

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Functional characterization of PEAK3/C19orf35 pseudokinase and its role in regulation of CrkII-dependent signaling

Permalink

<https://escholarship.org/uc/item/5mc9t9tr>

Author

Lopez, Mitchell

Publication Date

2019

Peer reviewed|Thesis/dissertation

Functional characterization of PEAK3/C19orf35 pseudokinase and its role in regulation of CrkII-dependent signaling

by
Mitchell Lopez

DISSERTATION

Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biochemistry and Molecular Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:

Natalia Jura

Natalia Jura

E438674A382B42F...

Chair

DocuSigned by:

Kevan Shokat

Kevan Shokat

DocuSigned by:

Roshanak Irannejad

Roshanak Irannejad

D006343F4F7646A...

Committee Members

I dedicate this to my family Luis and Renee Lopez, Annette Villa, Karen, Brad, Ivy and VanRyan Shaeffer, Luis and Fumiko Lopez, Joe Rambo, Darleen and Jack Sorg, and Gizmo.

Acknowledgments

Throughout my academic career I have had the great honor to learn from so many amazing people. All of these people, and so many more not mentioned, have seen something in me that I did not. My progression would not have been possible without the guidance, support, mentorship, and patience of so many.

First and foremost I would like to thank my thesis advisor Natalia Jura. You have taught me so much about dedication and compassion. Your drive and unwavering diligence to science inspires all of those that you meet. Thank you for taking your time with me and helping refine myself as a scientist. I appreciate how much you have supported me both in my triumphs and in my follies. I could not have imagined the opportunity to work with such a great scientist and mentor.

I would also like to thank my committee members Kevan Shokat and Roshanak Irannejad. I appreciate the support and guidance starting from my qualifying exam to my final committee meeting. I only wish that I could have had more time and chance to learn from you.

Thank you to all of the members of the Jura lab. You are all such an amazing group of people that has looked out for me in so many ways. Christopher Agnew was my first mentor in the lab and laid the foundation for my development in the lab but also continued to foster a learning relationship throughout our time together. Megan Lo, who is an incredible lab mate and collaborator. I truly appreciate your discipline and kindness as well as all that we have done together. Tarjani Thacker was always there for great advice and late-night comfort food. Jennifer Kung, thank you for the many conversations on pseudokinase biology and listening to my random musings. Karen Ruiz was always there to

give great advice on life with a sage-like deftness or the next podcast you didn't even know you needed. Nicole Michael-Frazier, thank you for kind encouragement. Devan Diwanji, whom I confided in when things were rough and laughed with later. Raph Trenker, whose exuberant sighs and infectious demeanor I will really miss. Erron Titus for our caffeinated rants on the state of the Union and other light-hearted topics. Ed Linossi who always had a cool and calculated approach to science and life, thank you for grounding me when I needed it. Hayarpi Torosyan, for the love of art, fruit snacks, and aliens. Jose Carmona-Negron, who will do great things for science and Porto Rico. Thank you all for the laughs, tears, and good times.

I would also like to thank my classmates! Thank you for taking me in being such a caring cohort. From the beginning you have taught me so much about science but also of companionship and comradery. I am truly amazed at how all of you have progressed and I am excited to see what great thing all of you will do.

To my family, Louis and Renee Lopez, Karen, Brad, Ivy and VanRyan Shaeffer, Luis and Fumiko Lopez, Joe Rambo, Darleen and Jack Sorg, you are the real reason that I am here and was able to succeed. You have believed in me at a time when no one would. I have learned no greater lessons than those I learned from you. Thanks to your love I have been able to flourish and grow into the person I am today.

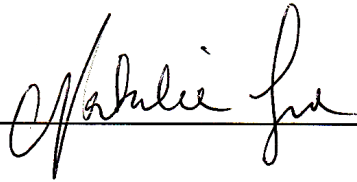
Finally, to the love of my life Annette Villa, thank you for all of your guidance and support. You have taught me love, compassion, and resilience. It is because you believed in me that I embarked on my academic odyssey and you have been my foundation throughout it all. Despite my late-nights in the library and later in the lab, you have stood by me and I am forever grateful.

Contributions

The text of this thesis/dissertation/manuscript is a reprint of the material as it appears in **PEAK3/C19orf35 pseudokinase, a new NFK3 kinase family member, inhibits CrkII through dimerization.**

Mitchell Lopez (MLL) and Megan Lo (ML) contributed to the design of the project, performed the research, analyzed the data, and wrote the manuscript. Specifically, MLL contributed the immunofluorescence microscopy data and analysis.

Natalia Jura, Ph.D.



11/14/19

Date

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

Mitchell Lopez

Abstract

Members of the New Kinase Family 3 (NKF3), PEAK1/SgK269 and Pragmin/SgK223 pseudokinases, have emerged as important regulators of cell motility and cancer progression. Here, we demonstrate for the first time that C19orf35 (PEAK3), a newly identified member of the NKF3 family, is a kinase-like protein evolutionarily conserved across mammals and birds and a novel regulator of cell motility. In contrast to its family members, which promote cell elongation when overexpressed in cells, PEAK3 overexpression does not have an elongating effect on cell shape but instead is associated with loss of actin filaments. Through an unbiased search for PEAK3 binding partners, we identified several regulators of cell motility, including the adaptor protein CrkII. We show that by binding to CrkII, PEAK3 prevents the formation of CrkII-dependent membrane ruffling. This function of PEAK3 is reliant upon its dimerization, which is mediated through a split helical dimerization (SHED) domain conserved among all NKF3 family members. Disruption of the conserved DFG motif in the PEAK3 pseudokinase domain also interferes with its ability to dimerize and subsequently bind CrkII, suggesting that the conformation of the pseudokinase domain might play an important role in PEAK3 signaling. Hence, our data identify PEAK3 as an NKF3 family member with a unique role in cell motility driven by dimerization of its pseudokinase domain.

TABLE OF CONTENTS

INTRODUCTION	1
CONSERVATION OF THE KINASE DOMAIN	1
<i>Conserved motifs and structural elements.....</i>	<i>1</i>
BIOLOGICAL AND STRUCTURAL CHARACTERISTICS OF PSEUDOKINASES	3
<i>Divergence of conserved catalytic motifs and the evolution of pseudokinases.....</i>	<i>3</i>
<i>Structural and functional relationship of the pseudokinase domain.....</i>	<i>5</i>
NEW KINASE FAMILY 3 PSEUDOKINASES.....	9
<i>NKF3 pseudokinases promote cellular migration</i>	<i>10</i>
<i>NKF3 scaffolding function facilitate RTK signaling</i>	<i>11</i>
<i>NKF domain architecture and structural domains.....</i>	<i>13</i>
CRK ROLE IN MOTILITY AND INTRACELLULAR SIGNALING	15
<i>Crk spliceoforms and mechanisms of inhibition</i>	<i>15</i>
<i>Crk-dependent GTPase activation and binding partner specificity</i>	<i>16</i>
<i>CrkL oncogene and it role in chronic myeloid leukemia</i>	<i>18</i>
<i>CrkL and CrkII phenotypic similarities and regulatory differences.....</i>	<i>18</i>
RESULTS	20
PEAK3 IS A DISTINCT MEMBER OF THE NKF3 FAMILY OF ATYPICAL PROTEIN KINASES.....	20
EVOLUTIONARY CONSERVATION OF PEAK3 AND ITS KINASE-LIKE DOMAIN.....	23
PEAK3 INTERACTOME.....	25
PEAK3 INTERACTS WITH CRKII VIA A PROLINE-RICH MOTIF/SH3 DOMAIN INTERACTION.....	26
NEGATIVE REGULATION OF CRKII BY PEAK3 REQUIRES THE C-TERMINAL DOMAIN.....	31

PREDICTED SHED DOMAIN IN PEAK3.....	32
PEAK3 DIMERIZES VIA THE SHED DOMAIN.....	33
DIMERIZATION OF PEAK3 IS NECESSARY FOR CRKII BINDING AND ITS NEGATIVE REGULATION.....	36
MUTATION OF THE DFG ASPARTATE IMPAIRS CRKII REGULATION BY PEAK3.....	39
DISCUSSION	40
MATERIALS AND METHODS.....	46
SUPPLEMENTAL FIGURES.....	53
WORKS CITED.....	67

List of Figures

FIGURE 1	22
FIGURE 2	28
FIGURE 3	30
FIGURE 4	32
FIGURE 5	35
FIGURE 6	38
SUPPLEMENTAL FIGURE 1.....	53
SUPPLEMENTAL FIGURE 2.....	54
SUPPLEMENTAL FIGURE 3.....	55
SUPPLEMENTAL FIGURE 4.....	56
SUPPLEMENTAL FIGURE 5.....	57
SUPPLEMENTAL FIGURE 6.....	58
SUPPLEMENTAL FIGURE 7.....	60
SUPPLEMENTAL FIGURE 8.....	61
SUPPLEMENTAL FIGURE 9.....	63
SUPPLEMENTAL FIGURE 10.....	64
SUPPLEMENTAL FIGURE 11.....	65

List of Tables

SUPPLEMENTAL TABLE 1	66
----------------------------	----

Introduction

Conservation of the kinase domain

Conserved motifs and structural elements

To date there are about 518 different kinases within the human genome, which is about 2% of the human genome. This is an appreciable amount of the genome dedicated to a single enzymatic reaction yet this super family of proteins participates in regulating a vast amount of cellular functions and are vital to development, homeostasis, and fitness. Kinases function as the messengers of the cell whose response can be initiated within seconds, such as the mitogenic MAPK signaling, and sustained, as in mTORC1 which regulates protein synthesis based on nutrient availability within the cell (1). In this manner kinases are the biochemical conduits that give rise to the network of signaling cascades that allow organisms to respond to a vast range of stimuli rapidly and with finite control. The function and regulation of kinases is tightly correlated to the conformation of the kinase domain and the auxiliary domains that accompany it.

The greater majority of eukaryotic kinases share a set of canonical structural features and amino acid motifs and are the defining features of a kinase. The definition of kinases has evolved over time with starting with Hunter and Hanks (2), who used sequence similarities to define a set of canonical motifs that were common amongst the known kinases. The kinome was further defined by Manning and colleagues (3), who utilized the expanding genomic databases to further classify other genes as kinases and atypical kinases. The Protein-Kinase Like fold was established and grouped genes that shared a similar tertiary structure (predicted or experimentally modeled) (2). Within this

classification is the superfamily of Eukaryotic Protein Kinase (ePK), which embodies most kinases, and Atypical Protein Kinases (aPK), which do not have common structural motifs but are different from ePK based on divergence from those canonical motifs that many kinases possess. The annotation of a gene as a potential kinase was based on the presence of conserved catalytic residues, structural characteristics that surround the catalytic site, and/or known catalytic function(2-4). The heterogeneous criteria allowed for the inclusion of proteins with a protein kinase-like fold but are otherwise sequentially divergent in specific defining motifs and also those enzymes that can phosphorylate substrates other than proteins, such as lipids or small molecules.

The key residues and motifs conserved in many kinases that comprise the active site where ATP binds are some of the defining features of many ePK. Not only are these residues highly conserved in many active kinases but also are critical for catalysis. For example, the activation loop near the substrate and ATP binding pocket usually contains phosphorylation site(s) that provides a negative charge that neutralizes other positively charged residues within the active site, specifically the HRD motif. The HRD motif is located in the catalytic loop that precedes the activation loop and serves as a proton acceptor from the hydroxyl group of the substrate during phosphotransfer. Another defining and catalytically important motif is the DFG motif which is just after the activation loop in the P+1 loop. The aspartate of the DFG provides charge support for Mg^{2+} ions that in turn provide stability for the α - and β -phosphates of ATP, as shown in the classic kinase model Protein Kinase A (PKA, PDBID: ATP1). Additionally, the rotamer conformation of the aspartate and phenylalanine residues of the DFG motif also cause a shift in the activation loop that allow substrates to access to the catalytic site of the kinase (5). These motifs and

loops are critical for ePK function and can have detrimental or hyperactivating effects on the catalytic function of kinases when they are mutated.

Biological and structural characteristics of pseudokinases

Divergence of conserved catalytic motifs and the evolution of pseudokinases

Within ePK and spread across the kinome families are proteins that have mutations in their catalytic sites yet still retain enough sequence similarities to be considered part of ePK and not aPK. These proteins are known as pseudokinases and are defined as having a protein-kinase like fold but have mutations in one or more catalytic residues including the catalytic lysine (Lys72), HxD (Asp166) motif, DFG (D184) motif and possibly mutations within the G-loop or activation loop (4). There are about 50 pseudokinases and can be found across the major families of kinome, while some are so unique that they form their own families of containing just a few members (4). The origin of pseudokinases had been hindered due to a lack of complete genomes from a diverse range of organisms outside of Eukaryota and across evolution but new global sequencing efforts provide insights about the origin of pseudokinases as well as their sequence diversity and function.

Pseudokinases have evolved multiple times in the course of evolution but the question remains whether they are derived from a once active enzyme or was it always a catalytically inactive protein. Phylogenetics reveals that pseudokinases can be found in practically every domain of life though one must recognize that many microbial kinases are quite different from eukaryotic kinases (6, 7). Pseudokinases represent 10% of all ePK, which suggests that they are not mere vestigial genes waiting to drift away. Interestingly, this percentage is a reoccurring value for many vertebrate organisms in that 10% of an

organism's kinome will be comprised of pseudokinases (6). Pseudokinases have emerged multiple time across evolution which suggests that the evolutionary pressures to maintain the otherwise highly conserved catalytic residues can relax in favor of other functions that may be specific for a given species' needs (8). As the physiology of organisms diversified and new biochemical processes developed, so did the diversity of pseudokinases.

Pseudokinases can be found as far back as bacteria and archaea but are most prevalent in eukaryotes. In phylogenetic analysis, many pseudokinases tend to group based on its given kingdom or even taxonomic domain e.g. pseudokinase families that were grouped based on sequence homology also tended to group by taxa as well. In some extreme cases entire pseudokinase families were found primarily in a single species, which can be seen for the Rig1 to Rig3 families in *Rhizophagus irregularis*. However *R. irregularis* is quite unique and interesting because 32% of its kinome are pseudokinases, which is an expansive kinome compared to other fungi analyzed (6). While some pseudokinase families are only extant in specific phyla, there are those examples that span across different taxonomical domains. For example the Two-Component System (TCS) pseudokinase family can be found in a broad range of bacteria, protists, fungi, and the phylum Aveolata (i.e. *Toxoplasma gondii*). While there is not a clear functional connection between these proteins, as many have not been functionally characterized, it is a clear example of a family of pseudokinases that have persisted through the tortuous path of evolution.

Structural and functional relationship of the pseudokinase domain

Both kinases and pseudokinases both have important functions that are independent of catalysis but rather rely on protein-protein interactions. These functions are critical for regulating a broad range of processes yet how pseudokinases are regulated is diverse, and in some cases, still remains elusive. However some hypothesize that the structural elements that are unique to protein kinases can offer some insight into how both the noncatalytic functions of kinases and pseudokinases are regulated.

The catalytic activity of many predicted pseudokinases has yet to be determined as there are many factors that contribute to the activity of a kinase including substrate specificity and cofactors. However Murphy and colleagues have proposed a methodology for determining if a pseudokinase can bind ATP or divalent ions using a thermal shift assay. This assay relies on the added thermal stability gained by binding ATP (or ATP analogs) and divalent cations and is an orthogonal assay to classic *in vitro* kinase assay because it does not rely on knowing the sequence of the specific substrate, which can be challenging (9). They conducted a survey of 35 different kinases to determine if the pseudokinases gained thermal stability when incubated with Mg^{2+}/Mn^{2+} , ATP, and some analogs, as well as some promiscuous drugs that bind in the catalytic site using recombinant protein (9). The pseudokinases tested were grouped into the following classes: those that cannot bind ATP or Mg^{2+}/Mn^{2+} (Class 1), ATP alone (Class 2), Mg^{2+}/Mn^{2+} only (Class 3), and those that bind both (Class 4). While this assay does not determine that an enzyme is catalytically active, it provides important biochemical information that is independent of cofactors or substrate specificity.

Many of the pseudokinases tested to date are catalytically inactive yet there are those that have been shown to be active or require ATP or Mg^{2+} for allosteric purposes. The well documented Her3 pseudokinase of the Epithelial Growth Factor Receptor (EGFR) family has been shown to bind both ATP and cations as well has been show to have some autophosphorylation activity despite lacking the catalytic aspartate of the HRD motif (HRN⁸³⁴) (9, 10). The catalytic activity of Her3 is quite low and may not be biologically relevant whereas its ability to allosterically activate other EGFR kinases is a much more potent role for Her3 (5). EGFR family kinases exemplifies how the noncatalytic functions of kinase can be just as vital as their catalytic function (reviewed in (11)).

With no lysine (K) or WNK 1-4 pseudokinases have alterations within the active site that accommodate for its specific function as a ion sensor. WNK1 received its namesake due to mutations in the $\beta 3$ lysine that participates in coordinating ATP (12), but has retained both the conserved DLG motif and HRD motif. The catalytic lysine in the $\beta 3$ has been instead compensated for by a lysine (K233) in the $\beta 2$ strand and functions similarly to the catalytic lysine in other kinases; K233 is thought to stabilizes ATP in the pseudoactive site (12). While other ion-sensing kinases, such as calcium calmodulin-dependent kinase, utilizes a separate calmodulin subunit, to bind and react to changes in calcium, WNK1 binds halides in the pseudokinase domain in a position where the catalytic lysine would typically point near the activation loop DLG. Thus the lack of $\beta 3$ lysine allows for halides to bind to WNK which causes dramatic changes in the conformation of WNK that make it inactive (13). WNK kinases are an example where the pseudokinase domain has taken on various alternative functions (i.e. binding halides), which cause global changes in the pseudokinase domain. Those changes in conformation found in Wnk pseudokinases show how critical the

conformation of the pseudokinase domain is important for catalysis yet there are others whose noncatalytic activity is dependent on the conformation of the pseudokinase domain.

The conformation of the pseudokinase domain of mixed lineage kinase domain like (MLKL) pseudokinase enables oligomerization both with itself as well as with receptor-interacting serine/threonine-protein kinase 3 (RIPK3). MLKL pseudokinase forms a homotetramer via an adjacent helical bundle (4HB) domain, but also forms an octomer with 4 units of RIPK3, which together form the necrosome holoenzyme that facilitates necroptosis. The activated holoenzyme is able to puncture holes into the membrane during necroptosis (14). While MLKL is devoid of Mg^{2+} binding (9), mutations to aspartate 351 (E351K) of GFE (DFG in PKA) inhibits MLKL ability to form a tetramer in solution as well as reduce its ability to puncture lipid membranes (14). This mutation causes an increased binding to ATP, which is inhibitory for MLKL oligomerization and function, and is also thought to change the conformation of the activation loop. Also within the activation loop are residues T357/S358 that when phosphorylated by RIPK3 cause changes in the conformation of the pseudokinase domain that allow the release the 4HB and leads to oligomerization and necroptosis (14, 15). MLKL exemplifies how the conformation of the pseudokinase domain is utilized as a mechanism of regulation and also how this can change the oligomeric state that is essential for binding to other proteins.

Kinase and pseudokinase activity and biological potency is dependent on the alignment of conserved hydrophobic residues known as spines. These residues are well conserved and their alignment results in specific conformational states with potentially unique biological capabilities. This is analogous to GTPases, which upon binding to GTP change conformation and is integral for its function (16-18). Similarly, some kinases have

non-catalytic functions, such as catalysis or allostery, that require a specific conformation that is achieved by binding to ATP (11). Hydrophobic residues from both the N and C-lobe coordinate into vertical linear columns or spines (19-21). The catalytic spine is created when the aromatic adenosine ring aligns in the back of the active site with other hydrophobic residues that align into a column or spine (19)(Reviewed in (22)). Additionally, the regulatory spine is near the front of the active site and aligns another set of hydrophobic residues that also form a spine that include the DFG-Phe and the H/YRD-Tyr. The alignment of these spines is critical for the catalytic and noncatalytic functions of protein kinases as shown with KSR; however structural models of some pseudokinases show how these spines are disrupted and are incapable of catalysis.

Kinase suppressor of Ras (KSR) is a pseudokinase with low catalytic ability and serves as a scaffold for RAF and MEK in RAS/MAP Kinase signaling. KSR participates in MAPK signaling by forming a side-by-side heterodimer with RAF kinases that allosterically activates RAF leading to increased MEK phosphorylation (23). KSR lacks the catalytic lysine (PKA K72) yet still retains some catalytic ability that has been shown in crystallographic models of KSR2 and *in vivo* (19, 24). When phenylalanine, an aromatic hydrophobic residue, is substituted for the alanine (A587F) in the catalytic spine, it mimics the adenosine aromatic ring of ATP and causes a conformation that is in the closed yet active state. This allows for the analysis of the noncatalytic functions of the pseudokinase domain can be parsed from the catalytic functions without compromising the specific conformation of an active kinase. The A587F mutation caused KRS to form a constitutive dimer with BRAF, which is not true for the WT KSR, yet was unable to cause phosphorylation of MEK showing the relevance of the catalytic activity of KSR in MAPK signaling (19). Further more

it also speaks to how the conformation of the pseudokinase domain can have dramatic effects on its ability to interact with other proteins, which in the case of KSR A587F, 'locks' the catalytic spine and causes KSR to adopt a conformation that allows for a constitutive dimer with BRAF.

