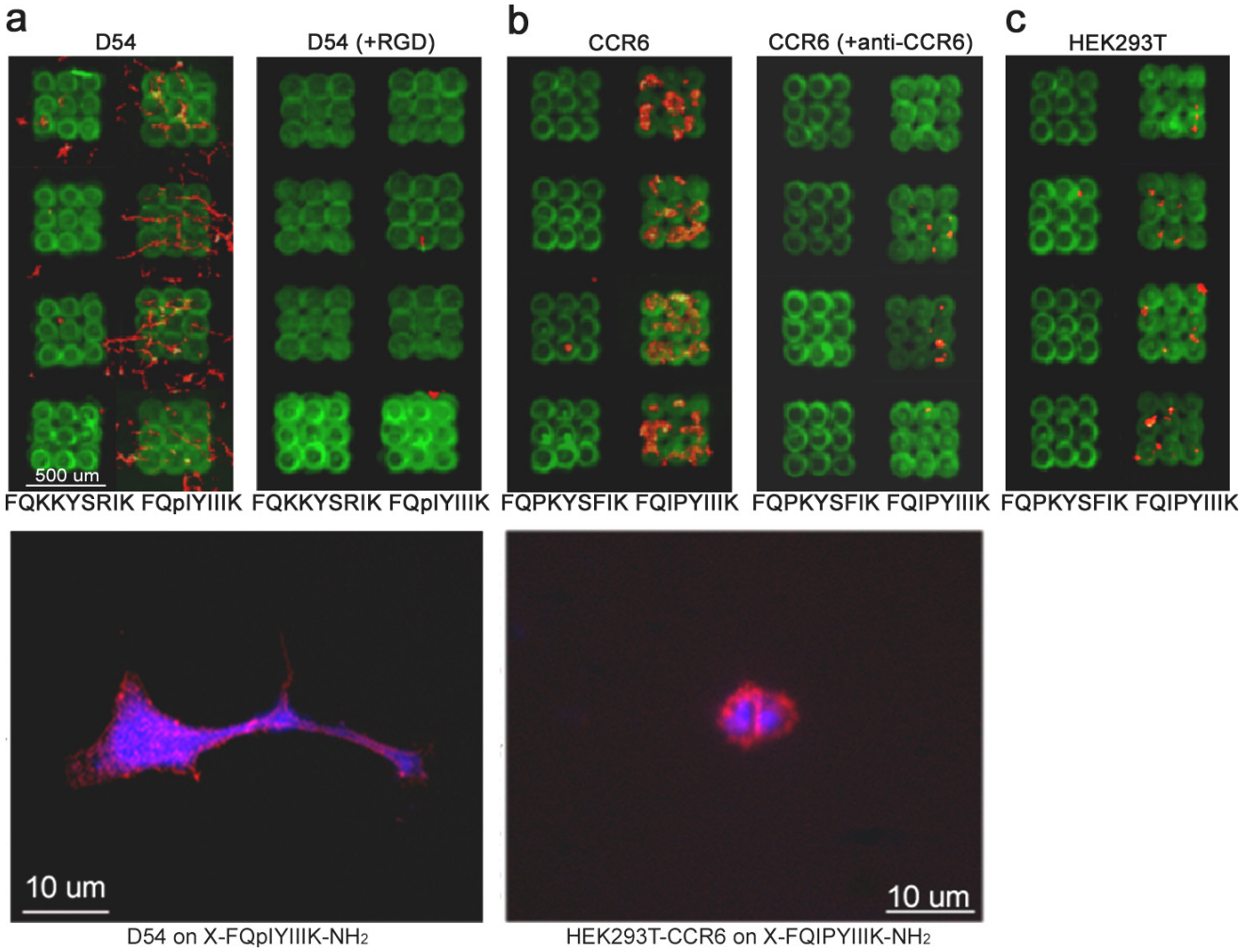

Figure S3: Verification of ligand binding. FITC (green = peptide ligand), DAPI (blue = nuclei staining of attached cells) and phalloidin (red = actin-filament staining of attached cells) visualization of: (**a**) D54 cells incubated on X-FQpIYIIIK-NH_2 and X-FQKKYSRIK-NH_2 surfaces (2 h) and cell growth continued for 16 h (left) and in the presence of soluble X-RGD-NH_2 as an integrin inhibitor (right). (**b**) HEK293T-CCR6 cells incubated on X-FQIPYIIIK-NH_2 and X-FQPKYSFIK-NH_2 surfaces (2 h) and cell growth continued for 16 h (left) and in the presence of soluble anti human CCR6 as a CCR6 inhibitor (right). (**c**) HEK293T cells incubated on X-FQIPYIIIK-NH_2 and X-FQPKYSFIK-NH_2 surfaces (2 h) and cell growth continued for 16 h. The cells were fixed (paraformaldehyde), the actin-filaments stained with phalloidin (red) and the nuclei stained with DAPI (blue).

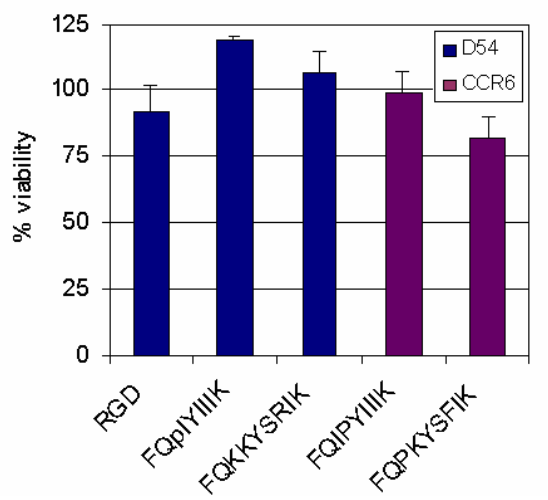

Figure S4: Cytotoxicity of hit-peptides assessed as cell viability by MTT assays. Cells were incubated with 100 μM delivery agent (24 h) the media was changed and the incubation continued for 24 h and absorbance was measured at 580 nm. Untreated cells are assumed to be 100% variable. Error bars indicate ± s.d. (n = 3).

References

1. C. Li, G. C. Tseng, W. H. Wong *Statistical analysis of gene expression microarray data; CRS Press* **2003**, 1-34.
2. N. Keiding, B. Morgan, T. Speed, P. van der Heijden *Interdisciplinary Statistics Series; CRC Press* **2008**.
3. N. Svensen, J. J. Díaz-Mochón, M. Bradley, *Tetrahedron Lett.* 2008, **49**, 6498.
4. J. J. Diaz-Mochon, L. Bialy, L. Keinicke, M. Bradley, *Chem. Commun.* 2005, **11**, 1384.