




ORIGINAL
ARTICLE

Structure of monomeric full-length ARC sheds light on molecular flexibility, protein interactions, and functional modalities

Erik I. Hallin*, Maria S. Eriksen*[†], Sergei Baryshnikov*[†],
Oleksii Nikolaienko*[†] , Sverre Grødem*[†], Tomohisa Hosokawa[‡],
Yasunori Hayashi[‡], Clive R. Bramham*^{†,1}  and Petri Kursula*^{§,1} 

*Department of Biomedicine, University of Bergen, Bergen, Norway

[†]K.G. Jebsen Centre for Neuropsychiatric Disorders, University of Bergen, Bergen, Norway[‡]Department of Pharmacology, Kyoto University Graduate School of Medicine, Kyoto, Japan[§]Faculty of Biochemistry and Molecular Medicine, University of Oulu, Oulu, Finland

Abstract

The activity-regulated cytoskeleton-associated protein (ARC) is critical for long-term synaptic plasticity and memory formation. Acting as a protein interaction hub, ARC regulates diverse signalling events in postsynaptic neurons. A protein interaction site is present in the ARC C-terminal domain (CTD), a bilobar structure homologous to the retroviral Gag capsid domain. We hypothesized that detailed knowledge of the three-dimensional molecular structure of monomeric full-length ARC is crucial to understand its function; therefore, we set out to determine the structure of ARC to understand its various functional modalities. We purified recombinant ARC and analyzed its structure using small-angle X-ray scattering and synchrotron radiation circular dichroism spectroscopy. Monomeric full-length ARC has a compact, closed structure, in which the oppositely charged N-terminal domain (NTD) and CTD are juxtaposed, and the flexible linker between them is not extended. The modeled structure of

ARC is supported by intramolecular live-cell Förster resonance energy transfer imaging in rat hippocampal slices. Peptides from several postsynaptic proteins, including stargazin, bind to the N-lobe, but not to the C-lobe, of the bilobar CTD. This interaction does not induce large-scale conformational changes in the CTD or flanking unfolded regions. The ARC NTD contains long helices, predicted to form an anti-parallel coiled coil; binding of ARC to phospholipid membranes requires the NTD. Our data support a role for the ARC NTD in oligomerization as well as lipid membrane binding. The findings have important implications for the structural organization of ARC with respect to distinct functions, such as postsynaptic signal transduction and virus-like capsid formation.

Keywords: activity-regulated cytoskeleton-associated protein, FRET, membrane binding, protein structure, small-angle X-ray scattering.

J. Neurochem. (2018) **147**, 323–343.

The activity-regulated cytoskeleton-associated protein (ARC) is required for long-term synaptic plasticity and memory formation. ARC binds to distinct protein partners in postsynaptic neuronal compartments to mediate a broad range of effects. In long-term depression, ARC interacts with components of the clathrin-mediated endocytosis machinery to

Abbreviations used: AMPA, α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid; ARC, activity-regulated cytoskeleton-associated protein; CA, capsid protein; CTD, C-terminal domain; DLS, dynamic light scattering; DOPC, 1,2-dioleoyl-*sn*-glycero-3-phosphocholine; DOPE, 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine; DOPG, 1,2-dioleoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol); DOPS, 1,2-dioleoyl-*sn*-glycero-3-phospho-*L*-serine; DTT, dithiothreitol; FLIM, fluorescence lifetime imaging; FRET, Förster resonance energy transfer; GKAP, guanylate kinase-associated protein; ITC, isothermal titration calorimetry; LTD, long-term depression; LTP, long-term potentiation; MA, matrix protein; mCh, mCherry; mT2, mTurquoise2; NMDA, N-methyl-D-aspartate; NSD, normalized spatial discrepancy; NTD, N-terminal domain; PDB, Protein Data Bank; PSD, postsynaptic density; SAXS, small-angle X-ray scattering; SEC-MALS, size exclusion chromatography – multi-angle light scattering; SRCD, synchrotron radiation circular dichroism spectroscopy; WAVE1, Wiskott–Aldrich syndrome protein family member 1.

Received May 9, 2018; revised manuscript received June 29, 2018; accepted July 12, 2018.

Address correspondence and reprint requests to Petri Kursula, Department of Biomedicine, University of Bergen, Jonas Lies vei 91, N-5009 Bergen, Norway. E-mail: petri.kursula@uib.no

¹These authors contributed equally to this work.

promote endocytosis of α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors (DaSilva *et al.* 2016). In long-term potentiation, ARC enables stabilization of nascent F-actin in dendritic spines (Messaoudi *et al.* 2007; Nair *et al.* 2017). ARC also acts in the nucleus to regulate gene transcription underlying homeostatic synaptic plasticity (Korb *et al.* 2013). Genetic variants of the ARC protein complex are associated with human general intelligence and schizophrenia risk (Fernández *et al.* 2017; Myrum *et al.* 2017). Thus, current evidence supports a view of ARC as a functionally versatile hub crucial for synaptic plasticity and aspects of human cognition (Nikolaienko *et al.* 2018).

Structural predictions and biochemical analyses suggest two folded regions flanking a central linker region in human ARC (Myrum *et al.* 2015). The C-terminal domain (CTD) of ARC contains two homologous domains, which may form a bilobar structure (Zhang *et al.* 2015). These small domains, the N- and C-lobe, show structural homology to the HIV Gag capsid protein CA. The N-lobe was co-crystallized with a peptide from the cytosolic tail of the AMPA receptor-binding protein stargazin, and the same binding pocket may accommodate other ligand peptides (Zhang *et al.* 2015), such as those from guanylate kinase-associated protein (GKAP), Wiskott-Aldrich syndrome protein family member 1 (WAVE1), and the glutamate receptor GluN2A. The interaction between ARC and GKAP is compatible with GKAP-driven activity-dependent reorganization of the postsynaptic density (PSD) structural network (Shin *et al.* 2012). WAVE1, a regulator of actin filament branching in dendritic spines, is a potential mediator of ARC in structural plasticity. Interaction with GluN2A further suggests a connection to N-methyl-D-aspartate (NMDA) receptors. The structure and function of the ARC N-terminal domain (NTD) remain unknown thus far.

Full-length recombinant ARC is capable of reversible self-oligomerization in water, while in salt-containing buffers, it is prone to aggregation (Byers *et al.* 2015; Myrum *et al.* 2015). Recently, ARC was shown to form virus-like capsid structures; these capsids contain RNA and are released from neurons in extracellular vesicles (Ashley *et al.* 2018;

Pastuzyn *et al.* 2018). Alternating oligomeric states may have functional relevance in ARC signalling and capsid formation, but the details of such mechanisms at the molecular level are lacking. Pure monomeric forms of full-length ARC have not been reported, while being indispensable for detailed structure-function studies.

We purified monomeric full-length human ARC and carried out structural and biophysical characterization using small-angle X-ray scattering (SAXS) and synchrotron radiation circular dichroism spectroscopy (SRCD). Analysis of full-length and truncated forms of ARC revealed a compact structure, in which the oppositely charged N- and C-terminal domains are juxtaposed. Peptide ligands bound specifically to the ARC N-lobe and did not induce structural changes within the lobes or flanking unfolded regions. Furthermore, the NTD of ARC contains two long positively charged helices and mediates binding to lipid membranes, possibly by interaction with phospholipid headgroups. The structural model was validated through Förster resonance energy transfer (FRET) imaging in hippocampal slices. Based on structural evidence, we posit distinct functional modalities for ARC in postsynaptic signal transduction and virus-like capsid formation.

Materials and methods

Notes on study design

The study did not involve pre-registration, randomization, or blinding. A 3D structure of monomeric full-length ARC, validated by live-cell FRET imaging, was considered the primary endpoint of the study. Analysis of ligand peptide and lipid membrane binding by ARC were secondary endpoints. A flowchart of the experiments carried out within this study is shown in Fig. 1. All chemicals were from Sigma-Aldrich (Saint Louis, MI, USA; RRID:SCR_008988), unless otherwise indicated. Synthetic lipids were from either Larodan (Malmö, Sweden) or Avanti Polar Lipids (Alabaster, AL, USA).

Ethics statement

This research was approved by Norwegian National Research Ethics Committee in compliance with EU Directive 2010/63/EU, ARRIVE guidelines. Persons involved in animal experiments have Federation

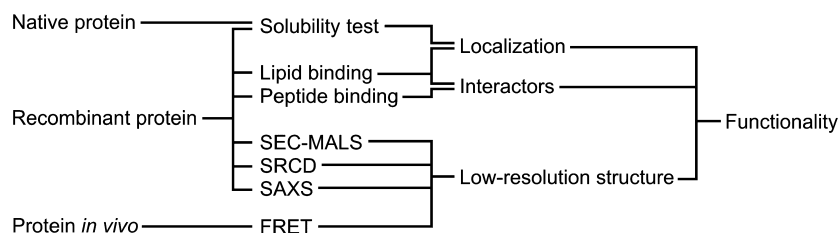


Fig. 1 Flowchart of the work carried out in the current study. The various starting points for experiments are shown, followed by techniques used, and the main outcomes of the experiments, from left

to right, respectively. All experiments aimed at uncovering the main goal, the molecular basis for the functionality of activity-regulated cytoskeleton-associated protein (ARC) (right).

of Laboratory and Animal Science Associations (FELASA) C course certificates and training. The rats used to obtain brain tissue were anesthetized using urethane and sacrificed without suffering.

Statistics

No statistical methods were employed to predetermine sample size of any of the presented experiments. Outliers in SAXS data were removed using standard protocols in the field, incorporated into automated data processing pipelines in essence, multiple consecutive X-ray exposures of the sample were automatically compared, and only the frames presenting no signs of radiation damage were averaged and used further in the analyses. Fits of models to the SAXS raw data (χ^2 values) and to each other (normalized spatial discrepancy; NSD) were calculated by the modelling programs as specified below. Normality of the data for statistical purposes was not assessed in this study, and no specific statistical software or tests were used.

Recombinant protein production

Full-length human ARC and six truncated forms (Table 1, Table S1, Fig. 2a) were expressed in *Escherichia coli* BL21(DE3) cells (RRID: NCBITaxon:511693), resulting in an N-terminally His-tagged protein cleavable by TEV protease. The constructs ARC₂₀₆₋₃₉₆, ARC₂₀₇₋₂₇₇, and ARC₂₇₇₋₃₇₀ were expressed with a maltose-binding protein using pETMBP_1a and ARC₁₇₀₋₃₉₆ with a ZZ-domain tag using pETZZ_1a (Bogomolovas *et al.* 2009; Myrum *et al.* 2015), both tags being located after the His-tag and removed during purification. Full-length ARC was expressed from pD441-SR (ATUM, Newark, CA, USA), and the remaining variants were produced using the pTH27 vector (Hammarström *et al.* 2006). For pETMBP_1a and pETZZ_1a ARC variants, restriction sites *NcoI* and *Acc65I* were used in cloning. pTH27 was used in homologous recombination-based cloning *via* the Gateway system (Thermo Fisher; RRID:SCR_008452).

For the expression of truncated ARC constructs, cells were grown at +37°C until an A₆₀₀ of 0.6 was reached. Protein expression was induced for 2–4 h at +30°C, using 1 mM isopropyl β-D-1-thiogalactopyranoside (Cat# I6758, Sigma-Aldrich). The cells were lysed in HEPES-buffered saline (HBS) (40 mM HEPES, 100 mM NaCl, pH 7.5) containing 0.1 mg/mL lysozyme (Cat# L6876, Sigma-Aldrich), using one freeze-thaw cycle followed by sonication. In the case of ARC₁₃₁₋₃₉₆, which contains one cysteine, 1 mM

dithiothreitol (DTT) was present in all of the purification steps until size exclusion chromatography (SEC).

The lysate was centrifuged at 16 000 g for 30 min at +4°C, and the soluble fraction was loaded onto an Ni-NTA resin (Cat# 30230, Qiagen, RRID: SCR_008539), which was washed with HBS containing 20 mM imidazole (Cat# I5513, Sigma-Aldrich). The bound protein was eluted with HBS containing 300 mM imidazole and cleaved with His-tagged TEV protease (van den Berg *et al.* 2006). The protein solution was dialyzed for 16–20 h at +4°C against a reservoir of 20 mM HEPES (pH 7.5), 100 mM NaCl, and 1 mM DTT. The dialyzed sample was passed through an Ni-NTA resin for removal of the TEV protease and the cleaved His-tag. The final purification step was SEC, using a Superdex S200 16/600 column (GE Healthcare, Chicago, IL, USA; RRID:SCR_000004) equilibrated with Tris-buffered saline (TBS) (20 mM Tris-HCl, pH 7.4, 150 mM NaCl) for all of the ARC truncation constructs except for ARC₂₇₇₋₃₇₀ and ARC₂₀₇₋₂₇₇, which were purified using a Superdex S75 16/600 column (GE Healthcare). The monomeric protein peak was collected and concentrated using spin concentrators. In the case of ARC₁₃₁₋₃₉₆, 1 mM DTT was added to the solution before concentration.