New Kinase Family 3 pseudokinases

The focus of my thesis pertains to a previously undescribed pseudokinase, C19orf35/PEAK3, and how it regulates cell morphology in the context of CrkII-dependent signaling. PEAK3, or C19orf35 as it was called until recently, had escaped annotation as a protein-kinase like macromolecule during the first edition of the kinome annotated by Manning and colleagues, possibly due to mutations of key catalytic residues, the addition of the SHED domain, and/or some low complexity insertions within the pseudokinase domain. While PEAK3 is predicted to have a protein-kinase like fold based on sequence similarity to the New Kinase Family 3 (NKF3) pseudokinases, the tertiary conformation of the domain I am referring to as the pseudokinase domain has not been solved. Our collaborator Krzysztof Pawłowski from the Warsaw University of Life Sciences utilized the Fold and Function Assignment algorithm(25) to identify C19orf35 as a putative pseudokinase based on its predicted tertiary structure, which is similar to pseudokinases PEAK1 and Pragmin. My interest in this protein was focused on deciphering the cellular processes that PEAK3/C19orf35 participates in, which domains are required for specific functions, and what protein binding partners PEAK3 interacts with to mediate that function. Moreover, my work puts PEAK3 into context with the growing body of work

pertaining to the NKF3 pseudokinases and provides the first functional characterization of PEAK3.

NKF3 pseudokinases promote cellular migration

Both PEAK1 and Pragmin are known to promote cell migration and participate in RTK signal transduction. The first studies of an NKF3 member characterized Pragmin as an effector in RhoA signaling and promoted RhoA activation (26). The increased activation of RhoA caused cells to shrink in size and inhibited neurite outgrowth. However, the role of Pragmin may be cell-type specific because others have shown that overexpression of Pragmin in gastric carcinoma, as well as pancreatic duct epithelial cells, causes cell elongation and an increase in their motile propensity (27), (28). While the mechanism and specific pathway has yet to be fully resolved, some have suggested that Pragmin sequesters C-Src kinase (Csk) away from the membrane, reducing the levels of inhibitory phosphorylation on Src Family kinases (SFK), which is agonistic towards cell motility (29). However others have found that interaction between Pragmin and Csk leads to an increase in the phosphorylation of downstream target proteins that is completely Src-independent (30). Furthermore, others have shown that overexpression of Pragmin leads to a decrease in phosphorylation of Src activation loop coupled with an increase in the inhibitory pY530, further corroborating that Pragmin promotes cell motility in a Src-independent manner (27). These conflicting results could be due to the molecular background of the many cell types used in these experiments yet one resounding result across many studies is the morphological changes that are caused by overexpression of Pragmin, which mimic those seen for PEAK1.

NKF3 scaffolding function facilitate RTK signaling

PEAK1 serves as an integral scaffold that mediates changes in RTK signaling output that promotes cell motility. PEAK1 was first discovered from phosphotyrosine enriched lysates isolated from pseudopodia, followed by 2D-liquid chromatography and tandem mass spectrometry, a process developed by John Yates' group known as MudPIT (31). Work by Wang and colleagues, followed by groups led by Richard Klemke, Jonathan Kelber, and Roger Daly, all made seminal discoveries in elucidating the role of PEAK1 in promoting cell migration and regulating gene expression. PEAK1 serve as a lynchpin for transitioning between the pro-mitogenic signaling complex to the pro-migratory complex downstream of EGF stimulation and PEAK1 bound to Shc is a hallmark of late-stage EGF signaling (32). There is tight temporal regulation of the PEAK1-Shc complex that only forms after sustained mitogen stimulation (eg. EGF), and is also dependent on PEAK1 phosphorylation, which only occurs 20 minutes after sustained stimulation (1, 32). These studies and more have highlighted how PEAK1 functions as a scaffolding protein that alters the effect of RTK signaling in a temporally regulated fashion.

The cellular response to mitogenic factors and cytokines can alter depending on the levels of PEAK1 expression and increase cells abilities to migrate and activate non-canonical signaling pathways. According to the Human Cancer Atlas and a growing body of research, PEAK1 is highly expressed in breast, pancreatic, and other types of cancer samples and also correlates with high-grade cancer, which presents physical hallmarks such as poorly differentiated abnormal tissue morphology and high in mesenchymal molecular markers (33-35). In congruence with the clinical observations, PEAK1 expression can promote cell proliferation and migration in various cancer models including

breast, non-small-cell lung carcinoma, and pancreatic cancers, all of which show increases in metastatic behavior and markers (34-37). PEA1 also increases Src activity by promoting activation loop phosphorylation (Tyr416) in the activation loop, as well PEA1 expression increases resistance against Src inhibitor AZM 475271 in these same cells (38). Therefore, increasing PEA1 expression is a mechanism by which cancer cells can bypass cell cycle regulation checkpoints and increase in proliferation.

Pragmin also increases mitogenic cell signaling but instead does it through different signaling cascades. While overexpression of Pragmin also causes an increase in cell proliferation and migration, some evidence point to a dependence on JAK/Stat activation more so than SFK (27). Inhibition of Stat3 completely abolishes the pro-migratory and proliferative effects of Pragmin where as inhibition of EGFR or Src had little effect on Stat3 phosphorylation, suggesting that Pragmin is able to sustain Stat3 activation regardless of EGFR and Src activation (27). Pragmin and Src signaling do have some feedback in that Pragmin does increase the levels of the inhibitory pY530 in Src due to potentiation of C-terminal Src Kinase (Csk) activity (29). However the mechanism and cellular effects of the increased Csk activity do have some confounding opinions, where some suggest a mechanism of sequestering Csk away from the membrane, others point to an increased focal adhesion turnover due to increased Csk-Pragmin signaling or RhoA GTPase mediated processes that promote cell migration (28-30). Taken together we find that PEA1 and Pragmin both have serious clinical implications due to their ability to imbue cancer cells with RTK inhibitor resistance and promote epithelial-to-mesenchymal transition (EMT).

PEA1 and Pragmin both promote cell migration by promoting actin polymerization and focal adhesion stability. Focal adhesions are large dynamic protein complexes that

fasten cells to their substrate via integrin, which is a membrane protein that binds the extracellular matrix, as well as paxillin and p130Cas adapter proteins on the intracellular side of the membrane that promote focal adhesion assembly (reviewed in (39)). PEAK1 promotes cell motility by increasing paxillin and p130Cas phosphorylation as well as binding to Crk proteins that promote actin polymerization via Rho GTPases Rac1 and Rap1 (31, 40, 41). PEAK1 knockdown causes dysregulation of focal adhesions by decreasing their lifetime yet also decreases the rate of dissociation of the complex (42). These results show how PEAK1 function is to regulate the dynamics of focal adhesions and reducing the cell's ability to quickly form and disassemble these complexes.

NKF domain architecture and structural domains

Together, PEAK1 and Pragmin work in concert to promote motility that is dependent on forming homotypic and heterotypic oligomers. Proteomic studies mentioned above have found PEAK1 and Pragmin are part of the same signaling complex and independent studies have all shown that PEAK1 and Pragmin often promote the same phenotypic changes in cell morphology. It was only until recently that the importance of oligomerization of NKF3 members was found to be integral for Stat3 activation, a downstream effector that had been individually shown for PEAK1 and Pragmin (27, 36, 43). The helical bundle near the N-terminal region of the pseudokinase domain is necessary not only for homotypic and heterotypic dimerization but also for Stat3 activation, which led to the canonical increase in cell migration (43). In fact much of the N-terminus was dispensable for Stat3 phosphorylation yet the N-terminal helices and pseudokinase domain proved to be integral (43). Once the structure of Pragmin, and later PEAK1, was elucidated did the community

discover how the N-terminal helix and a group of helices on the C-terminal region of the pseudokinase domain form a completely new structure that is unique amongst any known structure in the kinome.

NKF3 members are unified as a family by the sequences within the pseudokinase domain and the adjacent N- and C-terminal helices, which is necessary for much of their function. Those adjacent helices fold back onto the C-lobe of the pseudokinase domain into an 'XL' domain known as the Split Helical Dimerization (SHED) domain (44). The SHED domain helices α N1, α J and α K (Pragmin nomenclature), are held together by both polar and nonpolar residues that lock the intercept of the 'X' together and link the 'X' to the C-lobe of the pseudokinase domain. Notably in Pragmin, Phe1366 of the α J helix interacts with Asp1381 of the α N1 helix and all of which is surrounded by hydrophobic residues at the intercept of the 'X' (45). Additional support comes from Asp1381 from α K helix to the top of the α J helix Lys1365 and Lys1369 and is also shielded from the solvent by Phe1366 of α J. In this manner the α K helix serves to hold the α J helix near the C-lobe and α J holds the α N1 helix. In PEAK1 the forces that stabilize the SHED domain may differ since the residues at the intercept between α S (α N1 in Pragmin) and α K (α J in Pragmin) are largely nonpolar (44).

The structural models of PEAK1 and Pragmin revealed the unique conformation of the SHED domain and also how it mediates homo-dimerization. The symmetry mates in the crystal structures of both PEAK1 and Pragmin show how two SHED domains align in parallel to form a symmetrical twin 'X' (30, 44, 45). Structural homology algorithms, such as the Dali server and VAST, which align and assess structural similarities, have found that there are no known structures with the SHED domain geometry, though PTEN-induced

kinase1 (PINK1) does have a region that is somewhat similar (44, 46). The dimerization interface is also near the intercept of the single shed domain where Met1363 juts out from α N1 helix to form a dense hydrophobic interface in *trans* with the adjacent α N1 helix at Ser962, Leu963, and Leu966 along with the α J Met1364 and Ala1367. This hydrophobic region is located just above two Tyr959, one from each monomer, that hydrogen bond together. Many of the hydrophobic residues in this region are integral for the thermal stability and dimerization of Pragmin including Leu955, Leu966, and Met 1363. In addition, the dimer interface near the 'X' intercept is Leu966 that is surrounded by a hydrophobic pocket created by Met1363, Phe1366, Ala1367, all of which are in *cis*, and Val1371 in *trans* from the adjacent α J helix. Additionally there are many other nonpolar residues that protrude into hydrophobic pockets that provide the molecular force that holds the monomers together.

Crk role in motility and intracellular signaling

Crk spliceforms and mechanisms of inhibition

The adapter protein CrkII serves as critical link in signaling cascades downstream from focal adhesions as well as receptor kinases. Crk was originally discovered from a chicken tumor and was shown to be an oncogenic driver that increased the global levels of phosphorylation and was thus named Chicken tumor virus no. 10 regulator of kinases (47). Crk proteins are conserved in many different metazoans and Crk orthologs can be detected in distant species such as choanoflagellates (48). Since its discovery in the late 1980's, the molecular biology of Crk has been well characterized yet remains an active area of research

for many labs due to its role signal transduction as well as its conserved SH2 and SH3 domain architecture.

Crk has two spliceforms, with CrkI being a truncated form of the full length CrkII, and also a Crk paralog named Crk-like (CrkL) that likely arose from a gene duplication (41). CrkII and CrkL are comprised an SH2 domain at the N-terminal end followed by two tandem SH3 domains (N-terminal SH3 domain (SH3(N) and C-terminal SH3 domain (SH3(C)) that are separated by an important regulatory linker. CrkI is a truncated splice variant of CrkII that is missing the SH3(C) domain and regulatory linker suggesting that CrkI-dependent signaling is controlled through an alternative mechanism and not through the regulatory linker. CrkII/CrkL are inhibited by phosphorylation of Tyr221/207, respectively, at the regulatory linker between the tandem SH3 domains that causes an intramolecular binding of the SH2 domain to the linker (49, 50). It is through this mechanism of regulation that cells can control signaling that emanates from integrin and RTKs to GTPase activity.

Crk-dependent GTPase activation and binding partner specificity

Crk adapter proteins are the molecular link between RTK stimulation and phosphorylation and GTPase activation, which cause cell to migrate. Both Integrin and EGF stimulation causes CrkII to localize to the membrane where it associates with another adapter protein p130-Crk associated substrate (P130CAS) (51, 52). Once localized to the membrane, CrkII binds DOCK180, a guanine nucleotide exchange factor (GEF) for Rac1 GTPase (52). Overexpression experiments show that CrkII, DOCK180 and/or p130CAS increases the

motility and is dependent on either Rac1 or CDC42 but not RAS GTPases (52-54). DOCK180 is unusual in that it requires another cofactor, ELMO, which seems to potentiate DOCK180 and protect it from degradation (55), although ELMO does not bind to CrkII it is clearly needed to potentiate Rac1 activation by CrkII and DOCK180 (56). Rac1 promotes lamellipodia formation, and actin stress fiber formation whereas deletion of Rac1 causes cells to become thin and lose their stress fibers similar to those cells that have a CrkII knockdown (54, 57, 58). These experiments highlight how CrkII is able to link RTK stimulation to GTPase activation that causes cells to change morphology and specifically increase their motility through GTPase activation.

CrkII and CrkL are both SH2/SH3 adapter proteins, similar to Grb2 and Nck, yet bind to a distinct proline-rich motif and thus have a unique set of binding partners and signaling pathways. The binding motif of Crk proteins are highest with the motif found in C3G, a Rap1 GEF, and had a significantly lower binding affinity to Grb2 compared to CrkII (142 μ M vs. 1.82 μ M, respectively) (59). In fact Far Western blot analysis of an array of SH3 domains from different proteins, including Grb2, Nck, PI3K, Src, Fyn, Csk, Abl, and Plc- γ , were found incapable of binding the C3G peptide using this technique (59). The consensus motif is P₁-P₂-X₃-L₄-P₅-X₆-K₇, with L₄ and K₇ having the greatest detriment to the binding coefficient of any of the residues in the motif (60). Interestingly, the K₇ position of the SOS2 peptide that Grb2 binds is an arginine and mutating K₇ to R₇ in the C3G peptide dramatically increased the binding of Grb2 to the C3G peptide (142 μ M vs. 23 μ M, respectively), showing how those specific residues confer specificity for a given SH3 domain.

CrkL oncogene and its role in chronic myeloid leukemia

The adapter protein CrkL gained interest and notoriety as an oncogene in chronic myeloid leukemia due to its association with BCR-ABL and its increased phosphorylation in CML. The translocation of chromosome 9 and 22 is a hallmark of CML and creates the fusion oncogene BCR-ABL and circumstantially is also near the coding region of CrkL. The identification of CrkL came from genomic analysis near the t(9;22) translocation region, which is the regulatory region of CrkL. When the regulatory region is disrupted by the t(9;22) translocation, it causes an increase in CrkL expression (61). Similar to CrkII, CrkL binds paxillin, Cbl and p130Cas through its SH2 domain at YxxP motifs found in the substrate proteins (62),(63), (64). CrkL also binds proline-rich motifs via its N-terminal SH3 domain e.g. CrkL recruits BCR-ABL to the membrane by binding to the C-terminal tail of Abl via its N-terminal SH3 domain (65). Increased exogenous CrkL in leukemia mouse models that express the BCR-ABL fusion protein have a decreased lifespan compared to those only expressing BCR-ABL alone (66), suggesting that CrkL is able to potentiate the oncogenic effects of BCR-ABL. CrkL has identical domain architecture to CrkII and some redundant function *in vivo*, yet there are also non-redundant functions that are critical for development.

CrkL and CrkII phenotypic similarities and regulatory differences

CrkL and CrkII do have some redundancy in their signaling pathways and potential, yet there are some differences that make them unique. The redundancy between CrkII and CrkL can be seen *in vivo* where single knock out of CrkI/II or CrkL specifically within the kidney has no effect on kidney development, specifically the elongation of the podocyte

cells that create the epithelial filter barrier (67). However mice kidneys with the CrkI/II/CrkL double knockout have reduced functionality, including defects in podocyte extension that lead to albuminuria, suggesting that Crk and CrkL are necessary for kidney development and interestingly they are able to compensate for one another (67).

CrkII and CrkL both can bind GEF protein C3G and Dock180 where by they participate in GTPase activation that promotes cell motility that is critical during development (52, 68, 69). Cell area is also effected by CrkII/CrkL knockout, where cells are more rounded and smaller, however CrkII is better at rescuing cell morphology compared to CrkL, suggesting that CrkII is a more potent regulator of cell area than CrkL (70). Furthermore, both Crk^{-/-} and CrkL^{-/-} knockout mice die prenatally, suggesting that they are not completely redundant during development (71),(72),(73). Both Crk and CrkL are expressed ubiquitously in the heart during development and in both Crk^{-/-} and CrkL^{-/-} mouse models show signs of craniofacial and cardiac developmental defects. While there are clear similarities in expression location and phenotypic consequences for these paralogs, the structural orientation and ultimately the mechanism of regulation set Crk and CrkL apart.

CrkL and CrkII are both SH2/SH3 adapter proteins, yet the conformation and orientation of the SH2 and SH3 domains differs leading to dissimilarity in the accessibility of the SH3(N) domain. Both CrkII and CrkL have a regulatory tyrosine in the linker between the N and C terminal SH3 domains (Y221 and Y207, respectively). It has been well established that CrkII is inhibited by phosphorylation of Y221 due to an intramolecular interaction between the SH2 domain and the regulatory pY221 motif (49, 74). Phosphorylation of CrkL Y207 also causes its SH2 domain to bind the inhibitory linker

domain yet this does not seem to impede the SH3(N) from binding to its substrate (75). NMR structural analysis revealed the orientation of the SH3(N) domain with respect to the SH2 domain differs between CrkL and CrkII such that the binding site of CrkII SH3(N) is facing towards the SH2 domain in an occluded orientation yet the SH3(N) binding site of CrkL is facing outward toward the solution (75). The biochemical consequence reveals that the binding coefficient of CrkL to the Abl PPII peptide, which both Crk and CrkL bind via the SH3(N), is unaffected by the intramolecular interaction as it is with CrkII pY221 (75). While CrkL pY207 is still able to bind to Abl, the SH2 domain is bound to the regulatory linker and is indeed inhibitory due to changes in localization rather than through occlusion of the binding site.

Results

PEAK3 is a distinct member of the NKF3 family of atypical protein kinases.

We searched for distant homologs of protein kinases in the human proteome using the sensitive sequence comparison algorithm Fold and Function Assignment System (FFAS) (25), previously used by us to discover novel kinase-like proteins in humans (SELO and FAM69/DIA1 families (76, 77) and in bacteria and fungi (COTH (78)). The only novel protein with a significant FFAS score found that was not previously identified as kinase-like was a single uncharacterized protein, denoted as chromosome 19 open reading frame 35 (C19orf35). C19orf35 has the closest similarity to the PEAK1/SgK269 and Pragmin/SgK223 pseudokinases that together constitute the NKF3 family. Recently, Lecointre and colleagues reported the identity of C19orf35 as a new member of the NKF3

family, and C19orf35 was annotated in the Universal Protein Resource (Uniprot) as PEA3 (30). For consistency, from this point on, we refer to C19orf35 as PEA3. Using PEA1 or Pragmin as Blast queries results in only partial alignments covering fragments of the kinase-like domain. However, the FFAS method is capable of detecting similarity over the entire length of the putative kinase-like domain (Fig. 1A). The kinase-like domain of human PEA3 shares only approximately 26% sequence identity with human PEA1 and Pragmin due to low complexity regions (LCRs), which introduce long gaps in the pairwise alignments. This is most likely why the PEA3 sequence, although present in the Uniprot database since 2004, was not annotated as kinase-like until recently. Even now, standard protein domain annotation tools (CD-Search or Pfam HMM) do not report kinase similarity for PEA3.

LCRs, defined as areas of protein sequences with biased amino acid composition, located within the kinase-like domain are a distinct feature of NKF3 proteins (79). These regions are typically implicated in mediating protein-protein interactions (80). In NKF3 pseudokinases, the LCRs correlate with flexible regions, as judged by missing coordinates corresponding to these sequences in the structures of PEA1 and Pragmin. One example is the PAPA motif in Pragmin that is located between the HRD and NFL motifs, where the NFL motif corresponds to the DFG motif in active kinases (*SI Appendix*, Fig. S1A). The majority of the LCRs in PEA1, Pragmin, and PEA3 diverge in sequence from one another and are located in different regions. PEA3 stands out by having the largest portion of its kinase-like domain sequence (more than 20%) denoted as LCRs. Some of these motifs are relatively well-conserved in evolution, such as the PPGPPGSPGP motif that is immediately downstream of the DFG motif in PEA3 (*SI Appendix*, Fig. S1B).

