Supplemental Figures

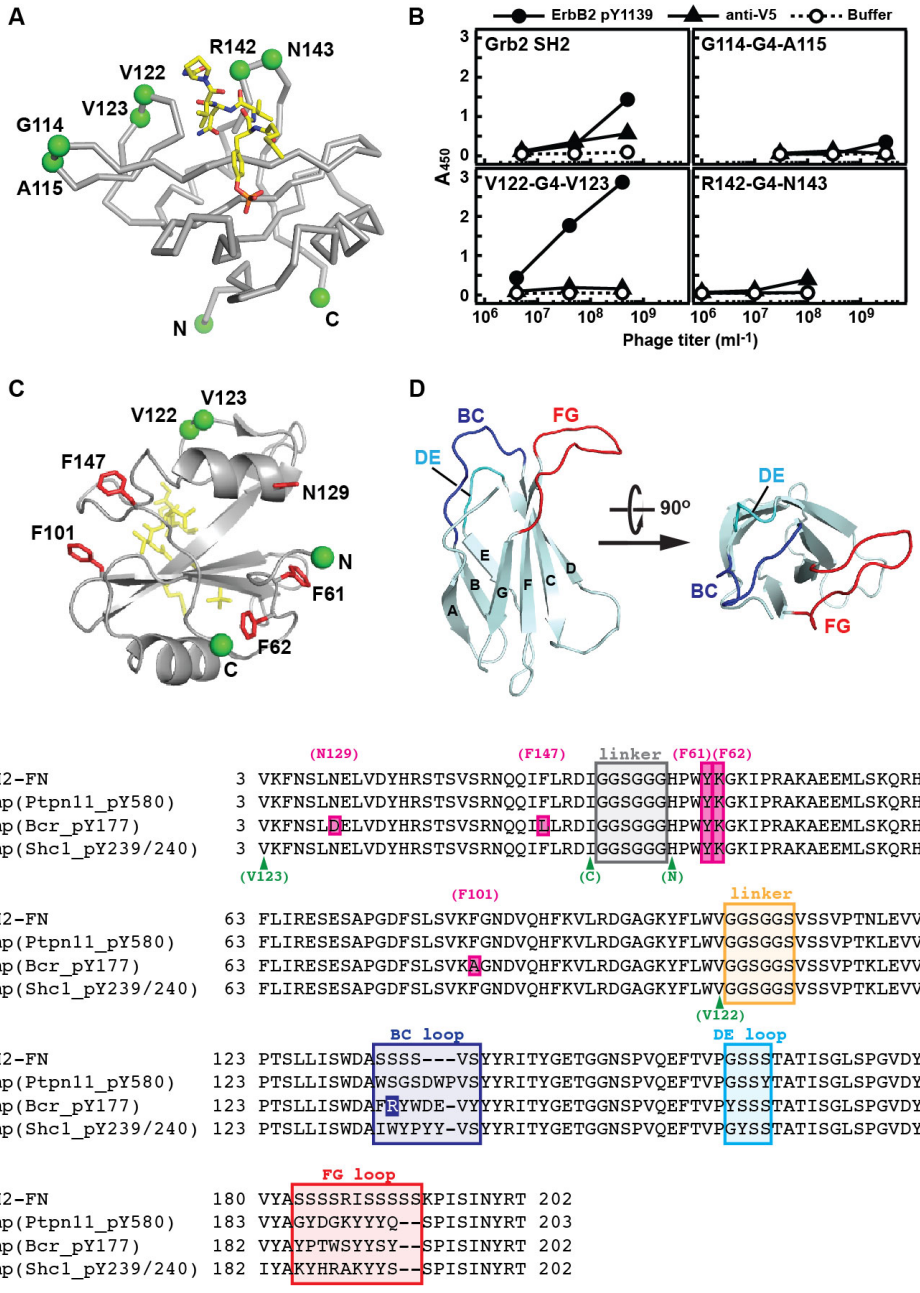
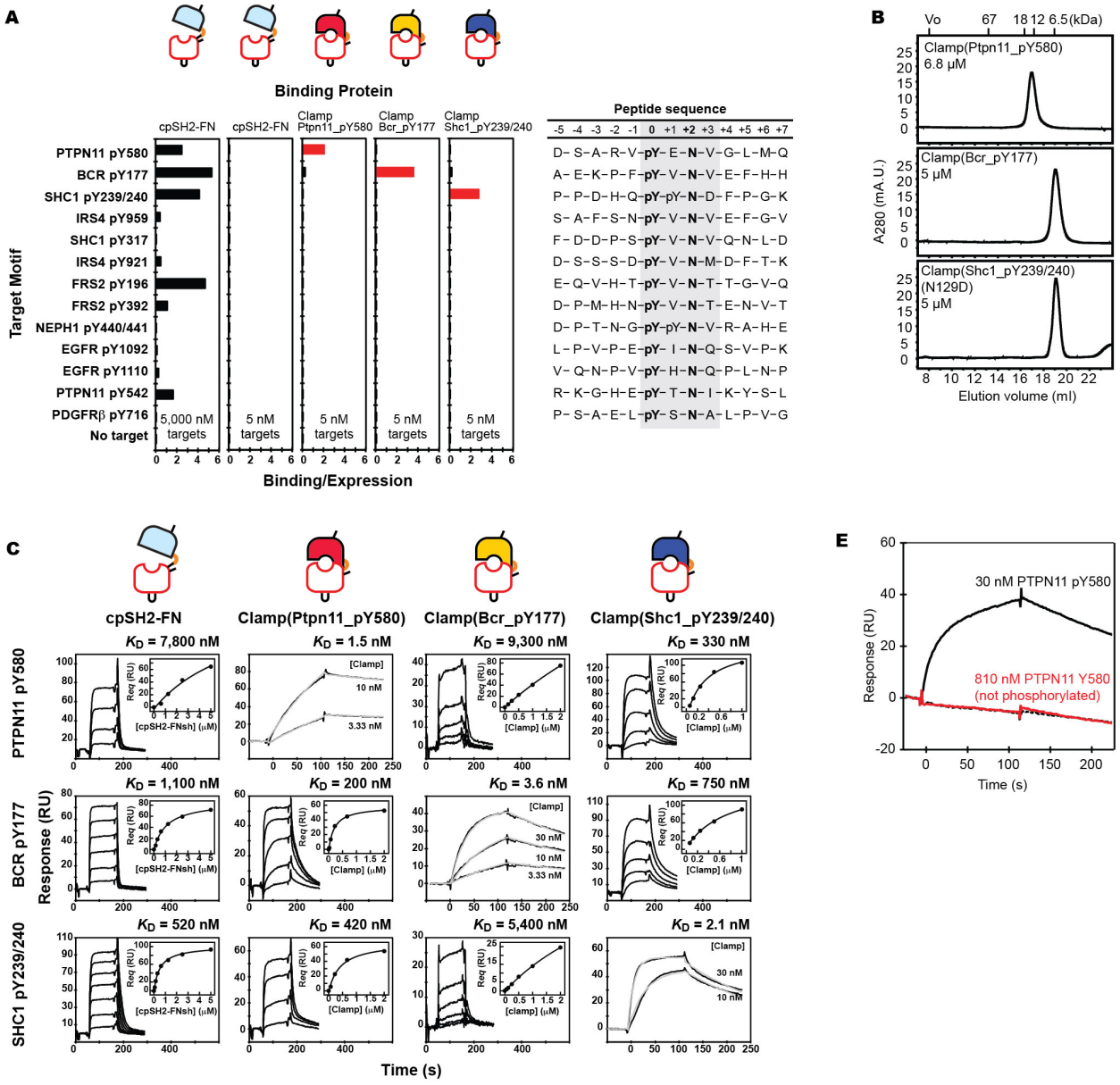


