The Pathogen Protein EspF_U Hijacks Actin Polymerization Using Mimicry and Multivalency

Nathan A. Sallee, Gonzalo M. Rivera, John E. Dueber, Dan Vasilescu, R. Dyche Mullins, Bruce J. Mayer & Wendell A. Lim

Supplementary Methods

Additional protein construction and purification. The EspF_U ΔLinker protein was assembled using an adapted BioBrick cloning strategy (for details see: parts.mit.edu). Briefly, we constructed a vector bearing a single synthetic repeat with EcoRI and BglII restriction sites upstream and BamHI and XhoI sites downstream. This vector was digested with BamHI and XhoI and the cut backbone was purified. Another sample of the same vector was digested with BglII and XhoI and the small fragment consisting of the repeat was purified. These two restriction fragments were then ligated using the compatible ends generated by BamHI and BglII, thereby generating a two-repeat construct. This process was repeated multiple times to make the six-repeat construct.

The GBD-C protein was made as described[7] by fusing human WASP residues 242-310 to residues 461-491 with a "GGSGGS" linker sequence. We made our GBD-EspF_U fusion by replacing WASP residues 466-491 in GBD-C with residues 1-17 from repeat 4. Both fusion proteins were assembled by two-step PCR.

EspF_U lacks cysteine residues, so for fluorescent labeling of repeat fragments, a "KCK" sequence was added to the N-terminus of the fragment. For the N-WASP WCA

(residues 392-501), the single cysteine at position 427 was labeled. For the N-WASP

GBD fragments, we made a C236S mutation that behaved like wildtype N-WASP in

assays of binding and activity. We then added a "KCK" sequence to the N-terminus of

C236S GBD fragments. These cysteine residues were labeled with the iodoacetamide

derivatives of the appropriate dyes: we used fluorescein for fluorescence polarization and

Oregon Green 488 for analytical ultracentrifugation (Invitrogen).

Circular dichroism. Circular dichroism spectra were measured of protein samples in 20

mM sodium phosphate pH 8.0 and 50 mM NaCl on an Aviv spectrometer. GBD-C and

GBD- EspF$_U$ 1-17 readings were done on samples at 20 μM concentration. The GBD

construct showed lower signal (due to its instability) and was assayed at 40 μM

concentration in order to get readings in the appropriate range. Temperature melts were

monitored at 222 nm from 10°C to 90°C in 2°C steps with one minute equilibration time.

Subsequent cooling of the samples back to 10°C showed that the unfolding of each

protein was reversible.

Fluorescence polarization binding assays. Fluorescein-labeled probe proteins were

added at 10-30 nM concentration, depending on the amount required to give the

appropriate fluorescence intensity. Series of samples were made with constant probe

concentration and increasing concentrations of unlabeled binding partner in 20 mM Tris

pH 8.0, 50 mM NaCl, 1 mM DTT and 0.1 mg/mL BSA. These samples were loaded into

a 384-well plate (Corning) and read in an Analyst HT plate reader (Molecular Devices)

(excitation: 485 nm; emission: 530 nm). All experiments were performed in triplicate and

the three data sets were fitted to binding curves using the Levenberg-Marquardt algorithm in the program ProFit (Quantum Soft). The mean and standard deviation of the K_d values from these three fits were reported.

Additional *In vivo* clustering assays. N-WASP localization experiments (figure S9) and images of EspF$_U$-transfected, unclustered cells (figure S10) were each acquired in two different cell lines: NIH 3T3s and N-WASP -/- mouse embryonic fibroblasts (MEFs)[36]. The NIH 3T3s were transfected with GFP-actin and the appropriate CD16/7-EspF$_U$ fusion protein, which was either clustered and detected with a rhodamine-tagged secondary antibody or left untreated. After fixation and permeabilization, endogenous N-WASP was immunostained using a rabbit polyclonal N-WASP antibody (provided by Marc Kirschner's lab) and detected with an Alexa 647-tagged secondary antibody, as previously described[26]. In the MEF experiments, the cells were transfected with GFP-N-WASP (or GFP alone in figure S10a) and the appropriate CD16/7-EspF$_U$ fusion protein, which was either clustered and detected with an Alexa 647-tagged secondary antibody or left untreated. After fixation and permeabilization, F-actin was stained with rhodamine-labeled phalloidin.

Comet Detector Algorithm. Details will be published elsewhere, but in brief this tool is an adaptation of the application for multiple particle detection and tracking from digital videos developed by Guy Levy at the Computational Biophysics Laboratory, ETH Zurich. The Comet Detector utilizes previously described algorithms[37, 38] for analysis of antibody clusters and actin comets in two channels of fluorescent images.

For detection of clusters, the algorithm assumes that the particles are circular in shape and that the size and intensity between particles vary by no more than one order of magnitude. The software allows the user to adjust the parameters relevant to the detection process: radius and threshold. The radius (in pixels) of the particles of interest (clusters) should be slightly larger than the apparent particle radius but smaller than the smallest distance between two particles. The radius is chosen as to minimize errors such as "collision" (two particles detected as one if radius is too large) and "left out intensity" (part of the particle being left out if radius is too small). The parameter threshold allows all the bright pixels in the upper n^{th} percentile of the image intensity to be considered as candidate particles. For detection of actin comets, a similar algorithm is used with the assumption that most of the intensity information is located at the "growing end", which is consistent with visual observations and modeling. In this case, the radius chosen must be large enough to include most of the comet body. Because of the "noisy" nature of the actin channel, true actin tails as well as other cytoskeletal structures (actin cables) are detected at any given threshold. However, spurious detections (such as actin cables) are then eliminated from the analysis during the pairing of information from both channels. The detection process on both channels consists of the following common steps; i) restoration of image, ii) location of candidate particle positions by finding local intensity maxima in the image filtered during the previous step and, iii) particle position correction based on the assumption that the center of mass of the particle is near (but not necessarily at) the point of local intensity maxima. After the detection of particles in each channel, a multi-pass algorithm is used to pair the detected clusters with the comets. We identified proximity as the most reliable pairing criterion. The algorithm identifies the comet

closest to a given cluster by searching over a progressively larger area around the cluster that is smaller or equal to the chosen comet diameter.

