PEAK3/C19orf35 pseudokinase, a new NFK3 kinase family member, inhibits CrkII through dimerization

Mitchell L. Lopez, Megan Lo, Jennifer E. Kung, Małgorzata Dudkiewicz, Gwendolyn M. Jang, John Von Dollen, Jeffrey R. Johnson, Nevan J. Krogan, Krzysztof Pawłowski, Natalia Jura

Supplementary Information (SI) Appendix

www.pnas.org/cgi/doi/10.1073/pnas.1906360116

Supplementary Materials and Methods

Multiple Sequence Alignment. The alignment was constructed using the Promals3D algorithm and manually adjusted using FFAS03 (1) pair-wise alignment results for human PEAK3 and the two homologs PEAK1 and Pragmin. Secondary structure assignments from PDB structure records for Pragmin (PDB ID: 5VE6) and PEAK1 (PDB ID: 6BHC) were added to the final alignment. Low complexity regions in human members of NKF3 family were identified using the SEG server (2).

Phylogenetic tree construction. 29 homologs were selected from vertebrates that possess representatives of the three subfamilies and from four non-vertebrate organisms with single NKF3 homologs (*Trichoplax adherens* (plocozoan)*, Strongylocentrotus purpuratus* (sea urchin), *Branchiostoma belheri* (lancelet), and *Acanthaster planci* (starfish). Multiple sequence alignment of the pseudokinase domains was built using the Promals3D algorithm (3) and was manually adjusted. The phylogenetic tree for was constructed using the PhyML method with aLRT statistics for calculating significance of branches (4). Branches with bootstrap values better than 70% were marked. Ancestral sequence reconstruction was performed using the Ancescon algorithm (5).

Sequence logos. Homologs were collected from the NCBI database by running the BLAST program with kinase domains of human PEAK1, Pragmin and PEAK3 as queries and maximum expect value of 1E-6. Redundancy in the resulting sequence set was removed using CD-hit algorithm at 95% identity level (6). Assignment of the resulting representative sequences to the three NKF3 subfamilies was verified by sequence clustering using the fastNJ algorithm. Multiple sequence alignment obtained using the Promals3D algorithm was split into three subfamily alignments with matched column numbering and used to create sequence logos with the WebLogo3 server (7). The logos included 103 Pragmin homologs, 98 PEAK1 homologs and 35 mammalian PEAK3 sequences. Avian PEAK3 sequences were not included due to substantial heterogeneity of the N-terminal regions.

Plasmids and cell culture. The PEAK3 gene was synthesized by GenScript and subcloned

into pcDNA4/TO. The wild-type CrkII plasmid was a generous gift from Scott Oakes. Mutations and deletions were introduced using Quikchange mutagenesis (Agilent). All constructs were verified via DNA sequencing (Elim Biopharm). HEK293 cells and COS-7 cells were cultured in Dulbecco's modified Eagle media (Life Technologies) supplemented with 10% FBS (Hyclone) and penicillin and streptomycin (Life Technologies). Epitope-tagged constructs were transiently transfected into cells using Lipofectamine 3000 (Invitrogen) according to the manufacturer's protocols. Cells were transfected for 24 hours prior to imaging or cell lysis.

Immunoprecipitation/mass-spectrometry. FLAG immunoprecipitations were done as previously described (8, 9). Specific details are as follows: 293T cells were transfected with 3xFLAG-tagged PEAK3 expression construct using PolyJet Reagent (SignaGen Laboratories) 20-24 hours after plating $1x10^7$ cells per 14.5 cm dish. 40 hours post- transfection, cells were dissociated and washed with 10 ml PBS +/- 10 mM EDTA, respectively, before centrifugation at ≥ 200xg, at 4°C for 5 minutes. Cell pellets were re- suspended in 1 ml of 0.5% Nonidet P-40 Substitute (Fluka Analytical) in IP buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA) supplemented with cOmplete mini EDTA-free protease and PhosSTOP phosphatase inhibitor cocktails (Roche), incubated on a tube rotator at 4°C for 30 minutes, and centrifuged at 3,500xg, 4°C for 20 minutes. Cell lysates, 20 ml anti-FLAG M2 magnetic beads (Sigma-Aldrich), and 2 mg 1xFLAG peptide (Sigma-Aldrich) in 0.3 ml IP buffer were incubated on a tube rotator at 4°C for 2 hours. After binding, FLAG beads were washed with 0.05% Nonidet P-40 Substitute in IP buffer (3 x 1 ml) and transferred to a new tube with a final wash in 1 ml IP buffer. Proteins were eluted by gently agitating FLAG beads with 30 ml of 0.05% RapiGest SF Surfactant (Waters Corporation) in IP buffer on a vortex mixer at room temperature for 30 minutes. FLAG-tagged protein expression and protein immunoprecipitation were assessed by western blot and silver stain, respectively, before submitting 10 ml eluate for mass spectrometry. Three independent biological replicates were performed for FLAG immunoprecipitations.

Mass spectrometry analysis. Purified proteins eluates were digested with trypsin for LC-MS/MS analysis. Samples were denatured and reduced in 2M urea, 10 mM NH4HCO3, 2

mM DTT for 30 min at 60°C, then alkylated with 2 mM iodoacetamide for 45 min at room temperature. Trypsin (Promega) was added at a 1:100 enzyme:substrate ratio and digested overnight at 37°C. Following digestion, samples were concentrated using C18 ZipTips (Millipore) according to the manufacturer's specifications. Digested peptide mixtures were analyzed by LC-MS/MS on a Thermo Scientific Velos Pro dual linear ion trap mass spectrometry system equipped with a Proxeon Easy nLC II high- pressure liquid chromatography and autosampler system. Samples were injected onto a pre-column (2 cm x 100 μm I.D. packed with ReproSil Pur C18 AQ 5 μm particles) in 0.1% formic acid and then separated with a one-hour gradient from 5% to 30% ACN in 0.1% formic acid on an analytical column (10 cm x 75 um I.D. packed with ReproSil Pur C18 AQ 3 μm particles). The mass spectrometer collected data in a data-dependent fashion, collecting one full scan followed by 20 collision-induced dissociation MS/MS scans of the 20 most intense peaks from the full scan. Dynamic exclusion was enabled for 30 seconds with a repeat count of 1. The resulting raw data was matched to protein sequences by the Protein Prospector algorithm (10). Data were searched against a database containing SwissProt Human protein sequences, concatenated to a decoy database where each sequence was randomized in order to estimate the false positive rate. The searches considered a precursor mass tolerance of 1 Da and fragment ion tolerances of 0.8 Da, and considered variable modifications for protein N-terminal acetylation, protein N-terminal acetylation and oxidation, glutamine to pyroglutamate conversion for peptide N-terminal glutamine residues, protein N-terminal methionine loss, protein N-terminal acetylation and methionine loss, and methionine oxidation, and constant modification for carbamidomethyl cysteine. Prospector data was filtered using a maximum protein expectation value of 0.01 and a maximum peptide expectation value of 0.05. Protein interactions were subsequently scored using the CompPASS algorithm (11).

Co-immunoprecipitation and Western blot analysis. HEK293 cells (2.5E6 cells) were seeded on 60 mm dishes and transfected the following day. 24 hours post transfection, cells were washed two times on ice with 1xPBS followed by lysis for 30 minutes on ice (0.5% Triton X-100, 0.5% NP-40, 150 mM NaCl, 50 mM Tris pH8.0, 1 mM NaF, 1 mM $Na(VO_4)_3$, 1 mM EDTA, cOmplete mini EDTA-free protease inhibitor cocktail (Roche)). Cells were scraped and clarified by centrifugation for 10 minutes at 15,000 rpm. The whole cell lysates were pre-

cleared with Protein A beads (Novex) for 30 minutes at 4°C. The pre-cleared lysates were then incubated with antibody/protein A complexed beads overnight at 4°C (anti-FLAG (mouse, Sigma), anti-HA (mouse, SCBT)). The beads were washed three times with lysis buffer. The bound proteins were eluted from the beads using SDS-loading buffer and were boiled at 95°C for 10 minutes prior to SDS/Page and analysis by Western blot. Samples were run on 12% acrylamide gels, and transferred onto PVDF membranes. Membranes were blocked in 5% milk in 1xTBS and 1% Tween-20 (TBS-T), followed by incubation with primary antibodies diluted in blocking buffer (anti- CrkII (rabbit, Proteintech), anti-FLAG (mouse, Sigma Aldrich; rabbit, CST), and anti-HA (mouse, SCBT)), followed by incubation with secondary antibodies diluted in blocking buffer (anti-mouse IgG Veriblot, Abcam; anti-IgG Veriblot, Abcam). ECL Western blotting detection reagent (GE) or ECL Prime (VWR) were used for detection.

