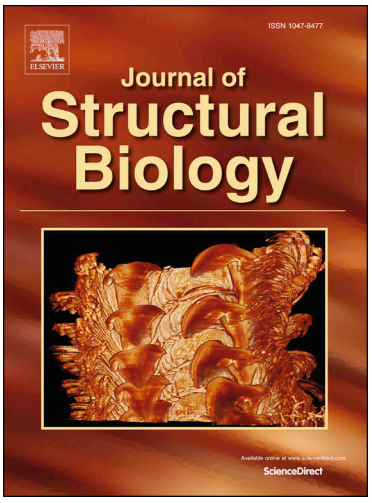
Journal Pre-proofs



Exaptation of two ancient immune proteins into a new dimeric pore-forming toxin in snails

M.L. Giglio, S. Ituarte, V. Milesi, M.S. Dreon, T.R. Brola, J. Caramelo, J.C.H. Ip, S. Maté, J.W. Qiu, L.H. Otero, H. Heras

PII: S1047-8477(20)30104-0
DOI: https://doi.org/10.1016/j.jsb.2020.107531
Reference: YJSBI 107531

To appear in: *Journal of Structural Biology*

Received Date: 6 February 2020
Revised Date: 1 May 2020
Accepted Date: 14 May 2020

Please cite this article as: Giglio, M.L., Ituarte, S., Milesi, V., Dreon, M.S., Brola, T.R., Caramelo, J., Ip, J.C.H., Maté, S., Qiu, J.W., Otero, L.H., Heras, H., Exaptation of two ancient immune proteins into a new dimeric pore-forming toxin in snails, *Journal of Structural Biology* (2020), doi: https://doi.org/10.1016/j.jsb.2020.107531

Exaptation of two ancient immune proteins into a new dimeric pore-forming toxin in snails

M.L. Giglio[a], S. Ituarte[a], V. Milesi[b], M.S. Dreon[a], T.R. Brola[a], J. Caramelo[c], J.C.H. Ip[d], S. Maté[a], J.W. Qiu[d], L.H. Otero[c,e*] and H. Heras[a*]

a. Instituto de Investigaciones Bioquímicas de La Plata "Prof. Dr. Rodolfo R. Brenner", INIBIOLP. CONICET CCT La Plata - Universidad Nacional de La Plata (UNLP), Facultad de Medicina,1900 La Plata, Argentina.
b. Instituto de Estudios Inmunológicos y Fisiopatológicos, IIFP. CONICET CCT La Plata – UNLP, Facultad de Ciencias Exactas, 1900 La Plata, Argentina.
c. Instituto de Investigaciones Bioquímicas de Buenos Aires, IIBBA. CONICET- Fundación Instituto Leloir, Av. Patricias Argentinas 435, C1405BWE Buenos Aires, Argentina.
d. Department of Biology, Hong Kong Baptist University, 224 Waterloo Road, Hong Kong, China.
e. Plataforma Argentina de Biología Estructural y Metabolómica PLABEM, Av. Patricias Argentinas 435, C1405BWE, Buenos Aires, Argentina.

***Co-corresponding authors:**

Horacio Heras
Instituto de Investigaciones Bioquímicas de La Plata (INIBIOLP), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) - Universidad Nacional de La Plata (UNLP),
60 y 120, 1900 La Plata, Argentina.
Tel: +54 (221) 482-4894
Email: h-heras@med.unlp.edu.ar
ORCID ID: 0000-0003-3379-0216

Lisandro H. Otero
Instituto de Investigaciones Bioquímicas de Buenos Aires (IIBBA), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) - Fundación Instituto Leloir, Patricias Argentinas 435, C1405BWE Buenos Aires, Argentina.
Tel: +054 (11) 5238-7500
Email: lotero@leloir.org.ar
ORCID ID: 0000-0002-5448-5483

Abstract

The Membrane Attack Complex-Perforin (MACPF) family is ubiquitously found in all kingdoms. They have diverse cellular roles, however MACPFs with pore-forming toxic function in venoms and poisons are very rare in animals. Here we present the structure of PmPV2, a MACPF toxin from the poisonous apple snail eggs, that can affect the digestive and nervous systems of potential predators. We report the three-dimensional structure of PmPV2, at 17.2 Å resolution determined by negative-stain electron microscopy and its solution structure by small angle X-ray scattering (SAXS). We found that PV2s differ from nearly all MACPFs in two respects: it is a dimer in solution and protomers combine two immune proteins into an AB toxin. The MACPF chain is linked by a single disulfide bond to a tachylectin chain, and two heterodimers are arranged head-to-tail by non-covalent forces in the native protein. MACPF domain is fused with a putative new Ct-accessory domain exclusive to invertebrates. The tachylectin is a six-bladed β-propeller, similar to animal tectonins. We experimentally validated the predicted functions of both subunits and demonstrated for the first time that PV2s are true pore-forming toxins. The tachylectin ¨B¨ delivery subunit would bind to target membranes, and then the MACPF ¨A¨ toxic subunit would disrupt lipid bilayers forming large pores altering the plasma membrane conductance. These results indicate that PV2s toxicity evolved by linking two immune proteins where their combined preexisting functions gave rise to a new toxic entity with a novel role in defense against predation. This structure is an unparalleled example of protein exaptation.

