Assembly and Molecular Architecture of the Phosphoinositide 3-Kinase p85α homodimer

Jaclyn LoPiccolo1,§, Seung Joong Kim4,§, Yi Shi5, Bin Wu3, Haiyan Wu1, Brian T. Chait5, Robert H. Singer3, Andrej Sali4, Michael Brenowitz2, Anne R. Bresnick2, and Jonathan M. Backer1,2

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

§ 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 p85α dimerizes by mechanisms not fully understood.

Results: p85α dimerization is driven by intermolecular interactions at both its N- and C-termini. A structural model of the p85α dimer reveals conformational diversity.

Conclusion: p85α dimerization is poised to link the heterologous assemblies that regulate PI3K signaling.

Significance: The p85α 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. p85α, 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 p85α 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 p85α that is functionally equivalent to native p85α. Analytical ultracentrifugation (AUC) studies showed that p85α 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 p85α 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 p85α dimer. Our study provides new insight into the structure and assembly of the p85α homodimer and suggests that this protein is a highly dynamic molecule whose conformational flexibility allows it to transiently associate with multiple binding proteins.

p85α (PIK3R1) is one of 5 regulatory subunits that inhibit and stabilize the p110 catalytic subunits of Class IA PI3 kinases (PI3K)(1). The p85α 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 p110α, p110β, and p110δ, 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 intramolecular and intermolecular interactions in addition to binding p110. p85α/p110 dimers are activated by binding of the p85α 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 p85β/p110 dimers (7). The iSH2 domains of both p85α and p85β bind the tumor suppressor BRD7, resulting in nuclear translocation of p85α and its sequestration from cytosolic p110 (8). Independently of p110, the N-terminus of p85α 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 p85α as well as the SH3-PR1-BCR fragment of p85α residues 1-333; Fig. 1B) has been reported, and peptides derived from the PR1 motif disrupt p85α dimerization (13). These results show that intermolecular SH3-PR1 interactions in the native protein are involved in p85α dimerization. In vivo dimerization has been demonstrated by reciprocal immunoprecipitation of differentially epitope-tagged p85α in several cell lines (11). A p85α mutation, identified in a human endometrial carcinoma, truncates p85α midway through the BCR domain at residue 160; expression of this mutant activates PI3K signaling by inhibiting homodimerization of endogenous p85α, thereby blocking its stabilization of PTEN. The pathological consequence of this mutation highlights the biological importance of the p85α dimer (11).

Because p85α 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 p85α 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 p85α, 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 p85α dimer. Our study provides new insight into p85α dimerization, suggesting that p85α 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 p85α—Wild-type human p85α 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 p85α. Hereafter, we refer to ‘cysteine-free p85α’ as p85α and the wild type protein as ‘native p85α’. Similarly, truncation mutants will be referred to as p85α with superscripts denoting the residues included in the fragment. The construct coding for p85α 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 μg/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 an 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. p85α 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 p85α sequence.

The resultant p85α 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 p85α 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 p85α (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 p85α were expressed and purified using the same protocol described for the full-length protein. For binding assays, native GST-p85α and GST-p85α were purified as described above except that p85α 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-p110α 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%o (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-p85α or native GST-p85α. 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 GTPγS (21) and incubated with rotation for 2 hr at 4 °C with p85α1-432. 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-p85α-nSH2 antibody and developed with ECL western blotting substrate (Pierce).

**Lipid kinase assay**—N-myc-p110α was immunopurified from transiently transfected HEK293T cells as above. Protein-G pellets were washed, incubated with bacterially purified GST-p85α or native GST-p85α, and assayed for lipid kinase activity as previously described (22).

**SH2-phosphopeptide binding**—GST-p85α-cSH2 and native GST-p85α-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 µM final, 3 x 10^6 CPM/assay) was incubated with 2 µg of GST-cSH2 domain in the absence or presence of 250 µM 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 (20, w) using these values.