Analytical ultracentrifugation. Oregon Green 488-labeled $EspF_U$ fragments and unlabelled N-WASP fragments were monitored at 495 nm and 280 nm, respectively. Labeled repeat fragments were present at sufficient concentration to give an absorbance reading of 0.3-0.4 at 495 nm (2-8 µM, depending on labeling efficiency). Unlabeled mini N-WASP was added at the indicated molar excesses. Samples of $EspF_U$ fragment alone, $EspF_U$ and N-WASP mixed, and N-WASP alone were spun at 10,000 and 20,000 rpm (in succession) for 22 hours at each speed (1 scan/hour). Equilibrium analytical ultracentrifugation experiments were performed at 20°C on a Beckman Optima XL-A ultracentrifuge with an An-60 Ti rotor. Data were processed using the program Reedit9 (Jeff Lary, National Analytical Ultracentrifuge Facility) and then fitted to effective reduced molecular weight (σ) values with WinNonlin[39]. Data and fitted curves were plotted and residuals were calculated using MATLAB.

Supplementary References

36.   Snapper, S. B. *et al.* N-WASP deficiency reveals distinct pathways for cell surface projections and microbial actin-based motility. *Nat. Cell Biol.* **3**, 897-904 (2001).

37.   Sbalzarini, I. F. & Koumoutsakos, P. Feature point tracking and trajectory analysis for video imaging in cell biology. *J. Struct. Biol.* **151**, 182-195 (2005).

38.   Crocker, J. C. & Grier, D. G. Methods of digital video microscopy for colloidal studies. *J. Colloid Interface Sci.* **179**, 298-310 (1996).

39.   Johnson, M. L., Correia, J. J., Yphantis, D. A. & Halvorson, H. R. Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J.* **36**, 575-588 (1981).
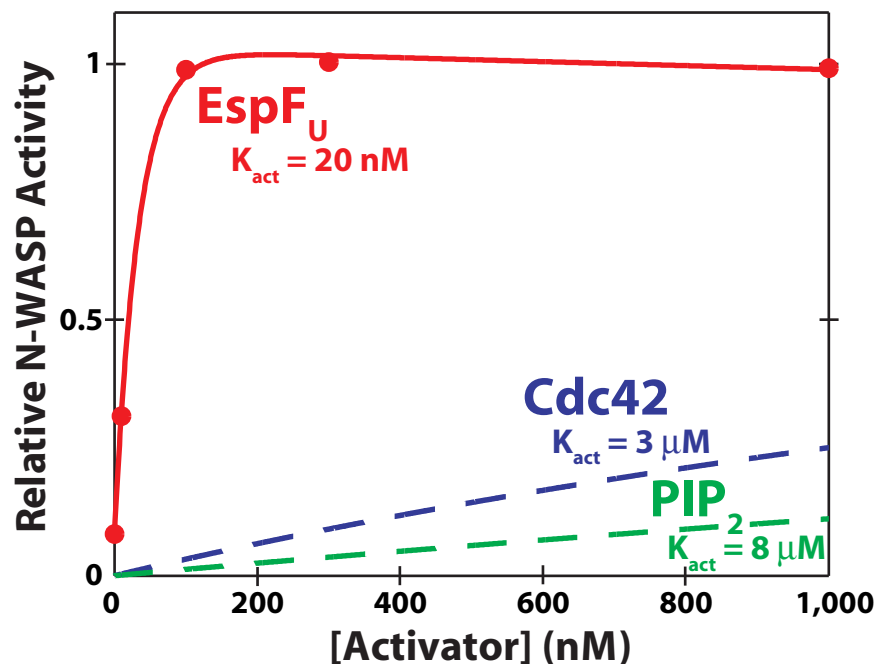
Figure S1 | EspF$_U$ is much more potent than endogenous activators of N-WASP. Activation data are fitted to a single exponential curve using MATLAB (MathWorks), with a small linear term subtracted (where necessary) to account for EspF$_U$ sequestration of actin and/or Arp2/3. This fit gives us a K$_{act}$ value for EspF$_U$ of 20 nM (concentration required for half-maximal N-WASP activity). This value is over 100-fold more potent than our previously-measured K$_{act}$ values for single endogenous activators like Cdc42 and PI(4,5)P$_2$ (3 μM and 8 μM, respectively[8]).

Figure S2 | Sample GST-pulldown results from mapping of binding sites. a, Testing GST fusions of EspF$_U$ repeat 4 fragments for ability to pulldown the N-WASP GBD. Amino acids 1-17 of the EspF$_U$ repeat constitute the minimal fragment that strongly interacts. **b,** Testing GST fusions of N-WASP GBD fragments for ability to pulldown EspF$_U$. Amino acids 228-270 of N-WASP constitute the minimal fragment that strongly interacts.
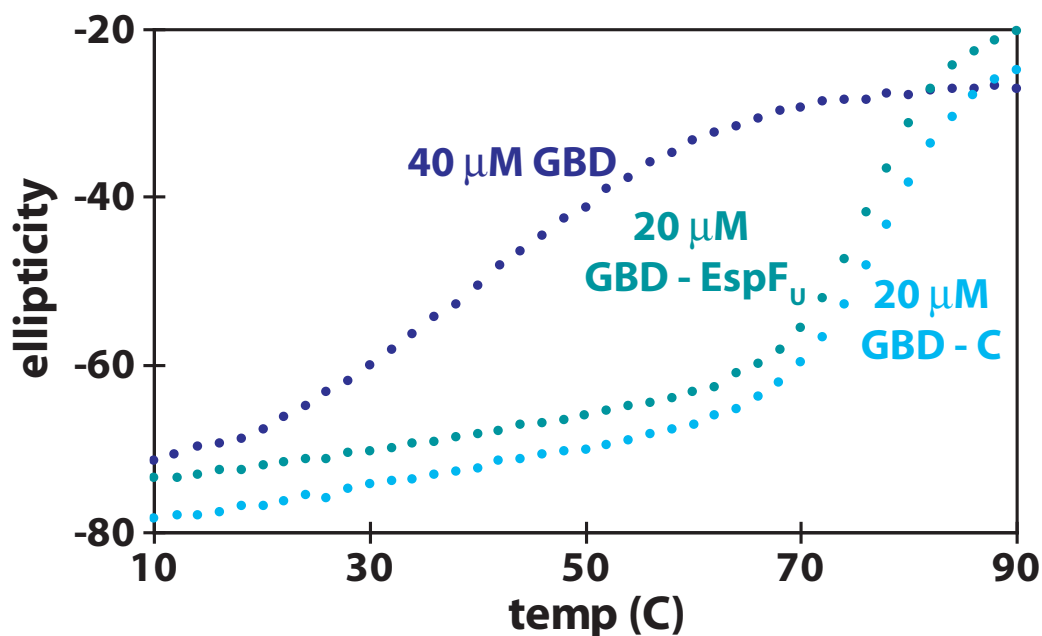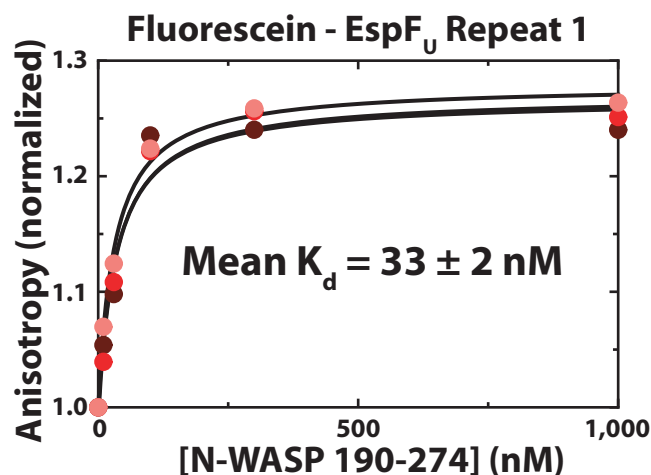
Figure S3 | The GBD fused to EspF$_U$ 1-17 forms a stably folded structure similar to autoinhibited WASP. Temperature melts of GBD proteins, monitored by circular dichroism. The isolated WASP GBD is unstable, as seen by its low melting temperature, shallow melting curve and weak ellipticity signal. Fusion of the C helix to the C-terminus of the GBD (GBD-C) results in a stably folded protein that approximates autoinhibited WASP. A fusion of EspF$_U$ 1-17 to the GBD (GBD-EspF$_U$) shows similar helical content and thermal stability to GBD-C, suggesting that it forms a structure similar to autoinhibited WASP.