Keywords: poisonous egg defense, lectin, MACPF, AB toxin, *Pomacea*

Introduction

The integrity of cellular membranes is crucial for life and the disruption of such integrity causes cell death. Animals have evolved many strategies for damaging membranes and pore formation by proteins is frequently used in toxic attack on cells, as it can lead to efficient disruption of cell metabolism or even cell death. Among these proteins, the largest group belongs to the Membrane Attack Complex and Perforin / Cholesterol- Dependent Cytolysins (MACPF/CDC) superfamily, with ubiquitous distribution in all kingdoms. Most characterized members of MACPF/CDC interact with membranes and form large pores eventually leading to cell death (hence the name: pore-forming toxins, PFTs). The varied functions of MACPFs include immunity, development as well as bacterial and parasite pathogenesis [1, 2]. In recent years, a small group of animal MACPF was also found playing a role in venoms and poisons of both vertebrate (the stonefish *Synanceia spp.*) [3] and invertebrates (Cnidaria [4, 5] and apple snails of the genus *Pomacea* [6-8]).

We focus on the PFTs from the poisonous eggs of *Pomacea* apple snails (Gastropoda: Ampullariidae). Among them, *Pomacea canaliculata* eggs contain the toxin perivitelin-2 (PcPV2), one of the most toxic egg proteins known [7]. PcPV2 is composed of two subunits, a MACPF chain, and a tachylectin-like chain (member of the F-type lectin family [9]), termed PcPV2-67 and PcPV2-31, respectively [6, 7]. Moreover, the egg fluid (PVF) of *Pomacea maculata*, a related species, also contains a PV2-67 and PV2-31 like proteins orthologous of the two PcPV2 subunits [10]. Also, PcPV2 can be included into the AB toxins, a small group of toxic proteins found in bacteria (e.g. botulinum neurotoxins) and plants (e.g. Type-2 RIP), that play a role in pathogenic processes and embryo defense, respectively. These AB toxins contain two

moieties, the “A” moiety that modifies some cellular target leading to cell death and the “B” moiety, which usually has a carbohydrate binding module (CBM), that recognizes glycans of the cell membrane and acts as a delivery subunit [11]. Some of the CBM properties of AB toxins have been recognized in PcPV2 as it can agglutinate erythrocytes and recognize intestinal cells [6], however, little is known about its sugar specificity and toxic mechanism. PcPV2 is unique among AB toxins in that not only the B but also the A moiety - a MACPF - has putative membrane binding capacity, although there is no experimental confirmation of the pore-formation capacity of PV2s. For instance, experiments with mice indicate that minute quantities of PcPV2 are lethal if they enter the bloodstream [6, 7]. Eggs extracts of *P. maculata* were also poisonous and caused lethal toxicity to mice by an unidentified factor. It was observed that after PVF inoculation, severe signs pointing to nervous disorders appeared while at longer periods, mice showed paralysis of the rear limbs and even death [12]. The poisonous eggs of *P. canaliculata* and *P. maculata*, have an additional line of defense advertising the noxiousness by a bright pink coloration that warns predators (aposematic coloration) (Fig. 1A). Apple snail defensive strategy pays off and as a result eggs have very few predators [13]. Here, we identified PmPV2 as the neurotoxic factor in *P. maculata* eggs and studied its structure and putative functions. We found that PmPV2 presents unique structural features compared with other MACPFs, and demonstrated that both subunits – tachylectin and MACPF – are functional, being the first experimental validation of the pore-forming capacity of PV2 toxins.

Results

Identification and toxic activity of PmPV2

We tracked the toxin of the PVF that causes lethal toxicity in mice by protein purification (Fig. 1B) and toxicity tests. A large oligomeric protein was subsequently identified through mass spectrometry, which indicated that the protein isolated was Perivitellin-2 (PmPV2) (Pma_3499_0.54 and Pma_3499_0.31 [14]) (Fig. S1).

Purified PmPV2 proved to cause the same neurological effects than previously reported for the whole *P. maculata* PVF, with a LD50-96h of 0.25 mg/kg after i.p. injection (Fig. S2). This pointed out that PmPV2 is responsible of the poisonous effect of snail eggs.

Structural features of PmPV2

Native PmPV2 is a ~162 kDa oligomeric glycoprotein that with an anionic detergent separate into a single band of ca. 98 kDa, which upon reduction dissociates into a heavy chain (PmPV2-67) and a light chain (PmPV2-31) (Fig. 1C, D, S3). The heterodimer is joined by a single disulfide bond between Cys161 (PmPV2-31) and Cys398 (PmPV2-67) as determined by mass spectrometry (Fig. 2A, Table S1). PmPV2-67 has two glycoforms of pI 5.22 and 5.38 while PmPV2-31 has a single form (pI 8.16) as determined in two-dimensional electrophoresis (2DE) (Fig. S3).

Homology modeling of the two subunits allowed us to obtain structures with a reasonable match to their templates (Fig. 2B, Fig S4). Pfam analysis indicates that PmPV2-31 chain has a lectin domain that belongs to the HydWA family (PF06462, E-value=$6.6e^{-5}$) (Fig. S5). This lectin-like domain was identified as structurally similar to carp fish egg lectin (4RUS, chain D), giving a 6 bladed β-propeller structure model (Fig. 2B).

