

Supplemental Material for the Manuscript:

Structural characterization of CAS SH3 domain selectivity and regulation reveals new CAS interaction partners

Authors: Jakub Gemperle¹, Rozálie Hexnerová², Martin Lepšík², Petr Tesina², Michal Dibus¹, Marian Novotný¹, Jan Brábek¹, Václav Veverka^{*2}, Daniel Rosel^{*1}

* Correspondence regarding NMR analysis should be directed to Vaclav Veverka (veverka@uochb.cas.cz), correspondence regarding other matters should be directed to Daniel Rosel (rosel@natur.cuni.cz)

¹Department of Cell Biology, Faculty of Science, Charles University in Prague, Czech Republic

²Institute of Organic Chemistry and Biochemistry of the CAS, Czech Republic

Supplemental figures and tables

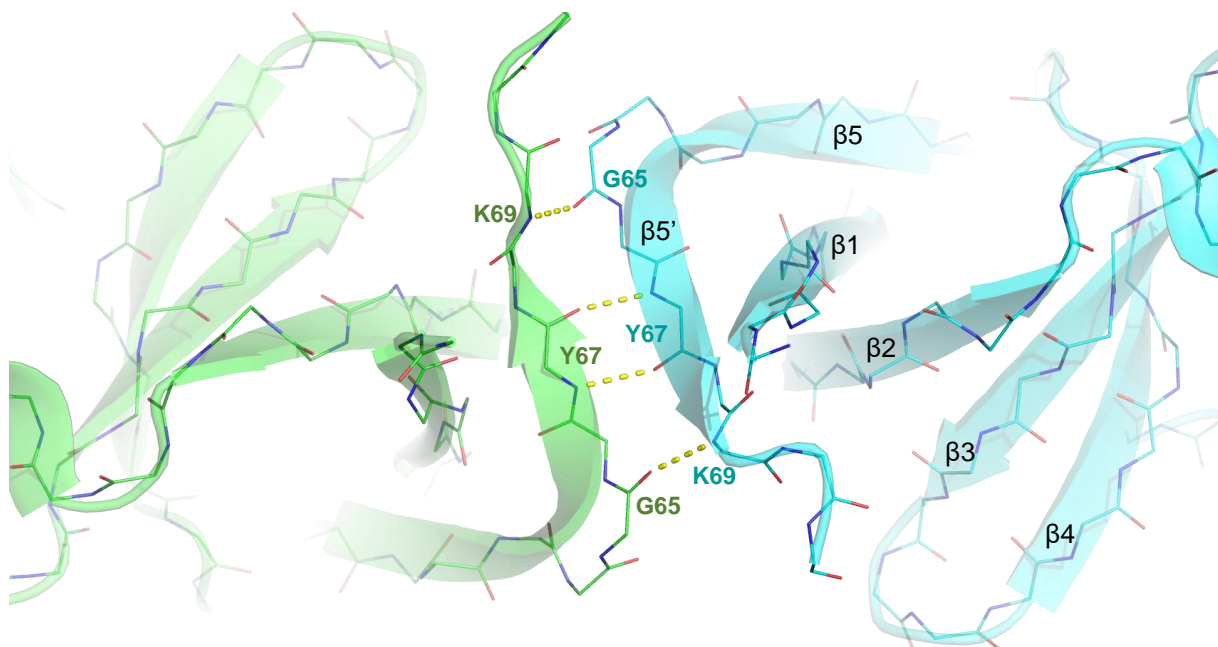


Figure S1. The crystallographic dimer (PDB code 1WYX) is stabilized by the crystal contacts that stimulate the formation of an additional two-stranded (labeled $\beta 5'$) beta-sheet at the dimer interface. The beta-sheet is stabilized by a network of four intra-main chain hydrogen bonds (highlighted in yellow).

