Supplementary Information

Versatile protein recognition by the encoded display of chemical elements on a constant macrocyclic scaffold

Yizhou Li[1, 2]*, Roberto De Luca[2], Samuele Cazzamalli[2], Francesca Pretto[3], Davor Bajic[2], Jörg Scheuermann[2]* & Dario Neri[2]*

Affiliation

1 Chongqing Key Laboratory of Natural Product Synthesis and Drug Research, School of Pharmaceutical Sciences, Chongqing University, 55 Daxuecheng South Road, Shapingba, Chongqing, 401331, P. R. China.

2 Department of Chemistry and Applied Biosciences, Swiss Federal Institute of Technology (ETH Zürich), Vladimir-Prelog-Weg 4, CH-8093 Zürich, Switzerland

3 Philochem AG, 8112 Otelfingen, Switzerland.

* Corresponding Authors

Tel: +41-44-6337401; e-mail: yizhouli@cqu.edu.cn; joerg.scheuermann@pharma.ethz.ch; dario.neri@pharma.ethz.ch

Table of Contents

1	**Abbreviations**	**S2**
2	**Materials and General Methods**	**S3**
3	**Library Synthesis, Purification, and Characterization**	**S4**
4	**Affinity Selections of the Library against Immobilized Target Proteins**	**S39**
5	**Affinity Measurement of Macrocycles Selected from the Library**	**S44**
6	**Immunofluorescence Performance of Selected PSA Binder**	**S81**
7	***In vivo* Performance of CAIX Binder**	**S82**
8	**Performance of Chemical Probes Developed from Selected CaM Binder**	**S85**
9	**References**	**S91**

.

1. Abbreviations.

BSA: bovine serum albumin

CBB: coomassie brilliant blue

DAPI: 4′,6-Diamidine-2′-phenylindole dihydrochloride

DCM: dichloromethane

DE: diversity element

DIPEA: *N*, *N*'-diisopropylethylamine

DMF: *N*, *N*'-dimethylformamide

DMSO: dimethyl sulfoxide

DPBS: Dulbecco's phosphate buffered saline

EDC: 1-ethyl-3- (3-dimethylaminopropyl) carbodiimide

EDC-HCl: 1-ethyl-3- (3-dimethylaminopropyl) carbodiimide hydrochloride

EDTA: ethylenediaminetetraacetic acid

FITC: Fluorescein isothiocyanate isomer I

Fmoc: 9-fluorenylmethyloxycarbonyl

HFIP: 1, 1, 1, 3, 3, 3-hexafluoro-2-propanol

HOAt: 1-hydroxy-7-azabenzotriazole

NHS: *N*-hydroxysuccinimide

PAGE: polyacrylamide gel electrophoresis

PBS: phosphate buffered saline

PyBOP: benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate

SE: succinimidyl ester

sNHS: *N*-hydroxysulfosuccinimide sodium salt

SPPS: solid phase peptide synthesis

TBTA: tris[(1-benzyl-1*H*-1, 2, 3-triazol-4-yl)methyl]amine

TEAA: triethylammonium acetate

TEA: triethylamine

TCEP-HCl: tris(2-carboxyethyl)phosphine hydrochloride

TFA: trifluoroacetic acid

Tfa: trifluoroacetate

TNBS: trinitrobenzenesulfonic acid

Tris-HCl: tris(hydroxylmethyl)aminomethane hydrochloride

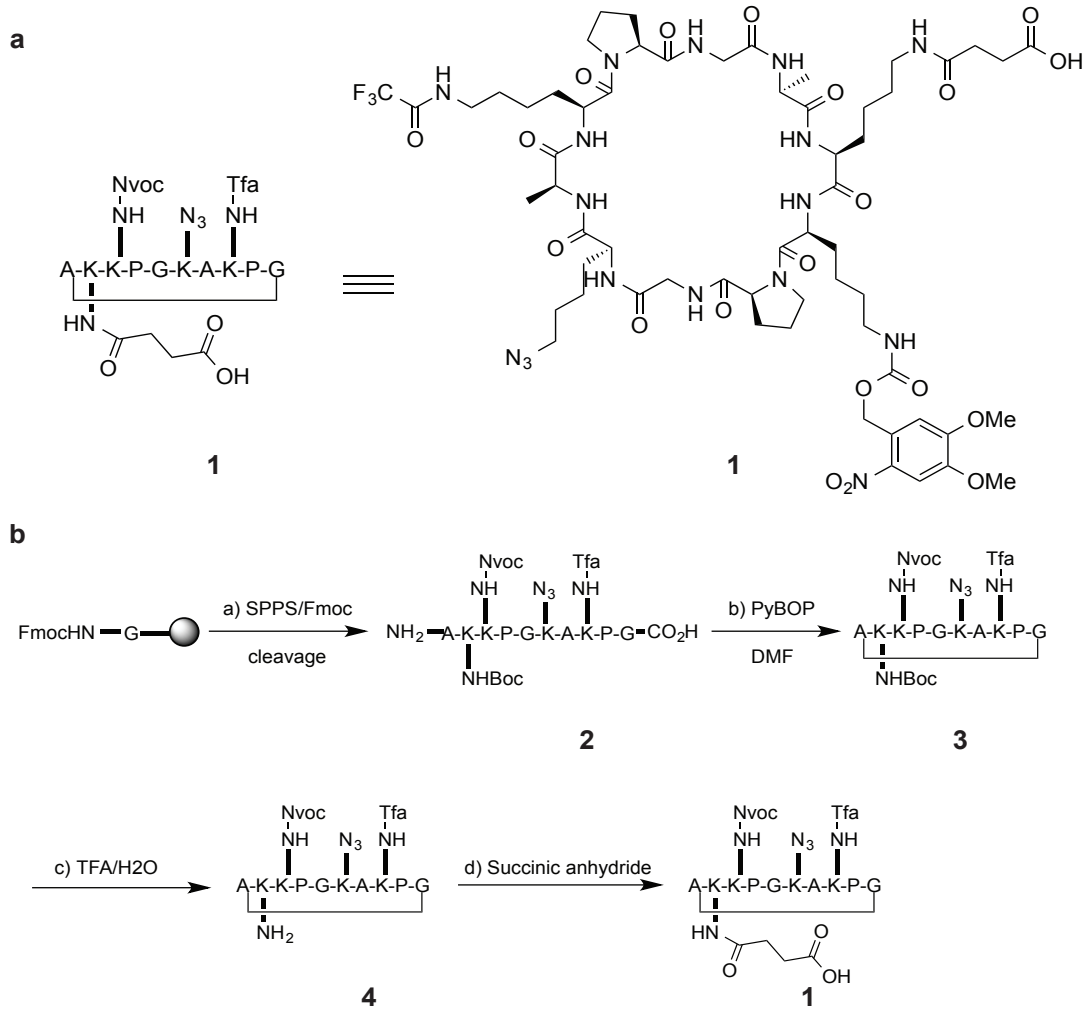
2. Materials and General Methods.

Unless otherwise noted, all reagents and solvents were purchased from commercial sources and used as received. Oligonucleotides were purchased from DNA Technology (Denmark) and IBA (Germany). All DNA sequences are written in 5'- to 3'- orientation unless otherwise noted. Carboxylic acids and terminal alkynes were purchased from several commercial suppliers including ABCR, ChemBridge, Sigma-Aldrich, TCI Europe, Alfa Aesar, Matrix Scientific, Enamine Store and Acros Organics. Water was purified with a Millipore Milli-Q system. All gel images were captured by a Bio-Rad Chemidoc image system. Nvoc-off and photo-crossing-linking experiments were conducted by a UVP CL-1000 Ultraviolet crosslinker at 365 nm with an intensity of approximately 100 μJ /cm^2.

3. Library Synthesis, Purification, and Characterization.

(a) Scaffold structures, synthesis, purification and characterization.

As discussed in the main text, three orthogonal amino protecting groups were incorporated in the cyclodecapeptide scaffold **1**. using *N*-α-Fmoc-ε-protected- *L*-Lys-OH. The cyclodecapeptide scaffold **1** was synthesized following a modified procedure reported by Boturyn and co-workers[1]. Assembly of all peptides was carried out using the Fmoc strategy manually in a glass reaction vessel fitted with a sintered glass frit. Coupling reactions were performed manually by using 2.0 equiv. of *N*-Fmoc-protected amino acid (relative to the resin loading) activated *in situ* with 2.0 equiv. of PyBOP and 5.0 equiv. of DIPEA in DMF for 60 min. The coupling efficiency in manual synthesis was asssessed by TNBS tests. Fmoc-protecting groups were removed by treatment with a piperidine/DMF solution (1:4) for 10 min. The process was repeated three times and the completeness of deprotection verified by UV absorption of the piperidine washings at 299 nm. Synthetic linear peptides were recovered directly upon acid cleavage. Before cleavage, the resin was washed thoroughly with DCM. The peptide was released from the resin using a cleavage solution of TFE/AcOH/DCM (2:1:7, 2 × 30 min).



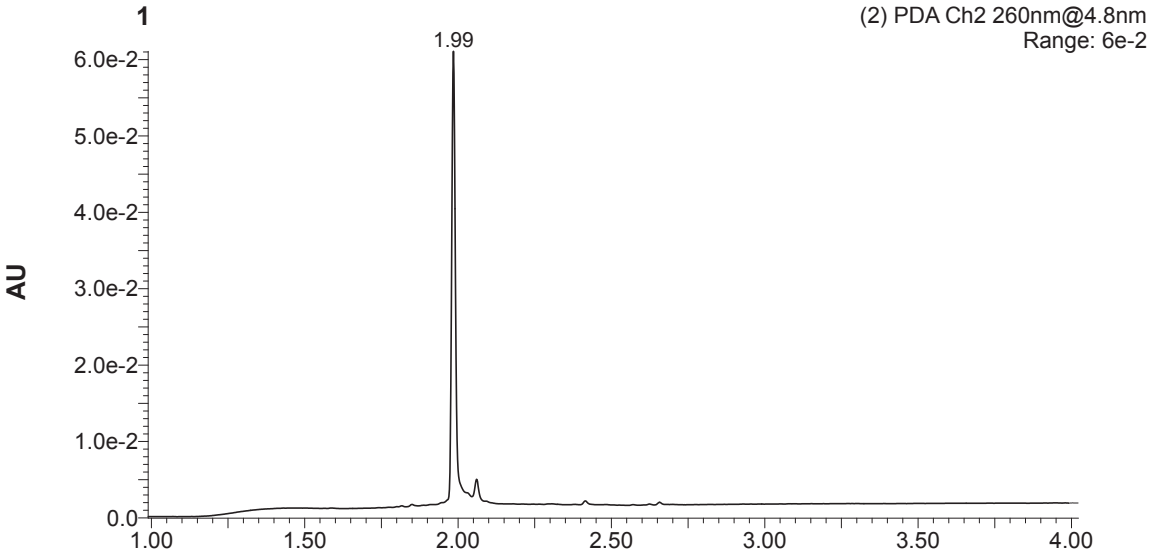
Supplementary Figure 1 | Structure of scaffold 1 for library construction. **a**, Structure of the cyclodecapeptide scaffold with three orthogonal amino protecting groups. All amino acid residues in the peptide sequence are in the *L*-form. **b**, Synthesis scheme of scaffold **1**. a) Fmoc-protected amino acid: 2.0 eq., PyPOB: 2.0 eq., DIPEA: 5.0 eq., SPPS; b) PyBOP: 1.0 eq. 25 ºC, 3 h; c) TFA/H_2O (20:1), 25 ºC, 1 h; d) Succinic anhydride: 1.0 eq., 25 ºC, 1 hour.

Synthesis of the linear decapeptide 2: the linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g, Rapp, Catalog: RA1213) using the general procedure in the following sequence: Fmoc-*L*-Pro-OH (Aldrich, Catalog: 47636), Fmoc-*L*-Lys(Tfa)-OH (Senn, Catalog: 100575), Fmoc-*L*-Ala-OH (Aldrich, Catalog: 531480), Fmoc-*L*-Lys(N_3)-OH (ChemPep, 101227), Fmoc-Gly-OH (Novabiochem, Catalog: 04-12-1001-25), Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Nvoc)-OH (Anaspec, Catalog: AS62574-1000), Fmoc-*L*-Lys(Boc)-OH (Fluka, 47624), Fmoc-*L*-Ala-OH. The peptide was released from the resin using a cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white-yellow powder after precipitation and washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. **HRMS (*m/z*, $C_{61}H_{94}F_3N_{17}O_{20}$, ESI):** calculated $[M+H]^+$: 1442.6891; found: 1442.6886.

Synthesis of the cyclodecapeptide 3: the linear decapeptide 2 (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 3 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1), followed by reverse-phase HPLC purification. The desired cyclodecapeptide **3** was recovered as a white-yellow powder after lyophilization. **HRMS (*m/z*, $C_{61}H_{92}F_3N_{17}O_{19}$, ESI):** calculated $[M+H]^+$: 1424.6786; found: 1424.6793.

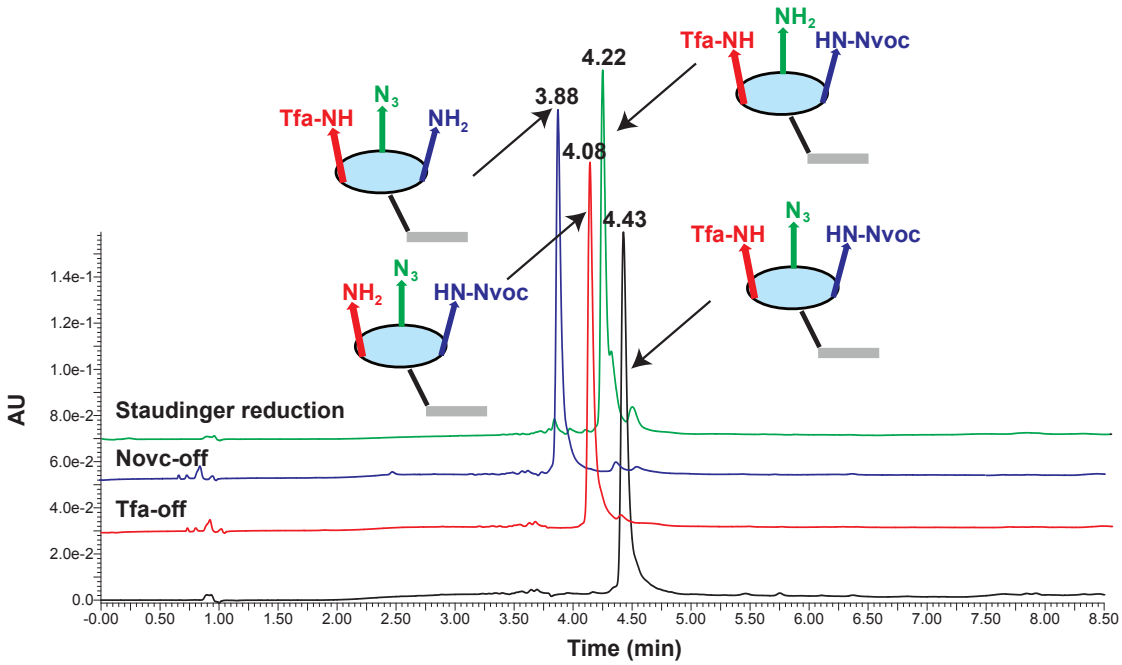
Synthesis of the cyclodecapeptide 4: to the cyclodecapeptide **3** was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **4** was recovered as a white-yellow powder after lyophilization. **HRMS (*m/z*, $C_{56}H_{84}F_3N_{17}O_{17}$, ESI):** calculated $[M+H]^+$: 1324.6261; found: 1324.6267.

Synthesis of the cyclodecapeptide 1: the linear decapeptide **4** (0.5 M) were dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA (5.0 equiv.). Succinic anhydride (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. The solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1), followed by reverse-phase HPLC purification. The desired cyclodecapeptide **1** was recovered as a white-yellow powder after lyophilization. **HRMS (*m/z*, $C_{60}H_{88}F_3N_{17}O_{20}$, ESI):** calculated $[M+H]^+$: 1424.6422; found: 1424.6416.



Supplementary Figure 2 | UPLC chromatogram of SC-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5% to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

(b) Testing the reactivity of the orthogonal protection groups.



Supplementary Figure 3 | UPLC chromatogram of scaffold-oligonucleotide conjugate and orthogonal deprotections. UPLC analyses were performed on a on a XBridge@ Oligonucleotide BEH C18 10 × 50 mm column at a flow rate of 0.5 mL/min with gradient: 0 % to 5 % B (0 to 0.5 minutes), 5 % to 50 % B (0.5 to 7 minutes), 50 % to 100 % B (7 to 7.1 minutes), 100 % B (7 to 8 minutes), 100 % to 0% B (8 to 10 minutes) (A= TEA 10 mM, HFIP 5 mM in water, B= MeOH), at 60 ºC. Detection by absorbance at 260 nm.

Conjugation of scaffold to oligonucleotide: To a solution of amino-modified oligonucleotide (code 1: 5'-amino-C6-GGAGCTTCTGAATTCTGTGTGCTGTTATGGCGAGTCCCATGGCGC -3'-OH, 100 nmol) in MOPS buffer (50 mM, pH 8.0, 0.5 M NaCl, 720 µL) was added a mixture of scaffold (cyclodecapeptide **1**) (60 mM, 450 µL), EDC (300 mM, 40 µL), HOAt (60 mM, 40 µL) and DIPEA (300 mM, 40 µL) in DMSO, previously activated for 15 minutes at room temperature. The reaction was agitated at room temperature for 16 h. The reaction solution was then treated with a second addition of freshly activated scaffold in DMSO (same activation mixture as above) and it was agitated for further 6 h at room temperature. Conjugation reactions were quenched with Tris-HCl (200 µL, 500 mM, pH 8.0) at 30 °C for 1 h.[2] After quenching, the conjugate was precipitated with ethanol before purification by HPLC. The separated and collected conjugate was vacuum-dried overnight, redissolved in H_2O (500 µL), quantified by UV absorption at 260 nm yielding a recovery of 45 %. The scaffold conjugate was characterized by UPLC-MS. Deconvoluted molecular mass: predicted: 15491; found: 15491.

Tfa deprotection: To the scaffold-oligonucleotide conjugate (10 nmol) ammonium hydroxide solution (25 % aq, 400 µL, Aldrich, Catalog: 30501) was added and the Tfa deprotection was allowed for 2h at 25 ºC. The Tfa-off conjugates were

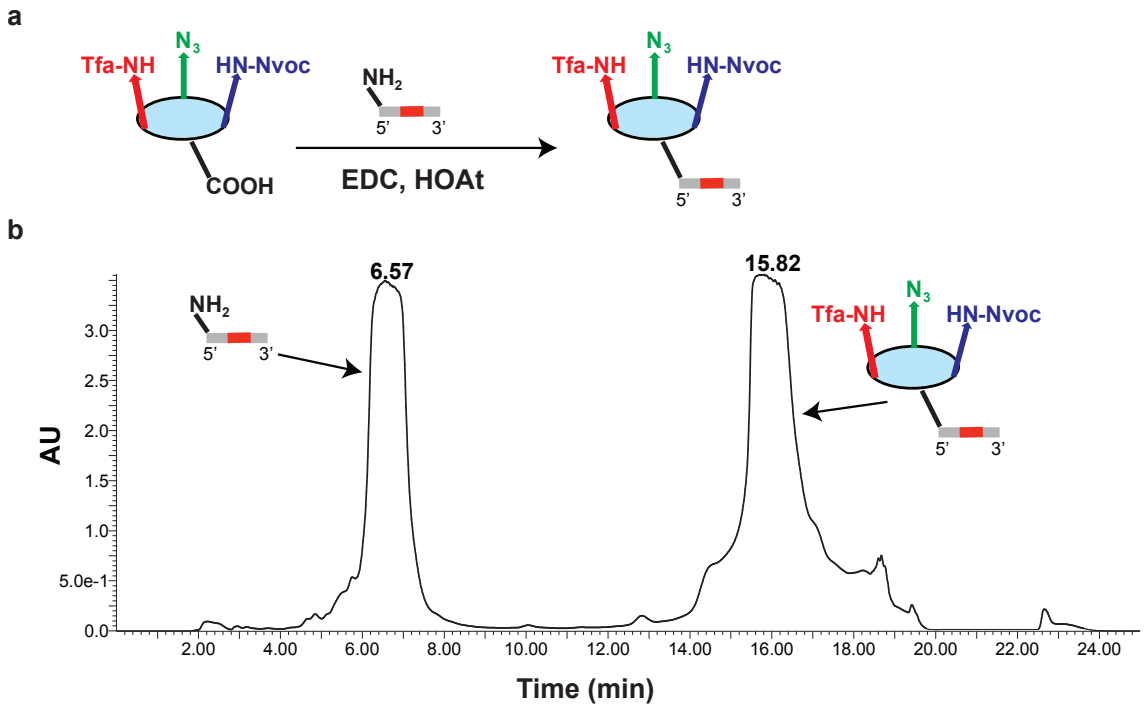
vacuum-dried overnight, redissolved in H_2O (200 µL), and quantified by UV absorption at 260 nm. The recovery yield was over 95 %. TFA-deprotection was analyzed by UPLC-MS. Deconvoluted molecular mass: predicted: 15395; found: 15395.

Novc-deprotection: Scaffold-oligonucleotide conjugate (10 nmol) in H_2O (500 µL) was subjected to irradiation at 365 nm for 60 min at 0 °C (on ice)[3]. The recovery yield was over 85 %. Novc deprotection was analyzed by UPLC-MS. Deconvoluted molecular mass: predicted: 15252; found: 15252.

Staudinger reduction: Scaffold-oligonucleotide conjugate (10 nmol) was dissolved in Tris-HCl (1 mL, 500 mM, pH 8.0), followed by addition of TCEP (20 mg, Aldrich, Catalog: 41996). The Staudinger reduction lasted for 12 h at 25 °C, the reduction product was isolated by ethanol precipitation and the pellet was dissolved in H_2O (200 µL). The recovery yield was over 90 % and the reduction was analyzed by UPLC-MS. Deconvoluted molecular mass: predicted 15465; found: 15465.

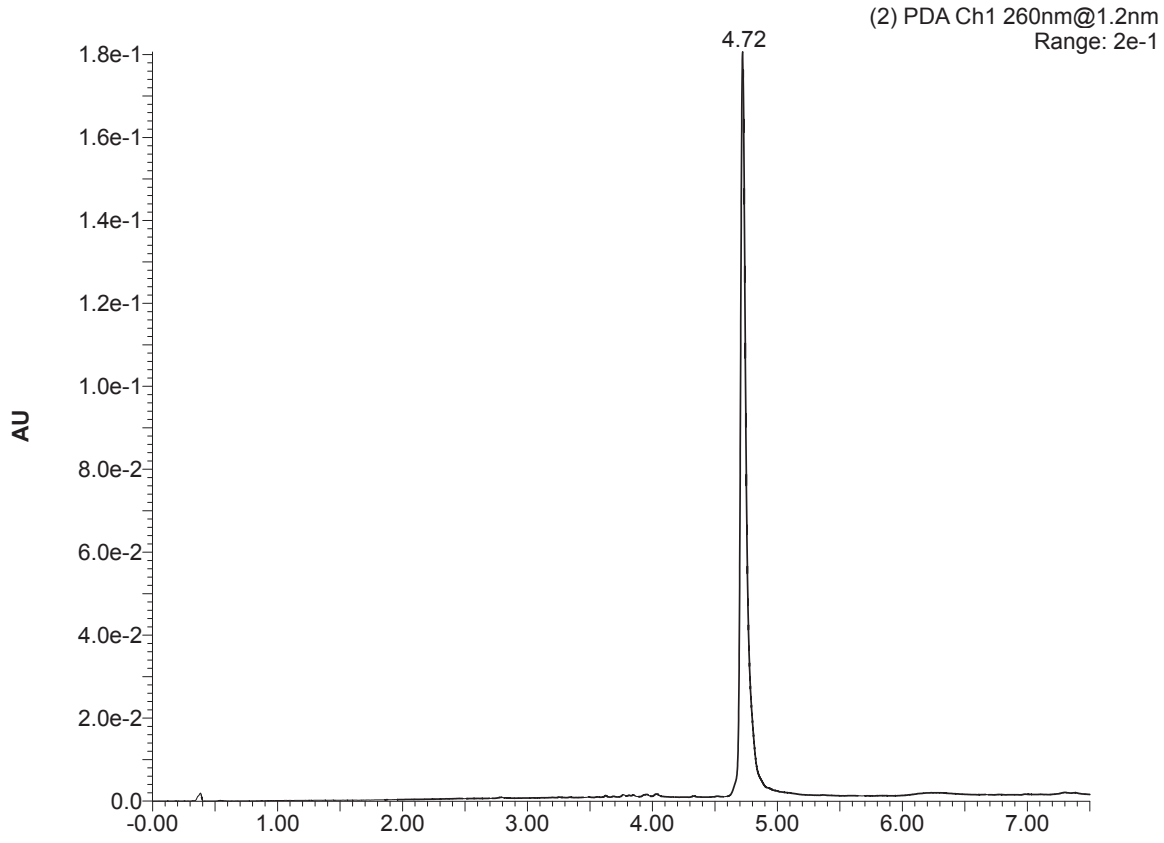
As shown in Supplementary Fig. 3, all three protection groups were orthogonally removed with high conversion (over 80 %), indicating the suitability of the designed scaffold **1** for library construction.

(c) Conjugation of scaffold to code 1 and deprotection of Tfa, purification and characterization.

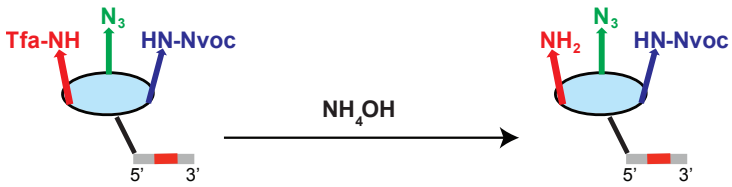


Supplementary Figure 4 | Conjugation of scaffold to code 1. **a**, Conjugation of scaffold to code 1 oligonucleotide by EDC and HOAt. **b**, Representative HPLC chromatogram of scaffold-code 1 conjugate. HPLC purifications were performed on a CT18-XTerra 10 × 150 mm column at a flow rate of 4 mL/min with gradient: 10 % B to 40 % B in 15 minutes, (A= TEAA 0.1 M in water, B= CH_3CN 80 % in water), at 25 ºC. Detection by absorbance at 260 nm.

To a solution of amino-modified oligonucleotide (code 1: 5'-amino-C6-GGAGCTTCTGAATTCTGTGTGCTG **NNNNNN**CGAGTCCCATGGCGC-3'-OH, **N** represents variable bases serving as code, 100 nmol) in MOPS buffer (50 mM, pH 8.0, 0.5 M NaCl, 720 µL) was added a mixture of scaffold (cyclodecapeptide **1**) (60 mM, 450 µL), EDC (300 mM, 40 µL), HOAt (60 mM, 40 µL) and DIPEA (300 mM, 40 µL) in DMSO, previously activated for 15 minutes at room temperature. The reaction was agitated at room temperature for 16 h. The reaction solution was then treated with a second addition of freshly activated scaffold in DMSO (same activation mixture as above) and it was agitated for further 6 h at room temperature. Conjugation reactions were quenched with Tris-HCl (200 µL, 500 mM, pH 8.0) at 30 °C for 1 h. After quenching, the conjugates were precipitated with ethanol before purifying by HPLC (See Supplementary Fig. 4). The separated and collected conjugates were vacuum-dried overnight, redissolved in H_2O (500 µL), and quantified by UV absorption at 260 nm. Obtained yields were around 20 - 50 %. All 283 conjugates were characterized by UPLC-MS (see Supplementary Fig. 5).



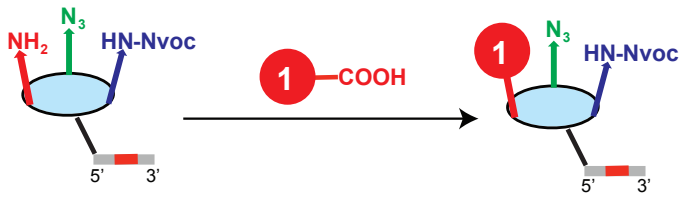
Supplementary Figure 5 | UPLC chromatogram of scaffold-code 1 conjugate after HPLC purification. UPLC analyses were performed on a on a XBridge@ Oligonucleotide BEH C18 10 × 50 mm column at a flow rate of 0.5 mL/min with gradient: 0 % to 5 % B (0 to 0.5 minutes), 5 % to 50 % B (0.5 to 7 minutes), 50 % to 100 % B (7 to 7.1 minutes), 100 % B (7 to 8 minutes), 100 % to 0 % B (8 to 10 minutes) (A= TEA 10 mM, HFIP 5 mM in water, B= MeOH), at 60 ºC. Detection by absorbance at 260 nm.



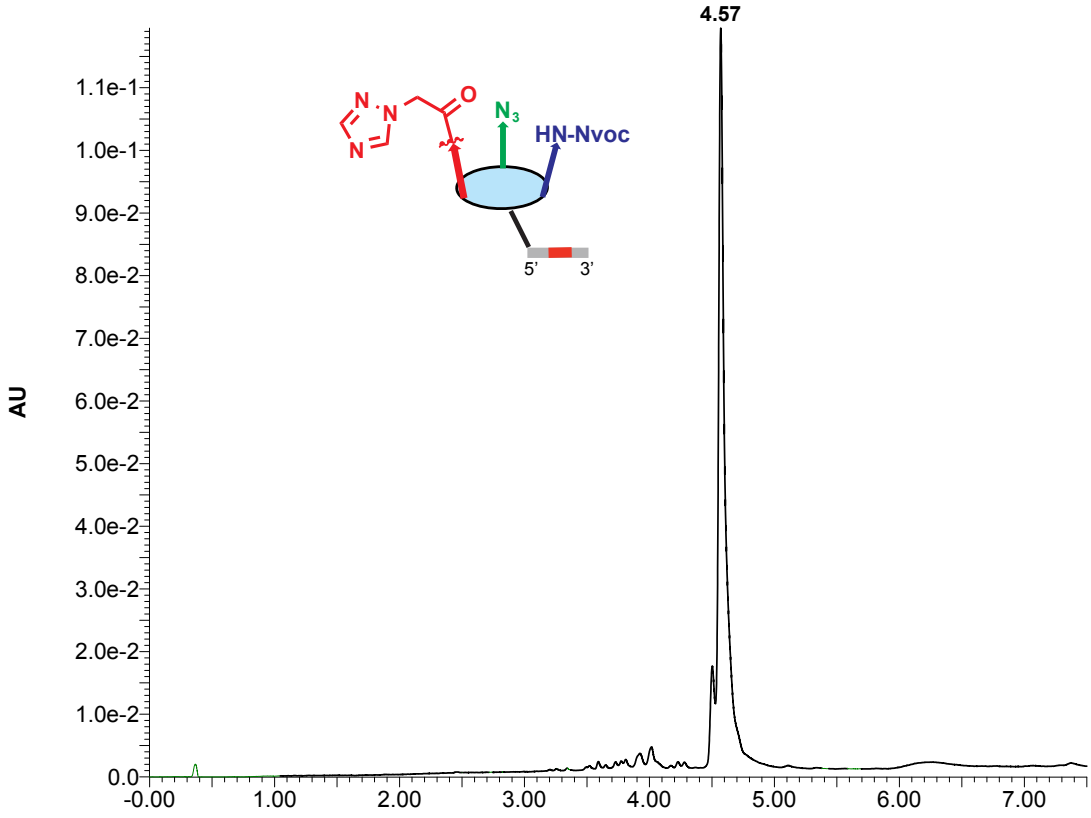
Supplementary Figure 6 | Deprotection of Tfa by ammonium hydroxide solution.

To scaffold-code 1 conjugates (10 nmol), ammonium hydroxide solution (25 % aq) was added and the Tfa deprotection reactions performed at 25 ºC for 2 h. The Tfa-off conjugates were vacuum-dried overnight, redissolved in H_2O (200 µL), and quantified by UV absorption at 260 nm. The recovery yields were over 95 %. All 283 conjugates were characterized by UPLC-MS.

(d) Conjugation of 281 carboxylic acids as 1st diversity elements (DE-1) and characterization.



Supplementary Figure 7 | Conjugation of 1st diversity elements.



Supplementary Figure 8 | Representative UPLC chromatogram of the 1st diversity element conjugate. UPLC analyses were performed on a on a XBridge@ Oligonucleotide BEH C18 10 × 50 mm column at a flow rate of 0.5 mL/min with gradient: 0 % to 5 % B (0 to 0.5 minutes), 5 % to 50 % B (0.5 to 7 minutes), 50 % to 100 % B (7 to 7.1 minutes), 100 % B (7 to 8 minutes), 100 % to 0 % B (8 to 10 minutes) (A= TEA 10 mM, HFIP 5 mM in water, B= MeOH), at 60 ºC. Detection by absorbance at 260 nm.

The conjugates (5 nmol) were immobilized on DEAE sepharose (0.1 mL of slurry, GE Heathcare, Catalog: 17-0709-01). The resin was washed with aq AcOH (3 × 0.5 mL, 10 mM), H_2O (3 × 0.5 mL) and DMSO (3 × 0.5 mL). To the resin-immobilized DNA-conjugate was added a solution of the corresponding carboxylic acid as diversity element (50 mM), EDC-

HCl (50 mM, Aldrich, Catalog: 03449) and HOAt (5 mM, Aldrich Catalog: 41996) in DMSO (0.5 mL). The slurry was agitated at 25 ºC for 2 h. The solution was removed, the resin was washed with DMSO (3 × 0.5 mL) and treated with freshly activated reaction solution. These steps were repeated to reach two coupling steps of 2 h each. The reaction solution was removed and the resin washed with DMSO (3 × 0.5 mL) and aq AcOH (3 × 0.5 mL, 10 mM). The DNA was eluted from the resin by incubation with aq AcOH (3 × 0.2 mL, 3 M) for 30 min each time[4]. The DNA-conjugates were isolated by ethanol precipitation, the pellets were redissolved in deionized water (500 µL), quantified by UV absorption at 260 nm. Recovery yields were around 80 - 90 %. 281 carboxylic acids and 283 DNA tags were employed as first diversity elements and corresponding barcodes (One extra tag was used to encode "Tfa-on" a further tag was used to encode "Tfa-off" [NH2]). All 283 conjugates were characterized by UPLC-MS and conversion yields were all over 80 %, based on the UPLC peak UV integral at 260 nm (see Supplementary Fig. 8). The analysis results are listed as follows:

Number	Smiles	Codon	Predicted (Da)	Found (Da)
1	C\C=C\CCC(O)=O	CACGTT	15435	15436
2	CC(C)CC(O)=O	ATCTAT	15422	15422
3	CCC(C)CC(O)=O	AGAATA	15495	15495
4	CC(C)CCC(O)=O	GTGAGA	15527	15528
5	OC(=O)CC1CCCC1	TCATTA	15449	15449
6	OC(=O)CC1CCCCC1	TTGACT	15479	15479
7	CN(C)CC(O)=O	ACTGTG	15465	15466
8	OCC(O)=O	GCCGCT	15399	15399
9	CC1=CC(C)=C(CCC(O)=O)C=C1	ATAGCT	15524	15524
10	OC(=O)CNC(=O)C1=CC=CC=C1	CACGCA	15495	15495
11	OC(=O)C[C@@H]1C[C@H]2CC[C@@H]1C2	TACCGG	15501	15501
12	CC1=CC=CC=C1CC(O)=O	TGGTTC	15503	15503
13	CC1=CC=C(CC(O)=O)C=C1	ACCACG	15466	15465
14	CC1=CC(CC(O)=O)=CC=C1	TGGTCA	15512	15512
15	OC(=O)CCS(=O)(=O)C1=CC=CC=C1	TGCAAC	15545	15545
16	CC1(C)[C@@H]2CC[C@@]1(C)[C@H](CC(O)=O)[C@H]2O	AAGTAC	15567	15567
17	CN(CC(O)=O)C(=O)C1=CC=CC=C1	AGCCGC	15525	15525
18	CC1=CC=C(SCC(O)=O)C=C1	ATCAAG	15537	15537
19	O[C@H](CC(O)=O)C1=CC=CC=C1	TCCTAG	15488	15488
20	COC1=CC(CCC(O)=O)=C(OC)C=C1	ACTATT	15531	15530
21	CC(O)(CC(O)=O)C(F)(F)F	CCTAAT	15478	15478
22	OCC1=CC=C(CC(O)=O)C=C1	AGTAAT	15536	15536
23	OC(=O)CC1=CC=C2OCOC2=C1	CTAAGT	15526	15526
24	COC1=CC(CC(O)=O)=CC=C1O	GAAGAT	15577	15577
25	COC1=CC(OC)=NC(CCC(O)=O)=N1	ACTGAC	15543	15543
26	OC(=O)CCC1=CC=CN=C1	GAATGC	15522	15522
27	OC(=O)CCCC1=CC=C(I)C=C1	AATTGG	15676	15676
28	COC1=CC2=C(NC(CC(O)=O)=C2)C=C1	ATTAGA	15575	15576
29	OC(=O)C[C@]12CC3CC(C[C@@](O)(C3)C1)C2	CGTGCG	15573	15573
30	OC(=O)CC1=CC(F)=C(F)C(F)=C1	TGGCGT	15568	15569
31	OC(=O)CCNC(=O)C1=CC=C(C=C1)[N+]([O-])=O	TTCTGA	15575	15576
32	CCOC1=CC(CC(O)=O)=CC=C1O	TCACTT	15493	15493
33	COC1=CC2=C(C=C1)C(=O)C(CC(O)=O)C2	GTCACT	15542	15543