The purification of full-length ARC (ARC₁₋₃₉₆) could not be performed using the protocol above, due to the formation of inclusion bodies during expression. The cells were induced as above, but kept at +37°C for the induction period. After lysis as above, with 1 mM DTT present, the insoluble fraction was washed with HBS containing 1 mM DTT and resuspended in 50 mM CAPS, 50 mM Tris, 10 mM DTT (pH 12). The sample was centrifuged at 16 000 g for 10 min at +25°C, and the soluble fraction was diluted 10-fold with HBS. For protein preparation at high pH, the dilution was made with 20 mM CAPS (pH 11) instead of HBS. The diluted protein solution was loaded onto an Ni-NTA resin, washed with HBS containing 1 mM DTT and 20 mM imidazole before elution with HBS containing 1 mM DTT and 300 mM imidazole. The eluted protein was loaded onto a Superdex S200 16/600 column, equilibrated with TBS, and the monomeric protein peak was collected. In the case of protein preparation at high pH, Tris was again replaced with 20 mM CAPS (pH 11). The monomeric fraction was supplemented with 1 mM DTT and concentrated. The His tag was not removed from full-length ARC.

Protein concentrations were determined by absorbance at 280 nm. Protein purity was analyzed with sodium dodecyl

Table 1 ARC constructs and their properties. The numbering is based on the human ARC sequence. Biophysical parameters obtained using SAXS, SEC-MALS, and DLS for the different recombinant ARC constructs. R_g^* is the R_g calculated for a perfect sphere with the *ab initio* volume of the protein

Protein name	Description	R_g (nm)	R_h (nm)	R_g/R_h	R_g/R_g^*	D_{max} (nm)	MW _{calc} (kDa)	MW _{MALS} (kDa)	<i>ab initio</i> volume/ MW _{MALS} (nm ³ /kDa)
ARC ₁₋₃₉₆ (pH 7)	Full-length ARC	3.58	N/A	N/A	1.43	14.37	47.9	50.1	2.78
ARC ₁₋₃₉₆ (pH 11)	Full-length ARC	3.85	N/A	N/A	1.51	15.0	47.9	52.8	2.86
ARC ₁₃₁₋₃₉₆	Linker region and CTD	3.71	3.69	1.01	1.98	14.01	30.2	34.4	1.72
ARC ₁₇₀₋₃₉₆	Half linker and CTD	3.03	3.25	0.93	1.60	12.20	26.4	27.8	2.17
ARC ₂₀₆₋₃₉₆	C-terminal domain	2.80	2.92	0.96	1.66	11.94	22.6	25.7	1.69
ARC ₂₀₆₋₃₆₁	CTD without C-tail	2.25	2.40	0.94	1.41	8.74	18.7	20.7	1.79
ARC ₂₀₇₋₂₇₇	N-lobe of CTD	1.43	1.92	0.74	1.24	6.13	8.9	9.7	1.37
ARC ₂₇₇₋₃₇₀	C-lobe of CTD	1.69	2.03	0.83	1.21	5.95	11.3	12.3	2.00

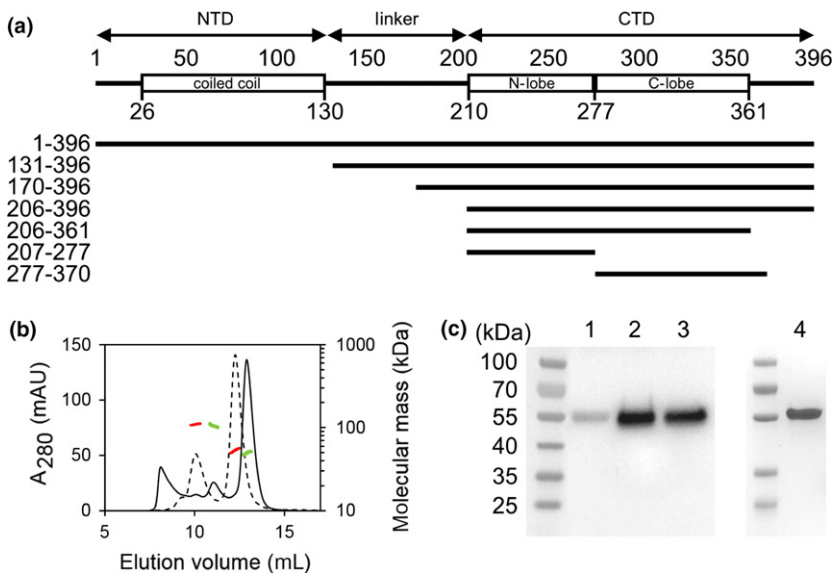


Fig. 2 Purification of soluble monomeric activity-regulated cytoskeleton-associated protein (ARC). (a) The recombinant ARC variants analyzed in this work. (b) SEC-MALS analysis of purified full-length ARC at pH 7 (solid chromatogram, mass shown in green) and 11 (dashed, mass shown in red). (c) Protein extraction from rat hippocampus analyzed on SDS-PAGE and western blotting using an anti-ARC antibody. 1 – The soluble fraction after homogenization at neutral pH; 2 – The insoluble fraction from neutral pH, re-solubilized in pH 12; 3 – The insoluble fraction solubilized in neutral pH with 2% SDS; 4 – Purified recombinant full-length ARC (Coomassie staining).

sulphate–polyacrylamide gel electrophoresis, giving a single strong Coomassie-stained band. Protein identity was confirmed using mass spectrometry of trypsin-digested in-gel samples, as described (Raasakka *et al.* 2015).

ARC-interacting peptides (Zhang *et al.* 2015) were synthesized by GenScript (Hong Kong, China; RRID:SCR_002891) with N-terminal acetylation and C-terminal amidation. The peptides were from Stargazin (RIPSYRYR), GKAP (TSPKFRSR), WAVE1 (RTPVVFVSP), and GluN2A (RNPLHNED).

Size exclusion chromatography – multi-angle light scattering (SEC-MALS)

Absolute molecular masses were determined with MALS, using either a DAWN Heleos (Wyatt), or a miniDAWN Treos MALS detector (for ARC₁₋₃₉₆ and ARC₂₀₇₋₂₇₇) (Wyatt). The SEC column used was either a Superdex S200 Increase 3.2/300 (GE Healthcare) or, in the case of ARC₁₋₃₉₆ and ARC₂₀₇₋₂₇₇, a Superdex S200 Increase 10/300 (GE Healthcare). The SEC-MALS systems were calibrated with ovalbumin, and the running buffer was TBS. Protein concentration was measured with an online refractometer, and the hydrodynamic radius with either an online quasi-elastic light scattering module or, in the case of ARC₂₀₇₋₂₇₇, a Zetasizer Nano (Malvern Instruments, UK). Data were processed and analyzed with ASTRA (Wyatt; RRID: SCR_016255).

Extraction of native ARC from rat brain tissue

One *Cornu Ammonis* region of the male Sprague Dawley rat (RRID: NCBITaxon:10116) hippocampus, corresponding to ~100 mg of tissue, was homogenized in 400 μ L of a buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1 mM EDTA, 1 mM DTT, and 0.1% Triton X-100, using a Potter-Elvehjem homogenizer and divided in two. The rats were 2–4 months of age and weighed between 250 and 300 g. The lysate was centrifuged at 21 000 g for 15 min at +4°C. The supernatants correspond to the soluble protein fraction, and the pellets were resuspended in 200 μ L of either the above buffer containing 2% sodium dodecyl sulphate (SDS) or a buffer containing 50 mM CAPS (pH 12), 150 mM NaCl, 1 mM

EDTA, 1 mM DTT, and 0.1% Triton X-100. The resuspended samples were centrifuged as above, and the supernatants correspond to proteins dissolved by SDS or high pH, respectively.

Supernatants were mixed with Laemmli sample buffer, heated at +95°C for 5 min and loaded on a 10% polyacrylamide gel. Proteins were separated by electrophoresis in TGS buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3, Bio-Rad) before being electro-transferred to a nitrocellulose membrane, using the Trans Blot Turbo transfer system (Bio-Rad). The membrane was blocked for 1 h in TBS-T (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1% (vol/vol) Tween 20) + 5% dry milk, prior to incubation with primary and secondary antibody for 1 h at +22°C. The membrane was washed 3 \times 5 min in TBS-T between incubations and before immunodetection using Pierce ECL Western Blotting Substrate (Thermo Scientific, Waltham, MA, USA), according to the manufacturer's instructions. Antibodies used were: polyclonal rabbit anti-ARC antibodies (RRID:AB_887694, Synaptic Systems, Germany; RRID:SCR_013612) and horseradish peroxidase-conjugated anti-rabbit secondary antibodies (RRID: AB_92641, Synaptic Systems, Germany; RRID:SCR_013612).

Circular dichroism spectroscopy

Folding of ARC was followed by SRCD. Truncated ARC was diluted into 20 mM phosphate buffer (pH 7.4) with 150 mM NaF and 0.5 mM DTT, while the full-length construct was desalted using a PD-10 column into the same buffer. For the sample at pH 11, phosphate was replaced with 20 mM CAPS (pH 11). Protein concentrations were 6–66 μ M. For mixtures of ARC₂₀₇₋₂₇₇ and ligand peptides, ARC₂₀₇₋₂₇₇ was at 67 μ M, and the peptides were present either in equimolar amounts or at a fourfold molar excess.

SRCD spectra were measured between 170 and 280 nm at +10°C, using a 100- μ m quartz cuvette, on the AU-CD beamline at the ASTRID2 synchrotron storage ring (ISA, Aarhus, Denmark). Deconvolutions were made with Dichroweb (Whitmore and Wallace 2004) using CDSSTR (Johnson 1999) and SP175 (Lees *et al.* 2006) reference databases to estimate secondary structure content.

CD data comparing the ARC CTD at neutral and high pH and the effect of lipids on the full-length ARC protein were collected using a J-

810 Spectropolarimeter (Jasco) and a 1-mm quartz cuvette at +20°C. The protein concentration for ARC CTD was 5 µM in a buffer consisting either of 20 mM phosphate (pH 7.4) or 10 mM CAPS (pH 11). The protein concentration of full length ARC was 1.5 µM in a buffer consisting of Tris-HCl (10 mM, pH 7.4), NaF (75 mM) with and without liposomes of 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC)/ 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPS) (1 : 1 ratio, 0.5 mM).

Small-angle X-ray scattering

SAXS data were collected on the BM29 beamline (Pernot *et al.* 2013) of ESRF (Grenoble, France), the B21 beamline at Diamond (Oxfordshire, UK), the P12 beamline (Blanchet *et al.* 2015) at European Molecular Biology Laboratory/DESY (Hamburg, Germany), and the SWING beamline at SOLEIL (Gif-sur-Yvette, France). The data used to generate models in the absence of bound peptide ligands were collected using a SEC-SAXS setup, where SAXS frames are collected as the protein elutes from an SEC column. The columns used were BioSEC-3 (Agilent; RRID: SCR_013575), Superdex S200 Increase 3.2/300 (GE Healthcare), or Superdex S200 Increase 10/300 (GE Healthcare), equilibrated with TBS. For full-length ARC at high pH, the latter column was used with 20 mM CAPS (pH 11) instead of Tris-HCl. For final model building, full-length ARC data from neutral pH was used.

SAXS measurements of protein/peptide mixtures (0.3–10 mg/mL in TBS) were made in batch mode with a 4-fold molar excess of peptide. For the ARC₂₀₇₋₂₇₇ and ARC₂₇₇₋₃₇₀ constructs with peptides, the buffer also contained 2 mM DTT and 2% glycerol. All SAXS measurements were done at +10°C.

SAXS data were processed using either ATSAS 2.8 (Franke *et al.* 2017) (RRID:SCR_015648) or FOXTROT3 (SOLEIL synchrotron). The collected frames were checked for radiation damage. The samples measured in batch mode were analyzed at different concentrations to avoid intermolecular events. SAXS models were generated using the programs DAMMIN (Svergun 1999), GASBOR (Svergun *et al.* 2001), CORAL (Petoukhov *et al.* 2012), and EOM (Bernadó *et al.* 2007; Tria *et al.* 2015) from the ATSAS 2.8 package. In the CORAL run, full-length ARC was modeled by simultaneously including the data from all constructs of different length. The molecular mass of samples in batch mode was based on either absolute scale or a bovine serum albumin standard.

Molecular dynamics simulations

The crystal structure of the ARC N-lobe domain (Protein Data Bank (PDB) entry 4X3H, residue 210–277), with or without the stargazin peptide (RIPSYRYR) was used for atomistic molecular dynamics simulations by GROMACS 5.1.4 (Abraham *et al.* 2015) (RRID: SCR_014565), using the GROMOS96 53A6 force field (Oostenbrink *et al.* 2004) under NPT (constant particle number, pressure, and temperature) conditions at 300 K with a time step of 2 fs. The protein was simulated in a cubic box (5 × 5 × 5 nm³) with water as solvent. The total charge of the system was neutralized using sodium ions. The R_g of the protein was calculated at intervals of 10 ps.

