

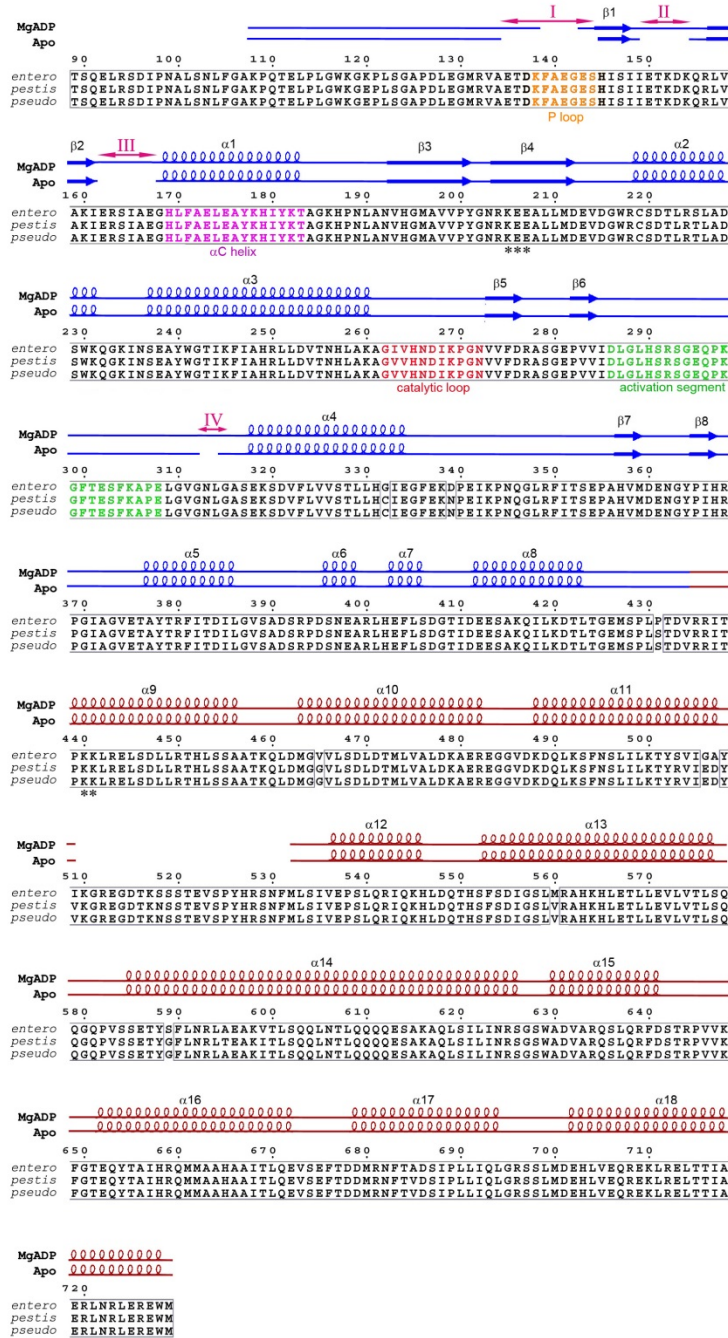
## SUPPLEMENTARY INFORMATION

### Mechanisms of *Yersinia* YopO kinase substrate specificity

*Wei Lin Lee, Pavithra Singaravelu, Sheena Wee, Bo Xue, Khay Chun Ang, Jayantha Gunaratne, Jonathan M Grimes, Kunchithapadam Swaminathan and Robert C Robinson*

<b>Data collection</b>	
Space group	P2 <sub>1</sub>
Cell dimensions	
a, b, c (Å)	108.73, 121.75, 118.58
( $\alpha$ , $\beta$ , $\gamma$ ) (°)	90.0, 94.9, 90.0
Resolution (Å)	20.0-2.93 (2.97-2.93)
R <sub>merge</sub>	0.125 (0.778)
I/ $\sigma$ (I)	11.9 (2.7)
Completeness (%)	99.9 (99.8)
Redundancy	6.8 (6.6)
<b>Refinement</b>	
Resolution (Å)	2.93-14.69 (2.93-2.97)
No. reflections	64,212
Rwork / Rfree	19.6/24.0
No. atoms	
Protein	14992
Ligand/ion	120
Water	9
B-factors overall (Å <sup>2</sup> )	75.56
Protein (Å <sup>2</sup> )	75.72
Ligand/ion (Å <sup>2</sup> )	56.03
Water (Å <sup>2</sup> )	57.93
R.m.s deviations	
Bond lengths (Å)	0.006
Bond angles (°)	1.272