The Nt region of the PmPV2-67 chain has a MACPF domain (PF01823, E-value= $7.e^{-24}$) with the conserved signature, the Cys residues and the 3 GlyGly sites, all assumed to be important for MACPFs membrane binding (Fig. S5). The best suitable template for the MACPF module was perforin-1 (3NSJ, chain A), an innate immune system protein. The structure modeled was a MACPF fold with their characteristic twisted and bent β-sheet core and its two flanking transmembrane hairpin helixes (TMH1/2) of 40 and 42 residues, respectively (Fig. 2B and Fig. S5). These two helix-clusters are amphiphilic, a known requirement to unfold and insert into membranes.

Although both PmPV2 subunits showed low identity with their templates (Fig. S4), this is in agreement with previous reports for both lectin and MACPF families [15-18]. On the other hand, the Ct region of the PmPV2-67 shares no obvious sequence similarity to any structurally or functionally characterized domain, preventing an accurate 3D modeling. However, a secondary structure prediction with a reasonable confidence was obtained (Fig. S6), showing 19 % α-helix and 30 % β-sheet structures.

Spectroscopic measurements provided further insight into the structure, indicating that PmPV2 does not have an absorbing prosthetic group (Fig. S7A). Protein tryptophan and other aromatic amino acids are buried in a non-aqueous and highly rigid environment (Fig. S7B,C). The secondary structure, with equivalent amounts of alpha helixes and beta sheets (Fig. S7D and Table S2), agrees with the predicted 3D model.

Negative-stain reconstruction of PmPV2

We used negative-stain transmission electron microscopy (TEM) to determine the oligomeric state and obtain low-resolution structural information. Single particles of PmPV2 were perfectly distinguishable from which a preliminary 3D map was obtained (Fig. S8A,B).

The ensuing 2D class averages (Fig. 2C) showcased a set of distinct well-defined projections revealing clear structural features. Consequently, an *ab-initio* 3D TEM map of PmPV2 was obtained from reference-free 2D class averages with no symmetry restraints, which was iteratively refined imposing a C2 symmetry. This symmetry type was also validated with symmetry test programs as specified in the "Methods" section. This symmetry was in good agreement with the experimental results mentioned above, where the native molecular mass estimated by LS almost doubled the 98-kDa estimated for each Lectin-MACPF dimeric subunit by SDS-PAGE, indicating two of these subunits are present in the PmPV2 structure. The final 3D map obtained (Fig. 2D) had a 17.2 Å resolution according to 0.143 FSC criteria (Fig. S8C), and showed size overall dimensions (180 Å x 95 Å), and volume (387.84 $Å^3$) consistent with a tetrameric assembly of PmPV2 subunits.

Despite the low-resolution, the tachylectin subunits with the typical donut-like shape (Fig. 2B,D,F), as well as the MACPF subunits with the characteristic planar structure (Fig. 2B,D,F), were perfectly recognizable revealing a head-to-tail quaternary rearrangement (Fig. 2F).

Accordingly, the docking of the MACPF and tachylectin models into the TEM-map, shows an antiparallel dimer-of-heterodimers assembly with a C2 symmetry (Fig. 2F). Both heterodimers are docked to each other by non-covalent forces between a tachylectin from one heterodimer and the MACPF of the other. The rest of the chain does not seem to be part of the tetramer assembly. Despite the PmPV2-67 Ct (IMAD, see below) linker domain is defined in the

TEM-map between the MACPF and tachylectin domains, no model information is available (Fig. 2F).

As a whole, structural data indicate that PmPV2 can be regarded as a dimer of heterodimers held together head-to-tail by non-covalent forces, being the subunits of each heterodimer linked by a single interchain disulfide bond.

SAXS of PmPV2

To analyze PmPV2 overall shape and size in solution, we used small angle X-ray scattering (SAXS). After merging the measurements from three different proteins concentrations, experimental data represented as log *I(q)* vs *q* showed the typical sigmoidal shape (Fig. S9A). SAXS analysis of the linearized region of Guinier indicated that PmPV2 has a gyration radius of 43.9 ± 0.3 Å (Fig. S9B). The Kratky plot [$I(q)q^2$ vs q] had a bell shape characteristic of globular proteins (Fig. S9C), where the higher q values did not reach cero, indicating the presence of flexible regions in the protein [19]. The pair distance distribution *P(r)* function had a maximum peak at 43.7 Å and a small shoulder at ~110 Å, indicating that PmPV2 is an anisometric protein with a D_{max} of 142.5 Å (Fig. S9D). These results are compatible with a 173 kDa particle, thus both SLS and SAXS yielded similar molecular weights. Several independent *ab-initio* runs yielded reproducible molecular shapes, and the average models generated are consistent with a dimeric state, in agreement with negative-stain TEM and SEC-SLS results (Fig. 2E and S10).The global shape of the low-resolution 3D model of PmPV2 obtained by SAXS of the protein in solution is in good concordance with the negative-stain TEM model (Fig. 2E,F and S10).

However, both TEM 3D model hands fit in the low-resolution SAXS model and, therefore, with the resolution attained at this stage no handedness can be discarded (Fig. S10)

PmPV2 is an active lectin and is able to form transmembrane pores

To analyze the activity of the tachylectin module we tested PmPV2 agglutinating capacity against rabbit red blood cells (RBC). PmPV2 above 0.8 mg/mL produced hemagglutination of RBC (Fig. 3A). To demonstrate that the lectin activity was responsible for the agglutination, and not some other process, a control test was performed adding different sugars to inhibit agglutination. This competition assay showed that PmPV2 hemagglutinating activity was strongly inhibited by aminated monosaccharaides, while other sugars had little or no effect (Fig. 3A).