Figure S1. Redesign of the Grb2 SH2 domain (cpSH2) and construction of pY-clamps, Related to Figure 1. **(A)** To identify the positions of new N and C termini that minimally impact the stability and function of the SH2 domain, we examined the effects of inserting Gly residues. The C**α** atoms of the original N and C terminus and the residues between which Gly_4 was inserted are shown as green spheres and labeled in the crystal structure of the Grb2 SH2 domain (gray) in complex with a peptide ligand (in yellow sticks) (PDB ID: 3N8M). **(B)** Phage ELISA (enzyme-linked immunosorbent assay) showing the expression level and binding to a pY-peptide of the three insertion mutants of the Grb2 SH2 domain. ELISA signals (A_{450}) are

plotted as a function of the phage titer added to the wells of a micro plate coated with pY peptide (ErbB2 pY1139), anti-V5 antibody or a buffer. **(C)** The positions of the surface mutations in the Grb2 SH2 domain (PDB ID: 3N8M) (Delorbe et al., 2010) that were mutated to increase the protein solubility of pY-clamps are shown as red sticks. The C**α** atoms of original N and C termini, and the new N and C termini in the cpSH2 (i.e. Val123 and Val122, respectively) are shown as green spheres and labeled. The pY peptide in yellow is located behind the SH2 domain from the reader. Note that all the mutated positions are located on the side of the SH2 domain opposite to the peptide-binding interface. **(D)** Schematic drawings of the structure of FN3 domain from two perspectives (PDB ID: 1FNF) (Leahy et al., 1996). The three loops that were diversified to construct the combinatorial library, termed BC, DE and FG loops, are shown in blue, cyan and red, respectively, and labeled. **(E)** The amino acid sequences of the pY-clamps and cpSH2-FN**.** The residue numbering of these proteins starts from the N-terminal Gly residue generated after cleavage with TEV protease, and the same numbering is used for the crystal structures of the complexes of clamp(Ptpn11_pY580) and clamp(Shc1_pY239/240). The regions corresponding to the linker between cpSH2 and FN3 and the BC, DE and FG loops in FN3 are boxed in orange, blue, cyan, and red, respectively, and labeled. The original N and C termini (labeled "N" and "C", respectively) of the SH2 domain, and Val122 and Val123, where the new termini for cpSH2 were created, are indicated with green arrowheads. The SH2 residues mutated in this study are boxed in magenta and identified according to the wild-type Grb2 SH2 sequence. The F61Y and F62K mutations were introduced in all pY-clamp constructs. The F147L and F101A mutations were introduced in clamp(Bcr_pY177) via library selection in which these positions were randomized and clones resistant to thermal treatment were selected using yeast surface display. The N129D and W134R mutations, found in the final clamp(Bcr_pY177) clone, probably due to PCR errors during library construction, are indicated. The N129D was also introduced in clamp(Shc1_pY239/240) when higher solubility was required.



D

	pY-Clamp K_D (in nM)			
Target	cpSH2-FN	Ptpn11_pY580	Bcr_pY177	Shc1_pY239/240
PTPN11 pY580	7,800 ± 5,600	1.5 (χ^2 = 0.496)	9,300 ± 2,300	330 ± 70
BCR pY177	1,100 ± 130	200 ± 40	3.6 (χ^2=0.376)	750 ± 100
SHC1 pY239/240	520 ± 60	420 ± 110	5,400 ± 940	2.1 (χ^2=2.09)

Figure S2. Binding properties of pY-clamps, Related to Figure 1. (A) Specificity of the selected pY-clamps and cpSH2-FN tested using yeast surface display. The mean fluorescence intensities of peptide binding were normalized with respect to the mean fluorescence intensity of the expression level. Data with 5 nM targets are shown with those for the intended targets in red. Note that binding was nearly undetectable for most cases. cpSH2-FN, the fusion construct before directed evolution, exhibited no detectable binding to pY peptides at the concentration of 5 nM (*second column from left*), and thus data with 5,000 nM targets are also shown (*left-most*

column). The amino acid sequences of the pY peptides are shown with the core pY-X-N-X motif shaded in gray. **(B)** Size exclusion chromatography characterization showing monodispersity of the pY-clamps. The pY clamps purified from *Escherichia coli* were subjected to size exclusion chromatography on a Superdex 200 10/300 GL column, equilibrated with 150 mM NaCl, 20 mM Tris HCl buffer, pH 7.5 and protein concentrations as indicated. Elution positions for molecular mass standards including bovine serum albumin (67 kDa), myoglobin (18 kDa), cytochrome C (12 kDa) and aprotinin (6.5 kDa), are indicated at the top. *mAU*, milliabsorbance units. Note that Clamp(Bcr_pY177) and clamp(Shc1_pY239/240) exhibited monodispersed profiles. Their retention volumes were greater than expected for their masses most likely due to weak interactions between the proteins with the column matrix. Clamp(Shc1_pY239/240) contained the N129D mutation that improved its stability (see also Figure S1E). **(C)** Surface plasmon resonance sensorgrams for the binding of the pY-clamps to pY motifs. Derived K_D values are also shown. The interactions of the pY-clamps with their intended targets were analyzed using kinetic data, where sensorgrams are shown in black and the global fitting of the 1:1 Langmuir binding model in gray. The other interactions were analyzed using equilibrium responses, and fitting of the 1:1 binding model is shown as insets. (**D**) Dissociation constants (K_D) of clamp-peptide interactions measured by surface plasmon resonance. K_D ± standard error in nM and χ^2 of global fitting for kinetic analysis are shown. The indicated errors indicate the errors from the least-squares fitting of the 1:1 binding model. (**E**) SPR sensorgrams of the Ptpn11_pY580 peptide and its unphosphorylated counterpart flowed over immobilized Clamp(Ptpn11_pY580), showing that the binding is dependent on phosphorylation of the peptide.

