SUPPLEMENTAL INFORMATION FOR:

Understanding the antagonism of Retinoblastoma Protein dephosphorylation by PNUTS provides insights into the PP1 Regulatory Code

Meng S. Choy[1], Martina Hieke[1], Ganesan Senthil Kumar[1], Greyson R. Lewis[2], Kristofer R. Gonzalez-DeWhitt[1], Rene P. Kessler[1], Benjamin J. Stein[1], Manuel Hessenberger[2], Angus C. Nairn[3], Wolfgang Peti[1,4] & Rebecca Page[2,*]

Inventory of Supplemental Information:

SI Materials and Methods

SI Materials and Methods

Cloning and expression. PP1α_{7-300} and PP1α_{7-330}, PP1β_{6-327} and PP1γ_{7-323} (isoform PP1γ1) were subcloned into the RP1B vector and produced as previously described (1). PNUTS DNA (*Rattus norvegicus*) encoding residues 309-433, 376-433, 376-453, 394-453 and 394-433, respectively, were cloned into a pETM30-GST or pETM30-GST-Nhe1 vector, that encodes an N-terminal His_6-tag followed by glutathione S-transferase (GST) and a TEV (tobacco etch virus) protease cleavage site. Single and double point mutations in PNUTS were introduced using the Quikchange mutagenesis kit (Agilent) following the manufacturer's protocol. All constructs were sequence verified. PNUTS was expressed in *E. coli* BL21 (DE3) CodonPlus-RIL cells (Agilent). Cells were grown in Luria Bertani Broth in the presence of selective antibiotics at 37°C to an OD_{600} of ~0.6, and expression was induced by the addition of 1 mM IPTG. Induction proceeded for ~20 h at 18°C prior to harvesting by centrifugation at 6,000 ×*g*. Cell pellets were stored at -80°C until purification.

For NMR measurements, expression of uniformly ^{15}N- or $^{15}N/^{13}C$-labeled $PNUTS_{309-433/376-453}$ was achieved by growing cells in M9 minimal media containing 1 g/L $^{15}NH_4Cl$ and/or 4 g/L [^{13}C]-*D*-glucose (CIL) as the sole nitrogen and carbon sources, respectively. Uniformly $^{2}H/^{15}N$-labeled PNUTS ([^{2}H,^{15}N]-$PNUTS_{394-433}$ and [^{2}H,^{15}N]-$PNUTS_{376-453}$) were expressed in M9 media supplemented with $^{15}NH_4Cl$ (1g/L) in 99% D_2O. Multiple rounds of D_2O adaptation were necessary for robust cell growth.

Purification of PP1α. Purification of PP1α_{7-300} for crystallization was performed as follows. About 10 – 15 g of cells expressing His_6-TEV-PP1α_{7-300} (1, 2) were lysed in PP1 Lysis Buffer (25 mM Tris pH 8.0, 700 mM NaCl, 5 mM imidazole, 1 mM $MnCl_2$, 0.1% Triton X-100) using high-pressure homogenization (Avestin C3) in the presence of EDTA-free protease inhibitor cocktail (Roche). The lysate was clarified by centrifugation at 100,000 ×*g* and loaded onto Ni^{2+}-NTA resin (Qiagen). Bound His_6-PP1 was washed with PP1 Buffer A (25 mM Tris pH 8.0, 700 mM NaCl, 5 mM imidazole, 1 mM $MnCl_2$), followed by a stringent wash using a buffer consisting of 94% PP1 Buffer A and 6% PP1 Buffer B (25 mM Tris pH 8.0, 700 mM NaCl, 0.25 M imidazole, 1 mM $MnCl_2$). The PP1α_{7-300} was eluted in PP1 buffer B and immediately purified using size exclusion chromatography (SEC, Superdex 75 26/60; 20 mM Tris pH 8.0, 500 mM NaCl, 0.5 mM TCEP). Peak fractions were incubated overnight with TEV protease at 4°C. The following day, the cleaved PP1α_{7-300} protein was loaded onto Ni^{2+}-NTA resin (Qiagen) and the flow-through collected. After concentration, the cleaved PP1α_{7-300} was purified using SEC (Superdex 75 26/60; 20 mM Tris pH 8.0, 500 mM NaCl, 0.5 mM TCEP). Fractions containing the PP1α_{7-300} were concentrated to 8 mg/mL and immediately used for crystallization trials.

Purification of PNUTS. Cells expressing PNUTS were lysed in lysis buffer (25 mM Tris pH 8.0, 500 mM NaCl, 5 mM imidazole, 0.1% Triton X-100) using high-pressure homogenization (Avestin C3) in the presence of EDTA-free protease inhibitor cocktail (Roche). The lysate was clarified by centrifugation at 50,000 ×*g* and loaded onto a HisTrap HP column (GE Healthcare). Elution was carried out in a 50 mM Tris pH 8.0, 500 mM NaCl buffer using a 5 – 500 mM imidazole gradient. Peak fractions were

pooled and dialyzed overnight at 4°C (20 mM Tris pH 7.0, 250 mM NaCl [$PNUTS_{394-433}$, $PNUTS_{394-453}$], pH 7.3 [$PNUTS_{309-433}$] or pH 8.0 [$PNUTS_{376-453,}$ $PNUTS_{376-433}$]) with TEV protease for His_6-tag cleavage. The following day, the cleaved protein was heat purified in two cycles. First, the dialyzed, cleaved protein was heated at 65°C for 15 min in a water bath. The supernatant was clarified by centrifugation and filtered (0.22 µm). After concentration to 4-5 mL, a second cycle of heat purification was performed at 80°C (15 min). The supernatant was collected, filtered (0.22 µm) and concentrated to 1 mM, flash frozen in liquid nitrogen and stored at -80°C until further use. For NMR spectroscopy, $^{13}C/^{15}N$-labeled PNUTS was purified as described above with one additional step using SEC (Superdex 75 26/60), concentrated to 1 mM, again heat purified (80°C; 15 min) to denature trace proteases, and used immediately (20 mM Na-phosphate pH 6.5, 50 mM NaCl, 0.5 mM TCEP, 10% D_2O).

Purification of the PNUTS:PP1α holoenzyme. Purification of the PNUTS:PP1α holoenzyme for crystallization trials was performed as follows. A ~10-15 g cell pellet expressing $PP1\alpha_{7-300}$ was lysed in PP1 Lysis Buffer (25 mM Tris pH 8.0, 700 mM NaCl, 5 mM imidazole, 1 mM $MnCl_2$, 0.1% Triton X-100), clarified by ultracentrifugation (100,000 *xg*) and immobilized on Ni^{2+}-NTA resin (Qiagen). Bound His_6-$PP1\alpha_{7-300}$ was washed with PP1 Buffer A (25 mM Tris pH 8.0, 700 mM NaCl, 5 mM imidazole, 1 mM $MnCl_2$), followed with a stringent wash using a buffer consisting of 95% PP1 Buffer A and 5% PP1 Buffer B (25 mM Tris pH 8.0, 700 mM NaCl, 250 M imidazole, 1 mM $MnCl_2$) and a low salt PP1 Buffer A wash (25 mM Tris pH 8.0, 150 mM NaCl, 5 mM imidazole, 1 mM $MnCl_2$) prior to incubation with $PNUTS_{394-433}$ for 1 h at 4°C. The complex was eluted in low salt PP1 Buffer B (25 mM Tris pH 8.0, 150 mM NaCl, 250 mM imidazole, 1 mM $MnCl_2$). Next, the complex was purified by SEC (Superdex 200 26/60; 20 mM Tris pH 8.4, 150 mM NaCl, 0.5 mM TCEP). Peak fractions were incubated overnight with TEV protease at 4°C. The following day, $PNUTS_{394-433}$:$PP1\alpha_{7-300}$ was incubated on Ni^{2+}-NTA resin for 1 hour to bind TEV protease and the cleaved His_6-tag. The flow-through was collected, concentrated to 10-12 mL and further purified by SEC (Superdex 75 26/60; 20 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM TCEP). Fractions containing the $PNUTS_{394-433}$:$PP1\alpha_{7-300}$ complex were concentrated to 6 mg/mL and immediately used for crystallization trials.

Isothermal titration calorimetry. Purification of free PP1 ($PP1\alpha_{7-330}$, $PP1\alpha_{7-300}$, $PP1\beta_{6-327}$, $PP1\gamma_{7-323}$; the long constructs result in the most stable, well-behaved PP1 for ITC studies) for ITC measurements was performed as follows. PP1 was lysed in PP1 Lysis Buffer (25 mM Tris pH 8.0, 700 mM NaCl, 5 mM imidazole, 1 mM $MnCl_2$, 0.1% Triton X-100), clarified by ultracentrifugation (100,000 *xg*) and immobilized on Ni^{2+}-NTA resin (Qiagen). Bound His_6-PP1 was washed with PP1 Buffer A (25 mM Tris pH 8.0, 700 mM NaCl, 5 mM imidazole, 1 mM $MnCl_2$), followed by a stringent wash containing 6% PP1 Buffer B (25 mM Tris pH 8.0, 700 mM NaCl, 250 mM imidazole, 1 mM $MnCl_2$). PP1 was eluted with PP1 Buffer B (25 mM Tris pH 8.0, 700 mM NaCl, 250 mM imidazole, 1 mM $MnCl_2$) and further purified using SEC (Superdex 75 26/60; GE Healthcare; ITC buffer: 20 mM Tris pH 7.5, 500 mM NaCl, 0.5 mM TCEP). Purified PP1 was used immediately for ITC measurements following purification. Purified $PNUTS_{394-433}$ was stored at -80 °C and was SEC purified (Superdex 75 10/300 GL) in ITC buffer before each ITC run. For

the $PNUTS_{376-453}$/$PP1\alpha_{7-330}$ and $PNUTS_{394-453}$/$PP1\alpha_{7-330}$ experiments, the purification of $PP1\alpha_{7-330}$ was carried out as described above except the Ni^{2+}-NTA bound PP1 was eluted into a stirring beaker with ITC buffer (20 mM Tris pH 7.5, 500 mM NaCl, 0.5 mM TCEP). $PNUTS_{376-453}$ or $PNUTS_{394-453}$ and $PP1\alpha_{7-330}$ were then dialyzed overnight into 20 mM Tris pH 7.5, 500 mM NaCl, 0.5 mM TCEP at 4°C. PNUTS (~30 μM) was titrated into $PP1\alpha_{7-330}$ (~2-4 μM) using a VP-ITC micro-calorimeter at 25ºC (Microcal, Inc.). Data were analyzed using NITPIC (3) and SEDPHAT (4).