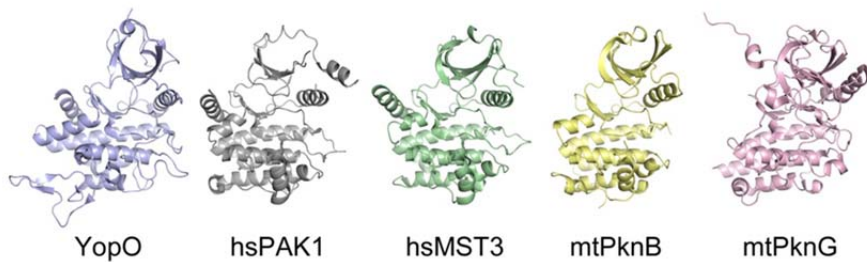
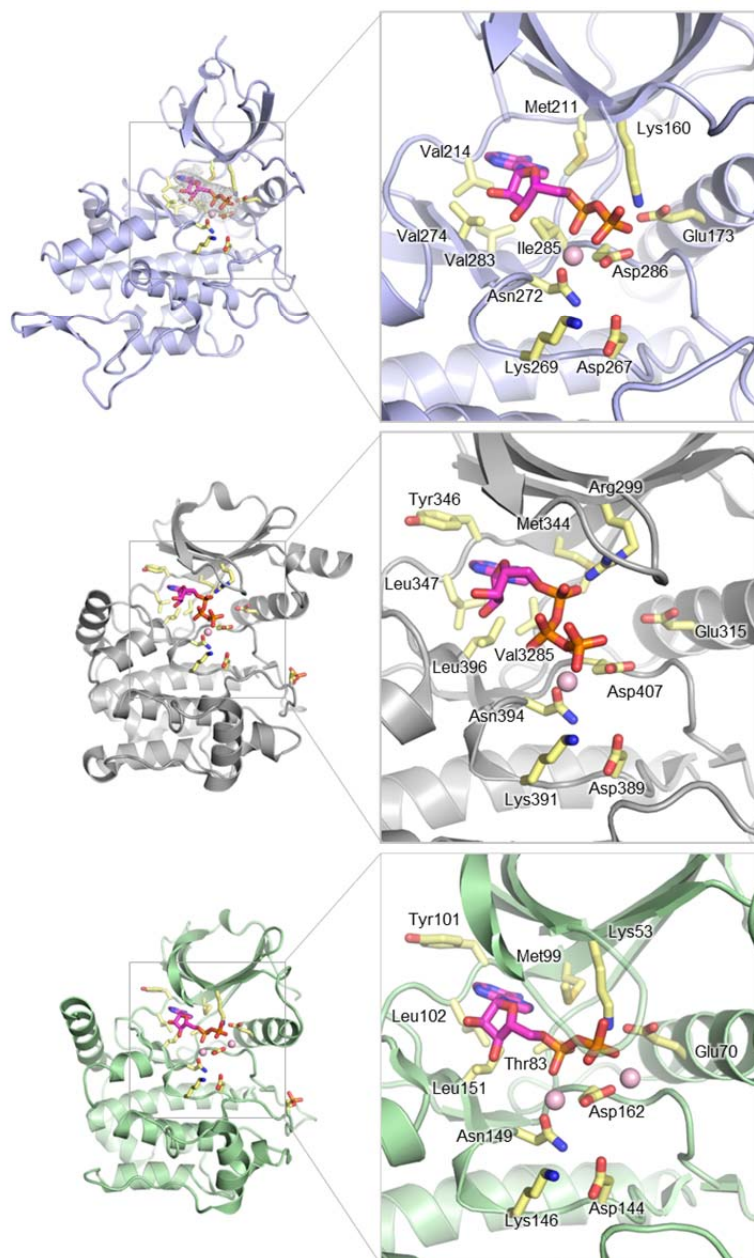
**Supplementary Table 1: Table of statistics associated with X-ray crystallographic analysis of the YopO-actin:MgADP complex.** Data collected from a single crystal. Values in parentheses are for the highest-resolution shell. The Rfree is calculated using 5% of the data.



**Supplementary Figure 1: The secondary structure alignment of the apo (PDB: 4CI6) and MgADP bound structures of the YopO-actin complex shown alongside the sequence alignment of YopO from different species of *Yersinia*. *Entero*, *pestis* and *pseudo* denote *Y. enterocolitica* (Uniprot Q93KQ6-1), *Y. pestis* (Uniprot Q9RI12-1) and *Y. pseudotuberculosis* (Uniprot Q05608-1), respectively. Residues marked with solid blue and yellow squares are involved in the interaction with actin and Rac1, respectively. Residues marked with solid black and grey squares are involved in interactions with Mg<sup>2+</sup> and ADP, respectively. The regions that were disordered in the apo structure, but are ordered in the MgADP structure, are numbered and labeled with pink arrows.**

**a**

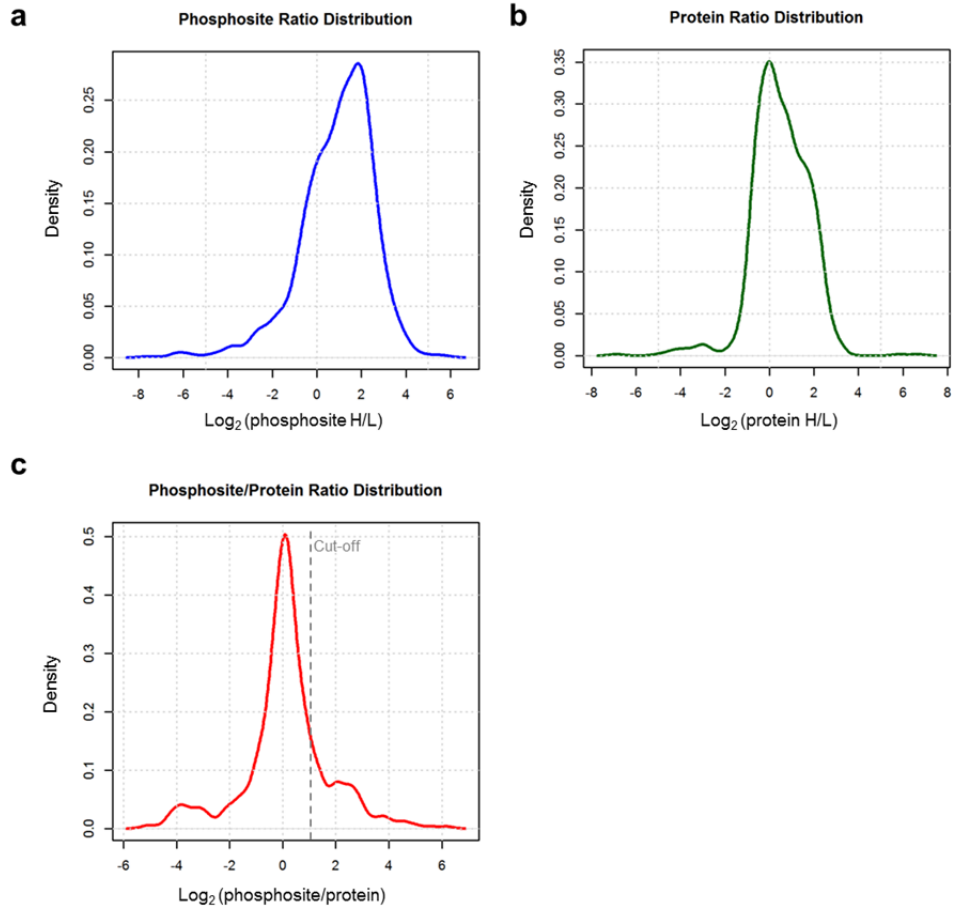
Species	PDB	Chain	Z	rmsd	lali	nres	%id	Description
<i>Homo sapiens</i>	4ZY5	A	23.5	2.6	244	281	16	PAK1 in complex with an inhibitor compound 17
	4QMX	A	23.4	2.8	250	288	17	MST3 in complex with saracatinib
<i>Mycobacteria tuberculosis</i>	1MRU	A	20.4	2.8	228	322	17	PknB in complex with ATP- $\gamma$ -S
	4Y0X	A	19.4	2.9	237	322	18	PknG in complex with ADP

