SUPPLEMENTARY FIGURES

Surface loops in a single SH2 domain are capable of encoding the spectrum of specificity of the SH2 family

Huadong Liu, Haiming Huang, Courtney Voss, Tomonori Kaneko, Wen Qin, Sachdev Sidhu and Shawn S-C. Li

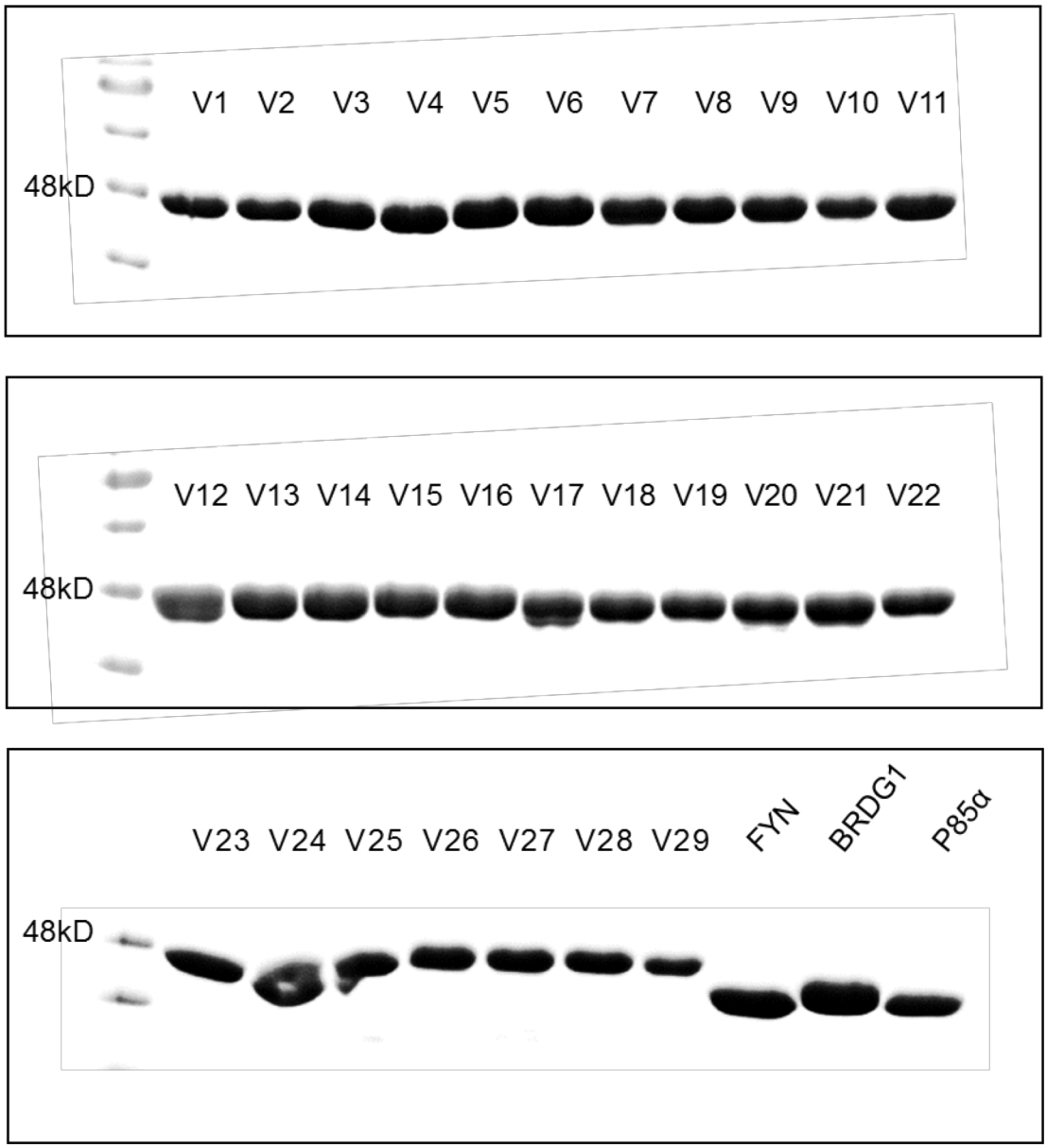

Figure S1: Coomassie blue staining SDS-PAGE gel of Fyn SH2 variants and wild-type SH2 domains from Fyn, BRGR1, or PI3K(p85α). The proteins, in $(His)_6$-GST fusion, were expressed in *E-coli* and purified by nickel column.

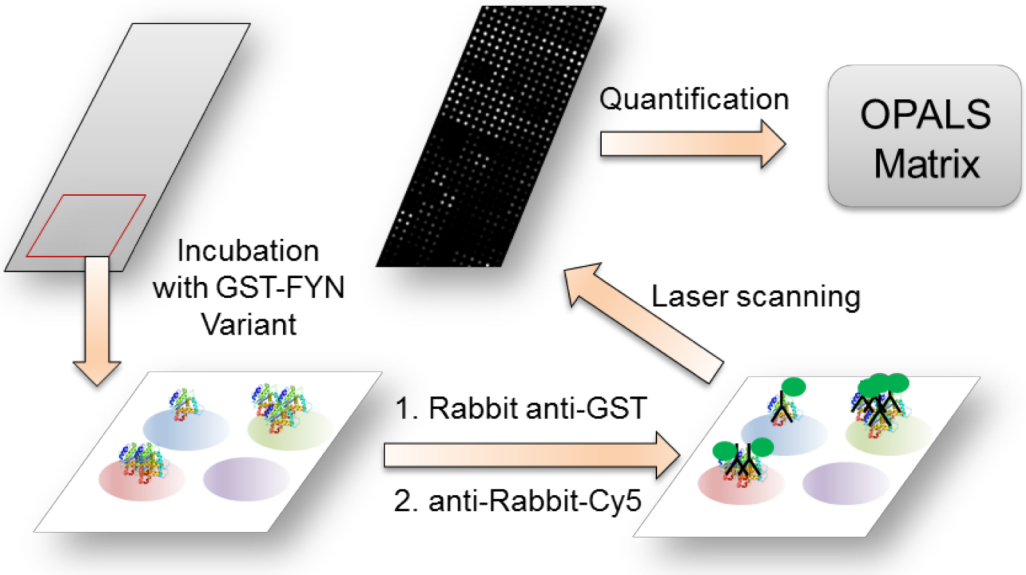

Figure S2: A flowchart of the Oriented Peptide Array Library (OPAL) strategy. OPAL slides were probed by different GST-SH2 domains or variants (mutants), followed by detection with anti-GST antibody and visualization with a Cy5-labeled secondary antibody.

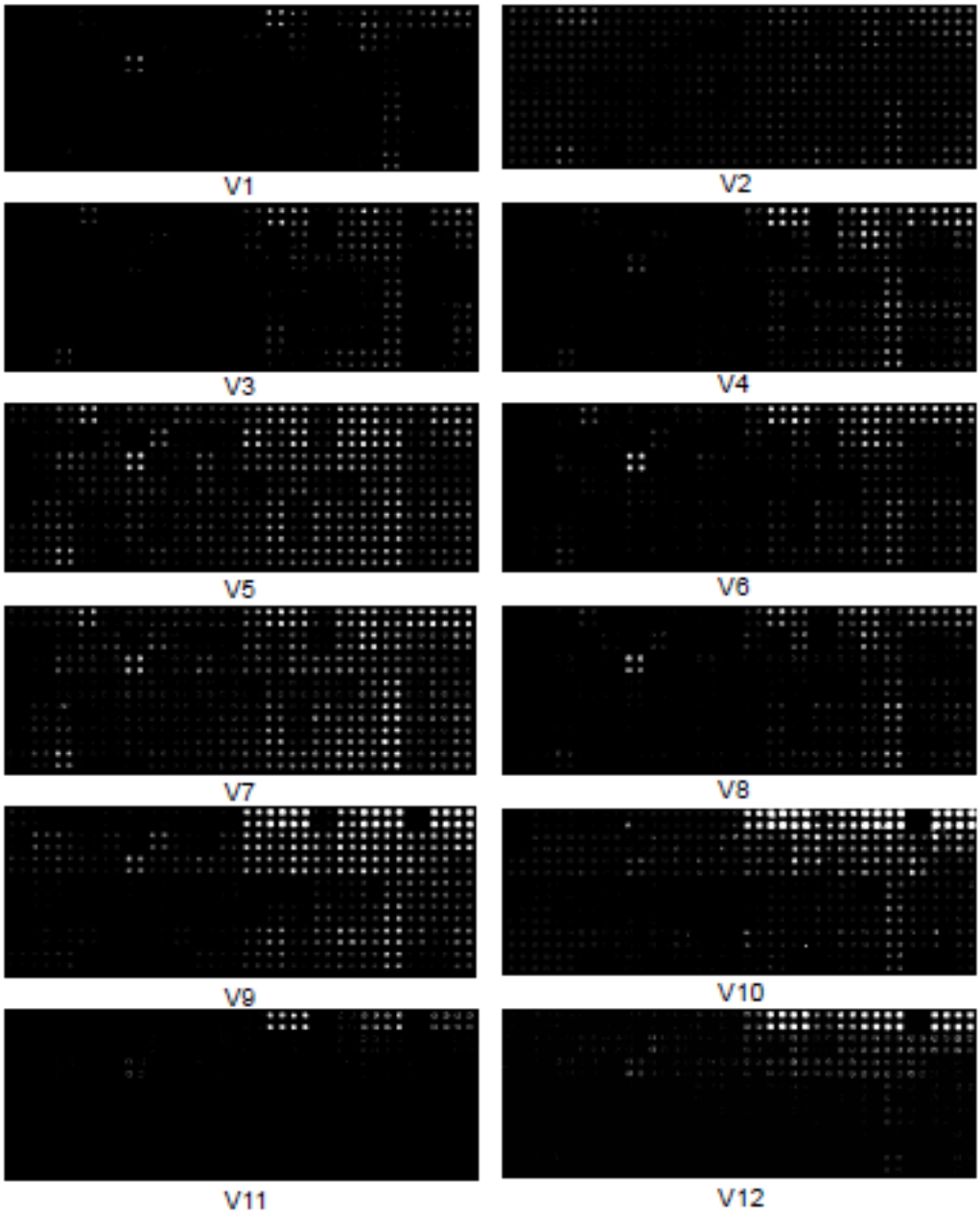

Figure S3a: Images of OPAL slides probed by different variants.

