



Cite this article: Ni T, Harlos K, Gilbert R. 2016 Structure of astrotactin-2: a conserved vertebrate-specific and perforin-like membrane protein involved in neuronal development. *Open Biol.* **6**: 160053. <http://dx.doi.org/10.1098/rsob.160053>

Received: 3 March 2016
Accepted: 7 April 2016

Subject Area:
biophysics/biochemistry/structural biology/
neuroscience

Keywords:
astrotactin-2, membrane attack complex-perforin protein, X-ray crystallography, biophysical interaction analysis, neural migration control

Author for correspondence:
Robert Gilbert
e-mail: gilbert@strubi.ox.ac.uk

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsob.160053>.

Structure of astrotactin-2: a conserved vertebrate-specific and perforin-like membrane protein involved in neuronal development

Tao Ni, Karl Harlos and Robert Gilbert

Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK

The vertebrate-specific proteins astrotactin-1 and 2 (ASTN-1 and ASTN-2) are integral membrane perforin-like proteins known to play critical roles in neurodevelopment, while ASTN-2 has been linked to the planar cell polarity pathway in hair cells. Genetic variations associated with them are linked to a variety of neurodevelopmental disorders and other neurological pathologies, including an advanced onset of Alzheimer's disease. Here we present the structure of the majority endosomal region of ASTN-2, showing it to consist of a unique combination of polypeptide folds: a perforin-like domain, a minimal epidermal growth factor-like module, a unique form of fibronectin type III domain and an annexin-like domain. The perforin-like domain differs from that of other members of the membrane attack complex-perforin (MACPF) protein family in ways that suggest ASTN-2 does not form pores. Structural and biophysical data show that ASTN-2 (but not ASTN-1) binds inositol triphosphates, suggesting a mechanism for membrane recognition or secondary messenger regulation of its activity. The annexin-like domain is closest in fold to repeat three of human annexin V and similarly binds calcium, and yet shares no sequence homology with it. Overall, our structure provides the first atomic-resolution description of a MACPF protein involved in development, while highlighting distinctive features of ASTN-2 responsible for its activity.

1. Introduction

Perforin-like proteins (PLPs) have been identified in all forms of cellular life except, currently, Archaeobacteria [1]. They represent a sub-branch of the largest known family of pore-forming proteins, the membrane attack complex-perforin/cholesterol-dependent cytolyisin (MACPF/CDC) family [2,3]. The CDCs were identified in Gram-positive bacteria [4,5] and their mechanism of action has been thoroughly studied using a combination of structural and mechanistic approaches [6,7]. Perforins were separately identified, first as part of the complement membrane attack complex (MAC) [8,9] and then in the form of perforin-1 [10,11], which delivers granzymes from cytotoxic cells into target antigen presenting cells. The solution of the structures of complement C8 α and a bacterial PLP revealed that MACPFs and CDCs belong to one homologous family of proteins [12,13], and that their identification as separate groupings was a historical accident [14].

The only basis for MACPF/CDC activity properly established so far is the oligomerization of many subunits in order to generate a pore-forming complex [15–19]. Binding to a target membrane serves the basic function of concentrating monomeric subunits on a planar substrate [14]. Then, concentration-dependent oligomerization occurs into a pre-pore complex before pore formation ensues [16,19–21]. It has been argued both that oligomerization to a complete ring of subunits is required for pore formation [7,22], and that incomplete rings (arcs) of MACPF/

CDC subunits can insert into membranes to form pores, with the size of any individual pore-forming assembly being determined by the availability of further monomers for incorporation into the pre-pore [5,6,18,19,23,24]. Recently, it became clear that indeed individual types of MACPF/CDC protein can form functional pores using oligomers of variable size depending on precisely such a kinetically determined mechanism [18,19,25–28]. The same protein can generate functionally important lesions of different sizes in target membranes depending on prevailing conditions such as the concentration of protein and also, for example, pH [18,19,29,30].

As a subset of MACPF/CDC proteins, we define PLPs as being proteins which have been identified by sequence homology as sharing common evolutionary ancestry with perforin-1 [3,31]. These proteins are involved in a wide range of biological processes, including not only immunity (MAC, perforin-1 and perforin-2) [32–34] but also cell invasion and egress by apicomplexan parasites [35–37] and organismal development [38–43]. Our focus here is on a PLP especially identified as being involved in neurodevelopment, astrotacin-2 (ASTN-2) [39,40]. In addition to its role in neural migration control, a recent report has indicated that ASTN-2 is also part of the mechanism of planar cell polarity (PCP) determination in hair cells via the activation of non-canonical Wnt signalling Frizzled-6 receptors [44]. In this paper, we report structures representing a large portion of ASTN-2. These provide a blueprint for understanding its activity and also that of ASTN-1, which is also neurodevelopmentally significant [38], and to some extent the bone morphogenetic protein and retinoic acid inducible neural-specific proteins (BRINPs) [41,42].

ASTN-1 was first discovered because antibodies against it block neuron–glial interactions *in vitro* [38]; it was subsequently shown to be expressed in post-mitotic neuronal precursors of the cerebellum, hippocampus, cerebrum and olfactory bulb with a role in the establishment of laminar structures [45]. ASTN-1 is directly responsible for the formation of neuron–glial fibre contacts in the cerebellum [45–48], and the discovery of ASTN-2 as its intracellular counterpart showed how ASTN-1 contacts might be recycled through the endosomal system to enable the forward migration of neuronal cells [39] (electronic supplementary material, figure S1).

ASTN-2 is most highly expressed in the cerebellum but also in the hippocampus, cortex and olfactory bulb [39,49]. The expression of ASTN-1 and -2 is differentially regulated in terms of cellular location [39] and developmental stage [50]. While ASTN-1 is well expressed on the cell surface and in neurons forming glial fibre contacts via its C-terminal ectodomain, ASTN-2 is expressed there only very weakly (with ASTN-1 present, on 0.22% of cells tested; in the absence of ASTN-1, on 0.07% of cells tested) [39]. Instead, ASTN-2 localizes mostly to vesicles inside neural cells, giving rise to a punctate antibody staining pattern in the soma and along neuronal processes [39]. In agreement, co-expression of ASTN-1 and fluorescently labelled ASTN-2 in HEK293T cells shows them to be co-localized to a subset of RhoB⁺ endosomes (as well as in non-RhoB⁺ vesicles) [39]. As cerebellar granule neurons grow, ASTN-2 is found co-localized with clathrin at the base of the leading process and opposite an interstitial junction with the glial fibre, which suggests its identification with coated vesicles (electronic supplementary material, figure S1) [39]. Intracellular imaging reveals the cycling of vesicles bearing ASTN-1 from the anterior pole of the neuronal soma and the base of the leading process,

into the cell, and down the leading process to form a new neuron–glial fibre junction towards its tip [39,40]. This has led to the hypothesis whereby ASTN-2 controls the recycling of ASTN-1-mediated contacts between the migrating neuron and glial fibre from the lagging to the leading edge of the moving cell [39] (electronic supplementary material, figure S1).

As might be expected given their role in neurodevelopment, data suggest that ASTN-1 and ASTN-2 are key to many aspects of basic mammalian neurobiology. For example, ASTN-1 is increased in rat brain following hippocampal injury, implying a role in repair processes [51], while ASTN-1 knockout mice display poorer balance and coordination than wild-type [47]. In humans, a comparative genomic hybridization study of a Russian cohort with intellectual disability identified an individual with a duplication of ASTN-1 showing multiple neurodevelopmental defects and delays as well as a number of non-neuronal phenotypic effects [52] which seem to map at least in part onto the PCP pathway [44,53].

Linkage of ASTN-2 to the development of the mammalian CNS is even stronger than that of ASTN-1. ASTN-2 has been implicated via genome-wide association studies in attention-deficit hyperactivity disorder (ADHD) [54], by copy number variant (CNV) analysis of a large human cohort in autism spectrum disorders (ASDs) [55] and in another CNV study in schizophrenia [56]. In one recent paper, a large set of neurodevelopmental disorder (NDD) subjects were compared with a set of population-based controls to identify 46 deletions and 12 duplications affecting ASTN-2; the NDD subjects demonstrated a variety of phenotypes including ASDs, ADHD, speech delay, anxiety and obsessive compulsive disorder [50]. In the same study, analysis of the spatio-temporal expression patterns of ASTN-1 and -2 in human brain samples from individuals of different ages showed ASTN-1 expression at consistently high levels, whereas ASTN-2 expression peaked in the early embryonic neocortex and postnatal cerebellar cortex [50]. Another recent paper identified a clinical link between ASTN-2-associated polymorphisms and the onset of Alzheimer's disease approximately 5 years earlier than the median [57].