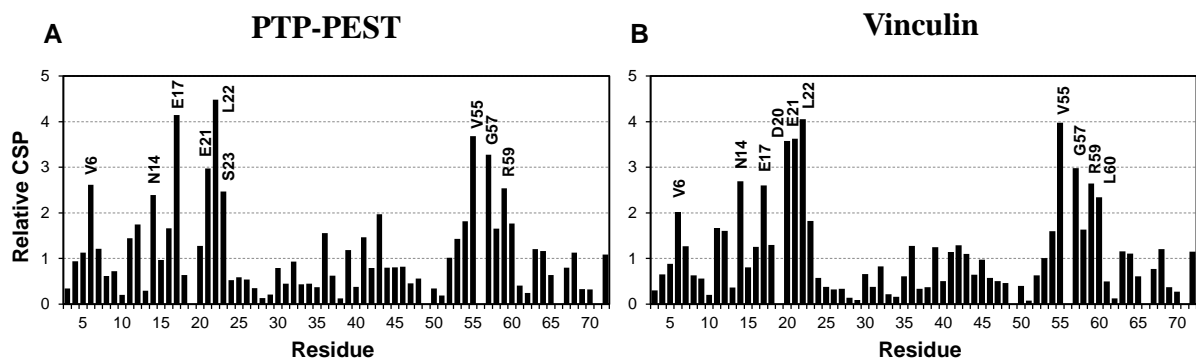


Figure S3. Relative chemical shift perturbations (CSP) of CAS SH3 induced by binding of **A** PTP-PEST peptide and **B** Vinculin peptide. The values were calculated from changes in positions of backbone amide resonances and divided by standard deviation. The distribution of induced changes is similar for both peptides, corresponding to highly overlapping binding regions derived from the structural data obtained for both complexes.

A

	Position
	...-10123456789...
FAK/FRNKb	PDPAAPPKKPPRPGAPGH
PYK2b	LEFTGPPQKPPRLGAQSI
C3G	TDEEVAPPKPPPLPGIRVV
Vinculin	LTDELAPPKPPPLPEGEVP
FAK/FRNKa	GGSD E APPKPSRPGYPSP
PYK2a	TAFQ E PPPKPSRPKYRPP
PTP-PEST	EKQD S PPPKPPRTRSCLV
MICA1	DPE M PPPKPPRSCSALA
	* **

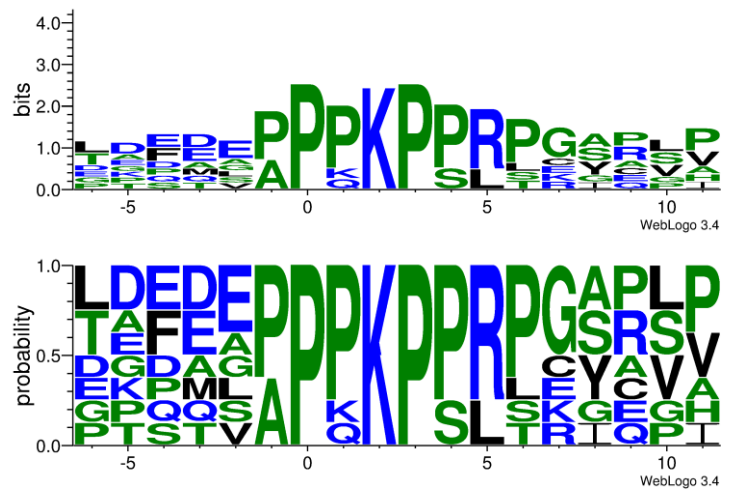
B

Figure S4. Enrichment of negatively charged amino acids within CAS SH3 binding ligands. **A** Multiple sequence alignment of eight high-affinity CAS SH3 binding sequences derived from known six human CAS SH3 binding proteins using CLUSTAL W. **B** Prolonged high-affinity CAS SH3 binding motif derived from sequences belonging to known human CAS SH3 interacting proteins. The x-axis shows the positions in relation to proline (position 0). The y-axis in the upper panel shows the bits. The graphs show the amino acid preference near the proline. The height of the letters is proportional to the frequency in the top eight protein sequences.