Considering the presence of a MACPF domain in the PmPV2-67 chain, we also evaluated the putative pore-forming activity of PV2 using Caco-2 cells, a cell line on which it binds to [6]. We examined membrane conductance changes by patch clamp techniques. Cells exposed to 29 nM PmPV2 (5 µg/mL) rapidly showed discrete current increments in a stepwise fashion that began to be detectable 2-3 min after the toxin was added (Fig. 3B and Fig. S11). This behavior lasted a few seconds and then the current stabilized at a final increased value respect to the control condition. From each discrete current jump, we calculated the conductance (G), obtaining a mean value of 1,116 ± 53 pS (n = 43 of six cells tested), and estimated a pore diameter (*d) of* 7.2 nm ($d = 2\sqrt{(Gh/\sigma\pi)}$) assuming a solution conductivity (σ) of 1.6 S. m^{-1} and a membrane thickness (h) of 5 nm. This experiment showed that PmPV2 had the capacity to form pores but did not give information on whether the lectin module was needed to recognize and

direct the toxin towards the membrane surface. To test this, the toxin was pre incubated with D-glucosamine before adding to the cells. After this treatment, the toxin was unable to change cell conductance, indicating that the lectin module was required for PmPV2 pore formation (Fig. 3B). This result suggested that PFT module was active and dependent on the presence of an active lectin for activity. In agreement with patch clamp results, TEM imaging of PmPV2 interaction with POPC/Cho liposomes captured pore-like structures with an inner diameter of 5.6 ± 0.16 nm (Fig. 3C).

Phylogenetic analysis revealed a novel MACPF accessory domain exclusive of invertebrates

BLASTp search of PmPV2-31 chain in NCBI non-redundant database revealed 22 similar sequences, mostly belonging to lectin families. All except one, a fish egg lectin-like protein from *Rhinatrema bivittatum*, belonged to invertebrates (Fig. S12A). Remarkably, the Cys161 involved in the disulfide linking of this subunit to the heavy chain, was only observed in *Pomacea* sequences (Fig. S12B). BLASTp analysis of PmPV2-67 chain showed 36 similar sequences scattered in vertebrates and invertebrates, 31 belonging to the MACPF family (Table S3). BLASTp searches of this sequence showed two conserved regions: an Nt-region containing the MACPF domain, and a Ct-region with a few matches with unknown proteins. A domain boundary prediction analysis by ThreaDom [20] indicated that PmPV2-67 has a relatively disorganized region between the MACPF domain and the Ct-region, suggesting the subunit is composed by two different domains. Analyzing the two regions separately, residues 1-335 (Nt-PmPV2-67) and 336-565 (Ct-PmPV2-67) unexpectedly, the Ct region matched 18 sequences exclusive of invertebrates (Fig. 4A), 13 associated with Ct-regions of MACPF-containing proteins

and 2 associated to a Notch domain (a domain involved in membrane interaction in vertebrate MACPF proteins) (Fig. 4B). Interestingly, phylogenetic analysis indicates an early diversification of MACPF PmPV2 like proteins in Mollusks (Fig. 4A). Multiple sequence alignment of the Ct region with the matching sequences revealed several conserved residues, in particular many Cys (Fig. S12C). We named this novel domain ¨Invertebrate MACPF Accessory Domain, IMAD¨. Notably, *P. canaliculata* and *P. maculata* IMADs contain binding site to the tachylectin chain trough a disulfide bridge, thus allowing the MACPF module attachment to the lectin. Other functions of IMAD in invertebrates remain to be investigated.

Discussion

PmPV2 structure and toxicity

The acquisition of venoms and poisons is a transformative event in the evolution of an animal, because it remodels the predator-prey interaction from a physical to a biochemical battle, enabling animals to prey on, or defend themselves against, much larger animals [21]. Here we report the initial functional and structural characterization of PmPV2, a toxin that, according to the experimental results on mice and cell cultures, would be a potential defense of apple snail embryos against predation [6-8].

Although not as potent as other snail toxins such as conotoxins [22], PmPV2 could be consider as “highly toxic”, similar to many snake venoms [23]. The toxin proved lethal to mice when it entered the bloodstream and those receiving sublethal doses displayed neurological signs similar to those caused by the PVF [12] or the PcPV2 toxin [7]. It is therefore unique among animals in that it has a dual function as both an enterotoxin [8] and as a neurotoxin (this study) affecting both the nervous and digestive system of mammals.

The general structural features of PmPV2, analyzed by spectroscopic methods and PAGE, were similar to those previously described for the orthologue PcPV2 orthologous [6, 7, 24]. Like PcPV2 [6, 7, 24], PmPV2 sequence indicate the presence of a lectin-like subunit (PmPV2-31) and a MACPF containing subunit (PmPV2-67), sharing 97 % and 96 % similarities, respectively [10]. In animals, lectins and MACPFs are ubiquitous and typically related with the innate immune system [1, 25], the main defense system against pathogens found in invertebrates [26]. The 3D homology models of both MACPF and lectin domains of PmPV2 agree with those found in the members of the respective families. The novelty here is that in

PV2s both chains are combined by a single disulfide bond forming lectin-MACPF heterodimers and two of these heterodimers are held together by non-covalent forces to form the native protein. Therefore, two structural features distinguish the PV2 toxins from the rest of the animal MACPFs (Fig. 5): (1) they are AB toxins, with a lectin B-chain that binds to cell surface glycans and a MACPF A-chain which kills target cells by forming membrane pores; (2) unlike other MACPF, PV2s are secreted as dimers.