FRET experiments

DNA constructs

FRET sensors for *in vivo* FRET imaging were assembled in the pcDNA3.1(+) vector (Cat# V79020, Thermo Fisher Scientific, RRID:

SCR_008452) behind the cytomegalovirus promoter. The original polylinker sequence between the *NheI* and *XhoI* restriction sites was replaced with a custom polylinker containing a Kozak sequence, start and stop codons, as well as additional restriction sites to facilitate subcloning (5'-GCTAGC-ACTAGT-ACC-ATG-ACCGGT-GCG ATCGC-A-GGATCC-GCGGCCGC-A-TCCGGA-TTAATTAA-A-TAA-CTCGAG-3'). The constructs encoding mTurquoise2 (pLifAct-mTurquoise2, Addgene (RRID:SCR_002037) plasmid #36201) (Goedhart *et al.* 2012), YPet (pCEP4YPet-MAMM, Addgene plasmid #14032) (Nguyen and Daugherty 2005), P2A sequence (pAAV-hSyn1-mRuby2-GSG-P2A-GCaMP6s-WPRE-pA, Addgene plasmid #50942) (Rose *et al.* 2016), and rat ARC (pGEX-4T-3-Arc) (Nikolaienko *et al.* 2017) have been described previously. The constructs encoding EGFP (pEGFP-N1) and mCherry (pmCherry-C1) were from Takara Bio (Mountain View, CA, USA; RRID:SCR_003960). A positive FRET control plasmid was obtained by introducing mTurquoise2 and YPet (or EGFP and mCherry) between *AgeI/AsiSI* and *BspEI/PacI* restriction sites, respectively. The resulting plasmid was later used to subclone the P2A sequence (negative FRET control), full-length ARC, the NTD of ARC (ARC₁₋₁₄₀), the ARC central linker (ARC₁₃₅₋₂₁₆), and the CTD of ARC (ARC₂₀₈₋₃₉₆) between *BamHI* and *NotI* restriction sites.

Slice culture preparation and DNA transfection

Transverse hippocampal slice cultures were prepared from male Wistar Hannover GALAS outbred rats (RRID:RGD_1566433) at P10 ± 1 and maintained for 12–21 days as f GASBOR described (Otmakhov *et al.* 2004; Gogolla *et al.* 2006) before DNA transfection by single-cell electroporation. Slices were prepared according to the interface method (De Simoni and Yu 2006). The rat pups were patted with ethanol-soaked paper and immediately decapitated using sharp surgical scissors. The skull was exposed by cutting into the skin along the midline of the head from the neck to the eyes before flipping it and pulling it over the skull. The skull was cut along the interaural line and the sagittal suture, and then peeled away using forceps. The brain was gently scooped out from the skull using a spatula and dropped into pre-chilled dissection solution (on ice). Once chilled (10 min), the dissected brains were decanted into a 60-mm dish resting on ice. Using a pair of small forceps, the cortices were separated from the brainstem, midbrain, and striatum. Each cortex was placed upside down, revealing the hippocampus. Curved forceps were used to gently flip the hippocampus over and cut it free from the occipital cortex region. After detachment from the cortex, larger blood vessels and excess tissue were carefully removed without stretching the hippocampus. The isolated hippocampi were extracted from the solution with a wide-bore Pasteur pipette and were lined up perpendicularly to the blade on the Teflon plate of the tissue chopper (McIlwain, RRID:SCR_015798). The hippocampi were chopped into 350–400 µm slices and rinsed off the teflon plate with chilled dissection solution. Slices suspended in dissection solution in a 60-mm petri dish were examined under a dissection microscope. Slices, which displayed clear, undamaged hippocampal structures, CA1–3 regions and dentate gyrus, were extracted, using a wide-bore pasteur pipette and placed on Millicell cell culture inserts (Cat# PICM03050, RRID:SCR_015799, Merck Millipore, RRID: SCR_001287) in 6-well plates with pre-warmed culture medium. For gene gun transfection, slices were prepared from Sprague Dawley rats at P8–10 and maintained for 7–10 days before transfection. The number of animals per cage was 4. The cage size

was $20 \times 36 \times 45$ cm, and the total number of animals used for this study was 50, with a body mass between 15 and 20 g. Animals were bred in-house in IVC IV cages containing two females under AAALAC International certification (nesting material, food and water *ad libitum*, $+22 \pm 1^\circ\text{C}$, 60–65% relative humidity, 12 h light/12 h dark alternate cycles).

DNA transfection was performed by single-cell electroporation (Haas *et al.* 2001, 2002) for ratiometric FRET experiments and by Helios Gene Gun (Bio-Rad; RRID:SCR_008426) for fluorescence lifetime imaging (FLIM)-FRET experiments. Single slices were transferred to a glass-bottom chamber (Leica) on a motorized shifting table (Luigs and Neumann, Germany), and continuously perfused (flow rate ~ 3 mL/min) with artificial cerebrospinal fluid of the following composition: 124 mM NaCl, 2.5 mM KCl, 2 mM CaCl₂, 1.5 mM MgCl₂, 1.25 mM NaH₂PO₄, 26 mM NaHCO₃, 10 mM D-glucose, balanced with Carbogen gas (95% O₂, 5% CO₂), pH 7.4 at $+22^\circ\text{C}$. Electrical pulse parameters for single-cell electroporation were pulse-width 1 ms, -10 V, at 200 Hz for 200–500 ms. After transfection, insert membranes (Millicell Cell Culture Insert, PICM03050, RRID:SCR_015799, Merck Millipore, RRID:SCR_001287) with slices were returned to the CO₂ incubator and maintained in culture medium until use. For transfection by gene gun, 1.6- μm gold microcarriers (Cat# 1652264, Bio-Rad, RRID:SCR_008426) were coated with plasmid DNA and fired directly into individual wells in 6-well culture dishes, using helium gas (O'Brien and Lummis 2006).

Imaging and data analysis

Imaging was performed 48 h after electroporation for all constructs except full-length ARC, which was imaged 24–28 h after electroporation. Slices were placed in the imaging chamber, and FRET ratios were determined after 15–25 min, using a Leica SP5 upright confocal microscope (DMI 6000 CS, Leica Microsystems, Germany; RRID:SCR_008960) with a water immersion objective HCX APO L 20 \times 1.0 W (Multiphoton, Leica Microsystems, Germany; RRID:SCR_008960). For excitation of the donor fluorophore, mTurquoise₂ (mT2), an Argon laser at 458 nm was used. Emission light was separated using an internal acousto optic beam splitter system, where 465–500 nm light was designated donor emission (mT2) and 525–620 nm light acceptor emission. The signal was captured in separate photomultiplier tubes using XYZ mode bidirectional scanning at 700 Hz, with a pinhole diameter of 180 μm for 25–32 focal steps (1.2–1.5 μm). Each focal plane image (512 \times 512 pixels, 0.5 μm pixel size) was averaged (Line average = 3, Frame average = 3, internal settings of LAS AF software, RRID:SCR_013673).

Here, two-photon FLIM-FRET imaging was performed on an Olympus FV1000 upright microscope through a 60X water immersion objective (NA=1) with an SPC-830 (Becker & Hickl) photon counting board for time-correlated single-photon counting. Photons were counted for 10–20 s, depending on expression level, in 64 \times 64 pixels. The donor (EGFP) was excited using a 910-nm 2-photon laser (Spectra-Physics), and emission was captured in an H7422-40 detector (Hamamatsu). Fluorescence lifetime was calculated in SPC-image software (Becker & Hickl), using mono-exponential curve fitting of photon distributions.

All image processing and analysis for ratiometric FRET experiments was performed in Fiji (RRID:SCR_002285) (Schindelin

et al. 2012). Raw image files (.lif, LAS AF software, RRID:SCR_013673) were imported, using the Bioformats importer plugin (Linkert *et al.* 2010) and split into two colour channels (donor, acceptor). Background signal was subtracted from both channels, using the rolling ball background subtraction function before the XYZ stacks were projected to two dimensions (XY), using a sum-intensity projection, and a threshold was manually applied to roughly segment the cells. The image calculator was used to divide the acceptor channel by the donor channel, yielding a 32-bit ratio image. Finally, the mean ratio for each soma was calculated. The data are presented as the mean \pm SD of the total sample size.

Negative and positive FRET control constructs were used to define the maximum and minimum FRET achievable for this imaging system and FRET pair. The negative control features the donor and acceptor separated by a P2A sequence, which yields cleavage in the polypeptide during translation. This ensures equimolar expression of donor and acceptor freely diffusing in the cell. In this disconnected state, the local fluorophore concentration is too low for detectable FRET to occur, such that any FRET signal detected is a result of donor/acceptor crosstalk (spectral bleed-through, direct excitation of acceptor) and background signal. In the positive control, the donor and acceptor fluorophores are separated by a 10-amino-acid sequence, short enough to ensure constitutive FRET.

Isothermal titration calorimetry

A MicroCal iTC200 (Malvern, UK) instrument was used to determine the affinities of peptides to the N-lobe (ARC₂₀₇₋₂₇₇) by isothermal titration calorimetry (ITC). The peptides (2.2 mM) were injected into the protein solution (0.22 mM) by 26 3- μL injections of the peptide solution, with one initial injection of 0.5 μL and a second 0.5- μL injection after 14 injections due to syringe refill. The protein and the peptides were in TBS buffer, and the experiments were done at $+25^\circ\text{C}$. The data were analyzed with MicroCal Origin 7 (RRID:SCR_002815), using a one-site-binding model.

Liposome co-sedimentation assay

Liposomes consisting of 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC), 1,2-dioleoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) (DOPG), 1,2-dioleoyl-*sn*-glycero-3-phospho-*L*-serine (DOPS) and 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine (DOPE) were prepared in various combinations by mixing the dissolved lipids in chloroform, drying using a freeze-dryer, and resuspending in TBS. After 15–20 h of agitation at $+25^\circ\text{C}$ and 7 freeze-thaw cycles, the suspension was passed through a 100-nm filter 11 times, resulting in 100-nm liposomes. Liposome size was confirmed by dynamic light scattering on a Zetasizer Nano instrument (Malvern, UK).

Co-sedimentation of ARC constructs was performed by mixing liposomes (0.5 mM DOPC/DOPS, 1 : 1) and protein (1.5 μM of ARC₁₋₃₉₆ and 1.9 μM of ARC₁₃₁₋₃₉₆ and ARC₂₀₆₋₃₉₆) in TBS. After 1 h at $+4^\circ\text{C}$, the solutions were centrifuged at 170 000 g for 60 min at $+4^\circ\text{C}$. The supernatant was removed and the pellet resuspended in TBS. Controls were made without liposomes, as well as with 130 mM sodium phosphate (pH 7.4) present. Samples were loaded on 4–20% gradient sodium dodecyl sulphate–polyacrylamide gel electrophoresis gels (Bio-Rad) and Coomassie-stained.

Co-sedimentation of full-length ARC (3.9 μM) with liposomes was also performed using various lipids (0.5 mM): DOPC/DOPG (1 : 1), DOPC/DOPS (1 : 1 and 4 : 1), DOPC/DOPE (1 : 1), and

DOPC alone. The protein/lipid ratio was varied by increasing the concentration of DOPC/DOPS (1 : 1), and the effect of NaCl concentration was tested using DOPC/DOPS (1 : 1, 0.5 mM) liposomes. The solubility of pelleted liposome/protein complexes (DOPC/DOPS, 1 : 1, 0.5 mM) after co-sedimentation with full-length ARC (3.9 μ M) was tested by resuspension in either TBS, pH 7.4 or 50 mM CAPS, 50 mM Tris, pH 12, followed by another centrifugation as described above.

Results

Purification of monomeric full-length ARC

In order to facilitate structural studies on full-length ARC and its domains, we set up recombinant expression systems for a number of ARC constructs (Fig. 2a, Table 1, Table S1). Bacterially expressed full-length ARC was insoluble under a variety of conditions, as reported before (Byers *et al.* 2015; Myrum *et al.* 2015). However, high pH made ARC soluble and suitable for structural analyses. In the presence of salt and reducing agent, a monomeric peak dominated the size exclusion chromatogram (Fig. 2b). The oligomeric state was determined with SEC-MALS and SAXS (Table 1, Fig. 2b). The monomeric state was preserved, when pH was reduced to a neutral value. If concentrated above \sim 1 mg/mL at neutral pH, ARC slowly aggregated, showing different oligomeric states. At high pH, monomeric ARC could be concentrated without aggregation.

All truncated variants lacking the N-terminal domain were soluble, monodisperse, and monomeric (Figure S1, Table 1), and they could be concentrated to $>$ 10 mg/ml. All these variants eluted from SEC as a single peak (Figure S1), and the molecular mass across the peak, determined by MALS, corresponded to a monomer (Table 1). Similarly, modelling the truncated constructs from the single peak by SEC-SAXS showed these variants are monomeric (see below). This indicates that the oligomerizing/aggregating property lies within the N-terminal domain. The NTD alone was insoluble and could not be studied in isolation.

To compare recombinant ARC to ARC *in vivo*, we extracted ARC from brain tissue using similar procedures. Native ARC from rat brain could, indeed, be solubilized at high pH, at levels comparable to SDS extraction (Fig. 2c). Tissue lysis at neutral pH solubilized only a minor fraction of ARC. This behavior of native ARC suggests that the protein is bound to an insoluble component – possibly a membrane – in the cell, through an interaction sensitive to pH and likely involving the N-terminal domain.

