

Supplementary Information

Synthesis and Screening of A DNA-Encoded Library of Non-Peptidic Macrocycles

Eric Koesema^{1†}, Animesh Roy^{1†}, Nicholas G. Paciaroni¹, Carlos Coito¹, and Thomas Kodadek^{1,2*}

¹Deluge Biotechnologies, 6671 W. Indiantown Rd., Suite 50-325, Jupiter, FL 33458.

²Department of Chemistry, The Scripps Research Institute, Jupiter, FL 33458.

Section I - Methods

General Synthetic Methods

The methods used for solid-phase peptoid synthesis and split-and-pool synthesis of solid-phase DELs have been well described previously.^{1,2} For these studies, Linker-1, used for library synthesis and FACS studies, was synthesized on 10 μm TentaGel M NH₂ resin (Rapp Polymere, 0.23 mmol/g). Linker-2, used for mass verification, and Linker-3, used for large-scale synthesis of fluoresceine-conjugated compounds, were synthesized on 160 μm TentaGel MB RAM resin (Rapp Polymere, 0.41 mmol/g). When applicable, HDNA was added via copper-catalyzed Click chemistry, as previously described.² For the library described herein, 60 mg of 10 μm Linker-1 resin was combined with 5 mg of 160 μm Linker-2 resin in a chromatography column (Bio-Rad) fritted with a 10 μm pore size frit (Mo Bi Tec). Following Fmoc deprotection (20% Piperidine in DMF, 2 x 10 min. at 25°C), N- α -Fmoc-S-2,4,6-trimethoxyphenylthio-L-cysteine (5 molar equivalents to bead reactive sites) was activated with DIC/Oxyma/2,4,6-Collidine (7.5/5/5 equivalents, in DMF) for 5 min. at RT prior to resin coupling (3 h, 37°C, 600 rpm). Following another Fmoc deprotection (20% Piperidine in DMF, 2 x 10 min. at 25°C), 2-(chloromethyl)oxazole-4-carboxylic acid was added using the same conditions. The resin was then aliquoted equally into 96 wells of a pre-wetted (DCM, 100 μl) hydrophobic MultiScreen Solvinert 96 Well Filter Plate (Millipore). Library synthesis occurred via successive iterations of amine displacement, chloroacid addition, and encoding DNA ligation, as previously described.¹ Following the final chloroacid addition, the STmp group of the cysteine is removed (100 mM DTT, 100 mM N-methylmorpholine, in DMF, 2 x 5 min. at 25°C, with agitation), and macrocyclization occurs via thioether formation (2X PBS, 37°C, 16 hours, with agitation).

Screening

10 copies of the library (~3 mg) were added to each of two microcolumns fritted with a large 10 μm pore size frit (Mo Bi Tec). The beads were then washed with (3x 500 μl) and equilibrated in PBST (RT, 18 hours, rotation), then blocked with ice-cold 50% Starting Block (Thermo Scientific) in PBST (4°C, 18 hours). Prior to the screens, the beads were washed thrice more with cold 50% Starting Block (Thermo Scientific) in PBST (3 x 500 μl). To one column, 100 nM A647-conjugated streptavidin (Molecular Probes), 1000 nM human IgG (Sigma, I4506), and 10 $\mu\text{g}/\text{ml}$ human Cot-I DNA (Invitrogen) was added, in cold PBST. To the other column, 100 nM A488-conjugated streptavidin (Molecular Probes), 1000 nM human IgG (Sigma, I4506), and 10 $\mu\text{g}/\text{ml}$ human Cot-I DNA (Invitrogen) was added, in cold PBST. The columns were then incubated for 18 hours at 4°C, with gentle agitation. The beads were then washed with cold 50% Starting Block in PBST (3 x 500 μl) prior to labeling the IgG. The IgG protein in each column was orthogonally labeled, relative to streptavidin, with either A488- or A647-conjugated anti-human IgG antibody (Invitrogen), at a concentration of 10 ng/ μl (150 μl) in cold 50% Starting Block, for 2 hours at 4°C with rotation. Beads were then washed 3x with cold PBST, followed by a 30 min. incubation in cold PBST at 4°C, with gentle agitation. This was repeated twice, then the beads were resuspended in 300 μl cold PBST and passed through a 35 μm cell strainer into FACS tubes (Falcon). Sorting was performed on a

BD Aria III FACS, collecting beads that clearly shifted away from the double negative (DN) population in either the green or red channels. Individual rescreening of resynthesized hits was done using the same general protocol, without Cot-I DNA, and analyzed on a BD Canto FACS. Data processing was done on FlowJo Software (FlowJo™ Software (for Windows 10) Version 10.6.1. Ashland, OR: Becton, Dickinson and Company; 2019).

NGS Analysis

Beads collected via FACS were wash several times with BTP-WB (10 mM bis-tris propane pH 7.6, 50 mM sodium chloride, 0.04% Tween-20) and transferred to individual 0.2 ml PCR tubes. Sample prep for NGS involves two, sequential PCR steps. For the first PCR, beads were amplified in a 40 µl reaction mix containing 1X SuperFi Buffer, 0.2 mM dNTP mix, 0.3 µM ABM062 primer, 0.3 µM ABM063 primer, and Platinum SuperFi DNA Polymerase (ThermoFisher, 0.02 U/µl). Thermal cyclization was 98°C, 30 s, (98°C, 10 s; 46°C, 10 s; 72°C, 15 s) x 25 cycles. PCR products were run on native PAGE (6% Acrylamide, run in 1X TBE, 110 V, 50 min.), along with Low MW DNA Ladder (NEB, N3233L). Bands running at 167 bp were extracted, placed into 40 µl ddH₂O, and incubated overnight at RT. The extracted DNAs were diluted 1000-fold and used as templates for the second amplifications (40 µl reactions, 1 µl DNA template, 1X SuperFi Buffer, 0.2 mM dNTP mix, 0.15 µM P5 universal forward 2 primer, 0.15 µM NGS barcode reverse primer, and Platinum SuperFi DNA Polymerase (ThermoFisher, 0.02 U/µl). Thermal cyclization was 98°C, 30 s, (98°C, 10 s; 60°C, 10 s; 72°C, 15 s) x 25 cycles. PCR products were run on native PAGE (as above). Bands running at 304 bp were extracted, placed into 40 µl ddH₂O, and incubated overnight at RT. Note, each sample was run on three gels and pooled for NGS analysis. DNAs were assessed by Qubit (ThermoFisher) and a bioanalyzer (Agilent 2100). NGS was done on the Illumina NextSeq 500. NGS decoding and structure elucidation were done as previously reported.² The NGS data are then analyzed for redundant hits (multiple beads displaying the same compound, differentiated only by the bead-specific barcode). Putative hits were selected based on a minimum redundancy of 3 beads in each of the screening replicates.

Hit Resynthesis

30 mg of 10 µm Linker-1 beads and 6 mg 160 µm Linker-2 beads were put into separate Mobicols with 10 µm pore-size frits. N-α-Fmoc-S-2,4,6-trimethoxyphenylthio-L-cysteine (5 molar equivalents to bead reactive sites) was activated with DIC/Oxyma/2,4,6-Collidine (7.5/5/5 equivalents) for 5 min. at RT prior to resin coupling (3 h, 37°C, 600 rpm). Following Fmoc deprotection (20% Piperidine in DMF, 2 x 10 min. at 25°C), 2-(chloromethyl)oxazole-4-carboxylic acid was added using the same conditions. Both the 10 µm and 160 µm resins were aliquoted, separately, into 44 wells of a pre-wetted (DCM, 100 µl) hydrophobic MultiScreen Solvinert 96 Well Filter Plate (Millipore). As above, synthesis occurred via successive iterations of amine displacement and chloroacid addition, but no DNA is ligated during resynthesis. Following the final chloroacid addition, the STmp group of the cysteine is removed (100 mM DTT, 100 mM N-methylmorpholine, in DMF, 2 x 5 min. at 25°C, with agitation), and macrocyclization occurs via thioether formation (2X PBS, 37°C, 16 hours, with agitation). For synthesis of linear analogs, the N-α-Fmoc-S-2,4,6-trimethoxyphenylthio-L-cysteine was replaced with N-Fmoc-L-methionine or N-Fmoc-L-alanine.

Assessment of Macrocyclization by MALDI

Immediately after macrocyclization, the hits resynthesized on 160 µm beads were washed with PBST (5 x 150 µl), then treated with 200 mM benzylbromide (150 µl, in PBST) for 2 hours at 37°C, 600 rpm. Beads were then washed 5 times with PBST, transferred to a polypropylene microtiter plate, washed with ddH₂O (2 x 125 µl/well), washed with 100% EtOH (2 x 150 µl/well) and allowed to air dry. Compounds were released from the resin with 90% TFA, 5% TIPS, 2.5% DCM, 2.5% ethanedithiol (40 µl, RT, 2 h) in a foil tape-sealed microtiter plate. The seal was then removed, and the samples were air dried in a fume hood,

O/N. The released compounds were dissolved in 5 μ l 50% ACN/H₂O, spotted on a MALDI plate with CHCA matrix (Sigma), and analyzed on a Bruker Microflex MALDI-TOF.