A literature search indicates that PV2s are the only reported animal toxins with a binary AB structure. Furthermore, these are the only AB toxins where the toxic moiety is a member of the MACPF family thus, instead of having toxicity by enzymatic activity to alter target cell metabolism, as all other AB toxins do [27], they affect cells by forming pores. From the functional point of view, the lectin moiety would increase MACPF targeting specificity as compared to other MACPFs that bind to membranes solely by protein-lipid interactions (Fig. 5) [28]. Remarkably, no dimeric arrangement was reported before for an AB toxin. Apart from PV2, the only MACPF disulfide bonded to another protein is the complement component C8α, where the C8γ domain is disulfide-linked to the MACPF domain (Fig. 5). Unlike PV2-31 lectin, C8γ lipocalin enhance C8α pore-formation by binding small hydrophobic molecules, though it is not assential for its activity [29].

In our work, SLS, SAXS and negative-stain TEM consistently indicate that native PV2 is a dimer of heterodimers. As far as we know, there is only a single report of another MACPF secreted as a structurally-stable water-soluble dimer [3] and not as monomers as the vast majority of MACPF (Fig. 5). Interestingly, this dimeric MACPF is also a cytotoxin, the fish stonustoxin (SNTX), which has a parallel subunit arrangement. However, unlike PV2, SNTX does

not have a lectin subunit, or even a carbohydrate-binding domain (Fig. 5) [3]. The reason for the dimeric arrangement is still unknown. SAXS and negative-stain TEM-derived models allowed a visual analysis of PmPV2, which revealed an antiparallel head-to-tail orientation of its protomers. In the negative-stain TEM 3D reconstruction, the tachylectin subunit appears like a donut, which agreed with the predicted β-propeller structure [30-33], whereas MACPF domain presents the characteristic flattened shape of the MACPF/CDC fold involved in oligomerization and pore formation [15]. The novel arrangement of PV2 reported in this study expands the structural diversity of MACPF PFTs in nature (Fig. 5).

Another interesting aspect of PmPV2 structure is the MACPF Ct domain. With few exceptions, most MACPFs have ancillary domains that either enhance binding affinity or selectivity in the protein membrane recognition process [2, 34] (Fig. 5). In this regard, the MACPF domain of PmPV2-67 subunit is fused with a novel Ct accessory domain. We found this Ct domain is conserved among many invertebrate MACPF-containing proteins, and phylogenetic analysis suggests that MACPF and Ct-domains may have already been combined in the last common ancestor of invertebrates. Our results suggest that this conserved domain, which we dubbed IMAD, is a new family of MACPF-accessory domains exclusive of invertebrates with a still unknown structure and a putative membrane recognition function. In *Pomacea*, the IMAD also contains the binding site to the tachylectin chain. The simultaneous interaction with several membrane components to attain higher binding affinity and specificity for the target cell has been reported for other MACPF/CDC PFTs [34]. This is another avenue of future research.

PmPV2 is a pore-forming toxin (PFT) delivered by a lectin

The presence of a MACPF domain in the primary structure of both PcPV2 and PmPV2 [6, 10, 35], suggested a putative pore-forming activity. Here, we confirmed for the first time that PV2s are indeed PFTs and that upon binding, they oligomerize into a complex that penetrates the target membrane. Patch clamp experiments also indicate that, once cells are perforated, the membrane oligomeric structures are stable. Besides, the discrete jumps in membrane conductance in a stepwise fashion are consistent with the pore-forming activity already reported for other PFTs [36, 37]. This was further supported by the identification in the predicted structure of amphipathic sequences in the TMH1/2 together with the typical MACPF/CDC fold required to form pores. Finally, the TEM images provided a visual confirmation of pore-like structures of ~6 nm inner diameter, which agrees with the pore size estimated by patch clamp measurements, and lies within the range reported for other MACPFs [1].

We demonstrated that, beside the pore forming activity, PmPV2 is also an active lectin with a primary specificity for aminated sugars. In this regard, CBMs in other AB toxins are found to function in delivering the toxic component of the protein to cell surfaces through glycan-CBM interactions [38]. As blocking the lectin activity inhibited the pore-forming capacity on biological membranes, we could suggest that the binding of the tachylectin subunit is a necessary step for the pore formation by the MACPF chain. However, further studies are needed to unveil the membrane binding and pore-formation mechanisms of this toxin.

In conclusion, we found that apple snail eggs have evolved a novel PFT, which, combined with other defenses of the eggs, would disable essential physiological systems in predators. We provide the first evidence that PV2 toxins from snail eggs are active PFTs. Apple snail PV2, however, differs in several structural respects from known MACPF PFTs as it is disulfide-linked to a lectin into an AB toxin arrangement and also because it is secreted as a dimer instead of a monomer in aqueous solutions. Linking two immune proteins in a new toxic entity massively accumulated in the eggs is likely to represent the key step for PV2 novel role in defense against predation, an unparalleled example of protein exaptation. To the best of our knowledge, this is the first description of an animal AB toxin directed toward cell membranes. Future work will look at whether there are differences in the pore structure and oligomerization mechanism between PV2s and other PFTs.