**b****c**

**Supplementary Figure 2: Identification of kinases structurally similar to YopO using Dali. (a)**

Table of structurally similar neighbours, showing two of each of the mammalian and bacterial kinases with the highest *Z* scores. *Z* score, the statistical significance of the similarity between the protein-of-interest and other neighbourhood proteins; rmsd, root mean square distance, root-mean-square deviation of C-alpha atoms in the least-squares superimposition of the structurally equivalent C-alpha atoms; lali, the number of structurally equivalent residues; nres, the total number of amino acids in the hit protein; %id, percentage of identical amino acids over structurally equivalent residues. (b) Visual comparison of structurally similar kinases. hs, *homo sapiens*; mt, *mycobacteria tuberculosis*. (c) Structural comparison of the catalytic clefts of YopO, PAK1 (PDB: 3Q53) and MST3 (PDB: 3A7J). The  $F_o-F_c$  electron density map at 3.0 Å resolution, contoured at 3.0  $\sigma$ , before MgADP was built is shown at the active site of the YopO kinase domain. ADP in YopO and MST3 and ATP in PAK1 are represented as sticks. Catalytic residues and residues that line the hydrophobic cleft are also represented as sticks. Mg (YopO and PAK1) and Mn (MST3) ions are shown as spheres (magenta).





**Supplementary Figure 4: Distributions of quantified (a) phosphosites, (b) proteins and (c) the phosphosite/protein ratio.** Phosphorylation sites above the peptide/protein ratio of 2 (equivalent to the log<sub>2</sub> value of 1) were considered as putative YopO phosphorylation sites.

	Position (mouse)	Position (human)	Modified sequence	Phosphosite H/L ratio normalized	Protein Ratio H/L normalized	Protein Ratio count	Phosphosite/protein ratio
<b>Filament elongator</b>							
EVL	S335	S337-p	SNSVEKPV <sup>(ph)</sup> SLLSR	31.57	0.94	14	33.45
	S356	S358-p	SPLQS <sup>(ph)</sup> QPHSR	3.45	0.94	14	3.65
	S360	S362-p	SPLQSQPHS <sup>(ph)</sup> R	3.70	0.94	14	3.92
VASP	S2	S2	(ac)S <sup>(ph)</sup> ETVICSSR	11.37	0.83	46	13.76
	S285-p	S289-p	ATQVGEKPPKDES <sup>(ph)</sup> ASEESEAR	1.88	0.83	46	2.27
	S317-p/ S318-p/ S319-p/ S320-p	S322-p/ S323-p/ S324-p/ S325-p	S <sup>(ph)</sup> <sup>#</sup> S <sup>(ph)</sup> <sup>#</sup> S <sup>(ph)</sup> <sup>#</sup> S <sup>(ph)</sup> <sup>#</sup> VTTSEAHPTPCSSDDSLER	1.67	0.83	46	2.01
DIAPH1	S1214	S1231	AGCAVTS <sup>(ph)</sup> LLASELTK	8.13	0.62	20	13.14
INF2	S1186	S1158-p	TEADSTSEGPEDEAQRGQS <sup>(ph)</sup> THLPR	3.54	0.54	9	6.53
SHOT1	S249-p	S249-p	RQS <sup>(ph)</sup> HLLQSSLPDQQLLK	11.43	1.09	4	10.51
<b>Actin turnover protein</b>							
CAP1	S254-p	S255-p	S <sup>(ph)</sup> ALFAQINQGESITHALK	12.14	1.53	57	7.93
	S22/ T24/ S25	S22/ T24/ S25-p	LEAVS <sup>(ph)</sup> <sup>#</sup> HT <sup>(ph)</sup> <sup>#</sup> S <sup>(ph)</sup> <sup>#</sup> DMHCGYGDSPSK	7.76	1.53	57	5.07
<b>Signalling adaptor</b>							
ABI1	S183-p	S183-p	TNPPTQKPPS <sup>(ph)</sup> PPVSGR	3.45	1.65	8	2.10
EPS8	S57	S57-p	DSVS <sup>(ph)</sup> SVSDVSQYR	8.57	3.35	11	2.56
FNBP1L (TOCA1)	S488-p / S489-p and T496-p /S501-p	S488-p / S489-p and T496-p /S501-p	RHS <sup>(ph)</sup> <sup>#</sup> S <sup>(ph)</sup> <sup>#</sup> DINHLVT <sup>(ph)</sup> <sup>#</sup> QGRES <sup>(ph)</sup> <sup>#</sup> PEGSYTDANQEV	15.39	5.75	4	2.67
<b>Scaffolding protein</b>							
MTSS1	S275-p/ S276-p	S271/ S272-p	S <sup>(ph)</sup> <sup>#</sup> S <sup>(ph)</sup> <sup>#</sup> VCSSLNSVNSSDSR	5.41	0.77	6	7.04
	S279-p	S275-p	SSVCS <sup>(ph)</sup> SLNSVNSSDSR	10.86	0.77	6	14.12
PDLIM5 (ENH)	T110 / S111-p	T110-p / S111-p	EVVKVPVITS <sup>(ph)</sup> PAVSK	13.38	5.56	4	2.41
WIPF1	S330-p	S340-p	NLS <sup>(ph)</sup> LTSSAPPLPSPGR	1.66	0.60	20	2.77
<b>Adherens junction protein</b>							
PTPN6 (SHP-2)	S534-p/ Y536-p	S534-p/ Y536-p	GQES <sup>(ph)</sup> <sup>#</sup> EY <sup>(ph)</sup> <sup>#</sup> GNITYPPAVR	9.95	3.55	5	2.80
RAVER1	S524/ S525-p	gap	IPLNPYLNHSLLLPS <sup>(ph)</sup> <sup>#</sup> S <sup>(ph)</sup> <sup>#</sup> NLAGK	21.16	1.32	9	16.03
	S637-p/ S641-p	gap	S <sup>(ph)</sup> <sup>#</sup> GGS <sup>(ph)</sup> <sup>#</sup> GGPLSHFYSGSPTS <sup>(ph)</sup> YFTSGLQAGLK	6.87	1.32	9	5.20
	S696	S531-p	AVGSSPMGSSEGLLGLGPGNGHS <sup>(ph)</sup> HLLK	9.36	1.32	9	7.11
	S709/ S711	gap	RS <sup>(ph)</sup> <sup>#</sup> FS <sup>(ph)</sup> <sup>#</sup> HLLPSPEPSPEGSYVGGHSQGLGGHYADSYLK	8.05	1.32	9	6.12
<b>DNA/RNA-binding protein</b>							
AHNAK	S217-p	S216-p	LPSGSGPAS <sup>(ph)</sup> PTTGSVAVDIR	3.24	1.50	7	2.16
EIF1	S2-p	S2-p	(ac)S <sup>(ph)</sup> AIQNLHSFDPFADASK	5.77	0.86	12	6.71
SF1	S14	S14-p	(ac)ATGANATPLDFPS <sup>(ph)</sup> K	9.79	1.29	14	7.57
	S272	S272	S <sup>(ph)</sup> ITNTTVCTK	9.10	1.29	14	7.03
UBAP2	S1018-p	S1004-p	GVS <sup>(ph)</sup> VSSGTGLPDMTGSVYNK	10.14	2.58	6	3.93
UBAP2L	S629-p	S609-p	RYPSSISSS <sup>(ph)</sup> PQK	6.49	2.54	15	2.55

