

Supporting Information

Ha and Boggon 10.1073/pnas.1717437115

SI Materials and Methods

Expression and Purification. PAK4 isoform-2 cDNA (Open Biosystems) encoding 426 aa (UniProt ID: O96013) (residues 1–426, termed PAK4^{FL}) was subcloned into a modified pET28 vector (Novagen) with an N-terminal hexa-histidine (6xHis) tag cleavable by tobacco etch virus (TEV) protease. Catalytic domain PAK4 residues 109–426 (termed PAK4^{cat}) was subcloned into the same expression vector. Because residues 121–426 of isoform-2 correspond exactly to residues 286–591 in isoform-1, for clarity we discuss kinase domain residues using isoform-1 numbering throughout. We also use isoform-1 numbering for residues in the N terminus (residues 1–119 are identical between isoform-1 and isoform-2). PAK4 residues 1–45 (termed PAK4^{N45}) were each subcloned into pMCSG9 vector (SnapGene) which has an N-terminal 6xHis tag followed by maltose binding protein cleavable by TEV protease. Human CDC42 (UniProt ID: P60953) cDNA (a gift of Joseph Schlessinger, Yale University, New Haven, CT) encoding either residues 1–177 or 1–181 was subcloned into pET22b vector (Novagen) with a C-terminal uncleavable 6xHis tag and Q61L mutation. Human RAC1 (UniProt ID: P63000) cDNA (a gift of Ruth Halaban, Yale University, New Haven, CT) encoding residues 1–181 was subcloned into pET22b vector having Q61L mutation with a C-terminal uncleavable 6xHis tag. Mutants were introduced using QuikChange (Agilent Technologies).

PAK4^{FL}, PAK4^{cat}, and PAK4^{N45} proteins were expressed in BL21-CodonPlus(DE3)RILP (Agilent Technologies) cells by induction with 0.5 mM isopropyl β -D-1-thiogalactopyranoside overnight at 18 °C. Harvested pellets were suspended in lysis buffer [20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM Tris(2-carboxyethyl)phosphine (TCEP), and 0.1 mM PMSF] and lysed by sonication. The supernatants were affinity-purified by HisTrap chelating column (GE) and then resolved over Resource S (GE) for (PAK4^{FL} or PAK4^{N45}) or Resource Q (GE) for (PAK4^{cat}, CDC42, or RAC1). Each protein was then loaded to a Superdex 75 10/300GL (GE) column and all elutions indicated monodisperse proteins.

For SAXS, samples of size-exclusion-purified PAK4^{FL} and PAK4^{FL}-CDC42 dialyzed against the same buffer (20 mM Tris, pH 8, 300 mM NaCl, and 1 mM TCEP). For ITC, Q61L mutants of CDC42 and RAC1 proteins were preincubated with 5 mM GTP and 10 mM MgCl₂ ahead of dialysis. All recombinant proteins were simultaneously dialyzed against the same buffer (PBS, pH 7.4). For kinase assays, following cell lysis proteins were purified by HisTrap chelating column (GE) and dialyzed against 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 1 mM DTT.

SAXS. We prepared samples of size-exclusion-purified PAK4^{FL} and PAK4^{FL}-CDC42 dialyzed against the same buffer (20 mM Tris, pH 8.0, 300 mM NaCl, and 1 mM TCEP). Final concentrations of 11, 5.6, and 2.8 mg/mL were prepared for PAK4^{FL} and 6.2, 3.1, and 1.5 mg/mL for PAK4^{FL}-CDC42. X-ray scattering was conducted at the LiX beamline at the National Synchrotron Light Source II (NSLS-II) and data were collected on a Pilatus 1M detector. Five individual 5-s exposures were collected for each concentration and for a buffer blank. Data integration, averaging, and buffer subtraction was conducted using pyXS (35). Following inspection of each exposure using Primus (36), radiation-damaged exposures were excluded. Exposures were merged together using pyXS and Guinier analysis performed using Primus to calculate R_g . Pair distribution functions $P(r)$ and forward scattering $I(0)$ were calculated using GNOM (37) and molecular weights estimated separately based on Porod volumes

calculated in Primus and excluded bead volumes of ab initio models from DAMMIF (38). DAMFILT envelopes were superposed with crystal structures using SUPCOMB (33), but because of the extent of loops unmodeled in the crystal structure no further superposition analyses were conducted. Dimensionless Kratky plots of q_2 vs. $I(q)$, where $q = q \times R_g$ and $I(q) = I(q)/I(0)$, were generated as described (21, 39). Porod–Debye plots of $q^4 \times I(q)$ vs. q^4 were generated as described (21).