For sedimentation velocity (SV) experiments, 300 µl of sample and an equal volume of buffer were loaded into two-sector cell assemblies. Three concentrations of each protein corresponding to A280 ≈ 0.1, 0.4, and 1.0 (equivalent to 1.2, 3.6, and 9.6 µM p85α 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 DCDT+ 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 µl 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 p85α EQ experiments were equilibrated at 8,000 and then 16,000 rpm; p85αL600 was equilibrated at 9,000 and then 16,000 rpm; p85αL432 was equilibrated at 10,000 and then 20,000 rpm; p85αL1-333 was equilibrated at 18,000 and then 25,000 rpm; p85αL76-322 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)**—p85α and native p85α were cloned into pEGFP-N1 using the EcoRI-BamHI sites. Native p85α was also used as a template for site-directed

---

4
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 $\lambda_{app}$ is measured by generalized Mandel’s Q parameter analysis (29). The brightness $\lambda_{EGFP}$ is obtained in calibration experiments by measuring cells transfected with EGFP.

**Chemical cross-linking analysis of p85α and p85α\(^{333}\) — p85α and p85α\(^{333}\) (Fig. 1B) were dialyzed into 20 mM HEPES pH 8.0 containing 20 mM NaCl, a solution condition that favors p85α dimerization (Fig. 3A). Twelve micromolar p85α was incubated with 0.5 mM disuccinimidyl suberate (DSS) for 15 min on ice with gentle agitation. Thirty micromolar p85α\(^{333}\) 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 p85α or p85α\(^{333}\) 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 p85α 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 μm 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 µm.

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=4πsin(θ)/λ, 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 Å⁻¹.

The ab initio shapes of the p85α₁,₃₃₃ and p85α 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 p85α and p85α₁,₃₃₃ dimers**—We computed ensembles of atomic multi-state models of the p85α₁,₃₃₃ dimer and the p85α 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 p85α, and 25 DST chemical cross-links for p85α₁,₃₃₃ were obtained as described above. HHpred (45,46) and ModWeb (47) identified the atomic structures of Homo sapiens p85α₂-82 [PDB, 315R (48), chain A], p85α₁,₁₁₅-₂₉₈ [PDB, 1PBW (49), chain A], p85α₂,₂₄₄₃₃ [PDB, 2IUG (50), chain A], p85α₂,₃₂₆₅₇₉ [PDB, 3HIZ and 3HHM (51), chain B], p85α₁,₃₃₃₄₃₃₁₀₀ [PDB, 2Y3A (52), chain B], p85α₂,₆₁₇₋₇₂ [PDB, 1IH9O (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 p85α dimer was built based on template structures (Stage 1) and SAXS profiles as follows. First, we built 100 atomic comparative models for p85α "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 p85α dimer. We added another copy of the p85α "monomer" at a random starting position for each sampling run, which resulted in an initial model of the p85α dimer. The SH3-binding PR1-like peptides were removed in the dimer model of p85α, to reflect the composition of the SAXS sample.

Next, a "monomer" model of p85α1-333 was obtained by removing residues of 334-724 in the best-scoring p85α "monomer" model (data not shown). Similarly, we added another copy of the p85α1-333 "monomer" at a random starting position for each sampling run, which resulted in an initial model of the p85α1-333 dimer. The SH3-binding PR1-like peptides were removed in the dimer model p85α1-333, 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 Cα 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 p85α, 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 p85α dimer and 25 DST chemical cross-links for the p85α1-333 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 p85α 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@p85α.1 − 519@p85α.1", "438@p85α.1 − 519@p85α.2", "438@p85α.2 − 519@p85α.1", and "438@p85α.2 − 519@p85α.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 p85α1-333 dimer, leading to a conclusion that the double intermolecular SH3-PR1 interactions are dominant in the p85α dimer.
Thus, we confined double intermolecular SH3-PR1 interactions in all restraint subsets of the p85α 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 p85α\(^{1-333}\) dimer (40.3% population) was further constrained during conformational sampling of the p85α 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 p85α and p85α\(^{1-333}\) 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 p85α\(^{1-333}\) dimer) and ~200,000 models (from 45 independent runs of the p85α 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 p85α\(^{1-333}\) and p85α 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 \(\chi\) 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 \(C_a - C_a\) 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 p85α\(^{1-333}\) and p85α 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 \(C_a - C_a\) 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 \(\chi\) 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 p85α—