ASTN-1 and ASTN-2 have been shown to interact directly with each other *in vitro* in a calcium-independent way with multiple regions of each protein apparently contributing to their interface [39]. Both ASTN-1 and ASTN-2 are integral membrane proteins in which a large C-terminal domain (extracellular for ASTN-1, endosome luminal for ASTN-2; electronic supplementary material, figure S1*b*) is complemented by an N-terminal cytosolic domain suspended between two transmembrane α -helices [39] (figure 1*a*). The C-terminus of each protein is known to be detectable on the outside of cells in which they have been expressed and this defines their membrane orientation (electronic supplementary material, figure S1*b*) [39]. We suggest that just as the bulk of ASTN-1 projects outwards from the plasma membrane the bulk of ASTN-2 will project into the lumen of cellular endosomes, making it topologically equivalent to ASTN-2 with two transmembrane α -helices providing a firm anchor in the membrane bilayer (figure 1*a*; electronic supplementary material, figure S1*b*). Thus, the endosomal regions of ASTN-2 comprise a smaller N-terminal domain and a larger C-terminal domain whose structure we report here and which we refer to as its *endodomain*. The ASTN-2 endodomain is exactly equivalent to the cell surface-exposed ectodomain of ASTN-1; and it is the first in a class of integral

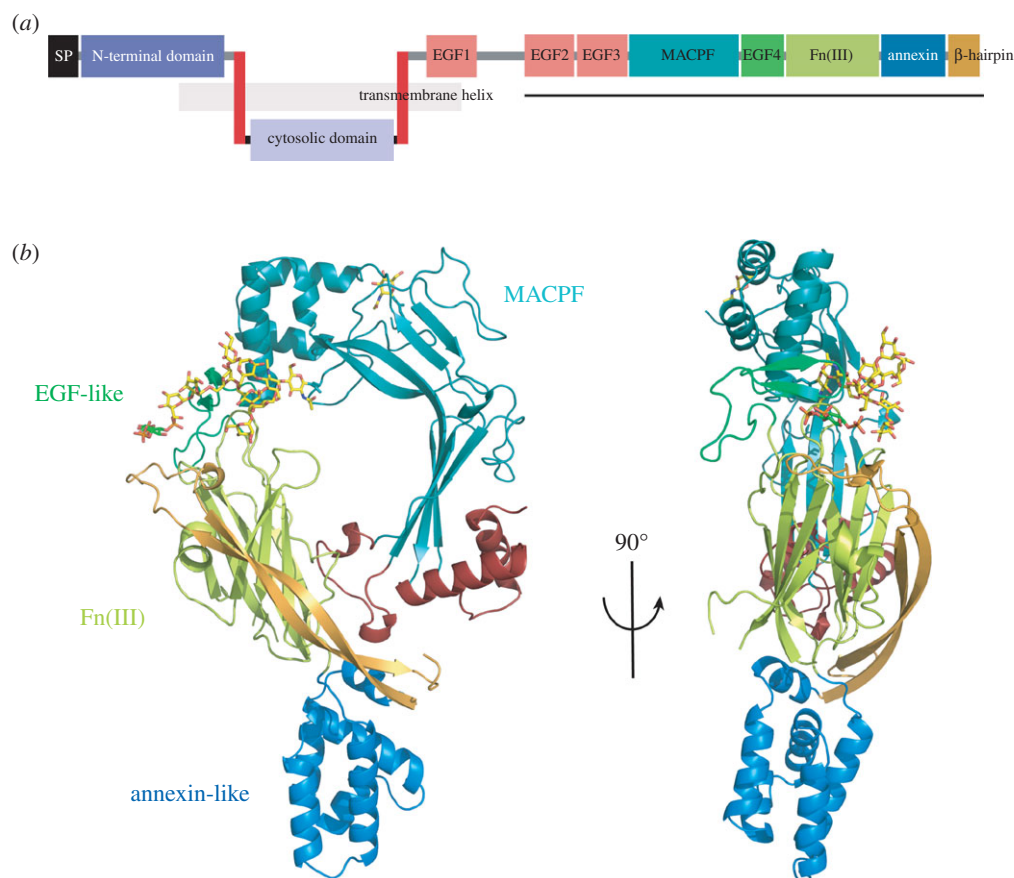


Figure 1. Overview of the structure of ASTN-2. (a) Schematic of the domain organization of ASTN-2 including its two transmembrane helices and the endodomain resolved here. (b) Two views of the ASTN-2 endodomain structure related by a 90° rotation, as shown. The different sub-domains (MACPF, EGF-like (EGF-4), Fn(III) and annexin-like) are highlighted and labelled directly; loop 1 and loop 2 are shown in red and the C-terminal β -hairpin packed back across the Fn(III) domain in yellow.

membrane PLPs (which includes perforin-2) to have its structure determined, though perforin-2 differs from ASTN-2 (and ASTN-1) in having only a single transmembrane helix [33,58].

In describing here the structure of the endodomain of ASTN-2, we provide a framework for understanding its role in cell migration and tissue development. We describe its unique combination of a canonical MACPF/CDC domain with a minimal epidermal growth factor (EGF)-like repeat (EGF-4), a previously unobserved type of fibronectin type III (Fn(III)) domain in which an additional two β -strands are folded across the core, and an annexin-like domain. We have determined structures of the endodomain at pH 7.5 and pH 5 and of the MACPF/CDC domain alone at pH 4, and find the regions of ASTN-2 resolved to be remarkably insensitive to acidification. We have shown that the ASTN-2 endodomain binds inositol phosphates at a binding pocket located between the EGF-4 and Fn(III) domains and that the annexin-like domain binds calcium, as might be expected. ASTN-2 appears to be unique to vertebrates and its polypeptide sequence is extremely conserved, with 53% identity between jawless fishes and humans (last common ancestor, 485 Ma [59]) (electronic supplementary material, figure S2). Sequences are much more diverse among fish species than among terrestrial vertebrates, as befits their extreme antiquity as a group; in fact the sequences of land-dwelling vertebrates are remarkably constrained. It is tempting to suggest that the distinctive challenges of terrestrial life, and the fundamental role apparently played by astrotactin molecules in vertebrate nervous systems, has particularly constrained diversity on land. In any case, the isolation of astrotactins to vertebrates suggests they do play a key role specifically in the

development of the vertebrate CNS, which provided the basis for the evolution of mammals leading, ultimately, to humans.

2. Results and discussion

2.1. Overall description of the ASTN-2 endodomain structure

We expressed and purified an ASTN-2 construct running from its second EGF-like repeat to the C-terminus (underlined in figure 1a) using HEK293 cells. Transient DNA transfection with the pHLsec vector previously described [60] enabled secretion into the cell medium, whence the protein was purified (see Material and methods for a full description of protein expression and purification). Figure 1b reports the overall structure of the ASTN-2 endodomain at pH 7.5 and comprising the MACPF/CDC, EGF-4, Fn(III) and annexin-like domains. The structure was solved using single isomorphous replacement with anomalous scattering (SIRAS) with a platinum-soaked derivative (see Material and methods for structure determination; see table 1 for data collection and refinement statistics). Although we verified that the expressed protein had remained intact such that EGF repeats 2 and 3 (EGF-2 and EGF-3) were present in the crystallized protein (electronic supplementary material, figure S3), they were not resolved in the structure, presumably due to disorder, though small-angle X-ray scattering (SAXS) data indicate that the domains are folded (see below and electronic supplementary material, figure S8c). The presence of EGF-2

Table 1. Structure of the endodomain of astrotactin-2. Statistics for the highest resolution shell are shown in parentheses.

data collection	ASTN-2 ₆₀₁₋₁₂₈₈ (native)	ASTN-2 ₆₀₁₋₁₂₈₈ (Pt-derivative)	ASTN-2 ₆₄₉₋₉₈₄	ASTN-2 ₆₄₉₋₁₂₈₈
space group	P1	P1 (Pt soak)	C121	P4 ₁ 2 ₁ 2
cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	101.24, 108.19, 111.36	101.59, 106.93, 112.65	98.7, 86.6, 108.36	103.9, 103.9, 304.2
α , β , γ (°)	87.18, 84.5, 65.67	87.29, 84.68, 65.66	90, 111.74, 90	90, 90, 90
resolution (Å)	98.58–3.16	112.16–5.0	60.77–3.63	85.8–5.22
<i>R</i> _{merge}	0.093 (0.890)	0.209 (1.111)	0.188 (1.547)	0.12 (3.047)
<i>R</i> _{pim}	0.036 (0.602)	0.046 (0.323)	0.026 (0.705)	0.027 (0.643)
<i>I</i> / σ	17.1(1.4)	17.1 (5.5)	19.3 (1.2)	17.3 (2.0)
completeness (%)	99.9(99.3)	99.9 (99.8)	99.4 (98.4)	100 (100)
redundancy	8.5 (3.4)	26.3 (25.5)	49.3 (6.8)	23.8 (21.4)
CC _{half}	0.998 (0.433)	0.969 (0.936)	0.999 (0.517)	0.999 (0.798)
refinement	ASTN-2 ₆₀₁₋₁₂₈₈ (native)		ASTN-2 ₆₄₉₋₉₈₄ pH 4	ASTN-2 ₆₄₉₋₁₂₈₈ pH 5
space group	P1		C121	P4 ₁ 2 ₁ 2
resolution (Å)	98.58–3.16		60.77–3.63	85.8–5.22
no. reflections	627 041 (18 219)		475 699 (4647)	164 862 (10 473)
unique reflections	73 362 (5382)		9653 (688)	6915 (490)
<i>R</i> _{work} / <i>R</i> _{free} (%)	23.3 (25.9)		27.7 (29.9)	32.8 (35.9)
no. atoms				
protein	17 842		4034	4456
ligand/ion	568		56	154
average <i>B</i> -factors				
protein	135.9		196	335
ligand/ion	224.1		217	348
RMSDs				
bond length (Å)	0.008		0.003	0.005
bond angles (°)	1.18		0.67	0.85

and EGF-3 in the protein that crystallized combined with their invisibility in the crystal structure indicates that they are not important in the folding of the regions of ASTN-2 we do resolve. This is supported by normal modes analysis of the ASTN-2 endodomain which suggests that it constitutes a remarkably tightly folded unit, and by the equivalent interaction of ASTN-2 with inositol phosphates with and without these domains (see below and electronic supplementary material, figure S10). In normal modes analysis, the first six modes are rigid-body translational and rotational modes in three dimensions, so the first non-trivial mode is mode 7. Normal modes 7–11 indicate only breathing motions within the endodomain, and highlight the combination of the MACPF and Fn(III) as a particularly robust region of the structure (electronic supplementary material, figure S4a). We also solved the structure of the endodomain at pH 5 and of the MACPF/CDC domain alone at pH 4 (electronic supplementary material, figure S4b); these structures are compared with the neutral pH endodomain structure (figure 1b) below.