Rigid body modeling was conducted using CORAL (40) using the crystal structure of PAK4^{cat}-PAK4^{N45}-CDC42. Thirty-two unmodeled N-terminal residues include the uncleaved N-terminal His tag sequence “MGSSHHHHHSSGLVPRGSHMENLYFQGFQFK-RKKRVEISAPSNFEHRVHT-GFDQHEQKFTGLPROWQSLIEESAR,” there were ninety-two unmodeled PAK4 residues (44–299) (PAK4 isoform 2 used is shorter than isoform 1, so we have maintained the numbering of the kinase domain for consistency with previous catalytic domain structural studies; see ref. 10 for details). Nine unmodeled C-terminal residues of CDC42 include the uncleavable His tag “PPEHHHHHH.” Data were cut off at $q_{\max} = 0.48 \text{ \AA}^{-1}$. Flexibility analysis conducted by the EOM method (25) used the crystal structure of PAK4^{cat}-PAK4^{N45}-CDC42. The 92-aa linker was modeled using RANCH and data were cut off at $q = 0.25 \text{ \AA}^{-1}$. Model superpositions were made using SUPCOMB (33).

PAK4 Crystallization and Data Collection.

PAK4^{FL}-CDC42 complex. Equal amounts of PAK4^{FL} and CDC42^{Q61L} were mixed together and copurified by Superdex 75 column. The eluted complex was concentrated to 5 mg/mL and incubated with 1 mM phosphoaminophosphonic acid-adenylate ester (AMP-PNP), 1 mM guanlyl imidodiphosphate (GMP-PNP), and 5 mM MgCl₂ before setting up crystallization drops. Cocrystals of PAK4^{FL}-CDC42^{Q61L} were obtained in 0.1 M Tris-HCl, pH 8.5, 50 mM Na₂SO₄, and 6% (vol/vol) PEG 6,000 at room temperature. The crystals were cryoprotected reservoir buffer plus 25–30% (vol/vol) glycerol. Data were collected at the Advanced Photon Source (APS) beamline 24-ID-E.

PAK4^{cat}-PAK4^{N45}-CDC42 complex. Separately purified PAK4^{cat}, PAK4^{N45}, and CDC42^{Q61L} were mixed in a 1:1:1 molar ratio concentrated to 5 mg/mL and incubated with 1 mM AMP-PNP, 1 mM GMP-PNP, and 5 mM MgCl₂ before setting up crystallization drops. Cocrystals of PAK4^{cat}-PAK4^{N45}-CDC42^{Q61L} were grown using hanging-drop vapor diffusion methodology by mixing a 1:1 volume ratio of purified PAK4^{cat}-PAK4^{N45}-CDC42^{Q61L} and reservoir solution containing 0.1 M Hepes, pH 7.5, and 22% (vol/vol) PEG 1,000 at room temperature. Cryoprotectant buffer contained reservoir buffer plus 25–30% (vol/vol) glycerol. Data were collected at the APS beamline 24-ID-E.

Structure Determination and Refinement.

