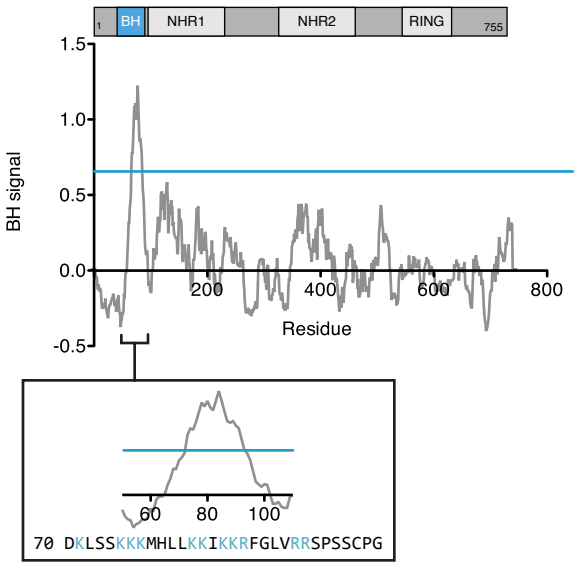
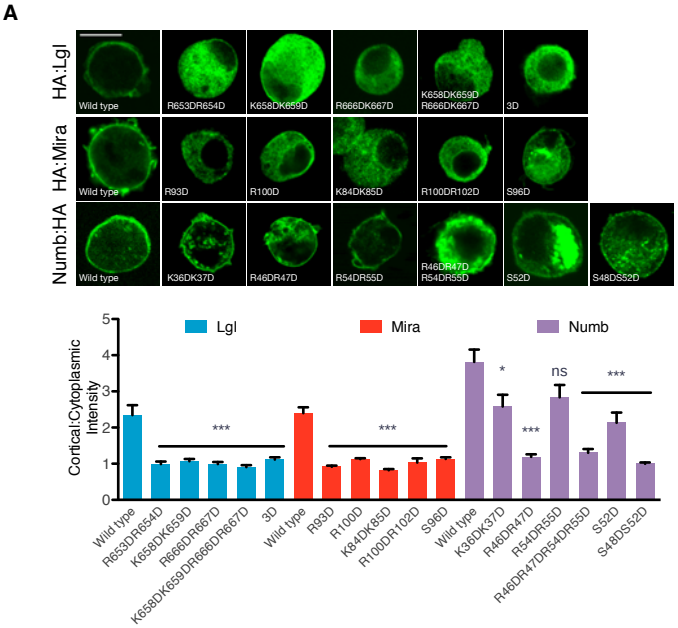
Supplemental Information Inventory

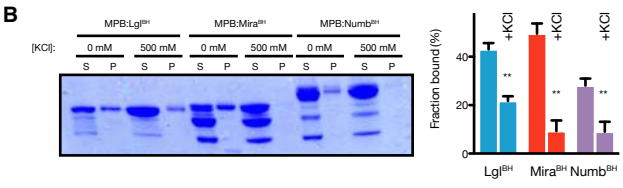
- Supplemental figures
- Supplemental figure legends
- Extended experimental procedures
- PRBH identification program code
- Supplemental References
- Supplemental Data Table 1 (“Supplemental Data Table 1.xlsx”)

