Generating Encoded Compound Libraries for Fabricating Microarrays as a High-throughput Protein Ligand Discovery Platform

Supporting Information

Volume ratios of partitioned layers were experimentally determined by peptide sequencing analysis of amino acids Leu, Nle and Ile, which had been conjugated to the outer layer, the intermediate layer and the interior core, respectively. **Figure S1** shows the HPLC trace of those residues cleaved from trilayered beads and the mole ratio of Leu, Nle and Ile was determined to be 1.8:1.2:1.0, very close to the experimental design (2:1:1 or 50%:25%:25%) by the concentrations of protecting groups (0.5 equiv. Fmoc and 0.25 equiv. Alloc) used to protect individual layers.

Branched poly-lysine "handles" compete effectively with amine residues on peptide compounds for site-specific immobilization. The second microarray (Microarray B) consisted of HA-K(K(K(NH_2)$_2$)$_2$)-SH, c-Myc K(K(K(NH_2)$_2$)$_2$)-SH, HA-K(K(K(NHAc)$_2$)$_2$)-SH, and c-Myc-K(K(K(NHAc)$_2$)$_2$)-SH. It was designed to compare the immobilization efficiency of peptides with branched poly-lysine handles to that of the same peptides but with acetylated amine residues on the poly-lysine. c-Myc has a free lysine side chain and a free N-terminal amino group. HA has only a free amino group on the N-terminal. The peptide control compounds were released from the beads with 10 μL of 2 mM TCEP in ddH_2O. The concentrations of the peptide solutions were estimated to be 60 μM. Each peptide solution was printed in quadruplet on an epoxy functionalized glass slide to form Microarray B.

Microarray B was incubated in a solution of mouse anti-c-Myc IgG antibody (Sigma-Aldrich, St. Louis, MO) at 200 nM in 1× PBS for 1 hour. **Figure S2(a)** shows the

change in OI-RD image. The anti-c-Myc antibody molecules bound to both the c-Myc with branched poly-lysine and the c-Myc with branched acetylated poly-lysines. However the antibody molecules captured by c-Myc with unblocked poly-lysines covered 1/3 of the printed area (assuming that IgG molecules lay on their sides on the surface, i.e., in side-on geometry), while the antibody molecules captured by c-Myc with acetylated poly-lysines only covered 1/8 of the printed area. It means that the branched poly-lysine handles on the c-Myc increase the number of functionally intact c-Myc molecules on an epoxy-coated glass surface by almost a factor of 3.

The same microarray was subsequently incubated in a solution of mouse anti-HA IgG antibody (Sigma-Aldrich, St. Louis, MO) at 4.4 nM in 1× PBS for 1 hour. **Figure S2(b)** shows the change in OI-RD image after incubation. Again mouse anti-HA IgG reacted with both the HA with a poly-lysines anchor and the HA with acetylated poly-lysines. Yet the amount of antibody molecules captured by HA with unblocked poly-lysines was 50% more than those captured by HA with acetyl-blocked poly-lysines.

These results show that branched poly-lysine residues on peptide compounds enhance immobilization of functionally intact peptides on epoxy-coated glass surface by many folds such that protein probes captured by these intact peptides cover a significant fraction of the printed surface. Clearly immobilization through free amino groups at the N-terminus or on the side chains of peptide compounds instead is either inefficient so that far fewer peptides become immobilized (in the case of HA) or is such that only a small fraction of immobilized peptides retain accessible affinity to solution-phase protein probes (in the case of c-Myc).

Optimal buffer for releasing compounds from beads and subsequent printing on epoxy-functionalized glass surface. The third microarray (Microarray C) was designed to explore the cleaving buffer for optimal compound release and surface immolization. It consisted of biotin-K(K(K(NH_2)$_2$)$_2$)-SH cleaved in 9 different buffers as follows: (1) 2 mM

TCEP in H_2O; (2) 2 mM TCEP and 100 mM $NaHCO_3$ in H_2O; (3) 2 mM TCEP in DMF; (4)2 mM TCEP and 2 mM DIEA in a mixture of DMF and H_2O (9:1 v/v); (5) 2 mM TCEP and 8 mM DIEA in a mixture of DMF and H_2O (9:1 v/v); (6) 2 mM TCEP and 20 mM DIEA in a mixture of DMF and H_2O (9:1 v/v); (7) 2 mM TCEP and 100 mM $NaHCO_3$ in a mixture of DMF and H_2O (7:3 v/v); (8) 2 mM TCEP and 100 mM $NaHCO_3$ in a mixture of DMF and H_2O (**1:1 v/v**); (9) 2 mM TCEP and 100 mM $NaHCO_3$ in a mixture of DMF with H_2O (3:7 v/v). We chose $NaHCO_3$ and DIEA to increase the pH for potential improvement in efficiency of amine-epoxy reactions on the glass surface. We adjusted the volume ratio of DMF to H_2O in order to better dissolve organic molecules. We deposited 10 μL of each buffer in a well that contains one OBOC bead so that the concentration of biotin-K(K(K(NH_2)$_2$)$_2$)-SH released from the bead was estimated roughly 75 μM. 10 replica of each solution were printed in a row on an epoxy functionalized surface to form Microarray C.

