**Supporting Information**

**Highly Sensitive In Vitro Selections for DNA-Linked Synthetic Small Molecules with Protein Binding Affinity and Specificity**

Jeffrey B. Doyon, Thomas M. Snyder, and David R. Liu*

Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138

**DNA synthesis and linker choice.** DNA oligonucleotides were synthesized on a PerSeptive Biosystems Expedite 8090 DNA synthesizer using standard phosphoramidite protocols. All reagents were purchased from Glen Research. The templates for the glutathione S-transferase (GST) selection were synthesized using a 5’-amino-modifier C12 and all other templates were synthesized using 5’-amino-modifier C5. A poly (ethylene glycol) (PEG) linker was used to connect the small molecules and DNA oligonucleotides in all but two cases as detailed below because this linker was found to offer the best combination of coupling efficiencies and low noncovalent association between small molecules and beads. In the two exceptions, linkers were chosen which closely mimic the linkers used in DNA-templated small molecule libraries generated in our group.

**Preparation of (1).** Glutathione was synthesized on the solid phase using standard Boc chemistry at room temperature. 200 mg PAM Resin (Advanced ChemTech) was swelled in 2 mL N,N-dimethylformamide (DMF) for 20 min. N-Boc-glycine (Sigma, 640 µmol, 112 mg), diisoproylcarbodiimide (570 µmol, 89 µL), and 4-dimethylaminopyridine (DMAP, 57 µmol, 7 mg) were added to the resin and stirred for 4 h. The resin was washed with DMF and then with DMF/CH₂Cl₂ (1:1). The N-Boc protecting group was removed using two 3 min washes of trifluoroacetic acid (TFA):m-cresol (95:5). The resin was then washed with DMF:CH₂Cl₂ (1:1) and DMF:pyridine (1:1). A solution of N-Boc-Cys(Fm)-OH (ChemImpex, 800 µmol, 320 mg), O-(7-Azabenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (Aldrich, 720 µmol, 274 mg), 2,6-lutidine (1.2 mmol, 131 µL) and N,N-diisopropylethylamine (DIPEA, 750 µmol, 131 µL) in 800 µL of 1-methyl-2-pyrrolidinone was stirred for 15 min and then added to the resin, stirring for 30 min. The resin was washed with DMF/CH₂Cl₂ (1:1). To remove the N-Boc protecting group on cysteine, a solution of TMS-OTf (2.8 mmol, 0.5 mL) and 2,6-lutidine (4.58 mmol, 0.5 mL) in 1.75 mL CH₂Cl₂ was added to the resin and stirred for 1 h. The resin was washed with methanol and then with DMF:CH₂Cl₂ (1:1). Fmoc-Glu-OFm (ChemImpex, 800 µmol, 438 mg) was coupled as described above. The fully protected glutathione was cleaved from the resin with a solution of trifluoromethanesulfonic acid:m-cresol:thioanisole:TFA (2:1:1:8), stirring for 1 h. The mixture was filtered and the filtrate was extracted into hexane. The crude extract was purified using preparative thin layer chromatography in hexane. The silica containing the crude product (Rf = 0.35) was washed extensively with hexane:ethyl acetate (4:1). The filtrate was isolated under vacuum to afford a yellowish solid. Yields for this synthesis were not optimized.

A solution of protected glutathione (1.1 µmol, 4 mg) in 90 µL DMF with N-hydroxysuccinimide (NHS, 11 µmol, 1.3 mg), dicyclohexylcarbodiimide (DCC, 11 µmol, 2.3 mg), and DMAP (5.7 µmol, 0.7 mg) was agitated for 1 h. The mixture was spun down and the supernatant was added to 5’-amino-terminated protected DNA on CPG beads. This mixture was agitated for 2 h and then the beads were washed with DMF, with CH₃CN, and dried with nitrogen.
Preparation of (2a). N-formyl-Met-Leu-Phe (MLF) was purchased from Sigma and coupled to 5’-amino-terminated protected DNA on CPG beads using the conditions described for (1).