Figure 1

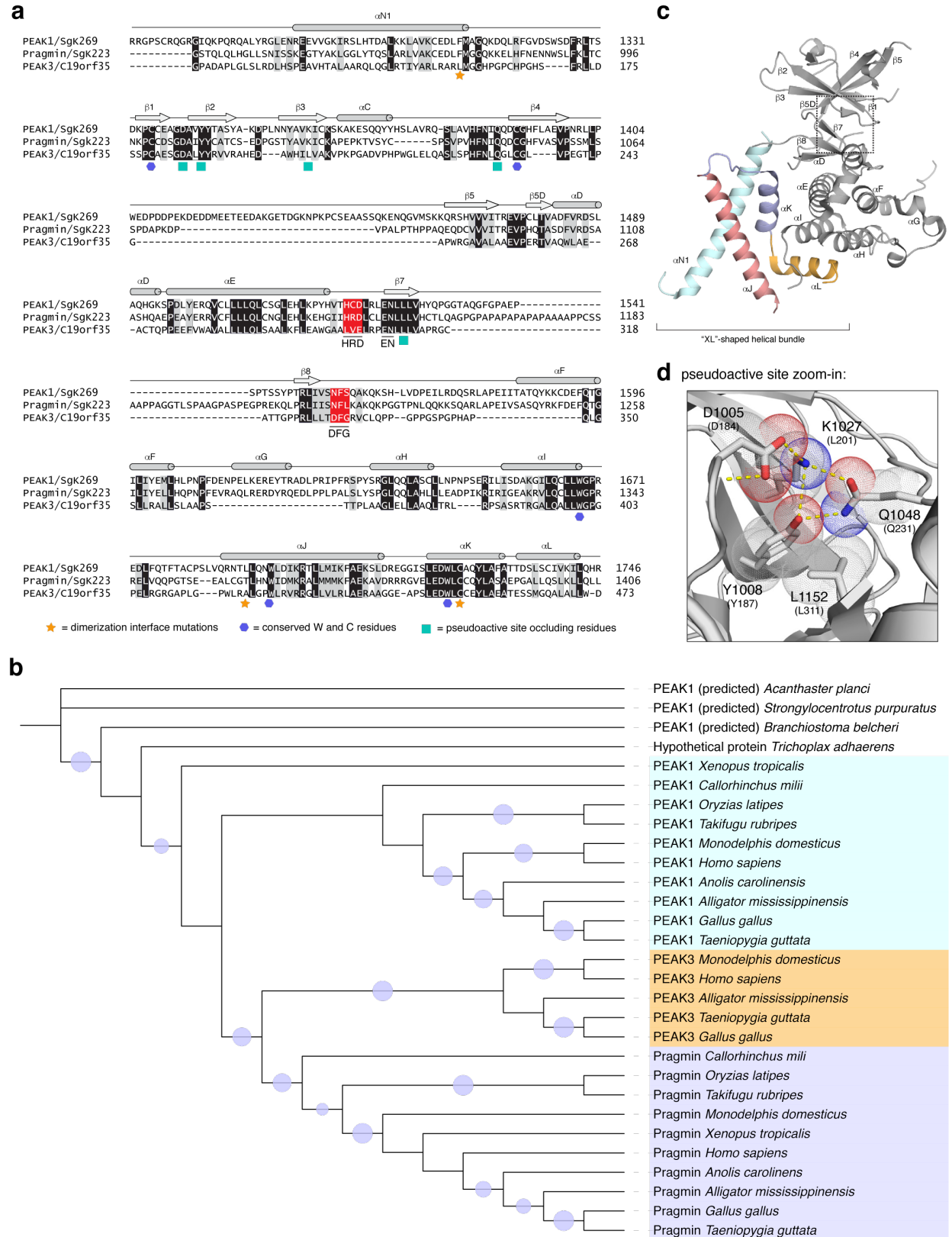


Figure 1. Identification of PEAK3 (C19orf35) as a homolog of PEAK1 and Pragmin. (a) Protein sequence alignment for the kinase domains of human PEAK1/SgK269, Pragmin/SgK223 and PEAK3 (C19orf35). Secondary structure elements are denoted based on the Pragmin structure (PDB ID: 5VE6), and sequence motifs corresponding to the canonical kinase active site motifs (HRD and DFG in classical kinases) are marked in red. Residues involved in dimerization, conserved W and C residues, and residues shown to occlude the pseudo-active site in Pragmin are marked by the indicated symbols. (b) Phylogenetic tree (PhyML) for selected NKF3 kinases. Branches with bootstrap values better than 70% are marked with circles. (c) Cartoon representation of the crystal structure of the Pragmin SHED domain/pseudokinase module (PDB ID: 5VE6). (d) Zoomed-in view of the pseudoactive site in Pragmin depicting residues that occlude the canonical nucleotide-binding pocket. Top numbering is for Pragmin residues (PDB ID: 5VE6), bottom in parentheses to predicted corresponding residues in PEAK3. The Pragmin D184 residue was modeled in the active site based on the structure of Pragmin (PDB: 6EWX).

Evolutionary conservation of PEAK3 and its kinase-like domain.

The NKF3 family likely appeared at the origin of Metazoa, which is indicated by the presence of homologs in sponges (*Amphimedon*) and placozoans (*Trichoplax*). Most invertebrates (e.g. cnidarians, echinoderms) contain a single member of the family, although in some lineages (e.g. insects and nematodes), the NKF3 family has seemingly been lost. The division into the PEAK1 and Pragmin subfamilies likely occurred between the emergence of chordates (the lancelet *Branchiostoma* has a single family member) and jawed vertebrates (*Gnathostomata* have both PEAK1 and Pragmin). The PEAK3 subfamily diverges from the Pragmin branch likely during the evolution of reptiles (e.g. it is found in crocodilians) and is present in birds and mammals (Fig. 1B). Interestingly, PEAK3 is missing in some reptile species such as snakes and lizards. The early members of the NKF3 family, i.e. proteins from sponges and placozoans, exhibit high sequence and structure conservation with those found in more complex metazoans (e.g. vertebrates), suggesting evolutionarily conserved functions of these proteins.

The kinase-like domain of PEAK3 is highly conserved in evolution and carries unique sequence alterations in key catalytic motifs found in active kinases when compared with PEAK1 and Pragmin. Interestingly, these alterations within the putative active site of PEAK3 vary markedly between species (Fig. 1A; *SI Appendix*, Fig. S1). Mammalian PEAK1 and Pragmin do not possess a conserved DFG motif. However, PEAK1 and Pragmin have intact HxD motifs: HRD in PEAK1 and HCD in Pragmin. In contrast, mammalian PEAK3 contains a conserved DFG motif, while the HxD motif is replaced by LxE. These unique features are not conserved in avian PEAK3 homologs, which have an NFF or SFF sequence in place of DFG, more closely resembling sequences present in PEAK1 (NFS) and in Pragmin (NFL). In PEAK1 and Pragmin, these motifs are almost perfectly conserved irrespective of a species. While HxD is also not conserved in avian PEAK3, the catalytic aspartate (contained within the HxD motif) is present within the QGD sequence that replaces the HxD motif.

In all species, PEAK3 has a conserved EN motif, corresponding to the EN sequence located in PKA at positions 170-171 that coordinates divalent cations (81). This feature is also present in PEAK1 and Pragmin (Fig. 1A; *SI Index*, Fig. S1). The catalytic lysine, equivalent to K72 in PKA is conserved in PEAK3 (K204 in PEAK3), but the glutamate (E91 in PKA) with which the catalytic lysine forms a salt bridge in active kinases is missing in PEAK3. In PEAK1 and Pragmin, the catalytic lysine is also present but is “hijacked” by interactions with three residues collectively termed the “inhibitory triad” that, in addition to other conserved residues, occlude the nucleotide binding pocket and prevent binding of ATP (Fig. 1C and 1D) (30, 44, 45). These residues are well-conserved in PEAK3 and are represented by D184, Y187, L201, Q231, and L311 (Fig. 1A and 1D). Collectively, the extent

of mutations in the key catalytic motifs strongly indicate that although PEAK3 diverges from other members of the NKF3 family, it still falls into the category of pseudokinases defined by Manning and colleagues as kinases that lack one or more of the canonical catalytic sequence motifs (3, 4). Hence all NKF3 kinases carry pseudokinase characteristics, and seem to derive from a common NKF3 pseudokinase ancestor in early Metazoans, as indicated by the sequence variability in the NKF3 proteins and reconstruction of the ancestral NKF3 sequence (Ancescon method (82)).

PEAK3 interactome.

With no prior insights into the function of PEAK3, we took an unbiased approach to identify its interacting partners using immunoprecipitation followed by mass spectrometry (IP/MS). Due to the current lack of a suitable antibody for detection of endogenous PEAK3, our analysis was conducted using a 3xFLAG-tagged human PEAK3 transiently expressed in HEK293T cells. Identified proteins that co-immunoprecipitated with PEAK3-3xFLAG can be categorized into several subgroups: (i) CrkII and CrkL, highly homologous adaptor proteins that regulate cell proliferation, adhesion and cytoskeletal integrity downstream from receptor tyrosine kinases and integrins (70, 83); (ii) 14-3-3 scaffold proteins (β , γ , η , and τ), which play diverse roles in signaling, including regulation of cell motility, survival and intracellular protein trafficking (84); (iii) guanine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs) for Rho family of small GTPases, including ASAP1 that participates in actin cytoskeletal dynamics and cell movement (85, 86); and (iv) several proteins connected to the regulation of cell death and survival, including SIAH1 ubiquitin

ligase (87), DRAK1 (88) and another 14-3-3 scaffold protein, 14-3-3 σ , also known as SFN (89) (*SI Appendix*, Table 1). Collectively, this analysis suggests that PEAK3 plays a role in a number of functions involved in cell proliferation, survival, and motility. The function of PEAK3 in cell motility, in particular, was underscored by the abundance of motility regulators in the immunoprecipitates.

PEAK3 interacts with CrkII via a proline-rich motif/SH3 domain interaction.

Given the documented roles of NKF3 family members in the regulation of cellular motility (27-29, 31, 42, 43, 90, 91), we focused our functional studies of PEAK3 on the adaptor protein CrkII, which was one of the most abundant PEAK3-interacting proteins identified by the IP/MS analysis measured as a high confidence score via the comparative proteomic analysis software suite (ComPASS) (92) (*SI Appendix*, Table 1). Primary sequence analysis of the N-terminal domain of PEAK3 revealed the presence of a putative CrkII-binding site, PPPLPK, located 71 residues upstream from the predicted kinase domain. This site closely resembles the consensus sequence present in known binding partners of CrkII, including the GEF proteins DOCK180 and C3G (Fig. 2A, B) (31, 59, 93). This putative CrkII-binding motif in PEAK3 is highly conserved across evolution, suggesting a potential importance in PEAK3 function (Fig. 2C). Interestingly, this sequence is also present in PEAK1 and Pragmin, and PEAK1 was previously shown to co-immunoprecipitate with CrkII (31). Using co-immunoprecipitation, we verified that transiently expressed PEAK3 in HEK293 cells is indeed able to bind both endogenously and exogenously expressed CrkII (Fig. 2D; *SI Appendix*, Fig. S2). Mutation of residues in the predicted CrkII-binding site (P56A, L59A,

and P60A in PPPLPK) completely abolished the ability of mutant PEA3 (PEAK3-3A) to co-immunoprecipitate with CrkII (Fig. 2D). CrkII is composed of an SH2 domain followed by two SH3 domains, termed N-terminal SH3 domain (SH3^N) and C-terminal SH3 domain (SH3^C). The inability of the PEA3-3A mutant to bind CrkII suggested that PEA3 interacts specifically with the SH3^N in CrkII, which has previously been shown to engage similar proline-rich motifs in other CrkII-binding partners (52, 59, 93, 94). Indeed, mutation of the SH3^N (CrkII-W170K) but not of the SH3^C domain (CrkII-W276K) rendered CrkII unable to bind PEA3 (Fig. 2E).

PEAK3 antagonizes CrkII-induced changes in cellular morphology. One well-characterized role of PEA1 and Pragmin is the regulation of cell morphology and migration. Both PEA1 and Pragmin localize to actin filaments and focal adhesions, induce cell elongation, and promote cell migration when transiently expressed in cells (27-29, 31, 34, 42, 43). In contrast, we found that PEA3 distributes diffusely throughout the cytoplasm of COS-7 cells and U2OS cells upon transient transfection, and had little to no effect on overall cell shape compared to vector-transfected control cells (Fig. 3A; *SI Appendix*, Fig. S3). However, while the majority of control cells retained prominent actin filaments that traversed the cell (16, 95). Those cells overexpressing PEA3 exhibited notably fewer stress fibers and possessed shorter, less organized actin filaments, mirroring the phenotype typically observed in cells in which the CrkII gene is knocked down (Fig. 3B) (96).

Figure 2

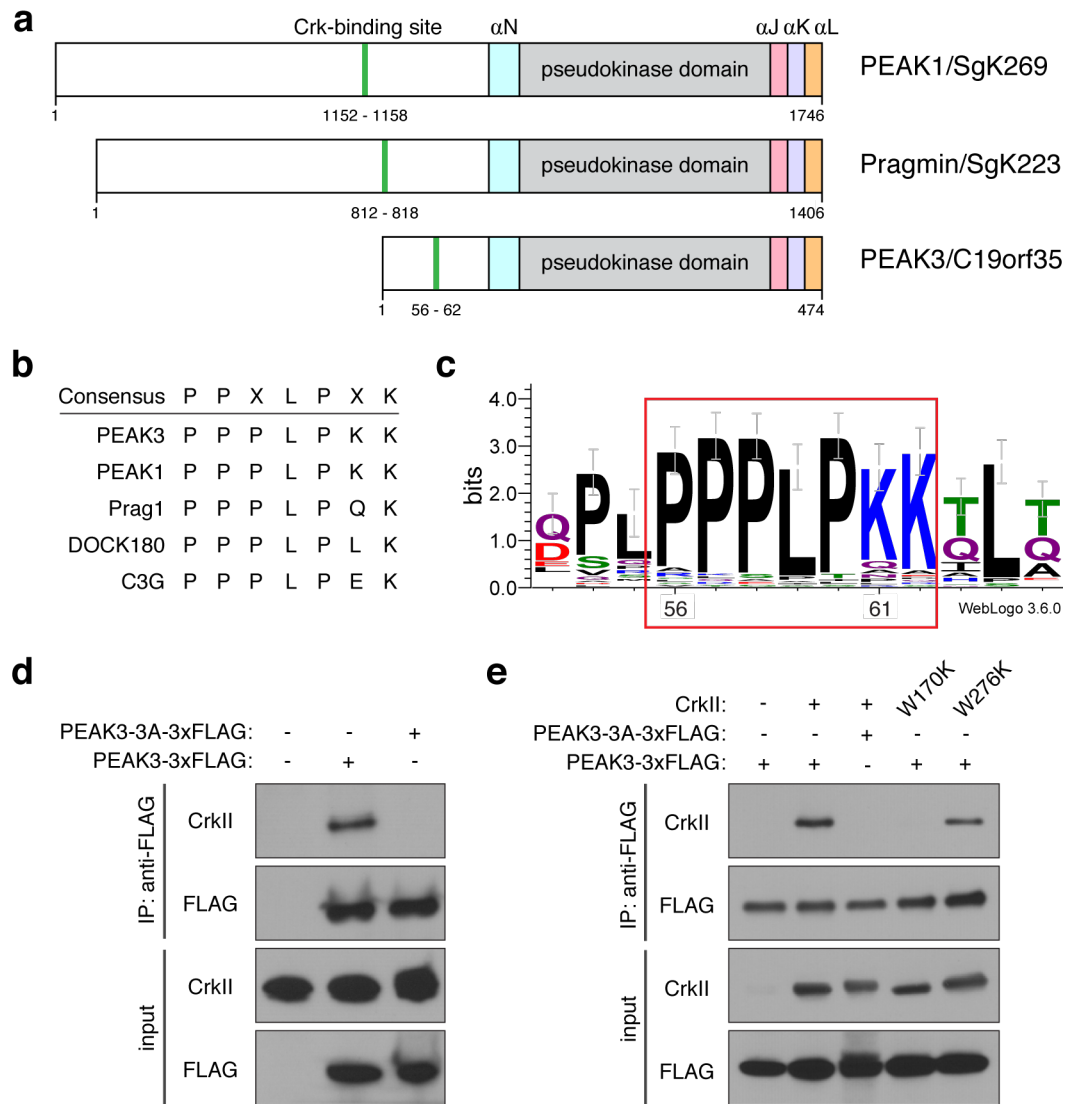


Figure 2. PEAK3 binds CrkII. (a) Schematic representation of PEAK1/SgK269, Pragmin/SgK223, and PEAK3/C19orf35 domain structure. The locations of the CrkII-binding sites and helical regions within the SHED domain are highlighted. (b) Consensus sequence of CrkII-binding sites in selected CrkII-binding partners. (c) Sequence logo depicting conservation of the CrkII-binding site in PEAK3 homologs. (d, e) Co-immunoprecipitation of FLAG-tagged wild type PEAK3 or a CrkII-binding mutant (PEAK3-3A) transiently expressed in HEK293 cells with either endogenous CrkII (d) or transiently expressed untagged CrkII variants carrying mutations in the SH3 domains (W170K in SH3^N and W276K in SH3^C) (e). Cell lysates were incubated with an anti-FLAG antibody, resolved by SDS/PAGE, and probed with the indicated antibodies by western blot. All co-immunoprecipitation data are representative of at least three independent experiments.

Endogenous CrkII participates in actin polymerization and positively regulates cell motility (51, 95-99). Exogenously expressed CrkII localizes to the cell cortex where it induces a spindle-shaped cell morphology with notable membrane extensions resembling lamellipodia or polarized membrane ruffles (50, 51, 70, 100-102). These morphological changes can be visualized robustly in COS-7 and U2OS cells (40, 51, 96, 103), hence we used these cell lines to measure the functional consequences of PEA3 overexpression on CrkII-dependent effects on cell morphology. Remarkably, cells co-expressing PEA3 and CrkII had markedly fewer membrane extensions and largely did not adopt a CrkII-dependent morphology (Fig. 3A; *SI Appendix*, Fig. S4A). In contrast, the PEA3-3A mutant, which does not interact with CrkII, was unable to interfere with the CrkII-dependent phenotype (Fig. 3A; *SI Appendix*, Fig. S4A).

To quantitatively compare these differences in cellular phenotypes, we developed a metric in which the effect of CrkII on cell morphology is measured as an increase in cell perimeter (Fig. 3C). While CrkII overexpression alone significantly increases cell perimeter, there is no change in cell perimeter when CrkII is co-expressed with wild type PEA3 (Fig. 3D; *SI Appendix*, Fig. S4B). PEA3-CA mutant has no effect on CrkII-dependent increase in cell perimeter, supporting a conclusion that PEA3 negatively regulates CrkII as a result of their direct interaction.

