Electronic Supplementary Material (ESI) for Chemical Science.

Supplementary Information

Table of Contents

1 List of abbreviations

AC: Average counts

ACN: Acetonitrile

c_{DEL}= DNA-encoded library concentration

DEL = DNA-encoded library

EF: Enrichment factor (counts / AC)

LC/MS: Liquid-Chromatography/Mass-Spectrometry

PBS: Phosphate buffered saline

RP-HPLC: Reversed Phase-High Performance Liquid Chromatography

TC: Total counts

TEAA: Triethanolamine acetate

2 Materials and Methods

2.1 Reagents

Unless otherwise noted, all reagents and solvents were purchased from commercial sources (ABCR, ACROS, Apollo scientific, Bachem, Enamine, Fluorochem, Sigma- Aldrich, and TCI) and used under the manufacturer's instructions. Oligonucleotides were purchased from DNA Technology (Denmark) and IBA (Germany). Water was purified with a Millipore Milli-Q system (Merck). Ligation buffer, DNA-Ligase and high-fidelity Phusion DNA polymerase were purchased from New England Biolabs. PCR purification and gel extraction kits were purchased from Qiagen.

2.2 Instruments and Gradients

Oligonucleotide derivatives were purified by semi-preparative **HT-Reversed Phase-High Performance Liquid Chromatography** (HT-RP-HPLC) on Waters XTerra® Shield RP18 (125 Å, 5 µm) column with 0.1 M TEAA pH = 7 in mQ millipore water (buffer A) and 0.1 M TEAA pH = 7 in mQ millipore water : ACN = 2 : 8 (buffer B) as mobile phase. Gradient (% of buffer B): 10% for 1 minute, 10% to 20% in 4 minutes, 20% to 70% in 10 minutes, 100% for 2 minutes.

Liquid-Chromatography/Mass-Spectrometry (LC/MS) characterization of oligonucleotides was performed with an Agilent 1260 Series LC equipped with an ACQUITY UPLC Oligonucleotide BEH C18 column (130 Å, 1.7 µm, 2.1 x 50 mm) coupled to an Agilent 6100 Series Single Quadrupole MS. The following gradient of eluent A (15 mM TEA, 400 mM HFIP in mQ H_2O) and eluent B (methanol) was applied at a flow rate of 0.4 mL/min and 60 °C column temperature: 0-0.2 min 95% A, 0.2-8.2 min 95% to 5% A, 8.2-8.7 min 5% A, 8.7-9.2 min 5 to 95% A, 9.2-13 min 95% A.

Nanodrop 2000 Spectrophotometer (Thermo Scientific) was used when specified for calculation of compounds concentration.

2.3 Software

Databases are managed by InstantJChem (ChemAxon). Selection Fingerprints are evaluated by MATLAB R2019b (mathworks).

3 Library synthesis, serial dilutions, characterization and purification

SO-DEL and NF-DEL library synthesis and HIT validation (HIT compound reported in **Table 1**) are reported in the literature.[1,2]

3.1 Calculation of library copies and dilutions

Copies per library member (input per selection experiment) were calculated as follows:

$$Input\ (per\ selection) = \frac{c_{DEL}\left(\frac{mol}{L}\right) x\ volume(L)\ x\ 6.022x10^{23}\left(\frac{molecules}{mol}\right)}{library\ size\ (molecules)}$$

Where:

Library concentrations were determined by using Beer-Lambert equation ($c = A_{260} \times \varepsilon^{-1} \times l^{-1}$, where $\varepsilon^{-1} \times l^{-1}$ = xxxx μM). DNA absorbance (A_{260}) was measured using Nanodrop instrument (**Section 2.2**) at λ=260 nm.

The library size of SO-DEL and SO-DEL-long is equal to 933,984 compounds while for NF-DEL library size is equal to 670,752 compounds.

Each DEL was properly diluted and aliquoted in order to have 10^7 copies as input in 10 μL as initial concentration for further serial dilutions.

Starting from 10^7 copies per library member per selection, 1:10 serial dilutions were performed to obtain all the needed inputs (from 10^7 to 10^2). Whenever explicated, selections were performed in duplicates as previously described.[3]

Each affinity selection experiment was performed with each DEL diluted in 100 μL of protein buffer containing 0.05% tween-20 and 20 μg/mL herring sperm DNA. Protein targets were immobilized on magnetic C1 streptavidin coated beads and selections were automated by King Fisher (Thermo Fisher) as previously reported[3] (more details on selection procedure in **Section 5**).

3.2 Encoding of SO-DEL-long

Code A: 5′-CTGTGTGCTG**XXXXXX**CGAGTCCCATGGCGC-3′

Code B: 5′-CGGATCGACG**YYYYYYY**GCGTCAGGCAGC-3′

Primer (L2C2): 5′-CGGATCGACGGTCTCACCACGGATCCATTCGATGCAGG-3′

Adapter L2C2: 5′-GTCGATCCGGCTGCCTGA-3′

Ligation site:

5′- …CTGTGTGCTG**XXXXXX**CGAGTCCCATGGCGCCGGATCGACG**YYYYYYY**GCGTCAGGCAGCCGGATCGACGGTCTCACCACGGATCCATTCGATGCAGG-3′
Adapter L2C2 3′AGTCCGTCGGCCTAGCTG

Final SO-DEL-long design:

5′- GGAGCTTCTGAATTCTGTGTGCTG**XXXXXX**CGAGTCCCATGGCGCCGGATCGACG**YYYYYYY**GCGTCAGGCAGCCGGATCGACGGTCTCACCACGGATCCATTCGATGCAGG-3′

New primer PCR1-b: 5′- CAGACGTGTGCTCTTCCGATCCGATATAGCCCTGCATCGAATGGATCCGTG-3′

3.3 SO-DEL ligation with L2C2

To 5 nmol of SO-DEL (96 µM), 50 µL of primer L2C2 (150 µM, 7.5 nmol, 1.5 eq.), 23 µL adaptor L2C2 (330 µM, 7.5 nmol, 1.5 eq.) and 25 µL 10x T4 DNA Ligase Reaction Buffer (New England Biolabs, #B0202S) were added. The mixtures were heated for 15 minutes at 75 °C and cooled at room temperature for 20 minutes to allow oligonucleotides annealing. T4 DNA ligase (New England Biolabs, #M0202L, 400,000 units/mL) was 8-fold diluted with 1x T4 DNA Ligase Reaction Buffer and 8 µL added to each ligation reaction (200 units). The ligation proceeded for 2 hours at room temperature and was subsequently quenched by heating for 10 min at 75 °C. The reaction was monitored via TBE Urea 15% gel (**see Figure S2**) then EtOH precipitated.[1,2] The precipitate was redissolved in 0.1 M TEAA and purified via RP-HPLC (performed in DNA denaturing conditions, column temperature = 60°C).[1,2]

