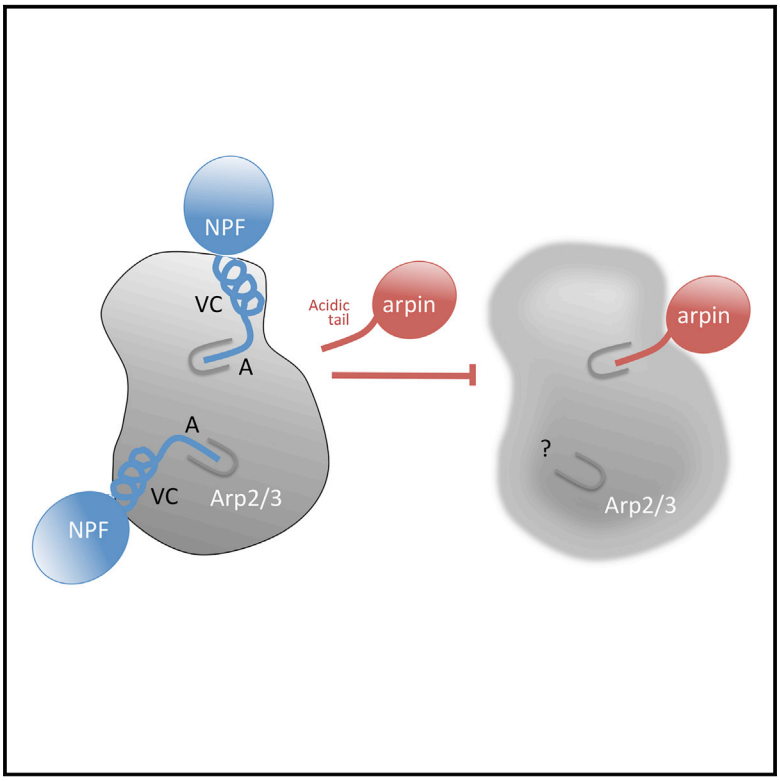
Article

Structure

Hybrid Structural Analysis of the Arp2/3 Regulator Arpin Identifies Its Acidic Tail as a Primary Binding Epitope

Graphical Abstract



Authors

Susan Fetics, Aurélien Thureau, Valérie Campanacci, ..., Alexis Gautreau, Javier Pérez, Jacqueline Cherfils

Correspondence

jacqueline.cherfils@ens-cachan.fr

In Brief

Arpin is a newly discovered regulator of actin polymerization, which steers cell migration by exerting a negative control on the Arp2/3 complex. Fetics et al. show that the C-terminal acidic tail of Arpin is a primary epitope and discuss how Arpin might use it to compete with VCA domains for binding to Arp2/3.

Highlights

- We combined X-ray crystallography, SAXS, and ITC to analyze the structure of Arpin
- The acidic C terminus of Arpin forms an elongated tail
- This acidic tail suffices to tether Arpin to interacting proteins with high affinity
- This suggests how Arpin may compete with VCA domains for binding to the Arp2/3 complex

Fetics et al., 2016, Structure *24*, 252–260
February 2, 2016
http://dx.doi.org/10.1016/j.str.2015.12.001

CellPress

CellPress

Structure
Article

Hybrid Structural Analysis of the Arp2/3 Regulator Arpin Identifies Its Acidic Tail as a Primary Binding Epitope

Susan Fetics,[1,2,6] Aurélien Thureau,[3,6] Valérie Campanacci,[2] Magali Aumont-Nicaise,[4] Irène Dang,[2,5] Alexis Gautreau,[2,5] Javier Pérez,[3] and Jacqueline Cherfils[1,2,*]

[1]Laboratoire de Pharmacologie et Biologie Appliquée, UMR 8113, CNRS-Ecole Normale Supérieure de Cachan, 61 Avenue du Président Wilson, 94235 Cachan Cedex, France

[2]Laboratoire d'Enzymologie et Biochimie Structurales, CNRS UPR3082, 91190 Gif-sur-Yvette, France

[3]Synchrotron SOLEIL, 91190 Gif-sur-Yvette, France

[4]Institut de Biochimie et Biophysique Moléculaire et Cellulaire, CNRS-Université Paris-Sud UMR 8619, 91400 Orsay, France

[5]Ecole Polytechnique-CNRS UMR7654, 91120 Palaiseau, France

[6]Co-first author

*Correspondence: jacqueline.cherfils@ens-cachan.fr

http://dx.doi.org/10.1016/j.str.2015.12.001

SUMMARY

Arpin is a newly discovered regulator of actin polymerization at the cell leading edge, which steers cell migration by exerting a negative control on the Arp2/3 complex. Arpin proteins have an acidic tail homologous to the acidic motif of the VCA domain of nucleation-promoting factors (NPFs). This tail is predicted to compete with the VCA of NPFs for binding to the Arp2/3 complex, thereby mitigating activation and/or tethering of the complex to sites of actin branching. Here, we investigated the structure of full-length Arpin using synchrotron small-angle X-ray scattering, and of its acidic tail in complex with an ankyrin repeats domain using X-ray crystallography. The data were combined in a hybrid model in which the acidic tail extends from the globular core as a linear peptide and forms a primary epitope that is readily accessible in unbound Arpin and suffices to tether Arpin to interacting proteins with high affinity.

INTRODUCTION

Cell migration is powered by the polymerization of actin meshworks at the cell leading edge. This process is activated by a regulatory cascade that involves the small GTPase Rac, the nucleation-promoting factor (NPF) WAVE complex, and the actin-branching complex Arp2/3, each promoting conformational changes into the next component that eventually result in the creation of a new actin branch (reviewed in Padrick and Rosen, 2010). Arpin was recently identified as a new player in this network, which mitigates the dynamics of actin polymerization through a Rac-Arpin-Arp2/3 inhibitory circuit that steers cell migration (Dang et al., 2013). Notably, depletion of Arpin increased the speed and directional persistence of migration in mammalian cells and in the ameba *Dictyostelium discoideum*, and its microinjection into fish keratocytes slowed migration speed and induced cells to turn (Dang et al., 2013). Thus, Arpin is a novel member in a growing repertoire of negative regulators of Arp2/3 functions (reviewed in Krause and Gautreau, 2014).

Arpin is conserved in many animals, including mammals, amphibia, reptiles, and fish, and is also found in gastropods, sponges, and ameba, although an Arpin sequence has not been detected so far in available bird or insect genomes (Figure S1A). It is a protein of about 25 kDa, characterized by the presence of the C-terminal acidic motif, which bears a striking resemblance to the acidic moiety of the VCA (verprolin-homology, central, and acidic motif) domain of NPFs (Dang et al., 2013) (Figure S1B). In WASP, WAVE, and other NPFs, the VCA is a regulatory element that binds to and activates the Arp2/3 complex in response to various inputs and tethers the complex to sites of actin branching (reviewed in Padrick and Rosen, 2010). The observation that full-length Arpin, but not a mutant lacking the acidic motif, impairs the ability of a VCA domain to pull down cellular Arp2/3 and inhibits actin polymerization promoted in vitro by a VCA peptide and the Arp2/3 complex, suggested that the acidic motif recognizes a VCA-binding site on Arp2/3 (Dang et al., 2013).