Preparation of (2b). MLF (10-100 µmol, 0.17 M) was dissolved in dry DMF with 1 equiv. 1-hydroxybenzotriazole (Novabiochem), 0.9 equiv. O-Benzotriazol-1-yl-N,N,N',N'-tetramethyluronium hexafluorophosphate (Aldrich), and 2.3 equiv. DIPEA. The solution was agitated at room temperature for 1 h and then added to a unique sequence of 5’-amino-terminated protected DNA on CPG beads. The mixture is agitated for 1 h at room temperature. The beads were washed with DMF, then with CH₃CN, and dried under nitrogen.

Preparation of (3). Fmoc-Lys(Mmt)-OH (Novabiochem) was attached to amino-terminated protected DNA on CPG beads using the method described for (2b). The Fmoc group was removed with three 2 min washes with 20% piperidine in DMF. The mixture was washed with DMF and then with CH₃CN. The α-amine was then capped with a solution of 5% 1-methylimidazole in acetic anhydride/pyridine/tetrahydrofuran (1:1:18) for 10 min at room temperature. The beads were washed with DMF and CH₃CN, then were treated with 3% trichloroacetic acid, 1% thioanisole in CH₂Cl₂ for 5 min at room temperature to remove the Mmt protecting group. The mixture was washed with CH₃CN and dried with nitrogen. Fmoc-Phg-OH (Novabiochem) was attached to the ε-amine of the Lys-linked DNA using the method described for (2b). After removal of the Fmoc protecting group, 4-carboxybenzenesulfonylamide (Aldrich) was attached to the beads using the method described for (2b). The beads were washed with DMF, then with CH₃CN, and dried with nitrogen.

Preparation of (4a, 4b). A 5’- biotin modified phosphoramidite (Glen Research) was used as the final monomer in the DNA synthesis.

Preparation of (5). Chymostatin (Sigma) was attached to amino-terminated protected DNA on CPG beads using the conditions described for (2b).

Preparation of (6). Antipain (Sigma, 1.5 µmol, 0.9 mg) was added to a 30 µL solution of 300 mM DCC and 300 mM NHS in DMF. After agitating for 1 h at room temperature, this solution was added to 45 µL of 5’-amino terminated DNA (~200-300 µM) in 0.1 M MES buffer pH 6.0. This DNA had previous been cleaved from the CPG beads and purified by HPLC as described in the next section. After 2 h, this solution was purified by gel filtration using Sephadex G-25 followed by reverse-phase HPLC.

Characterization of DNA-linked synthetic molecules. Small molecule-DNA conjugates were cleaved from the CPG beads with solution of methylamine: ammonium hydroxide (1:1) at 55 °C for 1 h. The solution was dried under vacuum and then purified by reverse phase HPLC using a TEAA/CH₃CN gradient and analyzed by MALDI-TOF mass spectrometry. Stock solution concentrations were determined using UV-Vis spectroscopy and serial dilutions were prepared for the selection experiments. Samples were stored in water at -20 °C.

Preparation of immobilized target proteins. NHS-activated Sepharose 4 Fast Flow (Amersham Pharmacia) was prepared according to the manufacturer’s protocol. Equine GST,
bovine carbonic anhydrase (CA), papain, MLCK-treated bovine chymotrypsin, and TPCK-treated bovine trypsin were purchased from Sigma. Typically, proteins were dissolved in PBS buffer pH 7.4-7.6 at concentrations of 20-100 µM. Protein concentrations were determined using UV-Vis spectrometry. Proteins were incubated with beads for 16 h at 4 °C. Beads were capped for 2 h with Tris buffer, washed extensively with the appropriate selection buffer containing 1 M NaCl, and then exchanged into the appropriate selection. Beads were stored for up to 1 month at 4 °C in a volume of selection buffer equal to the initial volume of beads used. Before use, papain beads were activated using a solution of 5.5 mM cysteine•HCl, 1.1 mM EDTA, and 0.067 mM β-mercaptoethanol for 30 min at 4 °C. Streptavidin magnetic particles (Roche) were washed 3x with selection buffer before use.