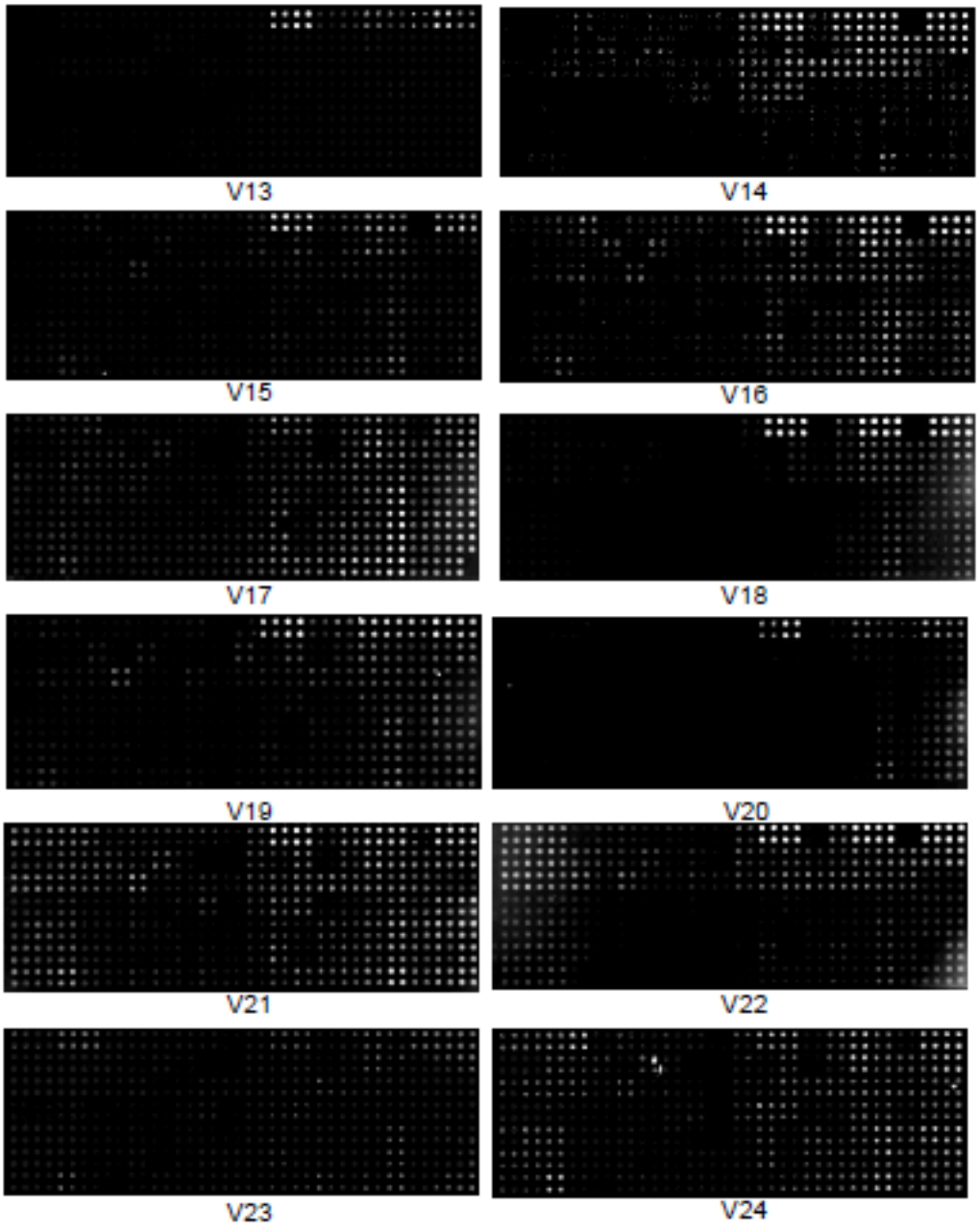

Figure S3b: Images of OPAL slides probed by different variants.

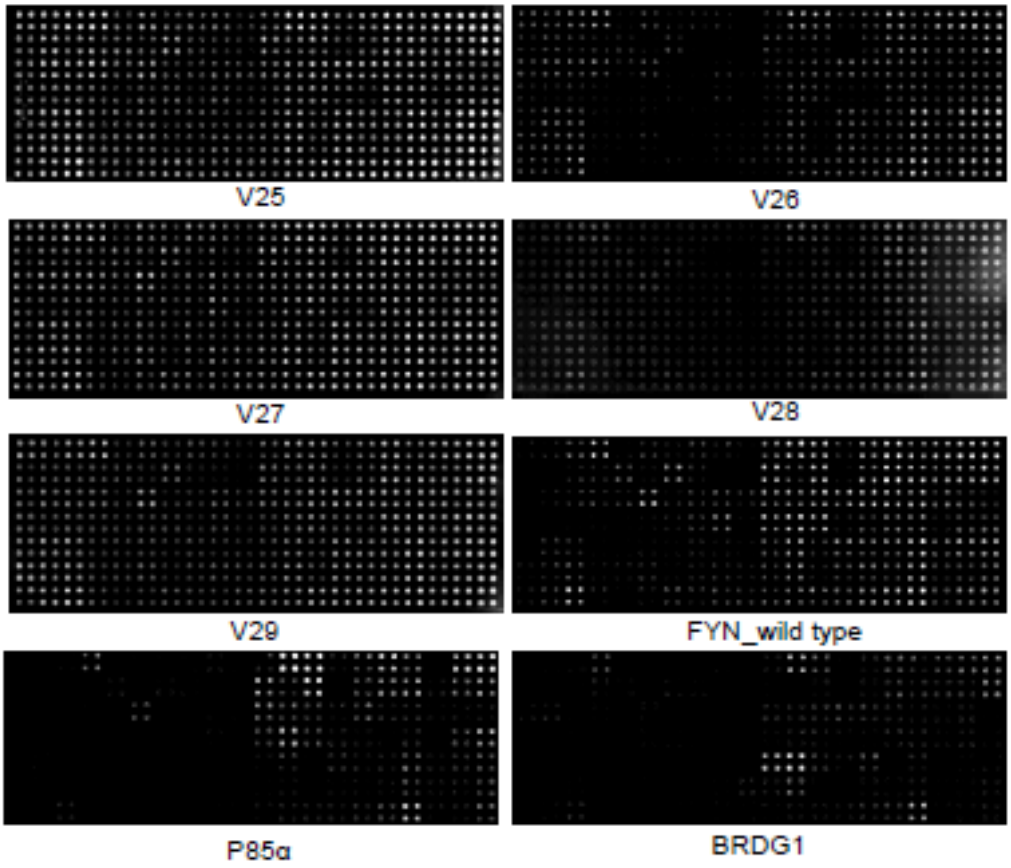

Figure S3c: Images of OPAL slides probed by different wt SH2 domains or variants.

P-1 Heat map

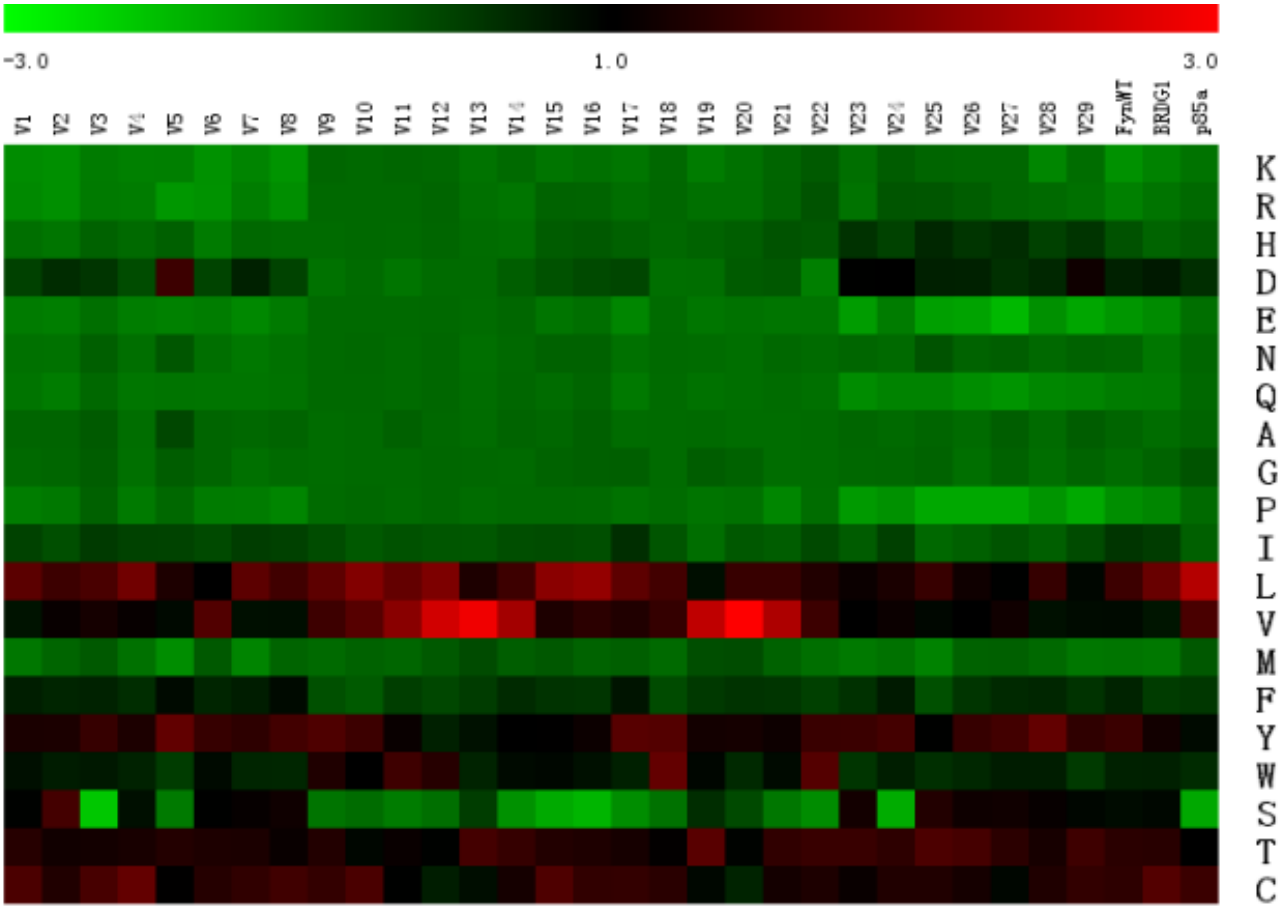

Figure S4a: Heat map of variant selectivity for positions P-1 based on the corresponding Z scores on OPAL. Color was generated according to the Z-score (i.e., green, small Z-score; red, large Z-score).

P+1 Heat map

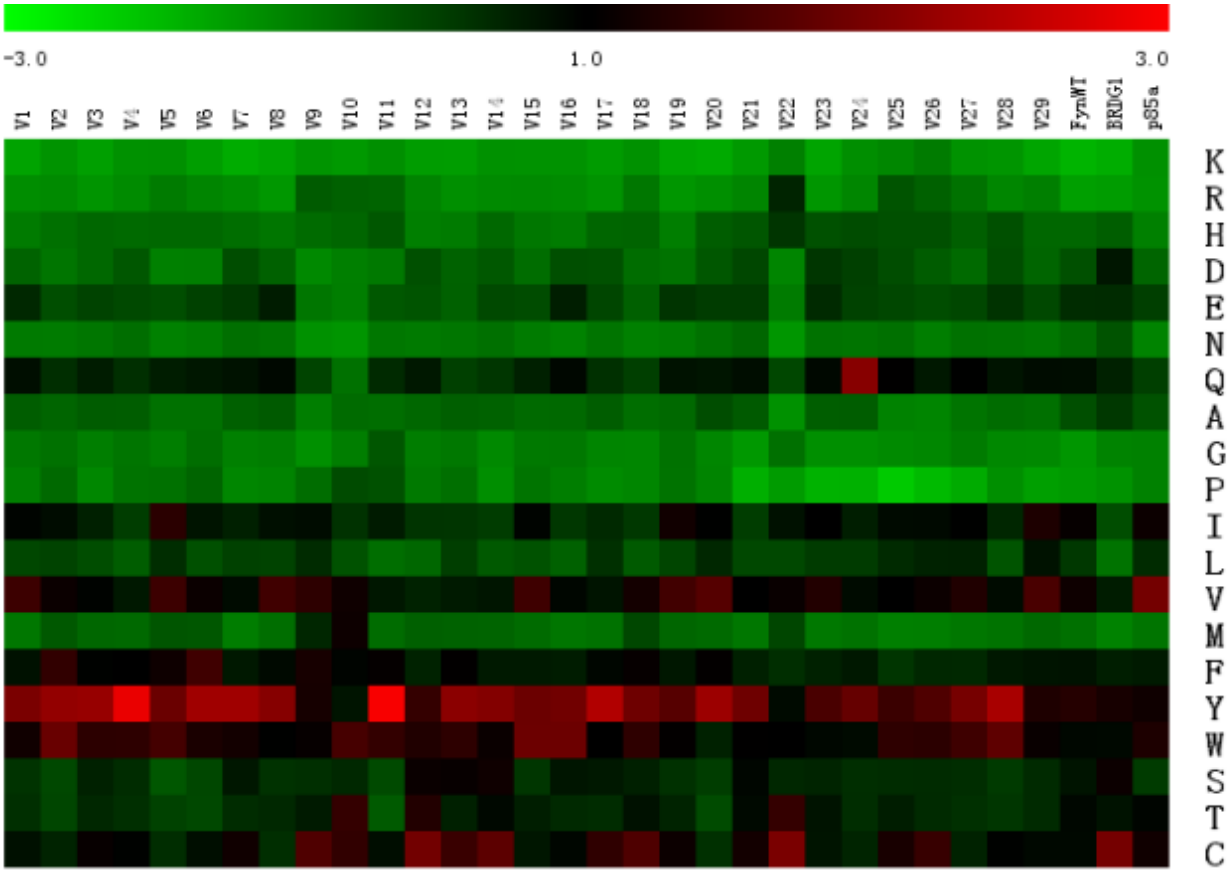

Figure S4b: Heat map of variant selectivity for positions P+1 based on the corresponding Z scores on OPAL. Color was generated according to the Z-score (i.e., green, small Z-score; red, large Z-score).

