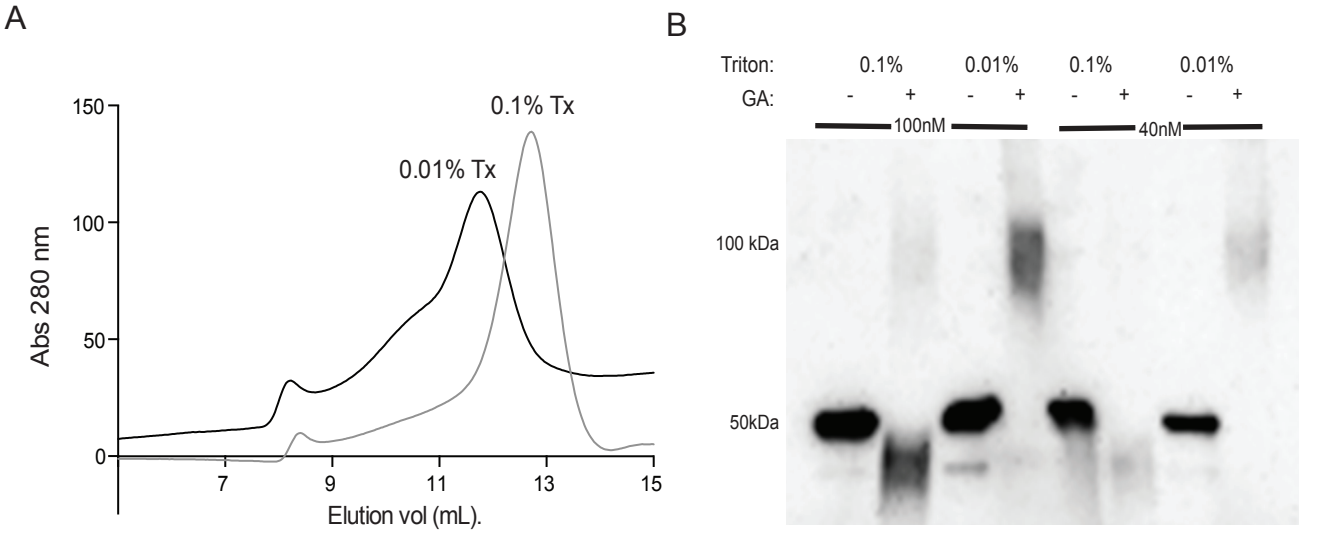
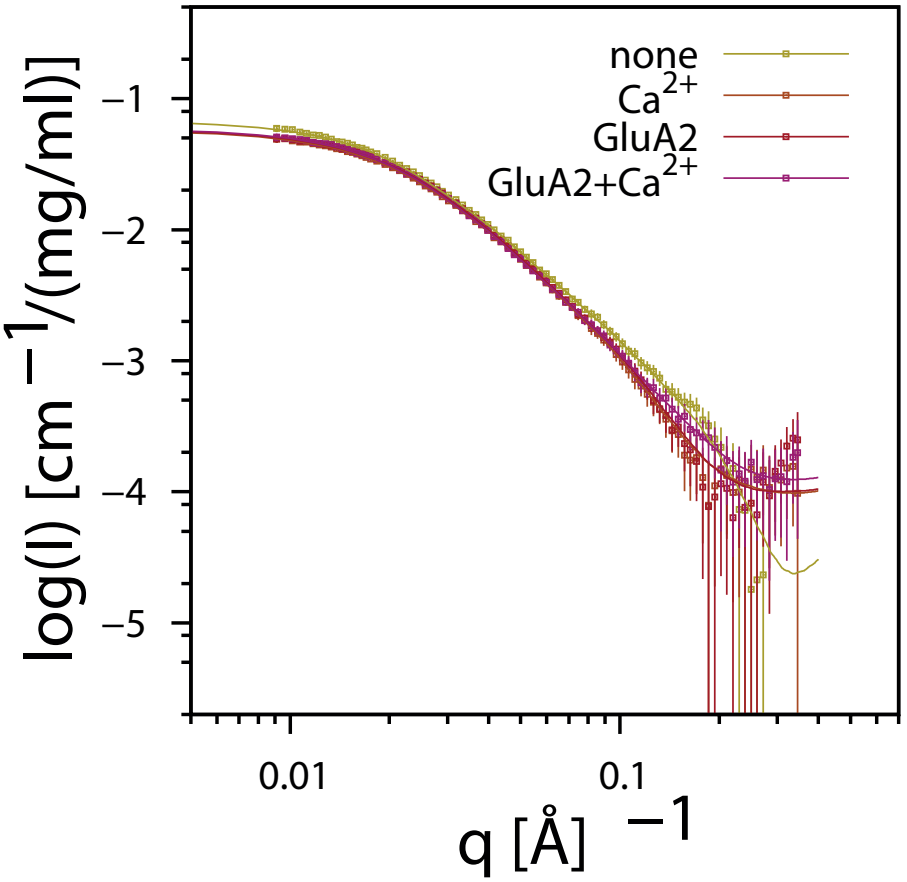
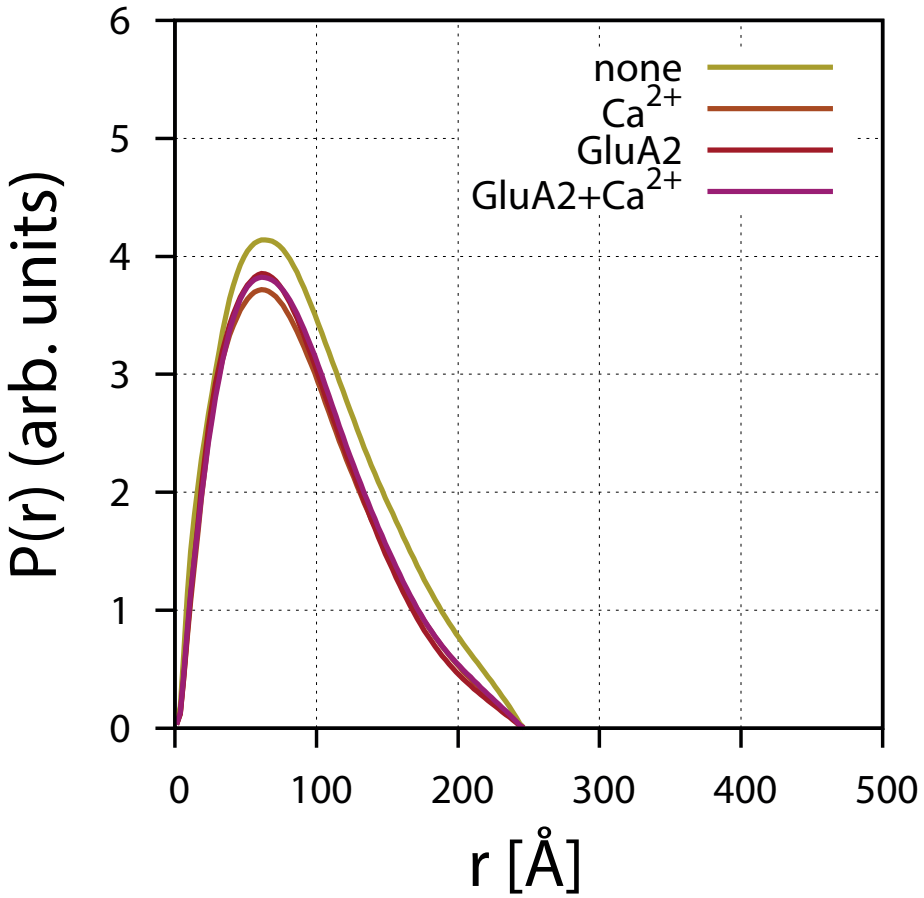
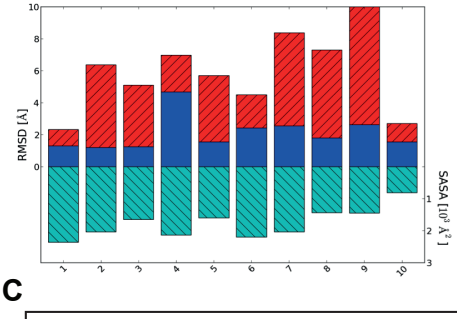