	10	12	14	17	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	59																								
CAS/P56945/9-60	-KALYD	V	A	E	S	-P	D	E	L	S	F	R	K	G	D	I	M	T	V	L	E	Q	D	T	Q	g	g	G	W	W	L	C	S	L	H	---	G	R	Q	G	I	V	P	G	N	L		
HCLS1/P14317/434-482	-VAVYDY	Q	G	E	G	-S	D	E	L	S	F	D	P	D	D	V	I	T	D	I	E	M	V	D	E	---	G	W	R	G	R	C	H	---	G	H	F	G	L	F	P	A	N	Y	V			
NCF2/P19878/246-294	-RVLFGF	V	P	E	T	-K	E	E	L	Q	V	M	P	G	N	I	V	F	V	L	K	K	G	N	D	---	N	W	A	T	V	M	F	N	---	G	Q	K	G	L	V	P	C	N	Y	L		
SH3D19/Q5HYK7/501-552	-VVLHDF	P	A	E	Q	-V	D	D	L	N	L	T	S	G	E	I	V	L	L	E	K	I	D	T	---	D	W	Y	R	G	N	C	R	---	N	Q	I	G	I	F	P	A	N	Y	V			
SH3D19/Q5HYK7/577-625	-VARFEY	I	G	E	Q	-K	D	E	L	S	F	S	E	G	E	I	I	L	K	E	Y	V	N	E	---	E	W	A	R	G	E	V	R	---	G	R	T	G	I	F	P	L	N	F	V			
SH3D19/Q5HYK7/667-715	e-ALHSF	T	A	E	T	-S	D	D	L	S	F	K	R	G	D	R	I	Q	I	L	E	R	L	D	S	---	D	W	C	R	G	R	L	Q	---	R	E	G	I	F	P	A	V	E	F			
SNX9/Q9Y5X1/6-57	-RVMYD	F	A	A	E	P	g	N	N	E	L	T	V	N	E	G	E	I	I	T	I	T	N	P	D	V	g	G	W	L	E	G	R	N	I	K	---	G	E	R	G	L	V	P	T	D	Y	V
SH3GL1/Q99961/312-360	-KALYD	F	E	P	E	N	-D	G	E	L	G	F	H	E	G	D	V	I	T	L	T	N	Q	I	D	E	---	N	W	Y	E	G	M	L	D	---	G	Q	S	G	F	F	P	L	S	Y	V	
CAP/Q9BX66/873-923	-IAKFN	F	N	G	E	T	-Q	V	E	M	S	F	R	K	G	E	R	I	T	L	L	R	Q	V	D	E	---	N	W	Y	E	G	R	I	P	G	T	-S	R	Q	G	I	F	P	I	T	Y	V
VINEXIN/O60504/460-510	-VAQYTF	K	G	L	-E	V	E	L	S	F	R	K	G	E	H	I	C	L	I	R	K	V	N	E	---	N	W	Y	E	G	R	I	T	G	T	-G	R	Q	G	I	F	P	A	S	Y	V		
SORBS2/O94875/944-994	-IAKYNF	N	A	D	T	-N	V	E	L	S	L	R	K	G	D	R	V	L	L	K	R	V	D	Q	---	N	W	Y	E	G	K	I	P	G	T	-N	R	Q	G	I	F	P	V	S	Y	V		
ASAP1/Q9ULH1/1073-1124	k-TIYD	C	A	N	-D	E	L	T	F	I	E	G	E	V	I	V	T	G	E	E	D	Q	---	E	W	W	I	G	H	I	E	G	Q	p	E	R	K	G	V	F	P	V	S	E	F			
ABI1/Q8IZP0/452-500	-VAIYD	Y	T	K	D	-D	E	L	S	F	M	E	G	A	I	I	V	I	K	K	N	D	---	G	W	Y	E	G	V	C	N	R	---	V	T	G	L	F	P	G	N	Y	V					