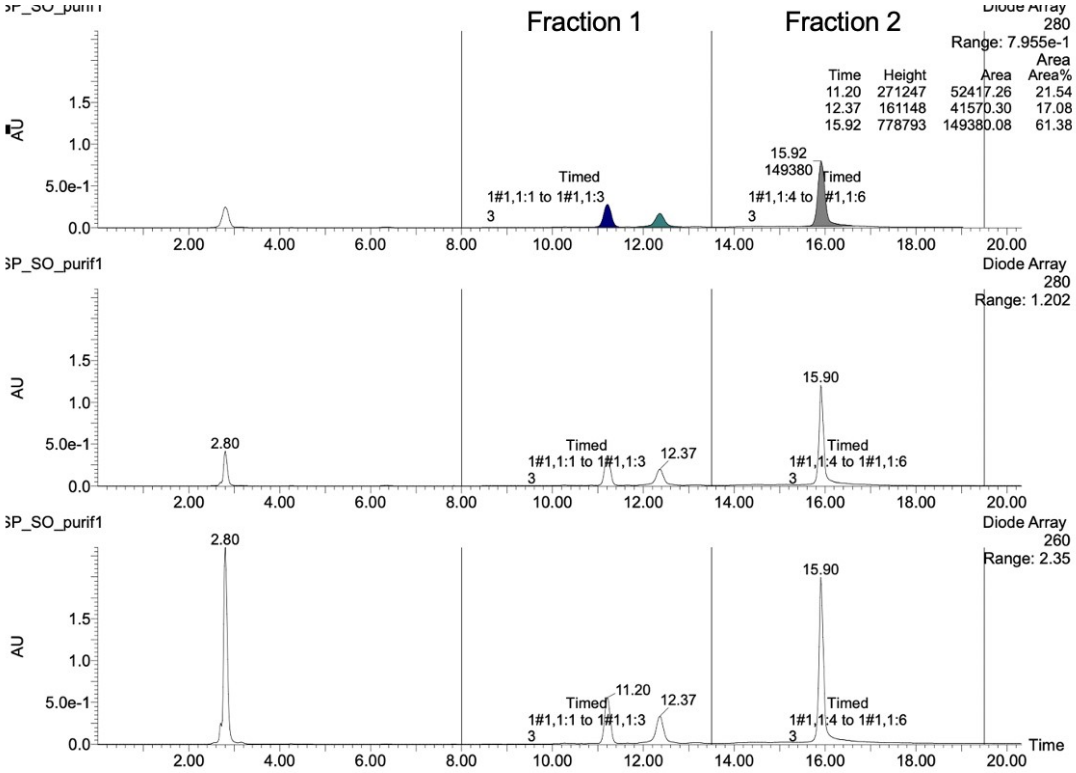

Figure S1: RP-HPLC purification of SO-DEL-long: RP-HPLC chromatogram of the elongated library. Fraction 1: non reacted Code L2C2 + Adapter L2C2; Fraction 2: SO-DEL-long.

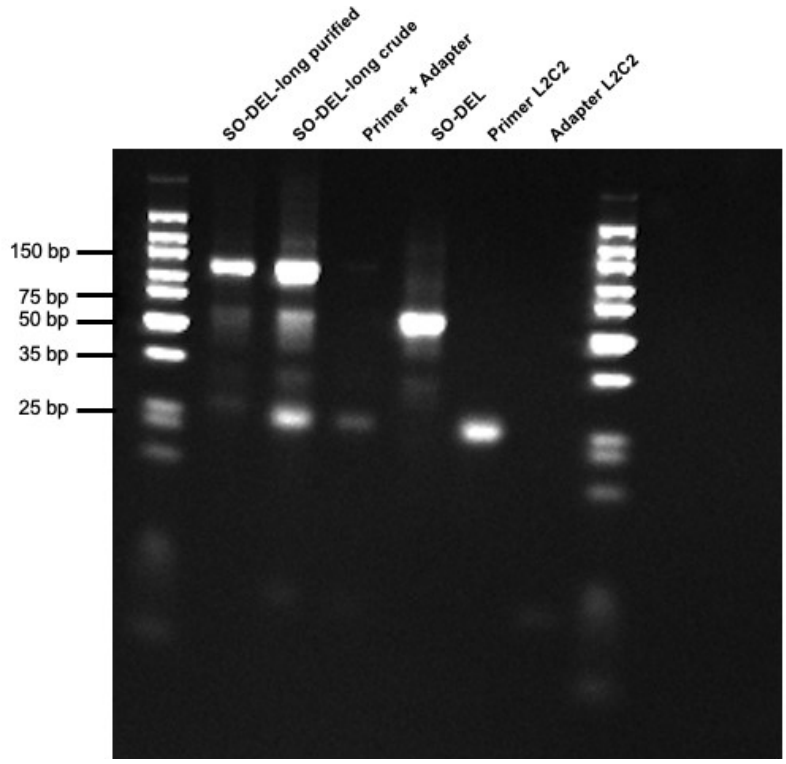

Figure S2: Characterization of SO-DEL-long: TBE Urea 15% gel of fractions obtained after HPLC purification of SO-DEL-long compared to crude after elongation, 1:1 mix Primer L2C2 and Adapter L2C2, SO-DEL, Primer L2C2 and Adapter L2C2.

4 Proteins for Affinity Selections

Cloning, expression and biochemical characterization of human CAIX and NSP14 are reported in the literature.[1,2] Human Serum Albumin was purchased as lyophilized powder from Sigma Aldrich (#A9511).
Biotinylation protocols, band shift assays and MS spectra before and after biotinylation (when available) are reported in the literature.[1,2] Protein specifications (buffer, MW, epsilon (280 nm, $M^{-1}cm^{-1}$) are reported in the **Table S1**.

4.1 List of proteins

Target	***Buffer***	***MW (Da)***	***ε (λ=280 nm) (M^{-1} cm^{-1})***	***Biotinylation***
CAIX	PBS, pH = 7.4	32732	34850	Chemical
NSP14	50 mM Tris-HCl, 200 mM NaCl, 5% (w/v) glycerol, 0.5 mM TCEP pH 8	61800 (monomer)	97750	Avi-tag
HSA	PBS, pH = 7.4	66561	3445	Chemical

Table S1: List and properties of proteins used for affinity selection. Abbreviations: CAIX: Carbonic Anhydrases IX; NSP14: Non-Structural Protein 14; HSA: Human Serum Albumin.

5 Selections

The selection protocol is reported in the literature.[3] Some changes were applied in order to optimize the protocol. They are described in detail in the paragraphs below.

The selection eluates are amplified by two rounds of PCR as previously reported[3] using the following primers:

PCR1-a: 5′–TACACGACGCTCTTCCGATCT***xxxxxx***GGAGCTTCTGAATTCTGTGTG–3′, where X represent a variable region which codify for the selection experiment.

PCR1-b: 5′–CAGACGTGTGCTCTTCCGATC***xxxxxx***CCTGCATCGAATGGATCCG–3'

PCR1-b (for SO-DEL-long): 5′–CAGACGTGTGCTCTTCCGATCCGATATAGCCCTGCATCGAATGGATCCGTG–3' (only this backward primer was used for SO-DEL-long selection campaign).

PCR2-a: 5′–AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT–3'

PCR2-b: 5′–CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC–3'

The PCR products were sequenced by Illumina high-throughput sequencing and the data obtained was processed and analyzed as previously reported.[3]

5.1 Amplifiability after PCR1: new design vs old design

The protocol for the first amplification step (PCR1) was optimized, implementing the standard protocol[1,2,3] with a lower input of PCR1 primers (15 pmol per primer instead of 30 pmol per primer) and using the following PCR program:

Step	*Temperature (°C)*	*Time*	←	#
#1	98	3 m 0 s		
#2	98	0 m 45 s		
#3	69	0 m 45 s		
#4	72	0 m 45 s	2	29
#5	72	5 m 0 s		
#6	4	Pause		

Table S2: Scheme of PCR1 cycle. Time is measured in minutes (m) and seconds (s). ← = going back to cycle number X, # = number or cycles.

The optimized PCR protocol allowed productive library amplification with the different inputs of DNA copies (see **Section 3.1**).

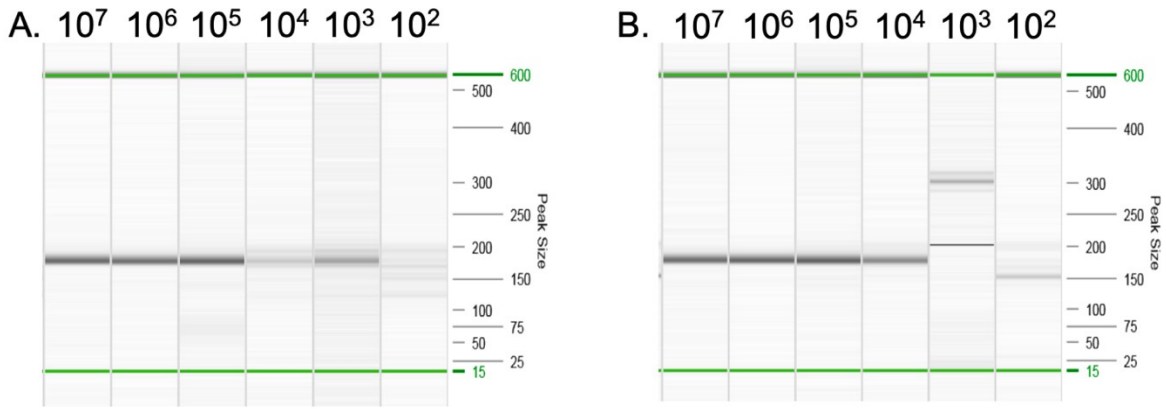

Figure S3: SO-DEL-long amplifiability using the optimized PCR1 protocol. A. Target: CAIX B. Target: NSP14.