2.2. MACPF and EGF-like domain structures

The N-terminus of the endodomain structure is formed by the canonical MACPF domain consisting of a central,

broken, antiparallel four-stranded β -sheet that possesses a distinctive central approximately 90° bend. Figure 2 shows a structure-based phylogenetic comparison of the ASTN-2 MACPF domain and other known equivalent structures from the MACPF/CDC superfamily. The most distinctive features of MACPF/CDC domains are two loops containing α -helices which link the core domain β -strands and which undergo refolding in pore-forming family members to generate a partial or complete β -barrel inserted in the membrane [16,61–63] (figure 2). In ASTN-2, the first such loop (loop 1, linking strands β 1– β 3 and β 4– β 5), comprises 19 residues and contains a single-turn 3_{10} -helix (η 1) and which packs against the Fn(III) domain interface (see below; see electronic supplementary material, figure S5 for sequence alignment and secondary structure). The equivalent region in perforin-1 contains 55 residues and two substantial α -helices; and in perfringolysin (and all CDCs) approximately 33 residues and three short multi-turn helices (electronic supplementary material, figure S6a). The loop 1 region of the ASTN-2 structure is therefore an outlier with respect to MACPF/CDC proteins as a whole, due to its brevity. The second loop (loop 2; linking strands β 6– β 7 and β 8– β 9) consists of 54 residues and contains two substantial α -helices (α 3 and α 4). Loop 2 is solvent exposed in contrast with the packing of

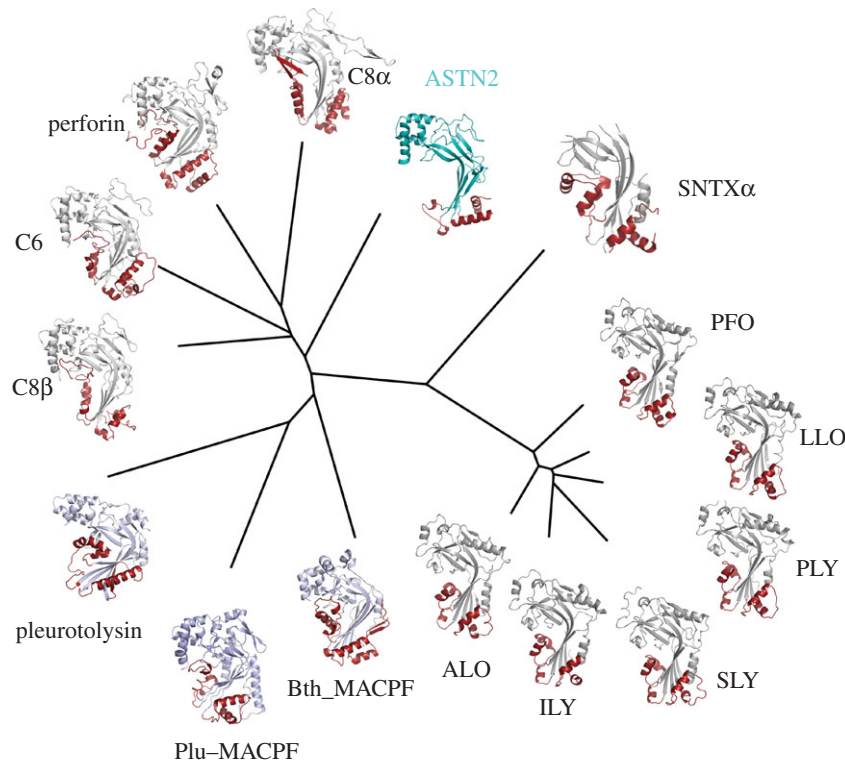


Figure 2. Structural phylogeny of known MACPF domain structures. The ASTN-2 MACPF domain core bent sheet is coloured cyan, with loop 1 and loop 2 in red; in all other cases the core sheet is coloured grey. From ASTN-2 clockwise the proteins shown are as follows: stonustoxin (SNTX α) from stonefish (PDB ID 4WVM); perfringolysin (PFO) from *Clostridium perfringens* (1PFO); listeriolysin (LLO) from *Listeria monocytogenes* (4CDB); pneumolysin (PLY) from *Streptococcus pneumoniae* (4QQA); suilysin (SLY) from *Streptococcus suis* (3HVN); intermedilysin (ILY) from *Streptococcus intermedius* (1S3R); anthrolysin (ALO) from *Bacillus anthracis* (3CQF); *Bacteroides thetaiotamicron* MACPF (Bth_MACPF) protein (3KK7); *Photobacterium luminescens* MACPF (Plu_MACPF) protein (2QP2); pleurotolysin from fungus *Pleurotus ostreatus*; complement C8 β from human (30JY); complement C6 from human (3T50); perforin from mouse (3NSJ) and complement C8 α from human (30JY).

loop 1 against the Fn(III) domain. The equivalent region to ASTN-2 loop 2 in perforin-1 contains 59 residues (four α -helices, two short, two long) and in perfringolysin 30 residues (three helices). Thus, while loop 1 of the ASTN-2 MACPF/CDC domain is unusually short, loop 2 is similar in length to that of perforin-1 and quite a bit longer than that of the CDCs. Most critically, however, the two MACPF domain loops are mismatched in length.

All MACPF/CDC proteins that form pores have to date been shown to do so using a β -stranded barrel in which each subunit contributes two β -strand hairpins to generate a structure similar to a bacterial outer membrane porin, but formed by multiple subunits rather than a single polypeptide chain [3,61]. This requires the two β -hairpins to be not significantly different in length to each other and that they are long enough to span a bilayer membrane. Two factors suggest that ASTN-2 loops 1 and 2 may not function in pore formation or that, if they do, the mechanism of pore formation by ASTN-2 differs significantly from that of the pore-forming members of the MACPF/CDC superfamily (e.g. perforin-1, C8 α , C6, pleurotolysin, perfringolysin, etc.; figure 2).

Firstly, the shortened ASTN-2 loop 1, packed against the Fn(III) repeat, is too short to span a bilayer within a β -barrel structure such as the CDCs, perforin and some PLPs are known to form [16,17,61,63,64]. Furthermore, while both perforin-1 and perfringolysin show an amphipathic sequence pattern in their respective loops 1 and 2, in ASTN-2 this pattern is lost (electronic supplementary material, figure S6a). This is most striking in loop 1, but even loop 2 is more hydrophobic than the equivalent regions of perforin-1 and perfringolysin, and rather than a hydrophobic–hydrophilic pattern alternating

every residue, the alteration seems to be more two-by-two. If loop 2 inserts into a membrane to form a pore this suggests it could do so as a helical structure rather than in the form of a β -hairpin—but whether ASTN-2 really forms membrane pores is unknown.

Secondly, while the exposed, unabbreviated ASTN-2 loop 2 is similar in length to the equivalent region of perforin-1, its large mismatch in length with loop 1 (ASTN-2 loop 1 is more than 30 residues shorter than loop 2) means that they could not pair in β -barrel formation. Yet for each other member of the MACPF/CDC superfamily, the two equivalent loops are similar in length to each other; this mismatch also suggests that ASTN-2 may not be a pore-forming protein, at least not in the way other MACPF/CD protein family members are.

The MACPF domain is completed by a single-turn 3_{10} helical turn (η 2) and a final α -helix (α 5; four turns) which are paired structurally with α 1 and α 2 to form a head region from which the MACPF domain β -strands extend. This head is linked to the Fn(III) domain by the EGF-4 domain module, which contains β 10 and β 11 and matches best to the integrin β ₃-subunit PSI domain EGF-like folds according to the SCOP database (family g.3.11.6) (electronic supplementary material, figure S6b); the two disulfide bridges in the EGF-4 domain superpose perfectly with those in the integrin β ₃ subunit PSI domain [65]. EGF repeats are also found acting as linker domains in perforin-1 itself and in complement C8 α and C8 β , at points which constitute hinges during the pre-pore to pore transition. To provide for comparison of the EGF repeats found in perforin, other MACPF proteins and other proteins identified as having similarly structured regions to the ASTN-2 EGF-4 repeat, we performed a structural phylogenetic analysis (electronic

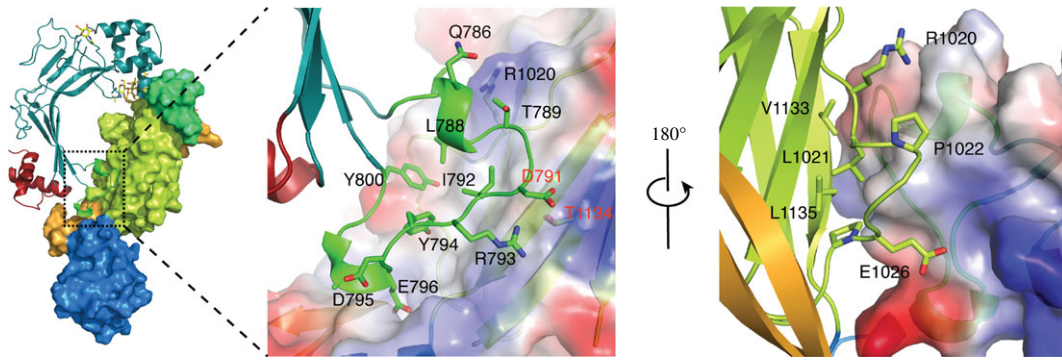


Figure 3. Close-up of the MACPF domain–Fn(III) domain interface. On the *left* is an overview in which the MACPF domain is shown in ribbon format and the EGF-4, Fn(III) and annexin-like domains with rendered surfaces. The domains are coloured as in figure 1. In the *centre* is a close-up view of the MACPF domain–Fn(III) interface in the same orientation with the Fn(III) surface rendered by charge from -5 kT/e (red) to $+5$ (blue). On the *right* is an equivalent view but rotated by 180° as shown, bringing the C-terminal β -hairpin packing across the Fn(III) domain into view.

supplementary material, figure S6*b*). The EGF-like repeats compared were those from perforin-1, C8 α and C8 β ; the PSI domain of integrin subunit β_3 ; and an EGF repeat from P-selectin as well as EGF-4 from ASTN-2. As shown, it seems that the EGF repeat found in perforin-1 itself and EGF-4 in ASTN-2 share a common line of ancestry, separate from the EGF repeats found in other immune system MACPFs (C8 α and C8 β). ASTN-2 also seems to be closer to the common point of origin of itself and perforin-1 and thus perhaps to represent a more ancient structural form; indeed the perforin-1 EGF repeat is rather an extreme structure compared with ASTN-2 EGF-4 and the EGF repeats from C8 α and C8 β .

The junction between the EGF-4 and Fn(III) domains contains density which is not part of the ASTN-2 polypeptide and which can plausibly be modelled as an inositol triphosphate bound within a surface cavity. The binding mode of this species and mutational and binding data which support its identification with a phosphorylated inositol are described below.