a

Ligand	K_d for EspF$_U$ Repeat 1	K_d for N-WASP WCA
N-WASP 228-270	4.49 ± 1.51 μM	
N-WASP 226-274	3.76 ± 0.62 μM	
N-WASP 214-274	18 ± 4 nM	
N-WASP 190-274	33 ± 2 nM	0.4 - 3.4 μM*
Monomeric Actin	7.58 ± 0.41 μM	1.26 ± 0.05 μM
Arp2/3 Complex	622 ± 160 nM	433 ± 52 nM

b

Figure S4 | Affinity comparisons between the EspF$_U$ repeat and the N-WASP WCA.
a, Binding of fluorescein-labeled EspF$_U$ repeat 1 or N-WASP WCA to N-WASP GBD fragments, monomeric actin or Arp2/3 was measured in fluorescence polarization assays. The mean K_d value is reported ± the standard deviation, n = 3. The WCA – GBD affinity (*) is a range of previously measured values for the corresponding WASP fragments[7]. **b,** A sample of binding data from the above table. The three data sets and their fitted curves are plotted.

a

b

```
460 TSGIVGALMEVMQKRSK 476    N-WASP C Helix
  1 LPDVAQRLMQHLAEHGI 17     EspFU Repeat 4
465 LNENPLRALYLIATNGT 481    PAK1 "C" Helix
```
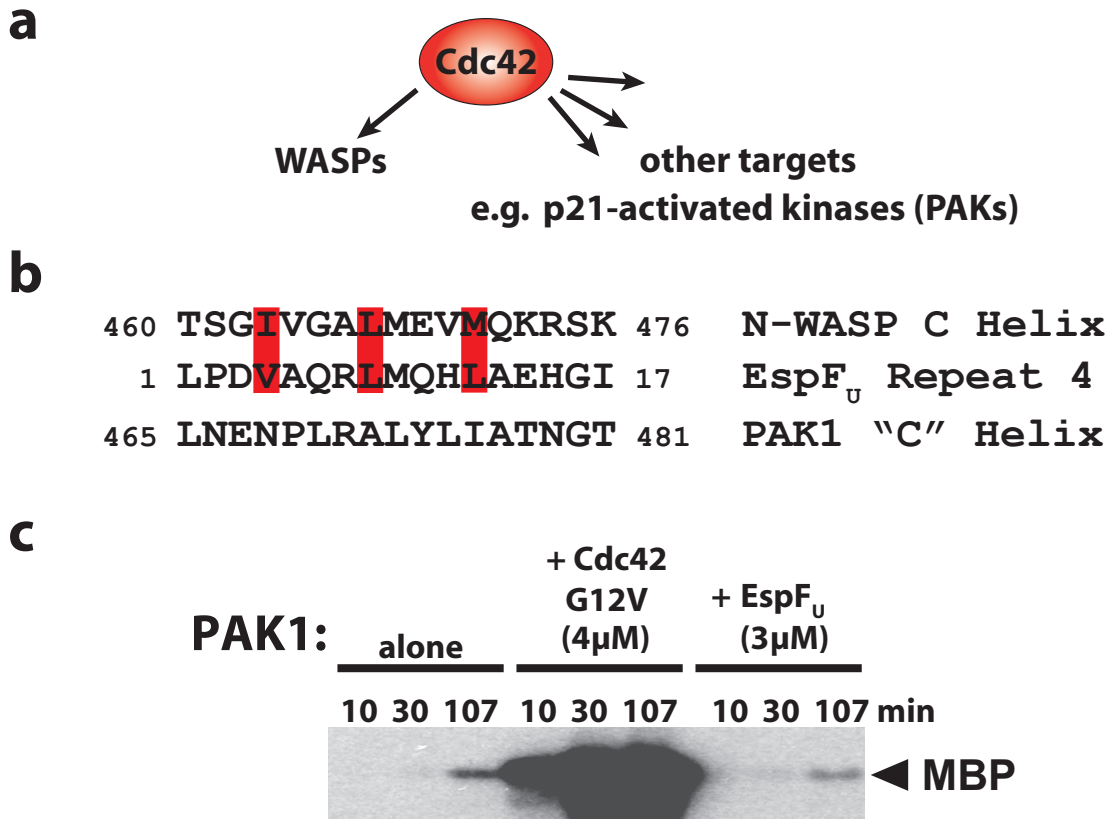
c