VCA regions are tightly regulated in order to control the dynamics of actin assembly in time and space. In WASP (Kim et al., 2000) and in the WAVE complex (Chen et al., 2010), the VCA is masked by intramolecular interactions, which contribute to maintaining the NPF in an inactive state. Upon activation of NPFs by Rac GTPases and other factors, the VCA motif becomes available to bind to actin and to the Arp2/3 complex (reviewed in Padrick and Rosen, 2010). Currently, the number and precise location of VCA-binding sites in the Arp2/3 complex continue to be investigated (Boczkowska et al., 2008, 2014; Gaucher et al., 2012; Padrick et al., 2011; Ti et al., 2011). A C-terminal acidic motif has also been identified in PICK1, which acts as a negative regulator of Arp2/3 (Rocca et al., 2008), although whether it binds to Arp2/3 directly is currently being debated (Madasu et al., 2015).

The presence of an acidic tail in Arpin raises the question of the extent to which its structural, regulatory, and functional properties mimic those of the acidic motif of the VCA regions. Currently little is known about the structure of Arpin. Previous

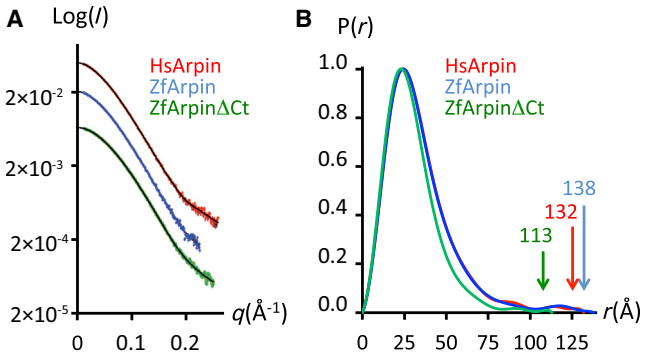


Figure 1. SAXS Analysis of Human and Fish Arpin

(A) SEC-SAXS scattering profiles of HsArpin (red), ZfArpin (blue) and ZfArpinΔCt (green). The scattering profiles calculated from the *P*(*r*) functions are superposed in black.

(B) *P*(*r*) distance distribution functions of HsArpin (red), ZfArpin (blue), and ZfArpinΔCt (green). D_{max} values are indicated. It should be noted that smaller D_{max} values resulted in a significantly poorer fit of the calculated profiles when compared with the experimental scattering profiles.

See also Figure S1, showing sequence alignments of Arpin proteins and acidic peptides from different species and SDS-PAGE of Arpin samples used in this study.

nuclear magnetic resonance analysis indicated that the 20 C-terminal residues in human Arpin have a high mobility in the picosecond to millisecond timescale, suggesting that they are exposed to the solvent and poorly structured (Dang et al., 2013). So far, Arpin has also resisted crystallographic analysis, likely because crystal growth is impaired by the flexibility of its C terminus. Small-angle X-ray scattering (SAXS) provides a powerful approach to analyze the conformation of partially unstructured proteins (reviewed in Bernado and Svergun, 2012; Rambo and Tainer, 2013; Receveur-Brechot and Durand, 2012). Here, we used a combination of SAXS to study the structure of unbound and protein-bound Arpin constructs from human and fish in solution, and X-ray crystallography to observe the conformation of its C-terminal acidic tail at high resolution. These data reveal that Arpin is comprised of an elongated globular core from which the C-terminal tail extends as a linear peptide, thus forming a primary epitope that is readily available to tether Arpin to interacting proteins.

RESULTS

SAXS Analysis of Human and Fish Arpin Proteins

The function of Arpin proteins in cell motility was previously investigated in human, fish, and ameba cells (Dang et al., 2013). None of these Arpin proteins produced crystals, which could be due to the flexibility of their acidic tails. Alternatively, we turned to SAXS analysis, which provides data on the shape of globular and flexible regions of proteins in solution. To increase the significance of our analysis, we conducted studies on both human Arpin (HsArpin) and zebrafish Arpin (ZfArpin), which have 59% sequence identity and similar acidic tails (^{211}EIREQGDGAEDEEWDD^{226} in HsArpin, ^{211}KSAAQGEGADD DEWDD^{226} in ZfArpin, see also Figures S1A and S1B). HsArpin, ZfArpin, and a ZfArpin mutant lacking the 16 C-terminal residues (ZfArpinΔCt hereafter) were expressed as recombinant proteins and purified to homogeneity (Figure S1C). Fusion tags used to facilitate purification were removed before structural analysis in order to eliminate any contribution to the scattering intensities. The SAXS profiles of HsArpin, ZfArpin, and ZfArpinΔCt were recorded on SWING beamline at synchrotron SOLEIL (Figure 1A) and analyzed with the FOXTROT (synchrotron SOLEIL) and ATSAS (Petoukhov et al., 2012) software suites (Table 1). Samples were separated based on their size and shape by online size-exclusion chromatography (SEC). The SEC elution profiles of all samples were symmetrical and devoid of shoulders and trailing tails, and the radius of gyration (R_g) values calculated for each image were stable throughout the elution peak (Figure S2), which is a strong indication that the samples were monodisperse. The scattering profiles of HsArpin and ZfArpin superpose very well, confirming that their structures are very similar. The R_g of HsArpin and ZfArpin estimated by Guinier analysis are 26.8 and 26.9 Å, respectively. These values are considerably larger than those expected for a globular protein of 25 kDa, as exemplified by chymotrypsinogen, a 25 kDa globular protein the R_g of which is 18 Å (Perkins et al., 1993). Accordingly, the distance distribution functions *P*(*r*) indicate that HsArpin and ZfArpin have maximum diameter values (D_{max}) of 132 and 138 Å, respectively (Figure 1B). Truncation of the C terminus from ZfArpin reduced its R_g value by 4 Å and its D_{max} value by 25 Å (Figure 1B). A similar SAXS analysis could not be carried out with HsArpin truncated at its C-terminal tail, because this construct was more prone to aggregation. These data indicate that the C terminus of Arpin proteins greatly contributes to their unusually large R_g and D_{max} values. However, the R_g and D_{max} values of ZfArpinΔCt remain larger than what would be expected for a mostly spherical protein, indicating that the shape of the globular core of Arpin proteins is non-spherical in solution.