Supplemental Figure 1

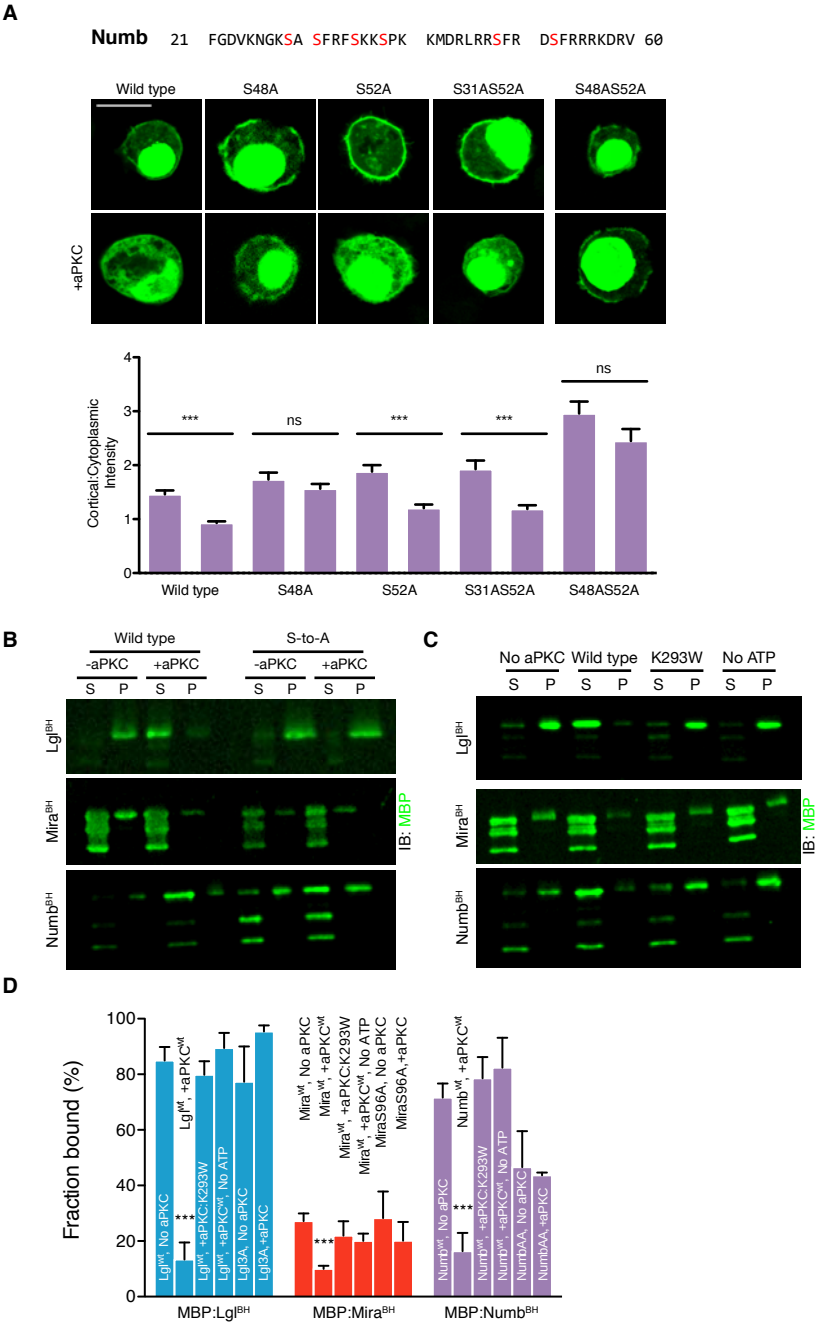


Supplemental Figure 2

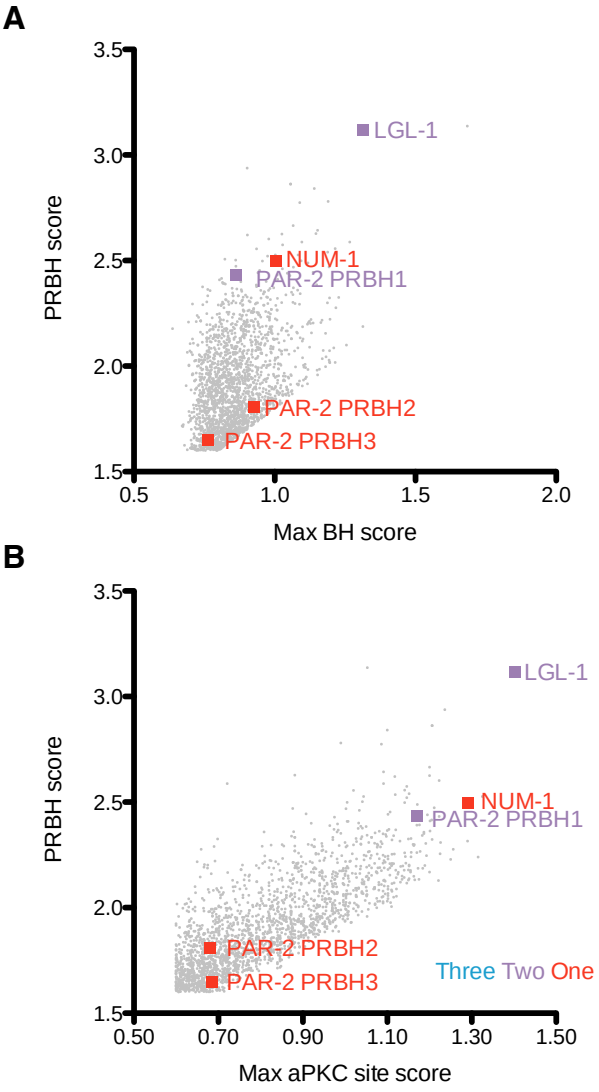




Supplemental Figure 3



Supplemental Figure 4



Supplemental Figure Legends

Supplemental Figure 1. Basic-Hydrophobic motif signal of Neur. Related to Figure 2.

The BH signal of Neur is plotted aligned to its domain architecture. The inset is a detail of Neur BH with basic residues colored blue. S residues within this region do not fit the aPKC consensus sequence which features a R in the -2 position and a hydrophobic residue in the +1 position, which makes it unlikely to be an aPKC-regulated PRBH motif. (Please see lower for further discussion of this). Abbreviations as follows: NHR, Neur homology region; RING, C_3HC_4 ring zinc finger.

Supplemental Figure 2. Charge mediates cortical targeting and phospholipid binding for Lgl, Mira, and Numb. Related to Figure 3.