Assessment of Macrocyclization by FACS

Immediately after macrocyclization, 0.25 mg of resin was transferred to a new, pre-wetted (DCM, 100 μ l) hydrophobic MultiScreen Solvintert 96 Well Filter Plate (Millipore), washed with PBST (5 x 150 μ l), treated with 3 mM mBBR (150 μ l/well, in 2X PBS) for 2 hours at 37°C, 600 rpm, washed (3 x 150 μ l BTP-WB, 3 x 150 μ l DMF, 3 x 150 μ l DCM, 3 x 150 μ l MeOH, 3 x 150 μ l PBST), and incubated in PBST O/N, RT, 300 rpm. The beads were resuspended in 300 μ l PBST, passed through a 35 μ m cell strainer into FACS tubes, and analyzed on a BD Canto FACS. Data processing was done on FlowJo Software (FlowJo™ Software (for Windows 10) Version 10.6.1. Ashland, OR: Becton, Dickinson and Company; 2019).

On-Bead Validation of Resynthesized Hits

For each hit to be analyzed, 0.25 mg of resin displaying the resynthesized hit were added to each of four mobicols fritted with a large 10 μ m pore size frit (Mo Bi Tec). The beads were then washed with (3x 500 μ l) and equilibrated in PBST (RT, 18 hours, rotation), then blocked with ice-cold 50% Starting Block (Thermo Scientific) in PBST (4°C, 18 hours). The beads were then washed 3x more with cold 50% Starting Block in PBST (3 x 500 μ l). To each column, 500 μ l A647-conjugated streptavidin (Molecular Probes) was added at the appropriate concentration, diluted in 50% Starting Block/PBST. The columns were then incubated for 18 hours at 4°C, with gentle agitation. Beads were then washed 3x with cold PBST, followed by a 30 min. incubation in cold PBST at 4°C, with gentle agitation. This was repeated twice, then the beads were resuspended in 300 μ l cold PBST and passed through a 35 μ m cell strainer into FACS tubes (Falcon). FACS analysis was performed on a BD Canto FACS. Data processing was done on FlowJo Software (FlowJo™ Software (for Windows 10) Version 10.6.1. Ashland, OR: Becton, Dickinson and Company; 2019).

Large-Scale Synthesis

Large-scale syntheses of Linker-3 and fluorescein-conjugated compounds was done using the same general protocol noted above, but with 0.1 g (0.04 nmol) 160 μ m TentaGel MB RAM resin and 3 equivalents of the acid monomers with 4.5/3/3 equivalents of DIC/Oxyma/2,4,6-Collidine. For linear analogs, the N- α -Fmoc-S-2,4,6-trimethoxyphenylthio-L-cysteine was replaced with N-Fmoc-L-alanine, and the terminal chloride was quenched with a solution of 0.5 M ethanethiol/0.5 M diisopropylethylamine (in DMF, incubated at room temperature for 6 hours).

Fluorescein conjugation

Tetrakis(triphenylphosphine)-palladium(0) (94.3 mg, 0.08 mmol, 2 equiv) was suspended in anhydrous DCM (1.5 mL) and mixed with phenylsilane (40.3 μ L, 0.32 mmol, 8 equiv). The mixture was sonicated for two minutes then added to the resin. The reaction was incubated at room temperature for 30 minutes and was gently vortexed every two minutes. The solution was drained, washed with DCM (5 x 2 mL), DMF (3 x 2 mL), 1% diisopropylethylamine in DMF (2 mL), 1% sodium diethyldithiocarbamate trihydrate in DMF (3 x 2 mL), DMF (2 x 2 mL) and DCM (3 x 2 mL). The resin was resubjected to the palladium alloc-deprotection conditions, followed by the washing protocol once more. 5(6)-carboxyfluorescein (46.1 mg, 0.12 mmol, 3 equiv) and oxyma (17.1 mg, 0.12 mmol, 3 equiv) were dissolved in DMF (0.75 mL) and was then added 2,4,6-trimethylpyridine (15.8 μ L, 0.12 mmol, 3 equiv) and diisopropylcarbodiimide (28.0 μ L, 0.18 mmol, 4.5 equiv). The solution was mixed at room temperature for one minute and then added to the resin. The reaction was mixed at 37 °C for 1.5 hours. The resin was washed with DMF (5 x 5 mL).

Resin cleavage and purification

The resin was washed with DCM (5 x 2 mL) and the crude compound was cleaved from the resin using a cleavage cocktail: 49% TFA, 49% DCM, 2% TIPS, 2 mL. The cocktail was added to resin, and the resin/syringe was covered with aluminum foil and was incubated at room temperature for one hour. The solution was dispensed into a 50 mL falcon tube, the resin was washed with DCM (3 x 2 mL) and each wash was pooled together in the tube. The TFA cleavage and DCM washing conditions were repeated once more and all solutions were pooled together, and the tube was covered in aluminum foil. The TFA/DCM solution was evaporated under a stream of argon, and the crude compound was precipitated with cold diethyl ether (15 mL) and pelleted by centrifugation (2 minutes, 3000 rpm). The pellet was resuspended in 50% MeCN/DI H₂O, filtered, and purified by preparative reverse-phase HPLC. Product fractions from the HPLC were analyzed for purity via UPLC. Pure fractions were pooled together, the MeCN was evaporated, the sample was frozen and lyophilized to afford a pure, lyophilized powders. Preparative HPLC was performed on a Waters HPLC using a 2545 binary gradient module equipped with a 2489 UV/vis detector and a X-Bridge BEH C19 OBD 19 x 250 mm prep column. LC-MS analysis was carried out on a Waters Aquity UPLC equipped with a BEH C18 2.1 x 100 mm x 1.7 μ m column, PDA detector, and an ESI SQ Detector 2 mass spectrometer.

Fluorescence Polarization

All FP assays were done in 1X PBS + 0.01% Tween-20 (PBST). Purified ligands were dissolved in PBST, and molarity confirmed by absorbance (495 nm, $\epsilon_{495} = 78,000 \text{ M}^{-1}\text{cm}^{-1}$) using a Nanodrop One^c (Thermo Scientific). Serial dilutions of streptavidin (high of 100 μ M, 2-fold serial dilutions in PBST) were mixed 1:1 with 10 nM ligand (20 μ l final volume), incubated for 30 min. at 22°C, then scanned on a Tecan SPARK plate reader. The K_D values were determined by nonlinear regression with Hill slope, using GraphPad Prism version 9.1.2 for Windows, GraphPad Software, La Jolla California USA. The competition assay was performed by pre-incubating 100 nM streptavidin with 50 nM DBT-3-FL for 5 min, then mixing 1:1 with a dilution series of DBT-3-nonFL (20 μ l final volume), incubating at 22°C for 30 min, then reading on the Tecan SPARK. The K_i was determined by nonlinear regression, One-fit K_i , using GraphPad Prism.

Chloroalkane penetration assay

HEK 293 cells stably expressing the HaloTag protein fused to GFP were cultured in DMEM with 10% FBS (Tet free). Cells were seeded in a 96-well plate 15 hours prior to the start of the experiment at 6×10^4 cells per well. The chloroalkane-linked macrocycles were diluted in the assay media to concentrations ranging from 0.02 to 40 μ M. To start the experiment, media was aspirated and 170 μ L of the media containing the compounds was added. The cells were incubated with the compounds for 3 hours at 37 °C in the presence of 5% CO₂. After that, the solution was removed by aspiration, and the cells were washed with assay media for 30 minutes followed by a 30-minute treatment with 20 μ M chloroalkane-tagged TAMRA dye in the assay media. Next, the cells were washed with media for 30 minutes. The media was removed, and cells were resuspended in 100 μ L DPBS.

The level of fluorescence in the cells was analyzed using LSR II flow cytometer (BD Biosciences) equipped with a BD High Throughput Sampler. The cells were analyzed in an HTS throughput mode using a sample volume of 10 μ L, flow rate of 1 μ L/s, 3 mixing cycles with a volume of 25 μ L, and a mixing speed of 200 μ L/s. These settings yielded an average number of $3\text{-}4 \times 10^3$ GFP-positive cells analyzed per well. The flow cytometry data was analyzed using FlowJo version 10.6.2 software and the CP₅₀ values were determined using GraphPad Prism 9 software.

Section II – Supplementary Figures

Figure S1 – Linker structures.