A
150
100
50
0
Abs 280 nm
0.1% Tx
0.01% Tx
7
9
11
13
15
Elution vol (mL).
B
Triton:
0.1%
0.01%
0.1%
0.01%
GA:
- + - + - + - +
100nM
40nM
100 kDa
50kDa


none
Ca2+
GluA2
GluA2+Ca2+
log(I) [cm−1/(mg/ml)]
−1
−2
−3
−4
−5
0.01
0.1
q [Å]−1


none
Ca2+
GluA2
GluA2+Ca2+
P(r) (arb. units)
6
5
4
3
2
1
0
0
100
200
300
400
500
r [Å]

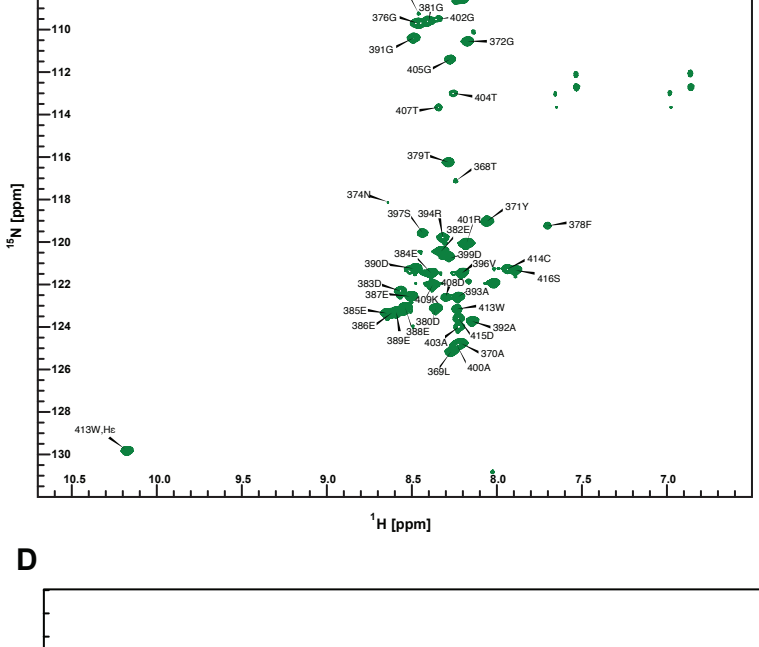
A

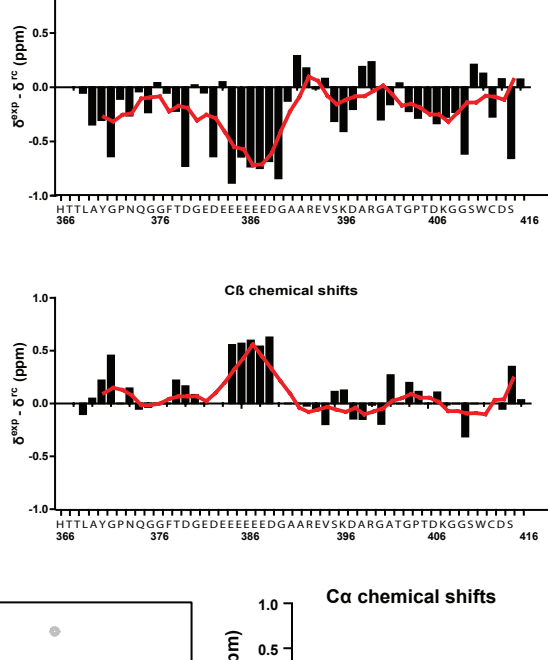


B

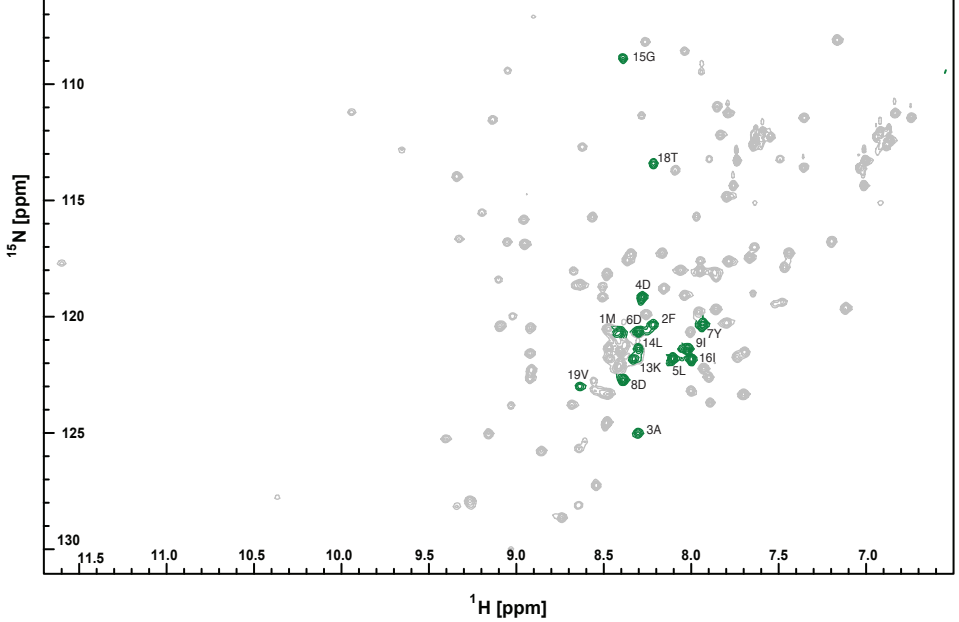
	RMSD [Å]	Hydrophobic SASA [Å²]	Buried hydrophobic surface [Å²]	Ratio
Amphyphysin	3.3	7075	2912	0.41
Endophilin	3.1	7628	2340	0.31
Arfaptin	2.6	6494	2215	0.34
CG-model	2.5	6325	2616	0.41

