Assembly and Molecular Architecture of the p85α homodimer

Jaclyn LoPiccolo^{1,§}, Seung Joong Kim^{4,§}, Yi Shi⁵, Bin Wu³, Haiyan Wu¹, Brian T. Chait⁵, Robert 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: *Molecular Architecture of the p85α 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: phosphatidylinositide 3-kinase (PI 3-kinase), phosphatidylinositol signaling, analytical ultracentrifugation, small-angle X-ray scattering (SAXS), molecular modeling, structural model

SUPPLEMENTAL METHODS

X-rav Small angle scattering (SAXS)—SAXS measurements of full-length and truncated p85a (Figures 3B and 4B; Supplemental Table S1, S2, and S3) 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 (1). 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 (2). The sample was placed at 1.7 meter from a MX225-HE (Rayonix, USA) CCD detector with a binned pixel size of 293 µm by 293 µm.

Up to 24 one-second exposures were made for each of the protein samples and buffers maintained at 10-25°C. Each of the resulting diffraction images was scaled using the transmitted beam intensity, azimuthally integrated by SASTool (3), 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].

Integrative multi-state modeling (Stage 2) Representation of subunits and translation of the data into spatial restraints

Building an initial model of the p85a and $p85\alpha^{1-333}$ dimers—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 (Supplemental Table S1) as follows. First, we built 100 atomic comparative models for p85α "monomer" in complex with the SH3-binding PR1-like peptide (RPLPPRP-GA), using MODELLER 9.14 (4) based on the crystal structures and the closest template structures (Stage 1 and Supplemental Table S6). 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 (5,6). Then, these 100 models were ranked by the x value of the fit. Second, the best-scoring "monomer" model (with a lowest x value) 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 $p85\alpha^{1-333}$ was obtained by removing residues of 334-724 in the best-scoring $p85\alpha$ "monomer" model. Similarly, we added another copy of the $p85\alpha^{1-333}$ "monomer" at a random starting position for each sampling run, which resulted in an initial model of the $p85\alpha^{1-333}$ dimer. The SH3-binding PR1-like peptides were removed in the dimer model $p85\alpha^{1-333}$, to reflect the composition of the SAXS sample.

Coarse-grained representation and spatial restraints—Domains were coarse-grained using beads representing individual residues; 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. Different subsets of the spatial restraints and constraints (*i.e.*, restraint subsets) were used for different sampling runs, to improve sampling efficiency (Supplemental tables S4 and S5).

The cross-link restraints were applied to the corresponding bead pairs, taking into account the ambiguity resulting from two identical p85 α subunits in a dimer; an ambiguous cross-link restraint considers all possible pair-wise assignments. For example, a restraint between residues 438 and 519 is evaluated (7) on distances:

"438@p85a.<u>I</u> - 519@p85a.<u>I</u>",
"438@p85a.<u>I</u> - 519@p85a.<u>I</u>",
"438@p85a.<u>2</u> - 519@p85a.<u>I</u>", and
"438@p85a.<u>2</u> - 519@p85a.<u>I</u>", followed by scoring only the least violated distance.

The excluded volume restraints were applied to each bead, using the statistical relationship between the volume and the residue that it covered (8-10).

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 (8-11).

231 DSS chemical cross-links for the p85α dimer and 25 DST chemical cross-links for the p85α¹⁻³³³ dimer were used to restrain the distances spanned by the cross-linked residues, corresponding to the restraint subset (Tables S4 and S5). 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 (PDB: 1PRL) (12), each one residing in a different subunit of the dimer. First, both single (restraint subsets 1 and 2 in Table S4) and double (restraint sub-

sets 3 and 4 in Table S4) intermolecular SH3-PR1 interactions were evaluated through the multi-state search for the $p85\alpha^{1-333}$ dimer, leading to a conclusion that the double intermolecular SH3-PR1 interactions are dominant in the $p85\alpha$ dimer. Thus, we confined double intermolecular SH3-PR1 interactions in all restraint subsets of the $p85\alpha$ dimer (Table S5).

4 homo-dimer cross-links between residues 438-438, 480-480, 530-530, and 567-567 were transformed to upperharmonic distance restraints (up to 20 Å), enforcing the parallel intermolecular orientation of two iSH2 domains (restraint subsets 1, 2, 3, 10, and 11 in Table S5). 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 (restraint subsets 4, 5, 6, 12, and 13 in Table S5). 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 (all restraint subsets in Table S5).

Lastly, a most populated state of the $p85\alpha^{1-333}$ dimer (40.3% population) was further constrained during conformational sampling of the $p85\alpha$ dimer, in selected restraint subsets 10 to 15 but not restraint subsets 1 to 9 (Table S5).

(Stage 3) Conformational sampling to produce a most parsimonious multi-state model consistent with all available data and information

For each of the restraint subsets, 2 to 3 independent sampling calculations were performed, each one starting with a random initial configuration (Supplemental Tables S4 and S5). 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 (10).

The multi-state SAXS score (13) 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 (5,6). The side chains of whole residues in each state were reconstructed using PULCHRA 3.06 (14) 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 $C_{\alpha} - C_{\alpha}$ distance of the corresponding residue pairs was smaller than a distance threshold of 35 Å, considering restraint ambiguity (above).

SUPPLEMENTAL FIGURE LEGENDS

FIGURE S1. Sedimentation equilibrium analysis of p85 α using SH3-binding proline rich peptide.

A) Weight average molecular weights (in kDa) obtained from sedimentation equilibrium of p85 α alone or in the presence of either 1 mM or 10 mM SH3-binding peptide in 20 mM NaCl (left) or 500 mM NaCl (right) at 10 °C.

B) Weight average molecular weights (in kDa) obtained from sedimentation equilibrium of p85α with 0.25, 0.5, or 1 mM SH3-binding peptide in 500 mM NaCl at 25 °C.

FIGURE S2. Sedimentation equilibrium analysis of p85α at 10 °C.