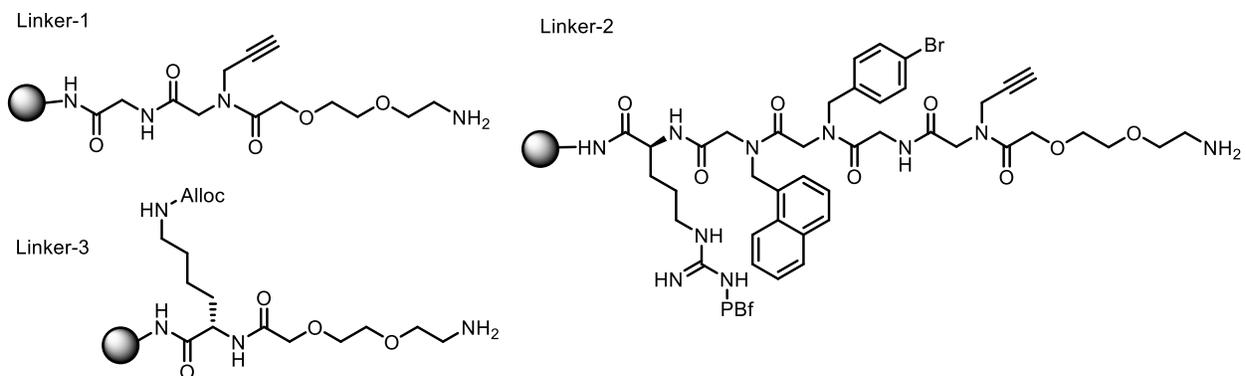
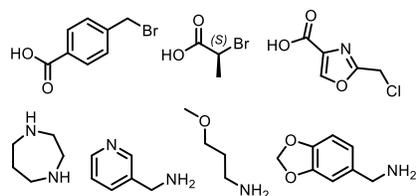
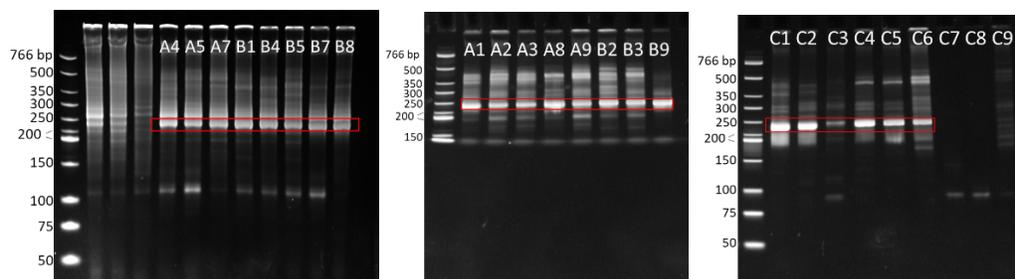


Figure S2 – Mini library synthesis and analysis.

A – A small set of acids and amines were put together in various combinations to form a “mini-library” of 27 compounds.



B – Native PAGE Gels for PCR. Products boxed in red were excised and sequenced.



C – Representative MALDI trace

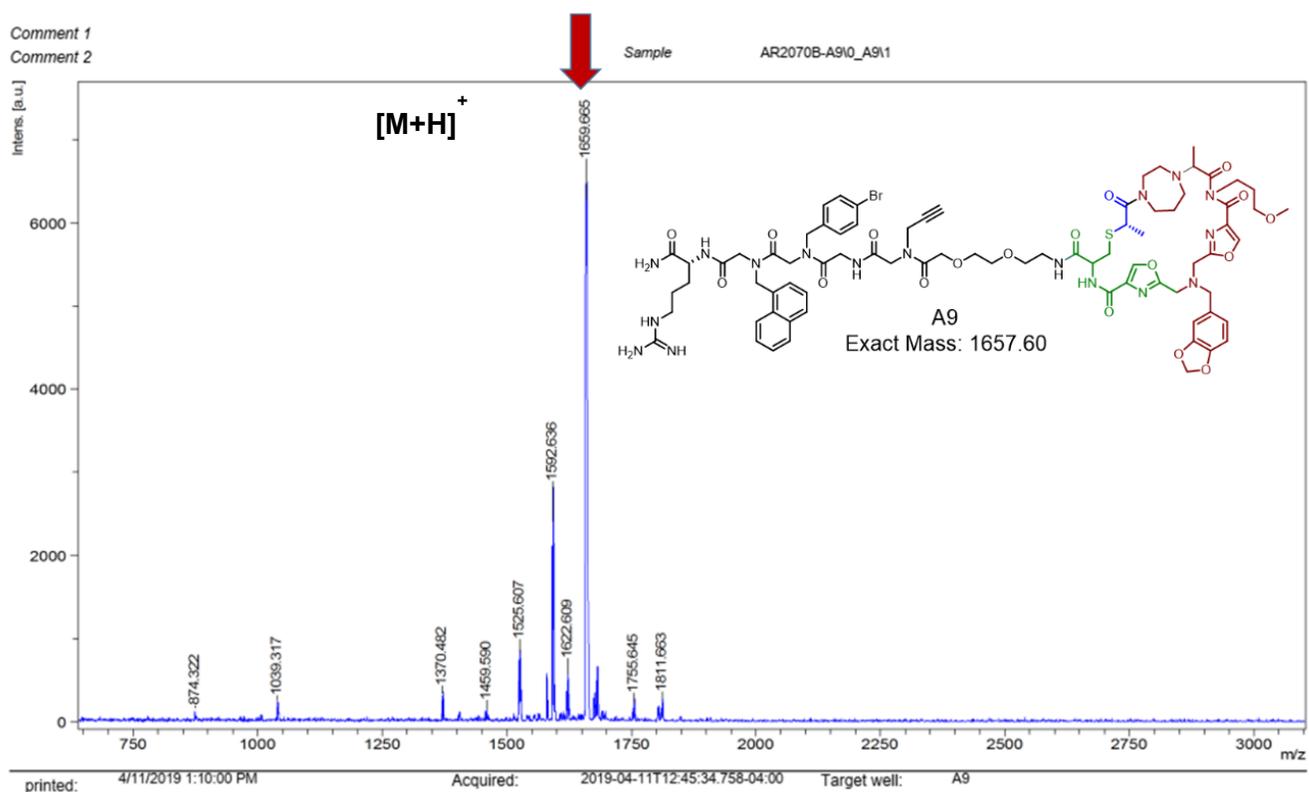
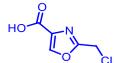
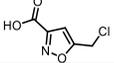
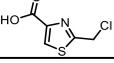
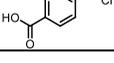
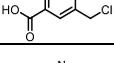
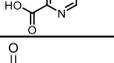
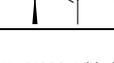
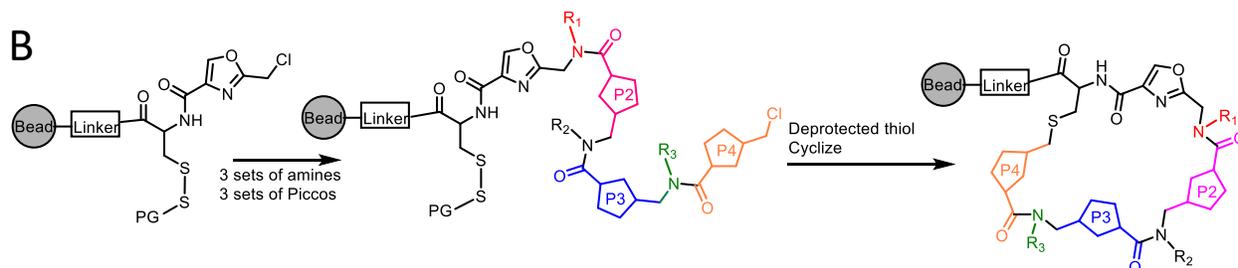


Figure S3 - Generation of “3.5-mer” and “2.5-mer” macrocycles in the same library. (A) During split and pool OBOC library synthesis, beads are aliquoted into each well of a 96-well filter plate. Backbone acids are added by row, followed by amine addition by column. (B) Most compounds in the library undergo 3 rounds of amination and acylation and form 3.5-mers. (C) By including a row that receives no PICCO (Row H), a portion of the library skips one round of amination/acylation, resulting in the omission of P2 and R2, and generation of a number of “2.5-mer” compounds in the library. PG = Protecting Group.

A

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B												
C												
D												
E												
F												
G												
H						No PICCO Added						

B



C

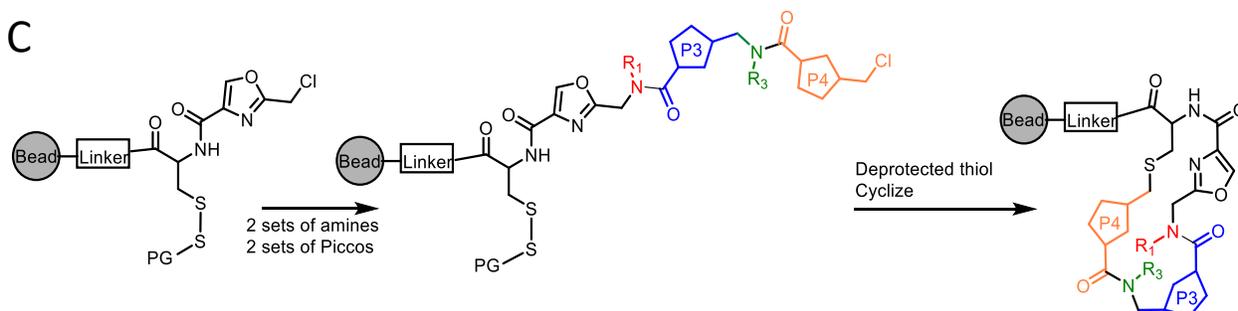


Figure S4 – Structures of linear and cyclic controls for mBBR assay. (A) Structure of the linear control PICCO, synthesized onto a linker that includes Cys(STMP) which, when deprotected, provides a free thiol for mBBR labeling. (B) Structure of the “cyclic” control PICCO. By replacing Cys-STmp with methionine, no free thiol exists and mBBR labeling cannot occur.