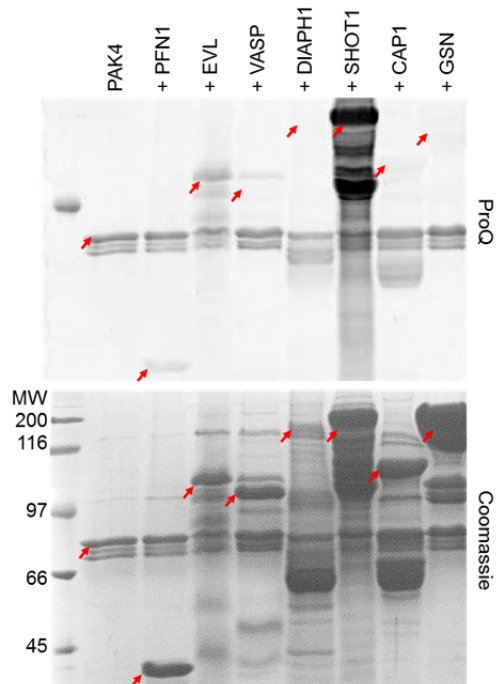
	S872-p/ S875	S852-p/ S855-p	DGS(ph) #LAS(ph) #NPYSGDLTK	9.64	2.54	15	3.79
<b>ZBP1</b>	T356-p	T371	GEVMESGDSEEPKKEDTGTSEAT(ph)PPR	7.72	1.96	7	3.95
	S384-p	S405	AMALGDSS(ph)PQTTEPVLR	5.04	1.96	7	2.57

**Supplementary Table 2: Phosphorylated peptides identified with KISS.** Identified peptides were subjected to cutoffs of mass spectrometry identification score of 44, protein ratio count of 4 and peptide/protein ratio of 2. All cysteines are carbamidomethylated. (ac) denotes acetylation and (ph) denotes phosphorylation. # denotes ambiguous identification in MS due to close proximity of serine/threonine pair.