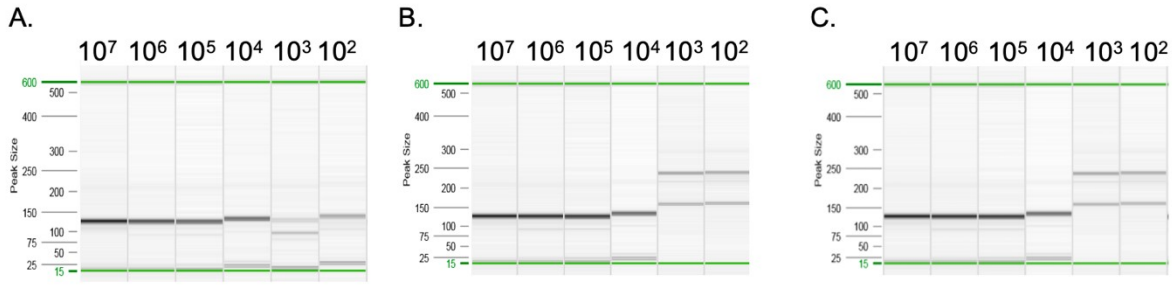

Figure S4: SO-DEL amplifiability using the optimized PCR1 protocol. A. Target: NSP14 B. Target: CAIX. C. HSA.

5.2 Selection fingerprints (duplicates and unselected SO-DEL-long)

HTS results were processed as previously reported[3] and the results are displayed in a three-dimensional fingerprint using MathWorks Matlab software. SO-DEL and NF-DEL were screened in duplicates (**Figure 2**, **Figure 3** and **Figure S5-S9**).
SO-DEL-long selections were performed in single experiments reported in the main manuscript (**Figure 4**).

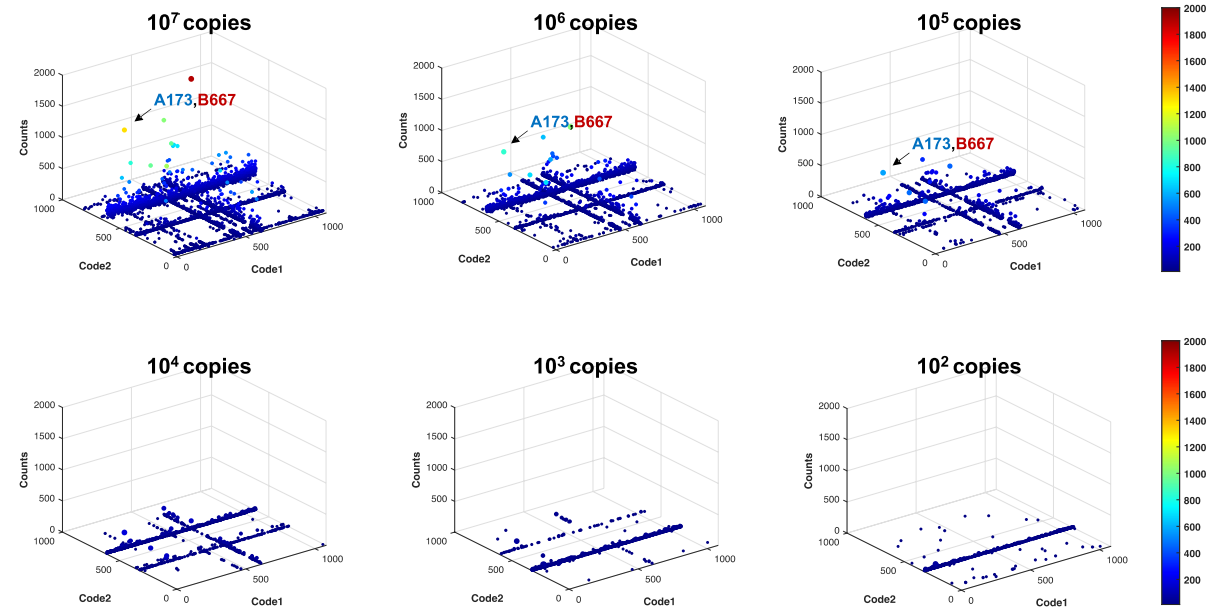

Figure S5: SO-DEL duplicate for CAIX. Combination A173/B667 was followed in the selections performed with different inputs. Command used for SO-DEL and SO-DEL-long sections against CAIX: include 1-1080; 1-856. Cut-off: 30.

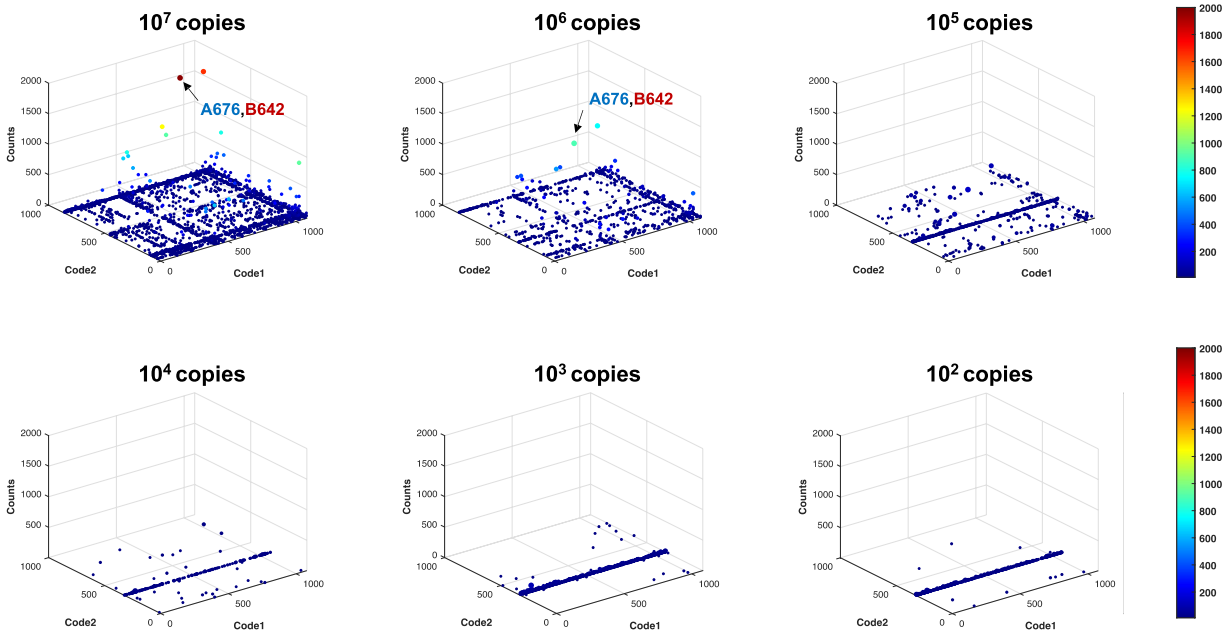

Figure S6: SO-DEL duplicate for HSA. Combination A676/B642 was followed in the selections performed with different inputs. Cut-off: 30.