The graphs show the distribution of $p85\alpha$ 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 $p85\alpha$ 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 fit to a monomer-dimer association model shown as the solid grey line. The residuals for the fit are shown by the symbols along the dotted line at 0.0.

FIGURE S3. Sedimentation equilibrium analysis of the p85α BCR domain.

Weight average molecular weights (kDa) obtained from sedimentation equilibrium of native and cysteine-free p85 α^{78-322} at 4 °C, 10 °C, and 37 °C. Dashed line indicates calculated MW of monomer (28.8 kDa).

FIGURE S4. Integrative multi-state modeling of the p85α dimers.

The approach proceeds through four stages: (1) gathering of data, (2) representation of subunits and translation of the data into spatial restraints, (3) conformational sampling to produce a 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) (8,15-20). Files for the input data, scripts, and output model structures in multiple states are available at http://salilab.org/p85.

FIGURE S5. Consistency between the chemical cross-links and the multi-state model of the $p85\alpha$ dimers

(Left) The green dots represent cross-linked residue pairs satisfied by the corresponding multistate 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 homo-dimer cross-linked residue pairs identified between two subunits in the dimer. As a result, the multi-state model of the p85 α ¹⁻³³³ dimer satisfied all 25 DST chemical cross-links (Top), and the multi-state model of the p85 α dimer satisfied 244 (95%) of the combined 256 (25 DST and 231 DSS) cross-links (Bottom).

(Right) Histograms of Euclidean C_{α} pair distances are shown for the chemical cross-links assessed against the multi-state models.

FIGURE S6. Consistency between the 25 DST chemical cross-links and the individual states of the $p85\alpha^{1-333}$ dimer

(1st and 2nd columns) Each of the 5 states in the multi-state model, along with population weights and domain labels, is shown. 25 DST chemical cross-links are visualized as blue (satisfied) and red (violated) lines connecting the corresponding residue pair in each state, using Xlink Analyzer (21). Colors were adjusted to distinguish individual domains in the dimer.

(3rd column) The green dots represent cross-linked residue pairs satisfied by the corresponding state of the $p85\alpha^{1-333}$ dimer 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 homo-dimer cross-linked residue pairs identified between two subunits in the dimer.

(4th column) Histograms of Euclidean C_{α} pair distances are shown for the chemical cross-links assessed against the corresponding state.

(5th column) We assessed the 25 DST chemical cross-links against the individual p85 α^{1-333} subunits in each of the 5 states. The diagonal insets show the cross-link evaluation maps within the same p85 α^{1-333} subunit itself (p85 α . $\underline{1}$ – p85 α . $\underline{1}$ or p85 α . $\underline{2}$ – p85 α . $\underline{2}$), thus exploring the intramolecular cross-links in the corresponding state. The off-diagonal insets show the cross-link evaluation maps between the opposing p85 α^{1-333} subunits (p85 α . $\underline{1}$ – p85 α . $\underline{2}$ or p85 α . $\underline{2}$ – p85 α . $\underline{1}$), thus exploring the inter-molecular cross-links in the corresponding state.

FIGURE S7. Consistency between the combined 256 (25 DST and 231 DSS) chemical cross-links and the individual states of the $p85\alpha$ dimer

(1st and 2nd columns) Each of the 5 states in the multi-state model, along with population weights and domain labels. Combined 256 chemical cross-links are visualized as blue (satisfied) and red (violated) lines connecting the corresponding residue pair in each state, using Xlink Analyzer (21). Colors were adjusted to distinguish individual domains in the dimer.

(3rd column) The green dots represent cross-linked residue pairs satisfied by the corresponding state of the p85 α dimer 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 homo-dimer cross-linked residue pairs identified between two subunits in the dimer.

(4th column) Histograms of Euclidean C_{α} pair distances are shown for the chemical cross-links assessed against the corresponding state.

(5th column) We assessed the combined 256 chemical cross-links against the individual p85 α subunits in each of the 5 states. The diagonal insets show the cross-link evaluation maps within the same p85 α subunit itself (p85 α . $\underline{1}$ - p85 α . $\underline{1}$ or p85 α . $\underline{2}$ - p85 α . $\underline{2}$), thus exploring the intramolecular cross-links in the corresponding state. The off-diagonal insets show the cross-link evaluation maps between the opposing p85 α subunits (p85 α . $\underline{1}$ - p85 α . $\underline{2}$ or p85 α . $\underline{2}$ - p85 α . $\underline{1}$), thus exploring the inter-molecular cross-links in the corresponding state.

SUPPLEMENTAL TABLE LEGENDS

For Tables S1-S3 - SAXS parameters obtained under the conditions favoring the dimer state were highlighted in brown; those obtained under monomer conditions in blue.

*Molecular weights (M.W.) were estimated using SAXS MOW (22) with a threshold of $Q_{max} = 0.2 \sim 0.3$ (Å⁻¹), depending on the data. †SAXS data has higher noise at low concentrations (~ 0.5 mg/mL) than at high concentrations.

TABLE S1. Summary of SAXS analysis for full-length p85a

Tables (A) - (C) summarize SAXS parameters of molecular weight (M.W.), radius of gyration (R_g), and maximum particle size (D_{max}) calculated from SAXS profiles of full-length p85 α , under various conditions; (A) in a low salt (20 mM NaCl) buffer at 10 and 25 °C; (B) in a high salt (500 mM) buffer at 25 °C; (C) in a high salt (500 mM NaCl) buffer at temperatures of 10°C and 25°C, with the SH3-binding PR1-like peptide.

At low salt, we observed predominantly p85 α dimer at protein concentrations of 12 μ M (1.0 mg/mL) at 10°C and 18 μ M (1.5 mg/mL) at 25°C. Saturating concentrations of the SH3-binding PR1-like peptide drove the equilibrium completely to monomer as measured by AUC.