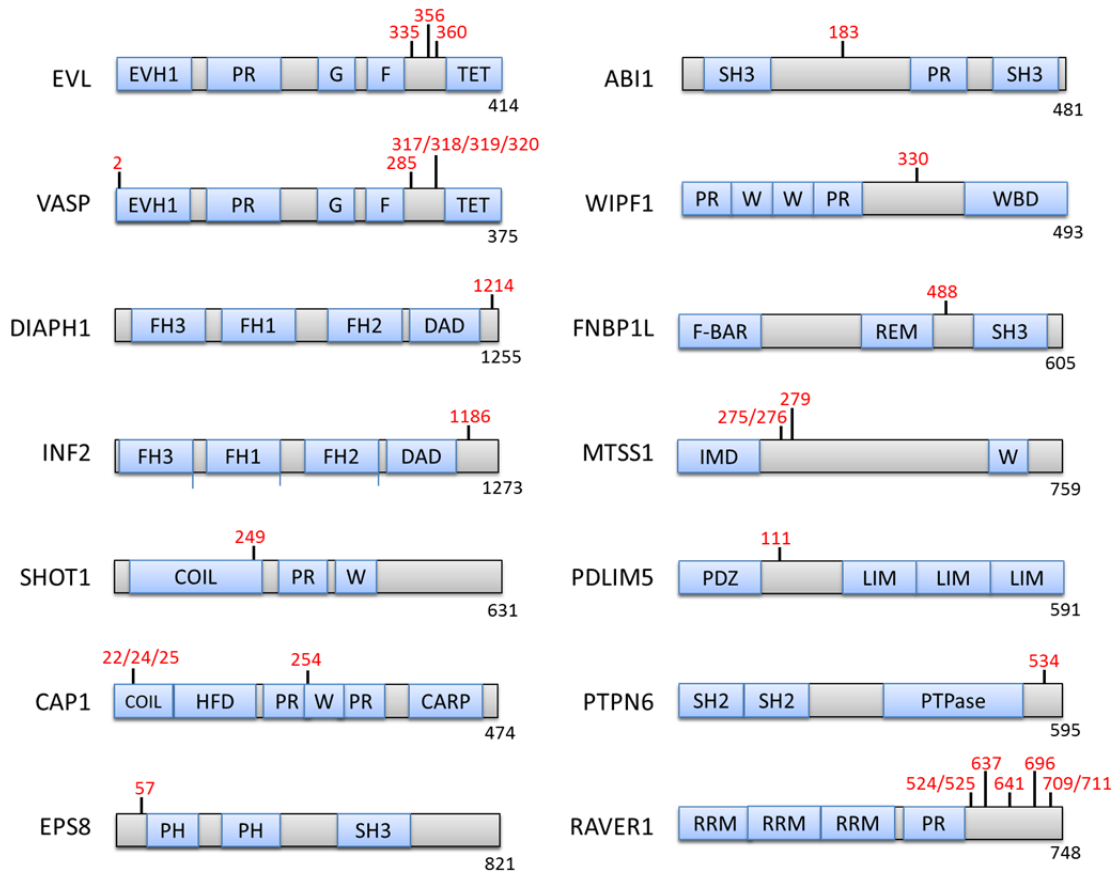
EVL		VASP		CAP1		SHOT1		DIAPH1	
Position in protein		Position in protein		Position in protein		Position in protein		Position in protein	
Mouse#	Human	Mouse#	Human	Mouse	Human#	Mouse#	Human	Mouse	Human#
T19/S20	T19/S20	S2	S2	S22	S22	T14	T14	S844	S861
S64	S64	T4	T4	S48	S49	S101	S101	S942	S959
S243	S245	S8	S8	S90	S91	P124/A125	S124/T125/	T1213	T1230
S249	S251	S9	S9	S254	S255	/T126/T128	T126/T128	S1214	S1231
S251	S253	S34	S34	S265	S266	S191	S191	S1218	S1235
S257	S259	S46	S46	S337	S338	S249	S249	T1242	T1259
S294	S296	P145	S149			S255	S255		
S300	S302	S241	S245			S256	S256		
S316/S317	S318/S319	S254	S258			S351	S351		
S327	S329	S285	S289			S375	S375		
S335	S337	S298	S303			S467	S467		
S356	S358	S300	S305			S473	S473		
S360	S362	S309	S314			S493	S493		
		T310	S315			T496	T496		
		S317/S318	S322/S323			S502	S502		
		/S319/S320	/S324/S325			S504	S504		
						S506	S506		
						S512	S512		
						T551/N552	T551/S552		
						/P553	/S553		
						S561	S561		
						S586	S586		
						S619	S619		

**Supplementary Table 3: Phosphorylated peptides identified by mass spectrometry analysis of *in vitro* phosphorylated YopO substrates.** Serines or threonines separated by “/” were too close in proximity to differentiate from the mass spectra to identify which site(s) was phosphorylated. Human EVL (Uniprot Q9UI08), VASP (Uniprot P50552) and SHOT1 (Uniprot A0MZ66) and mouse CAP1 (P40124) and DIAPH1 (Uniprot O08808) (residues 583–1255) were used. To facilitate comparison with the phosphorylation sites reported in the literature, corresponding residues in their mouse or human counterparts are shown alongside and denoted with a (#). Shown in grey are corresponding residues in mouse or human which are not conserved.