P+3 Heat map

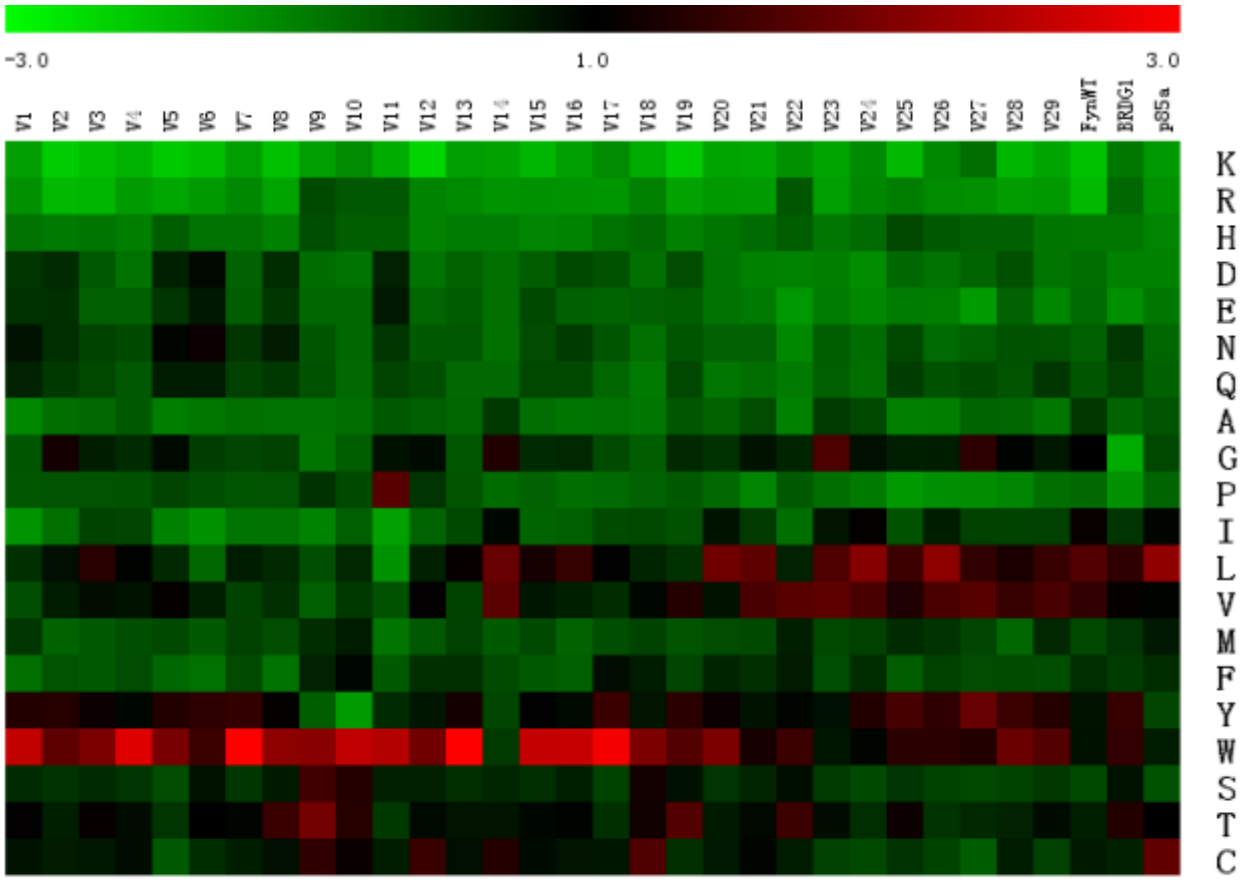

Figure S4c: Heat map of variant selectivity for positions P+1 based on the corresponding Z scores on OPAL. Color was generated according to the Z-score (i.e., green, small Z-score; red, large Z-score).

P+4 Heat map

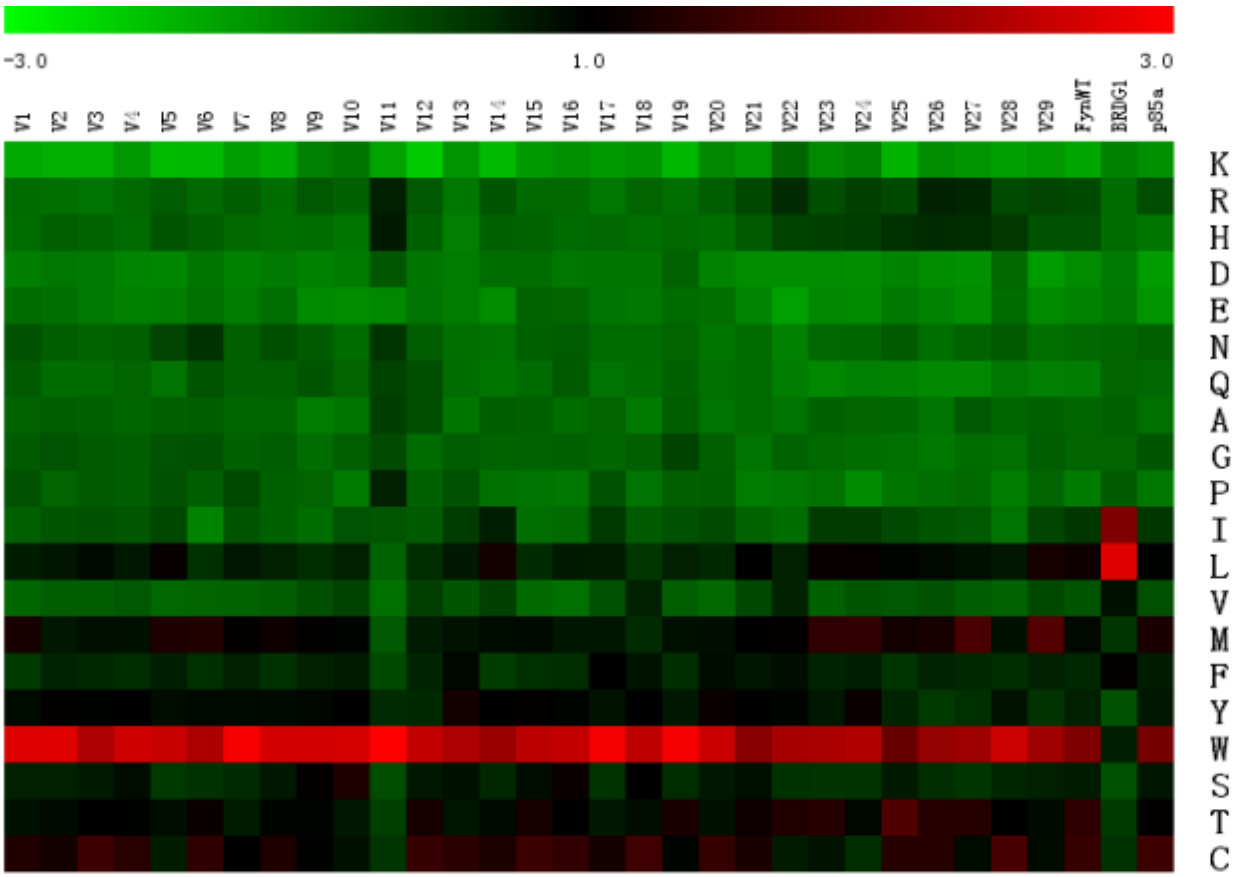

Figure S4d: Heat map of variant selectivity for positions P+4 based on the corresponding Z scores on OPAL. Color was generated according to the Z-score (i.e., green, small Z-score; red, large Z-score).

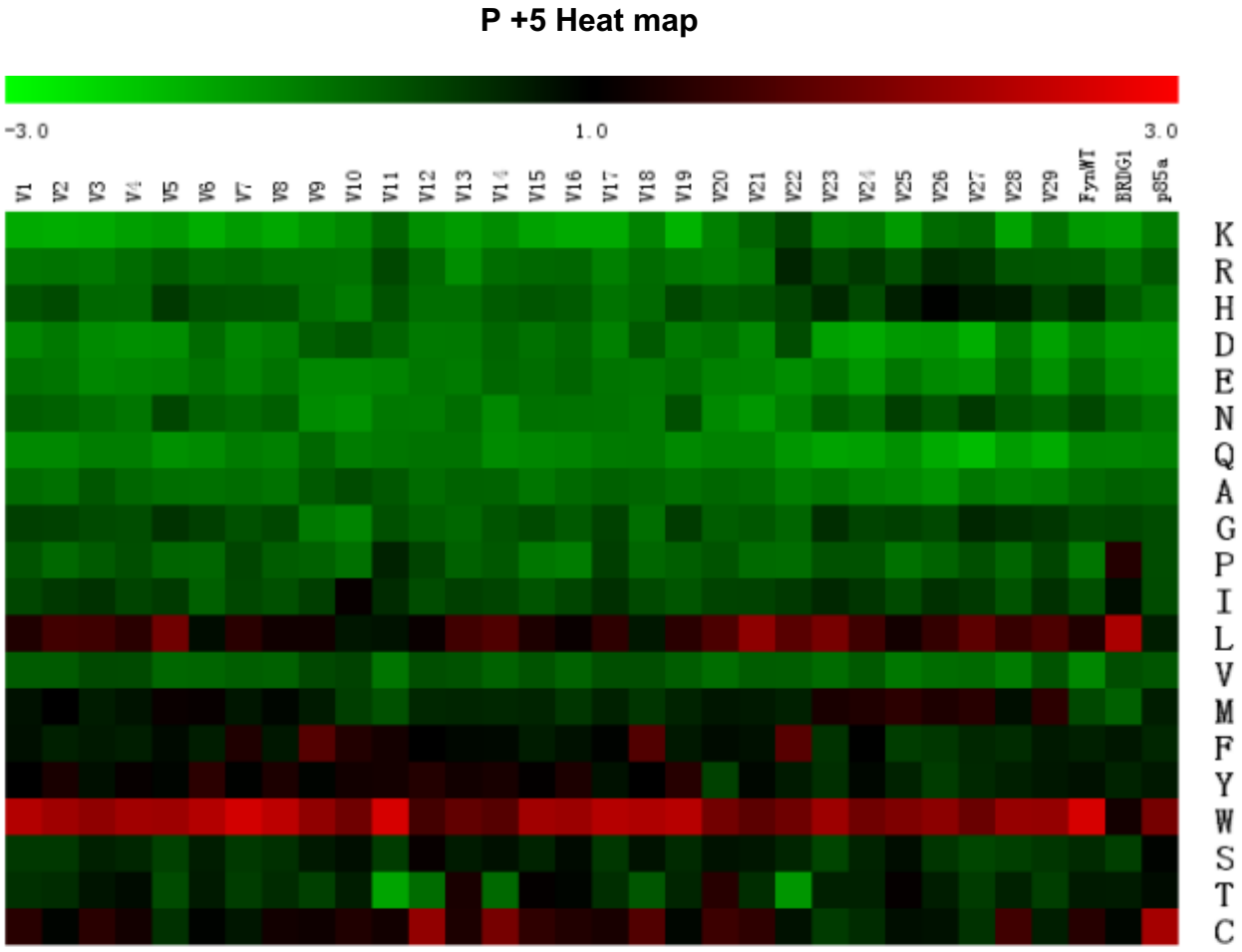

Figure S4e: Heat map of variant selectivity for positions P+5 based on the corresponding Z scores on OPAL. Color was generated according to the Z-score (i.e., green, small Z-score; red, large Z-score).