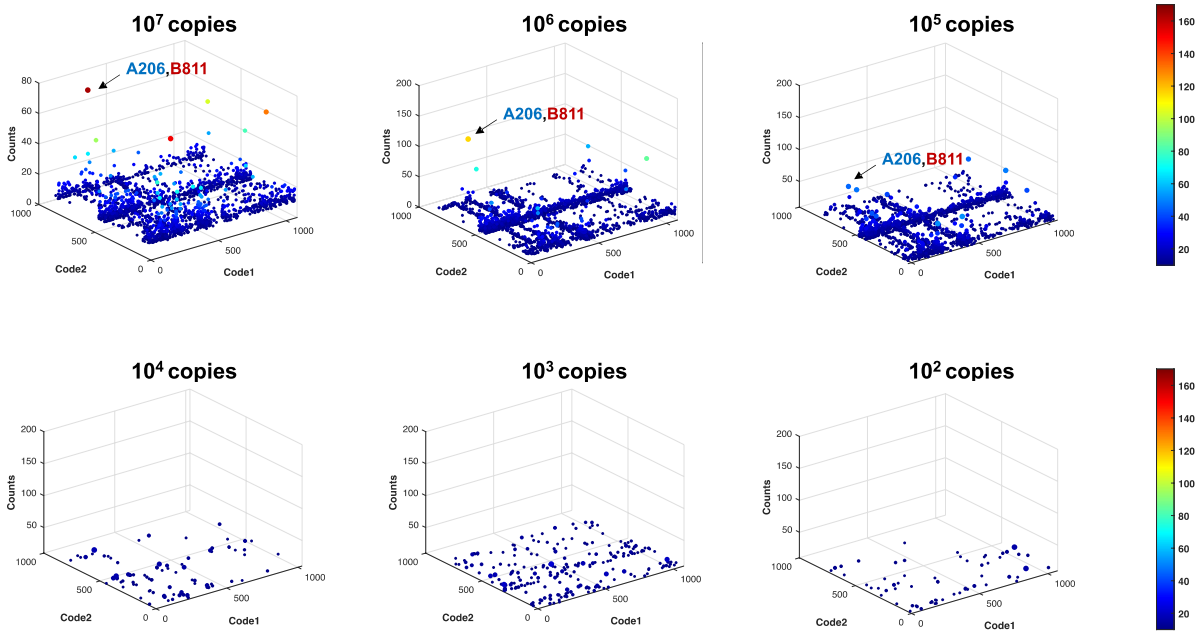

Figure S7: SO-DEL duplicate for NSP14. Combination A206/B811 was followed in the selections performed with different inputs. Cut-off: 30.

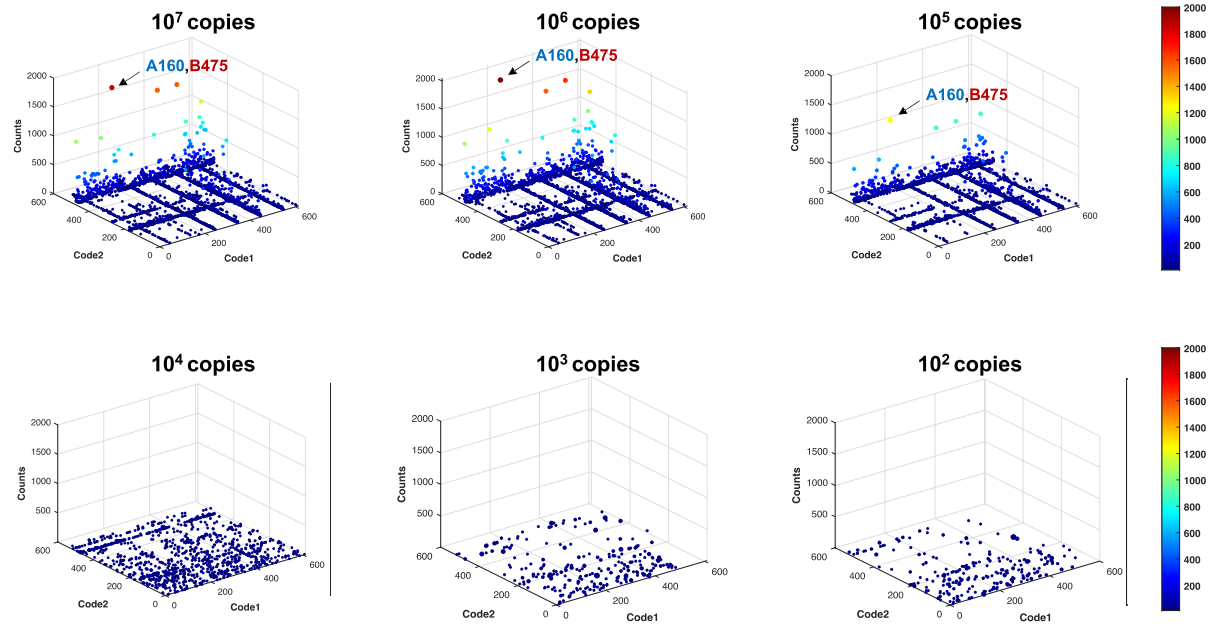

Figure S8: NF-DEL duplicate for CAIX. Combination A160/B475 was followed in the selections performed with different inputs. Cut-off: 30.

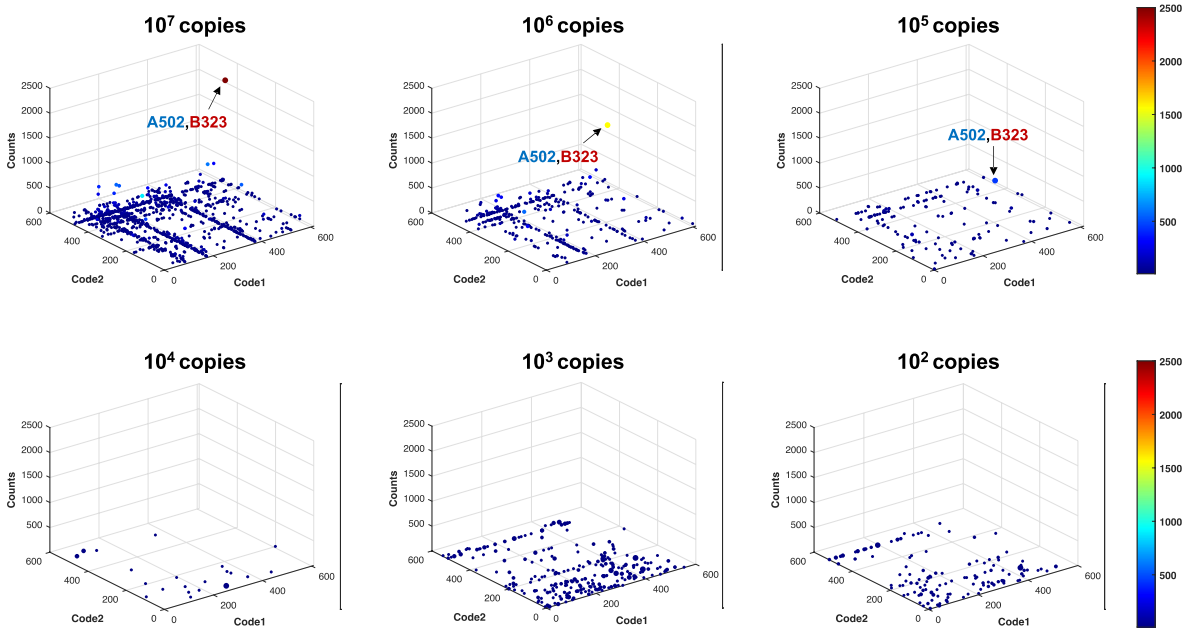

Figure S9: NF-DEL duplicate for HSA. Combination A502/B323 was followed in the selections performed with different inputs. Cut-off: 30.