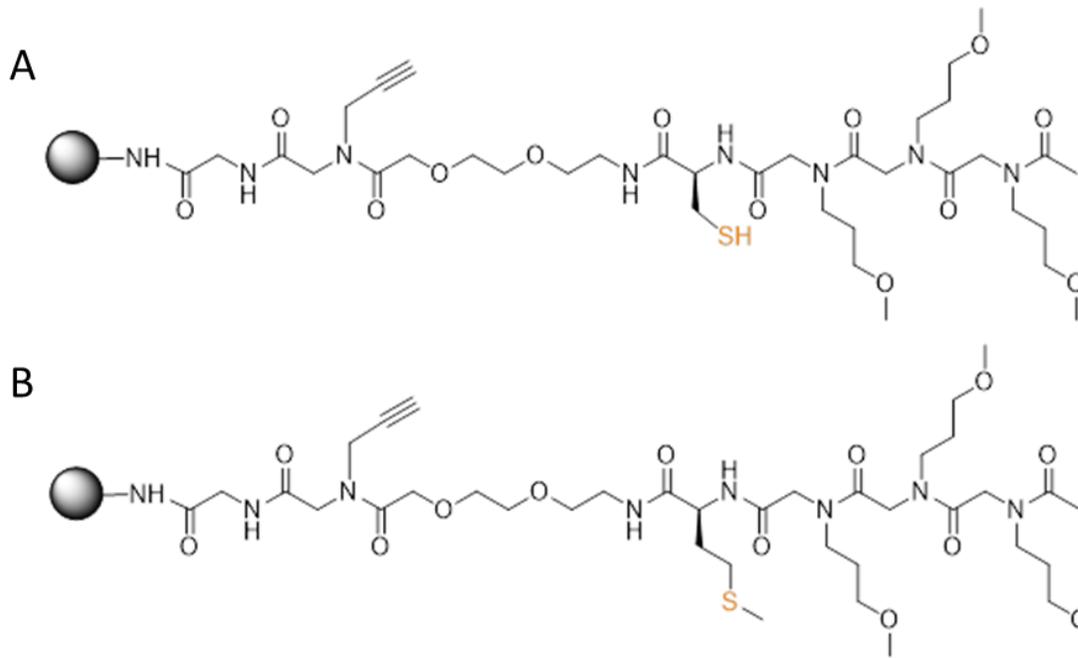


Figure S5 – Fluorescence polarization assays with DBT compounds against human IgG and Siglec-6. The DBT compounds show no evidence of binding to total human IgG or Siglec-6, demonstrating their selectivity for streptavidin.

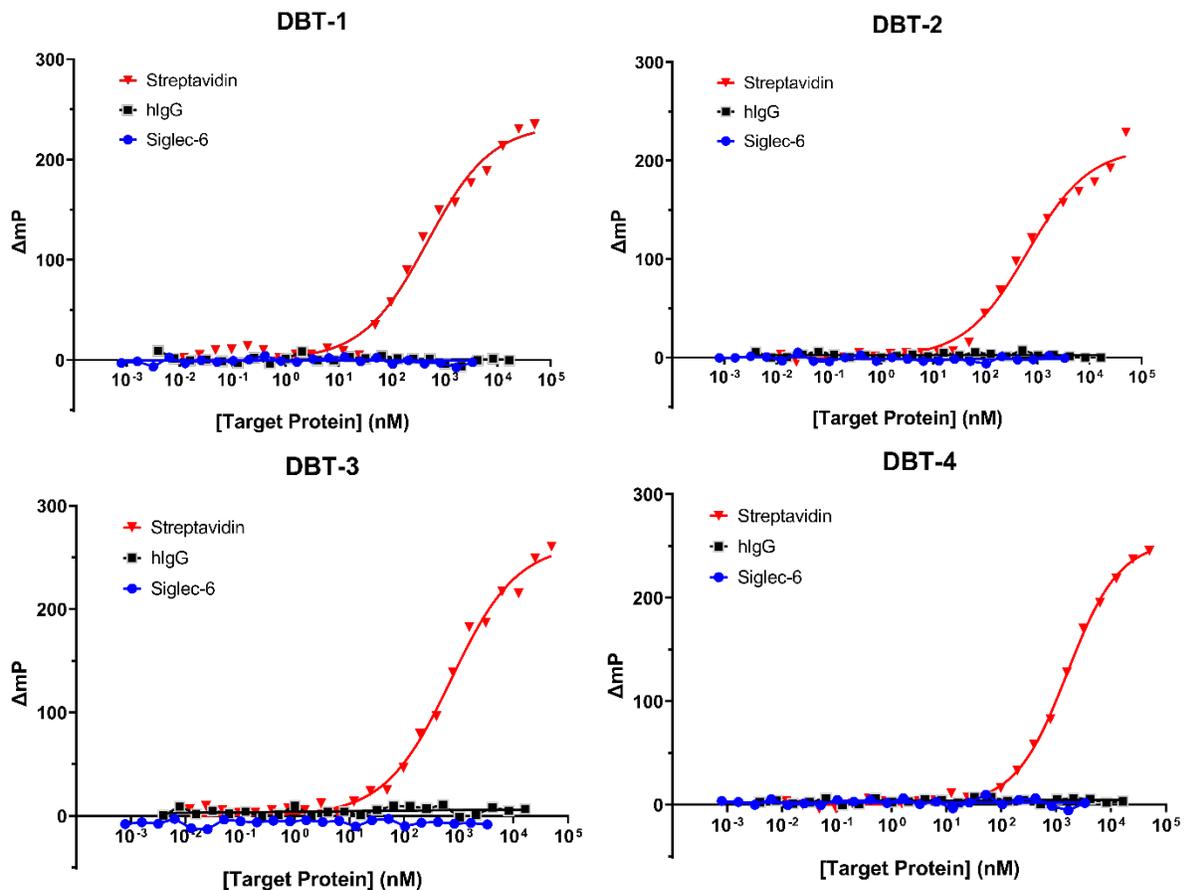


Figure S6 – Structures of macrocyclic hits and linear analogs used for FACS-based validation.