Figure S5. Alignment of human SH3 domains that are likely to bind a centrally located lysine in the ligand. Numbers indicate the primary amino acid sequence of the CAS SH3 domain. The alignment includes 13 unique SH3 domains from 11 proteins that include either Glu (yellow column) or Asp (green column) at the position equivalent to Glu17 in CAS protein and a conserved Glu that corresponds to CAS Glu21 (orange column). Unique sequence features of CAS, namely Asn14, Leu40 and Arg59, are also highlighted. Amino acids in small letters indicate positions in the alignment that are occupied in a small proportion of sequences.

A

Prot. name	Sequence	Interaction with SH3 p130Cas
C15oRf62	A ₇₈ PPKPPRL	?
C1oRf168	P ₂₇₈ PPKPSRP	?
CEP2	A ₁₂₆ PPKPPRL	?
DOK7	P ₃₁₇ PPKPLRP	?
FBN2	P ₃₃ PPKPPRP	?
GLIS2	P ₃₃₃ PPKPLLIP	?
MICAL1	P ₈₂₉ PPKPPRS	Known
PKN3	P ₅₀₉ PPKPPRL	?
PRAM1	P ₅₀₁ PAKPLLIP	?
FAK	A ₇₁₁ PPKPSRP	Known
PYK2	P ₇₁₃ PPKPSRP	Known
PTP-PEST	P ₃₃₃ PPKPPRT	Known
C3G	A ₂₆₆ PPKPLLIP	Known
SETD5	P ₆₀₆ PAKPSRP	?
Vinculin	A ₈₅₉ PPKPLLIP	Known
WNK1	A ₁₇₅₀ PSKPLLIP	?

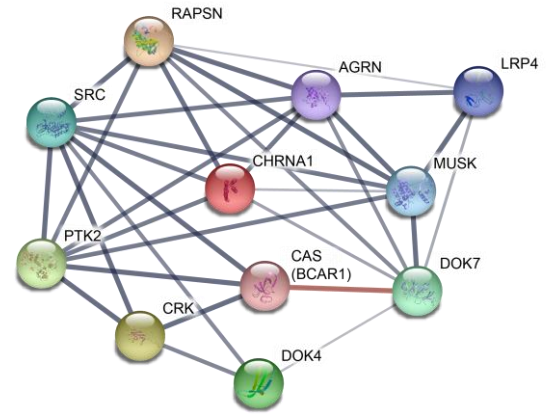
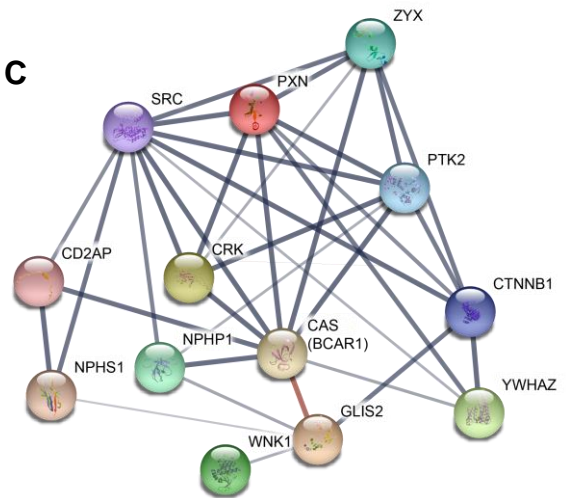
B**C**

Figure S6. Potential CAS/BCAR1 interaction partners. **A** Table of known and predicted CAS SH3 interacting proteins based on CAS SH3 binding motif. Six have already been shown to bind CAS. In proteins highlighted by gray boxes, we were able to find connections to the signaling processes involving CAS. **B** Signaling network of proteins that are important for acetylcholine receptor clustering in skeletal muscle with added CAS. The red line represents the new **CAS-DOK7** interaction. **C** Signaling network of proteins implicated in regulation of kidney function. The red line represents the new **CAS-GLIS2** interaction. Crosstalk is illustrated with STRING 10 (<http://version.10.string-db.org/>). Line thickness indicates the strength of data support, a red line represents a newly verified interaction.