34	OC(=O)CCC1CNC2=C1C=CC=C2	TTCACA	15497	15498
35	OC(=O)CC1=CC2=C(S1)C=CC(Cl)=C2	GTCTAA	15572	15573
36	COC1=CC=C(Br)C=C1CC(O)=O	AATCTC	15551	15510
37	OC(=O)CC1NC(=O)NC1=O	AATGGT	15544	15544
38	OC(=O)CCN1C(=O)OC2=C1C=CC=C2	TCTTAA	15528	15528
39	OC(=O)CC1=C(F)C(F)=C(F)C(F)=C1F	CTACAA	15541	15542
40	OC(=O)CC1=CC=C(OC2=CC=CC=C2)C=C1	CTTGGA	15590	15590
41	OC(=O)CCC1=CC=C(N1)C1=CC=CC=C1	CTCAGG	15562	15562
42	CC1=CN(CC(O)=O)C(=O)NC1=O	CCATAA	15499	15499
43	OC(=O)CC1=CC=C(Cl)N=C1	GACGTG	15559	15559
44	NS(=O)(=O)C1=CC=C(NC(=O)CCC(O)=O)C=C1	GAGACA	15652	15652
45	OC(=O)CC1C2=C(C=CC=C2)C2=C1C=CC=C2	GTGATT	15601	15601
46	OC(=O)CC1=CC=CC=C1OC1=CC=CC=C1	CGGCAC	15560	15560
47	NS(=O)(=O)C1=CC=C(NC(=O)CSCC(O)=O)C=C1	TGCGCG	15667	15668
48	CN(CC(O)=O)S(=O)(=O)C1=CC=CC=C1	GTACAT	15575	15576
49	NC(=O)NCCC(O)=O	GCACGA	15488	15488
50	OC(=O)CC1OC2=C(NC1=O)C=CC=C2	CTCGAT	15529	15529
51	OC(=O)CC1=CC2=C(N1)C=CC=C2	TAGCAC	15506	15506
52	CC1=C(CC(O)=O)NC2=C1C=CC=C2	ACACGG	15545	15545
53	COC1=C(CO)C=CC(OCC(O)=O)=C1	AGTTCT	15549	15549
54	CC1=C(CCC(O)=O)C=CC=C1	CTGGTG	15542	15542
55	OC(=O)CCC1=CC(=O)C2=C(O1)C=CC(F)=C2	GCGCAC	15568	15568
56	OC(=O)CCC1=NN=C(O1)C1=CC=CC=C1	CTCCGC	15501	15501
57	OC(=O)CCC1=NOC(=C1)C1=CC=C(Cl)C=C1	TGAGCA	15622	15623
58	OC(=O)CC1=CC=C2C=CC=CC2=C1	TGCCAG	15533	15533
59	OC(=O)CC1=CSC=C1	GAGAAG	15562	15562
60	OC(=O)CCC(=O)C1=CC=CS1	GTAGCA	15555	15555
61	COC1=CC(CC(O)=O)=CC=C1F	CATCTA	15490	15490
62	OCC1=CC=C(OCCC(O)=O)C=C1	TCGATT	15533	15533
63	OC(=O)CCN1C(=O)COC2=C1C=C(Cl)C=C2	GTTAGC	15617	15618
64	OC(=O)CCC#C	TATATC	15419	15419
65	OC(=O)CCN1C=NC2=C(C=CC=C2)C1=O	GAAGTC	15589	15590
66	CC1=NC2=C(C=CC=C2)N1CCC(O)=O	ATGTTA	15565	15565
67	OC(=O)CCC1=NC(=NO1)C1=CN=CC=C1	GCTCGC	15542	15542
68	OC(=O)CCN1C=CC(=O)NC1=O	TCTCTC	15457	15457
69	CC1=CC2=C(C=C1)C(CC(O)=O)C(=O)N2	TTAATA	15550	15550
70	COC1=C(C=CC=C1)C1=NOC(CCC(O)=O)=N1	TTGCTT	15576	15576
71	OC(=O)CCC1=NC(=NO1)C1=CC=CO1	TGAGGC	15595	15595
72	CC1=NC2=CC(F)=CC=C2C(=C1)C(O)=O	CCTCGT	15503	15504
73	CC1=NC2=CC=C(C=C2C=C1)C(O)=O	AACATT	15517	15517
74	OC(=O)C1CC2CCC1C2	AGGTGA	15551	15552
75	OC(=O)C1=C2C=CN=CC2=CC=C1	TCCAGC	15480	15481
76	OC(=O)C1=NC=CC2=CC=CC=C12	TCAGAT	15519	15519
77	OC(=O)C1=CNC(=O)NC1=O	CCAGCT	15463	15463
78	OC(=O)C1=CC=C2NC=CC2=C1	CACGAG	15517	15518
79	OC(=O)CC1=CNC2=CC=C(Br)C=C12	TATGGA	15640	15640
80	CN1N=C(C(O)=O)C2=CC=CC=C12	GCCTCA	15483	15483
81	OC(=O)C1=CC=C2NC=NC2=C1	TAGAGG	15573	15573
82	OC(=O)C1=NC2=C(C=CC=C2)N=C1	GTGTCT	15527	15528
83	OC(=O)C1=CC2=C(C=C1)N=CC=N2	CACCAA	15474	15474
84	CC1=NC(C)=C(CC(O)=O)C(O)=N1	TCACAG	15513	15513
85	OC(=O)CCNC(=O)NC12CC3CC(CC(C3)C1)C2	TTGAAC	15612	15612
86	OC(=O)C1=NNC(=C1)C1CC1	TAATGT	15513	15513
87	OC(=O)C1=CN2C=CSC2=N1	GCTATG	15530	15530
88	CSC1=NC=C(N1CC1=CC=CC=C1)C(O)=O	TACGTA	15509	15509
89	OC(=O)C1=CC=C(C=C1)N1C=CN=C1	CCTCAC	15455	15455
90	CC1=NC2=CC=C(C=C2N=C1C)C(O)=O	CGCCTT	15500	15500

91	OC(=O)CC1CCN(CC2=CC=CC=C2)CC1	GATAGT	15619	15620
92	OC(=O)C1=CC=C(C=C1)C1=NN=NN1	AGCGTC	15537	15537
93	OC(=O)C1=CC=C(CN2C=NN=N2)C=C1	GTCATA	15550	15550
94	CCN1N=C(C=C1C(O)=O)C(C)C	TGTTGT	15550	15551
95	CC(C)C1=CC(=NO1)C(O)=O	GTGTTG	15548	15549
96	OC(=O)C1=CC=C(CN2C=CC=N2)O1	TCTTGC	15505	15506
97	CC1(C)NC(=O)N(CC(O)=O)C1=O	CTATAT	15507	15507
98	OC(=O)CCCN1CC2=CC=CC=C2C1	TGCAGA	15576	15577
99	OC(=O)C1CCN(CC2=CC=CO2)CC1	ATGAGG	15620	15620
100	OC(=O)CCN1CCCCCC1=O	GTAACG	15556	15556
101	NC(=O)CN1CCCC(C1)C(O)=O	TAAGCT	15532	15532
102	OC(=O)C12CC3CC(CC(C3)(C1)N1C=NC=N1)C2	TACAAT	15577	15577
103	OC(=O)C1C2CC(C=C2)C1C(=O)NC1CC1	TCGAGA	15592	15592
104	OC(=O)C1=CN=C(N=C1)C1=CC=CN=C1	CCACTG	15508	15509
105	CC1=CC=C(C=C1N1CCNC1=O)C(O)=O	ATTCGC	15542	15543
106	OC(=O)[C@@H]1C[C@H]1C1=CC=CC=C1	CGTATA	15508	15508
107	CC1(C)C2CCC1(C(O)=O)C(=O)C2	GACAAC	15522	15522
108	CCCN(CCC)[C@@H](C)C(O)=O	TTGCCG	15511	15512
109	CN1[C@@H]([C@H](CC1=O)C(O)=O)C1=CN=CC=C1	TCGGCG	15583	15584
110	OC(=O)[C@@H]1CCC(=O)N1	GCGTTA	15491	15491
111	O[C@@H]1CC(=C[C@@H](O)[C@H]1O)C(O)=O	TGAACG	15545	15545
112	OCC(C(O)=O)C1=CC=CC=C1	TTACTC	15463	15463
113	OC(=O)[C@@H]1CSC(=O)N1	CTAGTA	15493	15493
114	CC(=O)N1C[C@H](O)C[C@H]1C(O)=O	CGATGT	15535	15535
115	CS(=O)(=O)C1=C2NC(=CC2=CC(F)=C1)C(O)=O	CCTTGG	15595	15595
116	CC(C)(O)C#CC1=CC=C(O1)C(O)=O	CCGATA	15525	15525
117	OC(=O)C1CN(CC2=CN=CC=C2)C(=O)C1	CTATCG	15542	15542
118	CC(C)CN1CC(CCC1=O)C(O)=O	TCTGCT	15512	15513
119	CC(=O)C1=C(C)N(CC(O)=O)N=C1C	CCTGAG	15543	15544
120	OC(=O)C1=CC=CC=C1N1CCC(=O)NC1=O	CCACCA	15510	15510
121	OC(=O)C1=NOC(=C1)C1=CC=CC=C1	TTCGTT	15518	15517
122	CC(N(C1=CC=CC=C1)S(C)(=O)=O)C(O)=O	GATCAC	15574	15574
123	COC1=CC=C(C=C1)C1=CC(=NN1)C(O)=O	AGTGAA	15613	15613
124	OC(=O)C1CCN(CC1)C(=O)C1=CC=C(C=C1)[N+]([O-])=O	ACGATC	15609	15609
125	OC(=O)CCCC1=NC(=NO1)C1=CC=NC=C1	TTCCAA	15539	15540
126	CC1=C(NC(=O)C2CCCO2)C=C(C=C1)C(O)=O	AAGAGC	15629	15629
127	CN(C)C(=O)N1CCC(CC1)C(O)=O	TCAACT	15506	15506
128	OC(=O)C1=CC=CO1	AGATAC	15467	15467
129	CC1=CC=C(S1)C(O)=O	CACTCG	15449	15449
130	CC1=C(SC(=N1)C1=CC=NC=C1)C(O)=O	TCCGGT	15558	15558
131	CC1=C(C=NN1C1=CC=CC=C1)C(O)=O	TGTCTT	15530	15531
132	CCOC1=C(C=CC=N1)C(O)=O	AGAGCG	15563	15563
133	OC(=O)C1=CN=C(O)C=C1	TGTCGA	15501	15501
134	CC1=NC=C(C=C1)C(O)=O	CGAGAG	15533	15534
135	OC(=O)C1=CC=C(Br)C=N1	CAGTCA	15533	15533
136	OC(=O)C1=CC=C(OC2=CC=C3OCOC3=C2)N=C1	GCGGAT	15646	15646
137	OC(=O)C1=C2C=CC=CC2=CC=C1	ATAACA	15511	15512
138	OC(=O)C1=CC=CC(OC2=CC=CC=C2)=C1	AGCTCA	15544	15544
139	OC(=O)C1=CC=C2OC(F)(F)OC2=C1	TGGCGT	15580	15580
140	CC(=O)NC1=CC=C(C=C1NC(C)=O)C(O)=O	CACATA	15551	15552
141	COC1=C(OC)C(=CC=C1)C(O)=O	CAGCAG	15577	15578
142	COC1=CC=C(C=C1OC)C(O)=O	CTACAA	15497	15497
143	OC(=O)C1=CC(F)=C(F)C=C1F	GTCCTG	15514	15514
144	OC(=O)C1=CC=C(C=C1)C(F)(F)F	CGCTTA	15512	15512
145	OC(=O)C1=C(F)C=C(Cl)C=C1	GATATA	15544	15544
146	OC(=O)C1=C(F)C(F)=CC=C1	ATTATG	15519	15520

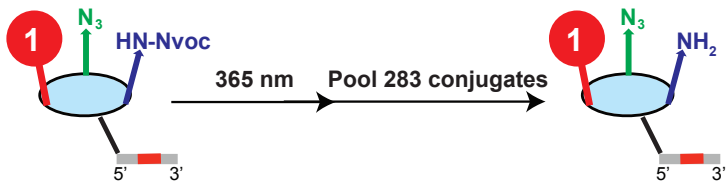
147	OC(=O)C1=CC(=C(F)C=C1)C(F)(F)F	CGATTG	15570	15570
148	OC(=O)C1=CC(F)=C(C=C1)C(F)(F)F	GATTCT	15545	15545
149	OC(=O)C1=CC=C(Cl)C=C1	GAGTGT	15558	15558
150	OC(=O)C1=CC=CC(Cl)=C1	ACCGCA	15472	15472
151	OC(=O)C1=C(Cl)C=CC=C1	GATACG	15527	15527
152	CC1=CC(=CC(C)=C1)C(O)=O	AGTCGT	15512	15512
153	OC(=O)C1=CC=C2C=CC=CC2=C1	ACGGAA	15552	15552
154	OC(=O)C1=NNC(=O)C=C1	TTCATC	15437	15437
155	CC1=NC2=CC=CC=C2N1CC(O)=O	GCATGT	15552	15553
156	OC(=O)C1=CN=CN=C1	TCCATA	15430	15430
157	OC(=O)C1=CSC(=N1)C1=CC=NC=C1	CGGTGG	15609	15609
158	CN(CCC(O)=O)C1CCN(C)C1	GCCAGT	15533	15534
159	OC(=O)C1=CC=CS1	CCAAGG	15484	15484
160	OC(=O)C1=CC=NN1	GAGCGG	15524	15524
161	OC(=O)C1=CNN=C1	CATGTG	15474	15474
162	OC(=O)C1=C(N=CC=N1)C(=O)N1CCCCC1	CAGGAA	15506	15506
163	CC(=O)NC1=CC(=CC(NC(C)=O)=C1)C(O)=O	TAGATA	15606	15606
164	OC(=O)CN1N=C2C=CC=CC2=N1	GCCTGG	15540	15540
165	CCN1CC2=C(C1=O)C(=CC=C2)C(O)=O	GCAGTG	15592	15592
166	CN1C(=O)NC(=O)C2=C(C=C(C)N=C12)C(O)=O	CATAGA	15590	15590
167	CCOCCOC1=CC=C(C=C1)C(O)=O	TCGCAT	15532	15532
168	CCN1CCCC(C)(C1)C(O)=O	CTACTT	15468	15468
169	NC(=O)C1(CC1)C(O)=O	ATCGCG	15476	15476
170	CN(CC(O)=O)C1=NC=NC2=C1NC=N2	TACCTC	15489	15489
171	OC(=O)C1=CC=C2N=CNC(=O)C2=C1	TAGTGA	15576	15576
172	CS(=O)(=O)C1=CC=CC(=C1)C(O)=O	TGTATG	15577	15578
173	OC(=O)C1=CC(=CC=C1)C#C	TACTAC	15452	15453
174	CS(=O)(=O)C1=CC=C(C=C1)C(O)=O	ATCCGA	15531	15531
175	OC(=O)C1=CC=CC=C1	ATTGCA	15468	15468
176	OC(=O)C1(CCC1)C1=CC=CC=C1	ATTCAT	15497	15497
177	CC1(CCCCC1)C(O)=O	TGACAA	15497	15498
178	CC1(CCCC=C1)C(O)=O	TGAATC	15486	15487
179	OC(=O)C1CN(CC2=CC=CC=C2)C1	ATAGAC	15546	15547
180	O[C@@H](C(O)=O)C1=CC=CC=C1	ATCAGT	15498	15498
181	O[C@H](C(O)=O)C1=CC=CC=C1	GATCGA	15523	15523
182	NC(=O)C1=CC=C(C=C1)C(O)=O	GTGCGC	15528	15529
183	CC1=CC=CC(=C1)C(O)=O	GCGCTG	15499	15499
184	CC1=CC=CC=C1C(O)=O	AGATGG	15547	15547
185	OC(=O)C1=CN=CC=C1	CCAGAC	15452	15452
186	OC(=O)C1CCCCC1	TGTGTA	15505	15506
187	CC(C)C(O)C(O)=O	ATGCCT	15440	15440
188	CNC(=O)C1=CC=C(C=C1)C(O)=O	TCGCGG	15542	15542
189	CC(=O)N1CCC(CC1)C(O)=O	ATCGGC	15518	15518
190	CC(=O)C1=CC=C(C=C1)C(O)=O	AAGCTG	15535	15535
191	COC1=C(OC)C=C(C=C1)C(=O)CCC(O)=O	ACTTAT	15518	15518
192	OC(=O)C1=CC(Br)=CC=C1	ACCTGT	15523	15523
193	NS(=O)(=O)C1=CC=C(C=C1)C(O)=O	GCGGCA	15573	15574
194	OC(=O)C1=CC=C(I)C=C1	TAGTTG	15625	15625
195	CC(=O)N[C@H](CC1=CC=CC=C1)C(O)=O	CAGAAC	15547	15548
196	CC(=O)N[C@@H](CC1=CC=C(O)C=C1)C(O)=O	AGGCAT	15594	15594
197	OC(=O)C1CCN(CC1)C(=O)NC1=CC=CC=C1	GTTCCA	15570	15570
198	OC(=O)C1=CC=C(C=C1)C#N	ACGACA	15487	15487
199	OC(=O)C1CCC1	CGGATC	15447	15447
200	CC1=NC(=CS1)C(O)=O	AGCACT	15474	15475
201	OC(=O)C1=CSC(=N1)C(F)(F)F	TATCCA	15503	15503
202	OC(=O)C1CC2=C(C1)C=CC=C2	CGCACG	15494	15494
203	OC(=O)C[C@H]1NC(=O)[C@@H](CC2=CC=CC=C2)N	CCATGC	15569	15570

	C1=O			
204	CC1=C(NC(=O)C2=CC=CO2)C=C(C=C1)C(O)=O	AACAGG	15625	15626
205	CC1=CC=C(NS(=O)(=O)C2=CC=C(C=C2)C(O)=O)C=C1	CATTAG	15637	15638
206	OC(=O)C1=CC(=C(Cl)C=C1)S(=O)(=O)N1CCOCC1	ATTAAC	15635	15636
207	CC1(C)CC2=C(O1)C(OCC(O)=O)=CC=C2	AGCATG	15593	15594
208	OC(=O)C(CC1=CC=CC=C1)NC(=O)C1=CC=CS1	TCGCCA	15582	15583
209	CC(C)CC(NC(N)=O)C(O)=O	GAATCA	15529	15530
210	NC(=O)N1CCCC1C(O)=O	TGATCT	15495	15496
211	OC(=O)C(CC1=CNC2=C1C=CC=C2)NS(=O)(=O)C1=CC=C(Cl)C=C1	AACTTA	15708	15709
212	CC1=CC=C(CN2CC(CC2=O)C(O)=O)C=C1	TCGGTC	15571	15572
213	CC1=C(SC2=C1C(=O)NC=N2)C(O)=O	ATAGTG	15596	15596
214	OC(=O)CN1C=NC=N1	CGCGGT	15490	15490
215	CC(C)CN(C(=O)CCC(O)=O)C1=C(N)N(CC(C)C)C(=O)NC1=O	AACTAG	15709	15710
216	CC(C)C(NC(=O)CC1=CC=CC=C1)C(O)=O	AGCGAG	15631	15631
217	OC(=O)C1CN(CCC2=CC=CC=C2)C(=O)C1	CAGTTC	15555	15555
218	CC(C)C(NC(=O)NC1=CC=CC=C1)C(O)=O	TCTCCG	15534	15535
219	OC1CC(N(C1)S(=O)(=O)C1=CC(Cl)=C(Cl)C=C1)C(O)=O	CCAGGA	15696	15696
220	CC(=O)NC1=NC(C(O)=O)=C(Br)C=C1	CAGTAT	15605	15605
221	OC(=O)C1=C(ON=C1)C1=CC=CC=C1	TGGTAG	15591	15592
222	CC1=NC(C(O)=O)=C(C)O1	CGCGTG	15504	15505
223	OC(=O)C1=CNN=C1C1=C(F)C=CC=C1	AAGACG	15586	15586
224	CN(C1=CC=C(C=C1)C(O)=O)S(=O)(=O)C1=CC=C(C)C=C1	ATGGAT	15691	15691
225	NC1=C(Br)C=C(C=N1)C(O)=O	ATATTC	15538	15538
226	OC(=O)C1=CC=C(C=C1)S(=O)(=O)NCC=C	TGATGA	15627	15628
227	OC(=O)CN1NC(=O)C=CC1=O	TGGAGC	15557	15557
228	CN1C=C(C(O)=O)C(=N1)C(F)F	GCGAGG	15588	15589
229	OC(=O)C1CN(C2CC2)C(=O)C1	CTTCGG	15507	15508
230	OC(=O)C1=CC2=C(N1)C=C(C=C2)C(F)(F)F	AGTAGG	15640	15640
231	OC(=O)C1=CC=C(C=C1)S(=O)(=O)NC1CC1	CTGACA	15572	15573
232	OC(=O)C1CCN(CC1)C(=O)C1=CC=CS1	CGTTAA	15585	15586
233	COC1=C(C=C(C=C1)C(O)=O)S(=O)(=O)NC1CC1	AAGTCT	15617	15618
234	NC(=O)NC1=CC(=CC=C1)C(O)=O	CATCCT	15462	15463
235	CC1=NC2=C(C=CC(=C2)C(O)=O)N1C1CC1	ACCAGA	15556	15557
236	CN1N=C2C=CC=CC2=C1C(O)=O	GATGTT	15553	15554
237	OC(=O)C1=C(C=C(Cl)C=C1)N1CCCC1=O	AGGAGT	15650	15651
238	CC1=NC(=C(S1)C(O)=O)C1=CC=CC=C1	CTGCTC	15517	15518
239	OC(=O)C1=CC2=C(N1)C=CC=C2F	GCTACT	15501	15501
240	CC1=C(C=C(C=C1)S(=O)(=O)NC1CC1)C(O)=O	TGTACA	15601	15602
241	OC(=O)C1=C(C=CC=C1)C1=CC=C(C=C1)C(F)(F)F	CGTAGT	15628	15629
242	OC(=O)C1=C(Cl)C=CC(O)=C1	TATCAT	15493	15494
243	CN1C2=C(N(CC(O)=O)C=N2)C(=O)N(C)C1=O	ACCAAT	15553	15554
244	CC1=C(C(O)=O)C(=NO1)C1=C(Cl)C=CC=C1F	GTTGTC	15608	15609
245	OC(=O)[C@@H]1CC(=O)NC(=O)N1	ATGTGC	15520	15520
246	CC1=C(C(=O)C(O)=O)C2=C(N1)C=CC=C2	CCGCCT	15486	15486
247	CC1=C(CCC(O)=O)C(=O)NC(=O)N1	CGACAT	15529	15529
248	NC(=O)NC(CC(O)=O)C1=CC=CS1	GCTGGA	15601	15601
249	OC(=O)C1(CCCC1)NC(=O)NCC1=CC=CC=C1	GCGTGC	15625	15626
250	CC(C)C(NC(=O)NCC1=CC=CC=C1)C(O)=O	CTCCAG	15557	15558
251	OC(=O)C1=CN=C(C=C1)N1C=NC=N1	GCTAAC	15521	15522
252	OC(=O)C1(CC1)C1=CC(Cl)=CC=C1	TTATCA	15517	15518
253	CC(C)CN1C(=O)NC(=O)C2=C1N=C(C=C2C(O)=O)C1CC1	CGTGTT	15656	15656
254	OC(=O)C1=NN(C(=O)C=C1)C1=CC=CC=C1	ACACTA	15531	15531

255	CC1=C(N=NN1C1=CC=C(F)C=C1)C(O)=O	ATGCAA	15576	15576
256	OC(=O)C1CCCN1C(=O)C1CC1	GAGCTC	15530	15530
257	CN1NC(=O)C2=C1NC(=O)C(CC(O)=O)=C2C	AGGAAC	15517	15517
258	OC(=O)CCC1=NNC(=O)NC1=O	CCGACG	15517	15517
259	CC1=C(CC(O)=O)C(=O)NC(N)=N1	ACTCAA	15498	15498
260	OC(=O)C1=NNC2=C1CCC2	CTTATT	15464	15464
261	OC(=O)C1=NNC2=C1CCCCC2	ACAGGC	15536	15536
262	CC1=C(C=NC=N1)C(O)=O	TTAAGG	15524	15524
263	CN1C2=C(NC(CCC(O)=O)=N2)C(=O)NC1=O	GACTAT	15584	15584
264	OC(=O)C1(CCOCC1)C1=CC=C(F)C=C1	GAACAG	15604	15604
265	OC(=O)C1CC1C(=O)N1CCN(CC1)C1=CC=CC=C1	CCTGCA	15581	15581
266	OC(=O)C1=CC(=CC=C1)N1NC(=O)C=CC1=O	CTGAAT	15578	15578
267	NC(=O)C1=CC=C(S1)C(O)=O	CGGCCG	15519	15519
268	OC(=O)C1CCCN1C(=O)C1=CC=C(Br)C=C1	GAGGAC	15694	15694
269	OC1CC(N(C1)C(=O)C1=CC=C(F)C=C1)C(O)=O	CACTGT	15575	15575
270	OC(=O)C1=CC=C(C=C1)C(F)F	CGACGG	15544	15544
271	OC(=O)CCN1C=CNC(=O)C1=O	TATAAG	15554	15554
272	OC(=O)CN1C=C2C=CC=CC2=N1	AGTTAG	15562	15562
273	OC(=O)C1=C2C=CC=CN2N=C1	ACTAGC	15493	15493
274	CC1=NC(=CN1)C(O)=O	TTGGTA	15503	15503
275	OC(=O)CC1(O)CCCCC1	CGCTAC	15465	15466
276	OC(=O)C1(CC1)C1=C(F)C=CC=C1	GTTAAT	15541	15541
277	CN1N=C(C(O)=O)C(Br)=C1C	ACGAAG	15599	155991
278	CC1=NC(=NO1)C1=CC(=CC=C1)C(O)=O	ATAATT	15549	15549
279	OC(=O)C1=C(Br)SC=N1	GTAGAG	15619	15619
280	COC1=C(OC)C=C(C=C1)C1(CCCC1)C(O)=O	GTTCGT	15603	15604
281	OC(=O)C1=CNC(=O)C(Br)=C1	CGCAGC	15550	1549
282	OC(=O)C(F)(F)F	TACTCA	15420	15420
283	[NH2]	CAGACT	15349	15349

Supplementary Table 1 | Structures and codons of the 1st diversity elements with deconvoluted predicted mass and found mass of the corresponding conjugates.

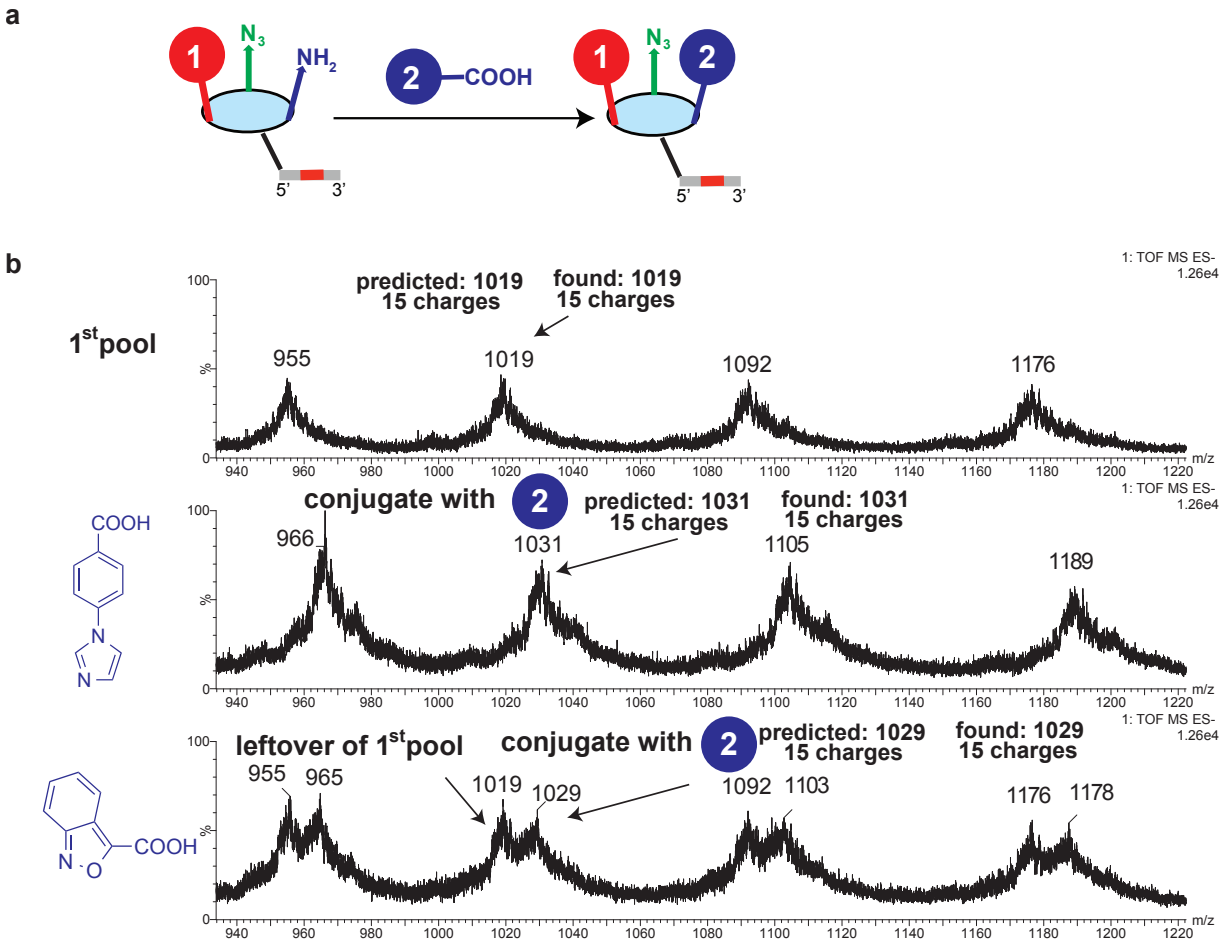
(e) Deprotection of Nvoc by UV irradiation and assembling of the 1st pool.



Supplementary Figure 9 | Deprotection of Nvoc by UV irradiation at 365 nm.

The individual conjugates with the first diversity elements were subjected to irradiation at 365 nm for 60 min at 0 °C (on ice). All 283 conjugates were characterized by UPLC-MS and conversion yields exceeded 95 % based on the UPLC UV peak integral of absorption at 260 nm. Equimolar amounts (3 nmol) of the 283 conjugates obtained as described above were combined to generate the 1st pool without further purification.

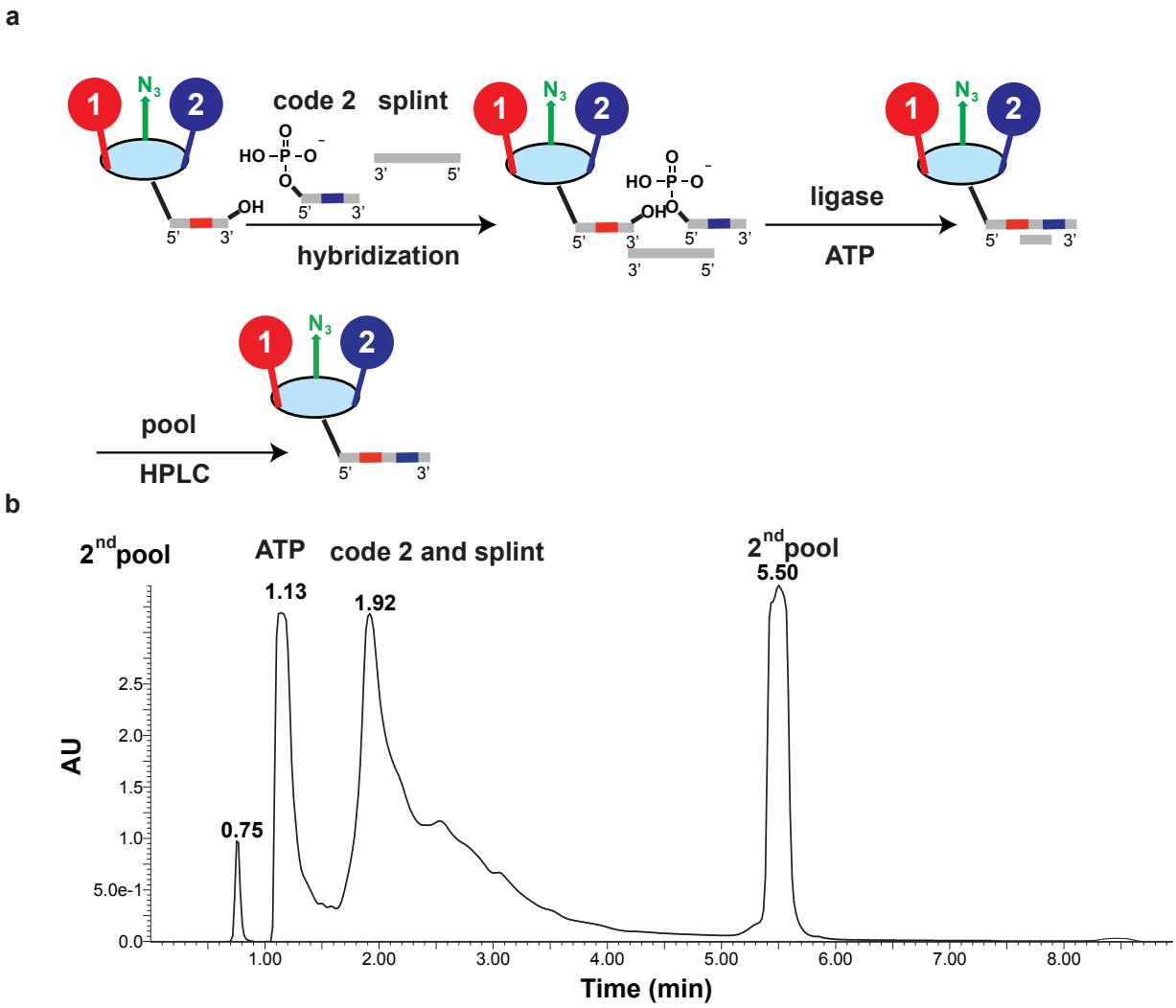
(f) Conjugation of 384 carboxylic acids as second diversity elements (DE-2), encoding by enzymatic ligation, assembling of the 2nd pool, purification and characterization.



Supplementary Figure 10 | Conjugation of the 2nd diversity elements. a, Conjugation of 2nd diversity elements to 1st pool. **b,** MS spectra of 1st pool, conjugates from 1st pool coupled to 4-(1*H*-imidazol-1-yl)benzoic acid and conjugates from 1st pool coupled to benzo[c]isoxazole-3-carboxylic acid, using 50 pmol for injection. MS analyses were performed on a Xevo G2-XS Q-TOF with electrospray ionization source. For 4-(1*H*-imidazol-1-yl)benzoic acid as diversity element, the conversion is over 80 % according to the peak intensity ratio between the conjugates and the starting pool. For benzo[c]isoxazole-3-carboxylic acid as diversity element, the conversion is ca. 40 % according to the peak intensity ratio between the conjugates and the starting pool.

The 1st pool was further split into 386 aliquots (2 nmol each) and conjugated with 384 carboxylic acids using the same conditions as described above. The DNA-conjugates were isolated by ethanol precipitation and the pellets redissolved in deionized water 200 (μL), the recovery of the conjugates was quantified by UV absorption at 260 nm and the yields determined to be around 60-90 %. All the 384 conjugates were characterized by UPLC-MS and conversion yields were all over

80 % based on the ratio of MS peak intensities of the starting pool (1st pool here) and the conjugates from the 2nd diversity elements (see Supplementary Fig. 10)[5].



Supplementary Figure 11 | Encoding of the 2nd diversity elements and purification. a Enzymatic encoding for the 2nd diversity elements and HPLC purification. **b** Representative HPLC chromatogram for the purification of the 2nd pool. HPLC purifications were performed on a CT18-XTerra 10 × 150 mm column at a flow rate of 4 mL/min with gradient: 10 % B (0 to 4 minutes), 60 % B (4.1 to 7 minutes), 10 % B (7.1 to 9 minutes), (A= TEAA 0.1 M in water, B= CH_3CN 80 % in water), at 65 ºC. Detection by absorbance at 260 nm.