A. Mutations to acidic residues reduce cortical localization of full length Lgl, Mira, and Numb. This Figure is related to Figure 3D but includes a larger set of point mutants. The localization of each full-length protein with point mutations was characterized by immunostaining transiently transfected S2 cells for the HA epitope tag. Localization was quantified as a cortical to cytoplasmic signal intensity ratio for at least 16 cells (mean ± SEM). Statistical testing was performed with t-test to compare the localization of each mutant to the wild type full-length control. A P-value of <0.05 and <0.0001 are marked with one and three asterisks, respectively. No statically significant difference is marked with ns. Scale bar, 10 μm.

B. High ionic strength disrupts BH phospholipid interactions. A Coomassie-stained gel from lipid vesicle-binding cosedimentation assays of vesicles with 4:1 PC:PS plus 10% PIP_2 is shown. Each BH motif was characterized as an MPB-fusion protein. The fraction of protein bound, quantified as the amount of protein in the pellet over total protein, was quantified in triplicate for each vesicle composition. The mean and SEM of the fraction bound are shown. Asterisks indicate $p < 0.01$. Significance was evaluated using a non-parametric t-test relative to low ionic strength.

Supplemental Figure 3. aPKC phosphorylation inhibits BH motif lipid binding for Lgl, Mira, and Numb. Related to Figure 4.

A. Identification of Numb BH residues involved in regulating aPKC-induced cortical

displacement. This figure is related to Figure 4B", but includes a larger collection of Numb S-to-A mutations. The sequence of Numb BH is displayed with potential phosphorylation sites highlighted in red. Representative images from transiently transfected S2 cells in the absence or presence of aPKC are displayed. Cells were transfected with EGFP-tagged BH motifs. aPKC expression was verified by immunostaining for PKCζ. The cortical to cytoplasmic intensity ratio was quantified and displayed as the mean ± SEM, and statistical significance was determined by a non-parametric t-test to compare each mutant to its singly transfected control. P values of <0.0001 are marked with three asterisks. No statistical difference is marked by ns. Scale bar, 10 μm.

B-D. aPKC kinase activity inhibits BH motif phospholipid binding. Images of representative immunoblots to characterize the effect of aPKC on Par substrate BH motif binding to PC/PS/PIP_2 vesicles are presented. Lipid-binding cosedimentation assays were performed in the presence of aPKC and ATP. For C, sample conditions varied as follows: No aPKC, ATP only; wild type, $aPKC^{259-606}$ and ATP; K293W, aPKC $K293W^{259-606}$ and ATP; No ATP, $aPKC^{259-606}$ but no ATP. All samples contained 4:1 PC:PS plus 10% PIP_2 and all BH motifs had the wild type sequence. Samples were analyzed by immunoblotting for the MBP tag. S marks the supernatant and P marks the pellet. The "S-to-A" mutations are: Lgl3A, MiraS96A and NumbS48AS52A. Quantification of aPKC's effect on BH motif vesicle binding displayed in A-B. The fraction of protein bound, quantified as the amount of protein in the pellet over total protein, was quantified in triplicate for each vesicle composition. The mean and SEM of the fraction bound are shown. Asterisks indicate $p < 0.0001$. Significance was evaluated using a non-parametric t-test relative to binding in the absence of aPKC.

Supplemental Figure 4. Distribution of PRBH scores in the *C. elegans* proteome. Related to Figure 5.

A-B. PRBH scores, maximum BH score, and the highest aPKC site score are displayed from *C. elegans* proteome. Each point marks a single putative PRBH motif and all identified PRBH motifs are displayed. Blue lines mark the PRBH threshold values. Sequences with scores less than these values are not candidate PRBH motifs. See also Supplemental Data Table 3.

Extended Experimental Procedures

Sequence IDs and Residue Numbers

The following sequences were characterized in this study:

Lethal giant larvae: Species, *Drosophila melanogaster;* Symbol, Lethal (2) giant larvae; FlyBase ID, FBgn0002121; Isoform, PA (length 1161 residues). Lgl BH motif constructs include residues 647-681. Lgl ΔBH motif constructs deletes residues 647-681.

Miranda: Species, *Drosophila melanogaster;* Symbol, Mira; FlyBase ID, FBgn0021776; Isoform, PA (length 829 residues). Mira BH constructs include residues 71-110 and the BH motif deletion construct removes residues 72-110.

Numb: Species, *Drosophila melanogaster;* FlyBase ID, FBgn0002973; Isoform, PA (length 556 residues). Numb BH constructs include residues 15-86 and the BH motif deletion construct removes residues 24-56.

Neuralized: Species, *Drosophila melanogaster;* Symbol, Neur; Flybase ID, FBgn0002932; Isoform, PA (length 754 residues); BH motif, residues 68-88.

Amer1: Species, *Homo sapiens;* NCBI reference sequence, NP_689637.3; PRBH, residues 154-199.

Casein Kinase I gamma 2: Species, *Homo sapiens;* Symbol, CKIγ-2; NCBI reference sequence, GenBank ID, AAB88627.1. PRBH motif, 369-415.