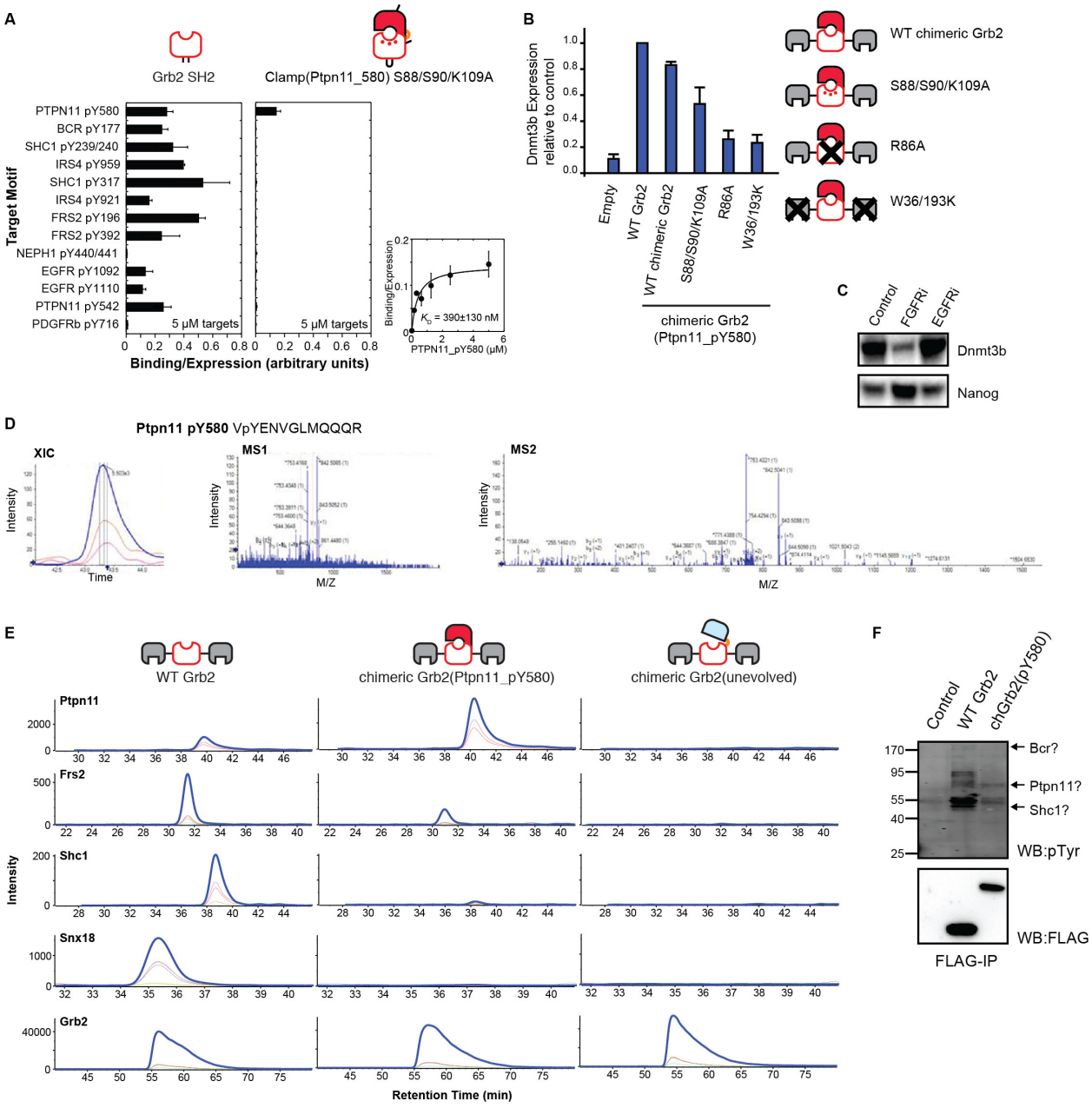


Figure S3. Properties of variants of chGrb2(Ptpn11_pY580) and evaluation of interactions of chimeric Grb2 proteins with cellular proteins, Related to Figures 3 and 4. (A) Yeast surface display analysis of the interactions of clamp(Ptpn11_pY580) containing the S88A, S90A and K109A mutations in the SH2 domain with the indicated pY-peptides. Binding signals with 5 μM targets normalized relative to the surface display level are plotted for the wild-type Grb2 SH2 (left) and for the mutant clamp. The right graph shows binding titration of clamp(Ptpn11_pY580) containing the S88A, S90A and K109A mutations with the Ptpn11_pY580 peptide. Note that the Binding/Expression values for data shown in this panel cannot be quantitatively compared with those in Figure S2A, because different lots of secondary reagents were used. **(B)** Dnmt3b levels in Grb2-/- mESCs following expression of WT Grb2 or chGrb2(Ptpn11_pY580) constructs, represented as the average ± standard deviation (n=3). **(C)** Immunoblot analysis of the effects of FGFR and EGFR inhibitors on mESC priming. WT mESCs were treated with vehicle control,

10μM SU5402 (FGFRi) or 3μM AG1478 (EGFRi) for 48 h before immunoblotting for Nanog and Dnmt3b. These treatments were previously shown to selectively inhibit the interaction of Grb2 with Ptpn11 and Shc1, respectively (see Findlay et al., 2013). **(D)** Mass-spectrometric identification of the Ptpn11 pY580 tryptic peptide in FLAG immunoprecipitates from chGrb2(Ptpn11_pY580) expressed in Grb2-/- mESCs. The extracted ion chromatogram (XIC), and fragmentation patterns for the precursor (MS1) and product ions (MS2) derived from phosphorylated pY580 peptide are shown. (**E**) The extracted ion chromatograms for proteins analyzed in Figure 4A. Numerical data are given in Table S5. (**F**) Immunoblot analysis of pY-containing proteins captured by WT Grb2 or chGrb2(Ptpn11_pY580) expressed in Grb2-/- mESCs.

Supplemental Tables

Table S1. Dissociation constants (K_D in nM) of pY-clamp/peptide interactions measured in the yeast display format, Related to Figure 1.