Figure S5 | EspF$_U$ is highly specific for WASPs - it does not activate other Cdc42 effectors like p21-activated kinases (PAKs). a, Cdc42, when GTP-bound, activates many downstream signaling proteins, including WASPs. Of these Cdc42 effectors, PAKs are the most closely related to WASPs - they have similar autoinhibited structures and Cdc42 activates them by a similar mechanism. **b,** Sequence alignment of the N-WASP C helix with the EspF$_U$ C mimic highlighting the three conserved hydrophobic residues that are critical for interaction with the GBD. The PAK helix that is structurally homologous to the C helix lacks three hydrophobic residues with the correct spacing. **c,** PAK kinase assay showing incorporation of radioactive phosphate into myelin basic protein (MBP). PAK1 alone is autoinhibited – Cdc42 activates it, but EspF$_U$ does not.
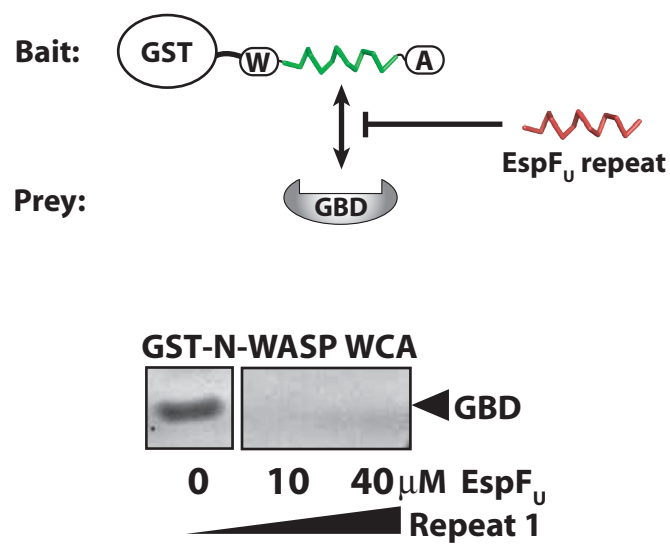
Figure S6 | A single EspF$_U$ repeat can disrupt the N-WASP autoinhibitory interaction. The interaction is shown *in trans* in a GST pulldown assay by incubating 1 µM GST-fused N-WASP WCA (on glutathione-agarose resin) with 10 µM N-WASP GBD. Addition of EspF$_U$ repeat 1 to the mixture disrupts the interaction.
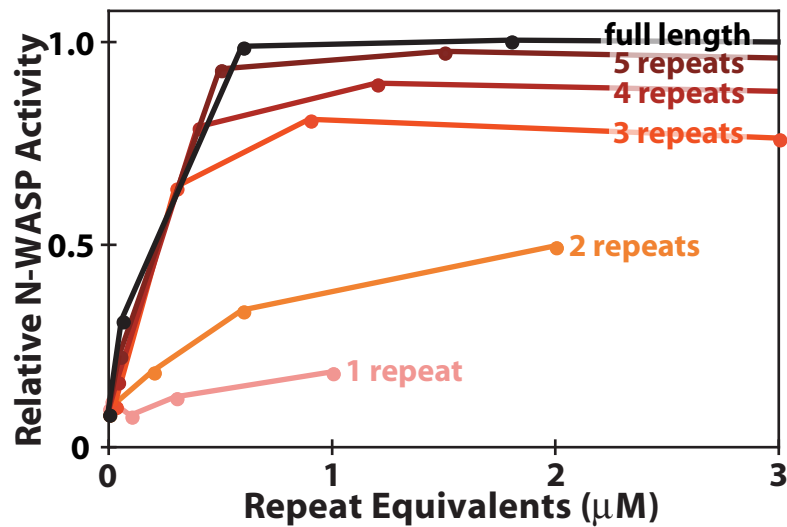
Figure S7 | Activities of EspF$_U$ fragments, scaled by concentration of repeats. Relative N-WASP activity was assessed with the pyrene-actin polymerization assay. The same data from figure 3A are presented here, but with the concentrations scaled by number of repeats in the EspF$_U$ fragment. For example, 1 µM of the 3-repeat fragment is considered 3 µM of repeats. The same trends apply in this representation of the data: one repeat can stimulate actin polymerization relatively weakly, 2 repeats have intermediate activity and 3 repeats are highly active.

a

EspF$_U$ ΔLinker

L P D V A Q R L M Q H L A E H G I	G S G S G S G S Repeat 1
L P D V A Q R L M Q H L A E H G I	G S G S G S G S Repeat 2
L P D V A Q R L M Q H L A E H G I	G S G S G S G S Repeat 3
L P D V A Q R L M Q H L A E H G I	G S G S G S G S Repeat 4
L P D V A Q R L M Q H L A E H G I	G S G S G S G S Repeat 5
L P D V A Q R L M Q H L A E H G I	G S G S G S G S Repeat 6

EspF$_U$ Repeat 4 Residues 1-17   **Synthetic Linker**
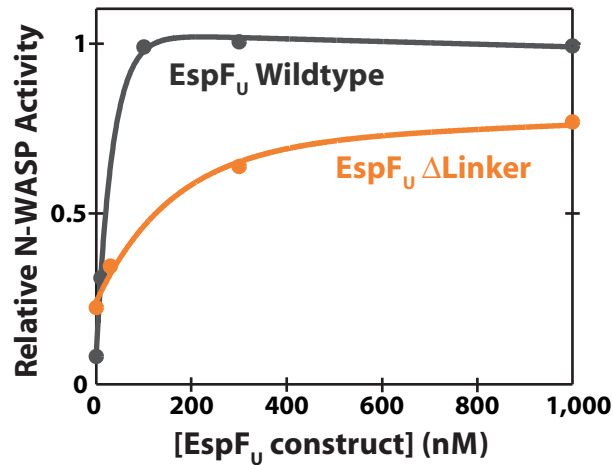
b