PAK4^{FL}-CDC42 complex. Crystallographic data were processed to 3.0- \AA resolution using the HKL2000 package (41). An initial molecular replacement solution was obtained using Phaser (42) using the previously determined crystal structures of PAK4 catalytic domain (PDB ID code 4FIF) (10) and CDC42 (PDB ID code 2ODB) as the search models. This yielded Z-scores for rotation and translation functions of 6.2 and 15.7, respectively, for PAK4 and 3.1 and 9.7 for CDC42. We noted, however, that CDC42 showed poor electron density. Therefore, for refinement we utilized only the coordinates for the PAK4 catalytic domain. We conducted refinement for the PAK4 using Phenix (43), model building in Coot (44), and model quality assessed using MolProbity (45). We then generated electron density maps and confirmed that the molecular

replacement solution for CDC42 correlated well with the unbiased difference density. As CDC42 showed poor electron density that did not allow positional refinement we fixed XYZ coordinates for PAK4 and CDC42. We then set the *B*-factor for CDC42 to the Wilson *B* and conducted one round of grouped *B*-factor refinement for CDC42, resulting in an average *B*-factor for CDC42 of 218.7 Å² (Table S2). The final *R* and *R*_{free} values for the PAK4^{FL}-CDC42 complex are 25.2% and 30.1%, respectively.

PAK4^{cat}-PAK4^{N45}-CDC42 complex. Crystallographic data were processed to 2.4-Å resolution using the HKL2000 package (41) and initial phases generated using molecular replacement using Phaser (42) and the previously determined crystal structures of PAK4 catalytic domain (PDB ID code 4FIF) (10) and CDC42 (PDB ID code 2ODB) as the search models. For the PAK4^{cat}-PAK4^{N45}-CDC42^{Q61L} structure this yielded Z-scores for rotation and translation functions of 8.7 and 6.4, respectively, for PAK4^{cat} and 3.4 and 13.1 for CDC42^{Q61L}. Refinement of the PAK4^{cat} and CDC42^{Q61L} was conducted using Refmac5 (46) with a maximum-likelihood target and TLS (translation, libration, screw). Model building was conducted in Coot (44) and during refinement interpretable electron density for PAK4^{N45} became clearly visible, which was built manually. Poor electron density is visible for the β and γ phosphates of AMP-PNP. Model quality assessed using MolProbity (45) yielded final *R* and *R*_{free} values of 22.1% and 26.4%, respectively. Crystallographic software was compiled by SBGrid (47) and structural figures generated using CCP4mg (48) and PyMOL. Buried surface area was calculated using the PISA server (49). RAC3-PAK4 CRIB (PDB ID code 2OV2) buries 2,028 Å² and CDC42-PAK6 CRIB (PDB ID code 2ODB) buries an average of 2,108 Å². A total of 3,152 Å² are buried between PAK4 and CDC42 in the PAK4^{cat}-PAK4^{N45}-CDC42 complex.

ITC. For ITC, a Nano-ITC (TA Instruments) was used. Purified proteins following preincubation with 5 mM MgCl₂ and 1 mM ATP for PAK4^{FL} or 1 mM GTP for GTPases were simultaneously dialyzed against PBS buffer (pH 7.4) three times. The reference cell was filled with water and the sample cell was filled with PAK4. Protein concentrations were determined using UV. For PAK4^{FL}-CDC42 interactions 14 μM of PAK4^{FL} was loaded in the sample cell and either 0.377 mM, 0.187 mM CDC42^{Q61L}, or 0.38 mM, 0.189 mM CDC42^{Q61L/E171R/L174R/E178R} was loaded in the syringe. For PAK4^{FL}-RAC1 interactions 14 μM of PAK4^{FL} was loaded in the sample cell and 0.312 mM and 0.159 mM RAC1^{Q61L} were loaded in the syringe. For PAK4^{cat} with PAK4^{N45}-CDC42^{Q61L} and PAK4^{cat} with PAK4^{N45} interactions 70 μM and