MAGUK p55 subfamily member 7: Species, *Homo sapiens;* Symbol, MPP7; NCBI Reference Sequence, NP_775767.2. PRBH, residues 289-383.

PIP82: Species, *Drosophila melanogaster*; FlyBase ID, FGgn0024943. PRBH motif, residues 400-450.

Molecular cloning

All molecular cloning was performed as previously described (Graybill et al., 2012). MBP-fusion proteins were made using pMal C2 vector (New England BioLabs). The pMT v5-HisA with a HA-tag or pTub (Lu and Prehoda, 2013) promoter was used for S2 cell transient expression. All proteins were tagged with EGFP, MBP, or the HA epitope tag at the NH2-terminus except for full-length Numb, which was COOH-terminally tagged. The Supplemental Experimental Procedures includes sequence IDs and residue numbers. The fragments of human MPP7 and fly PIP82 were cloned from synthetic gBlocks (IDT).

Cell culture and localization assays

Cell culture was performed as previously described (Lu and Prehoda, 2013). S2 cells were grown at 30°C in Schneider's media (Sigma-Aldrich) with 10% fetal bovine serum. S2 cells were transiently transfected with Effectene (QIAGEN) according to the manufacturer's protocol. Cells were seeded at ~2×10^6 cells/well in a 6-well plate and transfected 24 hours later with 0.5 μg plasmid. 24 hours after transfection, protein expression was induced for all pMT-transfected cells by addition of 0.5 mM $CuSO_4$. 24 hours post-induction, or 48 hours post-transfection for cells transfected with pTub plasmids, cells were plated on 12 mm glass coverslips in a 24-well plate for immunostaining and imaging.

For immunohistochemistry, cells were washed with PBS, and then fixed with 4% paraformaldehyde in PBS. Cells were washed with PBS with 0.1% saponin, washed three times with block [PBS, 0.1% saponin, and 1% bovine serum albumin (BSA)], and then blocked for 30 minutes with block. Primary antibody incubations were performed for either 1-2 hours at room temperature or overnight at 4°C. Cells were washed three times after immunostaining and between antibody staining steps. The following antibody dilutions were used: mouse anti-HA 1:1000 (Covance), rabbit anti-PKCζ 1:1000 (Santa Cruz Biotechnology), DyLight 488 Donkey anti-mouse 1:500 (Jackson Immunoresearch) and DyLight Donkey 649 anti-rabbit (Jackson Immunoresearch). The cell cortex was stained with Alexa Fluor 555-Phalloidin (1:500, Invitrogen). Coverslips were mounted using Vectashield Hardset Mounting Medium (Vector Laboratories). An Olympus Fluoview FV1000 BX61 with a PlanApo N 60×/1.42 oil and a Leica SP2 confocal microscope with a 63×/1.40-0.60 oil CS objective was used to acquire all images. The cortical-to-cytoplasmic intensities were quantified in ImageJ using the free-hand line tool and freehand selection tool to determine the average fluorescence intensities in each cellular area. Prism was used for statistical analysis of quantified localization data. Images were processed with ImageJ and the Adobe Suite.

Protein expression and purification

MBP-fusion proteins were expressed and purified from BL21 (DE3) cells, as previously described (Graybill et al., 2012). Cells were lysed by sonication, pelleted by centrifugation at 15,000 rpm, 30 min, 4°C in a JA-20 rotor and the soluble fraction was bound to amylose resin (New England BioLabs) for 45 min, 4°C. The resin was washed

with MBP lysis buffer (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 1 mM EDTA, 1 mM DTT), then eluted with MBP elution buffer (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 1 mM EDTA, 1 mM DTT, 5 mM maltose). Elutions were pooled for dialysis in storage buffer (20 mM HEPES, pH 7.5, 50 mM NaCl, 1 mM DTT) overnight at 4°C and concentrated with a 10 kDa molecular weight cut-off Vivaspin 20 concentrator (GE Healthcare). Purified proteins were frozen with liquid nitrogen and stored at -80°C.