Ab Initio Modeling of Arpin Protein Structures

The high-intensity SAXS data from the SEC peaks were averaged to generate ab initio models of HsArpin, ZfArpin, and ZfArpinΔCt. For each construct, 20 models were generated with DAMMIF, where the protein volume is filled with beads (Franke and Svergun, 2009), and with GASBOR, in which dummy residues are arranged as a polypeptide chain (Svergun et al., 2001). All models have a very good fit to the experimental SAXS data as measured by χ^2 values, and very good internal consistency as measured by normalized spatial discrepancy (NSD) values (Figures 2 and S3, Table 1). HsArpin and ZfArpin models consistently define a tadpole shape very similar for both proteins, in which individual models feature an elongated compact core (the head of the tadpole) and a linear extension (the tail fin) (Figures 2A and 2B). The models of HsArpin and ZfArpin generated by GASBOR also feature a non-spherical globular core with a linear extension (Figures S3A and S3B). As shown in Figures 2C and S3C, ZfArpinΔCt retains an elongated shape similar to that of the globular core of HsArpin and ZfArpin, but with a significantly smaller linear extension, compatible with the interpretation that the extension seen in the models of the full-length proteins corresponds to their C terminus tail. To visually estimate the departure of each of these proteins from globularity, we used the

Table 1. Statistics on SAXS Data Collection, Analysis, and Modeling

Data Collection Parameters						
Instrument				SWING beamline, equipped with a PCCD170170 Aviex Detector		
Beam geometry				Pinhole		
Wavelength (Å)				1.033		
q range (Å^{-1})				0.01–0.600		
Exposure time (s)				1.5 per frame		
Temperature (K)				295		
		ARD	HsArpin	ZfArpin	ZfArpinΔCt	ARD/HsArpin
Structural Parameters						
From Guinier fit	$I(0)$ (cm^{-1})	$0.043 \pm 2 \times 10^{-4}$	$0.050 \pm 5 \times 10^{-4}$	$0.041 \pm 7 \times 10^{-4}$	$0.067 \pm 4 \times 10^{-4}$	$0.072 \pm 5 \times 10^{-4}$
	R_g (Å)	17.70 ± 0.01	26.80 ± 0.07	26.89 ± 0.14	22.80 ± 0.04	33.00 ± 0.06
From $P(r)$	$I(0)$ (cm^{-1})	$0.044 \pm 1 \times 10^{-4}$	$0.050 \pm 1 \times 10^{-4}$	$0.041 \pm 1 \times 10^{-4}$	$0.067 \pm 1 \times 10^{-4}$	$0.072 \pm 1 \times 10^{-4}$
	R_g (Å)	18.22 ± 0.01	27.83 ± 0.08	27.94 ± 0.10	23.60 ± 0.06	33.49 ± 0.04
	D_{max} (Å)	58	132	138	113	126
Molecular Mass Determination						
Partial specific volume (cm^3 g^{-1})		0.733	0.727	0.721	0.728	0.730
Contrast (10^{10} cm^{-2})		2.941	3.045	3.129	3.022	2.999
Molecular mass (kDa) from $I(0)$		17.1	<27.4[a]	<30.8[a]	<25.5[a]	38.4
Molecular mass (kDa) from sequence		17.6	24.9	25.5	23.8	42.4
Modeling						
DAMMIF	χ^2	0.94	0.86	1.03	0.67	0.52
	NSD	0.54 ± 0.02	0.55 ± 0.02	0.53 ± 0.02	0.54 ± 0.02	0.63 ± 0.03
GASBOR	χ^2	–	1.4	1.19	1.19	–
	NSD	–	1.21 ± 0.04	1.21 ± 0.03	1.06 ± 0.03	–
BUNCH	χ^2	–	–	–	–	1.13
	NSD	–	–	–	–	1.26 ± 0.04
MONSA	χ^2	–	–	–	–	0.98/0.77/1.21
	NSD	–	–	–	–	0.52 ± 0.01

See also Figure S2, showing the $I(0)$ and R_g analysis of chromatography-coupled SAXS frames.

[a]Due to high concentration at the elution peak, the HPLC-UV detector was saturated and the concentrations of the sample were consequently underestimated. In that case, the molecular mass is overestimated; however, it remains close to the predicted molecular weight. We show in Figure S2 that R_g is constant across the peak, indicating that the sample is homogeneous and is devoid of aggregates.

dimensionless Kratky Plot (DLKP). The DLKP of a purely globular protein displays a maximum of about 1.1 at an abscissa of about 1.75, whereas both the maximum value and the abscissa increase as the protein is more elongated or partially unfolded (Durand et al., 2010; Receveur-Brechot and Durand, 2012). All Arpin constructs display a bell-shaped DLKP curve, indicating without ambiguity that they contain a folded domain (Figure 2D). However, the ZfArpinΔCt curve is closer to the theoretical curve of a globular protein than those of HsArpin and ZfArpin, which are shifted toward higher qR_g values. This analysis further substantiates the model in which the elongated or unfolded part is significantly shorter in ZfArpinΔC than in full-length Arpin proteins, and thus corresponds to the C-terminal tail.

SAXS Analysis of an Arpin-Ankyrin Repeats Complex

Previous work using different cellular models has shown that the acidic tail of Arpin proteins is a major component of their ability to interfere with Arp2/3 functions, which led to the hypothesis that it competes directly with the VCA of the WAVE complex for binding to the Arp2/3 complex (Dang et al., 2013). The above SAXS analysis of unbound Arpin proteins suggests that the C-terminal tail forms an elongated peptide which makes few interactions with the rest of the protein, raising the possibility that Arpin qualifies as a member of an important class of proteins that form protein-protein interactions via linear peptide epitopes (reviewed in Perkins et al., 2010). To investigate the binding properties of the C-terminal tail, we took advantage of the fact that HsArpin interacts with an ankyrin repeats domain (ARD) from Tankyrase 2 (I.D. and A.G., unpublished data), a multidomain poly(ADP-ribose)polymerase that recruits specific linear motifs in its substrates via its N-terminal ankyrin repeats (Guettler et al., 2011). This interaction was used here solely for structural analysis purposes, and its functional analysis will be reported elsewhere. SEC-SAXS data of an ARD construct from which the 6xHis-GST tag had been removed was recorded, which had an excellent fit to the theoretical scattering curve calculated from the

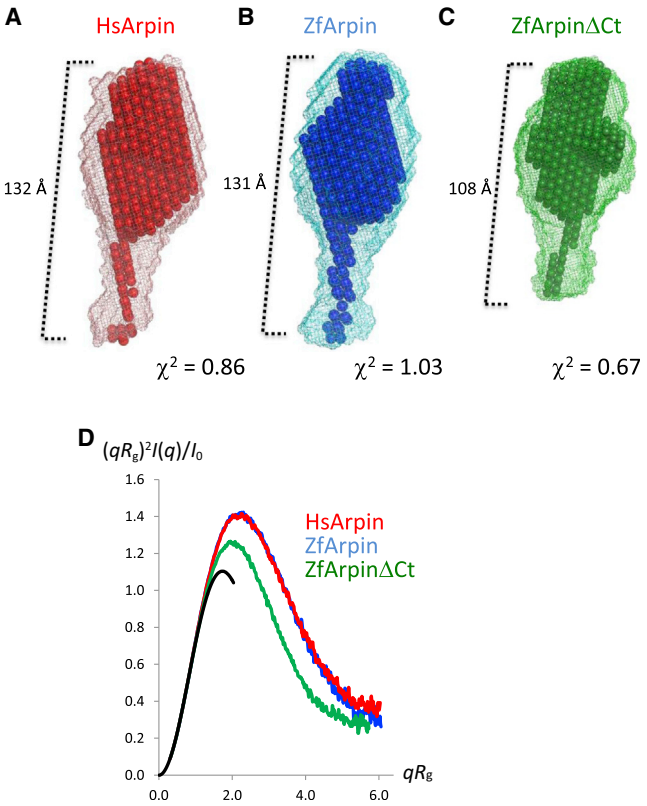


Figure 2. Ab Initio Modeling of Human, Fish, and Truncated Arpin

(A–C) Models of HsArpin (A), ZfArpin (B), and ZfArpinΔCt (C) were generated from high-intensity SAXS data using DAMMIF. For each Arpin construct, 20 independent models were generated and superimposed with DAMAVER. The largest dimension of the models (132, 131, and 108 Å, respectively) are in agreement with the D_{max} values estimated from the $P(r)$ functions. The volume occupied by the superposed models and the model with best χ^2 are shown in mesh and beads, respectively. All models are shown at the same scale. The largest dimension is shown by a ruler.