Crystallization and structure determination. Crystals of the $PNUTS_{394-433}$:$PP1\alpha_{7-300}$ holoenzyme were obtained using sitting drop (200 nL) vapor diffusion in 0.1 M Tris pH 7.8, 1.0 M LiCl, 18% PEG6000 at 4°C ($P3_221$) or 2% (v/v) tascimate, pH 8.0, 0.1 M Tris, pH 8.5, 16% (w/v) PEG3350 ($P4_12_12$). Crystals were cryo-protected by a 60 second soak in mother liquor supplemented with 30% glycerol and immediately flash frozen. X-ray data to 2.2 Å ($P3_221$) and 2.1 Å ($P4_12_12$) were collected at the National Synchrotron Light Source (BNL) beamline X25 at 100 K and a wavelength of 1.1 Å using a Pilatus 6M detector. The $PNUTS_{394-433}$:$PP1\alpha_{7-300}$ holoenzyme structures were solved by molecular replacement using Phaser as implemented in PHENIX (5) ($PP1\alpha_{7-300}$ (PDB: 3EA7, (1) was used as the search model). Clear electron density for the bound PNUTS was visible in the initial maps. The initial models of the $PNUTS_{394-433}$:$PP1\alpha_{7-300}$ holoenzymes were built using Phenix.AutoBuild, followed by iterative rounds of refinement in PHENIX (5) and manual building using Coot (6).

Crystals of PP1 (8 mg/mL $PP1\alpha_{7-300}$; 20 mM Tris pH 8.0, 500 mM NaCl, 0.5 mM TCEP) were obtained using the sitting drop (200 nL) vapor diffusion method at 4°C. Large, rod-shaped crystals formed in 0.1 M HEPES pH 7.0, 1.0 M LiCl, 20% (w/v) PEG6000. Crystals were cryo-protected by a 5 sec soak in mother liquor supplemented with 30% glycerol and immediately flash frozen in liquid nitrogen. A 1.45 Å dataset from a single $PP1\alpha_{7-300}$ crystal was collected at Beamline X25 at the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory. The structure was determined by molecular replacement using Phaser as implemented in PHENIX (5) (PDB: 3EA7, (1) was used as the search model). A solution was obtained in space group $P2_12_12_1$. The model was completed using iterative rounds of refinement in PHENIX (5) and manual building using Coot (6).

NMR measurements. All NMR measurements were performed at 298 K. The sequence-specific backbone assignment of unbound PNUTS (residues 376-453) was obtained from the following experiments performed on a Bruker Avance 500 MHz spectrometer with a HCN TCI z-gradient cryoprobe: 2D [^{1}H,^{15}N] HSQC, 3D HNCA, 3D HNCACB, 3D CBCA(CO)NH and 3D (H)CC(CO)NH (τ_m= 12 ms). The NMR spectra were processed with Topspin 1.3/2.1 (Bruker) and analyzed using the CARA software package (http://www.cara.nmr.ch). Assignments for $PNUTS_{376-453}$ were obtained for 82.5% of the backbone nuclei (^{13}Cα, Hα, ^{15}N, and H_N). Of the 77 expected backbone amide N-H pairs (4 prolines), 63 were identified. Accurate assignment was uncertain for 14 residues due to spectral overlap, low signal to noise and/or undetectable peaks. These residues are W401, K406, R408, E409, Y410, F411, Y412, F413, E414, L415, R420, I437, D440 and R441.

[^{2}H,^{15}N]-PNUTS$_{394-433}$ and [^{2}H,^{15}N]-PNUTS$_{376-453}$ were purified as previously described. Purification of the [^{2}H,^{15}N]-PNUTS$_{376-453}$:PP1α$_{7-330}$ complex was carried out identically as described for crystallization trials, up to the point where the complex bound to the Ni^{2+}-NTA column was eluted in low salt PP1 Buffer B (25 mM Tris pH 8.0, 150 mM NaCl, 250 mM imidazole, 1 mM $MnCl_2$). After this step, the eluted sample was immediately purified using SEC (Superdex 75 26/60 pre-equilibrated with NMR buffer: 20 mM Bis-Tris pH 6.5, 150 mM NaCl, 0.5 mM TCEP). Fractions containing the complex were pooled and concentrated using a 3 kDa cut off concentrator (Millipore). D_2O (10% v/v) was added to the sample prior to the NMR measurement. 2D [^{1}H,^{15}N] TROSY spectra of unbound [^{2}H,^{15}N] PNUTS$_{376-453}$ (100 μM), unbound [^{2}H,^{15}N] PNUTS$_{394-433}$ (100 μM) and the [^{2}H,^{15}N] PNUTS$_{376-453}$:PP1$_{7-330}$ complex (250 μM) were collected on a Bruker Avance IIIHD 850 MHz spectrometer with a HCN TCI z-gradient cryoprobe (298 K).

Dephosphorylation Reactions. Dephosphorylation of *p*NPP was initiated by addition of 0.1 μM holoenzyme to a reaction containing 2 mM *p*NPP in assay buffer (150 mM Bis-Tris, pH 6.5, 150 mM NaCl). Reactions were incubated at 30°C in a waterbath for 30 min and quenched by the addition of 1 M NaOH. The concentration of dephosphorylated *p*NP product was determined from the optical absorbance at 405 nm using a molar extinction coefficient of 18000 $M^{-1}cm^{-1}$.

Pull Down Assay. Four mutants of PNUTS$_{394-433}$ (*RVxF-mutant*, TVTW →TATA; *ΦΦ-mutant*, YF→AA; *R-mutants* R→A and R→E), were expressed and purified using the methods established for wt-PNUTS$_{394-433}$ with an additional Ni-NTA bead (Qiagen) incubation to ensure complete removal of any trace Ni-NTA interacting species prior to the assays. PP1α$_{7-330}$ was purified using Ni-NTA beads and SEC (20 mM Tris pH 7.5, 500 mM NaCl, 0.5 mM TCEP). 500 μL of pure PP1α$_{7-330}$ (at a concentration of 8 μM) was incubated with 25 μL bed volume of Ni-NTA beads for an hour with rocking (4 °C). The beads were pelleted by centrifugation (2000xg) and washed 3 times with 500 μL of SEC buffer. Wt- and mutant PNUTS were diluted to 12.5 μM in SEC buffer and incubated with PP1-saturated Ni-NTA beads (overnight at 4 °C). After pelleting the beads by centrifugation, the supernatant ("flow-through") was collected. The beads were then washed 3 times with SEC buffer. 30 μL of SDS loading buffer was then added to the beads and flow-through samples, after which both were boiled at 100 °C for 5 minutes. The samples were run on NuPAGE Novex 4 – 20% Bis-Tris gels, stained with SYPRO® Ruby protein gel stain (Life Technologies) according to manufacturer's protocols and the gels were scanned using a Typhoon 9410 scanner using blue laser at the excitation wavelength of 457 nm and emission filter of 610 nm (GeHealthcare) following destaining. Densitometry was performed using ImageQuant TL 7.0 (GeHealthcare).

Sequence variability within the ΦΦ motif. Sequence variability in the ΦΦ motif and Arg motifs was examined using JackHMMER (7) using the sequences of the PP1 interacting residues from spinophilin (residues 424-474), NIPP1 (158-216) and PNUTS (396-424). The sequences identified by JackHMMER were edited to remove duplicates

and/or incomplete sequences. This resulted in 125, 107 and 85 sequences from a diversity of organisms for spinophilin/neurabin, NIPP1 and PNUTS, respectively. Weblogo (8) was then used to generate logos for each family of aligned sequences.

Bioinformatics. A structure-based sequence alignment (namely, the PP1 holoenzymes complexes were aligned (6) using the coordinates of PP1 and the sequences of the PP1 interacting proteins aligned by hand based using the structural overlap of corresponding residues in the holoenzyme structures) was used to determine the number of residues between the motifs (5-8 between the RVxF and ΦΦ motifs; 8-9 between the ΦΦ- and Arg motifs). To determine the more general RVxF-ΦΦ-Arg motif, a logo was generated using the sequences from all three families of PP1 interacting proteins (317 sequences total).

ScanProsite (9) was used to identify additional PP1 interacting proteins that contain an RVxF-ΦΦ or an RVxF-ΦΦ-Arg motif. The primary sequences of 189 confirmed PP1 interacting proteins (10) were probed using the expressions "[RK]-x(0,1)-[IV]-x-[FW]-x(5,8)-[VIYFH]-[FIYRHNQSC]" and "[RK]-x(0,1)-[IV]-x-[FW]-x(5,8)-[VIYFH]-[FIYRHNQSC]-x(8,9)-R", where "[RK]-x(0,1)-[IV]-x-[FW]" defines the RVxF motif, "[VIYFH]-[FIYRHNQSC]" defines the ΦΦ motif (residues allowed at these positions determined using JackHMMER, as described above) and "R" defines the Arg motif; x(5,8) and x(8,9) represents the number of residues between the RVxF and ΦΦ motifs and ΦΦ and Arg motifs, respectively.

4 of the 189 PP1 binding proteins were not identified by the RVxF-filter "[RK]-x(0,1)-[IV]-x-[FW]": CASP2 ('H' at the [FW] position), I-2 ('Q' at the [IV] position), POLD3 ('L' at the [FW] position) and Rb ('L' at the [IV] position) and thus their sequences were analyzed manually for the presence of a ΦΦ and/or Arg motif. Default settings were used for the ScanProsite search with exception of 'greediness', which was disabled. Excluding the known interactors (PNUTS, NIPP1, spinophilin and neurabin), 72 sequences satisfied the ΦΦ-filter ([RK]-x(0,1)-[IV]-x-[FW]-x(5,8)-[VIYFH]-[FIYRHNQSC]) and 16 sequences satisfied the Arg-filter ([RK]-x(0,1)-[IV]-x-[FW]-x(5,8)-[VIYFH]-[FIYRHNQSC]-x(8,9)-R). These results were filtered to remove hits at sites other than experimentally determined RVxF sites, resulting in 51 and 10 proteins, respectively (52 and 11 with *CASP2* included).