2.3. Structures at pH 5 and pH 4

The movement of ASTN-2 through the endolysosomal system suggests that its activity might be regulated by pH. In an attempt to investigate the effect of reduced pH on the ASTN-2 structure, we obtained crystals of constructs consisting of EGF-3 to the C-terminus (residues 649–1288) at pH 5 and of EGF-3 plus the MACPF domain at pH 4 (residues 649–984), though in neither case was EGF-3 resolved in the resulting structures. The conformations of the lower pH forms of ASTN-2 are very similar to the equivalent regions of the neutral pH structure (electronic supplementary material, figure S4*b,c*) (RMSDs 0.23 Å (pH 5) and 1.52 Å (pH 4)). The region of greatest variation at pH 4 is loop 1 between β_3 and β_4 , where 3_{10} helix η_1 shifts approximately 2 Å towards the Fn(III) domain on acidification. Yet overall these regions of ASTN-2 seem remarkably insensitive to pH suggesting that, if pH is a triggering factor in ASTN-2 activity as it moves through the endolysosomal system, the sensor is not in these parts of the structure.

2.4. Fibronectin type III domain structure

The Fn(III) domain (β_{12} – β_{15} , π_1 , β_{16} – β_{19} with the start of β_{16} and the end of β_{19} tied together by a disulfide bond) follows directly on from EGF-4 and has the unexpected feature

that it is complemented by a 54-residue β -hairpin from the ASTN-2 C-terminus (β_{20} – β_{21}) that folds across the Fn(III) β -sandwich, clasping the edges of both sheets. As such it seems to lock the C-terminus back to the centre of the endo-domain, rigidifying the structure (electronic supplementary material, figure S4*a*). The loop between the hairpin strands projects towards EGF-4 and towards the inositol phosphate binding site, suggesting that this point in the structure may be a key point in its conformational regulation.

The MACPF domain contains two glycosylation sites; one at the top of the domain (N719) was trimmed back by EndoF1 (see Material and methods) to a single *N*-acetyl galactosamine residue. The other (N732) is occupied by a high-mannose glycan sidechain protected from glucosidase trimming that appears to assist in the stabilization of the MACPF/EGF-4/Fn(III) domains cassette (figure 1*b*; electronic supplementary material, figure S7). The principal contacts made by this sugar sidechain with EGF-4 and the Fn(III) surfaces are formed by the two *N*-acetyl galactosamine moieties and terminal mannoses of the glycan sidechain branches. The role of the glycan sidechain in stabilizing the MACPF/EGF-4/Fn(III) domains cassette is indicated by the effects of a N > Q mutation knocking out its sequon. Such a construct is not secreted during mammalian cell expression, suggesting it is misfolded.

2.5. Details of the interface between the MACPF and Fn(III) domains

The interface between the MACPF and Fn(III) domains is formed by the truncated loop 1 between strands β_3 and β_4 of the MACPF domain (figure 1*b*). The residues involved in the interactions are mostly charged (e.g. Arg793, Asp791, Glu796) though Phe800 projects towards the domain interface (figure 3). Despite the involvement of charged sidechains there do not seem to be direct electrostatic contacts formed between the domain surfaces, rather the presenting faces of the MACPF and the Fn(III) domains have a shape complementarity (figures 1*b* and 3). This suggests the possibility of breathing movements (supported by normal modes 8 and 9; electronic supplementary material, figure S4*a*) and even of a significant conformational change caused by movement of the MACPF away from the Fn(III) domain which might be enabled if the C-terminal β -hairpin ungrasps the Fn(III) domain. In an attempt to probe the functional significance of the MACPF/Fn(III) interface, we introduced a disulfide

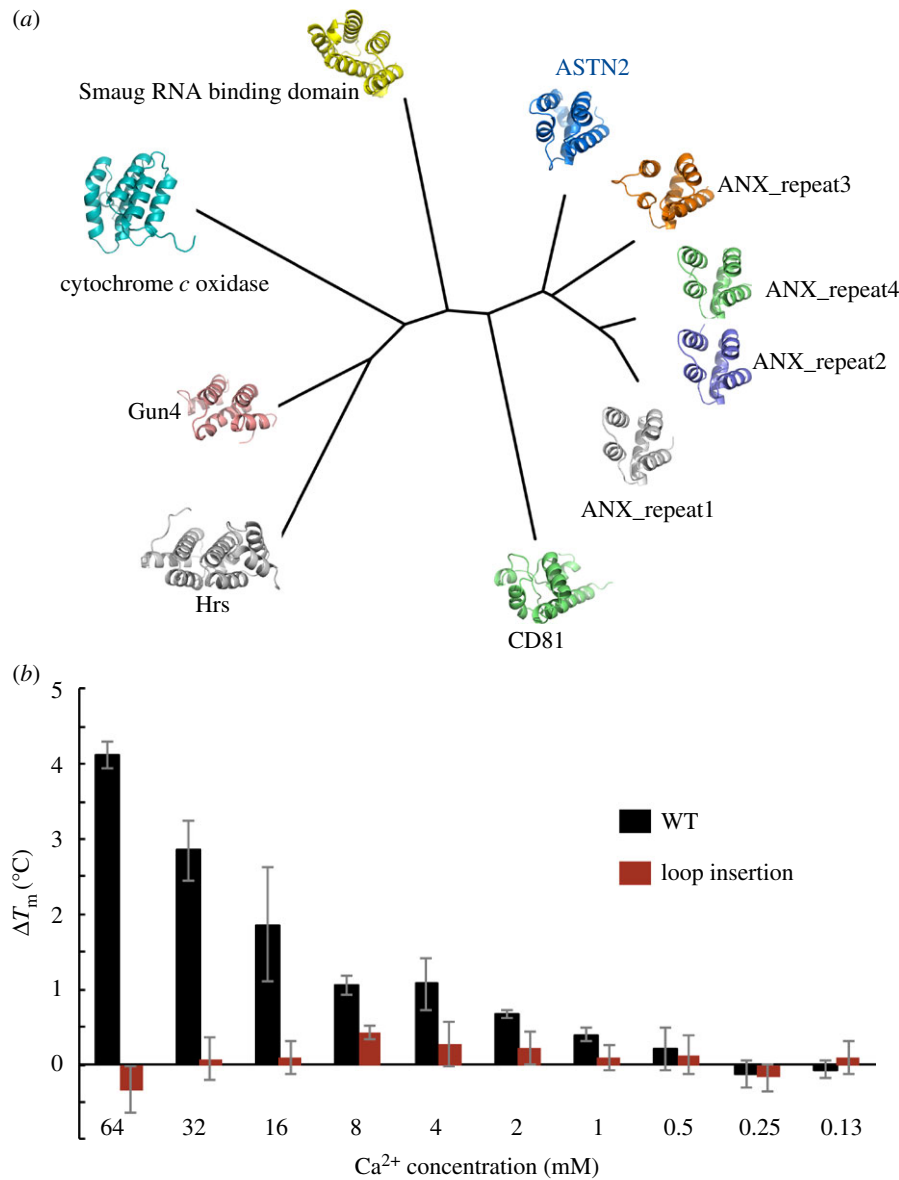


Figure 4. Annexin-like domain homology and calcium binding. (a) A structural phylogeny, constructed as described in the Material and methods, comparing the annexin-like domain of ASTN-2 with close structural homologues. Clockwise from the ASTN-2 domain as labelled, the structures shown are the repeats 3, 4, 2 and 1 from human annexin V (ANX, PDB ID 1AVH); the extracellular domain from human CD81 (1G8Q); a domain from Hrs1 from *Drosophila melanogaster* (1DVP); from Gun4 of *Thermosynechococcus elongates* (1Z3X); from bovine cytochrome c oxidase (10CR); and from Smaug RNA-binding protein from *Drosophila melanogaster* (10XJ). (b) Melting temperature shifts for the ASTN2₆₀₁₋₁₂₈₈, both as wild-type (WT) sequence and with the calcium-binding motif disrupted (loop inserted). Experiments were performed in triplicate and the mean \pm s.d. is reported.

bond to lock the two together (Asp791Cys and Thr1134Cys mutations); this resulted in a form of the protein which was secreted poorly from HEK293T cells, although SAXS data suggest that the protein that is secreted is folded like wild-type ASTN-2 (electronic supplementary material, figure S8).

Structural phylogenetic analysis of the Fn(III) domain excluding the long C-terminal β -hairpin (β 20– β 21) indicates that this domain is closest in structure to Fn(III) repeat 8 from human fibronectin itself, and occupies a single branch with this domain and a Fn(III) domain from tenascin (electronic supplementary material, figure S6c). It is certainly therefore a bona fide Fn(III) domain despite its unique complementation with the β 20– β 21 hairpin, for which there are no currently described structural parallels.

2.6. The annexin-like domain

The annexin-like domain is positioned between the Fn(III) domain and the long β -hairpin folded across it (β 20– β 21)

but, functionally, forms the tip of the ASTN-2 endodomain. An annexin-like domain is an unexpected feature of the ASTN-2 structure (α 6– α 11) which despite undetectable sequence homology with annexin itself superimposes very well on an annexin single repeat (figure 4a; see electronic supplementary material, figure S9 for a superimposition). Annexin itself consists of four tandem repeats of a common fold [66] and it is striking that the ASTN-2 annexin-like domain is closer in structure to human annexin repeat 3 (RMSD 1.45 Å) than human annexin repeat 3 is to repeat 1 (RMSD 1.58 Å). Thus, despite no sequence homology ASTN-2 and human annexin have very similar structures; their charge distributions, however, are very different (electronic supplementary material, figure S9b). For example, with the annexin V domains and the ASTN-2 annexin-like domain all oriented equivalently, the most-similar annexin V domain 3 presents a highly positively charged face whereas ASTN-2 presents a mixed face with regions of positive and negative charge.