**Selection Buffers.**

<table>
<thead>
<tr>
<th>Selection</th>
<th>Buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>GST</td>
<td>10 mM Tris pH 7.4, 0.1 M NaCl</td>
</tr>
<tr>
<td>CA</td>
<td>10 mM Tris pH 7.4, 0.1 M NaCl, 1 mM EDTA</td>
</tr>
<tr>
<td>Papain</td>
<td>50 mM Tris pH 7.4, 0.1 M NaCl, 10 mM CaCl₂</td>
</tr>
<tr>
<td>Trypsin</td>
<td>50 mM Tris pH 8.0, 0.1 M NaCl, 10 mM CaCl₂</td>
</tr>
<tr>
<td>Chymotrypsin</td>
<td>50 mM Tris pH 8.0, 0.1 M NaCl, 10 mM CaCl₂</td>
</tr>
<tr>
<td>Streptavidin</td>
<td>10 mM Tris pH 7.4, 0.1 M NaCl, 1 mM EDTA</td>
</tr>
</tbody>
</table>

**GST selection.** The amount of (1), the binding ligand, was varied between $10^3$ and $10^7$ molecules and (2a), the non-binding ligand, was used in $10^2$-$10^6$ molar excess. (1) and (2a) were added to 40 µL of GST beads and agitated at 4 °C for 1 h. The mixture was transferred to a 5.0 µm low-binding Durapore membrane spin filter (Millipore), washed with 2x 150 µL PBS pH 7.4, 1x 100 µL 0.1 M Tris pH 8.0, 0.5 M NaCl, and 1x150 µL PBS. The bound ligands were eluted by agitating the beads with 100µL 0.1 M glutathione (Sigma) at room temperature. The eluant was ethanol precipitated with 3 M sodium acetate and 1µL glycogen. The precipitate was used directly for PCR (see below).

**Carbonic anhydrase selection.** (2b), the non-binding ligand, and (3), the binding ligand, were added to 40µL of resuspended CA-linked beads and were diluted to 400 µL with selection buffer. Ratios were similar to those for the GST selection. The mixture was agitated at 4 °C for 1-2 h. Selections were then carried out at room temperature. Each mixture was transferred to a spin filter and washed 3x with 400 µL of wash buffer and 1x 400 µL with selection buffer. The resin was removed from the spin filter with 60 µL of selection buffer and the resulting beads were subjected to PCR.

**Papain selection.** (4a), the non-binding ligand, and (5) or (6), the binding ligands, were incubated with papain beads and selected as described for the carbonic anhydrase selection.

**Chymotrypsin selection.** (4a), the non-binding ligand, and (5), the binding ligand, were incubated with chymotrypsin beads and selected as described for the carbonic anhydrase selection.

**Trypsin selection.** (4a), the non-binding ligand, and (6), the binding ligand, were incubated with trypsin beads and selected as described for carbonic anhydrase.
Streptavidin selection. (3), the non-binding ligand, and (4b), the binding ligand, were incubated with 15 µL streptavidin magnetic particles and agitated at room temperature for 20 min. Using a MPC-S magnet (Dynal), the beads were washed 2x with 0.1 M NaOH, 1 mM EDTA (100-200 µL), 4x with wash buffer (100-200 µL), and 1x with selection buffer. Beads were resuspended in 15 µL dd H\(_2\)O.

Interestingly, the observed enrichment factors do not necessarily correlate with affinity. This lack of correlation suggests that the effective concentration of protein in each of the above selections exceeded the \(K_d\) of the relevant ligand-protein complexes. We speculate that the general hydrophobicity and flexibility of ligand binding sites within our protein targets causes some proteins to possess higher background binding to our negative controls than other proteins, regardless of their affinity for the ligands used in this study.