P+6 Heat map

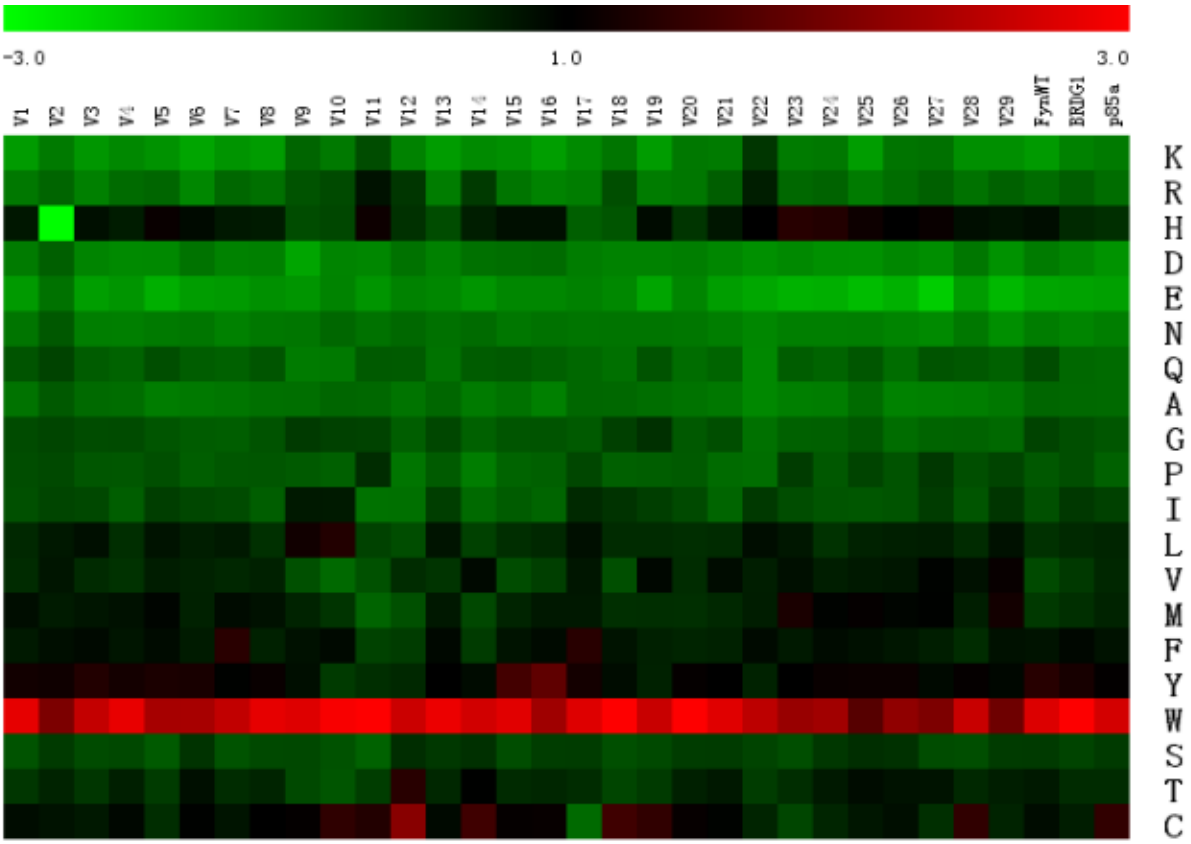

Figure S4f: Heat map of variant selectivity for positions P+6 based on the corresponding Z scores on OPAL. Color was generated according to the Z-score (i.e., green, small Z-score; red, large Z-score).

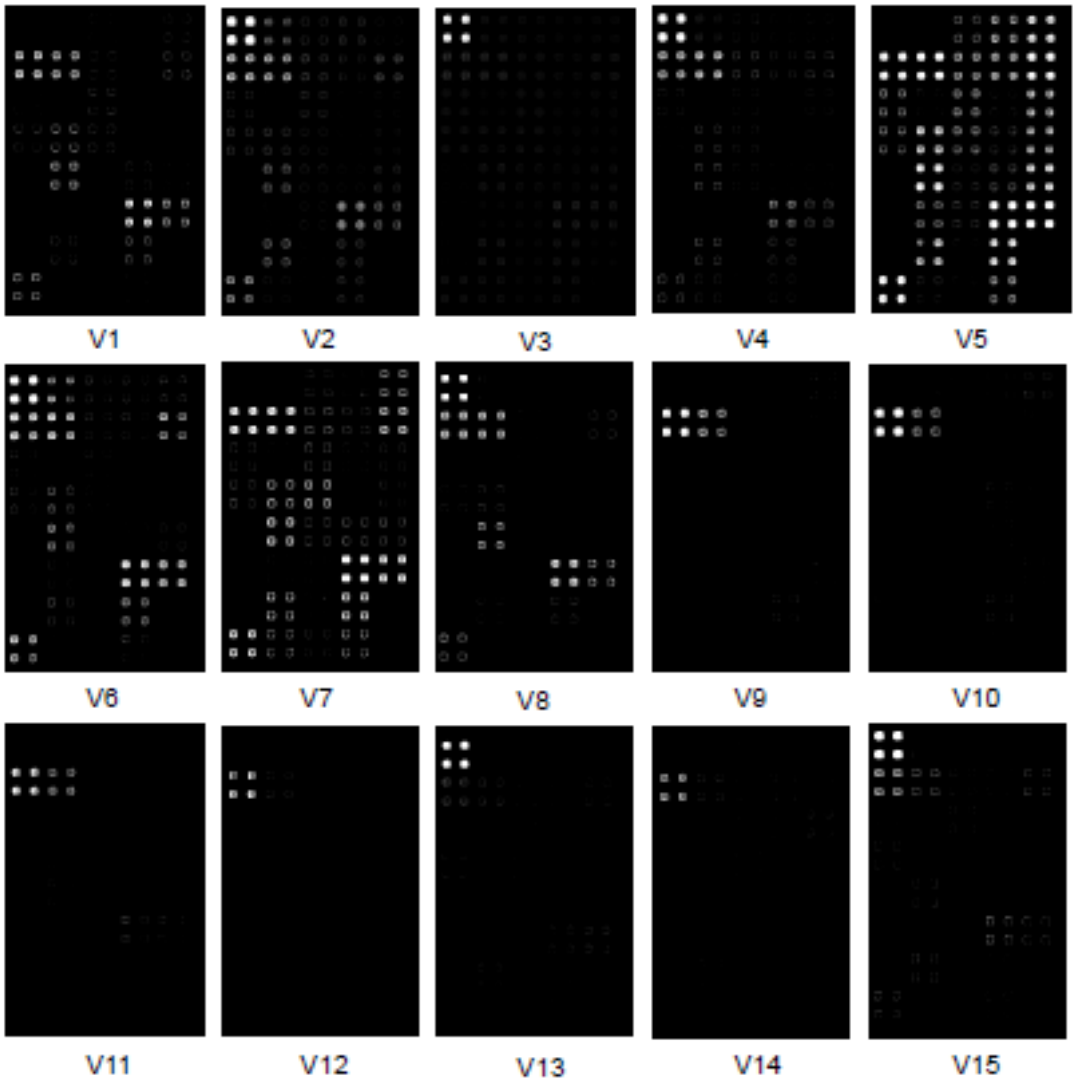

Figure S5a. Peptide ligand array probed by different variants. N-terminal biotin-labeled peptides were incubated with neutravadin. The resulting complex was printed on the slide in quadruplicates. Each slide was probed by a different GST fused FYN-SH2 variant or a control GST-SH2 domain. A rabbit-anti-GST antibody was used to detect the bound proteins, followed by imaging with a Cy5-labeled secondary antibody. Shown are images in a laser scanner.

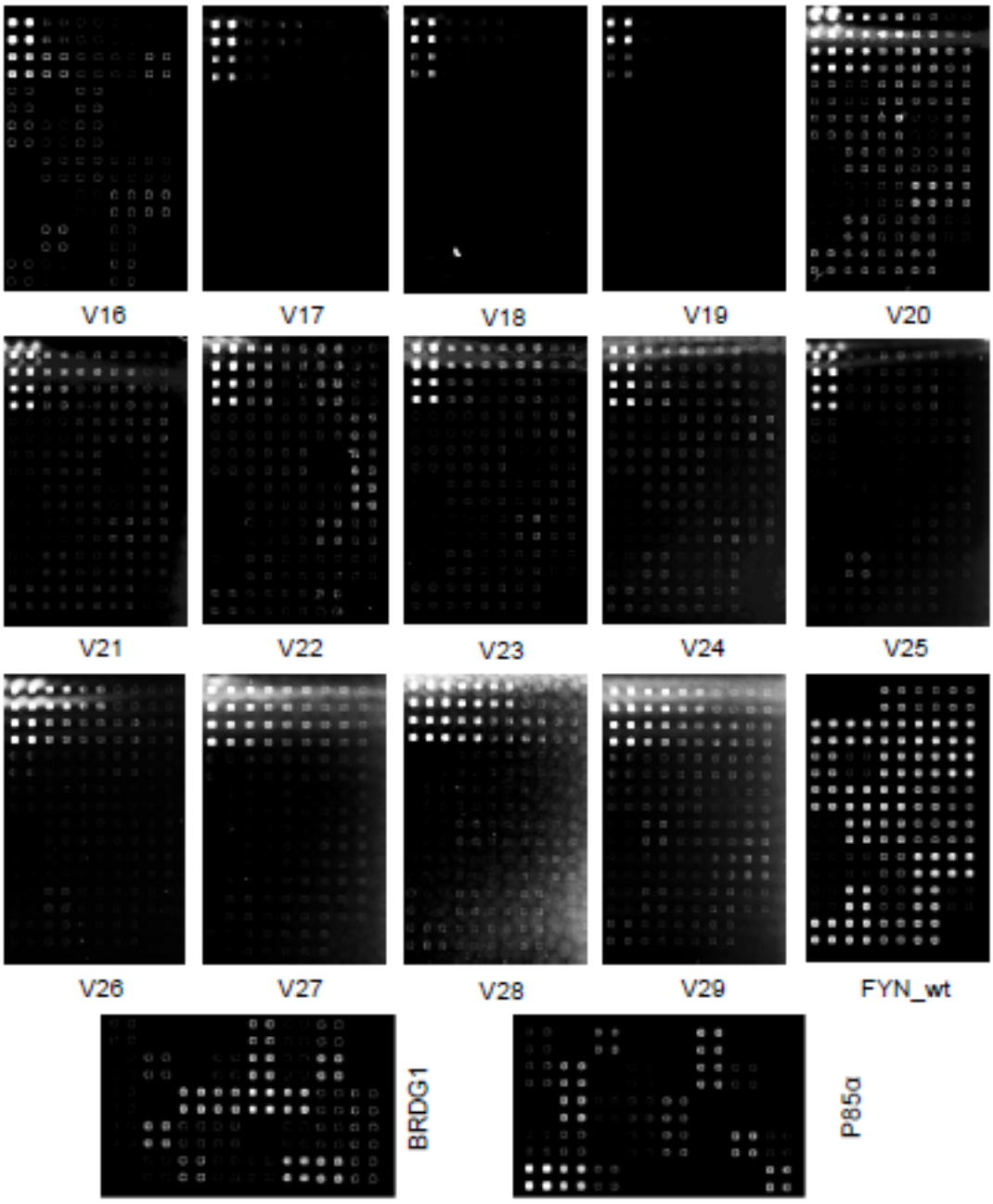

Figure S5b. Images of peptide ligand arrays probed by different Fyn SH2 variants or control wt SH2 domains.