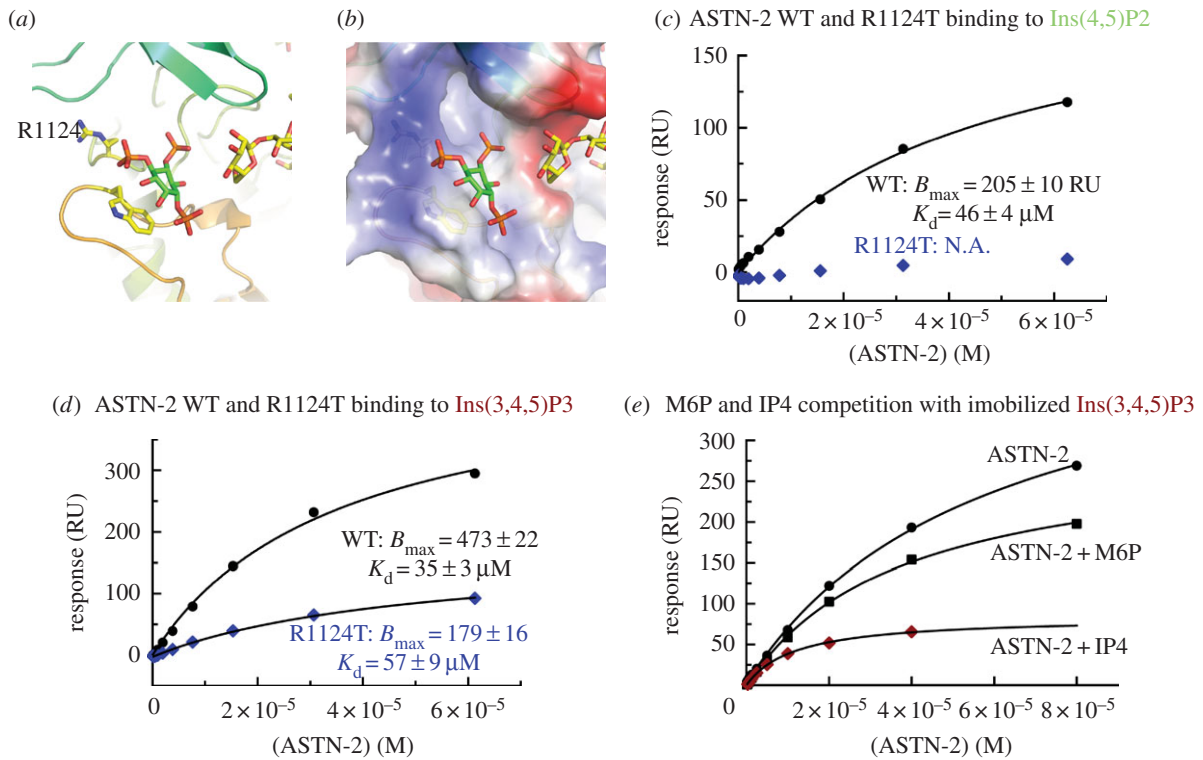


Figure 5. Inositol phosphate binding by ASTN-2. (a) Close-up of the inositol triphosphate modelled into extra density, one phosphate in electrostatic interaction with arginine 1124, a position occupied by a threonine in ASTN-1 and the inositol ring π -stacked with tryptophan 1259. (b) Surface charge representation of the inositol triphosphate binding pocket. (c) Wild-type (WT) ASTN-2 and an R1124T mutant interacting with immobilized inositol diphosphates, as measured by surface plasmon resonance (SPR). (d) Wild-type ASTN-2 and an R1124T mutant binding to immobilized inositol triphosphate, again measured by SPR. (e) A competition assay for the binding of ASTN-2 to the triphosphorylated inoside alone and with competition from $\text{Ins}(1,3,4,5)\text{P}_4$ (IP4) and mannose-6-phosphate (M6P) in solution. The affinity of ASTN-2 for the immobilized $\text{Ins}(3,4,5)\text{P}_3$ is much less affected by the presence of mannose-6-phosphate in solution than by the presence of free $\text{Ins}(1,3,4,5)\text{P}_4$.

Like annexin itself, the ASTN-2 annexin-like domain binds calcium, resulting in a stabilization of its structure as evidenced by an increased temperature of melting (figure 4b). This effect is confirmed by mutation of the putative calcium-binding residues within the annexin-like domain, which knock out the stabilizing effect of the presence of calcium (figure 4b). Annexin-like domains are known for their capacity to remodel membranes, triggered by calcium binding, and have also been suggested to be involved in the formation of pores in membranes [67]—both are possible biological roles of the ASTN-2 annexin-like domain. The MACPF/EGF-4/Fn(III) domains cassette is further stabilized by a disulfide bridge that caps the whole annexin-like fold, covalently linking its start (a linker region just before $\alpha 6$) to its end (a linker region in $\alpha 11$ which leads into $\beta 20$).

2.7. Inositol phosphate binding site

As mentioned above, the crystal structure contains density consistent with an inositol triphosphate species being bound in a shallow pocket formed by loops from EGF-4 and the β -hairpin slung across the Fn(III) domain. Initially, we identified difference density of unknown origin in this area which could be modelled as an inositol triphosphate, but faced a need to demonstrate its identity. We were assisted here by the observation that although ASTN-2 interacts with $\text{Ins}(3,4,5)\text{P}_3$ and $\text{Ins}(4,5)\text{P}_2$ in a surface plasmon resonance (SPR)-based experiment ($K_d = 30$ μ M) (figure 5 and electronic supplementary material, figure S10a), ASTN-1 does not (electronic supplementary material, figure S10b). A comparison of the sequences of ASTN-2 and ASTN-1 in this

region indicated that an arginine apparently interacting with $\text{Ins}(3,4,5)\text{P}_3$ in the ASTN-2 crystal structure is replaced by a threonine in ASTN-1. Mutation of ASTN-2 Arg1124 to Thr resulted in a form of the protein which, like ASTN-1, did not bind $\text{Ins}(4,5)\text{P}_2$ and had much-reduced affinity for $\text{Ins}(3,4,5)\text{P}_3$ (figure 5). It is possible that the binding of inositol phosphates is non-specific, for example that ASTN-2 binds specifically to mannose-6-phosphate instead. We therefore performed a competition assay (figure 5c and electronic supplementary material, figure S10c) to demonstrate that ASTN-2 makes a specific choice of $\text{Ins}(3,4,5)\text{P}_3$ over mannose-6-phosphate in our hands. We do not, however, know the biological significance of this interaction—which could be either that ASTN-2 binds phosphorylated phosphatidylinositol lipid species, or that it binds to inositol phosphate second messengers, or to another phosphorylated or similarly modified carbon ring system. The location of the binding pocket would permit either or both kinds of phosphorylated inositol species to be bound, though we are not aware of free inositol phosphates being found in either the extracellular or endosomal compartments.

3. Conclusion

Here we have described the structure of a major portion of ASTN-2, the first structure from an integral membrane MACPF/CDC domain-containing protein. ASTN-2 is closely related to not only ASTN-1 but also the BRINP proteins which, similarly, function in neural cell migration and guidance. Other integral membrane MACPF/CDC proteins

include perforin-2, a macrophage-specific form of perforin involved in bacterial killing [33,58], and torso-like, which plays a role in early development of the *Drosophila* embryo [43,68]. The structure we describe is of relevance for understanding ways in which both closely related proteins (ASTN-1 and the BRINPs) and also those more distantly related might function biologically. For example, we have shown that the MACPF/CDC domain of ASTN-2 does not equip it to form pores by a mechanism similar to that of perforin-1 and the complement MAC, or bacterial members of the family [3]. We have also shown how additional domains associated with the MACPF/CDC homologous region of ASTN-2 contribute to its structure and the basis on which it might act, by binding calcium (the annexin-like domain) and phosphorylated inositol species. Our structure reveals how different functions, such as membrane binding via an annexin-like domain, have been co-opted to enable the functioning of a modified MACPF/CDC domain in a novel context, as well as currently uniquely described mechanisms for bracing a protein structure by the folding of a long hairpin across the cleft of a Fn(III) domain. The stage is now set for further experiments investigating the details of ASTN-2 functioning in neuronal and other cells, including its role in PCP determination, and also for the determination of structures for full-length ASTN-2, for ASTN-1 and the BRINPs which, like ASTN-2, have modified MACPF/CDC domains seemingly lacking a basis on which to form pores such as other MACPF/CDC proteins form. Without an easy analogy to pore formation by perforin/MAC or the CDCs it is impossible to say why a MACPF/CDC domain is of advantage for the activity of ASTNs and BRINPs, but our structural studies provide an essential stepping stone along the way to an answer. One possibility is that ASTN-2 has a membrane fusion-type activity [14].

4. Material and methods

4.1. Design of constructs, cloning, mutagenesis and sequence analysis

The full-length cDNA clone of Human ASTN-2 (UniprotKB/Swiss-prot O75129, isoform 2) was obtained from Source Bioscience (UK) and the expression constructs reported here are all based on the pHLsec vector [60]. For production of the larger portion of ASTN-2's C-terminal endodomain (residues 601–1288), the construct was cloned in-frame with a Human Rhinovirus (HRV)-3C protease cleavage site, followed by a monoVenus and 8xhis tag. Other constructs (residues 701–1288 and residues 701–984) used in this study were all also cloned in-frame with hexahistidine. Site-directed mutagenesis was introduced by overlapping PCR [69]. Construct sequences were verified by DNA sequencing (Source Bioscience). Sequence alignments were performed using CLUSTAL OMEGA and displayed using ESPRIPT 3 [70] (electronic supplementary material, figure S5). Conservation of sequence was calculated using CONSURF [71] for plotting on the surface of the ASTN-2 structure (electronic supplementary material, figure S9c).

4.2. Protein expression and purification

The expression constructs were transiently transfected into HEK293T cells using polyethylenimine (PEI) as described

before [60]. All the recombinant proteins used in this study were expressed, secreted from HEK293T cells and purified from the media. For crystallization, the proteins were produced in the presence of 5 μ M of the class I α -mannosidase inhibitor, kifunensin, explaining the mannose-rich (unedited) glycan [72]. For all other experiments, the proteins were produced without kifunensin.