Target		pY-clamp			
	WT Grb2 SH2	cpSH2-FN	Ptpn11_pY580	Bcr_pY177	Shc1_pY239/240
PTPN11 pY580	1,270 ± 250	2,700 ± 2,300	**2.5 ± 0.5**	2,900 ± 800	270 ± 100
BCR pY177	3,220 ± 970	800 ± 430	102 ± 12	**0.42 ± 0.08**	130 ± 40
SHC1 pY239/240	1,690 ± 470	940 ± 470	960 ± 430	>10,000	**0.80 ± 0.17**
IRS4 pY959	2,370 ± 550	7,800 ± 8,100	6,200 ± 2,700	2,100 ± 610	1,900 ± 1,700
SHC1 pY317	1,090 ± 430	6,800 ± 5,700	4,800 ± 1,500	>10,000	1,700 ± 1,100
IRS4 pY921	1,280 ± 140	3,300 ± 2,100	600 ± 280	>10,000	540 ± 260
FRS2 pY196	710 ± 220	810 ± 240	464 ± 48	2,600 ± 1,200	1,350 ± 640
FRS2 pY392	750 ± 170	3,300 ± 2,700	680 ± 610	>10,000	730 ± 630
NEPH1 pY440/441	>10,000	>10,000	>10,000	>10,000	>10,000
EGFR pY1092	3,700 ± 590	3,800 ± 2,300	>10,000	>10,000	>10,000
EGFR pY1110	1,160 ± 450	4,100 ± 3,200	>10,000	>10,000	>10,000
PTPN11 pY542	2,540 ± 450	9,000 ± 5,300	420 ± 390	>10,000	1,800 ± 430
PDGFRb pY716	>10,000	>10,000	1,980 ± 100	>10,000	4,800 ± 2,200

* The indicated errors indicate the errors from the least-squares fitting of the 1:1 binding model.

Table S2. Dissociation constants (K_D) of pY-clamps to Ala-scanning peptides measured in the yeast display format, Related to Figure 2.

clamp(Ptpn11_pY580) to the PTPN11 pY580 motif		**clamp(Bcr_pY177) to the BCR pY177 motif**		**clamp(Shc_pY239/240) to the SHC1 pY239/240 motif**	
Mutation	K_D ± SD* (nM)	Mutation	K_D ± SD* (nM)	Mutation	K_D ± SD* (nM)
D575A	2.0 ± 0.02	E173A	3.9 ± 1.3	P234A	0.67 ± 0.12
S676A	2.7 ± 0.6	K174A	0.43 ± 0.04	P235A	0.55 ± 0.19
R578A	1.0 ± 0.08	P175A	2.5 ± 0.5	D236A	0.53 ± 0.09
V579A	3.8 ± 0.7	F176A	0.88 ± 0.38	H237A	1.0 ± 0.06
Y580A	>10,000	Y177A	>10,000	Q238A	0.58 ± 0.11
E581A	19 ± 2	V178A	1.9 ± 0.3	Y239A	5,750 ± 910
N582A	7600 ± 850	N179A	510 ± 110	Y240A	2,300 ± 1,100
V583A	410 ± 190	V180A	490 ± 130	N241A	3,000 ± 690
G584A	105 ± 17	E181A	780 ± 430	D242A	79 ± 9
L585A	8.0 ± 1.8	F182A	580 ± 190	F243A	1.2 ± 0.2
M586A	14 ± 1.8	H183A	930 ± 440	P244A	2.7 ± 0.3
Q587A	2.3 ± 0.4	H184A	1.4 ± 0.5	G245A	0.49 ± 0.09
Wild type	3.3 ± 0.7	Wild type	2.2 ± 0.4	K246A	0.59 ± 0.12
				Y239[a]	25,200 ± 3,800
				Y240[a]	760 ± 170
				Wild type	0.72 ± 0.16

*The indicated errors indicate the errors from the least-squares fitting of the 1:1 binding model.
[a]The Y239 and Y240 mutations of the SHC1 pY239/240 peptide lack the phosphorylation at 239 and 240, respectively. These data were acquired with chemically synthesized peptides.

Table S3. Data collection and refinement statistics for the crystal structures of pY-clamps, and their interface properties, Related to Figure 2.

	Clamp(Ptpn11_pY580)/ PTPN11 pY580 peptide complex	Clamp(Shc1_pY239/240)/ SHC1 pY239/240 peptide complex
Data collection statistics		
Space group	*C*2	$C222_1$
Cell parameters		
a, *b*, *c* (Å)	109.60, 34.04, 47.80	89.85, 92.19, 59.60
α, β, γ (°)	105.22	90.00
Beamline	APS-24ID-C	APS-23ID-B
Wavelength (Å)	0.9792	1.03316
Resolution (Å)	220-1.40 (1.45-1.40)	50.0-2.40 (2.44-2.40)
No. of reflections		
Observed	135,570	32,387
Unique	32,428 (3,095)	9,074 (375)
Completeness (%)	95.26 (92.8)	91.5 (76.7)
I/σ(I)	18.8 (3.5)	21.8 (3.5)
R_{merge} (%)	7.0 (41.3)	6.1 (31.4)
Refinement statistics		
Resolution range (Å)	52.88-1.4	32.17-2.41
Reflections used (test set)	32,418 (1653)	9059 (438)
R-factor (%)	17.4	21.3
R_{free} [1]	20.6	27.6
RMSD from ideal		
Bond lengths (Å)	0.006	0.009
Bond angles (°)	1.111	1.372
Protein residues	210	200
Water molecules	145	7
Mg^{2+}	1	0
Average B factor ($Å^2$)	24.2	44.0
Ramachandran plot statistics		
Favored (%)	96.6	93.1
Allowed (%)	3.4	6.4
Outliers (%)	0	0.5
Buried surface area[2] ($Å^2$)		
Peptide-pY clamp	1,558	1,571
Peptide-cpSH2 only	822	1,121
Surface complementarity[3] (*Sc*)	0.78	0.74

[1] R_{free} is the *R* factor for 5% of the reflections excluded from the refinement.

[2] Accessible surface areas buried in the interaction interfaces, which were calculated by using 'Protein interfaces, surfaces and assemblies' service PISA at the European Bioinformatics Institute (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html) (Krissinel and Henrick, 2007).

[3] Shape complementarity statistics calculated by using program CCP4 Sc (Lawrence and Colman, 1993).

Table S4. Complete list of Grb2 interacting proteins in Grb2-/- mESCs identified by IP-MS, related to Figure 3.