Purification of p85α 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 p85α for use in site-specific spin labeling studies (74). Therefore, we mutated each of the six cysteines in p85α (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 “p85α”, and the wild type protein as “native p85α”. p85α robustly expresses and is readily purified in milligram quantities (10-15 mg/L bacterial culture). Similar protein yield and stability is achieved with the p85α fragments that were studied (Fig. 1B). Creation of cysteine-free p85α 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 p85α relative to the native protein. Using immunoprecipitation and kinase assays, we found the two proteins showed similar binding to p110α and inhibition of its kinase activity (Figs. 2A & B). The cSH2 domains of both p85α 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 p85α has no appreciable effect on its biochemical activity.

p85α dimerization in vitro—The ability of p85α to self-assemble in vitro has been documented (13). A question that motivated our analysis of p85α 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 solutio conditions under which p85α is predominantly monomeric or dimeric. We utilized two modes of AUC analysis to characterize p85α 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 (Kd) of p85α and its fragments.

SV experiments with p85α analyzed by the time-derivative method revealed single peaks (Fig. 3A), whose S20,w values increased with increasing protein concentration and decreased with increasing temperature and salt concentration (Fig. 3B). These observations reveal that p85α 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 p85α 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 μM (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 (ΔH° = -23.7 and -14.3 kcal, respectively; Fig. 3C). The salt concentration substantially affects p85α dimerization.

For example, at 10 °C the values of Kd measured at 20 and 500 mM NaCl differ by almost
40 fold (Fig. 3F). p85α 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 p85α is almost completely dimeric. Conversely, p85α 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 (RLPPRPGA) that binds the p85α SH3 domain (75). A 25-fold molar excess of peptide (250 μM peptide vs. 9.6 μM 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 p85α is an alternative condition for the analysis of monomeric protein.

The results presented above suggest that the intermolecular SH3-PR1 interaction dominates p85α dimerization. To explore possible contributions of other domains to dimerization we analyzed the assembly of a series of p85α truncations (depicted in Fig. 1A). SEQ analysis of the p85α fragment containing only SH3-PR1-BCR-PR2 (p85α1-335) reveals ten-fold weaker dimerization compared to the full-length protein (Fig. 3F; $K_d = 30.0$ vs. 2.5 μM at 20 mM NaCl and 10 °C), loss of salt dependence, and diminished enthalpy ($\Delta H^o = -6.7$ 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 p85α dimerization and is the source of the reaction’s salt dependence (Figs. 3E & F, compare p85α with p85α1-660). These data show that cSH2 also contributes to p85α 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 p85α corresponding to residues 78 – 322, which contains the BCR domain flanked by the two proline-rich motifs. SEQ analysis shows that p85α78-322 does not dimerize; an identical result was obtained for native p85α78-322, regardless of salt concentration or temperature for either construct (data not shown). These data show that the BCR domain alone does not dimerize.

$p85\alpha$ 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-p85α, GFP-p85α or a monomeric GFP control. The brightness of the GFP fluorescence correlates linearly and positively with increasing concentrations of the GFP-p85α fusion protein. In contrast, the brightness of the monomeric GFP control is invariant with concentration. The slopes of the correlations for native p85α and p85α are identical and appreciably greater than that for GFP alone control (Fig. 5A). These data demonstrate that both native p85α and p85α dimerize in vivo in a concentration-dependent fashion.

To test the contribution of the SH3-PR1-BCR domains to p85α dimerization in vivo, we compared the FFS distributions of native GFP-p85α to those of a mutant p85α in which two key proline residues in PR1 and a methionine residue in the BCR domain were mutated to alanine (p85αPR1/M176A). These residues were selected based on studies showing that these mutants inhibit co-immunoprecipitation of tagged p85α constructs in intact cells1. The slope of brightness versus protein concentration was more than two-fold greater in U201S cells transfected with native GFP-p85α as compared to native GFP-p85αPR1/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 p85α dimerization, an observation consistent with our in vitro analyses.