Immunofluorescence analysis. COS-7 cells (7.0E4 per well) were plated onto glass coverslips and transfected the following day. 24 hours post transfection, cells were fixed in 3.7% formaldehyde in PBS for 1 hour at room temperature, permeabilized using 0.1% Triton-X in PBS for 5 minutes, and incubated with blocking buffer (1% BSA in PBS) for 5 minutes. Primary antibodies in the blocking buffer (anti-FLAG (rabbit, CST), anti-CrkII (mouse, SCBT)) were added for 1 hour at 37°C, followed by secondary antibodies (Alexa Fluor 568 anti-rabbit IgG (Life Technologies) and Alexa Fluor 488 anti-mouse IgG (Life Technologies)) for 1 hour at 37°C. Actin was stained using Alexa Fluor 674 Phalloidin (Life Technologies) for 20 minutes at 37°C. Images were acquired using a Nikon Elipse Ti equipped with a CSU-X1 spinning disc confocal and Andor Clara interline CCD camera with a Nikon Plan Apo 60X oil lens. The effects on actin stress fibers were blindly scored. Data is reported as a percent of the total (n=60 cells per group in each experiment). Cell perimeter calculations were generated by creating an ROI of each cell using Fiji and a Wacom tablet. Each cell was traced by hand using the Wacom tablet and the perimeter of the cell was calculated from the cell-shape vector. Data was analyzed using the Kruskal-Wallis test for the nonparametric comparison of the means followed by Dunn's Multiple Comparison Test for pairwise comparison between groups (12, 13).

Mapping sequence conservation onto known NKF3 structures. Structure mapping was performed considering sequence conservation within mammalian subfamily of PEAK3

homologs. Alignment conservation values (Jalview 1.18 (14)) were used for residue coloring (lowest conservation value (0) – white, highest conservation value (11) – red). Conservation scores were mapped onto PEAK1 dimer structure using UCSF Chimera (15). Conservation mapping onto the solved crystal structure was performed based on two pairwise alignments: (1) PEAK3 vs PEAK1 and (2) PEAK3 vs Pragmin. Structures for dimers of human Pragmin (PDB ID: 5VE6) and PEAK1 (PDB ID: 6BHC) were rendered using Pymol. Residues conserved in the alignments (strictly conserved or conservative replacements, as judged by positive BLOSUM62 matrix scores (16)), were colored according to BLOSUM62 scores as follows: yellow: BLOSUM62 values from 1 to 3; orange: 4-6; red (highest conservation): 7 to 11.

References

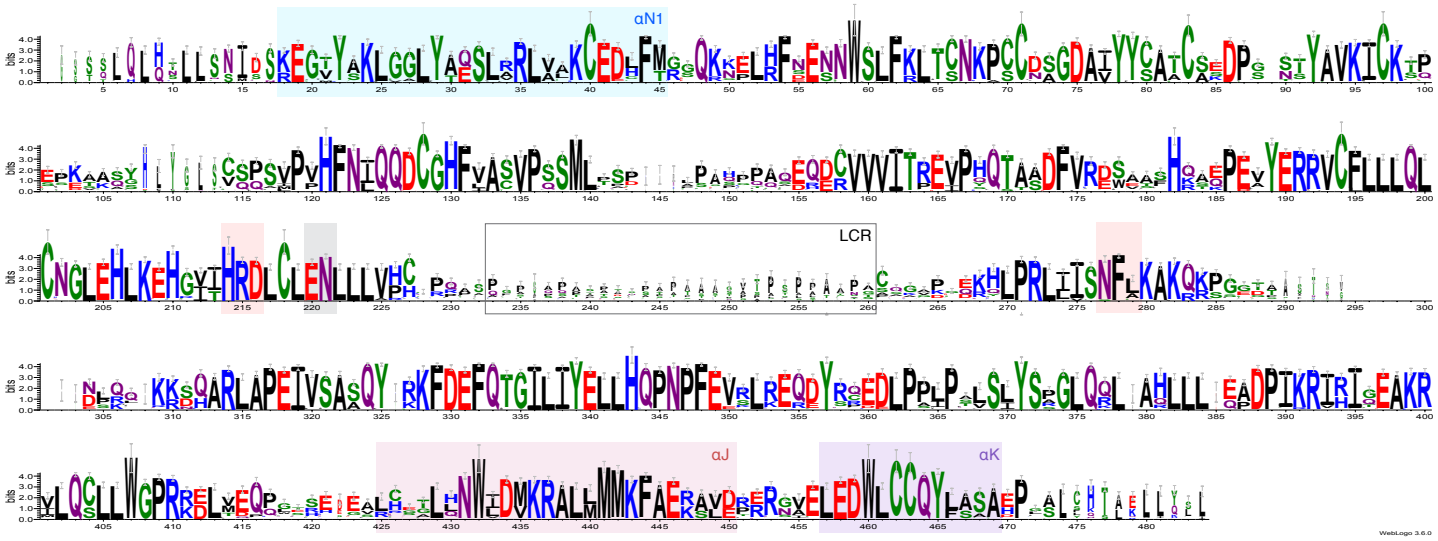
1. Jaroszewski L, Li Z, Cai X-H, Weber C, Godzik A (2011) FFAS server: novel features and applications. *Nucleic Acids Res* 39:W38–44.
2. Wootton JC (1994) Non-globular domains in protein sequences: automated segmentation using complexity measures. *Comput Chem* 18:269–285.
3. Pei J, Kim B-H, Grishin NV (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res* 36:2295–2300.
4. Guindon S, Delsuc F, Dufayard J-F, Gascuel O (2009) Estimating maximum likelihood phylogenies with PhyML. *Methods Mol Biol* 537:113–137.
5. Cai W, Pei J, Grishin NV (2004) Reconstruction of ancestral protein sequences and its applications. *BMC Evol Biol* 4:33.
6. Huang Y, Niu B, Gao Y, Fu L, Li W (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. Bioinformatics 26:680–682.
7. Crooks GE, Hon G, Chandonia J-M, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Res 14:1188–1190.
8. Jäger S et al. (2011) Global landscape of HIV-human protein complexes. Nature 481:365–370.
9. Jäger S et al. (2011) Purification and characterization of HIV-human protein complexes. Methods 53:13–19.
10. Clauser KR, Baker P, Burlingame AL (1999) Role of accurate mass measurement (+/- 10

ppm) in protein identification strategies employing MS or MS/MS and database searching. Anal Chem 71:2871–2882.