Figure S8 | The C-terminal portion of the repeat is not essential for activity. a, Sequence of EspF$_U$ ΔLinker, a chimeric EspF$_U$ protein consisting of six copies of the minimal GBD-binding peptide separated by eight-residue glycine-serine linkers instead of the 30 wildtype proline-rich residues. **b,** This chimera is active in the *in vitro* pyrene-actin polymerization assay, but not as active as wildtype EspF$_U$.
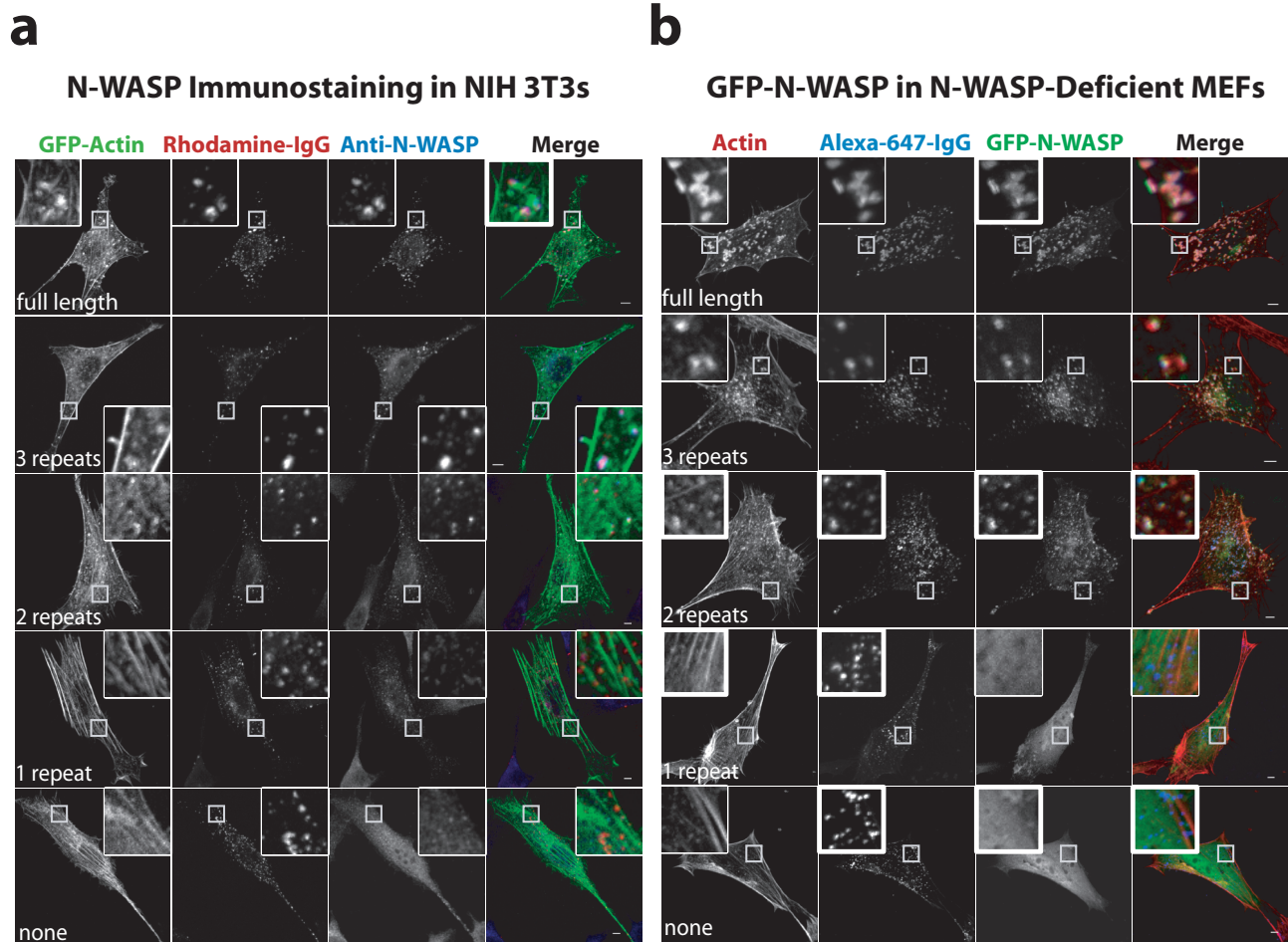
Figure S9 | N-WASP is recruited to sites of EspF$_U$-stimulated actin polymerization. a, The four images for each fragment show (left to right) GFP-actin fluorescence (green), rhodamine fluorescence from clusters of CD16/7 fusions (red), immunofluorescence staining of endogenous N-WASP (blue) and the overlay of these three images. Inset in each image is a portion of that cell (white box) shown at higher magnification. Scale bars indicate 5 μm. **b,** The four images for each fragment show (left to right) F-actin stained with rhodamine-phalloidin (red), Alexa 647 fluorescence from clusters of CD16/7 fusions (blue), GFP-N-WASP (green) and the overlay of these three images.
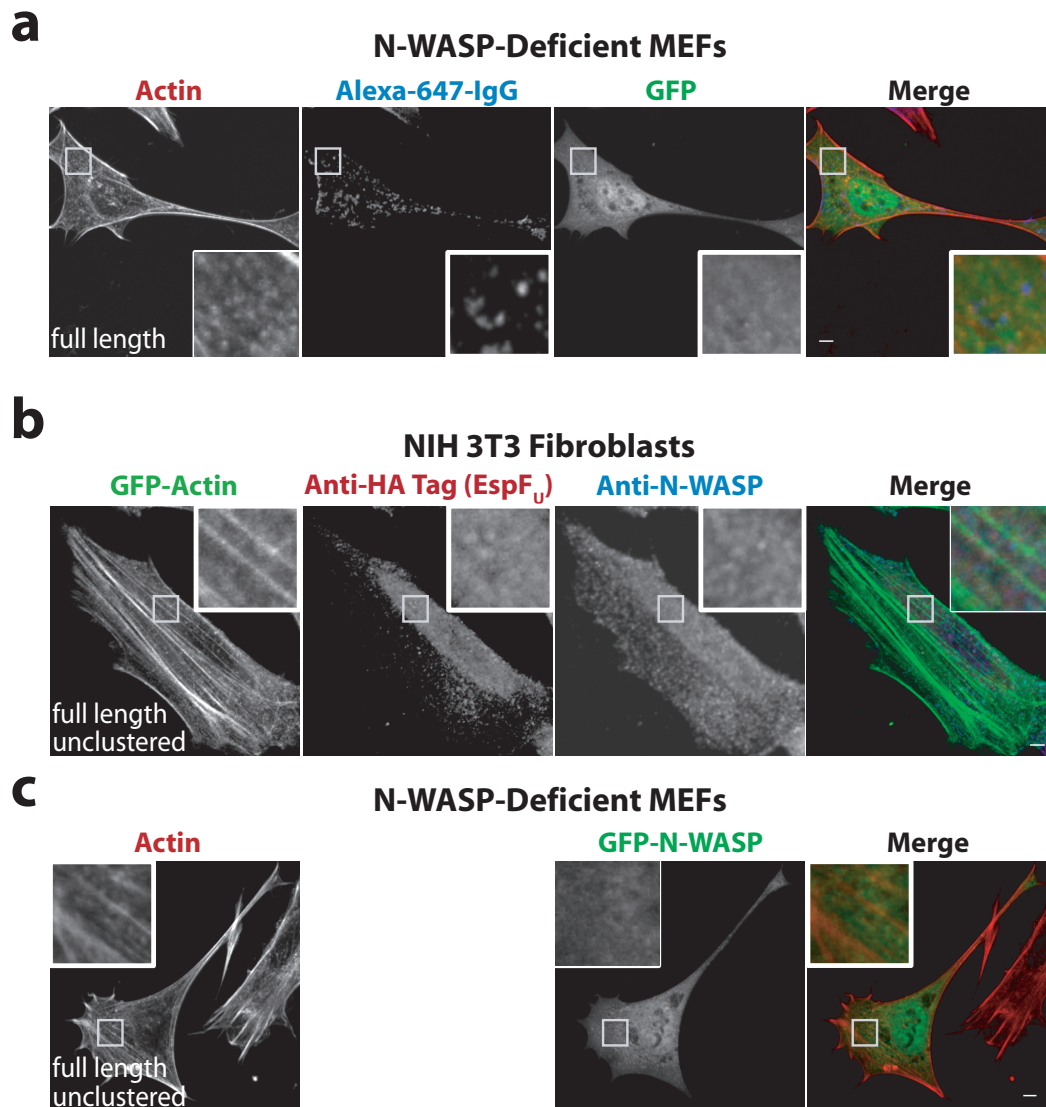
Figure S10 | N-WASP and clustering are both necessary for formation of actin structures. **a,** N-WASP -/- MEFs were transfected with GFP and CD16/7-EspF$_U$ full length, then clustered. In the absence of N-WASP, EspF$_U$ clustering fails to induce localized actin polymerization. Scale