Small angle X-ray scattering (SAXS)—SAXS is a measure of p85α assembly orthogonal to AUC that provides information about protein size and

1 G.B. Mills, personal communication.
shape (76). SAXS profiles of p85α 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 p85α dimer at protein concentrations of 12 μM (1.0 mg/mL) at 10°C (Fig. 7B and Table 1) and 18 μM (1.5 mg/mL) at 25°C. At the highest protein concentration analyzed (5.0 mg/mL or ~60 μM, 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 p85α1–333 (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, p85α78–322 (PR1-BCR-PR2) does not dimerize under any condition, confirming that the intermolecular SH3-PR1 interaction is critical for dimerization. SAXS analysis of p85α1–432 and p85α1–600 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 p85α dimerization.

Chemical cross-linking of p85α and p85α1–333 dimers—Chemical cross-linking of p85α1–333 and p85α 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 p85α1–333 and p85α dimers (Figs. 6D and 7D). The overall connectivity pattern of the DST cross-links for p85α1–333 is similar to that of the DSS cross-links for the same portions of full length p85α 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 p85α1–333 structural model in modeling the full-length protein that is presented below.

In full-length p85α, 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 3HI4 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 p85α 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 p85α 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 p85α 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 p85α and p85α1–333 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 p85α dimers in solution (63,67). We began with the p85α1–333 dimer (Fig. 1A) and tested whether or not the resulting ~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), χ > 13.9 and only 48% of the DST cross-links were satisfied. This result suggests that the p85α1–333 dimer is conformationally heterogeneous in solution.

We therefore computed multi-state models of the p85α1–333 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 p85α1−333 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 (\(D_{\text{max}}\)) of the major states ranged from 145 to 185 Å. \(D_{\text{max}}\) 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 p85α1−333 dimer, the full-length p85α 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 p85α dimer for up to 9 states, using the most populated state of the p85α1−333 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 p85α 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 (\(D_{\text{max}}\)) 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 p85α1−333 dimer.

Based on the multi-state model, the conformational dynamics of the p85α1−333 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 p85α 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 p85α dimerization is also seen in the multi-state model (Fig. 7C). In conclusion, it appears that both the p85α1−333 and full-length p85α dimers are highly dynamic molecules in solution, held together by the SH3-PR1 and cSH2 contacts, with maximal dimension of the molecules in solution (\(D_{\text{max}}\)) ranging from 110 to 250 Å and from 170 to 320 Å, respectively.

**DISCUSSION**

The five discrete protein-binding domains of p85α 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 p85α actively regulate its interactions with other signaling proteins. The fact that p85α dimerization is rapidly and freely reversible and its conformational states are highly heterogeneous would maximize its ability to interact with regulatory partners. p85α
dimerization occurs \textit{in vivo} in a concentration-dependent manner, demonstrating that the properties of p85α measured \textit{in vitro} are relevant to those in the cell (Fig. 5). If p85α’s partners preferentially bind to the monomer or dimer, the intracellular concentration of p85α, 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 p85α variant in which its six cysteines are mutated to residues of a similar chemical nature (\textit{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 p85α 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 p85α is a validated model for structural and biophysical analysis of the p85α 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 p85α is predominantly dimeric. Using these conditions, we obtained chemical cross-linking and SAXS data that were used for integrative structure determination of the p85α dimer. The flexible, elongated shape of the p85α 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 p85α 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 p85α 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 p85α dimer stability in cells deserves further study.

Although crystal structures have been solved for the five individual domains of native p85α (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 p85α-mediated regulation of PI3K signaling. Thus, we explored the structure and dynamics of the p85α 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 p85α dimer that defines the conformations and populations of individual states (Figures 6C and 7C). An important result is the conformational heterogeneity of the p85α 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 p85α\textsubscript{1-333} dimer, as it is more amenable than p85α to chemical cross-linking experiments, yet still exhibits the monomer/dimer equilibrium of the full-length protein. The multiple states of the p85α\textsubscript{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 p85α\textsubscript{1-333} fragment.