TABLE S2. Summary of SAXS analysis for $p85\alpha^{1-333}$ and $p85\alpha^{78-322}$

Tables (A) - (B) summarize SAXS parameters of molecular weight (M.W.), radius of gyration (R_g), and maximum particle size (D_{max}) calculated from SAXS profiles of p85 α^{1-333} , under various conditions; (A) in a low salt (20 mM) buffer at 10 °C, with or without the SH3-binding PR1-like peptide; (B) in a high salt (500 mM) buffer at 25 °C, with or without the SH3-binding PR1-like peptide. p85 α^{1-333} dimerizes at sample concentrations higher than 3.5 mg/mL (92.8 μ M), while it remains in a monomer state with the SH3-binding PR1-like peptide.

Tables (C) - (D) summarize SAXS parameters of molecular weight (M.W.), radius of gyration (R_g), and maximum particle size (D_{max}) calculated from SAXS profiles of p85 α^{78-322} , under various conditions; (C) in a low salt (20 mM) buffer at 10 °C; (D) in a high salt (500 mM) buffer at 10 °C. p85 α^{78-322} remains in a monomer state under any condition, confirming that the intermolecular SH3-PR1 interaction contributes to dimerization.

TABLE S3. Summary of SAXS analysis for p85 α^{1-432} and p85 α^{1-600}

Tables (A) - (B) summarize SAXS parameters of molecular weight (M.W.), radius of gyration (R_g), and maximum particle size (D_{max}) calculated from SAXS profiles of p85 α^{1-432} , under various conditions; (A) in a low salt (20 mM) buffer at 10 °C, with or without the SH3-binding PR1-like peptide; (B) in a high salt (500 mM) buffer at 10 °C, with or without the SH3-binding PR1-like peptide. p85 α^{1-432} dimerizes at high sample concentrations of ~5.0 mg/mL (102.5 μ M), while it remains in a monomer state with the SH3-binding PR1-like peptide.

Tables (C) - (D) summarize SAXS parameters of molecular weight (M.W.), radius of gyration (R_g), and maximum particle size (D_{max}) calculated from SAXS profiles of p85 α^{1-600} , under various conditions; (C) in a low salt (20 mM) buffer at 10 °C, with or without the SH3-binding PR1-like peptide; (D) in a high salt (500 mM) buffer at 10 °C, with or without the SH3-binding PR1-like peptide. p85 α^{1-600} dimerizes at sample concentrations higher than 2.0 mg/mL (28.6 μ M), while it remains in a monomer state with the SH3-binding PR1-like peptide.

Tables (A) - (D) show 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 $p85\alpha$ dimerization.

TABLES S4 AND S5. Restraint subsets for the $p85\alpha^{1-333}$ dimer (Table S4) and the full-length $p85\alpha$ dimer (Table S5), respectively

Different subsets of the spatial restraints and constraints (*i.e.*, restraint subsets) were used for different sampling runs, to improve conformational sampling efficiency. The conformational sampling produced $\sim 80,000$ (from 8 independent runs of the p85 α^{1-333} dimer, Table S4) and $\sim 200,000$ models (from 45 independent runs of the p85 α dimer, Table S5), respectively.

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 (PDB: 1PRL) (12), each one residing in a different subunit of the dimer. First, both single (restraint subsets 1 and 2 in Table S4) and double (restraint subsets 3 and 4 in Table S4) 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 (Table S5).

4 homo-dimer cross-links of 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 (restraint subsets 1, 2, 3, 10, and 11 in Table S5). 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 (restraint subsets 4, 5, 6, 12, and 13 in Table S5). 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 (all restraint subsets in Table S5).

Lastly, a most populated state of the $p85\alpha^{1-333}$ dimer (40.3% population) was further constrained during conformational sampling of the $p85\alpha$ dimer, in selected restraint subsets 10 to 15, but not restraint subsets 1 to 9 (Table S5).

TABLE S6. Representation of the p85a domains for integrative multi-state modeling

The domains of the p85 α were coarse-grained using beads representing individual residues and arranged into either a rigid body (column 4) or a flexible string (column 5) based on the available crystallographic structures or comparative models. The PR1 and PR2 motifs, and 2 linkers between rigid bodies are highlighted in red in columns 1, 3, and 5. The atomic structures of the p85 α domains have been previously determined through X-ray crystallography and NMR (column 3). Initial comparative models were built using MODELLER 9.14 (4) as described above, based on these atomic template structures.

TABLE S7. Consistency between the 25 DST chemical cross-links and the multi-state model of the $p85\alpha^{1-333}$ dimer

Euclidean C_{α} pair distances of the 25 DST chemical cross-links were calculated for the multi-state model of the p85 α^{1-333} dimer. Only a shortest pair distance among the 5 states was used for assessment of an observed chemical cross-link in the multi-state model (See also cross-link evaluation maps and histograms in Supplemental Figures S5 and S6). The blue cells in the table represent chemical cross-links that were satisfied by the multi-state model or the corresponding individual state, within the distance threshold of 35 Å. In contrast, the red cells represent chemical cross-links that violated the distance threshold of 35 Å.

TABLE S8. Consistency between the combined 256 (25 DST and 231 DSS) chemical cross-links and the multi-state model of the full-length p85a dimer

Euclidean C_{α} pair distances of the 25 DST and 231 DSS chemical cross-links (combined in the table) were calculated for the multi-state model of the full-length p85 α dimer. Only a shortest pair distance among the 5 states was used for assessment of an observed chemical cross-link in the multi-state model (See also cross-link evaluation maps and histograms in Supplemental Figures S5 and S7). The blue cells in the table represent chemical cross-links that were satisfied by the multi-state model or the corresponding individual state, within the distance threshold of 35

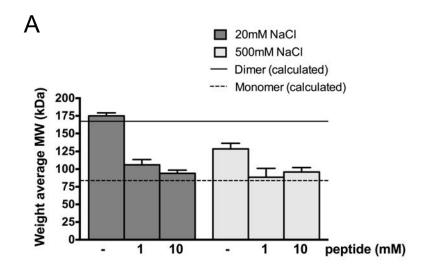
 $\hbox{Å}.$ In contrast, the red cells represent chemical cross-links that violated the distance threshold of 35 $\hbox{Å}.$