Gene ID	Gene Name	Protein ID	Number of Unique Peptides		
			WT Grb2	Grb2(Ptpn11_pY580)	Grb2(unevolved)
20662	Sos1	ENSMUSP00000067786	46	41	45
68524	Wipf2	ENSMUSP00000046991	22	22	27
13430	Dnm2	ENSMUSP00000072199	21	17	16
14784	Grb2	ENSMUSP00000021090	20	20	19
12402	Cbl	ENSMUSP00000041902	20	17	18
14388	Gab1	ENSMUSP00000034150	18	20	18
208650	Cblb	ENSMUSP00000110115	12	9	10
69710	Arap1	ENSMUSP00000081958	12	11	10
73178	Wasl	ENSMUSP00000031695	11	10	13
19247	Ptpn11	ENSMUSP00000058757	10	18	
330662	Dock1	ENSMUSP00000081531	9	2	
20663	Sos2	ENSMUSP00000044866	8	7	8
327826	Frs2	ENSMUSP00000020381	7	4	
140579	Elmo2	ENSMUSP00000073691	7	2	2
20416	Shc1	ENSMUSP00000103040	6		
170625	Snx18	ENSMUSP00000104864	5		
16331	Inpp5d	ENSMUSP00000044647	5	6	9
13429	Dnm1	ENSMUSP00000039033	3		
242915	Fam59b	ENSMUSP00000054208	3	1	
54324	Arhgef5	ENSMUSP00000031750	3	1	
13196	Asap1	ENSMUSP00000023008	3	6	10
19262	Ptpra	ENSMUSP00000028769	3		
215280	Wipf1	ENSMUSP00000092268	3	3	2
15505	Hsph1	ENSMUSP00000074392	3	2	4
72584	Cul4b	ENSMUSP00000059276	2		
381126	Fam59a	ENSMUSP00000048914	2	1	1
22631	Ywhaz	ENSMUSP00000022894	2	1	1
104015	Synj1	ENSMUSP00000023696	2	7	4
12469	Cct8	ENSMUSP00000026704	2		
N/A	N/A	ENSMUSP00000033594	2		
19656	Rbmxl1	ENSMUSP00000048153	1		6
268396	Sh3pxd2b	ENSMUSP00000044276	1	2	
16828	Ldha	ENSMUSP00000036386	1	2	
22630	Ywhaq	ENSMUSP00000100067		3	2
20779	Src	ENSMUSP00000029175		2	2
75273	Pelp1	ENSMUSP00000019065		2	
12465	Cct5	ENSMUSP00000022842		2	2
217843	Unc79	ENSMUSP00000082156		2	
27369	Dguok	ENSMUSP00000014698		2	
242481	Palm2	ENSMUSP00000030049			3
23983	Pcbp1	ENSMUSP00000054863			2
19701	Ren1	ENSMUSP00000092135			2
217866	Cdc42bpb	ENSMUSP00000042565			2

Proteins for which a minimum of 2 unique peptides were identified in at least one sample are shown. Rows shaded in blue show known ligands for SH3, those in red for SH2, and those in green for both SH3 and SH2. Grb2, which was the bait for IP, is shaded in yellow.

Table S5 (related to Figure 4. Extracted peptide ion transitions and integrated ion intensities used for the SWATH-MS analysis of the interactions of wild-type Grb2, clamp(Ptpn11_pY580) and clamp(unevolved) in mESC) is provided as a separate Excel file.

Supplemental Experimental Procedures

Antibodies and Chemicals. Anti-Nanog antibodies were from ReproCell Inc., anti-Dnmt3b antibodies from Imgenex, anti-Sos1 and anti-Shc1 antibodies from BD Biosciences, anti-Ptpn11, anti-Bcr and anti-Erk1/2 antibodies from Santa Cruz Biotechnology, anti-FLAG antibodies from Sigma-Aldrich and anti-ppErk and anti-Ptpn11 pY580 from Cell Signaling Technology. Anti-phospho-Shc (Tyr239/240) antibody (catalog number 2434) was from Cell Signaling Technology, which is known to cross-react with other pY-containing antigens (http://www.cellsignal.com/products/2434.html). Anti-phosphotyrosine (4G10) was generated in-house. Recombinant human FGF4 was from R&D Biosciences and SU5402 FGFR inhibitor was from Santa Cruz Biotechnology.

Preparation of phosphotyrosine (pY) peptides. We constructed genes for the target peptides fused to yeast small ubiquitin-like modifier (ySUMO) by utilizing a previously described expression vector (Huang et al., 2008). The genes encoded N-terminal His_{10} tag, a tobacco etch virus (TEV) protease cleavage site, ySUMO, a Cys-Gly linker and the 13-residue peptides. The ySUMO-fusion proteins were expressed in *E. coli* and purified by Ni-affinity chromatography as described previously (Huang et al., 2008). The ySUMO-peptide fusion proteins were treated with TEV protease to release the His_{10} tag. The His_{10} tag and TEV protease were then removed from the sample using Ni-affinity chromatography. The tag-cleaved ySUMO-peptide fusion proteins were biotinylated using maleimide-PEG_2-biotin (Thermo Scientific). The biotinylated ySUMO-peptide fusion proteins were phosphorylated with the kinase domain of EphA3 (produced from an expression vector provided by Dr. Sirano Dhe-Paganon, University of Toronto) in 20 mM Tris (pH 7.5), 5 mM $MgCl_2$, 5 mM $MnCl_2$ and 2 mM ATP by for 3 hrs at 37°C or for 16 hrs at 4°C. Essentially complete phosphorylation of the tyrosine residue in the peptide segment and of no other tyrosine residues in the fusion protein was verified by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). The Shc1 peptide prepared in this manner was fully phosphorylated at both Tyr residues. The reaction solution was then applied to a Ni-affinity column and followed by gel-filtration chromatography to remove the kinase and further purify the phosphorylated proteins.

The Ala-scanning mutants of the pY peptides were produced using a slightly different expression vector that encoded a His_6-tag and the Avi-Tag N-terminal to ySUMO. The peptides were fused C-terminal to ySUMO via a Gly_2 linker. These fusion genes were expressed in *E coli* BL21 (DE3) harboring an expression vector for BirA biotin ligase, and the proteins were purified by Ni-affinity chromatography. Phosphorylation of these proteins was carried out as described above. The phosphorylated and biotinylated ySUMO-peptide fusion proteins were further purified using a Resource Q anion-exchange column with a linear gradient of NaCl in 20 mM Tris HCl buffer (pH 7.0). Biotinylation and phosphorylation of ySUMO-peptide fusion proteins were verified to be essentially 100% by MALDI-TOF MS. Synthetic peptides were prepared using Fmoc (9-fluorenyl methoxycarbonyl) solid-phase chemistry on a Prelude Peptide synthesizer (Protein Technologies, Inc., Tucson AZ). Phosphotyrosine was incorporated using the N-fluorenylmethyloxycarbonyl-O-phospho-L-tyrosine derivative. Following purification by reverse-phase HPLC the authenticity of each peptide was confirmed by mass spectrometry