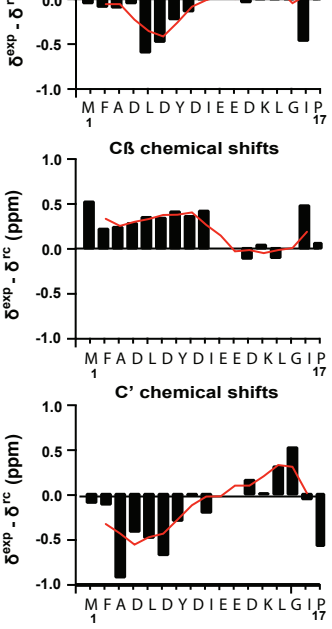
C

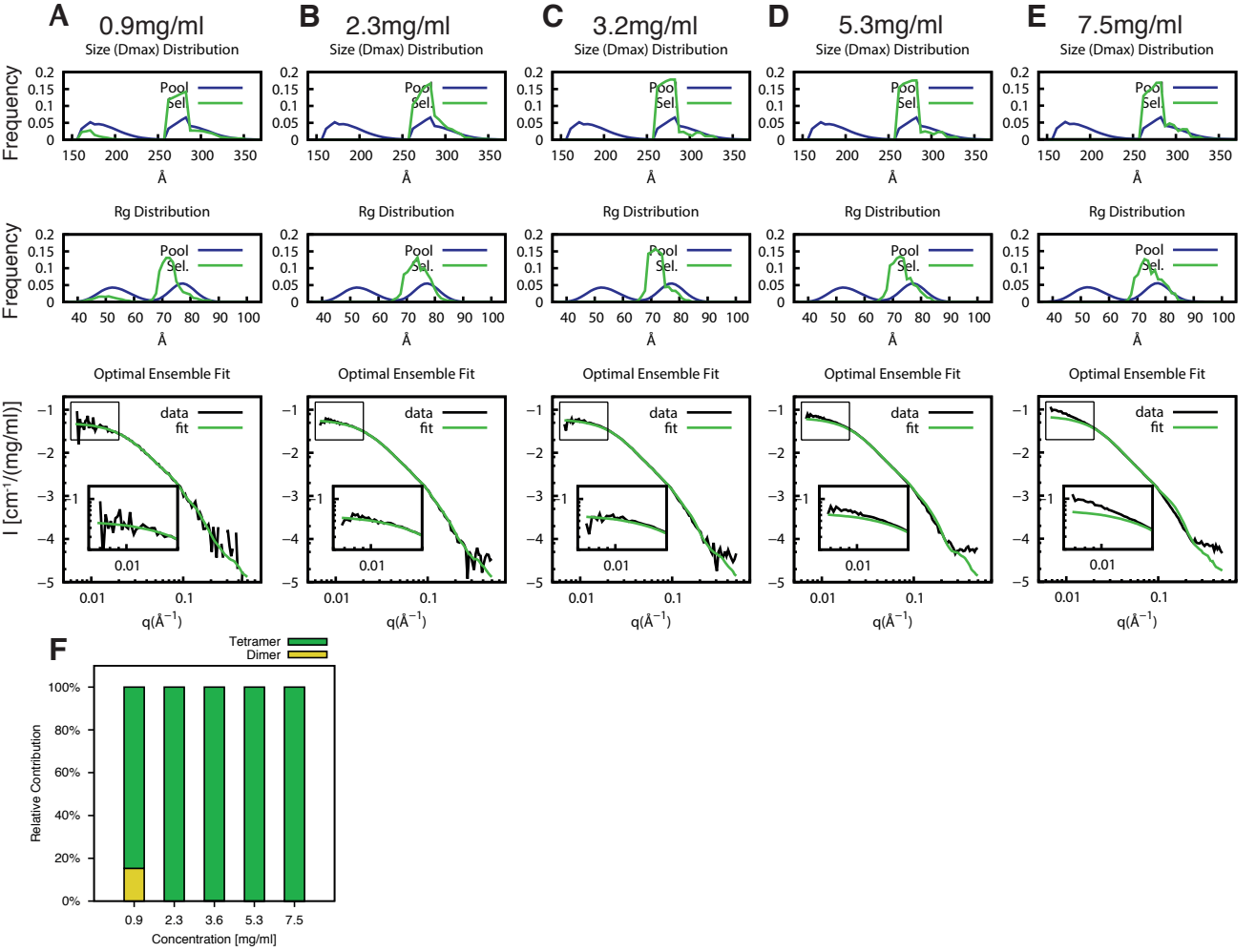


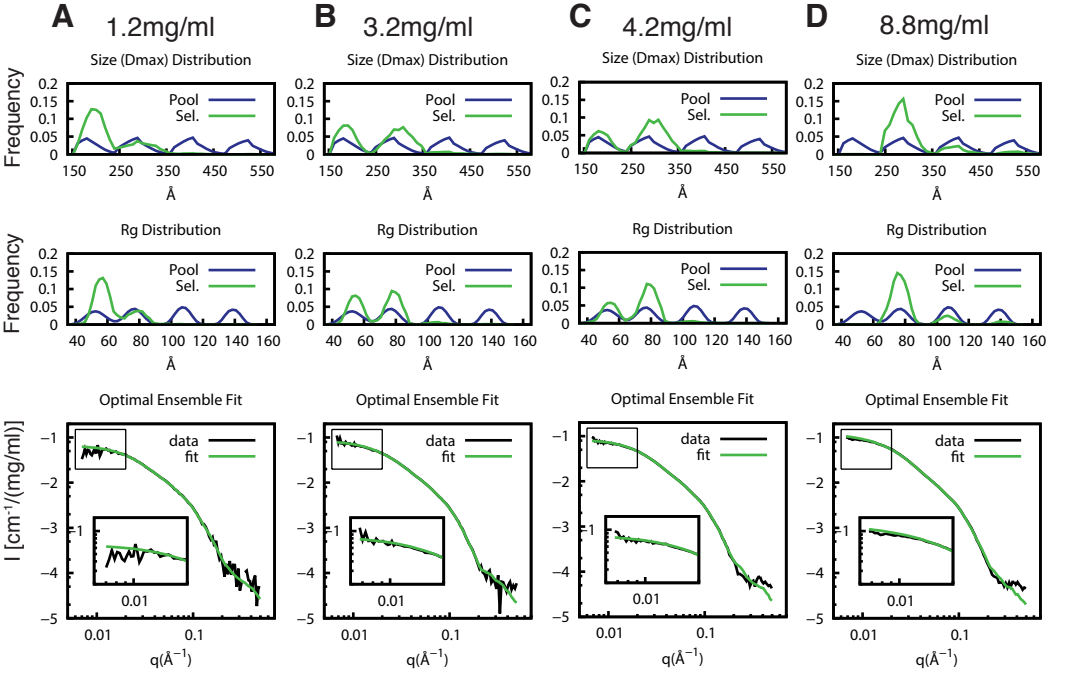


D






A
0.9mg/ml
B
2.3mg/ml
C
3.2mg/ml
D
5.3mg/ml
E
7.5mg/ml
Size (Dmax) Distribution
Frequency
Pool
Sel.
Å
Rg Distribution
Optimal Ensemble Fit
I [cm⁻¹/(mg/ml)]
data
fit
q(Å⁻¹)
F
Tetramer
Dimer
Relative Contribution
Concentration [mg/ml]


A
1.2mg/ml
B
3.2mg/ml
C
4.2mg/ml
D
8.8mg/ml
Size (Dmax) Distribution
Frequency
Pool
Sel.
Å
Rg Distribution
Optimal Ensemble Fit
I [cm-1/(mg/ml)]
data
fit
q(Å-1)

Dimer – PICK1^{LKV}					
ID	Combi. [mg/ml]	Dmax [Å]	Rg [Å]	I0 [cm^{-1}/(mg/ml)×10^{-2}]	χ^2
20vs36	3.2vs8.8	200	60.62 ± 0.054	1.000 ± 0.012	2.53
20vs38	3.2vs4.2	200	60.62 ± 0.054	1.000 ± 0.012	2.53
36vs38	8.8vs4.2	200	60.97 ± 0.039	1.033 ± 0.010	2.22
40vs20	1.2vs3.2	200	61.75 ± 0.074	1.023 ± 0.018	1.97
40vs36	1.2vs8.8	200	61.79 ± 0.076	1.024 ± 0.019	1.97
40vs38	1.2vs4.2	200	61.78 ± 0.075	1.024 ± 0.018	1.98
Mean	N/A	200	61.30 ± 0.011	1.012 ± 0.009	4.02