SUPPLEMENTAL REFERENCES

- 1. 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
- 2. 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
- 3. SASTool 2013.
- 4. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779-815
- 5. 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
- 6. 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
- 7. Erzberger, J. P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C. H., Cimermancic, P., Boehringer, D., Sali, A., Aebersold, R., and Ban, N. (2014) Molecular architecture of the 40SeIF1eIF3 translation initiation complex. *Cell* **158**, 1123-1135
- 8. Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprapto, 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
- 9. 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
- 10. 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
- 11. 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
- 12. 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
- 13. 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
- 14. Rotkiewicz, P., and Skolnick, J. (2008) Fast procedure for reconstruction of full-atom protein models from reduced representations. *J Comput Chem* **29**, 1460-1465
- 15. Alber, F., Forster, F., Korkin, D., Topf, M., and Sali, A. (2008) Integrating diverse data for structure determination of macromolecular assemblies. *Annual review of biochemistry* **77**, 443-477

- 16. 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
- 17. 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
- 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. Kosinski, J., von Appen, A., Ori, A., Karius, K., Muller, C. W., and Beck, M. (2015) Xlink Analyzer: software for analysis and visualization of cross-linking data in the context of three-dimensional structures. *Journal of structural biology* **189**, 177-183
- 22. 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

Figure S1



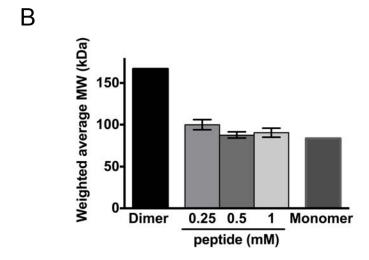


Figure S2

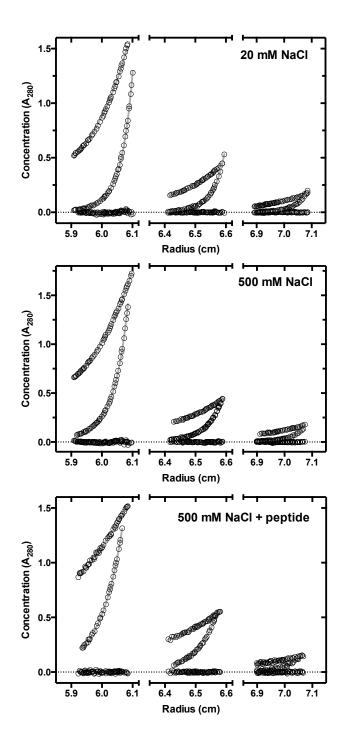
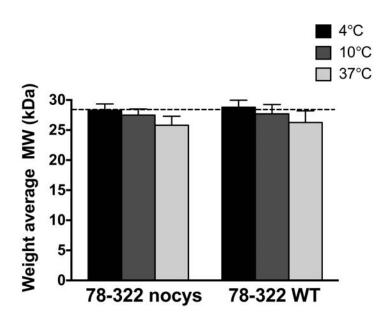


Figure S3



Experimental data

Statistical inference and physical principles

Steric

effect

2 units in a dimer

Excluded

Volume

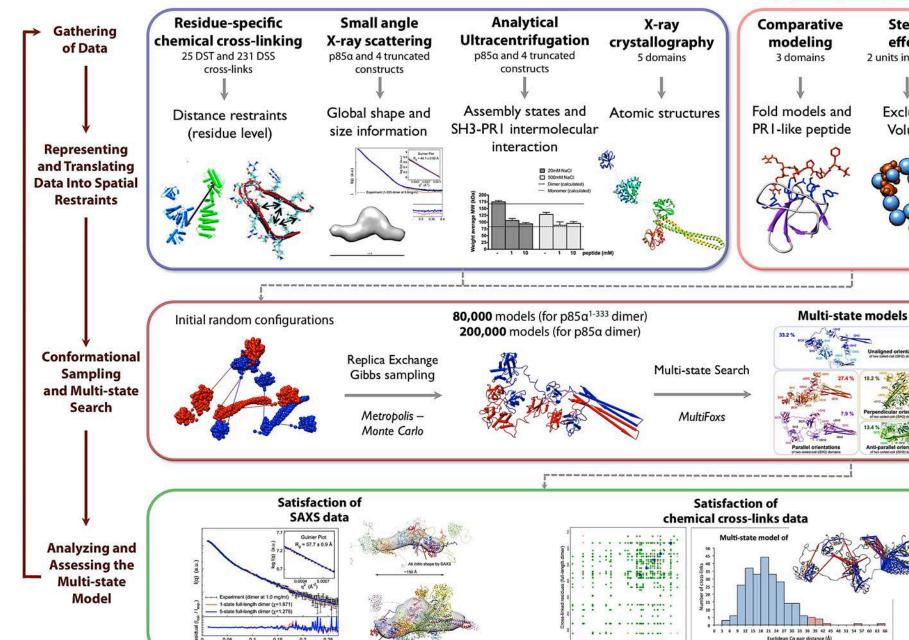
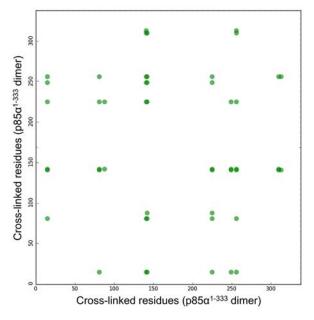
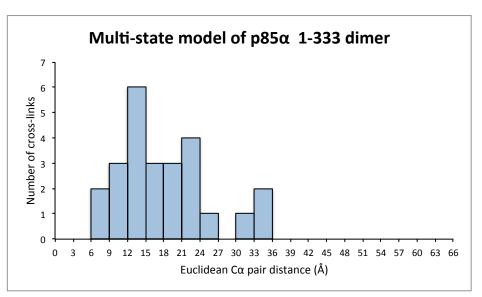