ARC has a helical N-terminal domain sensitive to pH

High-quality SRCD spectra show that full-length ARC contains mainly helical structure (Fig. 3a), and deconvolution of spectra from all truncated constructs, as well as selected difference spectra (Fig. 3b), allows a detailed mapping of structured regions (Fig. 3c). Secondary structure

predictions suggest the NTD is helical, while the linker region and C-terminal tail are predicted to be random coils. The CTD and the individual lobe structures give helical spectra (Fig. 3a), as expected from crystal structures of the two lobes (Zhang *et al.* 2015). The crystal structure of the N-lobe consists of 59% helix, 7% strand, and 34% random coil, similar to the secondary structure content observed by SRCD (Fig. 3c). The C-terminal tail and the linker region produce difference spectra corresponding to random coil (Fig. 3b). The NTD gives a strong helical signal, corresponding to $>$ 60% helix. The high 222/208 nm CD signal ratio of the NTD difference CD spectrum (Fig. 3b) suggests the presence of longer and more ideal helical structures (Greenfield 2006) than those in the CTD. A high 222/208 nm CD ratio is indicative of coiled-coil structure (Kammerer *et al.* 1998; Alfadhli *et al.* 2002; Steinmetz *et al.* 2007), which the NTD is predicted to contain.

Structural differences between full-length ARC at neutral and high pH were estimated through SRCD difference spectra, showing that most of the secondary structure of ARC remains intact at high pH (Fig. 3b and c). The lost helical structure is located in the NTD, as the CTD gave identical CD spectra at neutral pH and at pH 11 (Figure S2). Thus, a small fraction of the NTD helical structure unfolds at high pH, while ARC remains overall folded. The position of the monomeric peak of full length ARC in SEC shifted depending on pH, appearing slightly larger at high pH (Fig. 2b). This increase in the hydrodynamic radius is in line with subtle unfolding. The process is reversible, as the sample at neutral pH was originally solubilized under alkaline conditions.

The 3D structure of the ARC NTD is unknown, and its sequence homology to known protein structures is very low. Secondary structure predictions suggest the presence of two long helices, possibly forming an antiparallel coiled coil (Figure S3). *De novo* fold predictions were made with Robetta (Kim *et al.* 2004). This method gave models with antiparallel, elongated helices (Figure S3). Delta-Blast (Boratyn *et al.* 2012) suggests similarity to the FH2 domain of mDia1 (Nezami *et al.* 2010). This structure consists of two elongated helices. With RaptorX (Källberg *et al.* 2012), the best scoring match is to an SPX domain (Wild *et al.* 2016), again suggesting an elongated structure of two antiparallel helices (Figure S3). The NTD of ARC consisting of elongated helices is in good agreement with the results from both SRCD and SAXS (see below).

Solution structure of full-length ARC

The individual N- and C-lobes show SAXS profiles expected for globular domains (Fig. 4a). For the bilobar structure of the CTD, a dumbbell-like shape is evident; the N- and C-lobes are globular but separated (Fig. 4a and b). Full-length ARC has a compact, elongated shape with no signs of isolated globular domains. The maximum diameter of full-length ARC is similar to that of the C-terminal domain

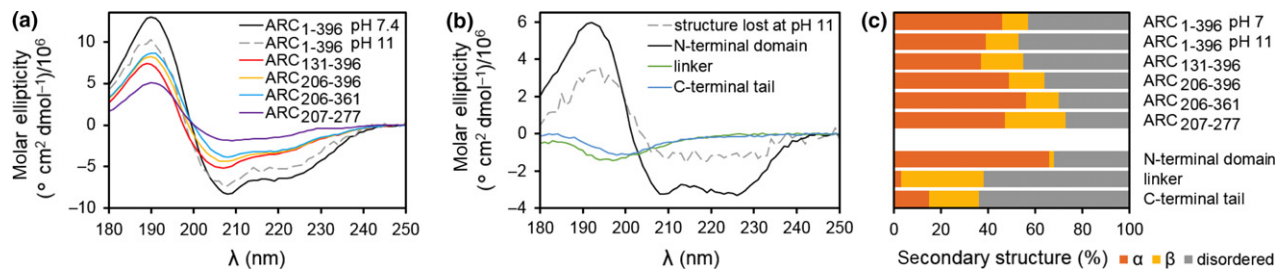


Fig. 3 Analysis of activity-regulated cytoskeleton-associated protein (ARC) folding. (a) synchrotron radiation circular dichroism spectroscopy (SRCD) analysis of ARC constructs. Note that in all CD spectra, ellipticity has been calculated using the molarity of the protein instead of peptide bonds; this is to facilitate comparison between constructs of different length. (b) Difference SRCD spectra. ARC₁₋₃₉₆ at neutral pH minus the

signal at high pH gives the spectrum of the structure that is lost at high pH, ARC₁₋₃₉₆ at neutral pH minus ARC₁₃₁₋₃₉₆ the spectrum of the N-terminal domain, ARC₁₃₁₋₃₉₆ minus ARC₂₀₆₋₃₉₆ the spectrum of the linker region, and ARC₂₀₆₋₃₉₆ minus ARC₂₀₆₋₃₆₁ the spectrum corresponding to the C-terminal tail. (c) Deconvoluted secondary structure contents for all constructs, as well as the segments deduced from difference spectra.

(Fig. 4b, Table 1), showing that the NTD is located near the bilobar structure of the CTD.

Dimensionless Kratky plots and distance distributions show a single maximum for the two lobes alone, as expected for a well-folded single-domain protein, and the peaks are located near the theoretical position for compact globular particles (Fig. 4b and c). The CTD constructs containing both lobes in tandem (ARC₂₀₆₋₃₆₁, ARC₂₀₆₋₃₉₆, ARC₁₇₀₋₃₉₃, ARC₁₃₁₋₃₉₆) show shoulders in such plots, typical for multidomain proteins. Kratky plots indicate a flexible C-terminal tail, and flexibility is increased in constructs with longer central linker regions (Fig. 4c). The Kratky plot for full-length ARC, on the other hand, indicates a high degree of folding, compact structure, and little overall flexibility. Thus, by interacting with the CTD, the NTD limits the flexibility of the linker region in the context of full-length ARC. Furthermore, the R_g of full-length ARC is slightly smaller than that of ARC₁₃₁₋₃₉₆ (Table 1), and the distance distribution plot shows less tailing for ARC₁₋₃₉₆ compared to ARC₁₃₁₋₃₉₆ (Fig. 4b). Thus, the linker region is stabilized by interactions between the N- and C-terminal domains in full-length ARC. A closed state with the NTD located close to the CTD reduces movements of the linker segment, resulting in a smaller particle diameter of full-length ARC compared to the construct without the NTD.

The available structural data provide no evidence for an open conformation of ARC, whereby the N- and C-terminal domains would be dissociated and connected by an extended linker. This theoretical open state would result in a much larger diameter. As the concentration of full-length ARC is increased and the protein starts to form aggregates, an apparent dimer peak appears in SEC (Fig. 2b); this peak could in theory correspond to an open state with a larger R_h . However, SEC-MALS clearly shows that this peak is a dimer and not an open monomer (Fig. 2b). A dimer of ARC could be a relevant first step towards the formation of capsid-like structures.

Various approaches were used to model the 3D structure of ARC based on SAXS data, including dummy atom and chain-like *ab initio* modelling, hybrid modelling coupled with loop building, and prediction of conformational ensembles. Dummy atom models for all constructs were first assembled. The models of the different truncations of ARC can be combined to build a full-length ARC structure (Fig. 4d). Full-length ARC is elongated but compact, and the N- and C-terminal halves are located close together. A comparison of full-length ARC with the model of ARC₁₃₁₋₃₉₆, in which the NTD is missing, shows an empty space along the elongated side of the full-length model. This volume corresponds to the predicted size and shape of the NTD.

Further truncations of the linker region and the C-terminal tail remove elongations on both sides of the models, leaving only the two lobes as a dumbbell shape, demonstrating that the lobe structures are two separated domains (Fig. 4d). Hybrid models of constructs corresponding to the CTD of ARC with and without the linker region and/or the C-terminal tail (Figure S4) show elongated flexible regions corresponding to the linker and C-terminal tail, in agreement with the SRCD difference spectra for these parts of the protein (Fig. 3b).

On one side of the dumbbell-like SAXS model corresponding to ARC₂₀₆₋₃₆₁, an extended feature can be seen, which is also present in the model of the N-lobe (ARC₂₀₇₋₂₇₇) (Fig. 4d). This extension could correspond to the more flexible N-terminal part of the N-lobe, supposed to bind peptides by folding onto them (Zhang *et al.* 2015). Chain-like models (Figure S4B) generally agreed with the dummy residue models, apart from the model of the N-lobe, which gave a bad fit at low q (Figure S3C). Ensemble optimization analysis (Figure S4C and D) of the SAXS data of ARC₂₀₇₋₂₇₇ gives a better fit, indicating a subpopulation of large particles of roughly twice the diameter. Importantly, this extension corresponds to the N-terminal part of the N-lobe that folds on

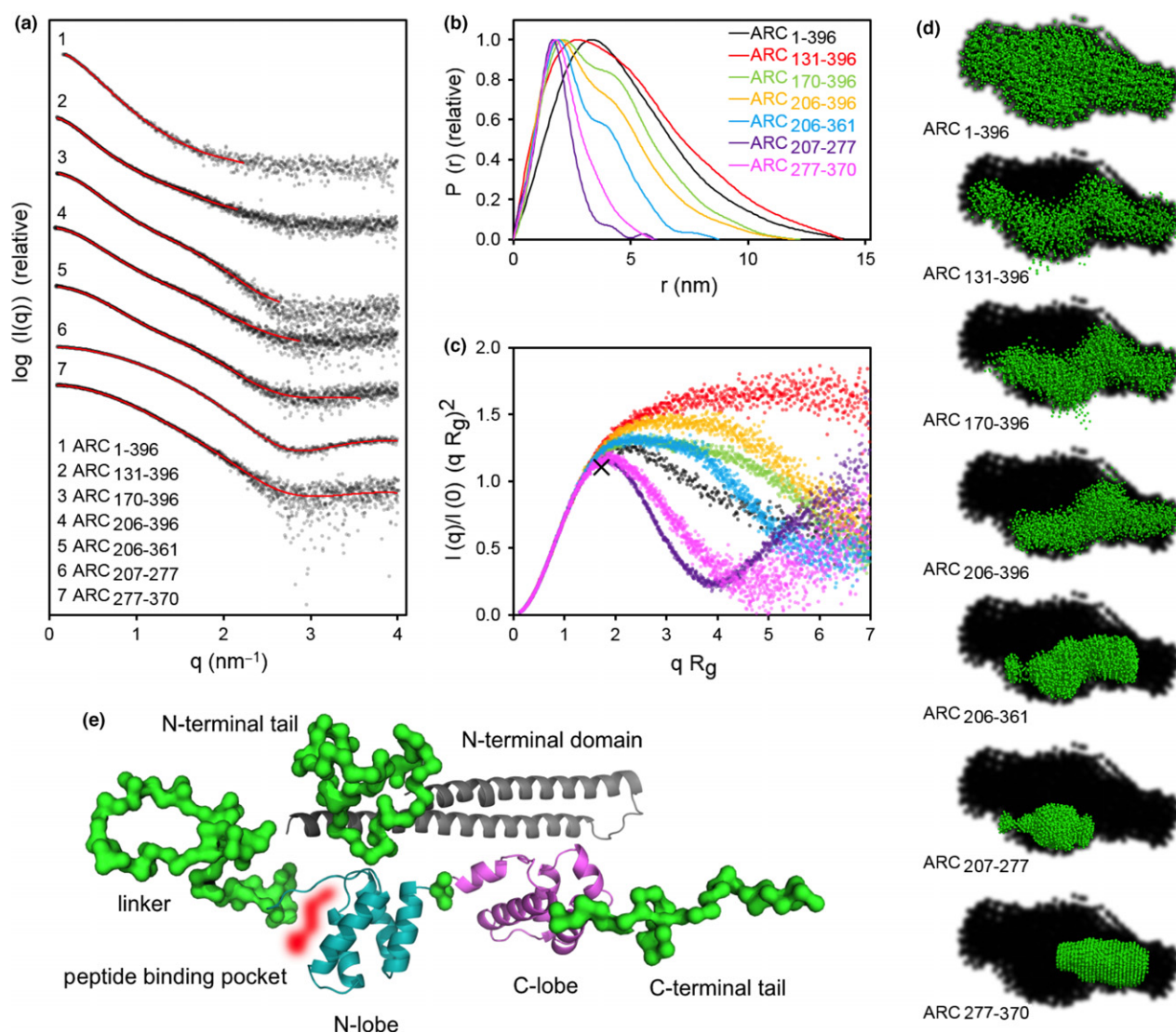


Fig. 4 Structural information from SAXS. (a) Scattering profiles. The data were shifted along the y axis for better visualization. Fits of *ab initio* dummy atom models (d) to the data are shown with red lines. (b) Distance distribution functions. (c) Dimensionless Kratky plots. The cross marks the position of the maximum expected for a rigid, fully spherical particle. (d) Dummy atom SAXS models (10 superimposed models per construct) shown in green on a black background, which represents the shape of full-

length activity-regulated cytoskeleton-associated protein (ARC) (ARC₁₋₃₉₆ $\chi^2 = 0.9$, NSD = 0.6; ARC₁₃₁₋₃₉₆ $\chi^2 = 1.3$, NSD = 0.8; ARC₁₇₀₋₃₉₆ $\chi^2 = 1.4$, NSD = 0.7; ARC₂₀₆₋₃₉₆ $\chi^2 = 1.4$, NSD = 0.6; ARC₂₀₆₋₃₆₁ $\chi^2 = 1.2$, NSD = 0.5; ARC₂₀₇₋₂₇₇ $\chi^2 = 1.4$, NSD = 0.5; ARC₂₇₇₋₃₇₀ $\chi^2 = 1.3$, NSD = 0.4). (e) A hybrid model based on the simultaneous use of SAXS data from all analyzed ARC constructs, the crystal structures of the two lobes, and a homology model of the NTD.

top of the stargazin peptide in the crystal structure (Zhang *et al.* 2015); in the unbound state, it may be flexible and extend away from the globular domain (Figure S4D,E). Note that this structural event was not large enough to alter the R_g value of ARC₂₀₇₋₂₇₇ in the presence and absence of peptides (see below). However, the effect on R_g is, in fact, expected to be minor upon peptide binding, as indicated by molecular dynamics simulations of the N-lobe crystal structure in the presence ($R_g = 2.0 \pm 0.4$ nm) and absence ($R_g = 2.0 \pm 0.5$ nm) of a bound stargazin peptide (Figure S5). Without

the peptide, the N-terminal segment of the N-lobe is flexible, but the average R_g does not differ from that of the peptide complex, which includes a larger globular particle.