The probability of these proteins containing functional ΦΦ and Arg motifs (i.e., ones that bind in the PP1 ΦΦ and Arg motif binding grooves) was further evaluated using homology, secondary structure prediction and disorder prediction. First, conservation of the arginine across different species was evaluated using multiple algorithms (PHMMER (11), BLAST (12), homologene (13)), as the arginine is perfectly conserved in spinophilin/neurabin and PNUTS. Next, secondary structure and protein disorder for these regions were predicted using psipred (14) and IUPRED (15, 16), respectively, as the currently known PP1 regulatory proteins that contain ΦΦ and Arg motifs are IDPs in their unbound state. This information was used to group the proteins predicted to contain an RVxF-ΦΦ-Arg motif as 'likely', 'somewhat likely', 'unclear' and 'unlikely'; only those proteins that are predicted to be an IDP were considered to have a functional RVxF-ΦΦ motif.

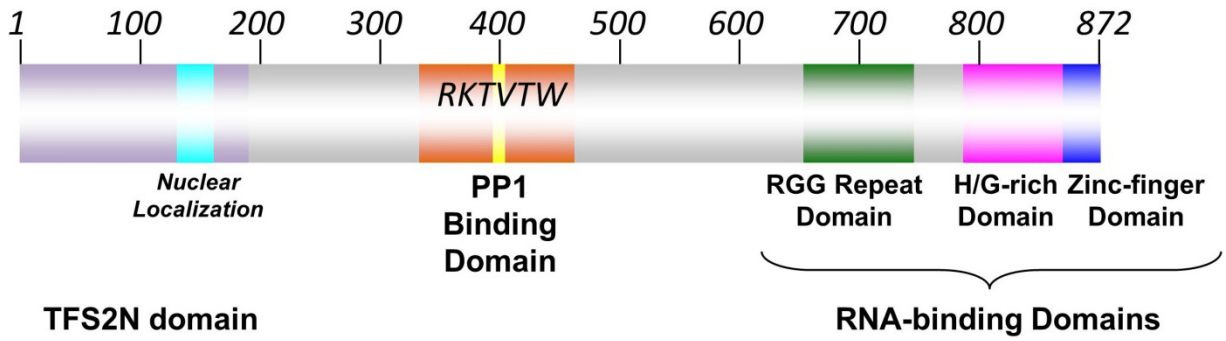

Figure S1. *Domain architecture of PNUTS.*

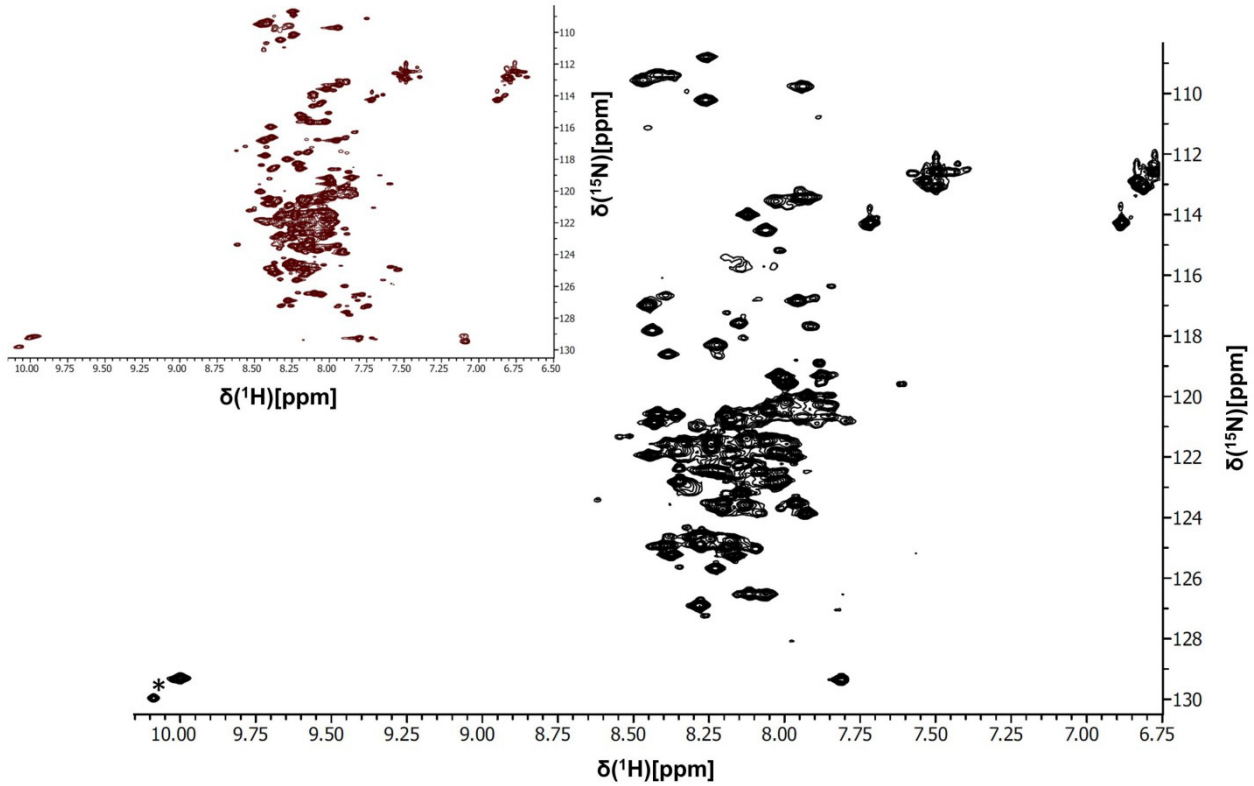

Figure S2. *2D [^{1}H,^{15}N] HSQC spectrum of ^{15}N labeled PNUTS$_{309-433}$:* ~230 µM in 20 mM Tris pH 6.8, 150 mM NaCl, 0.5 mM TCEP (PP1 compatible buffer), 10% D_2O, 298 K, 500 MHz [^{1}H]. Dramatically reduced dispersion in the ^{1}H chemical shift dimension is typical for intrinsically disordered proteins. As can be readily observed, the spectrum is plagued by slowly exchanging peaks, likely due to cis/trans isomerization of the 18 proline residues in the PNUTS$_{309-433}$ domain (this is most easily seen in the Trp NHε peaks labeled with '*'). *Insert*, *2D [^{1}H, ^{15}N] HSQC spectrum of PNUTS$_{309-433}$:* ~250 µM in 20 mM Na-phosphate pH 6.6, 50 mM NaCl, 0.5 mM TCEP (sequence-specific backbone assignment buffer), 10% D_2O, 298 K, 500 MHz [^{1}H]. Spectrum shows nearly twice the number of expected peaks. Thus, sequence-specific backbone assignments, some of which were uncertain, were only obtained for 47% of the nuclei.