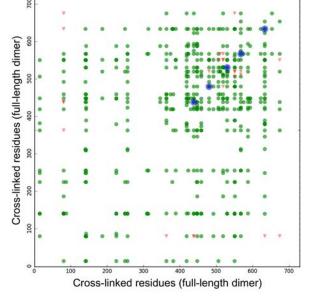
Figure S5

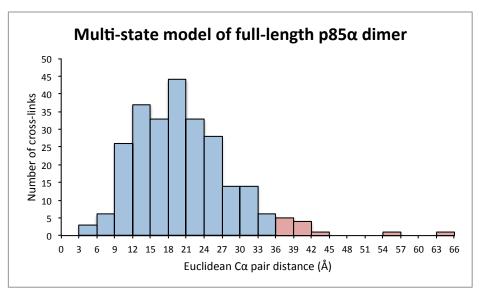
Multi-state model of 1-333 dimer

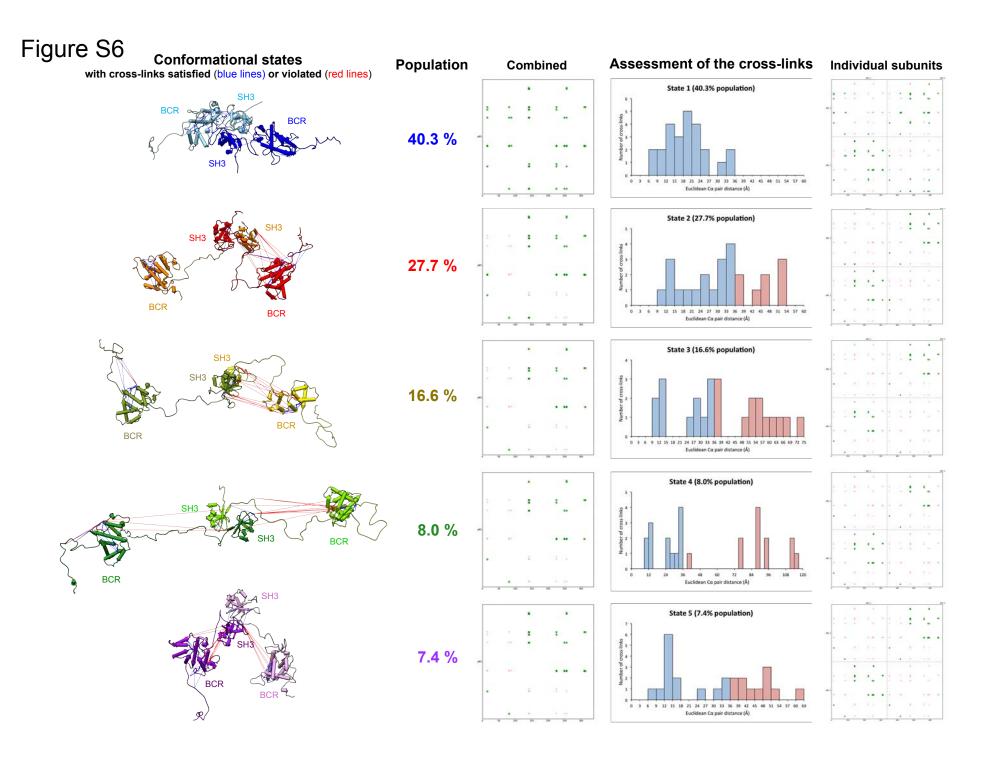




Multi-state model of Full-length dimer







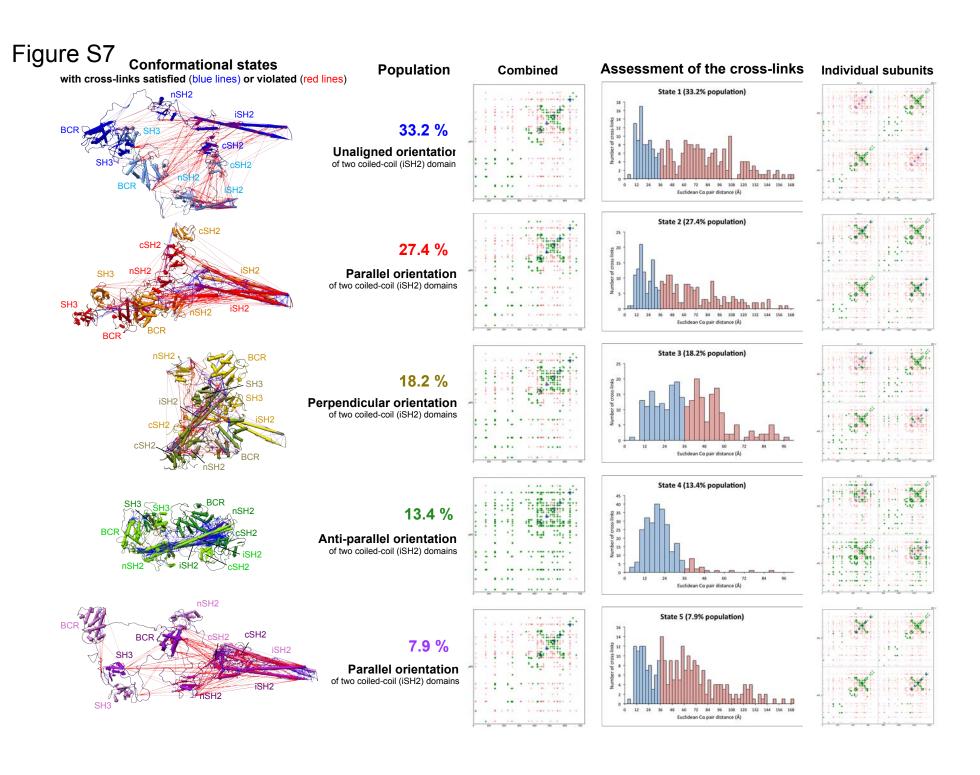


Table S1. Summary of SAXS analysis for the full-length p85α

Table A

Construct	N-CI	Concen	tration	10 °C				25 °C	M.W. from sequence		
Construct	NaCl	(mg/mL)	(µM)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	monomer	dimer
		0.5 †	6.0 †	141.0	171.6	51.3	131.2	189.8	56.6		
4 704		1.0	12.0	176.8	199.1	59.5	153.8	193.6	57.7		
1-724 full-length	20 mM	1.5	18.0	199.0	206.6	62.7	176.7	199.8	60.4	84 kDa	168 kDa
Tull-leligill		2.0	24.0	207.1	209.7	63.0	191.0	210.1	63.2		
		5.0	60.0	289.1	232.7	70.7	303.9	238.0	72.3		

Table B

Construct	NaCl	Concen	tration		25 °C		M.W. from	sequence
Construct	INACI	(mg/mL)	(µM)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	monomer	dimer
		0.5 †	6.0 †	96.4	174.7	54.0		
4 704	500 mM	1.0	12.0	107.4	192.2	58.9		
1-724 full-length		1.5	18.0	121.9	207.0	63.2	84 kDa	168 kDa
iuii-ieiigiii		2.0	24.0	120.3	202.8	63.2		
		5.0	60.0	139.0	221.1	66.9		

Table C

Construct	NaCl	Concen	tration	10 °C + SH3-binding peptide			25 °C + S	H3-binding	M.W. from sequence		
Construct	NaCi	(mg/mL)	(µM)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	monomer	dimer
		0.5 †	6.0 †	85.9	196.8	60.9	70.9	173.1	54.3		
4 704		1.0	12.0	84.1	178.7	55.8	80.0	175.9	54.6		
1-724 full-length	500 mM	1.5	18.0	92.3	191.8	57.1	83.9	190.1	55.8	84 kDa	168 kDa
Tull-leligill		2.0	24.0	90.3	183.7	55.9	88.1	184.1	55.4		
		5.0	60.0	104.9	195.2	58.2	97.6	196.9	58.3		

The SAXS parameters obtained under the conditions favoring the dimer state were highlighted in brown, while those obtained under the monomer condition in blue.

Radius of gyrations (Rg) were calculated in real space using DATGNOM in the ATSAS package.

^{*} Molecular Weights (M.W.) were estimated using SAXS MOW with a threshold of Qmax = 0.2~0.3 (1/Å), depending on the data.

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

Table S2. Summary of SAXS analysis for p85 α^{1-333} and p85 α^{78-322}