All His_6:aPKC$^{259-606}$ (kinase domain) constructs were expressed in HEK293F cells by transient transfection using 293fectin transfection reagent (Life Technology). Transfections were performed according to manufacturer's protocols. Expressions and protein purification of aPKC was performed as previously described (Graybill et al., 2012). Briefly, cells were cultured at 37°C, 8% CO_2, 125 rpm in Freestyle HEK293-F media (Invitrogen). Cells were cotransfected with pCMV His_6-aPKC$^{259-606}$ (wild type or K293W mutant) and pCMV GST:PDK1$^{161-500}$ (kinase domain) at ~1.0×10^6 cells/mL. Cells were harvested 24 hours post-transfection by centrifugation, resuspended in Ni^{2+} lysis buffer (50 mM NaH_3PO_4, 300 mM NaCl, 10 mM imidazole, pH adjusted to 8.0 with NaOH), then stored at -80°C. Cells were lysed by sonication, then lysates were pelleted by centrifugation at 15,000 rps, 30 min, 4°C in a JA-20 rotor. The supernatant was incubated with ammonium sulfate at a concentration of 45% (w/v) for 30 min, 4°C. Then, the solution was centrifuged at 15,000 rpm, 30 min, 4°C in a JA-20 rotor, the pellet was collected and resuspended in Ni^{2+} lysis buffer, then it was incubated with Ni^{2+}-nitrilotriacetic acid resin (Qiagen) for 45 min, 4°C. Resin were washed with Ni^{2+} lysis buffer, then protein was eluted with Ni^{2+} elution buffer (50 mM NaH_3PO_4, 300 mM NaCl, 250 mM imidazole, pH adjusted to 8.0 with NaOH). The eluted sample was dialyzed for 4 hours in storage buffer (20 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM DTT) using the SnakeSkin dialysis membrane (Thermo Scientific). Dialyzed proteins were clarified by centrifugation at 3,000 rpm, 30 min, 4°C. Clarified samples were concentrated using a 10 kDa molecular weight cut-off Vivaspin 20 concentrator (GE Healthcare). Purified proteins were frozen with liquid nitrogen and stored at -80°C.

Lipid binding cosedimentation assays

Lipids from the following sources were purchased from Avanti Polar Lipids: 1,2-dioleoyl-*sn*-glycero-3-phosphocholine, 1,2-Dioleoyl-*sn*-Glycero-3-[Phospho-L-Serine], L- (Egg,

Chicken), L-α-Phosphatidylinositol-4-phosphate (Brain, Porcine), L-a-α-Phosphatidic acid (Egg, chicken), Phosphatidlyinositol-4,5-bisphosphate (Brain, Porcine), and 1,2-dioleoyl-*sn*-glycero-3-[phosphoinositol-3,4,5-trisphosphate]. Giant unilamellar vesicles were prepared as described (Winters et al., 2005). Briefly, lipids were added to chloroform at the molar ratios specified, chloroform was removed by rotovap and placing under vacuum for 15 min. Giant unilamellar vesicles were formed by resuspending in 0.2 M sucrose at 0.5 mg/mL total lipid by incubation at 50°C for 4 hours or overnight. Vesicles were stored at 4°C and used within 3 days of generation.

Pelleting assays were performed as previously described (Prehoda et al., 2000). For pelleting assay without aPKC, proteins were pre-spun to clarify in TLA-100 rotor, 65,000 rpm, 20 min, 4°C. All lipid binding cosedimentation assays contained 1-6 μM MBP-tagged protein and 0.23 mg/mL total lipid in assay buffer (20 mM HEPES, pH 7.5, 50 mM NaCl, 1 mM DTT). After incubating proteins and vesicles at 30°C for 30 min, samples were pelleted by ultracentrifugation in a TLA-100 rotor at 65,000 rpm, 60 min, 4°C. The supernatants were removed, the pellets were washed, then the pellets were resuspended in a volume equal to the reaction volume. All samples were run on a 12.5% SDS-PAGE gel and stained with Coomassie Brilliant Blue for analysis. Gels were imaged using a scanner or the LiCOR. The band intensities were quantified using LiCOR ImageStudio and ImageJ. Gel images were processed with ImageJ and the Adobe Suite.

All pelleting assays containing aPKC were performed essentially as described above with the changes noted below. Reactions contained 0.1-0.5 μM MBP-tagged proteins and 0.034 μM aPKC in reaction buffer (500 mM Tris-HCl, pH 7.5, 100 mM NaCl, 100 μM ATP, 100 μM $MgCl_2$). Immunoblotting was performed to detect protein using mouse anti-MBP (1:1000, Santa Cruz Biotechnologies) and IRDye 800 CW goat anti-mouse (1:10,000, LiCOR). Membranes were blocked with TBS-T with 3% milk. Membranes were imaged with using a LiCOR and images were processed with the Adobe Suite. Band intensities were quantified with LiCOR ImageStudio and band intensity data was analyzed using Prism.

Protein kinase assays

Kinase assays were performed as previously described (Atwood and Prehoda, 2009; Graybill et al., 2012). Briefly, aPKC and substrates were incubated in assay buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$) at 30°C for 5 min before addition of 1 mM ATP doped with [γ-^{32}P]ATP (~1.0 X 10^5/nmol of ATP). The kinase reaction proceeded for 30 minutes before it was quenched by SDS loading dye. Samples were run on a 12.5% acrylamide SDS-PAGE gel, were dried, and exposed to a phosphor screen (Molecular Dynamics). The phosphor screen was imaged with a Storm 860. Images from the Storm 860 Molecular Imager were processed with ImageJ and the Adobe Suite.