Tetramer – PICK1^{LKV}					
ID	Combi. [mg/ml]	Dmax [Å]	Rg [Å]	I0 [cm^{-1}/(mg/ml)×10^{-2}]	χ^2
20vs36	3.2vs8.8	250	74.37 ± 0.368	0.991 ± 0.064	2.94
20vs38	3.2vs4.2	250	72.46 ± 5.086	1.085 ± 0.514	2.10
36vs38	8.8vs4.2	250	74.64 ± 1.853	0.974 ± 0.253	0.20
40vs20	1.2vs3.2	250	76.38 ± 3.203	0.950 ± 0.402	2.10
40vs36	1.2vs8.8	270	78.13 ± 0.677	1.028 ± 0.106	2.65
40vs38	1.2vs4.2	250	76.29 ± 0.506	1.028 ± 0.145	1.75
Mean	N/A	250	76.31 ± 0.381	1.033 ± 0.137	1.27

Supplemental figure legends

Figure S1. Monomeric PICK1 is obtained in TX 100 concentration above CMC (0.1%)

(A) Western blot showing crosslinking with GA of purified PICK1 at 100 and 40nM in 0.01 and 0.1% TX-100 demonstrating that in 0.01% TX-100 PICK1 can be crosslinked efficiently into dimers and higher orders species even at nanomolar concentration and that this crosslinking is absent in 0.1%TX-100. (B) SEC of PICK1 in 0.01% (black) and 0.1% TX-100 (gray) demonstrating a delayed elution of PICK1 at 0.1% TX-100 indicating monomeric PICK1 as the dominant species also at high concentrations. Relating to Figure 1

Figure S2: Calcium or PDZ binding peptide cause no overall structural rearrangement of PICK1.

SAXS data (top) and pair distance distribution functions by IFT (bottom) for PICK1, PICK1 + Ca^{2+} (50 μM free), PICK1 + GluA2 C10 peptide (100 μM) or both. No significant structural changes were detected. Relating to Figure 2

Figure S3. Definition of modules in PICK1

(A) Interface quality. Several measures of the quality of the interface formed in the 10 dimers with highest Contact Score (see methods), ordered from left to right (left is the best dimer according to the Contact Score). All measures are averages over the last 800 ns of the 1.6 μs simulation of each dimer. Shown in blue is the RMSD of the backbone beads relative to the initial structure of the dimer. Hashed red columns represent how symmetrical the dimer is, measured by an RMSD (see Methods, low RMSD means highly symmetrical). The SASA of hydrophobic residues buried in the dimer interface is depicted in cyan (see Methods). The dimer with the highest Contact Score is also the best dimer according to these measures, showing high stability (low RMSD) and symmetry and a large hydrophobic buried surface. (B) Comparison of different all-atoms models of the PICK1-BAR dimer.

Three homology models based on different templates (amphiphysin, endophillin and arfaptin 2) and the dimer selected from the CG-MD simulations. All values are averages over the second half of 40 ns-long AA-MD simulations. The RMSD is on the Cα atoms relative to the last frame of the trajectory. The Second column shows the SASA of hydrophobic resides for the dimer whereas the third one is the surface of hydrophobic residues buried in the dimer interface. The last column shows the ratio of these two surfaces, i.e. the fraction of hydrophobic surface that gets buried upon formation of the dimer. (C)^{1}H-^{15}N-HSQC NMR experiment of the PICK1 C-terminal (residues 366-416) shows signals from all the backbone N-H correlations. The spectrum displays relatively low dispersion in the proton dimension as seen for disordered proteins. (D) and (E) Cα and Cβ secondary chemical shift analysis using random coil chemical shifts by (Kjaergaard et al., 2011) as reference. The PICK1 C-terminal does not display any transient secondary structure, however, the stretch of Glu residues demonstrates somewhat extended nature, likely due to charge-charge repulsion. F)^{1}H-^{15}N-HSQC NMR experiment of the PICK1 N-terminal with the PICK1 PDZ domain linked to the gluA2 C-terminal shows signals from all the backbone N-H correlations. The 19 N-terminal residue resonances are shown in green and the remaining resonances from the PDZ domain and ligand are shown in grey. G, H and I) Cα, Cβ and C' secondary chemical shift analysis using random coil chemical shift by (Kjaergaard et al., 2011) as reference. The PICK1 N-terminal does not display any transient secondary structure. Relating to Figure 3.

Figure S4. PICK1^{WT} EOM analysis with dimer and tetramer structure pools.

Ensemble Optimization (EOM) on PICK1^{WT} concentration series (A) 0.9 mg/ml (B) 2.3 mg/ml (C) 3.2 mg/ml (D) 5.3 mg/ml and (E) 7.5 mg/ml, including dimers, and tetramers in the generated pool. Top panels show D_{max} and R_g of generated pools in blue and selected pools in green. Bottom panels show SAXS data and fit of optimal ensemble assuming only dimers and tetramers. It is clear that the low-q data at high concentrations cannot be described by a combination of dimers and tetramers (F)

Quantification of the fraction of dimers and tetramers at different concentrations of PICK1^{WT}. Relating to Figure 5.

Figure S5. EOM analysis including higher order oligomers on PICK1^{LKV} data.

Ensemble Optimization Method (EOM) fitting demonstrates a concentration dependent shift in relative distribution of dimers, tetramers, hexamers and octamers in PICK1^{LKV} samples (A) 1.2 mg/ml (B) 3.2 mg/ml (C) 4.2 mg/ml and (D) 8.8 mg/ml. Top panels and middle panels show D_{max} and R_g of the generated pools in blue and selected pools in green. Bottom panels show SAXS data and fit of optimal ensemble. Inclusion of hexamers and octamers still allow for obtaining good fits to the low-q data, but are not included much in the selected ensembles. Quantification of the fraction of dimers, tetramers, hexamers and octamers at different concentrations of PICK1^{LKV} is shown in Figure 6.

Table S1. Model-Independent parameters from IFT analysis of decomposed dimer A) and tetramer B) form factors. Relating to Figure 7

Supplemental Experimental Procedures

Cloning and protein purification

WT PICK1 and PICK1-LKV were cloned and expressed in *E. coli* as previously described (Madsen2005). The entire coding region of rat PICK1 (residues 2-416) is amplified from a pCINEO vector by PCR using *pfu* polymerase according to the instructions by the manufacturer (Stratagene, La Jolla, CA). The primers used introduce a 5' restriction site for MunI and 3' restriction site for AvrII and for the LKV construct this sequence was encoded in the 3' primer. The PCR fragment is cleaved with MunI and AvrII and cloned into the reading frame of the pET41a vector (Novagen, Madison, WI). From the above construct the last 52 C-terminal aa of PICK1 was amplified using

Phusion polymerase (Finnzymes) and cloned into pGEX-4T-2 (GE healthcare). For all constructs this produced a glutathione-S-transferase (GST) N-terminally fused to the expressed protein. The eYFP-PICK1 was described previously (Madsen et al., 2008) and the LKV mutation was introduced using a 3'-primer encoding the mutations.