89 μM of PAK4^{cat} was loaded in the sample cell and 0.852 mM and 1.436 mM of PAK4^{N45}-CDC42^{Q61L} complex were loaded in the syringe. For PAK4^{FL-R489A}-CDC42^{Q61L} interactions 89 μM of PAK4^{FL-R489A} was loaded in the sample cell and 0.933 mM, 0.356 mM, and 0.466 mM CDC42^{Q61L} were loaded in the syringe. For PAK4^{FL-R489A}-CDC42^{Q61L/E171R/L174R/E178R} interactions 21 μM of PAK4^{FL-R489A} was loaded in the sample cell and 0.435 mM and 0.162 mM CDC42^{Q61L/E171R/L174R/E178R} were loaded in the syringe. For CDC42^{Q61L}-PAK4^{N45} interactions 93 μM of CDC42^{Q61L} was loaded in the sample cell and 0.898 mM and 0.449 mM of PAK4^{N45} were loaded in the syringe. For RAC1^{Q61L}-PAK4^{N45} interactions 92 μM of RAC1^{Q61L} was loaded in the sample cell and 0.449 mM PAK4^{N45} was loaded in the syringe. All ITC experiments were conducted at 25 °C using a stir speed of 350 rpm and incremental titration of 20 injections, 2.5-μL volume, and 300-s intervals. Buffer data (PBS into PBS titration) were subtracted from the corresponding experiment before data fitting. Raw data were processed and integrated with NanoAnalyze software (TA Instruments). The first injection in each experiment was not considered for the analysis. Stoichiometry of interaction (*N*), dissociation constant (*K*_d), enthalpy changes (ΔH), and entropy changes (ΔS) were determined using NanoAnalyze software.

In Vitro PAK4 Kinase Assay. Kinase activity was performed using a radioactive with β-catenin as substrate. C-terminal 6xHis-tagged PAK4-FL and its polybasic mutant (R5A and R8A) were purified as described above, and purified CDC42 (1–181, Q61L: constitutive active mutant) and CDC42^{ELE-RRR} mutant (1–181, Q61L, E171R, L174R, and E178R) were added as effector proteins. Briefly, kinase assays were performed by adding 0.3–0.4 μM of kinase, 2 μM of β-catenin, with/without CDC42/CDC42^{ELE-RRR}, 0.12 μM of cold ATP, and 0.05 μCi of hot [³²P] ATP in Tris buffer (20 mM Tris-HCl, pH 8.0, 0.3 M NaCl, 1 mM TCEP, and 10 mM MgCl₂), total volume 25 μL. The reaction was conducted at 30 °C for 5 min, 10 min, and 30 min then stopped by addition of 5× sample buffer and analyzed by SDS/PAGE. Dried gels were analyzed by exposure to phosphor storage screen (GE Healthcare) followed by scanning using a Molecular Imager FX Pro Plus System (Bio-Rad) and quantification by optical densitometry Quantity One (Bio-Rad). Relative activities were compared with those measured for PAK4-FL. Measurements were calculated from five independent experiments. Significant differences were calculated by Prism 7 (GraphPad Software) with ANOVA analysis.