Figure 3

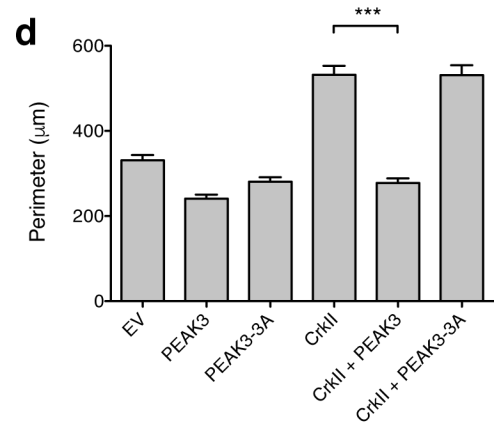
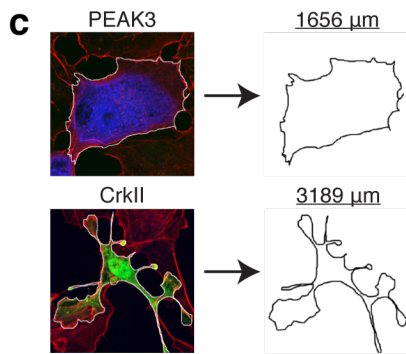
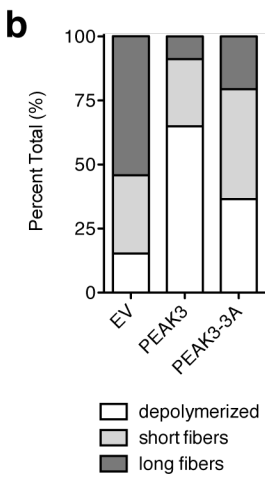
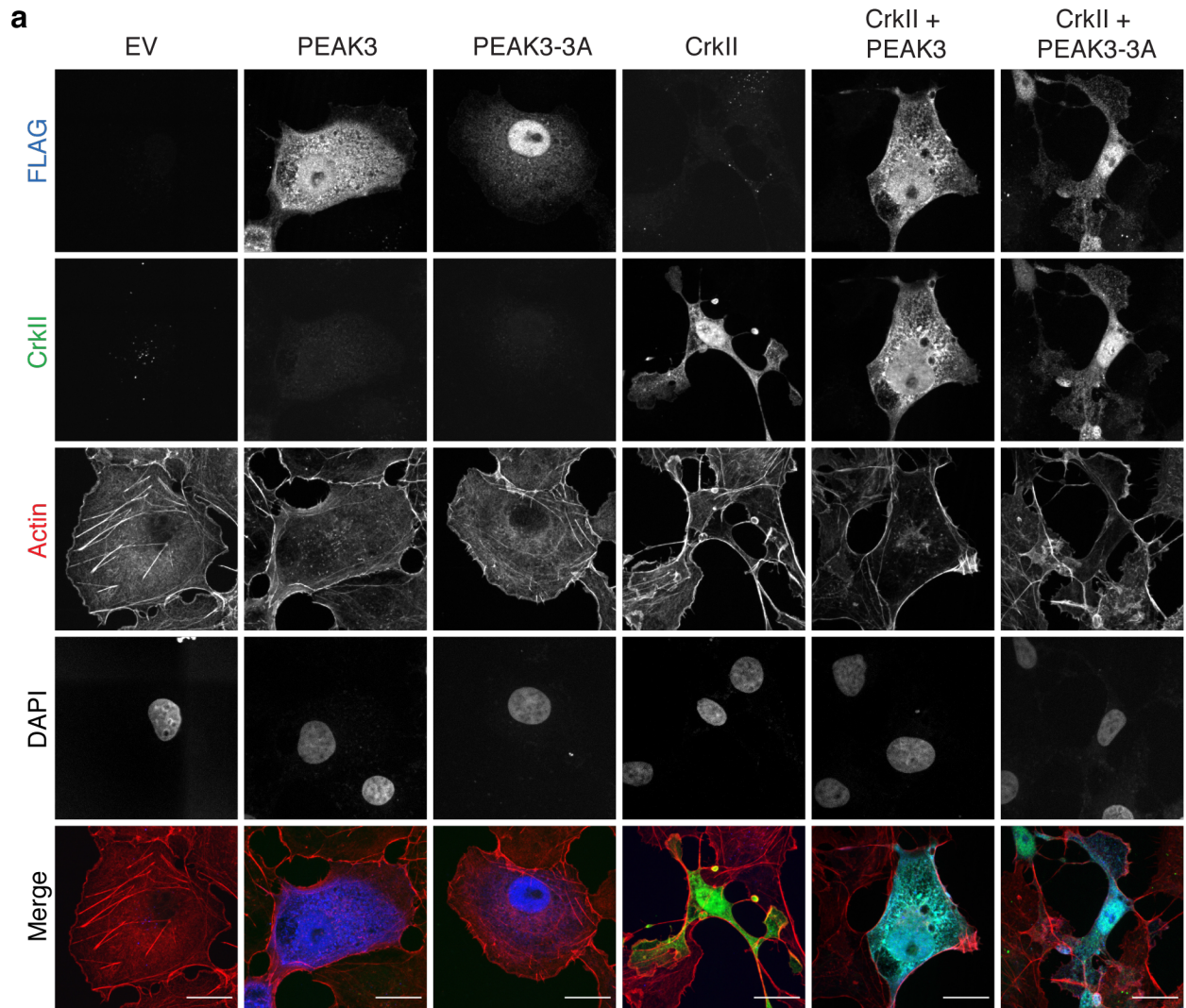


Figure 3. PEAK3 prevents CrkII-dependent lamellipodia-like membrane extensions and membrane ruffling. (a) Confocal microscopy imaging of COS7 cells transiently co-transfected with a FLAG-tagged wild type or a CrkII-binding deficient mutant of PEAK3 (PEAK3-3A) with either an empty vector or untagged CrkII. 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) Relative percentage of different actin fiber phenotypes measured in COS7 cells transiently transfected with PEAK3 and stained with Alexa Fluor 647-conjugated phalloidin (n = 60 cells per group). (c) Schematic illustrating perimeter calculation in representative cells from (b). (d) 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 Materials and Methods. Data represent the mean \pm SEM of three independent experiments (n = 20 cells in each experiment), *** p < 0.001.

Negative regulation of CrkII by PEAK3 requires the C-terminal domain.

The PEAK3 protein can be arbitrarily subdivided into two distinct domains: the N-terminal domain that contains the CrkII-binding site and the C-terminal domain that contains the pseudokinase domain. We exogenously expressed the N-terminal domain of PEAK3 to determine if it alone is sufficient for the negative regulation of CrkII. Unexpectedly, a construct containing only the N-terminal domain and missing the C-terminally-located pseudokinase domain (PEAK3 Δ PK) did not antagonize CrkII function (Fig. 4A and 4B; *SI Appendix*, Fig. S5). Even more surprisingly, the PEAK3 Δ PK mutant was also unable to interact with CrkII, despite the presence of the intact CrkII binding motif (Fig. 4C). Hence, while the CrkII binding motif is necessary to support binding of PEAK3 to CrkII, it is not sufficient for CrkII binding and inhibition, demonstrating that the C-terminal domain of PEAK3 plays an essential role in mediating the PEAK3/CrkII interaction.

Figure 4

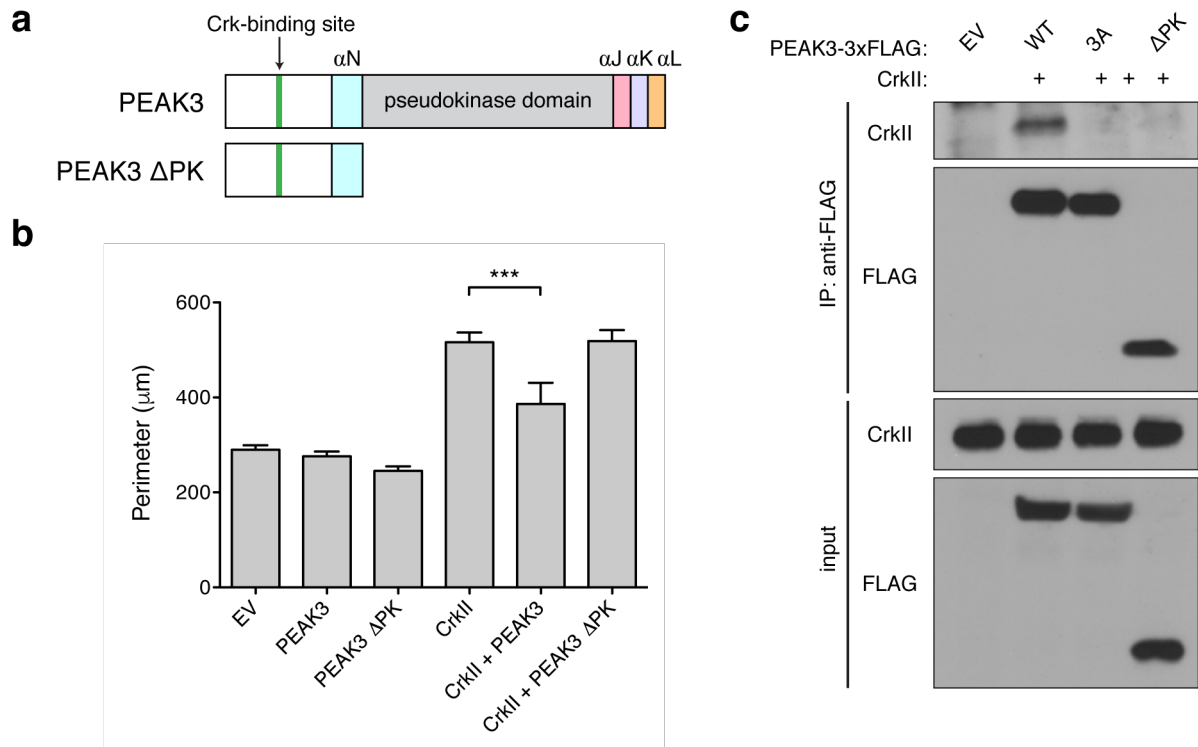


Figure 4. CrkII binding motif is insufficient for PEA3-dependent inhibition of CrkII. (a) Schematic representation of used PEA3 constructs. (b) Average perimeter of COS7 cells transiently expressing either wild type or the ΔPK variants of PEA3 with an empty vector or untagged CrkII. Data represent the mean ± SEM of three independent experiments (n = 20 cells in each experiment), *** p < 0.001. (c) Co-immunoprecipitation of FLAG-tagged wild type and ΔPK PEA3 variants, transiently expressed in HEK293 cells, with exogenously expressed, untagged CrkII. Cell lysates were incubated with an anti-FLAG antibody, resolved on SDS/PAGE, and probed with the indicated antibodies by western blot. Data is representative of three independent experiments.

Predicted SHED domain in PEA3.

Recent crystal structures of the PEA1 and Pragmin C-terminal domains have revealed identical dimer forms composed of the pseudokinase domains interacting through a set of helical bundles, termed the Split Helical Dimerization (SHED) domain (30, 43-45). The SHED domain is unique to NKF3 proteins and is comprised of the helix immediately N-terminal to the pseudokinase domain (αN helix) and three helices C-terminal to the

pseudokinase domain (α J, α K, α L helices) that form an "XL"-shaped helical bundle (lettering of helices reflect Pragmin nomenclature (45); PEAK1 helices are off-set by one letter from the α J helix (44)) (30, 44, 44, 45).

In PEAK3, the highest sequence similarity with PEAK1 and Pragmin, apart from the pseudoactive site, falls within the regions corresponding to the SHED domain. In the three C-terminal α -helices, sequence identity between PEAK3 and Pragmin is 34% and 32% between PEAK3 and PEAK1 (*SI Appendix*, Fig. S1 and S6). The most striking conserved sequence motifs include the tryptophan residue-containing motifs. These motifs include the EDWLCC sequence in the α K helix, WGP in the loop preceding the α J helix, and WL in the α J helix (Fig. 1A and 1C; *SI Appendix*, Fig. S1). In the PEAK1 and Pragmin crystal structures, these conserved motifs are involved in the interactions between the pseudokinase domain and the SHED domain (30, 44, 44, 45). PEAK3 has remarkably well-conserved sequences in these regions, indicating that the SHED domain is likely present in PEAK3 adopting a similar structure to the ones in PEAK1 and Pragmin (*SI Appendix*, Fig. S6). Thus, the SHED domain emerges as a unifying structural feature of the NKF3 family of pseudokinases.

PEAK3 dimerizes via the SHED domain.

PEAK1 and Pragmin form homo- and hetero-oligomers through two distinct dimer interfaces, one involving the SHED domain and another involving the α G helix/A-loop interface (30, 43, 43, 45). Mutation of these interfaces, especially of the hydrophobic interactions between the helices in the SHED domain, impairs the signaling ability of these pseudokinases (44, 44, 45). While the SHED domain in PEAK3 is highly similar to the SHED

domains of PEAK1 and Pragmin (Fig. 1A), the α G helix/A-loop interface is not significantly conserved in PEAK3. This lead us to hypothesize that PEAK3 too could dimerize via its putative SHED domain and that this might be critical for its interaction with CrkII.

To assess the ability of PEAK3 to homodimerize, we co-expressed HA-tagged and FLAG-tagged wild type PEAK3 variants in HEK293 cells and assessed their association by co-immunoprecipitation. These two differentially tagged PEAK3 constructs robustly co-immunoprecipitated (Fig. 5A; *SI Appendix*, Fig. S7A). Based on the crystal structures of Pragmin and PEAK1, we designed PEAK3 mutants that carried either individual substitutions of key residues in the predicted dimerization interface or deletion of one of the α N, α J, or α K helices in the SHED domain (Fig. 5B). These mutations/deletions abolished the ability of differentially tagged PEAK3 variants to co-immunoprecipitate (Fig. 5C and 5D; *SI Appendix*, Fig. S7B and S7C), while control mutations of residues located within the SHED domain helices but distal from the dimer interface had no effect (*SI Appendix*, Fig. S7D). These data suggest that PEAK3 self-associates through the SHED domain in a manner analogous to PEAK1 and Pragmin.

Figure 5

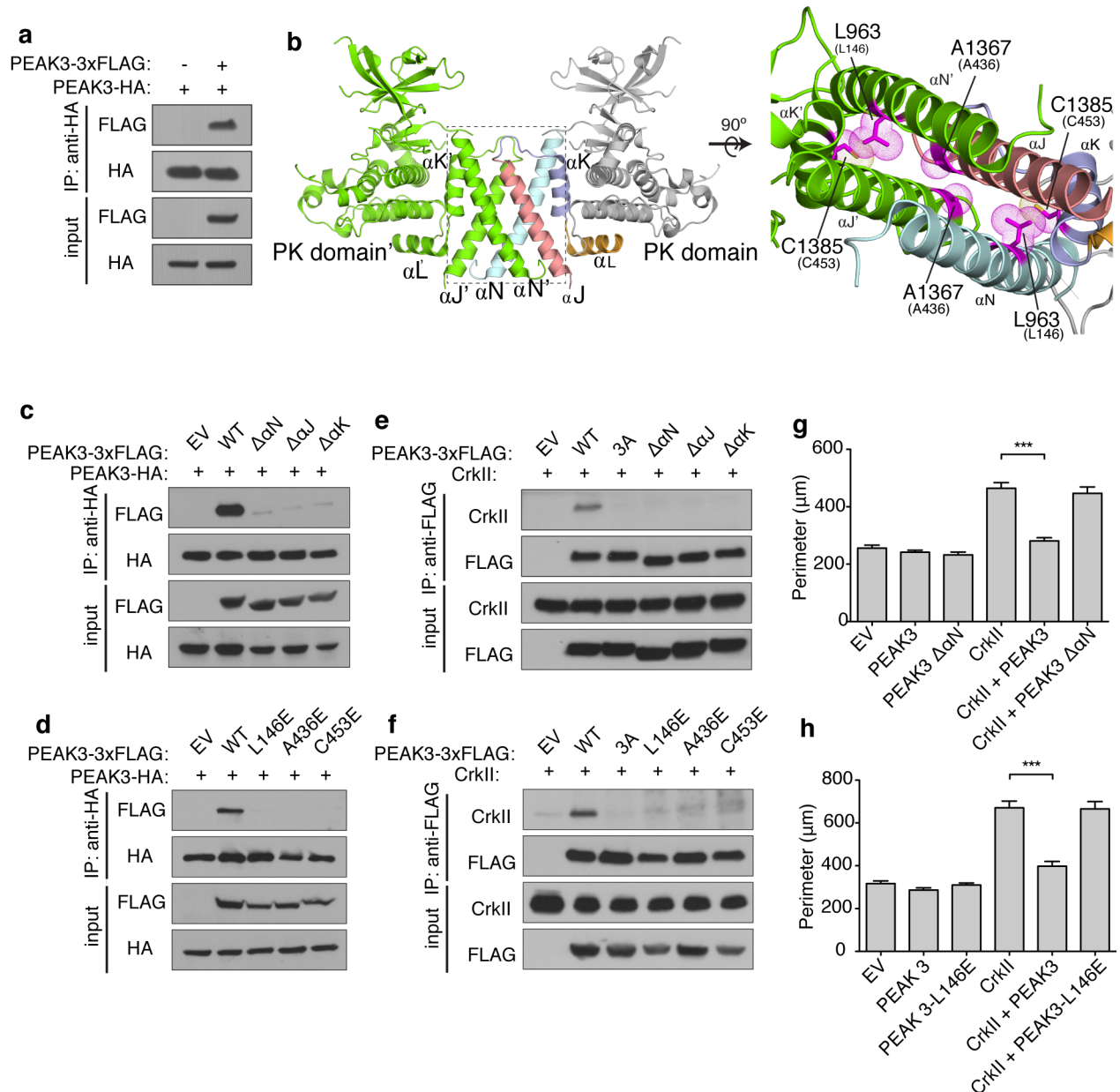


Figure 5. PEAK3 dimerization via the pseudokinase/SHED module is necessary for binding and inhibition of CrkII. (a) Co-immunoprecipitation of a FLAG-tagged and an HA-tagged variants of wild type PEAK3, transiently expressed in HEK293 cells. Cell lysates were incubated with an anti-FLAG antibody, resolved by SDS/PAGE, and probed with the indicated antibodies by western blot. (b) Left panel: cartoon representation of the Pragmin SHED/pseudokinase domain (PK) dimer (PDB ID: 5VE6). Right panel: zoomed in view of the dimerization interface depicting interactions between the helices within the SHED domain in Pragmin. Three residues

in the dimerization interface marked by sticks and dot representation were selected for mutagenesis in PEAK3 based on sequence homology. Top numbering corresponds to Pragmin (PDB ID: 5VE6), bottom in parentheses to PEAK3. **(c - f)** Co-immunoprecipitation of homodimerization deficient variants of PEAK3 with wild type PEAK3 (c, d) or with untagged CrkII (e, f), all transiently expressed in HEK293 cells. Cell lysates were incubated with an anti-HA (c, d) or anti-FLAG (e, f) antibody, resolved by SDS/PAGE, and probed with the indicated antibodies by western blot. All co-immunoprecipitation data is representative of three independent experiments. **(g, h)** Average perimeter of COS7 cells expressing wild type or homodimerization deficient variants of PEAK3 with an empty vector or untagged CrkII. Data represent the mean \pm SEM of three independent experiments (n = 20 cells in each experiment), *p < 0.05, *** p < 0.001.

Dimerization of PEAK3 is necessary for CrkII binding and its negative regulation.

While the C-terminal domain of PEAK3 does not contain a CrkII binding motif, it plays an essential role in CrkII binding and inhibition, as demonstrated by our finding that the construct missing this domain, PEAK3 Δ PK, was unable to engage with and antagonize CrkII-induced membrane ruffling (Fig. 4). To test if this role of the C-terminal domain is linked to its ability to mediate PEAK3 self-association through the SHED domain, we measured how mutations within the SHED domain dimer interface affect CrkII binding. All mutations that compromise PEAK3 dimerization also interfered with the ability of PEAK3 to interact with CrkII (Fig. 5E and 5F ; *SI Appendix*, Fig. S8A). Consequently, the PEAK3 dimerization mutants also failed to inhibit the formation of membrane extensions and membrane ruffles caused by CrkII overexpression (Fig. 5G and 5H; *SI Appendix*, Fig. S8B and S8C).

Our data thus far demonstrate that the N-terminal domain of PEAK3 containing the CrkII binding motif is unable to bind to or interfere with CrkII function when not dimerized by the C-terminal domain, which contains the pseudokinase domain. We wondered if the functional effect of the C-terminal domain-mediated dimerization could be mimicked by substitution of the C-terminal domain with an orthogonal domain that drives constitutive

PEAK3 dimerization. To test this, we fused the N-terminal region of PEAK3 (residues 1-131) containing the CrkII-binding site, but not the SHED domain or the pseudokinase domain, to a constitutively dimeric coiled coil domain (PEAK3-diCC) (Fig. 6A). While the monomeric N-terminal PEAK3 Δ PK construct is unable to bind CrkII (Fig. 4C), equivalent levels of the immunoprecipitated PEAK3-diCC fusion show notable CrkII binding (Fig. 6B). The ability of this minimal construct to restore CrkII binding underscores the importance of dimerization of the CrkII binding motif as a determinant of interaction between CrkII and PEAK3, and possibly also between CrkII and its other interaction partners. To our knowledge such property in known CrkII-binding proteins has not been described.

Despite the ability to interact with CrkII, PEAK3-diCC did not interfere with CrkII-dependent membrane ruffle formation in cells as observed for the wild type PEAK3 (Fig. 6C; *SI Appendix, S9*).

Figure 6

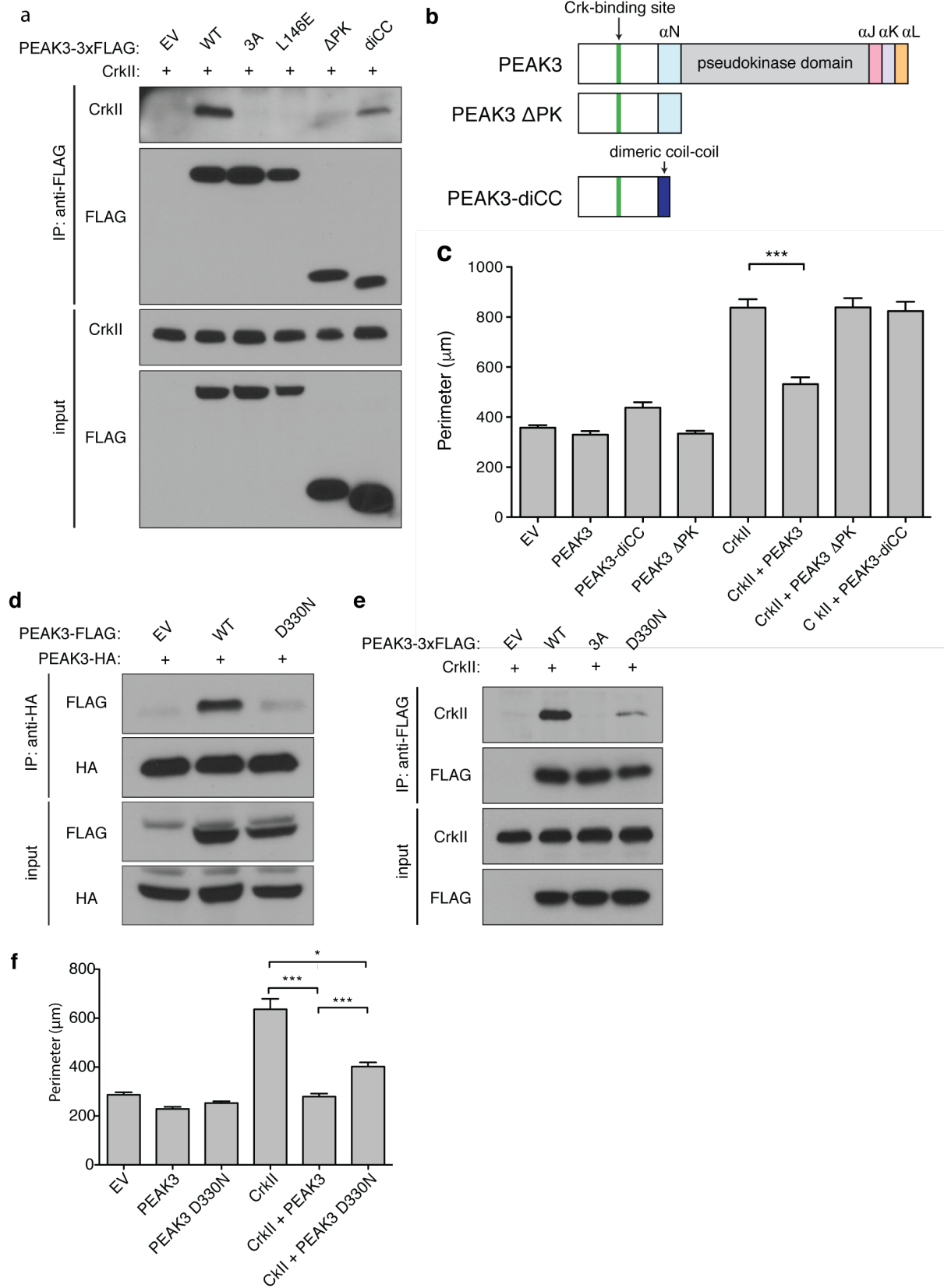


Figure 6. Role of PEAK3 pseudokinase domain in regulation of CrkII extends beyond dimerization. (a) Co-immunoprecipitation of untagged CrkII with FLAG-tagged dimerization mutants of PEAK3, transiently expressed in HEK293 cells. Cell lysates were incubated with an anti-FLAG antibody, resolved by SDS/PAGE, and then probed with the indicated antibodies by western blot. (b) Schematic representation of used PEAK3 constructs. (c) Average perimeter of COS7 cells transiently expressing either wild type or diCC variants of PEAK3 with an empty vector or untagged CrkII. Data represent the mean \pm SEM of three independent experiments (n = 20 cells in each experiment), *** p < 0.001. (d, e) Co-immunoprecipitation of the PEAK3 D330N variant with wild type PEAK3 (d) or untagged CrkII (e) transiently expressed in HEK293 cells. Cell lysates were incubated with an anti-FLAG antibody, resolved by SDS/PAGE, and then probed with the indicated antibodies by western blot. All co-immunoprecipitation data are representative of three independent experiments. (f) Average perimeter of COS7 cells transiently expressing either wild type or the D330N variant of PEAK3 with an empty vector or untagged CrkII. Data represent the mean \pm SEM of three independent experiments (n = 20 cells in each experiment), *** p < 0.001.