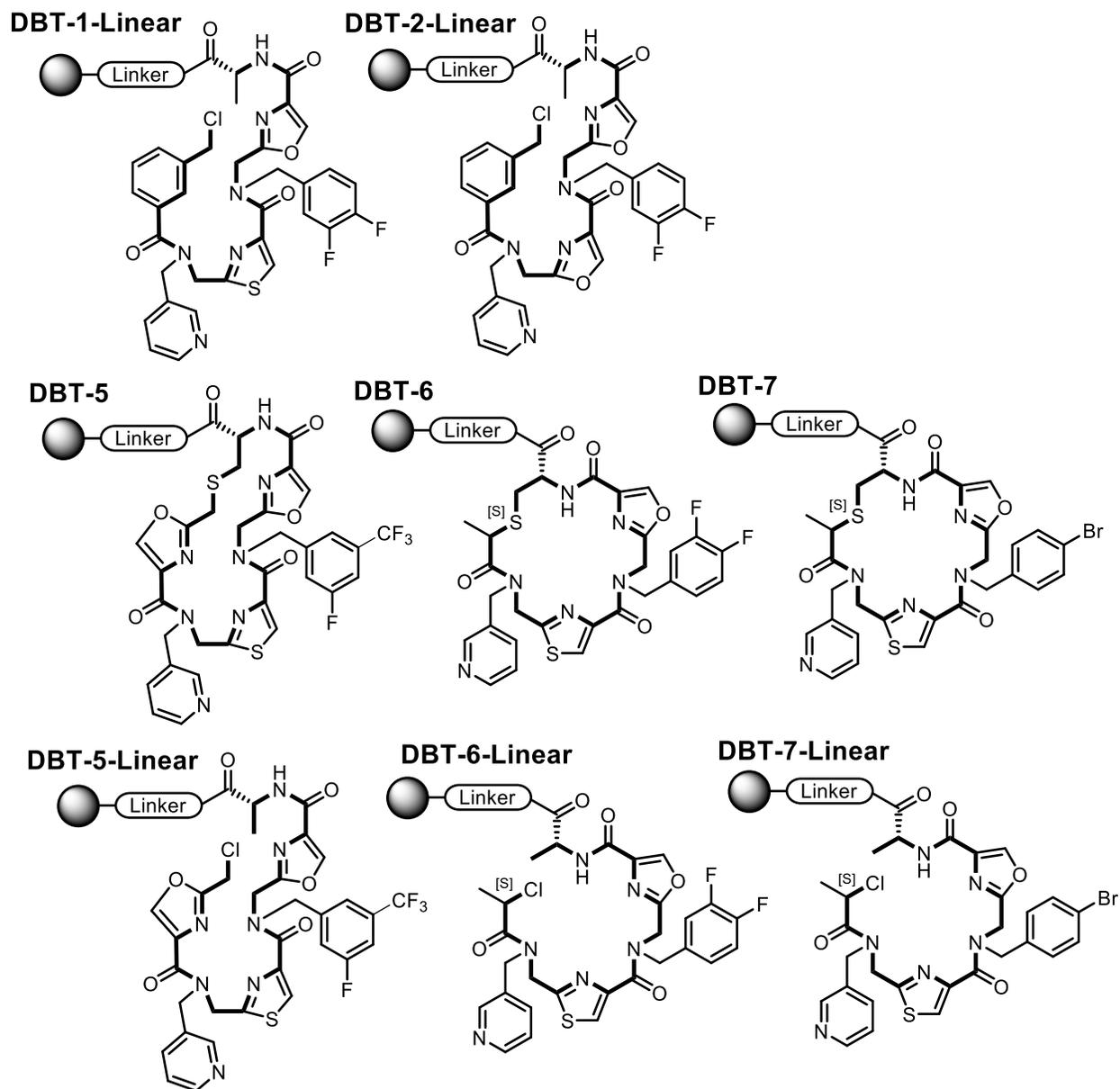


Figure S7 – Competition between the PICCO ligands and biotin for binding to SA-647. (A) Bead-displayed DBT-1 was incubated with 0.0, 0.08 or 0.4 nM SA-A647 in the presence of 5 mM biotin. No binding was observed. (B) Binding of the same beads with SA-A647 (0.4 nM) in the absence of biotin.

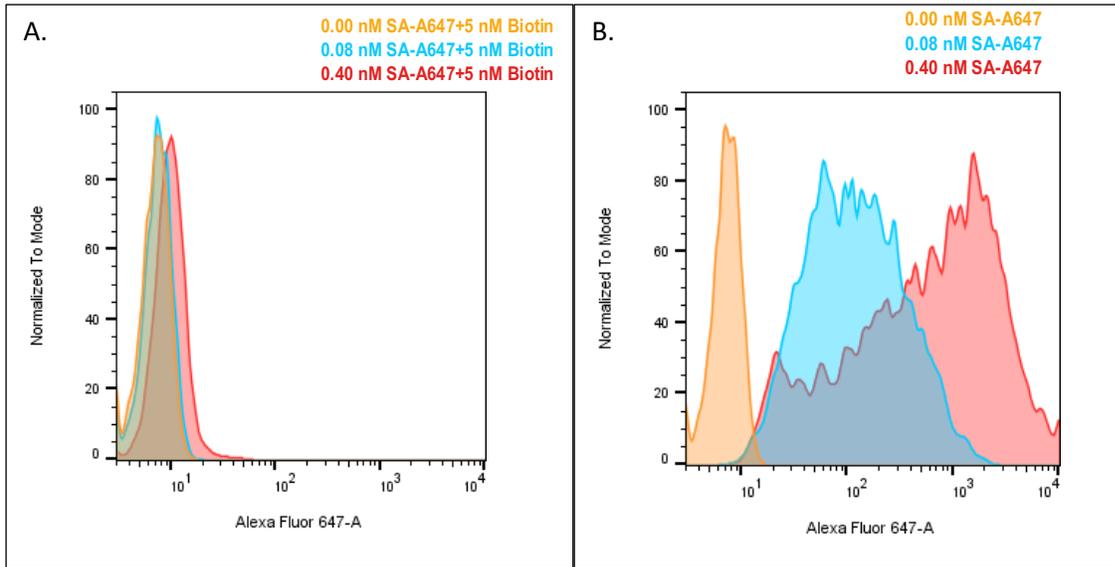


Table S1 – Streptavidin screening analysis. List of putative SA hits with min. redundancy of 3 in both screening replicates.

Number	Barcode IDs			Bead Counts per Screen	
	Pair1	Pair2	Pair3	SA-A488	SA-A647
1	13082403	15052606	17012808	5	5
2	13082404	15052602	17022808	6	4
3	13082404	15052602	17042808	3	7
4	13082406	15052601	17032808	3	6
5	13082408	15052607	17012808	3	6
6	13082403	15052610	17022808	4	4
7	13082406	15052604	17042808	4	4
8	13082408	15052610	17032808	5	3
9	13082403	15102603	17032808	3	5
10	13082406	15052606	17032808	3	5
11	13082404	15052607	17032808	4	3
12	13082404	15052608	17022808	4	3
13	13082407	15052609	17032808	4	3
14	13032408	15042609	17032810	3	4
15	13082403	15052601	17022808	3	4
16	13082404	15032605	17032808	3	4
17	13082406	15052603	17032808	3	4
18	13082408	15052609	17032808	3	4
19	13012409	15052605	17012808	3	3
20	13042408	15052608	17082808	3	3
21	13082403	15032608	17012808	3	3
22	13082403	15052609	17082808	3	3
23	13082404	15052604	17032808	3	3
24	13082404	15052605	17082808	3	3
25	13082404	15052609	17022808	3	3
26	13082404	15092601	17042808	3	3
27	13082404	15102601	17042808	3	3
28	13082405	15052609	17082808	3	3
29	13082406	15052610	17012808	3	3
30	13082407	15052610	17042808	3	3
31	13082408	15052603	17012808	3	3

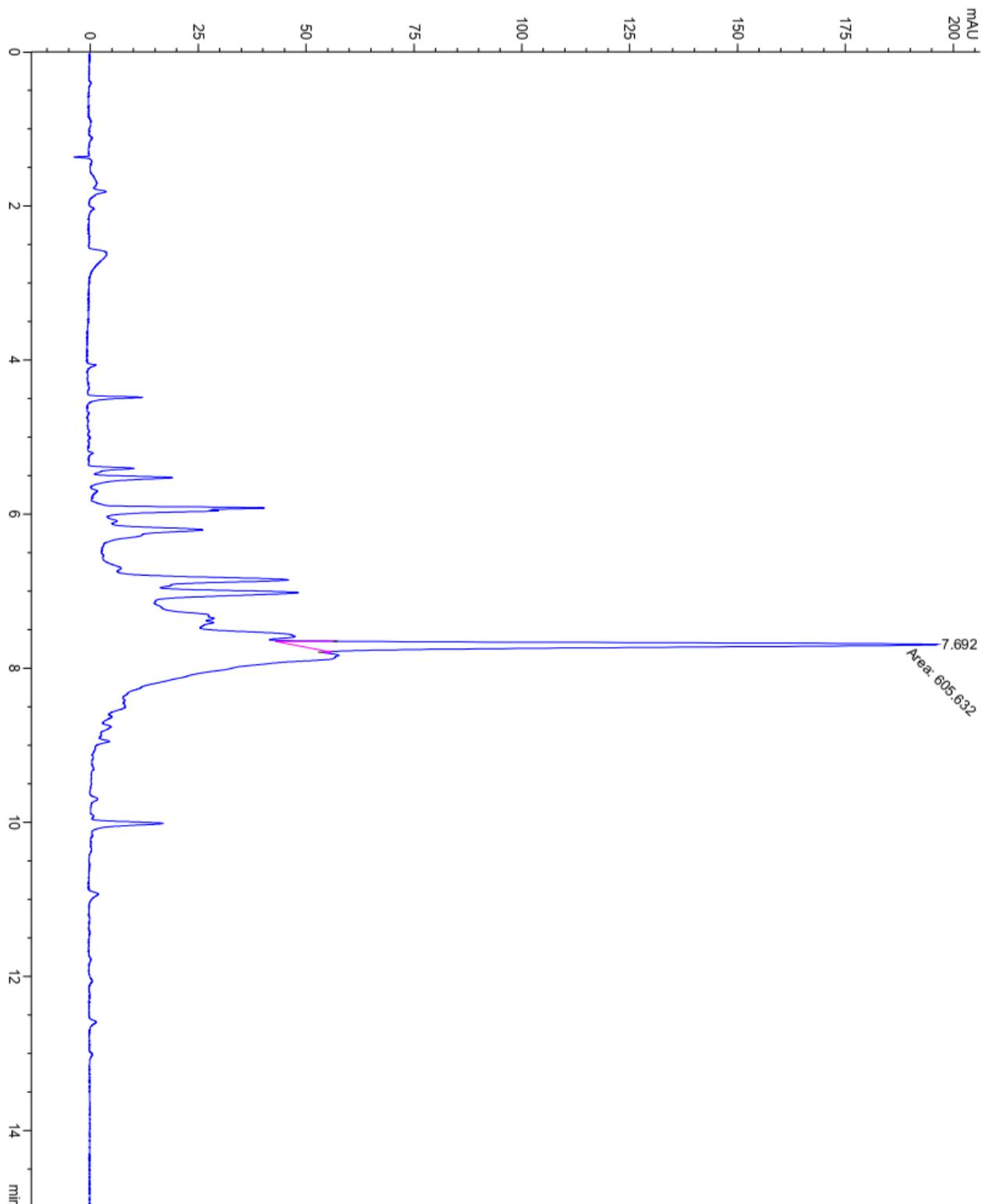
Table S2 – MALDI analysis of resynthesized streptavidin hits. Detected masses are highlighted.