Individual conjugates (1.2 nmol, in deionized water 120 µL) with second diversity elements were employed to splint-assisted enzymatic ligation without further purification. To these conjugates, 5'-phosphate-oligos (code 2: 5'-phosphate-CGGATCGACGNNNNNNNGCGTCAGGCAGC-3'-OH, 1.44 nmol, 1.2 equiv.) and splint (5'-OH-CGTCGATCCGGC GCCATGGG-3'-OH, 1.8 nmol, 1.5 equiv.) were added, followed with addition of T4 DNA ligase buffer (15 µL, New England BioLabs, Catalog: B0202S) and kept at 65 °C for 10 minutes. After the system cooled down to 22 °C, T4 DNA ligase (1 µL, 400 units, New England BioLabs, Catalog: M0202S) was added and splint-assisted enzymatic ligation reactions stood at 22 °C

for 6 h. The conversion yields were all over 80% based on denaturing PAGE analysis. 384 carboxylic acids and 386 DNA tags were employed as second diversity elements and corresponding barcodes (One extra tag was used to encode "Nvoc-on" and another tag was used to encode "Nvoc-off" [NH2]. Equimolar amounts (1.2 nmol) of the 386 conjugates obtained as described above were combined to generate the 2nd pool followed with ethanol precipitation. The recovered pellet was redissolved in deionized water (15 mL) and purified by HPLC to remove the splint and excess code 2 (see Supplementary Fig. 11). The HPLC purified 2nd pool was dried by lyophilization, redissolved in H_2O (10 mL), quantified the recovery conjugates by UV absorption at 260 nm and yielded a total of 418 nmol of conjugates. The MS-analysis after splint-assisted enzymatic ligation with code 2 is listed below:

Number	Smiles	Codon	Predicted (Da)	Found (Da)
1	CC\C=C/CC(O)=O	GTCTCAC	1014	1016
2	OC(=O)CCC1CCCC1	GTCGTAC	1017	1018
3	CC(C)(C)CC(O)=O	CTCATTG	1015	1015
4	CC(CC(O)=O)CC(C)(C)C	GTAGAGA	1021	1021
5	CC(CC(O)=O)C=C	GTTACCT	1015	1015
6	CC(=O)[C@@H]1C[C@@H](CC(O)=O)C1(C)C	AGTAATT	1020	1020
7	OC(=O)CCC=C	AGTGAGC	1017	1017
8	CN(C)CC(O)=O	ACTGATA	1016	1015
9	CCOCCC(O)=O	ACGTATA	1016	1017
10	OCC(O)=O	TTCTCCT	1011	1011
11	OC(=O)CNC(=O)C1=CC=CC=C1	AACTGCT	1018	1018
12	OC(=O)CCC1=CC=CC=C1	CACACAC	1015	1015
13	CCCCNC(=O)NCC(O)=O	CACGTGT	1018	1019
14	OC(=O)CCS(=O)(=O)C1=CC=CC=C1	CGAGGTG	1023	1026
15	CC1(C)[C@@H]2CC[C@@]1(C)[C@H](CC(O)=O)[C@H]2O	AAGCGAG	1022	1023
16	OC(=O)CCOC1=CC=CC=C1	GCGCATG	1019	1021
17	O[C@H](CC(O)=O)C1=CC=CC=C1	GTTGGTC	1019	1020
18	C[C@@H](NC(=O)CCC(O)=O)C1=CC=CC=C1	CTTCTCT	1017	1018
19	CN(C)C1=CC=C(CC(O)=O)C=C1	TTGCACG	1019	1019
20	COC1=CC=CC=C1CC(O)=O	GAGTAGA	1021	1021
21	COC1=CC(OC)=C(CCC(O)=O)C=C1	ATGTGAG	1022	1023
22	COC1=CC=CC(CCC(O)=O)=C1OC	AACGTAT	1020	1021
23	CC1=CC=CC=C1C(=O)NCC(O)=O	TCCGGCT	1018	1019
24	CC1=CC(=CC=C1)C(=O)NCC(O)=O	TGATGAT	1021	1021
25	CC1=CC=C(C=C1)C(=O)NCC(O)=O	TGTGGAC	1021	1022
26	CC(O)(CC(O)=O)C(F)(F)F	GTAGTGC	1020	1021
27	OCC1=CC=C(CC(O)=O)C=C1	GCAACAC	1017	1017
28	COC1=CC(CC(O)=O)=CC=C1O	AAGACCG	1019	1020
29	OC(=O)CCC1=CC2=C(OCC2)C=C1	AGAGAGA	1022	1023
30	COC1=CC=C(CC(O)=O)C(OC)=C1OC	TCGAGAT	1021	1022
31	COC1=CC(OC)=NC(CCC(O)=O)=N1	CCGACTT	1018	1019
32	OC(=O)CCC1=CC=CN=C1	TGAGATA	1019	1020
33	OC(=O)CCCC1=CC=C(I)C=C1	TTGGCGT	1024	1026
34	OC(=O)C[C@]12CC3CC(C[C@@](O)(C3)C1)C2	AATCCTC	1017	1018
35	COC1=CC(OC)=CC(CCC(O)=O)=C1	CACGTAC	1019	1019
36	CCOC1=CC(CC(O)=O)=CC=C1O	CACACGA	1018	1020
37	COC1=CC2=C(C=C1)C(=O)C(CC(O)=O)C2	CGTAACA	1020	1020

38	OC(=O)CCC1CNC2=C1C=CC=C2	AATTCCG	1018	1019
39	COC1=CC=C(Br)C=C1CC(O)=O	GCGTTAC	1021	1022
40	CC1=CC=C(F)C=C1CC(O)=O	CTCCATT	1015	1016
41	OC(=O)CC1NC(=O)NC1=O	CGCCGGT	1018	1018
42	OC(=O)CCN1C(=O)OC2=C1C=CC=C2	GTAAGAC	1021	1022
43	CC1=CN(CC(O)=O)C(=O)NC1=O	GCTGAAT	1020	1020
44	COC1=C(OC)C=C(CC(O)=O)C(Br)=C1	ATAAGGT	1025	1025
45	COC1=CC2=C(NC(CC(O)=O)=C2C)C=C1	ATCATTC	1018	1020
46	NS(=O)(=O)C1=CC=C(NC(=O)CCC(O)=O)C=C1	AGCGAGT	1024	1025
47	COC1=CC=C(CC(O)=O)C=C1O	CCAGACT	1017	1018
48	CN1C2N=CN(CC(O)=O)C2C(=O)N(C)C1=O	TGACCAG	1021	1023
49	OC(=O)CC1=CC=CC=C1OC1=CC=CC=C1	GCCTACA	1019	1020
50	NS(=O)(=O)C1=CC=C(NC(=O)CSCC(O)=O)C=C1	GCCTCGT	1023	1023
51	CN(CC(O)=O)S(=O)(=O)C1=CC=CC=C1	TGTCGTT	1021	1021
52	NC(=O)NCCC(O)=O	GTCTGAA	1018	1018
53	OC(=O)CCCC1=CC=C2OCCOC2=C1	CCGTACT	1019	1018
54	OC(=O)CC1OC2=C(NC1=O)C=CC=C2	AGGTGTC	1021	1022
55	OC(=O)CC1=CC2=C(N1)C=CC=C2	TGCCTGG	1019	1019
56	CC1=C(CC(O)=O)NC2=C1C=CC=C2	AGGATGC	1021	1021
57	COC1=C(CO)C=CC(OCC(O)=O)=C1	GTTATGC	1021	1021
58	COC1=CC(CC(O)=O)=CC(OC)=C1OC	GTAGGAA	1024	1024
59	OC(=O)CCC1=CC2=C(OCO2)C=C1	GTGTCGT	1020	1021
60	OC(=O)CCC1=NC(=NO1)C1=C(F)C=CC=C1	GCTCCTT	1019	1019
61	CC1=CC=C(C=C1)C(=O)CCC(O)=O	TTCTGAG	1020	1020
62	OC(=O)CCC(=O)C1=CC=CC=C1	TCATGGA	1019	1019
63	COC1=C(CCC(O)=O)C=CC=C1	ATCGTAA	1019	1019
64	OC(=O)CCC1=CC(=O)C2=C(O1)C=CC(F)=C2	GACTTAT	1021	1021
65	OC(=O)CCC1=CC(=O)C2=C(O1)C=CC(Br)=C2	CAACGTT	1023	1023
66	OC(=O)CCC1=NN=C(O1)C1=CC=CC=C1	CGATACT	1019	1020
67	OC(=O)CCC1=CC=CS1	TACGATG	1019	1019
68	OC(=O)CC1=CSC=C1	CCAGTGT	1017	1018
69	OC(=O)CCC(=O)C1=CC=CS1	CCTGGTG	1019	1020
70	COC1=CC=C(C=C1F)C(=O)CCC(O)=O	CCAGTTG	1020	1020
71	COC1=CC=C(C=C1)C(=O)CCC(O)=O	TCATCGT	1019	1019
72	OC(=O)CCC1=C(Cl)C=C(Cl)C=C1	ATATATC	1019	1020
73	OC(=O)CCC(=O)C1=CC=C(F)C=C1	GTGCCGA	1020	1021
74	OC(=O)CCC(=O)C1=CC(F)=CC(F)=C1	CAGACCA	1019	1020
75	OC(=O)CCC(=O)C1=CC=C(Cl)C=C1	CGATTGC	1020	1021
76	OC(=O)COC1=CC=C(Cl)C=C1	TACCTAC	1017	1017
77	OC(=O)CC1CCC=C1	GATGAGC	1018	1019
78	OC(=O)CCN1C(=O)COC2=C1C=C(Cl)C=C2	CAGGTTC	1022	1022
79	OC(=O)CCN1C(=O)COC2=C1C=CC=C2	ATAACTA	1020	1020
80	CCC1CCC(CC1)C(=O)NCC(O)=O	ACGTCCG	1019	1020
81	OC(=O)CCC1NC(=O)NC1=O	GTGCATA	1019	1020
82	CC1=NC2=C(C=CC=C2)N1CCC(O)=O	GAATCAA	1020	1021
83	OC(=O)CCC1=NC(=NO1)C1=CN=CC=C1	ACTTGCG	1020	1021
84	OC(=O)CCN1C=CC(=O)NC1=O	TGTTCGT	1019	1019
85	CC(=O)C1=C(C)N(CCC(O)=O)N=C1C	CCTCCGC	1017	1017
86	CC1=CC2=C(C=C1)C(CC(O)=O)C(=O)N2	CTCATAT	1018	1018
87	OC(=O)CCC(=O)NC1CCCCCC1	CTGAAGG	1022	1023
88	CN(C1CCCCC1)C(=O)CCC(O)=O	TCTAGCT	1019	1020
89	COC1=CC=CC=C1C1=NOC(CCCC(O)=O)=N1	TTCCTGT	1021	1021
90	OC(=O)CCC1=NC(=NO1)C1=CC=CO1	GACTGGA	1022	1022
91	CC1=C(C=CC=C1)C1=NOC(CCC(O)=O)=N1	TTAACCG	1020	1021
92	CC1=CC=C2N=C(C)C=C(C(O)=O)C2=C1	AGACTGA	1021	1021

93	OC(=O)C1=NC2=CC=CC=C2C=C1	ATCTTGC	1017	1018
94	OC(=O)C1=C2C=CC=NC2=CC=C1	CTAAGGC	1019	1019
95	OC(=O)C1=C2C=CN=CC2=CC=C1	ATCGCAT	1018	1018
96	OC(=O)C1=NC=CC2=CC=CC=C12	AAGTCCA	1018	1019
97	OC(=O)C1=CN=C2C=CC=CN2C1=O	CATTACG	1018	1019
98	OC(=O)C1=CC2=CC(F)=CC=C2N1	ATAGCCT	1018	1018
99	OC(=O)C1=CC=C2NC=CC2=C1	CCAGGTA	1018	1019
100	OC(=O)C1=CC=C2C=CNC2=C1	AGTAGTA	1020	1021
101	OC(=O)C1=CC2=C(N1)C=CC(O)=C2	TATGGAG	1021	1022
102	COC1=CC=C2NC(=CC2=C1)C(O)=O	AGCACGA	1020	1019
103	OC(=O)CN1C=C(C(=O)C2CC2)C2=CC=CC=C12	AATTGCA	1022	1022
104	CN1C=CC2=CC(=CC=C12)C(O)=O	CAGATTG	1019	1020
105	OC(=O)C1=CC2=CC=CC=C2N1	GTCCAAG	1018	1019
106	CN1N=C(C(O)=O)C2=CC=CC=C12	ATGCGCT	1018	1019
107	OC(=O)C1=CC=C2NC=NC2=C1	ACATAGT	1018	1019
108	OC(=O)C1=CC=C(CN2C=NC3=CC=CC=C23)C=C1	ATAGAGC	1023	1023
109	CC(CC1=CC=C2OCOC2=C1)C(O)=O	TATGTCG	1020	1021
110	OC(=O)C1=CC=C2OCCOC2=C1	TATCATT	1017	1018
111	OC(=O)C1=NC2=C(C=CC=C2)N=C1	CCTTCCG	1016	1016
112	OC(=O)C1=CC2=C(C=C1)N=CC=N2	TGCGTCG	1019	1019
113	CC1=NC(C)=C(CC(O)=O)C(O)=N1	AGAAGTG	1115	1115
114	OC(=O)C1=CN=C(N=C1)C1=CC=CS1	CGTAGGA	1022	1022
115	OC(=O)C1=CN=C(N=C1)N1CCOCC1	TCCGGTC	1019	1019
116	COC1=CC(C(O)=O)=C(OC)N=N1	ACGATCA	1018	1019
117	OC(=O)C1=C2C(=CC=C1)C(=O)C1=C2C=CC=C1	TCGTACA	1020	1020
118	OC(=O)CCNC(=O)NC12CC3CC(CC(C3)C1)C2	CTATTAT	1021	1023
119	OC(=O)C12CC3CC(CC(Cl)(C3)C1)C2	CGCAGGC	1020	1021
120	OC(=O)C1=NNC(=C1)C1CC1	TAGCTTC	1016	1017
121	OC(=O)C1=CC2=C(O1)C=CC=C2	CCTTCTC	1014	1015
122	COC1=CC=CC2=C1OC(=C2)C(O)=O	CTGGCCG	1019	1020
123	OC(=O)C1=CC2=C(OC=C2)C=C1	GTCGAGC	1019	1019
124	CC1=C(C=C(O1)S(=O)(=O)N1CCOCC1)C(O)=O	TAGATTA	1023	1024
125	OC(=O)C1=CN2C=CSC2=N1	TTGATAC	1018	1018
126	OC(=O)C1=CC=C(C=C1)N1C=CN=C1	GCACAAT	1019	1019
127	C[C@H]1CC[C@@H](CC1)C(O)=O	GCTTGAG	1019	1019
128	CC1=NC2=CC=C(C=C2N=C1C)C(O)=O	TACTTGG	1020	1022
129	OC(=O)CC1CCN(CC2=CC=CC=C2)CC1	CCACATA	1019	1021
130	OC(=O)C1CCN(CC1)C(=O)C1=CC=C(F)C=C1	GACAGTC	1022	1023
131	COC1=CC=CC(OC)=C1C(=O)N1CCC(CC1)C(O)=O	CGCGTTA	1023	1024
132	COC1=CC(=CC(OC)=C1)C(=O)N1CCC(CC1)C(O)=O	ATCTCCG	1022	1022
133	OC(=O)CN1C=C(Cl)C(=O)C(Cl)=C1	CTTGCAC	1019	1019
134	OC(=O)CN1C=C(I)C(=O)C(I)=C1	TGTCACT	1027	1027
135	CCCC(=O)C1=CN(CC(O)=O)C2=CC=CC=C12	GTGCGTG	1024	1024
136	OCCN1C=NC2=CC(=CC=C12)C(O)=O	ACGCATC	1018	1019
137	CC1=CC=C2NC(=O)C(CC(O)=O)C2=C1	GACGCGT	1021	1021
138	CC1=CC(=O)N(CC(O)=O)C2=CC=CC=C12	AGCGACG	1022	1022
139	OC(=O)C1=CC=C(C=C1)C1=NN=NN1	GCCGTAG	1113	1114
140	OC(=O)C1=CC(=CC=C1)C1=NN=NN1	CTCAGCA	1110	1111
141	OC(=O)C1=CC=C(CN2C=NN=N2)C=C1	CTTACCA	1017	1019
142	CC(C)C1=CC(=NO1)C(O)=O	GCAGGTG	1020	1021
143	CN1C=CC(=C1C)C1=CC(=NO1)C(O)=O	CCGGCTG	1020	1020
144	OCCC1=CN2N=C(C=C2N=C1)C(O)=O	CAACAAC	1018	1019

145	OC(=O)C1=CC=C(CN2C=CC=N2)O1	CGTTCAG	1019	1020
146	OC(=O)C1CCN(CC2=CC=CO2)CC1	CCACGAA	1019	1019
147	CC1=NN(CC(O)=O)C(=O)C2=CC=CC=C12	CGGAGAG	1023	1024
148	OC(=O)C12CC3CC(CC(C3)(C1)N1C=NC=N1)C2	GTCATGA	1022	1023
149	CC1=CC=C(C=C1)N(CC(O)=O)S(C)(=O)=O	ACTGACG	1022	1022
150	OC(=O)C1C2CC(C=C2)C1C(=O)NC1CC1	TGACGGA	1022	1023
151	OC(=O)C1=CN=C(N=C1)C1=CC=CN=C1	TCTTATT	1018	1019
152	CC1=CC=CN2C(CC(O)=O)=CN=C12	ATACTAC	1018	1018
153	CC1=CC=C(C=C1N1CCNC1=O)C(O)=O	TAGCCGT	1020	1021
154	OC(=O)[C@@H]1C[C@H]1C1=CC=CC=C1	TCCACGG	1017	1018
155	OC(=O)C1COC2=CC=CC=C2O1	CGGCTTC	1018	1018
156	CN1[C@@H]([C@H](CC1=O)C(O)=O)C1=CN=CC=C1	TGTGCTT	1020	1021
157	CC(C)[C@H]1CC[C@@H](CC1)C(O)=O	TTGTCTT	1017	1018
158	OC(=O)[C@@H]1CCC(=O)N1	ATAGTCA	1017	1017
159	O[C@@H]1CC(=C[C@@H](O)[C@H]1O)C(O)=O	TGGAGTA	1021	1021
160	OCC(C(O)=O)C1=CC=CC=C1	CGGATGG	1021	1021
161	OC(=O)[C@@H]1CSC(=O)N1	GCTACCA	1016	1016
162	CC1(C)C(C(O)=O)C1(C)C	ACAACGA	1017	1018
163	COC1=CC=C2OC(=CC2=C1)C(O)=O	TTGATGA	1021	1021
164	OC(=O)C1=CC2=CC(Cl)=CC=C2O1	GAGCCGC	1020	1020
165	CC(C)CN1CC(CC1=O)C(O)=O	GATTAAT	1020	1020
166	COC1=CC=CC(=C1)N1C=C(C=N1)C(O)=O	GATGCGG	1023	1023
167	OC(=O)C1CN(C2CCCC2)C(=O)C1	TAACGTA	1020	1019
168	CCCN1CCC(CC(O)=O)CC1	ACGCGTT	1019	1020
169	COC1=C2OCC(CC2=CC=C1)C(O)=O	CGGCCGT	1020	1019
170	CC1=CC(C)=C(C(O)=O)C(=O)N1	ATGGTTA	1020	1019
171	CCCN1N=CC(C(O)=O)=C1C	GTCGCCG	1018	1019
172	OC(=O)C1=CC=C(NC(=O)C2CCC=CC2)C=C1	TACGTCA	1021	1021
173	CCC(N1C=CC=N1)C(O)=O	AGCTCTT	1016	1017
174	CC(C)C1=NC(C)=C(S1)C(O)=O	CATTGTT	1018	1019
175	CC(C)C1=CC(=NC2=NC=NN12)C(O)=O	ACAGGAA	1021	1022
176	OC(=O)C1CCN(CC1)C(=O)C1CCCC1	TCTATGC	1019	1021
177	CCC(=O)N1CCCC(C1)C(O)=O	TAATACA	1018	1019
178	OC(=O)C1=CC=CC=C1N1CCC(=O)NC1=O	TACTCGA	1020	1021
179	OC(=O)C1=NOC(=C1)C1=CC=CC=C1	TGTCTAG	1020	1020
180	CC(CC(O)=O)N1N=C(C)C(C(C)=O)=C1C	TTCGCTC	1018	1019
181	OC(=O)CCCC1=NC(=NO1)C1=CC=NC=C1	GCTCAGT	1021	1021
182	OC(=O)C1=CC(NC(=O)C2CCCC2)=CC=C1	TACAGGA	1022	1023
183	CN(C)C(=O)N1CCC(CC1)C(O)=O	CCGGAAG	1021	1022
184	CC1=CC(C)=C2C(CC(O)=O)C(=O)NC2=C1	CCACGCT	1018	1018
185	CC1=C(C)N2C(S1)=NC=C(C(O)=O)C2=O	AGATAAG	1023	1023
186	OC(=O)C1CCCC1	ATCAGAC	1016	1016
187	OC(=O)C1=CC(Br)=C(Br)O1	ATCTCAA	1021	1022
188	CN(C)S(=O)(=O)C1=CC(C(O)=O)=C(C)O1	GTAACTC	1020	1021
189	OC(=O)C1=COC=C1	ACAGTAG	1017	1018
190	OC(=O)C1=C(Cl)SC(Cl)=C1	CGCTGTA	1019	1021
191	OC(=O)C1=CC=C(Cl)S1	GAGCCAT	1018	1019
192	CC1=NC(C)=C(S1)C(O)=O	ACTCTTG	1017	1018
193	CC1=C(SC(=N1)C1=CC=NC=C1)C(O)=O	TGTAGAG	1023	1024
194	CN1N=C(C=C1C(O)=O)C(C)(C)C	AGAACAA	1020	1020
195	CN1C=CC=C1C(O)=O	ATCTACT	1015	1015
196	OC(=O)C1=CC=C(N=C1)C(F)(F)F	GATAACG	1020	1021
197	CC1=CC=CN=C1C(O)=O	TTACGCT	1016	1016
198	OC(=O)C1=CN=CC(O)=C1	TAGACGA	1018	1019
199	OC(=O)C1=C(OC2=CC=CC=C2)N=CC=C1	AAGCATG	1021	1022

200	OC(=O)C1=CC=C(Br)C=N1	GACTTCA	1019	1019
201	OC(=O)C1=CC(=CN=C1)C1=CC=CS1	ATGTGTT	1021	1021
202	CC(C)OC1=CC=C(C=C1)C(O)=O	GTATGGA	1021	1022
203	OC(=O)C1=C2C=CC=CC2=CC=C1	CGGTTAA	1019	1021
204	OC(=O)C1=CC=C(OC2=CC=CC=C2)C=C1	TTACGAG	1021	1021
205	CC(=O)NC1=CC=C(C=C1NC(C)=O)C(O)=O	ATTCTGG	1022	1022
206	COC1=CC(=CC(OC)=C1)C(O)=O	CTATGCG	1019	1019
207	CN(C)C1=CC=CC(=C1)C(O)=O	CCGTTAT	1017	1017
208	COC1=CC(=CC(OC)=C1OC)C(O)=O	ACGCTGT	1020	1020
209	OC(=O)C1=CC(=CC=C1)C(F)(F)F	CTAGTGA	1020	1020
210	OC(=O)C1=CC(Cl)=C(F)C=C1	TTACAGG	1019	1020
211	OC(=O)C1=CC(Cl)=C(Cl)C=C1	GTAGCCA	1019	1020
212	OC(=O)C1=CC=C(F)C(F)=C1	CCTGTCG	1017	1016
213	OC(=O)C1=C(F)C(=CC=C1)C(F)(F)F	GCACGTC	1019	1020
214	OC(=O)C1=CC=C(Cl)C=C1	TTGGCTA	1018	1019
215	COC1=CC2=CC=C(C=C2C=C1)C(O)=O	TTAAGTG	1021	1022
216	NC(=O)COC1=CC=CC(=C1)C(O)=O	TAGGAAC	1021	1021
217	OC(=O)C1=NNC(=O)C=C1	GTGATGG	1020	1021
218	OC(=O)C1=CN=CN=C1	ACAAGAG	1018	1018
219	CN(CCC(O)=O)C1CCN(C)C1	ACATGAT	1019	1020
220	CC1(CC(O)=O)N2C=CC=CC2=NC1=O	TTGCCTG	1019	1019
221	CC1=NNC(=C1)C(O)=O	GAGTCTT	1017	1017
222	OC(=O)C1=CC=NN1	CGCGGAG	1018	1018
223	OC(=O)C1=CNN=C1	ATCTGTC	1015	1015
224	OC(=O)C1=C(N=CC=N1)C(=O)N1CCCCC1	CGTAGTG	1023	1023
225	CC1=CC=CC2=NC(=O)CC(C)(N12)C(O)=O	GAGCAGG	1023	1024
226	OC(=O)C1C2OC3(CN(CC4=NC=CC=C4)C(=O)C13)C=C2	AGCTGAA	1024	1025
227	CN1C(=O)C(=NC2=C1C=CC=C2)C(O)=O	TCGCCGC	1018	1018
228	CC1=NOC2=NC(=CC(C(O)=O)=C12)C1CC1	CGCCTGC	1019	1020
229	OC(=O)C1=CC2=C(NC(=O)C(=O)N2)C=C1	AACGACT	1112	1114
230	OC(=O)C1CCC=CC1	GCGTATT	1017	1017
231	CC1=NN2C(=C1)N=CC(C(O)=O)=C2C	CACGGTC	1018	1019
232	OC(=O)C1=CC=C(OC2=CC=CN=C2)O1	GAGCAAC	1020	1021
233	CS(=O)(=O)C1=CC=CC(=C1)C(O)=O	ATTAAGT	1020	1021
234	CS(=O)(=O)C1=CC=C(C=C1)C(O)=O	GAGTATG	1022	1023
235	CC(C)(C)C(O)=O	CCTTGGC	1014	1017
236	CC1(CC1)C(O)=O	CTTCCTC	1012	1012
237	CC1(CCCCC1)C(O)=O	TTGGTCG	1018	1019
238	OC(=O)C1CN(CC2=CC=CC=C2)C1	GTGGCTT	1020	1022
239	NC(=O)C1=CC=C(C=C1)C(O)=O	TCAACTC	1016	1016
240	CC1=CC=C(C=C1)C(O)=O	GTGAGTC	1018	1020
241	OC(=O)C1=CN=CC=C1	AATGATG	1018	1019
242	COC1=CC=C(C=C1)C(O)=O	CCAGCTC	1015	1016
243	COC1=CC=CC(=C1)C(O)=O	AACAAGG	1019	1019
244	OC(=O)C(O)(C1=CC=CC=C1)C1=CC=CC=C1	GCTGCCG	1021	1022
245	CNC(=O)C1=CC=C(C=C1)C(O)=O	TTGTCCG	1018	1019
246	CC(=O)N1CCC(CC1)C(O)=O	GCGTAGG	1021	1021
247	OC(=O)CN1C=NC2=C(C=CC=C2)C1=O	GAACCTG	1020	1020
248	CC(C)(O)C(O)=O	AGGTAGG	1019	1020
249	OC(=O)C1=CC=C(Br)O1	CAATATT	1018	1019
250	NS(=O)(=O)C1=CC=C(C=C1)C(O)=O	ATCACTG	1019	1019
251	OC(=O)C1=CC=C(I)C=C1	CAACGAG	1022	1023
252	CC(=O)N[C@H](CC1=CNC2=C1C=CC=C2)C(O)=O	CGAACGC	1021	1023
253	CC(=O)N[C@@H](CC1=CC=CC=C1)C(O)=O	GCTCGCT	1019	1019
254	CC(=O)N[C@H](CC1=CC=C(O)C=C1)C(O)=O	TAACTAG	1021	1021

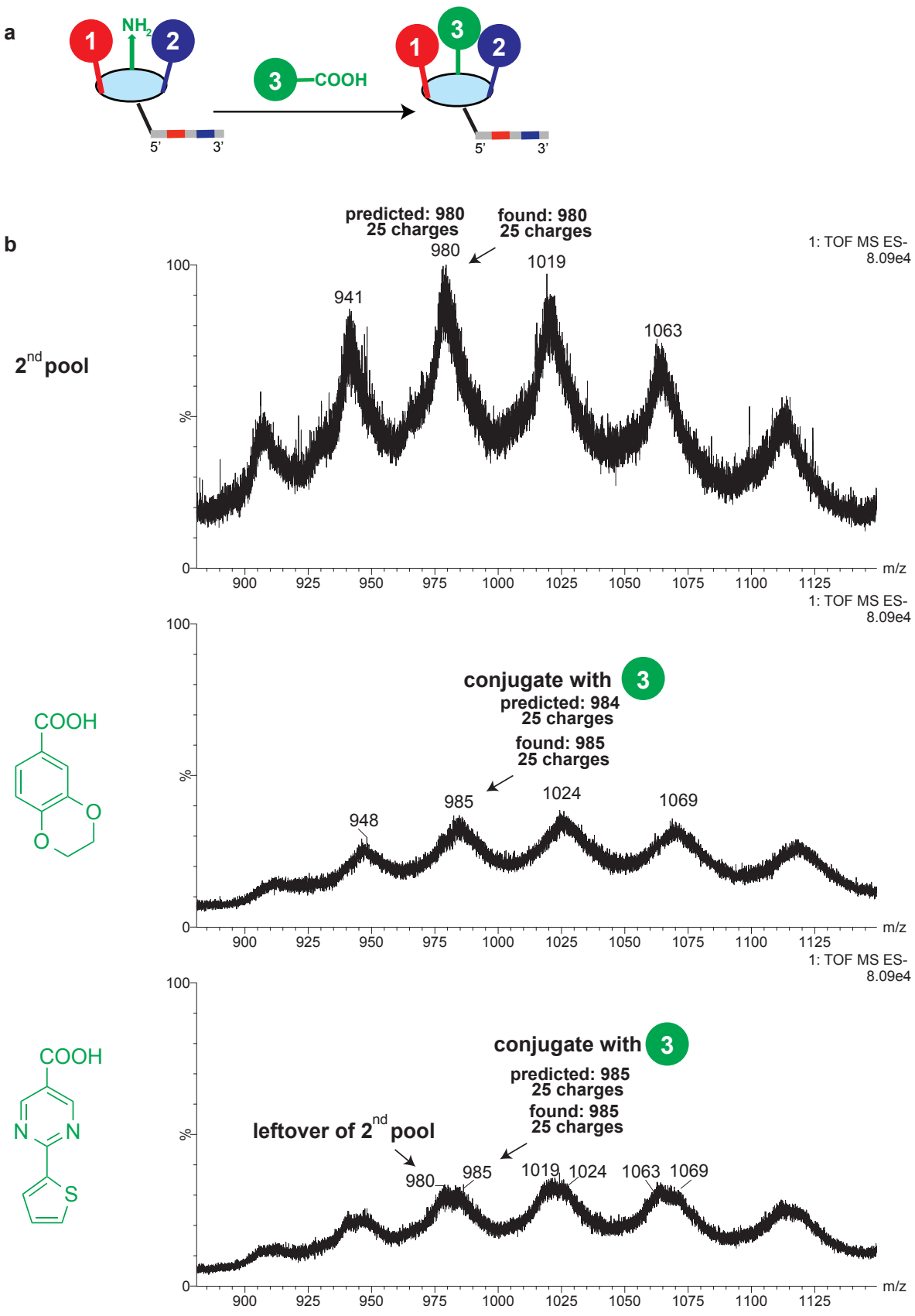
255	CC(=O)N[C@@H](CC1=CC=C(O)C=C1)C(O)=O	CTTCTAC	1018	1018
256	CC(=O)N[C@@H](CC(N)=O)C(O)=O	ATATGCT	1018	1018
257	CC(=O)N[C@@H](CCC(N)=O)C(O)=O	GCAAGCT	1019	1020
258	C[N](C)(C)CCCC(O)=O	CGTGATT	1018	1018
259	CC1=CC2=C(NC(=O)C2CC(O)=O)C=C1	CCGAATC	1018	1019
260	OC(=O)C1=CN=C(Br)C=C1	TCAGGCG	1021	1021
261	OC(=O)C1CCCN(C1)C(=O)NC1=CC=CC=C1	CATGCGT	1021	1022
262	OC(=O)C1=CC=C(C=C1)C#N	TGGAAGC	1019	1020
263	OC(=O)C1CCC1	CTTAACT	1014	1014
264	CC1=NC(=CS1)C(O)=O	AAGGCGT	1019	1019
265	OC(=O)C1CC2=C(C1)C=CC=C2	AGGTTCT	1018	1019
266	OC(=O)C[C@H]1NC(=O)[C@@H](CC2=CC=CC=C2)NC1=O	TTACAAT	1021	1022
267	CC1=C(NC(=O)C2=CC=CO2)C=C(C=C1)C(O)=O	CGACGAC	1021	1023
268	OC(=O)C1=CC(=CC=C1)S(=O)(=O)NC1=C(Cl)C=CC=C1	CGTGAAG	1119	1120
269	CC1CN(CC(C)O1)S(=O)(=O)C1=CC=C(C=C1)C(O)=O	ACACCGG	1023	1024
270	COC1=C(C=CC=C1)N(C)S(=O)(=O)C1=CC=C(C=C1)C(O)=O	AACCTTA	1023	1024
271	CC1(C)CC2=C(O1)C(OCC(O)=O)=CC=C2	AGTTCGG	1022	1023
272	OC(=O)C(CC1=CC=CC=C1)NC(=O)C1=CC=CS1	TGATTCT	1022	1023
273	CCC(C)C(NC(N)=O)C(O)=O	CTCGAGT	1018	1019
274	NC(=O)N1CCCC1C(O)=O	TACACTC	1015	1016
275	OC(=O)\C=C\C(=O)C1=CC(F)=C(F)C=C1	CGTGTAC	1020	1021
276	CC1=C(SC2=C1C(=O)NC=N2)C(O)=O	CCGCTGA	1112	1112
277	CC1=NN(C(C)=C1\C=C\C(O)=O)C1=CC=CC=C1	ACGCGCA	1021	1021
278	OC(=O)CN1C=NC=N1	GCCGCTT	1015	1016
279	COCCN1C(CCC(O)=O)=NC2=C1C(=O)NC(=O)N2CC(C)C	CACCTCT	1022	1022
280	CC1=CC=C(C=C1)S(=O)(=O)N(CC(O)=O)CC(O)=O	GCTTCAC	1114	1115
281	OC(=O)C1CN(CCC2=CC=CC=C2)C(=O)C1	GACCAGC	1021	1021
282	CC(C)C(NC(=O)NC1=CC=CC=C1)C(O)=O	TCTTCAG	1020	1020
283	OC1CC(N(C1)S(=O)(=O)C1=CC(Cl)=C(Cl)C=C1)C(O)=O	TAGTCGG	1027	1028
284	CC(C)NS(=O)(=O)C1=CC(=CC=C1)C(O)=O	GTGTCAA	1022	1023
285	CC(=O)NC1=NC(C(O)=O)=C(Br)C=C1	TAGGTCT	1022	1023
286	OC(=O)C1=C(ON=C1)C1=CC=CC=C1	GAACTAC	1019	1019
287	CC1=NC(C(O)=O)=C(C)O1	CTGTAGT	1018	1018
288	O[C@@H]1O[C@@H]([C@@H](O)[C@H](O)[C@H]1O)C(O)=O	CAATTGG	1020	1020
289	NC1=C(Br)C=C(C=N1)C(O)=O	ATACAGT	1020	1021
290	OC(=O)C1=CC=C(C=C1)S(=O)(=O)NCC=C	AAGCAAT	1022	1022
291	OC(=O)CN1NC(=O)C=CC1=O	GCGTGCG	1113	1113
292	CN1C=C(C(O)=O)C(=N1)C(F)F	CGGTATG	1020	1021
293	COC1=C(NC(C)=O)C=CC(=C1)S(=O)(=O)NC(CC1=CNC2=C1C=CC=C2)C(O)=O	TATCGGA	1030	1031
294	OC(=O)C1=CNN=C1C1CC1	TGGTAAC	1018	1019
295	OC(=O)C1CN(C2CC2)C(=O)C1	GCGGAGA	1021	1022
296	OC(=O)C1=C2N=CC(Br)=CN2N=C1	AGCTAAC	1021	1021
297	OC(=O)C1=CC2=C(N1)C=C(C=C2)C(F)(F)F	TGCCTTC	1019	1019
298	OC(=O)C1=CC=C(C=C1)S(=O)(=O)NC1CC1	TGCCGGC	1021	1023
299	OC(=O)C1CCN(CC1)C(=O)C1=CC=CS1	ACTGCTC	1019	1020
300	OC(=O)C1=C(Br)C=CC(=C1)S(=O)(=O)NC1CC	TCCTATA	1023	1023

	1			
301	COC1=C(C=C(C=C1)C(O)=O)S(=O)(=O)NC1CC1	CACAACG	1021	1022
302	NC(=O)NC1=CC(=CC=C1)C(O)=O	CAGTACA	1018	1019
303	CC1=NC2=C(C=CC(=C2)C(O)=O)N1C1CC1	ACCGGAC	1020	1020
304	OC(=O)C1=C(N=CC=N1)C1=NC2=C(S1)C=CC=C2	GCAGATT	1023	1023
305	CN1N=C2C=CC=CC2=C1C(O)=O	CCACCAC	1015	1015
306	OC(=O)C1=C(C=C(Cl)C=C1)N1CCCC1=O	AGAAGCA	1023	1023
307	CC1=NC(=C(S1)C(O)=O)C1=CC=CC=C1	TGCGATC	1020	1021
308	CC1=CC=C(CN2N=CC3=C2N=C(C)C=C3C(O)=O)C=C1	AGGCAAC	1024	1023
309	OC(=O)C1=CC2=C(N1)C=CC=C2F	CATCAGC	1017	1018
310	OC(=O)C1=NN(C(=O)NC1=O)C1=CC=CC=C1	GTTAGAT	1115	1116
311	CC(O)C(NS(=O)(=O)C1=CC(C)=C(C)C=C1)C(O)=O	CTGCAGC	1022	1023
312	CC1=C(C=C(C=C1)S(=O)(=O)NC1CC1)C(O)=O	TCTTCCT	1019	1019
313	NC1=C(N=CC=N1)C(O)=O	ATCTGGT	1017	1018
314	OC(=O)C1=C(C=CC=C1)C1=CC=C(C=C1)C(F)(F)F	TGAACGG	1024	1025
315	OC(=O)C1=C(Cl)C=CC(O)=C1	CGCCACA	1016	1018
316	CC1=NC2=C(C=C(O)C=C2)C(=O)N1CC(O)=O	TCATAGC	1020	1020
317	CC(C)[C@H](N1CC2=C(C=CC=C2)C1=O)C(O)=O	ACCGATC	1019	1021
318	CN(C)CC1=CNC2=C1C=CC(=C2)C(O)=O	ACCGTCA	1019	1019
319	CN1C2=C(N(CC(O)=O)C=N2)C(=O)N(C)C1=O	TACTAAG	1021	1022
320	CC1=C(C(O)=O)C(=NO1)C1=C(Cl)C=CC=C1F	GCACTAA	1021	1021
321	OC(=O)[C@@H]1CC(=O)NC(=O)N1	TAGGTGG	1021	1021
322	CC1=C(C(=O)C(O)=O)C2=C(N1)C=CC=C2	ATCAGCT	1019	1019
323	CC(C)N1N=CC2=C1N=C(C)C=C2C(O)=O	ATTGTAG	1022	1022
324	OC(=O)C1=CNN=C1C1=CC=C(F)C=C1	TCACTCT	1017	1017
325	CC1=C(CCC(O)=O)C(=O)NC(=O)N1	AACGCTA	1112	1112
326	NC(=O)NC(CC(O)=O)C1=CC=CS1	TATGGCT	1021	1021
327	OC(=O)C1(CCCC1)NC(=O)NCC1=CC=CC=C1	CCTTGAA	1021	1022
328	CC(C)CC(NC(=O)NCC1=CC=C(C)C=C1)C(O)=O	CTCGCAG	1022	1022
329	CC(C)C(NC(=O)NCC1=CC=CC=C1)C(O)=O	AGAGAAT	1024	1024
330	OC(=O)C1=CN=C(C=C1)N1C=NC=N1	TCCTACG	1017	1018
331	CC1=C(C=NN1C1=NNC(=O)C=C1)C(O)=O	AGCTCGC	1112	1113
332	CC1=NC2=C(C=NN2C(C)=C1)C(O)=O	CAAGCCT	1018	1018
333	OC(=O)C1(CC1)C1=CC(Cl)=CC=C1	ATGTAGC	1020	1022
334	CCCN1C(=O)NC(=O)C2=C1N=C(C=C2C(O)=O)C1CC1	GAAGGCT	1118	1119
335	CC(C)CN1C(=O)NC(=O)C2=C1N=C(C=C2C(O)=O)C1CC1	TCATGAC	1023	1024
336	OC(=O)C1=NN(C(=O)C=C1)C1=CC=CC=C1	TTCTAGC	1019	1021
337	OC(=O)C1=CNN=N1	AACGTGA	1110	1110
338	CC1=C(N=NN1C1=CC=C(F)C=C1)C(O)=O	ACAATTA	1020	1020
339	OC(=O)C1CCCN1C(=O)C1CC1	GAATGTC	1020	1020
340	CN1NC(=O)C2=C1NC(=O)C(CC(O)=O)=C2C	CAGAGAT	1115	1116
341	OC(=O)CCC1=NNC(=O)NC1=O	TCGCCAT	1110	1110
342	OC(=O)C1=CC=C(NC(=O)NC2CC2)C=C1	GAATGGT	1023	1024
343	CC1=C(CC(O)=O)C(=O)NC(N)=N1	ATGCCGG	1113	1113
344	OC(=O)C1=CC=C(NC1=O)C1=CC=CC=C1	CTTGATA	1113	1113
345	OC(=O)C1=NNC2=C1CCC2	GATCGGC	1018	1020
346	OC(=O)C1=NNC2=C1CCCCC2	GATAAGT	1021	1022
347	OC(=O)CCC(NC(=O)NC1=CC=C(F)C=C1)C(O)=O	ACCACTC	1113	1113

348	CC1=C(C=NC=N1)C(O)=O	GCCAAGA	1018	1018
349	CC(NC(=O)C1=CC=C(Br)S1)C(O)=O	CGACTAA	1022	1023
350	OC(=O)C1=CC(NC(=O)NC2CC2)=CC=C1	CAACAGG	1021	1022
351	CN1C2=C(NC(CCC(O)=O)=N2)C(=O)NC1=O	CGCTCAC	1111	1111
352	OC(=O)C1(CCOCC1)C1=CC=C(F)C=C1	TGGATAG	1023	1023
353	OC(=O)C1CC1C(=O)N1CCN(CC1)C1=CC=CC=C1	ACGGCAT	1023	1024
354	OC(=O)C1=CC(=CC=C1)N1NC(=O)C=CC1=O	TGCAAGT	1115	1115
355	NC(=O)C1=CC=C(S1)C(O)=O	AATAATA	1019	1019
356	OC(=O)C1CCCN1C(=O)C1=CC=C(Br)C=C1	TGTAGGC	1025	1026
357	OC1CC(N(C1)C(=O)C1=CC=C(F)C=C1)C(O)=O	ACTAACA	1021	1022
358	OC(=O)C1=CC=C(C=C1)C(F)F	CATATAC	1017	1017
359	OC(=O)CCN1C=CNC(=O)C1=O	CCTCGGT	1062	1063
360	OC(=O)C1=NNC(=O)C1	GACTCCG	1060	1060
361	OC(=O)CN1C=C2C=CC=CC2=N1	CGCTATC	1017	1017
362	OC(=O)C1=C2C=CC=CN2N=C1	ATCGCGA	1018	1019
363	CC1=NC(=CN1)C(O)=O	CGAGAGC	1018	1018
364	OC(=O)CCC(NC(=O)NC1=CC=C(C=C1)C#N)C(O)=O	AAGAGGA	1120	1121
365	OC(=O)CC1(O)CCCCC1	CGCAATT	1017	1018
366	OC(=O)C1(CC1)C1=C(F)C=CC=C1	TGGTACG	1020	1022
367	CN1N=C(C(O)=O)C(Br)=C1C	ACGTCGA	1021	1021
368	CC1=C(Br)C=NN1CC(O)=O	CCAAGGT	1021	1021
369	CC1=NC(=NO1)C1=CC(=CC=C1)C(O)=O	TTGTGAC	1020	1021
370	OC(=O)CN(CC1=CC=CC=C1)CC1=CC=CC=C1	GCGGTGT	1024	1024
371	OC(=O)C1=C(Br)SC=N1	GTACTGG	1021	1022
372	COC1=C(OC)C=C(C=C1)C1(CCCC1)C(O)=O	TCGTCTC	1019	1020
373	CC1=CC=C(C)C(CC(O)=O)=C1	TACCACT	1016	1016
374	OC(=O)CC1=CC=C2OCOC2=C1	AGTCTCA	1018	1018
375	CC1=C(CC(O)=O)N=C(O1)C1=CC=CC=C1	TGTTGCT	1020	1021
376	OC(=O)CC1=COC2=C1C=CC=C2	TGGCAAT	1019	1020
377	OC(=O)CCC1=NC=C(O1)C1=CC=CC=C1	CGCTCCG	1018	1020
378	OC(=O)CCC1=NC=C(O1)C1=CC=C(Cl)C=C1	AAGGTTG	1024	1025
379	CC1=C(OCC(O)=O)C=CC(Cl)=C1	GCTACAT	1019	1019
380	OC(=O)CCC1=CC=C(O1)C1=CC=CS1	TGTCTTA	1020	1022
381	COC1=CC2=C(C=C1)C(CC(O)=O)=CO2	CACGAAT	1019	1020
382	CC1=C(CC(O)=O)C2=CC(C)=CC(C)=C2N1	GCGATGC	1021	1023
383	CC1=NNC(C(O)=O)=C1Br	CGCTGCT	1019	1020
384	OC(=O)C1=C(Br)C(=NN1)C1CC1	CAGTGAG	1067	1068
385	[NH]C(OCC1=C([N+]([O-])=O)C=C(OC)C(OC)=C1)=O	ACCGCGT	1011	1012
386	[NH2]	ATATCCA	1010	1011

Supplementary Table 2 | Structures and codons of the 2nd diversity elements with predicted mass and found mass of the corresponding conjugates at 24 charge state ions.