Design and generation of the circularly permutated Grb2 SH2 domain (cpSH2). General design strategies for creating affinity clamps have been described (Huang et al., 2008; Koide and Huang, 2013). The N and C termini of the Grb2 SH2 domain are located on the opposed side of the peptide-binding interface (Figures 1C and S1). Thus, in order to connect the FN3 domain to Grb2 SH2 domain such that the three diversified loops of FN3 are proximal to the peptide-binding site of the SH2 domain, we first needed to relocate the SH2 termini by circular permutation (Figure 1C). To identify a potential location for the new termini that does not substantially destabilize the SH2 domain, we first designed three mutant SH2 domains in which a four-glycine (G_4) linker into the loops proximal to the peptide-binding site using a synthetic gene for the human Grb2 SH2 domain. These mutants were then displayed on phage and their peptide-binding function was analyzed using phage ELISA (Figure S1B). The mutant in which the G_4 linker was inserted between Val122 and Val123 (the residue numbering according to that of the full-length human Grb2 protein, UniProt ID P62993) was displayed on the phage surface and exhibited the strongest binding to pY peptide of these mutants (Figure S1B). Therefore, we constructed a gene encoding a circular permutant with Val123 as the new N terminus and Val122 as the new C terminus (cpSH2) fused to FN3 with a GGSGGG linker (Figure S1E). We noted two surface-exposed Phe residues (F61 and F62) located on the opposite side of the peptide-binding site. Because they may contribute to aggregation, we replaced the two residues with Tyr and Lys, respectively (Figures S1C and S1E).

Construction and sorting of combinatorial yeast surface display libraries. Yeast surface display experiments were performed according to previously published procedures with modifications (Koide et al., 2007b; Koide et al., 2012). The DNA fragment encoding the cpSH2 domain containing the F60Y and F61K mutations described above, a GGSGGS linker and the FN3 domain was cloned into a yeast-display vector, pGalAgaCamR (Koide et al., 2007b), using the *Nco*I and *Xho*I sites. The resulting yeast display vector of cpSH2-FN was linearized by endonuclease digestion of a *Bam*HI site within the GGSGGS linker portion and of a *Xho*I site 3' to the FN3 segment. Yeast strain EBY100 was transformed with a mixture of the linearized vector and PCR-amplified DNA fragments for the genes encoding an FN3 monobody library (Wojcik et al., 2010). Correctly reconstituted vectors encoded the fusion protein for Aga2-cpSH2-FN-V5 tag in which FN3 loops are diversified. The transformants were selected in tryptophan-deficient media and the resulting library contained ~2×10^7 clones.

Sorting of a yeast display library by magnetic bead capture was performed as follows. The transformed yeast cells were grown in the SD-CAA media at 30°C for 24-48 hrs, and then protein expression was induced by growing the cells in the SG-CAA media at 30°C for 24 h, as described previously (Boder and Wittrup, 1997). The induced yeast cells were incubated with a biotinylated pY peptide whose concentration was set at ~1/10 × K_D or ~1/100 × K_D of cpSH2-FN in BSS-T buffer (137 mM NaCl, 8.10 mM Na_2HPO_4, 2.68 mM KCl, 1.47 mM KH_2PO_4, pH 7.4, 1 mg/ml bovine serum albumin and 0.1 % (w/v) Tween 20) at 4°C for 30-60 min. After washing with BSS-T, the cells were mixed with streptavidin-coated magnetic beads and incubated at 4°C for 30-60 min, and cells bound to the pY peptide were captured using a magnet stand. After washing with BSS-T, the yeast cells captured by the magnetic beads were used directly to inoculate YPD media and the cells were grown at 30°C.

Subsequent rounds of library sorting were performed using a BD FACSAria flow cytometer. Yeast cells displaying clamp proteins were prepared as described above. The cells were washed with BSS-T and incubated with a biotinylated pY peptide and anti-V5 antibody in BSS-T at 4°C for 30-60 min. Cells were then collected by centrifugation and washed with BSS-T before incubating with secondary reagents, streptavidin conjugated with PE and anti-rabbit IgG-DyLight 650 conjugate, at 4°C for 30-60 min. Cells were washed with BSS-T before FACS sorting. After several rounds of FACS sorting, individual clones were prepared and analyzed. Their amino acid sequences were deduced from DNA sequencing.

Affinity measurements using yeast display. Affinity measurement of the pY clamps using yeast display was carried out as described previously (Koide et al., 2012). K_D values were determined from plots of the mean fluorescent intensity against pY peptide concentration by fitting the 1:1 binding model using Igor Pro software (WaveMetrics).

Expression and Purification of pY clamps. The genes for pY clamps and cpSH2-FN were cloned in an expression vector, pHFT2 (Huang et al., 2006; Koide et al., 2007a), and expressed in *E. coli* BL21 (DE3) cells. This vector expresses a cloned gene with His_{10}, FLAG tag and a TEV cleavage site fused to the N-terminus. Protein expression was induced using auto-induction media for 24 h at 30°C (Studier, 2005). Proteins were purified with Ni-affinity chromatography following standard protocols. The N-terminal tag was cleaved with TEV protease and the cleaved protein was purified with Ni-affinity chromatography. For surface plasmon resonance experiments, the proteins were further purified on a Resource S cation-exchange column to further reduce the level of uncleaved proteins containing the His_{10}-tag.

Improvement of the solubility of clamp(Bcr_pY177) with surface mutations. Clamp(Bcr_pY177) in the original form showed limited solubility. Visual inspections of the Grb2 SH2 structure identified additional surface-exposed hydrophobic residues, Phe101 and Phe147. We diversified these positions with the NHK codon (where N = A/C/G/T, H = A/C/T and K = G/T, encoding Phe, Tyr, His, Lys, Asp, Glu, Asn, Gln, Ser, Thr, Leu, Ile, Met, Val, Ala and Pro) in the yeast display format to produce a library. The yeast cells harboring the library were incubated at 42°C for 45 min and, after cooling, a clone retaining binding to 1 nM BCR pY177 peptide was recovered that contained two additional substitutions, N129D in the cpSH2 domain and W134R in the BC loop of the FN3 domain, as likely due to errors introduced during PCR amplification (Figure S1E). This clone showed improved solubility and thus used for further characterization of clamp(Bcr_pY177).