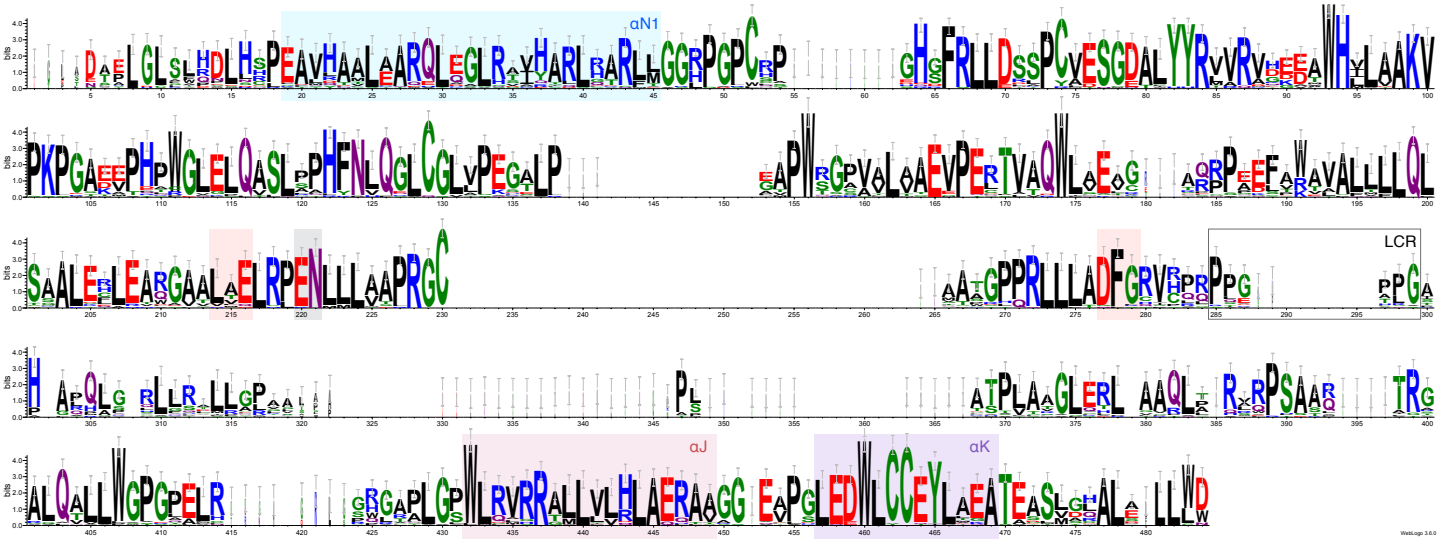
11. Sowa ME, Bennett EJ, Gygi SP, Harper JW (2009) Defining the human deubiquitinating enzyme interaction landscape. Cell 138:389–403.
12. Kruskal WH, Wallis WA (1952) Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association* 47:583–621.
13. Dunn OJ (1961) Multiple Comparisons among Means. *Journal of the American Statistical Association* 56:52–64.
14. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench.Bioinformatics 25:1189–1191.
15. Pettersen EF et al. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612.
16. Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA 89:10915–10919.

Supplementary Figure 1, Lopez, Lo et al.

a Pragmin/SgK223



b PEAK3/C19orf35



c PEAK1/SgK269

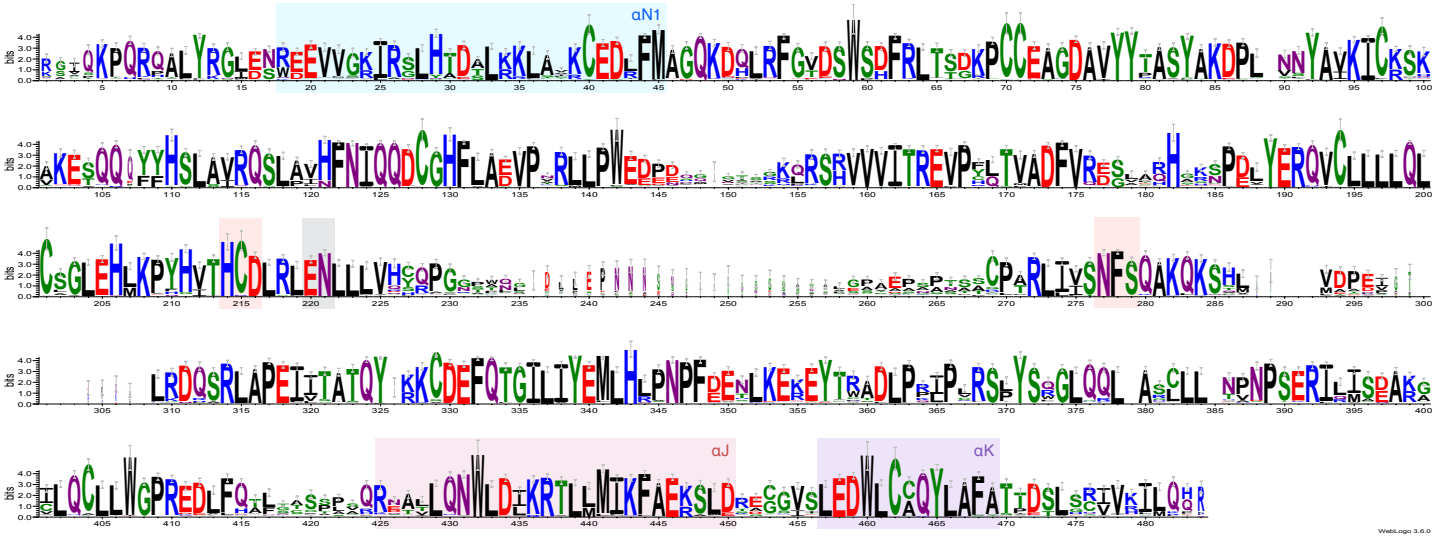


Fig. S1. Sequence conservation in the NKF3 family. Aligned sequence logos for (A)Pragmin/SgK223 sequences, (B) mammalian PEAK3/C19orf35 sequences, and (C)PEAK1/SgK269 sequences. The motifs corresponding to the canonical active site HRD encompass residues 214-216 and the motifs corresponding to the canonical DFG encompass residues 277-279 (highlighted red). The conserved EN motifs encompass residues 220-221 (highlighted grey). The conserved SHED domain helices correspond to residues 18/19-45(αN1 - highlighted blue), residues 425-250 (Pragmin, PEAK1) or 432-449 (PEAK3) (αJ - highlighted magenta), and residues 457-469 (αK - highlighted purple). Key LCR areas are boxed. Multiple sequence alignment (Promals3D) of the NKF3 family split into three subfamily alignments with matched column numbering. Sequence logos created using the WebLogo3 server.

Supplementary Table 1, Lopez, Lo et al.

Hit #	Name	Uniprot ID	ComPASS score
1	C19orf35	Q6ZS72	711.90724
2	c-Crk	P46108	557.91605
3	14-3-3 protein eta	Q04917	161.16351
4	SIAH1	Q8IUQ4	161.05648
5	14-3-3 protein sigma	P31947	140.33994
6	14-3-3 protein beta/alpha	P31946	45.04511
7	14-3-3 protein gamma	P61981	33.21708
8	14-3-3 protein epsilon	P62258	32.00987
9	FGD3	Q5JSP0	31.58578
10	14-3-3 protein zeta/delta	P63104	26.21800
11	14-3-3 protein theta	P27348	23.02257
12	STK17A	Q9UEE5	12.89484
13	CrkL	P46109	11.57874
14	ZFP791	Q3KP31	11.16726
15	HSPA1A/HSPA1B	P08107	9.53939
16	ASAP1	Q9ULH1	9.11803
17	HSP7C	P11142	9.01850
18	Tubulin beta-5 chain	P07437	8.02081
19	HSPA2	P54652	7.74915
20	Tubulin beta-4B chain	P68371	7.59386
21	Tubulin beta-2B chain	Q9BVA1	6.97615
22	Tubulin beta-4A chain	P04350	6.92820
23	Tubulin beta-2A chain	Q13885	6.87992
24	Tubulin alpha-1B chain	P68363	6.75771
25	HSPA1L	P34931	6.73300

Table S1. List of top interactors of PEAK3 as identified by IP/MS. Top interactors of PEAK3 identified by the IP/MS analysis, including their Uniprot ID and abundance score. The details of the analysis are described in Methods.

Supplementary Figure 2, Lopez, Lo et al.

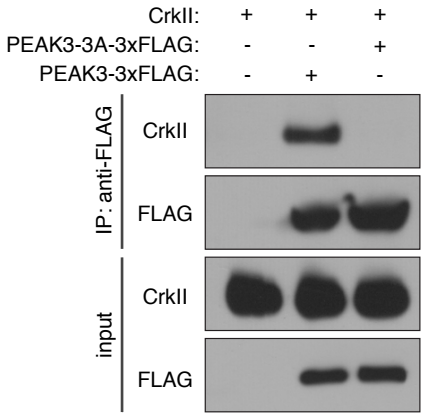


Fig. S2. Exogenously expressed PEAK3 interacts with CrkII. Coimmunoprecipitation of FLAG-tagged wild-type PEAK3, transiently expressed in HEK293 cells, with exogenously expressed, untagged CrkII. Protein levels were detected with the indicated antibodies by Western blot.. Data are representative of three independent experiments.

Supplementary Figure 3, Lopez, Lo et al.