Wash Buffers.
Carbonic anhydrase: 10 mM Tris pH 7.4, 0.25-0.5 M NaCl
Papain: 50 mM Tris pH 7.4, 0.5 M NaCl, 1 mM EDTA
Chymotrypsin: 50 mM Tris pH 8.0, 0.5 M NaCl, 10 mM CaCl\(_2\)
Trypsin: 50 mM Tris pH 8.0, 0.5 M NaCl, 10 mM CaCl\(_2\)
Streptavidin: 10 mM Tris pH 7.4, 1.0 M NaCl, 1 mM EDTA

Iterated carbonic anhydrase selection. 10\(^8\) molecules of (3) and 10\(^{11}\) molecules of (2b) were incubated with 40 µL carbonic anhydrase beads for 1 h and then selected as described. After the first round of selection, 5 µL of resuspended agarose beads were removed for PCR. 6 M guanidinium\(\cdot\)HCl, 10 mM EDTA (40 µL) was added to the beads and the mixture was heated to 90 °C for 15 min. The beads were filtered away using a Wizard Minicolumn (Promega). The filtrate was buffer exchanged into selection buffer using a Centrisep Spin Column (Princeton Separations). A new aliquot of carbonic anhydrase beads was added to the eluted templates. After a second round of selection, the agarose beads were suspended in 30 µL of H\(_2\)O and 15 µL were used for PCR. The PCR products were digested with Hind III, generating the results below.

The triple iteration selection was carried out essentially as described above with a few minor changes. The prepared CA beads were incubated with ZnSO\(_4\) (1 mM) for 1 h and then washed
extensively with selection buffer containing 2 M NaCl. The beads were exchanged back into selection buffer and used directly for the iterated selection. $10^9$ molecules of (3) and $10^{15}$ molecules of (4b) were added to the beads and selected as described above. After the first round of selection, 3 µL aliquot was removed for PCR. A second round of selection was carried out as described above and an 8 µL aliquot of beads was removed for PCR. After a third round of selection, the resulting beads were removed from the spin filter using 30 µL of dd H$_2$O and 15 µL of resuspended beads were used for PCR. The results following Hind III digestion are shown below:

![Diagram of molecular structures and selection results](image)

**Papain affinity and papain specificity selections.**

**Affinity selection:** 6x$10^9$ molecules of (6), 2.3x$10^{10}$ molecules of (5), and 1.4x$10^{11}$ molecules of (4a) were added to 40 µL papain beads for 1 h. The beads were washed with papain wash buffer (3 x 100 µL) and once with 100 µL papain selection buffer. The beads were removed from the spin filter with 30 µL of dd H$_2$O. A 3 µL aliquot of resuspended beads was removed for PCR. The DNA conjugates were eluted from the beads by adding 70 µL 6 M guanidinium•HCl and heating the mixture to 90°C for 15 min. The eluted material was buffer exchanged as described in the iterated carbonic anhydrase selection. After a second round of selection, the agarose beads were removed from the spin filter using 30 µL H$_2$O and 15 µL of resuspended beads were used for PCR.

**Specificity selection:** The amounts of antipain, chymostatin, and biotin listed above were added to 40 µL chymotrypsin agarose beads in chymotrypsin selection buffer and incubated for 1 h. The beads were spun down and the flow through was added to 40 µL fresh chymotrypsin beads and incubated for 1 h. The beads were spun down and 15 µL of 100 µM chymotrypsin in papain selection buffer was added to the flow through and then incubated for 1 h. This solution was
added to 40 µL of papain beads and selected as described above. The small molecule-DNA conjugates were eluted and buffer exchanged as described, incubated with 15 µL 100 µM chymotrypsin for 1 h and then subjected to a second round of selection. The beads were removed from the spin filter with 30 µL of H₂O and 15 µL were used for PCR. The results following restriction digestion with the appropriate enzyme are shown below.

**Contamination controls.** Due to the high sensitivity of these experiments, two important contamination controls were used throughout these studies. First, each selection was carried out as described above except no ligand-DNA conjugates were added to the protein-linked beads, which allowed us to test for buffer contamination and any cross-contamination among samples. Secondly, a PCR reaction in which no material from the selection was added was used to test for contamination in primers, dNTPs, and PCR buffers.