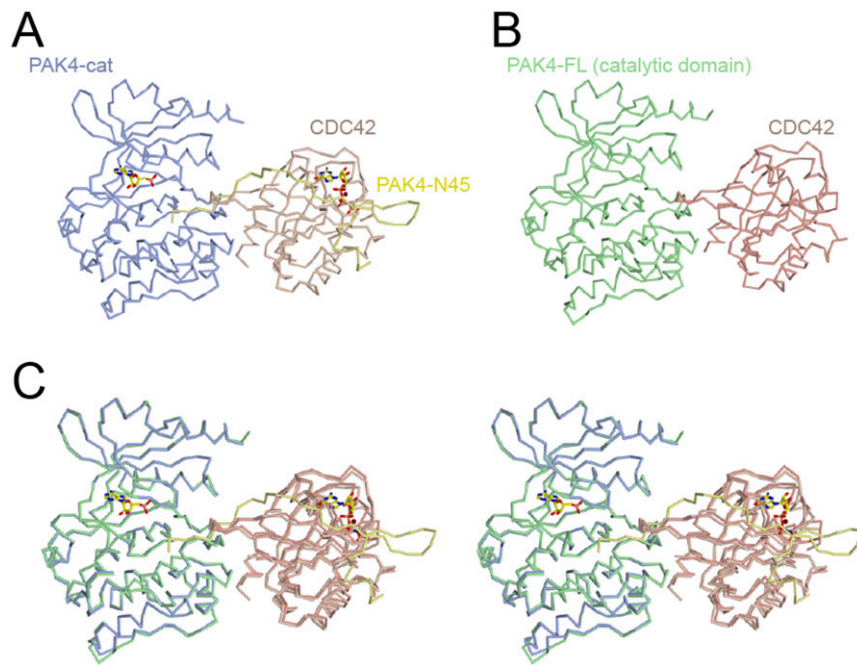


Fig. S4. Comparison of PAK4^{FL}-CDC42 and PAK4^{cat}-PAK4^{N45}-CDC42 structures. (A) Ribbon diagram of refined PAK4^{cat}-PAK4^{N45}-CDC42 structure. (B) Ribbon diagram of PAK4^{FL}-CDC42 molecular replacement solution. (C) Stereo view of the superposition of the two structures. Superposition is conducted for the kinase domain C-lobes only, as this accentuates any differences.