(D) DLKPs of HsArpin (red), ZfArpin (blue), ZfArpinΔCt (green), and theoretical curve for a purely globular protein (black). All proteins have a bell-shaped curve characteristic of the presence of a globular domain, but have a peak abscissa and height higher than those of a purely globular domain indicating the presence of an elongated and/unfolded element.

See also Figure S3, showing ab initio modeling of Arpin proteins using GASBOR.

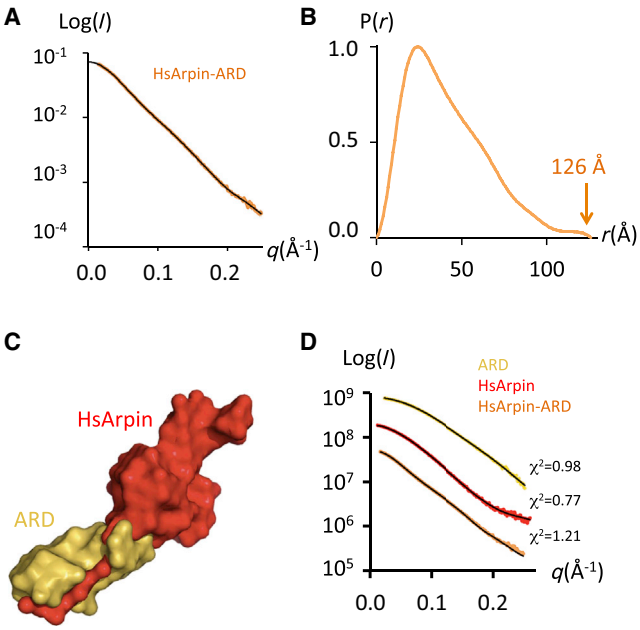


Figure 3. SAXS Analysis of Arpin-ARD Complex

(A) Experimental SEC-SAXS scattering profile of the HsArpin-ARD complex. The scattering profile calculated from the $P(r)$ functions is superposed in black.

(B) $P(r)$ distance distribution function.

(C) A representative ab initio envelope of the HsArpin-ARD complex generated by MONSA from the high-intensity SAXS data of unbound HsArpin, unbound ARD, and the HsArpin-ARD complex. HsArpin is shown in red, the ARD in yellow.

(D) Fit of the experimental scattering intensities of the ARD (yellow), HsArpin (red), and the HsArpin-ARD complex (orange) using the calculated scattering intensities of the MONSA model with the lowest χ^2. The model accounts well for the experimental SAXS data of both the unbound proteins and the complex. See also Figure S4, showing the SAXS analysis of the unbound ARD and the SEC profile of the HsArpin-ARD complex.

crystallographic structure of the unbound ARD (Guettler et al., 2011) (Figures S4A and S4B). The HsArpin-ARD complex was purified by SEC (Figure S4C) and characterized by synchrotron SEC-SAXS (Figures 3A, 3B, and S2, Table 1). The distance distribution function $P(r)$ of the complex has a small shoulder in its right-hand side, which indicates that the two proteins may arrange as separated domains in the complex. The $P(r)$ also lacks trailing high values, which contrasts with the $P(r)$ functions of unbound Arpin proteins. The experimental SAXS curves of unbound HsArpin, unbound ARD, and the HsArpin-ARD complex were combined with the program MONSA (Svergun, 1999) to generate ab initio models of the HsArpin-ARD complex. A model with an excellent fit to the experimental data could be generated with this approach (Figures 3C and 3D). The resulting envelope has an elongated globular core that lacks the linear extension seen in unbound Arpin structures, which indicates that the C terminus tail is no longer exposed and suggests that it is directly involved in tethering Arpin to the ARD. However, the precise position of the C terminus tail and the contribution of the globular core of Arpin to the binding interface cannot be resolved in this envelope.

Hybrid Model of the Arpin-ARD Complex from Combined X-Ray Crystallography and SAXS

Hybrid structural models that incorporate partial atomic structures can be reconstituted by combining SAXS and crystallographic data. We solved the crystal structure of the C-terminal peptide from HsArpin (residues 211–222, HsArpin^{Ct} hereafter) in complex with the ARD at 1.8-Å resolution (Figure 4A, crystallographic statistics are given in Table 2). The Arpin peptide adopts an extended conformation in the complex and forms extensive interactions with the second and third ankyrin repeats of the ARD (Figure 4B). This conformation is similar to that of peptides previously crystallized in complex with this ARD (Guettler et al., 2011), with root-mean-square deviation (RMSD) values in the 2-Å range (Figure 4C and Table S1), the functional significance of which is currently being investigated and will be described elsewhere. Notably, the