We computed the multi-state model of the full-length p85α dimer using this SH3-BCR conformational state as a constraint, in conjunction with cross-linking and SAXS data. The conformational heterogeneity that characterizes the p85α 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 p85α 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 p85α dimer. This conformational flexibility may allow iSH2 to form heterologous interactions within the context of the p85α 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 p85α1-333 or p85α 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 p85α dimer (Fig. 3A).

A study recently published by Cheung et al has generated a model of the p85SH2-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 Kd measured for their His-PR1-BCR construct is 162.9 ± 41.4 µM, indicating very weak assembly. In contrast, we did not detect any dimerization of an untagged PR1-BCR-PR2 construct (p85α78-322) at three temperatures and two solution conditions. Not surprisingly, the presence of SH3-binding peptide had no effect on p85α78-322 oligomerization unlike its complete inhibition of dimerization of p85α 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 p85α dimer stabilization in the contact 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 p85α in its dimeric form.

Another issue for future work will be whether the same interactions that drive p85α homodimerization occur when p85α is bound to p110. Although there is one report of p85/p110 heterotetramers in cultured cells (88), these higher order complexes have not been reported during purification and gel filtration analysis of recombinant p85/p110 dimers. Structures of the nSH2-iSH2/p110α and iSH2-cSH2/p110β 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 p85α homodimer, it is not clear that binding of p110 to the iSH2 domain would prevent the formation of (p85/p110) homotetramers. 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 (P41RR01209), 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

Assembly of the p85α homodimer


Assembly of the p85a homodimer

Assembly of the p85α homodimer

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 p85α and domain-containing fragments.

A) The six cysteines of wild-type human p85α were mutated (C146S, C167S, C498S, C656S, C659V, and C670L) to generate cysteine-free recombinant p85α, heretofore referred to as p85α. Four 85α truncations were also prepared.

B) Coomassie-stained SDS-PAGE gel of purified p85α and fragments.

FIGURE 2. Characterization of cysteine-free p85α.

A) Immunoblot analysis of myc-p110α binding to GST alone, GST-p85α or native GST-p85α.

B) Kinase activity of p110α purified from HEK293T cells in the presence of GST-p85α or native GST-p85α. Data is the average ± SEM from three independent experiments.

C) Binding of GST-cSH2 domains from p85α or native p85α to a photoactivatable 125I-labeled tyrosine phosphopeptide, in the absence or presence of unlabeled phosphopeptide.

D) Immunoblot analysis of GST-p85α1-432 binding to recombinant Rac1 loaded without or with GTPγS.

FIGURE 3. p85α dimerization in vitro measured by analytical ultracentrifugation (AUC).

A) Sedimentation coefficient distribution, g(s*), obtained from sedimentation velocity analysis of p85α at three concentrations (1.2, 3.6, and 9.5 µM) in 20 mM NaCl containing buffer at 4 °C. All sedimentation coefficient (S, in Svedbergs) values presented in the paper have been corrected to 20 °C and the density of water (S20,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 p85α concentration (absorbance at 280 nm) at 4, 15, 37, and 15 °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 p85α (C) p85α1-333 (D) in 20 mM and 500 mM NaCl containing buffer.

E) Weight average molecular weights (kDa) obtained from sedimentation equilibrium analysis of p85α 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 p85α 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 p85α at 10 °C.**

The graphs show the distribution of p85α 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 p85α were analyzed (1.2, 3.6, and 9.6 µM). 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. p85α dimerization in vivo measured by fluorescence fluctuation spectroscopy (FFS).**

A) U201S cells were transfected with GFP-p85α or native GFP-p85α. 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-p85α or native GFP-p85α PR1/M176A. Brightness was measured as a function of concentration and compared to a monomeric GFP control.

**FIGURE 6. Structure and dynamics of the p85α₁-333 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 p85α₁-333 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 p85α₁-333 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 p85α₁-333 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 p85α\textsuperscript{1-333} dimer satisfied all 25 DST chemical cross-links.