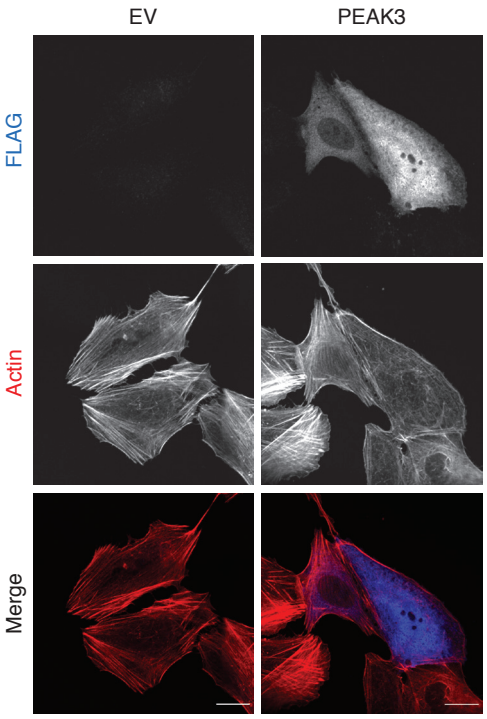


Fig. S3. Exogenous PEAK3 expression leads to a reduction of actin stress fibers. Confocal microscopy imaging of U2OS cells transiently co-transfected with a FLAG-tagged wild-type PEAK3. PEAK3 was detected with anti-FLAG antibody (blue) and F-actin with Alexa Fluor 647-conjugated phalloidin staining (red), (Scale bars, 20 μm).

Supplementary Figure 4, Lopez, Lo et al.

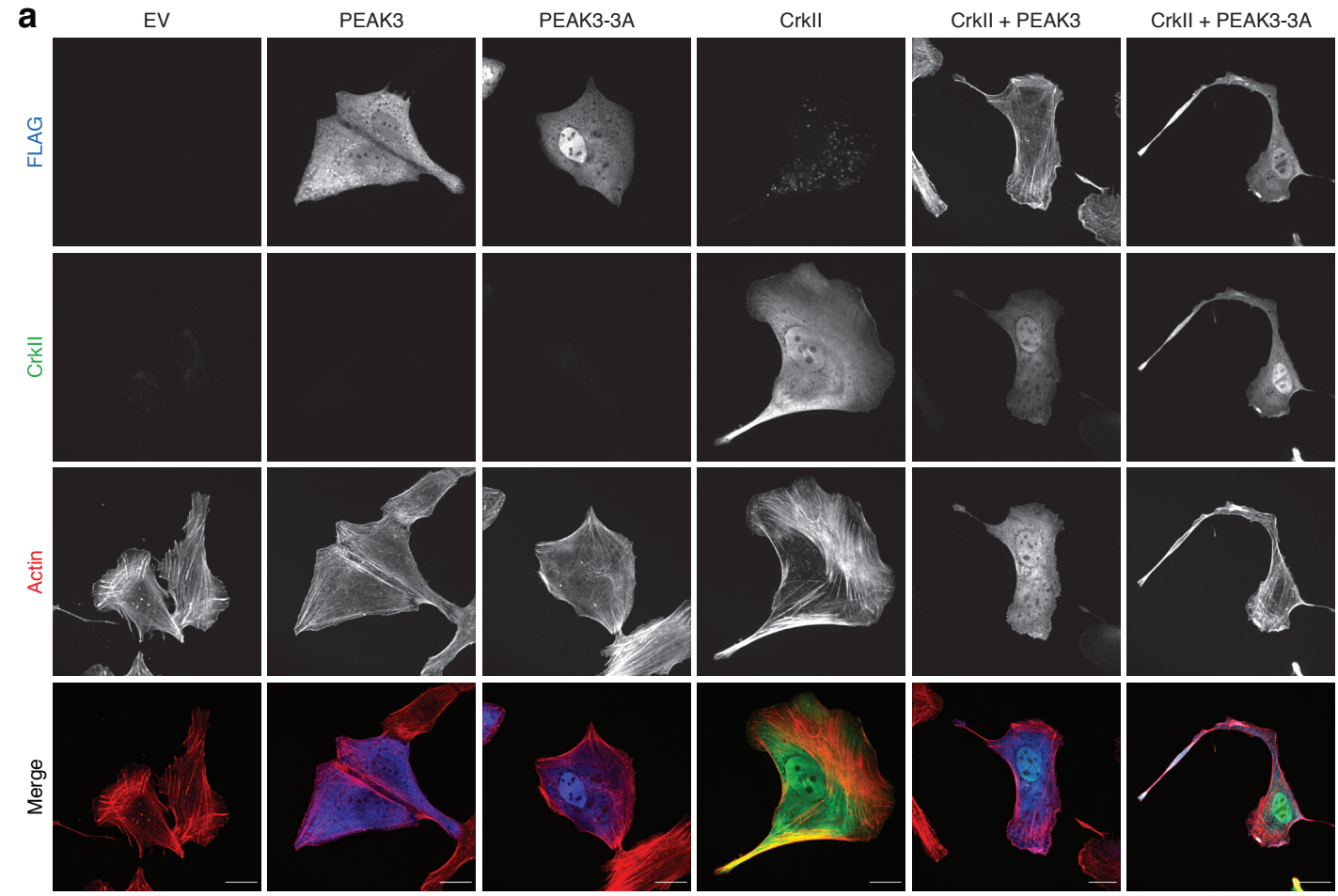


Fig. S4. PEAK3 antagonizes CrkII-dependent cell ruffling in U2OS cells. *(A)* Confocal microscopy imaging of U2OS cells transiently co-transfected with a FLAG-tagged wild-type or a CrkII-binding deficient (3A) mutant of PEAK3 (PEAK3-3A) with either an empty vector or untagged CrkII. CrkII was detected with anti-CrkII antibody (green), PEAK3 with anti-FLAG antibody (blue) and F-actin with Alexa Fluor 647- conjugated phalloidin staining (red). All scale bars correspond to 20 µm. *(B)* Average perimeter of COS7 cells expressing either wild type or a CrkII-binding deficient (PEAK3-3A) variants of PEAK3 with an empty vector or untagged CrkII, quantified as described in Methods. Data represent the mean ± SEM of 3 independent experiments (n = 20 cells in each experiment), **P*< 0.05.

Supplementary Figure 5, Lopez, Lo et al.