Methods

Eggs collection and PmPV2 purification

Adult females of *Pomacea maculata* were collected in the Parana River in San Pedro (33°30'35.97'' S; 59°41'52.86'' W), Buenos Aires province, Argentina and kept in the laboratory (Collection permit number DI-2018-181-GDEBA-DAPYAMAGP, Government of the Buenos Aires Province). Eggs were collected within 24 h of laid and kept at -20 °C until processed. Pools of three clutches were homogenized in ice-cold 20 mM Tris-HCl, pH 7.4, keeping a 3:1 v/w buffer:sample ratio as previously described [7]. The crude homogenate was sequentially

centrifuged at 10,000 xg for 30 min and at 100,000 xg for 50 min to obtain the egg perivitelline fluid, PVF.

PmPV2 was obtained following the method described for PcPV2 [39]. Briefly, PVF was ultracentrifuged in a NaBr (density = 1.28 g/ml) gradient at 207,000 xg for 22 h at 10 °C. Then, PmPV2 fraction was purified by high performance liquid chromatography (HPLC) using a Mono Q™ 10/100 GL (GE Healthcare Bio-Sciences AB) column using a gradient of NaCl in 20 mM Tris-HCl buffer, pH 8.5; and by size-exclusion chromatography in a Superdex 200 10/300 GL (GE Healthcare Bio-Sciences AB) column. Purity was checked by electrophoresis in 4-20 % polyacrylamide gels.

Protein content was determined either by the method of Lowry [40] using Bovine Serum Albumin (BSA) as standard, or using PmPV2 molar extinction coefficient at 280 nm, ε^{280nm} (*see below*).

Protein identification and determination of disulfide bonds

To identify PmPV2, 10 µg of purified PmPV2 were first separated by SDS-PAGE and then visualized with colloidal Coomassie Brilliant Blue method. The 98-kDa band was sliced, alkylated with iodacetamide, and digested in-gel with mass spectrometry grade trypsin (Perkin-Elmer).

Peptides were desalted with Sep-Pak C18 cartridges (Waters, Milford, USA) and dried using SpeedVac concentrator (Eppendorf, Hamburg, Germany). Dried samples were reconstituted using 0.1 % formic acid for analysis using LTQ-Orbitrap Elite coupled to an Easy-nLC (Thermo Fisher, Bremen, Germany) with 80 min LC gradient: 5 min in 98% solution A (0.1% formic acid in H2O), 35 min in 7 - 20% solution B (0.1% formic acid in acetonitrile), 20 min in 20

- 35% solution B, 10 min in 35 - 90% solution B, 10 min in 90% solution B. The MS data were captured within a range of 500 to 1800 m/z. The ten most abundant multiple-charged ions with a signal threshold >500 counts were selected for fragmentation under high-energy collision-induced dissociation (HCD; 2.0 m/z of solation width 10 ms of activation time, 40% of normalized collision energy).

Raw data were converted to .mgf files using Proteome Discoverer 1.3.0.339 (Thermo Finnigan, CA) and the MS files were searched against Ampubase database [41].

To identify the disulfide bond between the two PmPV2 subunits, the MS files were searched against custom proteinPmPV2 databases [Pma_3499_0.31, Pma_3499_0.54 and Pma_3499_0.24, which were found in PmPVF [14] using the pLink-SS incorporated into pLink 2.3.5 [42-44] with the cross linkage search of disulfide bond and default parameters.

Toxicity tests

All studies performed with animals were carried out in accordance with the Guide for the Care and Use of Laboratory Animals [45] and were approved by the ‘‘Comité Institucional de Cuidado y Uso de Animales de Experimentación’’ of the School of Medicine, UNLP (Assurance No. P08-01-2013). Animals were obtained from the Experimental Animals Laboratory of the School of Veterinary Science, UNLP. Groups of five female BALB/cAnN mice (body weight: 16 ± 1.1 g) were injected intraperitoneally (i.p.) with a single dose of 200 µL of PBS buffer (1.5 mM NaH_2PO_4, 8.1 mM Na_2HPO_4, 140 mM NaCl, 2.7 mM KCl, pH 7.4) or the same volume of a serial dilution of five concentrations of PmPV2. Median lethal dose (LD50) was determined by a lethality test 96 h after injection, statistical analysis was performed by PROBIT using EPA-Probit

analysis program v1.5 statistical software of the US Environmental Protection Agency (US EPA), based on Finney's method [46].

Mass determination

Molecular weight of native PmPV2 in solution was determined by light scattering using a Precision Detectors PD2010 90° light scattering instrument tandemly connected to a high-performance liquid chromatography and an LKB 2142 differential refractometer. 100 µL of the sample were injected into a Superdex 200 GL 10/300 (GE Healthcare) column, and the chromatographic runs were performed with a buffer containing 20 mM Tris-HCl pH 7.5, 250 mM NaCl under isocratic conditions at a flow rate of 0.4 mL/min at 20 °C. The concentration of the injected sample was 1.35 mg/ml. The MW of each sample was calculated relating its 90° light scattering and refractive index (RI) signals and comparison of this value with the one obtained for BSA (MW 66.5 kDa) as a standard using the software Discovery32. The reported MW values are an average between the values relating RI and UV with scattering.