This discrepancy possibly reflects the weaker binding between CrkII and PEAK3-diCC compared to the wild type PEAK3 (Fig. 6B). Alternatively, or concurrently, the C-terminal pseudokinase/SHED module in PEAK3 might play a role in antagonizing CrkII signaling that extends beyond serving as a dimerization domain for the CrkII binding motif.

Mutation of the DFG aspartate impairs CrkII regulation by PEAK3.

NKF3 kinases have evolved multiple sequence alterations within their pseudokinase domains relative to active kinases, which are predicted to render them catalytically inactive. One notable difference between PEAK3 and PEAK1 and Pragmin is conservation of the DFG motif in PEAK3. The DFG motif is present in a number of human pseudokinases, although its importance for their function is unclear (4). In active kinases, the aspartate residue within the DFG motif plays a critical role in catalysis by coordinating Mg²⁺ ions (104). In a subset of these kinases, the DFG motif is found to adopt two distinct conformations, DFG-in and DFG-out, which are coupled to conformational changes in other regions of the kinase domain (105). The interactions made by the DFG aspartate are key to

these changes (106). It is therefore possible that in pseudokinases that conserve the DFG motif, such as PEAK3, the conformation of this motif could be coupled to functional conformational changes in the other parts of the pseudokinase domain.

We examined the importance of the DFG motif in PEAK3's function as a negative regulator of CrkII by mutating the DFG aspartate to asparagine (PEAK3 D330N) and testing the ability of this mutant to homodimerize, interact with CrkII, and inhibit CrkII-dependent membrane ruffles in cells. Interestingly, PEAK3 D330N did not dimerize as efficiently as wild type PEAK3 (Fig. 6D). Consistent with our observation that PEAK3 dimerization is necessary for CrkII binding, PEAK3 D330N exhibited markedly impaired binding to CrkII (Fig. 6E) and did not inhibit CrkII-dependent morphological changes to cell shape to the same extent as wild type PEAK3 (Fig. 6F; *SI Appendix*, Fig. S10). These data show that the integrity of the DFG motif is essential for PEAK3 function as a negative regulator of CrkII.

Discussion

The PEAK1 and Pragmin pseudokinases have grown in prominence due to the key roles they play in the regulation of cellular motility and oncogenesis (107). Here, we present evidence that PEAK3 is a close homolog of PEAK1 and Pragmin that likely has evaded annotation as a kinase due to the high LCR content within its pseudokinase domain. Despite this difference, PEAK3 retains features that are characteristic of the NKF3 family. These include the residues defined as the inhibitory triad, which occlude the nucleotide-binding pocket. Together with the mutations of several catalytic residues within the kinases domain, these features define PEAK3 as a pseudokinase. Based on sequence analysis and mutagenesis studies, we also demonstrate that PEAK3 self-associates through a conserved

SHED domain that flanks its pseudokinase domain, similar to PEAK1 and Pragmin. Hence, the presence of the SHED domain and its role in oligomerization are defining and unique features of NKF3 kinases.

Phylogenetic analysis of the NKF3 family shows that the ancestral NKF3, which likely appeared in the ancestor of Metazoans, was most similar to PEAK1 and was already a pseudokinase. This protein had an NFS motif instead of a DFG although it retained the HxD motif, which became HCD in human PEAK1, and in some PEAK1 and Pragmin homologs - a canonical HRD motif. It is intriguing that mammalian (but not avian) PEAK3 proteins have apparently reverted to the DFG motif while their HxD drifted into the LxE motif. Thus at present, our analysis supports a hypothesis that NKF3 family is an example of a kinase-like family that evolved originally as pseudokinases, similar to examples discussed by Kannan and co-workers (6).

Our unbiased search for PEAK3 binding partners reveals a propensity for PEAK3 to interact with regulators of cell motility, mirroring documented roles of PEAK1 and Pragmin (27-29, 31, 42, 43, 43, 90). While both Pragmin and PEAK1 are known to contain proline-rich CrkII binding motifs, and PEAK1 was shown to bind CrkII (31), our study is the first to describe a functional link between CrkII signaling and an NKF3 family member to our knowledge. We show here that PEAK3 regulates CrkII in a manner contrasting that of other NKF3 family members: PEAK3 inhibits CrkII while PEAK1 and Pragmin stimulate pro-motile signaling in cells and would be therefore expected to potentiate CrkII function. The underlying mechanism(s) for these differences is unclear but could reflect the distinct domain structure of PEAK3 compared to other NKF3 family members. Both PEAK1 and Pragmin contain large N-terminal regions that likely encode numerous unique functions

that are absent in PEA3. If PEA1 and Pragmin do indeed promote CrkII signaling, this function may be encoded within their N-terminal domain that is missing in PEA3.

We further demonstrate that the SHED domain-mediated dimerization of PEA3 is essential for its ability to bind CrkII. The emerging role of dimerization in binding seems to stem from the necessity to dimerize the CrkII-binding motifs themselves, as efficient CrkII binding can be recapitulated when these sites in PEA3 are brought together by an orthogonal dimerization module. Since the CrkII binding site in PEA3 represents the canonical proline-rich motif found in other known CrkII binding partners, we anticipate that these proteins might also regulate their interaction with CrkII through dimerization. Some of the known CrkII interactors, such as PEA1 and Pragmin, are known to exist as dimers. Interestingly, CrkII and its highly related homolog CrkL, which we also identified as a binding partner of PEA3 in the IP/MS analysis, have been shown to form dimers (108). Hence, the inherent property of both CrkII and some of its binding partners to oligomerize might be an essential mechanism for regulation of their mutual interactions and signaling.

While necessary for CrkII binding, dimerization of the CrkII binding motif is not sufficient for CrkII inhibition by PEA3 in the absence of the SHED and pseudokinase domains. These findings suggest the SHED-pseudokinase domain module plays an important role in CrkII inhibition beyond supporting PEA3 self-association. One such function could be mediating PEA3 heterodimerization with PEA1 and Pragmin, which we predict could occur based on sequence similarities within their SHED domains. Since all NFK3 family members have CrkII-binding sites, all possible NFK3 homo- and heterodimers could efficiently bind CrkII in theory. Given the opposing phenotypic outcomes between PEA3 and other NFK3 members on cell motility, heterodimerization of PEA3 with

PEAK1 or Pragmin could interfere with their positive signaling properties. The outstandingly shorter length of the N-terminal domain in PEAK3 relative to PEAK1 and Pragmin suggests that PEAK3 may have evolved to act as a dominant negative regulator of PEAK1 and Pragmin, and antagonizes their functions encoded by the N-terminal domains. The regulatory role of PEAK3 for other NKF3 pseudokinases would be consistent with the later appearance of PEAK3 in evolution compared to the other two family members. Future studies are needed to parse out the functional relationship of PEAK3 with its family members.

An additional unique feature of PEAK3, which distinguishes it from PEAK1 and Pragmin, is the presence of an intact DFG motif. Our data show that mutation of the DFG motif affects the ability of PEAK3 to homodimerize and subsequently interact with CrkII. The importance of the DFG motif for PEAK3 function is intriguing due to the critical role of this motif in kinase catalysis. While at present we cannot rule out that PEAK3 might be catalytically active, PEAK3 is an unlikely candidate for an active kinase based on its poor conservation of other residues in the putative active site. Rather, the loss-of-function effect of the DFG mutation suggests that it is a resulting conformational change within the pseudoactive site of PEAK3 that influences its dimerization and, consequently, its function. DFG mutations have previously been shown to affect the oligomerization of other kinases and pseudokinases. Notably, mutation of the DFG aspartate (D161N) in the kinase RIPK3 induces oligomerization and assembly of RIPK3 into a multimeric apoptotic complex (107, 109). Other kinase-inactivating mutations, such as those of the β 3 lysine or catalytic aspartate, do not affect RIPK3 oligomerization, indicating that the D161N mutation stabilizes a specific conformation of RIPK3 that promotes oligomerization (109). The

opposite effect is observed in the pseudokinase MLKL. In MLKL, the DFG motif is replaced by a GFE motif, and mutation of the GFE glutamate (E351K) prevents MLKL oligomerization. This mutation was proposed to stabilize a conformation that it is not permissive for the release of the adjacent 4-helix bundle (4HB) domain which drives oligomerization (14). The SHED domain in NKF3 kinases, composed of helices flanking the pseudokinase domain, maintains close contacts with the pseudokinase domain in the crystal structures of PEAK1 and Pragmin. It is therefore likely that conformational changes within the pseudokinase domain are sensed by the SHED domain and can ultimately affect dimerization.

Potential regulation of PEAK3 dimerization through conformational changes in its pseudokinase domain presents an exciting opportunity for pharmacological modulation of PEAK3 oligomerization, as previously achieved in RIPK3 (109, 110). While RIPK3 inhibitors target the ATP-binding site, this will likely not work for PEAK3, whose nucleotide-binding pocket is predicted to be occluded. However, recent studies on another pseudokinase, TRIB1, point to an important role of protein-protein interactions in regulating the conformation of the pseudoactive site. Like in PEAK3, the nucleotide-binding site in TRIB1 is highly occluded and inaccessible to ligands (9). Binding of the transcription factor C/EBP α to the C-lobe of TRIB1 alters the conformation of the pseudoactive site, including its equivalent DFG motif (SLE in TRIB1), which rotates to a semi-active position upon C/EBP α binding (111). The TRIB1 studies demonstrate that distal binding events can have global effects on the conformation of the pseudoactive site. Further studies can reveal if such interactions exist in PEAK3 and whether they can be leveraged for the pharmacological manipulation of its function.

The therapeutic relevance of targeting PEAK3 remains to be determined, but there is clear therapeutic potential (110). While at present there are no known prevalent disease-associated mutations in PEAK3, *PEAK3* mRNA levels are significantly elevated in acute myeloid leukemia (AML) patient samples relative to other cancer types (*SI Appendix*, Fig. S11)(112, 113). The E3 ubiquitin ligase SIAH1, which is a therapeutic target in AML (87), was identified as a PEAK3 binding partner in our IP/MS analysis. SIAH1 targets for degradation the oncogenic protein FMS-like tyrosine kinase 3 with internal tandem duplication mutation (FLT3-ITD), a mutant FLT3 variant detected in 40% of AML cases (114).

It is possible that PEAK3 interferes with this process through direct interaction with SIAH1. CrkL, which too was identified in our IP/MS screen, is one of the downstream effectors of the FLT3 signaling axis that contributes to leukemogenesis (115). Crk family proteins are known to play key roles in cancer invasion and migration by integrating and amplifying extracellular signals (83). Genetic knockdown of CrkII specifically decreases the cell migration and malignant potential of multiple human cancer cells including lung, breast, and ovarian cancers (83, 96, 116). Mechanistically, inhibition of CrkII leads to reduced or stochastic F-actin networks and reduction in lamellipodia (96, 99, 116), mirroring the morphological changes we observe when PEAK3 is overexpressed in cells. Therefore, pharmacological targeting of PEAK3 could prove useful in AML and potentially additionally types of cancers that are susceptible to inhibitors that target cellular motility pathways.

Materials and Methods

Multiple Sequence Alignment. The alignment was constructed using the Promals3D algorithm and manually adjusted using FFAS03 (25) pairwise 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, often missing in the PDB structures, were identified using the SEG server (79).

Phylogenetic tree construction. 29 sequences, representing vertebrate and non-vertebrate animals, were used to construct the phylogenetic tree. Homologs were selected from vertebrates that possess representatives of the three subfamilies and from four non-vertebrate organisms with single NKF3 homologs (picozoan *Trichoplax adherens*, sea urchin *Strongylocentrotus purpuratus*, lancelet *Branchiostoma belheri* and starfish *Acanthaster planci*). Multiple sequence alignment of the pseudokinase domains was built using the Promals3D algorithm (117) with manual adjustments. The phylogenetic tree for the set of 29 sequences was constructed using the PhyML method with aLRT statistics for calculating significance of branches (118). Branches with bootstrap values better than 70% were marked. Ancestral sequence reconstruction was performed using the Ancestron algorithm (28).

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 (119). 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 (120). 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 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 (121, 122). 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 1×10^7 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 $\geq 200 \times g$, 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 cComplete mini EDTA-free protease and PhosSTOP phosphatase inhibitor cocktails (Roche), incubated on a tube rotator at 4°C for 30 minutes, and centrifuged at $3,500 \times g$, 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 NH_4HCO_3 , 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 μ m I.D. packed with ReproSil Pur C18 AQ 5 μ m 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 μ m I.D. packed with ReproSil Pur C18 AQ 3 μ m 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 (123). 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 (92).

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, HEK293 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 $\text{Na}(\text{VO}_4)_3$, 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 before being transferred onto PVDF membranes for Western blot analysis. Following transfer, membranes were blocked in 5% milk diluted in 1xTBS and 1% Tween-20 (TBS-T) for 30 minutes at room temperature. Membranes were then incubated with primary antibodies diluted in blocking buffer overnight at 4°C (anti-CrkII (rabbit, Proteintech), anti-FLAG (mouse, Sigma Aldrich; rabbit, CST), and anti-HA (mouse, SCBT)). Primary antibodies were washed off 3 times for 5 minutes using TBS-T. Membranes were incubated with secondary antibodies diluted in blocking buffer for 2 hours at room temperature (anti-mouse IgG Veriblot, Abcam; anti-IgG Veriblot, Abcam) then washed 4 times for 5-8 minutes using TBS-T. Proteins were then visualized using ECL Western blotting detection reagent (GE) or ECL prime (VWR).

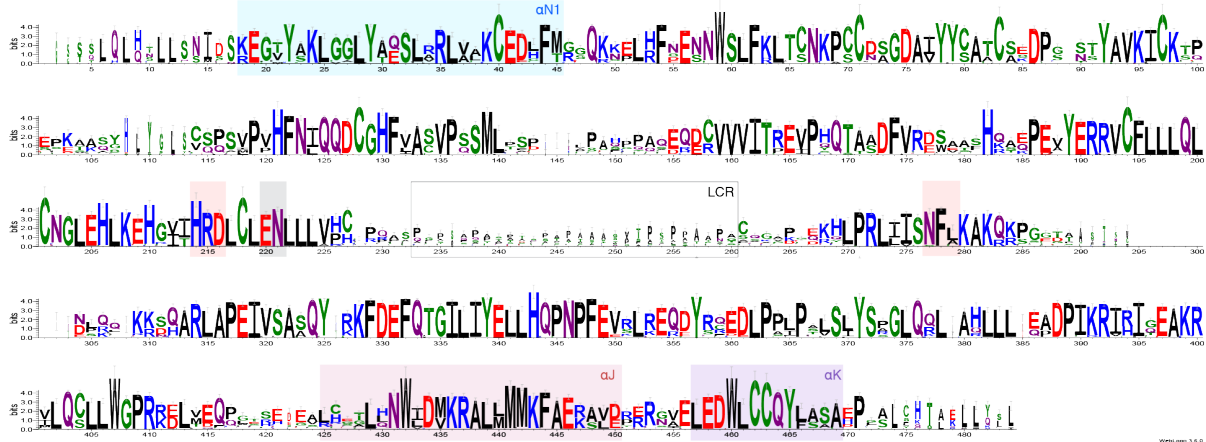
Immunofluorescence analysis. COS-7 cells (7.0×10^4 per well) were plated onto glass coverslips and transfected the following day. 24 hours post transfection, cells were fixed in 3.7% formaldehyde in PBS for 1 hour at room temperature or overnight at 4°C, and subsequently permeabilized using 0.1% Triton-X in PBS for 5 minutes, and incubated with a blocking buffer (1% BSA in PBS) for 5 minutes. The 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 of PEAK3 and CrkII-binding mutant (PEAK3-3A) on actin stress fibers were blindly scored. Cells were binned based on the extent of visible actin fibers within the cytosolic region of the cell into 3 categories: (1) long, prominent fibers that traversed over 50% of the cell, (2) short, stochastic fibers, or (3) no appreciable amount of polymerized actin. 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, as shown in Fig. 3c. 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 (124, 125).

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 (126)) 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 (127). 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 (128)), were colored according to BLOSUM62 scores as follows: yellow: BLOSUM62 values from 1 to 3; orange: 4-6; red (highest conservation): 7 to 11.

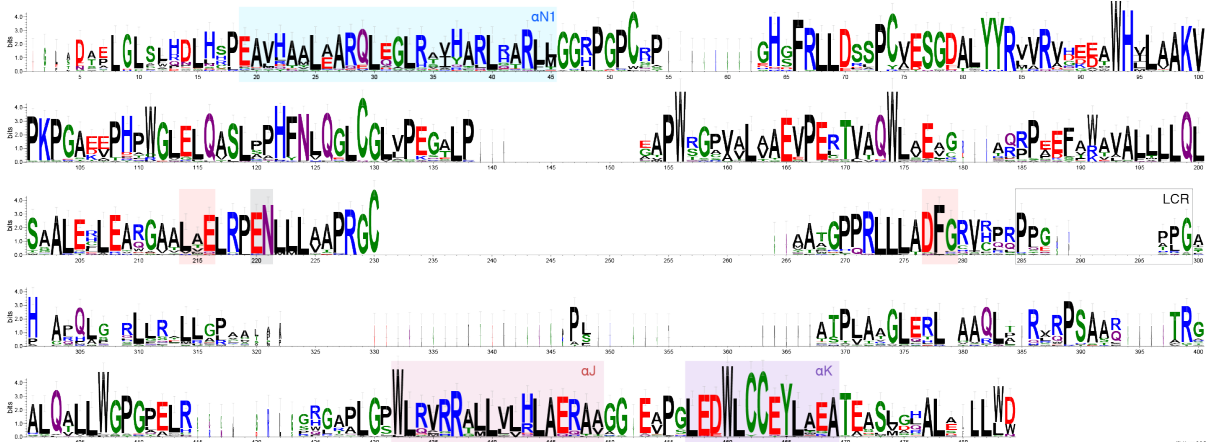
Supplemental Figures

Supplemental Figure 1

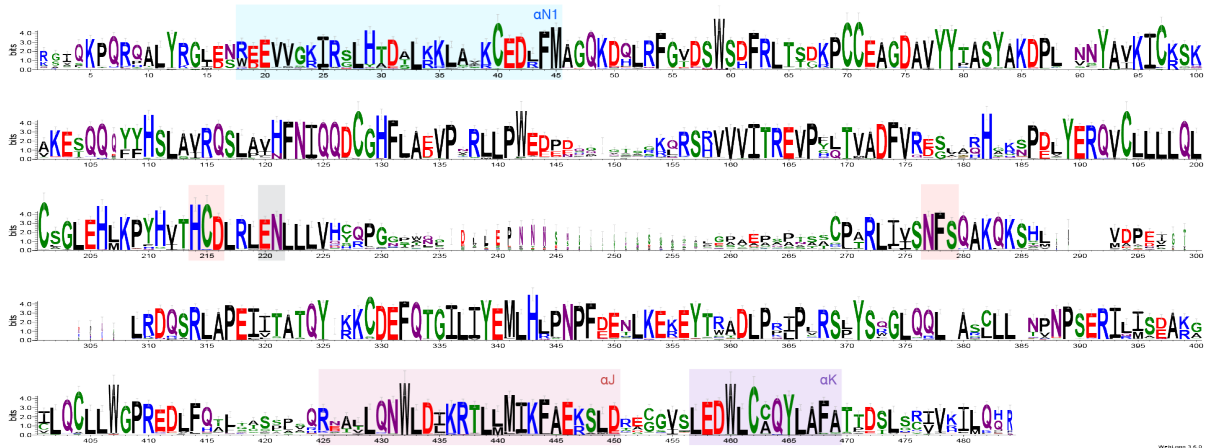
a Pragmin/Sgk223



b PEAK3/C19orf35

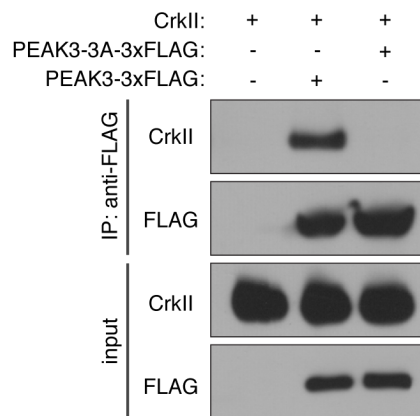


c PEAK1/Sgk269



Supplementary Figure 1. 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 (α N1 - highlighted blue), residues 425-250 (Pragmin, PEAK1) or 432-449 (PEAK3) (α J - highlighted magenta), and residues 457-469 (α K - 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.