A hybrid model of full-length ARC was generated (Fig. 4e), based on SAXS data from all constructs, and using the crystal structures of the N- and C-lobe together with the predicted helical structure of the N-terminal domain simultaneously for modelling. This model positions all different domains of ARC, in good agreement with the *ab initio* models. In the model, the coiled coil of the NTD lies

on top of the bilobar CTD; the N and C termini of ARC lie at opposite ends. The N and C termini and the linker are likely to be flexible chains, but the central linker is restricted by the arrangement of the folded NTD and CTD.

For understanding the molecular properties of different ARC constructs, the R_g values from SAXS, the R_h values from dynamic light scattering, R_g values corresponding to a sphere with the observed molecular volume from SAXS (R_g^*), and the molecular volume/mass ratio were compared (Table 1). For a globular particle, one would expect ratios of $R_g/R_h = 0.775$, $R_g/R_g^* = 1.0$, and volume/mass = $1.212 \text{ nm}^3/\text{kDa}$ (Erickson 2009). From SAXS experiments, a volume/mass ratio of $1.7 \text{ nm}^3/\text{kDa}$ for proteins has been empirically obtained (Petoukhov *et al.* 2012), possibly reflecting the level of hydration and a typical amount of flexibility in a globular protein. Deviations from theoretical values are signs of non-globularity, and flexible regions take up a proportionally large volume. An analysis of these parameters further confirms that the two lobes are globular, while the addition of the linker and C-terminal tail results in non-globular structures. Full-length ARC is more globular than the truncated constructs with disordered regions, but its volume/mass ratio suggests that the flexible regions, indeed, do take up a large volume in solution also in the context of the full-length protein.

Taken together, experimentally derived 3D models of full-length monomeric ARC allow relative positioning of all domains of ARC, showing that the N- and C-terminal halves

of the protein are both elongated and interact with each other along the long axis of full-length ARC. Moreover, full-length ARC is less flexible than expected, and the data provide no evidence for an open conformation.

Corroboration of the structural model of ARC in live neurons

In order to validate the structural model of full-length ARC, we carried out FRET experiments in live neurons with constructs carrying donor and acceptor fluorophores at the N and C termini, respectively (Fig. 5 and Figure S6). FRET efficiency reflects the distance between the pair of fluorophores, provided that the relative orientation of the fluorophores is constant and the distance is less than $\sim 10 \text{ nm}$, which is close to the maximum dimension determined for full-length ARC by SAXS (Table 1).

Two FRET methods were applied to ARC fragments expressed in cultured rat hippocampal slices (Fig. 5a and b). Ratiometric FRET (Fig. 5c and d) estimates FRET efficiency nonlinearly from fluorescence intensity, using an mTurquoise2 (mT2)-YPet pair. FLIM-FRET (Fig. 5e and f) provides true FRET efficiency based on the reduced fluorescence lifetime of the donor in the presence of the acceptor, using an EGFP-mCherry (mCh) pair. Both methods show a weak interaction between fluorophores fused to opposite ends of full-length ARC or the CTD (ARC₂₀₈₋₃₉₆), suggesting a distance approaching $\sim 10 \text{ nm}$. The NTD (ARC₁₋₁₄₀), on the other hand, shows FRET similar to the

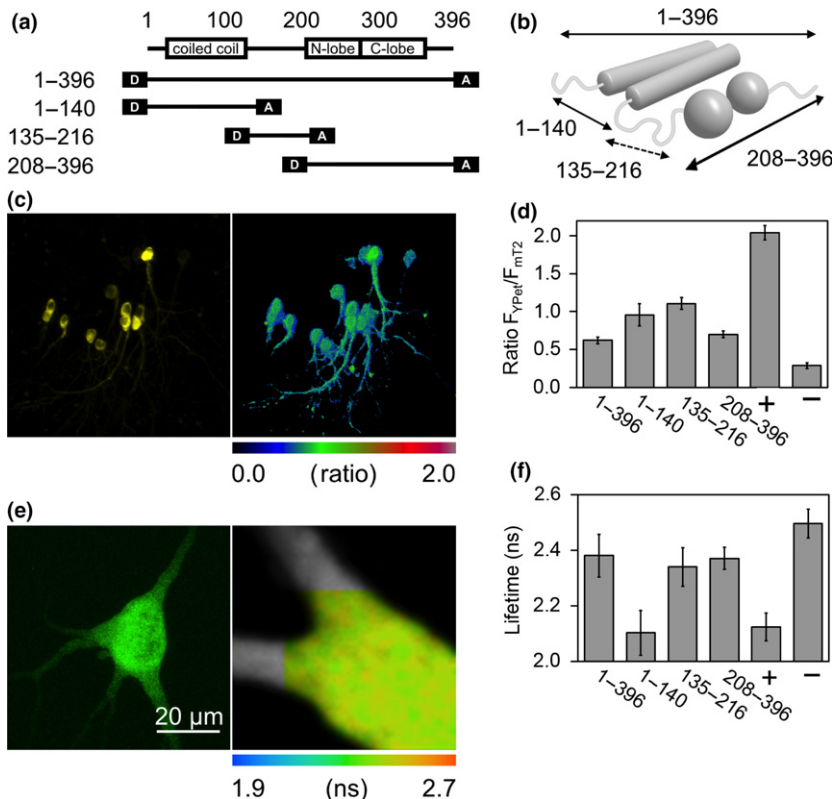


Fig. 5 Förster resonance energy transfer (FRET) analysis of activity-regulated cytoskeleton-associated protein (ARC) constructs in hippocampal slices. (a) The constructs used in FRET experiments. D, donor; A, acceptor. (b) A schematic model of ARC, showing the distances measured with FRET. (c) Representative fluorescence (left) and ratiometric FRET (right) images of neurons expressing full-length ARC. (d) Quantification of the ratiometric data (mean \pm SD). +, positive control; -, negative control. A high ratio is a sign of a short distance. Number of cells analyzed (n), from left to right: $n = 50, 20, 170, 3, 21, 17$. (e) Representative fluorescence (left) and fluorescence lifetime imaging (FLIM)-FRET (right) images of a neuron expressing full-length ARC. (f) Quantitative analysis of FLIM-FRET data (mean \pm SD). A short lifetime is a sign of a short distance. Number of cells analyzed (n), from left to right: $n = 41, 12, 48, 59, 23, 22$.

positive control, indicating a short distance between the termini. These results corroborate the SAXS-based model of ARC, including a large distance between the N and C termini in full-length ARC and in the CTD, but a short distance between termini of the NTD (Fig. 5b). The latter observation fits well with an antiparallel coiled-coil NTD structure formed of two long helices.

The central linker (ARC₁₃₅₋₂₁₆) alone gives rise to high ratiometric FRET efficiency between YPet and mT2, but between EGFP and mCh, as quantified by FLIM, the efficiency is not significantly different from full-length and ARC₂₀₈₋₃₉₆. This disparity is likely caused by the flexible nature of the linker in isolation (Myrum *et al.* 2015), coupled with the ability of YPet to form heterodimers with other members of the GFP family (Ohashi *et al.* 2007; Kolossov *et al.* 2011) – providing, in effect, a measure of flexibility rather than distance. In line with this observation, a recent study combining SAXS and FRET showed decoupling of R_g and end-to-end distance in intrinsically disordered molecules (Fuertes *et al.* 2017). Taken together, our FRET experiments validate the experimentally derived 3D molecular structure of full-length ARC.

Selective ligand peptide binding does not cause large-scale conformational changes

The binding of ligand proteins to the ARC N-lobe in the CTD, next to the linker region, could affect ARC overall conformation and trigger altered functional properties. Peptides known to bind the isolated N-lobe (Zhang *et al.* 2015) (Fig. 6a) indeed caused changes in ARC X-ray scattering profiles at low q (Fig. 6b), but this could not be linked to clear structural rearrangements (Figure 6C,S7A). When peptides were mixed with ARC constructs and analyzed with SAXS, R_g of ARC did not change. However, an increase was seen in forward scattering $I(0)$ for all constructs containing the N-lobe (Fig. 6b and c). This increase in molecular mass of the particle is an indication of complex formation between the N-lobe and the peptides.

The N- and C-lobe are structurally homologous (Fig. 6d); only the N-lobe has been previously tested for peptide binding. The peptide-induced increase in SAXS forward scattering (Fig. 6c) occurred for constructs containing the N-lobe (ARC₁₃₁₋₃₉₆, ARC₁₇₀₋₃₉₆, ARC₂₀₆₋₃₉₆, ARC₂₀₆₋₃₆₁ and ARC₂₀₇₋₂₇₇), but not for the C-lobe alone (ARC₂₇₇₋₃₇₀), showing that the peptides specifically interact with the N-lobe. The peptide-binding pocket of the N-lobe is not conserved in the C-lobe (Fig. 6d), and the function of the structurally conserved C-lobe remains enigmatic.

The increase in $I(0)$ was dependent on the peptide used (Fig. 6b and c) and similar responses from all constructs were obtained. Regions outside the ARC N-lobe do not appear to affect peptide binding; this is likely not the case for full-length protein ligands. In SAXS experiments, the

stargazin peptide had the strongest effect, followed by GKAP and WAVE1. The GluN2A peptide did not affect forward scattering. These experiments could not be done with full-length ARC, as it requires an SEC-SAXS setup, which is not compatible with rather low-affinity peptide complexes.

SRCD is a sensitive method to detect complex formation even in the absence of apparent secondary structure changes (Cowieson *et al.* 2008). We used SRCD spectroscopy to assess the effect of peptide binding on ARC. After subtraction of the signal of the free peptide, subtle differences can be observed in the spectra (Fig. 6e, S7B). Difference spectra (Figure 6e, S7B) show a maximum near 190 nm and a minimum near 200 nm for all peptides except GluN2A, indicative of a similar binding event by the WAVE1, GKAP, and stargazin peptides. In the crystal structure (Zhang *et al.* 2015), the stargazin peptide forms a short β strand between two β strands of the N-lobe (Fig. 6d). It is unlikely that the N-lobe can form a stable β sheet in the absence of a bound ligand peptide, and a change in secondary structure should accompany peptide binding. This change should reflect the loss of random coil structure and the gain of β structure. The observed difference spectra reflect overall changes in the protein and peptide components upon complex formation, and while they are reproducible between peptides and experiments, they cannot be used to pinpoint exactly the kind of change occurring.

SAXS and SRCD indicate that the WAVE1, GKAP, and stargazin peptides form a complex with the ARC N-lobe, while the GluN2A peptide does not. ITC was further used to obtain binding affinities (Table 2, Fig. 6f). The peptide with the highest affinity was stargazin; binding was also observed for WAVE1 and GKAP. The affinity of the GluN2A peptide was too low to be measured.

Binding determinants of ARC ligand peptides

To evaluate differences between ARC ligand peptides, we modeled them based on the stargazin peptide complex (Figure S7C). All peptides can be placed in the binding pocket without severe steric clashes. Most hydrogen bonds between the N-lobe and the stargazin peptide are formed between mainchain groups, indicating that peptide binding may have low sequence specificity. The proposed binding motif is Px(Y/F/H) (Zhang *et al.* 2015), but upon close inspection of the crystal structure, the binding pocket in fact covers eight residues of the peptide (Figure S7C). The sidechain interactions are clearly related to peptide affinity.