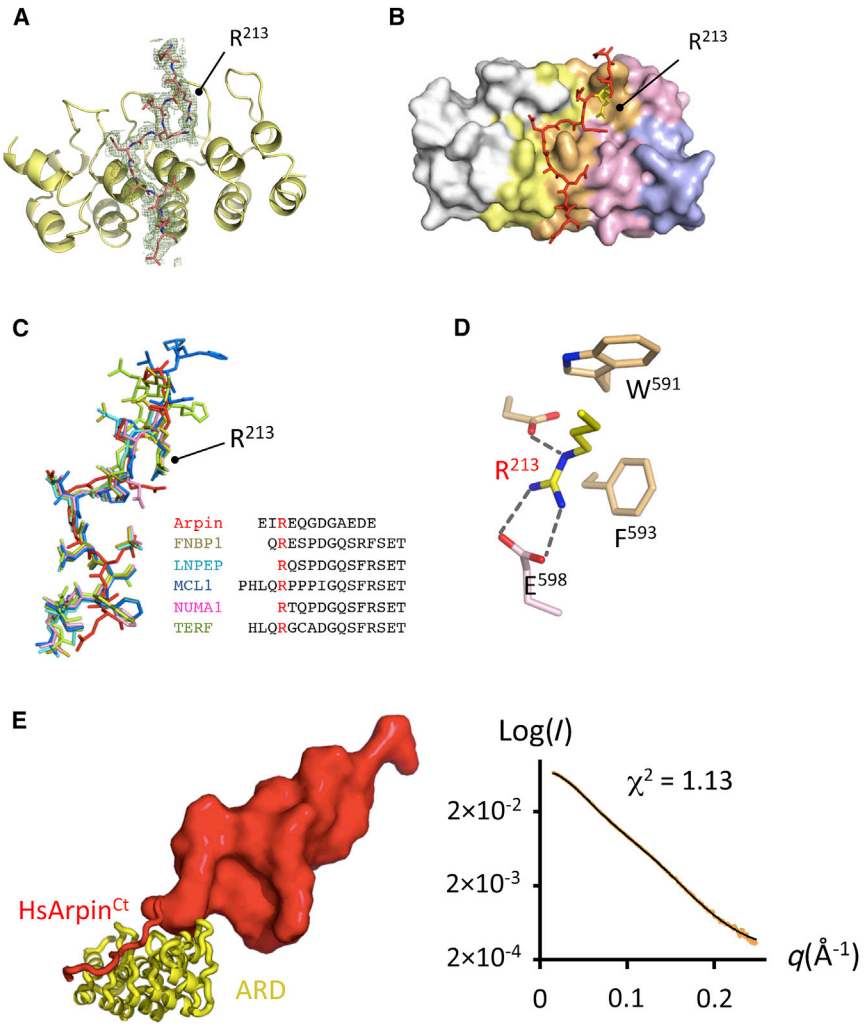


Figure 4. Hybrid Crystallography and SAXS Model of the HsArpin^{Ct}-ARD Complex

(A) $F_o - F_c$ electron density contoured at 3σ of the HsArpin^{Ct}-ARD complex, from which the HsArpin^{Ct} peptide has been omitted. The refined crystallographic structure of the peptide is superposed. The ARD is in yellow.

(B) Crystallographic structure of the HsArpin^{Ct}-ARD complex. The ARD is shown in surface representation, with each ankyrin repeat in a different color. HsArpin^{Ct} is shown in red.

(C) Superposition of the HsArpin^{Ct} peptide (in red) to peptides from ARD-substrate peptide complexes (Guettler et al., 2011). Note that the invariant arginine has a highly conserved position. PDB entry codes of the ARD-peptide complexes shown are: TERF, PDB: 3TWS; MCL1, PDB: 3TWU; NUMA1, PDB: 3TWV; LNPEP, PDB: 3TWW; FNBP1, PDB: 3TWX.

(D) Close-up view of Arg^{213} in the HsArpin^{Ct}-ARD complex. Hydrogen bonds are indicated with dotted lines.

(E) Hybrid model generated by BUNCH using the crystallographic structure of the HsArpin^{Ct}-ARD complex and the SAXS data of the HsArpin-ARD complex. The fit between the experimental SAXS data and the theoretical data calculated from the model are shown.

See also Table S1, which gives the RMSD values between Arpin and other ARD-bound peptides.

conserved arginine found in all peptides (Arg^{213} in Arpin) forms the same set of hydrogen bonds and non-polar interactions with the ARD as seen in all other peptides (Figure 4D). Next, we combined the crystallographic coordinates of the HsArpin^{Ct}-ARD complex and the SAXS data of the HsArpin-ARD complex using BUNCH (Petoukhov and Svergun, 2005). The resulting hybrid model of the HsArpin-ARD complex is composed of the SAXS envelope corresponding to ArpinΔCt and the atomic coordinates of the HsArpin^{Ct}-ARD complex (Figure 4E). The model depicts an arrangement in which the acidic tail of Arpin is sequestered by the ARD, and the globular core of Arpin interacts with the ARD via an interface of restricted area. This analysis suggests that the C-terminal tail of Arpin is a major component of the tethering of Arpin to binding proteins.

The C-Terminal Tail of Arpin Is a Primary Binding Epitope

To analyze quantitatively the hypothesis that the acidic tail of Arpin is a primary epitope, we determined the binding affinities of HsArpin and of the HsArpin^{Ct} peptide for the ARD by isothermal titration calorimetry (ITC) (Figures 5A and 5B, Table 3). The dissociation constants of HsArpin and the HsArpin^{Ct} peptide are 0.65 ± 0.05 and 0.32 ± 0.02 μM, respectively, with thermodynamic parameters that are dominated by the enthalpic contribution ΔH in both cases, indicating that the acidic tail provides most of the binding energy with little or no contribution from the globular core. To further confirm that the interaction of the peptide in the crystal is relevant to the interaction of full-length Arpin, we mutated Arg^{213}, a residue that establishes a network of interactions with the ARD and is conserved in other peptides that bind to this ARD (Figures 4C and 4D). The Arg213Ala mutation reduced the dissociation constant of HsArpin by 10-fold, and that of HsArpin^{Ct} by more than 20-fold. Interestingly, ZfArpin lacks an equivalent arginine in its C terminus, and accordingly no interaction with the ARD was detected by ITC (Figure S5A). Mutation of Ala^{213} to Arg in ZfArpin resulted in partial recovery of association with the ARD (Figure S5B). We conclude from these experiments that the C-terminal acidic tail of Arpin is a primary binding epitope that binds tightly on its own to the ARD, and that the globular core of Arpin does not contribute significantly to this interaction.

DISCUSSION

We analyzed the structure and interactions of Arpin by combined SAXS, X-ray crystallography, and ITC as a step toward understanding its negative regulation of Arp2/3 functions. We showed that Arpin has a globular core with an ellipsoid shape from which

Table 2. Crystallographic Statistics of the Arpin Peptide-ARD Structure

	PDB: 4Z68
Data Collection	
X-Ray source	PROXIMA-2A
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	29.54, 43.80, 107.64
α, β, γ (°)	90, 90, 90
Temperature (K)	100
Wavelength (Å)	0.9791
Resolution (Å)	40.57–1.86 (1.97–1.86)
No. of unique reflections	12,297 (1893)
Redundancy	7.8 (7.9)
Completeness (%)	99.5 (97.0)
$\langle I/\sigma(I)\rangle$	17.8 (5.0)
$CC_{1/2}$ (%)	99.9 (92.9)
R_{meas} (%)	9.8 (57.2)
Refinement	
Resolution range (Å)	40.57–1.86
No. of reflections	12,294
R_{work}/R_{free}	16.98/21.66
RMSD	
Bond length (Å)	0.006
Bond angles (°)	1.041
Ramachadran plot	
Favored (%)	100
Allowed (%)	0
Number of outliers	0

the acidic tail extends as an elongated peptide, and that this tail is a linear primary epitope that suffices to tether Arpin to interacting proteins. The robustness of our structural analysis was ensured by control quality of protein samples, internal validations of ab initio models, cross-validations between human and zebrafish Arpin models, mutagenesis, and thermodynamic analysis of interaction parameters. The acidic tail of Arpin was previously shown to be a critical element in Arp2/3 signaling. However, studies of other proteins that function in the Arp2/3 pathway showed that the presence of an acidic peptide in their C termini does not necessarily imply that these peptides bind to the Arp2/3 complex as a primary epitope, as shown by SAXS analysis for the PICK1 protein (Madasu et al., 2015). Our analysis of the structure and interactions of Arpin reveals that its acidic tail has structural and energetic features that enable it to tether Arpin to interacting proteins with high affinity. These characteristics support a resemblance of the acidic tail of Arpin to the acidic tail of the VCA regions of NPFs for binding to the Arp2/3 complex. In addition, the full accessibility of the acidic tail in unbound Arpin indicates that Arpin is constitutively active, hence its regulation, if any, must occur by mechanisms other than an inhibitory conformational change.