**FIGURE 7. Structure and dynamics of the full-length p85α 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 \textit{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 p85α 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 (\(D_{\text{max}}\)) 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 p85α 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 p85α 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 p85α 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 (\(D_{\text{max}}\)), and Porod volume calculated from SAXS profiles of p85α full-length and p85α\textsuperscript{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 gyration (\(R_g\)) 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 \(Q_{\text{max}} = 0.2–0.3\) (1/Å), depending on the data. Native molecular weights of the p85α monomer and dimer are 84 and 168 kDa, respectively. In contrast, molecular weights of the p85α\textsuperscript{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>
<thead>
<tr>
<th>Protein</th>
<th>Concentration (mg/mL)</th>
<th>Concentration (µM)</th>
<th>20mM NaCl (10°C)</th>
<th>Native Molecular Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MW* estimation (kDa)</td>
<td>$D_{max}$ (Å)</td>
<td>$R_g$ (Real) (Å)</td>
</tr>
<tr>
<td>p85α</td>
<td>0.5†</td>
<td>6.0†</td>
<td>141.0</td>
<td>171.6</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>12.0</td>
<td>176.8</td>
<td>199.1</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>18.0</td>
<td>199.0</td>
<td>206.6</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>24.0</td>
<td>207.1</td>
<td>209.7</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>60.0</td>
<td>289.1</td>
<td>232.7</td>
</tr>
<tr>
<td>p85α¹⁻³³³</td>
<td>0.5†</td>
<td>13.3†</td>
<td>60.0</td>
<td>151.5</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>19.9</td>
<td>67.5</td>
<td>150.8</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>26.5</td>
<td>69.2</td>
<td>154.3</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>53.0</td>
<td>72.4</td>
<td>157.1</td>
</tr>
<tr>
<td></td>
<td>3.5</td>
<td>92.8</td>
<td>77.5</td>
<td>161.2</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>132.5</td>
<td>75.9</td>
<td>155.1</td>
</tr>
</tbody>
</table>
Figure 1

A

SH3  BCR  nSH2  iSH2  cSH2
PR1  PR2

C146S  C167S  C498S  C656S  C670L

1  85  115  298  322  430  600  617  724

Full-length

1-600

1-432

1-333

78-322

B

Figure 1
Figure 2

A

GST  GST-p85α  GST-p85α

native

B

percent kinase activity (p110α)

-  p85α native  p85α

C

p85α-cSH2

native  p85α-cSH2

-  +  -  +

Cold phosphopeptide

D

p85α (1-444)

-  +

GTP
Figure 3

A

B

E

C

D

F

20mM NaCl
500mM NaCl
Peptide
Dimer (calculated)
Monomer (calculated)

Weight average MW (kDa)

p85 1-600 1-432 1-333 78-322

Kd

20mM NaCl
500mM NaCl

p85 1-600 1-432 1-333
Figure 5

A

- **GFP-p85α native**
- **GFP-p85α**
- **GFP**

B

- **GFP-p85α native**
- **GFP-p85α PR1-M176A**
- **GFP**
- **PR1/M176A**
- **WT**
Figure 6

A. Ab initio shape by SAXS

B. Guinier Plot
\[ R_g = 44.1 \pm 0.62 \text{\AA} \]

C. Major conformational states
- p85α-333 dimer
  - 40.3%
  - 16.6%
  - 27.7%

D. Cross-linked residues (p85α-333 dimer)
- 8.0%
- 7.4%
Figure 7

A

Ab initio shape by SAXS

~200 Å

B

Guinier Plot

$R_g = 57.7 \pm 0.9 \text{ Å}$

Residual $l_{\text{cat}} / l_{\text{exp}}$

Experiment (dimer at 1.0 mg/ml)

1-state full-length dimer ($\chi = 1.671$)

5-state full-length dimer ($\chi = 1.275$)

C

Unaligned orientation of two coiled-coil (ISH2) domains

33.2 %

27.4 %

7.9 %

Perpendicular orientation of two coiled-coil (ISH2) domains

18.2 %

Parallel orientations of two coiled-coil (ISH2) domains

13.4 %

D

Cross-linked residues (full-length dimer) vs. Cross-linked residues (full-length dimer)