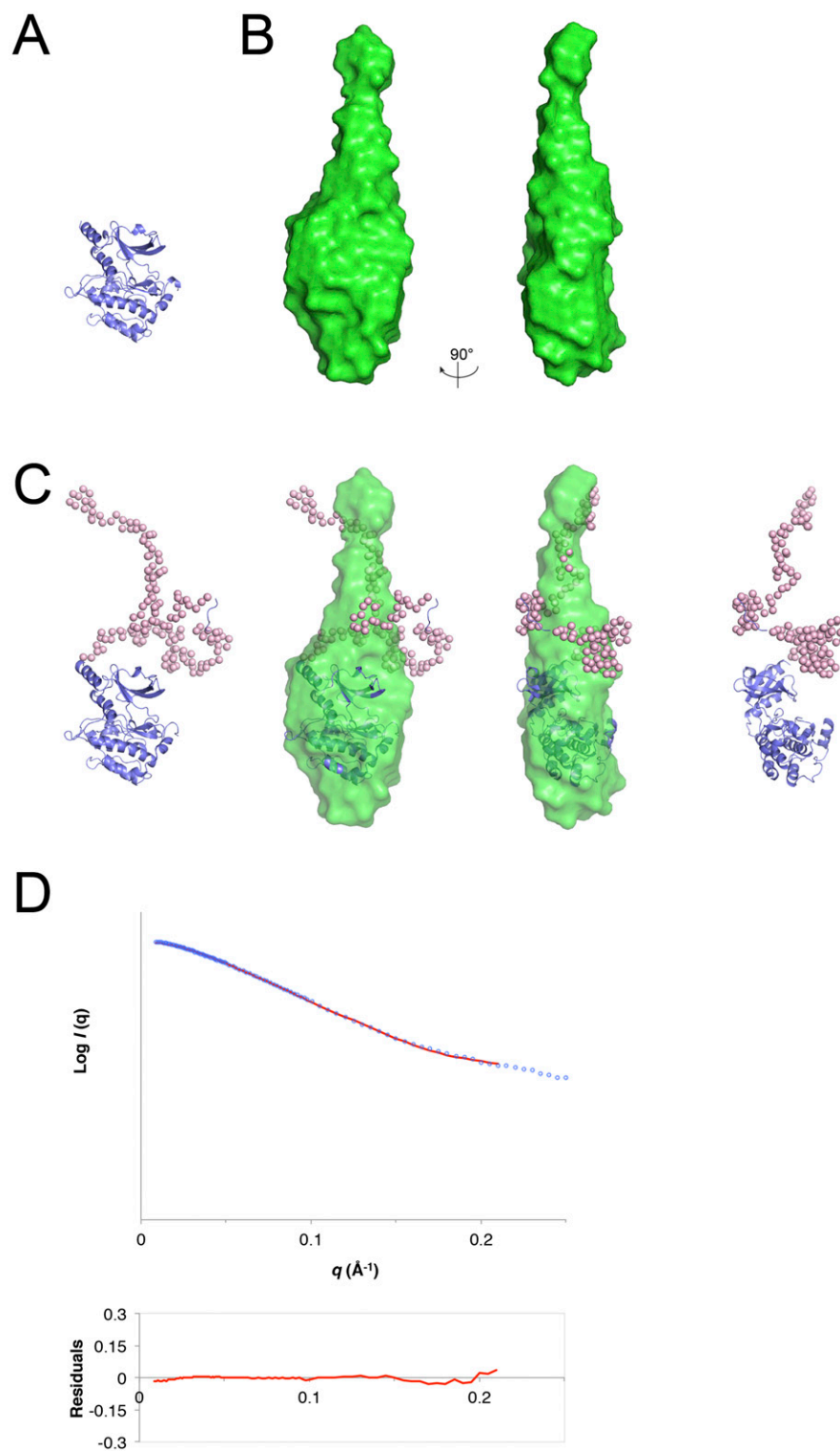


Fig. 55. SAXS analysis and rigid body modeling for PAK4-FL. (A) Crystal structure of PAK4^{cat} in blue. (B) Averaged SAXS molecular envelope generated using DAMMIF (DAMFILT envelopes shown) for PAK4^{FL}. (C) Rigid body analysis. CORAL model of PAK4^{FL} with the best statistical fit to the experimental data. Interdomain dummy atoms are shown as pink spheres. Superposition with the DAMMIF envelope is shown in the central pair. (D) Rigid body analysis. Experimental SAXS data (logarithmic scale) are shown fitted to the theoretical scattering profile of the best rigid body (red line). Residuals are shown below.