Currently, the atomic details of the structure and interactions of the acidic tail of VCAs remain incompletely understood. In the absence of activating factors, the VCA of N-WASP folds as a compact structure in which the binding sites for activating elements are masked by intramolecular interactions, while the acidic tail is disordered and most likely exposed (Kim et al., 2000). Likewise, the VCA of WAVE1 in the WAVE complex is autoinhibited when folded as an irregular, loosely packed structure that forms intramolecular interactions and contacts with other subunits in the complex, but in which the acidic tail is disordered (Chen et al., 2010). Thus, it is conceivable that the acidic tail of autoinhibited NPFs remains to some extent accessible, similar to the acidic tail of Arpin. How exactly the acidic tail of NPFs interacts with and activates the Arp2/3 complex is also being debated. A body of evidence indicates that the Arp2/3 complex has two binding sites for the acidic region, each with different actin-branching activities and binding energies (Boczkowska et al., 2014; Padrick et al., 2011; Ti et al., 2011). SAXS analysis of an actin- and VCA-bound Arp2/3 complex provided data for a model in which the acidic tail of the VCA binds as an extended structure at the interface between the Arp3 and ARPC1 subunits (Boczkowska et al., 2008). Consistent with this structural study, two distinct yet synergetic binding sites for acidic regions were identified on Arp3 and ArpC1, the Arp3 site appearing to have the greatest impact for activation (Padrick et al., 2011). Direct interaction of the C-terminal tripeptide of a VCA region with the Arp3 subunit was observed by X-ray crystallography, while the rest of the acidic tail was not ordered in this complex (Ti et al., 2011). The latter study suggested an intriguing scenario in which the high-affinity acidic tail-binding site of the Arp2/3 complex would contribute negatively to actin branch formation, while the low-affinity site located on the Arp3 subunit would have a positive contribution. A related scenario proposed that activating elements, including the VCA and the Cdc42 GTPase, form a regulatory module in order to hold the Arp2/3 complex in an inactive conformation awaiting other activation inputs (Prehoda et al., 2000). The VCAs of N-WASP and WAVE were recently shown by ITC to bind to the low- and high-affinity binding sites of the Arp2/3 complex, with dissociation constants in the 20–40 and 0.2–1 μM range, respectively (Boczkowska et al., 2014). Interestingly, the affinity of Arpin for the ARD is similar to that of the high-affinity VCA-binding site, and at least one order of magnitude greater than that of the low-affinity binding site. In light of this still fragmentary knowledge, our identification of the acidic tail of Arpin as an exposed primary epitope with the potential to bind to partners with affinities in the low micromolar range suggests that Arpin could compete for either of the VCA-binding sites and that the acidic tails of Arpin and WAVE may be available to bind simultaneously to the Arp2/3 complex. Ultimately, this may bring about an exquisite balance between linked positive and negative interactions regulating actin branch formation. Progress in understanding this complex crosstalk between epitopes of similar structures will require more information on the nature, conformation, and affinity of the Arpin-binding site(s) within the Arp2/3 complex.

EXPERIMENTAL PROCEDURES

Cloning, Protein Production, and Purification

Full-length human Arpin (HsArpin), full-length *Danio rerio* (zebrafish) Arpin (ZfArpin), truncated zebrafish Arpin (residues 1–210; ZfArpinΔCt), and the

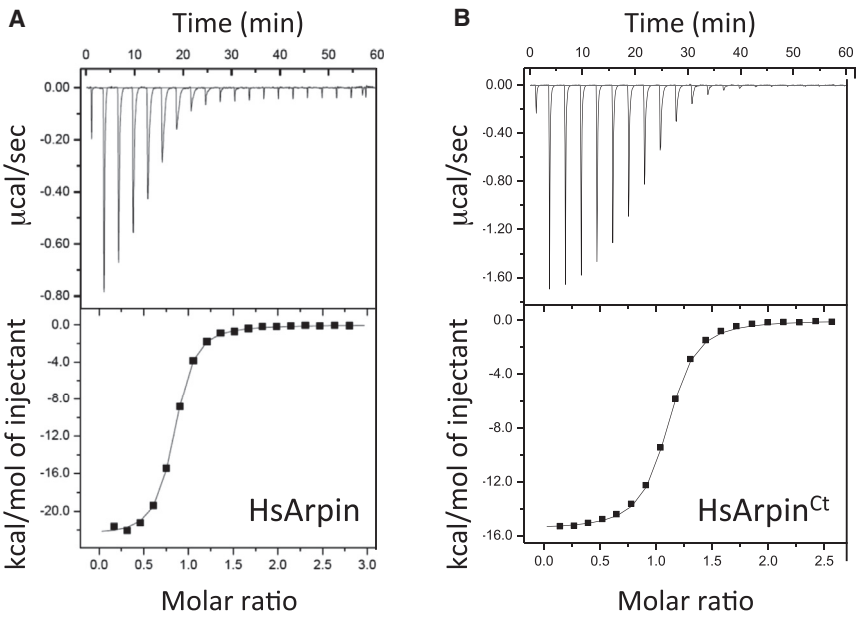


Figure 5. ITC Analysis of the Interaction of HsArpin Proteins and Peptides with the ARD

In these experiments, the ARD was in the syringe and HsArpin (A) and HsArpin^{Ct} (B) were in the cell. The raw thermogram data are in the upper panel, the binding isotherms are in the lower panel.

See also Figure S5, showing the ITC analysis of the interaction of ZfArpin with the ARD.

ARC4 ankyrin repeat domain of human Tankyrase2 (residues 489–649) were subcloned into the pETG30A vector (a kind gift from Arie Geerlof, EMBL), resulting in proteins fused with an N-terminal 6xHis-GST tag followed by a tobacco etch virus (TEV) protease cleavage site. The ZfArpin^{A213R} and HsArpin^{R213A} mutants were made using a two-step PCR procedure to introduce mutations followed by recombination cloning in the pETG-30A vector. All clones were verified by sequencing (GATC Biotech).

All proteins were produced in Rosetta(DE3)pLysS *Escherichia coli* strains. Expression was induced by 0.5 mM isopropyl β-D-1-thiogalactopyranoside at 37°C and carried on overnight at 20°C. Cells were harvested by centrifugation and resuspended in lysis buffer containing a protease inhibitor cocktail. The buffer for Arpin constructs was 50 mM Tris (pH 8), 100 mM NaCl, 2.5 mM $MgCl_2$, 5 mM imidazole, 2 mM BME; for the ARD, 50 mM Tris (pH 7.5), 500 mM NaCl, 5 mM imidazole, 2 mM BME. Cells were sonicated, centrifuged at 20,000 × *g* for 35 min, and the supernatant was filtered using a 0.22 µm filter. All proteins were first purified through an Ni-NTA affinity column (GE Healthcare, AKTA) using 10 and 25 mM imidazole for washing and elution, followed by desalting. The 6xHis-GST tag was removed by cleavage with recombinant 6xHis-tagged TEV protease (1:25 w/w ratio) at 4°C overnight, and separated from untagged proteins by a second Ni-NTA passage. Purification was polished by SEC. The purity of all proteins was >90% as assessed by SDS-PAGE (Figure S1C). Typical yield for purified Arpin constructs was 20 mg/l culture. All proteins were concentrated to 15–20 mg/ml, flash frozen in liquid N_2 and stored at −80°C.