Т	a	bl	le	P

Construct	NaCl	Concen	tration	10 °C			10 °C + SH3-binding peptide			M.W. from sequence	
Construct	NaCi	(mg/mL)	(µM)	M.W. (kDa)*	Dmax (Å)	Rg (Å)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	monomer	dimer
		0.5 †	13.3 †	60.0	151.5	43.3	38.3	120.9	34.6		
		0.75	19.9	67.5	150.8	43.1					
1-333	20 mM	1.0	26.5	69.2	154.3	44.1	37.5	121.1	34.9	38 kDa	75 kDa
1-333	20 IIIIVI	2.0	53.0	72.4	157.1	44.9	40.8	124.5	34.9	30 KDa	75 KDa
		3.5	92.8	77.5	161.2	46.3	45.5	123.8	35.4		
		5.0	132.5	75.9	155.1	44.3	46.4	122.4	35.0		

Table B

_	Table B											
	Construct	NaCl	Concen	tration	25 °C			25 °C + S	H3-binding	M.W. from sequence		
	Construct	Naci	(mg/mL)	(µM)	M.W. (kDa)*	Dmax (Å)	Rg (Å)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	monomer	dimer
			0.5 †	13.3 †	64.4	158.1	43.7					
			0.75	19.9	57.6	164.0	46.1	33.7	128.5	37.3		
	1-333	500 mM	1.0	26.5				36.1	139.9	38.2	38 kDa	75 kDa
	1-333	500 IIIIVI	2.0	53.0	69.2	156.0	44.6	42.8	128.5	36.7	JO KDA	75 KDa
			3.5	92.8	75.9	167.0	47.7	47.1	139.8	39.9		
			5.0	132.5	80.9	170.5	48.7	49.9	142.9	40.8		

Table C

Construct	NaCl	Concen	tration		10 °C		M.W. from	sequence
Construct	INACI	(mg/mL)	(µM)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	monomer	dimer
		0.5 †	17.4 †	31.0	75.8	24.8		
		1.0	34.7	31.2	88.8	26.3		
78-322	20 mM	1.5	52.1	31.7	83.7	25.8	28 kDa	55 kDa
		2.0	69.4	33.3	129.8	28.1		
		5.0	173.5	34.2	100.3	28.1		

Table D

Construct	NaCl	Concen	tration		10 °C		M.W. from	sequence
Construct	INACI	(mg/mL)	(µM)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	monomer	dimer
		0.5 †	17.4 †	31.7	83.3	26.6		
		1.0	34.7	32.0	101.7	28.0		ı
78-322	500 mM	1.5	52.1	31.2	97.3	27.9	28 kDa	55 kDa
		2.0	69.4	30.8	87.4	26.9		
		5.0	173.5	31.8	95.7	27.4		

The SAXS parameters obtained under the conditions favoring the dimer state were highlighted in brown, while those obtained under the monomer condition in blue.

Radius of gyrations (Rg) were calculated in real space using DATGNOM in the ATSAS package.

^{*} Molecular Weights (M.W.) were estimated using SAXS MOW with a threshold of Qmax = 0.2~0.3 (1/Å), depending on the data.