Surface plasmon resonance (SPR) measurement. SPR measurements were carried out on a Biacore 3000 instrument. A ySUMO-pY peptide fusion protein, described above, was immobilized on a Ni-NTA chip in such a way that gave approximately 100 response units of immobilization. For measurement of the binding of clamp(Shc1_pY239/240) to pY-peptides, His_{10}-tagged ySUMO protein was immobilized on the control surface at approximately the same level as that of ySUMO-pY peptide fusion proteins on the main surface to prevent non-specific interaction of this clamp with the surface. Sensorgrams were collected as pY clamps were flowed at various concentrations in 10 mM HEPES (pH 7.4), 150 mM NaCl, 50 μM EDTA, and

0.005% (w/v) Tween20 at a flow rate of 30 µl min^{-1}. The kinetic data were analyzed using a 1:1 binding model with BIAevaluation (GE Healthcare Life Sciences). For weak interactions (K_D > 0.1µM), sensorgrams were collected at a flow rate of 5 µl min^{-1}, and K_D values were calculated from plots of equilibrium response values against pY clamp concentrations by fitting the 1:1 binding model using Igor Pro software (WaveMetrics).

X-ray crystallography. To purify the pY-clamps in complex with pY peptide, His-tagged ySUMO-peptide proteins were phosphorylated, treated with SUMO-hydrolase (Malakhov et al., 2004) to cleave the ySUMO portion and mixed with the tag-cleaved pY-clamps. The mixture of a pY-clamp, a pY-peptide, SUMO hydrolase and kinase was applied to a Ni-affinity column and the flow through fraction, corresponding to the pY-clamp/peptide complex, was collected. For clamp Ptpn11_580, further purification was carried out using Superdex 75 gel filtration chromatography (GE Healthcare Life Sciences) in 20 mM Tris (pH 7.5), 150 mM NaCl.

The clamp(Ptpn11_pY580)/PTPN11 pY580 peptide complex was concentrated to ~5 mg/ml and crystallized in 30% (w/v) PEG3000, 0.2 M $MgCl_2$, 0.1 M Tris (pH 8.5) at 20°C using hanging-drop vapor-diffusion method. The crystals were cryoprotected with a mixture of 80% mother liquid and 20% ethylene glycol and flash-frozen in liquid nitrogen. Clamp(Shc1_pY239/240) in complex with the SHC1 pY239/240 peptide was concentrated to ~10 mg/ml and crystallized in 1.66 M ammonium sulfate, 2.32% PEG400, 0.1 M HEPES (pH 7.5) at 20°C. The crystals were frozen in a mixture of 85% mother liquid and 15% glycerol as a cryoprotectant.

Diffraction data were collected at the Advanced Photon Source beamlines (Argonne National Laboratory). Data sets were processed with HKL2000 (Otwinowski and Minor, 1997). The structure of the clamp(Ptpn11_pY580)/PTPN11 pY580 peptide complex was determined at 1.4 Å by the molecular replacement method using the atomic coordinates of Grb2 SH2 (PDB code: 3N8M) (Delorbe et al., 2010) and a monobody without the loop regions (PDB code: 3UYO) (Koide et al., 2012) as search models. The program MOLREP in the CCP4 program suite (Vagin and Teplyakov, 2000) located one SH2 domain and one FN3 domain in the asymmetric unit. The clamp Shc1_239/240 complex was determined by molecular replacement using PHASER (McCoy et al., 2007) in CCP4 with the separated coordinates of the cpGrb2 SH2 domain and the FN3 domain from the Clamp Ptpn11_580 structure as search models. The models of engineered loops, linker and pY peptide were manually built into the electron density maps following molecular replacement. Manual model fitting was performed using Coot (Emsley and Cowtan, 2004), followed by refinement with REFMAC5 (Murshudov et al., 1997) and Phenix.refine (Adams et al., 2002). Data collection and refinement statistics are listed in Table S3. Molecular graphics were generated using PyMol (www.pymol.org). Solvent-accessible surface areas were calculated using PISA in CCP4 (Krissinel and Henrick, 2007).

mESC culture and transfection. mESCs were cultured under standard conditions on gelatin coated plates in media containing LIF, 10% ES Cell qualified Fetal Calf Serum (Gibco) and 5% Knockout Serum Replacement (Invitrogen). Grb2-/- (Cheng et al., 1998) and Fgf4-/- (Wilder et al., 1997) mESCs were described previously. Where applicable, mESCs were starved of LIF and FCS for 2 h prior to stimulation with recombinant FGF4.

The Polymerase chain reaction (PCR) was used to insert each pY clamp into full length Grb2 replacing the wild-type SH2 domain. These chimeric Grb2 genes and wild-type Grb2 were cloned in to an expression vector driven by the beta-actin promoter (pCAGGS; Findlay et al. 2013). The monitored expression levels were similar to endogenous Grb2 levels found in mESCs (Figure 3A). Grb2-/- mESCs were transfected using Lipofectamine LTX (Invitrogen) according to the manufacturer's instructions, and selected with puromycin for 48 h before lysis in 1X Cytometric Bead Array assay buffer (BD Biosciences). Nanog and Dnmt3b levels in transfected mESCs were determined by immunoblotting and subsequent quantification. Immunostaining with Nanog and Dnmt3b antibodies was carried out according to the manufacturer's instructions, prior to analysis by confocal microscopy.

Immunoprecipitation for immunoblotting. A 10-cm dish of transfected Grb2-/- mESCs per sample was lysed in FLAG IP-MS lysis buffer (Bisson et al., 2011), and triple-FLAG tagged proteins immunoprecipitated for 2 h with 10 µl of FLAG-M2 Agarose (Sigma-Aldrich). Immunoprecipitates were washed with 3 x 1 ml lysis buffer prior to immunoblot analysis.

Immunoprecipitation for mass-spectrometry. Two 15 cm dishes of Grb2-/- mESCs transfected with full-length chGrb2 proteins per sample were starved of LIF and FBS for 2 h to maintain autocrine signaling, lysed in FLAG IP-MS lysis buffer (Bisson et al., 2011), and triple-FLAG tagged proteins immunoprecipitated for 2 h with 25 µl of FLAG-M2 Agarose (Sigma-Aldrich). Immunoprecipitates were washed with 3 x 1 ml lysis buffer followed by 2 x 1 ml of 20 mM Tris-HCl, pH 7.5. Proteins were eluted with 50 mM phosphoric acid, and subjected to solid-phase trypsin digestion. Briefly, SCX 30 µm beads (PolyLC Inc.) were packed into a TopTip column (Glygen Corp), samples loaded and washed with 10 mM KH_2PO_4 pH 3.0 then 30 µL HPLC-grade water. Samples were reduced using TCEP (Sigma-Aldrich) then alkylated and digested using sequencing-grade trypsin (Promega) and 10 mM iodoacetamide (Sigma-Aldrich) for 2 h at 37°C. Digested peptides were eluted with 200 mM ammonium bicarbonate pH 8.0, acidified using formic acid, and stored at -40°C until mass spectrometric analysis.