**PCR conditions, gel electrophoresis, and relative DNA quantitation.** Templates surviving the selection were amplified using PCR. All reactions contained 1 µM of each primer and 250 µM of each dNTP (Promega). For the GST selection, the precipitated DNA was used in the PCR
reaction and amplified with Platinum Taq (Invitrogen). PCR conditions were step 1: 94 °C, 2'; step 2: 94 °C, 30 s; step 3: 55 °C, 1'; step 4: 72 °C, 30 s; step 5: go to step 2 x 29; step 6: 72 °C, 5'; step 7: hold at 4 °C. For all other selections, the agarose beads (3-15 µL) were used directly in the PCR reaction with Taq polymerase (Promega). PCR conditions were step 1: 94 °C, 2’ step 2: 94 °C, 30s; step 3: 55 °C, 1'; step 4: 72 °C, 30 s; step 5: go to step 2, x 24; step 6: 4 °C.

PCR products were digested for 1-2 h with the restriction enzymes (New England Biolabs, 5-10 units) that digest the ligand-encoding DNA. Digestion products were analyzed by electrophoresis on 3% agarose gels and quantitated by ethidium bromide staining and densitometry on a Stratagene Eagle Eye II system.

**Enrichment calculations.** Enrichment ratios are calculated as the ratio of the fraction of binding ligand surviving the selection as determined by restriction digestion to the fraction of binding ligand entering the selection as determined by the known concentrations of the stock solutions. The fraction of binding ligand relative to total DNA was calculated in moles (correcting for changes in molecular weight that resulted from restriction digestion). PCR amplification of mixtures of templates can favor those templates that exist in lowest abundance at the start of the PCR reaction. This phenomenon, referred to as dynamic range compression PCR, likely benefits enrichments arising from large excesses of non-ligand and is documented in Chapter 5 of Urbach, J. M. *Ph.D. Thesis* 1999, Harvard University, Advisor: J. Szostak.

**DNA sequences of templates and primers.**
Restriction endonuclease cleavage sites are underlined.
DNA sequences for GST selections:
GSH-template (1): 5'- GCC TCT GCG ACC GTT CGG AAG CTT CGC GAG TTG CCC AGC GCG *(Hind III)*
MLF-template (2a): 5’- GCC TCT GCG ACC GTT CGG GAA TTC CGC GAG TTG CCC AGC GCG *(Eco RI)*
Primer 1: 5’-GCC TCT GCG ACC GTT CGG
Primer 2: 5’-CGC GCT GGG CAA CTC GCG

DNA sequences for CA selections:
Phenyl sulfonamide-template (3): 5’-CGA TGC TAG CGA AGG AAG CTT CCA CTG CAC GTC TGC *(Hind III)*
MLF-template (2b): 5’-CGA TGC TAG CGA AGG GAA TTC CCA CTG CAC GTC TGC *(Eco RI)*
Biotin-template (4b): 5’-CGA TGC TAG CGA AGG GAA TTC CCA CTG CAC GTC TGC *(Eco RI)*
Primer 1: 5’-CGA TGC TAG CGA AGG
Primer 2: 5’-GCA GAC GTG CAG TGG

DNA sequences for protease selections:
Chymostatin-template (5): 5’-GCA GTC GAC TCG ACC GAG TCT CCA CTG CAC GTC TGC *(Bam HI)*
Antipain-template (6): 5’-GCA GTC GAC TCG ACC CAG CTG GGC TAC GAC GTG CAC *(Pvu II)*
Biotin-template (4a): 5’-GCA GTC GAC TCG ACC AAG CTT GGC TAC GAC GTG CAC
(\textit{Hind} III)
Primer 1: 5’-GCA GTC GAC TCG ACC
Primer 2: 5’-GTG CAC GTC GTA GCC

Complete structures of synthetic groups linked to DNA.