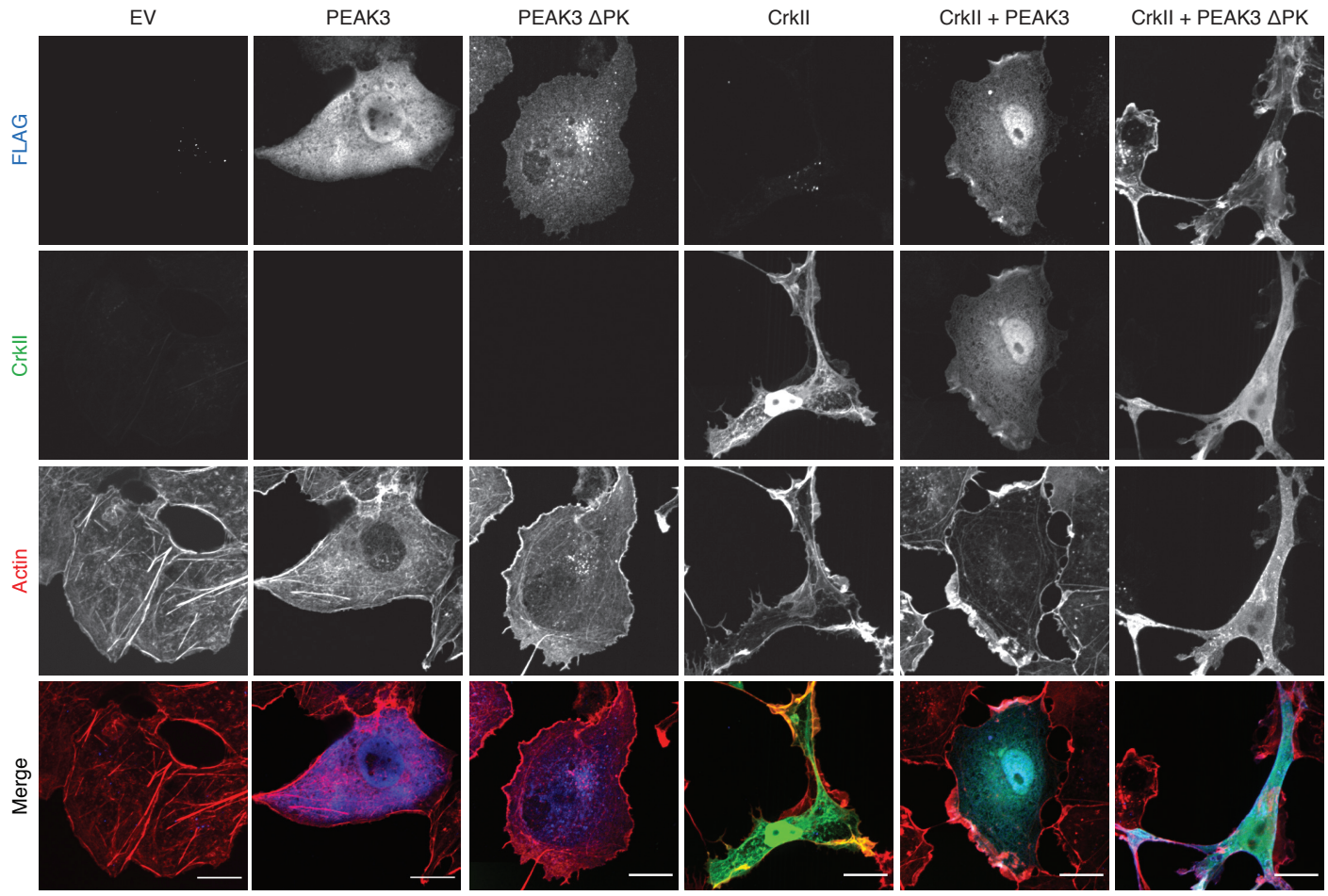


Fig. S5. CrkII binding motif is insufficient to drive inhibition of CrkII function. Confocal microscopy imaging of COS-7 cells transiently co-transfected with a FLAG-tagged wild type or ΔPK mutant of PEAK3 (PEAK3 ΔPK) with either an empty vector or untagged CrkII. CrkII was detected with anti-CrkII antibody (green), PEAK3 with anti-FLAG antibody (blue) and F-actin with Alexa Fluor 647-conjugated phalloidin staining (red), (Scale bars, 20 μm)..

Supplementary Figure 6, Lopez, Lo et al.

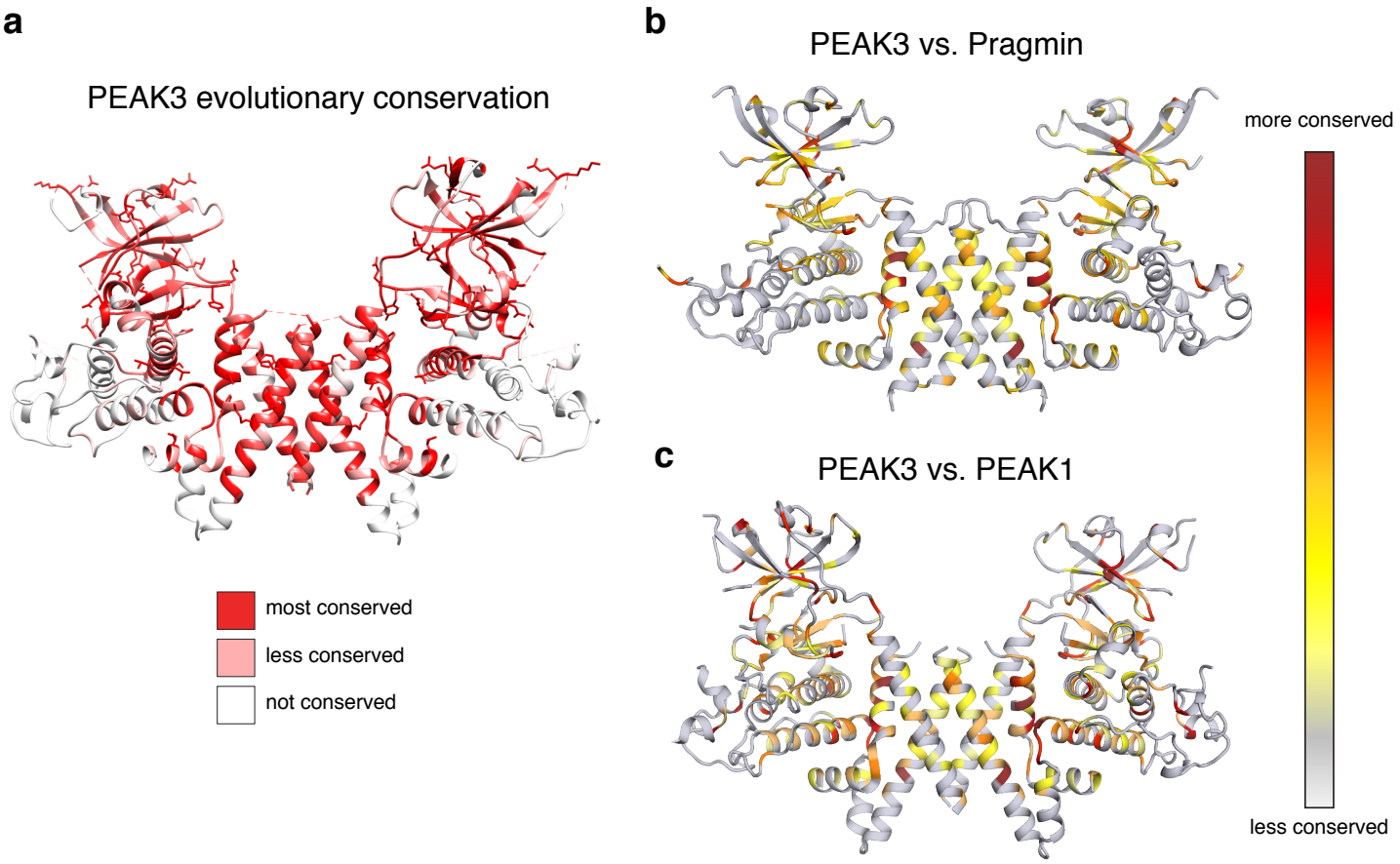


Fig. S6. Sequence conservation in the NKF3 family mapped onto 3D structures. *(A)* Sequence conservation within the identified mammalian PEAK3 homologs across evolution mapped onto the PEAK1 structure. Alignment conservation values are represented by different coloring (lowest conservation = white, highest conservation = red). *(B, C)* Similarity between human Pragmin and PEAK3 *(B)* or human PEAK1 and PEAK3 *(C)* mapped onto the three-dimensional structures (*B*, Pragmin, PDB: 5VE6; *C*, PEAK1, PDB: 6BHC). Pairwise alignments of PEAK3 with Pragmin *(B)* or PEAK1 *(C)* were considered. Residues conserved in the alignments (strictly conserved or conservative replacements, as judged by positive BLOSUM62 matrix scores), rendered as ribbons and colored according to the BLOSUM62 scores as follows: yellow: BLOSUM62 values from 1 to 3; orange: 4-6; red (highest conservation): 7 to 11.

Supplementary Figure 7, Lopez, Lo et al.