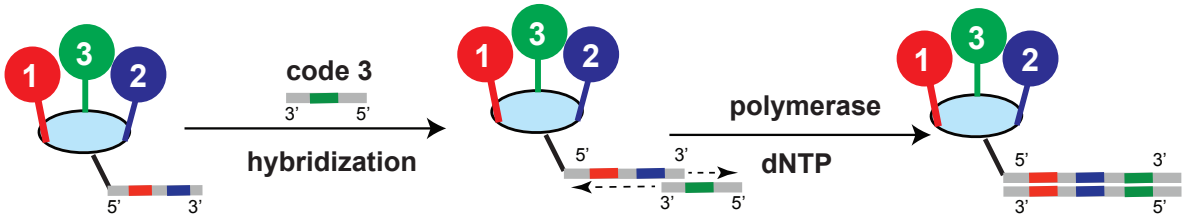
(g) Conjugation of 185 carboxylic acids and 136 alkynes as third diversity elements (DE-3), encoding by enzymatic polymerization, assembling of the 3rd Pool, purification and characterization.



Supplementary Figure 12 | Conjugation of the 3rd diversity elements. a, Conjugation of 3rd diversity elements to the 2nd pool. **b,** representative MS spectra of 2nd pool, conjugates from 2nd pool + 2,3-dihydrobenzo[*b*][1,4]dioxine-6-carboxylic acid and conjugates from 2nd pool + 2-(thiophen-2-yl)pyrimidine-5-carboxylic acid, 50 pmol used for injection. MS analyses were performed on Xevo G2-XS Q-TOF

with electrospray ionization source. For 2,3-dihydrobenzo[*b*][1,4]dioxine-6-carboxylic acid as diversity element, the conversion was over 80 % according to the peak intensity ratio between the conjugates and the starting pool. For 2-(thiophen-2-yl)pyrimidine-5-carboxylic acid as diversity element, the conversion was about 50 % according to the peak intensity ratio between the conjugates and the starting pool.

A part of the 2nd pool (100 nmol) was divided into 136 (0.7 nmol each) aliquots. These aliquots (0.7 nmol each) were immobilized on DEAE speharose (50 µL of slurry, GE Heathcare, Catalog: 17-0709-01). The resin was washed with aq AcOH (3 × 0.5 mL, 10 mM), H_2O (3 × 0.5 mL) and DMSO: H_2O: *t*BuOH = 4 : 3 : 1 (3 × 0.5 mL). To the resin-immobilized DNA-conjugate was added to a solution of the corresponding alkyne diversity element (20 mM), TBTA (10 mM, Aldrich, Catalog: 678937), $CuSO_4$ (2.5 mM, Aldrich, Catalog: 451657) and ascorbate (10 mM, Aldrich, Catalog: 41996) in DMSO: H_2O: *t*BuOH = 4 : 3 : 1 (0.2 mL). The slurry was agitated at 25 ºC for 4 h. The reaction solution was removed and the resin washed with DMSO: H_2O: *t*BuOH = 4 : 3 : 1 (6 × 0.5 mL), aq EDTA (3 × 0.5 mL, 50 mM, FisherBio, Catalog: BP2482-500) and aq AcOH (3 × 0.5 mL, 10 mM). The DNA was eluted from the resin by incubation with aq AcOH (3 × 0.2 mL, 3 M) for 30 min each time. The DNA-conjugates were isolated by ethanol precipitation and the pellets redissolved in deionized water (100 µL), the recovery of conjugates quantified by UV absorption at 260 nm, resulting recovery yields were around 80 - 90 %. A part of the 2nd pool (152 nmol) was dissolved by Tris-HCl (10 mL, 500 mM, pH 8.0), followed with addition of TECP-HCl (350 mg). The Staudinger reduction lasted for 12 h at 25 °C, reduction products were isolated by ethanol precipitation and the pellets redissolved in deionized water (5 mL). The reduction products were further divided into 185 (0.7 nmol each) aliquots conjugated with 185 carboxylic acids using the same condition as described above. The DNA-conjugates were isolated by ethanol precipitation and the pellets redissolved in deionized water (200 µL), quantified the recovery conjugates by UV absorption at 260 nm and recovery yields were around 40-60 %. All the 324 conjugates were characterized by UPLC-MS and conversion yields were all over 80 % based on the MS peak intensities of the starting pool (2nd pool here) and the conjugates from 3rd diversity elements (see Supplementary Fig. 12).



Supplementary Figure 13 | Encoding of the 3rd diversity elements and purification.

Individual conjugates with the third diversity elements were subjected to polymerase extension without further purification. To these conjugates (250 pmol), oligos (code 3: 5'-OH-GCTCTGCACGGTCGC**NNNNNNN**GCTGCCTGACGC -3'-OH, 300 pmol, 1.2 equiv.) were added, followed with addition of NEB Buffer 2 (40 μL, New England BioLabs, Catalog: B7002S) and kept at 50 °C for 10 minutes. After the system cooled down to 22 °C, a solution of 2'-deoxynucleotide triphosphates (330 μM, 40 μL, Catalog: N0446S) and Klenow Fragment 3'-> 5' *exo*- (2 μL, 10 units, New England BioLabs, Catalog: M0212S) was added. The polymerase extension reaction was allowed to proceed at 22 °C for 1 hour (see Supplementary Fig. 13). The conversion yields were all over 80% based on the UV integral of absorption at 260 nm by UPLC and denaturing PAGE analysis. 185 carboxylic acids and 136 alkynes and 324 DNA tags were employed as third diversity elements and corresponding barcodes, respectively (Two extra tags were used to encode "N_3-on" and another tag was used to encode "N_3-off" [NH2]). Equimolar amounts (250 pmol) of the 324 conjugates obtained as described above were combined to generate the 3rd pool as the final **ETH-YL** library with 35 million members. The analysis results after polymerase extension reaction with code 3 are listed as follows:

Number	Smiles	Codon	Predicted (Da)	Found (Da)
1	Cl.CN(CC#C)CC1=CC=CC=C1	GAGGTCG	1026	1024
2	CN1C=NC=C1C#C	TTCGCCG	1024	1025
3	CC(O)(C=C)C#C	GGATATA	1024	1024
4	NC1=CC(=CC=C1)C#C	ACCTAGG	1024	1026
5	OC(C#C)(C1=CC=CC=C1)C(F)(F)F	TGTATGA	1028	1029
6	CCN(CC)CC#C	GTTAACG	1024	1026
7	C#CC1=CSC=C1	ACTGTTC	1024	1024
8	C#CC1=CC=CN=C1	TACATAA	1024	1023
9	FC(F)(F)C1=C(C=CC=C1)C#C	ATGCGCC	1027	1027
10	CN(C)C1=CC=C(C=C1)C#C	CAACCAC	1026	1025
11	OC(=O)C1=CC=CC(=C1)C#C	TCGCGGT	1119	1118
12	OC(C#C)(C1=CC=CC=C1)C1=CC=CC=C1	ATGGCAC	1028	1029
13	NC(CC#C)C(O)=O	CGCGAAG	1118	1119
14	NC1=CC=C(C=C1)C#C	ACAGCAT	1024	1025
15	C#CC1=NC=CC=C1	TGGTTAG	1024	1024
16	NC1=CC=CC=C1C#C	GACAGGT	1024	1026
17	NS(=O)(=O)C1=NN=C(NC(=O)CCCC#C)S1	GGTAACT	1076	1076
18	C#CC1=CCCCC1	CGGATTA	1024	1025
19	C[C@]12CC[C@H]3[C@@H](CCc4cc(O)ccc34)[C@@H]1CC[C@@]2(O)C#C	GCCAAGA	1032	1032
20	ClC1=CC=CC(Cl)=C1C#C	CAATGGC	1027	1026
21	CC1=CC=C(NC(=O)NCC#C)C=C1	GTTCCGT	1027	1027
22	OC(=O)C1=CC(=CN=C1)C#C	TCGCCAC	1172	1174
23	O=C1N(CC#C)C2=C(C=CC=C2)C1=O	CACAACA	1027	1026
24	CCC(C)(O)C#C	CGCGCCT	1024	1023
25	C#CC1=CC=NN1	GAGTAGT	1023	1024

26	BrC1=NC=C(OCC#C)C=C1	ACATCGG	1028	1027
27	CNC1=CC=C(OCC#C)C=C1	TCTCGTA	1026	1025
28	BrC1=CC2=C(C=C1)N(CC#C)C(=O)C2=O	CTACAAT	1031	1031
29	CN1C=C(C=N1)C#C	CTCCGAA	1024	1023
30	COC1=C(Br)C=C(C=C1)C#C	TAGTCCT	1028	1029
31	CC(NCC#C)C1=CC2=C(OCC(=O)N2)C=C1	GGTTGGC	1029	1030
32	ClC1=C(C=CC=C1)C#C	ACTACAG	1025	1026
33	NC(=O)C1=CC(=CN=C1)C#C	TCACAAT	1026	1027
34	COC1=C(F)C=C(C=C1)C#C	GAGCAAG	1026	1025
35	O=C(NCC#C)NC1CC1	TGGCGCT	1025	1026
36	C#CC1=CC(=CC=C1)C1=CC=CC=C1	CTGAGCC	1027	1028
37	COC1=CC(C#C)=C(Cl)C=C1	GCCTCGT	1027	1028
38	Cl.NCCC(O)CCC#C	TGTTACG	1025	1025
39	C#CCN1C2=C(C=CC=C2)C2=C1C=CC=C2	TAATTCG	1028	1030
40	OC1=CC(OCC#C)=CC=C1	ATGTAAC	1026	1026
41	FC1=CC(=CC(F)=C1)C#C	TGTTCCA	1025	1026
42	C#CCN1C=NC=N1	CCATCTT	1024	1025
43	NC(=O)NC1=CC(=CC=C1)C#C	GCGCATT	1026	1027
44	C#CCN1C=CC2=C1C=CC=C2	TTACTCC	1026	1028
45	CC1=CC(NC(=O)NCC#C)=CC=C1	CTGCTTC	1027	1028
46	COC1=C(C=C(Cl)C=C1)C#C	CGTGAGC	1027	1026
47	FC(F)(F)C1=NC(OCC#C)=CC=C1	GGTCCAC	1028	1028
48	O=C1NC(CC#C)C(=O)N1	TCCACAG	1119	1120
49	CCN1C=C(C=N1)C#C	GATGAAT	1025	1024
50	FC(F)OC1=C(C=CC=C1)C#C	GTCCGGT	1027	1028
51	CC1=NC(=CC=C1)C#C	CGAGATA	1024	1024
52	CC(O)CC#C	TAGCCAA	1023	1024
53	Cl.NC(=N)NCC#C	TTCGTTA	1024	1024
54	Cl.C#CCN1C=CN=C1C1=CC=CS1	AAGTTAA	1027	1027
55	OC(=O)CC(=O)NCC#C	ACGAAGC	1172	1173
56	OC(=O)[C@@H]1CCCN1CC#C	TAAGCTG	1119	1120
57	BrC1=CC2=C(OCC(=O)N2CC#C)C=C1	TTGTCAT	1031	1031
58	BrC1=C(OCC#C)C=CC=N1	CGCCACA	1028	1027
59	O=C1CCCCN1CC#C	TCGTCGG	1025	1027
60	C#CC1=C2CCCCC2=CC=C1	ATATATA	1026	1027
61	C#CC1=NC2=C(C=CC=C2)N=C1	ACGGCGG	1026	1026
62	NC(=S)NC1=CC(=CC=C1)C#C	GGCTGAT	1027	1024
63	C#CCNC1CCCC1	CTAATGA	1025	1027
64	CN(CC#C)C1=NC=C(N)N=C1	CGGATGC	1026	1027
65	OC(=O)C1=CC(=CC=C1)S(=O)(=O)NCC#C	GTCAGAT	1236	1237
66	C#CC1=CN=CN=C1	GGCGACC	1024	1026
67	OC(=O)C1(CC#C)CCC1	CCAGCTC	1025	1024
68	O=C1NCCN1CC#C	AGGCGGC	1025	1025
69	COC1=NC(=CC=C1)C#C	AGCATTA	1025	1025
70	CC1=NC(=CS1)C#C	ACAATCG	1025	1026
71	C#CCN1C=CN=C1	AATGTAC	1024	1024
72	ClC1=CC2=C(C=C1)N(CC(=O)NCC#C)C(=O)C2=O	GGTTGAA	1031	1031
73	O=C(NCC#C)C1=CNC(=O)C=C1	GATAATT	1027	1027
74	NC1=C(C=CC=C1)C(=O)NCC#C	TAGACAT	1027	1026

75	C#CC1=CC=NC=C1	CCTCCGC	1024	1026
76	NC1=C(F)C(F)=C(C#C)C(F)=C1F	GCAGTAC	1027	1028
77	CC1=NC2=C(C=C1)C=C(C=C2)C#C	ATGTTCC	1027	1027
78	Cl.NC(CC#C)C(N)=O	GCATAGG	1024	1025
79	FC1=C(C=CC(Br)=C1)C(=O)C#C	GGCCTCT	1029	1028
80	CC[C@]12CC[C@H]3[C@@H](CCC4=CC(=O)CC[C@H]34)[C@@H]1CC[C@@]2(O)C#C	CTGAACA	1033	1032
81	CS(=O)(=O)NCCC#C	TATTATA	1026	1026
82	CC(O)(C#C)C1=C(Br)C=CC=C1	GTATGAC	1029	1030
83	CC(=O)C(CC#C)C(C)=O	ATTCTCC	1025	1025
84	CNC(=O)C#C	TAACCGG	1023	1024
85	C#CC1=CN=CS1	ATAGCTG	1024	1026
86	C#CC1=NC=NC=C1	TCAGTTG	1024	1024
87	ClC1=C(Cl)C(=CC=C1)C#C	GCGTAAT	1027	1028
88	OC(CC1=CC=CC=C1)C#C	GGTGCAT	1026	1027
89	COCC(N)C#C	ACCGGCC	1024	1023
90	NC(C#C)C1=CC=C(Br)C=C1	GCTAAGG	1028	1027
91	C#CC1=NC=CS1	ACACTAT	1024	1024
92	CC(C)(NS(C)(=O)=O)C#C	AGGTGTC	1026	1026
93	OC(CC#C)C1CC1	TGCTTAC	1024	1024
94	Cl.C#CCNCC1CC1	CCGGTAG	1024	1023
95	COCCN(C)CC#C	GAGGCAA	1025	1025
96	ClC1=CC(=CC(Cl)=C1)C#C	AGGATAA	1027	1027
97	C#CC1=NC2=C(S1)C=CC=C2	CTCACCT	1026	1028
98	NC(CO)CC#C	TAGCTTA	1024	1025
99	CC1=C(OCC#C)C=CC=C1	CTAGCAC	1026	1026
100	ClC1=C(NC(=O)NCC#C)C=CC=C1	AACGTCG	1028	1027
101	FC1=C(F)C=C(NC(=O)NCC#C)C=C1	CACGGAC	1028	1028
102	OC(=O)C1=CC=C(C=C1)S(=O)(=O)NCC#C	GAATCTG	1236	1236
103	CC1=CC(NC(=O)NCC#C)=NO1	CCGGACT	1027	1027
104	O=S(=O)(NCC#C)C1=CC=CC=C1	TTGGATG	1028	1030
105	O=C(NCC#C)NC1=CC=CC=C1	TACGAAG	1027	1026
106	BrC1=CC2=C(OCCC2NCC#C)C=C1	TCCTATT	1031	1031
107	OC(=O)C1(CCC2=C1C=CC=C2)NC(=O)NCC#C	AAGTGGC	1124	1124
108	ClC1=CC=C(NC(=O)NCC#C)C=C1	ATCAATG	1028	1028
109	COC1=CC(=NC=C1)C#C	GATAACC	1025	1025
110	FC1=C(NC(=O)NCC#C)C=CC=C1	AGCTAAT	1028	1030
111	NC(C#C)C1CCCCC1	ATACAGC	1025	1026
112	CC1=CC(NC(=O)CNCC#C)=CC=C1	TCAGGAA	1028	1027
113	FC1=C(C=CN=C1)C#C	TAGCTAC	1025	1024
114	CC(=O)O[C@]1(CCC2C3CCC4=CC(=O)CC[C@@H]4C3CC[C@@]12C)C#C	ACAACTC	1034	1034
115	Cl.C#CCN1C=NC2=C1C=CC=C2	CACATTC	1026	1027
116	OC(=O)CCCCC#C	GATCATC	1171	1119
117	BrC1=CC=C(C=C1)C#C	TACGGAT	1027	1027
118	C#CC1=C2C=CC3=CC=CC4=CC=C(C=C1)C2=C34	TGTGGTG	1029	1029
119	NCC#C	ACCAGGC	1022	1023
120	OCCCC#C	GACCACA	1023	1023
121	COC1=CC=C(C#C)C(C)=C1	CATGACG	1026	1025
122	C#CC1=CC=C2N=CC=NC2=C1	GGTGCCG	1026	1026

123	NCCOCC#C	CTTAGGA	1024	1024
124	OC1(CCCCC1)C#C	TGGCATA	1025	1024
125	NC1=CC=CC=C1NC(=O)C1=CC=CC(=C1)C#C	TAATGAC	1029	1030
126	C#CC1=CC=CC=C1	TGGTCAC	1024	1023
127	C#CC1=CC=C(C=C1)C1=CC=CC=C1	CCTGGCC	1027	1028
128	FC(F)(F)C1=CC(=CC(=C1)C#C)C(F)(F)F	TACTCGG	1030	1031
129	CC(C)(C)C1=NN(C(=O)O1)C1=C(Cl)C=C(Cl)C(OCC#C)=C1	GTTGAGA	1034	1035
130	OC1(C#C)C2=C(C=CC=C2)C2=C1C=CC=C2	CGCTAGC	1028	1029
131	CC1(C)OC2=C(C=CC=C2)C2OC(C)(CCC12)C#C	TACGGTC	1030	1031
132	CC1(C)CC(O)(CC(C)(C)N1)C#C	GCCGGAG	1027	1026
133	O=C1NC=C(C#C)C(=O)N1	GGTCGTG	1119	1118
134	CC1=C(C)C2=C(N)N=C(SCC#C)N=C2S1	CTTACCG	1030	1030
135	C#CC1=C2C=CNC2=NC=N1	GCACAGC	1026	1026
136	NCCC(=O)N1CC2=C(C=CC=C2)C#CC2=C1C=CC=C2	GCGAATG	1031	1032
137	CC(C)CC(O)=O	AAGCGCT	1023	1022
138	OC(=O)CC1CCCC1	CAGCGTC	1024	1023
139	OC(=O)CC1CCCCC1	ATACTGT	1024	1025
140	CN(C)CC(O)=O	CAGAATA	1023	1022
141	OCC(O)=O	ATGCATT	1022	1022
142	OC(=O)C[C@@H]1C[C@H]2CC[C@@H]1C2	GGCGTAG	1025	1026
143	CC1(C)[C@@H]2CC[C@@]1(C)[C@H](CC(O)=O)[C@H]2O	CGCGTAC	1027	1028
144	COC1=CC(OC)=C(CCC(O)=O)C=C1	ATCGGTA	1027	1026
145	OC(=O)CCCC1=CC=C(I)C=C1	GTCTAAT	1031	1029
146	OC(=O)C[C@]12CC3CC(C[C@@](O)(C3)C1)C2	ATGACGA	1027	1030
147	COC1=CC=C(Br)C=C1CC(O)=O	CGCACGC	1029	1027
148	OC(=O)CC1NC(=O)NC1=O	AGTCTGG	1025	1025
149	OC(=O)CCN1C(=O)OC2=C1C=CC=C2	ACGCATC	1028	1028
150	NS(=O)(=O)C1=CC=C(NC(=O)CCC(O)=O)C=C1	TCGTGTG	1030	1030
151	COC1=CC=C(CC(O)=O)C=C1O	GGATGGA	1026	1028
152	NS(=O)(=O)C1=CC=C(NC(=O)CSCC(O)=O)C=C1	GTTATCA	1031	1030
153	CN(CC(O)=O)S(=O)(=O)C1=CC=CC=C1	GTGTCCT	1028	1027
154	NC(=O)NCCC(O)=O	TCCGAGC	1024	1024
155	OC(=O)CC1OC2=C(NC1=O)C=CC=C2	GCAGACG	1121	119
156	CC1=C(CC(O)=O)NC2=C1C=CC=C2	ATCTTGA	1026	1025
157	OC(=O)CCC1=NN=C(O1)C1=CC=CC=C1	GTCGTAA	1028	1027
158	OC(=O)CCN1C(=O)COC2=C1C=CC=C2	ACTGTCA	1028	1027
159	OC(=O)CC(C1=CC=CO1)C1=CC=CC=C1	GAAGTCT	1028	1027
160	CC1=NC2=C(C=CC=C2)N1CCC(O)=O	AGCTGTA	1027	1028
161	COC1=C(C=CC=C1)C1=NOC(CCC(O)=O)=N1	CTGACTC	1029	1027
162	OC(=O)CCC1=NC(=NO1)C1=CC=CO1	GGCAGAC	1027	1028
163	OC(=O)CC1=COC2=C1C=CC=C2	AGTGATT	1028	1027
164	CC1=CC=C2N=C(C)C=C(C(O)=O)C2=C1	ATATTCG	1027	1027
165	OC(=O)C1=CC=C2N=CC=CC2=C1	GCTGGTG	1026	1025
166	OC(=O)C1=CN(C2CC2)C2=CC(Cl)=C(F)C=C2C1=O	CCACTCT	1030	1029
167	OC(=O)C1=C2C=CN=CC2=CC=C1	CTGGCGG	1026	1023
168	OC(=O)C1=CC=C2NC=CC2=C1	TTGTGAC	1025	1026
169	COC1=CC=C2NC(=CC2=C1)C(O)=O	ACGGTAA	1026	1026
170	CN1C=CC2=CC(=CC=C12)C(O)=O	GCGTGCA	1026	1026

171	OC(=O)C1=CC2=CC=CC=C2N1	CTCTAGT	1025	1024
172	OC(=O)C1=CC=C2NC=NC2=C1	TCGAACG	1025	1022
173	OC(=O)C1=CC=C2OCCOC2=C1	GACGGCA	1026	1024
174	OC(=O)C1=NC2=C(C=CC=C2)N=C1	TCTAAGC	1026	1024
175	OC(=O)C1=CC2=C(C=C1)N=CC=N2	GGAAGTC	1026	1026
176	OC(=O)C1=CN=C(N=C1)C1=CC=CS1	GCCAATC	1027	1026
177	OC(=O)C1=NNC(=C1)C1CC1	CAACTCC	1025	1024
178	OC(=O)C1=CN2C=CSC2=N1	GTAAGCC	1026	1025
179	COC1=CC(=CC(OC)=C1)C(=O)N1CCC(CC1)C(O)=O	ACAAGTG	1031	1030
180	CC(C)C1=CC(=NO1)C(O)=O	GTATAAG	1025	1027
181	OC(=O)C12CC3CC(CC(C3)(C1)N1C=NC=N1)C2	AGTGGCG	1029	1027
182	CC1=CC=C(C=C1)N(CC(O)=O)S(C)(=O)=O	CGAGGAT	1029	1028
183	OC(=O)C1C2CC(C=C2)C1C(=O)NC1CC1	GGTACGC	1028	1025
184	OC(=O)C1=CN=C(N=C1)C1=CC=CN=C1	ATGGTTC	1027	1028
185	OC(=O)[C@@H]1C[C@H]1C1=CC=CC=C1	CGAAGCC	1025	1025
186	OC(=O)C1COC2=CC=CC=C2O1	GAGGATC	1026	1020
187	CC1(C)C2CCC1(C(O)=O)C(=O)C2	ATCTCCA	1026	1026
188	CN1[C@@H]([C@H](CC1=O)C(O)=O)C1=CN=CC=C1	CGATCAT	1028	1028
189	CC(C)[C@H]1CC[C@@H](CC1)C(O)=O	CTGCGAG	1026	1025
190	OC(=O)[C@@H]1CCC(=O)N1	CATAGGT	1024	1025
191	O[C@@H]1CC(=C[C@@H](O)[C@H]1O)C(O)=O	TGAATTG	1026	1026
192	OCC(C(O)=O)C1=CC=CC=C1	TAGTTGT	1025	1025
193	CC(=O)N1C[C@H](O)C[C@H]1C(O)=O	AGTCCAT	1026	1026
194	COC1=CC=CC(=C1)N1C=C(C=N1)C(O)=O	GCGGTAT	1028	1029
195	OC(=O)C1CN(CC2=CN=CC=C2)C(=O)C1	GACGCTG	1028	1027
196	OC(=O)C1CN(C2CCCC2)C(=O)C1	AACACCA	1027	1026
197	CC(=O)C1=C(C)N(CC(O)=O)N=C1C	AGAGACG	1027	1027
198	CC(C)C1=NC(C)=C(S1)C(O)=O	TCTCTAG	1026	1026
199	OC(=O)C1CCN(CC1)C(=O)C1CCCC1	TTACGTA	1028	1026
200	OC(=O)C1=CC=CC=C1N1CCC(=O)NC1=O	AAGAGAA	1028	1028
201	OC(=O)C1=NOC(=C1)C1=CC=CC=C1	TAGGATA	1071	1072
202	COC1=CC=C(C=C1)C1=CC(=NN1)C(O)=O	CCTCCTT	1028	1030
203	OC(=O)CCCC1=NC(=NO1)C1=CC=NC=C1	CCTAACC	1028	1027
204	OC(=O)C1=CC(NC(=O)C2CCCC2)=CC=C1	GGAACAG	1028	1028
205	CN(C)C(=O)N1CCC(CC1)C(O)=O	TAGAACC	1027	1027
206	OC(=O)C1CCCC1	GTGCATG	1023	1024
207	CN(C)S(=O)(=O)C1=CC(C(O)=O)=C(C)O1	ATTATAT	1028	1028
208	OC(=O)C1=COC=C1	GAGACGT	1023	1023
209	OC(=O)C1=CC=C(Cl)S1	ATGTATG	1025	1024
210	CC1=C(SC(=N1)C1=CC=NC=C1)C(O)=O	TCGGCAA	1028	1028
211	CN1N=C(C=C1C(O)=O)C(C)(C)C	TTACTAT	1026	1025
212	CC1=C(C=NN1C1=CC=CC=C1)C(O)=O	TCCGCGT	1027	1026
213	CN1C=CC=C1C(O)=O	TGTGATA	1024	1024
214	CN1C(=CC2=C1C=CO2)C(O)=O	ACCGATA	1025	1025
215	OC(=O)C1=CC=C(N=C1)C(F)(F)F	TACATGC	1026	1025
216	CC1=CC=CN=C1C(O)=O	AATTATG	1024	1022
217	OC(=O)C1=CN=C(O)C=C1	ACCACAT	1024	1023
218	OC(=O)C1=CC=C(Br)C=N1	AGCACAG	1027	1026
219	COC1=C(OC)C(=CC=C1)C(O)=O	TGAACAT	1026	1026

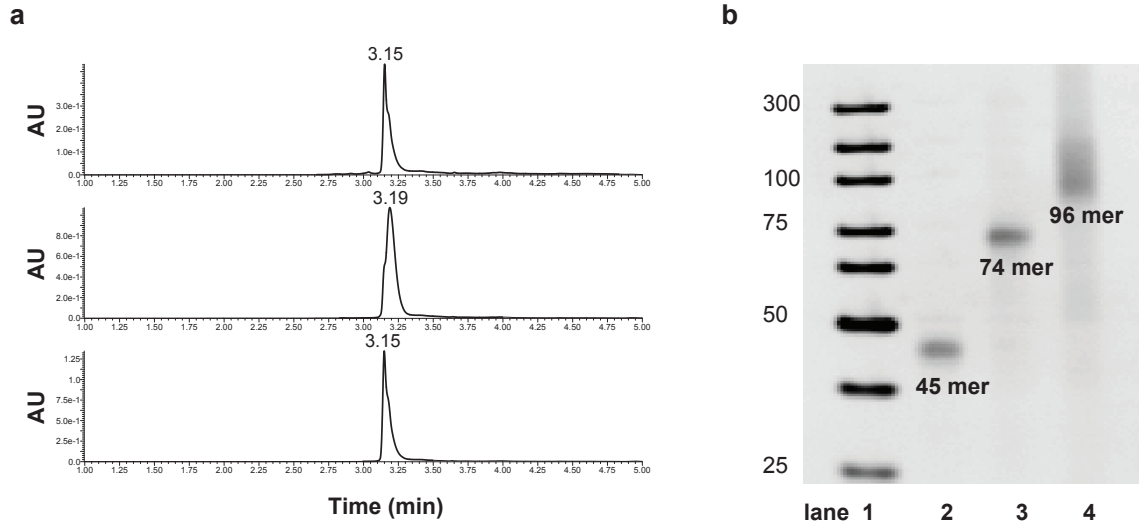
220	COC1=CC=C(C=C1OC)C(O)=O	TATTCAT	1026	1027
221	COC1=CC(=CC(OC)=C1)C(O)=O	CGTCTTG	1026	1026
222	CN(C)C1=CC=CC(=C1)C(O)=O	AGAGCAG	1025	1026
223	OC(=O)C1=CC(F)=C(F)C=C1F	TCTATCG	1026	1023
224	OC(=O)C1=CC=C(C=C1)C(F)(F)F	GATGTCC	1026	1026
225	OC(=O)C1=C(Cl)C=C(Cl)C=C1	TGGAGCG	1026	1027
226	OC(=O)C1=C(F)C=CC(=C1)C(F)(F)F	TATCCTG	1027	1027
227	OC(=O)C1=CC=C(Cl)C=C1	AGCGCGG	1025	1024
228	CCOC(=O)CNC1=C(C=CC=C1)C(O)=O	TCTGCTA	1028	1026
229	CCOC(=O)CNC1=CC(=CC=C1)C(O)=O	CCTATGT	1028	1028
230	OC(=O)C1=NNC(=O)C=C1	ACTCGCC	1024	1024
231	OC(=O)C1=CN=CN=C1	GGTGGTA	1024	1023
232	CN(CCC(O)=O)C1CCN(C)C1	CCTTATA	1026	1024
233	CC1(CC(O)=O)N2C=CC=CC2=NC1=O	TGACGTT	1027	1027
234	CCN1N=C(C)C=C1C(O)=O	CATACGC	1025	1026
235	OC(=O)C1=CC=NN1	CCTAGAG	1023	1023
236	OC(=O)C1=CNN=C1	CATTCAC	1023	1022
237	CC1=CC=C(C(O)=O)C(O)=N1	GTATTCA	1118	1120
238	OC(=O)C1=C(N=CC=N1)C(=O)N1CCCCC1	GACCTCG	1028	1027
239	CN1C=NC2=C(C(C)=C(S2)C(O)=O)C1=O	GGCCGTA	1028	1028
240	CC(=O)NC1=CC(=CC(NC(C)=O)=C1)C(O)=O	TTCAACG	1028	1028
241	CC1=NOC2=NC(=CC(C(O)=O)=C12)C1CC1	AGAGGAC	1028	1028
242	CN1C(=O)NC(=O)C2=C(C=C(C)N=C12)C(O)=O	TGTATCT	1122	1122
243	OC(=O)C1=CC2=C(NC(=O)C(=O)N2)C=C1	CGGCGGT	1121	1122
244	OC(=O)C1CCC=CC1	AATTACA	1024	1023
245	CC1=NN2C(=C1)N=CC(C(O)=O)=C2C	CTGTTAG	1026	1026
246	CC1=C(C(O)=O)C(C)=NO1	TTACCTG	1024	1024
247	CN(CC(O)=O)C1=NC=NC2=C1NC=N2	TCGTATC	1121	1120
248	OC(=O)C1=CC=C2N=CNC(=O)C2=C1	TGCCGCC	1120	1119
249	OC(=O)C1=CC=C(OC2=CC=CN=C2)O1	TACCGAC	1027	1028
250	CC1(CC1)C(O)=O	CGTAAGG	1023	1023
251	CC1(CCCCC1)C(O)=O	GGATTAT	1024	1023
252	CC1(CCCC=C1)C(O)=O	CAATATT	1024	1024
253	NC(=O)C1=CC=C(C=C1)C(O)=O	TAATAAG	1025	1023
254	COC1=CC=CC(=C1)C(O)=O	GTTGCGG	1025	1025
255	CC(=O)C1=CC=C(C=C1)C(O)=O	CATAGAC	1025	1026
256	OC(=O)C1=CC=C(Br)O1	GCGATCT	1026	1026
257	NS(=O)(=O)C1=CC=C(C=C1)C(O)=O	AACTATC	1027	1026
258	OC(=O)C1=CC=C(I)C=C1	GCGCACA	1029	1027
259	CC(=O)N[C@H](CC1=CC=C(O)C=C1)C(O)=O	GAAGAGA	1028	1027
260	NS(=O)(=O)C1=NN=C(NC(=O)CCC(O)=O)S1	CGCCTGG	1030	1028
261	CC1=CC2=C(NC(=O)C2CC(O)=O)C=C1	ACGACTT	1027	1027
262	OC(=O)C1CCN(CC1)C(=O)NC1=CC=CC=C1	GCCTTCA	1029	1030
263	OC(=O)C1CCCN(C1)C(=O)NC1=CC=CC=C1	GGAGTAA	1029	1029
264	CC1=NC(=CS1)C(O)=O	TTATATG	1024	1026
265	OC(=O)C1=CSC(=N1)C(F)(F)F	GCGAGTT	1027	1027
266	OC(=O)C1CC2=C(C1)C=CC=C2	ACCGTAT	1025	1025
267	CC1(C)CC2=C(O1)C(OCC(O)=O)=CC=C2	AAGTCCG	1028	1027
268	OC(=O)C(CC1=CC=CC=C1)NC(=O)C1=CC=CS1	TGATTCA	1030	1030