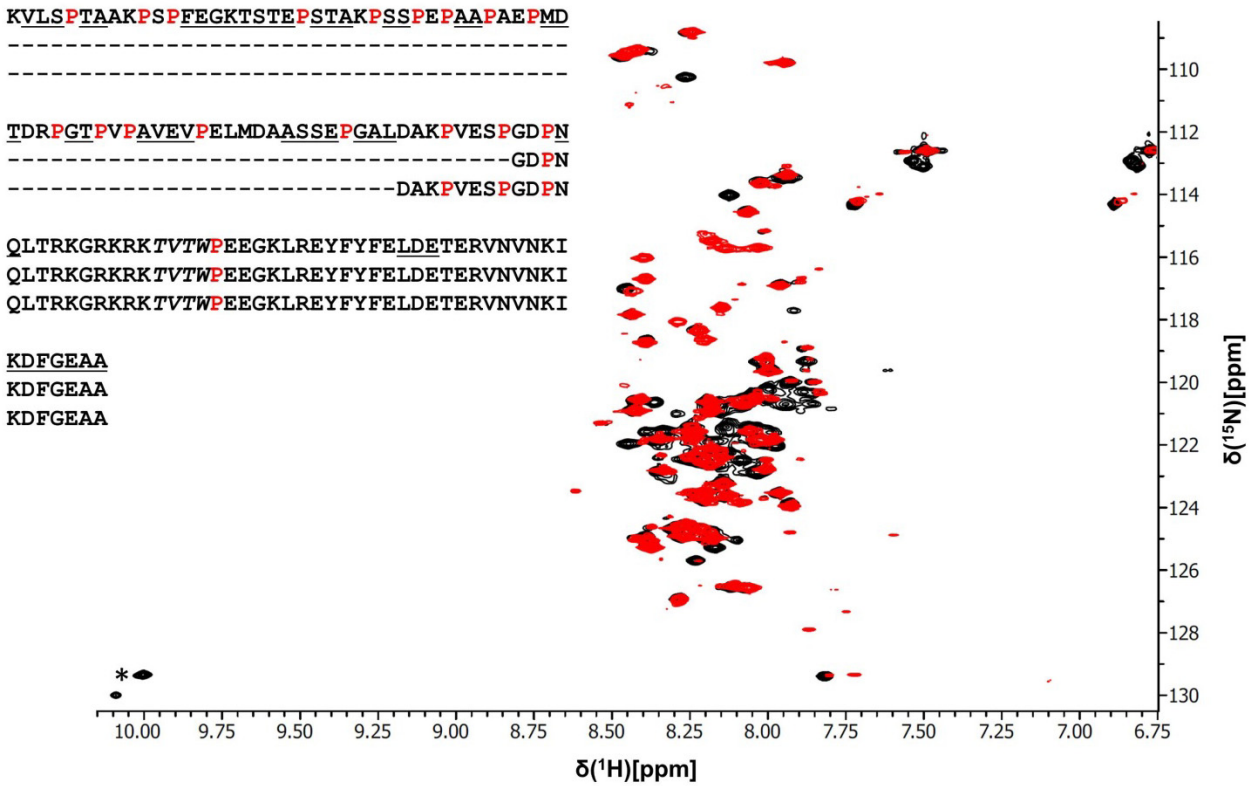

Figure S3. *2D [^{1}H,^{15}N] HSQC spectra of ^{15}N labeled PNUTS$_{309-433}$ alone and in complex with PP1α$_{7-330}$.* The 2D [^{1}H,^{15}N] HSQC spectrum of ^{15}N labeled PNUTS$_{309-433}$ alone is in black: ~230 μM in 20 mM Tris pH 6.8, 150 mM NaCl, 0.5 mM TCEP (PP1 compatible buffer), 10% D_2O, 298 K, 500 MHz [^{1}H]. The 2D [^{1}H,^{15}N] HSQC spectrum of ^{15}N labeled PNUTS$_{309-433}$ in complex with PP1α$_{7-330}$ is in red: ~100 μM in PP1 compatible buffer, 10% D_2O; 500 MHz [^{1}H]. A number of peaks are missing in the ^{15}N labeled PNUTS$_{309-433}$:PP1 α_{7-330} spectrum due to a direct interaction with PP1 and thus significantly reduced transverse relaxation rates; the Trp NHε peaks are labeled with '*'. Insert: primary sequences of PNUTS$_{309-433}$, PNUTS$_{383-433}$ and PNUTS$_{376-433}$, showing the increased number of prolines between residues 309 to 380. The sequence specific backbone assignment was achieved for all underlined residues and was difficult due to significant chemical shift overlap as well as the large number doubled peaks due to proline cis/trans isomerization. Nevertheless, analysis of these spectra enabled us to conclude that only residues beyond 376, consistent with previously reported biochemical data (17), are most likely to interact directly with PP1.

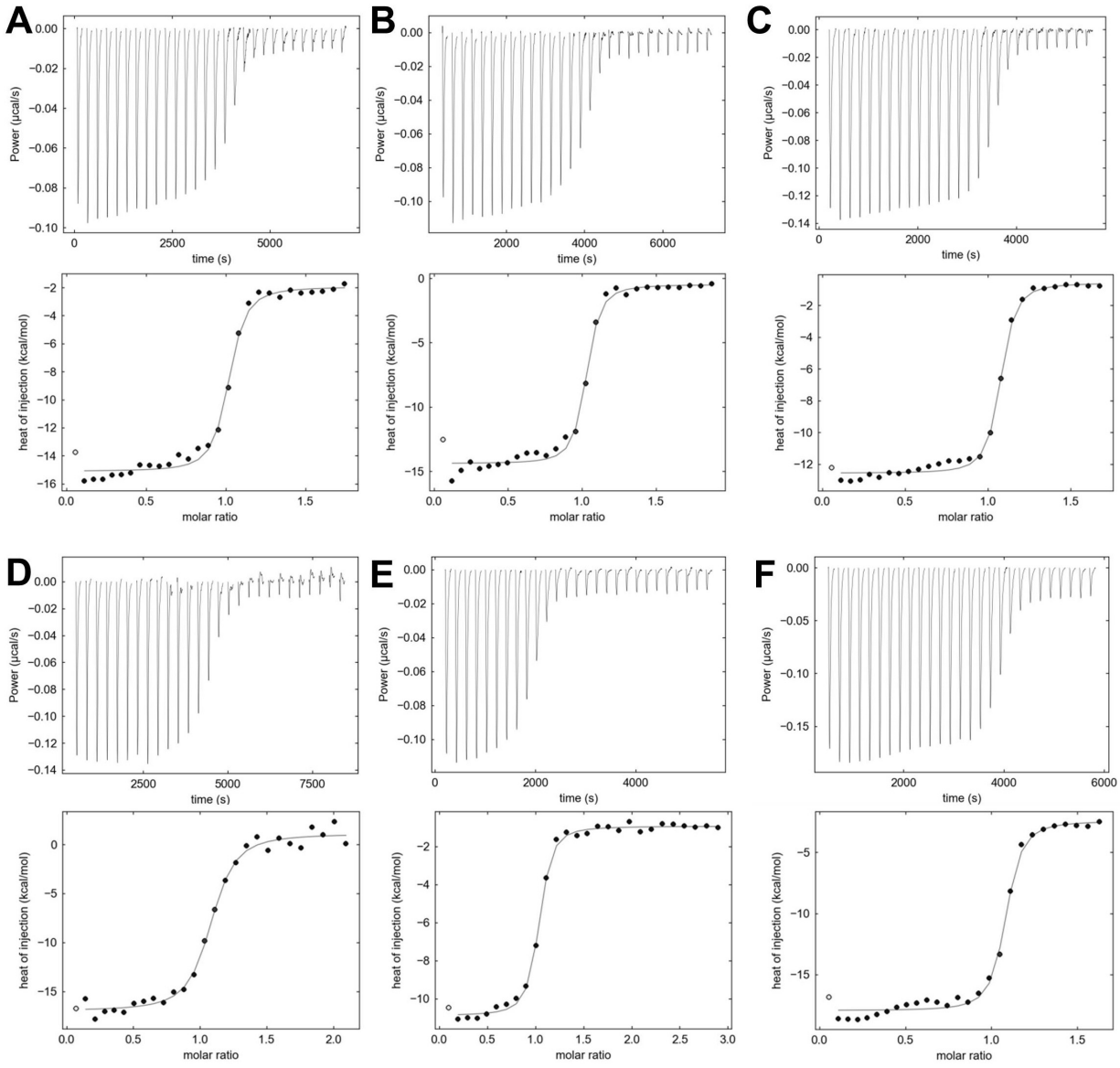

Figure S4. *Isothermal titration calorimetry of PNUTS with PP1.* (**A**) $PNUTS_{376-453}$:$PP1\alpha_{7-330}$ (K_d, 8.7 ± 0.8 nM); (**B**) $PNUTS_{394-453}$:$PP1\alpha_{7-330}$ (K_d, 9.8 ± 3.5 nM); (**C**) $PNUTS_{394-433}$:$PP1\alpha_{7-330}$ (K_d, 9.3 ± 1.9 nM), (**D**) $PNUTS_{394-433}$:$PP1\alpha_{7-300}$ (K_d, 21.2 ± 0.1 nM), (**E**) $PNUTS_{394-433}$:$PP1\beta_{6-327}$ (K_d, 10.7 ± 2.6 nM), (**F**) $PNUTS_{394-433}$:$PP1\gamma_{7-323}$ (K_d, 8.8 ± 0.3 nM). Measurements for each interaction repeated between two and four times.

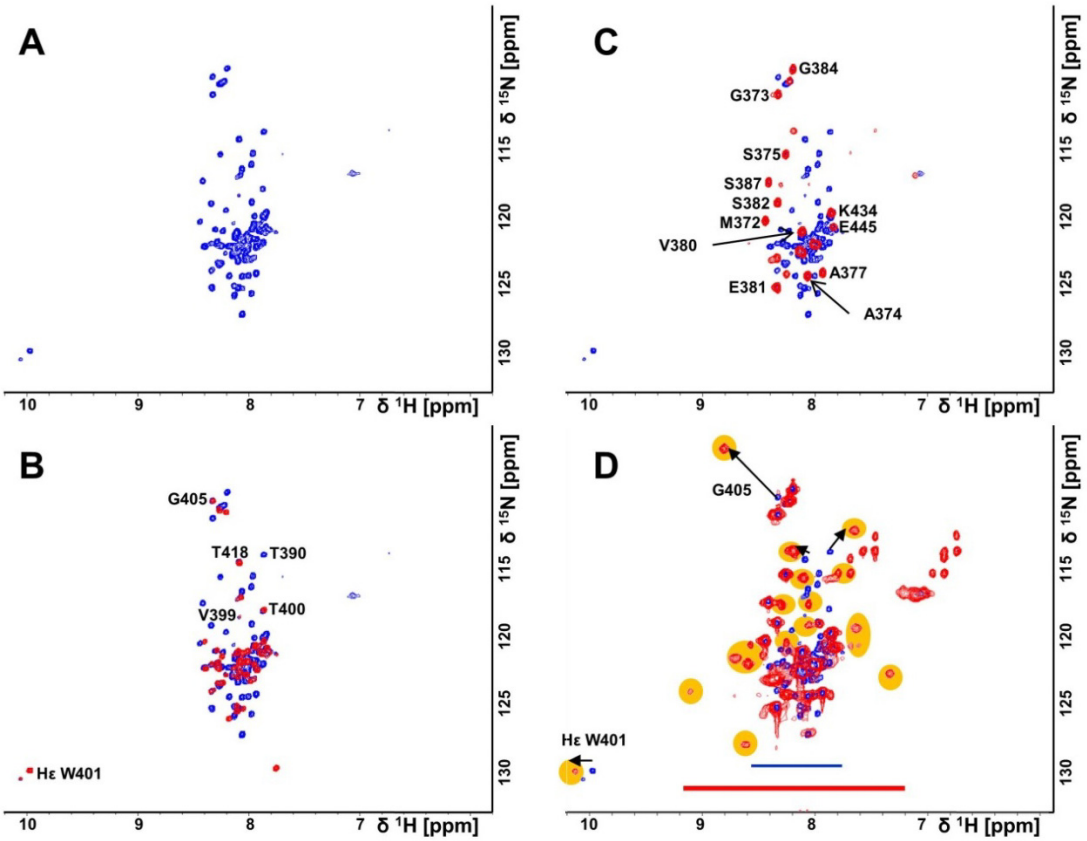

Figure S5. *Identifying the direct interaction of PNUTS with PP1 using NMR spectroscopy.* **A.** 2D [^{1}H,^{15}N] TROSY spectrum of 100 µM unbound [^{2}H,^{15}N] PNUTS$_{376-453}$ (20 mM Bis-Tris pH 6.5, 150 mM NaCl, 0.5 mM TCEP, 10% D_2O; 298 K; 850 MHz [^{1}H]); **B.** Overlay of the 2D [^{1}H,^{15}N] TROSY spectrum of unbound [^{2}H,^{15}N] PNUTS$_{376-453}$ (blue) and unbound [^{2}H,^{15}N] PNUTS$_{394-433}$ (red). Same conditions as in (A). A few residues are annotated; as expected, most peaks overlap perfectly. **C.** Overlay of the 2D [^{1}H,^{15}N] TROSY spectrum of unbound [^{2}H,^{15}N] PNUTS$_{376-453}$ (blue) and 250 µM [^{2}H,^{15}N] PNUTS$_{376-453}$:PP1α$_{7-330}$ complex (red). Same conditions as in (A). Only flexible, i.e. highly intense, peaks are shown and annotated from the [^{2}H,^{15}N] PNUTS$_{376-453}$:PP1α$_{7-330}$ complex sample – they all belong to residues that stay flexible upon complex formation with PP1 and belong to residues N-terminal to residue 390 and C-terminal to residue 424 (not all peaks are annotated for clarity). **D.** Same as (C) but now also weaker peaks are shown, while the intensity of the unbound peaks was held constant. Blue/red bars highlight the significant change (more than double) in ^{1}H chemical shift dispersion in PNUTS upon binding PP1. Yellow circles highlight peaks that belong to PNUTS residues when bound to PP1. A few arrows show the most likely changes. As expected, the sidechain Hε W401 (the 'F' of the RVxF-motif in PNUTS), which exists in two slowly exchanging conformations in the unbound state (2 peaks), is a single peak in the PP1 bound form, as it becomes deeply buried in the PP1 RVxF binding pocket.