The first residue of the stargazin peptide is an arginine, which is also present in WAVE1 and GluN2A. It is bound between Asp210 and Gln212 (Figure S7C), suggesting that a large polar residue is favored at this position. The second residue in stargazin is an isoleucine, located in a more hydrophobic environment; the other peptides have small

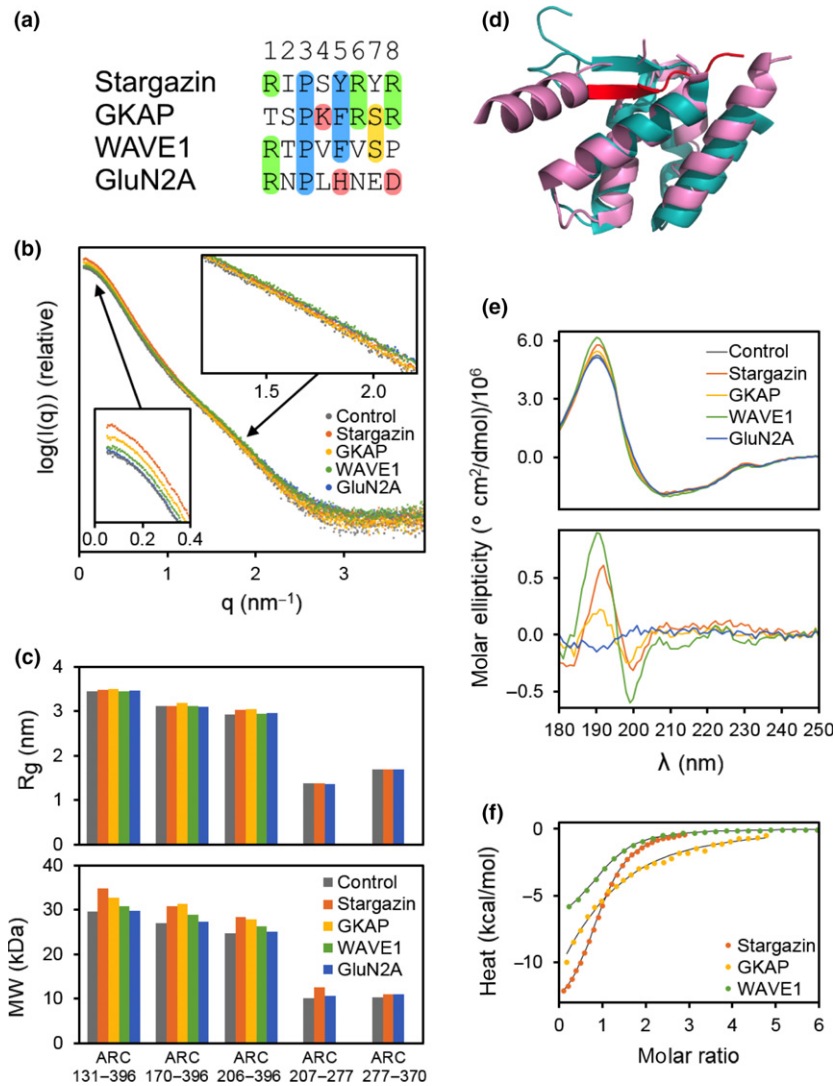


Fig. 6 Peptide binding by activity-regulated cytoskeleton-associated protein (ARC). (a) Alignment of the ligand peptide sequences. The consensus motif is coloured in blue, Arg residues in green, and other residues discussed in the text in pink and yellow. (b) SAXS scattering curves for ARC₁₃₁₋₃₉₆ with and without peptides. Similar effects were seen with all constructs containing the N-lobe. (c) Structural properties of truncated forms of ARC measured with SAXS before and after the addition of peptides. Top, R_g obtained through Guinier approximation; bottom, molecular weight of the particle calculated using $I(0)$. Note that the normalization of data for $I(0)$ calculation at 1 mg/ml was done using the protein concentration. As the peptide mass was not taken into

account in this correction, the observed increase in mass may seem higher than expected. Hence, the result should be considered qualitative (binding/no binding). (d) Superposition of the N- (cyan, PDB entry 4X3H) and C-lobe (pink, PDB entry 4X3X). The stargazin peptide bound to the N-lobe is shown in red. (e) synchrotron radiation circular dichroism spectroscopy (SRCD) spectra of ARC₂₀₇₋₂₇₇ with and without peptides (top) and difference spectra, in which the spectrum of ARC₂₀₇₋₂₇₇ has been subtracted (bottom). The spectra given by the peptides alone have been subtracted. The peptide:protein molar ratio was 4 : 1. Plots at a ratio of 1 : 1 show the same features (Figure S7B). (f) Calorimetric titration of the N-lobe with the 3 ligand peptides.

polar residues in this position. Proline at the third position is the only residue present in all four peptides (Fig. 6a); it lies in a hydrophobic pocket, and its carbonyl group makes a hydrogen bond to the side chain of His245. This interaction, reminiscent of proteins binding to Pro-rich sequences (Kursula *et al.* 2008), is likely to be a major determinant for peptide binding, and the conserved rigid proline residue

defines local backbone conformation in the peptide. The fourth residue in the stargazin peptide is a solvent-exposed serine, unlikely to play a role in specificity. The aromatic residue at position five in stargazin, GKAP, and WAVE1 is in a hydrophobic pocket, making stacking interactions. GluN2A has a histidine at this position, which may explain the observed lack of binding. Arginine residues located at

position six and eight of stargazin and GKAP are both near Glu215, allowing salt bridge formation. In the case of GluN2A, this region is instead negatively charged. The stargazin peptide has a tyrosine at position 7, which can stack against Tyr274. These differences explain both the higher affinity of stargazin towards ARC as well as the lack of GluN2A binding.

The N-terminal domain of ARC binds to phospholipid membranes

ARC may interact with polyanionic surfaces, such as phospholipid membranes, RNA, or microfilaments, especially due to the high positive charge of the NTD. A recent study showed lipid membrane binding and palmitoylation of ARC (Barylko *et al.* 2017). To investigate, which domain of ARC has affinity towards membranes, *in vitro* co-sedimentation assays were performed. Full-length ARC co-sedimented with DOPC:DOPS liposomes (Fig. 7a), showing that ARC binds to negatively charged lipid membranes. The NTD was required for ARC binding to liposomes. Co-sedimentation of full-length ARC was prevented by phosphate (Fig. 7a), which may compete with protein binding to phospholipids. High salt concentration also prevented co-sedimentation. The results suggest that the interaction between ARC and the membrane is mediated by the positively charged NTD binding to phospholipid headgroups.

The testing of different phospholipid headgroups showed that co-sedimentation did not occur with phosphatidylcholine alone, but the presence of an equimolar amount of DOPS, DOPG, or DOPE promoted binding (Fig. 7b). Considering the chemical difference between DOPC and DOPE, it thus appears that the methyl groups on phosphatidylcholine, which make its headgroup less polar than that of phosphatidylethanolamine and allows less hydrogen bonding, prevents ARC binding to the membrane surface. To observe pH dependence of ARC–membrane interactions in this system with purified components, the proteolipid pellet was extracted at neutral and high pH. Only high-pH buffer resolubilized ARC from the lipid pellet (Fig. 7b), suggesting that, during ARC extraction from recombinant sources or tissue (see above), the membrane interaction must be broken to make ARC soluble. CD spectroscopy indicated that ARC secondary structure content is not affected by the presence of liposomes (Figure S2). Taken together, ARC binds to phospholipid membranes, and this binding is affected by the lipid composition as well as the buffer conditions.

Discussion

Understanding the molecular function of ARC in different scenarios related to synaptic plasticity, modulation of the protein networks in the PSD, and the formation of virus-like capsids requires detailed knowledge on its 3D structure,

interactions, and flexibility. The availability of pure, monomeric, full-length human ARC allows new levels of structural and biophysical characterization.

Structure of full-length ARC

Full-length ARC is folded and compact, with the helical N-terminal domain located on top of the C-terminal lobes, which are separate entities (Fig. 4e). Importantly, no extended linker was observed between the NTD and CTD, and a closed conformation is supported by both X-ray and dynamic light scattering analyses. The N- and C-terminal domains are both elongated, and our data indicate contact between them, although higher-resolution methods will be needed to discern the actual orientations of the domain surfaces with respect to one another.

The structural properties of the model derived from SAXS and SRCD are further supported by intramolecular FRET analysis of ARC expressed in live neurons in cultured hippocampal tissue slices (Fig. 5). Both ratiometric and FLIM-FRET data are compatible with a compact arrangement of full-length ARC, in which the CTD and NTD are both elongated. In the isolated CTD, the N- and C-termini lie at opposite ends, whereas the termini of the NTD are in close vicinity, consistent with an antiparallel coiled-coil structure of the NTD. This structural arrangement of the NTD could be critical for interactions of ARC with the polyanionic surfaces of membranes, the actin cytoskeleton, or RNA.

The compact full-length ARC deviates from globularity, as evidenced by the R_g/R_g^* ratio, and it has significant flexible segments, highlighted by the large molecular volume/mass ratio (Table 1). While the NTD and CTD remain in close contact, the linker and the termini of ARC are flexible, providing an apparent increase in molecular volume in *ab initio* modelling. While ARC has been hypothesized to undergo conformational changes and different states of oligomerization during its functional cycle, it is clear that factors not present in our experiments will be required to trigger conformational changes. Thus, it remains to be seen whether binding events or post-translational modifications can cause opening of the compact structure of ARC.

Oligomeric status and membrane interactions

ARC has been proposed to alternate between monomeric and oligomeric/aggregated forms (Byers *et al.* 2015; Myrum *et al.* 2015). Recently, ARC was shown to build virus-like capsid structures (Ashley *et al.* 2018; Pastuzyn *et al.* 2018), which are packed into membrane vesicles. Here, we have shown that aggregation tendency and insolubility of recombinant ARC are mediated by the N-terminal domain alone. Upon truncation of this domain, solubility was high, and monodisperse monomers were observed. The C-terminal half of ARC has no propensity to oligomerize on its own. It is possible that the oligomerization property of the ARC NTD plays a role in capsid-like structure formation.

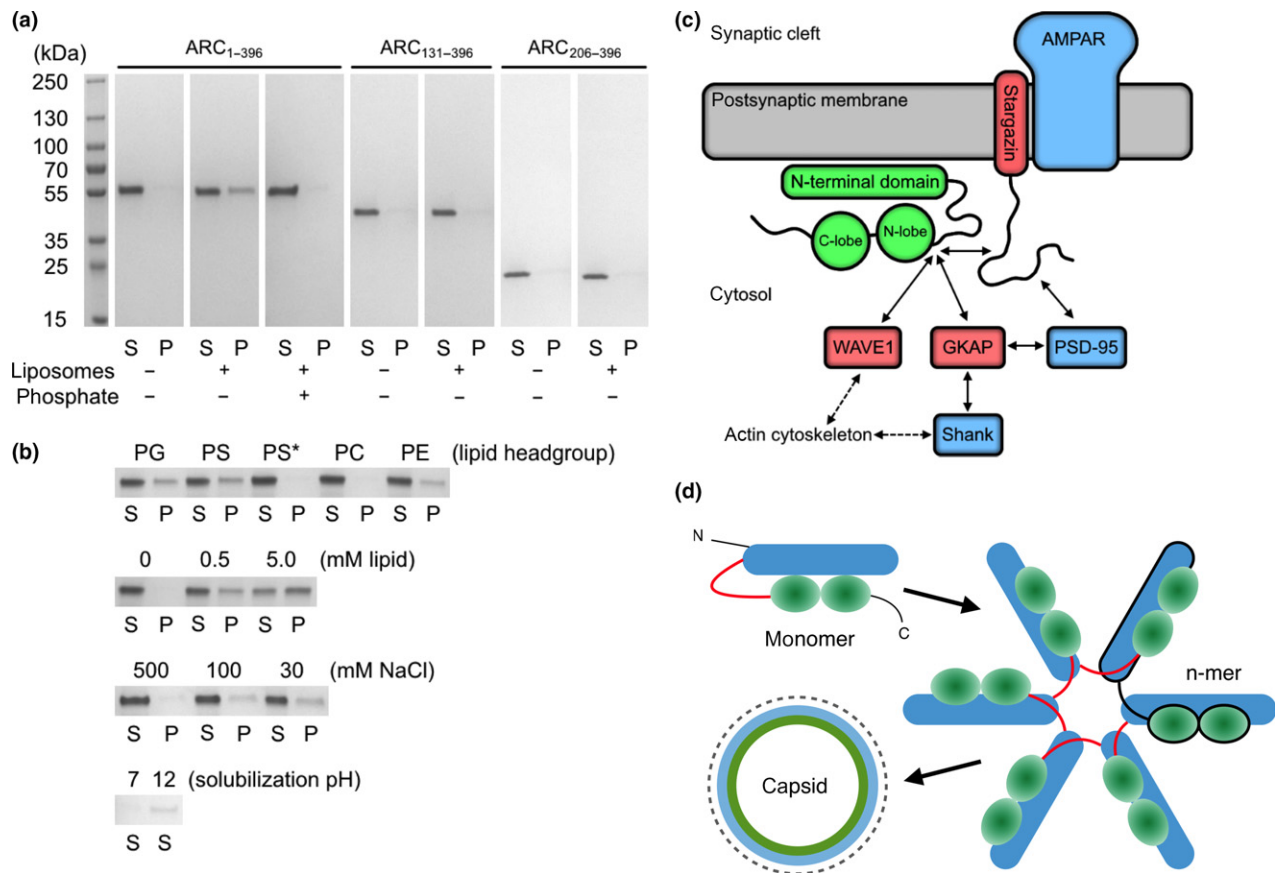


Fig. 7 Activity-regulated cytoskeleton-associated protein (ARC) association with lipid membranes. (a) Co-sedimentation of ARC with liposomes. (b) Effects of lipid headgroup, lipid and salt concentration, and pH on ARC co-sedimentation. (c) A model for the localization of ARC. ARC may bind the cytosolic side of the postsynaptic membrane through its NTD; the N-lobe will be in close vicinity to the cytoplasmic tail of stargazin, allowing an indirect connection to the α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor. Other interactions are likely to compete with this interaction, eventually making ARC a central player in the postsynaptic density (PSD) protein network.