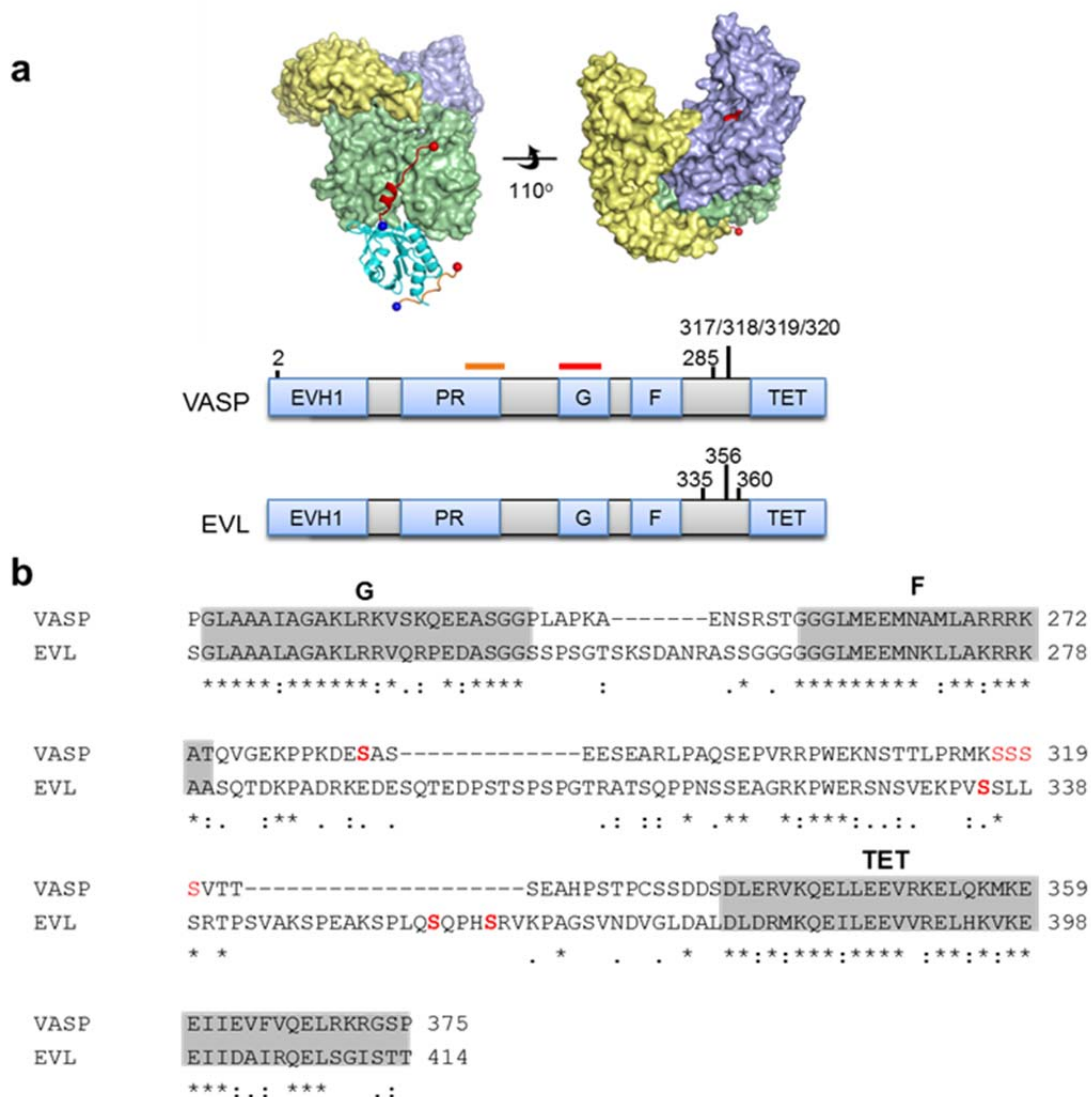


**Supplementary Figure 5: *In vitro* phosphorylation of actin-binding proteins by PAK4.**

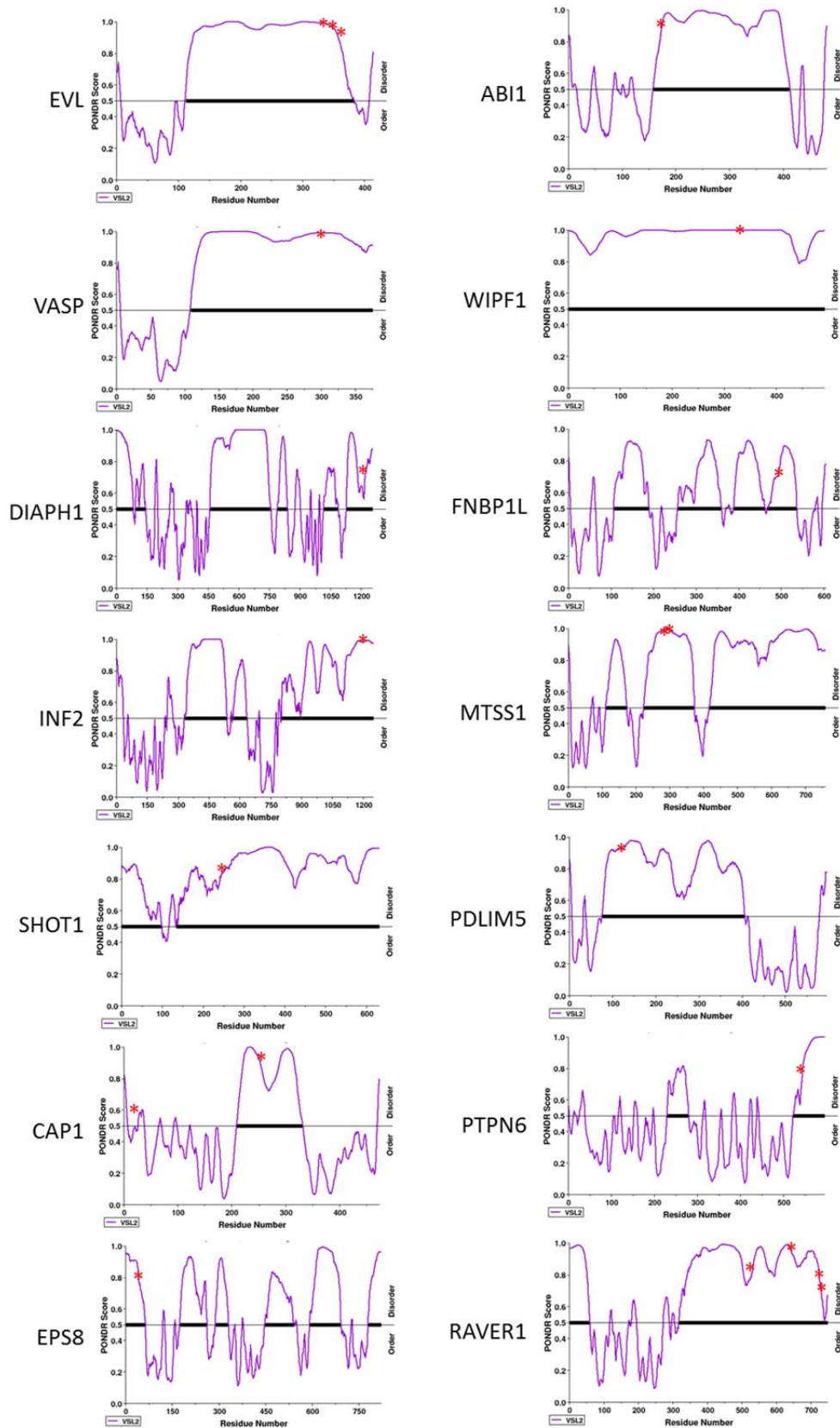
The SDS-PAGE was imaged with ProQ Diamond phosphoprotein staining for phosphorylation followed by Coomassie staining for total protein. 3.8  $\mu\text{M}$  of PAK4 was incubated with 15.2  $\mu\text{M}$  of the respective substrates in 50 mM Hepes pH 7.5, 1 mM ATP, 50 mM KCl, 10 mM  $\text{MgCl}_2$  and 1 mM DTT at 30°C for 30 mins. PFN1, profilin; GSN, gelsolin.