The media containing the secreted recombinant proteins were harvested 5–6 days after transfection, then filtered through 0.22 μ m filters and dialysed against phosphate-buffered saline (pH 7.4) overnight before loading onto a pre-equilibrated HisTrap HP column (GE Healthcare, 5 ml) in a cold room. The column with proteins bound was washed with 10 column volumes of Tris buffer (20 mM Tris–HCl pH 7.5 500 mM NaCl) before being eluted with 10–500 mM imidazole with a gradient concentration in the same Tris buffer. GST-3C was added into the eluates to remove the C-terminal monoVenus-8xhis where necessary. The glycan chains of proteins for crystallization were trimmed using Endo-F1. Subsequently, the eluates were buffer-exchanged into low salt buffer (20 mM Tris–HCl, pH 8.5, 50 mM NaCl) and loaded onto a pre-equilibrated Q column (GE Healthcare, 5 ml) and eluted with the same buffer containing a higher concentration of NaCl. This step efficiently removed albumin contamination. Finally, proteins eluted from the Q column were applied to a size-exclusion chromatography column (Superdex 200 16/600, GE Healthcare) and homogeneous proteins were pooled together, concentrated to about 10 mg ml⁻¹, flash-frozen in liquid nitrogen and stored at –80°C until further use.

4.3. Crystallization, data collection and processing

Crystallization screening was carried out using commercially available crystallization reagents by vapour diffusion methods in sitting drop format in 96-well plates. For crystallization of ASTN₂₆₀₁₋₁₂₈₈, 200 nl sitting drops of protein solutions (12 mg ml⁻¹ in 10 mM HEPES pH 7.5, 150 mM NaCl) were mixed with 100 nl of precipitant (12% PEG 20000, 150 mM KSCN, 0.1 M Bis-tris propane pH 7.5–8.5) and equilibrated against 90 μ l of crystallization condition. The small needle-like crystals usually appeared after a week and microseeding with the needle-like crystals yielded reasonably-sized crystals [73]. For crystallization of ASTN-₂₆₄₉₋₉₈₄, the protein was concentrated to 15 mg ml⁻¹ and crystallized in 0.1 M citric buffer, pH 4.0–4.5, 20% PEG 1000. ASTN-₂₆₄₉₋₁₂₈₈ was crystallized in 0.1 M MES, pH 5.0–6.0, 5% PEG 6000. Platinum derivative crystals of ASTN₂₆₀₁₋₁₂₈₈ were prepared by soaking the crystals with concentrated PtCl₄ for 1 h; 25% (vol/vol) glycerol prepared in mother liquor was used as cryoprotectant for both native and platinum-soaked crystals, which were subsequently flash-frozen in liquid nitrogen. Diffraction data were collected at 100 K at Diamond Light Source (Didcot, UK) on beamlines I02, I03, I04, I04-1 and I24. Although most of the crystals diffracted poorly with a high-resolution limit up to 3.6 Å, a native dataset with higher resolution (3.16 Å) was eventually achieved by screening a number of crystals. An inverse-beam data collection strategy was employed to accurately record the anomalous signal from platinum derivative crystals. Diffraction data were indexed, integrated and scaled using the Xia2 pipeline [74].

4.4. Structure determination and refinement

The structure of ASTN-2 endodomain was solved by single isomorphous replacement with anomalous scattering (SIRAS). A high-redundancy platinum derivative dataset was achieved by combining 21 datasets selected from seven crystals. The Pt sites were initially identified using HKL2MAP [75] with SHELXC [76], which were subsequently fed into PHENIX.autosol [77]. The resolution cut-off was set to 6.5 Å and solvent content to 0.6 to obtain initial phases. Density modification with non-crystallographic symmetry (NCS) averaging and phase extension to the full resolution of the native dataset resulted in an interpretable map by RESOLVE [78]. A polyalanine model was initially built into the electron density map manually in COOT [79], followed by refinement in PHENIX.refine [77] with external (Pt-SAD) phase restraints. The electron density of amino acid residue side chains appeared after two cycles of model building and refinement. Further density modification in RESOLVE with NCS calculated from the polyalanine model resulted in an excellent electron density map, which could then be fed into PHENIX.autobuild. The model was completed by manual rebuilding in COOT and refinement in REFMAC5 and PHENIX. The crystal structure of the MACPF domain C-terminus at pH 5 and of the MACPF domain alone at pH 4 were solved by molecular replacement using copy A of the structure solved at pH 7.5 in PHASER and refinement was carried out in PHENIX [77]. Surface electrostatic potentials for ASTN-2 and other proteins were calculated using APBS [80]. The crystallographic statistics are listed in table 1 and all models were validated with MOLPROBITY software [81]. Data collection and structure determination will be described in detail elsewhere.

4.5. Surface plasmon resonance

The SPR experiments were performed using a Biacore T200 machine (GE Healthcare Life Sciences) at 20°C in 10 mM HEPES, pH 7.5, 150 mM NaCl, PtdIns(3,5)P₂ (C-35B6a), PtdIns(4,5)P₂ (C-45B6a), PtdIns(3,4,5)P₃ (C-39B6a) and Ins(1,3,4,5)P₄ (Q-1345) were purchased from Echelon Biosciences. To immobilize the biotinylated inositol phosphate onto the sensor chip, a BIAcore CM5 chip (GE Healthcare Life Sciences) was first derivatized with streptavidin following the manufacturer's instructions, and inositol phosphates were then injected on to channels 2 and 4 of the biosensor surface, leaving channels 1 and 3 as empty controls. The analyte with twofold serial dilutions was applied at a flow rate of 20 µl min⁻¹ for 180 s followed by 300 s of dissociation time. The biosensor chip was regenerated by 0.1% SDS after each running cycle. To perform the competition assay, the analyte (ASTN-2₇₀₁₋₁₂₈₈) was incubated with a 10-fold molar concentration of Ins(1,3,4,5)P₄ and mannose-6-phosphate for 45 min, respectively, before being serially diluted in running buffer. The data were fit with the 1:1 Langmuir adsorption model ($B = B_{\max}C/(K_d + C)$, where B is the response of bound analyte and C is the concentration of the analyte in the sample) to calculate the dissociation constant (K_d) using BIAcore BIAANALYSIS software.

4.6. Thermofluor

Thermofluor experiments were conducted using a real-time PCR machine [82]. Identical protein samples (4 µg) with

SYPRO orange stain (Life Technologies S-6650, final concentration 3×) were mixed with serial concentrations of CaCl₂ (0.13–64 µM) in 10 mM HEPES, pH 7.5, 150 mM NaCl. Fluorescence measurements were recorded from 298 to 372 K with a 1 K temperature increase each cycle. The melting temperature T_m (the midpoint of the unfolding transition) was calculated from the melting curve.

4.7. Small-angle X-ray scattering

SAXS data were collected on B21 at the Diamond Light Source (Didcot, UK). The measurements were carried out at 293 K in 10 mM HEPES, pH 7.5, 150 mM NaCl buffer with a momentum transfer range of $0.004 \text{ \AA}^{-1} < q < 0.45 \text{ \AA}^{-1}$, where $q = 4\pi\sin(\theta)/\lambda$ and 2θ is the scattering angle. SAXS datasets for WT ASTN-2 were collected at four concentrations: 4.92 mg ml⁻¹, 2.49 mg ml⁻¹, 1.27 mg ml⁻¹ and 1 mg ml⁻¹, and the disulfide bond locking mutant was measured at 1 mg ml⁻¹. Eighteen frames of measurements were recorded and the frames without radiation damage were averaged. The scattering intensity from buffer alone was subtracted from the averaged data to obtain the protein scattering in solution. Data reduction was carried out using the ATSAS package [83]. For wild-type ASTN-2, a merged dataset was obtained by combining the low-angle part of the low-concentration dataset with the high-angle part of the high-concentration dataset. The radius of gyration (R_g) was calculated from a Guinier plot using AUTORG [83]. Particle distance distribution function $P(r)$ was calculated in GNOM [84] using the low-resolution range of the dataset ($0.01 \text{ \AA}^{-1} < q < 0.15 \text{ \AA}^{-1}$). An *ab initio* model was then calculated from the $P(r)$ function using DAMMIF and DAMMIN [85]. EGF-2 and EGF-3 domain structures were predicted using PHYRE2 [86] using tandem EGF domains of Del-1 as a threading model (PDB code: 4D90 [87]). The crystal structure of ASTN-2 was then fitted manually in PUMOL.

4.8. Analytical ultracentrifugation

Sedimentation velocity experiments were performed using a Beckman Optima XL-I analytical ultracentrifuge equipped with absorbance and interference optics. Double-sector 3 mm centrepieces were used with protein samples at 1 mg ml⁻¹ and using absorbance measurements at 280 nm. Experiments were performed at 20°C, taking sample distribution scans every 6 min. Data were analysed using SEDFIT software using the $c(s,f/f_0)$ method of interpretation to generate sample distributions in s (sedimentation coefficient) without assuming a particular number of species [88]. The resulting distributions were curve fit in PROFIT (QuantumSoft, Uetikon am See, CH).

4.9. Structural phylogeny calculation

Superimposition of homologous protein structures was performed using SHP [89] as previously reported [90]. The phylogenetic tree was calculated using a pairwise evolutionary distance matrix determined from the superimposed domains. The tree representation was generated using the programs FITCH and DRAWTREE as part of the PHYLIP package [91].

4.10. Normal modes analysis

Normal modes analysis was performed using the online server ELNÉMO [92]. The vibrational movements of the

structure were visualized by superimposition of models within each of the six lowest frequency non-trivial normal modes in CHIMERA [93].

Data accessibility. The structures reported in this paper have been deposited in the RCSB protein databank, PDB IDs 5J67 (pH 7.5), 5J68 (pH 5) and 5J69 (pH 4).

Authors' contributions. T.N. and R.G. planned the work, performed the experiments, analysed the data and wrote the paper; K.H. performed experiments and commented on the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Funding. T.N. was supported by the Chinese Scholarships Council and by the Medical Research Council (grant MR/N000331/1); K.H. was supported by the UK Medical Research Council (grant MR/N00065X/1) and the Wellcome Trust (grant no. 090532/Z/09/Z); the Division of Structural Biology is a part of the Wellcome Trust Centre for Human Genetics, Wellcome Trust Core grant no. 090532/Z/09/Z.

Acknowledgements. T.N. was a Chinese Scholarships Council-University of Oxford scholar within the Nuffield Department of Medicine. We gratefully acknowledge advice from Drs Kamel El-Omari and Jingshan Ren, Diamond Light Source beamline staff, and the support of Dr David Staunton with use of the Oxford Molecular Biophysics Suite within the university's Department of Biochemistry.