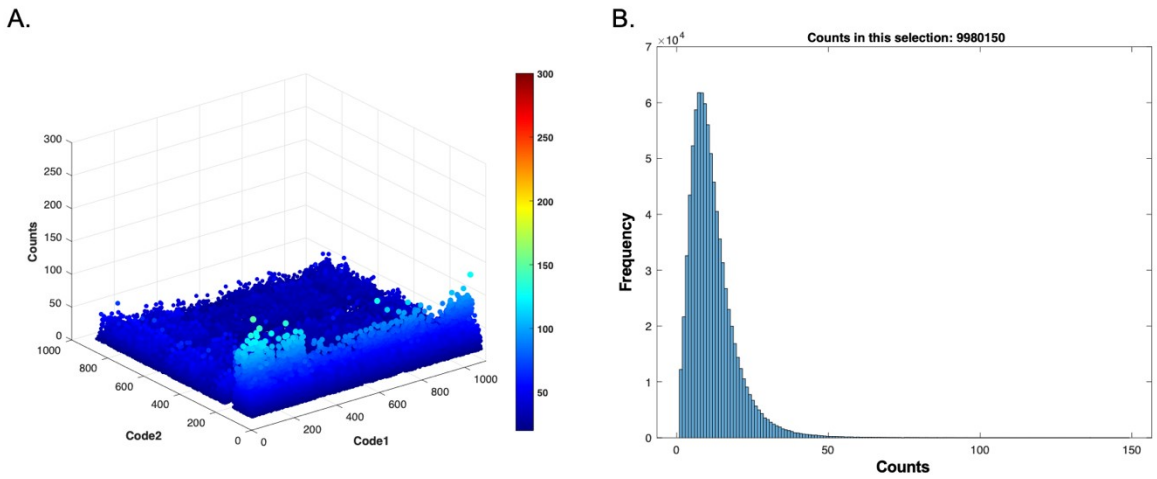

Figure S10: SO-DEL-long: A. unselected (naïve) library. Each library member is represented by a dot for which the identity is given by the coordinate (x/y = codeA/codeB). Dot color and size corresponds to the normalized sequence counts as depicted by the jet color bar. B. Count distribution of the unselected library. Total counts = 9,980,150, average count = 10.7, cut-off applied = 20.

5.2.1 Enrichment factors

The enrichment factors (EF) for each building block combinations were calculated using the following equations:

$$(1)\ AC(sel) = \frac{TC\ (sel)}{library\ size} = \frac{TC\ (sel)}{933{,}984(SO)\ or\ 335{,}376\ (NF)}$$

$$(2) EF\left(BB_A;BB_B\right) = \frac{counts\left(BB_A;BB_B\right)}{AC(sel)}$$

Where AC and TC are the average number of counts and the total number of counts for a specific selection, respectively. All TC and calculated AC values are reported in the **Table S3**.
EF for the enriched combinations is reported in the **Table S4**.

Sel ID	*Antigen*	*Library*	*DEL-Size*	*DEL-Input*	*TC (sel)*	*AC (sel)*
#1	CAIX	SO-DEL	933,984	10^7	711,625	0.76
#2	CAIX	SO-DEL	933,984	10^6	533,480	0.57
#3	CAIX	SO-DEL	933,984	10^5	321,772	0.34
#4	CAIX	SO-DEL	933,984	10^4	133,422	0.14
#5	CAIX	SO-DEL	933,984	10^3	104,114	0.11
#6	CAIX	SO-DEL	933,984	10^2	243,023	0.26
#7	CAIX	SO-DEL	933,984	10^7	779,010	0.83
#8	CAIX	SO-DEL	933,984	10^6	496,096	0.53
#9	CAIX	SO-DEL	933,984	10^5	315,158	0.34
#10	CAIX	SO-DEL	933,984	10^4	143,720	0.15
#11	CAIX	SO-DEL	933,984	10^3	100,444	0.11
#12	CAIX	SO-DEL	933,984	10^2	196,675	0.21
#13	HSA	SO-DEL	933,984	10^7	532,959	0.57
#14	HSA	SO-DEL	933,984	10^6	319,840	0.34
#15	HSA	SO-DEL	933,984	10^5	168,505	0.18
#16	HSA	SO-DEL	933,984	10^4	75,927	0.08
#17	HSA	SO-DEL	933,984	10^3	70,564	0.08
#18	HSA	SO-DEL	933,984	10^2	188,624	0.20
#19	HSA	SO-DEL	933,984	10^7	556,812	0.60
#20	HSA	SO-DEL	933,984	10^6	304,482	0.33
#21	HSA	SO-DEL	933,984	10^5	223,417	0.24
#22	HSA	SO-DEL	933,984	10^4	20,878	0.02
#23	HSA	SO-DEL	933,984	10^3	73,309	0.08
#24	HSA	SO-DEL	933,984	10^2	129,790	0.14
#25	NSP14	SO-DEL	933,984	10^7	670,938	0.72
#26	NSP14	SO-DEL	933,984	10^6	670,255	0.72
#27	NSP14	SO-DEL	933,984	10^5	408,279	0.44
#28	NSP14	SO-DEL	933,984	10^4	281,450	0.30
#29	NSP14	SO-DEL	933,984	10^3	239,258	0.26
#30	NSP14	SO-DEL	933,984	10^2	206,254	0.22
#31	NSP14	SO-DEL	933,984	10^7	591,899	0.63

#32	NSP14	SO-DEL	933,984	10^6	599,424	0.64
#33	NSP14	SO-DEL	933,984	10^5	497,356	0.53
#34	NSP14	SO-DEL	933,984	10^4	179,706	0.19
#35	NSP14	SO-DEL	933,984	10^3	138,321	0.15
#36	NSP14	SO-DEL	933,984	10^2	174,315	0.19
#37	CAIX	NF-DEL	335,376	10^7	287,183	0.85
#38	CAIX	NF-DEL	335,376	10^6	299,259	0.90
#39	CAIX	NF-DEL	335,376	10^5	200,525	0.60
#40	CAIX	NF-DEL	335,376	10^4	88,806	0.26
#41	CAIX	NF-DEL	335,376	10^3	41,336	0.12
#42	CAIX	NF-DEL	335,376	10^2	40,268	012
#43	CAIX	NF-DEL	335,376	10^7	375,098	1.12
#44	CAIX	NF-DEL	335,376	10^6	301,906	0.90
#45	CAIX	NF-DEL	335,376	10^5	197,443	0.58
#46	CAIX	NF-DEL	335,376	10^4	89,138	0.26
#47	CAIX	NF-DEL	335,376	10^3	39,624	0.12
#48	CAIX	NF-DEL	335,376	10^2	38,276	0.12
#49	HSA	NF-DEL	335,376	10^7	346,441	1.04
#50	HSA	NF-DEL	335,376	10^6	186,562	0.56
#51	HSA	NF-DEL	335,376	10^5	70,007	0.20
#52	HSA	NF-DEL	335,376	10^4	15,236	0.04
#53	HSA	NF-DEL	335,376	10^3	40,211	0.12
#54	HSA	NF-DEL	335,376	10^2	33,113	0.10
#55	HSA	NF-DEL	335,376	10^7	342,768	1.02
#56	HSA	NF-DEL	335,376	10^6	197,053	0.58
#57	HSA	NF-DEL	335,376	10^5	88,381	0.26
#58	HSA	NF-DEL	335,376	10^4	54,103	0.16
#59	HSA	NF-DEL	335,376	10^3	37,303	0.12
#60	HSA	NF-DEL	335,376	10^2	34,271	0.10
#61	CAIX	SO-DEL-long	933,984	10^7	2,214,514	2.37
#62	CAIX	SO-DEL-long	933,984	10^6	1,967,700	2.11
#63	CAIX	SO-DEL-long	933,984	10^5	1,853,613	1.98
#64	CAIX	SO-DEL-long	933,984	10^4	618,659	0.66
#65	CAIX	SO-DEL-long	933,984	10^3	198,662	0.21
#66	CAIX	SO-DEL-long	933,984	10^2	148,835	0.16
#67	NSP14	SO-DEL-long	933,984	10^7	1,662,888	1.78
#68	NSP14	SO-DEL-long	933,984	10^6	1,926,131	2.06
#69	NSP14	SO-DEL-long	933,984	10^5	1,797,496	1.92
#70	NSP14	SO-DEL-long	933,984	10^4	1,635,530	1.75
#71	NSP14	SO-DEL-long	933,984	10^3	979,571	1.05
#72	NSP14	SO-DEL-long	933,984	10^2	337,933	0.36
#73	Naïve	SO-DEL-long	933,984		9,980,150	10.69

Table S3: total counts and calculated average counts for all selection experiments.