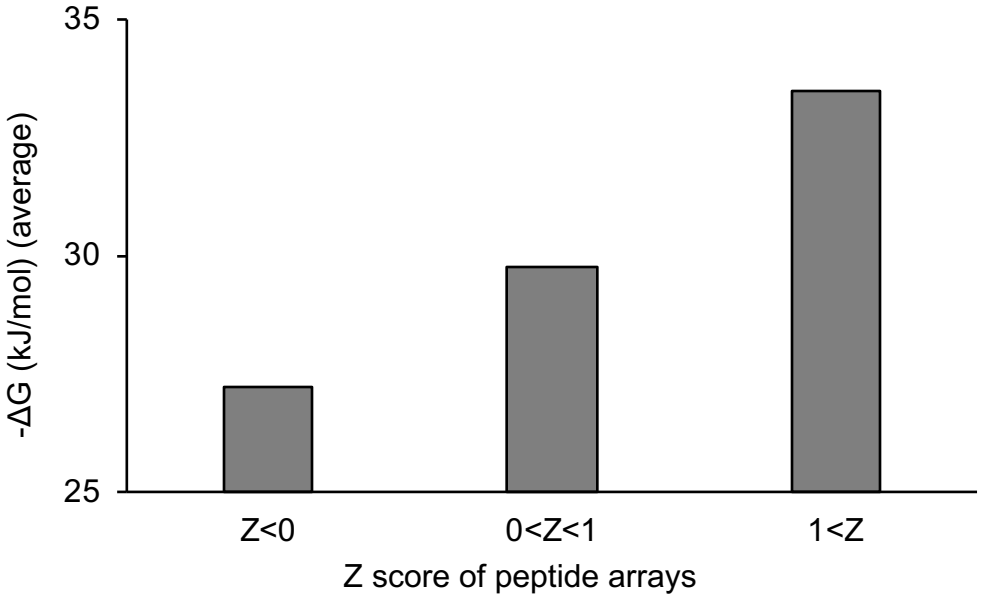

Figure S6: In-solution peptide-variant binding affinities correlate with the peptide array data. The peptide-SH2 variant binding free energy (ΔG) were calculated from the corresponding equilibration constant (Kd) derived from florescence polarization measurements according to the equation ΔG^o =-RTln(Kd). Shown is a plot of the average - ΔG^o value for SH2 variants grouped according to their Z scores on the peptide array. The peptide array data matches well with in-solution binding assay- that is, the higher Z score the greater the binding affinity.

FynWT

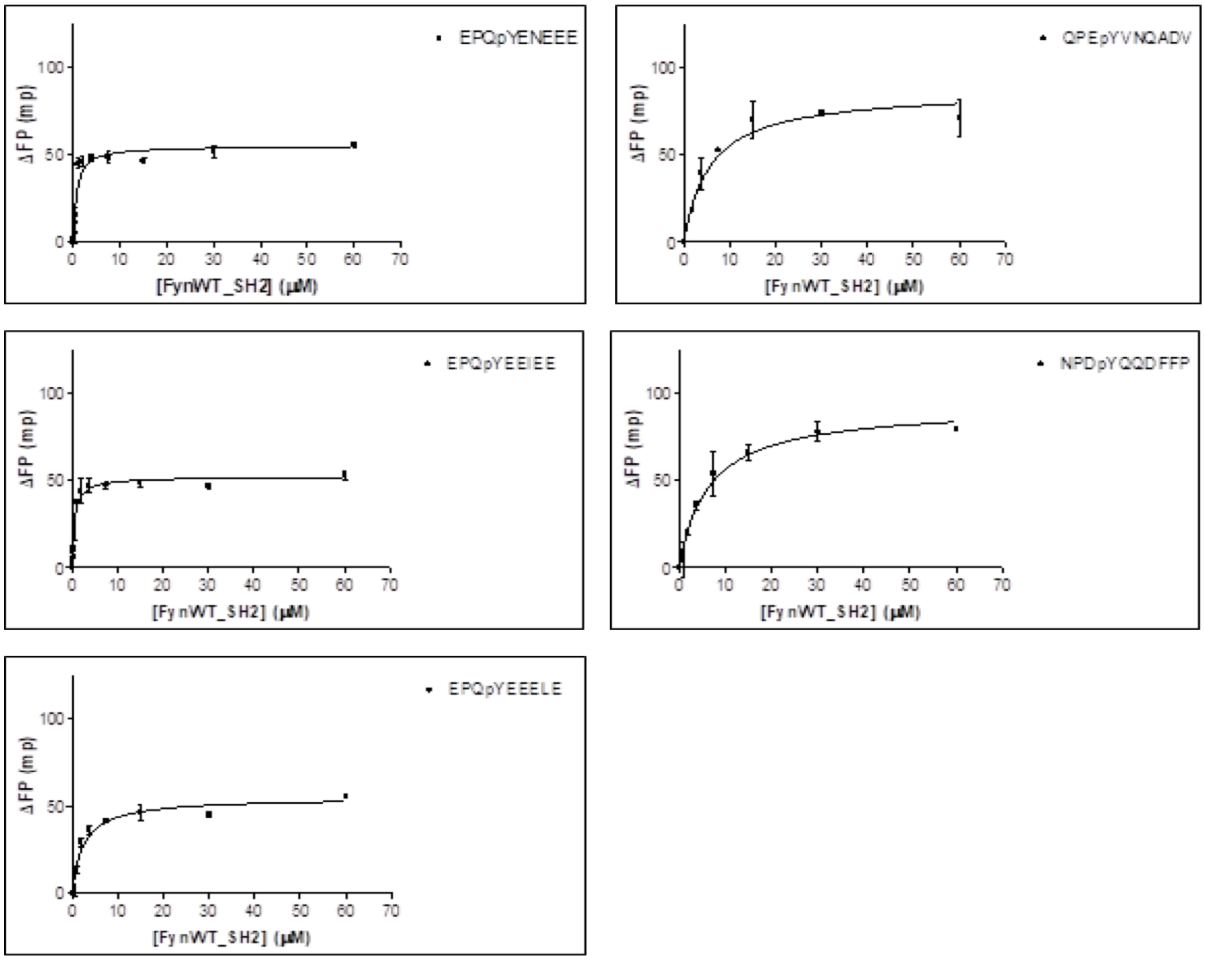

Figure S7a: Representative binding curves for the wt Fyn SH2 domain to different peptides obtained from fluorescence polarization measurements (see also Table 1)

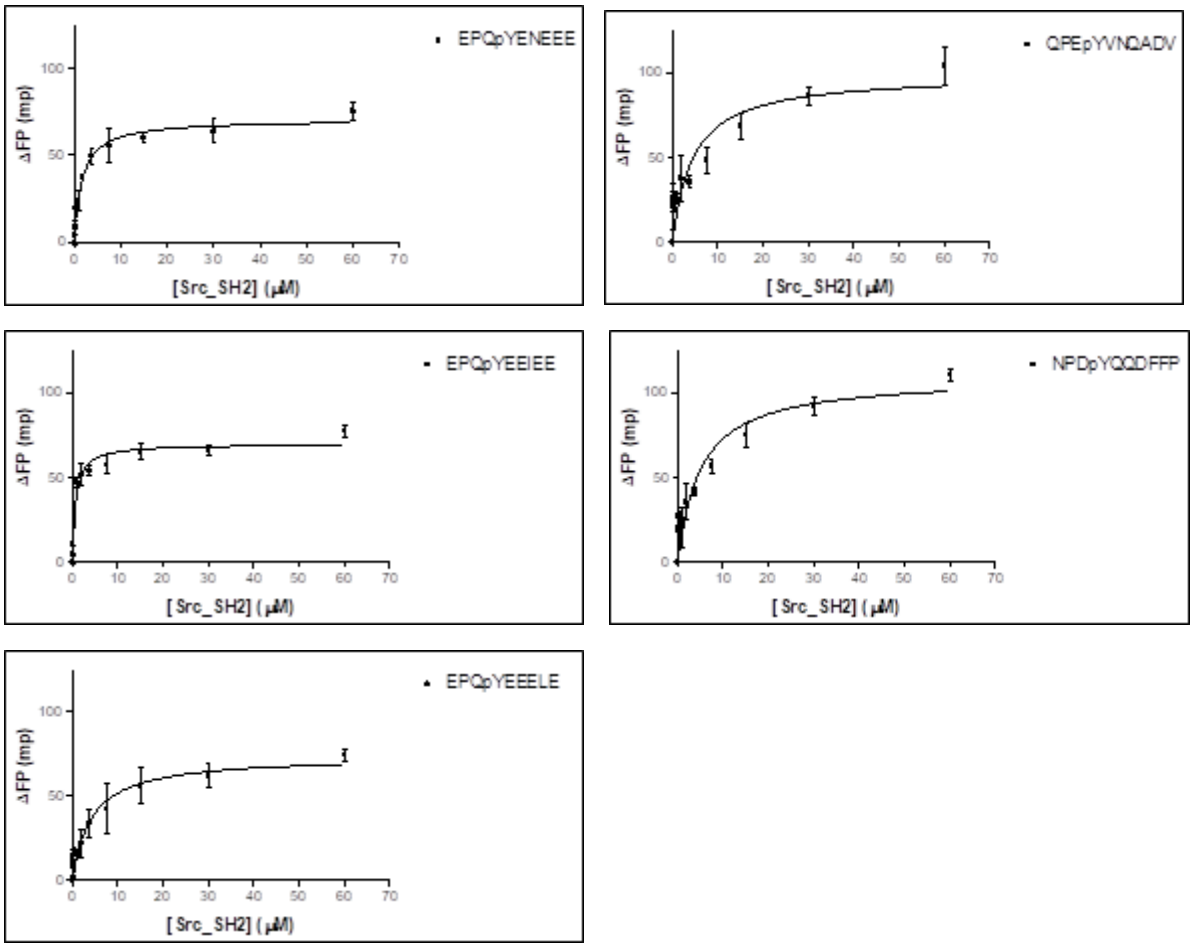

Figure S7b: Representative binding curves for the wt Src SH2 domain to different peptides (see also Table 1)