269	CC1=NN(C(C)=C1\C=C\C(O)=O)C1=CC=CC=C1	CCTGATT	1029	1027
270	OC(=O)CN1C=NC=N1	TGGAGTT	1024	1023
271	OC(=O)C1CN(CCC2=CC=CC=C2)C(=O)C1	CACCATA	1028	1027
272	CC(=O)NC1=NC(C(O)=O)=C(Br)C=C1	GCGTGGC	1029	1029
273	OC(=O)C1=C(ON=C1)C1=CC=CC=C1	CATGGTC	1026	1026
274	CC1=NC(C(O)=O)=C(C)O1	TTGTGCA	1024	1025
275	CC(C)[C@H](N(C)C)C(O)=O	GAGCGTA	1025	1022
276	O[C@@H]1O[C@@H]([C@@H](O)[C@H](O)[C@H]1O)C(O)=O	TTAGGAC	1027	1025
277	NC1=C(Br)C=C(C=N1)C(O)=O	AGCAGCA	1028	1027
278	OC(=O)CN1NC(=O)C=CC1=O	TCCATCT	1119	1119
279	OC(=O)C1=C(ON=C1)C1CC1	CAGAAGC	1025	1025
280	OC(=O)C1CN(C2CC2)C(=O)C1	TGCGGCG	1026	1026
281	CN1N=NC2=C1N=CC(=C2)C(O)=O	CAGACAA	1026	1024
282	OC(=O)C1=CC=C(C=C1)S(=O)(=O)NC1CC1	GTTATTC	1029	1028
283	OC(=O)C1CCN(CC1)C(=O)C1=CC=CS1	GTCAAGC	1028	1030
284	OC(=O)C1=C(Br)C=CC(=C1)S(=O)(=O)NC1CC1	TCACGCA	1032	1033
285	COC1=C(C=C(C=C1)C(O)=O)S(=O)(=O)NC1CC1	AAGGCGT	1030	1029
286	CC1=NC2=C(C=CC(=C2)C(O)=O)N1C1CC1	ATAATTG	1028	1028
287	OC(=O)CC(CC(O)=O)C1=C(Br)C=CC=C1	CAGCACA	1124	1126
288	CC(O)C(NS(=O)(=O)C1=CC(C)=C(C)C=C1)C(O)=O	AGTGTAG	1030	1031
289	CC1=C(C=C(C=C1)S(=O)(=O)NC1CC1)C(O)=O	TGCGCAT	1029	1029
290	NC1=C(N=CC=N1)C(O)=O	CCGCCTG	1024	1026
291	OC(=O)C1=C(Cl)C=CC(O)=C1	CTCCAAC	1026	1025
292	COC1=NN2C(CCC(O)=O)=NN=C2C=C1	TGACGGC	1028	1028
293	CN(C)CC1=CNC2=C1C=CC(=C2)C(O)=O	ACTCCTA	1028	1029
294	CC1=C(C(O)=O)C(=NO1)C1=C(Cl)C=CC=C1F	AGGCTGT	1029	1030
295	OC(=O)[C@@H]1CC(=O)NC(=O)N1	CATTGCT	1025	1027
296	CC1=NOC2=C1C(=CC(=N2)C1CC1)C(O)=O	GCTTAAC	1028	1028
297	NC(=O)NC(CC(O)=O)C1=CC=CS1	GCTAGGA	1027	1028
298	OC(=O)C1=NN(C(=N1)C1=CC=CS1)C1=CC=C(F)C=C1	TTATGAT	1031	1031
299	OC(=O)C1=CC2=C(N=C1)N(C1CC1)C(=O)NC2=O	AGCGCCA	1029	1027
300	CC1=C(C=NN1C1=NNC(=O)C=C1)C(O)=O	GACTGAA	1121	1121
301	CC1=NC2=C(C=NN2C(C)=C1)C(O)=O	TTCATGA	1026	1025
302	OC(=O)C1(CC1)C1=CC(Cl)=CC=C1	CCAATGG	1027	1027
303	CCCN1C(=O)NC(=O)C2=C1N=C(C=C2C(O)=O)C1CC1	TAAGAGC	1124	1124
304	OC(=O)C1=C(N=CC=N1)C(=O)NCC1=CC=C(F)C=C1	CACAGCT	1030	1030
305	OC(=O)CCC(NC(=O)NC1=CC=C(F)C=C1)C(O)=O	CTTCGCT	1124	1124
306	CC(NC(=O)C1=CC=C(Br)S1)C(O)=O	ACGATTC	1030	1030
307	CN1N=C(C)C(Br)=C1C(O)=O	TGGCTAT	1028	1028
308	OC(=O)C1CC1C(=O)N1CCN(CC1)C1=CC=CC=C1	GTGTTAC	1030	1028
309	OC(=O)C1=CC(=CC=C1)N1NC(=O)C=CC1=O	CTAGACG	1122	1120
310	NC(=O)C1=CC=C(S1)C(O)=O	GAGTCTA	1026	1024
311	OC(=O)C1CCCN1C(=O)C1=CC=C(Br)C=C1	AGAGTCT	1031	1030
312	OC(=O)C1=CC=C(C=C1)C(F)F	AAGATTA	1026	1029
313	OC(=O)CCN1C=CNC(=O)C1=O	TGGCTCA	1026	1028
314	OC(=O)C1=NNC(=O)C1	CCGTACC	1024	1021
315	OC(=O)C1=C2C=CC=CN2N=C1	AAGGTGG	1025	1025
316	CC1=NC(=CN1)C(O)=O	TAAGATT	1024	1023
317	CC1=NC(C(O)=O)=C(Cl)C=N1	CCACACC	1026	1026

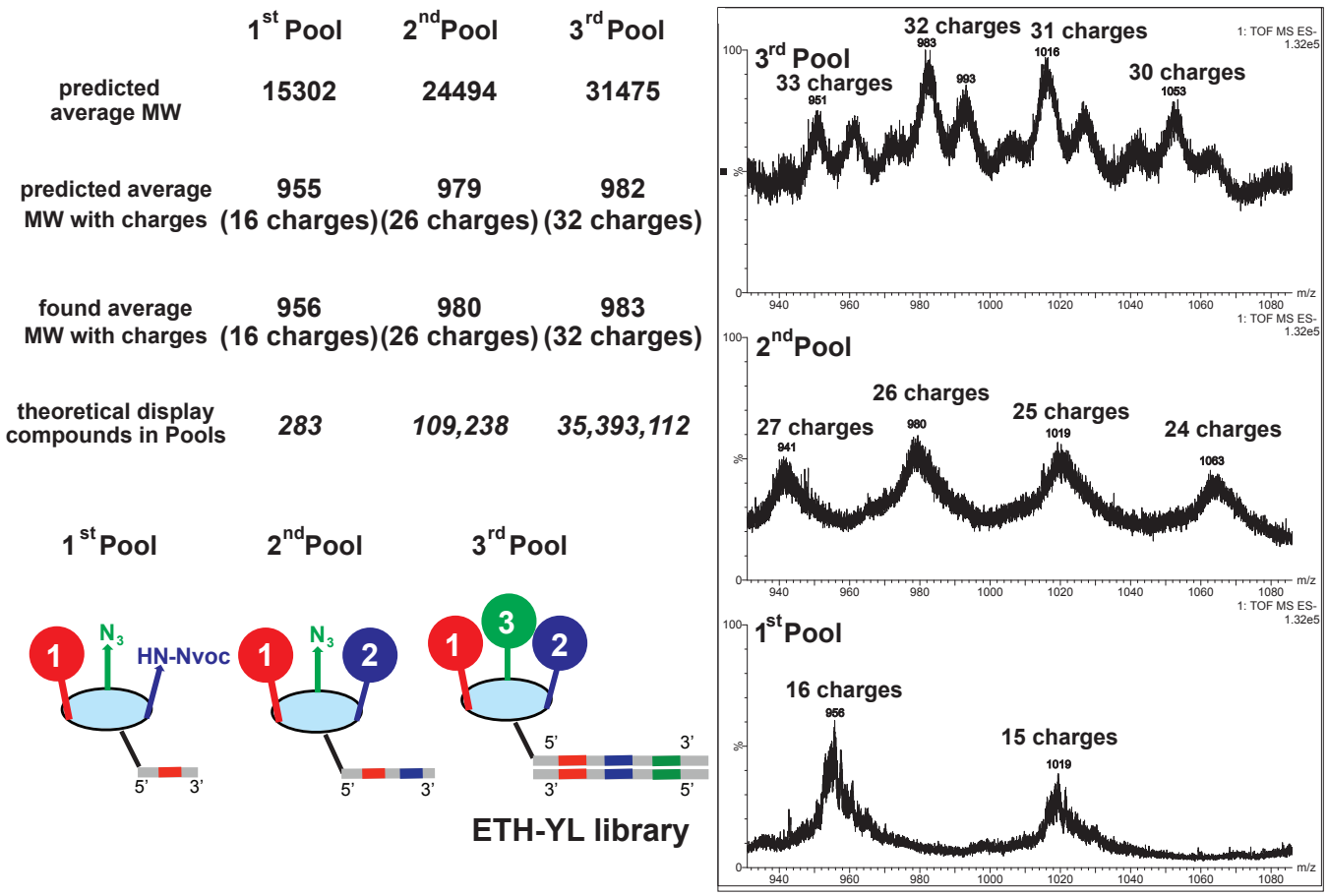
318	CC1=C(Br)C=NN1CC(O)=O	CCAATTA	1028	1028
319	OC(=O)C1=C(Br)SC=N1	TGACCAC	1027	1028
320	OC(=O)C1=CN=C(N=C1)C(F)(F)F	GAATGCC	1027	1027
321	N3	ACATTAA	1022	1022
322	N3	GTGAGAG	1025	1025
323	CC(O)=O	TAGTATT	1021	1021
324	[NH2]	GGCGTTA	1019	1019

Supplementary Table 3 | Structures and codons of the 3rd diversity elements with predicted mass and found mass of corresponding conjugates as 31 charge state ions.

(h) Characterization of ETH-YL library with 35 million displayed compounds.



Supplementary Figure 14 | Characterization of ETH-YL library with UPLC and gel electrophoresis. a UPLC chromatograms of the 1st, 2nd and 3rd pool. UPLC analyses were performed on a on a XBridge@ Oligonucleotide BEH C18 10 × 50 mm column at a flow rate of 0.5 mL/min with gradient: 0 % to 5 % B (0 to 0.5 minutes), 5 % B (0.5 to 2 minutes), 5 % to 60 % B (2 to 2.6 minutes), 60 % B (2.6 to 4 minutes), 60 % to 0 % B (4 to 5 minutes) (A= TEA 10 mM, HFIP 5 mM in water, B= MeOH), at 60 ºC. Detection by absorbance at 260 nm. **b** 1st, 2nd and 3rd pools were analyzed by denaturing gel electrophoresis. Lane1, Marker; lane 2, 1st pool; lane3, 2nd pool; lane4, 3rd pool.



Supplementary Figure 15 | Summary of predicted average molecular weight (MW) of and found average MW at different charge states for 1st, 2nd and 3rd pool. MS spectra of 1st, 2nd and 3rd pools, 50 pmol injection. MS analyses were performed on a Xevo G2-XS Q-TOF with electrospray ionization source.

As shown in Supplementary Fig. 14 and Fig. 15, pools of library at different synthesis stages were analyzed by UPLC and electrophoresis, indicating the reliable encoding and reaction efficiency of the **ETH-YL** library.

4. Affinity Selections of the Library against Immobilized Target Proteins.

(a) Target proteins for selections (CAIX, HRP, TNKS 1, HSA, AGP, CaM, PSA, L19TNF and TNF), biotinylation of target proteins and characterization.

Protein [a]	Abbr.	MS Predicted	MS Found	Quart. Structure	Supplier	Catalog
Carbonic Anhydrase IX	CAIX	29 kDa (monomer)	29044 Da	Dimer	aa 120-397 [c]	-
Horseradish Peroxidase	HRP	43 kDa	43175 Da	Monomer	Thermo Fisher	31490
Tankyrase 1 [b]	TNKS 1	29 kDa	28711 Da	Monomer	aa 1106-1325-Bioh [c]	-
Albumin from human serum	HSA	67 kDa	66559 Da	Monomer	Sigma	A3782
Alpha 1 Acid Glycoprotein, human Plasma	AGP	19 kDa	18186 Da	Monomer	Athens Research & Technology	16-16-010700
Calmodulin	CaM	17 kDa	16707 Da	Monomer	EnzoLifeScience	BML-SE325-0001
Prostate Specific Antigen Protein	PSA	28 kDa	28431 Da	Monomer	MyBio Source	MBS173180
L19 antibody-tumor necrosis factor fusion portein	L19-TNF	44 kDa (monomer)	43959 Da	Trimer	Philogen	-
tumor necrosis factor	TNF	18 kDa (monomer)	17802 Da	Trimer	BL21(DE3)	-

Supplementary Table 4 | Summary of proteins used in affinity-based selections. [a] All proteins have the sequence of the human protein. [b] Tankyrase 1 was biotinylated enzymatically using a Bir A tag. [c] The expression of CAIX, TNKS 1[6], L19-TNF[7] and TNF was reported previously.

Target proteins (with exception of Tankyrase 1) were chemically biotinylated with EZ-LinkTM NHS-LC-Biotin (typically with 3-5 equiv. to target protein, Thermo Fisher Scientific, Catalog: 21336) for affinity screening according to supplier's instructions followed by dialysis and characterization by UPLC-MS.

Protein [a]	Abbr.	MS Predicted	MS Found
biotinylated Carbonic Anhydrase IX	bCAIX	29384 Da (monomer)	29366 Da
botinylated Horseradish Peroxidase	bHRP	43515 Da	43514 Da
biotinylated Albumin from human serum	bHSA	66900 Da	66901 Da

biotinylated human Alpha 1 Acid Glycoprotein, human Plasma	bAGP	18526 Da	18525 Da
biotinylated human Calmodulin	bCaM	17047 Da	17046 Da
biotinylated Prostate Specific Antigen Protein	bPSA	28771 Da	28770 Da
biotinylated L19 antibody-tumor necrosis factor fusion portein	bL19-TNF	44299 Da (monomer)	44300 Da
biotinylated tumor necrosis factor	bTNF	18273 Da (monomer)	18273 Da

Supplementary Table 5 | Characterization of the biotinylated proteins used in selections.

(b) Affinity screening.

Portein	Buffer	pH	Beads
CAIX	PBS [a]	7.4	Dynabeads MyOne™ Steptavidin C1 Thermo Fisher Scientific, Catalog: 65001
HRP	PBS	7.4	Dynabeads MyOne™ Steptavidin C1
TNKS 1	20 mM Hepes, 200 mM NaCl, 0.5 mM TCEP	7.5	Dynabeads MyOne™ Steptavidin M270 Thermo Fisher Scientific, Catalog: 65305
HSA	PBS	7.4	Dynabeads MyOne™ Steptavidin C1
AGP	PBS	7.4	Dynabeads MyOne™ Steptavidin T1 Thermo Fisher Scientific, Catalog: 65601
CaM	DPBS[b]	7.4	Dynabeads MyOne™ Steptavidin C1
PSA	PBS	7.4	Dynabeads MyOne™ Steptavidin T1
L19-TNF	PBS	7.4	Dynabeads MyOne™ Steptavidin C1
TNF	PBS	7.4	Dynabeads MyOne™ Steptavidin C1

Supplementary Table 6 | Summary of selection conditions. a, PBS (Life technologies, Catalog: 10010-015). **b**, DPBS was made from 10 × dulbecco's phosphate buffered saline (sigma, Catalog: D1283) and pH was adjusted to 7.4.

Affinity selections were performed using a KingFisher magnetic particle processor (Thermo Fisher Scientific). Streptavidin-coated magnetic beads (0.1 mg) were resuspended in buffer (100 µl, as summarized in table supplementary table 6) and subsequently incubated with biotinylated target protein (100 µl, 1 µM) for 30 min with continuous gentle mixing. Protein-coated beads were washed three times with buffer + Tween (200 µl, 0.05% v/v Tween 20) that was supplemented with biotin (100 µM) in order to block remaining binding sites of streptavidin, and subsequently incubated with the library (100 µl, 5 nM

total concentration, in buffer + Tween) for 1 h with continuous gentle mixing. After removing unbound library members by washing five times with buffer + Tween (200 µl), beads carrying bound library members were resuspended in elution buffer (100 µl, 10 mM Tris-Cl, pH 8.5, Qiagen, Catalog: 19086) and the DNA-compound conjugates separated from the beads by heat denaturation of streptavidin and the target protein (95 °C for 5 min). The DNA of eluted library members was quantified by qPCR[8] and amplified by PCR (25-35 cycles), introducing at the same time additional, selection-specific DNA barcodes, and submitted to Illumina® high-throughput DNA sequencing (HiSeq 2500, Functional Genomics Center Zurich). Selection results were decoded by an inhouse-written C++ program and visualized using MatLab software version R2013a (8.1.0.604, MathWorks)[9].

All PCRs were conducted at 95 °C for 3 min, and then 25 - 35 cycles of 95 °C for 45 s (denature), 65 °C for 45 s (anneal/extend), 72 °C for 45 s, finally 72 °C for 5 min.

Selection-specific Primers:

5'- TACACGACGCTCTTCCGATCT**NNNNNN**GGAGCTTCTGAATTCTGTGTG -3’

5'- CAGACGTGTGCTCTTCCGATC**NNNNNN**GCTCTGCACGGTCGC -3’

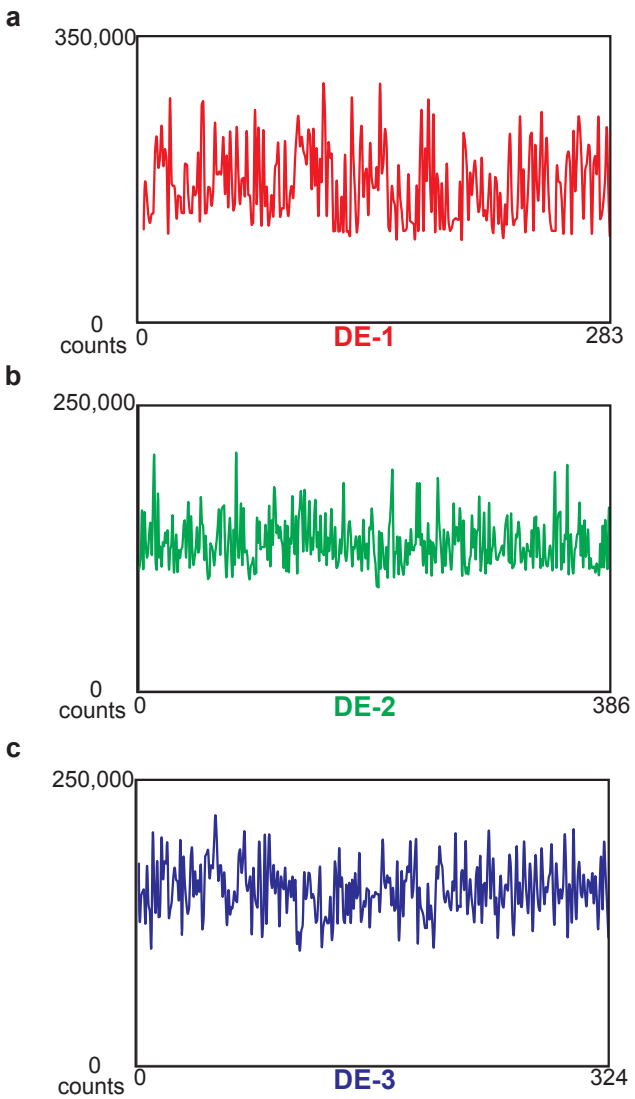
Illumina® Sequencing Primer:

5'- AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT -3’

5'- CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC -3’

Note: Colored sequences are selection-specific DNA barcodes. **N** representing variable bases serving as code.

(c) High-throughput sequencing of the library pool prior to selection.



Supplementary Figure 16 | Distribution of DE-1, DE-2 and DE-3 of the library pool prior to selection. a, Distribution of **DE-1,** the average is 169,047 counts. **b**, Distribution of **DE-2,** the average is 129,851 counts. **c**, Distribution of **DE-3,** the average is 153,926 counts.

(d) Enrichment factor calculation.

$$(1) \qquad EF(i, j, k) = SC(i, j, k)\left\{\sum_{x}^{283}\left[\sum_{y}^{386}\left(\sum_{z}^{324} SC(x, y, z)\right)\right]\right\} \text{ X 283 X 386 X 324}$$

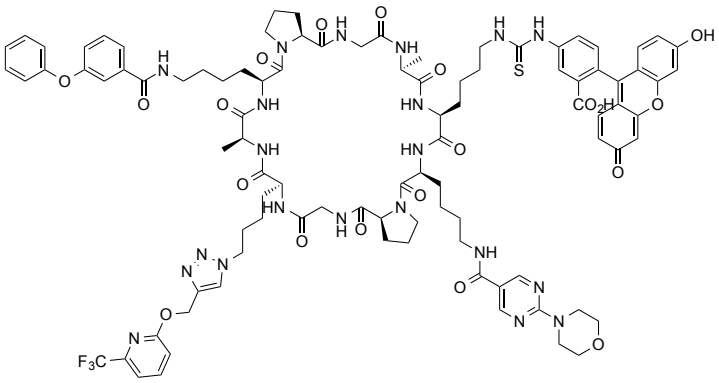
where $EF(i, j, k)$ is the enrichment (i, j and k define the number of the diversity element at **DE-1**, **DE-2** and **DE-3**), $SC(i, j, k)$ is the sequence count. Sum of $SC(i, j, k)$ is the total sequence counts of one affinity-based selection.

Target	Compound	DE-1	DE-2	DE-3	Total sequence counts	Sequence counts	Enrichment
HSA	**HSA-1**	**138**	**115**	**47**	3,344,294	129	1365-fold
HSA	**HSA-2**	**95**	**38**	**261**	3,344,294	50	529-fold
AGP	**AGP-1**	**36**	**376**	**203**	1,929,066	209	3845-fold
AGP	**AGP-2**	**125**	**39**	**163**	1,929,066	109	2000-fold
CaM	**CaM-2**	**241**	**314**	**323**	1,931,086	29	532-fold
PSA	**PSA-1**	**205**	**182**	**17**	2,244,812	121	1908-fold
L19-TNF	**TNF-1**	**20**	**361**	**106**	2,353,006	284	4272-fold
L19-TNF	**TNF-2**	**170**	**65**	**54**	2,353,006	189	2842-fold
TNF	**TNF-1**	**20**	**361**	**106**	2,357,282	18	270-fold
TNF	**TNF-2**	**170**	**65**	**54**	2,357,282	**n.d.**	**n.d.**

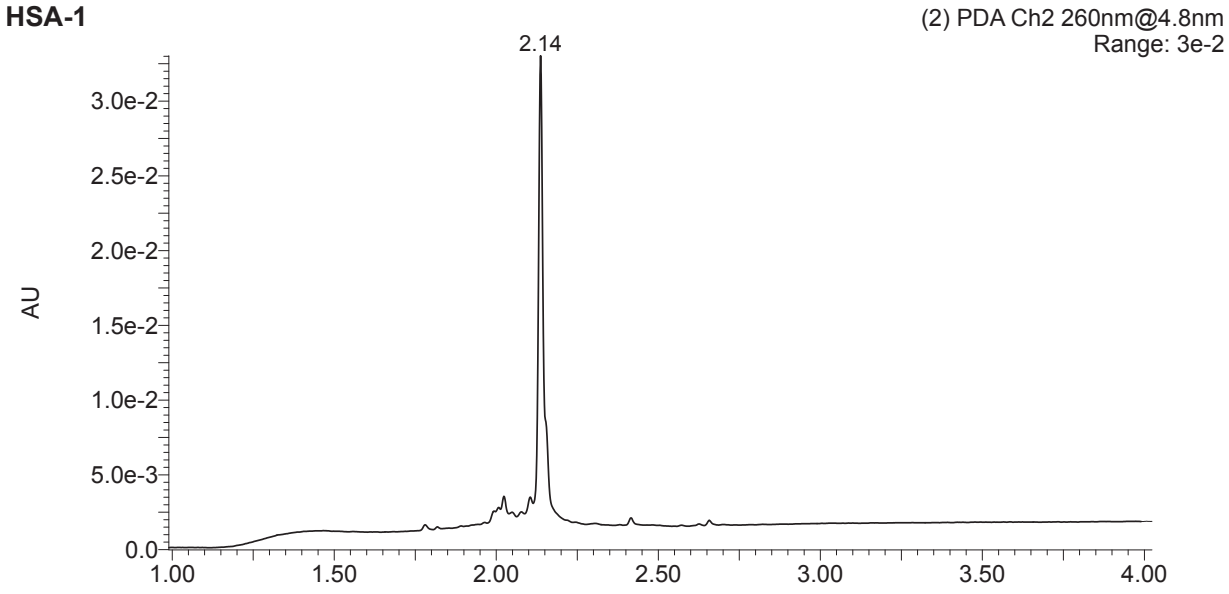
Supplementary Table 7 | Summary of enrichment. n.d. means not determined

5. Affinity Measurement of Macrocycles Selected from the Library.

(a) Affinity determination of HSA binders by fluorescence polarization measurement.



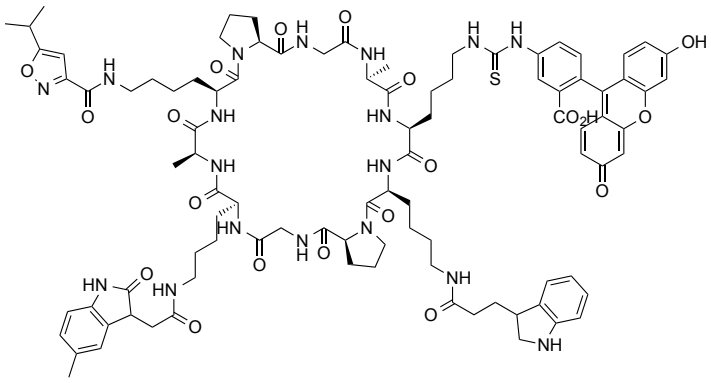
Supplementary Figure 17 | Structure of the HSA binder HSA-1.



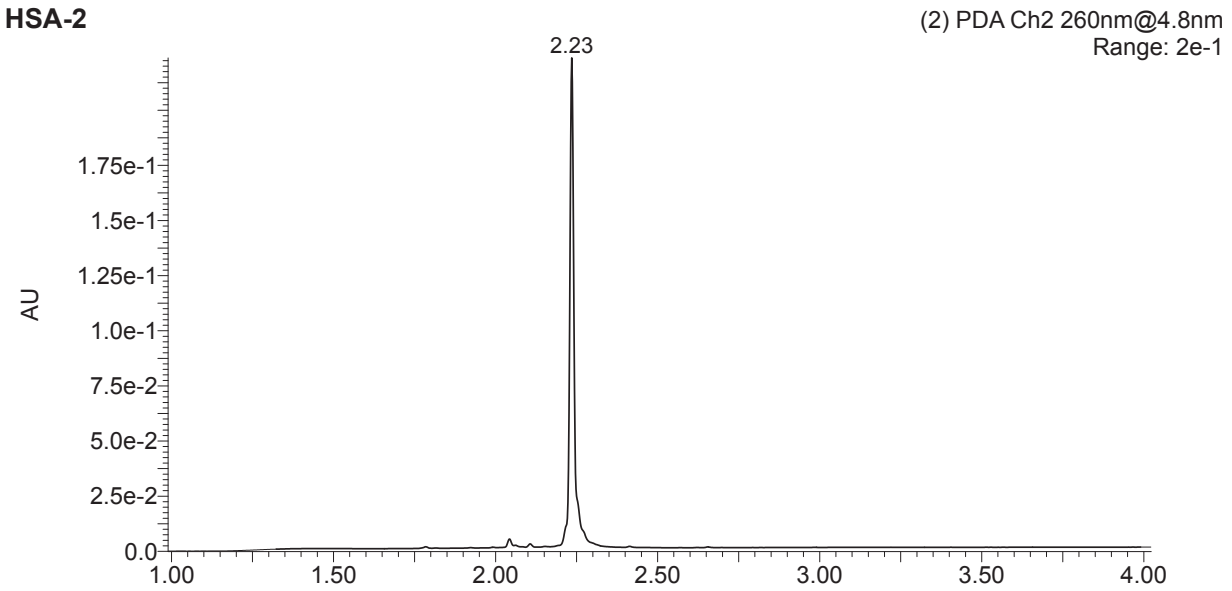
Supplementary Figure 18 | UPLC chromatogram of HSA-1. UPLC analysis was performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

HSA-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH (Senn, Catalog: 100569), Dde-off (hydroxylamine HCl salt 1.3 g, imidazole 900 mg in 6 mL of NMP : DCM = 5 : 1), DE-1-BB138 (2 equiv., TCI, Catalog: P1253-5G), Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(N_3)-OH, DE-3-BB47 (2 equiv., Enamine, Catalog: EN300-

192983; CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-2-BB115 (2 equiv., Maybridge, Catalog: CC69501DA), Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **HSA-1-amino** was recovered as a white powder after lyophilization. **HSA-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv., TCI, F0026-1G) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification. The desired HAS binder **HSA-1** was recovered as a yellow powder after lyophilization. **HRMS (*m/z*, $C_{96}H_{111}F_3N_{21}O_{20}S$, ESI):** calculated $[M+H]^+$: 1966.7897; found: 1966.8037.



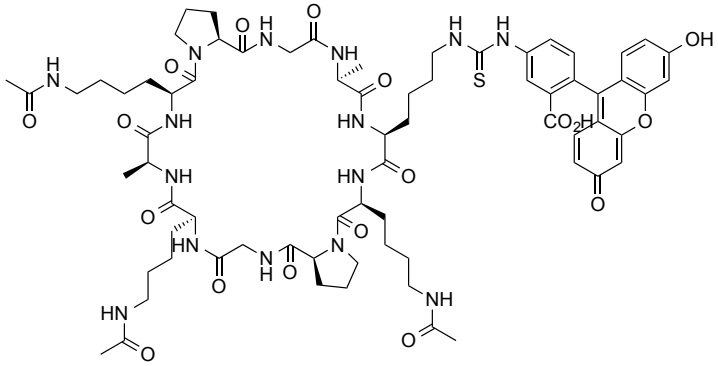
Supplementary Figure 19 | Structure of the HSA binder HSA-2.



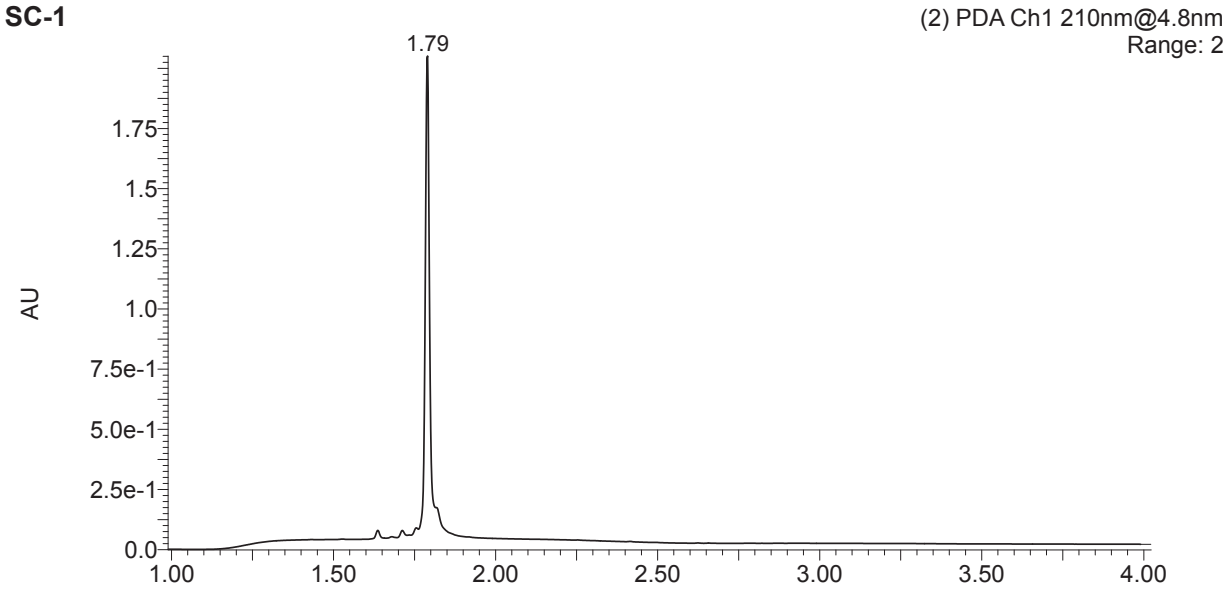
Supplementary Figure 20 | UPLC chromatogram of HSA-2. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

HSA-2 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB95 (2 equiv., ChemBridge, Catalog: 4003153), Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(Dde)-OH, DE-3-BB261 (2 equiv., SigmaAldrich, Catalog: CBR01145), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-2-BB38 (2 equiv., SigmaAldrich, Catalog: CDS003058), Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **HSA-2-amino** was recovered as a white powder after lyophilization. **HSA-2-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired

HSA binder **HSA-2** was recovered as a yellow powder after lyophilization. **HRMS (*m/z*, $C_{94}H_{117}N_{18}O_{20}S$, ESI):** calculated $[M+H]^+$: 1849.8412; found: 1849.8433.



Supplementary Figure 21 | Structure of the scaffold control SC-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.



Supplementary Figure 22 | UPLC chromatogram of SC-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 210 nm.

Scaffold control SC-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Ac)-OH (2 equiv., Senn, Catalog: 101317), Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(Ac)-OH, Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Ac)-OH, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage

solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **SC-1-amino** was recovered as a white powder after lyophilization. **SC-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired scaffold control **SC-1** was recovered as a yellow powder after lyophilization. **HRMS (*m/z*, $C_{71}H_{96}N_{15}O_{18}S$, ESI):** calculated $[M+H]^+$: 1478.6779; found: 1478.6789.

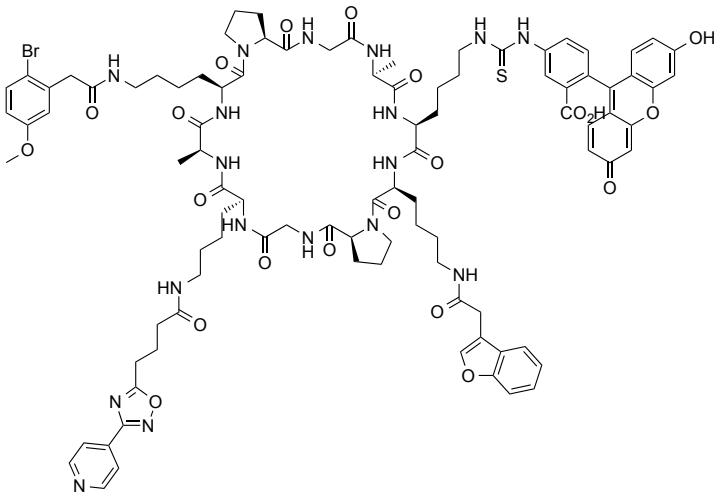
Fluorescence polarization measurement with HSA binder: Freshly dissolved fluorescein labelled macrocycle (7.5 µL, final concentration 50 nM, final DMSO content adjusted to 1% in PBS) was incubated at 22 °C for 10 min in a black 384-well plate (Greiner, non-binding) in PBS (pH 7.4, Life technologies, Catalog: 10010-015) with increasing concentrations of HSA to a final volume of 15 µL. The fluorescence anisotropy was measured on a Spectra Max Paradigm multimode plate reader (Molecular Devices). Experiments were performed in triplicate and the mean anisotropy values fitted to equation (3) using KaleidaGraph 4.1.3 (Synergy Software),

(2) $$A = \frac{1}{2}\left\{([P]_0 + [L]_0 + K_D) - \sqrt{([P]_0 + [L]_0 + K_D)^2 - 4[P]_0[L]_0}\right\}$$

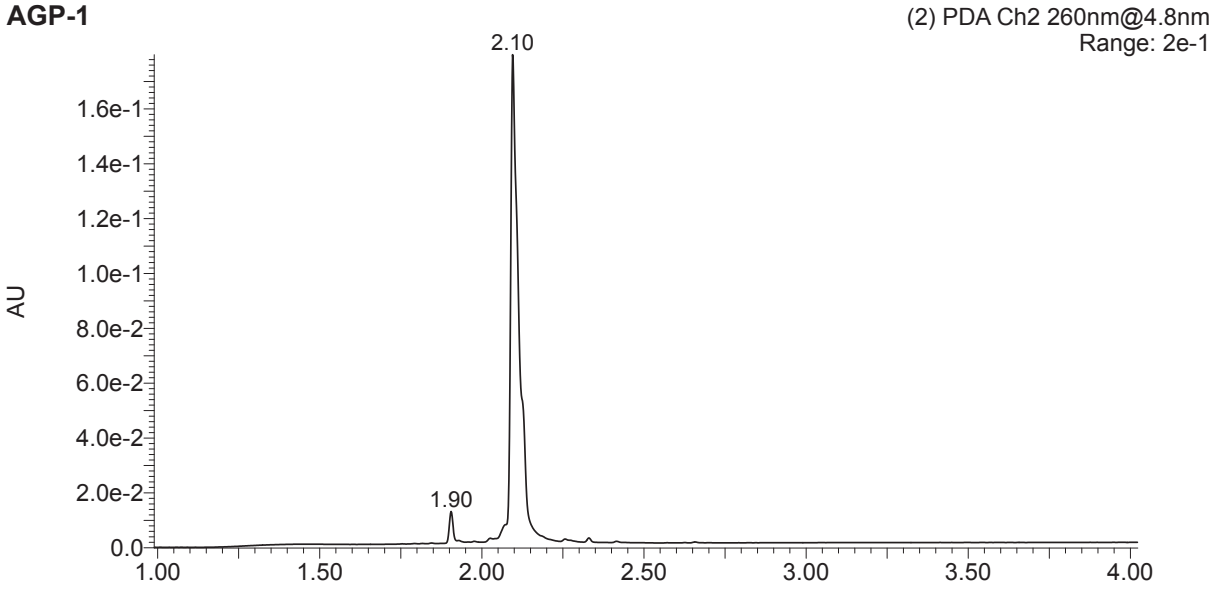
(3) $$AI = \alpha + \beta A$$

where AI is anisotropy, α, β are fitting parameters, $[P]_0$ is total protein concentration, $[L]_0$ is total concentration of the fluorescently labeled binder, K_D is the dissociation constant.

(b) Affinity determination of AGP binders by fluorescence polarization measurement.



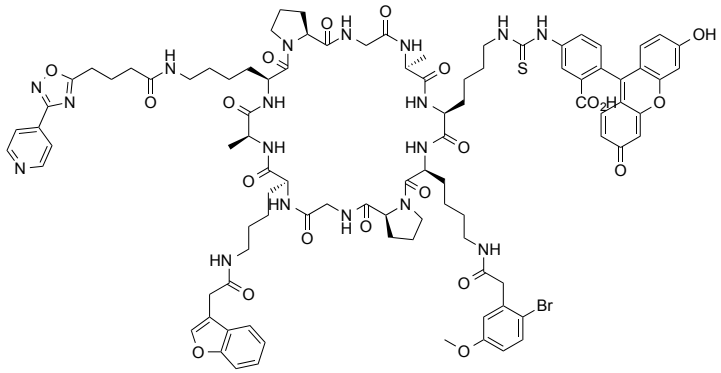
Supplementary Figure 23 | Structure of the AGP binder AGP-1.