Cmpd	Expected Masses			
	Cyclic	Linear	Cyclic (+BenzBr)	Linear (+BenzBr)
1	1739.80	1776.25	1829.92	1866.38
2	1817.81	1854.26	1907.93	1944.39
3	1714.78	1751.23	1804.90	1841.36
4	1584.54	1620.99	1674.66	1711.12
5	1627.45	1663.90	1717.57	1754.03
6	1593.62	1630.07	1683.74	1720.20
7	1650.60	1687.05	1740.72	1777.18
8	1627.45	1663.90	1717.57	1754.03
9	1611.42	1647.87	1701.54	1738.00
10	1549.54	1585.99	1639.66	1676.12
11	1643.59	1680.04	1733.71	1770.17
12	1565.62	1602.07	1655.74	1692.20
13	1600.62	1637.07	1690.74	1727.20
14	1638.72	1675.17	1728.84	1765.30
15	1584.52	1620.97	1674.64	1711.10
16	1577.53	1613.98	1667.65	1704.11
17	1631.72	1668.17	1721.84	1758.30
18	1501.56	1538.01	1591.68	1628.14
19	1643.52	1679.97	1733.64	1770.10
20	1574.42	1610.87	1664.54	1701.00
21	1634.52	1670.97	1724.64	1761.10
22	1531.52	1567.97	1621.64	1658.10
23	1691.68	1728.13	1781.80	1818.26

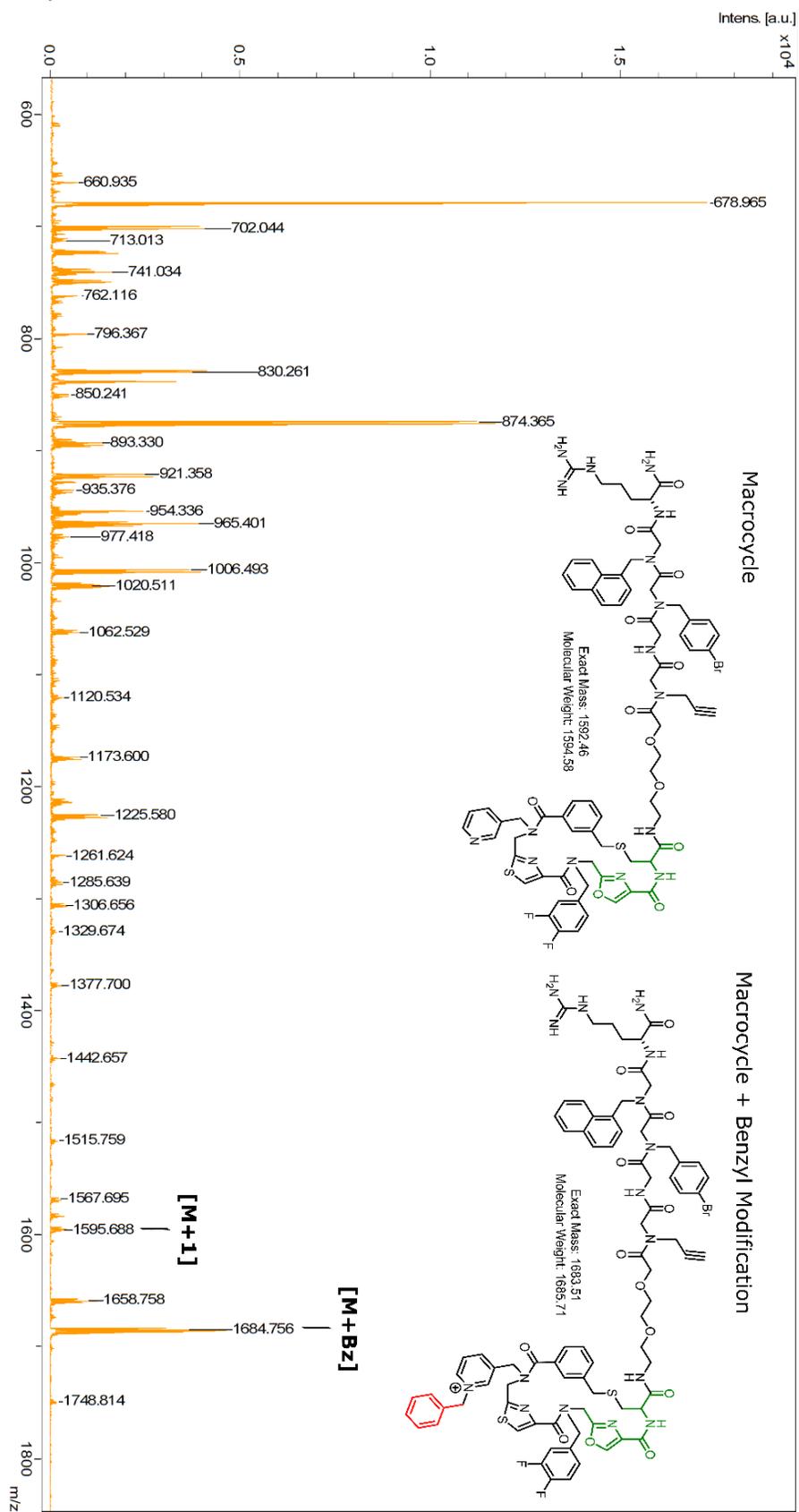
Section III – Spectrometry Data

Macrocyclization of compound 3, shown in Figure 1.

LC chromatogram at 254 nm.