Computational identification of PRBH motifs

The PRBH algorithm was implemented using the Python programming language with standard Python libraries including numpy and matplotlib. EMBL-EBI reference proteome files from *Homo sapiens, D. melanogaster, C. elegans,* and *S. cerevisiae*. The BH scoring algorithm used the previously described scores for BH character (Brzeska et al., 2010): A, -0.17; C, 0.24; D, -1.23; E, -2.02; F, 1.13; G, -0.01; H, -0.17; I, 0.31; K, 2; L, 0.56; M, 0.23; N, -0.42; P, -0.45; Q, -0.58; R, 2; S, -0.13; T, -0.14; V, -0.07; W, 1.85; Y, 0.94. Each amino acid residue was assigned a BH score for over a 19 residue BH window (i.e. the amino acid and its 9 NH_2- and COOH-terminal neighbors; unless specified otherwise). Sequences of adjacent residues with BH scores over a BH threshold of 0.6 were identified as peaks. The amino acid sequence of the peak, as well as, the maximum BH score and area over the BH threshold were compiled in a peak list. For human Numb and CKIγ-2, the BH window was reduced to 15 residues such that BH motifs at the extreme NH_2- and COOH-terminus were identified. BH scores for residues 9 residues or less from the NH_2- or COOH-terminus cannot be scored. Reducing the BH window allows residues closer to the NH_2- or COOH-terminus to be scored.

aPKC consensus sequences were identified by scoring the residues flanking S/T residues. The positional scoring was based on the previously described aPKC consensus sequence (Wang et al., 2012). This scoring metric does not strictly follow the consensus sequence from Wang and coworkers such that Ser 96, a verified aPKC

phosphosites on Mira, fit the consensus. Hydrophobic residues were score as follows, except where noted otherwise: A, 0.05; F, 0.2; I, 0.2; L, 0.2; M, 0.25; V, 0.1; W, 0.1; Y, 0.1. Basic residues were score as follows, except where noted otherwise: R, 0.1; K, 0.07; H, 0.04; D, -0.06; E, -0.06. Residues in the +2 position were scored as follows: R, 0.35; K, 0.15; H, 0.05. Residues in the +3 position were scored as follows: F, 0.25; L, 0.25; M, 0.25; I, 0.15; V, 0.1; W, 0.1; Y, 0.1; A, 0.05; R, 0.1; K, 0.1; H, 0.05. The +1 position was scored using the hydrophobic residue metric, and a penalty of -0.4 was given for residues that were not hydrophobic. Points were awarded for basic or hydrophobic residues from the -9 to the -4 position. Sequences with aPKC scores above the aPKC threshold of 0.6 were compiled in an aPKC site list. For proteins with multiple aPKC sites, each was included in the aPKC site list with its corresponding site score.

The aPKC site list and the BH site list were compared, and the PRBH algorithm analyzed sequences occurring in both lists. A PRBH score was computed for sequences occurring in both lists by using the following calculation: ‘Maximum BH score’ + ‘Maximum aPKC site score’ + 0.2 * ‘aPKC site number in the PRBH motif’. Three additional tests were performed on PRBH sequences: first, the area of the BH motif must be less than 400X the number of aPKC phosphosites. This eliminated many highly basic sequences with few candidate sites. Such sequences were not desired, as phosphorylation is not likely to alter BH character dramatically. Second, when a BH window size of 19 was used, BH motifs with areas less than 90 were discarded. This parameter was not used for identification of PRBH motifs at a protein’s NH_2- or COOH-terminus because the BH scoring metric reduces the BH signal for these peaks due to the scoring process. Third, PRBH scores less than the PRBH threshold of 1.6 were discarded. Sequences that passed these criteria were compiled in a PRBH sequence list that included the following: protein names, PRBH score, the PRBH sequence, maximum BH score, BH area, aPKC site scores, and aPKC site sequences.

