SUPPORTING INFORMATION for "Host-Pathogen Interaction Profiling Using Self-Assembling Human Protein Arrays" by Yu, Decker, et al.



Supplementary Figure S1. Expression of N-terminal and C-terminal SidM HaloTag constructs using human HeLa cell-based expression system. $50ng/\mu L$ of plasmid DNA was employed and expression was executed at 30°C for 2 hr. $50-100 \mu g/ml$ LidA and SidM were produced using human HeLa cell-based expression system and quantified using purified HaloTag protein as the standard.



Supplementary Figure S2. Reproducibility of target detection by LidA in two independent experiments. The binding of HaloTag-LidA to identical NAPPA arrays was detected with Alexa660-labeled Halo-ligand in two independent experiments on different days. Protein slides were artificially colored red (day 1) and green (day 2) and merged in order to visualize the reproducibility of target candidate detection (yellow). The entire set of arrays yielded reproducible results; one representative array is shown here.



Supplementary Figure S3. Validation of bait protein capture on NAPPA array. To confirm protein display on NAPPA produced with the human HeLa cell-based expression system, we probed one of the arrays with anti-GST tag antibody followed by Tyramide Signal Amplification (PerkinElmer). The results, shown in the form of a heat map, indicate that the spot signals on the array were significantly increased over those on a control slide lacking T7 polymerase (not shown) suggesting that most of the human ORF clones were successfully transcribed, translated, and then captured on the array. The heat map indicates signal intensity ranging from blue (low) to red (high).



Supplementary Figure S4. To examine the effect of differential nucleotide binding on the GTPase targets, an array comprised of select Rab GTPases, plus the negative controls RhoA and Cdc42, was synthesized with HeLa cell-based IVTT. Spotted proteins were then stripped of nucleotide and left 'unloaded', or loaded with either GDP or GTP γ S to render the Rab proteins inactive or active, respectively. Arrays were subjected to the PPI assay using LidA as bait. The results are shown in the form of a heat map indicating signal intensity ranging from blue (low) to yellow (high).



Supplementary Figure S5. Input of HaloTag protein, HaloTag-LidA and HaloTag-SidM used in the bead-based pull-down assay. $50ng/\mu L$ of plasmid DNA was employed and expression was performed at 30°C for 2 hr.



Supplementary Figure S6. Interface residues conserved within LidA-binding Rabs. (A) Cartoon illustration of Rab1(1-176, light blue) bound by LidA(224-599; light grey) (34). Residues Ile76 and Thr75 (cluster L3), Lys61 (L2), Tyr8 (L1), and Lys103 (L4) are shown as sticks. Note that all four clusters are located within or adjacent to the LidA-Rab1 interface. Figure generated using the PyMOL Molecular Graphics System, Schrödinger, LLC.



Supplementary Figure S7. Intracellular localization of exogenous mCherry-LidA. COS1 cells were transiently transfected with a plasmid encoding mCherry-LidA and GFP-tagged target candidates, and protein colocalization 16 hours after transfection was determined by fluorescence microscopy. Scale bar, 1μ m. The broadly dispersed fluorescent signal for mCherry-LidA yielded inconclusive results for colocalization with selected host targets.