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

Table S3. Summary of SAXS analysis for p85 α^{1-432} and p85 α^{1-600}

٦	Гα	h	le	Δ
	ıa	υ	ᆫ	_

Construct	NaCl	Concen	tration	10 °C			10 °C + S	H3-binding	M.W. from sequence		
Construct	Naci	(mg/mL)	(µM)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	monomer	dimer
		0.5 †	10.3 †	68.2	137.7	42.7	51.0	135.0	39.6		
		1.0	20.5	74.7	157.9	47.2	51.1	134.9	39.5		
1-432	20 mM	1.5	30.8	78.3	166.2	48.8	51.3	134.1	39.3	49 kDa	98 kDa
		2.0	41.0	81.4	172.1	49.2	51.4	137.6	39.7		
		5.0	102.5	93.9	186.5	53.4	52.5	138.1	40.1		

Table B

Construct	NaCl	Concen	tration	10 °C			10 °C + S	H3-binding	M.W. from sequence		
Construct	INACI	(mg/mL)	(µM)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	monomer	dimer
		0.5 †	10.3 †	72.4	173.5	50.5	48.1	132.3	42.3		
		1.0	20.5	66.7	166.3	51.1	51.4	153.0	45.7		
1-432	500 mM	1.5	30.8	72.3	181.1	54.0	53.5	158.8	46.0	49 kDa	98 kDa
		2.0	41.0	77.0	181.7	54.4	53.0	157.1	46.7		
		5.0	102.5	84.7	188.7	56.4	51.8	160.5	45.9		

Table C

Construct	NaCl	Concen	tration	10 °C			10 °C + S	H3-binding	M.W. from sequence		
Construct	NaCi	(mg/mL)	(µM)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	monomer	dimer
		0.5 †	7.2 †	97.3	159.4	48.8	69.3	138.4	44.5		
		1.0	14.3	109.1	178.0	54.0	79.8	167.0	48.6		
1-600	20 mM	1.5	21.5	126.1	195.1	58.1	82.9	176.7	50.0	70 kDa	140 kDa
		2.0	28.6	130.5	203.1	59.2	85.2	181.2	51.2		
		5.0	71.5	175.8	234.9	69.1	93.9	188.1	54.1		

Table D

Construct	NaCl	Concentration			10 °C		10 °C + S	H3-binding	M.W. from sequence		
Construct	INACI	(mg/mL)	(µM)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	<i>M.W.</i> (kDa)*	Dmax (Å)	Rg (Å)	monomer	dimer
		0.5 †	7.2 †	90.4	188.3	57.7	83.3	199.2	56.3		
		1.0	14.3	113.9	213.2	63.6	81.2	181.8	53.8		
1-600	500 mM	1.5	21.5	117.2	223.0	65.6	84.5	191.7	56.2	70 kDa	140 kDa
		2.0	28.6	121.2	221.7	65.6	82.5	187.9	55.9		
		5.0	71.5	135.8	236.5	70.5	83.8	197.3	57.5		

The SAXS parameters obtained under the conditions favoring the dimer state were highlighted in brown, while those obtained under the monomer condition in blue.

Radius of gyrations (Rg) were calculated in real space using DATGNOM in the ATSAS package.

^{*} Molecular Weights (M.W.) were estimated using SAXS MOW with a threshold of Qmax = 0.2~0.3 (1/Å), depending on the data.

 $[\]dagger$ SAXS data has a higher noise at low concentrations (~0.5 mg/mL) than at high concentrations.

Table S4. Restraint subsets for the p85 $\alpha^{1\text{-}333}$ dimer

Subset index	Sampling index	DST 1-333 XLs	Single SH3-PR1	Double SH3-PR1	Number of runs	Models generated per each run	Models generated in total
1	1-2		Y		2	9,800	19,600
2	3-4	Υ	Y		2	9,800	19,600
3	5-6			Y	2	9,800	19,600
4	7-8	Υ		Υ	2	9,800	19,600

Sum 8 78,400

Table S5. Restraint subsets for the full-length p85 α dimer

Subset index	Sampling index	DST 1-333 XLs	DSS 1-724 XLs	Parallel iSH2 XLs	Anti-parallel iSH2 XLs	Homodimer cSH2 XL	Starting Model	Double SH3-PR1	Number of runs	Models generated per each run	Models generated in total
1	1 - 3			Υ		Υ	Random	Υ	3	10,000	30,000
2	4 - 6	Υ		Υ		Υ	Random	Y	3	1,000	3,000
3	7 - 9		Υ	Υ		Υ	Random	Υ	3	500	1,500
4	11 - 13				Υ	Υ	Random	Y	3	10,000	30,000
5	14 - 16	Υ			Υ	Υ	Random	Υ	3	1,000	3,000
6	17 - 19		Υ		Υ	Υ	Random	Y	3	500	1,500
7	21 - 23					Υ	Random	Y	3	10,000	30,000
8	24 - 26	Υ				Υ	Random	Υ	3	1,000	3,000
9	27 - 29		Υ			Υ	Random	Υ	3	500	1,500
10	31 - 33			Υ		Υ	1-333, 40.3%	Υ	3	10,000	30,000
11	37 - 39		Υ	Υ		Υ	1-333, 40.3%	Y	3	500	1,500
12	41 - 43				Y	Υ	1-333, 40.3%	Y	3	10,000	30,000
13	47 - 49		Υ		Υ	Υ	1-333, 40.3%	Y	3	500	1,500
14	51 - 53					Υ	1-333, 40.3%	Y	3	10,000	30,000
15	57 - 59		Υ			Υ	1-333, 40.3%	Υ	3	500	1,500

Sum 45 198,000

Table S6. Representation of the $p85\alpha$ domains for integrative multi-state modeling