References

- Anderlueh G, Kisovec M, Krasevec N, Gilbert RJ. 2014 Distribution of MACPF/CDC proteins. *Subcell. Biochem.* **80**, 7–30. (doi:10.1007/978-94-017-8881-6_2)
- Anderlueh G, Gilbert RJC. 2014 *MACPF/CDC proteins: agents of defence, attack and invasion*. Dordrecht, The Netherlands: Springer.
- Gilbert RJC. 2015 Perforins. In *Electrophysiology of unconventional channels and pores* (ed. A Delcour), pp. 289–312. Dordrecht, The Netherlands: Springer.
- Bernheimer AW. 1947 Comparative kinetics of hemolysis induced by bacterial and other hemolysins. *J. Gen. Physiol.* **30**, 337–353. (doi:10.1085/jgp.30.4.337)
- Bhakdi S, Tranum-Jensen S, Ziegoleit A. 1985 Mechanism of membrane damage by streptolysin-O. *Infect. Immun.* **47**, 52–60.
- Gilbert RJ. 2005 Inactivation and activity of cholesterol-dependent cytolysins: what structural studies tell us. *Structure* **13**, 1097–1106. (doi:10.1016/j.str.2005.04.019)
- Tweten RK. 2005 Cholesterol-dependent cytolysins, a family of versatile pore-forming toxins. *Infect. Immun.* **73**, 6199–6209. (doi:10.1128/IAI.73.10.6199-6209.2005)
- Amiguet P, Brunner J, Tschopp J. 1985 The membrane attack complex of complement: lipid insertion of tubular and nontubular polymerized C9. *Biochemistry* **24**, 7328–7334. (doi:10.1021/bi00346a046)
- Borsos T, Dourmashkin RR, Humphrey JH. 1964 Lesions in erythrocyte membranes caused by immune haemolysis. *Nature* **202**, 251–252. (doi:10.1038/202251a0)
- Young JD, Hengartner H, Podack ER, Cohn ZA. 1986 Purification and characterization of a cytolytic pore-forming protein from granules of cloned lymphocytes with natural killer activity. *Cell* **44**, 849–859. (doi:10.1016/0092-8674(86)90007-3)
- Young LH, Joag SV, Zheng LM, Lee CP, Lee YS, Young JD. 1990 Perforin-mediated myocardial damage in acute myocarditis. *Lancet* **336**, 1019–1021. (doi:10.1016/0140-6736(90)92486-2)
- Hadders MA, Beringer DX, Gros P. 2007 Structure of C8alpha-MACPF reveals mechanism of membrane attack in complement immune defense. *Science* **317**, 1552–1554. (doi:10.1126/science.1147103)
- Rosado CJ *et al.* 2007 A common fold mediates vertebrate defense and bacterial attack. *Science* **317**, 1548–1551. (doi:10.1126/science.1144706)
- Gilbert RJ. 2016 Protein–lipid interactions and non-lamellar lipidic structures in membrane pore formation and membrane fusion. *Biochim. Biophys. Acta* **1858**, 487–499. (doi:10.1016/j.bbame.2015.11.026)
- Gilbert RJ, Jimenez JL, Chen S, Tickle IJ, Rossjohn J, Parker M, Andrew PW, Saibil HR. 1999 Two structural transitions in membrane pore formation by pneumolysin, the pore-forming toxin of *Streptococcus pneumoniae*. *Cell* **97**, 647–655. (doi:10.1016/S0092-8674(00)80775-8)
- Tilley SJ, Orlova EV, Gilbert RJ, Andrew PW, Saibil HR. 2005 Structural basis of pore formation by the bacterial toxin pneumolysin. *Cell* **121**, 247–256. (doi:10.1016/j.cell.2005.02.033)
- Law RH *et al.* 2010 The structural basis for membrane binding and pore formation by lymphocyte perforin. *Nature* **468**, 447–451. (doi:10.1038/nature09518)
- Metkar S *et al.* 2015 Perforin oligomers form arcs in cellular membranes: a locus for intracellular delivery of granzymes. *Cell Death Differ.* **22**, 74–85. (doi:10.1038/cdd.2014.110)
- Praper T, Sonnen AF, Viero G, Kladnik A, Froelich CJ, Anderlueh G, Dalla Serra M, Gilbert RJ. 2011 Human perforin employs different avenues to damage membranes. *J. Biol. Chem.* **286**, 2946–2955. (doi:10.1074/jbc.M110.169417)
- Heuck AP, Hotze EM, Tweten RK, Johnson AE. 2000 Mechanism of membrane insertion of a multimeric beta-barrel protein: perfringolysin O creates a pore using ordered and coupled conformational changes. *Mol. Cell.* **6**, 1233–1242. (doi:10.1016/S1097-2765(00)00119-2)
- Hotze EM, Wilson-Kubalek EM, Rossjohn J, Parker MW, Johnson AE, Tweten RK. 2001 Arresting pore formation of a cholesterol-dependent cytolysin by disulfide trapping synchronizes the insertion of the transmembrane beta-sheet from a prepore intermediate. *J. Biol. Chem.* **276**, 8261–8268. (doi:10.1074/jbc.M009865200)
- Tweten RK, Hotze EM, Wade KR. 2015 The unique molecular choreography of giant pore formation by the cholesterol-dependent cytolysins of Gram-positive bacteria. *Annu. Rev. Microbiol.* **69**, 323–340. (doi:10.1146/annurev-micro-091014-104233)
- Gilbert RJ. 2002 Pore-forming toxins. *Cell Mol. Life Sci.* **59**, 832–844. (doi:10.1007/s00018-002-8471-1)
- Gilbert RJ, Dalla Serra M, Froelich CJ, Wallace MI, Anderlueh G. 2014 Membrane pore formation at protein–lipid interfaces. *Trends Biochem. Sci.* **39**, 510–516. (doi:10.1016/j.tibs.2014.09.002)
- Leung C *et al.* 2014 Stepwise visualization of membrane pore formation by suliyisin, a bacterial cholesterol-dependent cytolysin. *eLife* **3**, e04247. (doi:10.7554/eLife.04247)
- Podobnik M, Marchiorretto M, Zanetti M, Bavdek A, Kisovec M, Cajnko MM, Lunelli L, Dalla Serra M, Anderlueh G. 2015 Plasticity of listeriolysin O pores and its regulation by pH and a unique histidine. *Sci. Rep.* **5**, 9623. (doi:10.1038/srep09623)
- Sonnen AF, Pitzko J, Gilbert RJC. 2014 Incomplete pneumolysin oligomers form membrane pores. *Open Biol.* **4**, 140044. (doi:10.1098/rsob.140044)
- Gilbert RJC, Sonnen AF. 2016 Measuring kinetic drivers of pneumolysin pore structure. *Eur. Biophys. J.* **45**, 365–376. (doi:10.1007/s00249-015-1106-x)
- Birmingham CL, Canadien V, Kaniuk NA, Steinberg BE, Higgins DE, Brumell JH. 2008 Listeriolysin O allows *Listeria monocytogenes* replication in macrophage vacuoles. *Nature* **451**, 350–354. (doi:10.1038/nature06479)
- Czuczman MA, Fattouh R, van Rijn JM, Canadien V, Osborne S, Muise AM, Kuchroo VK, Higgins DE, Brumell JH. 2014 *Listeria monocytogenes* exploits efferocytosis to promote cell-to-cell spread. *Nature* **509**, 230–234. (doi:10.1038/nature13168)
- Ponting CP. 1999 Chlamydial homologues of the MACPF (MAC/perforin) domain. *Curr. Biol.* **9**, R911–R913. (doi:10.1016/S0960-9822(00)80102-5)
- McCormack R, de Armas L, Shiratsuchi M, Podack ER. 2013 Killing machines: three pore-forming proteins of the immune system. *Immunol. Res.* **57**, 268–278. (doi:10.1007/s12026-013-8469-9)
- McCormack RM *et al.* 2015 Perforin-2 is essential for intracellular defense of parenchymal cells and phagocytes against pathogenic bacteria. *eLife*. **4**, e06508. (doi:10.7554/eLife.06508)
- Sonnen AF, Henneke P. 2014 Structural biology of the membrane attack complex. *Subcell. Biochem.* **80**, 83–116. (doi:10.1007/978-94-017-8881-6_6)