Chromatography-Coupled SAXS

SAXS experiments were conducted on the SWING beamline at the SOLEIL synchrotron (λ = 1.033 Å). The Aviex charge-coupled device detector was positioned to collect useful data in the *q* range 0.01–0.430 Å^{-1} ($q = 4\pi\sin\theta\lambda^{-1}$, where 2θ is the scattering angle). All solutions were flowed at a fixed temperature (15°C) through a quartz capillary with a diameter of 1.5 mm and a wall thickness of 10 µm, and placed in a vacuum chamber. Samples at a concentration of 7–10 mg/ml were dialyzed overnight against a buffer containing 50 mM HEPES (pH 7.5), 100 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP), then injected into a high-performance liquid chromatography (HPLC) SEC (SEC-3 300 Å, Agilent) using an Agilent HPLC system, and eluted directly into the SAXS flow-through capillary cell at a flow rate of 0.2 ml/min essentially as described in Biou et al. (2010). Frames were collected continuously with a duration of 1.5 s and a dead time of 0.5 s between frames. All proteins eluted as a single peak. Frames corresponding to the high-intensity fractions of the peak and having constant R_g (within noise) were averaged. A large number of frames collected before the void volume were averaged and subtracted from the signal of the protein to account for buffer scattering. Data reduction to absolute units, frame averaging, and subtraction were done using the FOXTROT program (synchrotron SOLEIL). All subsequent data processing was carried out with the ATSAS suite (Petoukhov et al., 2012). The forward scattering $I(0)$ and the radius of gyration R_g were derived by the Guinier approximation $I(q) = I(0)\exp(-q^2R_g^{2/3})$ for $qR_g < 1.3$ using PRIMUS. The pair-distance distribution functions, $P(r)$, and the maximal dimension of the macromolecule, D_{max} were calculated with GNOM. SAXS data collection is summarized in Table 1. Data for the ARD, HsArpin, ZfArpin, ZfArpinΔCt, and the HsArpin-ARD complex have been deposited into the SASBDB database (http://www.sasbdb.org/) with entry codes SASBDB: SASBT2, SASBU2, SASBV2, SASBW2, and SASBX2, respectively.

Ab Initio Modeling

All software used for ab initio modeling was from the ATSAS suite (Petoukhov et al., 2012). To model the structures of the unbound proteins, 20 independent models were generated with either DAMMIF (Franke and Svergun, 2009) in slow mode or GASBOR (Svergun et al., 2001), compared and aligned with SUPCOMB, and averaged with DAMAVER (Petoukhov et al., 2012) to determine common structural features and representative shapes. The theoretical scattering curve of the unbound ARD was calculated based on its known crystallographic structure (Guettler et al., 2011) and compared with the experimental SAXS curve using the program CRYSOL (Petoukhov et al., 2012). To model the envelope of the HsArpin-ARD complex, 20 independent models were generated by fitting together the scattering curves of HsArpin, the

Table 3. Analysis of the Affinity of Arpin Proteins for the ARD Analyzed by ITC

	n	K_d (µM)	Δ*H* (kcal/mol)	−*T*Δ*S* (kcal/mol)	Δ*G* (kcal/mol)
HsArpin	1	0.65 ± 0.05	−16.8 ± 0.1	8.6	−8.2
HsArpin^{Ct}	0.8	0.32 ± 0.02	−23.5 ± 0.2	15.2	−8.3
HsArpin^{R213A}	0.8	28 ± 5	−6.75 ± 1.35	0.65	−6.1
HsArpin$^{Ct\text{-}R213A}$	0.9	11.8 ± 1.4	−9 ± 0.8	2.4	−6.6
ZfArpin		Not measurable			
ZfArpin^{A213R}	1.2	22.8 ± 2.5	−3.1 ± 0.5	−3.1	−6.2

ARD, and the ARD-HsArpin complex with MONSA (Svergun, 1999), which uses simulated annealing to model a complex comprised of two components (here, the ARD and HsArpin) from an ensemble of beads inside a defined search volume. To generate the hybrid HsArpin-ARD model, the scattering data of the HsArpin-ARD complex were restrained by the crystal structure of the HsArpin^{Ct}-ARD complex using the program BUNCH (Petoukhov and Svergun, 2005). The residue count was adjusted by adding residues lacking electronic density in the ARD (two residues in the N terminus) and in the Arpin peptide (five residues in the C terminus) with MODELLER using the automodel class (Webb and Sali, 2014). The unknown region (ArpinΔCter) was represented by an appropriate number of beads. Ab initio modeling statistics are summarized in Table 1.

Crystallization and Crystallographic Analysis

The ARD-HsArpin^{Ct} complex was prepared in the presence of a 1.6 M excess peptide at a final protein concentration of 20 mg/ml. Initial sitting drop screens yielded small crystals that diffracted at low resolution on PROXIMA1 beamline (synchrotron SOLEIL), which were subsequently optimized with the hanging-drop vapor diffusion method in 2.2 M ammonium sulfate and 7% (v/v) isopropanol at 18°C. Crystals were harvested, supplemented with 30% PEG 400 for cryoprotection, and flash frozen in liquid nitrogen. Diffraction data were collected at 100 K on the PROXIMA2 microfocus beamline (synchrotron SOLEIL) and processed with XDS (Kabsch, 2010) using the XDSME scripts (Pierre Legrand, PROXIMA1 beamline, SOLEIL). Crystals belong to space group $P2_12_12_1$ and diffracted at a resolution of at least 1.86 Å. The structure was solved by molecular replacement with PHENIX (Adams et al., 2010) using the crystal structure of unbound ARD (Guettler et al., 2011), PDB: 3TWQ as a model. Refinement was carried out with PHENIX and graphical building with COOT (Emsley et al., 2010). Quality of the model was assessed using MolProbity. Data collection and structure refinement statistics are given in Table 2. Coordinates have been deposited with the PDB (PDB: 4Z68).

Characterization of Arpin-ARD Interactions

The complexes between Arpin constructs and the ARD were prepared by incubation in a 1:2.3 molar ratio at 4°C overnight and purified by SEC (Figure S4C). Peptides corresponding to residues 211–222 from HsArpin (EIREQGDGAEDE) and the corresponding EIAEQGDGAEDE mutant peptide were ordered from GL Biochem. Dissociation constants were measured by ITC using an ITC200 device (Malvern, Microcal). Samples were dialyzed against 50 mM HEPES (pH 7.5), 110 mM NaCl, and 1 mM TCEP 4°C. In all experiments, the ARD was in the syringe at 500 μM and Arpin samples were in the cell at 50 μM. All measures were done at least in duplicate. The data were fitted using a single-site binding model.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and one table and can be found with this article online at http://dx.doi.org/10.1016/j.str.2015.12.001.