Polyacrylamide gel electrophoresis (PAGE)

Native and subunit composition of PmPV2 were determined by PAGE in 4-20% gradient polyacrylamide gels using Mini-Protean II System (Bio Rad Laboratories, Inc., Hercules, CA). Non-native conditions were performed using 0.1% sodium dodecyl sulfate (SDS), 0.5% dithiothreitol (DTT) and β-mercaptoethanol. Low and high molecular weight markers (GE Healthcare Bioscience, Uppsala, Sweden) were run in parallel. Gels were stained using Coomassie Brillant Blue G-250. Glycosylation was detected by PAS staining following the

McGuckin and McKenzie [47] method modified by Streitz et al [48], using a commercial Schiff reagent (BioPack). Further analysis was performed by two-dimensional electrophoresis gels (2-DE) in an Ettan IPGphor 3 system (GE Healthcare), as previously described [39] using 60 µg of PmPV2.

Spectroscopic analysis

Absorbance: Absorption spectra of PmPV2 (0.64 mg/mL in 20 mM Tris-HCl, 150 mM NaCl buffer, pH 7.5) were recorded between 240 and 700 nm. Ten spectra of three independent pools were measured and averaged. Forth-derivative operation was applied to analyze the relative contribution of different aromatic residues [49].

The molar extinction coefficient of denatured PmPV2 was experimentally determined by measuring the absorbance at 280 nm of a solution of 720 µg of lyophilized protein in 6 M guanidinium hydrochloride (GnHCl), following equation 1:

$$C = \frac{Abs}{\varepsilon} \qquad [1]$$

where *C* is the protein concentration (in $mg.mL^{-1}$), *Abs* the absorbance at a given wavelength (in nm), ε the molar extinction coefficient (in $mg^{-1}.mL$). To determine the molar extinction coefficient of the native PmPV2, the absorbance of the native and the denatured protein were measured at identical protein concentrations. Since the concentrations are equal, we combined the equation 1 of the two solutions to obtain the native molar extinction coefficient:

$$\varepsilon_{nat} = \frac{(Abs_{nat})(\varepsilon_{den})}{(Abs_{den})} \qquad [2]$$

where ε is the molar extinction coefficient (in mg^{-1}.mL), *Abs* the absorbance at a given wavelength (in nm), subscript *nat* refers to native protein and subscript *den* refers to denatured protein.

All these experiments were performed using an Agilent 8453 UV/Vis diode array spectrophotometer (Agilent Technologies).

Fluorescence: Fluorescence emission spectra of PmPV2 (65 µg/mL) in PBS buffer were recorded in scanning mode in a Perkin-Elmer LS55 spectrofluorometer (Norwalk). Protein was exited at 280 nm (4 nm slit) and emission recorded between 275 and 437 nm. Fluorescence measurements were performed in 10 mm optical-path-length quartz-cells. The temperature was controlled at 25±1 °C using a circulating-water bath.

Circular dichroism: Spectra of PmPV2 (70–140 µM) were recorded on a Jasco J-810 spectropolarimeter using quartz cylindrical cuvettes of 1-mm or 10-mm path lengths for the far-UV (200–250 nm) and near-UV (250–310 nm) regions, respectively. Data were converted into molar ellipticity $[\theta]_M$ ($deg.cm^2$. $dmol^{-1}$) using a mean residue weight value of 115.5 g/mol for PmPV2.

Proportions of different secondary structures were also obtained using CD spectra in DichroWeb [50] software using Contin and K2d algorithms.

Small angle X-ray scattering (SAXS): Synchrotron SAXS data from solutions of PmPV2 in 20 mM Tris, pH 7 were collected at the SAXS2 beam line at the Laboratório Nacional de Luz Sincrotron

(Campina, Brazil) using MAR 165 CDD detector at a sample-detector distance of 1.511 m and at a wavelength of λ = 0.155 nm. Scattering was recorded as $I_{(q)}$ vs q, where s = 4πsinθ/λ, and 2θ is the scattering angle. Protein concentrations ranging between 0.8 and 2 mg/ml were measured at 20 °C, collecting five successive 300 second frames. The data were normalized to the intensity of the transmitted beam and radially averaged; the scattering of the solvent-blank was subtracted. The low angle data collected at lower concentration were merged with the highest concentration high angle data to yield the final composite scattering curve, using ATSAS 2.8.4-1 software [51]. *Ab-initio* shape determination was performed using DAMIFF online (https://www.embl-hamburg.de/biosaxs/dammif.html) [52] and the resulting damstart.pdb file was used to refine the model using DAMIN [53] with default parameters except for symmetry (P2 imposed). Raw data, fits and models were deposited in SASBDB repository (SASDEN3). (https://www.sasbdb.org/data/SASDEN3/ohuzme8q9a/).

Bioinformatic analysis

PmPV2-31 and PmPV2-67 subunit sequences were annotated as Pma_3499_0.54 and Pma_3499_0.31, respectively, by Sun et al. [14]. PmPV2 related-sequences from different organisms were retrieved from NCBI non-redundant database by BLASTp set as default (threshold E-value=1e-5) and aligned using MUSCLE multiple alignment tool (https://www.ebi.ac.uk/Tools/msa/muscle/) for homology analysis. Phylogenetic analysis was performed using MrBayes v.3.2.6 software, with four chains of 100,000 generations. The tree was sampled every 100 generations, and the final burnin value was set to 20,000. The standard deviation of the split frequencies fell below 0.05. Trees were visualized by FigTree v.1.4.3.