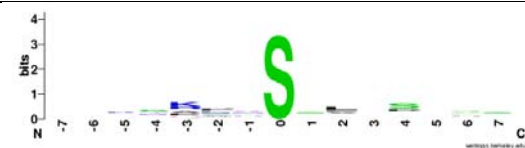
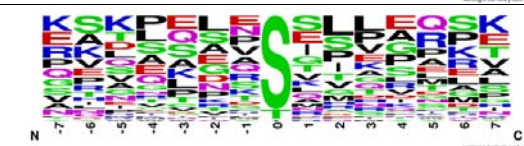

**Supplementary Figure 6: Domain organization showing position of validated YopO phosphorylation sites.** EVH1, enabled/VASP homology 1 domain; PR, poly-proline; G, G-actin binding domain; F, F-actin binding domain; Tet, tetramerization domain; FH1, 2, 3, formin homology domains 1, 2 and 3; DAD, diaphanous autoregulatory domain; COIL, coiled coil; ; W, Wiskott-Aldrich syndrome homology region 2 (WH2); HFD, helical folded domain; CARP,  $\beta$ -sheet/ $\beta$ -helix (CARP) domain; PH, Pleckstrin homology; SH2 and 3, src homology-2 and 3; WBD, WASP-binding domain; REM, Ras exchanger motif; IMD, IRSp53 and MIM homology domain; PTPase, protein tyrosine phosphatase.



**Supplementary Figure 7. Model of binding of EVL and VASP to YopO:actin for phosphorylation.** Model of the quaternary complex of YopO–actin–profilin–VASP fragment derived from PDB code 2PBD, in which the poly-proline sequence is shown in orange and the G-actin binding domain is in red. The N- and C-termini are shown as blue and red spheres respectively. The model illustrates how the binding of VASP to YopO:actin may allow the region C-terminal of the G-actin binding domain to span into the catalytic cleft for phosphorylation. The similar domain structures of EVL and VASP suggest that they may share a similar mode of binding, though this does not explain the phosphorylation observed on Ser2 of VASP. Mouse orthologs are shown here with the KISS phosphorylation sites shown in bold red. Residues shown in red but not in bold are ambiguously identified and either one of the four serines could be phosphorylated. G, G-actin binding domain; F, F-actin binding domain; Tet, tetramerization domain.



**Supplementary Figure 8: Prediction of disordered regions on YopO substrates using PONDR.** The prediction was performed on mouse orthologs using the VSL2 predictor algorithm. Residues phosphorylated by YopO are indicated by red asterisks.

	<i>In vitro</i> phosphorylation sites	KISS sites
<b>EVL</b>	QQVVI NYSI VKGLKY PEDASGGSSPSGTSK GSSPSGTSKSDANRA SPSGTSKSDANRASS KSDANRASGGGGGGL AEKKEDESQMEDPST ESQMEDPSTSPSPGT GRKPWERSNSVEKPV NSVEKPVSSI LSRTP EAKSPLOSQPHSRMK PLOSQPHSRMKPAGS	NSVEKPVSSLLSRTP EAKSPLOSQPHSRVK PLOSQPHSRVKPAGS
<b>VASP</b>	GTGPOAFSRVQI YHN YHNPTANSFRVFRV KROQPGPSEHI ERRV VSKQEEAAGGPTAPK PKAESGRSGGGGLME EKTPKDESANQEEPE EARVPAQSESVRRPW RVPAQSESVRRPWK RRPWKNSTTLPRMK RPWEKNSTTLPRMKS	MSETVI CSS EKPPKDESASEESEA
<b>DIAPH1</b>	SKTAQNLSI FLGSFR NI KPEI VSVTAACEE RKAGCAVTSLLASEL KAGCAVTSLLASELT AVTSLLASELTKDDA KKSEGVPTI LEEAKE	KAGCAVTSLLASELT
<b>INF2</b>	N. D.	DEAORGQSTHLPRTG
<b>SHOT1</b>	EKQLQLI TSLKEOAI NKTLKRI SMLYMAKL OEKTVLNSEVLEORK QNKLKROSHLLLQSS QSHLLLQSSI PDQQL SHLLLQSSI PDQQLL RVNQSENSVPPPPPP NPI RSLMSMI RKRSH TSSRSLKSLDPENSE KSLDPENSE TELERI VTAEADSSSPTGI LA EADSSSPTGI LATSE PTGI LATSEKSMPLV GI LATSEKSMPLV LATSEKSMPLVLSV KSMPLVLSVSSVTKT KLEGCTSSKVTFFPP KVTFQPPSSI GCRKK VVLDPVSTHEPOTK OPENKEDSI ENVRET	QNKLKROSHLLLQSS
<b>CAP1</b>	PYVQAFDSSLANPVA RALLATASQCQPAG SDDSASRSALFAQI N AQI NQGESI THALKH VENQENVSNLVI DDT	SDDSASRSALFAQI N
<b>EPS8</b>	N. D.	NYARDSVSSVSDVSQ
<b>ABI1</b>	N. D.	PPTQKPPSPVSGRG
<b>FBNP1L</b>	N. D.	IRGDRRHSSDI NHLV
<b>MTSS1</b>	N. D.	SRKSSVCSSLNSVNS
<b>PDLIM5</b>	N. D.	VKPVPI TSPAVSKVT
<b>CTNND1</b>	N. D.	DPRRRLRSYEDMI GE
<b>PTPN6</b>	N. D.	N. A.
<b>Sequence logo</b>		
<b>Frequency plot</b>		

**Supplementary Figure 9: Analysis of substrate phosphorylation sequence determinants.**

Phosphorylation sites including 7 residues up and downstream of the phosphoacceptor as identified by *in vitro* and KISS phosphorylation were used in the sequence analysis. Ambiguously determined phosphorylation sites were left out from this table and not included in logo generation. *In vitro phosphorylation* was performed using human EVL, VASP, SHOT1 and mouse CAP1 and DIAPH1 and sequences shown are of their respective orthologs. KISS was performed on mouse macrophage-like cell line Raw264.7 and consequently shown sequences were from mouse. The phosphorylated residue is positioned at 0, while positive and negative positions represent amino acids on the COOH- and NH<sub>2</sub>-terminal side of the phosphorylated residue respectively. These plots are generated by WebLogo<sup>2</sup>. N.D., not determined. N.A., not applicable, due to ambiguous identification.