bars indicate 5 μm. **b,** NIH 3T3 fibroblasts transfected with GFP-actin and CD16/7-EspF$_U$ full length (HA-tagged) were not treated with clustering antibodies. After fixation, the EspF$_U$ fusion was stained with anti-HA followed by rhodamine-labeled IgG. In the absence of clustering, endogenous N-WASP shows a diffuse cytoplasmatic/nuclear staining with some accumulation in the perinuclear region; prominent F-actin structures are lacking. **c,** Unclustered CD16/7-EspF$_U$ full length in MEFs. GFP-N-WASP shows a diffuse cytosolic/nuclear staining in the absence of clustering.
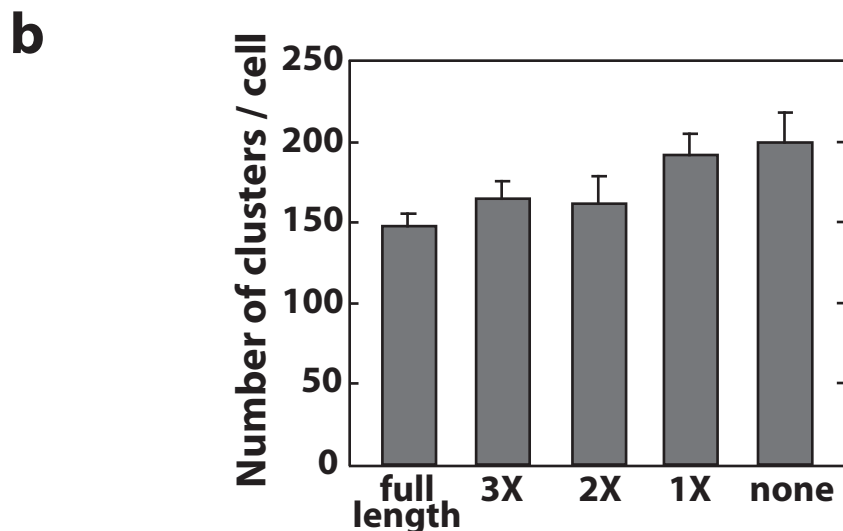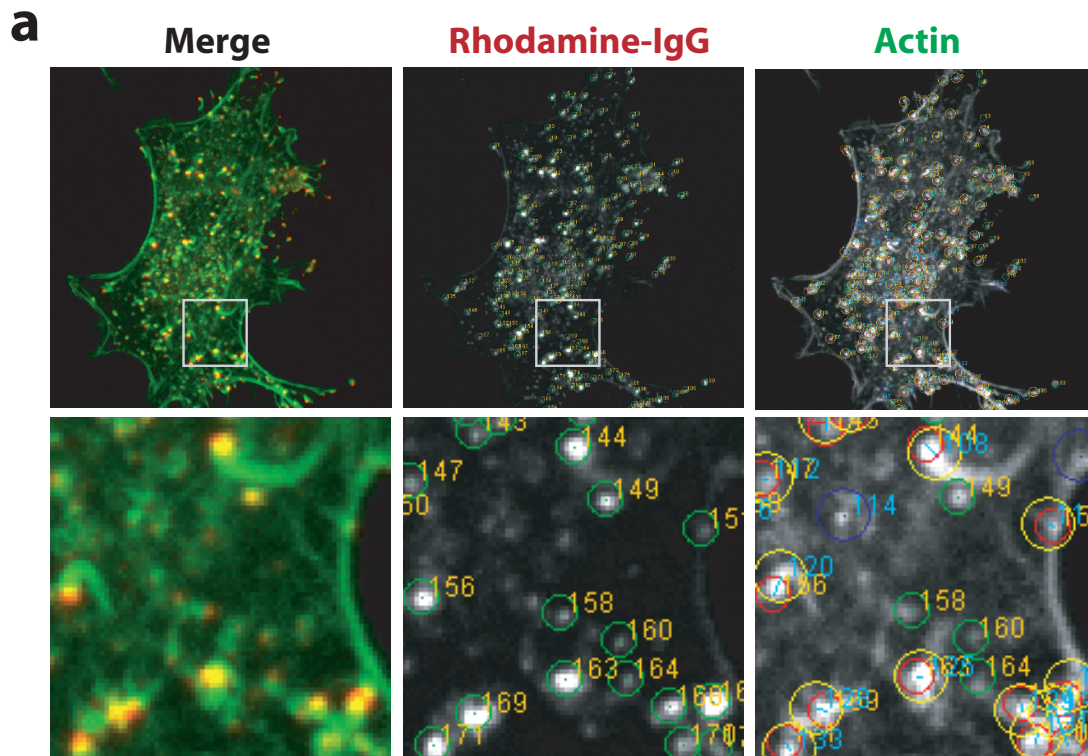
Figure S11 | Quantification of *in vivo* clustering data. a, Shown is a representative image of a cell co-transfected with actin-GFP and CD16/7-Nck. A section of the cell (white square) is shown at higher magnification in the bottom panels. The left panels show the merged image depicting antibody-induced aggregates (red) and the locally induced F-actin (green). The center panels show the aggregates detected by the algorithm in the red channel (green circles). The right panels show F-actin in the green channel with the detected green and red clusters superimposed. Red and yellow circles represent matched antibody clusters and actin patches, respectively. Green and blue circles are orphan antibody clusters and actin patches, respectively. **b,** Similar numbers of clusters per cell