Sel ID	Antigen	DEL format	DEL- Input	HIT (Axx, Byy)	counts	EF
#1	CAIX	SO-DEL	10^7	A173,B667	1256	1650
#2	CAIX	SO-DEL	10^6	A173,B667	968	1697
#3	CAIX	SO-DEL	10^5	A173,B667	590	1715
#4	CAIX	SO-DEL	10^4	A173,B667	140	981
#5	CAIX	SO-DEL	10^3	A173,B667	15	135
#6	CAIX	SO-DEL	10^2	A173,B667	0	0
#7	CAIX	SO-DEL	10^7	A173,B667	1312	1575
#8	CAIX	SO-DEL	10^6	A173,B667	848	1598
#9	CAIX	SO-DEL	10^5	A173,B667	577	1712
#10	CAIX	SO-DEL	10^4	A173,B667	186	1210
#11	CAIX	SO-DEL	10^3	A173,B667	57	531
#12	CAIX	SO-DEL	10^2	A173,B667	0	0
#13	HSA	SO-DEL	10^7	A676,B642	1807	3170
#14	HSA	SO-DEL	10^6	A676,B642	851	2488
#15	HSA	SO-DEL	10^5	A676,B642	179	993
#16	HSA	SO-DEL	10^4	A676,B642	53	653
#17	HSA	SO-DEL	10^3	A676,B642	11	146
#18	HSA	SO-DEL	10^2	A676,B642	0	0
#19	HSA	SO-DEL	10^7	A676,B642	1975	3317
#20	HSA	SO-DEL	10^6	A676,B642	902	2770
#21	HSA	SO-DEL	10^5	A676,B642	144	603
#22	HSA	SO-DEL	10^4	A676,B642	5	224
#23	HSA	SO-DEL	10^3	A676,B642	7	89
#24	HSA	SO-DEL	10^2	A676,B642	3	22
#25	NSP14	SO-DEL	10^7	A206,B811	169	236
#26	NSP14	SO-DEL	10^6	A206,B811	107	149
#27	NSP14	SO-DEL	10^5	A206,B811	76	174
#28	NSP14	SO-DEL	10^4	A206,B811	13	43
#29	NSP14	SO-DEL	10^3	A206,B811	0	0
#30	NSP14	SO-DEL	10^2	A206,B811	0	0
#31	NSP14	SO-DEL	10^7	A206,B811	77	122
#32	NSP14	SO-DEL	10^6	A206,B811	116	181
#33	NSP14	SO-DEL	10^5	A206,B811	46	86
#34	NSP14	SO-DEL	10^4	A206,B811	18	94
#35	NSP14	SO-DEL	10^3	A206,B811	0	0
#36	NSP14	SO-DEL	10^2	A206,B811	0	0
#37	CAIX	NF-DEL	10^7	A160,B475	1843	2168
#38	CAIX	NF-DEL	10^6	A160,B475	2024	2249
#39	CAIX	NF-DEL	10^5	A160,B475	1245	2075
#40	CAIX	NF-DEL	10^4	A160,B475	22	85
#41	CAIX	NF-DEL	10^3	A160,B475	5	42
#42	CAIX	NF-DEL	10^2	A160,B475	0	0

#43	CAIX	NF-DEL	10^7	A160,B475	2329	2079
#44	CAIX	NF-DEL	10^6	A160,B475	1980	2200
#45	CAIX	NF-DEL	10^5	A160,B475	1095	1888
#46	CAIX	NF-DEL	10^4	A160,B475	13	50
#47	CAIX	NF-DEL	10^3	A160,B475	12	100
#48	CAIX	NF-DEL	10^2	A160,B475	3	25
#49	HSA	NF-DEL	10^7	A502,B323	2468	2373
#50	HSA	NF-DEL	10^6	A502,B323	1567	2798
#51	HSA	NF-DEL	10^5	A502,B323	456	2280
#52	HSA	NF-DEL	10^4	A502,B323	0	0
#53	HSA	NF-DEL	10^3	A502,B323	3	25
#54	HSA	NF-DEL	10^2	A502,B323	0	0
#55	HSA	NF-DEL	10^7	A502,B323	2314	2269
#56	HSA	NF-DEL	10^6	A502,B323	1067	1840
#57	HSA	NF-DEL	10^5	A502,B323	452	1738
#58	HSA	NF-DEL	10^4	A502,B323	5	31
#59	HSA	NF-DEL	10^3	A502,B323	0	0
#60	HSA	NF-DEL	10^2	A502,B323	4	40
#61	CAIX	SO-DEL-long	10^7	A173,B667	5337	2252
#62	CAIX	SO-DEL-long	10^6	A173,B667	5429	2573
#63	CAIX	SO-DEL-long	10^5	A173,B667	4575	2311
#64	CAIX	SO-DEL-long	10^4	A173,B667	657	995
#65	CAIX	SO-DEL-long	10^3	A173,B667	4	19
#66	CAIX	SO-DEL-long	10^2	A173,B667	19	119
#67	NSP14	SO-DEL-long	10^7	A206,B811	238	134
#68	NSP14	SO-DEL-long	10^6	A206,B811	163	79
#69	NSP14	SO-DEL-long	10^5	A206,B811	104	54
#70	NSP14	SO-DEL-long	10^4	A206,B811	78	45
#71	NSP14	SO-DEL-long	10^3	A206,B811	56	53
#72	NSP14	SO-DEL-long	10^2	A206,B811	1	3

Table S4: Number of counts and calculated enrichment factors (EF) for the building block combination **A122, B198** in all FAP selection experiments and in control experiments (no protein and unselected DELs).

6 Full chemical structures of binders

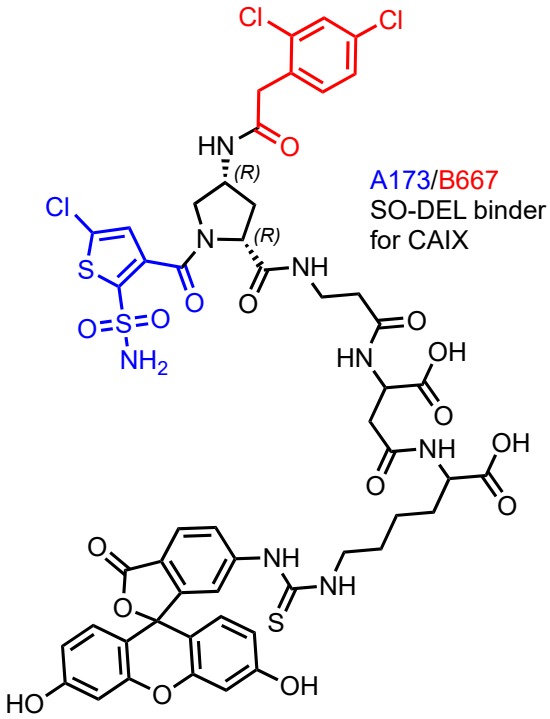

Figure S11: Full chemical structure of SO-DEL A173/B667 binder for CAIX.

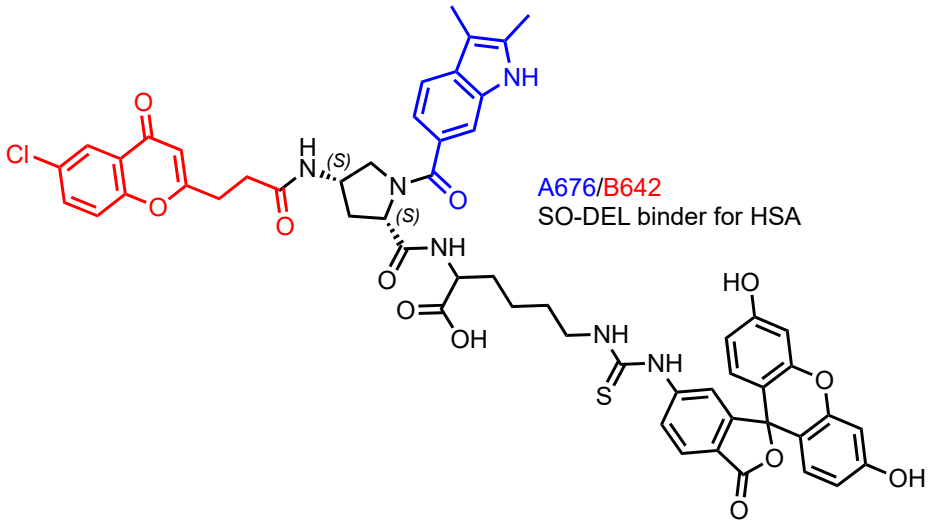

Figure S12: Full chemical structure of SO-DEL A676/B642 binder for HSA.