AUTHOR CONTRIBUTIONS

S.F. purified proteins, performed the crystallographic, mutagenesis, and SAXS experiments, and solved the crystal structure; S.T. performed and analyzed the SAXS experiments; V.C. assisted in protein preparation and crystallographic experiments; M.A.-N. performed the ITC experiments; I.D. and A.G. identified the interaction between Arpin and the ARD; J.P. supervised the SAXS analysis; J.C. designed and coordinated the study and wrote the paper with input from the other authors.

ACKNOWLEDGMENTS

This work was supported by a grant from the Agence Nationale de la Recherche to J.C. and A.G. We are grateful to Andrew Thompson, William Sheppard, and the scientific staff at PROXIMA1 and PROXIMA2 at the synchrotron SOLEIL (Gif-sur-Yvette, France) for making the macromolecular crystallography beamlines available to us and for their help and excellent advice; to François Peurois (LPBA, CNRS-ENS Cachan) for help with protein sample analysis; and to Tracy Bellande and Armelle Vigouroux for making the LEBS/IMAGIF crystallization platform available to us. We thank Dominique Durand (IBBMC, CNRS-Université Paris-Sud) for helpful discussions.

Received: May 28, 2015
Revised: November 24, 2015
Accepted: December 1, 2015
Published: January 7, 2016

REFERENCES

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Bernado, P., and Svergun, D.I. (2012). Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. Mol. Biosyst. *8*, 151–167.

Biou, V., Aizel, K., Roblin, P., Thureau, A., Jacquet, E., Hansson, S., Guibert, B., Guittet, E., van Heijenoort, C., Zeghouf, M., et al. (2010). SAXS and X-ray crystallography suggest an unfolding model for the GDP/GTP conformational switch of the small GTPase Arf6. J. Mol. Biol. *402*, 696–707.

Boczkowska, M., Rebowski, G., Petoukhov, M.V., Hayes, D.B., Svergun, D.I., and Dominguez, R. (2008). X-ray scattering study of activated Arp2/3 complex with bound actin-WCA. Structure *16*, 695–704.

Boczkowska, M., Rebowski, G., Kast, D.J., and Dominguez, R. (2014). Structural analysis of the transitional state of Arp2/3 complex activation by two actin-bound WCAs. Nat. Commun. *5*, 3308.

Chen, Z., Borek, D., Padrick, S.B., Gomez, T.S., Metlagel, Z., Ismail, A.M., Umetani, J., Billadeau, D.D., Otwinowski, Z., and Rosen, M.K. (2010). Structure and control of the actin regulatory WAVE complex. Nature *468*, 533–538.

Dang, I., Gorelik, R., Sousa-Blin, C., Derivery, E., Guerin, C., Linkner, J., Nemethova, M., Dumortier, J.G., Giger, F.A., Chipysheva, T.A., et al. (2013). Inhibitory signalling to the Arp2/3 complex steers cell migration. Nature *503*, 281–284.

Durand, D., Vives, C., Cannella, D., Perez, J., Pebay-Peyroula, E., Vachette, P., and Fieschi, F. (2010). NADPH oxidase activator p67(phox) behaves in solution as a multidomain protein with semi-flexible linkers. J. Struct. Biol. *169*, 45–53.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Franke, D., and Svergun, D.I. (2009). DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. J. Appl. Crystallogr. *42*, 342–346.

Gaucher, J.F., Mauge, C., Didry, D., Guichard, B., Renault, L., and Carlier, M.F. (2012). Interactions of isolated C-terminal fragments of neural Wiskott-Aldrich syndrome protein (N-WASP) with actin and Arp2/3 complex. J. Biol. Chem. *287*, 34646–34659.

Guettler, S., LaRose, J., Petsalaki, E., Gish, G., Scotter, A., Pawson, T., Rottapel, R., and Sicheri, F. (2011). Structural basis and sequence rules for substrate recognition by Tankyrase explain the basis for cherubism disease. Cell *147*, 1340–1354.

Kabsch, W. (2010). XDS. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Kim, A.S., Kakalis, L.T., Abdul-Manan, N., Liu, G.A., and Rosen, M.K. (2000). Autoinhibition and activation mechanisms of the Wiskott-Aldrich syndrome protein. Nature *404*, 151–158.

Krause, M., and Gautreau, A. (2014). Steering cell migration: lamellipodium dynamics and the regulation of directional persistence. Nat. Rev. Mol. Cell Biol. *15*, 577–590.

Madasu, Y., Yang, C., Boczkowska, M., Bethoney, K.A., Zwolak, A., Rebowski, G., Svitkina, T., and Dominguez, R. (2015). PICK1 is implicated in organelle motility in an Arp2/3 complex-independent manner. Mol. Biol. Cell *26*, 1308–1322.

Padrick, S.B., and Rosen, M.K. (2010). Physical mechanisms of signal integration by WASP family proteins. Annu. Rev. Biochem. *79*, 707–735.

Padrick, S.B., Doolittle, L.K., Brautigam, C.A., King, D.S., and Rosen, M.K. (2011). Arp2/3 complex is bound and activated by two WASP proteins. Proc. Natl. Acad. Sci. USA *108*, E472–E479.

Perkins, S.J., Smith, K.F., Kilpatrick, J.M., Volanakis, J.E., and Sim, R.B. (1993). Modelling of the serine-proteinase fold by X-ray and neutron scattering and sedimentation analyses: occurrence of the fold in factor D of the complement system. Biochem. J. *295*, 87–99.

Perkins, J.R., Diboun, I., Dessailly, B.H., Lees, J.G., and Orengo, C. (2010). Transient protein-protein interactions: structural, functional, and network properties. Structure *18*, 1233–1243.

Petoukhov, M.V., and Svergun, D.I. (2005). Global rigid body modeling of macromolecular complexes against small-angle scattering data. Biophys. J. *89*, 1237–1250.

Petoukhov, M.V., Franke, D., Shkumatov, A.V., Tria, G., Kikhney, A.G., Gadja, 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.

Prehoda, K.E., Scott, J.A., Mullins, R.D., and Lim, W.A. (2000). Integration of multiple signals through cooperative regulation of the N-WASP-Arp2/3 complex. Science *290*, 801–806.

Rambo, R.P., and Tainer, J.A. (2013). Super-resolution in solution X-ray scattering and its applications to structural systems biology. Annu. Rev. Biophys. *42*, 415–441.

Receveur-Brechot, V., and Durand, D. (2012). How random are intrinsically disordered proteins? A small angle scattering perspective. Curr. Protein Pept. Sci. *13*, 55–75.

Rocca, D.L., Martin, S., Jenkins, E.L., and Hanley, J.G. (2008). Inhibition of Arp2/3-mediated actin polymerization by PICK1 regulates neuronal morphology and AMPA receptor endocytosis. Nat. Cell Biol. *10*, 259–271.

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.

Ti, S.C., Jurgenson, C.T., Nolen, B.J., and Pollard, T.D. (2011). Structural and biochemical characterization of two binding sites for nucleation-promoting factor WASp-VCA on Arp2/3 complex. Proc. Natl. Acad. Sci. USA *108*, E463–E471.

Webb, B., and Sali, A. (2014). Protein structure modeling with MODELLER. Methods Mol. Biol. *1137*, 1–15.