Constructs	Amino acid sequences	No. of residues between FKHV and phosphoacceptor serine
G1-ΔDIAPH1	GPGRPMVVEHPEFLKAGKEPGLQIWRVEKFDLVPVPTNLYGDFFTGDAYVILKTVQLRNGNLQYDLHYWLGNECSQDESGAAAIFTVQLDDYLNGRAVQHREVQGFESATFLGYFKSGLKYKKGAVASG <b>FKHV</b> VPNEVVVQRLKAGCAVTSLLASELT	17
G1-15-ΔDIAPH1	GPGRPMVVEHPEFLKAGKEPGLQIWRVEKFDLVPVPTNLYGDFFTGDAYVILKTVQLRNGNLQYDLHYWLGNECSQDESGAAAIFTVQLDDYLNGRAVQHREVQGFESATFLGYFKSGLKYKKGAVASG <b>FKHV</b> VPNEVVVQRLGGAGAGAGGAAGGAGKAGCAVTSLLASELT	32
G1-22-ΔDIAPH1	GPGRPMVVEHPEFLKAGKEPGLQIWRVEKFDLVPVPTNLYGDFFTGDAYVILKTVQLRNGNLQYDLHYWLGNECSQDESGAAAIFTVQLDDYLNGRAVQHREVQGFESATFLGYFKSGLKYKKGAVASG <b>FKHV</b> VPNEVVVQRLGGAGAGAGGAGAGAGGAAGGAGKAGCAVTSLLASELT	39
G1-30-ΔDIAPH1	GPGRPMVVEHPEFLKAGKEPGLQIWRVEKFDLVPVPTNLYGDFFTGDAYVILKTVQLRNGNLQYDLHYWLGNECSQDESGAAAIFTVQLDDYLNGRAVQHREVQGFESATFLGYFKSGLKYKKGAVASG <b>FKHV</b> VPNEVVVQRLGGAGAGAGGAAGGAGGGAGAGAGGAGAGGAGGAGKAGCAVTSLLASELT	47
G1-36-ΔDIAPH1	GPGRPMVVEHPEFLKAGKEPGLQIWRVEKFDLVPVPTNLYGDFFTGDAYVILKTVQLRNGNLQYDLHYWLGNECSQDESGAAAIFTVQLDDYLNGRAVQHREVQGFESATFLGYFKSGLKYKKGAVASG <b>FKHV</b> VPNEVVVQRLGGAGAGAGGAAGGAGGGAGAGGGAGAGGGAGAGGAGGAGKAGCAVTSLLASELT	53
G1-15-PKCa	GPGRPMVVEHPEFLKAGKEPGLQIWRVEKFDLVPVPTNLYGDFFTGDAYVILKTVQLRNGNLQYDLHYWLGNECSQDESGAAAIFTVQLDDYLNGRAVQHREVQGFESATFLGYFKSGLKYKKGAVASG <b>FKHV</b> VPNEVVVQRLGGAGAGAGGAAGGAGKRRRRRKGSKKSELT	32
G1-15-PKA	GPGRPMVVEHPEFLKAGKEPGLQIWRVEKFDLVPVPTNLYGDFFTGDAYVILKTVQLRNGNLQYDLHYWLGNECSQDESGAAAIFTVQLDDYLNGRAVQHREVQGFESATFLGYFKSGLKYKKGAVASG <b>FKHV</b> VPNEVVVQRLGGAGAGAGGAAGGAGKAGKKRFSIFDEL	32
G1-15-PAK	GPGRPMVVEHPEFLKAGKEPGLQIWRVEKFDLVPVPTNLYGDFFTGDAYVILKTVQLRNGNLQYDLHYWLGNECSQDESGAAAIFTVQLDDYLNGRAVQHREVQGFESATFLGYFKSGLKYKKGAVASG <b>FKHV</b> VPNEVVVQRLGGAGAGAGGAAGGAGKAGCKRTSLLASELT	32
G1-15-CK2	GPGRPMVVEHPEFLKAGKEPGLQIWRVEKFDLVPVPTNLYGDFFTGDAYVILKTVQLRNGNLQYDLHYWLGNECSQDESGAAAIFTVQLDDYLNGRAVQHREVQGFESATFLGYFKSGLKYKKGAVASG <b>FKHV</b> VPNEVVVQRLGGAGAGAGGAAGGAGKAGCAVTSLLDDEL	32
G1-15-Poly-G	GPGRPMVVEHPEFLKAGKEPGLQIWRVEKFDLVPVPTNLYGDFFTGDAYVILKTVQLRNGNLQYDLHYWLGNECSQDESGAAAIFTVQLDDYLNGRAVQHREVQGFESATFLGYFKSGLKYKKGAVASG <b>FKHV</b> VPNEVVVQRLGGAGAGAGGAAGGAGKGGGGGGSGGGGGT	32

**Supplementary Table 4. Protein sequences of the artificial substrates.** The residues GPGRP are left after 3C protease cleavage. The FKHV motif, shown in bold, is the gelsolin equivalent of the LKKT motif in WH2 domains<sup>3</sup>. The poly-glycine-alanine linker is shown in red.

## References

- 1 Pei, J. & Grishin, N. V. PROMALS: towards accurate multiple sequence alignments of distantly related proteins. *Bioinformatics* **23**, 802-808 (2007).
- 2 Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: A Sequence Logo Generator. *Genome Res* **14**, 1188-1190 (2004).
- 3 Irobi, E. *et al.* Structural basis of actin sequestration by thymosin-β4: implications for WH2 proteins. *Embo J* **23**, 3599 (2004).