Figure S3 shows the change in OI-RD image after Microarray C was incubated in a solution of F_{ab} fragments of anti-biotin mouse IgG at 87 nM in 1× PBS for 1 hour. Clearly adding organic base DIEA to the cleaving buffer did not improve the overall release and immobilization efficiency of biotin-K(K(K(NH_2)$_2$)$_2$)-SH, based on the amount of captured mouse IgG fragments. In contrast, adding inorganic base $NaHCO_3$ to the cleaving buffer dramatically improved the release and immobilization efficiency of biotin-K(K(K(NH_2)$_2$)$_2$)-SH so that the amount of captured mouse IgG molecules increased by more than a factor of 3. In fact the captured IgG molecules were more than a full monolayer in side-on geometry, indicating that some of the captured IgG molecules were in end-on geometry. However the cleaving buffer of 2 mM TCEP and 100 mM $NaHCO_3$ in ddH_2O does not wet the epoxy-coated glass surface well. It is important for real-time microarray detection that compound solutions for printing wets the surface of a functionalized solid support so that the immobilized compounds form a uniform layer with a suitably large diameter. Adding DMF to 2 mM TCEP

and 100 mM $NaHCO_3$ in ddH_2O improved the wetting of the cleaving buffer on epoxy-coated glass surface. Among the four buffers with $NaHCO_3$, 2 mM TCEP and 100 mM $NaHCO_3$ in DMF and H_2O at 1:1 (v/v) yielded the most uniform layer of immobilized biotin-K(K(K(NH_2)$_2$)$_2$)-SH and the amount of captured mouse IgG molecules by such a biotin layer fully covered the printed surface in side-on geometry. From this study we picked 2 mM TCEP and 100 mM $NaHCO_3$ in DMF and H_2O at volume ratio of 1:1 as the optimal buffer for subsequent OBOC compound release and printing.

Table S1. 30 amino acids for building blocks X_1 and X_2. The full spelling of the unnatural amino acids is as follows. (1) Acpc: 2-Amino-1-cyclopentane carboxylic acid; (2) Dpr: 2,3-Diaminopropionic acid; (3) Phg: Phenylglycine; (4) Chg: Cyclohexylglycine; (5) Aad: 3-Aminoadipic acid; (6) D-4-Pal: (4-Pyridyl)-D-alanine; (7) D-2-Thi: β-thienyl-D-alanine; (8) Bpa: 4-Benzoyl-phenylalanine; (9) Ach: Aminocyclohexane carboxylic acid; (10) Aic: 2-Aminoindane-2-carboxylic acid; (11) D-HoCit: D-Homocitrulline; (12) Nal-1: β-(1-Naphthyl)-alanine.

Name	M.W. (Da)	Name	M.W.	Name	M.W.	Name	M.W.
Gly	297.3	His	619.7	D-Thr	397.5	Aic	399.44
Ala	311.3	Chg	379.45	Ile	353.4	Phe (4-Me)	401.45
Acpc	323.34	Aad	439.4	Asn	596.7	Tyr	459.6
Dpr	426.45	D-Phe	387.4	Asp	411.5	D-HoCit	411.4
Ser	383.4	D-4-Pal	388.42	Ach	365.42	Tyr (Me)	417.45
Pro	337.4	D-2-Thi	393.46	D-Gln	610.7	Trp	526.6
Val	339.4	Arg	662.8	D-Glu	425.5	Nal-1	437.49
Phg	373.4	Bpa	491.53				

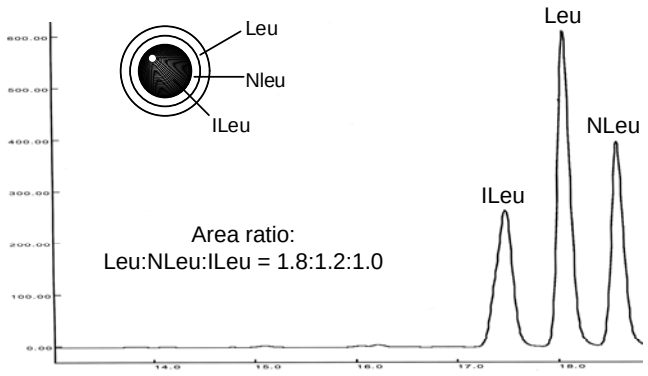


Figure S1. HPLC trace of the amino acid residues cleaved from trilayered beads via Edman degradation in peptide sequencing analysis. The trilayered beads were conjugated with three different amino acids, respectively (the outer layer: Leucine; the intermediate layer: NorLeucine (Nle); the interior core: Isoleucine (Ile)).