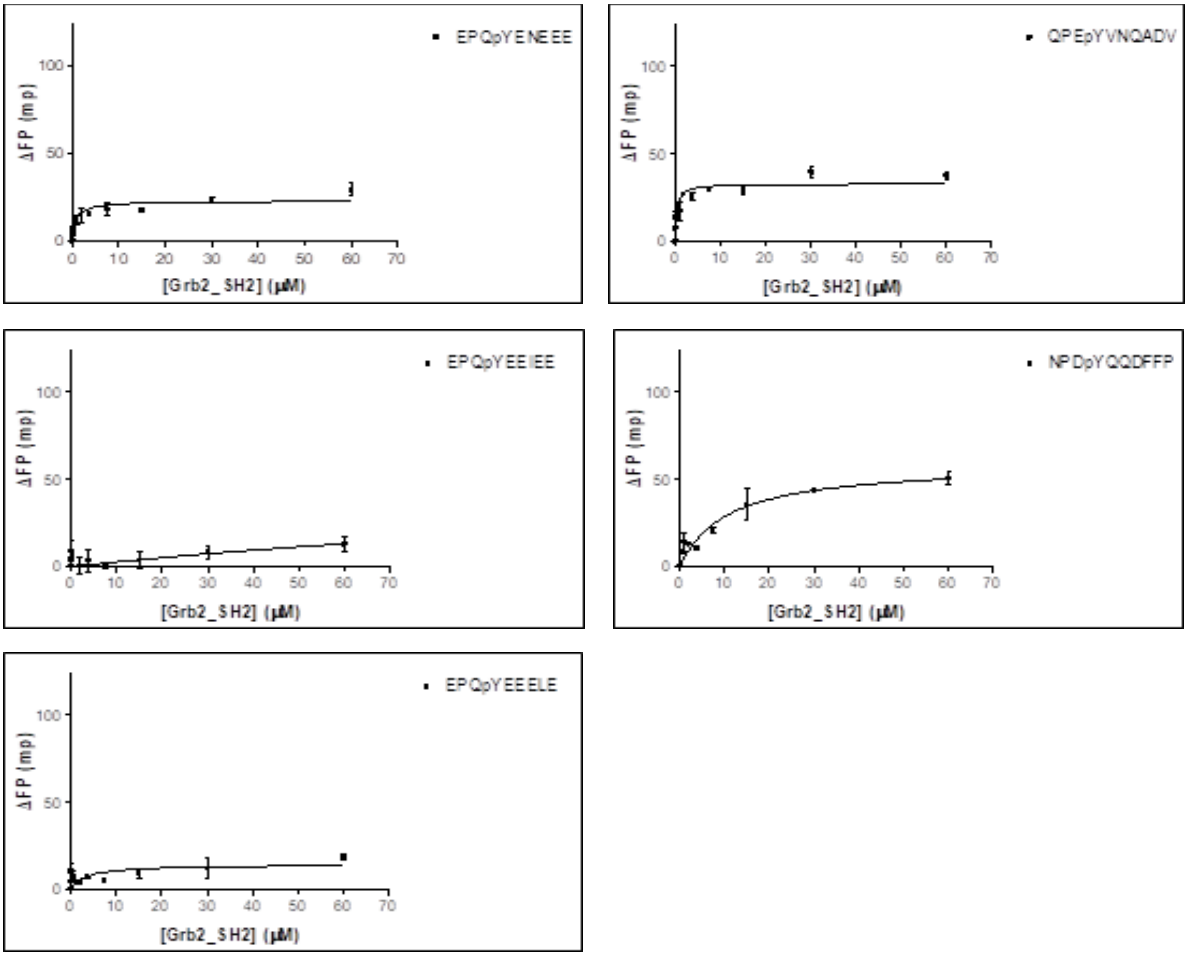

Figure S7c: Representative binding curves for the wt Grb2 SH2 domain to different peptides (see also Table 1)

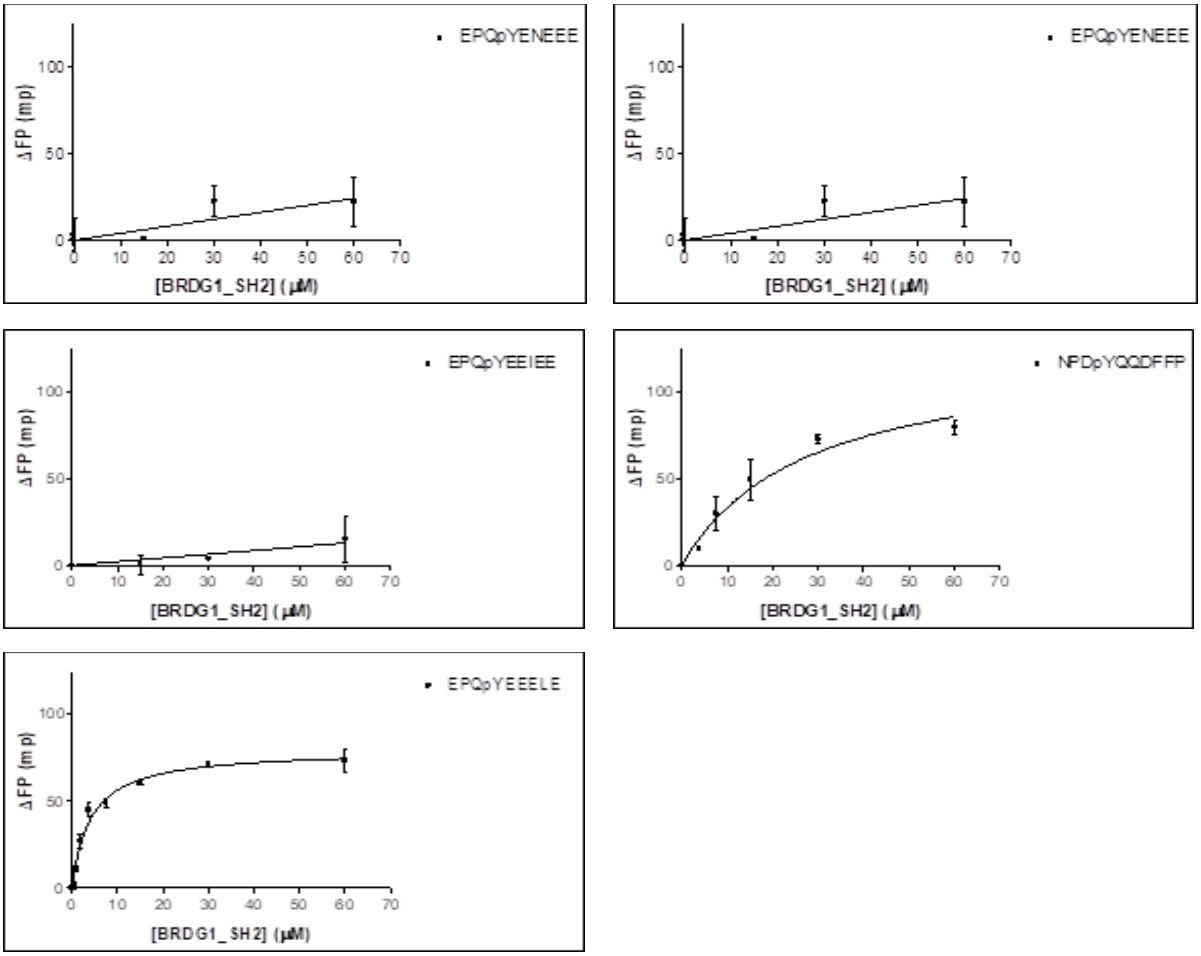

Figure S7d: Representative binding curves for the wt BRDG1 SH2 domain to different peptides (see also Table 1)

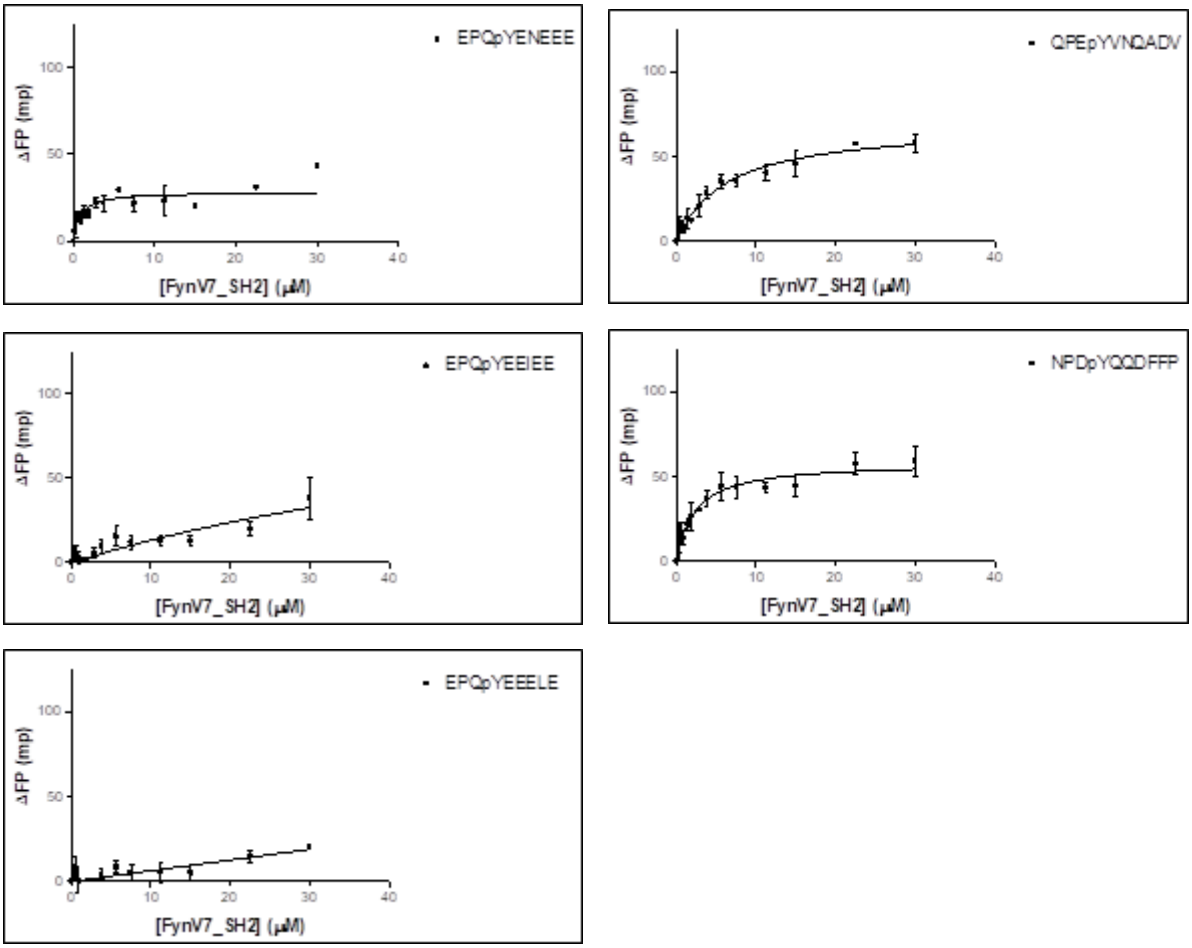

Figure S7e: Representative binding curves for the Variant #7 to different peptides (see also Table 1)

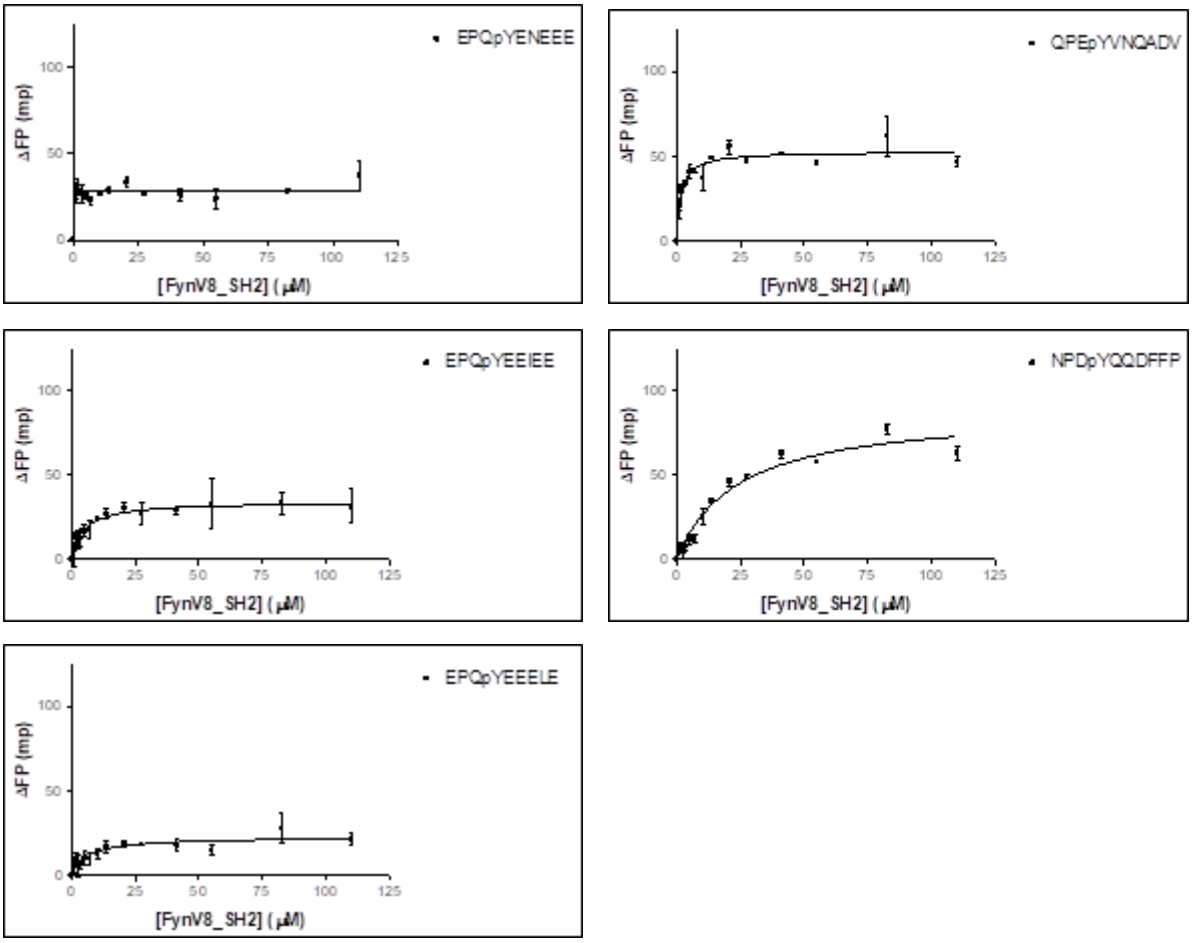

Figure S7f: Representative binding curves for the Variant #8 to different peptides (see also Table 1)