One prominent feature of the NTD is its high isoelectric point of almost 10, something seen for many peripheral membrane proteins (Han *et al.* 2013). Native ARC extracted from brain tissue appears insoluble after homogenization, being assumed to bind to a larger component (Lyford *et al.* 1995). Low solubility may be caused by membrane binding by the N-terminal domain, which could start to aggregate at high protein concentrations. When the NTD forms a helical structure, this will result in an amphipathic arrangement, whereby 24 of the 25 positively charged residues may be located on the same face (Figure S3A,E); the opposite side would be hydrophobic. The partial disruption or neutralization of this polarized helical structure, for example, by high pH, could be required to solubilize ARC from membrane surfaces.

Phosphate prevented ARC from co-sedimenting with membranes. Interestingly, the SPX domain, a remote

homologue of the ARC NTD, functions as a signalling molecule that binds phosphate and inositol polyphosphates (Wild *et al.* 2016). Furthermore, capsid formation by ARC was shown to be promoted by phosphate (Pastuzyn *et al.* 2018). Phosphate clearly affects the functional properties of ARC, and likely the positive surface of the NTD plays a role.

Table 2 ARC-peptide-binding parameters obtained with ITC

Peptide	K_d (μM)	ΔH (kcal/mol)	$-\Delta S$ (kcal/mol)
Stargazin	33.3 ± 0.8	-14.2 ± 0.04	+8.1
GKAP	248 ± 30	-21.4 ± 0.24	+16.4
WAVE1	45.7 ± 1.4	-7.4 ± 0.04	+1.4
GluN2A	N/A	N/A	N/A

homologue of the ARC NTD, functions as a signalling molecule that binds phosphate and inositol polyphosphates (Wild *et al.* 2016). Furthermore, capsid formation by ARC was shown to be promoted by phosphate (Pastuzyn *et al.* 2018). Phosphate clearly affects the functional properties of ARC, and likely the positive surface of the NTD plays a role.

In addition, co-sedimentation was prevented by high salt concentration, and ARC could be re-solubilized from the lipid pellet at high pH (Fig. 7b). All the results, including the effects of lipid composition on binding, point towards the importance of polar interactions between ARC and phospholipid headgroups. In a recent paper, recombinant ARC was shown to bind liposomes, and a weak interaction was also observed with phosphatidylcholine alone (Barylko *et al.* 2017). It must be pointed out that the recombinant ARC used in the latter experiments had not been purified by SEC, and therefore, it is likely to have consisted of mostly soluble high-molecular-weight aggregates. In the future, it will be important to determine how phospholipid headgroups affect the formation of ARC-containing endosomes or capsid structures.

The size, shape, membrane-binding capability, and presence of long helices in the ARC NTD are all features of BAR (Bin/Amphiphysin/Rvs) domains, which bind to and destabilize the lipid membrane to promote curvature, using positively charged residues (Zimmerberg and McLaughlin 2004). This activity is often involved in endocytosis, and the N-terminal domain of ARC could participate in endocytosis by introducing curvature stress into the membrane surrounding AMPA receptors. In addition, ARC might be involved in recruiting other proteins, such as GKAP and WAVE1, close to membranes (Fig. 7c).

Interaction with ligand proteins

The peptide interaction site (Zhang *et al.* 2015) is located in the globular N-lobe, adjacent to the linker region. Binding to ligand peptides could have structural effects on ARC. From SAXS measurements, it was clear that a larger complex was formed after the addition of binding peptides, and a small structural effect could be seen with SRCD. However, SAXS did not reveal restructuring of the linker segment or conformational changes upon peptide binding. When binding a full-length interaction partner, more of the linker region and other domains of ARC are likely to be involved in target protein binding, possibly increasing binding affinity.

Our binding experiments were made with peptides of 8 residues, chosen to match the entire peptide-binding site in the crystal structure (Zhang *et al.* 2015). In contrast to stargazin, WAVE1, and GKAP, binding of the GluN2A peptide to the N-lobe could not be observed. This is likely to be caused by poor conservation of the binding motif. The low affinity for this peptide does not rule out an interaction between ARC and GluN2A. The cytoplasmic domain of GluN2A contains several regions with a PxY/F motif, more similar to the stargazin peptide sequence than the sequence suggested earlier to bind ARC (Zhang *et al.* 2015) and analyzed here.

ARC must be located close to the postsynaptic membrane to interact with the short cytoplasmic tail of stargazin (Fig. 7c). Stargazin regulates AMPA receptor trafficking, mobility, and channel properties. Taking into account the binding of multiple stargazin units to the AMPA receptor

(Zhao *et al.* 2016), it is possible that ARC surrounds the AMPA receptor, while attached to the membrane through the NTD (Fig. 7c). The C terminus of stargazin binds to PDZ (PSD-95/Discs large/Zona occludens-1) domains from several PSD proteins (Chen *et al.* 2000). Therefore, binding of ARC to stargazin might alter stargazin interactions within the PSD scaffold, with impacts on AMPA receptor behavior. ARC interaction with GKAP is highly relevant in this respect. GKAP is an important hub in the PSD protein scaffold, which will compete with stargazin for ARC binding. GKAP also binds to PSD-95 (Kim *et al.* 1997), and the GKAP C terminus is bound by Shank proteins further down in the PSD assembly (Naisbitt *et al.* 1999; Ponna *et al.* 2018). Thus, ARC binding to the postsynaptic membrane and its ligand proteins may have complex effects on the molecular composition and signalling properties of the PSD.

Implications of ARC structure for its functional modalities

It is likely that different functions of ARC are related to its various forms and environments, including free monomeric ARC, ARC associated with membranes, endosomes, actin filaments, or nuclear bodies, as well as virus-like ARC capsids. The connections between various ARC-binding partners, its post-translational modifications (phosphorylation, SUMOylation, and palmitoylation), membrane binding, and ARC assembly into ordered homo-oligomers, including capsids, are to a large extent still enigmatic.

The high-resolution structure of the ARC-based capsids (Ashley *et al.* 2018; Pastuzyn *et al.* 2018) is not known, including their internal symmetry and details of molecular organization. In the HIV capsid, the Gag capsid protein CA forms both penta- and hexameric assemblies. In both cases, the CA-NTD forms the core of the structure (Pornillos *et al.* 2009; Zhao *et al.* 2013), while the CTD mediates additional contacts to monomeric NTD domains in the CA oligomer (Figure S8). The CA-CTD also keeps the penta- and hexamers together through homodimerization (Zhao *et al.* 2013). It is therefore surprising and noteworthy that the ARC N and C lobes, which are both homologous to the CA-CTD, show no signs of homodimerization. In contrast, we have shown that homo-oligomerization of ARC depends on the N-terminal region of the protein.

The flexible linker between the N- and C-terminal domains in CA is very short. While we observed the ARC NTD and CTD in contact with each other in the case of monomeric ARC in solution, it is perfectly conceivable that the CTD of ARC could similarly contact neighbouring ARC monomer NTDs within the ARC capsid structure (Fig. 7d). The long flexible linker of ARC would give much conformational freedom in this assembly, possibly allowing domain swapping upon oligomer formation. In line with this hypothesis, the NTD and CTD of ARC were both shown to be required for proper capsid assembly (Pastuzyn *et al.* 2018). We hypothesize that in the ARC capsid-like structures, the NTD

forms oligomeric cores linked together through the lobe structures in the CTD. The importance of ARC having two lobe domains homologous to the CA-CTD is currently unknown.

In HIV capsids, the matrix protein (MA) binds to lipid membranes. MA has properties similar to the ARC NTD, including high helical content, positive charge, myristoylation, and phosphoinositide as well as membrane binding (Saad *et al.* 2006). The importance of the ARC NTD in membrane binding and oligomeric assembly suggests a dual functional role analogous to both the Gag MA and the CA protein. The interplay between membrane interactions, protein ligand binding to the ARC N-lobe, and capsid formation is an important outstanding issue. It seems unlikely that ligand proteins would specifically get packed into capsids as cargo. However, the presence of ARC ligand proteins and/or a lipid membrane may affect or regulate capsid formation.

Acknowledgments and conflict of interest disclosure

We acknowledge access to and support at SAXS and SRCD beamlines at EMBL/DESY, ESRF, SOLEIL, DIAMOND, and ASTRID2. The project was supported by the Research Council of Norway (TOPPFORSK grant to C.R.B.). The authors declare no conflicts of interest.

All experiments were conducted in compliance with the ARRIVE guidelines.

Institutional approval

Institutional approval was not required for this study; the experiments were approved at the national level.

Open Science Badges

This article has received a badge for *Open Materials* because it provided all relevant information to reproduce the study in the manuscript. The complete Open Science Disclosure form for this article can be found at the end of the article. More information about the Open Practices badges can be found at <https://cos.io/our-services/open-science-badges/>.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. SEC analysis of truncated ARC constructs shows single peaks.

Figure S2. Additional CD spectra.

Figure S3. The NTD is predicted to form a coiled-coil structure.

Figure S4. Additional SAXS modeling.

Figure S5. R_g as a function of time in molecular dynamics

simulations of the ARC N-lobe with (blue) and without (red) the stargazin peptide.

Figure S6. FRET experiments.

Figure S7. ARC N-lobe – peptide interactions.

Figure S8. Comparison of the ARC lobes to the HIV capsid protein.

Table S1. ARC constructs analyzed in this work.

References

- Abraham M. J., Murtola T., Schulz R., Páll S., Smith J. C., Hess B. and Lindahl E. (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25.
- Alfadhli A., Steel E., Finlay L., Bächinger H. P. and Barklis E. (2002) Hantavirus nucleocapsid protein coiled-coil domains. *J. Biol. Chem.* **277**, 27103–27108.
- Ashley J., Cordy B., Lucia D., Fradkin L. G., Budnik V. and Thomson T. (2018) Retrovirus-like gag protein Arc1 Binds RNA and traffics across synaptic boutons. *Cell* **172**, e11.
- Barylko B., Wilkerson J. R., Cavalier S. H., Binns D. D., James N. G., Jameson D. M., Huber K. M. and Albanesi J. P. (2017) Palmitoylation and Membrane Binding of Arc/Arg3.1: A Potential Role in Synaptic Depression. *Biochemistry*
- van den Berg S., Löfdahl P. A., Härd T. and Berglund H. (2006) Improved solubility of TEV protease by directed evolution. *J. Biotechnol.* **121**, 291–298.
- Bernadó 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.
- Blanchet C. E., Spilotros A., Schwemmer F. *et al.* (2015) Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J. Appl. Crystallogr.* **48**, 431–443.
- Bogomolovas J., Simon B., Sattler M. and Stier G. (2009) Screening of fusion partners for high yield expression and purification of bioactive viscotoxins. *Protein Expr. Purif.* **64**, 16–23.
- Boratyn G. M., Schäffer A. A., Agarwala R., Altschul S. F., Lipman D. J. and Madden T. L. (2012) Domain enhanced lookup time accelerated BLAST. *Biol Direct* **7**, 12.
- Byers C. E., Barylko B., Ross J. A. *et al.* (2015) Enhancement of dynamin polymerization and GTPase activity by Arc/Arg3.1. *Biochim. Biophys. Acta* **1850**, 1310–1318.
- Chen L., Chetkovich D. M., Petralia R. S., Sweeney N. T., Kawasaki Y., Wenthold R. J., Brecht D. S. and Nicoll R. A. (2000) Stargazin regulates synaptic targeting of AMPA receptors by two distinct mechanisms. *Nature* **408**, 936–943.
- Cowieson N. P., Miles A. J., Robin G., Forwood J. K., Kobe B., Martin J. L. and Wallace B. A. (2008) Evaluating protein:protein complex formation using synchrotron radiation circular dichroism spectroscopy. *Proteins* **70**, 1142–1146.
- DaSilva L. L., Wall M. J., P de Almeida L., Wauters S. C., Januário Y. C., Müller J. and Corrêa S. A. (2016) Activity-regulated cytoskeleton-associated protein controls AMPAR endocytosis through a direct interaction with clathrin-adaptor protein 2. *eNeuro* **3**, e0144-15.2016.
- De Simoni A. and Yu L. M. (2006) Preparation of organotypic hippocampal slice cultures: interface method. *Nat. Protoc.* **1**, 1439–1445.
- Erickson H. P. (2009) Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biol. Proced Online* **11**, 32–51.