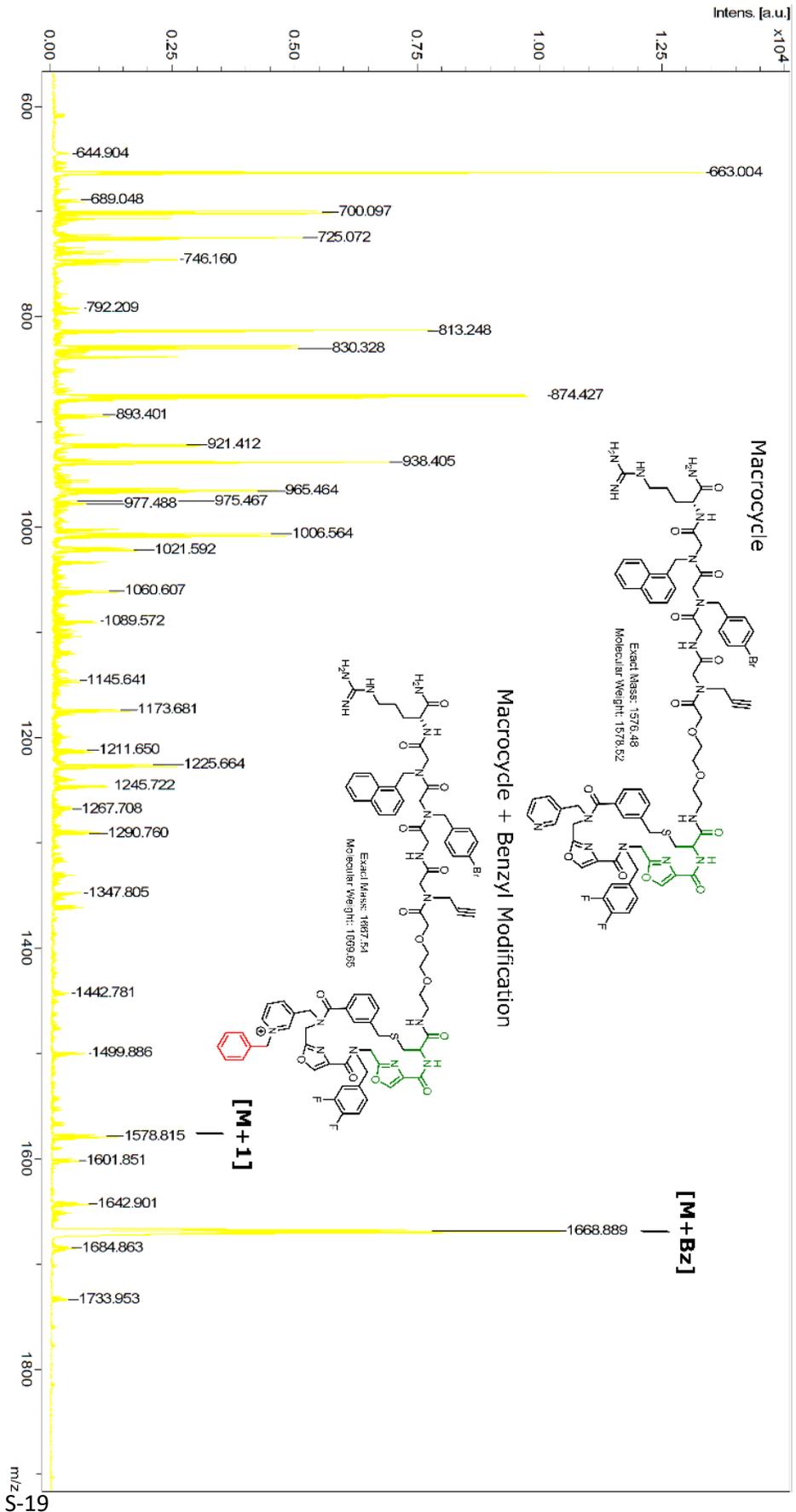


Resynthesized hits

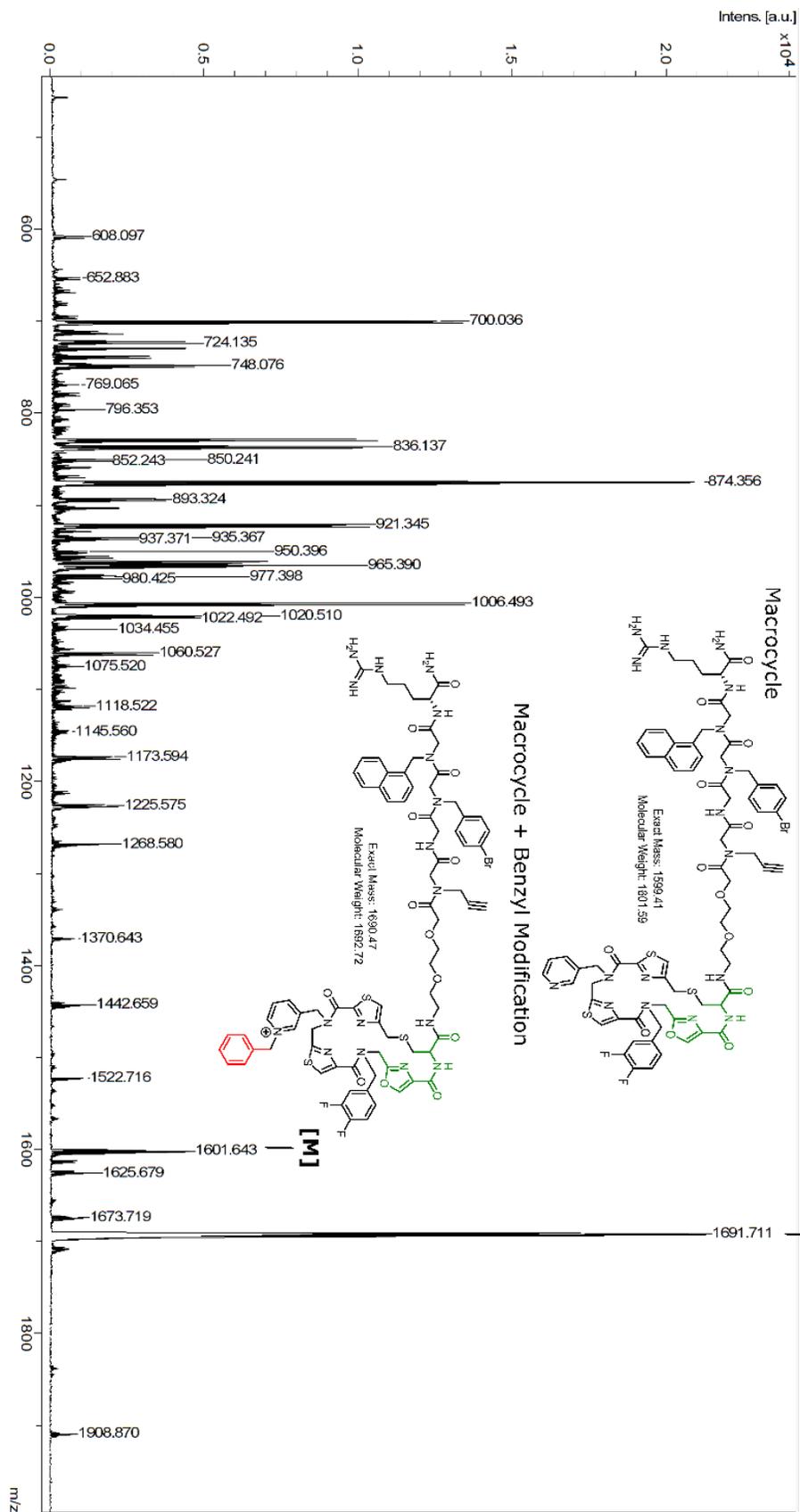


DBT-1

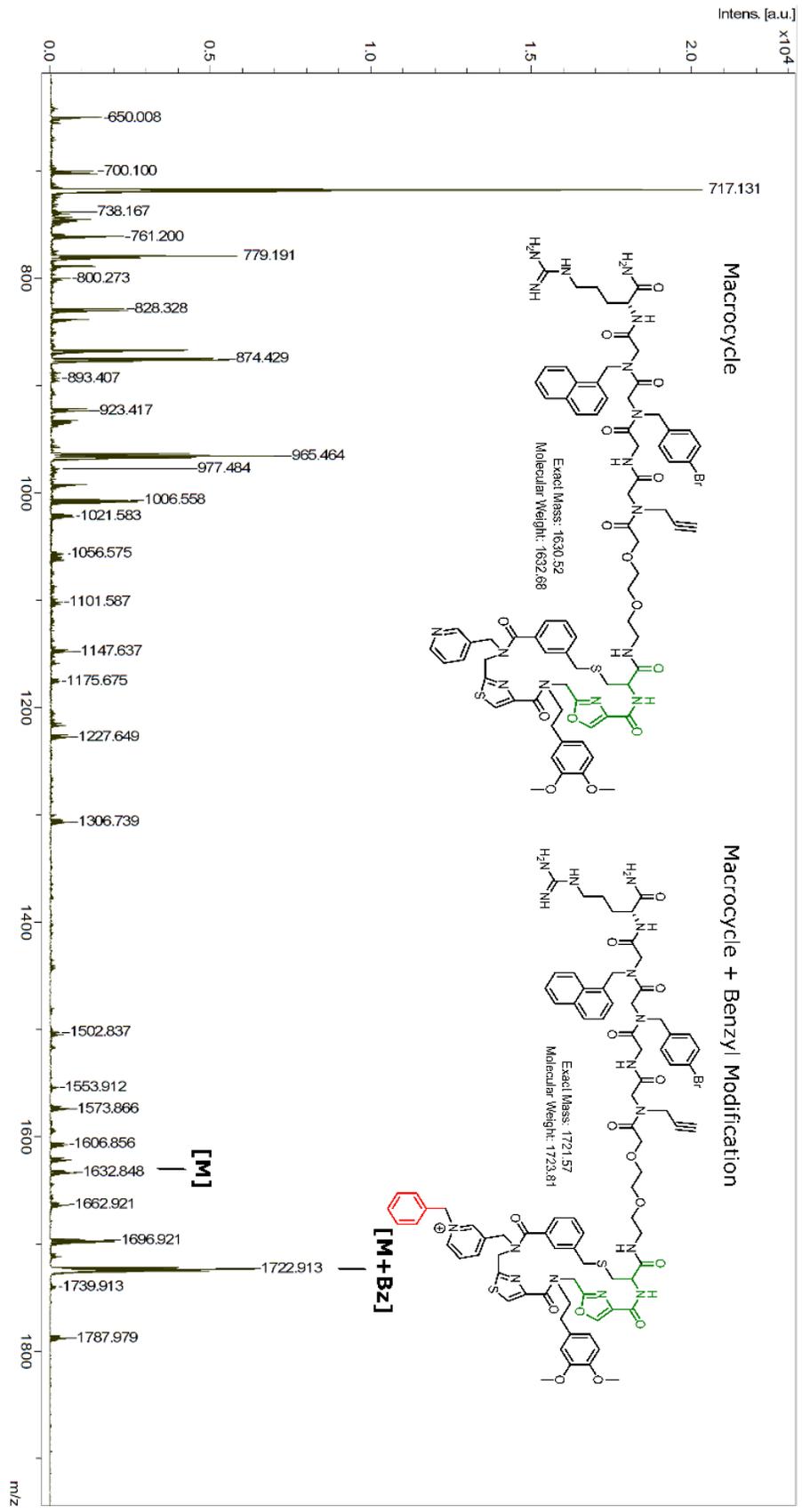
DBT-2



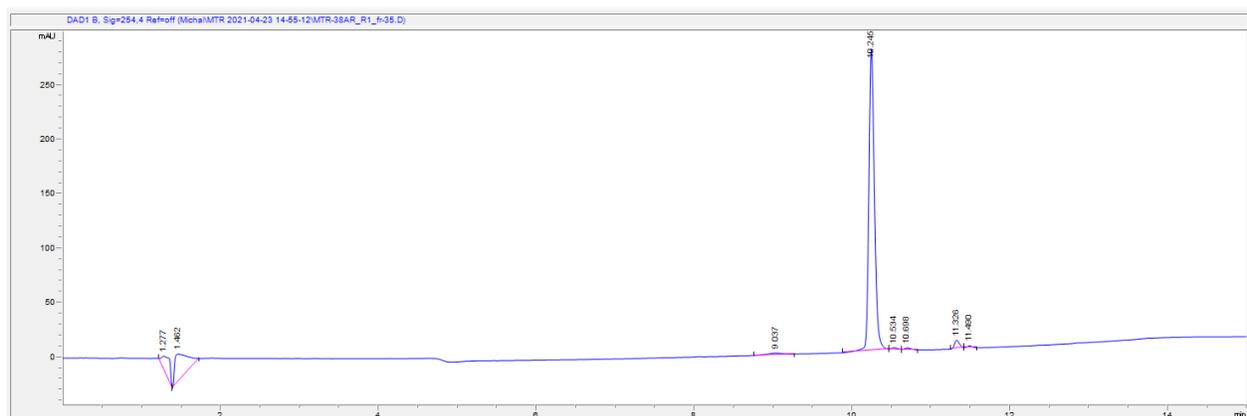
DBT-3



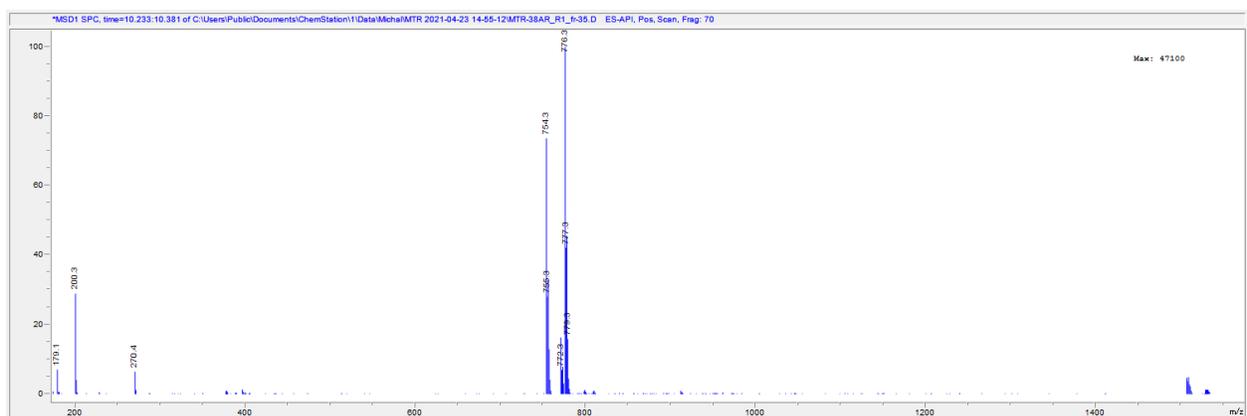
DBT-4



Section IV. Characterization of chloroalkane-tagged macrocycles KMR-7, -8 and -9.



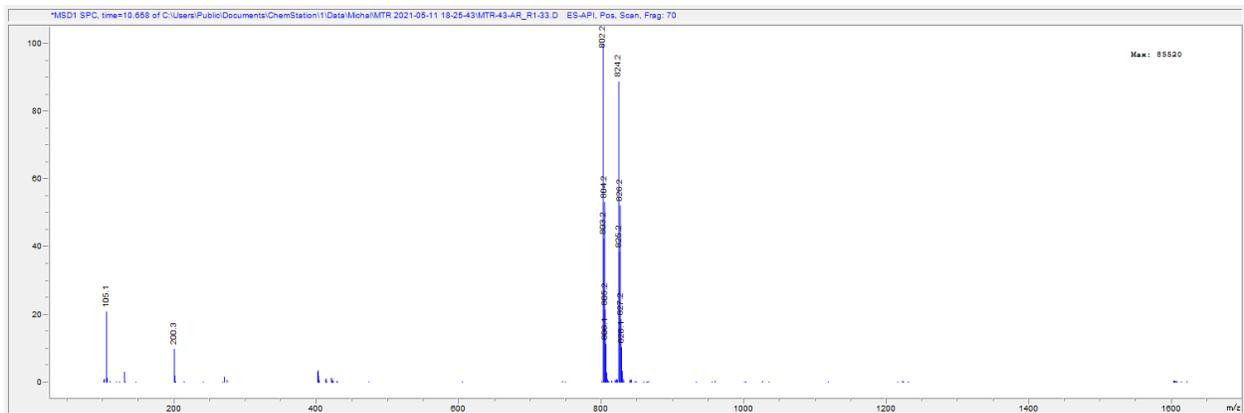
LC-MS trace of KMR-7.



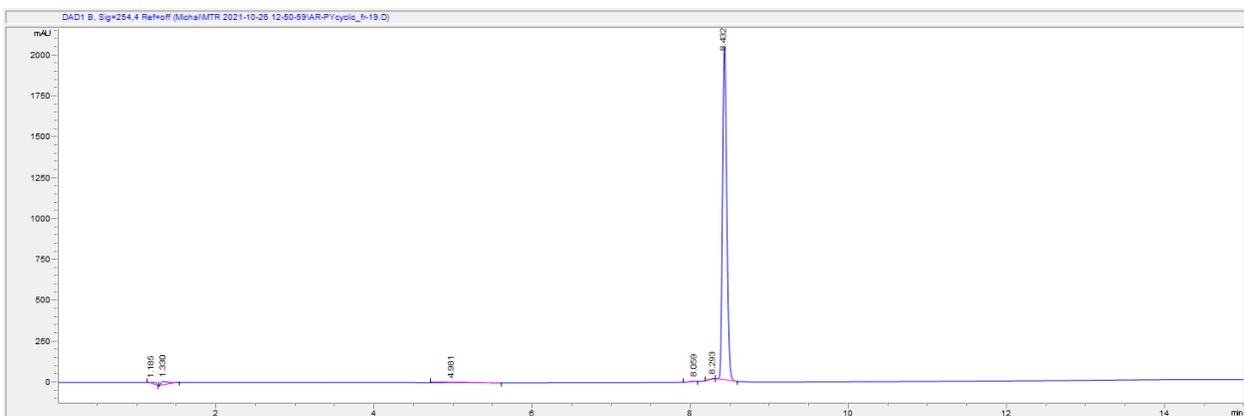
Mass spectrum of KMR-7.



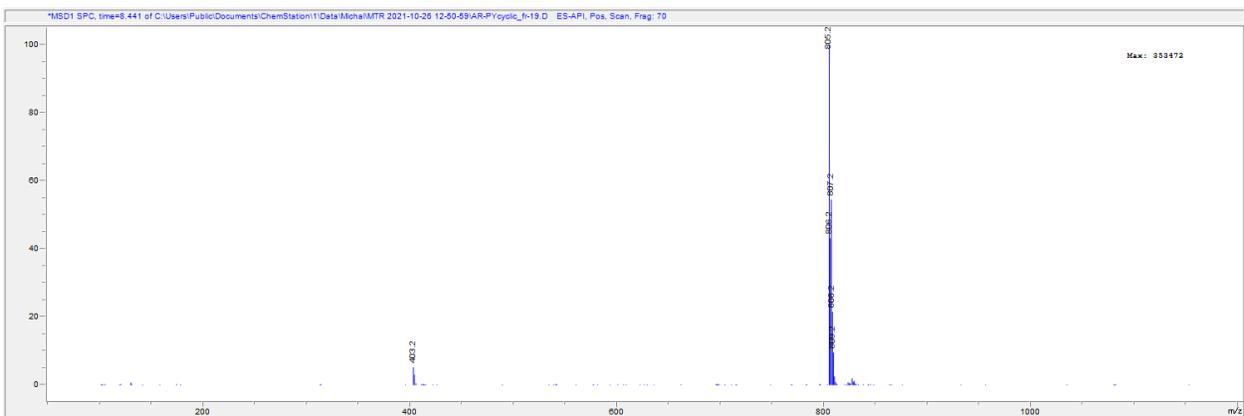
LC-MS trace of KMR-8.



Mass spectrum of KMR-8.



LC-MS trace of KMR-9.



Mass spectrum of KMR-9.

References

1. MacConnell, A. B.; McEnaney, P. J.; Cavett, V. J.; Paegel, B. M., DNA-Encoded Solid-Phase Synthesis: Encoding Language Design and Complex Oligomer Library Synthesis. *ACS Comb Sci* **2015**, *17*, 518-534.
2. Mendes, K.; Malone, M. L.; Ndungu, J. M.; Suponitsky-Kroyter, I.; Cavett, V. J.; McEnaney, P. J.; MacConnell, A. B.; Doran, T. M.; Ronacher, K.; Stanley, K.; Utset, O.; Walzl, G.; Paegel, B. M.; Kodadek, T., High-throughput identification of DNA-encoded IgG ligands that distinguish active and latent mycobacterium tuberculosis infections. *ACS Chem. Biol.* **2017**, *19*, 234-243.