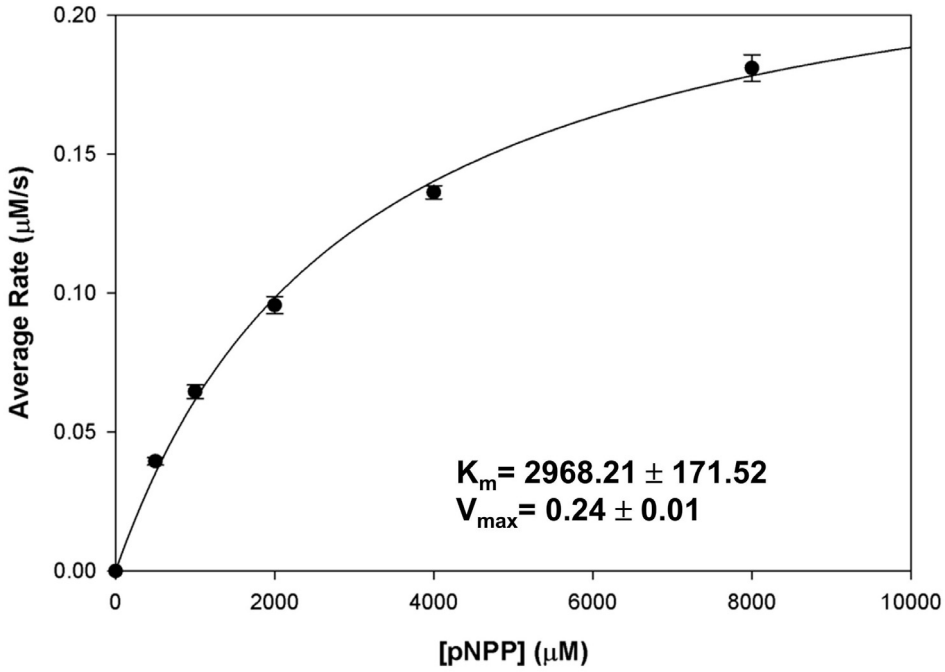

Figure S6. *The PNUTS:PP1 holoenzyme is catalytically active*. Dephosphorylation of the *p*-Nitrophenyl Phosphate (*p*NPP) substrate by the $PNUTS_{376-433}$:$PP1\alpha_{7-300}$ holoenzyme. Michaelis-Menten kinetic parameters were determined by measuring initial reaction rates at various *p*NPP concentrations.

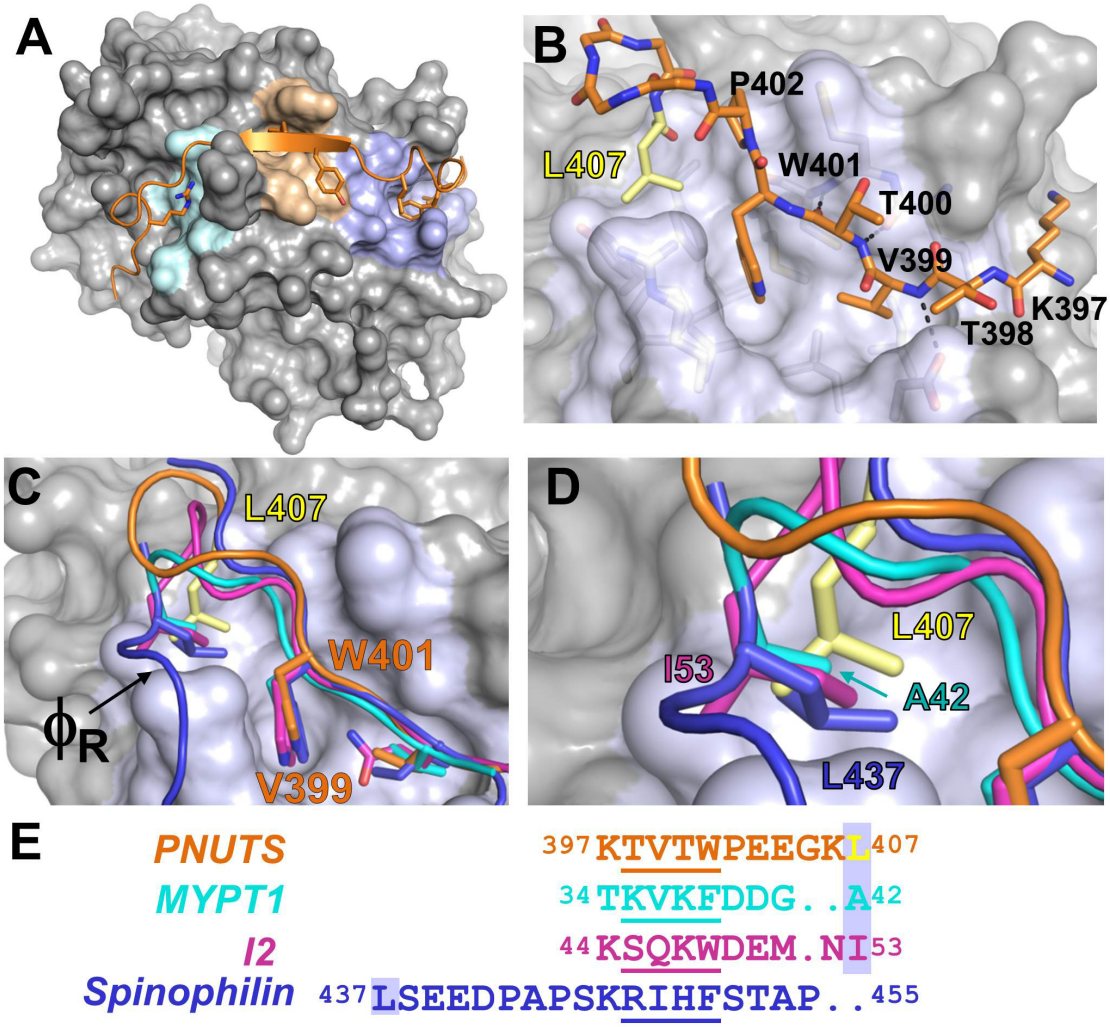

Figure S7. *Extension of the RVxF binding pocket.* (**A**) Cartoon representation of $PNUTS_{394-433}$ (orange) and $PP1\alpha_{7-300}$ (grey surface). PP1 binding pockets are shaded: RVxF (light blue), ΦΦ (orange), Arg (light green). (**B**) PNUTS residues ^{397}KTVTWPEEGKL^{407} bind the PP1 RVxF binding pocket. PNUTS residues are shown as orange sticks, except residue $Leu407_{PNUTS}$, which is yellow (side chains of $Glu403_{PNUTS}$, $Glu404_{PNUTS}$ and $Lys406_{PNUTS}$ omitted for clarity); PP1 residues illustrated as a transparent surface with PP1 residues that comprise the RVxF binding pocket shown as sticks and colored in light blue. (**C**) Overlay of the RVxF binding motifs from PNUTS (orange), MYPT1 (cyan), I-2 (pink) and spinophilin (blue). PP1 interacting protein residues that bind the extended RVxF pocket, which we have named the Φ_R pocket, are shown as sticks in the corresponding colors, except $Leu407_{PNUTS}$, in yellow; PNUTS residues are labeled. (**D**) Close-up of (C), with the residues that structurally overlap with $Leu407_{PNUTS}$ labeled. (**E**) Sequence alignment of the RVxF motifs shown in (C); residues that structurally overlap with $Leu407_{PNUTS}$ are shaded. The RVxF motif of each sequence is underlined.