The GST fused proteins were expressed in *E. coli* BL21(DE3) pLysS (Novagen). The transformed bacteria are grown to OD_{600} 0.6 and expression of the fusion protein is induced with isopropyl-β-D1-thiogalactopyranoside (25 mg/ml) overnight at 30ºC. The bacteria were lysed by freezing and thawing in TBS buffer containing [50 mM Tris pH 7.4, 125 mM NaCl, 1mM DTT, 1 vol% TX-100]. The lysate was cleared by centrifugation (rotor SS-34, 18000 rpm, 48000 x g, 30 min). The supernatant was incubated with glutathione-sepharose beads (Amersham Biosciences) under slow rotation for 90 minutes at 4°C. The beads were pelleted at 3000 g for 5 minutes and washed in TBS buffer containing [50 mM Tris pH 7.4, 125 mM NaCl, 1 mM DTT, 0.01 vol% TX-100 or 0.04 % foscholin 12] by three batch washes. The protein was separated from the GST domain by cleavage with thrombin protease (Novagen) in the above wash buffer at 4ºC overnight. Purified proteins were stored at 4º C until use the same day.

Protein Expression and Purification for NMR: Bacteria transformed with the pGEX 4T-2 vector containing the PICK1 C-terminal GST fusion protein insert or the GST-N-terminal-PDZ-GluA2 (Erlendsson et al., 2014), were grown in M9 minimal media for ^{15}N and ^{13}C isotopic labeling using $^{15}NH_4SO_4$ and ^{13}C-glucose as sole sources of nutrition. The GST fusion protein was purified as described for PICK1 full length above. This yielded final concentrations of 750 μM and 430 μM, respectively. All solutions were prepared in TBS buffer (50 mM Tris, 125 mM NaCl, 1mM DTT, 0.01 vol% TX-100, pH 7.4).

Size exclusion chromatography

Size exclusion chromatography (SEC) was performed on an Äkta purifier (GE Healthcare) with a BioSep-SEC-S2000 size exclusion column (Phenomenex) or a superdex 200 10/300 GL size exclusion column (GE Healthcare). The mobile phase was 50 mM Tris buffer (pH 7.4) containing 125 mM NaCl, 0.01 vol%- 0.1 vol% reduced TX-100. Flow rate 0.3-0.5 mL/min.

Analytical ultra-centrifugation

Sedimentation velocity experiments were performed on a Beckman XL-1 analytical ultracentrifuge at 4°C and 50,000 rpm. The sedimentation was measured as absorbance at a wavelength of 280 nm. The partial specific volumes were calculated according to the additivity scheme (Makhatadze et al., 1990) and data was analyzed using the c(S) model implemented by SEDFIT (Brown and Schuck, 2006).

Calcium sensitivity

Before SAXS experiments HEDTA (Sigma) was added to a final concentration of 0.5 mM and the proteins were centrifuged at 100,000 G for 30 min. Additional components i.e. $CaCl_2$ and peptides were added 20 min. before measurements. Peptides were added to a final concentration of 100 µM and $CaCl_2$ was added giving a free concentration of 50 µM as calculated using the Webmaxc standard program http://www.stanford.edu/~cpatton/webmaxc/webmaxcS.html.

Confocal microscopy

On the day of imaging the media was exchanged for medium without phenol red (Invitrogen) and allowed to equilibrate for 30 min. Cells were visualized using a Zeiss LSM 510 confocal laser-scanning microscope using an oil immersion 63x objective. YFP was excited with the 488 nm laser line from an argon–krypton laser, and the emitted light was detected using a 505–550 nm band pass filter. Image treatment was done in ImageJ.

Förster Resonance Energy Transfer (FRET) measurements

For C-terminal fusion of fluorophore to PICK1, the protein was expressed in the pEYFP-N1 or pECFP-N1 vectors (Clontech). N-terminal fusions are as described for YFP-PICK1 above. eYFP and CFP fused PICK1 was expressed in COS7 cells using transient transfection (lipo2000, Invitrogen), expression of eYFP together with eCFP was used as a negative control. 24-48 hours after transfection the FRET signal was measured in a spectrofluorometer (Fluoromax-4, Jobin Yvon Technology) from 900.000 cells suspended in PBS. The FRET value was calculated using the following equation N_{FRET} = (FRET-YFP*BT-CFP*BT)/(YFP*CFP)^{½}, BT= Bleedthrough. In addition, the cytosolic FRET signal was measured in single cells with an epifluorescence microscope (Carl Zeiss TM210, Germany) using the "three-filter method" described in (Xia and Liu, 2001). This assay confirmed the observations from the spectrofluorometer measurements. The N_{FRET} for PICK1-eYFP + PICK1-eCFP (0.067 ± 0.007) was significantly higher than for eYFP + eCFP (0.034 ± 0.006) $P<0.01$ in one-way ANOVA with Bonferroni's post test, whereas there was no significant difference between eYFP-PICK1 + eCFP-PICK1 (0.034 ± 0.009) and eYFP + eCFP, n=4-5 independent experiments.

SAXS data acquisition

Acquisition: The scattering experiments were carried out at the EMBL X33 beamline (C. E. Blanchet, 2012) at the DORIS storage ring (DESY, Hamburg) following standard procedures. Samples were loaded using the available automatic sample changer and keeping both sample and measurement temperature at 4°C. Each sample was measured in 4×30s while the sample was oscillated in the beam to minimize potential radiation damage. Background buffers were measured prior and subsequent to each sample such that the averaged background scattering could be subtracted from the scattering of the sample. A concentration series was measured for each protein. As a part of the experiments, it was verified that no signs of radiation damage to the protein were present. The SAXS data were collected on a two-dimensional MAR345 image plate detector and with the applied instrument

settings a momentum transfer range $0.008 < q < 0.496$ Å-1 ($q = 4\pi \sin\theta/\lambda$, where θ is the half scattering angle and λ is the X-ray wavelength) was obtained. Additional measurements were carried out at MAX-lab (Lund, Sweden) on the beamline I711 (Knaapila et al., 2009) using the MAR165 CCD detector. Here data were obtained in the q-range 0.01–0.33 Å^{-1}. During both experiments, data points were azimuthally averaged and re-binned to logarithmically equidistant points. Absolute scale calibration of the scattering intensity into units of scattering cross section per unit volume (1/cm) was carried out using water as a reference (D. Orthaber, 2000). This calibration was double-checked against freshly prepared samples of bovine serum albumin (Svergun, 2007) with known protein concentrations (~4 mg/ml). Finally, data were normalized by the protein concentration to units of [cm^{-1}/(mg/ml)]. Sample concentration was measured on a nanodrop (Thermo Scientific) in duplicate.

Initial SAXS data analysis

Due to the intrinsic experimental uncertainty in concentration measurements, scattering data were re-scaled slightly internally to ensure overlap in the intermediate to high-q region insensitive to concentration effects ($0.06 < q < 0.1$). For direct comparison of samples, log/log and log/lin plots of scattering profiles were investigated by visual inspection. Guinier plots were used to estimate reliable data range at low q-values, extrapolate the forward scattering, *I(0)*, and determine radius of gyration, R_g. The recorded scattering profiles were transformed into real space representations in terms of the pair distance distribution function *p(r)*, by means of indirect Fourier transformation using a modified version of Glatter's original procedure (Glatter., 1977; Pedersen et al., 1994), which also determines the forward scattering, *I(0)*, radius of gyration, R_g, along with the maximal internal distance, D_{max}.

The average molecular mass pr. scattering unit, M, was estimated from the forward scattering:

$$I(0) = c\Delta\rho_m^2 MP(0) \Rightarrow M = \frac{I(0)}{c\Delta\rho_m^2},$$

where c is the concentration, $\Delta\rho_m$ is the excess scattering length pr. unit mass of proteins relative to the solvent. A value of 2.0×10^{10} cm/g was applied for the calculations. *P(q)* is the form factor, which

is unity at $q = 0$. The oligomerization factor is the average mass divided by the monomeric mass: $M/M_{Monomer}$.