		11 Seg	gments
5 Domains	Representations	5 structured segments	6 disordered segments
		begins ends at at	begins ends at at
	1-1: flexible string of a bead (DISOPRED) 2-82: Rigid Body,	2 - 82	1 - 1
2	PDB, 3I5R_A (100% seq identity) PDB, 1PRL_C (30% seq identity)		
	83-116: flexible string of beads (DISOPRED)		83 - 116
	117-298: Rigid Body, PDB, 1PBW_A (100% seq id)	117 - 298	
	299-323: flexible string of beads (DISOPRED)		299 - 323
	324-427: Rigid Body, PDB, 2IUG_A (100% seq id) PDB, 3HIZ_B (100% seq id) PDB, 3HHM_B (100% seq id)	324 - 427	
	428-438: flexible string of beads (DISOPRED)		428 - 438
Simmer Charles and	439-599: Rigid Body, PDB, 2V1Y_B (100% seq id) PDB, 3HIZ_B (100% seq id) PDB, 3HHM_B (100% seq id) PDB, 2Y3A_B (73% seq id)	439 - 599	
	600-615: flexible string of beads (DISOPRED)		600 - 615
	616-720: Rigid Body, PDB, 1H9O_A (100% seq id) PDB, 2Y3A_B (73% seq id) 721-724: flexible string of beads (DISOPRED)	616 - 720	721 - 724
	5 Domains The state of the sta	1-1: flexible string of a bead (DISOPRED) 2-82: Rigid Body, PDB, 3ISR_A (100% seq identity) PDB, 1PRL_C (30% seq identity) PDB, 1PRL_C (30% seq identity) PDB, 1PRL_C (30% seq identity) R3-116: flexible string of beads (DISOPRED) 117-298: Rigid Body, PDB, 1PBW_A (100% seq id) PDB, 3HPBW_A (100% seq id) PDB, 3HIZ_B (100% seq id) PDB, 2Y3A_B (73% seq id) FDB, 2Y3A_B (73% seq id) R1H9O_A (100% seq id) PDB, 2Y3A_B (73% seq id)	1-1: flexible string of a bead (DISOPRED)

Table S7. Consistency between the 25 DST chemical cross-links and the multi-state model of the p85 α^{1-333} dimer

	25 DS	T cross-lir	nked pairs	Min	Minimum Cα pair distance [Å] for each state								
	Residue 1	Residue 2	Shortest Pair Distance [Å] in the Multi- state Model	State 1	State 2	State 3	State 4	State 5					
	15	81	11.1	26.8	28.4	11.1	11.5	15.8					
	15	225	21.6	21.6	43.9	59.3	77.7	41.8					
	81	225	20.9	20.9	51.6	73.0	94.1	38.8					
	88	225	19.3	19.3	53.1	48.5	77.1	44.8					
	141	15	21.9	21.9	31.2	52.8	89.6	50.3					
	141	81	21.3	21.3	35.1	56.8	112.0	36.1					
	141	225	32.9	32.9	32.9	32.9	32.9	32.9					
	142	15	20.8	20.8	33.1	51.4	88.7	53.8					
p85a	142	81	17.5	17.5	38.2	56.1	111.4	39.2					
1-333	142	88	13.5	13.5	51.1	38.9	89.3	47.1					
dimer	142	225	34.8	34.8	34.8	34.8	34.8	34.8					
aimer	249	15	15.5	15.5	37.8	62.3	89.7	48.5					
	249	141	11.6	11.6	11.6	11.6	11.6	11.6					
	249	142	13.4	13.4	13.4	13.4	13.4	13.4					
	249	225	24.3	24.3	24.3	24.3	24.3	24.3					
	256	15	11.6	11.6	45.4	63.3	94.0	60.3					
	256	81	22.2	22.2	46.2	67.3	114.9	48.6					
	256	141	14.3	14.3	14.3	14.3	14.3	14.3					
	256	142	13.2	13.2	13.2	13.2	13.2	13.2					
	256	225	34.6	34.6	34.6	34.6	34.6	34.6					
	310	141	13.1	21.0	16.4	36.5	34.6	13.1					
	310	142	12.6	19.7	18.5	36.8	35.6	12.6					
	310	256	7.0	7.7	23.7	34.3	24.9	7.0					
	313	141	17.0	17.8	24.1	28.5	40.9	17.0					
	313	256	6.8	6.8	30.1	29.4	29.3	14.2					

Table S8. Consistency between the combined 256 (25 DST and 231 DSS) chemical cross-links and the multi-state model of the full-length p85α dimer

	Combined 2 231 DSS) cr			Min	Minimum Cα pair distance (Å) for each state Combined 256 (25 DST and 231 DSS) cross-linked pairs I Shortast Pay								Min	Minimum Cα pair distance [Å] for each state				
		sidue 2	Shortest Pair Distance [Å] in the Multi- state Model	State 1	State 2	State 3	State 4	State 5		Residue 1	Residue 2	Shortest Pair Distance [Å] in the Multi- state Model	State 1	State 2	State 3	State 4	State :	
Full- length pB5a dimer	81	11 25 25 25 15 15 25 25 25 15 15 25 25 25 15 15 25 25 25 15 15 25 25 25 15 15 25 25 25 15 15 25 25 25 15 15 25 25 25 15 15 25 25 25 15 15 25 25 25 15 15 25 25 25 25 15 15 25 25 25 25 15 15 25 25 25 25 25 25 25 25 25 25 25 25 25	114 114 115	202 213 214 214 214 214 214 214 214 214 214 214	265 242 243 244 244 245 245 245 245 245 245 245 245	202 216	28.5 28.6 28.6 28.6 28.6 28.6 28.6 28.6 28.6	18.8 18.8	Full-length p85a dimer	448 448 448 448 448 448 448 448 448 448	423 480 480 480 480 480 480 480 480 480 480	102 102	45.6 45.2 45.2 45.2 45.2 45.2 45.2 45.2 45.2	202 402 403 403 403 403 403 403 403 403 403 403	407 202 203 203 203 203 203 203 203 203 203	182 267 267 267 267 267 267 267 267 267 26	38.6 41.2 31.5 31.5 31.5 31.5 31.5 31.5 31.5 31.5	