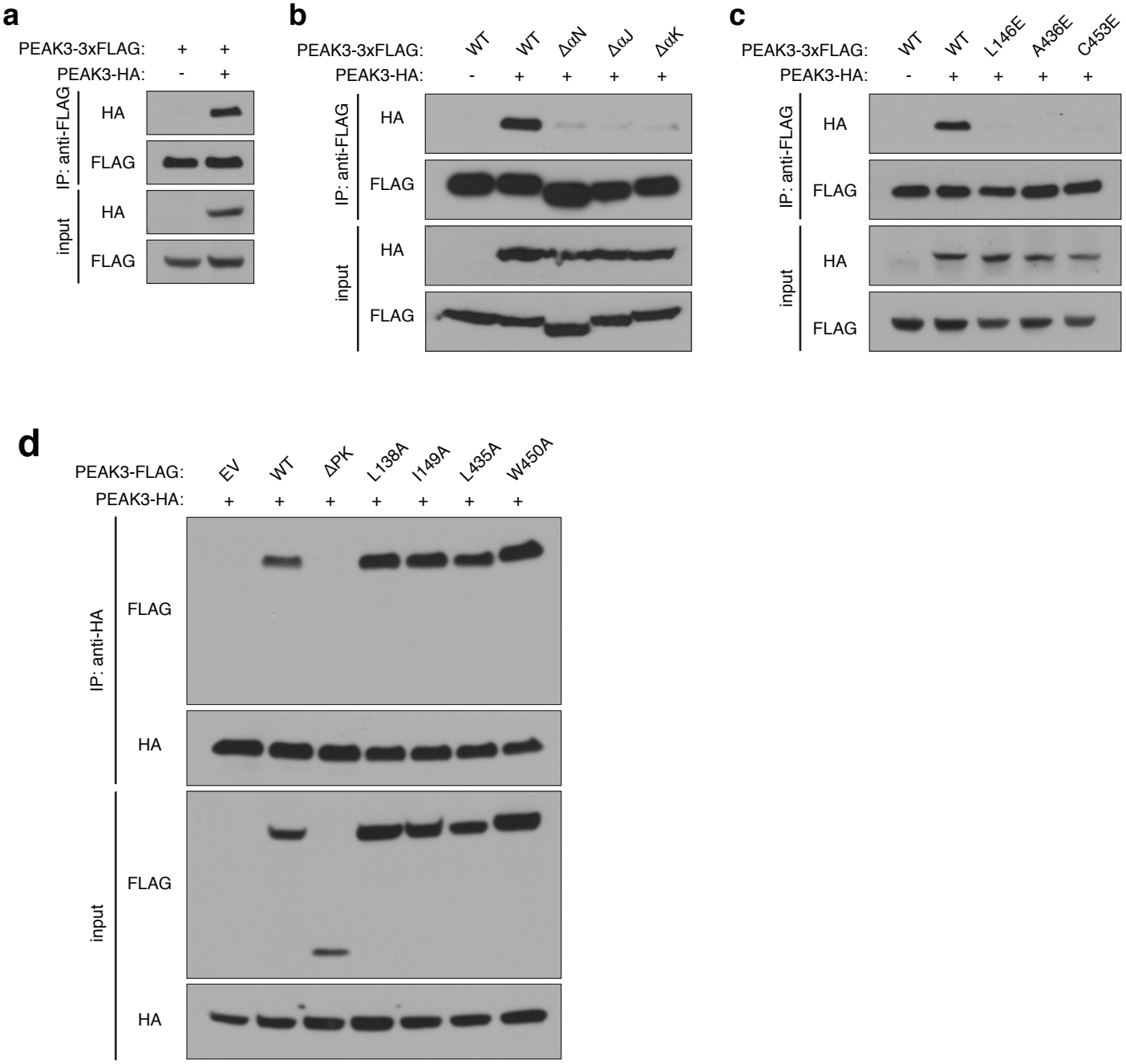
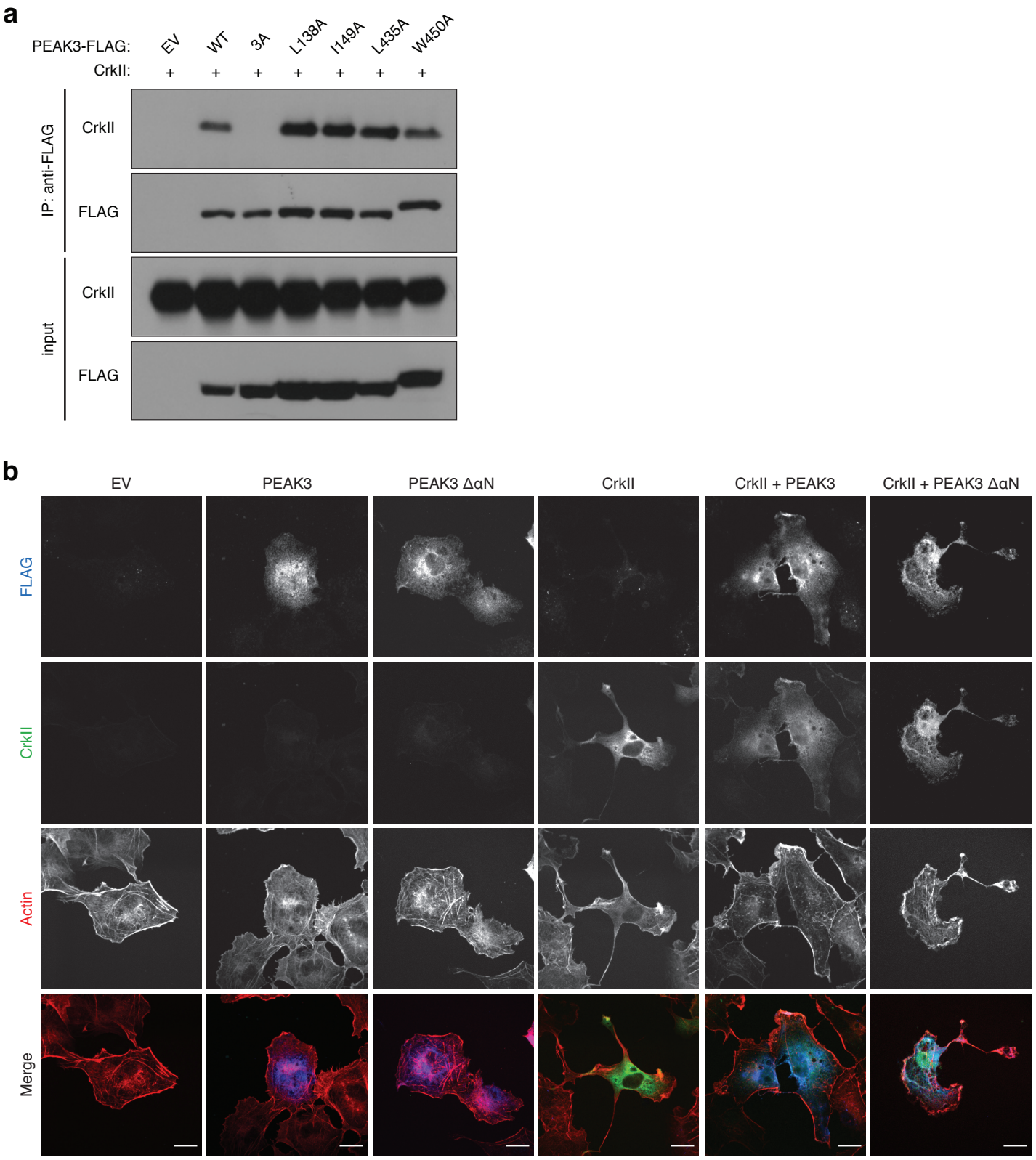


Fig. S7. Mutations within the putative SHED domain disrupts the ability of PEAK3 to dimerize. *(A - C)* Co-immunoprecipitation FLAG-tagged wild-type PEAK3 with either HA-tagged PEAK3 *(A)* or mutant variants where either one of three alpha helices thought to be important for dimerization is deleted *(B)* or residues in the helical interface is mutated *(C)*, transiently expressed in HEK293 cells. Proteins levels were detected with the indicated antibodies by Western blot.. *(D)* Co-immunoprecipitation of an HA-tagged wild-type construct of PEAK3 with a FLAG-tagged wild-type or mutants, transiently expressed in HEK293 cells. Proteins levels were detected with the indicated antibodies by Western blot.. All coimmunoprecipitation data are representative of at two independent experiments.

Supplementary Figure 8a, Lopez, Lo et al.



Supplementary Figure 8c, Lopez, Lo et al.