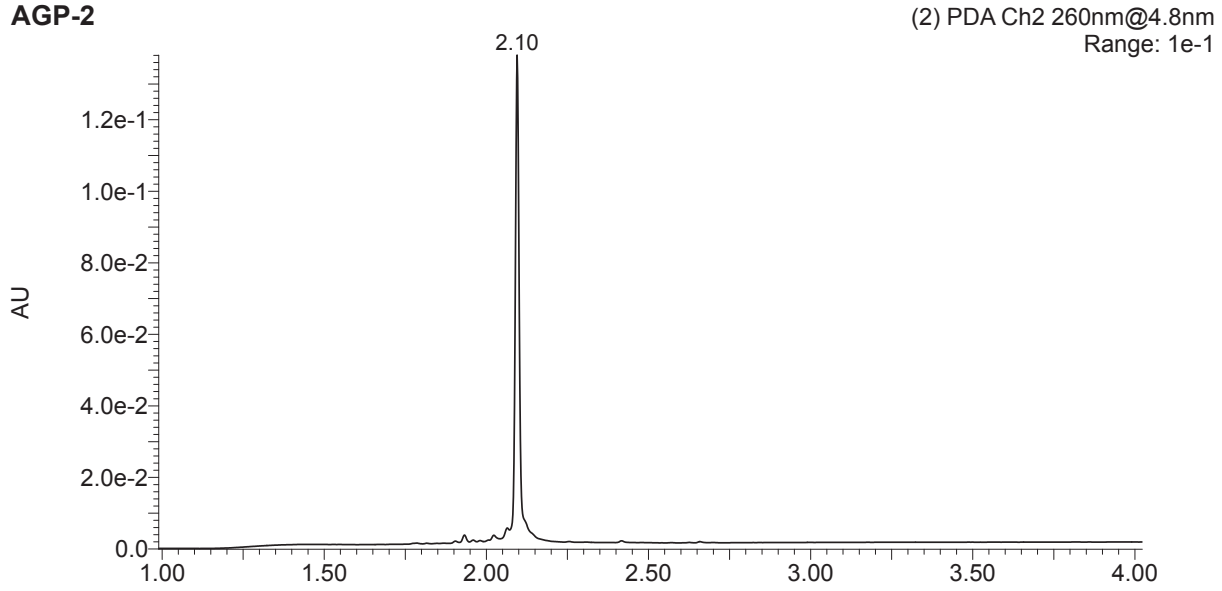
Supplementary Figure 24 | UPLC chromatogram of AGP-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

AGP-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB36 (2 equiv., SigmaAldrich, Catalog: CDS023568), Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-3-BB203 (2 equiv., ChemBridge, Catalog: 9071001), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-2-

BB376 (2 equiv., Apollo, Catalog: 225074), Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **AGP-1-amino** was recovered as a white powder after lyophilization. **AGP-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv., TCI,) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired AGP binder **AGP-1** was recovered as a yellow powder after lyophilization. **HRMS (*m/z*, $C_{95}H_{112}BrN_{18}O_{21}S$, ESI):** calculated $[M+H]^+$: 1951.7154; found: 1951.7183.



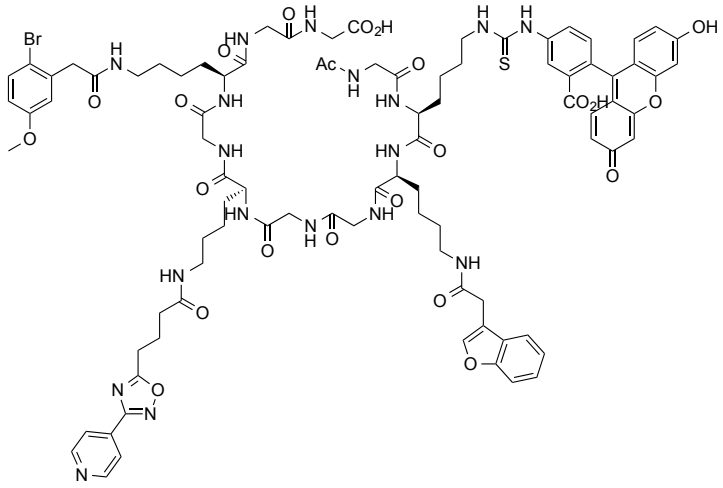
Supplementary Figure 25 | Structure of the AGP binder AGP-2.



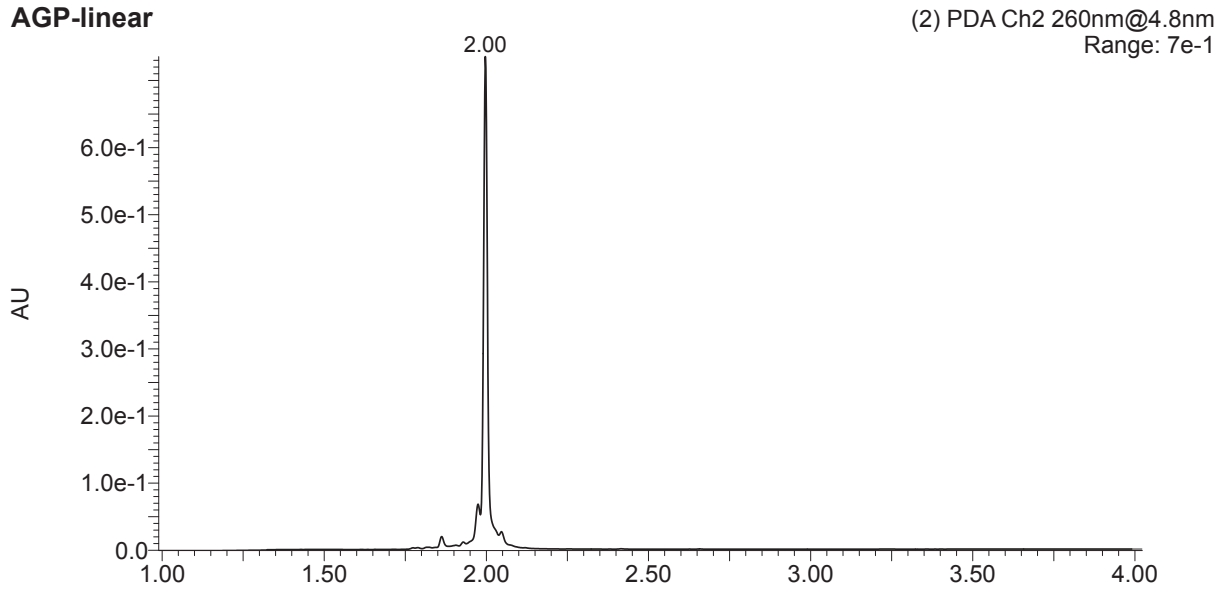
Supplementary Figure 26 | UPLC chromatogram of AGP-2. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

AGP-2 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB125 (2 equiv., ChemBridge, Catalog: 9071001), Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-3-BB163 (2 equiv., Apollo, Catalog: 225074), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-2-BB39 (2 equiv., SigmaAldrich, Catalog: CDS023568), Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **AGP-2-amino** was recovered as a white powder after lyophilization. **AGP-2-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.,) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired

AGP binder **AGP-2** was recovered as a yellow powder after lyophilization. **HRMS (*m/z*, $C_{95}H_{112}BrN_{18}O_{21}S$, ESI):** calculated $[M+H]^+$: 1951.7154; found: 1951.7194.



Supplementary Figure 27 | Structure of the linear AGP binder AGP-linear.



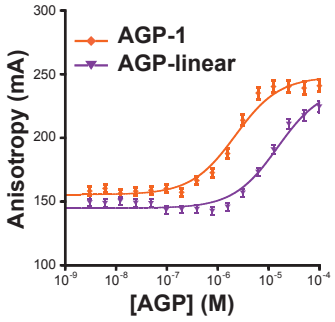
Supplementary Figure 28 | UPLC chromatogram of AGP-linear. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

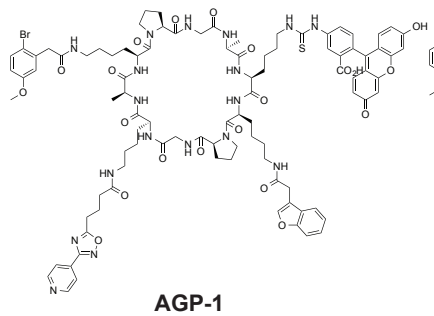
Linear AGP binder AGP-linear synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-Gly-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB36 (2 equiv., SigmaAldrich, Catalog: CDS023568), Fmoc-Gly -OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-3-BB203 (2 equiv., ChemBridge, Catalog: 9071001), Fmoc-Gly-OH, Fmoc-Gly -OH, Fmoc-*L*-

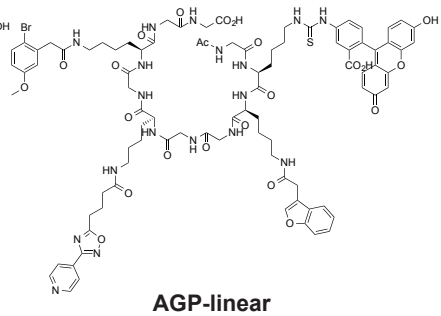
Lys(Dde)-OH, Dde-off, DE-2-BB376 (2 equiv., Apollo, Catalog: 225074), Fmoc-*L*-Lys(Boc)-OH, Ac-Gly-OH(2 equiv., SigmaAldrich, Catalog: A16300). The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The desired decapeptide **AGP-1-linear-amino** was recovered as a white powder after lyophilization. **AGP-1-linear-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired linear AGP binder **AGP-1-linear** was recovered as a yellow powder after lyophilization. **HRMS (*m/z*, $C_{89}H_{104}BrN_{18}O_{23}S$, ESI):** calculated $[M+H]^+$: 1903.6426; found: 1903.6477.

Fluorescence polarization measurement with AGP binder: Freshly dissolved fluorescein labeled macrocycles (7.5 µL, final concentration 50 nM, final DMSO content adjusted to 1% in PBS) were incubated at 22 °C for 10 min in a black 384-well plate in PBS (pH 7.4) with increasing concentrations of AGP to a final volume of 15 µL. The fluorescence anisotropy was measured on a Spectra Max Paradigm multimode plate reader. Experiments were performed in triplicate and the mean anisotropy values fitted to equation (3) using KaleidaGraph 4.1.3.

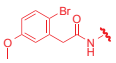
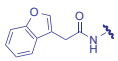
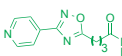
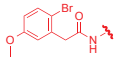
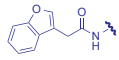
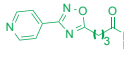
(a)





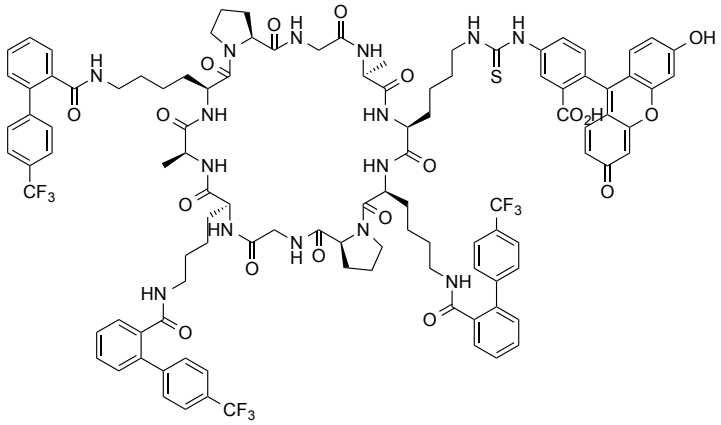


(b)

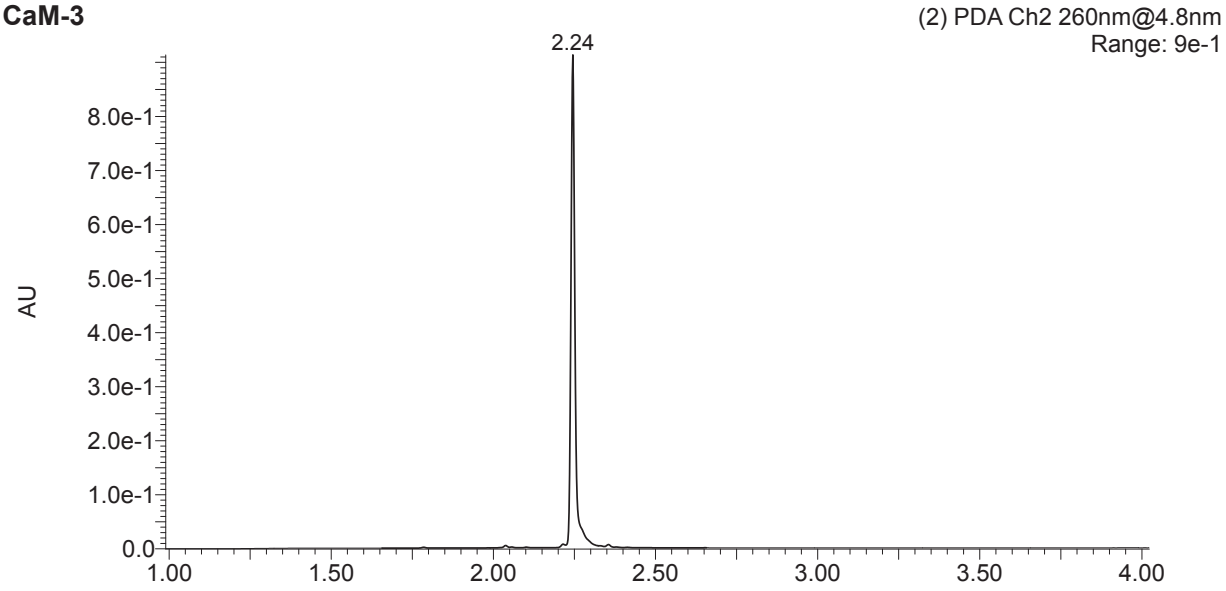
Target	Compound	Enrichment	*K*d (µM)	DE-1		DE-2		DE-3	
AGP	AGP-1	3845-fold	2.2 ± 0.4		36		376		203
AGP	AGP-linear	n.d.	15 ± 3		36		376		203

Supplementary Figure 29 | Comparison of cyclized APG binder and linear APG binder. a, Structure of the linear AGP binder **AGP-1** and **AGP-1-linear**. **b,** Summary of *K*d determination by fluorescence polarization.

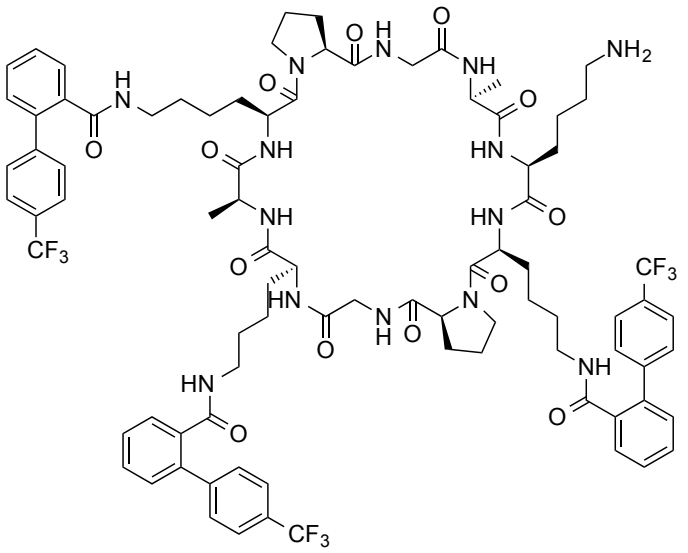
(c) Affinity determination of CaM binders by fluorescence polarization measurements.



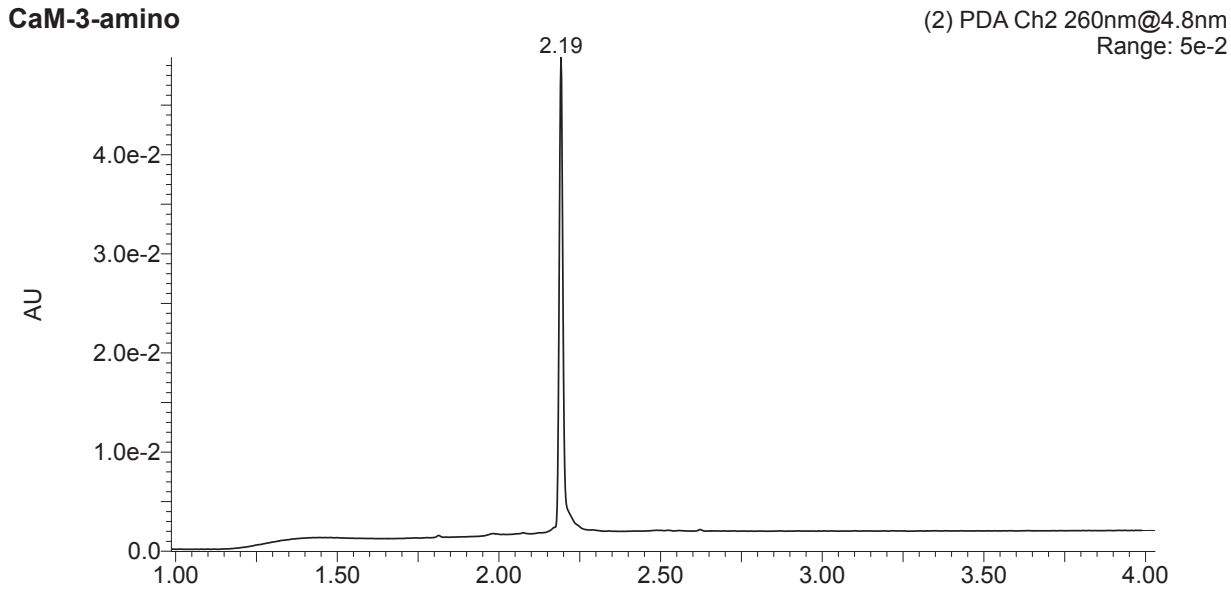
Supplementary Figure 30 | Structure of the CaM binder CaM-3.



Supplementary Figure 31 | UPLC chromatogram of CaM-3. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.



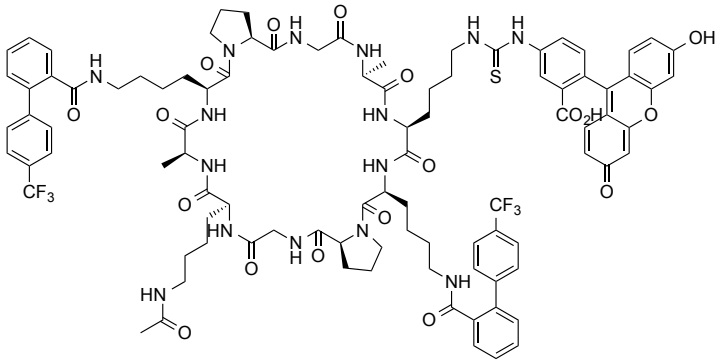
Supplementary Figure 32 | Structure of the CaM binder CaM-3-amino.



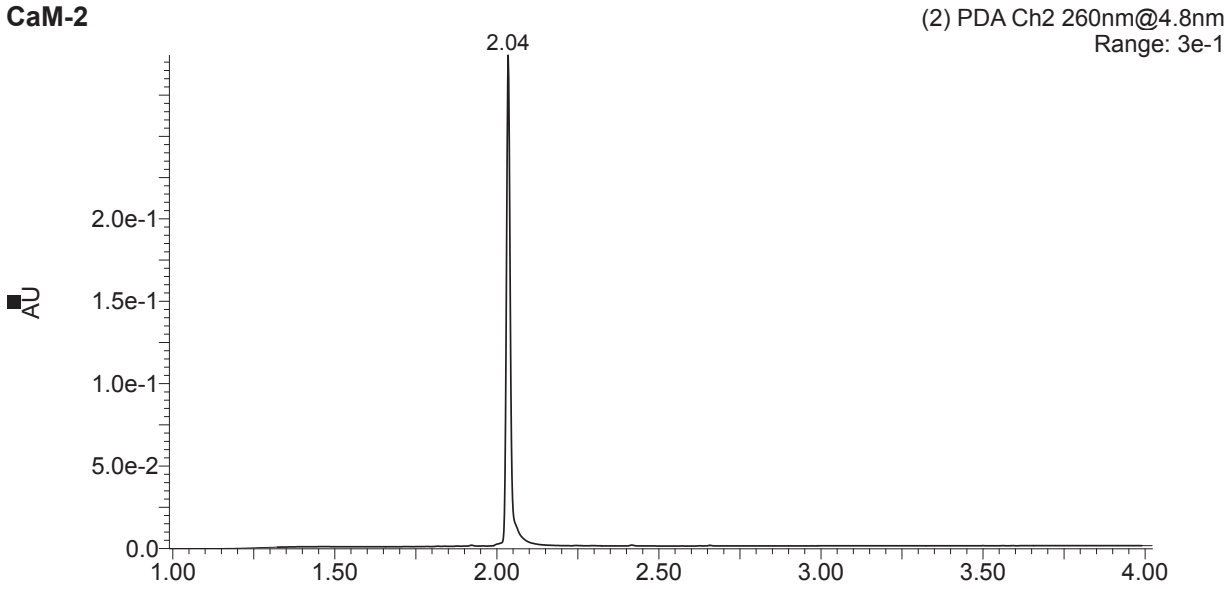
Supplementary Figure 33 | UPLC chromatogram of CaM-3-amino. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

CaM-3 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB241 (2 equiv., SigmaAldrich, Catalog: 346357), Fmoc-*L*-Ala-OH, , Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB241 (2 equiv., SigmaAldrich, Catalog: 346357), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-2-BB314 (2 equiv., SigmaAldrich, Catalog: 346357), Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM)

was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **CaM-3-amino** was recovered as a white powder after lyophilization. **HRMS (*m/z*, $C_{86}H_{100}F_9N_{14}O_{13}$, ESI):** calculated $[M+H]^+$: 1707.7451; found: 1707.7021. **CaM-3-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv., TCI,) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired CaM binder **CaM-3** was recovered as a yellow powder after lyophilization. **HRMS (*m/z*, $C_{107}H_{111}F_9N_{15}O_{18}S$, ESI):** calculated $[M+H]^+$: 2096.7808; found: 2096.7821.

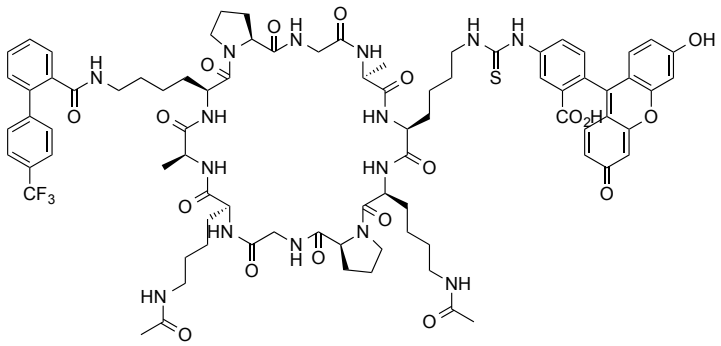


Supplementary Figure 34 | Structure of the CaM binder CaM-2.

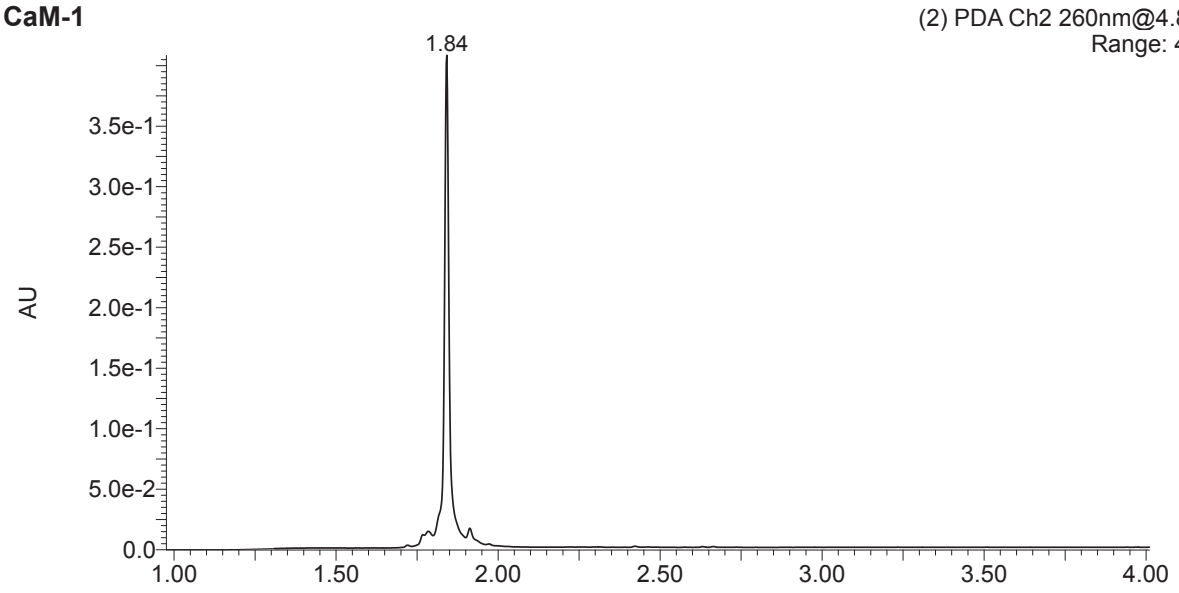


Supplementary Figure 35 | UPLC chromatogram of CaM-2. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100% B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

CaM-2 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB241 (2 equiv.), Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(Ac)-OH, Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-2-BB314 (2 equiv.), Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **CaM-2-amino** was recovered as a white powder after lyophilization. **CaM-2-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv., TCI,) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired CaM binder **CaM-2** was recovered as a yellow powder after lyophilization. **HRMS (*m/z*, $C_{95}H_{106}F_6N_{15}O_{18}S$, ESI):** calculated $[M+H]^+$: 1890.7465; found: 1890.7519.



Supplementary Figure 36 | Structure of the CaM binder CaM-1.



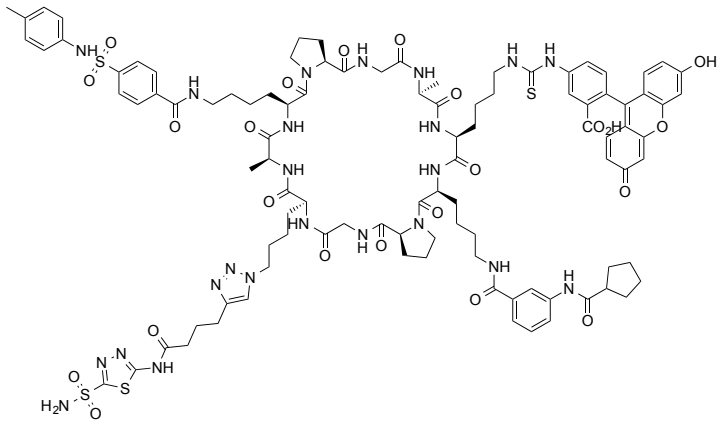
Supplementary Figure 37 | UPLC chromatograms of CaM-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

CaM-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB241 (2 equiv., SigmaAldrich, Catalog: 346357), Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(Ac)-OH, Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Ac)-OH, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **CaM-1-amino** was recovered as a white powder after lyophilization. **CaM-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv., TCI,) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The

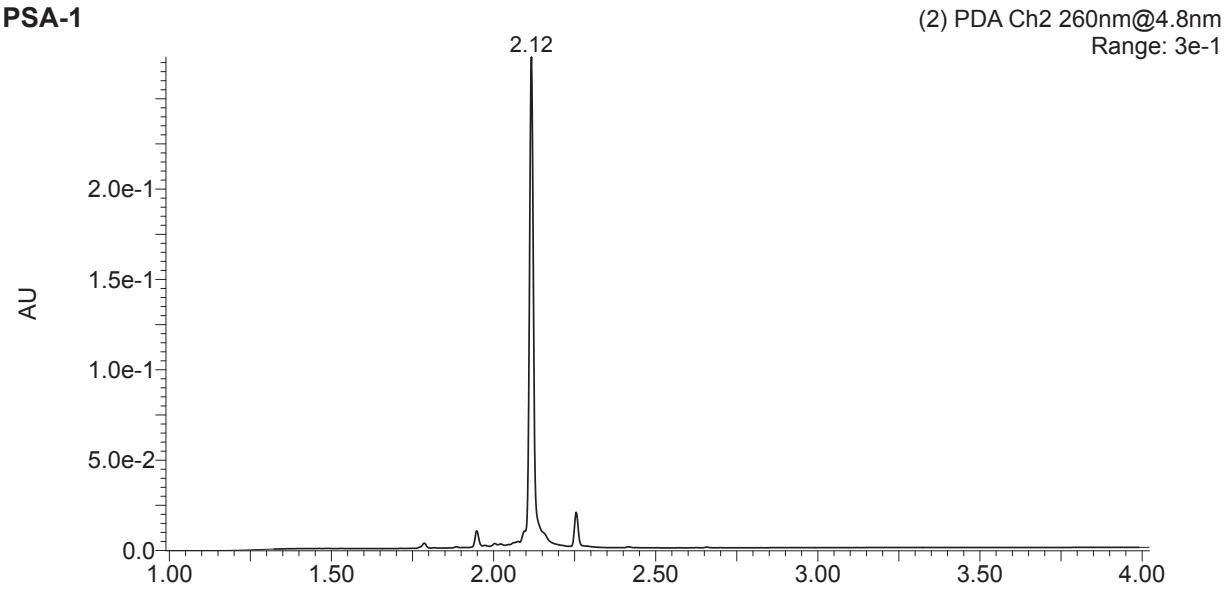
desired CaM binder **CaM-1** was recovered as a yellow powder after lyophilization. **HRMS (*m/z*, $C_{83}H_{101}F_3N_{15}O_{18}S$, ESI):** calculated $[M+H]^+$: 1684.7122; found: 1684.7213.

Fluorescence polarization measurement with CaM binder: Freshly dissolved fluorescein labeled macrocycles (7.5 µL, final concentration 75 nM, final DMSO content adjusted to 1% in DPBS were incubated at 22 °C for 10 min in a black 384-well plate in DPBS (pH 7.4) with increasing concentrations of CaM to a final volume of 15 µL. The fluorescence anisotropy was measured on a Spectra Max Paradigm multimode plate reader. Experiments were performed in triplicate and the mean anisotropy values fitted to equation (3) using KaleidaGraph 4.1.3.

(d) Affinity determination of PSA binders by fluorescence polarization measurements.



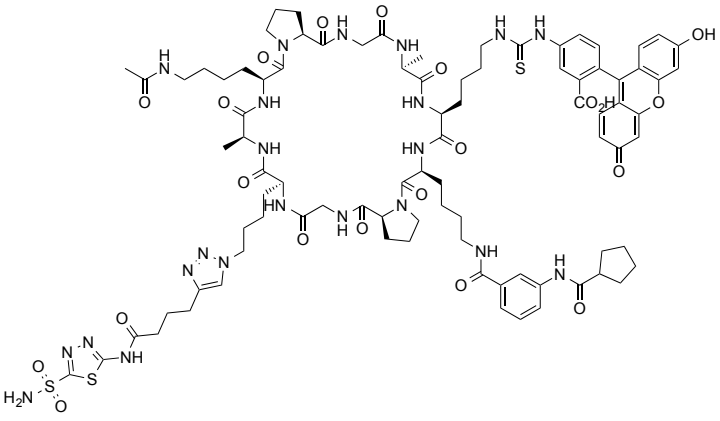
Supplementary Figure 38 | Structure of the PSA binder PSA-1.



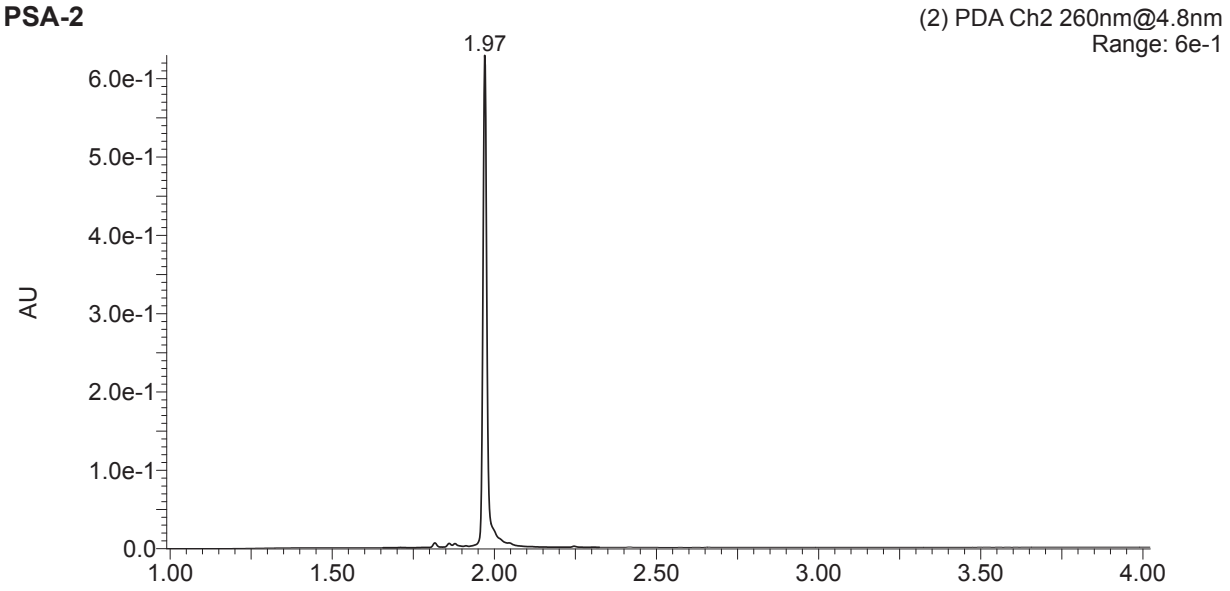
Supplementary Figure 39 | UPLC chromatogram of PSA-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 %

formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

PSA-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB205 (2 equiv., Enamine, Catalog: EN300-00311), Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(N_3)-OH, DE-3-BB17 (2 equiv., *N*-(5-sulfamoyl-1,3,4-thiadiazol-2-yl)hex-5-ynamide[10]); CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-2-BB182 (2 equiv., ChemBridge, Catalog: 9071750), Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **PSA-1-amino** was recovered as a white powder after lyophilization. **PSA-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired PSA binder **PSA-1** was recovered as a yellow powder after lyophilization. **HRMS (*m/z*, $C_{100}H_{122}N_{23}O_{23}S_4$, ESI):** calculated $[M+H]^+$: 2140.7967; found: 2140.7993.



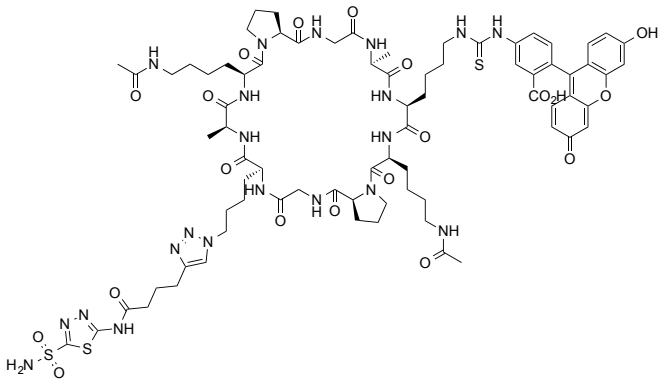
Supplementary Figure 40 | Structure of the PSA binder PSA-2.



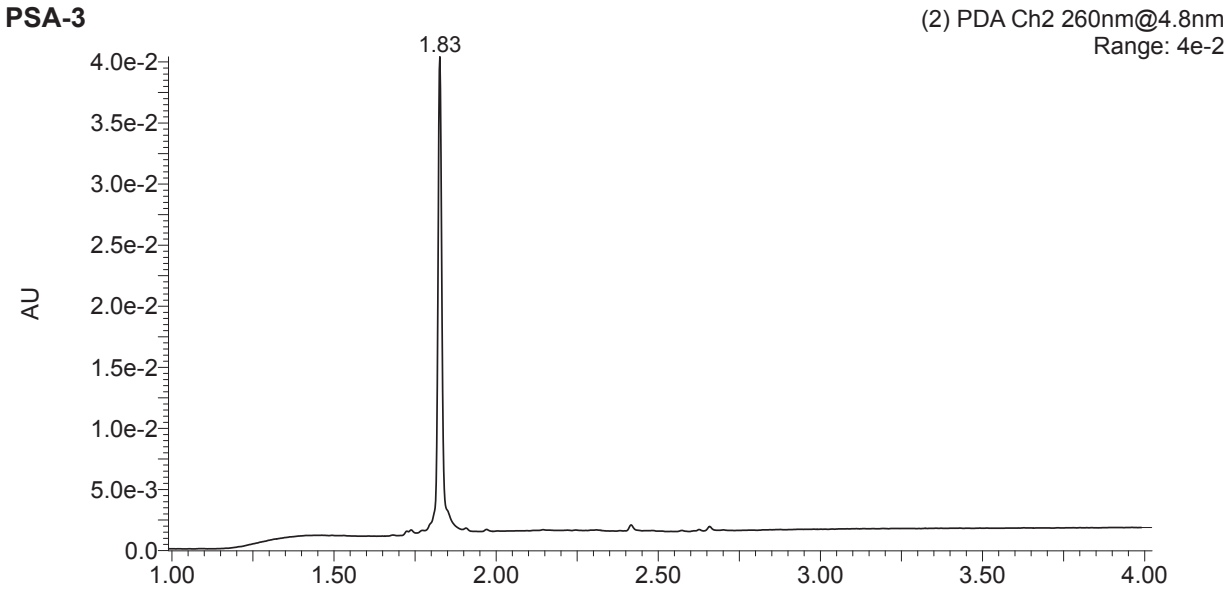
Supplementary Figure 41 | UPLC chromatogram of PSA-2. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

PSA-2 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Ac)-OH, Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(N_3)-OH, DE-3-BB17 (2 equiv., *N*-(5-sulfamoyl-1,3,4-thiadiazol-2-yl)hex-5-ynamide); CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-2-BB182 (2 equiv.), Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **PSA-2-amino** was recovered as a white powder after lyophilization. **PSA-2-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired PSA

binder **PSA-2** was recovered as a yellow powder after lyophilization. **HRMS (*m/z*, $C_{88}H_{113}N_{22}O_{21}S_3$, ESI):** calculated $[M+H]^+$: 1909.7613; found: 1909.7715.



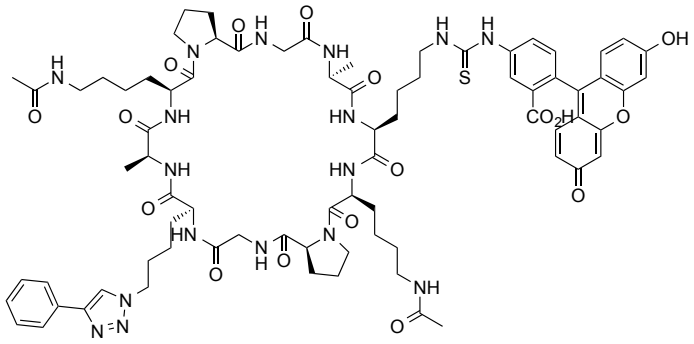
Supplementary Figure 42 | Structure of the PSA binder PSA-3.