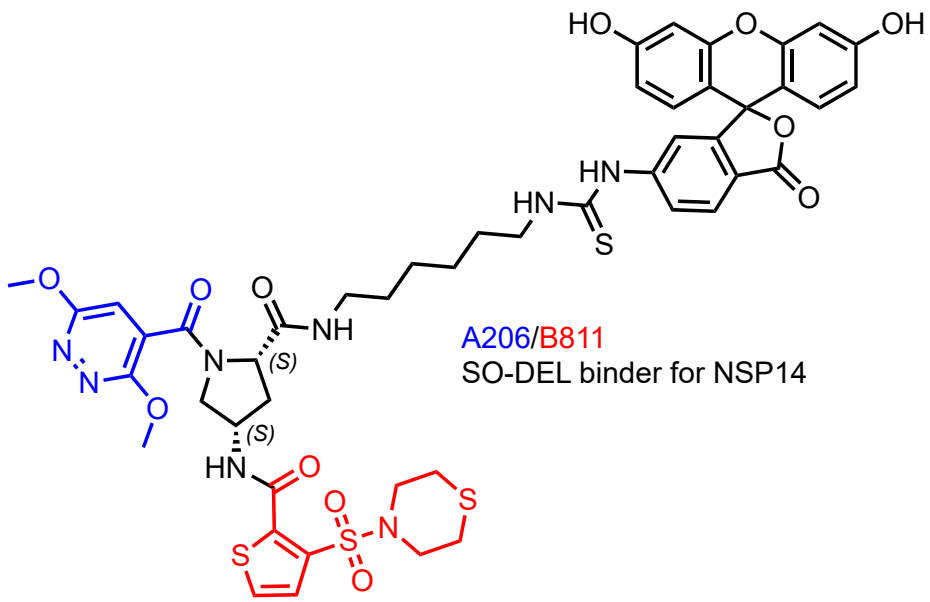

Figure S13: Full chemical structure of SO-DEL A206/B811 binder for NSP14.

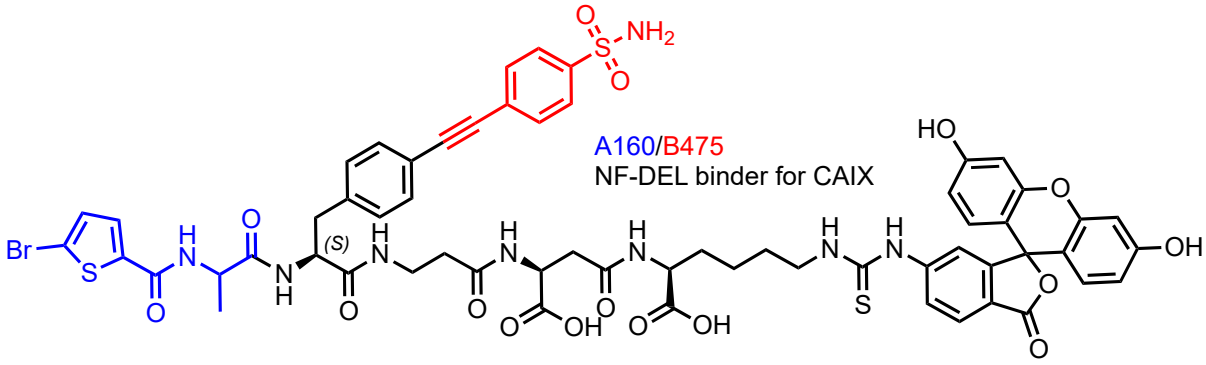

Figure S14: Full chemical structure of NF-DEL A160/B475 binder for CAIX.

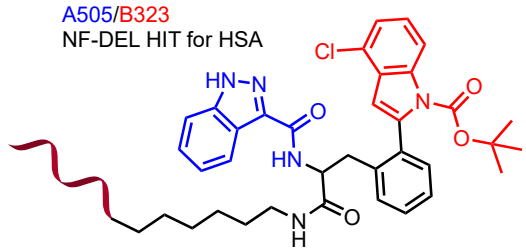

Figure S15: Full chemical structure of NF-DEL A505/B323 HIT for HSA.

7 Impact of library input on hit discovery rate

7.1 SO-DEL

DEL Input	CAIX (1)	CAIX (2)	HSA (1)	HSA (2)	NSP14 (1)	NSP14 (2)
10^7	990	951	233	229	19	12
10^6	993	977	108	100	17	11
10^5	697	733	21	17	6	5
10^4	106	120	1	0	0	3
10^3	0	6	1	4	n.a.*	n.a.*
10^2	2	1	1	0	n.a.*	n.a.*

Table S5: number of counts per selection (above a fixed threshold = 50) per library input for SO-DEL against CAIX, HSA and NSP14. n.a.* values are referring to selections showing a high background noise with low average counts. CAIX = number of counts per selection against CAIX. HSA = number of counts per selection against HSA. NSP14 = number of counts per selection against NSP14. (1) and (2) indicate that the experiment was performed in duplicate.

7.2 NF-DEL

DEL Input	CAIX (1)	CAIX (2)	HSA (1)	HSA (2)
10^7	634	801	162	151
10^6	707	719	61	57
10^5	455	411	7	10
10^4	0	0	0	0
10^3	0	0	0	0
10^2	0	0	0	0

Table S6: number of counts per selection (above a fixed threshold = 50) per library input for NF-DEL against CAIX and HSA. CAIX = number of counts per selection against CAIX. HSA = number of counts per selection against HSA. (1) and (2) indicate that the experiment was performed in duplicate.

7.3 SO-DEL-long

DEL Input	CAIX	NSP14
10^7	1000	614
10^6	999	134
10^5	921	140
10^4	979	47
10^3	517	53
10^2	297	n.a.*

Table S7: number of counts per selection (above a fixed threshold = 50) per library input for SO-DEL-long against CAIX and NSP14. n.a.* values are referring to selections showing a high background noise with low average counts. CAIX = number of counts per selection against CAIX. NSP14 = number of counts per selection against NSP14.

8 Impact of PCR1 on the efficiency of selection recovery

8.1 NF-DEL, SO-DEL and SO-DEL-long: PCR1 serial diluitions

NF-DEL, SO-DEL and SO-DEL-long were serially diluted and efficacy of PCR1 was tested.