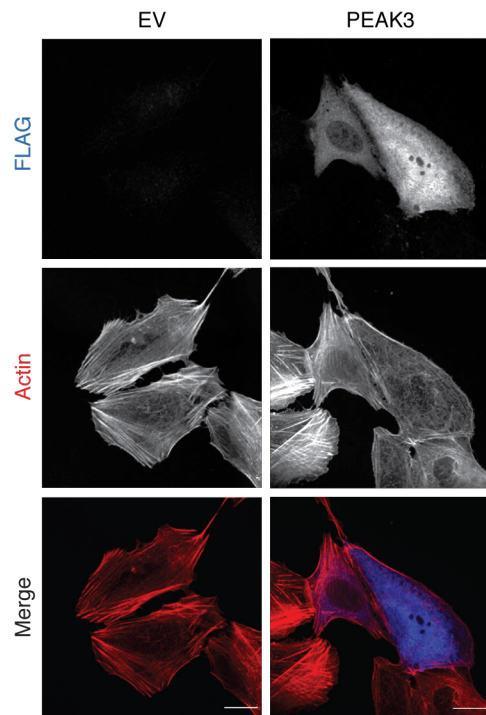
Supplemental Figure 2



Supplementary Figure 2. Exogenously expressed PEAK3 interacts with CrkII.

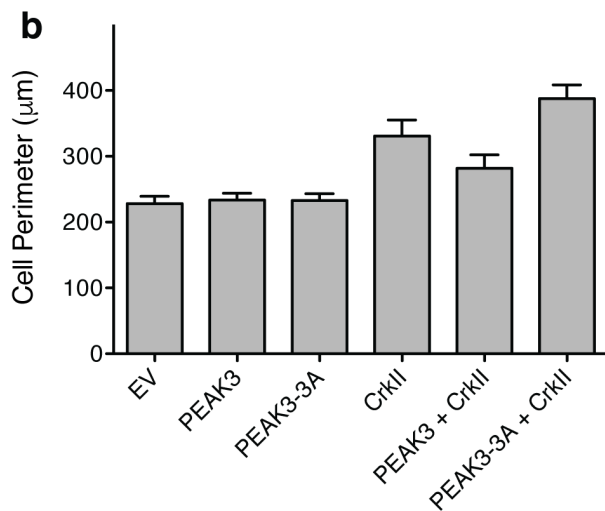
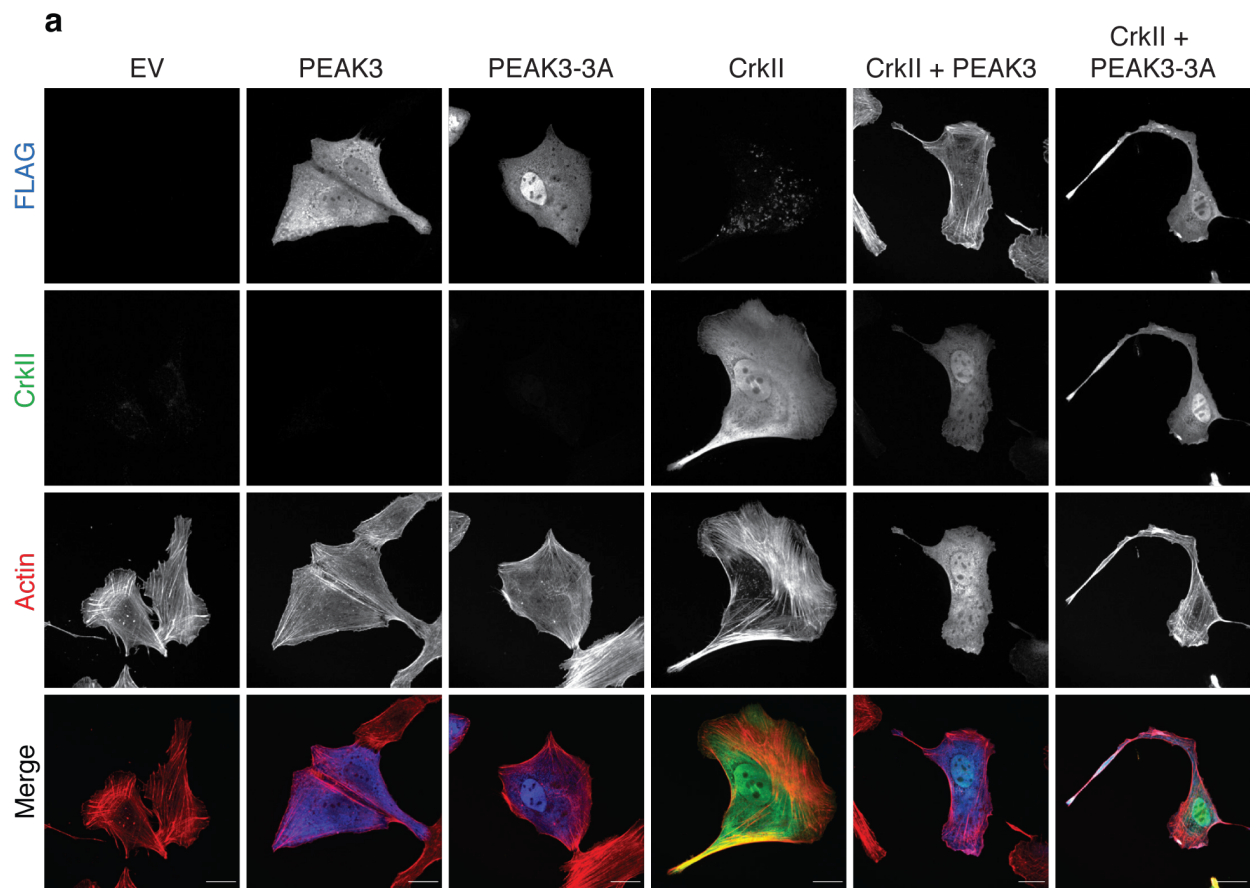
Coimmunoprecipitation of FLAG-tagged wild type PEAK3, transiently expressed in HEK293 cells, with exogenously expressed, untagged CrkII. Cell lysates were incubated with an anti-FLAG antibody, resolved on SDS/PAGE, and probed with the indicated antibodies by Western blot. Data is representative of three independent experiments.

Supplemental Figure 3



Supplementary Figure 3. 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). All scale bars correspond to 20 μm .

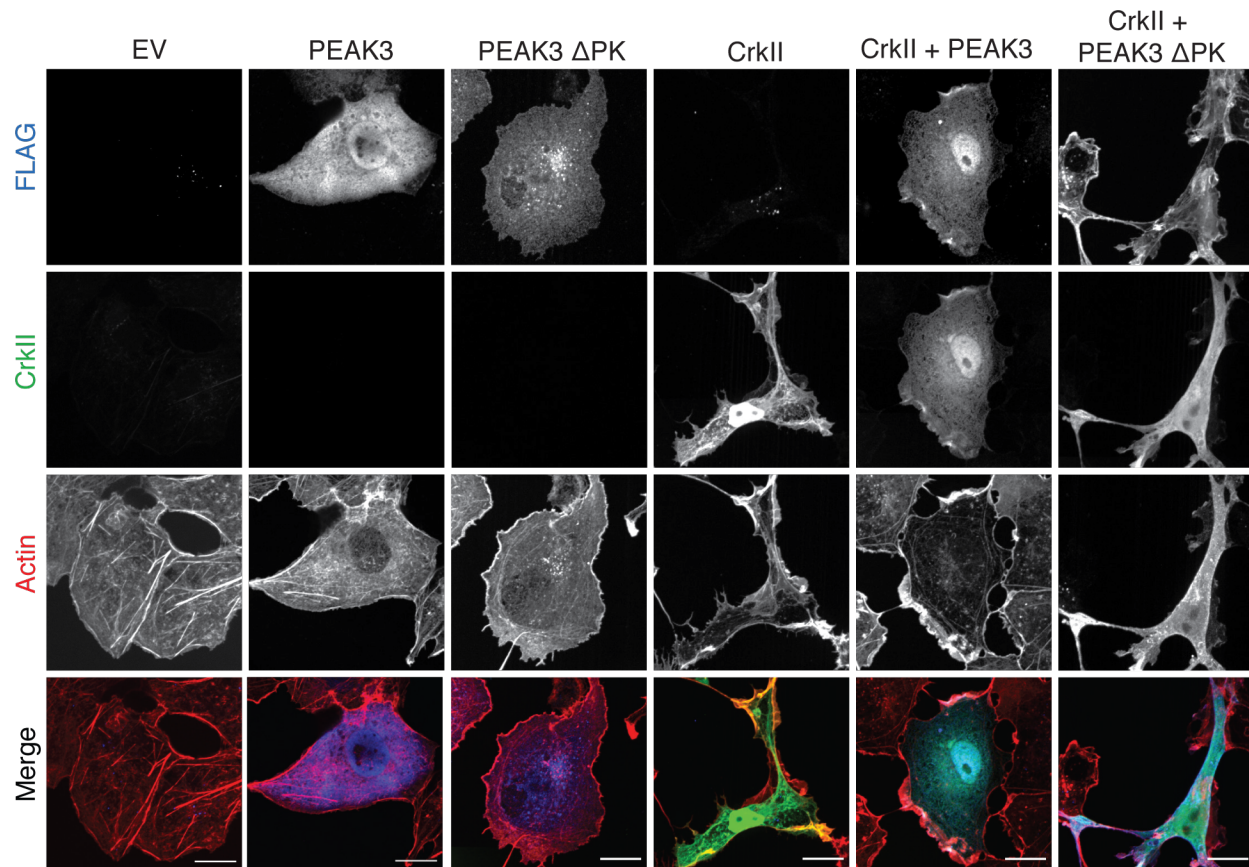
Supplemental Figure 4



Supplementary Figure 4. 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 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 \pm SEM of three independent experiments (n = 20 cells in each experiment).

Supplemental Figure 5

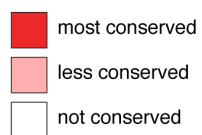
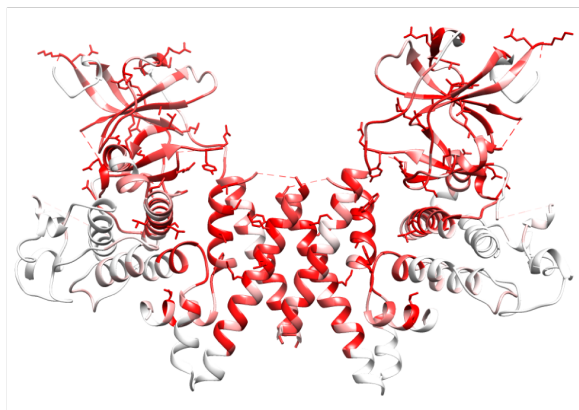


Supplementary Figure 5. CrkII binding motif is insufficient to drive inhibition of CrkII function. Confocal microscopy imaging of COS-7 cells transiently co-transfected with a FLAG-tagged wild type or Δ PK mutant of PEAK3 with either an empty vector or untagged CrkII. CrkII was detected with anti-CrkII 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 .

Supplemental Figure 6

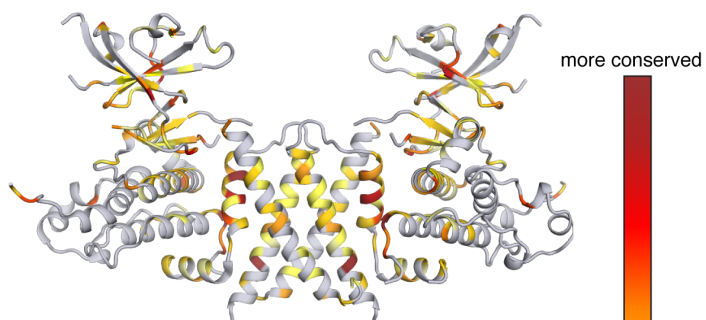
a

PEAK3 evolutionary conservation



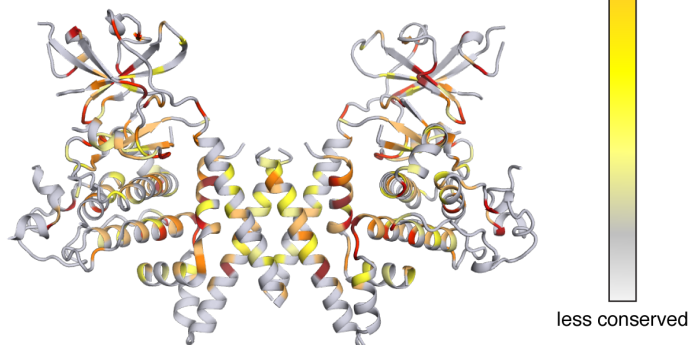
b

PEAK3 vs. Pragmin



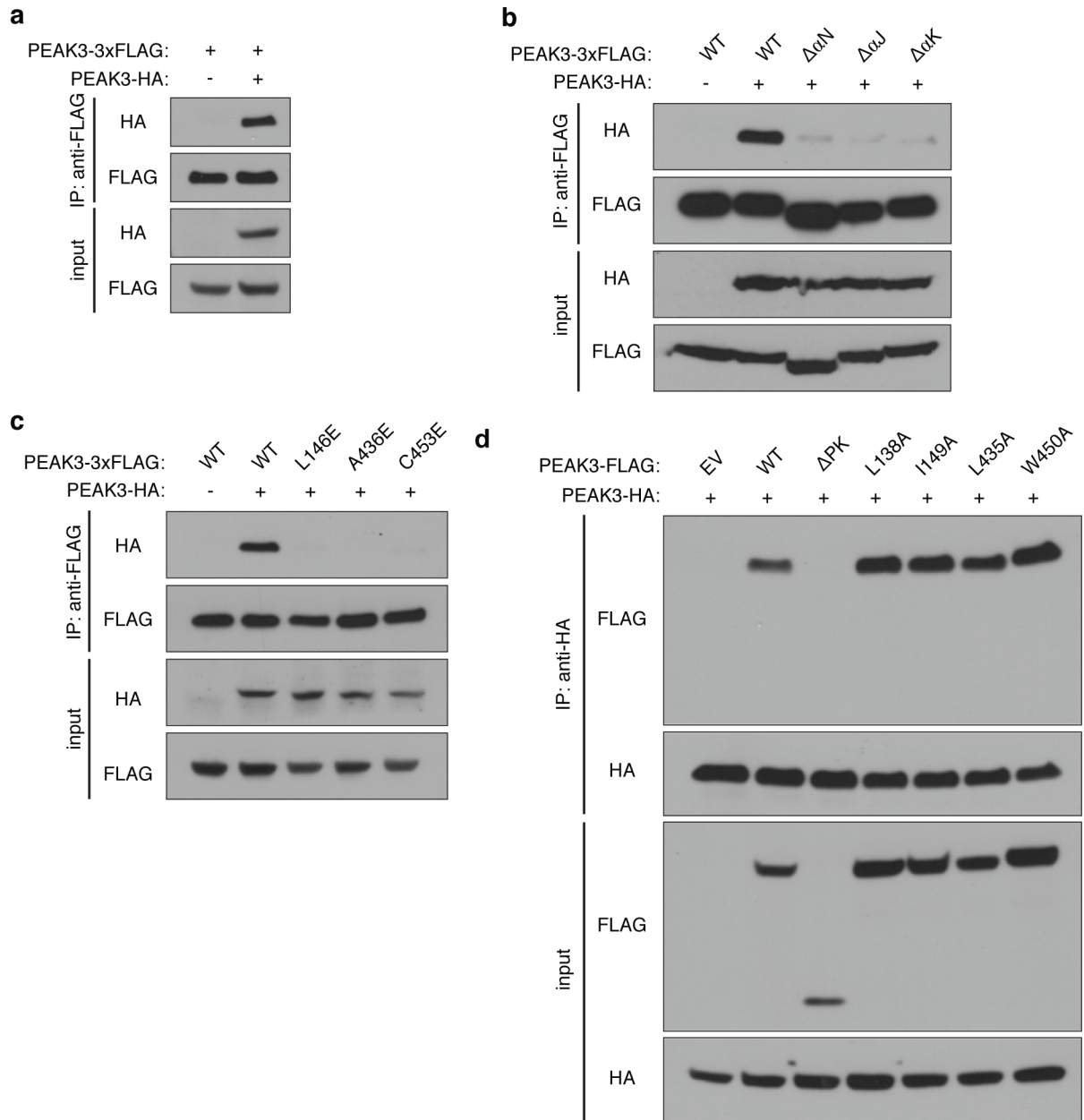
c

PEAK3 vs. PEAK1



Supplementary Figure 6. 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.

Supplemental Figure 7

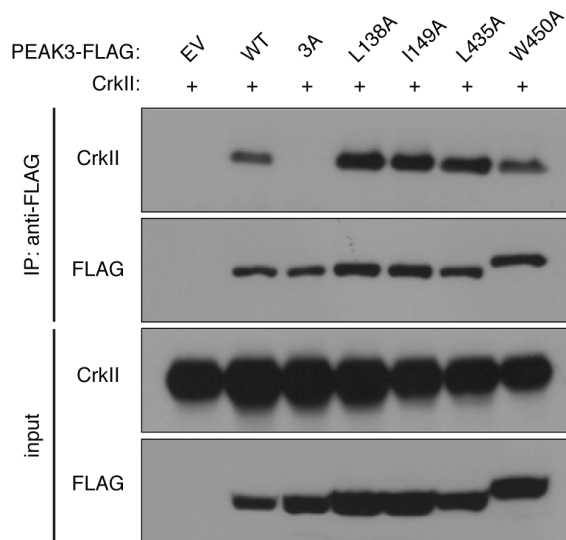


Supplementary Figure 7. Mutations within the putative SHED domain disrupts the ability of PEAK3 to dimerize. (a - c) Co-immunoprecipitation FLAG-tagged WT PEAK3 with either HA-tagged PEAK3 (a) or mutant variants where either one of three alpha helices thought to be important for dimerization is deleted (b) or residues in the helical interface is mutated (c), transiently expressed in HEK293 cells. Cell lysates were incubated with an anti-FLAG antibody, resolved by SDS/PAGE, and then probed with the indicated antibodies by Western blot. **(d)** Co-immunoprecipitation of an HA-tagged WT construct of PEAK3 with a FLAG-tagged WT or mutants, transiently expressed in HEK293 cells. Cell lysates were incubated with an anti-FLAG antibody, resolved by SDS/PAGE, and then probed with the indicated antibodies by Western blot. All coimmunoprecipitation data are

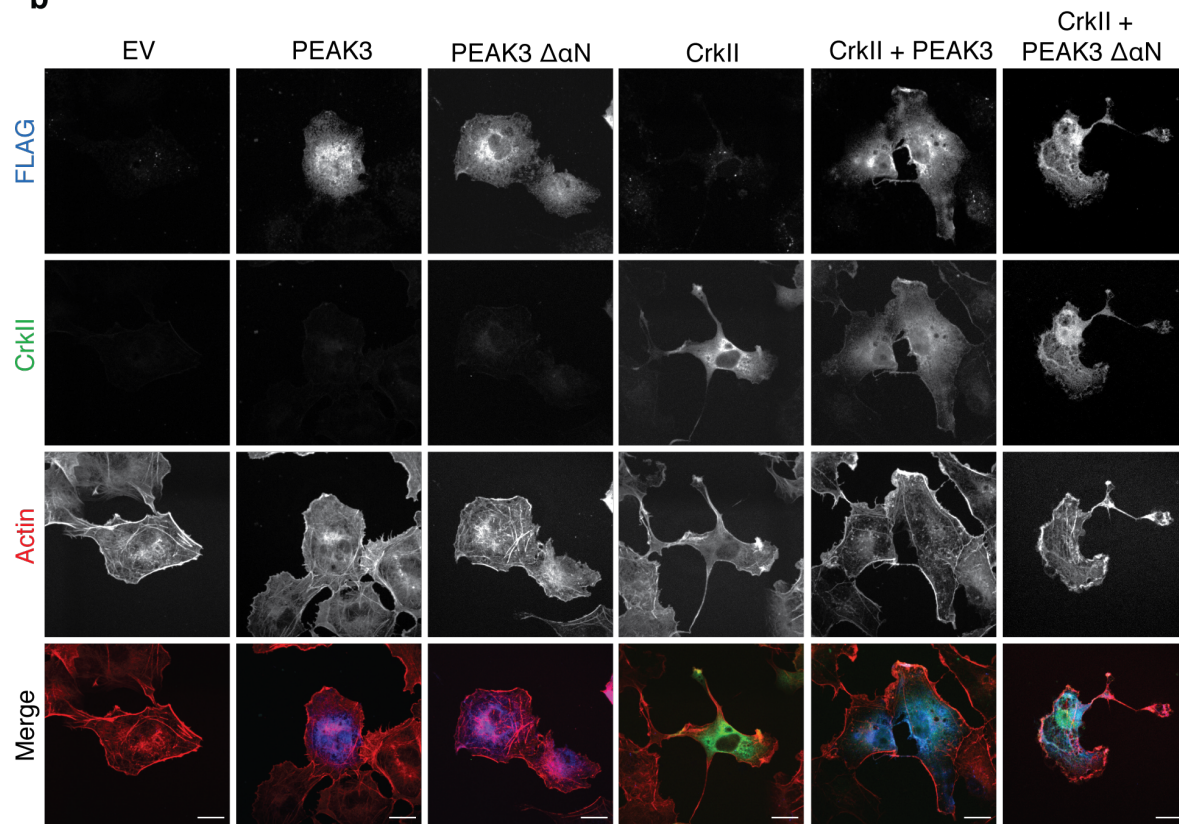
representative of at two independent experiments.