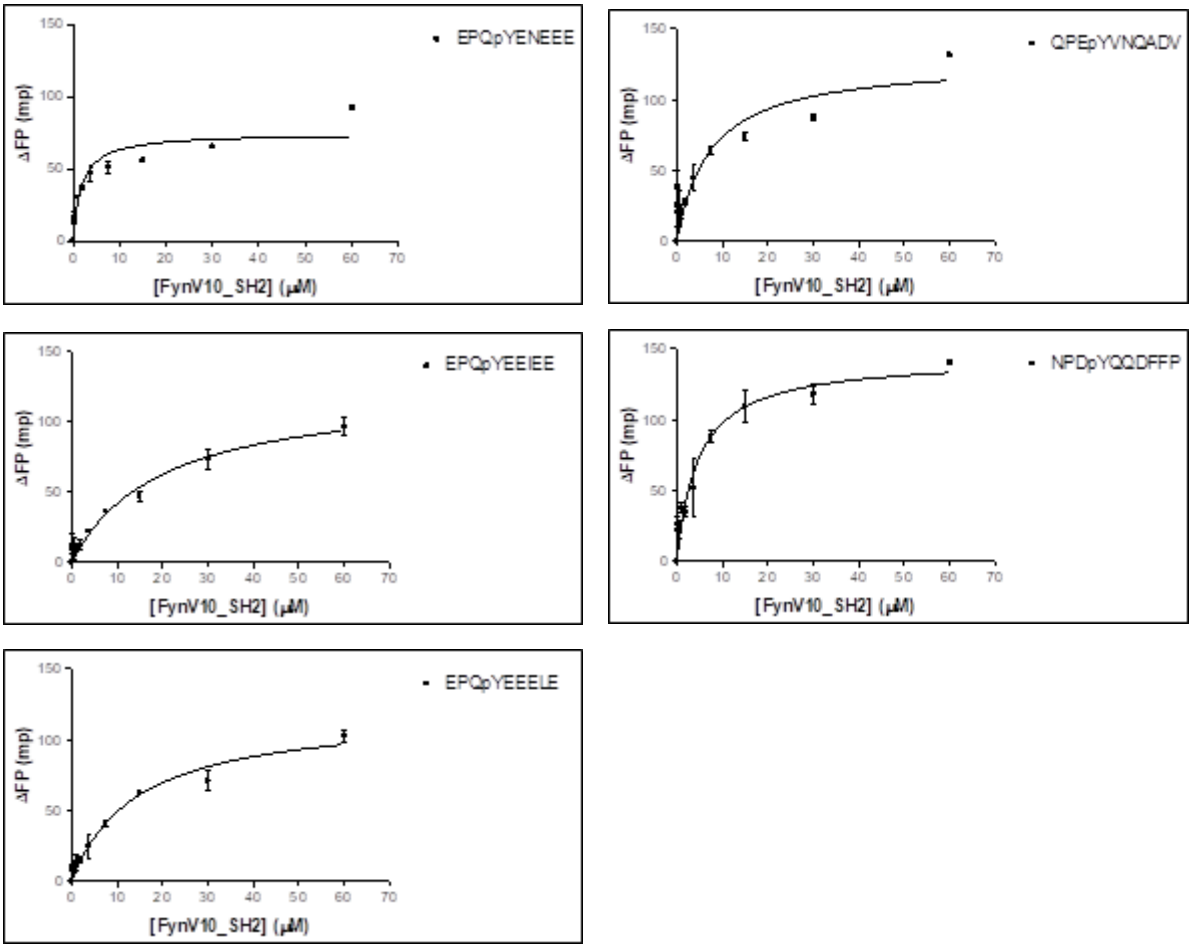

Figure S7g: Representative binding curves for the Variant #10 to different peptides (see also Table 1)

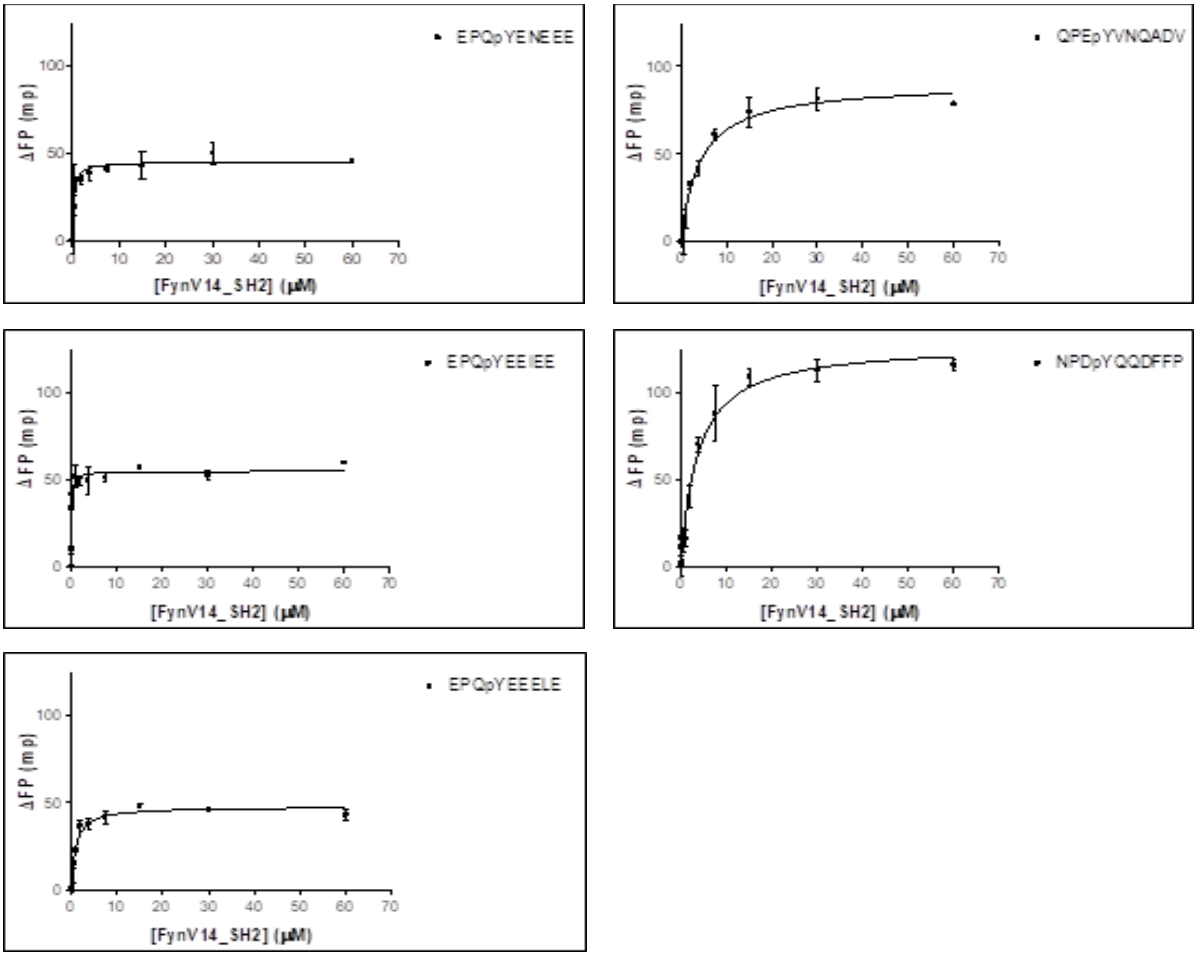

Figure S7h: Representative binding curves for the Variant #14 to different peptides (see also Table 1)

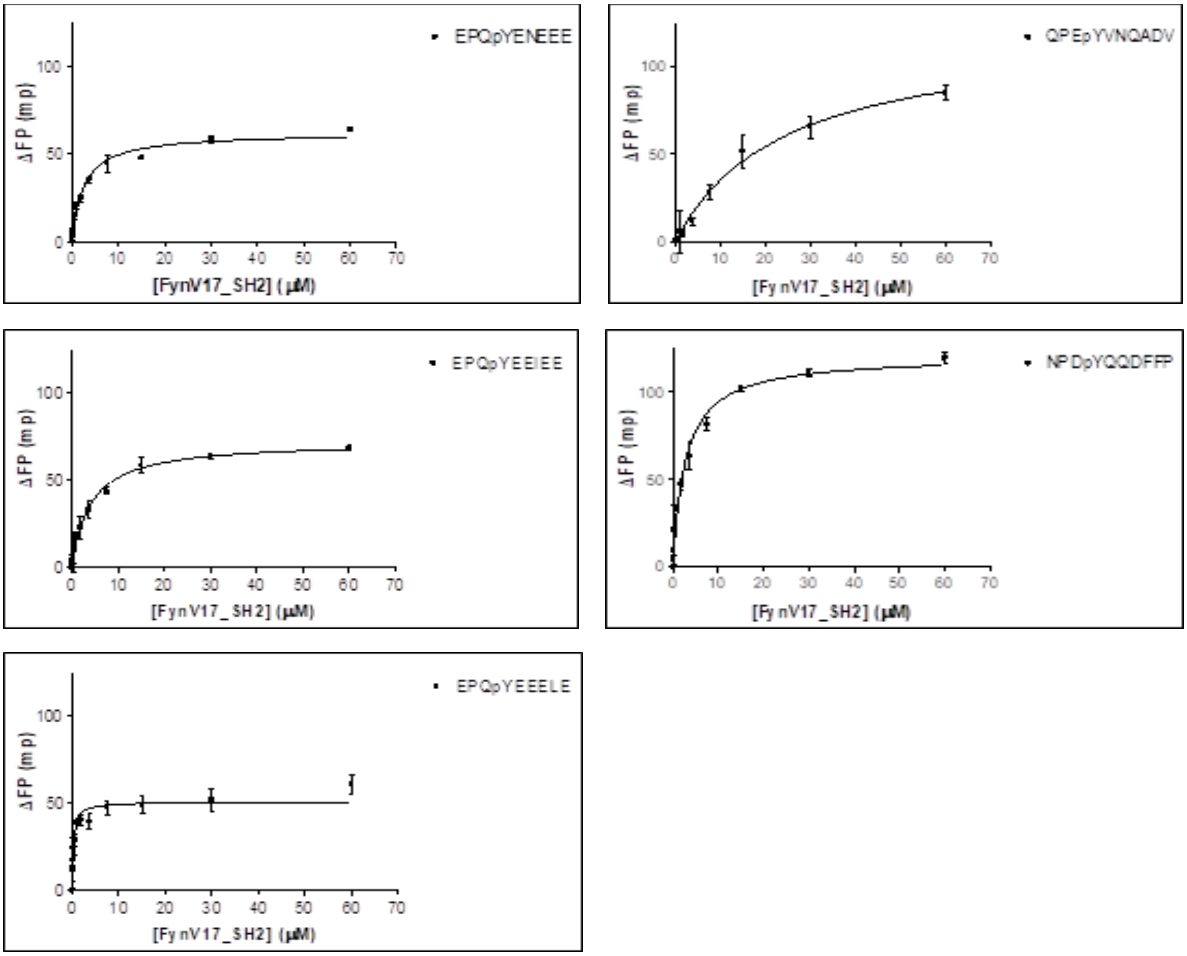

Figure S7i: Representative binding curves for the Variant #17 to different peptides (see also Table 1)