Definition of individual structural modules of PICK1 for EOM

We and others have previously solved the structure of the PICK1 PDZ domain, stabilized by a C-terminal fusion sequence that docks into the peptide binding groove, by either X-ray crystallography or NMR (Elkins et al., 2007; Erlendsson et al., 2014; Pan et al., 2007). We used the most well-defined structure, which includes the last ten residues of the dopamine transporter C-terminus stabilizing the PDZ binding groove (PDB 2LUI) (see Figure 3A)(Erlendsson et al., 2014).

There is no structural information on the PICK1 BAR domain and the purified isolated BAR domain is highly unstable precluding structural determination. Consequently, we turned to computational approaches to generate a molecular model of the PICK1 BAR dimer. An all-atom model of the PICK1-BAR domain was built as described below (*Molecular Dynamics Simulation* section), using the homologous BAR domain in Arfaptin-2 as a template (Tarricone et al., 2001). This model was then transformed into a Martini-based coarse-grained representation and used in 5 independent dimerization simulations, each containing 10 PICK1-BAR monomers and run for 600 ns. The 10 most probable dimer interfaces were selected from these simulations using a score based on the frequency of contacts observed for each residue during the simulations, as described in *Molecular Dynamics Simulation* section. The selected dimers were each simulated separately with coarse-grained MD simulation for an additional 1.6 μs to assess the feasibility of the interfaces. Figure S3A shows measures of the stability and quality of these dimers root mean square deviation (RMSD), symmetry, solvent-accessible surface areas (SASA)). It reveals that the dimer with the highest *Contact Score* (number 1) also appears to be best according to these various measures. Indeed the BAR dimer scoring the highest according to the *Contact Score* is very stable, with a backbone

RMSD of 1.3 Å, and is characterized by the largest hydrophobic surfaces buried in the interface (2390 $Å^2$) and the highest symmetry among all 10 dimers.

This dimer was therefore designated as the "best" model of the PICK1-BAR dimer and was further studied in separate simulations for comparison with other models (homology models of the dimer based on arfaptin 2, amphiphysin (Casal et al., 2006) and endophilin (Weissenhorn, 2005)) using all-atom MD simulation. The equilibrated structure of this dimer (after 80 ns of AA-MD) is shown in Figure 3B. These simulations confirmed that the dimer selected on the basis of the *Contact Score* indeed compared favorably to the BAR dimers built by simple homology modeling (using amphiphysin, endophilin and arfaptin 2 as templates), most notably in terms of stability and the SASA of hydrophobic residues (see Fig S3B). Taken together, the computational data indicate that the dimer structure that emerged from the CG-MD simulations is a good model of the PICK1-BAR dimer and can be used in the interpretation of the SAXS data. For more information about the procedures see the *Molecular Dynamics Simulations* section below.

NMR

To gain insight into the structural arrangement of the terminal parts of the protein, we determined the backbone chemical shifts for the N-terminal (residue 1-17) linked to the PDZ domain (Figure S3C) and the isolated C-terminal (residue 366-416) (Figure S3D) of PICK1 using NMR spectroscopy. The 1H-15N-HSQC spectra showed little if any dispersion, indicating that neither of the regions comprises any higher secondary or tertiary structure at neutral pH. Examples are illustrated in Figure 3C.

All samples were prepared containing 0.01 mM NaN_3, 0.25 mM DSS and 10% D_2O. Backbone resonances of PICK1 C-terminal and the N-terminal-PDZ-GluA2 were assigned by heteronuclear correlation experiments including $^{1}H,^{15}N$-HSQC, CBCANH, CBCA(CO)NH, HNCA, HN(CO)CA, HNCO and HN(CA)CO. A Varian Unity 800MHz spectrometer with a cryoprobe at 15ºC was used for all experiments. All data were processed using NMRPipe (Delaglio et al., 1995) and assigned in CCPNMR Analysis (Vranken et al., 2005).

Molecular Dynamics simulations

A molecular model of the PICK1-BAR domain was built with *homology modeling*. Briefly, an alignment of homologous proteins for which a crystal structure is available was generated with HHPred (Soding et al., 2005). The BAR domain from the human Arfaptin 2 emerged as the closest homologue to the PICK1-BAR domain (23% sequence identity, compared to 16% for amphiphysin, the second best template) and its structure (PDB ID: 1I49, (Tarricone et al., 2001)) was therefore used as template for the homology modeling. The alignment was manually refined before generating 20 models using modeler (Fiser and Sali, 2003). When aligned on Arfaptin 2, the sequence of the PICK1 BAR domain contains a somewhat longer loop (compared to Arfaptin 2) between Helix 2 and Helix 3 (residues 272 to 297), which was modeled as helical in order to elongate the coiled-coil

structure of the BAR domain. The predicted structures were all very similar (average pairwise RMSD ~3Å) and the final model was chosen based on the energy criteria from modeller.

A PICK1-BAR dimer was also modeled according to the Arfaptin 2 crystal structure. Moreover, we used two other homologues to the PICK1-BAR, the Amphiphysin and Endophilin BAR domains, to build alternative models for the PICK1-BAR monomer and dimer, using the same approach as described above. The resulting three dimeric structures (Arfaptin, amphiphysin and endophilin based) were used for comparison with a model obtained from coarse-grained molecular dynamics simulations (see below and results).

All-atom molecular dynamics (AA-MD) simulations: AA-MD simulations were performed using NAMD (Phillips et al., 2005). All dimer constructs were hydrated in a rectangular water box (typical size ~90 Å x 75 Å x 115 Å containing ~155,000 atoms) with 150 mM NaCl and simulated, using the CHARMM27 force field with CMAP corrections (Mackerell et al., 2004) and periodic boundary conditions, in the NPT ensemble with isotropic pressure coupling and a timestep of 2 fs. All systems were initially equilibrated for 1.5 ns during which the backbone of the protein was restrained by a harmonic potential with decreasing force constant (from 1 to 0.1 kcal/mol/Å).

Coarse-grained molecular dynamics (CG-MD) simulations: CG-MD simulations used the Martini force-field with ELNEDYN (Periole, 2009) and were performed using the GROMACS software) (Lindahl et al., 2001). Here too periodic boundary conditions and isotropic pressure coupling were used and the simulations were run with a time step of 10 fs. A 10 ns initial equilibration using position restraints for the protein with a force constant of 1 kJ/mol/Å was performed for all systems. Simulation times reported include the standard correction factor of 4 for the MARTINI force field.

Dimerization of PICK1-BAR monomers: Dimerization of the PICK1-BAR monomers was simulated using CG-MD simulations. Systems consisting of 10 PICK1-BAR monomers solvated in a cubic water box with a length of 240 Å (~120,000 CG beads) were built. The monomers were placed in pairs separated by ~15 Å and in different relative positions (combination of rotations by ~20 degrees

and displacements of ~15 Å from the Arfaptin-2 dimer interface found in the crystal structure) to facilitate the formation of interfaces. Five independent simulations, of 600 ns each, were performed from two different initial configurations. Of the dimers formed in these multiple trajectories, 10 were chosen for further study (see below). Each of them was solvated in a cubic water box of size 180 Å corresponding to ~50k CG beads and was simulated for 1.6 µs.