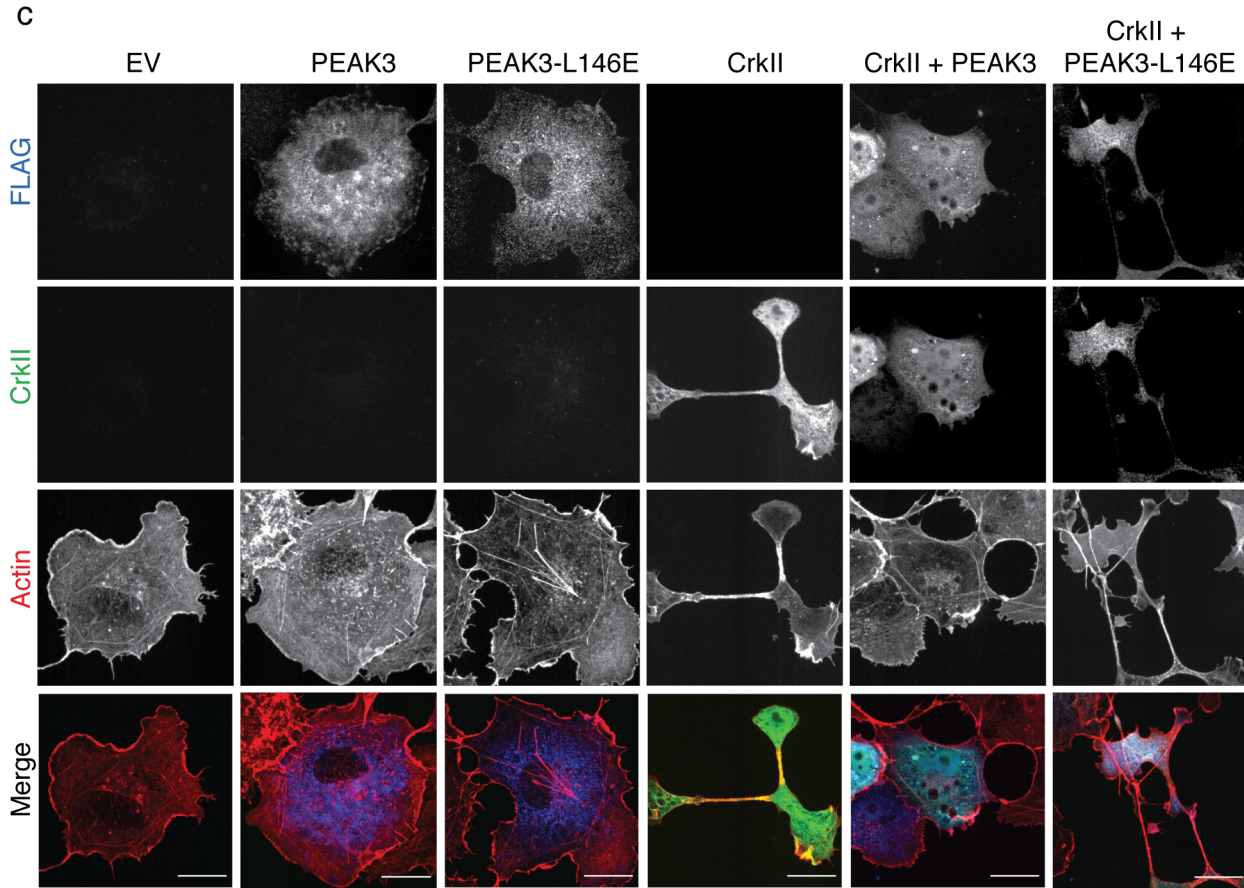
Supplemental Figure 8

a



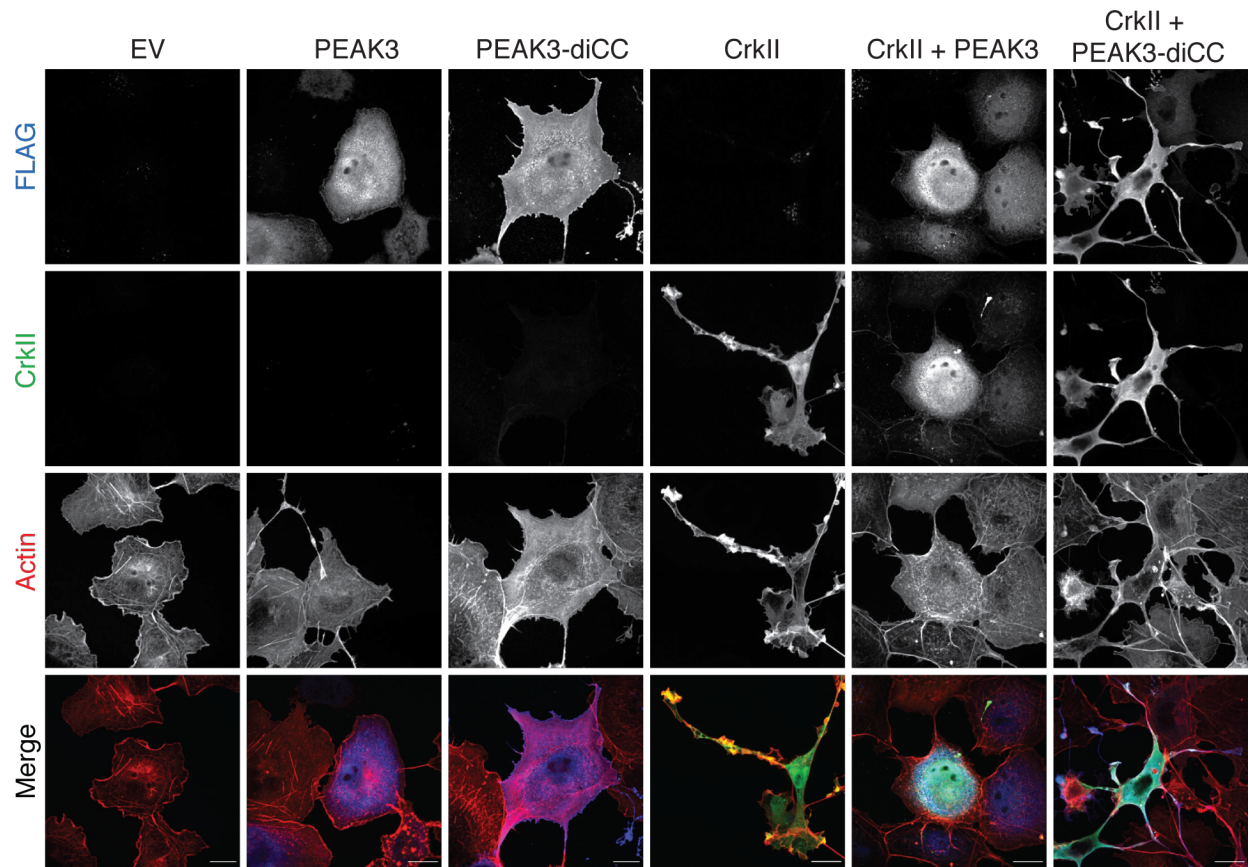
b





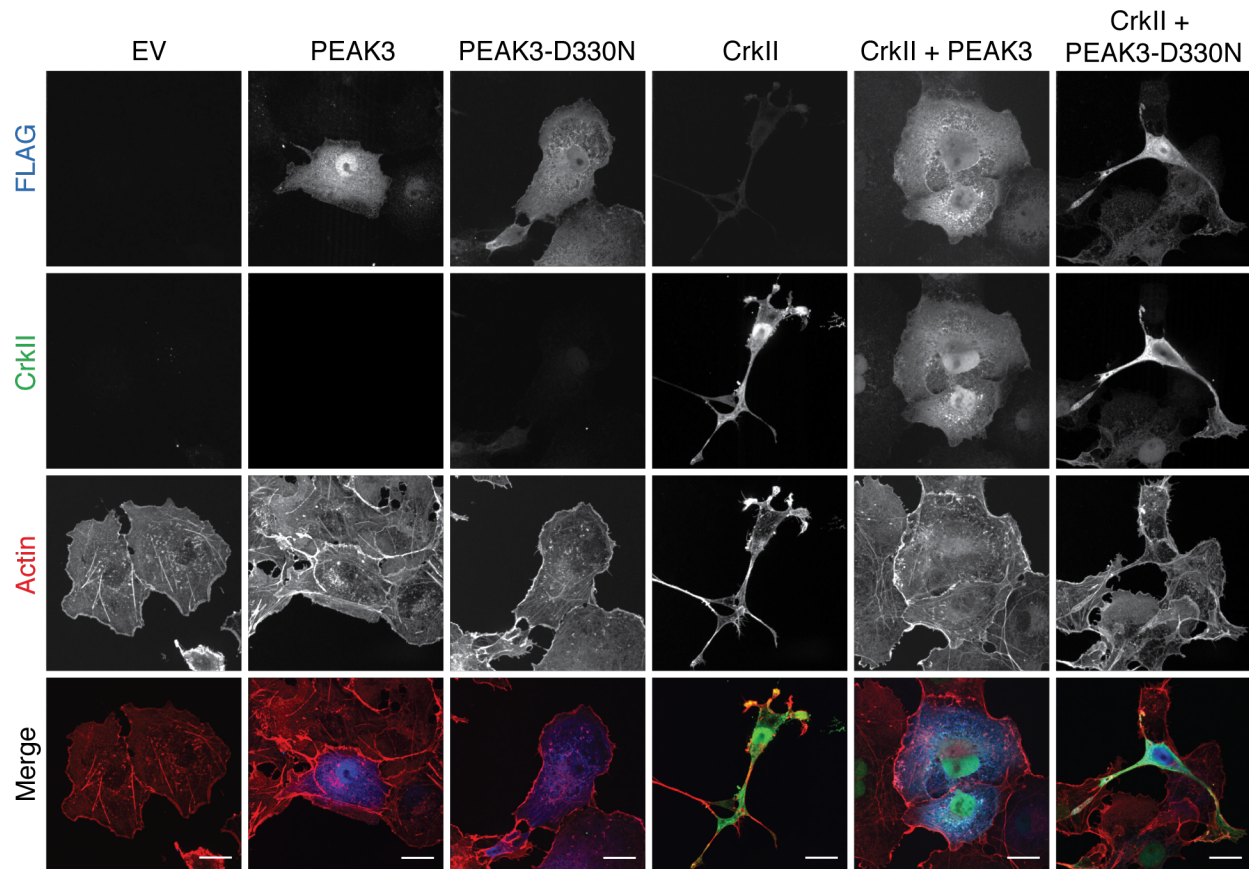
Supplementary Figure 8. Homotypic association of PEAK3 is required for its interaction with CrkII. **(a)** Co-immunoprecipitation of untagged CrkII with FLAG- tagged control dimerization mutants of PEAK3, transiently expressed in HEK293 cells. Cell lysates were incubated with an anti-FLAG antibody, resolved by SDS/PAGE, and then probed 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 $\Delta\alpha N1$ mutant of PEAK3 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 . **(c)** Confocal microscopy imaging of COS-7 cells transiently co-transfected with a FLAG-tagged wild type or L146E dimerization mutant of PEAK3 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 .

Supplemental Figure 9



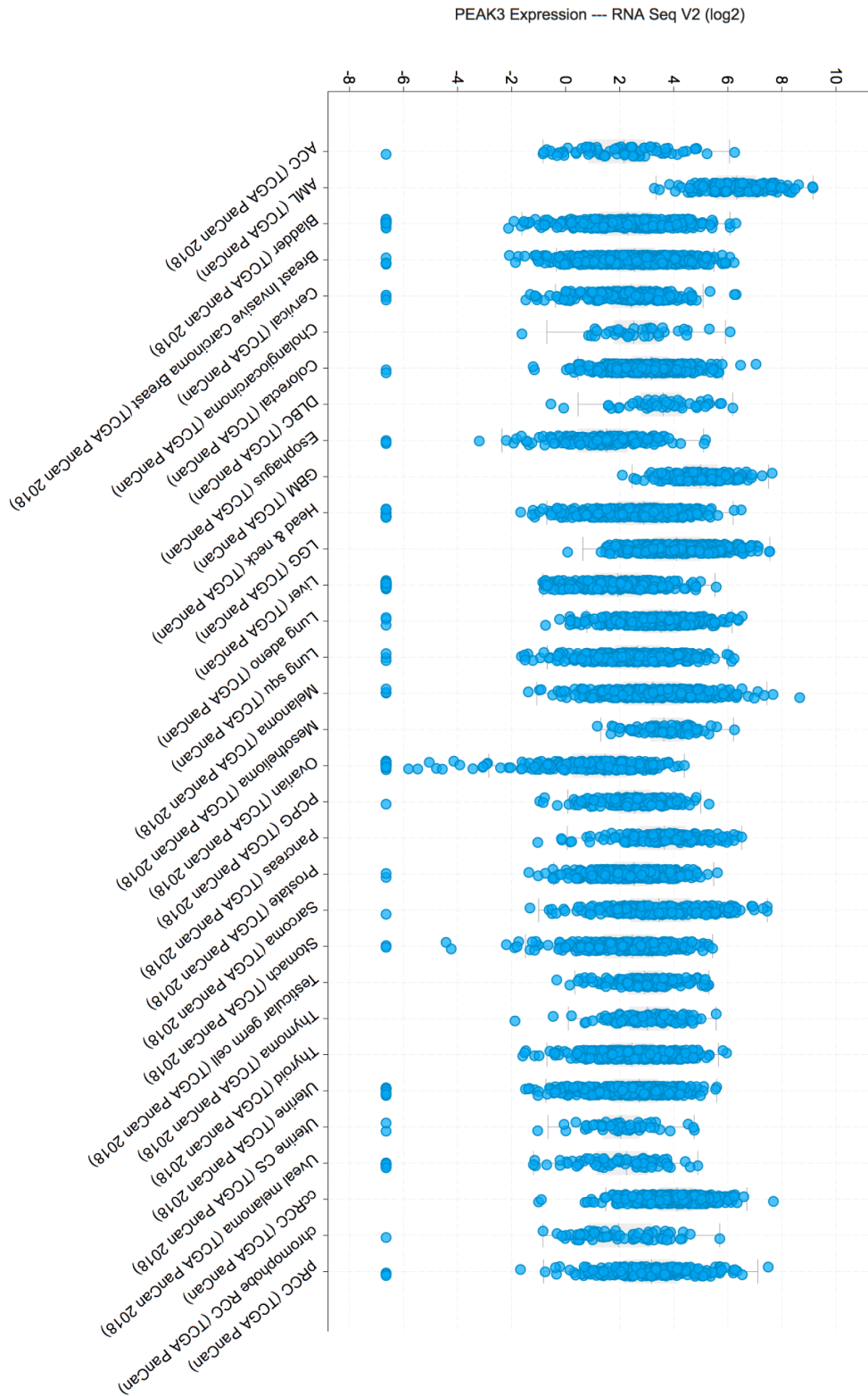
Supplementary Figure 9. Constitutive dimerization of the CrkII binding site is not sufficient to antagonize CrkII function. Confocal microscopy imaging of COS-7 cells transiently co-transfected with a FLAG-tagged wild type or diCC mutant of PEA3 with either an empty vector or untagged CrkII. CrkII was detected with anti-CrkII antibody (green), PEA3 with anti-FLAG antibody (blue) and F-actin with Alexa Fluor 647-conjugated phalloidin staining (red). All scale bars correspond to 20 μm .

Supplemental Figure 10



Supplemental Figure 10. Mutation of the DFG motif in PEA3 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 PEA3 with either an empty vector or untagged CrkII. CrkII was detected with anti-CrkII antibody (green), PEA3 with anti-FLAG antibody (blue) and F-actin with Alexa Fluor 647-conjugated phalloidin staining (red). All scale bars correspond to 20 μ m.

Supplemental Figure 11



Supplementary Figure 11. PEAK3 mRNA levels in patient-derived AML cell lines. PEAK3 mRNA levels in patient-derived cancer cell lines from the Cancer Genome Atlas Project.

Supplementary Table 1. 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.

Supplemental Table 1

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

Works Cited

1. Johnson H, Lescarbeau RS, Gutierrez JA, White FM (2013) Phosphotyrosine profiling of NSCLC cells in response to EGF and HGF reveals network specific mediators of invasion. *J Proteome Res* 12:1856–1867.
2. Hanks SK, Hunter T (1995) Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J* 9:576–596.
3. Manning G, Plowman GD, Hunter T, Sudarsanam S (2002) Evolution of protein kinase signaling from yeast to man. *Trends Biochem Sci* 27:514–520.
4. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S (2002) The protein kinase complement of the human genome. *Science* 298:1912–1934.
5. Jura N, Shan Y, Cao X, Shaw DE, Kuriyan J (2009) Structural analysis of the catalytically inactive kinase domain of the human EGF receptor 3. *Proc Natl Acad Sci USA* 106:21608–21613.
6. Kwon A et al. (2019) Tracing the origin and evolution of pseudokinases across the tree of life. *Sci Signal* 12.
7. Kannan N, Taylor SS, Zhai Y, Venter JC, Manning G (2007) Structural and functional diversity of the microbial kinome. *PLoS Biol* 5:e17.
8. Shaw AS, Kornev AP, Hu J, Ahuja LG, Taylor SS (2014) Kinases and pseudokinases:

- lessons from RAF. *Mol Cell Biol* 34:1538–1546.
9. Murphy JM et al. (2014) A robust methodology to subclassify pseudokinases based on their nucleotide-binding properties. *Biochem J* 457:323–334.
 10. Shi F, Telesco SE, Liu Y, Radhakrishnan R, Lemmon MA (2010) ErbB3/HER3 intracellular domain is competent to bind ATP and catalyze autophosphorylation. *Proc Natl Acad Sci USA* 107:7692–7697.
 11. Kung JE, Jura N (2016) Structural Basis for the Non-catalytic Functions of Protein Kinases. *Structure/Folding and Design* 24:7–24.
 12. Xu B et al. (2000) WNK1, a novel mammalian serine/threonine protein kinase lacking the catalytic lysine in subdomain II. *J Biol Chem* 275:16795–16801.
 13. Piala AT et al. (2014) Chloride sensing by WNK1 involves inhibition of autophosphorylation. *Sci Signal* 7:ra41.
 14. Petrie EJ et al. (2018) Conformational switching of the pseudokinase domain promotes human MLKL tetramerization and cell death by necroptosis. *Nature Communications* 9:2422.
 15. Sun L et al. (2012) Mixed lineage kinase domain-like protein mediates necrosis signaling downstream of RIP3 kinase. *Cell* 148:213–227.
 16. Ridley AJ (1996) Rho: theme and variations. *Current Biology* 6:1256–1264.
 17. Hodge RG, Ridley AJ (2016) Regulating Rho GTPases and their regulators. *Nature*

Publishing Group 17:496–510.