are detected with all EspF$_U$ constructs. This shows that all of the EspF$_U$ proteins have similar expression levels and efficiencies of membrane-targeting and clustering. Mean values are reported $\pm$ the standard error of the mean, n = 10.
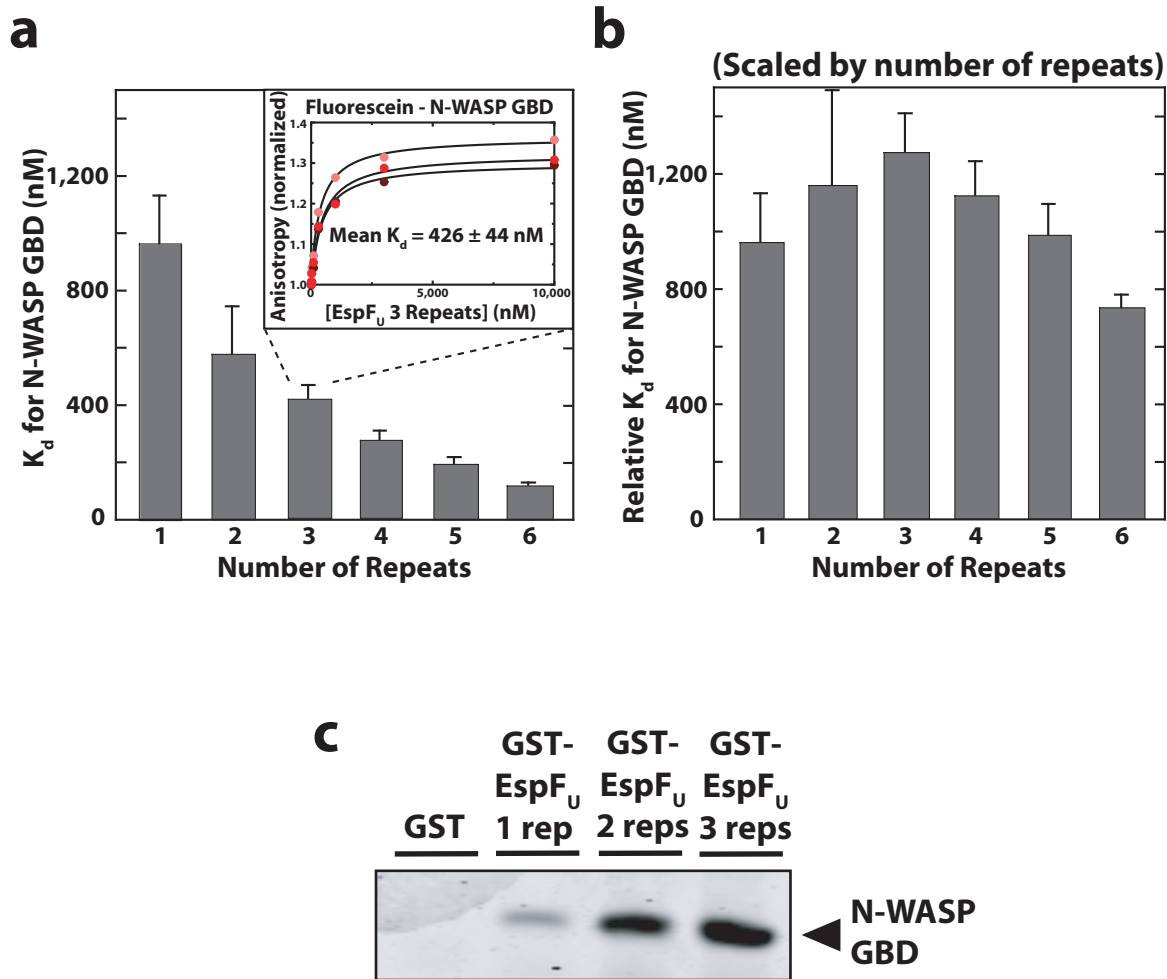
Figure S12 | Each EspF$_U$ repeat can bind an N-WASP GBD simultaneously. a, Using a fluorescence polarization binding assay, we measured the affinities of EspF$_U$ fragments for fluorescein-labeled N-WASP GBD (residues 226-274). Plotted is the measured K$_d$ value vs. the number of repeats in the EspF$_U$ fragment (full-length is considered 6 repeats here) and we see a roughly linear correlation. The mean K$_d$ value is reported ± the standard deviation, n = 3. Inset in the graph are the three sets of binding data for the three-repeat fragment and their fitted curves. **b,** The same graph, but now the K$_d$ values are scaled by the number of repeats (i.e. binding sites) in the construct, so the K$_d$ and standard deviation of a three-repeat fragment, for example, are multiplied by three. This plot shows similar affinity values per binding site across the different fragments, suggesting that the various sites bind to the GBD independently of one another. **c,** GST-pulldown assays showing that EspF$_U$ fragments with increasing numbers of repeats bind to increasing quantities of N-WASP GBD. In each lane, 1 µM GST-fused bait protein was incubated with 10 µM GBD.