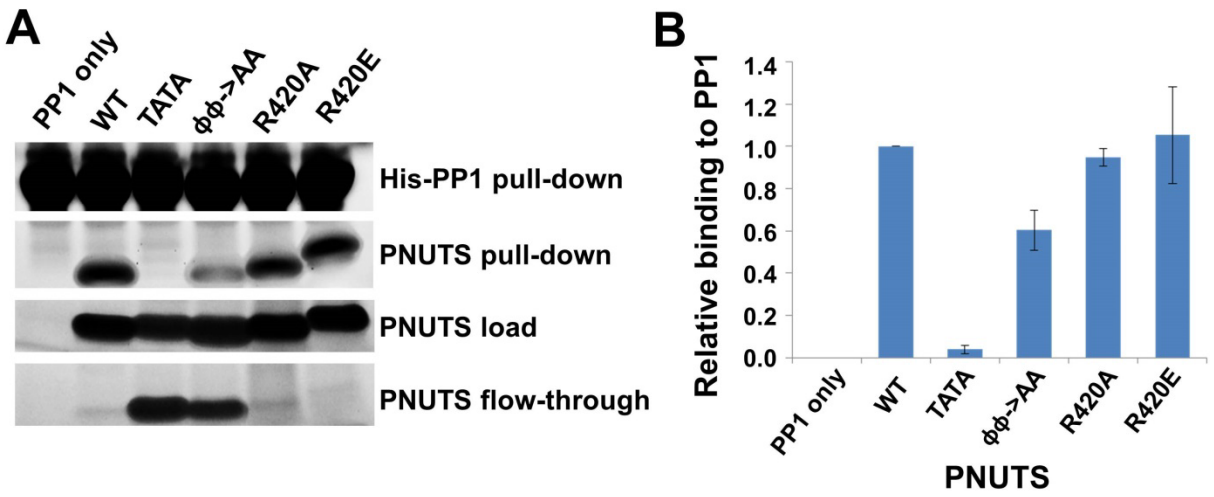

Figure S8. *PNUTS is anchored to PP1 via the RVxF and ΦΦ motifs.* (**A**) The binding of PNUTS WT and mutants (RVxF: TVTW→TATA; ΦΦ: YF→AA; Arg: R→A and R→E) to PP1α$_{7-330}$ was determined using a pull-down assay. Pull-down samples, loaded PNUTS samples and flow-through samples were subjected to SDS-PAGE, fixed, stained overnight with SYPRO Ruby protein gel stain and then scanned using Typhoon 9410. (**B**) Densitometry analysis of gels as shown in (**A**) to determine the amount of PNUTS pulled-down by PP1α$_{7-330}$, normalized to WT. The experiment was repeated three times. Error bars, s.d.

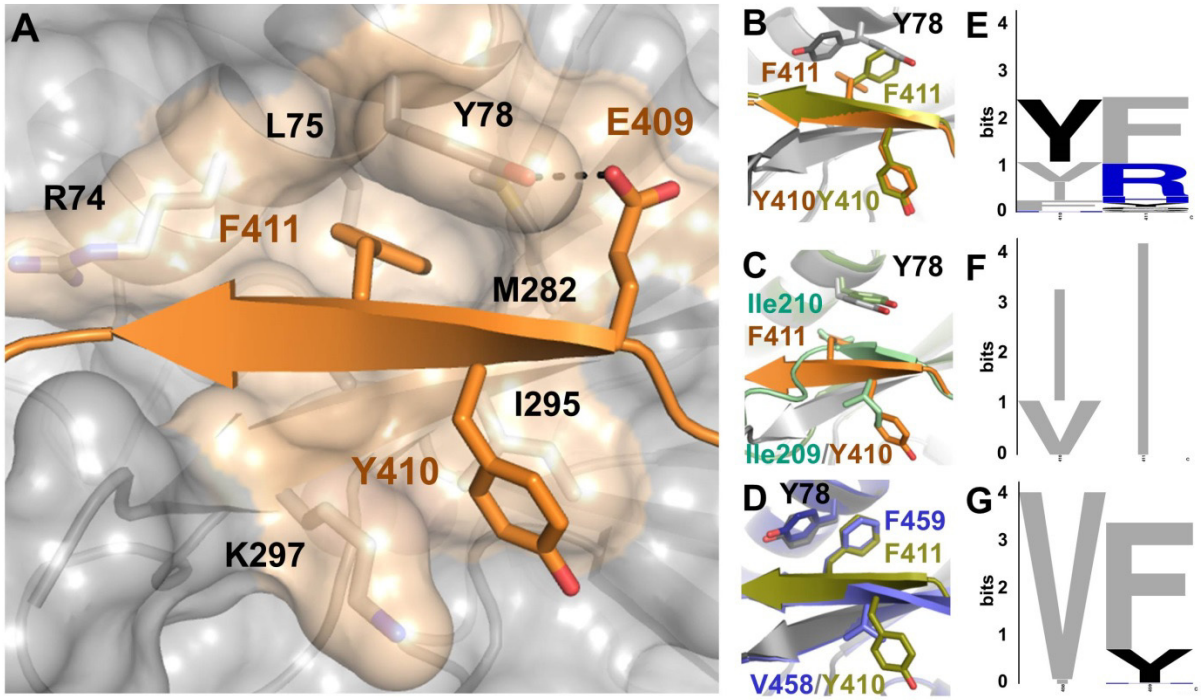

Figure S9. *Pliability of the ΦΦ motif binding pocket*. (**A**) The interactions between PNUTS (orange) and PP1 (gray surface) at the ΦΦ motif binding pocket. PP1 residues (beige) that interact directly with the PNUTS ΦΦ motif residues (orange) shown as sticks and labeled. (**B**) Overlay of the PNUTS:PP1 complexes from the $P3_221$ (orange/light grey) and $P4_12_12$ (yellow-green/dark grey) crystal forms. ΦΦ motif residues $Tyr410_{PNUTS}$, $Phe411_{PNUTS}$ and $Tyr78_{PP1}$ are shown as sticks and labeled. (**C**) Overlay of the PNUTS:PP1 ($P3_221$; same colors as in B) and NIPP1:PP1 (light green/green) complexes. ΦΦ motif residues (PNUTS, Tyr410/Phe411; NIPP1, Ile209/Ile210) and $Tyr78_{PP1}$ are shown. (**D**) Overlay of the PNUTS:PP1 ($P4_12_12$; same colors as in B) and spinophilin:PP1 (light blue/dark blue) complexes. ΦΦ motif residues (PNUTS, Tyr410/Phe411; spinophilin, Val458/Phe459) and $Tyr78_{PP1}$ are shown. Sequence logos depicting the sequence conservation of the ΦΦ motif in PNUTS (**E**), NIPP1 (**F**) and spinophilin (**G**). Homologs identified using JackHMMER; hydrophobic residues (gray), aromatic residues (black), positively charged residues (blue).

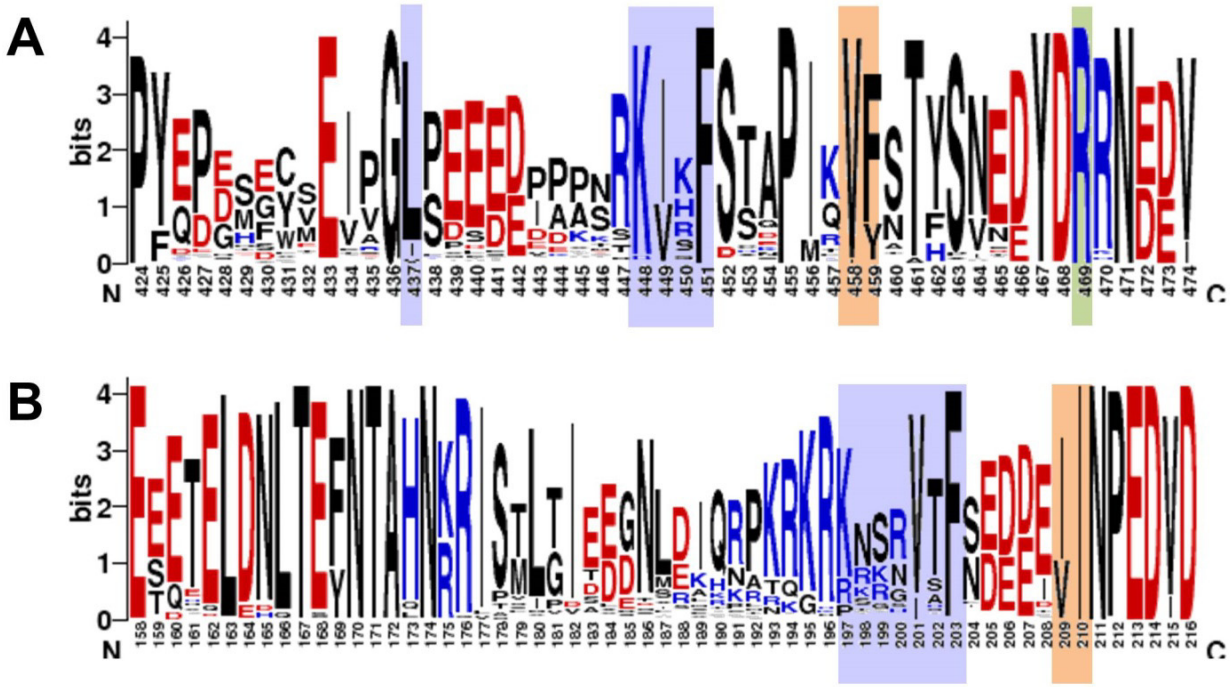

Figure S10. *The PNUTS ΦΦ motif is diverse.* Position weight matrix for aligned sequences identified using JackHMMER and plotted as a sequence logo. Positively charged residues blue, negatively charged residues red, hydrophobic residues black and the rest grey. Shaded regions highlight residues that bind in the RVxF pocket (light blue), the ΦΦ-pocket (beige) and the Arg-pocket (green). (**A**) Spinophilin/neurabin; (**B**) NIPP1 (PNUTS is shown in **Fig. 1E**).

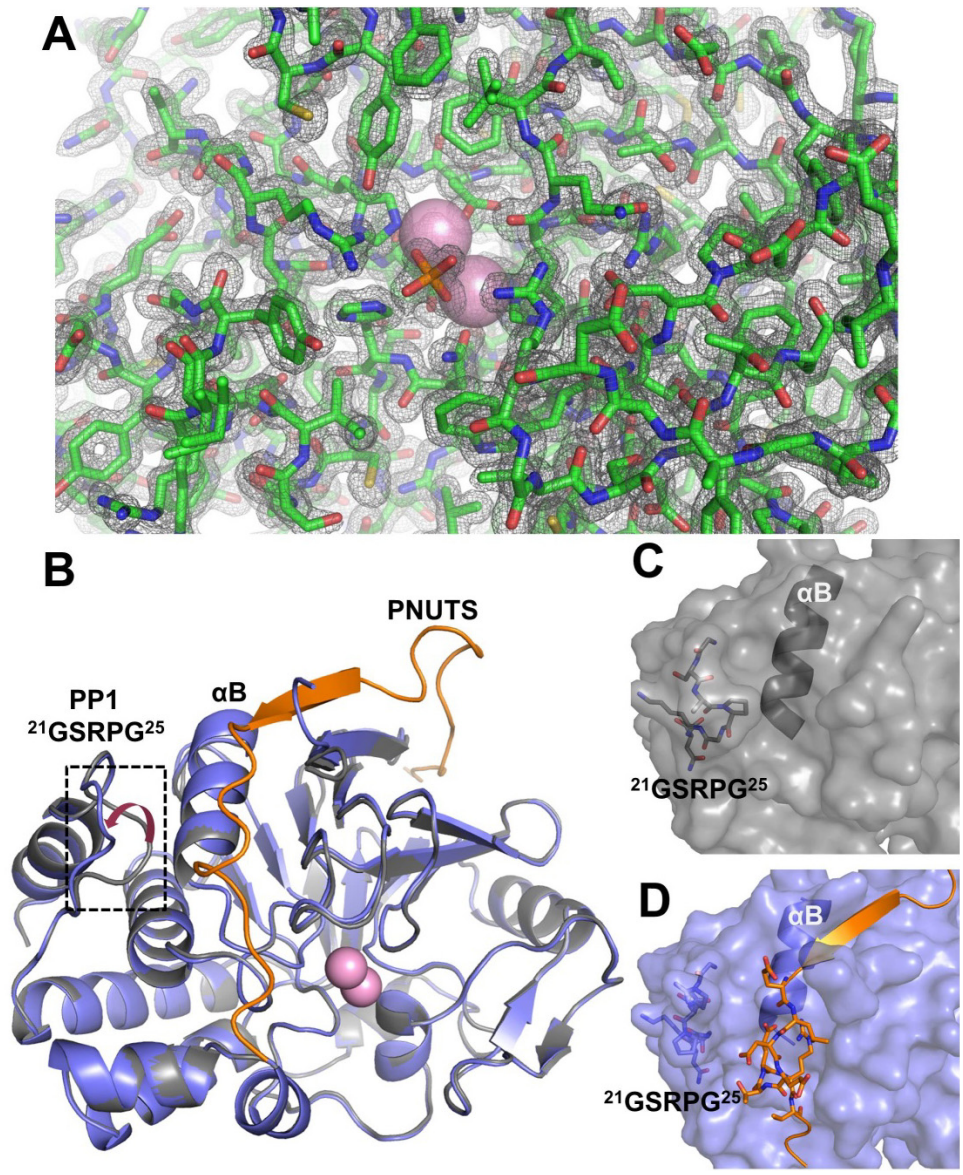

Figure S11. *Conformational changes of loop $^{21}GSRPG^{25}$ in PP1 bound to PNUTS compared to free PP1*. (**A**) High-resolution PP1 structure (1.45 Å; green sticks with electron density map represented by mesh) including a phosphate molecule (orange) and two Mn^{2+}-ions (pink spheres) at the active site. (**B**) 1.45 Å structure of free PP1 (grey) and the PNUTS:PP1 complex (purple) superimposed on PP1 showing the near perfect alignment between the two PP1 molecules. The primary exception is the $^{21}GSRPG^{25}$ loop (magenta arrow and boxed). (**C**) In free PP1, the $^{21}GSRPG^{25}$ loop is in a closed conformation. (**D**) In the PNUTS:PP1 complex, the $^{21}GSRPG^{25}$ loop shifted outward away from α-helix αB to accommodate PNUTS in the PP1 C-terminal groove.

Table S1. *Crystallographic data collection and refinement statistics.*

	PNUTS:PP1α$_{7-300}$	PNUTS:PP1α$_{7-300}$	PP1α$_{7-300}$
Data Collection			
Space group	$P3_221$	$P4_12_12$	$P2_12_12_1$
Cell dimensions			
a,b,c (Å)	130.8, 130.8, 47.7	92.4, 92.4, 199.3	65.7, 77.6, 133.0
α, β, γ (°)	90, 90, 120	90, 90, 90	90, 90, 90
Copies/UA[a]	1	2	2
Resolution (Å)	50.0 – 2.20 (2.24 – 2.20)	50.0 – 2.10 (2.14 – 2.10)	50.0 – 1.45 (1.48 – 1.45)
Unique reflections	24,217	51,308	120,264
R_{merge}	0.07 (0.66)	0.08 (0.62)	0.05 (0.27)
I/σI	19.3 (3.0)	22.5 (3.8)	18.5 (3.0)
Completeness (%)	100.0 (100.0)	100.0 (99.9)	99.4 (93.1)
Redundancy	9.2 (7.3)	11.7 (6.2)	5.8 (2.9)
Refinement			
Resolution (Å)	43.98 – 2.20	49.85 – 2.10	46.93 – 1.45
R_{work}[b]/R_{free}[c]	15.1/18.5	17.2/20.2	15.0/16.7
No. atoms			
Protein	2598	5071	4768
Water	159	315	621
Mn^{2+} Ions	2	4	4
B factor			
Protein	38.0	48.6	15.2
Water	41.1	49.8	28.7
Mn^{2+} Ions	32.1	29.7	11.5
rms deviations			
Bond length (Å)	0.010	0.003	0.013
Bond angle (°)	1.18	0.807	1.54
Ramachandran Plot[d]			
core (%)	95.4	96.1	96.1
allowed (%)	99.7	100.0	99.8
disallowed (%)[e]	0.3	0.0	0.2
PDB Code	4MOY	4MP0	4MOV

Highest resolution shell is shown in parentheses.

[a] Asymmetric unit.

[b] $R_{factor} = 100 \times \Sigma |F_P - F_{P(calc)}| / \Sigma F_P$.

[c] R_{free} was calculated from 5% of the data.

[d] Determined using MolProbity (18)

[e] Cys273

Table S2. *The ΦΦ binding residues are more highly conserved within, versus between, families of PP1 regulators*

Protein	# homologs	ΦΦ sequence[o]
NIPP1	**107**	II (72), VI (35)
Spinophilin	**125**	VF (96), VY (25), IF (2), IY (1), VH (1)
PNUTS	**85**	YF (47), VR (12), IR (11), FH (3), FF/FY (2), IQ/VY/FR/VS/VC/IN/HH/IH (1)

ºSequence variability in the ΦΦ motif was characterized using JackHMMER (7) using the sequences of the PP1 interacting residues from spinophilin (residues 424-474), NIPP1 (residues 158-216) and PNUTS (residues 396-424). The sequences identified by JackHMMER were edited to remove duplicates and/or incomplete sequences, resulting in 125, 107 and 85 independent sequences from a diversity of organisms for spinophilin/neurabin, NIPP1 and PNUTS, respectively. The residues that constitute the ΦΦ motifs in these families of PP1 interaction proteins are listed; the number of incidences of a particular ΦΦ sequence is indicated in parenthesis. For example, only two ΦΦ sequences were observed in the 107 NIPP1 sequences (II, VI), while 5 were observed in the 125 spinophilin/neurabin sequences (VF, VY, IF, IY, VH). The only ΦΦ sequence that was found in more than one regulator (VY) is underlined. Of the three PP1 regulatory proteins experimentally verified to contain a ΦΦ motif, NIPP1's is the most highly conserved and is found as only one of two sequences, II or VI. The spinophilin ΦΦ motif is also highly conserved, with 97% of the sequences being either VF or VY. In contrast, the PNUTS ΦΦ motif is considerably more variable, being a YF, VR or IR in 82% of the sequences, but also present as FH, FF, FY, IQ, VY, FR, VS, VC, IN, HH, IH. Together, this analysis suggests that within a PP1 regulatory family, the ΦΦ motif is largely conserved and that while the residues found most often in the Φ_A (Val, Ile, Tyr) and Φ_B (Phe, Ile, Tyr) positions are hydrophobic, other types of amino acids can be accommodated at these positions.

Table S3. *PP1 interacting proteins predicted to contain a functional RVxF-ΦΦ motif*

Gene	Protein	UniProt ID	Sequence[1]	IDP[2]
Confirmed RVxF-X_{5-8}-ΦΦ motif				
PPP1R10	PNUTS	Q96QC0	KSVTWPEEGKLREYFYFELDETER	Y
PPP1R9A	Neurabin	Q9ULJ8	RKIKFSSAPIKVFNTYSNEDYDR	Y
PPP1R9B	Spinophilin	Q96SB3	RKIHFSTAPIQVFSTYSNEDYDR	Y
PPP1R8	NIPP1	Q12972	SRVTFSEDDEIINPEDVDPSVG	Y
Potential RVxF-X_{5-8}-ΦΦ motif				
AATK	LMTK1	Q6ZMQ8	KAVSFFDDVTVYLFDQESPTR	L
AKAP1	AKAP149	Q92667	KGVLFSSKSAEVCKQDSPFSRVP	L
AKAP11	AKAP220	Q9UKA4	KKVQFAEALATHILSLATEMAAS	A
ANKRD42	SARP	Q8N9B4	KKVHFGSIHDAVRAGDVKQLSEI	U
APC	APC	P25054	KHVSFTPGDDMPRVYCVEGTPINFS	L
CASC5	KNL1	Q8NG31	RRVSFADTIKVFQTESHMKIVR	L
CASP2[3]	Caspase-2	Q9IB67	SKVHHGSFPLPVQESTLSRPGR	L
CCDC8	CCDC8	Q9H0W5	KTVRFQTPGRFSWFCKRRRAF	L
CDCA2	Repo-Man	Q69YH5	KRVTFGEDLSPEVFDESLPANTPL	L
CEP192	CEP192	Q8TEP8	KHVTFENHRIVSPKNSDLKNTS	L
CHCHD6	CHCHD6	Q9BRQ6	RRVSFGVDEEERVRVLQGVRLSEN	L
CSMD1	CSMD1	Q96PZ7	KAVRFDTTLNTVCTVV	L
CSRNP2	CSRNP2	Q9H175	KNVRFDQVTVYYFARRQGFTSVP	L
CSRNP3	CSRNP3	Q8WYN3	KNVHFSCVTVYYFTRRQGFTSVP	L
DDX31	DEAD box protein 31	Q9H8H2	RRVSWAKKALQSFIQAYATYPRELK	A
DLG2	Chapsyn-110	Q15700	KTVKFNAKPGVIDSKGSFNDKR	L
FARP1	FERM	Q9Y4F1	KKVQFERKHSKIHSIRSLASQPTE	L
FKBP15	FK506BP15	Q5T1M5	RRVKFARDSGSDGHSVSSRDSAAPS	L
GPATCH2	GPATCH2	Q9NW75	RMVHFSPDSHHHDHWFSPGAR	L
GRXCR1	glutaredoxin	A8MXD5	RKVRFRIASSHSGRVLKEVYED	L
IKZF1	Ikaros	Q13422	CRVLFLDHVMYTIHMGCHGFRDPF	L
ITPR1	IP3R1	Q14643	KTVTFEEHIKEEHNMWHYLCFIVL	L
ITPR3	IP3R3	Q14573	KTVSFEEHIKLEHNMWNYLYFIVL	L
KCNA6	KCNA6	P17658	RRVRFFDPLRNEYFFDRNRPSFDAI	A
LMTK2	KPI-2	Q8IWU2	KAVTFFDDVTVYLFDQETPTKE	L
LMTK3	LMTK3	Q96Q04	KMVSFHGDVTVYLFDQETPTNE	L
MAP1B	MAPB1	P46821	RSVNFSLTPNEIKVSAEAEVAPVSP	L
MCM7	MCM7	P33993	RSVRFSEAEQRCVSRGFTPAQFQA	A
NONO	P54nrb	Q15233	RVRFACHSASLTVRNLPQYVSNE	U
OPN3	Opsin-3	Q9H1Y3	KKVTFNSSSIIFIITSDESLSVDD	L
PCIF1	PCIF1	Q9H4Z3	KVVKWNVEDTFSWLRKDHSASK	A
PLCL1	PRIP-I	Q15111	KTVSFSSMPSEKKISSANDCISFMQ	A
PPP1R3A	GM	Q16821	RRVSFADSFGFNLVSVKEFDCWELP	L

PPP1R15A	GADD34	O75807	RKVRFSEKVTVHFLAVWAGPAQA	L
PPP1R15B	FLJ14744/CReP	Q5SWA1	KKVTFLEEVTEYYISGDEDRKGPWE	L
PPP1R18	Phostensin	Q6NYC8	LKISFSETALETTYQYPSESSVLEE	L
PPP1R26/ KIAA0649	1A6/DRIM BP	Q5T8A7	KKVRFSTAQTHFLEQLGGLRR	L
PPP1R27/ DYSFIP1	DYSFIP1	Q86WC6	RSVRFPNDVLFLDHIRQGDLEQVGR	U
PPP1R35	LOC221908	Q8TAP8	RQVRFRLTPPSPVRSEPQPAVPQE	L
PPP1R37/ LRRC68	LLRC68	O75864	KRVTFPSDEDIVSGAVEPKDPWR	A
RPGRIP1L	RPGRIP1L	Q68CZ1	KKVSFVDIMPHQSDETSPPPED	L
SFPQ	PSF	Q86VG2	LRVRFATHAAALSVRNLSPYVSNEL	U
SLC7A14	Solute Carrier 7A14	Q8TBB6	RRVQWGAAWYAMHSRILRTKPVES	L
SMARCB1	SNF5	Q12824	KPVKFQLEDDGEFYMIGSEVGNYL	A
SPOCD1	SPOCD1	Q6ZMY3	KMVSFNSKVEKRYYQPDDRRPNVPL	L
SPRED1	SPRED1	Q7Z699	RHVSFQDEDEIVRINPRDILIR	L
TMEM132D	TMEM132D	Q14C87	KRVKFTTFTAVSSDDEYPTRNS	L
TNS1	Tensin 1	Q9HBL0	GKVEFVFSYGPEKIQGMEHLENGPS	L
TRIM42	Trim42	Q8IWZ5	RKVTFSTHSLGNQHIYQRSSSMLSF	L
TSKS	TSKS	Q9UJT2	KAVSFHGVEPQMSHQPMHWCLNLKR	L
WWC1	KIBRA	Q8IX03	KVEFLLLEGATGFRPSGCITTIHEDE	L
ZFYVE1	ZFYVE1	Q9HBF4	KVVSFLLVDENEEIQVTNEEDFIRK	U

[1]Sequence that corresponds to the potential RVxF-ΦΦ motif
[2]IDP, intrinsically disordered protein; Y, IDP demonstrated experimentally; L, likely an IDP based on PSI-PRED and IUPRED predictions; U, unlikely an IDP based on PSI-PRED and IUPRED predictions; A, ambiguous whether it is an IDP based on PSI-PRED and IUPRED predictions
[3]Xenopus Casp2

Table S4. *PP1 interacting proteins predicted to contain a functional RVxF-ΦΦ-R motif*

Gene	Protein	UniProt ID	RVxF-ΦΦ-R Sequence[1]	IDP[2]	Arg cons[3]	Likelihood
Confirmed RVxF-X_{5-8}-ΦΦ-X_{8-9}-R motif						
PPP1R10	PNUTS	Q96QC0	KSVTWPEEGKLREYFYFELDETER	Y	R	**Confirmed experimentally**
PPP1R9A	Neurabin	Q9ULJ8	RKIKFSSAPIKVFNTYSNEDYDR	Y	R	**Confirmed experimentally**
PPP1R9B	Spinophilin	Q96SB3	RKIHFSTAPIQVFSTYSNEDYDR	Y	R	**Confirmed experimentally**
Potential RVxF-X_{5-8}-ΦΦ-X_{8-9}-R motif[4]						
AATK	LMTK1	Q6ZMQ8	KAVSFFDDVTVYLFDQESPTR	L	R	**Unclear;** although 'R' conserved in LMTK1, it is a 'K' in LMTK2 and an 'N' in LMTK3
CASC5	KNL1	Q8NG31	RRVSFADTIKVFQTESHMKIVR	L	R/K	**Somewhat Likely;** 'K' can likely substitute for 'R' in the PP1 Arg motif binding pocket
CASP2	Caspase-2	Q9IB67	SKVHHGSFPLPVQESTLSRPGR	L	R	**Unclear;** RVxF motif only present in Xenopus
DLG2	Chapsyn-110	Q15700	KTVKFNAKPGVIDSKGSFNDKR	L	R/G	**Unclear;** 'R' somewhat conserved but some organisms (i.e., rat) have a 'G'
GPATCH2	GPATCH2	Q9NW75	RMVHFSPDSHHHDHWFSPGAR	L	R	**Likely;** IDP and 'R' conserved
PPP1R26/ KIAA0649	1A6/DRIM BP	Q5T8A7	KKVRFSTAQTHFLEQLGGLRR	L	R/H/Q/ GK	**Not likely;** 'R' not conserved
PPP1R27/ DYSFIP1	DYSFIP1	Q86WC6	RSVRFPNDVLFLDHIRQGDLEQVGR	U	R	**Not likely;** 'R' predicted to be in a structured domain
PPP1R37/ LRRC68	LLRC68	O75864	KRVTFPSDEDIVSGAVEPKDPWR	A	R/K/I/Q /E/F/S	**Not likely;** 'R' not conserved and may be structured
SPRED1	SPRED1	Q7Z699	RHVSFQDEDEIVRINPRDILIR	L	R	**Unclear;** although 'R' conserved in SPRED1, it is a 'T' in SPRED2
TSKS	TSKS	Q9UJT2	KAVSFHGVEPQMSHQPMHWCLNLKR	L	R	**Likely;** IDP and 'R' conserved
ZFYVE1	SFYVE1	Q9HBF4	KVVSFLLVDENEEIQVTNEEDFIR	U	R/K/E/ S	**Not likely;** 'R' not conserved and likely structured

[1]Sequence that corresponds to the potential RVxF-ΦΦ-R motif
[2]IDP, intrinsically disordered protein; Y, IDP demonstrated experimentally; L, likely an IDP based on PSI-PRED and IUPRED predictions; U, unlikely an IDP based on PSI-PRED and IUPRED predictions; A, ambiguous whether it is an IDP based on PSI-PRED and IUPRED predictions.
[3]Conservation of the Arg motif arginine analyzed using BLAST, PHMMER and homologene.

Supplemental References

1. Kelker MS, Page R, & Peti W (2009) Crystal structures of protein phosphatase-1 bound to nodularin-R and tautomycin: a novel scaffold for structure-based drug design of serine/threonine phosphatase inhibitors. *J Mol Biol* 385(1):11-21.
2. Peti W & Page R (2007) Strategies to maximize heterologous protein expression in Escherichia coli with minimal cost. *Protein Expr Purif* 51(1):1-10.
3. Keller S*, et al.* (2012) High-precision isothermal titration calorimetry with automated peak-shape analysis. *Analytical chemistry* 84(11):5066-5073.
4. Houtman JC*, et al.* (2007) Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Protein science : a publication of the Protein Society* 16(1):30-42.
5. Adams PD*, et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 2):213-221.
6. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 4):486-501.
7. Johnson LS, Eddy SR, & Portugaly E (2010) Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC bioinformatics* 11:431.
8. Crooks GE, Hon G, Chandonia JM, & Brenner SE (2004) WebLogo: a sequence logo generator. *Genome research* 14(6):1188-1190.
9. de Castro E*, et al.* (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic acids research* 34(Web Server issue):W362-365.
10. Heroes E*, et al.* (2013) The PP1 binding code: a molecular-lego strategy that governs specificity. *The FEBS journal* 280(2):584-595.
11. Finn RD, Clements J, & Eddy SR (2011) HMMER web server: interactive sequence similarity searching. *Nucleic acids research* 39(Web Server issue):W29-37.
12. Camacho C*, et al.* (2009) BLAST+: architecture and applications. *BMC bioinformatics* 10:421.
13. Sayers EW*, et al.* (2012) Database resources of the National Center for Biotechnology Information. *Nucleic acids research* 40(Database issue):D13-25.
14. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 292(2):195-202.
15. Dosztanyi Z, Csizmok V, Tompa P, & Simon I (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21(16):3433-3434.
16. Dosztanyi Z, Csizmok V, Tompa P, & Simon I (2005) The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J Mol Biol* 347(4):827-839.
17. Kim YM*, et al.* (2003) PNUTS, a protein phosphatase 1 (PP1) nuclear targeting subunit. Characterization of its PP1- and RNA-binding domains and regulation by phosphorylation. *J Biol Chem* 278(16):13819-13828.

18. Chen VB*, et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 1):12-21.