CAS SH3

Amino acid sequence after thrombin cleavage:

GSMKYLNVLAKALYDNVAESPDELSFRKGDIMTVLERDTQGLDGWWLCSLHGRQGIVPGNRLKILVGMYDKKPAG^{EFIVTD}

CAS SH3 Vinculin

Amino acid sequence after thrombin cleavage:

GSMKYLNVLAKALYDNVAESPDELSFRKGDIMTVLERDTQGLDGWWLCSLHGRQGIVPGNRLKILVGMYDKKPAG^{SGGSGS}
GLTDELAPPKPLPEGEV

CAS SH3 PTP-PEST

Amino acid sequence after thrombin cleavage:

GSMKYLNVLAKALYDNVAESPDELSFRKGDIMTVLERDTQGLDGWWLCSLHGRQGIVPGNRLKILVGMYDKKPAG^{SGGSGS}
GSEKQDSPPP KPPRTRSCL

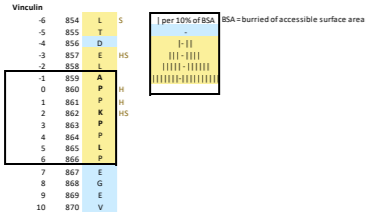
Figure S7. Primary sequences of the CAS SH3 domain chimeras. The first two amino acids (GS) are residuals from the thrombin cleavage site. Blue indicates the extended sequence from cloning vector ended by introduction of a stop codon from *EcoRI* restriction site. Red represents the inserted sequence of the linker. Sequences from the ligands are underlined.

Supplementary table 1. NMR restraints and structural statistics

	<i>SH3</i>		<i>PTP-PEST</i>		<i>Vinculin</i>	
<i>Non-redundant distance and angle constraints</i>						
Total number of NOE constraints	1165		1998		2483	
Short-range NOEs						
Intra-residue ($i = j$)	208		426		494	
Sequential ($ i - j = 1$)	305		468		585	
Medium-range NOEs ($1 < i - j < 5$)	122		183		267	
Long-range NOEs ($ i - j \geq 5$)	530		921		1137	
Torsion angles	100		80		88	
Hydrogen bond restrains	-		-		-	
Total number of restricting constraints	1265		2078		2571	
Total restricting constraints per restrained residue	17.6		23.9		28.9	
<i>Residual constraint violations</i>						
Distance violations per structure						
0.1 – 0.2 Å	3.45		9.25		8.40	
0.2 – 0.5 Å	2.22		3.65		3.80	
> 0.5 Å	0		0		0	
r.m.s. of distance violation per constraint	0.02 Å		0.02 Å		0.02 Å	
Maximum distance violation	0.50 Å		0.50 Å		0.50 Å	
Dihedral angle violations per structure						
1 – 10 °	2.33		1.23		0.9	
> 10 °	0		0		0	
r.m.s. of dihedral violations per constraint	0.51 °		0.46 °		0.31 °	
Maximum dihedral angle violation	5.00 °		5.00 °		4.90 °	
<i>Ramachandran plot summary from Procheck</i>						
Most favoured regions	96.4%		92.3%		93.1%	
Additionally allowed regions	3.6%		7.5%		6.7%	
Generously allowed regions	0.0%		0.1%		0.1%	
Disallowed regions	0.0%		0.0%		0.0%	
<i>r.m.s.d. to the mean structure</i>						
	<i>ordered</i> ¹	<i>all residues</i>	<i>ordered</i> ¹	<i>all residues</i>	<i>ordered</i> ¹	<i>all residues</i>
All backbone atoms	0.3 Å	6.7 Å	0.3 Å	2.7 Å	0.2 Å	2.4 Å
All heavy atoms	0.9 Å	6.9 Å	0.7 Å	3.0 Å	0.5 Å	2.6 Å