35. Amino R *et al.* 2008 Host cell traversal is important for progression of the malaria parasite through the dermis to the liver. *Cell Host Microbe* **3**, 88–96. (doi:10.1016/j.chom.2007.12.007)
36. Tavares J, Formaglio P, Thiberge S, Mordelet E, Van Rooijen N, Medvinsky A, Menard R, Amino R. 2013 Role of host cell traversal by the malaria sporozoite during liver infection. *J. Exp. Med.* **210**, 905–915. (doi:10.1084/jem.20121130)
37. Kafsack BF, Pena JD, Coppens I, Ravindran S, Boothroyd JC, Carruthers VB. 2009 Rapid membrane disruption by a perforin-like protein facilitates parasite exit from host cells. *Science* **323**, 530–533. (doi:10.1126/science.1165740)
38. Edmondson JC, Liem RK, Kuster JE, Hatten ME. 1988 Astrotactin: a novel neuronal cell surface antigen that mediates neuron–astroglial interactions in cerebellar microcultures. *J. Cell Biol.* **106**, 505–517. (doi:10.1083/jcb.106.2.505)
39. Wilson PM, Fryer RH, Fang Y, Hatten ME. 2010 Astn2, a novel member of the astrotactin gene family, regulates the trafficking of ASTN1 during glial-guided neuronal migration. *J. Neurosci.* **30**, 8529–8540. (doi:10.1523/JNEUROSCI.0032-10.2010)
40. Solecki DJ. 2012 Sticky situations: recent advances in control of cell adhesion during neuronal migration. *Curr. Opin. Neurobiol.* **22**, 791–798. (doi:10.1016/j.conb.2012.04.010)
41. Kawano H *et al.* 2004 Identification and characterization of novel developmentally regulated neural-specific proteins, BRINP family. *Mol. Brain Res.* **125**, 60–75. (doi:10.1016/j.molbrainres.2004.04.001)
42. Kobayashi M, Nakatani T, Koda T, Matsumoto K, Ozaki R, Mochida N, Takao K, Miyakawa T, Matsuoka I. 2014 Absence of BRINP1 in mice causes increase of hippocampal neurogenesis and behavioral alterations relevant to human psychiatric disorders. *Mol. Brain* **7**, 12. (doi:10.1186/1756-6606-7-12)
43. Johnson TK, Crossman T, Foote KA, Henstridge MA, Saligari MJ, Forbes Beadle L, Herr A, Whisstock JC, Warr CG. 2013 Torso-like functions independently of Torso to regulate *Drosophila* growth and developmental timing. *Proc. Natl Acad. Sci. USA* **110**, 14 688–14 692. (doi:10.1073/pnas.1309780110)
44. Chang H, Cahill H, Smallwood PM, Wang Y, Nathans J. 2015 Identification of astrotactin2 as a genetic modifier that regulates the global orientation of mammalian hair follicles. *PLoS Genet.* **11**, e1005532. (doi:10.1371/journal.pgen.1005532)
45. Zheng C, Heintz N, Hatten ME. 1996 CNS gene encoding astrotactin, which supports neuronal migration along glial fibers. *Science* **272**, 417–419. (doi:10.1126/science.272.5260.417)
46. Fishell G, Hatten ME. 1991 Astrotactin provides a receptor system for CNS neuronal migration. *Development* **113**, 755–765.
47. Adams NC, Tomoda T, Cooper M, Dietz G, Hatten ME. 2002 Mice that lack astrotactin have slowed neuronal migration. *Development* **129**, 965–972.
48. Stitt TN, Hatten ME. 1990 Antibodies that recognize astrotactin block granule neuron binding to astroglia. *Neuron* **5**, 639–649. (doi:10.1016/0896-6273(90)90218-5)
49. Allen Institute for Brain Science. 2015 Allen mouse brain atlas. See <http://mouse.brain-map.org/experiment/show/70305401>.
50. Lionel AC *et al.* 2014 Disruption of the ASTN2/TRIM32 locus at 9q33.1 is a risk factor in males for autism spectrum disorders, ADHD and other neurodevelopmental phenotypes. *Hum. Mol. Genet.* **23**, 2752–2768. (doi:10.1093/hmg/ddt669)
51. Price M *et al.* 2003 Seven cDNAs enriched following hippocampal lesion: possible roles in neuronal responses to injury. *Mol. Brain Res.* **117**, 58–67. (doi:10.1016/S0169-328X(03)00285-7)
52. Kashevarova AA *et al.* 2014 Array CGH analysis of a cohort of Russian patients with intellectual disability. *Gene* **536**, 145–150. (doi:10.1016/j.gene.2013.11.029)
53. Wang Y, Nathans J. 2007 Tissue/planar cell polarity in vertebrates: new insights and new questions. *Development* **134**, 647–658. (doi:10.1242/dev.02772)
54. Lesch KP *et al.* 2008 Molecular genetics of adult ADHD: converging evidence from genome-wide association and extended pedigree linkage studies. *J. Neural Transm.* **115**, 1573–1585. (doi:10.1007/s00702-008-0119-3)
55. Glessner JT *et al.* 2009 Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* **459**, 569–573. (doi:10.1038/nature07953)
56. Vrijenhoek T *et al.* 2008 Recurrent CNVs disrupt three candidate genes in schizophrenia patients. *Am. J. Hum. Genet.* **83**, 504–510. (doi:10.1016/j.ajhg.2008.09.011)
57. Wang KS *et al.* 2014 Polymorphisms within ASTN2 gene are associated with age at onset of Alzheimer's disease. *J. Neural Transm.* **122**, 701–708. (doi:10.1007/s00702-014-1306-z)
58. McCormack RM, Lyapichev K, Olsson ML, Podack ER, Munson GP. 2015 Enteric pathogens deploy cell cycle inhibiting factors to block the bactericidal activity of Perforin-2. *eLife* **4**, e06505. (doi:10.7554/eLife.06505)
59. Benton MJ. 2005 *Vertebrate palaeontology*, 3rd edn. New York, NY: John Wiley.
60. Aricescu AR *et al.* 2006 Eukaryotic expression: developments for structural proteomics. *Acta Crystallogr.* **D62**, 1114–1124. (doi:10.1107/S0907444906029805)
61. Lukyanova N *et al.* 2015 Conformational changes during pore formation by the perforin-related protein pleurotolysin. *PLoS Biol.* **13**, e1002049. (doi:10.1371/journal.pbio.1002049)
62. Shatursky O, Heuck AP, Shepard LA, Rossjohn J, Parker MW, Johnson AE, Tweten RK. 1999 The mechanism of membrane insertion for a cholesterol-dependent cytolysin: a novel paradigm for pore-forming toxins. *Cell* **99**, 293–299. (doi:10.1016/S0092-8674(00)81660-8)
63. Rebol CF, Whisstock JC, Dunstone MA. 2014 A new model for pore formation by cholesterol-dependent cytolysins. *PLoS Comput. Biol.* **10**, e1003791. (doi:10.1371/journal.pcbi.1003791)
64. Gilbert RJ, Mikelj M, Dalla Serra M, Froelich CJ, Anderlugh G. 2013 Effects of MACPF/CDC proteins on lipid membranes. *Cell Mol. Life Sci.* **70**, 2083–2098. (doi:10.1007/s00018-012-1153-8)
65. Xiong JP, Stehle T, Goodman SL, Arnaout MA. 2004 A novel adaptation of the integrin PSI domain revealed from its crystal structure. *J. Biol. Chem.* **279**, 40 252–40 254. (doi:10.1074/jbc.C400362200)
66. Gerke V, Creutz CE, Moss SE. 2005 Annexins: linking Ca²⁺ signalling to membrane dynamics. *Nat. Rev. Mol. Cell Biol.* **6**, 449–461. (doi:10.1038/nrm1661)
67. Lemmon MA. 2008 Membrane recognition by phospholipid-binding domains. *Nat. Rev. Mol. Cell Biol.* **9**, 99–111. (doi:10.1038/nrm2328)
68. Stevens LM, Beuchle D, Jurcsak J, Tong X, Stein D. 2003 The *Drosophila* embryonic patterning determinant torsolike is a component of the eggshell. *Curr. Biol.* **13**, 1058–1063. (doi:10.1016/S0960-9822(03)00379-8)
69. Heckman KL, Pease LR. 2007 Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protoc.* **2**, 924–932. (doi:10.1038/nprot.2007.132)
70. Robert X, Gouet P. 2014 Deciphering key features in protein structures with the new ENDScript server. *Nucleic Acids Res.* **42**, W320–W324. (doi:10.1093/nar/gku316)
71. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. 2010 ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, W529–W533. (doi:10.1093/nar/gkq399)
72. Nettleship J. 2012 Structural biology of glycoproteins. In *Glycosylation* (ed. S Petrescu), pp. 41–62. Rijeka, Croatia: InTech.
73. Bergfors T. 2003 Seeds to crystals. *J. Struct. Biol.* **142**, 66–76. (doi:10.1016/S1047-8477(03)00039-X)
74. Winter G. 2010 xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190. (doi:10.1107/S0021889809045701)
75. Pape T, Schneider G. 2004 HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J. Appl. Crystallogr.* **37**, 843–844. (doi:10.1107/S0021889804018047)
76. Sheldrick GM. 2010 Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr.* **D66**, 479–485.
77. Adams PD *et al.* 2010 PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr.* **D66**, 213–221.
78. Terwilliger T. 2004 SOLVE and RESOLVE: automated structure solution, density modification and model building. *J. Synchrotron. Radiat.* **11**, 49–52. (doi:10.1107/S0909049503023938)
79. Emsley P, Cowtan K. 2004 Coot: model-building tools for molecular graphics. *Acta Crystallogr.* **D60**, 2126–2132. (doi:10.1107/S0907444904019158)

80. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001 Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA* **98**, 10 037–10 041. (doi:10.1073/pnas.181342398)
81. Chen VB *et al.* 2010 MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr.* **D66**, 12–21. (doi:10.1107/S0907444909042073)
82. Nettleship JE, Brown J, Groves MR, Geerlof A. 2008 Methods for protein characterization by mass spectrometry, thermal shift (ThermoFluor) assay, and multiangle or static light scattering. *Methods Mol. Biol.* **426**, 299–318. (doi:10.1007/978-1-60327-058-8_19)
83. Petoukhov MV *et al.* 2012 New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350. (doi:10.1107/S0021889812007662)
84. Svergun DI. 1992 Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503. (doi:10.1107/S0021889892001663)
85. Franke D, Svergun DI. 2009 DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42**, 342–346. (doi:10.1107/S0021889809000338)
86. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015 The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858. (doi:10.1038/nprot.2015.053)
87. Schurpf T, Chen Q, Liu JH, Wang R, Springer TA, Wang JH. 2012 The RGD finger of Del-1 is a unique structural feature critical for integrin binding. *FASEB J.* **26**, 3412–3420. (doi:10.1096/fj.11-202036)
88. Brown PH, Schuck P. 2006 Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophys. J.* **90**, 4651–4661. (doi:10.1529/biophysj.106.081372)
89. Stuart DI, Levine M, Muirhead H, Stammers DK. 1979 Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 Å. *J. Mol. Biol.* **134**, 109–142. (doi:10.1016/0022-2836(79)90416-9)
90. Riffel N, Harlos K, Iourin O, Rao Z, Kingsman A, Stuart D, Fry E. 2002 Atomic resolution structure of Moloney murine leukemia virus matrix protein and its relationship to other retroviral matrix proteins. *Structure* **10**, 1627–1636. (doi:10.1016/S0969-2126(02)00896-1)
91. Felsenstein J. 1997 An alternating least squares approach to inferring phylogenies from pairwise distances. *Syst. Biol.* **46**, 101–111. (doi:10.1093/sysbio/46.1.101)
92. Suhre K, Sanejouand YH. 2004 Elnemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic Acids Res.* **32**, W610–W614. (doi:10.1093/nar/gkh368)
93. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004 UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612. (doi:10.1002/jcc.20084)