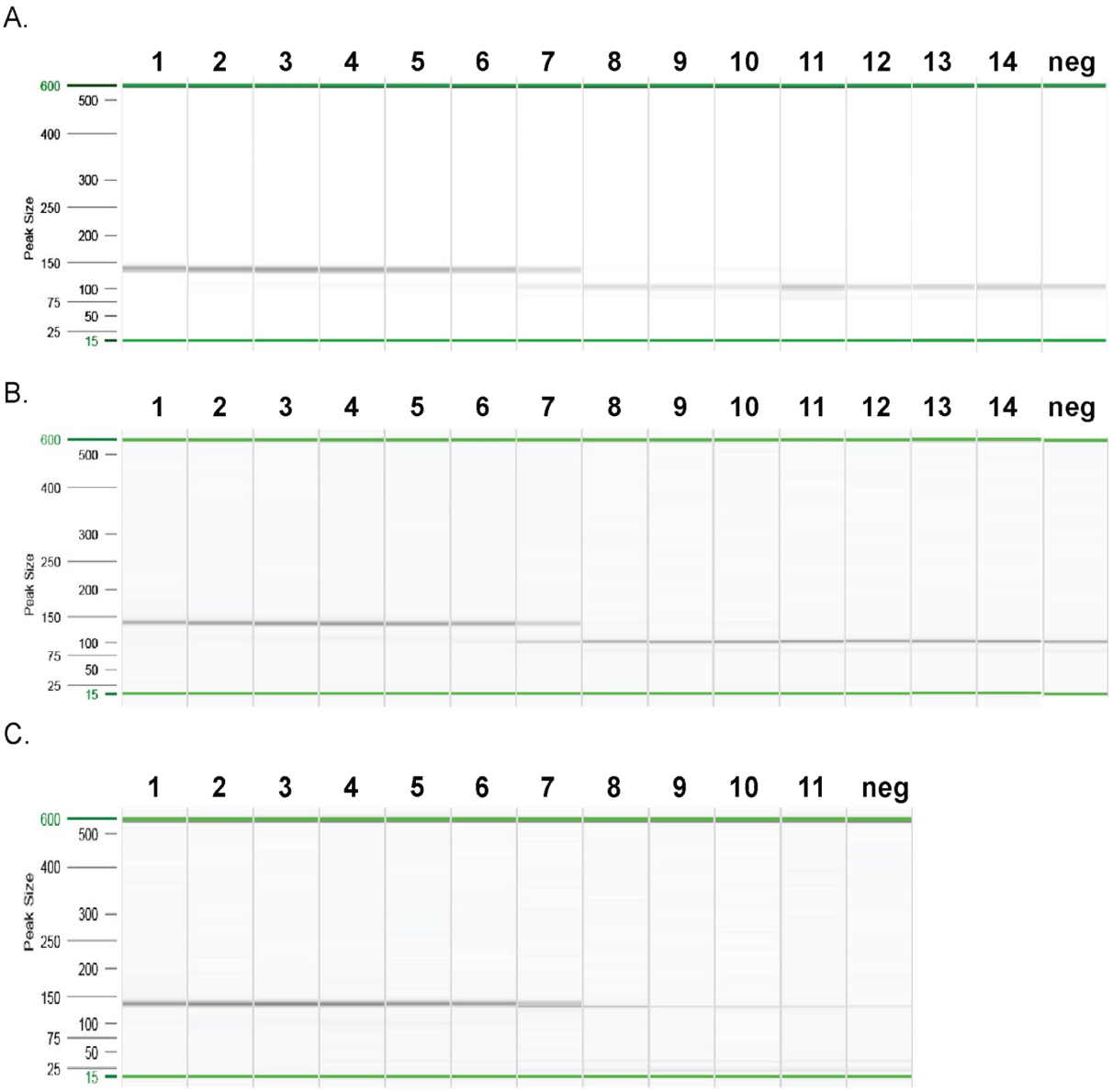

Figure S16: PCR1 performed with different inputs of copies of library (corresponding to different input of number of molecules) with A. NF-DEL naïve; B. SO-DEL naïve and C. SO-DEL-long naïve. Water was added as a negative control (neg). Corresponding copies of library and number of molecules per each gel column are described in **Table S8** and **Table S9**.

reference	copies of library	library size	tot. Number of molecules
1	3.0E+06	335'376	1.0E+12
2	3.0E+05	335'376	1.0E+11
3	3.0E+04	335'376	1.0E+10
4	3.0E+03	335'376	1.0E+09
5	3.0E+02	335'376	1.0E+08
6	3.0E+01	335'376	1.0E+07
7	3.0E+00	335'376	1.0E+06
8	3.0E-01	335'376	1.0E+05
9	3.0E-02	335'376	1.0E+04
10	3.0E-03	335'376	1.0E+03
11	3.0E-04	335'376	1.0E+02
12	3.0E-05	335'376	1.0E+01
13	3.0E-06	335'376	1.0E+00
14	3.0E-07	335'376	1.0E-01
neg	water control		0

Table S8: Copies of library and total number of molecules corresponding to each input of PCR1 gel showed in **Figure S16** (panel A) for NF-DEL (library size 335'376).

reference	copies of library	library size	tot. Number of molecules
1	1.0E+06	933'984	1.0E+12
2	1.0E+05	933'984	1.0E+11
3	1.0E+04	933'984	1.0E+10
4	1.0E+03	933'984	1.0E+09
5	1.0E+02	933'984	1.0E+08
6	1.0E+01	933'984	1.0E+07
7	1.0E+00	933'984	1.0E+06
8	1.0E-01	933'984	1.0E+05
9	1.0E-02	933'984	1.0E+04
10	1.0E-03	933'984	1.0E+03
11	1.0E-04	933'984	1.0E+02
12	1.0E-05	933'984	1.0E+01
13	1.0E-06	933'984	1.0E+00
14	1.0E-07	933'984	1.0E-01
neg	water control		0

Table S9: Copies of library and total number of molecules corresponding to each input of PCR1 gel showed in **Figure S16** (panel B and C) for SO-DEL and SO-DEL-long (library size 933'984).

8.2 PCR1 with 52 cycles: increasing the number of cycles did not change the profile of gels

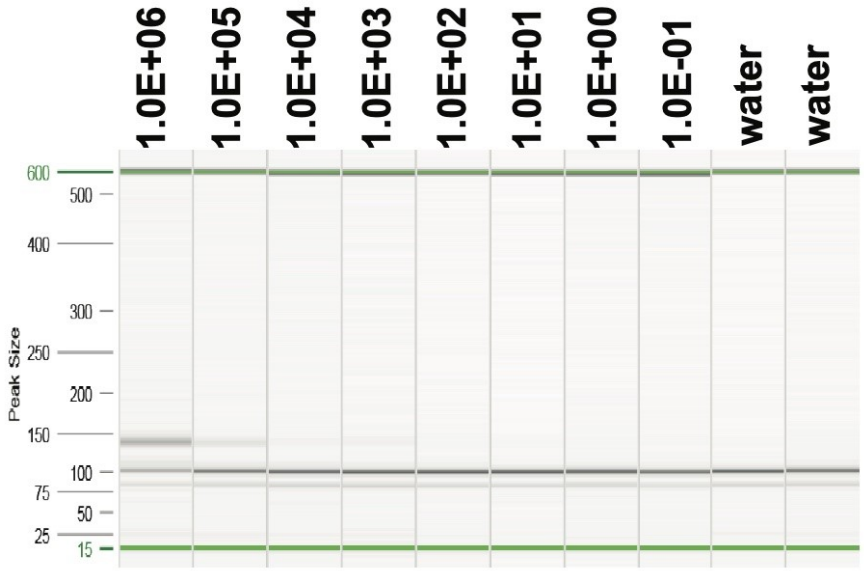

Figure S17: PCR1 performed with NF-DEL naïve (serially diluted from 10^6 to 10^{-1} number of molecules) and water negative controls with PCR1 cycles = 52.

9 Example of a cost estimation for library synthesis

	Universal	Codes 1 [(a)]	Adapt. 1	Codes 2[(a)]	Adapt. 2	Codes 3[(a)]	Adapt 3	Product
Amount (nmol)	10000	3000	3000	600	600	120	120	16
Cost of goods (euro/nmol)	0.2	2	0.7	2	0.7	2	0.7	
Total cost of goods (euro)	2000	6000	2100	1200	420	240	84	12044

Table S10: estimated encoding costs (oligonucleotides only) for the synthesis of 100 x 10^5 copies (16 nmol) of a 3 building blocks library containing 1'000'000'000 compounds (1'000 x 1'000 x 1'000). (a) The costs for amounts of oligonucleotides lower than 10 nmol each are underestimated (the costs per nanomole refer to 10 nmol scale production).

10 References

[1] Favalli N, Bassi G, Pellegrino C, et al. Stereo- and regiodefined DNA-encoded chemical libraries enable efficient tumour-targeting applications. *Nat Chem*. 2021;13(6):540-548. doi:10.1038/s41557-021-00660-y

[2] Oehler S, Lucaroni L, Migliorini F, et al. A DNA-encoded chemical library based on chiral 4-amino-proline enables stereospecific isozyme-selective protein recognition. *Nat. Chem.* 2023. doi.org/10.1038/s41557-023-01257-3

[3] Decurtins W, Wichert M, Franzini RM, et al. Automated screening for small organic ligands using DNA-encoded chemical libraries. *Nat Protoc*. 2016;11(4):764-780. doi:10.1038/nprot.2016.039