The PRBH Motif Identification Program

```
#!/usr/bin/env python

class fasta:
    def __init__(self,desc,seq):
        self.desc = desc
        self.seq = seq

def fasta_file(it):
    """
    generator that returns an iterator over the records in a FASTA format file
    """
    d = fasta('','')
    for i in it:
        if i[0] == '>':
            if len(d.seq) > 0:
                yield d
                d = fasta('','')
            d.desc = i[1:].strip()
        else:
            d.seq += i.strip()
    if len(d.seq) > 0:
        yield d

hydRes = {'A':0.05,'F':0.2,'I':0.2,'L':0.2,'M':0.25,'V':0.1,'W':0.1,'Y':0.1}
basicRes = {'R':0.1,'K':0.07,'H':0.04,'D':-0.06,'E':-0.06,'Y':0.0}
M3Res =
{'F':0.25,'L':0.25,'M':0.25,'I':0.15,'V':0.1,'W':0.1,'Y':0.1,'A':0.05,'R':0.1,'
K':0.1,'H':0.05}
M2Res = {'R':0.35,'K':0.15,'H':0.05}

class prbh_protein:
    def __init__(self,fasta,prbhs):
        self.fasta, self.prbhs = (fasta,prbhs)

class prbh_motif:
    def __init__(self,bh,apkcs,score):
        self.bh, self.apkcs, self.score = (bh,apkcs,score)
    def __eq__(self,other):
        return self.bh == other.bh and (self.score - other.score) < 0.00001 and
self.apkcs[0] == other.apkcs[0]

class bh_motif:
    def __init__(self,start,end,height,area):
        self.start, self.end, self.height, self.area = (start,end,height,area)
    def __eq__(self,other):
        return self.start == other.start and (self.height - other.height) <
0.00001 and (self.area - other.area) < 0.00001

class aPKC_site:
    def __init__(self,pos,score,seq):
        self.pos, self.score, self.seq = (pos,score,seq)
    def __eq__(self,other):
        return self.pos == other.pos and (self.score - other.score) < 0.00001
and self.seq == other.seq

def max_seq(seq,score_dict):
    hyd = 0.0
    for i in seq:
        hyd = max(hyd,score_dict.get(i,0.0))
    return hyd

def calc_aPKC_sites(seq, threshold=0.6):
    sites = []
    for pos,char in enumerate(str(seq).upper()):
        if char == 'S' or char == 'T':
            score = 0.0
            # pos + 1 hydrophobic
            if pos+1 < len(seq):
                score += hydRes.get(seq[pos+1],-0.4) # hydScore or -.4 if not
hydrophobic
            # pos +2 - +6 basic
```

```
            basicTemp = 0.0
            for i in range(2,7):
                if pos+i+1 < len(seq):
                    basicTemp += basicRes.get(seq[pos+i],0.0)
            score += basicTemp * 0.8
            # pos -2 usually basic - very often arginine. -0.2 penality if not
basic.
            if pos-2 > 0:
                score += M2Res.get(seq[pos-2],0.)*1.4
            # pos -3: highest scoring hydrophobic, or basic residue with a
lower score. No penality for others.
            if pos-3 > 0:
                score += M3Res.get(seq[pos-3],0.)
            # pos -4 to -9: highest scoring basic and highest score
hydrophobic.
            hyd = max_seq(seq[max(0,pos-9):max(0,pos-4)],hydRes)
            bas = max_seq(seq[max(0,pos-9):max(0,pos-4)],basicRes)
            score += (hyd + bas) * 0.5
            if score > threshold:
                sites.append(aPKC_site(pos,score,seq[max(0,pos-
7):min(pos+17,len(seq))]))
    return sites

bhDict = {'O':-.17, 'U':0,'X':0,'A': -.17, 'C':0.24, 'D':-1.23, 'E': -2.02,
'F':1.13, 'G':-.01, 'H':-.17, 'I':.31, 'K':2,'L':.56,'M':.23,'N':-.42,'P':-
.45,'Q':-.58,'R':2, 'S':-.13,'T':-.14,'V':-.07,'W':1.85,'Y':.94}

def running_bh_sum(seq, window_size):
    pos = 0
    if len(seq) < window_size:
        return
    running_sum = sum([bhDict.get(x,0) for x in seq[:window_size]])
    yield running_sum
    for c in seq[window_size:]:
        running_sum += bhDict.get(c,0) - bhDict.get(seq[pos],0)
        pos += 1
        yield running_sum

def calc_bh(seq, window_size=19, threshold=0.6):
    motifs = []
    height = 0.0
    area = 0.0
    startRes = -1

    for pos,bsum in enumerate(running_bh_sum(seq,window_size)):
        bh_score = bsum/window_size
        if bh_score > threshold:
            if startRes == -1: #new BH
                startRes = pos
            height = max(bh_score,height)
            area += bsum
        elif startRes != -1: #close out BH
                motifs.append(bh_motif(startRes,pos+window_size,height,area))
                startRes = -1
                height = 0.0
                area = 0.0
    return motifs

def calc_bh_score(seq,window_size=19):
    scores = []
    for pos,bsum in enumerate(running_bh_sum(seq,window_size)):
        bh_score = bsum/window_size
        scores.append(bh_score)
    return scores

def calc_prbh(seq,threshold=1.6):
    prbhs = []
    bhs = []
    sites = calc_aPKC_sites(seq)
    if sites:
        bhs = calc_bh(seq)
    for bh in bhs:
```

```
        phos = []
        for apkc in sites:
            if apkc.pos > bh.start and apkc.pos < bh.end:
                phos.append(apkc)
        if len(phos):
            score = bh.height + max([x.score for x in phos]) + len(phos)*0.2
            if bh.area < len(phos) * 400.:
                if bh.area/90. > 1:
                    if score > threshold:
                        prbhs.append(prbh_motif(bh,phos,score))
    return prbhs

class tagged_prbh(object):
    def __init__(self,prbh,fasta):
        self.prbh, self.fasta = (prbh,fasta)
    def __eq__(self,other):
        return self.prbh.score-other.prbh.score < 0.00001 and
self.prbh.apkcs[0].seq == other.prbh.apkcs[0].seq
    def __hash__(self):
        return hash((self.prbh.score, self.prbh.apkcs[0].seq))

def fasta_prbh(fasta):
    import gzip
    if fasta[-2:] == 'gz':
        fi = gzip.GzipFile(fasta,'r')
    else:
        fi = open(fasta,'r')

    hits = []
    with fi as f:
        for rec in fasta_file(f):
            prbhs = calc_prbh(rec.seq)
            for i in prbhs:
                hits.append(tagged_prbh(i,rec))
    return hits

def bh_motifs_string(seq,motifs):
    hit_string = list(' '*len(seq))
    motif_string = ""
    for a in motifs:
        width = a.end - a.start
        motif_string += ' '*(a.start - len(motif_string))
        motif_string += "{:*^width}".format("{:.1f}".format(a.height))
    return motif_string

def apkc_sites_strings(seq,sites):
    hit_string = list(' '*len(seq))
    score_string = list(' '*len(seq))
    for a in sites:
        hit_string[a.pos] = '|'
    hit_string = "".join(hit_string)
    for a in sites:
        s = "{:.1f}".format(a.score)
        score_string = score_string[:a.pos-1]+list(s)+list(' '*(len(seq)-
a.pos+1-len(s))) #score_string[a.pos-1+len(s):-len(s)+2]
    score_string = "".join(score_string)
    return hit_string,score_string

human = fasta_prbh('/Path_to_input_file/9606.fasta.gz')
fly = fasta_prbh('/Path_to_input_file /7227.fasta.gz')
worm = fasta_prbh('/Path_to_input_file /6239.fasta.gz')
sponge = fasta_prbh('/Path_to_input_file
/Amphimedon_queenslandica.Aqu1.26.pep.all.fa')
with_name = ((human,"human"),(fly,"fly"),(worm,"worm"),(sponge,"sponge"))



def max_site(apkcs):
    maxs = 0.0
    for i in apkcs:
        maxs = max(maxs,i.score)
    return maxs
```

```
human_s = sorted(human,key=lambda rec: rec.prbh.score, reverse=True)
fly_s = sorted(fly,key=lambda rec: rec.prbh.score, reverse=True)
worm_s = sorted(worm,key=lambda rec: rec.prbh.score, reverse=True)
sponge_s = sorted(sponge,key=lambda rec: rec.prbh.score, reverse=True)


sorted_prbh = (human_s,fly_s,worm_s,sponge _s)
with open('/File_for_results.txt', 'w+') as f1:
    for k in sorted_prbh:
        for i in k:
            s = ''
            for j in i.prbh.apkcs:
                s += "{}*{}*".format(j.score,j.seq)
            print
>>f1,'*',i.fasta.desc.split('|')[2],'*',i.prbh.score,'*',i.prbh.bh.area,'*',i.p
rbh.bh.height,'*',i.fasta.seq[i.prbh.bh.start:i.prbh.bh.end],'*',len(i.prbh.apk
cs),'*',s
```