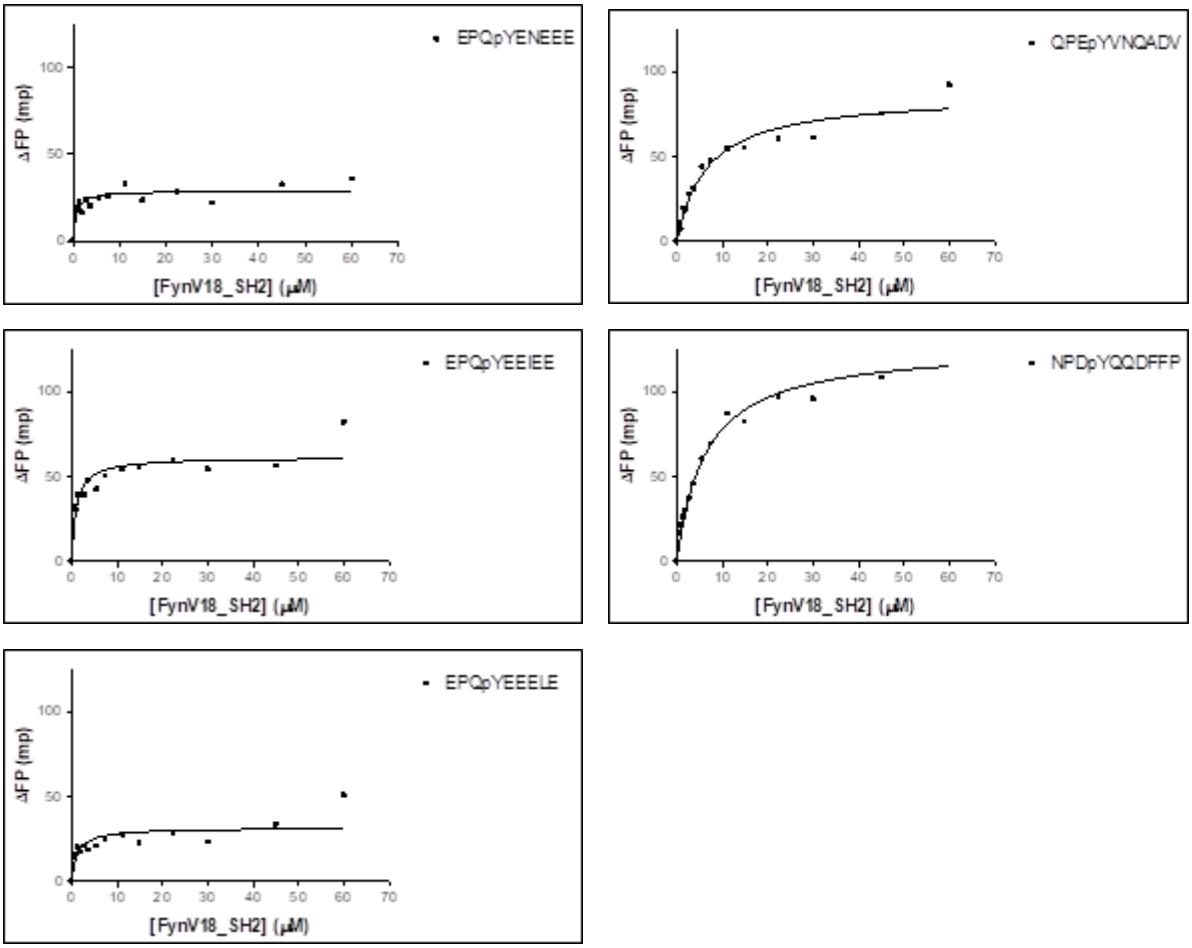

Figure S7j: Representative binding curves for the Variant #18 to different peptides (see also Table 1)

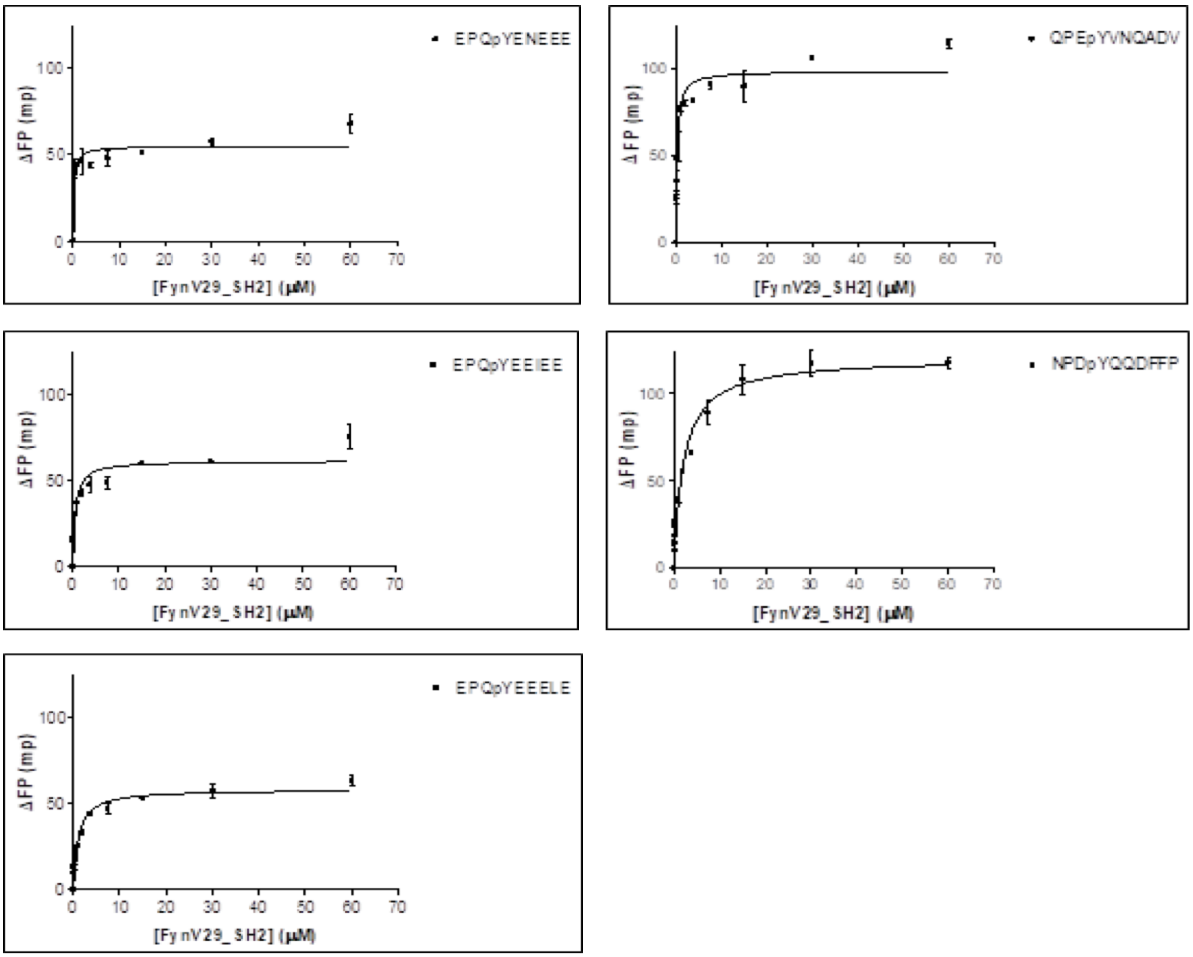

Figure S7k: Representative binding curves for the Variant #29 to different peptides (see also Table 1)

Supplementary Table 1. A count of unique phage variants identified by different bait peptides

Bait peptide	Sequence	Motif	Total variants	Library 3+3	Library 3+x
pYEEI-variant_2N#	EPQpYE**N**EEE	P+2N	0	0	0
pYEEI-variant_3I#	EPQpYEE**I**EE	P+3I	11	5	6
pYEEI-varient_4L#	EPQpYEEE**L**E	P+4L	14	8	6
VEGFR1-pY1213#	DVRpYV**N**AAKF	P+2N	0	0	0
ShcA-pY239#	DHQpYY**N**DAPG	P+2N	20	8	12
β2-adrenoreceptor-pY350#	SKApYG**N**GASS	P+2N	0	0	0
PDGFRβ-pY716#	AELpYS**N**AAPV	P+2N	14	8	6
ErbB2-pY1139#	QPEpYV**N**QADV	P+2N	2	1	1
TIE2-pY1102	RKTpYV**N**TTLY	P+2N	0	0	0
FCERB-pY219	DRVpYEE**L**NIYS	P+3L	0	0	0
SIG11-pY668	TTEpYSE**I**KIHT	P+3I	5	5	0
CD79A-pY188	ENLpYEG**L**NLDD	P+3L	14	7	7
CEA20-pY578	ESIpYEV**L**GMQQ	P+3L	4	3	1
TRAF7-pY275*	QDTpYETH**L**ET	P+4L	2	2	0
MALT1-pY470*	RNDpYDDT**I**PI	P+4I	3	0	3
RSKL-pY423	YQHpYDLD**L**KD	P+4L	0	0	0
B-raf-pY85*	YEEpYTSK**L**DA	P+4L	0	0	0
EGFR-pY869	EKEpYHAEGGK		0	0	0
EGFR-pY915	SKPpYDG**IP**AS	P+3I, P+4P	0	0	0
EGFR-pY944#	IDVpYMI**MV**KA	P+3M, P+4V	0	0	0
EGFR-pY978	PQRpYLV**I**QGD	P+3I	4	2	2
EGFR-pY998	SNFpYRA**LM**DE	P+3L, P+4M	3	3	0
EGFR-pY1016	ADEpYLI**P**QQG	P+3P	6	6	0
EGFR-pY1069	LQRpYSSD**P**TG	P+4P	0	0	0
EGFR-pY1092	VPEpYI**N**QSVP	P+2N	8	0	8
EGFR-pY1110	NPVpYH**N**Q**P**LN	P+2N, P+4P	15	8	7
EGFR-pY1125	DPHpYQD**P**HST	P+3P	1	1	0
EGFR-pY1138	NPEpYL**N**T**V**QP	P+2N, P+4V	7	6	1
EGFR-pY1172	NPDpYQQD**F**FP	P+4F	8	7	1
EGFR-pY1197	NAEpYLR**V**APQ	P+3V	0	0	0
ErbB4-pY1035	PPIpYTSRARI		0	0	0
ErbB4-pY1066	QFVpYRDGGFA		0	0	0
ErbB4-pY1208	EPLpYL**N**TFAN	P+2N	11	9	2
ErbB4-pY1221	KAEpYLKNNIL		0	0	0
ErbB4-pY1301	PPPpYRHRNTV		0	0	0
		Sum	**152**	**89**	**63**

Notes:

These peptides contain variations from the wt sequences to avoid confusion in classification; Each peptide contains an N-terminal ahx-ahx (where ahx is 6-aminohexanoic acid) linker for attachment of biotin and immobilization.

* These pTyr sites are not reported as natural phosphorylation sites.

EGFR 869, 915 & 944 sites (*italicized*) are located in the kinase domain.

The peptide classification is based on the presence of an Asn at the P+2 or a hydrophobic residue at the P+3 or P+4 position.