¹ Residues with sum of phi and psi order parameters > 1.8

[CAS SH3] Vinculin structure number	01 Wt	02 Wt	03 Wt	04 Wt	05 Wt	06 Wt	07 Wt	08 Wt	09 Wt	10 Wt	11 Wt	12 Wt	13 Wt	14 Wt	15 Wt	16 Wt	17 Wt	18 Wt	19 Wt	20 Wt	21 Wt	22 Wt	23 Wt	24 Wt	25 Wt	26 Wt	27 Wt	28 Wt	29 Wt	30 Wt	31 Wt	32 Wt	33 Wt	34 Wt	35 Wt	36 Wt	37 Wt	38 Wt	39 Wt	40 Wt						
6 LEU 83	0.00	19.58	0.00	0.00	51.71	0.00	0.00	11.89	0.00	0.00	0.00	0.00	0.00	11.96	0.00	0.00	28.28	0.00	0.00	25.44	0.00	0.00	3.56	0.00	0.00	0.00	0.00	0.00	4.46	42.67	5.94	0.00	14.32	0.00	0.00	8.53	64.78	7.29	4.66	0.00						
-4 THR 84	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	18.15	0.00	18.41	0.00	2.46	0.67	18.92	0.00	0.00	0.00	0.00	0.00	28.76	32.54	4.40	0.00	0.00	0.00	0.00	0.00	1.60	35.11	78.17	0.00	0.00					
-3 ASP 85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
-2 LEU 87	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
-1 ALA 88	51.01	49.59	51.88	54.55	48.00	49.60	53.04	53.99	49.60	52.58	50.00	44.60	54.78	50.73	49.84	50.01	53.90	53.81	48.60	50.90	55.27	52.66	48.71	49.33	49.33	51.94	50.64	49.28	51.62	57.00	48.09	52.29	48.52	51.48	48.11	48.74	50.49	47.88	50.55	0.00	0.00					
0 PRO 89	76.75	78.14	79.81	86.40	81.58	77.83	77.42	80.52	81.25	73.43	80.48	83.54	82.11	74.07	80.39	78.75	79.23	82.41	78.83	80.12	83.47	80.60	80.12	81.03	81.03	77.21	81.90	77.16	85.89	80.63	81.39	85.03	82.26	81.48	77.69	78.86	82.70	78.31	75.31	0.00	0.00					
1 PRO 90	18.08	25.55	17.65	20.91	26.66	25.67	23.65	24.31	21.73	17.66	23.25	24.12	23.22	23.87	22.11	24.55	24.14	25.08	21.87	19.88	24.85	25.08	21.20	21.20	21.20	19.36	20.66	24.34	23.91	20.66	24.34	22.68	26.07	17.00	18.78	24.25	22.18	22.62	16.92	22.22	0.00	0.00				
2 LYS 91	122.64	128.59	126.02	123.19	128.57	129.77	128.75	129.66	136.52	138.04	125.81	109.73	126.13	125.48	134.22	125.81	128.70	127.68	127.37	132.06	127.64	126.81	126.81	126.81	126.81	119.65	125.36	130.33	130.75	127.07	131.04	125.52	128.54	134.20	135.20	137.71	123.18	133.98	126.87	126.87	0.00	0.00				
3 PRO 92	77.91	82.32	79.01	78.13	82.14	78.13	78.13	79.41	78.68	78.69	82.15	79.24	78.46	77.66	82.51	78.50	79.17	79.76	83.90	80.37	80.37	80.37	80.37	80.37	80.37	80.37	79.77	83.90	78.70	80.37	79.47	77.20	73.11	79.63	81.36	80.16	81.02	79.12	85.76	77.63	82.84	79.53	0.00	0.00		
4 PRO 93	6.12	6.39	5.52	6.61	6.72	5.76	6.01	6.01	6.01	6.37	7.21	6.14	5.15	5.88	6.50	5.65	6.50	5.76	6.47	5.88	6.57	7.15	5.64	6.57	6.57	6.57	6.57	6.57	6.57	6.57	6.57	6.57	6.57	6.57	6.57	6.57	6.57	6.57	6.57	6.57	6.57	6.57	6.57			
5 LEU 94	90.84	95.51	90.50	93.67	92.84	97.55	94.86	93.34	94.02	92.53	93.00	94.50	97.36	96.32	99.01	92.36	91.51	99.00	90.16	93.89	97.19	95.20	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	90.50	
6 PRO 95	1.43	1.53	1.45	1.50	1.49	1.52	1.43	1.50	1.48	1.49	1.51	1.40	1.38	1.42	1.48	1.51	1.44	1.58	1.44	1.50	1.53	1.52	1.45	1.41	1.41	1.39	1.47	1.45	1.43	1.49	1.48	1.41	1.49	1.48	1.41	1.49	1.48	1.42	1.49	1.34	1.39	1.39	1.39	1.39	1.39	1.39
7 GLU 96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	14.60	9.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
8 GLY 97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
9 GLU 98	27.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
10 VAL 99	0.86	15.89	0.00	0.00	0.00	0.00	0.00	5.40	0.00	0.00	0.00	0.00	0.00	51.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				