18. Takahashi M et al. (2008) Sequential activation of Rap1 and Rac1 small G proteins by PDGF locally at leading edges of NIH3T3 cells. *Genes Cells* 13:549–569.
19. Hu J et al. (2011) Mutation that blocks ATP binding creates a pseudokinase stabilizing the scaffolding function of kinase suppressor of Ras, CRAF and BRAF. *Proc Natl Acad Sci USA* 108:6067–6072.
20. Kornev AP, Taylor SS (2010) Defining the conserved internal architecture of a protein kinase. *Biochim Biophys Acta* 1804:440–444.
21. Taylor SS, Shaw A, Hu J, Meharena HS, Kornev A (2013) Pseudokinases from a structural perspective. *Biochem Soc Trans* 41:981–986.
22. Roskoski R (2015) A historical overview of protein kinases and their targeted small molecule inhibitors. *Pharmacol Res* 100:1–23.
23. Rajakulendran T, Sahmi M, Lefrançois M, Sicheri F, Therrien M (2009) A dimerization-dependent mechanism drives RAF catalytic activation. *Nature* 461:542–545.
24. Brennan DF et al. (2011) A Raf-induced allosteric transition of KSR stimulates phosphorylation of MEK. *Nature* 472:366–369.
25. Jaroszewski L, Li Z, Cai X-H, Weber C, Godzik A (2011) FFAS server: novel features and applications. *Nucleic Acids Research* 39:W38–44. Available at: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=21715>

387&retmode=ref&cmd=prlinks.

26. Tanaka H, Katoh H, Negishi M (2006) Pragmin, a novel effector of Rnd2 GTPase, stimulates RhoA activity. *J Biol Chem* 281:10355–10364.
27. Tactacan CM et al. (2015) The pseudokinase SgK223 promotes invasion of pancreatic ductal epithelial cells through JAK1/Stat3 signaling. *Mol Cancer* 14:139.
28. Senda Y, Murata-Kamiya N, Hatakeyama M (2016) C-terminal Src kinase-mediated EPIYA phosphorylation of Pragmin creates a feed-forward C-terminal Src kinase activation loop that promotes cell motility. *Cancer Sci* 107:972–980.
29. Fatemeh Safari, Murata-Kamiya N, Saito Y, Hatakeyama M (2011) Mammalian Pragmin regulates Src family kinases via the Glu-Pro-Ile-Tyr-Ala (EPIYA) motif that is exploited by bacterial effectors. *PNAS* vol. 108:14938e14943.
30. Lecointre C et al. (2018) Dimerization of the Pragmin Pseudo-Kinase Regulates Protein Tyrosine Phosphorylation. *Structure/Folding and Design* 26:545–554.e4.
31. Wang Y et al. (2010) Pseudopodium-enriched atypical kinase 1 regulates the cytoskeleton and cancer progression [corrected]. *Proc Natl Acad Sci USA* 107:10920–10925.
32. Zheng Y et al. (2013) Temporal regulation of EGF signalling networks by the scaffold protein Shc1. *Nature* 499:166–171.
33. Kelber JA et al. (2012) KRas induces a Src/PEAK1/ErbB2 kinase amplification loop that drives metastatic growth and therapy resistance in pancreatic cancer. *Cancer*

Research 72:2554–2564.

34. Fujimura K et al. (2014) A hypusine-eIF5A-PEAK1 switch regulates the pathogenesis of pancreatic cancer. *Cancer Research* 74:6671–6681.
35. Ding C et al. (2018) Overexpression of PEAK1 contributes to epithelial-mesenchymal transition and tumor metastasis in lung cancer through modulating ERK1/2 and JAK2 signaling. *Cell Death Dis* 9:802.
36. Croucher DR et al. (2013) Involvement of Lyn and the atypical kinase Sgk269/PEAK1 in a basal breast cancer signaling pathway. *Cancer Research* 73:1969–1980.
37. Fujimura K, Wang H, Watson F, Klemke RL (2018) KRAS oncoprotein expression is regulated by a self-governing eIF5A-PEAK1 feed-forward regulatory loop. *Cancer Research* 78:1444–1456.
38. Agajanian M, Runa F, Kelber JA (2015) Identification of a PEAK1/ZEB1 signaling axis during TGF β /fibronectin-induced EMT in breast cancer. *Biochem Biophys Res Commun* 465:606–612.
39. Wozniak MA, Modzelewska K, Kwong L, Keely PJ (2004) Focal adhesion regulation of cell behavior. *Biochim Biophys Acta* 1692:103–119.
40. Kain KH, Klemke RL (2001) Inhibition of cell migration by Abl family tyrosine kinases through uncoupling of Crk-CAS complexes. *J Biol Chem* 276:16185–16192.
41. Feller SM (2001) Crk family adaptors-signalling complex formation and biological

- roles. *Oncogene* 20:6348–6371.
42. Bristow JM, Reno TA, Jo M, Gonias SL, Klemke RL (2013) Dynamic phosphorylation of tyrosine 665 in pseudopodium-enriched atypical kinase 1 (PEAK1) is essential for the regulation of cell migration and focal adhesion turnover. *Journal of Biological Chemistry* 288:123–131.
 43. Liu L et al. (2016) Homo- and Heterotypic Association Regulates Signaling by the SgK269/PEAK1 and SgK223 Pseudokinases. *Journal of Biological Chemistry* 291:21571–21583.
 44. Ha BH, Boggon TJ (2017) The crystal structure of pseudokinase PEAK1 (Sugen Kinase 269) reveals an unusual catalytic cleft and a novel mode of kinase fold dimerization. *Journal of Biological Chemistry*.
 45. Patel O et al. (2017) Structure of SgK223 pseudokinase reveals novel mechanisms of homotypic and heterotypic association. *Nature Communications*:1–15.
 46. Holm L (2019) Benchmarking Fold Detection by DaliLite v.5. *Bioinformatics*.
 47. Birge RB, Fajardo JE, Mayer BJ, Hanafusa H (1992) Tyrosine-phosphorylated epidermal growth factor receptor and cellular p130 provide high affinity binding substrates to analyze Crk-phosphotyrosine-dependent interactions in vitro. *J Biol Chem* 267:10588–10595.
 48. Shigeno-Nakazawa Y et al. (2016) A pre-metazoan origin of the CRK gene family and co-opted signaling network. *Nature Publishing Group*:1–14.

49. Abassi YA, Vuori K (2002) Tyrosine 221 in Crk regulates adhesion-dependent membrane localization of Crk and Rac and activation of Rac signaling. *EMBO J* 21:4571–4582.
50. Saleh T et al. (2015) Cyclophilin A promotes cell migration via the Abl-Crk signaling pathway. *Nat Chem Biol* 12:117–123.
51. Klemke RL et al. (1998) CAS/Crk Coupling Serves as a “Molecular Switch” for Induction of Cell Migration. *J Cell Biol* 140:961–972. Available at: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=9472046&retmode=ref&cmd=prlinks>.
52. Kiyokawa E, Yuko Hashimoto SK, Sugimura H, Kurata T, Matsuda M (1998) Activation of Rac1 by a Crk SH3-binding protein, DOCK180. *GENES DEVELOPMENT* 12:3331–3336:1–7.
53. Okada S, Matsuda M, Anafi M, Pawson T, Pessin JE (1998) Insulin regulates the dynamic balance between Ras and Rap1 signaling by coordinating the assembly states of the Grb2-SOS and CrkII-C3G complexes. *EMBO J* 17:2554–2565.
54. Dolfi F et al. (1998) The adaptor protein Crk connects multiple cellular stimuli to the JNK signaling pathway. *Proc Natl Acad Sci USA* 95:15394–15399.
55. Makino Y et al. (2006) Elmo1 inhibits ubiquitylation of Dock180. *Journal of Cell Science* 119:923–932.
56. Tina L Gumienny et al. (2001) CED-12/ELMO, a Novel Member of the

- CrkII/Dock180/ Rac Pathway, Is Required for Phagocytosis and Cell Migration. 1–15.
57. Guo F, Debidda M, Yang L, Williams DA, Zheng Y (2006) Genetic deletion of Rac1 GTPase reveals its critical role in actin stress fiber formation and focal adhesion complex assembly. *J Biol Chem* 281:18652–18659.
 58. Lamorte L, Kamikura DM, Park M (2000) A switch from p130Cas/Crk to Gab1/Crk signaling correlates with anchorage independent growth and JNK activation in cells transformed by the Met receptor oncoprotein. *Oncogene* 19:5973–5981.
 59. Knudsen BS et al. (1995) Affinity and specificity requirements for the first Src homology 3 domain of the Crk proteins. *EMBO J* 14:2191–2198.
 60. Matsuda M et al. (1996) Interaction between the amino-terminal SH3 domain of CRK and its natural target proteins. *J Biol Chem* 271:14468–14472.
 61. Hoeve ten J et al. (1994) Cellular interactions of CRKL, and SH2-SH3 adaptor protein. *Cancer Research* 54:2563–2567.
 62. Salgia R et al. (1995) CRKL links p210BCR/ABL with paxillin in chronic myelogenous leukemia cells. *J Biol Chem* 270:29145–29150.
 63. Sattler M, Salgia R (1998) Role of the adapter protein CRKL in signal transduction of normal hematopoietic and BCR/ABL-transformed cells. *Leukemia* 12:637–644.
 64. Gelkop S, Babichev Y, Isakov N (2001) T Cell Activation Induces Direct Binding of the Crk Adapter Protein to the Regulatory Subunit of Phosphatidylinositol 3-Kinase

- (p85) via a Complex Mechanism Involving the Cbl Protein. *Journal of Biological Chemistry* 276:36174–36182.
65. Uemura N, Salgia R, Ewaniuk DS, Little MT, Griffin JD (1999) Involvement of the adapter protein CRKL in integrin-mediated adhesion. *Oncogene* 18:3343–3353.
 66. Hemmeryckx B et al. (2001) Crkl enhances leukemogenesis in BCR/ABL P190 transgenic mice. *Cancer Research* 61:1398–1405.
 67. George B et al. (2014) Crk1/2 and CrkL form a hetero-oligomer and functionally complement each other during podocyte morphogenesis. *Kidney Int* 85:1382–1394.
 68. Ichiba T et al. (1997) Enhancement of guanine-nucleotide exchange activity of C3G for Rap1 by the expression of Crk, CrkL, and Grb2. *J Biol Chem* 272:22215–22220.
 69. Nishihara H et al. (2002) DOCK2 associates with CrkL and regulates Rac1 in human leukemia cell lines. *Blood* 100:3968–3974.
 70. Park T-J, Curran T (2014) Essential roles of Crk and CrkL in fibroblast structure and motility. *Oncogene* 33:5121–5132.
 71. Park T-J, Boyd K, Curran T (2006) Cardiovascular and craniofacial defects in Crk-null mice. *Mol Cell Biol* 26:6272–6282.
 72. Moon AM et al. (2006) Crkl deficiency disrupts Fgf8 signaling in a mouse model of 22q11 deletion syndromes. *Dev Cell* 10:71–80.
 73. Guris DL, Duester G, Papaioannou VE, Imamoto A (2006) Dose-dependent

- interaction of Tbx1 and Crkl and locally aberrant RA signaling in a model of del22q11 syndrome. *Dev Cell* 10:81–92.
74. Feller SM, Knudsen B, Hanafusa H (1994) c-Abl kinase regulates the protein binding activity of c-Crk. *EMBO J* 13:2341–2351.
 75. Jankowski W et al. (2012) Domain organization differences explain Bcr-Abl's preference for CrkL over CrkII. *Nat Chem Biol* 8:590–596.
 76. Dudkiewicz M, Lenart A, Pawłowski K (2013) A novel predicted calcium-regulated kinase family implicated in neurological disorders. *PLoS ONE* 8:e66427.
 77. Dudkiewicz M, Szczepińska T, Grynberg M, Pawłowski K (2012) A novel protein kinase-like domain in a selenoprotein, widespread in the tree of life. *PLoS ONE* 7:e32138.
 78. Nguyen KB et al. (2016) Phosphorylation of spore coat proteins by a family of atypical protein kinases. *Proc Natl Acad Sci USA* 113:E3482–E3491. Available at: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=27185916&retmode=ref&cmd=prlinks>.
 79. Wootton JC (1994) Non-globular domains in protein sequences: automated segmentation using complexity measures. *Computers & chemistry* 18:269–285. Available at: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=7952898&retmode=ref&cmd=prlinks>.

80. Coletta A et al. (2010) Low-complexity regions within protein sequences have position-dependent roles. *BMC Systems Biology* 2010 4:1 4:19. Available at: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=20385029&retmode=ref&cmd=prlinks>.
81. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. (1991) Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. 253:407–414. Available at: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=1862342&retmode=ref&cmd=prlinks>.
82. Cai W, Pei J, Grishin NV (2004) Reconstruction of ancestral protein sequences and its applications. *BMC Evol Biol* 4:33.
83. Birge RB, Kalodimos C, Inagaki F, Tanaka S (2009) Crk and CrkL adaptor proteins: networks for physiological and pathological signaling. *Cell Commun Signal* 7:13.
84. Fu H, Subramanian RR, Masters SC (2000) 14-3-3 proteins: structure, function, and regulation. *Annu Rev Pharmacol Toxicol* 40:617–647.
85. Brown MT et al. (1998) ASAP1, a phospholipid-dependent arf GTPase-activating protein that associates with and is phosphorylated by Src. *Mol Cell Biol* 18:7038–7051.
86. Chen P-W et al. (2016) The Arf GTPase-activating Protein, ASAP1, Binds Nonmuscle Myosin 2A to Control Remodeling of the Actomyosin Network. *Journal of Biological*

Chemistry 291:7517–7526.

87. Kramer OH, Stauber RH, Bug G, Hartkamp J, Knauer SK (2013) SIAH proteins: critical roles in leukemogenesis. *Leukemia* 27:792–802.
88. Farag AK, Roh EJ (2019) Death-associated protein kinase (DAPK) family modulators: Current and future therapeutic outcomes. *Med Res Rev* 39:349–385.
89. Lodygin D, Hermeking H (2005) The role of epigenetic inactivation of 14-3-3sigma in human cancer. *Cell Res* 15:237–246.
90. Safari F, Murata-Kamiya N, Saito Y, Hatakeyama M (2011) Mammalian Pragmin regulates Src family kinases via the Glu-Pro-Ile-Tyr-Ala (EPIYA) motif that is exploited by bacterial effectors. 108:14938–14943. Available at: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=21873224&retmode=ref&cmd=prlinks>.
91. Liu L et al. (2016) Homo- and Heterotypic Association Regulates Signaling by the SgK269/PEAK1 and SgK223 Pseudokinases. *Journal of Biological Chemistry* 291:21571–21583.
92. Sowa ME, Bennett EJ, Gygi SP, Harper JW (2009) Defining the human deubiquitinating enzyme interaction landscape. *Cell* 138:389–403.
93. Kiyokawa E, HASHIMOTO Y, Kurata T, Sugimura H, Matsuda M (1998) Evidence That DOCK180 Up-regulates Signals from the CrkII-p130 CasComplex. *Journal of Biological Chemistry* 273:24479–24484. Available at:

<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=9733740&retmode=ref&cmd=prlinks>.

94. Lamorte L, Royal I, Naujokas M, Park M (2002) Crk adapter proteins promote an epithelial-mesenchymal-like transition and are required for HGF-mediated cell spreading and breakdown of epithelial adherens junctions. *Mol Biol Cell* 13:1449–1461.
95. Ridley AJ et al. (2003) Cell migration: integrating signals from front to back. *Science* 302:1704–1709.
96. Rodrigues SP et al. (2005) CrkI and CrkII function as key signaling integrators for migration and invasion of cancer cells. *Mol Cancer Res* 3:183–194.
97. Petit V et al. (2000) Phosphorylation of tyrosine residues 31 and 118 on paxillin regulates cell migration through an association with CRK in NBT-II cells. *J Cell Biol* 148:957–970.
98. Tang DD, Zhang W, Gunst SJ (2005) The adapter protein CrkII regulates neuronal Wiskott-Aldrich syndrome protein, actin polymerization, and tension development during contractile stimulation of smooth muscle. *J Biol Chem* 280:23380–23389.
99. Nakashima N et al. (1999) The functional role of CrkII in actin cytoskeleton organization and mitogenesis. *J Biol Chem* 274:3001–3008.
100. Vallés AM, Beuvin M, Boyer B (2004) Activation of Rac1 by Paxillin-Crk-DOCK180 Signaling Complex Is Antagonized by Rap1 in Migrating NBT-II Cells. *Journal of*

Biological Chemistry 279:44490–44496.

101. Kobashigawa Y et al. (2007) Structural basis for the transforming activity of human cancer-related signaling adaptor protein CRK. *Nat Struct Mol Biol* 14:503–510.
102. Antoku S, Saksela K, Rivera GM, Mayer BJ (2008) A crucial role in cell spreading for the interaction of Abl PxxP motifs with Crk and Nck adaptors. *Journal of Cell Science* 121:3071–3082.
103. Jacquemet G et al. (2019) Filopodome Mapping Identifies p130Cas as a Mechanosensitive Regulator of Filopodia Stability. *Current Biology* 29:202–216.e7.
104. Huse M, Kuriyan J (2002) The conformational plasticity of protein kinases. *Cell* 109:275–282.
105. Jura N et al. (2011) Catalytic Control in the EGF Receptor and Its Connection to General Kinase Regulatory Mechanisms. *Molecular Cell* 42:981–986.
106. Shan Y et al. (2009) A conserved protonation-dependent switch controls drug binding in the Abl kinase. *Proc Natl Acad Sci USA* 106:139–144.
107. Newton K et al. (2014) Activity of protein kinase RIPK3 determines whether cells die by necroptosis or apoptosis. *Science* 343:1357–1360.
108. Harkiolaki M, Gilbert RJC, Jones EY, Feller SM (2006) The C-Terminal SH3 Domain of CRKL as a Dynamic Dimerization Module Transiently Exposing a Nuclear Export Signal. *Structure* 14:1741–1753.

109. Mandal P et al. (2014) RIP3 Induces Apoptosis Independent of Pronecrotic Kinase Activity. *Molecular Cell* 56:481–495. Available at:
<http://dx.doi.org/10.1016/j.molcel.2014.10.021>.
110. Kung JE, Jura N (2019) Prospects for pharmacological targeting of pseudokinases. *Nat Rev Drug Discov* 18:501–526.
111. Jamieson SA et al. (2018) Substrate binding allosterically relieves autoinhibition of the pseudokinase TRIB1. *Sci Signal* 11. Available at:
<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=30254053&retmode=ref&cmd=prlinks>.
112. Gao J et al. (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 6:p11. Available at:
<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=23550210&retmode=ref&cmd=prlinks>.
113. Cerami E et al. (2012) The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data: Figure 1. *Cancer Discov* 2:401–404. Available at:
<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22588877&retmode=ref&cmd=prlinks>.
114. Buchwald M et al. (2010) Ubiquitin conjugase UBCH8 targets active FMS-like tyrosine kinase 3 for proteasomal degradation. *Leukemia* 24:1412–1421.

115. Hartman AD et al. (2006) Constitutive c-jun N-terminal kinase activity in acute myeloid leukemia derives from Flt3 and affects survival and proliferation. *Exp Hematol* 34:1360–1376.
116. Linghu H et al. (2006) Involvement of adaptor protein Crk in malignant feature of human ovarian cancer cell line MCAS. *Oncogene* 25:3547–3556.
117. PROMALS3D: a tool for multiple protein sequence and structure alignments. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. 36:2295–2300. Available at:
<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=18287115&retmode=ref&cmd=prlinks>.
118. Estimating maximum likelihood phylogenies with PhyML. (2009) Estimating maximum likelihood phylogenies with PhyML. 537:113–137. Available at:
<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=19378142&retmode=ref&cmd=prlinks>.
119. CD-HIT Suite: a web server for clustering and comparing biological sequences. (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. 26:680–682. Available at:
<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=20053844&retmode=ref&cmd=prlinks>.
120. WebLogo: a sequence logo generator. (2004) WebLogo: a sequence logo generator. 14:1188–1190. Available at:

- <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=15173120&retmode=ref&cmd=prlinks>.
121. Global landscape of HIV-human protein complexes. (2011) Global landscape of HIV-human protein complexes. 481:365–370. Available at:
<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22190034&retmode=ref&cmd=prlinks>.
122. Purification and characterization of HIV-human protein complexes. (2011) Purification and characterization of HIV-human protein complexes. 53:13–19. Available at:
<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=20708689&retmode=ref&cmd=prlinks>.
123. Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. (1999) Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. 71:2871–2882. Available at:
<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=10424174&retmode=ref&cmd=prlinks>.
124. Use of Ranks in One-Criterion Variance Analysis (1952) Use of Ranks in One-Criterion Variance Analysis. 47:583–621. Available at:
<https://www.jstor.org/stable/pdf/2280779.pdf?refreqid=excelsior%3A6a2545ce15a6d980eb4043a2bdfbfdce>.

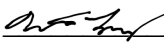
125. Multiple Comparisons among Means (1961) Multiple Comparisons among Means. 56:52–64. Available at:
<http://www.tandfonline.com/doi/abs/10.1080/01621459.1961.10482090>.
126. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. 25:1189–1191. Available at:
<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=19151095&retmode=ref&cmd=prlinks>.
127. UCSF Chimera--a visualization system for exploratory research and analysis. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. 25:1605–1612. Available at:
<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=15264254&retmode=ref&cmd=prlinks>.
128. Amino acid substitution matrices from protein blocks. (1992) Amino acid substitution matrices from protein blocks. 89:10915–10919. Available at:
<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=1438297&retmode=ref&cmd=prlinks>.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

DocuSigned by:

BB5E91D3CD2D4EC...

Author Signature

11/9/2019

Date