- Fernández E., Collins M. O., Frank R. A. W. *et al.* (2017) Arc requires PSD95 for assembly into postsynaptic complexes involved with neural dysfunction and intelligence. *Cell Rep.* **21**, 679–691.
- Franke D., Petoukhov M. V., Konarev P. V. *et al.* (2017) ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212–1225.
- Fuertes G., Banterle N., Ruff K. M. *et al.* (2017) Decoupling of size and shape fluctuations in heteropolymeric sequences reconciles discrepancies in SAXS vs. FRET measurements. *Proc. Natl Acad. Sci. USA* **114**, E6342–E6351.
- Goedhart J., von Stetten D., Noirclerc-Savoie M., Lelimosin M., Joosen L., Hink M. A., van Weeren L., Gadella T. W. and Royant A. (2012) Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. *Nat. Commun.* **3**, 751.
- Gogolla N., Galimberti L., DePaola V. and Caroni P. (2006) Preparation of organotypic hippocampal slice cultures for long-term live imaging. *Nat. Protoc.* **1**, 1165–1171.
- Greenfield N. J. (2006) Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876–2890.
- Haas K., Sin W. C., Javaherian A., Li Z. and Cline H. T. (2001) Single-cell electroporation for gene transfer in vivo. *Neuron* **29**, 583–591.
- Haas K., Jensen K., Sin W. C., Foa L. and Cline H. T. (2002) Targeted electroporation in *Xenopus* tadpoles in vivo—from single cells to the entire brain. *Differentiation* **70**, 148–154.
- Hammarström M., Woestenenk E. A., Hellgren N., Härd T. and Berglund H. (2006) Effect of N-terminal solubility enhancing fusion proteins on yield of purified target protein. *J. Struct. Funct. Genomics* **7**, 1–14.
- Han H., Myllykoski M., Ruskamo S., Wang C. and Kursula P. (2013) Myelin-specific proteins: a structurally diverse group of membrane-interacting molecules. *BioFactors* **39**, 233–241.
- Johnson W. C. (1999) Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins* **35**, 307–312.
- Källberg M., Wang H., Wang S., Peng J., Wang Z., Lu H. and Xu J. (2012) Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* **7**, 1511–1522.
- Kammerer R. A., Schulthess T., Landwehr R., Lustig A., Engel J., Aebi U. and Steinmetz M. O. (1998) An autonomous folding unit mediates the assembly of two-stranded coiled coils. *Proc. Natl Acad. Sci. USA* **95**, 13419–13424.
- Kim E., Naisbitt S., Hsueh Y. P., Rao A., Rothschild A., Craig A. M. and Sheng M. (1997) GKAP, a novel synaptic protein that interacts with the guanylate kinase-like domain of the PSD-95/SAP90 family of channel clustering molecules. *J. Cell Biol.* **136**, 669–678.
- Kim D. E., Chivian D. and Baker D. (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* **32**, W526–W531.
- Kolossov V. L., Spring B. Q., Clegg R. M., Henry J. J., Sokolowski A., Kenis P. J. and Gaskins H. R. (2011) Development of a high-dynamic range, GFP-based FRET probe sensitive to oxidative microenvironments. *Exp. Biol. Med. (Maywood)* **236**, 681–691.
- Korb E., Wilkinson C. L., Delgado R. N., Lovero K. L. and Finkbeiner S. (2013) Arc in the nucleus regulates PML-dependent GluA1 transcription and homeostatic plasticity. *Nat. Neurosci.* **16**, 874–883.
- Kursula P., Kursula I., Massimi M., Song Y. H., Downer J., Stanley W. A., Witke W. and Wilmanns M. (2008) High-resolution structural analysis of mammalian profilin 2a complex formation with two physiological ligands: the formin homology 1 domain of mDial1 and the proline-rich domain of VASP. *J. Mol. Biol.* **375**, 270–290.
- Lees J. G., Miles A. J., Wien F. and Wallace B. A. (2006) A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics* **22**, 1955–1962.
- Linkert M., Rueden C. T., Allan C. *et al.* (2010) Metadata matters: access to image data in the real world. *J. Cell Biol.* **189**, 777–782.
- Lyford G. L., Yamagata K., Kaufmann W. E., Barnes C. A., Sanders L. K., Copeland N. G., Gilbert D. J., Jenkins N. A., Lanahan A. A. and Worley P. F. (1995) Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-associated protein that is enriched in neuronal dendrites. *Neuron* **14**, 433–445.
- Messaoudi E., Kanhema T., Soulé J., Tiron A., Dageyte G., da Silva B. and Bramham C. R. (2007) Sustained Arc/Arg3.1 synthesis controls long-term potentiation consolidation through regulation of local actin polymerization in the dentate gyrus in vivo. *J. Neurosci.* **27**, 10445–10455.
- Myrum C., Baumann A., Bustad H. J. *et al.* (2015) Arc is a flexible modular protein capable of reversible self-oligomerization. *Biochem. J.* **468**, 145–158.
- Myrum C., Nikolaienko O., Bramham C. R., Haavik J. and Zayats T. (2017) Implication of the APP Gene in Intellectual Abilities. *J. Alzheimers Dis.* **59**, 723–735.
- Nair R. R., Patil S., Tiron A., Kanhema T., Panja D., Schiro L., Parobczak K., Wilczynski G. and Bramham C. R. (2017) Dynamic Arc SUMOylation and selective interaction with F-actin-binding protein drebrin A in LTP consolidation in vivo. *Front Synaptic Neurosci.* **9**, 8.
- Naisbitt S., Kim E., Tu J. C., Xiao B., Sala C., Valtschanoff J., Weinberg R. J., Worley P. F. and Sheng M. (1999) Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and cortactin. *Neuron* **23**, 569–582.
- Nezami A., Poy F., Toms A., Zheng W. and Eck M. J. (2010) Crystal structure of a complex between amino and carboxy terminal fragments of mDial1: insights into autoinhibition of diaphanous-related formins. *PLoS ONE* **5**, e12992.
- Nguyen A. W. and Daugherty P. S. (2005) Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nat. Biotechnol.* **23**, 355–360.
- Nikolaienko O., Eriksen M. S., Patil S., Bito H. and Bramham C. R. (2017) Stimulus-evoked ERK-dependent phosphorylation of activity-regulated cytoskeleton-associated protein (Arc) regulates its neuronal subcellular localization. *Neuroscience* **360**, 68–80.
- Nikolaienko O., Patil S., Eriksen M. S. and Bramham C. R. (2018) Arc protein: a flexible hub for synaptic plasticity and cognition. *Semin. Cell Dev. Biol.* **77**, 33–42.
- O'Brien J. A. and Lummis S. C. (2006) Biolistic transfection of neuronal cultures using a hand-held gene gun. *Nat. Protoc.* **1**, 977–981.
- Ohashi T., Galiacy S. D., Briscoe G. and Erickson H. P. (2007) An experimental study of GFP-based FRET, with application to intrinsically unstructured proteins. *Protein Sci.* **16**, 1429–1438.
- Oostenbrink C., Villa A., Mark A. E. and van Gunsteren W. F. (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **25**, 1656–1676.
- Otmakhov N., Tao-Cheng J. H., Carpenter S., Asrican B., Dosemeci A., Reese T. S. and Lisman J. (2004) Persistent accumulation of calcium/calmodulin-dependent protein kinase II in dendritic spines after induction of NMDA receptor-dependent chemical long-term potentiation. *J. Neurosci.* **24**, 9324–9331.
- Pastuzyn E. D., Day C. E., Kearns R. B. *et al.* (2018) The neuronal gene arc encodes a repurposed retrotransposon gag protein that mediates intercellular RNA transfer. *Cell* **172**, e18.
- Pernot P., Round A., Barrett R. *et al.* (2013) Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **20**, 660–664.
- Petoukhov M. V., Franke D., Shkumatov A. V., Tria G., Kikhney A. G., Gajda M., Gorba C., Mertens H. D., Konarev P. V. and Svergun D. I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350.

- Ponna S. K., Ruskamo S., Myllykoski M., Keller C., Boeckers T. M. and Kursula P. (2018) Structural basis for PDZ domain interactions in the post-synaptic density scaffolding protein Shank3. *J. Neurochem.* **6**, 449–463.
- Pornillos O., Ganser-Pornillos B. K., Kelly B. N., Hua Y., Whitby F. G., Stout C. D., Sundquist W. I., Hill C. P. and Yeager M. (2009) X-ray structures of the hexameric building block of the HIV capsid. *Cell* **137**, 1282–1292.
- Raasakka A., Myllykoski M., Laulumaa S., Lehtimäki M., Härtlein M., Moulin M., Kursula I. and Kursula P. (2015) Determinants of ligand binding and catalytic activity in the myelin enzyme 2',3'-cyclic nucleotide 3'-phosphodiesterase. *Sci. Rep.* **5**, 16520.
- Rose T., Jaepel J., Hübener M. and Bonhoeffer T. (2016) Cell-specific restoration of stimulus preference after monocular deprivation in the visual cortex. *Science* **352**, 1319–1322.
- Saad J. S., Miller J., Tai J., Kim A., Ghanam R. H. and Summers M. F. (2006) Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proc. Natl Acad. Sci. USA* **103**, 11364–11369.
- Schindelin J., Arganda-Carreras I., Frise E. *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682.
- Shin S. M., Zhang N., Hansen J., Gerges N. Z., Pak D. T., Sheng M. and Lee S. H. (2012) GKAP orchestrates activity-dependent postsynaptic protein remodeling and homeostatic scaling. *Nat. Neurosci.* **15**, 1655–1666.
- Steinmetz M. O., Jelesarov I., Matousek W. M., Honnappa S., Jahnke W., Missimer J. H., Frank S., Alexandrescu A. T. and Kammerer R. A. (2007) Molecular basis of coiled-coil formation. *Proc. Natl Acad. Sci. USA* **104**, 7062–7067.
- Svergun D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **76**, 2879–2886.
- Svergun D. I., Petoukhov M. V. and Koch M. H. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.* **80**, 2946–2953.
- Tria G., Mertens H. D., Kachala M. and Svergun D. I. (2015) Advanced ensemble modeling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207–217.
- Whitmore L. and Wallace B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **32**, W668–W673.
- Wild R., Gerasimaite R., Jung J. Y. *et al.* (2016) Control of eukaryotic phosphate homeostasis by inositol polyphosphate sensor domains. *Science* **352**, 986–990.
- Zhang W., Wu J., Ward M. D., Yang S., Chuang Y. A., Xiao M., Li R., Leahy D. J. and Worley P. F. (2015) Structural basis of arc binding to synaptic proteins: implications for cognitive disease. *Neuron* **86**, 490–500.
- Zhao G., Perilla J. R., Yufenyuy E. L. *et al.* (2013) Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **497**, 643–646.
- Zhao Y., Chen S., Yoshioka C., Bacongus I. and Gouaux E. (2016) Architecture of fully occupied GluA2 AMPA receptor-TARP complex elucidated by cryo-EM. *Nature* **536**, 108–111.
- Zimmerberg J. and McLaughlin S. (2004) Membrane curvature: how BAR domains bend bilayers. *Curr. Biol.* **14**, R250–R252.

Open Practices Disclosure

Manuscript Title: Structure of monomeric full-length ARC sheds light on molecular flexibility, protein interactions, and functional modalities

Corresponding Author: Petri Kursula

Articles accepted to *Journal of Neurochemistry* after 01.2018 are eligible to earn badges that recognize open scientific practices: publicly available data, material, or preregistered research plans. Please read more about the badges in our *author guidelines and Open Science Badges page*, and you can also find information on the Open Science Framework [wiki](#).

Please check this box if you are interested in participating.

To apply for one or more badges acknowledging open practices, please check the box(es) corresponding to the desired badge(s) below and provide the information requested in the relevant sections. To qualify for a badge, you must provide a URL, doi, or other permanent path for accessing the specified information in a public, open-access repository. **Qualifying public, open-access repositories are committed to preserving data, materials, and/or registered analysis plans and keeping them publicly accessible via the web in perpetuity.** Examples include the Open Science Framework ([OSF](#)) and the various Dataverse networks. Hundreds of other qualifying data/materials repositories are listed at <http://re3data.org/>. Preregistration of an analysis plan must take place via a publicly accessible registry system (e.g., [OSF](#), [ClinicalTrials.gov](#) or other trial registries in the [WHO Registry Network](#), institutional registration systems). **Personal websites and most departmental websites do not qualify as repositories.**

Authors who wish to publicly post third-party material in their data, materials, or preregistration plan must have the proper authority or permission agreement in order to do so.

There are circumstances in which it is not possible or advisable to share any or all data, materials, or a research plan publicly. For example, there are cases in which sharing participants' data could violate confidentiality. If you would like your article to include an explanation of such circumstances and/or provide links to any data or materials you have made available—even if not under conditions eligible to earn a badge—you may write an alternative note that will be published in a note in the article. Please check this box if you would like your article to include an alternative note and provide the text of the note below:

Alternative note:

Open Data Badge

1. Provide the URL, doi, or other **permanent path** for accessing the data in a **public, open-access repository**:

Confirm that there is sufficient information for an independent researcher to reproduce **all of the reported results**, including codebook if relevant.

Confirm that you have registered the uploaded files so that they are **time stamped** and cannot be age.

Open Materials Badge

1. Provide the URL, doi, or other **permanent path** for accessing the materials in a **public, open-access repository**: all relevant information is provided in the manuscript.

Confirm that there is sufficient information for an independent researcher to reproduce **all of the reported methodology**.

Confirm that you have registered the uploaded files so that they are **time stamped** and cannot be age.

Preregistered Badge

1. Provide the URL, doi, or other **permanent path** to the registration in a **public, open-access repository***:

2. Was the analysis plan registered prior to examination of the data or observing the outcomes? If no, explain.**

3. Were there additional registrations for the study other than the one reported? If yes, provide links and explain.*

*No badge will be awarded if (1) is not provided, **or** if (3) is answered “yes” without strong justification

**If the answer to (2) is “no,” the notation DE (Data Exist) will be added to the badge, indicating that registration postdates realization of the outcomes but predates analysis.

By signing below, authors affirm that the above information is accurate and complete, that any third-party material has been reproduced or otherwise made available only with the permission of the original author or copyright holder, and that publicly posted data do not contain information that would allow individuals to be identified without consent.

Date: Tornio, Finland, July 17th 2018

Name: Petri Kursula

Signature: 