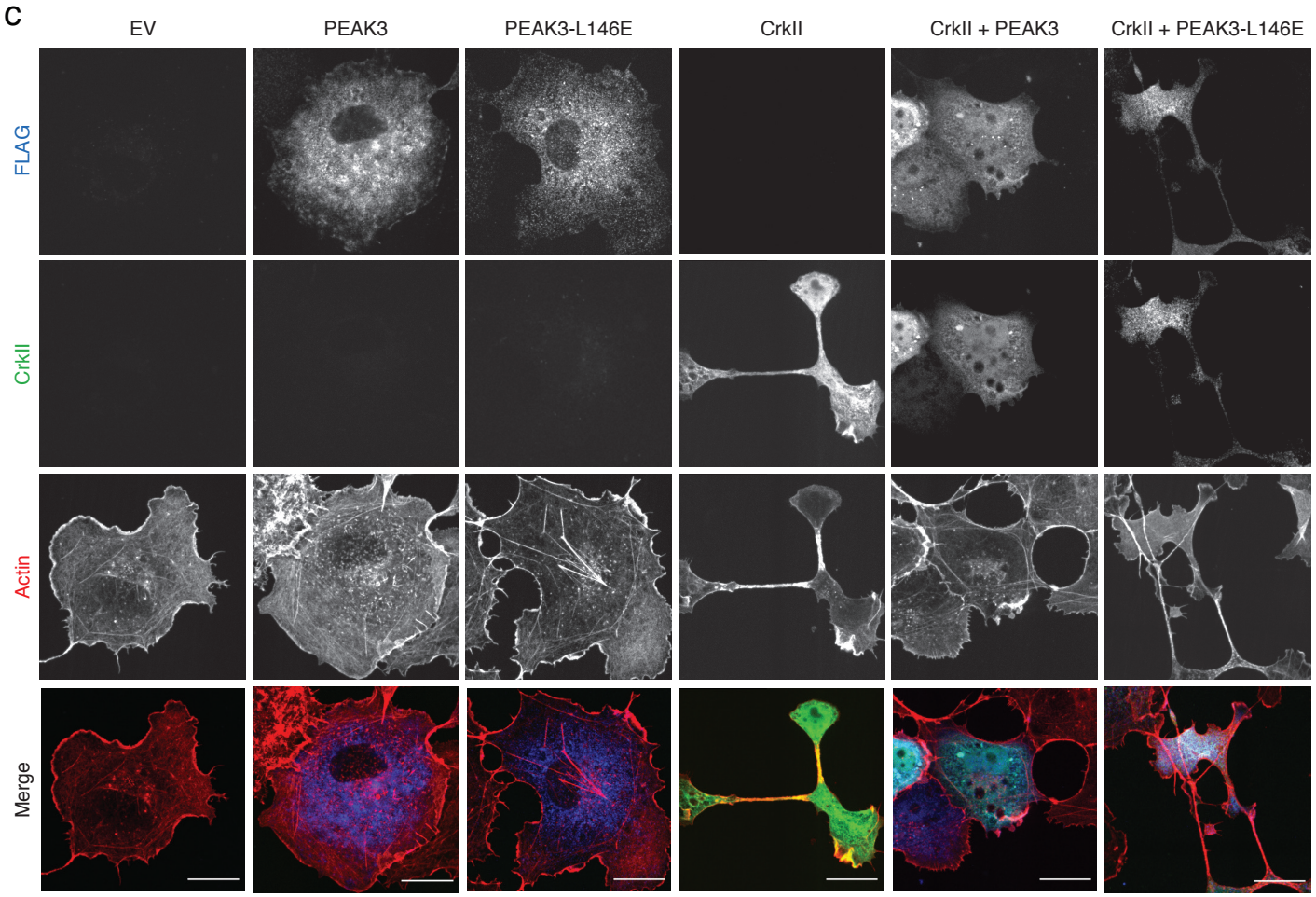


Fig. S8. Homotypic association of PEAK3 is required for its interaction with CrkII. *(A)* Co-immunoprecipitation of untagged CrkII with FLAG-tagged control dimerization mutants of PEAK3, transiently expressed in HEK293 cells. Proteins levels were detected with the indicated antibodies by Western blot.. Data is representative of three independent experiments. *(B)* Confocal microscopy imaging of COS-7 cells transiently co-transfected with a FLAG-tagged wild-type or ΔαN1 mutant of PEAK3 (PEAK3 ΔαN) with either an empty vector or untagged CrkII. CrkII was detected with anti-CrkII antibody (green), PEAK3 with anti-FLAG antibody (blue) and F-actin with Alexa

Supplementary Figure 9, Lopez, Lo et al.

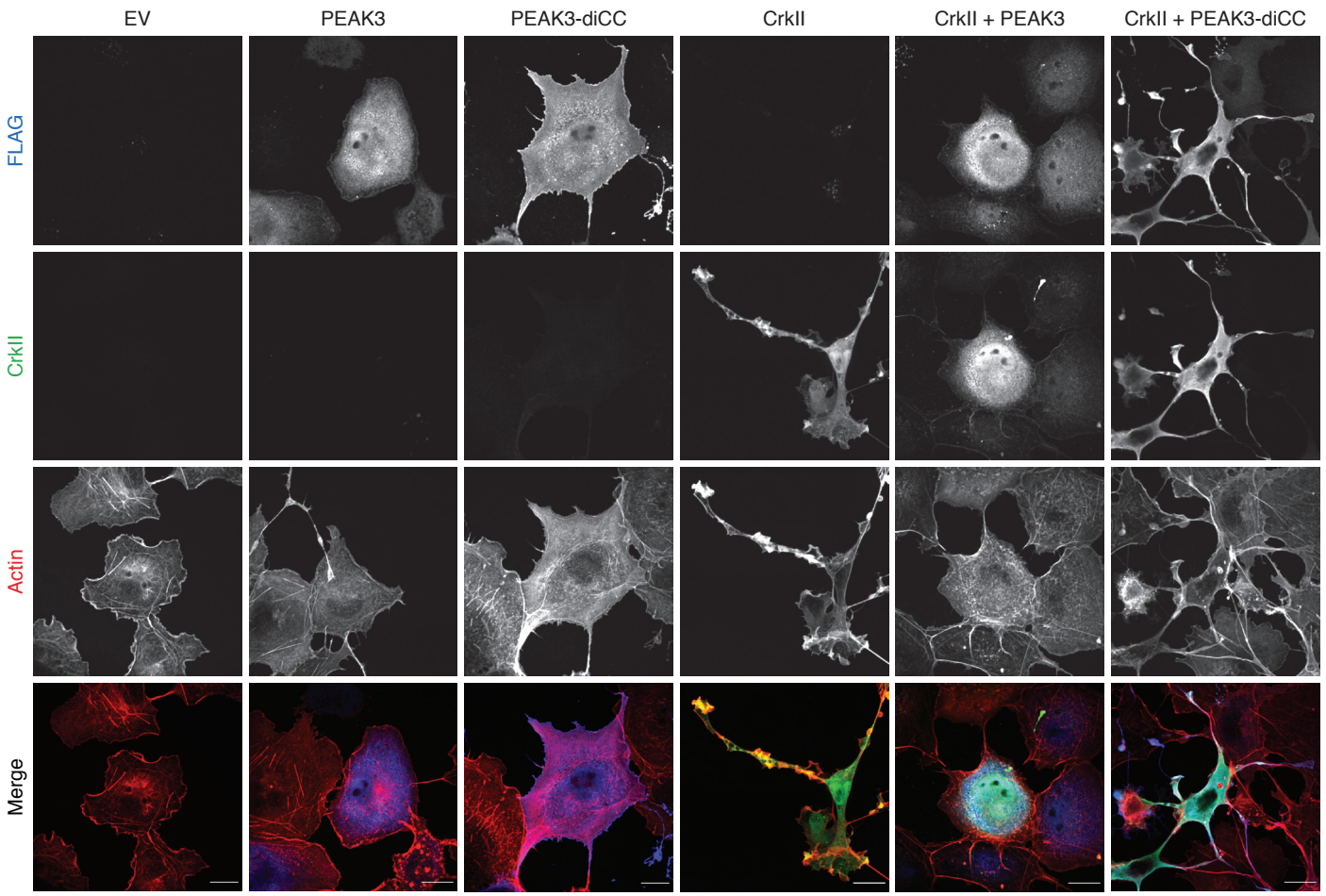


Fig. S9. Constitutive dimerization of the CrkII binding site is not sufficient to antagonize CrkII function. Confocal microscopy imaging of COS-7 cells transiently co-transfected with a FLAG-tagged wild type or diCC mutant of PEAK3 (PEAK3 diCC) with either an empty vector or untagged CrkII. CrkII was detected with anti-CrkII antibody (green), PEAK3 with anti-FLAG antibody (blue) and F-actin with Alexa Fluor 647-conjugated phalloidin staining (red), (Scale bars, 20 μm).

Supplementary Figure 10, Lopez, Lo et al.