SWATH-MS analysis. All data were acquired on a TripleTOF 5600 instrument (AB SCIEX) in line with a NanoLC-Ultra 1D plus system (Eksigent). Chromatographic separation was carried out using a trap and elute workflow on a cHiP-LC column 75 µm x 15 cm (Eksigent) packed with ReproSil-Pur 3 µm C18-AQ (Dr. Maisch HPLC GmbH). Peptides were bound on the trap chip at 2% acetonitrile (ACN) with 0.1% formic acid and eluted over an 85 min gradient increasing to 30% ACN. The MS1 TOF survey scan was set from 400 to 1300 Da with an accumulation time of 250 ms at 30 K resolution, 20 candidate ions (50 ms each) were selected for collision induced dissociation (CID) per cycle of 2.3 s and former precursor ions were excluded for 15 s. The product ion MS/MS scan range was set between 100 and 1250 Da with an accumulation time of 50 ms at ~15K resolution. Following peptides injection (~250 ng), each SWATH acquisition was followed by ACN column washes and re-equilibration with a standard tryptic mixture of 30 fmol BSA (Bruker-Michrom) and 60 fmol α-casein (Sigma-Aldrich) to minimise carryover. Prior to SWATH sample acquisition, a TOF MS calibration was done on 10 selected BSA and α-casein peptides; MS/MS calibration was done for one of the 10 peptides to ensure mass accuracy between samples. Each SWATH acquisition records consecutive high-resolution MS/MS spectra

of all precursor ions that filter through 33 sequential windows of 25 amu spanning the mass range of 100-1250 amu using 100 ms for each SWATH (+ 250 ms TOF scan) for 3.4 s cycle time on a 2-30% organic 85 min liquid chromatography gradient.

All of the spectral and chromatographic data from all identified Grb2 binding proteins obtained were consolidated into a spectral library using ProteinPilot (AB SCIEX) and UniProtKB/Swiss-Prot protein database and were imported into Peakview (1.2.01 research version, AB SCIEX). Peakview correlated both peptide identification and LC retention times to extract specific MS/MS transition data for each peptide with an assigned confidence of ≥95%. The top 3 abundant ion transitions from the top ranked peptides for each protein were used to retrieve quantitative data (in counts/sec) using a 0.05 amu extraction width over a ±5 min LC time and visualised using MarkerView (AB SCIEX). For each individual sample, the same peptide transitions were summed into peptides and the peptides were summed into proteins. After normalizing each sample/condition to the amount of Grb2, the extracted ions for selected proteins were visualized using principle components analyses within Markerview.

Supplemental References

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr *58*, 1948-1954.

Bisson, N., James, D.A., Ivosev, G., Tate, S.A., Bonner, R., Taylor, L., and Pawson, T. (2011). Selected reaction monitoring mass spectrometry reveals the dynamics of signaling through the GRB2 adaptor. Nat Biotechnol *29*, 653-658.

Boder, E.T., and Wittrup, K.D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. Nat Biotechnol *15*, 553-557.

Cheng, A.M., Saxton, T.M., Sakai, R., Kulkarni, S., Mbamalu, G., Vogel, W., Tortorice, C.G., Cardiff, R.D., Cross, J.C., Muller, W.J.*, et al.* (1998). Mammalian Grb2 regulates multiple steps in embryonic development and malignant transformation. Cell *95*, 793-803.

Delorbe, J.E., Clements, J.H., Whiddon, B.B., and Martin, S.F. (2010). Thermodynamic and Structural Effects of Macrocyclization as a Constraining Method in Protein-Ligand Interactions. ACS Med Chem Lett *1*, 448-452.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.

Huang, J., Koide, A., Makabe, K., and Koide, S. (2008). Design of protein function leaps by directed domain interface evolution. Proc Natl Acad Sci U S A *105*, 6578-6583.

Huang, J., Koide, A., Nettle, K.W., Greene, G.L., and Koide, S. (2006). Conformation-specific affinity purification of proteins using engineered binding proteins: Application to the estrogen receptor. Protein Expr Purif *47*, 348-354.

Koide, A., Gilbreth, R.N., Esaki, K., Tereshko, V., and Koide, S. (2007a). High-affinity single-domain binding proteins with a binary-code interface. Proc Natl Acad Sci U S A *104*, 6632-6637.

Koide, A., Tereshko, V., Uysal, S., Margalef, K., Kossiakoff, A.A., and Koide, S. (2007b). Exploring the capacity of minimalist protein interfaces: interface energetics and affinity maturation to picomolar K_D of a single-domain antibody with a flat paratope. J Mol Biol *373*, 941-953.

Koide, A., Wojcik, J., Gilbreth, R.N., Hoey, R.J., and Koide, S. (2012). Teaching an Old Scaffold New Tricks: Monobodies Constructed Using Alternative Surfaces of the FN3 Scaffold. J Mol Biol *415*, 393-405.

Koide, S., and Huang, J. (2013). Generation of high-performance binding proteins for Peptide motifs by affinity clamping. Methods Enzymol *523*, 285-302.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J Mol Biol *372*, 774-797.

Lawrence, M.C., and Colman, P.M. (1993). Shape complementarity at protein/protein interfaces. J Mol Biol *234*, 946-950.

Leahy, D.J., Aukhil, I., and Erickson, H.P. (1996). 2.0 A crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and synergy region. Cell *84*, 155-164.

Malakhov, M.P., Mattern, M.R., Malakhova, O.A., Drinker, M., Weeks, S.D., and Butt, T.R. (2004). SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. J Struct Funct Genomics *5*, 75-86.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*, 658-674.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta. Crystallogr. D. Biol. Crystallogr. *53*, 240-255.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods Enzymol. *276*, 307-326.

Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. Protein Expr Purif *41*, 207-234.

Vagin, A., and Teplyakov, A. (2000). An approach to multi-copy search in molecular replacement. Acta Crystallogr D Biol Crystallogr *56*, 1622-1624.

Wilder, P.J., Kelly, D., Brigman, K., Peterson, C.L., Nowling, T., Gao, Q.S., McComb, R.D., Capecchi, M.R., and Rizzino, A. (1997). Inactivation of the FGF-4 gene in embryonic stem cells alters the growth and/or the survival of their early differentiated progeny. Dev Biol *192*, 614-629.

Wojcik, J., Hantschel, O., Grebien, F., Kaupe, I., Bennett, K.L., Barkinge, J., Jones, R.B., Koide, A., Superti-Furga, G., and Koide, S. (2010). A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain. Nat Struct Mol Biol *17*, 519-527.