Scoring of dimer interfaces: Scoring of the dimer interfaces from the CG-MD simulations (see above) was based on pairwise contacts in the protein-protein interface. Specifically, we first extracted all the interfaces formed during the multiple simulations. Then for each residue r_i in the monomer, we computed the fraction p_i of interfaces containing that residue in the contacting surface. Finally we used those p_i to score each individual interface from the simulations using:

$$S = \sum_{\{i,j\}} \frac{p_i p_j}{d_{ij}} \theta(10 - d_{ij}),$$

where the sum runs over all pairs of residues (i in the first protomer and j in the second one), d_{ij} is the minimal distance between residues i and j (in Ångström) and θ is the Heaviside step function defined as:

$$\theta(x) = \begin{cases} 0 \text{ if } x < 0 \\ 1 \text{ if } x > 0 \end{cases}$$

This score has been validated on multiple test simulations of different model systems for which the dimer interface is known, including on the dimerization of Arfaptin BAR domains. For all test cases, and notably for Arfaptin, this score was shown to yield a strong correlation with objective measures of the correctness of an interface (as measured for example by the RMSD towards the known dimer structure) and performed much better than standard measures such as RMSD and buried hydrophobic surface (manuscript under preparation). The symmetry of the dimers obtained from CG-MD was measured with a Root Mean Square Deviation (RMSD) of the backbone (BB) beads of the dimer aligned on itself but with the two monomers swapped. Specifically we made a copy of the dimer

(denoting the two monomers in the original dimer as M_1 and M_2 and in the copy as M'_1 and M'_2) and found the alignment minimizing the RMSD defined by:

$$\left(\frac{1}{N}\sum_{i=1}^{N}(\vec{r}_i^{M_1}-\vec{r}_i^{M_2'})^2+(\vec{r}_i^{M_2}-\vec{r}_i^{M_1'})^2\right)^{1/2},$$

with N the number of residues in each monomer. This RMSD will be 0 for a perfectly symmetrical dimer and increase with the asymmetry in the dimer. All the analyses of the MD simulations (including the implementation of the *Contact Score*), and snapshots thereof, were made using Openstructure (Biasini et al., 2013) and the plotting library Matplotlib (Hunter, 2007).

Solvent-Accessible Surface Areas (SASA): SASAs were obtained with the MSMS software (Sanner et al., 1996) and using a probe radius of 1.4Å for AA-MD and 2.35 Å for CG-MD. The hydrophobic surface buried in the dimer interfaces (see Results) was obtained by summing the SASA of hydrophobic residues in each monomer, and subtracting from the result the SASA of hydrophobic residues in the dimer.

SAXS Analysis: Ensemble Optimization Method (EOM)

With EOM, large pools of structures are generated, composed by rigid components known from other structural studies and user-defined flexible parts assigned with random structures. The theoretical scattering curve from each structure is calculated, and ensembles are optimized towards sample data, enabling an analysis of the degree of flexibility of the protein. Due to the uniqueness problem, there is no guarantee that the resulting optimal ensembles represent the actual structures in the sample, but are rather considered representative of possible structures or often visited conformations. EOM has previously been used to describe the structural component of flexible proteins and has been used extensively for characterization of intrinsically disordered proteins (Mylonas et al., 2008). The method is also applicable for a description of multidomain proteins, exhibiting significant interdomain flexibility (Bernado et al., 2010) and can be applied either based solely on SAXS data or

in combination with NMR data as in the example above. Samples of even more complex composition can now be analysed, since the new version of the program allows the incorporation of oligomers in the pool of models (Bernado et al., 2007; Petoukhov and Svergun, 2012).

Populations of 10.000 structures each were generated by Ranch (Bernado et al., 2007; Petoukhov and Svergun, 2012). Scattering curves were calculated for each structure [from 0.0001 to 0.5Å^{-1}] using 15 spherical harmonics. Optimal ensembles were fitted to data with the program Gajoe (Bernado et al., 2007; Petoukhov and Svergun, 2012), with 2000 generations and 100 repeats.

SAXS Analysis: Two-component Decomposition

Assuming that each sample consists of only dimers and tetramers, a special case of singular value decomposition may be applied to recover the monomer and dimer form factors from the experimental data.

For a sample consisting of more than one scattering component, the total scattering intensity at scattering vector q can generally be written as (Pedersen et al., 1994)

$$I(q) = \sum_i \Delta\rho_m^2 c_i M_i P_i(q)$$

where $\Delta\rho_m$ is the excess scattering length pr. unit mass of proteins relative to the solution, c_i is the concentration, M_i is the molecular mass, and $P_i(q)$ is the form factor of component i. In the present case with two components and the scattering intensity normalized to 1 mg/ml, this reduces to

$$\begin{aligned} \frac{I(q)}{c_M M_M} &= \Delta\rho_m^2 \frac{c_D M_D}{c_M M_M} P_D(q) + \Delta\rho_m^2 \frac{c_T M_T}{c_M M_M} P_T(q) \\ &= \Delta\rho_m^2 2\frac{c_D}{c_M} P_D(q) + \Delta\rho_m^2 4\frac{c_T}{c_M} P_T(q) \end{aligned}$$

where index M, D and T denote monomer, dimer and tetramer, respectively. We have normalized by the total amount of monomer, which can be measured by absorption spectroscopy. In the last equality, we use that M_D=2M_M and M_T=4M_M.

The factors c_D/c_M and c_T/c_M can be calculated for each sample through a few considerations of particle conservation (c_M=2c_D+4c_T) and the definition of average molecular mass:

$$M_{av} = \frac{c_M M_M}{c_D + c_T}$$

leading to

$$\frac{c_D}{c_M} = \frac{2M_M}{M_{av}} - \frac{1}{2} = \frac{2\Delta\rho^2}{\frac{I(0)}{c_M M_M}} - \frac{1}{2} \quad \text{and}$$

$$\frac{c_T}{c_M} = \frac{1}{2} - \frac{M_M}{M_{av}} = \frac{1}{2} - \frac{\Delta\rho^2}{\frac{I(0)}{c_M M_M}}$$

where we use that we can calculate the oligomerization factor M_{av}/M_M from $I(0)/c_M M_M$ as described previously.

Since the form factors of the dimers and tetramers should not change with concentration or between samples, the normalized scattering from sample i, $I_i'(q) = I_i(q)/(c_{M,i}\, M_M)$, from a series of samples with varying concentration can be described by

$$I_i'(q) = a_i P_D(q) + b_i P_T(q)$$

where

$$a_i = \Delta\rho_m^2 \left(\frac{4\Delta\rho_m^2}{I_i'(0)} - 1 \right) \quad \text{and} \quad b_i = \Delta\rho_m^2 \left(2 - \frac{4\Delta\rho_m^2}{I_i'(0)} \right)$$

By combining two samples, denoted sample i and j, with different concentration of monomer, and consequently different $I_i'(q)$, we can treat this as two equations with two unknowns. Expressing P_D in terms of i and P_T in terms of j, by substitution we get

$$P_D(q) = \frac{\frac{1}{a_i} I_i'(q) - \frac{b_i}{a_i b_j} I_j'(q)}{1 - \frac{a_j b_i}{a_i b_j}} \quad \text{and} \quad P_T(q) = \frac{\frac{1}{b_j} I_j'(q) - \frac{a_j}{a_i b_j} I_i'(q)}{1 - \frac{a_j b_i}{a_i b_j}}$$

This allows us to calculate the form factors for each 'q', yielding an estimated 'experimental' form factor for each pair of samples. The analytical expression enables the calculation of error bars for the final calculated form factors using standard rules for error propagation.

Substituting and rewriting we get, for a given sample with index *i*,

$$I_i(q) = \left(2M_{Dimer}\Delta\rho^2 - I_{0,i}\right)P_{Dimer}(q) + \left(-2M_{Dimer}\Delta\rho^2 + 2I_{0,i}\right)P_{Tetramer}(q),$$

where P_{Dimer} and $P_{Tetramer}$ are assumed independent of index *i*, i.e. of sample.