BAD AVG G *	OOD		
Rab1	1	MNPEYDYLFKLLLIGDSGVGKSCLLLRFADDTYTENYISTIGVDFKIRTIELDGKTIKL	59
Rab2B	1	MTYAYLFKYIIIGDTGVGKSCLLLQFTDKRFQPVHDLTIGVEFGARMVNIDGKQIKL	57
Rab4B	1	MAEDRHFLFKFLVIGSAGTGKSCLLHOFIENKFKODSNHTIGVEFGSRVVNVGGKTVKL	59
Rab8a	1	MAKTYDYLFKLLLIGDSGVGKTCVLFRFSEDAFNSTFISTIGIDFKIRTIELDGKRIKL	59
Rab8b	1	MAKTYDYLFKLLLIGDSGVGKTCLLFRFSEDAFNTTFISTIGIDFKIRTIELDGKKIKL	59
Rab10	1	MAKKTYDLLFKLLLIGDSGVGKTCVLFRFSDDAFNTTFISTIGIDFKIKTVELOGKKIKL	60
Rab11b	1	MGTRDDEYDYLFKVVLIGDSGVGKSNLLSRFTRNEFNLESKSTIGVEFATRSIOVDGKTIKA	62
Rab13	1	MAKAYDHLFKLLLIGDSGVGKTCLIIRFAEDNFNNTYISTIGIDFKIRTVDIEGKKIKL	59
Rab27a	1	MSDGDYDYLIKFLALGDSGVGKTSVLYQYTDGKFNSKFITTVGIDFREKRV <mark>VYRA</mark> SGPDGATGRGQRIHL	70
Rab35	1	MARDYDHLFKLLIIGDSGVGKSSLLLRFADNTFSGSYITTIGVDFKIRTVEINGEKVKL	59
cons	1	* *:*::*:*:*::::::: *:*::*::	72
Rab1	60	OTWDTAGOERFRTTTSSYYRGABGIIVVYDVTDOESYANVKOWLOEIDRYAS-ENVNKLLVGNKSDLTTKKV	130
Rab2B	58	OIWDTAGQESFRSITRSYYRGAAGALLVYDITRRETFNHLTSWLEDAROHSS-SNMVIMLIGNKSDLESRRD	128
Rab4B	60	OIWDTAGOERFRSVTRSVYRGAAGALLVYDITSRETYNSLAAWLTDARTLAS-PNIVVILCGNKKDLDPERE	130
Rab8a	60	OIWDTAGOERFRTITTAYYRGAMGIMLVYD ITNEKSFDN IRNWIRN I EEHAS - ADVEKMILGNKCDVNDKRO	130
Rab8b	60	QIWDTAGQERFRTITTAYYRGAMGIMLVYDITNEKSFDNIKNWIRNIEEHAS-SDVERMILGNKCDMNDKRQ	130
Rab10	61	QIWDTAGQERFHTITTSYYRGAMGIMLVYDITNGKSFENISKWLRNIDEHAN-EDVERMLLGNKCDMDDKRV	131
Rab11b	63	QIWDTAGQERYRAITSAYYRGAVGALLVYDIAKHLTYENVERWLKELRDHAD-SNIVIMLVGNKSDLRHLRA	133
Rab13	60	QVWDTAGQERFKTITTAYYRGAMGIILVYDITDEKSFENIQNWMKSIKENAS-AGVERLLLGNKCDMEAKRK	130
Rab27a	71	QLWDTAGQERFRSLTTAFFRDAMGFLLLFDLTNEQSFLNVRNWISQLQMHA <mark>YCE</mark> NPDIVLCGNKSDLEDQRV	142
Rab35	60	QIWDTAGQERFRTITSTYYRGTHGVIVVYDVTSAESFVNVKRWLHEINQN <mark>C</mark> <mark>D</mark> DVCRILVGNKNDDPERKV	129
cons	73	*:****** ::::* :::*:: :: : *:	144
Rab1	131	VDNTTAKEFADSLGIPFLETSAKNATNVEQAFMTMAAEIKKRMGPGAA <mark>S</mark> <mark>GGE-RPN</mark> <mark>LKI</mark>	188
Rab2B	129	<mark>VKREEGEAFAREHGLIFMETSAKTACNVEEAFINTAKEIHRKIQQGLFD</mark> <mark>VHNEANGI</mark> -KIGPQQSIST	195
Rab4B	131	VTFLEASRFAQENELMFLETSALTGENVEEAFLKCARTILNKIDSGELDP <mark>ERMGSGI</mark> -QYGDASLRQL	197
Rab8a	131	<mark>VSKERGEKLALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKKLEG</mark> <mark>NSPQGSN</mark> <mark>QGVKI</mark>	191
Rab8b	131	VSKERGEKLAIDYGIKFLETSAKSSANVEEAFFTLARDIMTKLNRKMNDSNSAGAGGPVKI	191
Rab10	132	VPKGKGEQIAREHGIRFFETSAKANINIEKAFLTLAEDILRKTPVKEPNSENVDI	186
Rabilb	134	VPTDEARAFAEKNNLSFIETSALDSTNVEEAFKNILTEIYRIVSOKQIADCAAHDESPGNNVVDI	198
Rab13	142	VQREQADKLAREHGIRFFETSAKSSMNVDEAFSSLARDILLKSGGRRSGNGRKPPSTDL	189
Rabz/a	143	VREEEATALAEKIGTPIFETSAANGTNISQATEMLIDLIMKRMERCVDKSWI-PEGVVRSNGHAST	207
Rad35	130	VETEDAIRFAGQMGIQLFETSAKENVNVEEMFNCITELVLKAKKDNLARQQQQQQQNDVVKL	190
cons	145	* . :* . : :**** *:.: : : :	216
Rab1	189	<mark>DSTPV</mark> <mark>KPAGGGCC</mark> 201	
Rab2B	196	SVGPSASQRNSRDIGSNSGCC 216	
Rab4B	198	R <mark>QPRSAQAVAPQPCGC</mark> 213	
Rab8a	192	TPDQQKRSSFFRCVLL 207	
Rab8b	192	TENRSKKTSFFRCSLL 207	
Rab10	187	SSGGGVTGWKSKCC 200	
Rabilb	199	SVPP-TTDGQKPNKLQCCQNL 218	
Rabij	190	KNTNKCSLG 203	
Rab2/a	208	DULSEEKEKGACGU 221	
RdD33	191		
cons	217	237	

*

Supplementary Figure S8. Primary sequence alignment of Rab GTPases AMPylated by SidM. Sequence alignment was performed using Tcoffee (68). Numbers indicate amino acid residues (shown in single letter code) of each protein. Color coding indicates conservation level ranging from blue (low) to red (high). Tyrosine-77 in Rab1 and the equivalent residues in the other Rabs

are shown in white. Rab27A is the only candidate lacking a tyrosine or equivalent residue available for AMPylation (serine, threonine) at that position. Instead, Rab27A contains a ten amino acid insertion (residues 52-61) not present in the other candidates that could provide an alternative site for AMPylation.