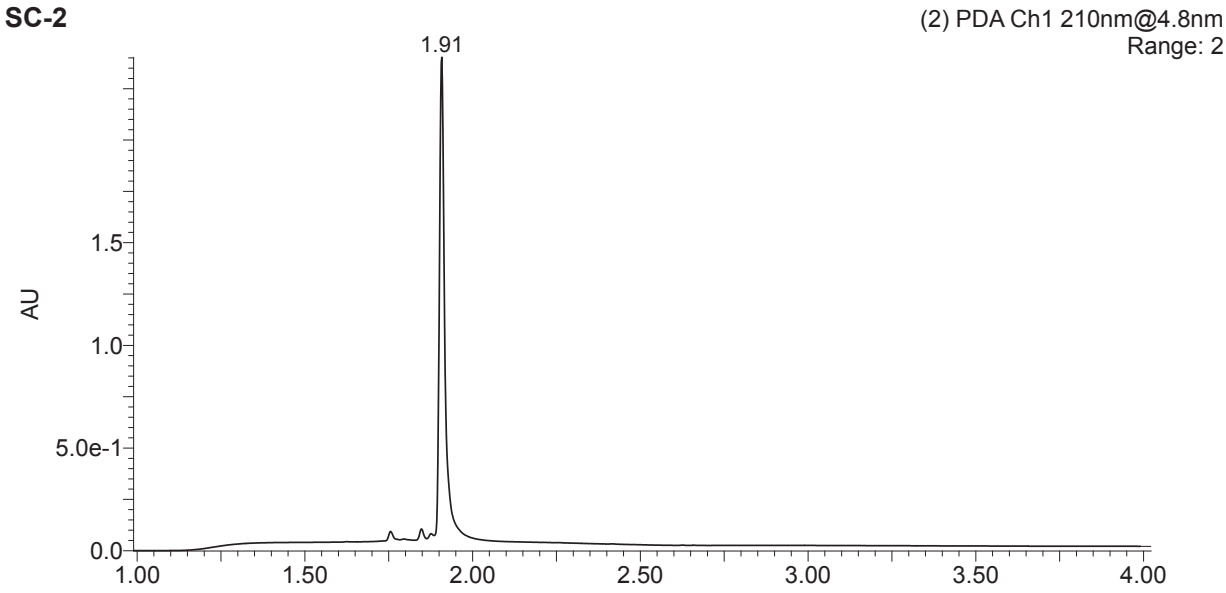
Supplementary Figure 43 | UPLC chromatogram of PSA-3. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

PSA-3 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Ac)-OH, Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(N_3)-OH, DE-3-BB17 (2 equiv., *N*-(5-sulfamoyl-1,3,4-thiadiazol-2-yl)hex-5-ynamide); CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Ac)-OH, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved

by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **PSA-3-amino** was recovered as a white powder after lyophilization. **PSA-3-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired PSA binder **PSA-3** was recovered as a yellow powder after lyophilization. **HRMS (*m/z*, $C_{77}H_{102}N_{21}O_{20}S_3$, ESI):** calculated $[M+H]^+$: 1736.6772; found: 1736.6813.



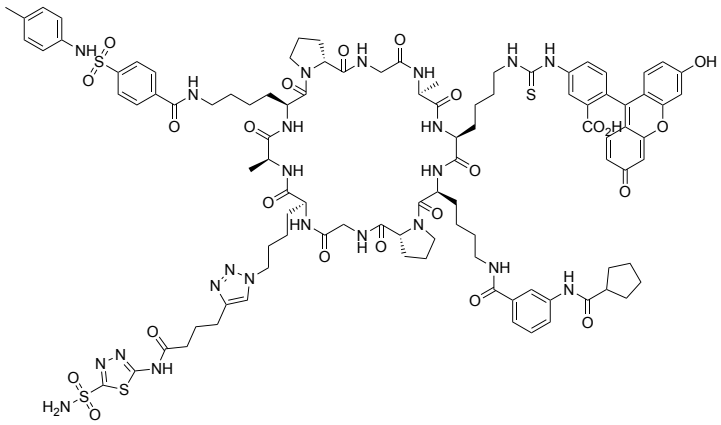
Supplementary Figure 44 | Structure of the scaffold control SC-2.



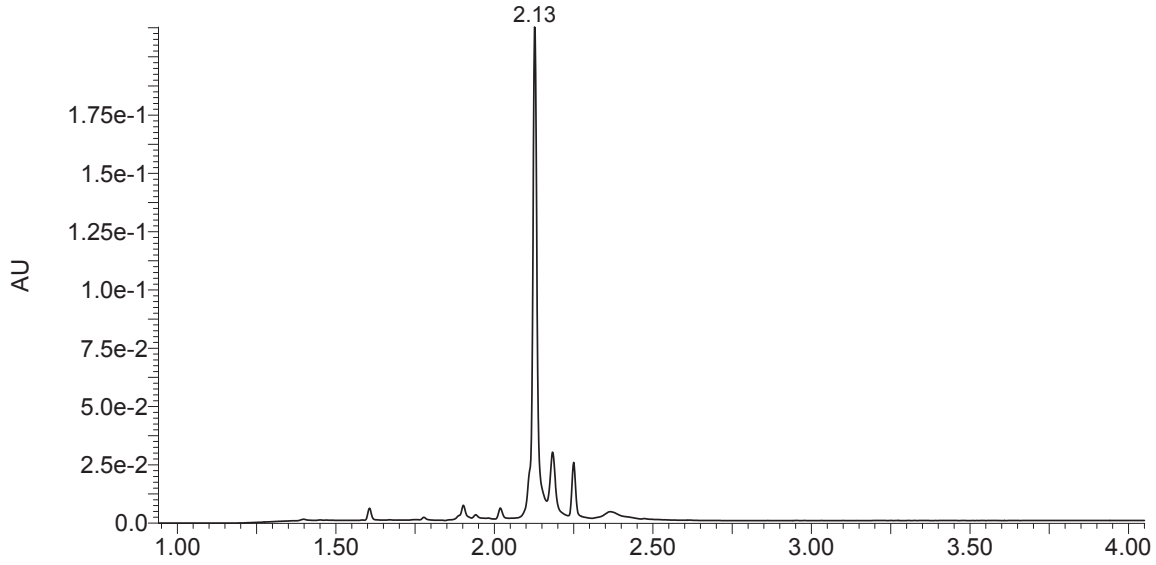
Supplementary Figure 45 | UPLC chromatograms of SC-2. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow

rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 210 nm.

Scaffold control SC-2 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Ac)-OH, Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(N_3)-OH, DE-3-BB17 (2 equiv., Acros, Catalog: 15246-0250, CuI, 10%; TBTA, 10% in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Ac)-OH, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **SC-2-amino** was recovered as a white powder after lyophilization. **SC-2-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired scaffold control **SC-2** was recovered as a yellow powder after lyophilization. **HRMS (*m/z*, $C_{77}H_{98}N_{17}O_{17}S$, ESI):** calculated $[M+H]^+$: 1564.7047; found: 1564.6655.



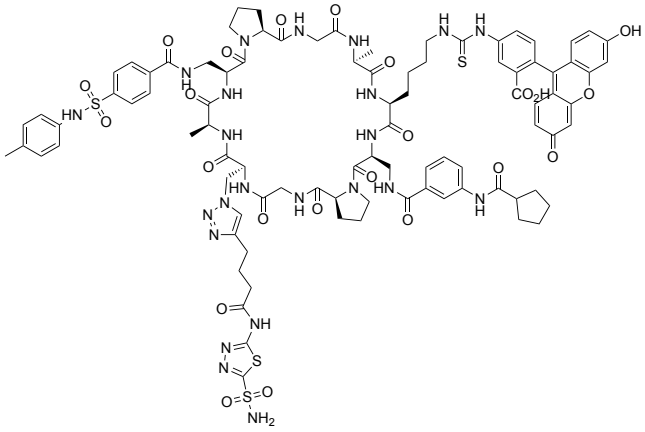
Supplementary Figure 46 | Structure of the PSA binder D-PSA-1.



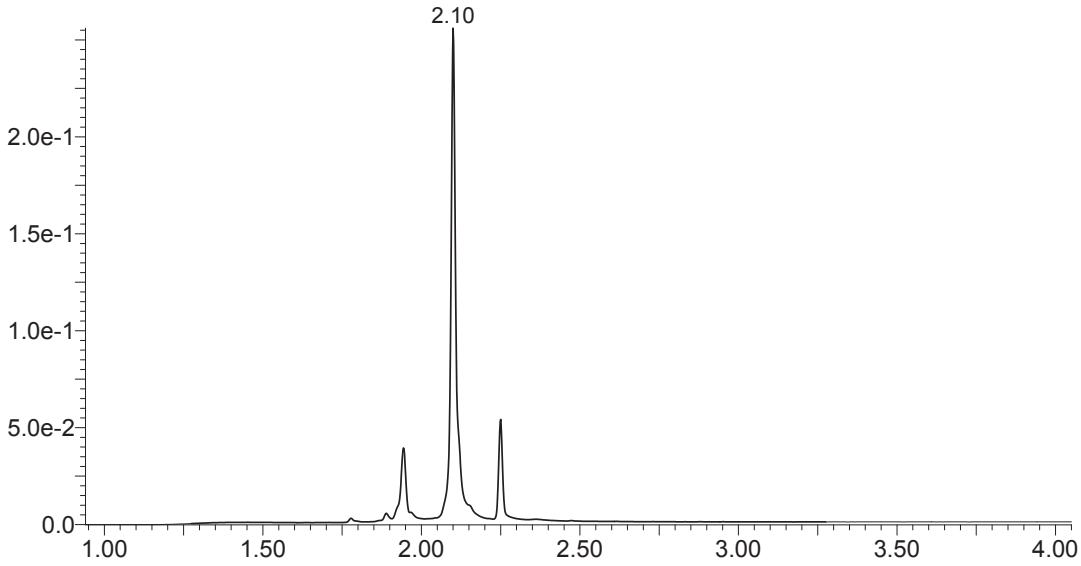
Supplementary Figure 47 | UPLC chromatogram of D-PSA-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

D-PSA-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*D*-Pro-OH (2 equiv., SigmaAldrich, Catalog: 47532), Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB205, Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(N_3)-OH, DE-3-BB17; CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*D*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-2-BB182, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **D-PSA-1-**

amino was recovered as a white powder after lyophilization. **D-PSA-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired PSA binder **D-PSA-1** was recovered as a yellow powder after lyophilization. **HRMS (m/z, $C_{100}H_{122}N_{23}O_{23}S_4$, ESI):** calculated $[M+H]^+$: 2140.7967; found: 2140.7910.



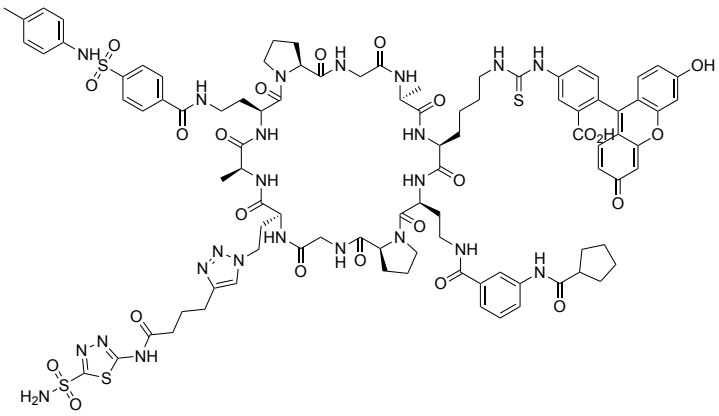
Supplementary Figure 48 | Structure of the PSA binder DAP-PSA-1.



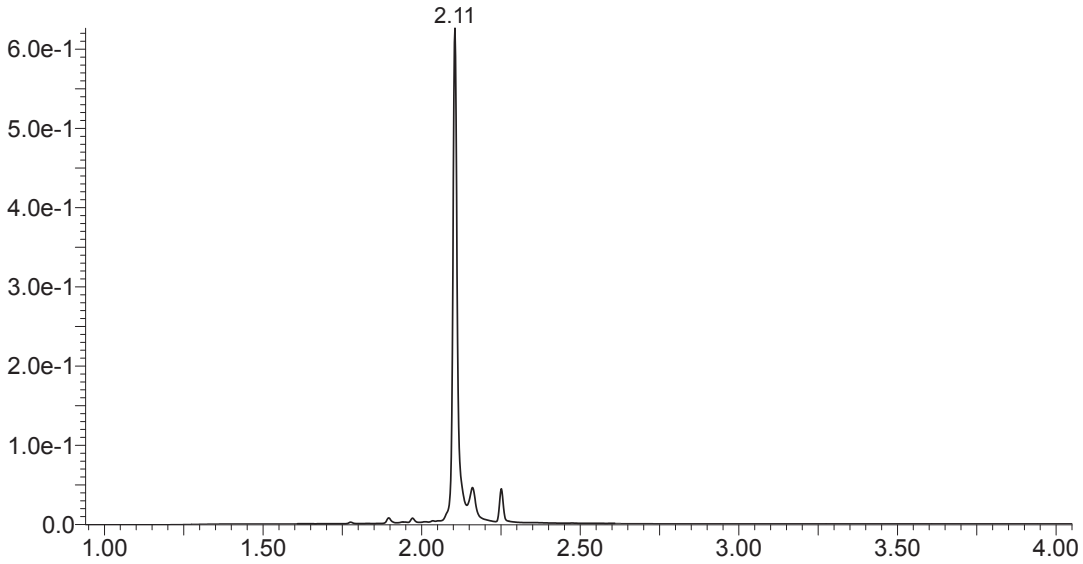
Supplementary Figure 49 | UPLC chromatogram of DAP-PSA-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

DAP-PSA-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Dap(Dde)-OH

(2 equiv., Iris, Catalog: FAA1462), Dde-off, DE-1-BB205, Fmoc-*L*-Ala-OH, Fmoc-*L*-Dap(N_3)-OH(2 equiv., Iris, Catalog: FAA1820), DE-3-BB17; CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Dap(Dde)-OH, Dde-off, DE-2-BB182, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a yellow powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **DAP-PSA-1-amino** was recovered as a white powder after lyophilization. **DAP-PSA-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired PSA binder **DAP-PSA-1** was recovered as a yellow powder after lyophilization. **HRMS (m/z, $C_{91}H_{104}N_{23}O_{23}S_4$, ESI):** calculated $[M+H]^+$: 2014.6658; found: 2014.6520.



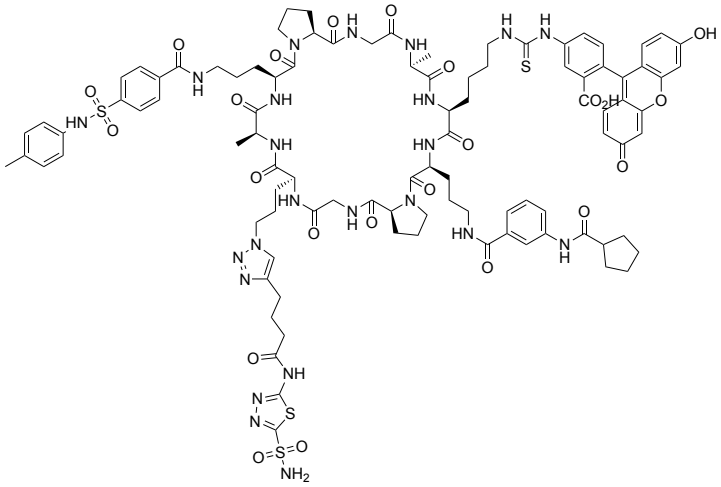
Supplementary Figure 50 | Structure of the PSA binder DAB-PSA-1.



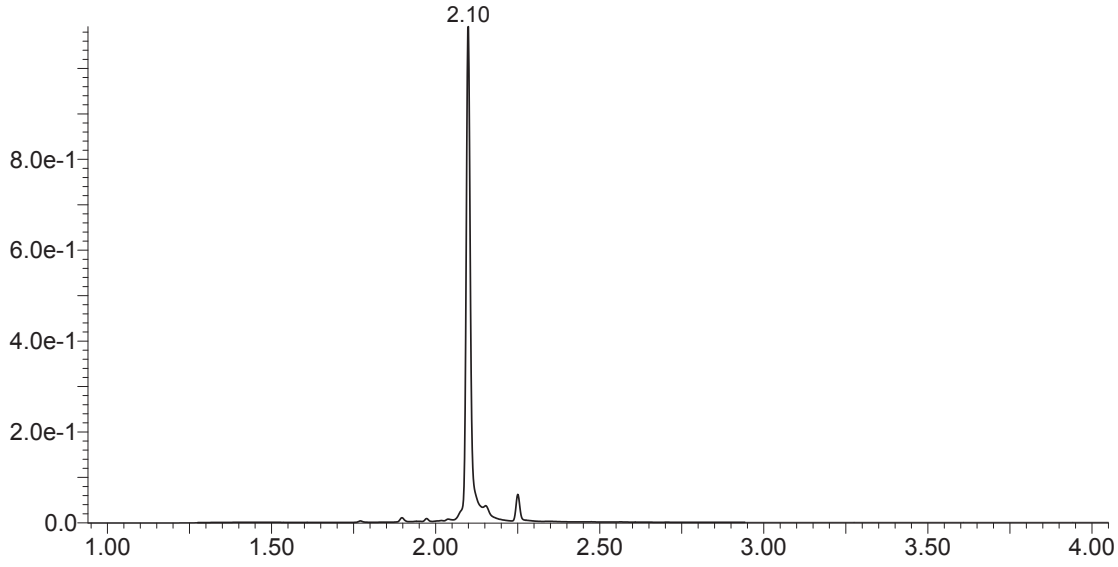
Supplementary Figure 51 | UPLC chromatogram of DAB-PSA-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

DAB-PSA-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Dab(Dde)-OH (2 equiv., Iris, Catalog: FAA1365), Dde-off, DE-1-BB205, Fmoc-*L*-Ala-OH, Fmoc-*L*-Dab(N_3)-OH(2 equiv., Iris, Catalog: FAA6620), DE-3-BB17; CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Dab(Dde)-OH, Dde-off, DE-2-BB182, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **DAB-PSA-1-amino** was recovered as a white powder after lyophilization. **DAB-PSA-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC

purification was performed. The desired PSA binder **DAB-PSA-1** was recovered as a yellow powder after lyophilization. **HRMS (m/z, $C_{94}H_{110}N_{23}O_{23}S_4$, ESI):** calculated $[M+H]^+$: 2056.7028; found: 2056.7057.



Supplementary Figure 52 | Structure of the PSA binder ORN-PSA-1.



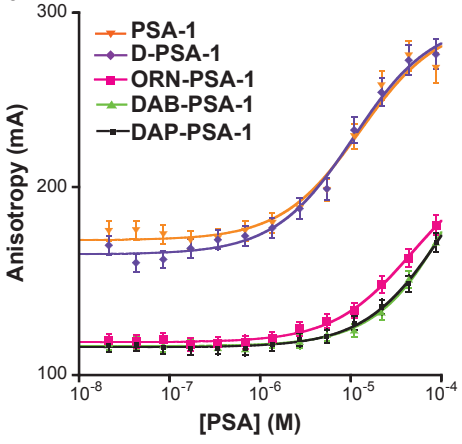
Supplementary Figure 53 | UPLC chromatogram of ORN-PSA-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

ORN-PSA-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Orn(Dde)-OH (2 equiv., Iris, Catalog: FAA1365), Dde-off, DE-1-BB205, Fmoc-*L*-Ala-OH, Fmoc-*L*-Orn(N_3)-OH(2 equiv., Iris, Catalog: FAA6620), DE-3-BB17; CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-

L-Pro-OH, Fmoc-*L*-Orn(Dde)-OH, Dde-off, DE-2-BB182, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **ORN-PSA-1-amino** was recovered as a white powder after lyophilization. **ORN-PSA-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired PSA binder **ORN-PSA-1** was recovered as a yellow powder after lyophilization. **HRMS (m/z, $C_{97}H_{116}N_{23}O_{23}S_4$, ESI):** calculated $[M+H]^+$: 2098.7497; found: 2098.7439.

Fluorescence polarization measurement with PSA binder: Freshly dissolved fluorescein labeled macrocycles (5 µL, final concentration 50 nM, final DMSO content adjusted to 1% in PBS) were incubated at 22 °C for 10 min in a black 384-well plate in PBS (pH 7.4) with increasing concentrations of PSA to a final volume of 10 µL. The fluorescence anisotropy was measured on a Spectra Max Paradigm multimode plate reader. Experiments were performed in triplicate and the mean anisotropy values fitted to equation (3) using KaleidaGraph 4.1.3.

a



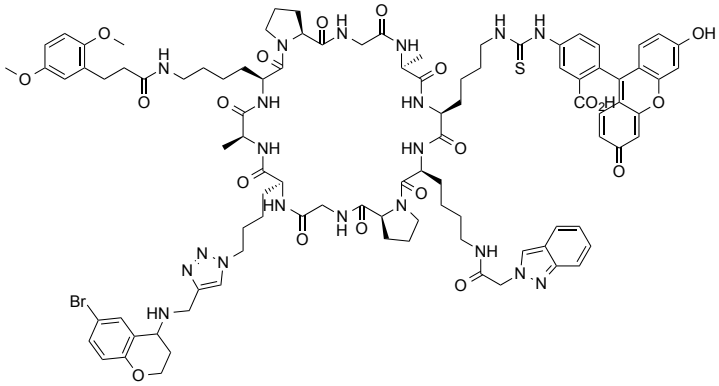
b

Target	Compound	Enrichment	*Kd* (µM)
PSA	PSA-1	1908-fold	13 ± 3
PSA	D-PSA-1	n.d.[1]	11 ± 2
PSA	ORN-PSA-1	n.d.[1]	47 ± 5
PSA	DAB-PSA-1	n.d.[1]	109 ± 18
PSA	DAP-PSA-1	n.d.[1]	161 ± 41

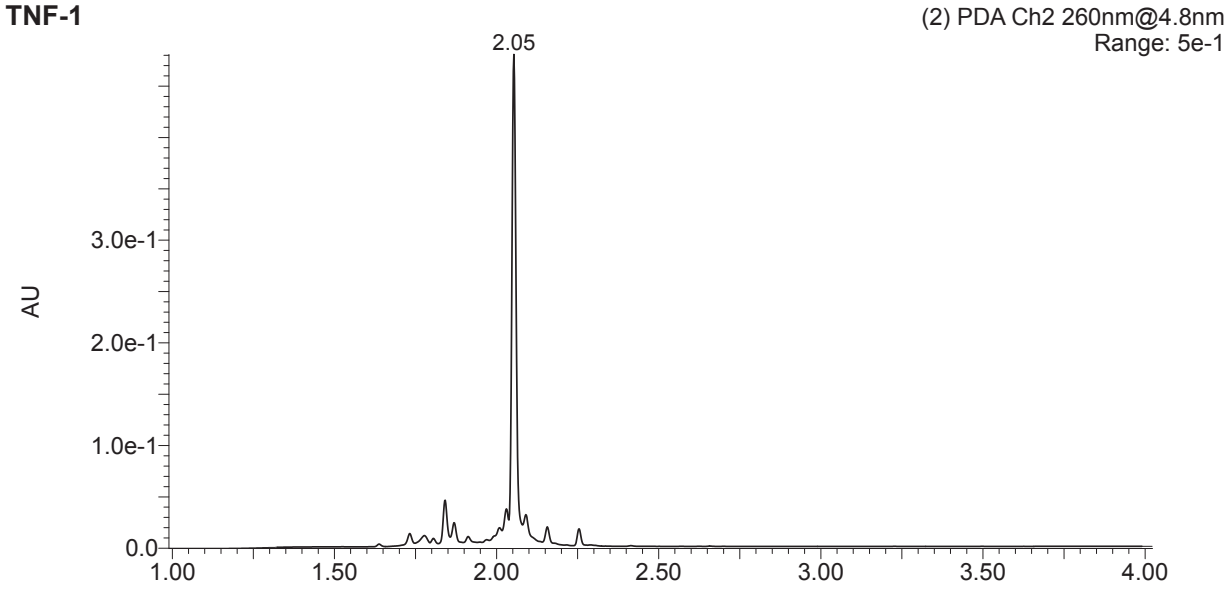
Supplementary Figure 54 | Binder validation of selected cyclic peptides against PSA. a, fluorescence polarization measurements of

selected fluorescently-labeled synthetic cyclic peptides against PSA. Error bars indicate the standard deviation of three measurements. **b**, Enrichments and dissociation constants of synthesized cyclic peptides, chemical structures and corresponding identification numbers (**ID**) of the three diversity elements. For calculation of enrichment see Supplementary Table 7. **n.d.** = not determined.[1] = not included in the library.

(e) Affinity determination of L19-TNF binders by fluorescence polarization measurement.



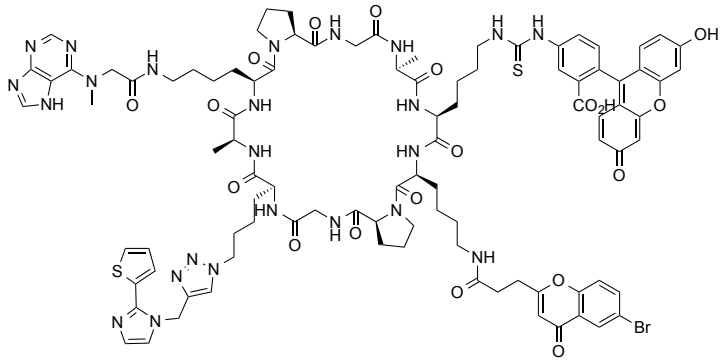
Supplementary Figure 55 | Structure of the L19-TNF binder TNF-1.



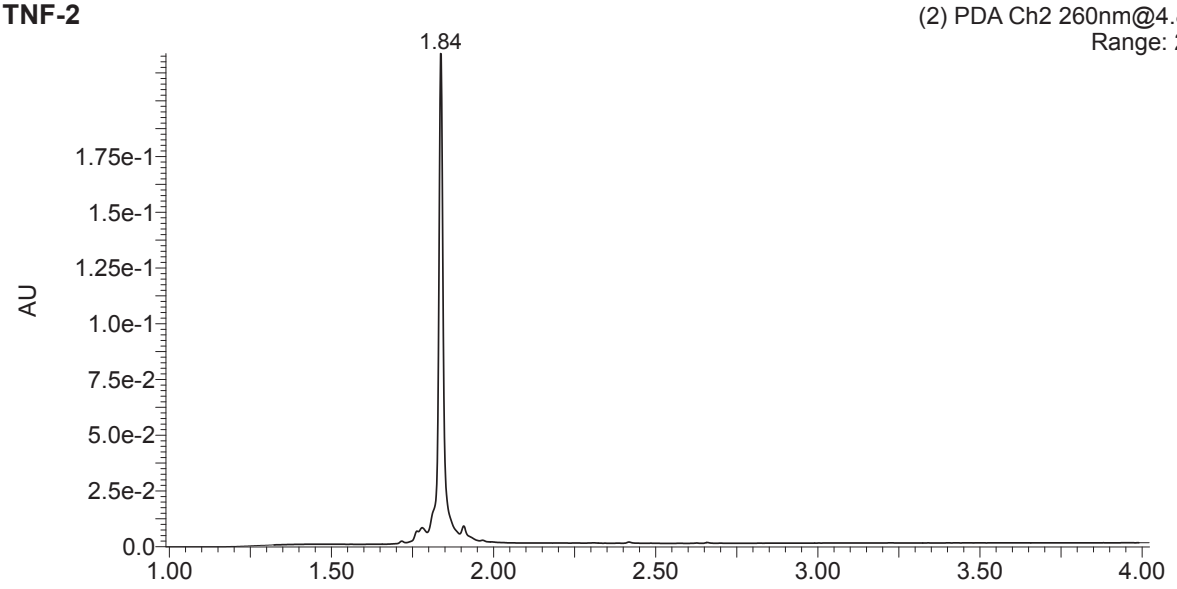
Supplementary Figure 56 | UPLC chromatogram of TNF-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

TNF-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB20 (2 equiv., SigmaAldrich, Catalog: 575658), Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(N_3)-OH, DE-3-BB106 (2 equiv.,

Enamine, Catalog: EN300-75317); CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-2-BB361 (2 equiv., Enamine, Catalog: EN300-71681), Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **TNF-1-amino** was recovered as a white powder after lyophilization. **TNF-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired L19-TNF binder **TNF-1** was recovered as a yellow powder after lyophilization. **HRMS (m/z, $C_{97}H_{118}BrN_{20}O_{20}S$, ESI):** calculated $[M+H]^+$: 1993.7735; found: 1993.7762.



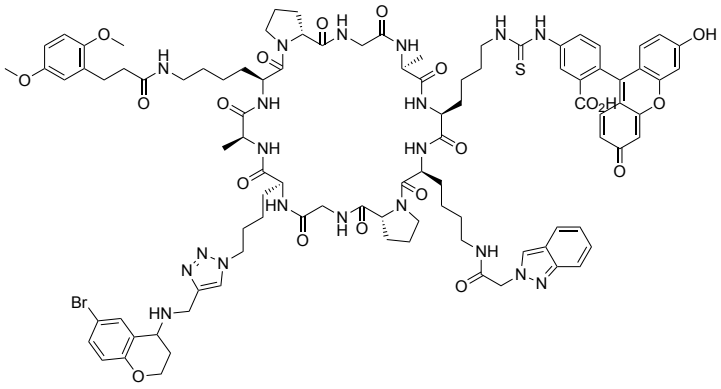
Supplementary Figure 57 | Structure of the L19-TNF binder TNF-2.



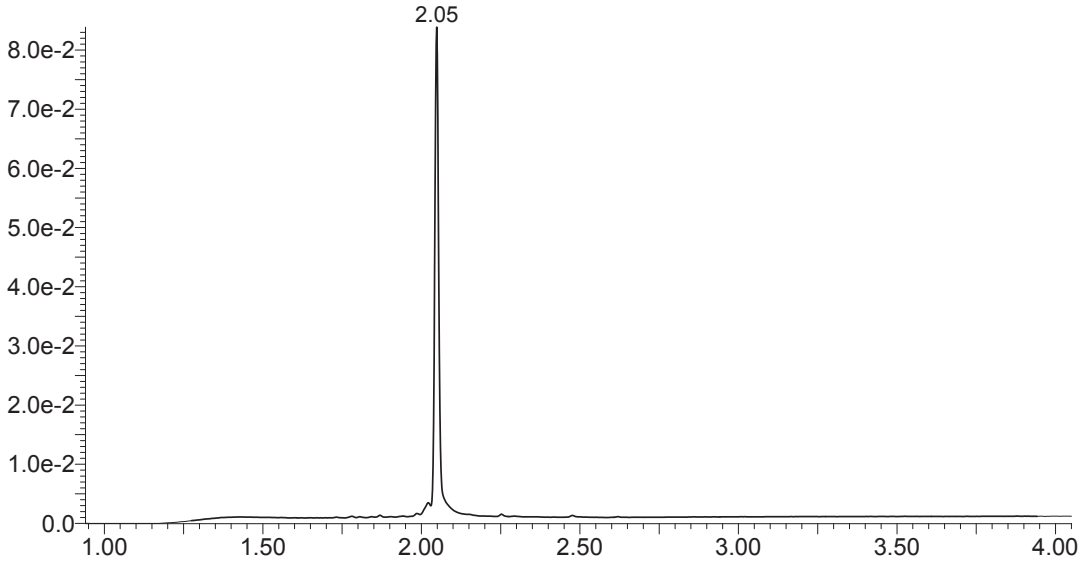
Supplementary Figure 58 | UPLC chromatogram of TNF-2. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

TNF-2 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB170 (2 equiv., Enamine, Catalog: EN300-31422), Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(N_3)-OH, DE-3-BB54 (2 equiv., Enamine, Catalog: EN300-73177); CuI, 10%; TBTA, 10% in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-2-BB65 (2 equiv., Aurora Building Blocks, Catalog: A00.670.978), Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **TNF-2-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed.

The desired L19-TNF binder **TNF-2** was recovered as a yellow powder after lyophilization. **HRMS (m/z, $C_{95}H_{110}BrN_{24}O_{19}S_2$, ESI):** calculated $[M+H]^+$: 2033.7004; found: 2033.7035.



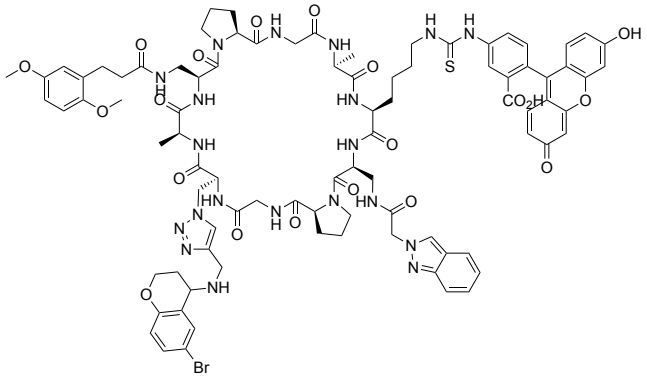
Supplementary Figure 59 | Structure of the L19-TNF binder D-TNF-1.



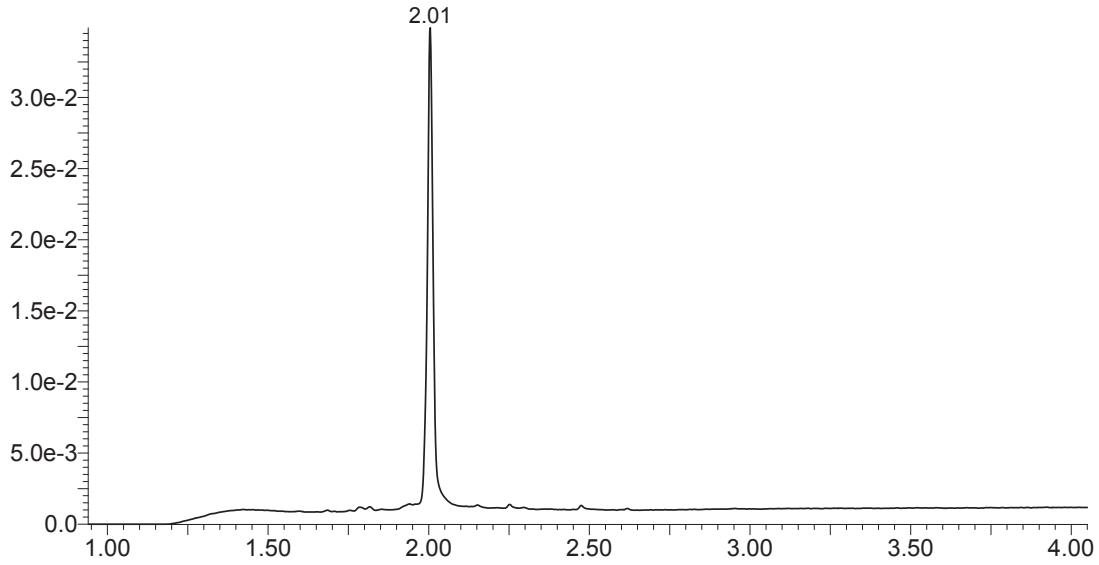
Supplementary Figure 60 | UPLC chromatogram of D-TNF-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

D-TNF-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*D*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB20, Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(N_3)-OH, DE-3-BB106; CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*D*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-2-BB361, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by

CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **D-TNF-1-amino** was recovered as a white powder after lyophilization. **D-TNF-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired L19-TNF binder **D-TNF-1** was recovered as a yellow powder after lyophilization. **HRMS (m/z, $C_{97}H_{118}BrN_{20}O_{20}S$, ESI):** calculated $[M+H]^+$: 1993.7735; found: 1993.7607.

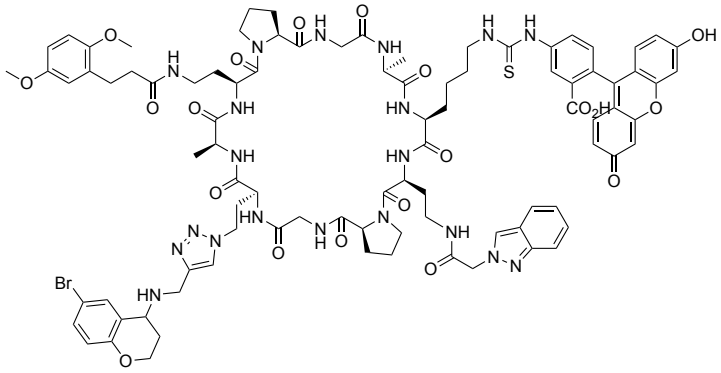


Supplementary Figure 61 | Structure of the L19-TNF binder DAP-TNF-1.

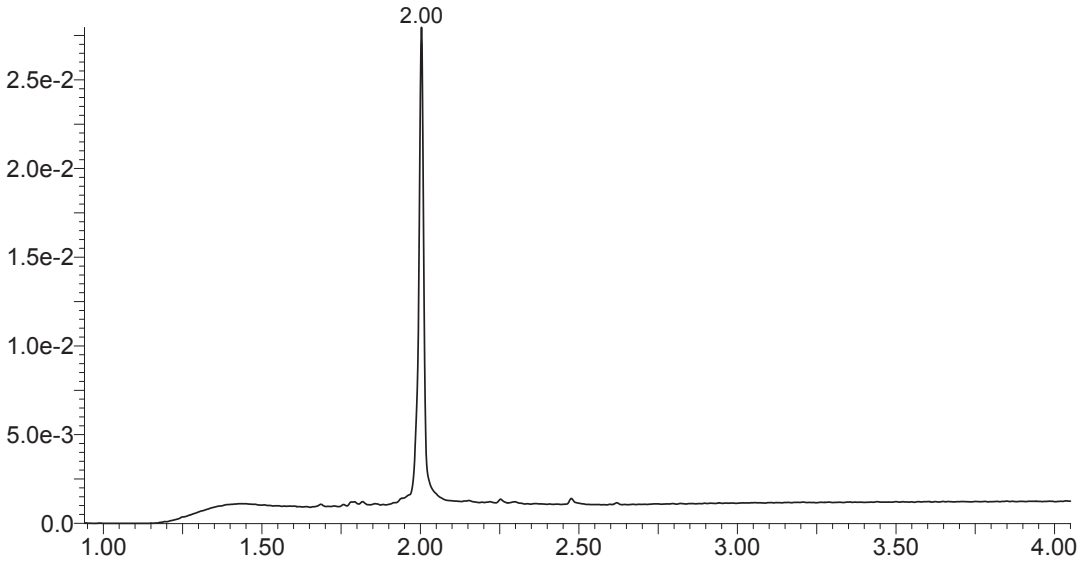


Supplementary Figure 62 | UPLC chromatogram of DAP-TNF-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

DAP-TNF-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Dap(Dde)-OH, Dde-off, DE-1-BB20, Fmoc-*L*-Ala-OH, Fmoc-*L*-Dap(N_3)-OH, DE-3-BB106; CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Dap(Dde)-OH, Dde-off, DE-2-BB361, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **DAP-TNF-1-amino** was recovered as a white powder after lyophilization. **DAP-TNF-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired L19-TNF binder **DAP-TNF-1** was recovered as a yellow powder after lyophilization. **HRMS (m/z, $C_{88}H_{100}BrN_{20}O_{20}S$, ESI):** calculated $[M+H]^+$: 1867.6327; found: 1867.6407.



Supplementary Figure 63 | Structure of the L19-TNF binder DAB-TNF-1.