Supplementary table S6. Interface analysis of CAS SH3-Vinculin (ligand part). Hydrogen bonds and nonpolar contacts between CAS SH3 and its ligands were analyzed with PSA with default settings. Prior to analysis, the linker sequence (residues 75-82) was removed from the chimeric proteins. Solvent accessible residues are colored yellow and inaccessible residues are blue. ΔG represents free energy of dissociation (dissociation energy effect, kcal/mol). BSA means Buried Surface Area (Å²) and is displayed with bars, one bar per 10%. HSOC indicates residues that contain the across-interface hydrogen bond, salt bridge, or disulfide bond atoms. The corresponding table cells have a brown background and contain letter H for a hydrogen bond, S for a salt bridge, and HS for a combination of hydrogen bond and salt bridge. http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html

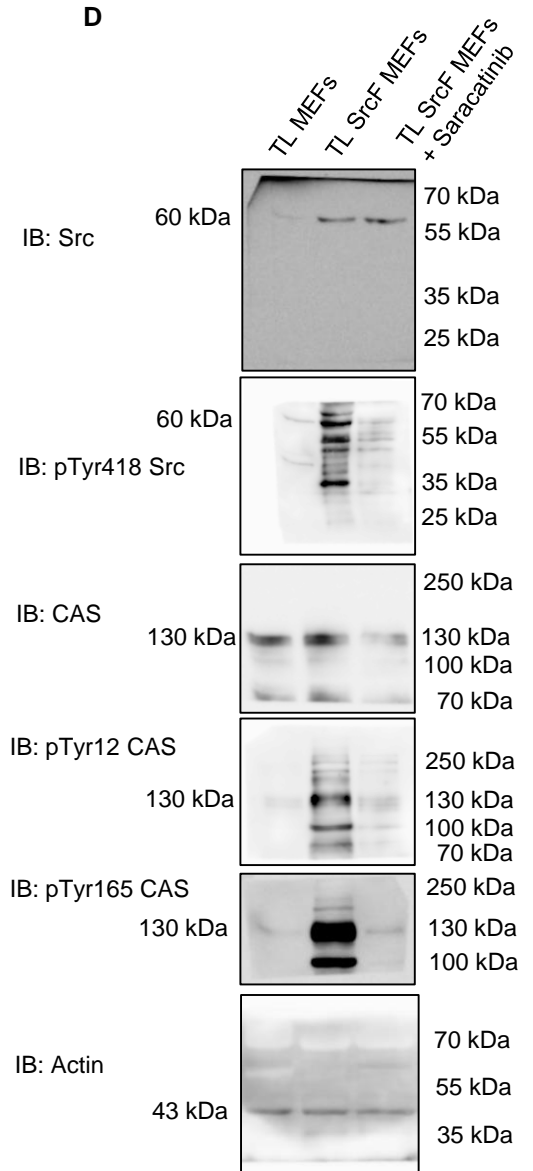
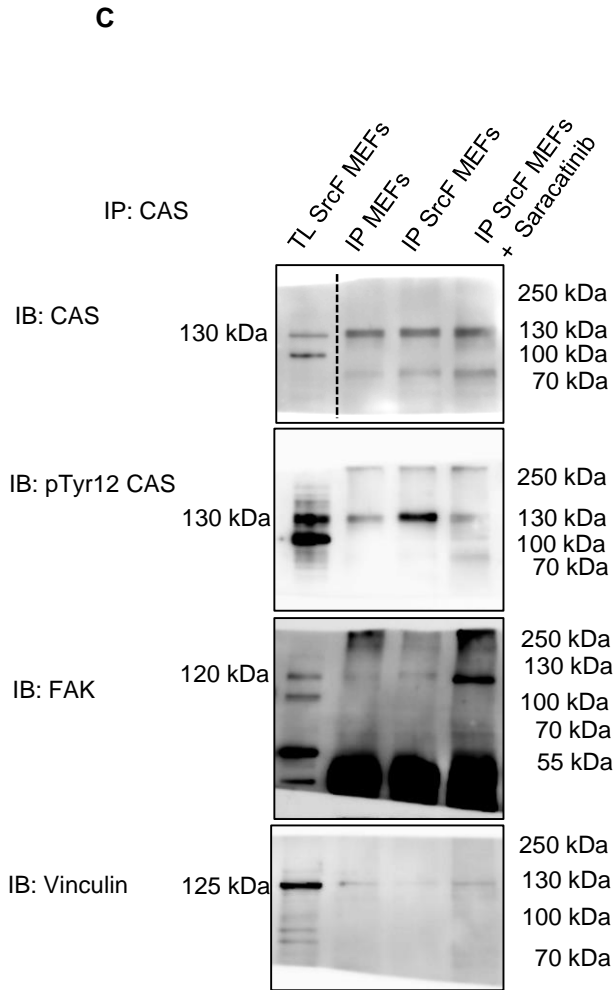
CAS_SH3-ligand_variant	Int.ene	Disp	H-bond
CAS-PTP_PEST-A12	-115.3	-55.8	-24.5
CAS-PTP_PEST-E12	-114.5	-58.3	-25.2
CAS-PTP_PEST-F12	-110.7	-60.7	-24.0
CAS-PTP_PEST-Y12	-124.6	-61.9	-30.8
CAS-PTP_PEST-pY12	-103.2	-57.8	-25.9

CAS_SH3-ligand_variant	Int.ene	Disp	H-bond
CAS-Vinculin-A12	-89.4	-56.5	-19.0
CAS-Vinculin-E12	-93.1	-61.8	-19.6
CAS-Vinculin-F12	-102.8	-63.5	-19.0
CAS-Vinculin-Y12	-104.1	-63.5	-19.1
CAS-Vinculin-pY12	-83.7	-57.5	-18.5

Supplementary table S8. Interaction energies – Quantum mechanics.

Interaction energies (Int. ene) and their dispersion (Disp) and hydrogen-bonding (H-bond) components of CAS vs. PTP-PEST/Vinculin binding for the wild-type, Tyr12 phosphorylated, and Tyr12 mutated CAS variants. All values are in kcal/mol.

Uncropped immunoblots for Figure 7:



Uncropped immunoblots for Figure 8:

