

Assembly and Molecular Architecture of the Phosphoinositide 3-Kinase p85 α homodimer

Jaclyn LoPiccolo^{1,§}, Seung Joong Kim^{4,§}, Yi Shi⁵, Bin Wu³, Haiyan Wu¹, Brian T. Chait⁵, Robert H. Singer³, Andrej Sali⁴, Michael Brenowitz², Anne R. Bresnick², and Jonathan M. Backer^{1,2}

¹Department of Molecular Pharmacology, ²Department of Biochemistry, ³Department of Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, NY 10461, USA

⁴Department of Bioengineering and Therapeutic Sciences and Department of Pharmaceutical Chemistry, California Institute for Quantitative Biosciences, Byers Hall, 1700 4th Street, Suite 503B, University of California, San Francisco, San Francisco, CA 94158, USA

⁵Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, NY 10065, USA

Running title: *Assembly of the p85 α homodimer*

[§]These authors contributed equally to this work.

To whom correspondence should be addressed: Jonathan M. Backer, MD, Department of Molecular Pharmacology, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA, Tel: 718-430-2153; E-mail: jonathan.backer@einstein.yu.edu; Anne R. Bresnick, PhD., Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA, Tel: 718 430-2741; Email: anne.bresnick@einstein.yu.edu; Michael Brenowitz, Ph.D., Department of Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY, USA, 10461, Tel: 718 430-3179, Email: michael.brenowitz@einstein.yu.edu.

Keywords: phosphoinositide 3-kinase (PI 3-kinase), phosphatidylinositol signaling, analytical ultracentrifugation, small-angle X-ray scattering (SAXS), molecular modeling, structural model

Background: The Class IA PI 3-kinase regulatory subunit p85 α dimerizes by mechanisms not fully understood.

Results: p85 α dimerization is driven by intermolecular interactions at both its N- and C-termini. A structural model of the p85 α dimer reveals conformational diversity.

Conclusion: p85 α dimerization is poised to link the heterologous assemblies that regulate PI3K signaling.

Significance: The p85 α dimer is a dynamic participant in PI3K signaling.

ABSTRACT

Phosphoinositide 3-kinases (PI3Ks) are a family of lipid kinases that are activated by growth factor and G-protein coupled receptors, and propagate

intracellular signals for growth, survival, proliferation, and metabolism. p85 α , a modular protein consisting of 5 domains, binds and inhibits the enzymatic activity of Class IA PI3Ks. Here, we describe the structural states of the p85 α dimer, based on data from *in vivo* and *in vitro* solution characterization. Our *in vitro* assembly and structural analyses have been enabled by the creation of cysteine-free p85 α that is functionally equivalent to native p85 α . Analytical ultracentrifugation (AUC) studies showed that p85 α undergoes rapidly reversible monomer-dimer assembly that is highly exothermic in nature. In addition to the documented SH3-PR1 dimerization interaction, we identified a second intermolecular interaction mediated by cSH2 domains at the C-terminal end of the polypeptide. We have

demonstrated *in vivo* concentration-dependent dimerization of p85 α using fluorescence fluctuation spectroscopy (FFS). Finally, we have defined solution conditions under which the protein is predominantly monomeric or dimeric, providing the basis for small angle X-ray scattering (SAXS) and chemical cross-linking structural analysis of the discrete dimer. These experimental data have been used for the integrative structure determination of the p85 α dimer. Our study provides new insight into the structure and assembly of the p85 α homodimer and suggests that this protein is a highly dynamic molecule whose conformational flexibility allows it to transiently associate with multiple binding proteins.

p85 α (PIK3R1) is one of 5 regulatory subunits that inhibit and stabilize the p110 catalytic subunits of Class IA PI3 kinases (PI3K)(1). The p85 α protein consists of 5 domains (Fig. 1): an N-terminal Src-homology 3 (SH3) domain, a BCR homology Rac/Cdc42-binding domain flanked by two proline-rich motifs (PR1 and PR2), and two Src-homology 2 domains (nSH2 and cSH2) that flank the coiled-coil inter-SH2 domain (iSH2). The iSH2 domain binds to the N-terminal adaptor-binding domain (ABD) of p110 α , p110 β , and p110 δ , yielding a stable low-activity heterodimer. Activation of p85/p110 heterodimers by receptor tyrosine kinases (RTKs) requires binding of the nSH2 and cSH2 domains to phosphorylated YXXM motifs in the RTKs themselves, or in adaptors like IRS-1, which relieves inhibitory contacts between the SH2 domains and p110 (2). In cells, activated p85/p110 heterodimers phosphorylate PI(4,5)P₂ to generate PI(3,4,5)P₃. The lipid phosphatase PTEN antagonizes PI3-kinase signaling by dephosphorylating PI(3,4,5)P₃ (3).

Given their modular structure, it is not surprising that p85 subunits participate in intra- and intermolecular interactions in addition to binding p110. p85 α /p110 dimers are activated by binding of the p85 α BCR domain to GTP-bound Rac (4) and Cdc42 (5) as well as the binding of proline-rich domains to SH3 domains from Src-family kinases (6). The binding of the influenza protein NS1 similarly activates p85 β /p110 dimers (7). The iSH2 domains of both p85 α and p85 β bind the tumor suppressor BRD7, resulting in

nuclear translocation of p85 α and its sequestration from cytosolic p110 (8). Independently of p110, the N-terminus of p85 α binds PTEN, protecting the latter from degradation and negatively regulating PIP3 production in cells (9-11).

The proline-rich motifs that flank the BCR domain both contain consensus SH3-binding sequences (12). Dimerization of p85 α as well as the SH3-PR1-BCR fragment of p85 α (residues 1-333; Fig. 1B) has been reported, and peptides derived from the PR1 motif disrupt p85 α dimerization (13). These results show that intermolecular SH3-PR1 interactions in the native protein are involved in p85 α dimerization. *In vivo* dimerization has been demonstrated by reciprocal immunoprecipitation of differentially epitope-tagged p85 α in several cell lines (11). A p85 α mutation, identified in a human endometrial carcinoma, truncates p85 α midway through the BCR domain at residue 160; expression of this mutant activates PI3K signaling by inhibiting homodimerization of endogenous p85 α , thereby blocking its stabilization of PTEN. The pathological consequence of this mutation highlights the biological importance of the p85 α dimer (11).

Because p85 α plays a central regulatory role in the PI3K signaling pathway, characterization of its structure and assembly dynamics is crucial to understanding its function in both normal physiology and disease. Our study combines *in vivo* and *in vitro* solution characterization of p85 α dimerization with integrative multi-state modeling of its global architecture using complementary analytical approaches (14-20). We have characterized the reversible monomer-dimer equilibrium of p85 α , showing that dimerization is mediated by multiple domain contacts. Further, we have defined solution conditions under which the protein is monomeric or predominantly dimeric, and used these conditions for small angle X-ray scattering (SAXS) and chemical cross-linking studies that have informed the structural modeling of the p85 α dimer. Our study provides new insight into p85 α dimerization, suggesting that p85 α is a highly dynamic molecule whose conformational flexibility allows it to efficiently exchange among multiple binding partners.

EXPERIMENTAL PROCEDURES

Expression and purification of cysteine-free p85 α —Wild-type human p85 α was cloned into the pGEX-6P-1 bacterial expression vector (GE Healthcare) using the BamHI-EcoRI sites. Its six cysteines were mutated (C146S, C167S, C498S, C656S, C659V, and C670L) to generate cysteine-free p85 α . Hereafter, we refer to ‘cysteine-free p85 α ’ as p85 α and the wild type protein as ‘native p85 α ’. Similarly, truncation mutants will be referred to as p85 α with superscripts denoting the residues included in the fragment. The construct coding for p85 α was expressed in BL21-CodonPlus competent cells (Agilent Technologies), which were induced overnight with 0.4 mM IPTG (isopropyl β -D-1-thiogalactopyranoside) at 25 °C. The cells were harvested by centrifugation and the pellets resuspended on ice in lysis buffer (PBS containing 4 mM DTT, 2 mM EDTA, 2 mM PMSF, 2.5 U/ml Pierce universal nuclease for cell lysis (Thermo Scientific) and Roche cOmplete, Mini Protease Inhibitor Tablets (Roche Diagnostics)).

The cells were lysed by sonication in an ice bath using a Branson Sonicator with a microprobe tip at output level 5 for 30 s followed by 30 s on ice, for 5 cycles. The lysate was brought to 1% Triton-X, incubated for 20 min at 4 °C on a rotating wheel, and centrifuged at 15,000 rpm in a SS-34 rotor for 20 min at 4 °C. The supernatant was incubated with Pierce glutathione agarose (Thermo Scientific) for 2 - 4 h at 4 °C on a rotating wheel. The resin was washed 3 times by resuspension in 10 column volumes of 50 mM Tris, 150 mM NaCl, pH 8.0. p85 α was cleaved from the GST with PreScission protease (GE Healthcare) overnight at 4 °C in cleavage buffer (50 mM Tris, 150 mM NaCl, 1 mM EDTA, 1 mM DTT at pH 8.0) on a rotating wheel. The resin was transferred to a chromatography column and supernatant containing the cleaved protein was collected, along with 2 washes of 1 column volume cleavage buffer each. The PreScission cleavage reaction leaves five residues (GPLGS) at the N-terminus preceding the p85 α sequence.

The resultant p85 α was dialyzed overnight into Mono Q buffer (20 mM Tris, 20 mM NaCl, pH 8.0), and loaded onto a Mono Q 10/100 GL anion exchange column (GE Healthcare) and eluted with a 0 - 350 mM NaCl gradient over 40-column

volumes. The peak fractions were analyzed by SDS-PAGE, pooled, concentrated and loaded onto a HiLoad 26/60 Superdex 200 prep grade gel filtration column (GE Healthcare) equilibrated in gel filtration buffer (50 mM Tris, 300 mM NaCl, pH 8.0). Fractions were analyzed by SDS-PAGE and those containing >95% pure p85 α were pooled and concentrated for use or storage at -80 °C. Stored protein was thawed on ice and centrifuged using a TLA-120.2 rotor (Beckman Coulter) at 80,000 rpm for 15 min at 4 °C prior to use. Protein concentrations were measured using UV absorbance at 280 nm (Nanodrop 2000 UV-Vis Spectrophotometer, Thermo Scientific) and corresponding extinction coefficients calculated from protein sequences using ExPASy ProtParam (<http://web.expasy.org/protparam/>). Protein mass was confirmed via mass spectrometry to be 83,938.6 Da (based on the sequence the calculated mass is 83,951.6 Da).

Truncated fragments of p85 α (Fig. 1) were expressed from the corresponding cDNAs that were synthesized by PCR and ligated into pGEX-6P-1 using the BamHI-EcoRI sites. All constructs were verified by sequencing. The truncated fragments of p85 α were expressed and purified using the same protocol described for the full-length protein. For binding assays, native GST-p85 α and GST-p85 α were purified as described above except that p85 α was not cleaved from the glutathione agarose. SDS-PAGE and Coomassie staining were used to quantitate bead-bound GST fusion proteins. Beads were either stored at 4 °C for up to 1 week or frozen in 50% glycerol at -20°C.

In vitro binding assays—HEK293T cells were transfected with wild-type N-myc-p110 α using Fugene HD for 48 hr. Cells were lysed on ice in lysis buffer [20 mM Tris HCl, (pH 8.1), 137 mM NaCl, 1 mM MgCl₂, 1 mM CaCl₂, 10% (v/v) glycerol, Nonidet P-40; 150, M vanadate; 1 mM phenylmethylsulfonyl fluoride, Roche tablet, and phosphatase cocktails (Sigma)], followed by incubation on a rotating wheel for 20 min at 4 °C and centrifugation at 13,000 rpm for 10 min. The clarified supernatant was then incubated with rotation for 2 hr at 4 °C with glutathione Sepharose beads or beads complexed with GST-p85 α or native GST-p85 α . The beads were washed

3 times with PBS containing 1% NP-40, once with PBS, and boiled in 2X Laemmli sample buffer for analysis by SDS-PAGE. Membranes were blotted with an N-myc antibody (Cell Signaling Technologies) and developed with ECL western blotting substrate (Pierce).

For Rac binding, GST-Rac was bacterially purified and bound to glutathione Sepharose, followed by loading without or with GTP γ S (21) and incubated with rotation for 2 hr at 4 °C with p85 α ¹⁻⁴³². The glutathione Sepharose beads were washed 3 times with PBS containing 1% NP-40, once with PBS, and boiled in 2X Laemmli sample buffer for analysis by SDS-PAGE. Membranes were blotted with in-house anti-p85 α -nSH2 antibody and developed with ECL western blotting substrate (Pierce).

Lipid kinase assay—N-myc-p110 α was immunopurified from transiently transfected HEK293T cells as above. Protein-G pellets were washed, incubated with bacterially purified GST-p85 α or native GST-p85 α , and assayed for lipid kinase activity as previously described (22).

SH2-phosphopeptide binding—GST-p85 α -cSH2 and native GST-p85 α -cSH2 constructs (residues 617-724) were expressed in BL-21 *E. coli*, processed as above, and partially purified by elution with 20 mM glutathione from glutathione Sepharose beads. A tyrosine phosphopeptide containing the photo-activatable amino acid Benzoyl Phenylalanine (Bpa) (Gly-Asp-Gly-pTyr-BPA-Pro-Met-Ser-Pro-Lys-Ser) was N-terminally labeled with [¹²⁵I]Bolton-Hunter reagent and desalted by chromatography on Sephadex G-10. The labeled peptide (4.7 μ M final, 3 x 10⁶ CPM/assay) was incubated with 2 μ g of GST-cSH2 domain in the absence or presence of 250 μ M unlabeled peptide. The samples were irradiated on ice with a 200 nm UV lamp at a distance of 1 cm for 1 hr, boiled in Laemmli sample buffer and analyzed by SDS-PAGE and autoradiography.

Analytical ultracentrifugation (AUC)—AUC studies were conducted with a Beckman Optima XL-I centrifuge using the AN-60Ti rotor and the absorption optics set to 280 nm. Protein samples

were run in either the low or high salt SAXS buffers described below at temperatures ranging from 4 to 37 °C. An SH3-binding peptide (RPLPPRPGA) used to inhibit dimerization was synthesized by GenScript and kept at -20 °C as a concentrated stock solution that was thawed and diluted into the buffer appropriate for each experiment. Sednterp version 20120828 Beta (<http://sednterp.unh.edu>) was used to calculate the partial specific volume of the proteins from their sequence and the density and viscosity of the buffers. The sedimentation parameters were corrected to standard conditions ($2\theta, w$) using these values.

For sedimentation velocity (SV) experiments, 300 μ l of sample and an equal volume of buffer were loaded into two-sector cell assemblies. Three concentrations of each protein corresponding to $A_{280} \approx 0.1, 0.4, \text{ and } 1.0$ (equivalent to 1.2, 3.6, and 9.6 μ M p85 α monomer) were analyzed for each solution condition. Sixty to 70 scans were collected over the course of a sedimentation velocity run. A subset of scans, beginning with those where a clear plateau was evident between the meniscus and the boundary, were selected for time-derivative analysis using DCDDT+ *ver.* 2.4.2 (23,24). The SV experiments presented in this paper were conducted at 42,000 rpm. For sedimentation equilibrium (SEQ) experiments, 120 μ l each of sample and buffer were loaded into six-channel cell assemblies at the concentrations noted above. Data was taken following initial equilibration for 24 hr at the indicated speed and then following a second 24 hr equilibration at higher speed. Scans were also taken after 22 hr so that equilibration could be confirmed. The p85 α EQ experiments were equilibrated at 8,000 and then 16,000 rpm; p85 α ¹⁻⁶⁰⁰ was equilibrated at 9,000 and then 16,000 rpm; p85 α ¹⁻⁴³² was equilibrated at 10,000 and then 20,000 rpm; p85 α ¹⁻³³³ was equilibrated at 18,000 and then 25,000 rpm; p85 α ⁷⁸⁻³²² was equilibrated at 20,000 and then 30,000 rpm. The equilibrium experiments were globally analyzed using HeteroAnalysis *ver.* 1.1.58 (25,26).

Fluorescence fluctuation spectroscopy (FFS)—p85 α and native p85 α were cloned into pEGFP-N1 using the EcoRI-BamHI sites. Native p85 α was also used as a template for site-directed

mutagenesis of proline residues 96 and 99 to alanine and methionine 176 to alanine (PR1/M176A). FFS experiments were performed on a home-built two-photon fluorescence fluctuation microscope described previously (27), which is composed of an Olympus IX-71 and a mode locked Ti:Sapphire laser (Chameleon Ultra, Coherent). A 60x Plan-Apo oil immersion objective (NA=1.4, Olympus) was used to focus the laser and collect the fluorescence. Two short-pass filters (ET680sp-2p8, Chroma) eliminate scattered laser light. A band pass emission filter further filtered the fluorescence (FF01-525/50-01, semrock). An avalanche photodiode detector (APD) (SPCM-AQR-14, Perkin-Elmer) was directly connected to a data acquisition card (FLEX02, Correlator.com). The recorded photon counts were stored and later analyzed with programs written in IDL (ITT Exelis). The normalized brightness b (28) is defined as $b = \lambda_{app}/\lambda_{EGFP}$, which measures the oligomerization of the target protein. The sample apparent brightness λ_{app} is measured by generalized Mandel's Q parameter analysis (29). The brightness λ_{EGFP} is obtained in calibration experiments by measuring cells transfected with EGFP.

Chemical cross-linking analysis of p85 α and p85 α^{1-333} – p85 α and p85 α^{1-333} (Fig. 1B) were dialyzed into 20 mM HEPES pH 8.0 containing 20 mM NaCl, a solution condition that favors p85 α dimerization (Fig. 3A). Twelve micromolar p85 α was incubated with 0.5 mM disuccinimidyl suberate (DSS) for 15 min on ice with gentle agitation. Thirty micromolar p85 α^{1-333} was incubated with a 50X molar excess of disuccinimidyl tartrate (DST) freshly prepared in 100% DMSO for 15 min at room temperature with gentle agitation. The DSS and DST cross-linking reactions were quenched with the addition of Tris to a final concentration of 100 mM. Dimer formation was assessed using SDS-PAGE followed by Coomassie staining.

After electrophoresis, the gel region corresponding to p85 α or p85 α^{1-333} dimer was excised, and digested in-gel with trypsin to release the cross-linked peptides. The samples were processed and analyzed by mass spectrometry as previously described (30,31). Briefly, the cross-linked target proteins were in-gel digested twice

with trypsin (1:10 w/w). After proteolysis, cross-linked peptides were extracted, dried in a Speedvac (Savant), and desalted/purified on a C18 solid phase extraction column (Waters). Next, the cross-linked peptides were enriched by size exclusion chromatography and analyzed by an Orbitrap Q Exactive (QE) Plus mass spectrometer (Thermo Fisher). The QE plus instrument was directly coupled to an EasyLC system (Thermo Fisher). The cross-linked peptides were loaded onto an Easy-Spray column heated at 35 °C (C18, 3 mm particle size, 200 Å pore size, and 50 mm X 15 cm, Thermo fisher) and eluted using a 120 min LC gradient (2% - 10% B, 0-6 min; 10% - 35% B, 6-102 min; 35% - 100% B, 102-113 min; followed by equilibration, where mobile phase A consisted of 0.1% formic acid and mobile phase B consisted of 0.1% formic acid in acetonitrile) at a flow rate of ~300 nl/min. The spray voltage was 2.0 kV and the top 10 most abundant ions (with a charge stage of 4-7) were selected and fragmented by HCD (normalized HCD energy 28).

The raw data were transformed to MGF (mascot generic format) and searched by pLink (32). An initial MS1 search window of 5 Da was allowed to cover all isotopic peaks of the cross-linked peptides. The data were automatically filtered using a mass accuracy of MS1 \leq 10 ppm (parts per million) and MS2 \leq 20 ppm of the theoretical monoisotopic (A0) and other isotopic masses (A+1, A+2, A+3, and A+4) as specified in the software. An additional search parameter was methionine oxidation as a variable modification. A maximum of two trypsin missed-cleavage sites was allowed. The initial search results were obtained using a default 5% false discovery rate (FDR), as predicted by a target-decoy search strategy. In our analysis, we treated the 5% FDR as a rough initial filter of the raw data. We then manually inspected all of the crosslink MS/MS spectra that were assigned by the software and applied additional filters to remove potential false positive identifications from the dataset (30,31). This processing resulted in the removal of ~30% of the crosslinking data from the 5% FDR results, ensuring high-quality data for structural modeling.

Small angle X-ray scattering (SAXS)—SAXS profiles of full-length and truncated p85 α were measured at concentrations ranging from 0.5 to 5.0

mg/mL in either 20 mM NaCl, 20 mM HEPES, 5% glycerol at pH 8.0 or 500 mM NaCl, 20 mM HEPES, 5% glycerol at pH 7.5 at temperatures of 10 and 25 °C. The glycerol protects the proteins from radiation damage during X-ray exposure (33). All solutions were filtered through 0.1 μ m membranes (Millipore).

SAXS measurements were carried out at Beamline 4-2 of the Stanford Synchrotron Radiation Lightsource (SSRL) in the SLAC National Accelerator Laboratory. At SSRL, the beam energy and current were 11 keV and 500 mA, respectively. A silver behenate sample was used to calibrate the q-range and detector distance. Data collection was controlled with Blu-Ice (34). We used an automatic sample delivery system equipped with a 1.5 mm-diameter thin-wall quartz capillary within which a sample aliquot was oscillated in the X-ray beam to minimize radiation damage (35). The sample was placed at 1.7 m from a MX225-HE (Rayonix, USA) CCD detector with a binned pixel size of 293 by 293 μ m.

Up to 24 one-second exposures were used for each sample and buffers maintained at 10-25°C. The SAXS profile of each buffer was obtained under the same conditions and subtracted from the protein SAXS profile. Each of the resulting diffraction images was scaled using the transmitted beam intensity, azimuthally integrated by SASTool (36), and averaged to obtain fully processed data in the form of intensity versus q [$q=4\pi\sin(\theta)/\lambda$, θ =one-half of the scattering angle; λ =X-ray wavelength].

The buffer subtracted SAXS profiles were initially analyzed using the ATSAS package (37) to calculate radius of gyration (R_g), maximum particle size (D_{max}), pair distribution function ($P(r)$), and Porod volume (Table 1). The molecular weight (MW_{SAXS}) of each SAXS sample was estimated using SAXS MOW (38) with a threshold of $Q_{max} = 0.2 - 0.3 \text{ \AA}^{-1}$.

The *ab initio* shapes of the p85 α^{1-333} and p85 α dimers (Figs. 6A and 7A, respectively; gray envelope) were generated from the corresponding dimer SAXS profile by running DAMMIF (39) 20 times, and then refined through additional 50 DAMMIN (40) runs followed by superposition and averaging with DAMAVER (41).

Integrative multi-state modeling of p85 α and p85 α^{1-333} dimers—We computed ensembles of atomic multi-state models of the p85 α^{1-333} dimer and the p85 α dimer based on SAXS profiles, chemical cross-links, the assembly state determined by AUC, and crystal structures of its domains and homologs. A “multi-state model” is a model that specifies multiple discrete structural states of the system, all of which are required to explain the input information (19,42-44). In contrast, in an ensemble of models, any single model explains the input information. This approach proceeds through four stages: (1) data gathering, (2) representation of subunits and translation of the data into spatial restraints, (3) conformational sampling to produce the most parsimonious multi-state model consistent with all available data and information, and (4) analysis and assessment of the multi-state model. Our protocol was scripted using the Python Modeling Interface (PMI), version 2f82087, a library for modeling macromolecular complexes based on our open-source Integrative Modeling Platform package release 2.5 (<http://salilab.org/imp/>)(14-20). Files for the input data, modeling scripts, output model structures in multiple states, and additional figures and tables are available at <http://salilab.org/p85>. The procedure for each stage is summarized below.

Stage 1: Data gathering. SAXS profiles, assembly states, 231 DSS chemical cross-links for p85 α , and 25 DST chemical cross-links for p85 α^{1-333} were obtained as described above. HHpred (45,46) and ModWeb (47) identified the atomic structures of *Homo sapiens* p85 α^{2-82} [PDB, 3I5R (48), chain A], p85 $\alpha^{115-298}$ [PDB, 1PBW (49), chain A], p85 $\alpha^{324-433}$ [PDB, 2IUG (50), chain A], p85 $\alpha^{326-579}$ [PDB, 3HIZ and 3HHM (51), chain B], p85 $\alpha^{431-600}$ [PDB, 2V1Y (52), chain B], p85 $\alpha^{617-721}$ [PDB, 1H9O (53), chain A], and homologs [PDB, 1PRL (54), chain C and 2Y3A (55), chain B]. DomPRED (56), PSIPRED (57,58), and DISOPRED (59) predicted domain boundaries, secondary structure segments, and disordered regions, respectively. The intermolecular interaction between SH3 and PR1 domains was confirmed by its disruption by the SH3-binding peptide (RPLPPRPGA).

Stage 2: Subunit representation and translation of the data into spatial restraints. The size and

shape information contained in SAXS profiles can be used to improve the accuracy of atomic comparative models. An initial atomic model of the p85 α dimer was built based on template structures (Stage 1) and SAXS profiles as follows. First, we built 100 atomic comparative models for p85 α "monomer" in complex with the SH3-binding PR1-like peptide (RPLPPRPGA) (data not shown), using MODELLER 9.14 (60) based on the crystal structures and the closest template structures.

The theoretical SAXS profile and the χ value of the fit to the experimental "monomer" SAXS profile were calculated for each of the 100 comparative models, using FoXS (61,62). Then, these 100 models were ranked by the χ value of the fit. Second, the best-scoring "monomer" model (with a lowest χ value; data not shown) was used as a template for building an initial model of the p85 α dimer. We added another copy of the p85 α "monomer" at a random starting position for each sampling run, which resulted in an initial model of the p85 α dimer. The SH3-binding PR1-like peptides were removed in the dimer model of p85 α , to reflect the composition of the SAXS sample.

Next, a "monomer" model of p85 α ¹⁻³³³ was obtained by removing residues of 334-724 in the best-scoring p85 α "monomer" model (data not shown). Similarly, we added another copy of the p85 α ¹⁻³³³ "monomer" at a random starting position for each sampling run, which resulted in an initial model of the p85 α ¹⁻³³³ dimer. The SH3-binding PR1-like peptides were removed in the dimer model p85 α ¹⁻³³³, to reflect the composition of the SAXS sample.

Domains were coarse-grained using beads representing individual residues and arranged into either a rigid body or a flexible string based on the available crystallographic structures or comparative models (63). The coordinates of a bead were those of the corresponding C α atoms. In a rigid body, the beads have their relative distances constrained during conformational sampling in Stage 3, whereas in a flexible string the beads are restrained by the sequence connectivity, as described later in this section. The residues in the rigid bodies and flexible strings corresponded to 87% and 13% of p85 α , respectively.

With this representation, the information gathered in Stage 1 was converted into spatial restraints and constraints. We used different subsets of the spatial restraints and constraints (*i.e.*, restraint subsets) for different sampling runs, to maximize sampling efficiency. 231 DSS chemical cross-links for the p85 α dimer and 25 DST chemical cross-links for the p85 α ¹⁻³³³ dimer were used to construct a Bayesian scoring function that restrained the distances spanned by the cross-linked residues (64). The cross-link restraints were applied to the corresponding bead pairs, taking into account the difficulty of distinguishing intermolecular versus intramolecular cross-links in two identical p85 α subunits in the dimer; an ambiguous cross-link restraint considers all possible pair-wise assignments. For example, a restraint between residues 438 and 519 is evaluated (64) from the distances: "438@p85 α .1 - 519@p85 α .1", "438@p85 α .1 - 519@p85 α .2", "438@p85 α .2 - 519@p85 α .1", and "438@p85 α .2 - 519@p85 α .2", followed by scoring only the least violated distance.

Four homo-dimer cross-links between residues 438-438, 480-480, 530-530, and 567-567 were transformed to upper-harmonic distance restraints (up to 20 Å), enforcing the parallel intermolecular orientation of two iSH2 domains. In contrast, 4 chemical cross-links of 438-519, 438-530, 447-519, and 447-530 were transformed to upper-harmonic distance restraints (up to 20 Å), enforcing the anti-parallel intermolecular orientation of two iSH2 domains. In addition, a homo-dimer cross-link of 633-633 was transformed to upper-harmonic distance restraints (up to 20 Å), enforcing the intermolecular interaction of two cSH2 domains.

Notably, we applied 5 upper-harmonic distance restraints on residues 14-92, 51-92, 54-92, 70-92, and 73-92 (up to 13.5 Å) to retain the intermolecular interaction sites between SH3 and PR1 domains, each one residing in a different subunit of the dimer. A crystal structure of the SH3 domain bound to a poly-proline peptide (PDB, 1PRL) was used as a template for interaction site identification (54). Initially, both single and double intermolecular SH3-PR1 interactions were evaluated through the multi-state search for the p85 α ¹⁻³³³ dimer, leading to a conclusion that the double intermolecular SH3-PR1 interactions are dominant in the p85 α dimer.

Thus, we confined double intermolecular SH3-PR1 interactions in all restraint subsets of the p85 α dimer.

The excluded volume restraints were applied to each bead, using the statistical relationship between the volume and the residue that it covered (15,30,65). We applied the sequence connectivity restraint, using a harmonic upper-bound function of the distance between consecutive beads in a subunit, with a threshold distance equal to four times the sum of the radii of the two connected beads. The bead radius was calculated from the excluded volume of the corresponding bead, assuming standard protein density (15,30,65,66). Lastly, the most populated state of the p85 α ¹⁻³³³ dimer (40.3% population) was further constrained during conformational sampling of the p85 α dimer, in selected restraint subsets.

Stage 3: Conformational sampling to produce the most parsimonious multi-state model consistent with all available data and information. The initial dimer models of p85 α and p85 α ¹⁻³³³ were subjected to conformational sampling using Replica Exchange Gibbs sampling, based on the Metropolis Monte Carlo algorithm (63,67). The Monte Carlo moves included random translation and rotation of rigid bodies (up to 0.5 Å and 0.02 radians, respectively) and random translation of individual beads in the flexible segments up to 0.5 Å. For each of the restraint subsets, 2 to 3 independent sampling calculations were performed, each one starting with a random initial configuration. 4 to 16 replicas were used with temperatures ranging between 1.0 and 2.5. A model was saved every 10 Gibbs sampling steps, each consisting of a cycle of Monte Carlo steps that moved every rigid body and flexible bead once (30). The sampling produced ~80,000 (from 8 independent runs of the p85 α ¹⁻³³³ dimer) and ~200,000 models (from 45 independent runs of the p85 α dimer) that were submitted for subsequent multi-state analysis. The entire sampling procedure took 1 week on a cluster of ~400 computational cores.

The resulting ~80,000 and ~200,000 models obtained for the p85 α ¹⁻³³³ and p85 α dimers, respectively, were pruned to identify multi-state models that satisfied both the experimental SAXS profiles and the chemical cross-links datasets. MultiFoXS (68,69) was used to prune the datasets

with a composite score defined as a sum of the “*multi-state SAXS score*” (70) and the “*multi-state cross-links score*”.

The *multi-state SAXS score* (70) is the χ value for the comparison of the *multi-state SAXS profile* to the experimental profile; the *multi-state SAXS profile* is a weighted average of the theoretical SAXS profiles for the selected subset of states, each one calculated using FoXS (61,62). The side chains of whole residues in each state were reconstructed using PULCHRA 3.06 (71) for higher accuracy in the theoretical SAXS profiles.

The *multi-state cross-link score* is a negative value of the proportion of chemical cross-links satisfied in the selected subset of states; a cross-link restraint was considered to be satisfied by the subset if the minimum C α – C α distance of the corresponding residue pairs was smaller than a distance threshold of 35 Å, considering restraint ambiguity (above).

Independent fitting of subsets ranging from 1 to 9 states showed that 5 states of each protein was sufficient to account for both the experimental SAXS profiles and the chemical cross-links datasets.

Stage 4: Analysis and assessment of the multi-state model. The most populated state in the multi-state model was used as a reference for rigid body least-squares superposition of the remaining states. The multi-state models of p85 α ¹⁻³³³ and p85 α dimers were visualized with UCSF Chimera (Figs. 6C and 7C, respectively) (72). The TM-scores for each pair of the individual states were calculated using the corresponding web-server (73).

The multi-state model was assessed for how well it satisfied the data from which it was computed, including chemical cross-links, excluded volume, sequence connectivity, and SAXS profiles. We validated the multi-state model against each of the chemical cross-links; a cross-link restraint was considered to be satisfied by the multi-state model if the minimum C α - C α distance of the corresponding residue pairs was less than 35 Å. The excluded volume and sequence connectivity restraints were considered to be satisfied by an individual state if their combined score was less than 100. Finally, The χ value and the residual plot were used for the comparison of the “*multi-state SAXS profile*” to the experimental profile.

RESULTS

Characterization of cysteine-free p85 α —Purification of p85 α from bacterial expression systems is hampered by poor protein stability (unpublished data). We noted substantial improvements in protein purity and yield while preparing a cysteine-free mutant of p85 α for use in site-specific spin labeling studies (74). Therefore, we mutated each of the six cysteines in p85 α (two in the BCR domain, one in the iSH2 domain, and three in the cSH2 domain) to serine, leucine or valine (Fig. 1A), depending on whether the cysteines were predicted to participate in hydrophobic interactions based on crystal structures of isolated domains. As stated in Methods, the resultant protein is herein referred to as “p85 α ”, and the wild type protein as “native p85 α ”. p85 α robustly expresses and is readily purified in milligram quantities (10-15 mg/L bacterial culture). Similar protein yield and stability is achieved with the p85 α fragments that were studied (Fig. 1B). Creation of cysteine-free p85 α enabled the solution assembly and structural studies described in this paper.

We performed several assays to demonstrate the functional integrity of each cysteine-containing domain of p85 α relative to the native protein. Using immunoprecipitation and kinase assays, we found the two proteins showed similar binding to p110 α and inhibition of its kinase activity (Figs. 2A & B). The cSH2 domains of both p85 α proteins are comparably labeled by a photoactivatable ¹²⁵I-YXXM phosphopeptide; labeling is almost completely eliminated by competition with unlabeled peptide, showing that the reaction is specific (Fig. 2C). Lastly, the interaction of the small GTPase Rac1 with the p85¹⁻⁴³² (4) shows the expected GTP dependence, demonstrating that the BCR domain is unaffected by mutation of its cysteine residues (Fig. 2D). Taken together, these results show that substitution of the cysteine residues in p85 α has no appreciable effect on its biochemical activity.

p85 α dimerization in vitro—The ability of p85 α to self-assemble *in vitro* has been documented (13). A question that motivated our analysis of p85 α dimerization is whether it can play a role in the assembly of the various heterologous complexes that regulate PI3K

signaling. A second motivation was to enable structure and assembly studies by identifying solution conditions under which p85 α is predominantly monomeric or dimeric. We utilized two modes of AUC analysis to characterize p85 α self-assembly. Sedimentation velocity (SV) analysis measures the sedimentation rate and yields two coefficients, sedimentation (S) and diffusion (D), whose ratio S/D can provide the molecular weight. However, since S and D are dependent on both molecular size and shape, their ratio can yield erroneous molecular weights. Erroneous molecular weight values calculated from S/D can also result from protein self-association. Therefore we also conducted sedimentation equilibrium (SEQ) analyses, whose results are not dependent on molecular shape, to determine the molecular weight, stoichiometry, and dissociation constant (K_d) of p85 α and its fragments.

SV experiments with p85 α analyzed by the time-derivative method revealed single peaks (Fig. 3A), whose $S_{(20,w)}$ values increased with increasing protein concentration and decreased with increasing temperature and salt concentration (Fig. 3B). These observations reveal that p85 α undergoes rapidly reversible self-association at micromolar protein concentrations and that the assembly interaction is exothermic with an electrostatic component. The molecular weight values calculated from S/D from these experiments (Fig. 3A) proved to be erroneous upon comparison with the SEQ analyses described below (data not shown).

SEQ analysis of p85 α was conducted as a function of temperature at both low and high salt conditions at initial protein concentrations ranging from 1.2 to 9.6 μ M (Figs. 3C & D). The data are consistently described as a monomer – dimer equilibrium (Figs. 3E and 4). Adding higher order species does not improve the fit of the assembly models to the data (analysis not shown). The temperature dependence of dimerization is exothermic, as observed by SV; the linear Van't Hoff plots reveal different enthalpies at low and high salt ($\Delta H^\circ = -23.7$ and -14.3 kcal, respectively; Fig. 3C). The salt concentration substantially affects p85 α dimerization.

For example, at 10 °C the values of K_d measured at 20 and 500 mM NaCl differ by almost

40 fold (Fig. 3F). p85 α remains a mixture of monomer and dimer at 500 mM NaCl with the equilibrium biased towards monomer at the higher salt. Under conditions of low temperature and low salt p85 α is almost completely dimeric. Conversely, p85 α is predominantly monomeric at high temperature and high salt. Studies conducted at approximate physiological salt concentration (140 mM NaCl) are also well described by the monomer – dimer equilibrium with a K_d value comparable to that measured at the high salt concentration (data not shown).

To evaluate the contribution of intermolecular SH3-PR1 interactions to dimerization, we examined dimer formation in the presence a 9 amino acid SH3-binding peptide containing a proline-rich consensus sequence (RPLPPRPGA) that binds the p85 α SH3 domain (75). A 25-fold molar excess of peptide (250 μ M peptide vs. 9.6 μ M protein) drives the equilibrium completely to monomer regardless of the temperature or salt concentration (data not shown). This reaction occurs presumably by competition of the peptide with the PR1 domains for intermolecular SH3 domain binding. Thus, peptide-bound p85 α is an alternative condition for the analysis of monomeric protein.

The results presented above suggest that the intermolecular SH3-PR1 interaction dominates p85 α dimerization. To explore possible contributions of other domains to dimerization we analyzed the assembly of a series of p85 α truncations (depicted in Fig. 1A). SEQ analysis of the p85 α fragment containing only SH3-PR1-BCR-PR2 (p85 α^{1-333}) reveals ten-fold weaker dimerization compared to the full-length protein (Fig. 3F; $K_d = 30.0$ vs. 2.5 μ M at 20 mM NaCl and 10 °C), loss of salt dependence, and diminished enthalpy ($\Delta H^\circ = -6.7$ and -11.6 kcal, respectively; Fig. 3D). Analysis of a series of more refined truncations shows that it is the cSH2 domain that contributes to p85 α dimerization and is the source of the reaction's salt dependence (Figs. 3E & F, compare p85 α with p85 α^{1-600}). These data show that cSH2 also contributes to p85 α dimerization.

Published studies suggest that the BCR domain is potentially part of the dimer interface (13,49). To test whether the BCR domain itself dimerizes, we expressed and purified a fragment of p85 α corresponding to residues 78 – 322, which

contains the BCR domain flanked by the two proline-rich motifs. SEQ analysis shows that p85 α^{78-322} does not dimerize; an identical result was obtained for native p85 α^{78-322} , regardless of salt concentration or temperature for either construct (data not shown). These data show that the BCR domain alone does not dimerize.

p85 α dimerization in vivo—Fluorescence fluctuation spectroscopy (FFS) was used to measure dynamic changes in protein association *in vivo*. U201S cells were transfected with native GFP-p85 α , p85 α or a monomeric GFP control. The brightness of the GFP fluorescence correlates linearly and positively with increasing concentrations of the GFP-p85 α fusion protein. In contrast, the brightness of the monomeric GFP control is invariant with concentration. The slopes of the correlations for native p85 α and p85 α are identical and appreciably greater than that for GFP alone control (Fig. 5A). These data demonstrate that both native p85 α and p85 α dimerize *in vivo* in a concentration-dependent fashion.

To test the contribution of the SH3-PR1-BCR domains to p85 α dimerization *in vivo*, we compared the FFS distributions of native GFP-p85 α to those of a mutant p85 α in which two key proline residues in PR1 and a methionine residue in the BCR domain were mutated to alanine (p85 $\alpha^{\text{PR1/M176A}}$). These residues were selected based on studies showing that these mutants inhibit co-immunoprecipitation of tagged p85 α constructs in intact cells¹. The slope of brightness versus protein concentration was more than two-fold greater in U201S cells transfected with native GFP-p85 α as compared to native GFP-p85 $\alpha^{\text{PR1/M176A}}$, consistent with the mutant exhibiting reduced dimerization *in vivo* (Fig. 5B). Since mutation of PR1 and the BCR domain did not completely abolish dimerization, these data indicate that additional domains contribute to p85 α dimerization, an observation consistent with our *in vitro* analyses.

Small angle X-ray scattering (SAXS)—SAXS is a measure of p85 α assembly orthogonal to AUC that provides information about protein size and

¹ G.B. Mills, personal communication.

shape (76). SAXS profiles of p85 α and its fragments (Fig. 1) were measured under conditions shown by AUC analysis to favor monomer or dimer. At low salt, we observed by SAXS predominantly p85 α dimer at protein concentrations of 12 μ M (1.0 mg/mL) at 10°C (Fig. 7B and Table 1) and 18 μ M (1.5 mg/mL) at 25°C. At the highest protein concentration analyzed (5.0 mg/mL or \sim 60 μ M, greater than that analyzed by AUC), oligomers of higher order than dimer were observed. Saturating concentrations of the SH3-binding peptide drove the equilibrium completely to monomer as measured by AUC.

Our SAXS analysis of p85 α ¹⁻³³³ (SH3-PR1-BCR-PR2) is likewise consistent with the AUC results. At high protein concentrations, the protein is fully dimeric at low salt and low temperature (Fig. 6B and Table 1) and salt independent. In contrast, p85 α ⁷⁸⁻³²² (PR1-BCR-PR2) does not dimerize under any condition, confirming that the intermolecular SH3-PR1 interaction is critical for dimerization. SAXS analysis of p85 α ¹⁻⁴³² and p85 α ¹⁻⁶⁰⁰ is consistent with AUC data showing that the salt dependence of dimerization resides within the cSH2 domain (residues 617 to 724). The analysis also suggests that the iSH2 domain (residues 430 to 600) makes a small (two-fold) contribution to p85 α dimerization.

Chemical cross-linking of p85 α and p85 α ¹⁻³³³ dimers—Chemical cross-linking of p85 α ¹⁻³³³ and p85 α dimers with DST and DSS, respectively, was carried out under solution conditions favorable to dimer formation. The chemical cross-linking reactions were optimized for mass spectrometric analysis. The covalent adducts were analyzed by mass spectrometry. Our cross-linking and mass spectrometric analysis revealed 25 and 231 unique cross-linked residue pairs, respectively, for the p85 α ¹⁻³³³ and p85 α dimers (Figs. 6D and 7D). The overall connectivity pattern of the DST cross-links for p85 α ¹⁻³³³ is similar to that of the DSS cross-links for the same portions of full length p85 α showing that they provide complementary information on similar conformers. This result demonstrates that the absence of the C-terminal domains does not affect the conformational states sampled by SH3-PR1-BCR-PR2 and supports the use of the p85 α ¹⁻³³³ structural model in modeling the full-length protein that is presented below.

In full-length p85 α , cross-links were identified that are consistent with both parallel and antiparallel orientations of the two coiled-coil (iSH2) domains in the dimer. Homodimer cross-links between residues 438-438, 480-480, 530-530, and 567-567 are consistent with the parallel intermolecular orientation, according to the crystal structures [PDB 3HIZ and 3HHM (49)] of an isolated iSH2 domain. In contrast, cross-links between residues 438-519, 438-530, 447-519, and 447-530 are compatible with the anti-parallel orientation. These results suggest that the iSH2 domains in p85 α can dock in either orientation upon dimerization. Furthermore, many lysine residues formed multiple crosslinks spanning over the N- and C-terminal domains (e.g. K81, K142, K187, and K256; Figs. 6D and 7D), which is also consistent with the p85 α dimer being highly dynamic in solution.

We also identified chemical cross-links between residues 633-633 in the cSH2 domains. While it is impossible to distinguish intermolecular from intramolecular cross-links between non-identical residues in two identical p85 α subunits based on the cross-linking data alone, cross-links between identical residues can only be intermolecular. Moreover, the 633-633 cross-link is also consistent with the AUC observation that the cSH2 domain contributes to dimerization.

Integrative structure determination of p85 α and p85 α ¹⁻³³³ dimers in multiple states—We carried out conformational sampling using Replica Exchange Gibbs sampling based on the Metropolis Monte Carlo algorithm to study the structure and dynamics of p85 α dimers in solution (63,67). We began with the p85 α ¹⁻³³³ dimer (Fig. 1A) and tested whether or not the resulting \sim 80,000 models were consistent with the experimental SAXS profile and the 25 chemical cross-links. Both the SAXS profile and the cross-linking dataset were not simultaneously explained by any single sampled conformational state. For example, in the best-scoring single-state model (light red in Fig. 6B), $\chi > 13.9$ and only 48% of the DST cross-links were satisfied. This result suggests that the p85 α ¹⁻³³³ dimer is conformationally heterogeneous in solution.

We therefore computed multi-state models of the p85 α ¹⁻³³³ dimer for up to 9 states, using

MultiFoXS (68,69). The results show that a weighted combination of 5 states is sufficient to explain the experimental dimer SAXS profile within its noise ($\chi = 2.623$; blue in Fig. 6B) and all of the 25 chemical cross-links within a distance threshold of 35 Å (Fig. 6D). The best-scoring multi-state model of the p85 α^{1-333} dimer consists of 3 major states (conformations) with population weights of 40.3% (blue), 27.7% (red), and 16.6% (yellow), and 2 minor states with population weights of 8.0% (green) and 7.4% (purple) (Fig. 6C). The maximum particle sizes (D_{max}) of the major states ranged from 145 to 185 Å. D_{max} of the minor states are ~ 110 and ~ 250 Å. The root-mean-square deviation (RMSD) and the TM-scores (73) for each pair of the 5 states range from 18.3 to 36.1 Å and from 0.55 to 0.57, respectively, indicating highly heterogeneous folds within the multi-state model.

Similarly to the p85 α^{1-333} dimer, the full-length p85 α dimer is conformationally heterogeneous in solution; neither the SAXS profile nor the cross-linking dataset were simultaneously explained by any single sampled conformational state. For the best-scoring single-state model (light red in Fig. 7B), $\chi > 1.671$ and only 40% of the combined 256 DST cross-links were satisfied. We therefore computed multi-state models of the p85 α dimer for up to 9 states, using the most populated state of the p85 α^{1-333} dimer (40.3%, colored blue in Fig. 6C) as an additional constraint in the selected restraint subsets. This model accommodates all of the SAXS and cross-linking data. The results show that a weighted combination of 5 states is sufficient to explain the experimental dimer SAXS profile within its noise ($\chi = 1.275$; blue in Fig. 7B) and 95% of the combined 256 chemical cross-links within a distance threshold of 35 Å (Fig. 7D). The best-scoring multi-state model of the p85 α dimer consists of 3 major states with population weights of 33.2% (blue), 27.4% (red), and 18.2% (yellow), and 2 minor states with population weights of 13.4% (green) and 7.9% (purple) (Fig. 7C). The maximum particle sizes (D_{max}) of the 5 states were ranged from 170 to 320 Å. The root-mean-square deviation (RMSD) and the TM-scores (73) for each pair of the 5 states range from 26.8 to 65.6 Å and 0.27 to 0.43, respectively, indicating more conformational heterogeneity than the p85 α^{1-333} dimer.

Based on the multi-state model, the conformational dynamics of the p85 α^{1-333} dimer appear to be dominated by the relative intramolecular motions of the SH3 and BCR domains, connected by the PR1 motif linkers (residues 85 to 115) (Fig. 6C). The maximal displacement between the SH3 and BCR domains in the multi-state model is ~ 100 Å. Notably, the BCR domains do not appear to contact each other directly, in agreement with experimental results showing that the BCR domain is monomeric (residues 78-322; Fig. 3A). Reciprocal intermolecular SH3-PR1 interactions were identified in each of the 5 states, consistent with this interaction as a critical mediator of dimerization.

A large degree of heterogeneity is observed in the p85 α dimer, particularly in the coiled-coil (iSH2) domains, as well as the neighboring nSH2 and cSH2 domains. Importantly, the multi-state model indicates that the two iSH2 domains orient relative to one another in parallel, anti-parallel, or even perpendicular orientations (Fig. 7C). The heterogeneity in iSH2 domain orientation is supported by LOGICOIL calculations, a coiled-coil oligomerization state prediction program (77). The configurations of the intermolecular SH3-PR1 interaction sites were in agreement with our AUC results (Fig. 3A). Similarly, two BCR domains were not in contact with each other, which is also consistent with the AUC data. The previously discussed contribution of the cSH2-cSH2 contacts to p85 α dimerization is also seen in the multi-state model (Fig. 7C). In conclusion, it appears that both the p85 α^{1-333} and full-length p85 α dimers are highly dynamic molecules in solution, held together by the SH3-PR1 and cSH2 contacts, with maximal dimension of the molecules in solution (D_{max}) ranging from 110 to 250 Å and from 170 to 320 Å, respectively.

DISCUSSION

The five discrete protein-binding domains of p85 α allow it to function as a scaffolding protein and mediate the activity of multiple signaling pathways. Our studies have begun to explore whether the physical properties of p85 α actively regulate its interactions with other signaling proteins. The fact that p85 α dimerization is rapidly and freely reversible and its conformational states are highly heterogeneous would maximize its ability to interact with regulatory partners. p85 α

dimerization occurs *in vivo* in a concentration-dependent manner, demonstrating that the properties of p85 α measured *in vitro* are relevant to those in the cell (Fig. 5). If p85 α 's partners preferentially bind to the monomer or dimer, the intracellular concentration of p85 α , and hence its dimerization state, could play a role in modulating the activity of its binding partners in the cell.

These studies are enabled by the design, expression, and purification of a p85 α variant in which its six cysteines are mutated to residues of a similar chemical nature (*i.e.*, similar degree of hydrophobicity as inferred from crystal structures). Removal of surface cysteines to stabilize protein expression and purification is a commonly used technique; exposed cysteines have pKa values comparable to physiological pH (78), and thus are highly responsive to fluctuations in physiological and environmental conditions (79). Importantly, our cysteine-free p85 α behaves identically to the native protein in all functional assays, as all cysteine-free domains retain their ability to interact with known binding partners (Fig. 2). Thus, cysteine-free p85 α is a validated model for structural and biophysical analysis of the p85 α protein. Moreover, this validated model will serve as a platform for selective labeling with probes in future structure and dynamics studies.

We used our understanding of the chemical nature of the self-assembly equilibrium to identify conditions under which p85 α is predominantly dimeric. Using these conditions, we obtained chemical cross-linking and SAXS data that were used for integrative structure determination of the p85 α dimer. The flexible, elongated shape of the p85 α dimer states visualized in the multi-state model (Fig. 7C) highlights the importance of using sedimentation equilibrium (13) and SAXS to determine accurate molecular weights independent of molecular shape in studies of this molecule and its complexes.

Our sedimentation equilibrium analyses underscore the importance of the previously identified intermolecular SH3-PR1 interactions (13) at physiological conditions (Fig. 3). Our studies also demonstrate novel intermolecular electrostatic interactions mediated by the cSH2 domain, which stabilize the p85 α dimer. AUC, SAXS and chemical cross-linking results all support the presence of a cSH2 dimer contact. The structural consequence of the N-terminal (SH3-

PR1) and C-terminal (cSH2-cSH2) contacts is that the p85 α polypeptide is essentially pinned at each end in the dimer, allowing the intervening domains to exhibit substantial conformational flexibility (Fig. 7C). The relative contribution of the SH3-PR1 and cSH2 dimer contacts to p85 α dimer stability in cells deserves further study.

Although crystal structures have been solved for the five individual domains of native p85 α (49-51,53,80-84), the protein has not been amenable to either X-ray crystallography, presumably because of the protein's flexibility, or NMR spectroscopy, due to its size. The orientation and spatial relation of each domain relative to the others is unknown, although this knowledge is critically important to defining the molecular mechanism(s) of p85 α -mediated regulation of PI3K signaling. Thus, we explored the structure and dynamics of the p85 α dimers by applying an integrative modeling approach that relies on data from orthogonal experimental methods at different levels of resolution. The result is a multi-state model of the p85 α dimer that defines the conformations and populations of individual states (Figures 6C and 7C). An important result is the conformational heterogeneity of the p85 α dimer, which is evident in the structural diversity of the multiple states that comprise the multi-state model.

We addressed the issue of conformational heterogeneity by first modeling the p85 α ¹⁻³³³ dimer, as it is more amenable than p85 α to chemical cross-linking experiments, yet still exhibits the monomer/dimer equilibrium of the full-length protein. The multiple states of the p85 α ¹⁻³³³ dimer that constitute the multi-state model differ in the relative orientation of the SH3 and BCR domains. Notably, the BCR domains do not directly contact each other. In the most populated (40.3%) state, both SH3 domains interact *in trans* with the PR1 regions of the opposing p85 α ¹⁻³³³ fragment.

We computed the multi-state model of the full-length p85 α dimer using this SH3-BCR conformational state as a constraint, in conjunction with cross-linking and SAXS data. The conformational heterogeneity that characterizes the p85 α dimer is mainly due to the diverse configurations of the coiled-coil (iSH2) domains as well as the neighboring nSH2 and cSH2 domains. The two iSH2 domains in the p85 α dimer are oriented parallel, anti-parallel, or even

perpendicularly, relative to each other. The variety of orientations identified in the multi-state model is consistent with AUC data showing that the iSH2 domains do not appear to significantly stabilize the p85 α dimer. This conformational flexibility may allow iSH2 to form heterologous interactions within the context of the p85 α dimer. In contrast, the cSH2 domains contact each other in all 5 states of the multi-state model, consistent with their contribution to dimerization as observed by AUC (Fig. 3E).

In agreement with our AUC results showing lack of BCR dimerization in solution, the BCR domains do not appear to interact with each other in the multi-state model of either the p85 α ¹⁻³³³ or p85 α dimers. Although the two BCR domains are juxtaposed within one of the 5 conformational states (Fig. 7C; 27.4%, red and orange) we have no evidence that these domains stabilize the p85 α dimer (Fig. 3A).

A study recently published by Cheung *et al* has generated a model of the p85^{SH3-BCR} fragment as a dimer bound to PTEN and small GTPase Rab5 (85). In this model, the dimer interface consists of SH3-PR1 and BCR-BCR contacts, the latter based on a previously published crystal structure of the BCR domain that shows a BCR homodimer (49). The dimerization K_d measured for their His-PR1-BCR construct is $162.9 \pm 41.4 \mu\text{M}$, indicating very weak assembly. In contrast, we did not detect any dimerization of an untagged PR1-BCR-PR2 construct (p85 α ⁷⁸⁻³²²) at three temperatures and two solution conditions. Not surprisingly, the presence of SH3-binding peptide had no effect on p85 α ⁷⁸⁻³²² oligomerization unlike its complete inhibition of dimerization of p85 α and the deletion fragments that contain the SH3 domain. It is possible that the presence of a His tag in the constructs studied by Cheung *et al.* enhanced the

dimerization of the PR1-BCR fragment, as purification of His-tagged proteins using nickel chromatography can result in oligomerization (86). Even if BCR-BCR contacts do weakly contribute to p85 α dimer stabilization in the context of the full-length protein, the role of contacts seen in the BCR crystal structure is uncertain. A DiMoVo score (87) of 0.366 in the published structure [PDB, 1PBW] suggests a crystallographic dimer, not a biological one. It remains possible that binding of additional partners, such as PTEN, may bring the BCR domains in closer proximity to one another and further stabilize p85 α in its dimeric form.

Another issue for future work will be whether the same interactions that drive p85 α homodimerization occur when p85 α is bound to p110. Although there is one report of p85/p110 heterotetramers in cultured cells (88), these higher order complexes have not reported during purification and gel filtration analysis of recombinant p85/p110 dimers. Structures of the nSH2-iSH2/p110 α and iSH2-cSH2/p110 β heterodimers show that the ABD, C2 and kinase domains of p110 make extensive contacts with the iSH2 domain, whereas the nSH2 and cSH2 domains contact the helical, C2 and kinase domains (83,84). These contacts would likely preclude a number of the conformational states derived from the SAXS and cross-linking data in this study. However, given the modest energetic contributions of the iSH2 and cSH2 domains to the formation of the p85 α homodimer, it is not clear that binding of p110 to the iSH2 domain would prevent the formation of (p85/p110)₂ heterotetramers. Analysis of p85/p110 heterotetramers, and identification of conditions that promote heterotetramer formation, will be an important next step in these studies.

Acknowledgements: This work was supported by NIH grants R01 GM112524 (JMB and ARB), R01 NS83085 (RHS), and T32 GM007288 (JL). Portions of this research were carried out at the SSRL, SLAC National Accelerator Laboratory operated for DOE by Stanford University. The SSRL SMBP is supported by the DOE Office of Biological and Environmental Research, by the National Institutes of Health, NCRR, Biomedical Technology Program (P41RR001209), and by NIGMS, National Institutes of Health (P41GM103393). We thank Drs. David Cowburn, Gary Gerfen, and Mark Girvin (Albert Einstein Medical School) for helpful discussions. We thank Drs. Gordon Mills (MD Anderson) and Stefan Arold (King Abdullah University of Science and Technology) for sharing unpublished data. We also thank Tsutomu Matsui and Thomas M. Weiss at SSRL (SLAC National Accelerator Laboratory) for their assistance with collecting SAXS data, and Dina Schneidman-Duhovny (UCSF) for discussions on SAXS analysis.

Conflict of interest: The authors declare they have no conflicts of interest with contents of this manuscript.

Author contributions: JL purified all proteins, conducted and analyzed the AUC experiments with the assistance of MB, and performed the chemical crosslinking. SJK conducted and analyzed the SAXS experiments, and performed the molecular modeling. JL and SJK wrote the manuscript. YS conducted and analyzed the chemical cross-linking and mass spectrometry experiments. BW conducted and analyzed the FFS experiments. HW created and characterized no-cys p85. AS, RS, and BTC provided intellectual input. ARB, MB, and JMB edited the manuscript, and conceived and supervised the project.

REFERENCES

1. Backer, J. M. (2010) The regulation of class IA PI 3-kinases by inter-subunit interactions. *Curr Top Microbiol Immunol* **346**, 87-114
2. Burke, J. E., and Williams, R. L. (2015) Synergy in activating class I PI3Ks. *Trends Biochem Sci* **40**, 88-100
3. Maehama, T., and Dixon, J. E. (1998) The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. *The Journal of biological chemistry* **273**, 13375-13378
4. Tolia, K. F., Cantley, L. C., and Carpenter, C. L. (1995) Rho family GTPases bind to phosphoinositide kinases. *The Journal of biological chemistry* **270**, 17656-17659
5. Zheng, Y., Bagrodia, S., and Cerione, R. A. (1994) Activation of phosphoinositide 3-kinase activity by Cdc42Hs binding to p85. *J Biol Chem* **269**, 18727-18730
6. Pleiman, C. M., Hertz, W. M., and Cambier, J. C. (1994) Activation of phosphatidylinositol-3' kinase by Src-family kinase SH3 binding to the p85 subunit. *Science* **263**, 1609-1612
7. Hale, B. G., Batty, I. H., Downes, C. P., and Randall, R. E. (2008) Binding of influenza A virus NS1 protein to the inter-SH2 domain of p85 suggests a novel mechanism for phosphoinositide 3-kinase activation. *The Journal of biological chemistry* **283**, 1372-1380
8. Chiu, Y. H., Lee, J. Y., and Cantley, L. C. (2014) BRD7, a tumor suppressor, interacts with p85 α and regulates PI3K activity. *Mol Cell* **54**, 193-202
9. Rabinovsky, R., Pochanard, P., McNear, C., Brachmann, S. M., Duke-Cohan, J. S., Garraway, L. A., and Sellers, W. R. (2009) p85 Associates with unphosphorylated PTEN and the PTEN-associated complex. *Mol Cell Biol* **29**, 5377-5388
10. Chagpar, R. B., Links, P. H., Pastor, M. C., Furber, L. A., Hawrysh, A. D., Chamberlain, M. D., and Anderson, D. H. (2010) Direct positive regulation of PTEN by the p85 subunit of phosphatidylinositol 3-kinase. *Proc Natl Acad Sci U S A* **107**, 5471-5476
11. Cheung, L. W., Hennessy, B. T., Li, J., Yu, S., Myers, A. P., Djordjevic, B., Lu, Y., Stemke-Hale, K., Dyer, M. D., Zhang, F., Ju, Z., Cantley, L. C., Scherer, S. E., Liang, H., Lu, K. H., Broaddus, R. R., and Mills, G. B. (2011) High frequency of PIK3R1 and PIK3R2 mutations in endometrial cancer elucidates a novel mechanism for regulation of PTEN protein stability. *Cancer discovery* **1**, 170-185
12. Kapeller, R., Prasad, K. V., Janssen, O., Hou, W., Schaffhausen, B. S., Rudd, C. E., and Cantley, L. C. (1994) Identification of two SH3-binding motifs in the regulatory subunit of phosphatidylinositol 3-kinase. *The Journal of biological chemistry* **269**, 1927-1933
13. Harpur, A. G., Layton, M. J., Das, P., Bottomley, M. J., Panayotou, G., Driscoll, P. C., and Waterfield, M. D. (1999) Intermolecular interactions of the p85 α regulatory subunit of phosphatidylinositol 3-kinase. *J Biol Chem* **274**, 12323-12332
14. Russel, D., Lasker, K., Webb, B., Velazquez-Muriel, J., Tjioe, E., Schneidman-Duhovny, D., Peterson, B., and Sali, A. (2012) Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. *PLoS biology* **10**, e1001244
15. Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprpto, A., Karni-Schmidt, O., Williams, R., Chait, B. T., Rout, M. P., and Sali, A. (2007) Determining the architectures of macromolecular assemblies. *Nature* **450**, 683-694
16. Alber, F., Forster, F., Korkein, D., Topf, M., and Sali, A. (2008) Integrating diverse data for structure determination of macromolecular assemblies. *Annual review of biochemistry* **77**, 443-477
17. Webb, B., Lasker, K., Schneidman-Duhovny, D., Tjioe, E., Phillips, J., Kim, S. J., Velazquez-Muriel, J., Russel, D., and Sali, A. (2011) Modeling of proteins and their assemblies with the integrative modeling platform. *Methods in molecular biology* **781**, 377-397

18. Ward, A. B., Sali, A., and Wilson, I. A. (2013) Biochemistry. Integrative structural biology. *Science* **339**, 913-915
19. Molnar, K. S., Bonomi, M., Pellarin, R., Clinthorne, G. D., Gonzalez, G., Goldberg, S. D., Goulian, M., Sali, A., and DeGrado, W. F. (2014) Cys-scanning disulfide crosslinking and bayesian modeling probe the transmembrane signaling mechanism of the histidine kinase, PhoQ. *Structure* **22**, 1239-1251
20. Schneidman-Duhovny, D., Pellarin, R., and Sali, A. (2014) Uncertainty in integrative structural modeling. *Current opinion in structural biology* **28**, 96-104
21. Fritsch, R., de Krijger, I., Fritsch, K., George, R., Reason, B., Kumar, M. S., Diefenbacher, M., Stamp, G., and Downward, J. (2013) RAS and RHO families of GTPases directly regulate distinct phosphoinositide 3-kinase isoforms. *Cell* **153**, 1050-1063
22. Yu, J., Zhang, Y., McIlroy, J., Rordorf-Nikolic, T., Orr, G. A., and Backer, J. M. (1998) Regulation of the p85/p110 phosphatidylinositol 3'-kinase: stabilization and inhibition of the p110 α catalytic subunit by the p85 regulatory subunit. *Molecular and cellular biology* **18**, 1379-1387
23. Philo, J. S. (2006) Improved methods for fitting sedimentation coefficient distributions derived by time-derivative techniques. *Anal Biochem* **354**, 238-246
24. Stafford, W. F., 3rd. (1992) Boundary analysis in sedimentation transport experiments: a procedure for obtaining sedimentation coefficient distributions using the time derivative of the concentration profile. *Analytical biochemistry* **203**, 295-301
25. Cole, J. L. (2004) Analysis of heterogeneous interactions. *Methods Enzymol* **384**, 212-232
26. Cole, J. L., Lary, J. W., T, P. M., and Laue, T. M. (2008) Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods in cell biology* **84**, 143-179
27. Wu, B., Chao, J. A., and Singer, R. H. (2012) Fluorescence fluctuation spectroscopy enables quantitative imaging of single mRNAs in living cells. *Biophys J* **102**, 2936-2944
28. Chen, Y., Wei, L. N., and Muller, J. D. (2003) Probing protein oligomerization in living cells with fluorescence fluctuation spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 15492-15497
29. Sanchez-Andres, A., Chen, Y., and Muller, J. D. (2005) Molecular brightness determined from a generalized form of Mandel's Q-parameter. *Biophysical journal* **89**, 3531-3547
30. Shi, Y., Fernandez-Martinez, J., Tjioe, E., Pellarin, R., Kim, S. J., Williams, R., Schneidman-Duhovny, D., Sali, A., Rout, M. P., and Chait, B. T. (2014) Structural characterization by cross-linking reveals the detailed architecture of a coatomer-related heptameric module from the nuclear pore complex. *Molecular & cellular proteomics : MCP* **13**, 2927-2943
31. Cevher, M. A., Shi, Y., Li, D., Chait, B. T., Malik, S., and Roeder, R. G. (2014) Reconstitution of active human core Mediator complex reveals a critical role of the MED14 subunit. *Nature structural & molecular biology* **21**, 1028-1034
32. Yang, B., Wu, Y. J., Zhu, M., Fan, S. B., Lin, J., Zhang, K., Li, S., Chi, H., Li, Y. X., Chen, H. F., Luo, S. K., Ding, Y. H., Wang, L. H., Hao, Z., Xiu, L. Y., Chen, S., Ye, K., He, S. M., and Dong, M. Q. (2012) Identification of cross-linked peptides from complex samples. *Nat Methods* **9**, 904-906
33. Kuwamoto, S., Akiyama, S., and Fujisawa, T. (2004) Radiation damage to a protein solution, detected by synchrotron X-ray small-angle scattering: dose-related considerations and suppression by cryoprotectants. *J. Synchrotron Rad.* **11**, 462-468
34. McPhillips, T. M., McPhillips, S. E., Chiu, H. J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M., and Kuhn, P. (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *Journal of synchrotron radiation* **9**, 401-406
35. Martel, A., Liu, P., Weiss, T. M., Niebuhr, M., and Tsuruta, H. (2012) An integrated high-throughput data acquisition system for biological solution X-ray scattering studies. *Journal of synchrotron radiation* **19**, 431-434

36. SASTool 2013. SasTool. 2013, <http://ssrl.slac.stanford.edu/~saxs/analysis/sastool.htm>
37. Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V., and Svergun, D. I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *Journal of Applied Crystallography* **45**, 342-350
38. Fischer, H., Neto, M. D., Napolitano, H. B., Polikarpov, I., and Craievich, A. F. (2010) Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *Journal of Applied Crystallography* **43**, 101-109
39. Franke, D., and Svergun, D. I. (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *Journal of Applied Crystallography* **42**, 342-346
40. Svergun, D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophysical journal* **76**, 2879-2886
41. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography* **36**, 860-864
42. Carter, L., Kim, S. J., Schneidman-Duhovny, D., Stohr, J., Poncet-Montange, G., Weiss, T. M., Tsuruta, H., Prusiner, S. B., and Sali, A. (2015) Prion Protein-Antibody Complexes Characterized by Chromatography-Coupled Small-Angle X-Ray Scattering. *Biophys J* **109**, 793-805
43. Sali, A., Berman, H. M., Schwede, T., Trewhella, J., Kleywegt, G., Burley, S. K., Markley, J., Nakamura, H., Adams, P., Bonvin, A. M., Chiu, W., Peraro, M. D., Di Maio, F., Ferrin, T. E., Grunewald, K., Gutmanas, A., Henderson, R., Hummer, G., Iwasaki, K., Johnson, G., Lawson, C. L., Meiler, J., Marti-Renom, M. A., Montelione, G. T., Nilges, M., Nussinov, R., Patwardhan, A., Rappsilber, J., Read, R. J., Saibil, H., Schroder, G. F., Schwieters, C. D., Seidel, C. A., Svergun, D., Topf, M., Ulrich, E. L., Velankar, S., and Westbrook, J. D. (2015) Outcome of the First wwPDB Hybrid/Integrative Methods Task Force Workshop. *Structure* **23**, 1156-1167
44. Pelikan, M., Hura, G. L., and Hammel, M. (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering. *General physiology and biophysics* **28**, 174-189
45. Soding, J. (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21**, 951-960
46. Soding, J., Biegert, A., and Lupas, A. N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**, W244-248
47. Pieper, U., Webb, B. M., Dong, G. Q., Schneidman-Duhovny, D., Fan, H., Kim, S. J., Khuri, N., Spill, Y. G., Weinkam, P., Hammel, M., Tainer, J. A., Nilges, M., and Sali, A. (2014) ModBase, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res* **42**, D336-346
48. Batra-Safferling, R., Granzin, J., Modder, S., Hoffmann, S., and Willbold, D. (2010) Structural studies of the phosphatidylinositol 3-kinase (PI3K) SH3 domain in complex with a peptide ligand: role of the anchor residue in ligand binding. *Biol. Chem.* **391**, 33-42
49. Musacchio, A., Cantley, L. C., and Harrison, S. C. (1996) Crystal structure of the breakpoint cluster region-homology domain from phosphoinositide 3-kinase p85 alpha subunit. *Proc Natl Acad Sci U S A* **93**, 14373-14378
50. Nolte, R. T., Eck, M. J., Schlessinger, J., Shoelson, S. E., and Harrison, S. C. (1996) Crystal structure of the PI 3-kinase p85 amino-terminal SH2 domain and its phosphopeptide complexes. *Nat Struct Biol* **3**, 364-374
51. Mandelker, D., Gabelli, S. B., Schmidt-Kittler, O., Zhu, J., Cheong, I., Huang, C. H., Kinzler, K. W., Vogelstein, B., and Amzel, L. M. (2009) A frequent kinase domain mutation that changes the interaction between PI3K α and the membrane. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 16996-17001
52. Hersleth, H. P., Hsiao, Y. W., Ryde, U., Gorbitz, C. H., and Andersson, K. K. (2008) The crystal structure of peroxymyoglobin generated through cryoradiolytic reduction of myoglobin compound III during data collection. *The Biochemical journal* **412**, 257-264

53. Pauptit, R. A., Dennis, C. A., Derbyshire, D. J., Breeze, A. L., Weston, S. A., Rowsell, S., and Murshudov, G. N. (2001) NMR trial models: experiences with the colicin immunity protein Im7 and the p85 α C-terminal SH2-peptide complex. *Acta crystallographica. Section D, Biological crystallography* **57**, 1397-1404
54. Feng, S., Chen, J. K., Yu, H., Simon, J. A., and Schreiber, S. L. (1994) Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science* **266**, 1241-1247
55. Zhang, X., Vadas, O., Perisic, O., Anderson, K. E., Clark, J., Hawkins, P. T., Stephens, L. R., and Williams, R. L. (2011) Structure of lipid kinase p110 β /p85 β elucidates an unusual SH2-domain-mediated inhibitory mechanism. *Mol Cell* **41**, 567-578
56. Marsden, R. L., McGuffin, L. J., and Jones, D. T. (2002) Rapid protein domain assignment from amino acid sequence using predicted secondary structure. *Protein science : a publication of the Protein Society* **11**, 2814-2824
57. Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**, 195-202
58. Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K., and Jones, D. T. (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res* **41**, W349-357
59. Ward, J. J., McGuffin, L. J., Bryson, K., Buxton, B. F., and Jones, D. T. (2004) The DISOPRED server for the prediction of protein disorder. *Bioinformatics* **20**, 2138-2139
60. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779-815
61. Schneidman-Duhovny, D., Hammel, M., and Sali, A. (2010) FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res* **38**, W540-544
62. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A., and Sali, A. (2013) Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys J* **105**, 962-974
63. Shi*, Y., Fernandez-Martinez*, J., Tjioe*, E., Pellarin*, R., Kim*, S. J., Williams, R., Schneidman, D., Sali, A., Rout, M. P., and Chait, B. T. (2014) Structural characterization by cross-linking reveals the detailed architecture of a coatamer-related heptameric module from the nuclear pore complex. *Molecular & cellular proteomics : MCP* **13**, 2927-2943
64. Erzberger, J. P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C. H., Cimermanic, P., Boehringer, D., Sali, A., Aebersold, R., and Ban, N. (2014) Molecular architecture of the 40S eIF1eIF3 translation initiation complex. *Cell* **158**, 1123-1135
65. Shen, M. Y., and Sali, A. (2006) Statistical potential for assessment and prediction of protein structures. *Protein science : a publication of the Protein Society* **15**, 2507-2524
66. Fernandez-Martinez, J., Phillips, J., Sekedat, M. D., Diaz-Avalos, R., Velazquez-Muriel, J., Franke, J. D., Williams, R., Stokes, D. L., Chait, B. T., Sali, A., and Rout, M. P. (2012) Structure-function mapping of a heptameric module in the nuclear pore complex. *J Cell Biol* **196**, 419-434
67. Rieping, W., Habeck, M., and Nilges, M. (2005) Inferential structure determination. *Science* **309**, 303-306
68. Carter, L., Kim, S. J., Schneidman-Duhovny, D., Stöhr, J., Poncet-Montange, G., Weiss, T. M., Tsuruta, H., Prusiner, S. B., and Sali, A. (2015) Prion protein-antibody complexes characterized by chromatography-coupled small-angle X-ray scattering. *Biophysical Journal* **in press**
69. Lawler, E. L., and Wood, D. E. (1966) Branch-and-Bound Methods - a Survey. *Oper Res* **14**, 699-&
70. Kim*, S. J., Fernandez-Martinez*, J., Sampathkumar*, P., Martel, A., Matsui, T., Tsuruta, H., Weiss, T., Shi, Y., Markina-Inarrairaegui, A., Bonanno, J. B., Sauder, J. M., Burley, S. K., Chait, B. T., Almo, S. C., Rout, M. P., and Sali, A. (2014) Integrative structure-function mapping of the nucleoporin Nup133 suggests a conserved mechanism for membrane anchoring of the nuclear pore complex. *Molecular & cellular proteomics : MCP* **13**, 2911-2926
71. Rotkiewicz, P., and Skolnick, J. (2008) Fast procedure for reconstruction of full-atom protein models from reduced representations. *J Comput Chem* **29**, 1460-1465

72. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612
73. Zhang, Y., and Skolnick, J. (2004) Scoring function for automated assessment of protein structure template quality. *Proteins* **57**, 702-710
74. Fu, Z., Aronoff-Spencer, E., Backer, J. M., and Gerfen, G. J. (2003) The structure of the inter-SH2 domain of class IA phosphoinositide 3-kinase determined by site-directed spin labeling EPR and homology modeling. *Proc Natl Acad Sci U S A* **100**, 3275-3280
75. Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W., and Schreiber, S. L. (1994) Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* **76**, 933-945
76. Schneidman-Duhovny, D., Kim, S. J., and Sali, A. (2012) Integrative structural modeling with small angle X-ray scattering profiles. *BMC structural biology* **12**, 17
77. Vincent, T. L., Green, P. J., and Woolfson, D. N. (2013) LOGICOIL--multi-state prediction of coiled-coil oligomeric state. *Bioinformatics* **29**, 69-76
78. Marino, S. M., and Gladyshev, V. N. (2010) Cysteine function governs its conservation and degeneration and restricts its utilization on protein surfaces. *J Mol Biol* **404**, 902-916
79. Marino, S. M., and Gladyshev, V. N. (2012) Analysis and functional prediction of reactive cysteine residues. *The Journal of biological chemistry* **287**, 4419-4425
80. Koyama, S., Yu, H., Dalgarno, D. C., Shin, T. B., Zydowsky, L. D., and Schreiber, S. L. (1993) Structure of the PI3K SH3 domain and analysis of the SH3 family. *Cell* **72**, 945-952
81. Hoedemaeker, F. J., Siegal, G., Roe, S. M., Driscoll, P. C., and Abrahams, J. P. (1999) Crystal structure of the C-terminal SH2 domain of the p85alpha regulatory subunit of phosphoinositide 3-kinase: an SH2 domain mimicking its own substrate. *J Mol Biol* **292**, 763-770
82. Booker, G. W., Breeze, A. L., Downing, A. K., Panayotou, G., Gout, I., Waterfield, M. D., and Campbell, I. D. (1992) Structure of an SH2 domain of the p85 alpha subunit of phosphatidylinositol-3-OH kinase. *Nature* **358**, 684-687
83. Huang, C. H., Mandelker, D., Schmidt-Kittler, O., Samuels, Y., Velculescu, V. E., Kinzler, K. W., Vogelstein, B., Gabelli, S. B., and Amzel, L. M. (2007) The structure of a human p110alpha/p85alpha complex elucidates the effects of oncogenic PI3Kalpha mutations. *Science* **318**, 1744-1748
84. Miled, N., Yan, Y., Hon, W. C., Perisic, O., Zvelebil, M., Inbar, Y., Schneidman-Duhovny, D., Wolfson, H. J., Backer, J. M., and Williams, R. L. (2007) Mechanism of two classes of cancer mutations in the phosphoinositide 3-kinase catalytic subunit. *Science* **317**, 239-242
85. Cheung, L. W., Walkiewicz, K. W., Besong, T. M., Guo, H., Hawke, D. H., Arold, S. T., and Mills, G. B. (2015) Regulation of the PI3K pathway through a p85alpha monomer-homodimer equilibrium. *Elife* **4**
86. Sprules, T., Green, N., Featherstone, M., and Gehring, K. (1998) Nickel-induced oligomerization of proteins containing 10-histidine tags. *BioTechniques* **25**, 20-22
87. Bernauer, J., Bahadur, R. P., Rodier, F., Janin, J., and Poupon, A. (2008) DiMoVo: a Voronoi tessellation-based method for discriminating crystallographic and biological protein-protein interactions. *Bioinformatics* **24**, 652-658
88. Perez-Garcia, V., Redondo-Munoz, J., Kumar, A., and Carrera, A. C. (2014) Cell activation-induced phosphoinositide 3-kinase alpha/beta dimerization regulates PTEN activity. *Molecular and cellular biology* **34**, 3359-3373
89. Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D., Konarev, P. V., and Svergun, D. I. (2012) New developments in the program package for small-angle scattering data analysis. *Journal of applied crystallography* **45**, 342-350

FOOTNOTES

This work was supported by NIH grants R01 GM112524 (JMB and ARB), R01 NS83085 (RS), and T32 GM007288 (JL).

The abbreviations used are: PI3K, phosphoinositide 3'-kinase; AUC, analytical ultracentrifugation; FFS, fluorescence fluctuation spectroscopy; SAXS, small angle X-ray scattering, SEQ, sedimentation equilibrium; SV, sedimentation velocity; PTEN, phosphatase and tensin homologue deleted on chromosome ten; PR, proline-rich.

FIGURE LEGENDS

FIGURE 1. Cysteine-free p85 α and domain-containing fragments.

A) The six cysteines of wild-type human p85 α were mutated (C146S, C167S, C498S, C656S, C659V, and C670L) to generate cysteine-free recombinant p85 α , heretofore referred to as p85 α . Four 85 α truncations were also prepared.

B) Coomassie-stained SDS-PAGE gel of purified p85 α and fragments.

FIGURE 2. Characterization of cysteine-free p85 α .

A) Immunoblot analysis of myc-p110 α binding to GST alone, GST-p85 α or native GST-p85 α .

B) Kinase activity of p110 α purified from HEK293T cells in the presence of GST-p85 α or native GST-p85 α . Data is the average \pm SEM from three independent experiments.

C) Binding of GST-cSH2 domains from p85 α or native p85 α to a photoactivatable ¹²⁵I-labeled tyrosine phosphopeptide, in the absence or presence of unlabeled phosphopeptide.

D) Immunoblot analysis of GST-p85 α ¹⁻⁴³² binding to recombinant Rac1 loaded without or with GTP γ S.

FIGURE 3. p85 α dimerization *in vitro* measured by analytical ultracentrifugation (AUC).

A) Sedimentation coefficient distribution, g(s*), obtained from sedimentation velocity analysis of p85 α at three concentrations (1.2, 3.6, and 9.5 μ M) in 20 mM NaCl containing buffer at 4 $^{\circ}$ C. All sedimentation coefficient (*S*, in Svedbergs) values presented in the paper have been corrected to 20 $^{\circ}$ C and the density of water (*S*_{20,w}) to allow comparison of data from experiments performed in different buffers and temperatures.

B) Plot of sedimentation velocity data showing *S* as a function of p85 α concentration (absorbance at 280 nm) at 4, 15, 37, and 15 $^{\circ}$ C in 300 mM NaCl containing buffer.

C and D) V'ant Hoff plots relating the association constant (*K*_a) as a function of temperature for full-length p85 α (C) p85 α ¹⁻³³³ (D) in 20 mM and 500 mM NaCl containing buffer.

E) Weight average molecular weights (kDa) obtained from sedimentation equilibrium analysis of p85 α and corresponding fragments in buffer containing either 20 mM NaCl or 500 mM NaCl, and at 500 mM

NaCl in the presence of 1 mM SH3-binding peptide. Solid and dashed lines indicate the dimer and monomer molecular weights calculated from sequence, respectively.

F) Dissociation constant (K_d) values for each p85 α construct in low and high salt, derived from fitting sedimentation the equilibrium data to a monomer-dimer association model. Values are represented as joint confidence intervals of one standard deviation.

FIGURE 4. Sedimentation equilibrium analysis of p85 α at 10 °C.

The graphs show the distribution of p85 α protein concentration during sedimentation equilibrium analysis at 10 °C obtained by measuring the absorption at 280 nm versus the radial distance from the center of the rotor. Three concentrations of full-length p85 α were analyzed (1.2, 3.6, and 9.6 μ M). The equilibrium protein concentration distributions measured at each of the two rotor speeds are shown (upper curve: 8,000 rpm; lower curve: 16,000 rpm) with the best global fit to a monomer-dimer association model shown as the solid line in a) 20 mM NaCl and b) 500 mM NaCl, and c) fit to molecular weight of a single species in 500 mM NaCl in the presence of proline-rich peptide. The residuals for the fit are shown by the symbols along the dotted line at 0.0.

FIGURE 5. p85 α dimerization *in vivo* measured by fluorescence fluctuation spectroscopy (FFS).

A) U201S cells were transfected with GFP-p85 α or native GFP-p85 α . Brightness was measured as a function of concentration and compared to a monomeric GFP control. The data sets were not significantly different as indicated by their common regression line.

B) U201S cells were transfected with native GFP-p85 α or native GFP-p85 α PR1/M176A. Brightness was measured as a function of concentration and compared to a monomeric GFP control.

FIGURE 6. Structure and dynamics of the p85 α ¹⁻³³³ homodimer revealed through an integrative modeling approach.

A) The best-scoring multi-state model is composed of 3 major states with population weights of 40.3% (blue), 27.7% (red), and 16.6% (yellow), and 2 minor states with population weights of 8.0% (green) and 7.4% (purple). The most populated state (blue) was used as a reference for rigid body least-squares superposition of the remaining four states. The *ab initio* shape (represented as a gray envelope) computed from the experimental SAXS profile was also superposed for comparison.

B) Comparison of the experimental SAXS profile (black) of the p85 α ¹⁻³³³ dimer with the calculated SAXS profiles from the single-state ($\chi=13.90$, red) and the 5-state ($\chi=2.623$, blue) models. The lower plot shows the residuals (calculated intensity/experimental intensity) of each calculated SAXS profile. The upper inset shows the SAXS profiles in the Guinier plot with an R_g fit of 44.1 ± 0.62 Å. The maximum particle size (D_{max}) was ~ 150 Å (determined experimentally, Table 1).

C) Each of the 5 states in the multi-state model along with population weights and domain labels is shown. Colors were adjusted to distinguish individual domains in the dimer. The conformational dynamics of p85 α ¹⁻³³³ dimer appears to be dominated by the relative intra-molecular motions of the SH3 and BCR domains, connected by the PR1 motif linkers.

D) Consistency between the 25 DST chemical cross-links and the multi-state model of the p85 α ¹⁻³³³ dimer. The green dots represent cross-linked residue pairs satisfied by the multi-state model within the

distance threshold of 35 Å. The multi-state model of the p85 α ¹⁻³³³ dimer satisfied all 25 DST chemical cross-links.

FIGURE 7. Structure and dynamics of the full-length p85 α homodimer revealed through an integrative modeling approach.

A) The best-scoring multi-state model is composed of 3 major states with population weights of 33.2% (blue), 27.4% (red), and 18.2% (yellow), and 2 minor states with population weights of 13.4% (green) and 7.9% (purple). The most populated state (blue) was used as a reference for rigid body least-squares superposition of the remaining four states. The *ab initio* shape (represented as a gray envelope) computed from the experimental SAXS profile was also superposed for comparison.

B) Comparison of the experimental SAXS profile (black) of the p85 α dimer with the calculated SAXS profiles from the single-state ($\chi=1.671$, red) and the 5-state ($\chi=1.275$, blue) models. The lower plot shows the residuals (calculated intensity/experimental intensity) of each calculated SAXS profile. The upper inset shows the SAXS profiles in the Guinier plot with an R_g fit of 57.7 ± 0.9 Å. The maximum particle size (D_{max}) was ~ 200 Å (determined experimentally, Table 1).

C) Each of the 5 states in the multi-state model, along with population weights and domain labels is shown. Colors were adjusted to distinguish individual domains in the dimer. A large degree of heterogeneity was observed in the full-length p85 α dimer, particularly in the coiled-coil (iSH2) domains as well as the neighboring nSH2 and cSH2 domains. The two iSH2 domains are oriented in multiple conformations relative to one another (e.g., parallel, anti-parallel, and perpendicular).

D) Consistency between the combined 256 (25 DST and 231 DSS) chemical cross-links and the multi-state model of the full-length p85 α dimer. The green dots represent cross-linked residue pairs satisfied by the multi-state model within the distance threshold of 35 Å. The red dots represent cross-linked residue pairs that violated the distance threshold of 35 Å. The blue dots represent 5 homodimer chemical cross-links identified on the same residues between two subunits in the dimer. The multi-state model of p85 α dimer satisfied 244 (95%) of the combined 256 chemical cross-links.

TABLE 1. SUMMARY OF SAXS ANALYSIS.

Table 1 summarizes SAXS parameters of molecular weight ($M.W.$), radius of gyration (R_g), maximum particle size (D_{max}), and Porod volume calculated from SAXS profiles of p85 α full-length and p85 α ¹⁻³³³ samples, under the low salt condition (20 mM NaCl) at 10°C. The SAXS parameters obtained under the conditions favoring the dimer state are highlighted in gray scale. Radius of gyrations (R_g) were calculated using DATGNOM and AUTORG in the ATSAS package (89) in real and reciprocal space, respectively. Porod volumes were calculated using DATPOROD in the ATSAS package (89). Additional tables for other SAXS samples are available at <http://salilab.org/p85>.

* Molecular Weights ($M.W.$) were estimated using SAXS MOW (38) with a threshold of $Q_{max} = 0.2\sim 0.3$ ($1/\text{Å}$), depending on the data. Native molecular weights of the p85 α monomer and dimer are 84 and 168 kDa, respectively. In contrast, molecular weights of the p85 α ¹⁻³³³ monomer and dimer are 38 and 75 kDa, respectively.

† SAXS data has a higher noise at low concentrations (~ 0.5 mg/mL) than at high concentrations.

Table 1

| Protein | Concentration | | 20mM NaCl (10°C) | | | | | Native Molecular Weight | |
|-------------------------------|---------------|--------------|------------------------------------|---|--|--|---------------------------------------|-------------------------|----------------|
| | (mg/mL) | (μ M) | <i>MW</i> * estimation (kDa) | <i>D</i> _{max} (\AA) | <i>R</i> _g (<i>Real</i>) (\AA) | <i>R</i> _g (<i>Reciprocal</i>) (\AA) | Porod volume (\AA^3) | Monomer (kDa) | Dimer (kDa) |
| p85 α | 0.5 † | 6.0 † | 141.0 | 171.6 | 51.3 | 53.1 \pm 3.3 | 259,815 | 84 | 168 |
| | 1.0 | 12.0 | 176.8 | 199.1 | 59.5 | 57.7 \pm 0.9 | 374,029 | | |
| | 1.5 | 18.0 | 199.0 | 206.6 | 62.7 | 62.8 \pm 1.3 | 475,310 | | |
| | 2.0 | 24.0 | 207.1 | 209.7 | 63.0 | 63.7 \pm 0.7 | 490,793 | | |
| | 5.0 | 60.0 | 289.1 | 232.7 | 70.7 | 70.8 \pm 1.0 | 688,412 | | |
| p85 α ¹⁻³³³ | 0.5 † | 13.3 † | 60.0 | 151.5 | 43.3 | 37.3 \pm 3.3 | 81,499 | 38 | 75 |
| | 0.75 | 19.9 | 67.5 | 150.8 | 43.1 | 41.0 \pm 2.0 | 83,867 | | |
| | 1.0 | 26.5 | 69.2 | 154.3 | 44.1 | 43.8 \pm 1.5 | 87,460 | | |
| | 2.0 | 53.0 | 72.4 | 157.1 | 44.9 | 44.0 \pm 1.0 | 91,517 | | |
| | 3.5 | 92.8 | 77.5 | 161.2 | 46.3 | 46.3 \pm 1.2 | 97,002 | | |
| | 5.0 | 132.5 | 75.9 | 155.1 | 44.3 | 44.1 \pm 0.6 | 95,438 | | |

Figure 1

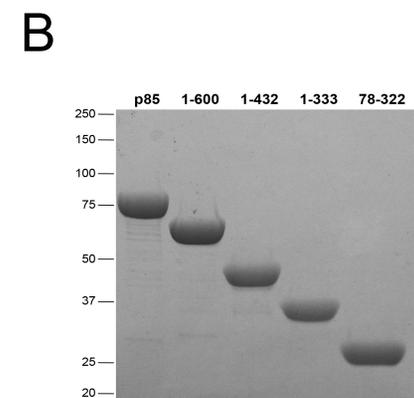
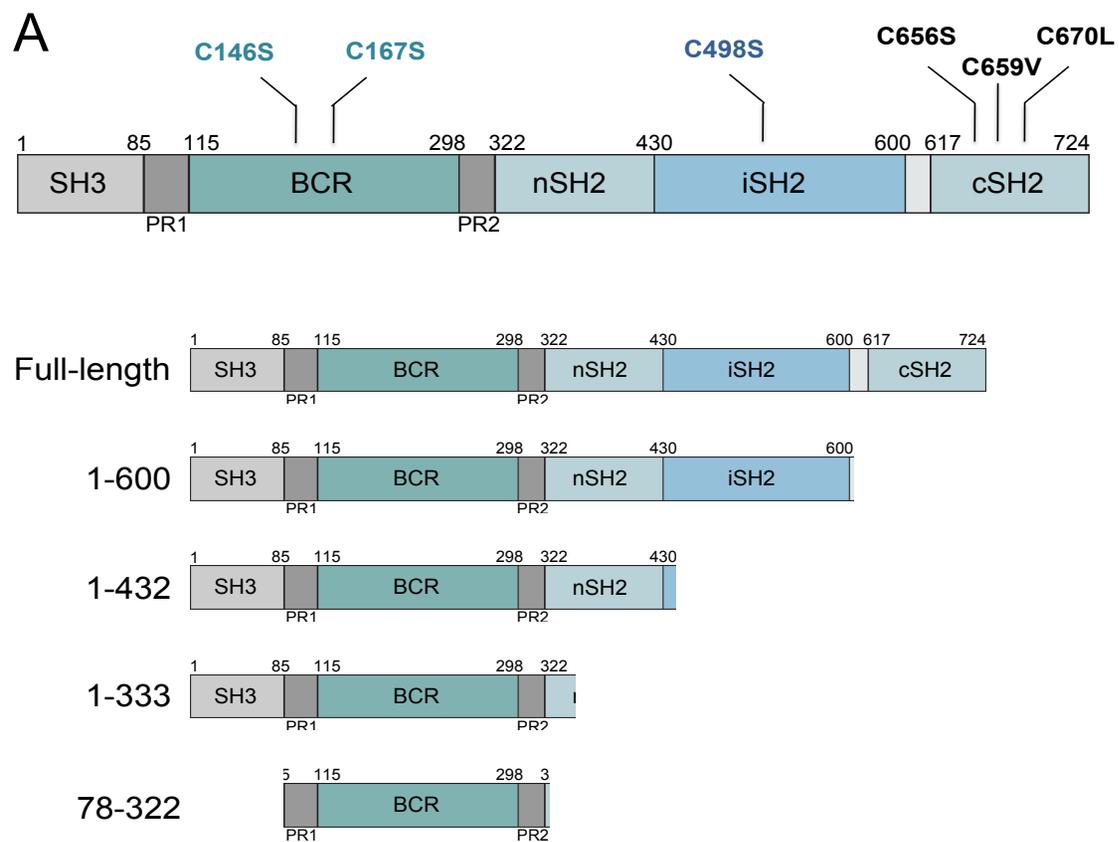


Figure 2

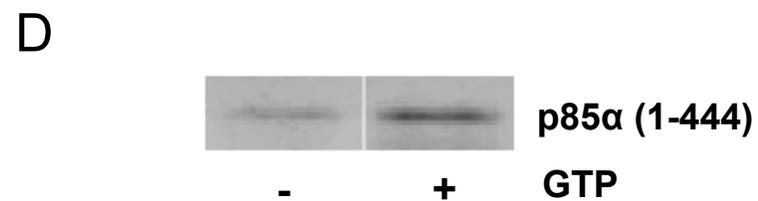
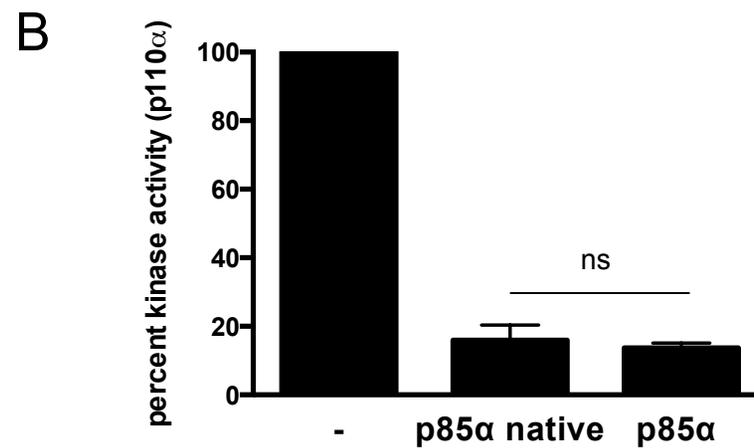
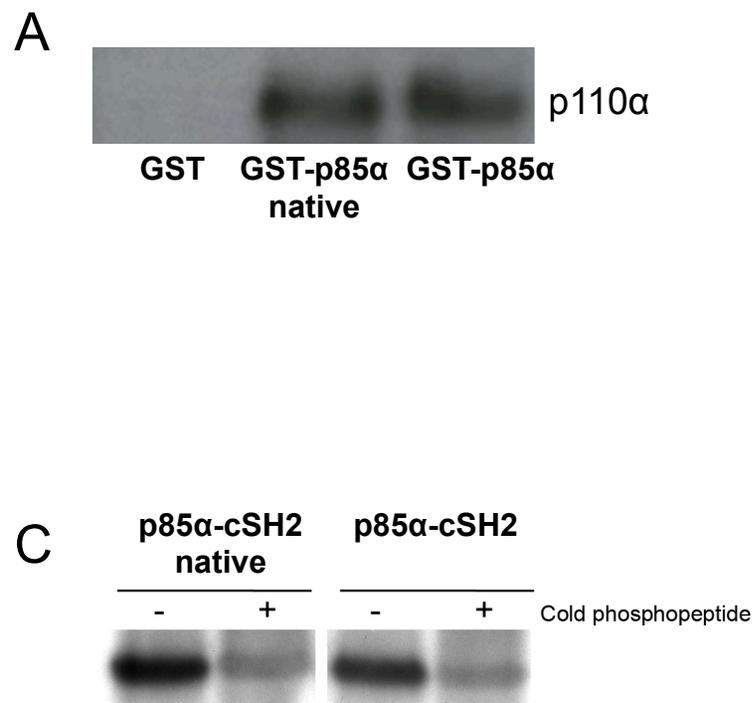


Figure 3

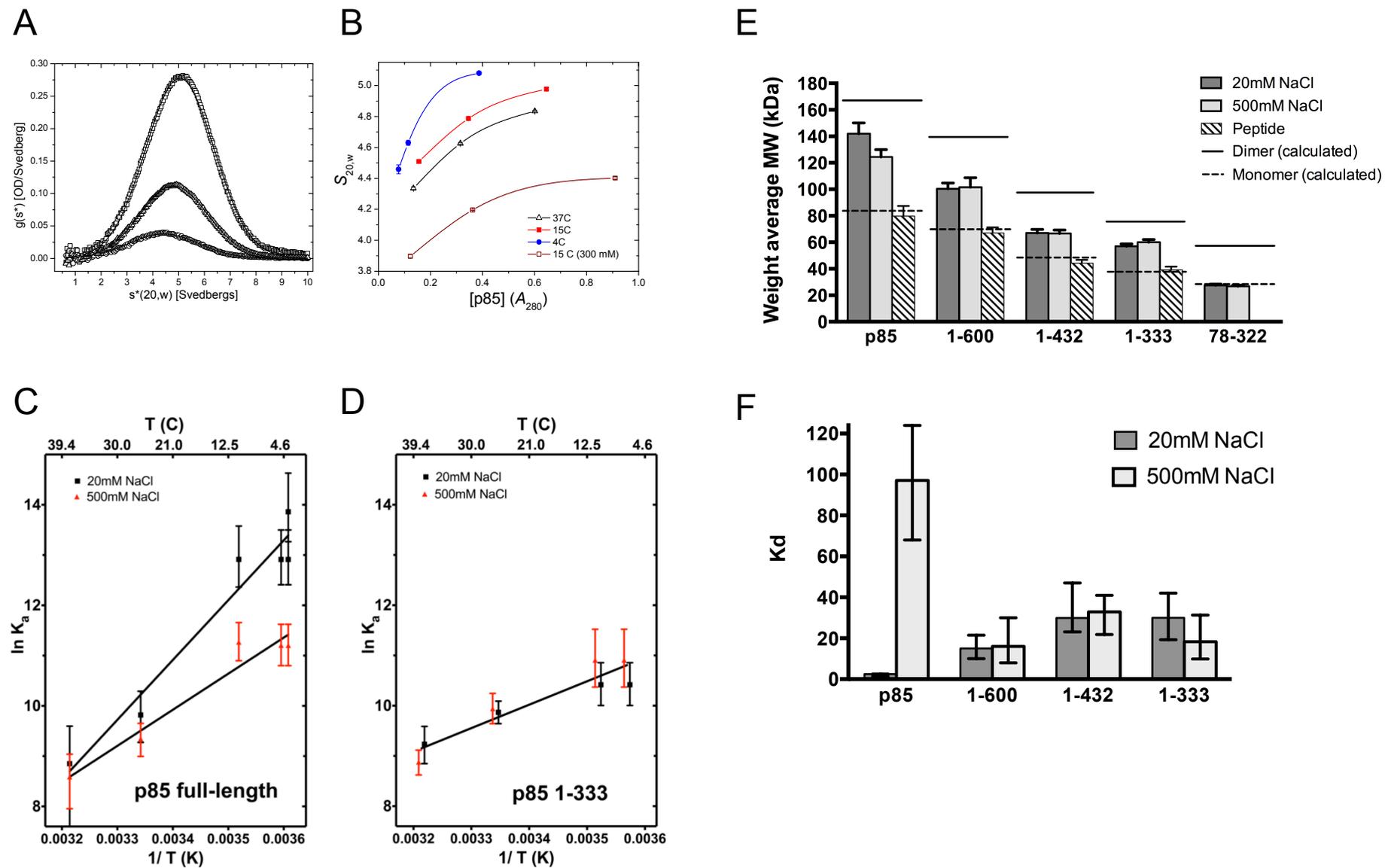


Figure 4

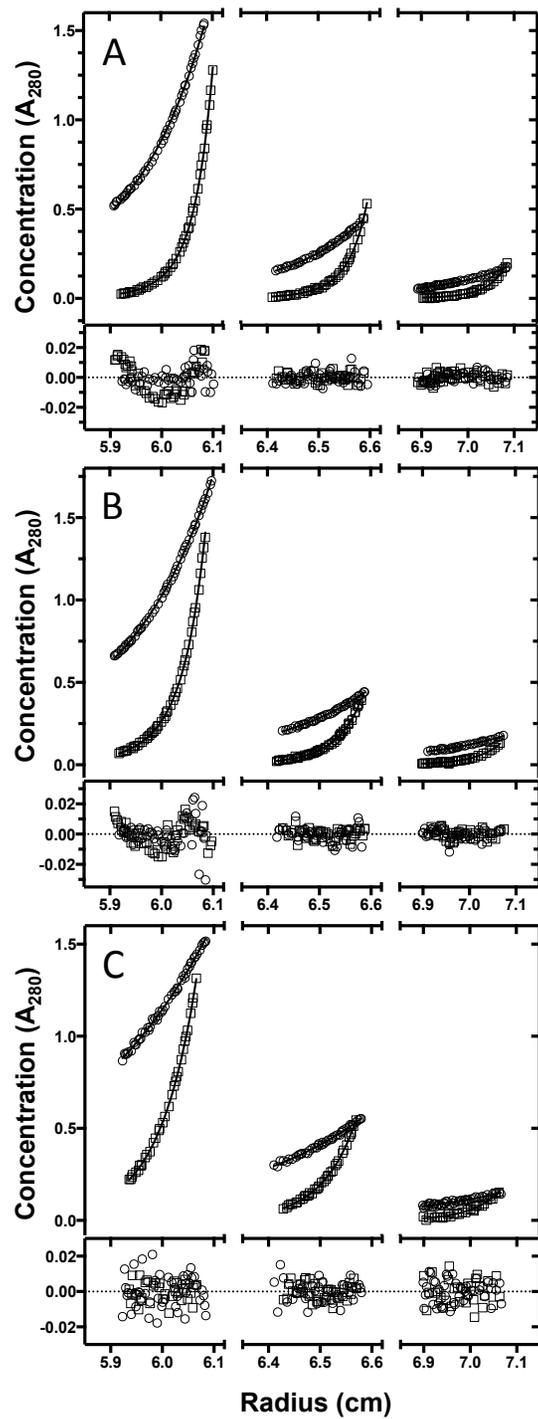


Figure 5

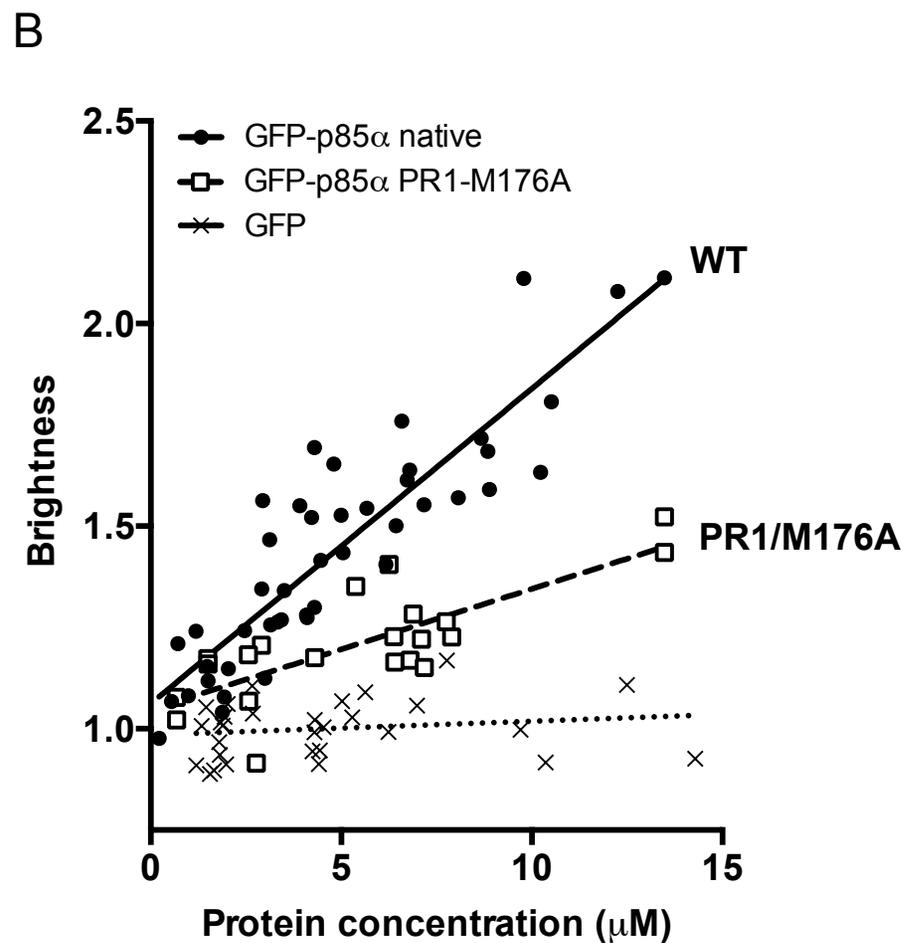
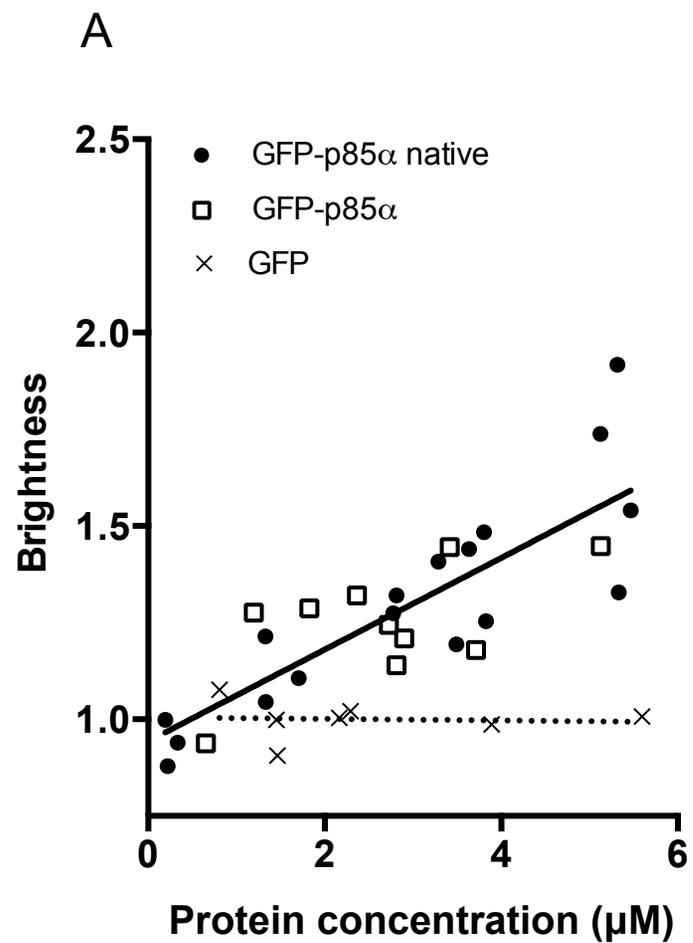


Figure 6

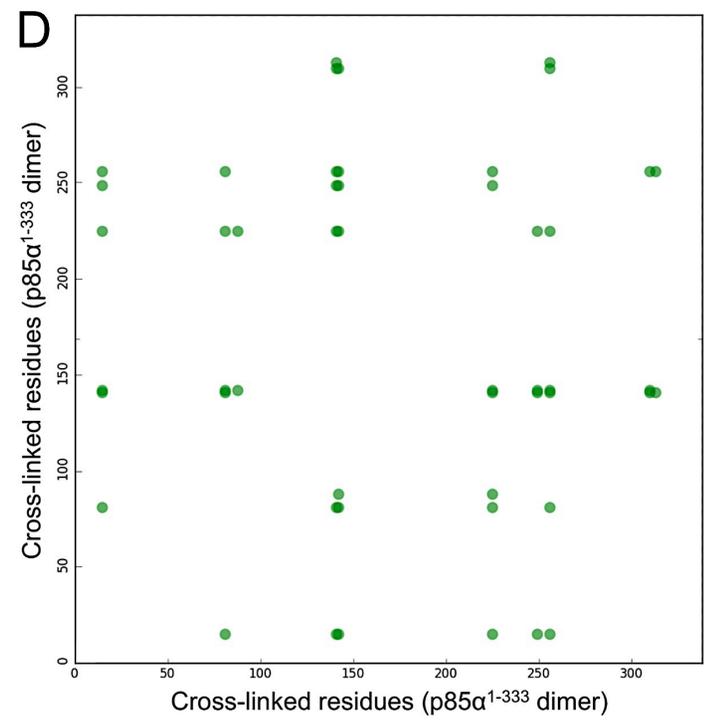
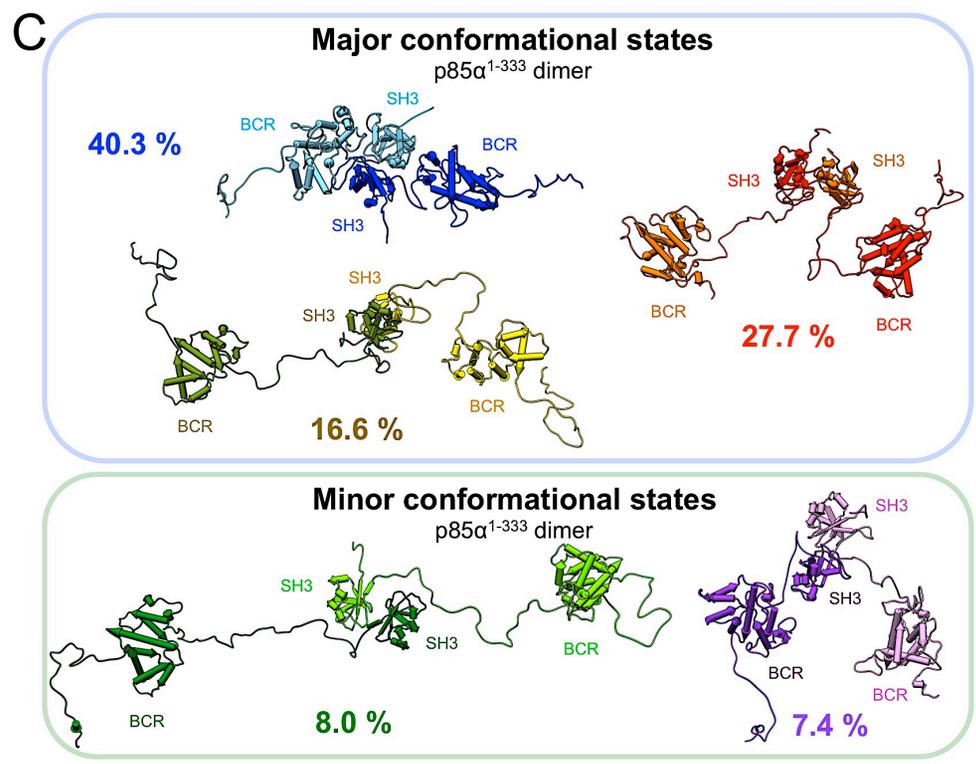
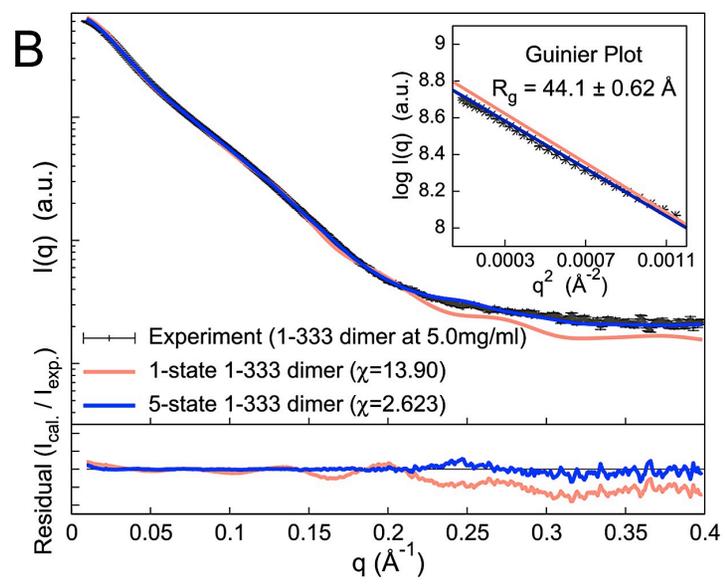
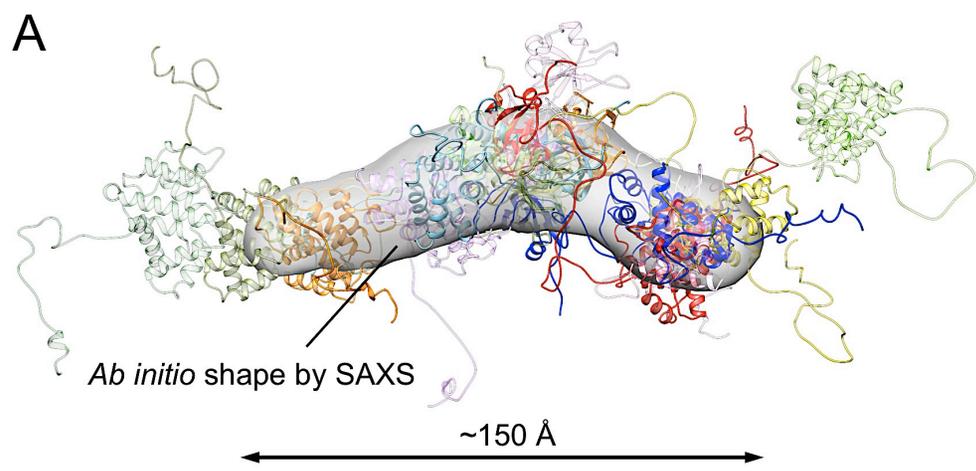


Figure 7