a 2 EspF$_U$ Repeats

+ 2X mini N-WASP   5X mini N-WASP

Absorbance (495nm)

Residuals

Radius

σ = 2.64   σ = 3.52

c 2 EspF$_U$ Repeats:

b

	σ values
mini N-WASP (mNW) alone:	1.46
EspF$_U$ 2 repeats alone:	0.59
Predicted 1 EspF$_U$: 1 mNW complex:	2.05
Predicted 1 EspF$_U$: 2 mNW complex:	3.51

d 3 Repeats Data:

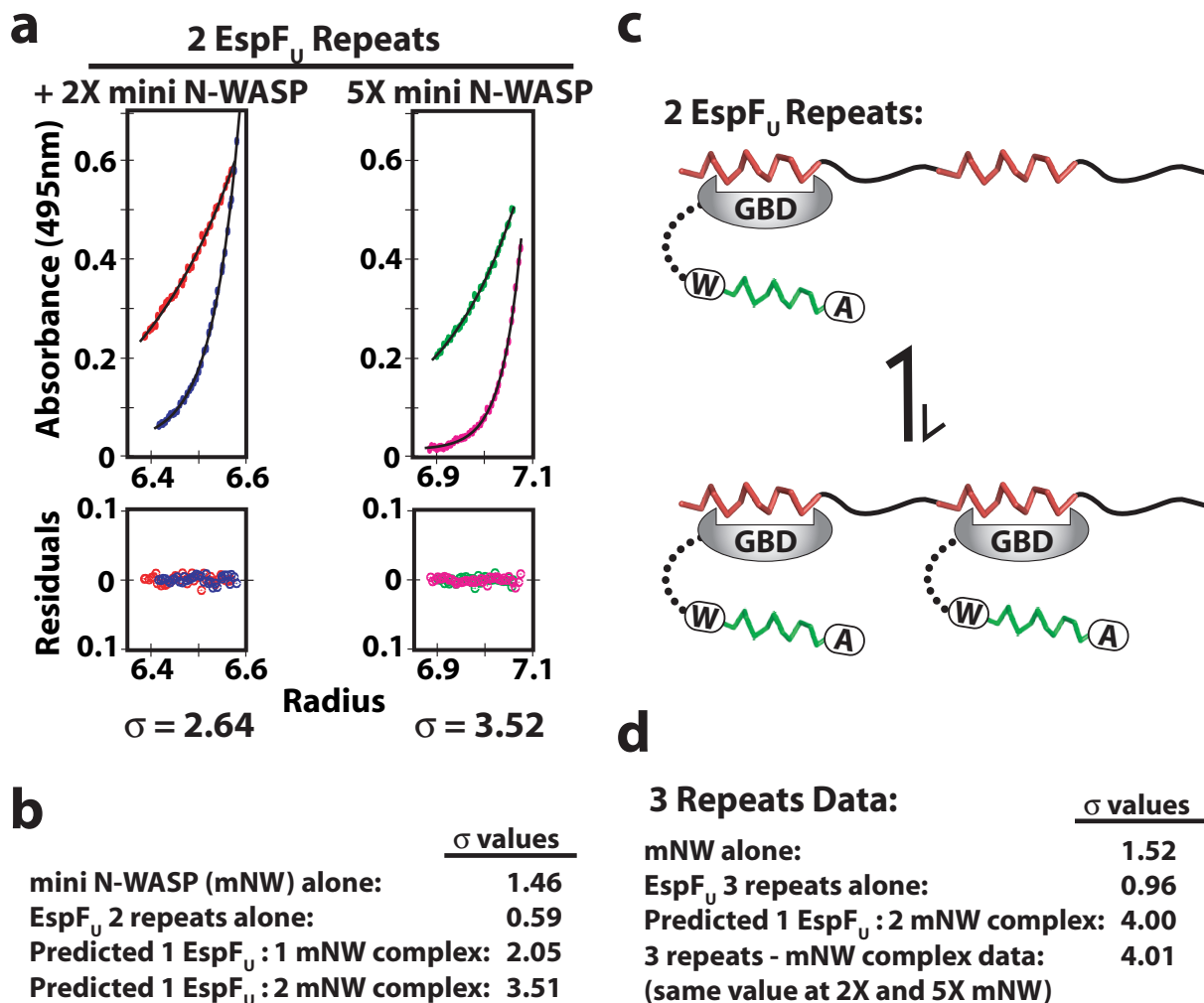|	σ values
mNW alone:	1.52
EspF$_U$ 3 repeats alone:	0.96
Predicted 1 EspF$_U$: 2 mNW complex:	4.00
3 repeats - mNW complex data:	4.01
(same value at 2X and 5X mNW)	

Figure S13 | A two-repeat fragment of EspF$_U$ has high affinity for one mini N-WASP protein, but low affinity for a second. a, Analytical ultracentrifugation (AUC) of the complex of a two-repeat fragment with 2-fold (2X) or 5-fold (5X) excess mini N-WASP (mNW). In each panel, the top curve is at 10,000 rpm and the bottom at 20,000 rpm. Shown below are the residuals for the fits to the data. The two conditions fit to different effective reduced molecular weight (σ) values (bottom). **b,** The individual EspF$_U$ and mNW proteins were analyzed and their σ values are summarized, along with predicted values for complexes. At 2X mNW, the σ value from a is closer to the 1:1 complex. At 5X mNW, we see just the 1:2 complex. **c,** These data show that binding of a second mNW to the two-repeat fragment is disfavored, possibly due to steric effects. This likely explains the lower activity of the two-repeat construct in assays of actin polymerization. **d,** Summary of data from AUC experiments on a three-repeat fragment of EspF$_U$, showing that it binds two mNW proteins at both 2X and 5X mNW. The high activity of three repeats in actin assays correlates with its ability to bind two mNWs more effectively. At even higher concentrations of mNW, we found that the repeats would saturate with mNW proteins (data not shown). We tried to similarly analyze full-length EspF$_U$, but the complex with mNW precipitated in our AUC experiments.