Defining $a_i = 2M_i\Delta\rho^2 - I_{0,i}$, and $b_i = -2M_i\Delta\rho^2 + 2I_{0,i}$ we can then rewrite, for two samples with index *i* and *j*,

$$I_i(q) = a_i \cdot P_{Dimer}(q) + b_i \cdot P_{Tetramer}(q),$$
$$I_j(q) = a_j \cdot P_{Dimer}(q) + b_j \cdot P_{Tetramer}(q)$$

Expressing P_{Dimer} in terms of *i*, and $P_{Tetramer}$ in terms of *j*, we can treat this as two equations with two unknowns, and solve for P_{Dimer} and $P_{Tetramer}$:

$$P_{Dimer}(q) = \frac{\frac{I_i(q)}{a_i} - \frac{b_i I_j(q)}{a_i b_j}}{1 - \frac{b_i a_j}{a_i b_j}}$$

$$P_{Tetramer}(q) = \frac{I_j(q) - a_j \cdot P_{Dimer}(q)}{b_j}$$

These equations are calculated at each '*q*', yielding an estimated decomposed 'experimental' Dimer and Tetramer form factor, from each pair of sample-measurements. Error bars are propagated using the standard rules for error propagation.

Supplemental References

Bernado, P., Modig, K., Grela, P., Svergun, D.I., Tchorzewski, M., Pons, M., and Akke, M. (2010). Structure and Dynamics of Ribosomal Protein L12: An Ensemble Model Based on SAXS and NMR Relaxation. Biophys J *98*, 2374-2382.

Bernado, P., Mylonas, E., Petoukhov, M.V., Blackledge, M., and Svergun, D.I. (2007). Structural characterization of flexible proteins using small-angle X-ray scattering. J Am Chem Soc *129*, 5656-5664.

Biasini, M., Schmidt, T., Bienert, S., Mariani, V., Studer, G., Haas, J., Johner, N., Schenk, A.D., Philippsen, A., and Schwede, T. (2013). OpenStructure: an integrated software framework for computational structural biology. Acta Crystallogr D Biol Crystallogr *69*, 701-709.

Brown, P.H., and Schuck, P. (2006). Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. Biophysical journal *90*, 4651-4661.

C. E. Blanchet, A.V.Z., A. G. Kikhney, D. Franke, P. V. Konarev, W. Shang, R. Klaering, B. Robrahn, C. Hermes, F. Cipriani, D. I. Svergun, and M. Roessle (2012). Instrumental setup for high-throughput small- and wide-angle solution scattering at the X33 beamline of EMBL Hamburg. Journal of Applied Crystallography *45*.

Casal, E., Federici, L., Zhang, W., Fernandez-Recio, J., Priego, E.M., Miguel, R.N., DuHadaway, J.B., Prendergast, G.C., Luisi, B.F., and Laue, E.D. (2006). The crystal structure of the BAR domain from human Bin1/amphiphysin II and its implications for molecular recognition. Biochemistry *45*, 12917-12928.

D. Orthaber, A.B., and O. Glatter (2000). SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. Journal of Applied Crystallography *33*.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR *6*, 277-293.

Elkins, J.M., Papagrigoriou, E., Berridge, G., Yang, X., Phillips, C., Gileadi, C., Savitsky, P., and Doyle, D.A. (2007). Structure of PICK1 and other PDZ domains obtained with the help of self-binding C-terminal extensions. Protein Sci *16*, 683-694.

Erlendsson, S., Rathje, M., Heidarsson, P.O., Poulsen, F.M., Madsen, K.L., Teilum, K., and Gether, U. (2014). PICK1 (protein interacting with C-kinase 1) binding promiscuity relies on unconventional PDZ (PSD-95/Discs-large/ZO-1 homology) binding modes for non-class II PDZ ligands. The Journal of biological chemistry.

Fiser, A., and Sali, A. (2003). Modeller: generation and refinement of homology-based protein structure models. Methods Enzymol *374*, 461-491.

Glatter., O. (1977). new method for the evaluation of small-angle scattering data. Journal of Applied Crystallography *10*.

Hunter, J.D. (2007). Matplotlib: A 2D Graphics Environment. Computing in Science and Engineering,.

Kjaergaard, M., Brander, S., and Poulsen, F.M. (2011). Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. J Biomol NMR *49*, 139-149.

Knaapila, M., Svensson, C., Barauskas, J., Zackrisson, M., Nielsen, S.S., Toft, K.N., Vestergaard, B., Arleth, L., Olsson, U., Pedersen, J.S., *et al.* (2009). A new small-angle X-ray scattering set-up on the crystallography beamline I711 at MAX-lab. J Synchrotron Radiat *16*, 498-504.

Lindahl, E., Berk, H., and Spoel, D.v.d. (2001). GROMACS 3.0: a Package for Molecular Simulation and Trajectory Analysis. Molecular Modeling Annual *7*.

Mackerell, A.D., Jr., Feig, M., and Brooks, C.L., 3rd (2004). Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. Journal of computational chemistry *25*, 1400-1415.

Madsen, K.L., Eriksen, J., Milan-Lobo, L., Han, D.S., Niv, M.Y., mmendrup-Johnsen, I., Henriksen, U., Bhatia, V.K., Stamou, D., Sitte, H.H., *et al.* (2008). Membrane localization is critical for activation of the PICK1 BAR domain. Traffic *9*, 1327-1343.

Makhatadze, G.I., Medvedkin, V.N., and Privalov, P.L. (1990). Partial molar volumes of polypeptides and their constituent groups in aqueous solution over a broad temperature range. Biopolymers *30*, 1001-1010.

Mylonas, E., Hascher, A., Bernado, P., Blackledge, M., Mandelkow, E., and Svergun, D.I. (2008). Domain conformation of tau protein studied by solution small-angle X-ray scattering. Biochemistry *47*, 10345-10353.

Pan, L., Wu, H., Shen, C., Shi, Y., Jin, W., Xia, J., and Zhang, M. (2007). Clustering and synaptic targeting of PICK1 requires direct interaction between the PDZ domain and lipid membranes. The EMBO journal *26*, 4576-4587.

Pedersen, J.S., Hansen, S., and Bauer, R. (1994). The aggregation behavior of zinc-free insulin studied by small-angle neutron scattering. Eur Biophys J *22*, 379-389.

Periole, X., Marco Cavalli, Siewert-Jan Marrink, and Marco A. Ceruso (2009). Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. J Chem Theory Comput *5*.

Petoukhov, M.V., and Svergun, D.I. (2012). Applications of small-angle X-ray scattering to biomacromolecular solutions. Int J Biochem Cell Biol *45*, 429-437.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. Journal of computational chemistry *26*, 1781-1802.

Sanner, M.F., Olson, A.J., and Spehner, J.C. (1996). Reduced surface: an efficient way to compute molecular surfaces. Biopolymers *38*, 305-320.

Soding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res *33*, W244-248.

Svergun, E.M.a.D.I. (2007). Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. Journal of Applied Crystallography *40*.

Tarricone, C., Xiao, B., Justin, N., Walker, P.A., Rittinger, K., Gamblin, S.J., and Smerdon, S.J. (2001). The structural basis of Arfaptin-mediated cross-talk between Rac and Arf signalling pathways. Nature *411*, 215-219.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins *59*, 687-696.

Weissenhorn, W. (2005). Crystal structure of the endophilin-A1 BAR domain. JMolBiol *351*, 653-661.

Xia, Z., and Liu, Y. (2001). Reliable and global measurement of fluorescence resonance energy transfer using fluorescence microscopes. BiophysJ *81*, 2395-2402.