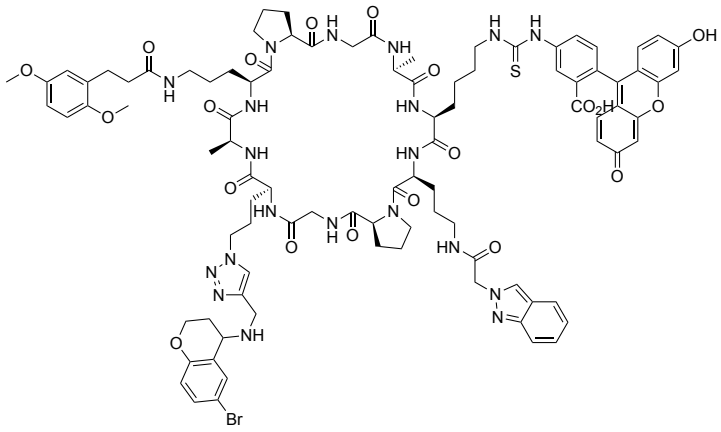


Supplementary Figure 64 | UPLC chromatogram of DAB-TNF-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

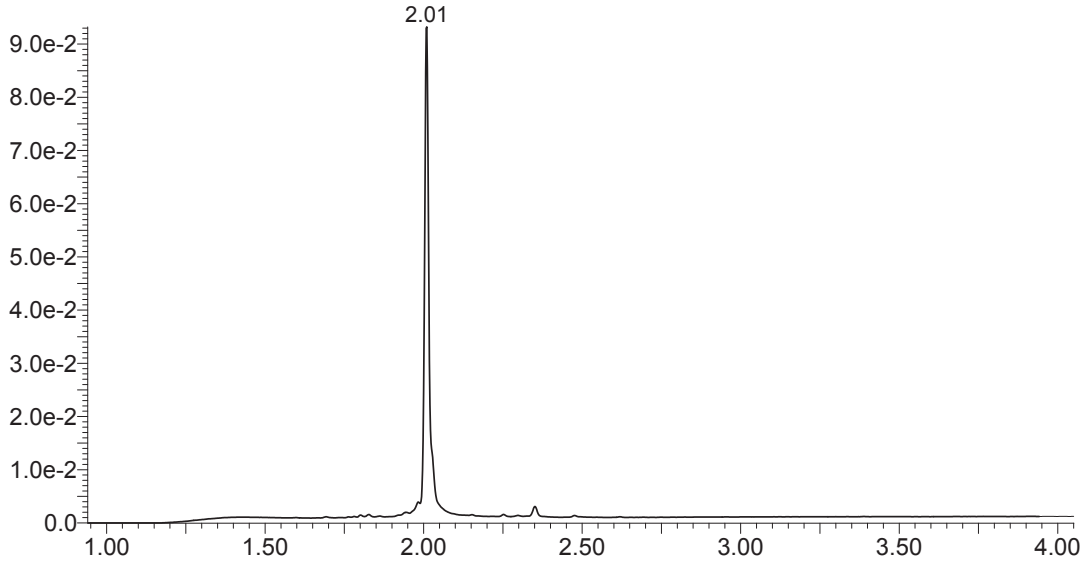
DAB-TNF-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Dab(Dde)-OH (2 equiv., Iris, Catalog: FAA1365), Dde-off, DE-1-BB20, Fmoc-*L*-Ala-OH, Fmoc-*L*-Dab(N_3)-OH(2 equiv., Iris, Catalog: FAA6620), DE-3-BB106; CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Dab(Dde)-OH, Dde-off, DE-2-BB361, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **DAB-TNF-1-amino** was recovered as a white powder after lyophilization. **DAB-TNF-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC

purification was performed. The desired L19-TNF binder **DAB-TNF-1** was recovered as a yellow powder after lyophilization.

HRMS (m/z, $C_{91}H_{106}BrN_{20}O_{20}S$, ESI): calculated $[M+H]^+$: 1909.6763; found: 1909.6758.



Supplementary Figure 65 | Structure of the L19-TNF binder ORN-TNF-1.



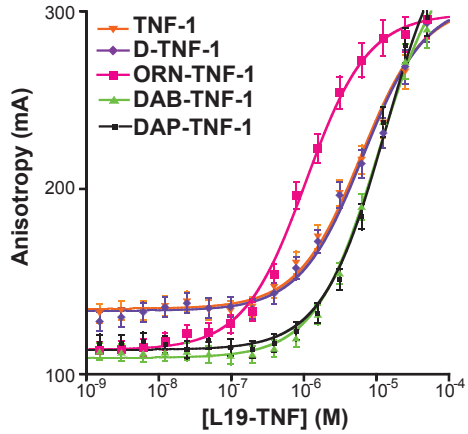
Supplementary Figure 66 | UPLC chromatogram of ORN-TNF-1. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

ORN-TNF-1 synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Orn(Dde)-OH (2 equiv., Iris, Catalog: FAA1365), Dde-off, DE-1-BB20, Fmoc-*L*-Ala-OH, Fmoc-*L*-Orn(N_3)-OH(2 equiv., Iris, Catalog: FAA6620), DE-3-BB106; CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Orn(Dde)-OH, Dde-off, DE-2-BB361, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Ala-OH. The peptide was released

from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **ORN-TNF-1-amino** was recovered as a white powder after lyophilization. **ORN-TNF-1-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired L19-TNF binder **ORN-TNF-1** was recovered as a yellow powder after lyophilization. **HRMS (m/z, $C_{94}H_{112}BrN_{20}O_{20}S$, ESI):** calculated $[M+H]^+$: 1951.7266; found: 1951.7267.

Fluorescence polarization measurement with L19-TNF binder: Freshly dissolved fluorescein labeled macrocycles (7.5 µL, final concentration 50 nM, final DMSO content adjusted to 2% in PBS) were incubated at 22 °C for 10 min in a black 384-well plate in PBS (pH 7.4) with increasing concentrations of L19-TNF to a final volume of 15 µL. The fluorescence anisotropy was measured on a Spectra Max Paradigm multimode plate reader. Experiments were performed in triplicate and the mean anisotropy values fitted to equation (3) using KaleidaGraph 4.1.3.

a



b

Target	Compound	Enrichment	*Kd* (µM)
L19-TNF	TNF-1	4272-fold	6.1± 0.8
L19-TNF	D-TNF-1	n.d.[1]	6.5±1
L19-TNF	ORN-TNF-1	n.d.[1]	1.0 ±0.7
L19-TNF	DAB-TNF-1	n.d.[1]	11±1
L19-TNF	DAP-TNF-1	n.d.[1]	13±2

Supplementary Figure 67 | Binder validation of selected cyclic peptides against L19-TNF. a, fluorescence polarization measurements of selected fluorescently-labeled synthetic cyclic peptides against PSA. Error bars indicate the standard deviation of three measurements. **b**,

Enrichments and dissociation constants of synthesized cyclic peptides, chemical structures and corresponding identification numbers (**ID**) of the three diversity elements. For calculation of enrichment see Supplementary Table 7. **n.d.** = not determined.[1] = not included in the library.

(f) Affinity determination of TNF binders by fluorescence polarization measurements.

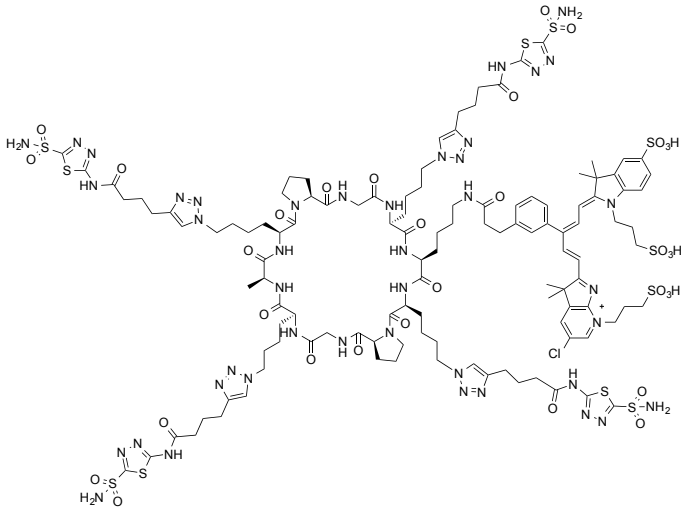
Fluorescence polarization measurements with TNF binder: Freshly dissolved fluorescein labeled macrocycles (7.5 µL, final concentration 50 nM, final DMSO content adjusted to 2% in PBS) were incubated at 22 °C for 10 min in a black 384-well plate in PBS (pH 7.4) with increasing concentrations of TNF to a final volume of 15 µL. The fluorescence anisotropy was measured on a Spectra Max Paradigm multimode plate reader. Experiments were performed in triplicate and the mean anisotropy values fitted to equation (3) using KaleidaGraph 4.1.3.

6. Immunofluorescence Performance of Selected PSA Binder.

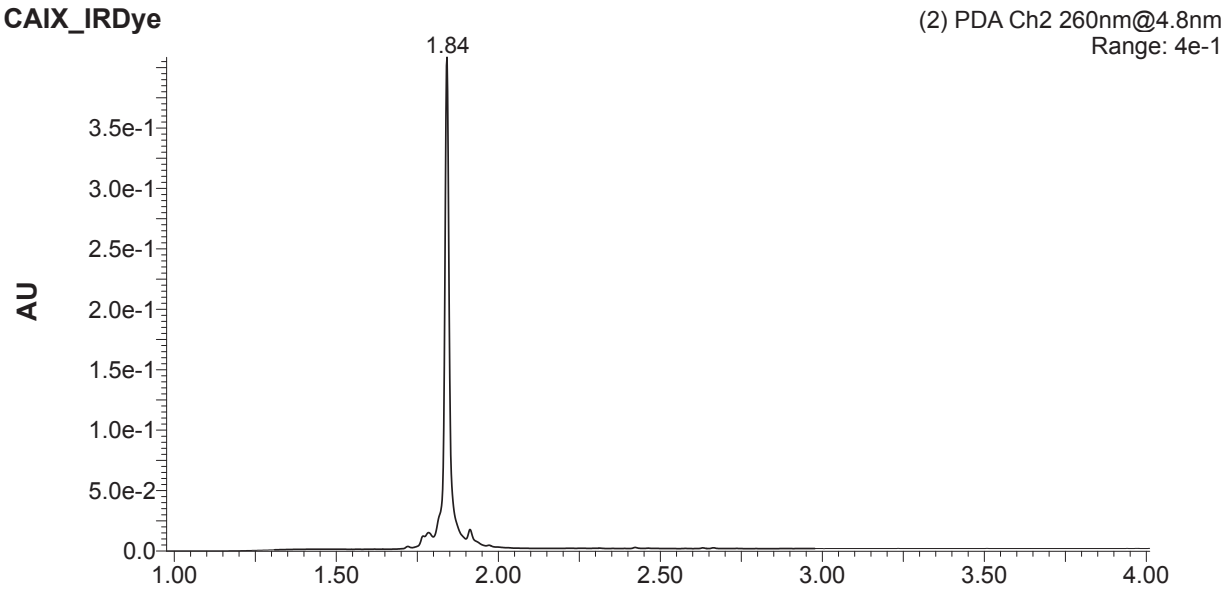
Immunofluorescence staining of PSA-1 and SC-2 on OCT-embedded snap frozen sections: OCT-embedded frozen sections were defrosted and fixed in acetone (- 20 °C, 10 min). The sections were blocked with 20 % FCS (fetal calf serum, Invitrogen) in PBS (pH = 7.4) for 30 min. Sections were then washed (PBS, 5 min) before staining cell nuclei with DAPI (SigmaAldrich, Catalog: D9542). After two additional rounds of washing, **PSA-1** (100 µL, 0.5 µM in PBS with 5 % DMSO) or **SC-2** (100 µL, 0.5 µM in PBS with 5 % DMSO, as control) were then added to sections and incubated at room temperature for 2 h under dark. For detection, rabbit anti-FITC antibody (1:1000, Bio-Rad, Catalog: 4510-7804, 1h) was added to sections followed by goat anti-rabbit antibody Alexa 488 (1:500, Invitrogen, Catalog: A11008, 1h). After additional washing steps (PBS, 5 min), sections were dried at room temperature under dark and mounted with DAKO fluorescent mounting medium (SigmaAldrich, Catalog: M1289). Slides were analyzed with Axioskop2 mot plus microscope (Zeiss).

Immunofluorescence staining of anti-PSA antibody to OCT-embedded snap frozen sections: OCT-embedded frozen sections were defrosted and fixed in acetone (- 20 °C, 10 min). The sections were blocked with 20 % FCS (fetal calf serum, Invitrogen) in PBS (pH = 7.4) for 30 min. Sections were then washed (PBS, 5 min) before staining cell nuclei with DAPI. After two additional rounds of washing, Mouse anti-PSA antibody (1: 25, DAKO, Catalog: M0750, PBS with 2 % BSA) was then added to sections and incubated at room temperature for 2 h under dark. For detection, goat anti-mouse antibody Alexa 488 (1:5000, Invitrogen, Catalog: A11029, 1h) was added to sections. After additional washing steps (PBS, 5 min), sections were dried at room temperature under dark and mounted with DAKO fluorescent mounting medium. Slides were analyzed with Axioskop2 mot plus microscope.

7. *In vivo* Performance of CAIX Binder.



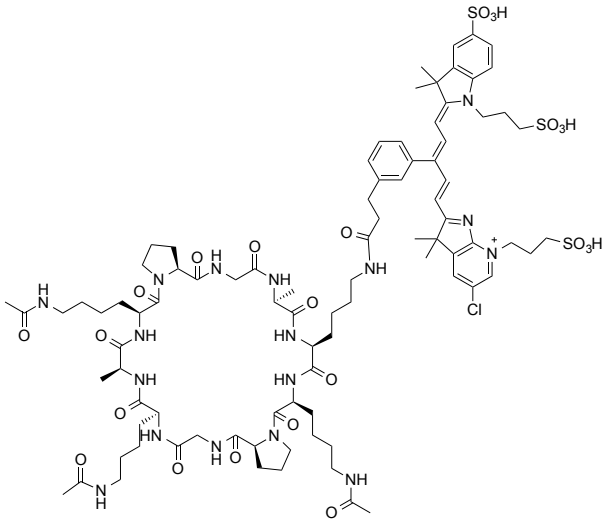
Supplementary Figure 68 | Structure of the CAIX binder IRDye conjugate CAIX-IRDye.



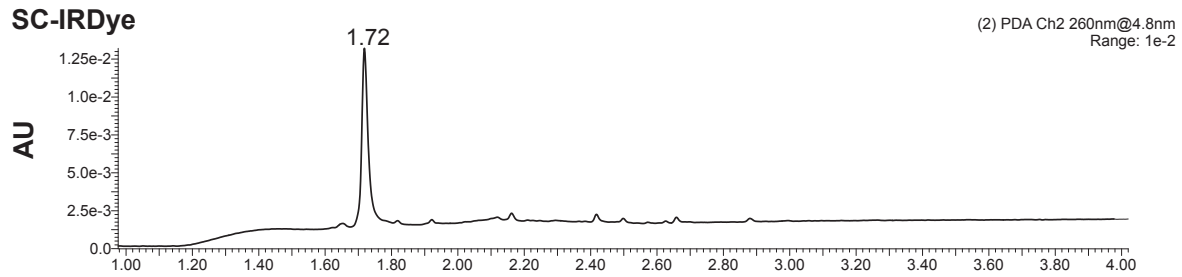
Supplementary Figure 69 | UPLC chromatogram of CAIX-IRDye. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

CAIX-IRDye conjugate synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(N_3)-OH, Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(N_3)-OH, Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(N_3)-OH, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*-Lys(N_3)-OH, DE-3-BB17 (8 equiv., *N*-(5-sulfamoyl-1,3,4-thiadiazol-2-yl)hex-5-ynamide; CuI, 10 %; TBTA, 10 % in 6 mL of THF : DCM = 4 : 1 overnight at 25 °C). The peptide was released from the resin using cleavage

solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **CAIX-IRDye-amino** was recovered as a white powder after lyophilization. **CAIX-IRDye-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). IRDye@680 RD NHS (0.9 equiv., LI-COR, Catalog: P/N 929-70050) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired CAIX binder **CAIX-IRDye** was recovered as a purple powder after lyophilization. **HRMS (*m/z*, $C_{118}H_{160}ClN_{42}O_{32}S_{11}^+$, ESI):** calculated $[M+H]^{2+}$: 1531.9398 found: 1532.4447.



Supplementary Figure 70 | Structure of the scaffold control IRDye conjugate SC-IRDye.



Supplementary Figure 71 | UPLC chromatogram of SC-IRDye. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

Scaffold control IRDye conjugate SC-IRDye synthesis: SC-1-amino (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). IRDye@680 RD NHS (0.9 equiv., LI-COR, Catalog: P/N 929-70050) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired **SC-IRDye** was recovered as a purple powder after lyophilization. **HRMS (*m/z*, $C_{89}H_{127}ClN_{17}O_{23}S_3^+$, ESI):** calculated $[M]^+$: 1932.8136, found: 1932.8158.

Cell Cultures: The human renal cell carcinoma cell line SKRC-52 was kindly provided by Professor E. Oosterwijk (Radbound University Nijmegen Medical Centre, Nijmegen, The Netherlands) and tested for CAIX expression by immunofluorescence. Cells were cultured in adhesion in RPMI medium (Invitrogen) supplemented with fetal calf serum (10%, FCS, Invitrogen) and Antibiotic-Antimycotic (1%, AA, Invitrogen) at 37 °C and 5% CO_2. For passaging, cells were detached using Trypsin-EDTA 0.05% (Invitrogen) when reaching 90% confluence and re-seeded at a dilution of 1:6.

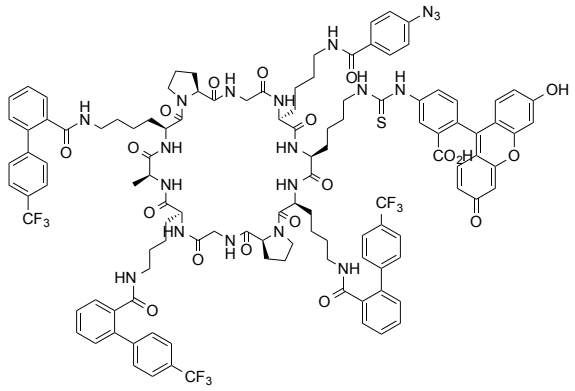
Animal Studies: All animal experiments were conducted in accordance with Swiss animal welfare laws and regulations under the license number 27/2015 granted by the Veterinäramt des Kantons Zürich.

Implantation of Subcutaneous SKRC-52 Tumors: SKRC-52 cells were grown to 80% confluence and detached with Trypsin-EDTA 0.05% (Life Technologies). Cells were washed with Hank's Balanced Salt Solution (HBSS, pH 7.4) once, counted and re-suspended in HBSS to a final concentration of 3.4×10^7 cells/ml. Aliquots of 5×10^6 cells (120 µl of a suspension) were injected subcutaneously in the right flank of female athymic BALB/c nu/nu mice (6-8 weeks of age, Janvier).

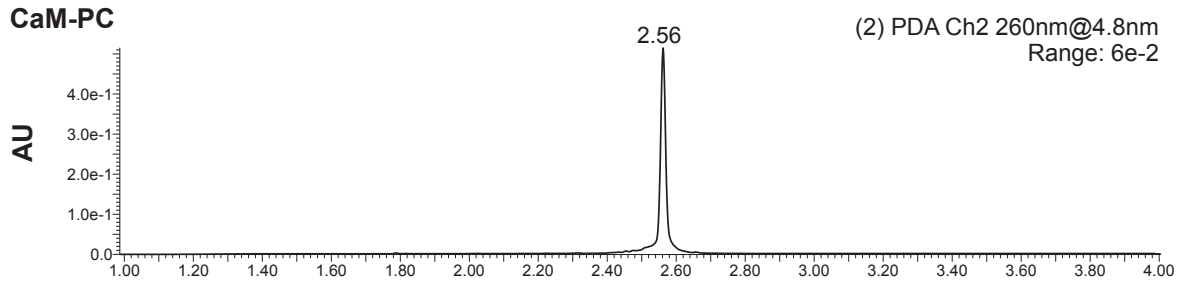
IVIS Imaging: Female BALB/c nude mice bearing subcutaneous SKRC-52 tumors were injected intravenously with compound **CAIX-IRDye** or **SC-IRDye** (3 nmol), dissolved in sterile PBS, containing 10% of DMSO to increase the solubility. Mice were anesthetized with isoflurane and fluorescence images acquired on an IVIS Spectrum imaging system (Xenogen, exposure 1s, binning factor 8, excitation at 675 nm, emission filter at 720 nm, f number 2, field of view 13.1). Images were taken after 5 min, 1 h, 3 h and 7 h after compounds administration. Food and water was given ad libitum during that period. At 24 h time point, animals were sacrificed by cervical dislocation and pictures of organs were taken under the IVIS camera (Xenogen, exposure 1s, binning factor 8, excitation at 675 nm, emission filter at 720 nm, f number 2, field of view 13.1).

8. Performance of Chemical Probes Developed from Selected CaM binder.

(a) Synthesis of CaM specific probe and scaffold control probe.



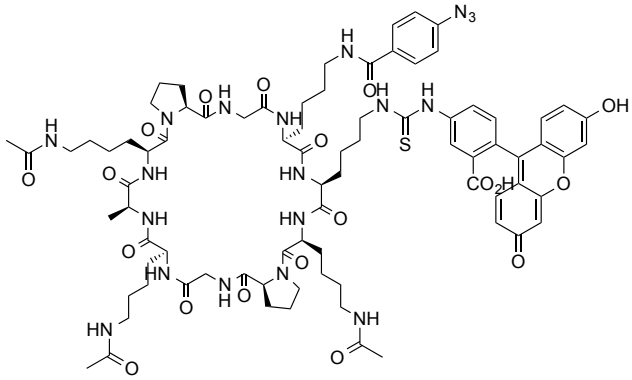
Supplementary Figure 72 | Structure of the CaM specific probe CaM-PC.



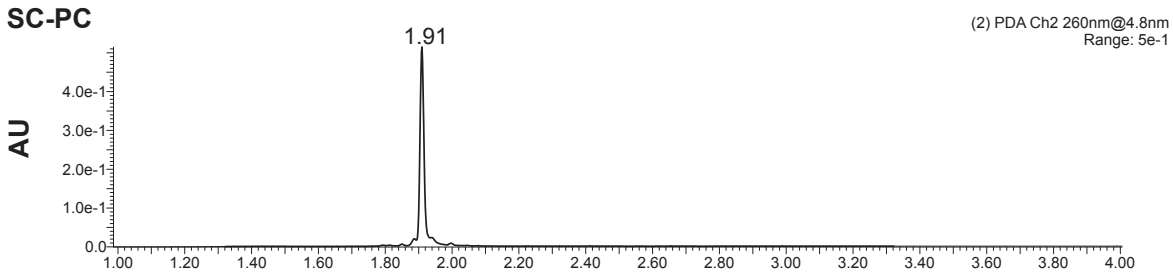
Supplementary Figure 73 | UPLC chromatogram of CaM-PC. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

CaM specific probe CaM-PC synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB241 (2 equiv., SigmaAldrich, Catalog: 346357), Fmoc-*L*-Ala-OH, , Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-1-BB241 (2 equiv., SigmaAldrich, Catalog: 346357), Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Dde)-OH, Dde-off, DE-2-BB314 (2 equiv., SigmaAldrich, Catalog: 346357), Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*- Fmoc-*L*-Lys(Dde)-OH, Dde-off, 4-azidobenzoic acid (2 equiv., TCI, Catalog: A0930),. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after

precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **CaM-PC-amino** was recovered as a white powder after lyophilization. **CaM-PC-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv., TCI,) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired CaM binder **CaM-PC** was recovered as a white powder after lyophilization. **HRMS (*m/z*, $C_{117}H_{121}F_9N_{19}O_{19}S$, ESI):** calculated $[M+H]^+$: 2298.8663; found: 2298.8625.



Supplementary Figure 74 | Structure of the scaffold control probe SC-PC.

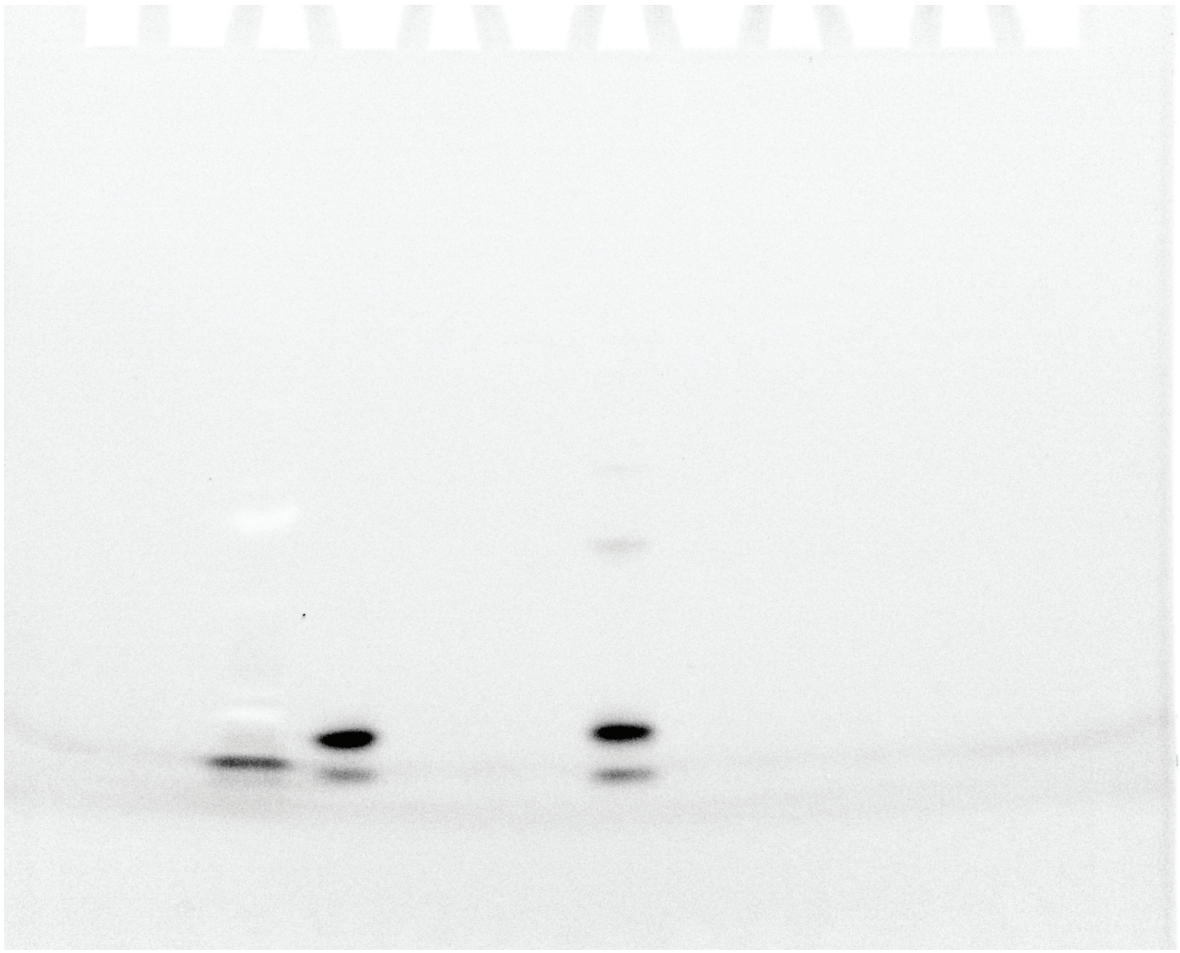


Supplementary Figure 75 | UPLC chromatogram of SC-PC. UPLC analyses were performed on a BEH C18 2.1 × 50 mm column at a flow rate of 0.6 mL/min with gradient: 5 % B (0 to 0.5 minutes), 5 % to 100 % B (0.5 to 4 minutes), 100 % B (4 to 6 minutes), (A= 0.1 % formic acid in water, B= CH_3CN with 0.1 % formic acid), at 40 ºC. Detection by absorbance at 260 nm.

Scaffold control probe SC-PC synthesis: The linear decapeptide was assembled on Fmoc-Gly-2-chlorotritylchloride® Tenta gel-resin (0.1 mmol, loading of 0.17 mmol/g) using the general procedure with following sequence: Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Ac)-OH (2 equiv., Senn, Catalog: 101317), Fmoc-*L*-Ala-OH, Fmoc-*L*-Lys(Ac)-OH, Fmoc-Gly-OH, Fmoc-*L*-Pro-OH, Fmoc-*L*-Lys(Ac)-OH, Fmoc-*L*-Lys(Boc)-OH, Fmoc-*L*- Fmoc-*L*-Lys(Dde)-OH, Dde-off, 4-azidobenzoic acid (2 equiv., TCI, Catalog: A0930),. The peptide was released from the resin using cleavage solution of TFE/AcOH/DCM (2:1:7). Linear protected peptide was obtained as a white solid powder after precipitation, washing with diethyl ether, dissolved by CH_3CN/H_2O (1:1) and lyophilized. The linear decapeptide (5 mM) was dissolved in DMF and the pH was adjusted to 9 by addition of DIPEA. PyBOP (1.0 equiv.) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired cyclodecapeptide was recovered as a white powder after lyophilization. To the cyclodecapeptide was added 1 mL H_2O and 20 mL TFA and the solution stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1). The desired cyclodecapeptide **SC-PC-amino** was recovered as a white powder after lyophilization. **SC-PC-amino** (0.5 M) was dissolved in DMF and DIPEA was added (5.0 equiv.). FITC (1.2 equiv., TCI,) was added and the solution stirred at 25 ºC for 1 h. Solvent was removed under reduced pressure and the residue dissolved in CH_3CN/H_2O (1:1) and reverse-phase HPLC purification was performed. The desired CaM binder **SC-PC** was recovered as a white powder after lyophilization. **HRMS (*m/z*, $C_{81}H_{106}N_{19}O_{19}S$, ESI):** calculated $[M+H]^+$: 1680.7633; found: 1680.7703.

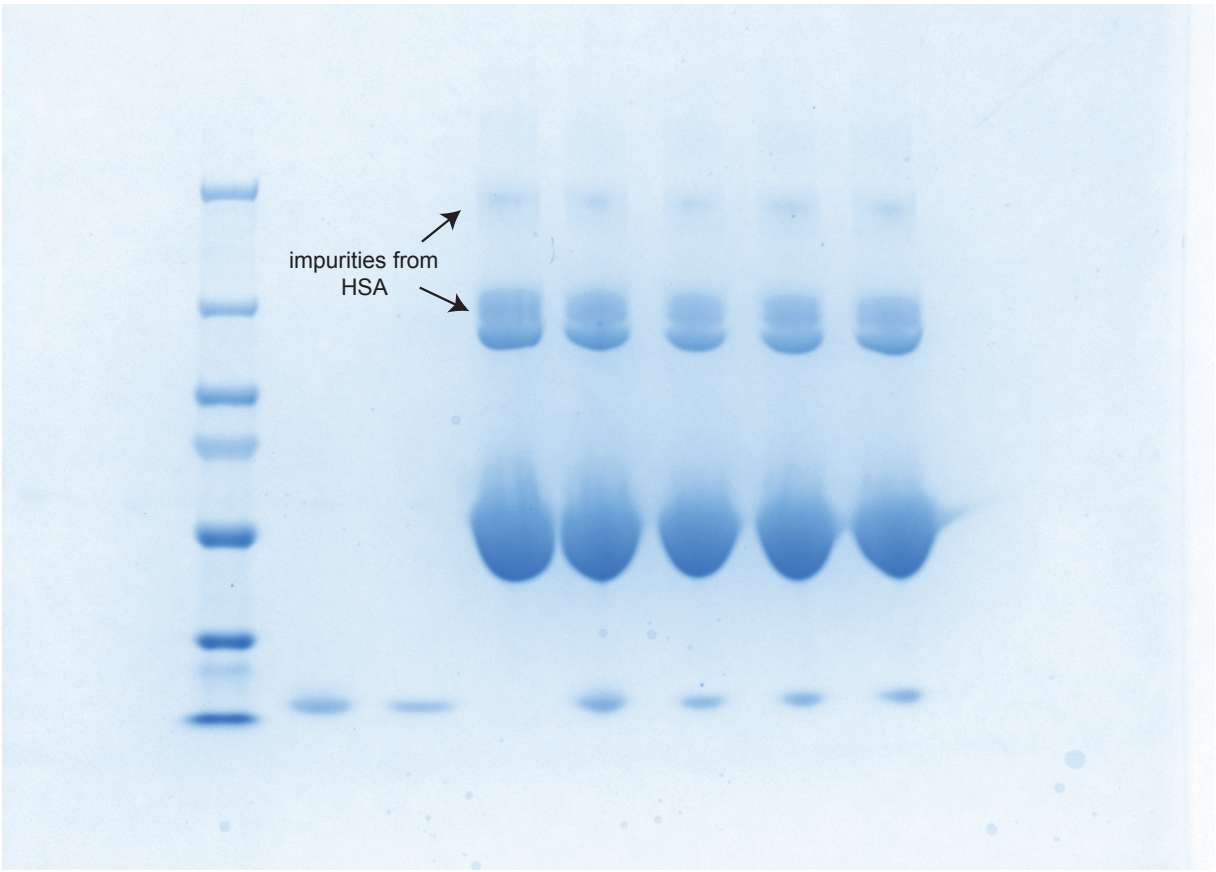
(b) CaM labelling experiments.

In a 96-well plate, HSA (15 μM) spiked with CaM (3 μM) in DPBS (with 10 % DMSO) was incubated with the probe (**CaM-PC** or **SC-PC**) at 10 μM for 10 min at 0 °C (on ice) with gentle shaking at 300 μL final volume[11]. The samples were irradiated under 365 nm at 0 °C (on ice) for 15 min. Protein samples were purified by Vivaspin 500 centrifugal concentrators (Sigma, Catalog: Z614025) and mixed with 5 × SDS loading buffer, boiled at 95 °C for 5 min, and separated by SDS-PAGE (12%). The gel was washed with methanol in water (50 %) for 30 min and visualized for in-gel fluorescence with a Bio-Rad Chemidoc image system. The gel was stained with Coomassie Brilliant Blue for 10 min and destained before visualizing with a Bio-Rad Chemidoc image system.



marker	1	2	3	4	5	6	7
CaM	+	+	−	+	+	+	+
HSA	−	−	+	+	+	+	+
CaM-PC	+	+	+	+	+	−	+
SC-PC	−	−	−	−	−	+	−
hv	+	−	+	+	−	+	+

Supplementary Figure 76 | Full gel imaging with FITC detection. left lane, marker; lane 1, CaM (3 μM)/**CaM-PC** (10 μM), *hv*; lane 2, same as lane 1, no *hv*; lane 3, HSA (15 μM)/**CaM-PC** (10 μM), *hv*; lane 4, CaM (3 μM)/HSA (15 μM)/**CaM-PC** (10 μM), *hv*; lane 5: same as lane 4, no *hv*; lane 6, same as lane 4 with scaffold control **SC-PC** instead of **CaM-PC**; lane 7, same as lane 4, with addition of soluble competitor **CaM-3-amino** (100 μM).



marker	1	2	3	4	5	6	7
CaM	+	+	−	+	+	+	+
HSA	−	−	+	+	+	+	+
CaM-PC	+	+	+	+	+	−	+
SC-PC	−	−	−	−	−	+	−
hv	+	−	+	+	−	+	+

Supplementary Figure 77 | Full gel imaging with Coomassie Brilliant Blue detection. left lane, marker; lane 1, CaM (3 μM)/**CaM-PC** (10 μM), *hv*; lane 2, same as lane 1, no *hv*; lane 3, HSA (15 μM)/**CaM-PC** (10 μM), *hv*; lane 4, CaM (3 μM)/HSA (15 μM)/**CaM-PC** (10 μM), *hv*; lane 5: same as lane 4, no *hv*; lane 6, same as lane 4 with scaffold control **SC-PC** instead of **CaM-PC**; lane 7, same as lane 4, with addition of soluble competitor **CaM-3-amino** (100 μM).

(c) Mass spectrometry analysis of covalent adduct of CaM-probe to CaM.

In a 96-well plate, CaM (10 μM) in DPBS (with 10 % DMSO) was incubated with the probe (**CaM-PC**) at 10 μM for 10 min at 0 °C (on ice) with gentle shaking at 300 μL final volume. The samples were irradiated under 365 nm at 0 °C (on ice) for 15 min. Protein samples were purified by Vivaspin 500 centrifugal concentrators and dissolved by 100 μL PBS. 10 μL of the dissolved samples were injected into a Xevo G2-XS Q-TOF with electrospray ionization source. In detail, data were acquired for 0.5 sec scan time in continuum mode over *m/z* range from 500 to 4000 Da (desolvation gas temperature = 500 °C,

desolvation gas flow rate = 1000 L/h, capillary voltage = 3 kV, sampling cone voltage = 25 V). MaxEnt 1 software was used to deconvolute the multiple charge states.

9. References.

1. Galibert, M., Renaudet, O., Dumy, P. & Boturyn, D. Access to Biomolecular Assemblies through One-Pot Triple Orthogonal Chemoselective Ligations. *Angew. Chem. Int. Ed.* **50**, 1901-1904 (2011).
2. Li, Y. *et al.* Optimized Reaction Conditions for Amide Bond Formation in DNA-Encoded Combinatorial Libraries. *ACS. Comb. Sci.* **18**, 438-443 (2016).
3. Li, Y., Zhao, P., Zhang, M., Zhao, X. & Li, X. Multistep DNA-templated synthesis using a universal template. *J. Am. Chem. Soc.* **135**, 17727-17730 (2013).
4. Franzini, R. M. *et al.* Systematic evaluation and optimization of modification reactions of oligonucleotides with amines and carboxylic acids for the synthesis of DNA-encoded chemical libraries. *Bioconjug. Chem.* **25**, 1453-1461 (2014).
5. Clark, M. A. *et al.* Design, synthesis and selection of DNA-encoded small-molecule libraries. *Nat. Chem. Biol.* **5**, 647-654 (2009).
6. Franzini, R. M. *et al.* Identification of structure-activity relationships from screening a structurally compact DNA-encoded chemical library. *Angew. Chem. Int. Ed.* **54**, 3927-3931 (2015).
7. Halin, C. *et al.* Synergistic Therapeutic Effects of a Tumor Targeting Antibody Fragment, Fused to Interleukin 12 and to Tumor Necrosis Factor α. *Cancer Res.* **63**, 3202-3210 (2003).
8. Li, Y., Zimmermann, G., Scheuermann, J. & Neri, D. Quantitative PCR is a Valuable Tool to Monitor the Performance of DNA-Encoded Chemical Library Selections. *ChemBioChem* **18**, 848-852 (2017).
9. Decurtins, W. *et al.* Automated screening for small organic ligands using DNA-encoded chemical libraries. *Nat. Protoc.* **11**, 764-780 (2016).
10. Krall, N. *et al.* A Small-Molecule Drug Conjugate for the Treatment of Carbonic Anhydrase IX Expressing Tumors. *Angew. Chem. Int. Ed.* **53**, 4231-4235 (2014).
11. Zhao, P. *et al.* Selection of DNA-encoded small molecule libraries against unmodified and non-immobilized protein targets. *Angew. Chem. Int. Ed.* **53**, 10056-10059 (2014).