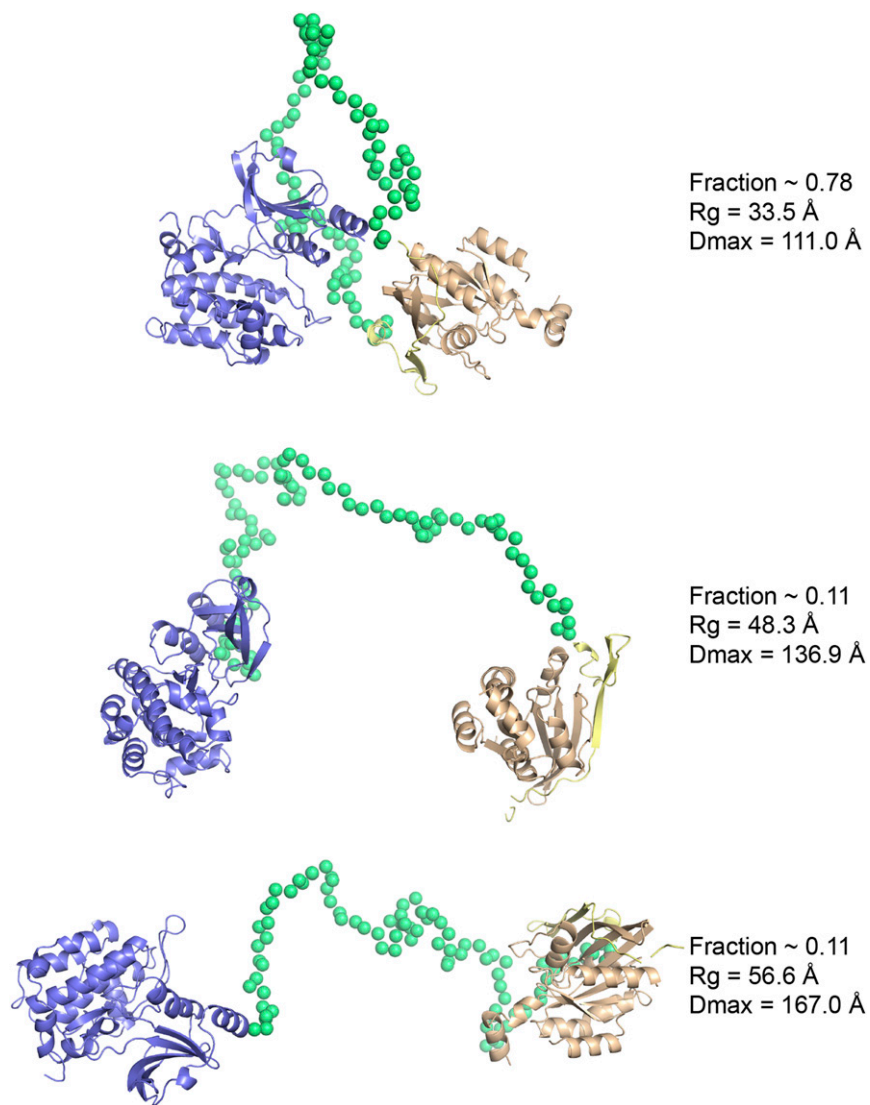


Fig. S6. Composition of the best EOM ensemble model. Three PDB models of PAK4-CDC42 were used in the best EOM model, these are shown, with the fraction comprised in the model, the R_g and D_{max} for each PDB model. As the complex in solution is flexible, this composition corresponds to the best fit to the experimental data under the parameters used.

Table S1. SAXS data collection and refinement statistics

	PAK4 ^{FL}	PAK4 ^{FL} -CDC42
Data collection parameters		
Beamline	NSLS-II LiX	NSLS-II LiX
Beam geometry	300- × 300-μm point source	300- × 300-μm point source
Detector	Pilatus 1M	Pilatus 1M
Beam wavelength, eV	10,790	10,790
Temperature, K	295	295
Protein concentration, mg/mL	11.7/5.9/3.0	6.3/3.1/1.6
Q range, Å ⁻¹	0.009–0.244/0.009–0.253/0.009–0.256	0.009–0.256/0.009–0.256/0.009–0.256
No. of exposures	5/5/5	5/5/5
No. of exposures averaged	5/5/5	5/5/5
Exposure time, s	5/5/5	5/5/5
Structural parameters		
<i>I</i> (0), cm ⁻¹ (reciprocal)	50.5/23.1/11.0	54.9/26.3/12.6
<i>R_g</i> , Å (reciprocal)	38.2/36.2/36.9	39.4/37.4/38.2
<i>I</i> (0), cm ⁻¹ (real)	50.5 ± 0.8/23.1 ± 0.7/11.0 ± 0.6	54.9 ± 0.8/26.3 ± 0.8/12.6 ± 0.8
<i>R_g</i> , Å (real)	38.5 ± 1.4/36.5 ± 2.6/37.3 ± 4.4	39.6 ± 1.3/37.6 ± 2.8/38.5 ± 5.9
<i>D_{max}</i> , Å	172/172/171	169/170/173
Model building		
Theoretical molecular mass, kDa	50.9	71.8
Porod volume	86,294	106,337
Molecular mass from Porod volume, kDa	50.7	62.5
No. of ab initio models calculated	20	20
Best model <i>P</i> value*	0.0088	0.0055
Best adjusted model <i>P</i> value*	0.1865	0.1164
Molecular mass from excluded volume, kDa	56.7	77.9
Rigid body model best χ^2 value	0.42	0.83
EOM ensemble χ^2 value	n.d.	0.13
EOM ensemble <i>R_g</i> , Å	n.d.	37.8
Software		
Data reduction and averaging	pyXS	pyXS
Data processing	Primus	Primus
Ab initio model generation	DAMMIF	DAMMIF
Rigid body model	CORAL	CORAL
Interdomain flexibility	EOM	EOM
Three-dimensional graphics representation	Pymol	Pymol

n.d., not determined.

*CorMap *P* values from DAMMIF.