Supplemental References

Atwood, S.X., and Prehoda, K.E. (2009). aPKC phosphorylates Miranda to polarize fate determinants during neuroblast asymmetric cell division. Curr. Biol. CB *19*, 723–729.

Brzeska, H., Guag, J., Remmert, K., Chacko, S., and Korn, E.D. (2010). An experimentally based computer search identifies unstructured membrane-binding sites in proteins: application to class I myosins, PAKS, and CARMIL. J. Biol. Chem. *285*, 5738–5747.

Graybill, C., Wee, B., Atwood, S.X., and Prehoda, K.E. (2012). Partitioning-defective Protein 6 (Par-6) Activates Atypical Protein Kinase C (aPKC) by Pseudosubstrate Displacement. J. Biol. Chem. *287*, 21003–21011.

Lu, M.S., and Prehoda, K.E. (2013). A NudE/14-3-3 pathway coordinates dynein and the kinesin Khc73 to position the mitotic spindle. Dev. Cell *26*, 369–380.

Prehoda, K.E., Scott, J.A., Mullins, R.D., and Lim, W.A. (2000). Integration of multiple signals through cooperative regulation of the N-WASP-Arp2/3 complex. Science *290*, 801–806.

Wang, C., Shang, Y., Yu, J., and Zhang, M. (2012). Substrate recognition mechanism of atypical protein kinase Cs revealed by the structure of PKCι in complex with a substrate peptide from Par-3. Struct. Lond. Engl. 1993 *20*, 791–801.

Winters, M.J., Lamson, R.E., Nakanishi, H., Neiman, A.M., and Pryciak, P.M. (2005). A membrane binding domain in the ste5 scaffold synergizes with gbetagamma binding to control localization and signaling in pheromone response. Mol. Cell *20*, 21–32.

Supplemental Data Table 1 (filename: "Supplemental Data Table 1.xlsx"). PRBH scores and descriptions. Related to Figure 5 and Supplemental Figure 4.

Supplemental Data Table 1 lists the PRBH motifs identified in the proteomes of human, fly (1A), worm (1B), and sponge (1C), respectively (1D). The PRBH score, maximum BH score, aPKC site scores, aPKC site sequences, and PRBH sequence also are given for each PRBH motif.