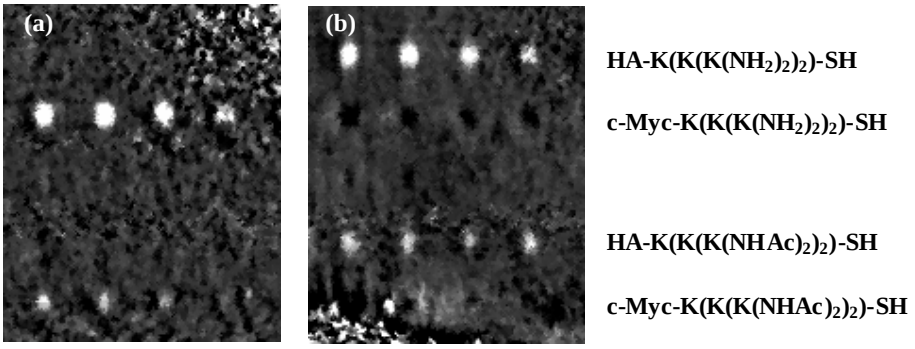
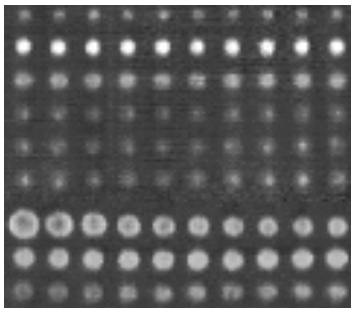


Figure S2. Change in OI-RD image of Microarray B, consisting of HA-K(K(K(NH_2)$_2$)$_2$)-SH, c-Myc-K(K(K(NH_2)$_2$)$_2$)-SH, HA- K(K(K(NHAc)$_2$)$_2$)-SH and c-Myc-K(K(K(NHAc)$_2$)$_2$)-SH, after incubation in (a) mouse anti-c-Myc IgG at 200 nM in 1× PBS for 1 hour; and then (b) in mouse anti-HA IgG at 4.4 nM in 1× PBS for 1 hour. HA is a peptide of sequence YPYDVPDYA with only one free amino group at the N-terminus. c-Myc is a peptide of sequence EQKLISEEDL with two free amino groups, one at the N-terminus and the other on the side chain of the lysine residue. All four columns in both (a) and (b) were printed at the same target concentration of 60 μM.



2 mM TCEP in H_2O
2 mM TCEP and 100 mM $NaHCO_3$ in H_2O
2 mM TCEP in DMF
2 mM TCEP and 2 mM DIEA in DMF/ H_2O (9:1 v/v)
2 mM TCEP and 8 mM DIEA in DMF/ H_2O (9:1 v/v)
2 mM TCEP and 20 mM DIEA in DMF/ H_2O (9:1 v/v)
2 mM TCEP and 100 mM $NaHCO_3$ in DMF: H_2O (7:3 v/v)
2 mM TCEP and 100 mM $NaHCO_3$ in DMF: H_2O (5:5 v/v)
2 mM TCEP and 100 mM $NaHCO_3$ in DMF: H_2O (3:7 v/v)

Figure S3. Change in OI-RD image of Microarray C, consisting of biotin-K(K(K$(NH_2)_2)_2$)-SH in 9 different cleaving buffers, after incubation in a solution of F_{ab} fragments of mouse anti-biotin IgG at 87 nM in 1× PBS for 1 hour. All 10 columns are printed at the same biotin-K(K(K$(NH_2)_2)_2$)-SH concentration of 75 μM.

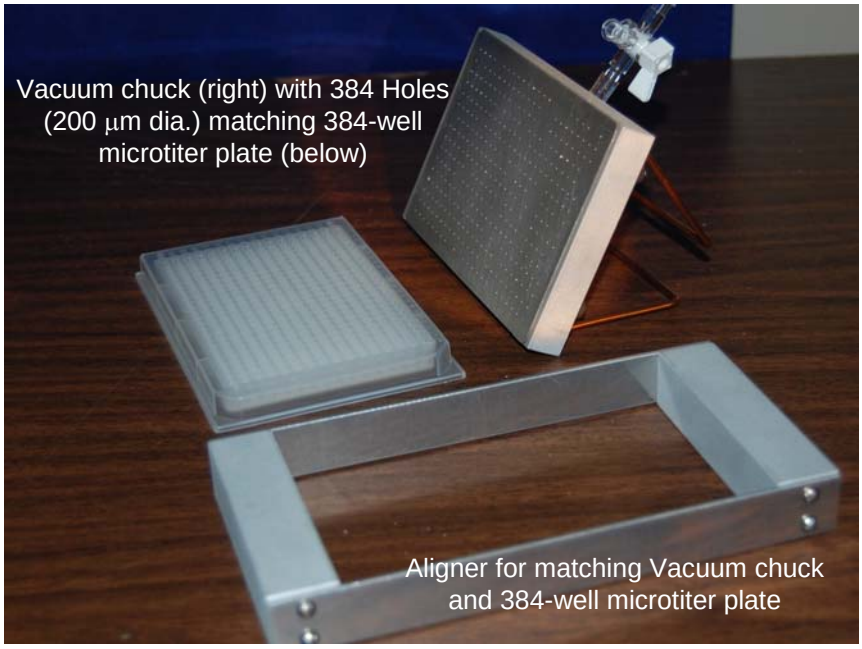


Figure S4. A single-bead placement device for distributing 384 beads into 384 wells of a microtiter plate. It contains a vacuum chuck with 384 drilled holes of 200 μm diameter for picking up and releasing beads simultaneously.