Table S2. Crystallographic data collection and refinement statistics

	Cocrystallized complex	
	PAK4 ^{cat} -PAK4 ^{N45} -CDC42	PAK4 ^{FL} -CDC42
PDB ID code	SUPK	SUPL
Data collection		
Space group	<i>P</i> 2 ₁	<i>P</i> 6 ₃
X-ray source and detector	APS 24-ID-E ADSC Q315	APS 24-ID-E ADSC Q315
Wavelength, Å	0.97920	0.97922
Unit cell		
a, b, c, Å	58.5, 61.6, 85.8	141.7, 141.7, 62.1
α, β, γ, °	90, 91.6, 90	90, 90, 120
Resolution range, * Å	50.00–2.40 (2.49–2.40)	50.00–3.00 (3.11–3.00)
No. of unique reflections	24,144	14,433
Completeness, * %	100 (100)	99.9 (100)
<i>R</i> _{pimr} , * %	7.4 (91.2)	3.4 (70.7)
CC _{0.5} , % (in high-resolution bin)	31.1	58.8
Mean < >/<σ >*	12.0 (1.8)	20.3 (1.6)
Wilson <i>B</i> -factor	47.3	77.2
Redundancy	10.9 (10.6)	18.0 (17.2)
Refinement statistics		
Resolution range, * Å	50.0–2.4 (2.46–2.4)	43.6–3.0 (3.23–3.0)
<i>R</i> _{factor} , * %	22.1 (22.9)	25.2(36.3)
Free <i>R</i> _{factor} , * %	26.4 (32.1)	30.1 (47.3)
Free <i>R</i> reflections, * %	4.6 (4.1)	5.0 (4.7)
Free <i>R</i> reflections, * no.	1,097 (68)	726 (129)
Residues built		
PAK4	5–43, 300–588	299–589
CDC42	2–133, 137–177	2–178
No. of water molecules	25	1
Mean <i>B</i> -factor, Å ²		
Protein (PAK4)	60.6	73.4
Protein (PAK4 ^{cat})	56.2	—
Protein (PAK4 ^{N45})	88.8	—
AMP-PNP	62.9	—
Protein (CDC42)	84.2	218.7
GMP-PNP	83.9	—
Mg ²⁺	76.6	—
H ₂ O	48.6	—
Model statistics		
Rmsd bond lengths, Å	0.008	0.012
Rmsd bond angles, °	1.375	1.557
Ramachandran plot, % (favored/allowed/disallowed)	96.4/3.6/0	95.4/4.1/0.4

*Indicates high-resolution shell.

Table S3. Thermodynamic properties of the interaction between PAK4 and small GTPases RAC1 and CDC42

Sample cell	Syringe	K_d , μM	N	ΔH , kJ/mol	ΔS , J/mol·K	ΔG , kJ/mol
PAK4 ^{FL}	CDC42	0.2	0.91	-81.2	-145.1	-37.9
PAK4 ^{FL}	CDC42	0.5	0.95	-81.6	-152.4	-36.2
PAK4 ^{FL}	CDC42 ^{ELE-RRR}	1.0	0.92	-91.2	-191.1	-34.3
PAK4 ^{FL}	CDC42 ^{ELE-RRR}	1.1	0.88	-100.0	-211.5	-37.0
PAK4 ^{FL}	RAC1	3.2	1.29	-38.5	-24.0	-31.3
PAK4 ^{FL}	RAC1	2.8	0.70	-40.8	-30.7	-31.7
PAK4 ^{cat}	PAK4 ^{N45} :CDC42	142.0	1.18	14.3	121.7	-21.9
PAK4 ^{cat}	PAK4 ^{N45} :CDC42	66.7	1.51	12.4	121.6	-23.8
CDC42	PAK4 ^{N45}	0.98	0.60	-103.9	-232.7	-34.55
CDC42	PAK4 ^{N45}	1.43	0.53	-96.52	-211.8	-33.36
CDC42	PAK4 ^{N45}	1.58	0.56	-102.4	-232.3	-33.11
RAC1	PAK4 ^{N45}	7.7	0.82	-101.2	-241.7	-29.18
RAC1	PAK4 ^{N45}	12.5	0.90	-111.1	-278.7	-28.00
PAK4 ^{FL-R489A}	CDC42	1.58	0.69	-59.37	-88.04	-33.12
PAK4 ^{FL-R489A}	CDC42	1.60	1.33	-116.2	-278.9	-33.08
PAK4 ^{FL-R489A}	CDC42	0.71	1.04	-56.21	-70.87	-35.08
PAK4 ^{cat}	PAK4 ^{N45}	No heat	No heat	No heat	No heat	No heat

Data from each of the independent ITC experiments are shown. Experiments conducted at 298 K. No heat indicates no measurable heat was observed from the titration.