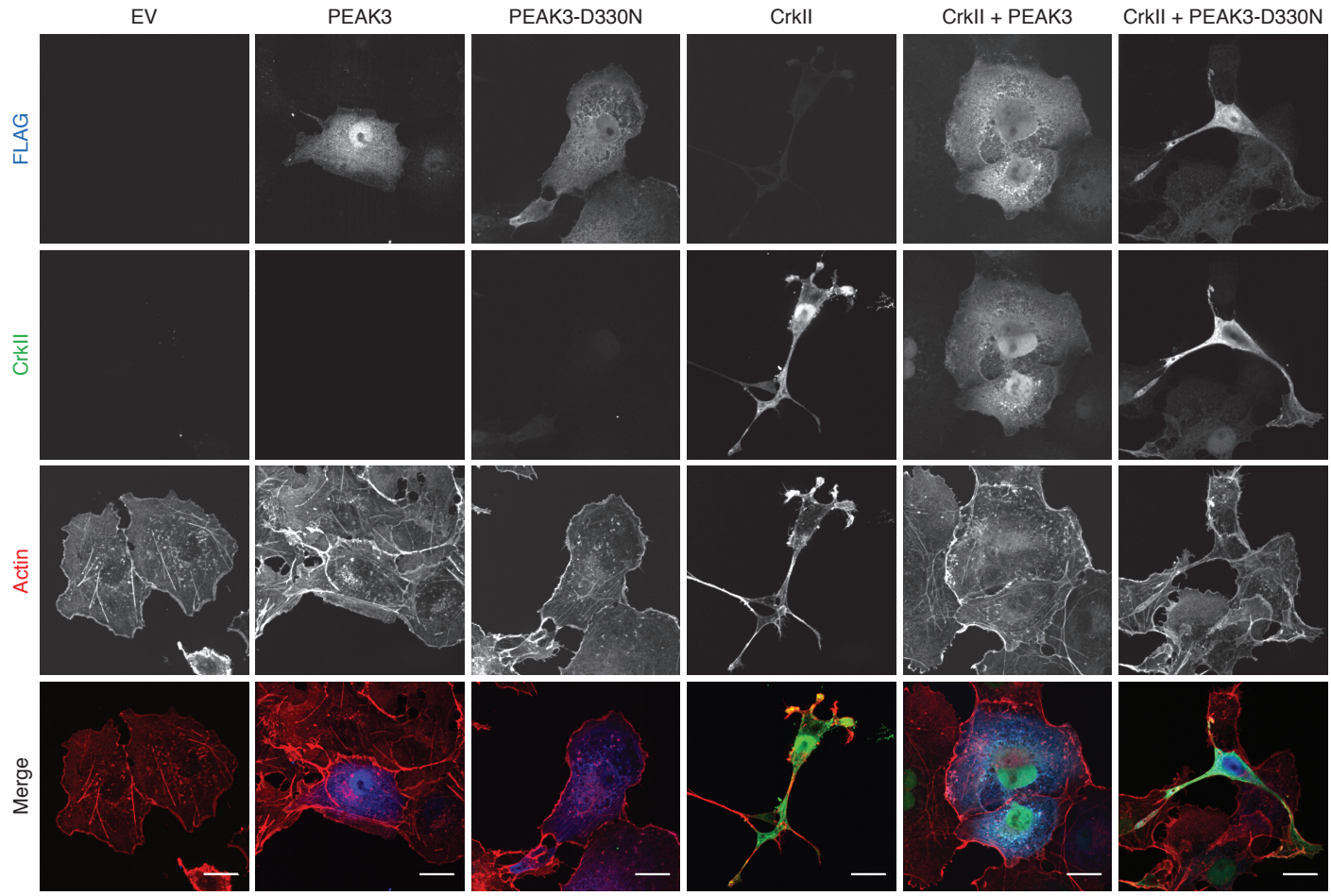


Fig. S 10. Mutation of the DFG motif in PEAK3 diminishes its ability to rescue CrkII-dependent morphology. Confocal microscopy imaging of COS-7 cells transiently co-transfected with a FLAG-tagged wild type or D330N mutant of PEAK3 (PEAK3 D330N) with either an empty vector or untagged CrkII. CrkII was detected with anti- CrkII antibody (green), PEAK3 with anti-FLAG antibody (blue) and F-actin with Alexa Fluor 647-conjugated phalloidin staining (red), (Scale bars, 20 μm).

Supplementary Figure 11, Lopez, Lo et al.

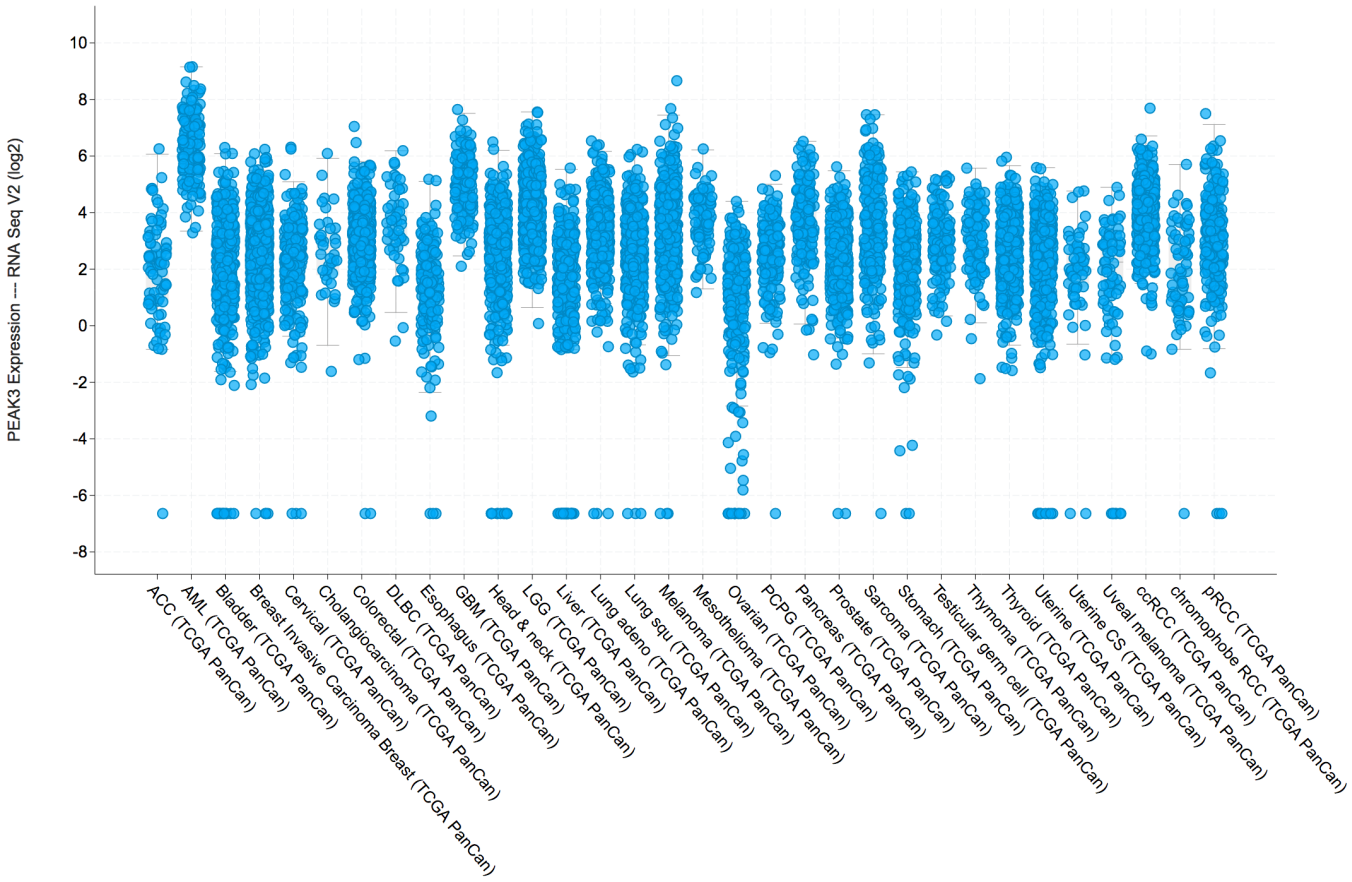


Fig. S11. PEAK3 mRNA levels in patient-derived AML cell lines. PEAK3 mRNA levels in patient-derived cancer cell lines from the Cancer Genome Atlas Project.