Three-dimensional structures of PmPV2 subunits were predicted by homology modeling using Phyre2 software, which applied a profile-profile alignment algorithm [54], and pdb files were visualized using UCSF Chimera 1.14 [55]. Pair-wise sequence alignment between PmPV2 subunits and its templates were performed using LALIGN server [56], whereas structural alignments were made using the DALI server [57]. Finally, the secondary structure of PmPV2-67 C-terminal region was predicted using PSIPRED server v.4.0 [58] set as default.

Negative-stain TEM data acquisition of PmPV2, image processing, single-particle reconstruction, and refinement

PmPV2 protein samples were suspended in 20 mM Tris-HCl, 150 mM NaCl, pH 8.5 buffer at 0.05 mg/ml and kept on ice before grid preparation (higher concentrations caused oligomerizarion of the samples on the grids). Then, 3 µl of sample was loaded on ultrathin holey-carbon-supported grids, previously pretreated with a glow discharge system for TEM grids during 50 s, under a pressure of 37 Pa. The samples were incubated with the grids 1 min, blotted by filter papers, and then stained with uranyl acetate 2% (w/v) for 30 s. The excess of stain was removed by blotting. PmPV2 TEM analysis were performed at LNNano-CNPEM, Brazil (proposal ID 24346). Data acquisition was performed using a Talos F200C (Thermo Fisher) operated at 200 KV with a FEI BM-Ceta direct electron detector model. Data acquisition was performed on a grid, using at a nominal magnification of 73,000 X, corresponding to a calibrated pixel size of 2.02 Å per pixel and a defocus range of -2.0 to -4.0 µm. This defocus range was suitable for obtaining a good contrast balance, which allowed us to accurately visualize, pick and align particles from the specimen. A total number of 60 micrographs were recorded with an average

electron dose per image of 20 e- per Å2. Estimation of CTF, particle picking, 2D classification, reconstruction of an *ab-initio* model, and refinement were executed using the software cisTEM [59] (Fig. S13). Briefly, after estimating CTF, an initial template-free particle picking was performed. The preliminary set of picked single particles (20,115 particles) was first exposed to an initial 2D classification resulting in 19 classes (15,317 particles). Subsequently, 2D class averages were used for getting a preliminary *ab-initio* 3D map with no symmetry restriction (Fig. S14A). The first cycles of refinement were performed in the same symmetry condition (Fig. S14B). To assess the symmetry, the map was carefully examined in UCSF-CHIMERA where a 2-fold rotation axis (C2) was evident. Therefore, C2 symmetry was imposed for a set of particles (n= 15,317) and the following iterative refinement cycles, which allowed a better 3D map reconstruction (Fig. S14C).

To validate C2 symmetry, the maps C1 and C2 were firstly aligned to the C2 symmetry axis with e2proc3d.py from EMAN2.3 processing suite [60]. Then, the symmetry of the maps were examined using the Map Symmetry tool from PHENIX [61], where a C2 symmetry was unambiguosly assigned (score=1.41, CC=1.0). Moreover, C2 symmetry was validated with the program ProSHADE from CCP-EM platform [62] , where this symmetry type was detected in the C1 *ab-initio* map with a peak height (average) of the correlation after rotation for all the symmetry rotation = 0.7 (Table S4).

The MACPF and lectin 3D models were manually adjusted as rigid bodies considering both hands of the map using UCSF CHIMERA [55] and guided by the template crystals. After fitting of the models in one half of the dimeric complex, the other half was then independently fitted into the density map. Additionally, the ring-like structure of the lectin subunit was

positioned using the interdomain (IMAD/lectin) disulfide bond as reference. Thus, the cysteine residue 161 from lectin domain was orientated towards the IMAD domain during the manual docking. To improve fitting, the models were refined by real-space refinement against their respective maps in PHENIX [61]. Figures were generated using UCSF-CHIMERA. The negative-stain TEM datasets are available in the wwPDB repository (https://deposit.wwpdb.org/deposition/) with accession code EMD-21097.

Lectin activity of PmPV2

Rabbit blood samples were obtained from animal facilities at Universidad Nacional de La Plata by cardiac puncture and collected in sterile Alsever's solution (100 mM glucose, 20 mM NaCl, and 30 mM sodium citrate, pH 7.2). Prior to use, red blood cells (RBC) were washed by centrifugation at 1,500 xg for 10 min in PBS. Hemagglutinating and hemolytic activity were assayed using a two-fold serial dilution of PmPV2 (3.4 mg/mL) as previously described [6]. Primary specificity was determined by a competition assay. Erythrocytes were incubated with PmPV2 (0.87 mg/mL) in the presence of 0.1 M of D-mannose, D-galactose, D-galactosamine, N-acetyl-D-galactosamine, D-glucose, D-glucosamine, N-acetyl-D-glucosamine or D-fucose. PmPV2 concentration was selected as the concentration providing visible agglutination in previous analysis. All monosaccharides were purchased from Sigma-Aldrich.

Microscopic analyses of PmPV2 pores

Preparation of small unilamellar vesicles (SUVs): Multilamellar vesicles were prepared by mixing synthetic 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and cholesterol (Cho)

(Avanti Polar Lipids, Birmingham, AL, USA) dissolved in HPLC-grade chloroform/methanol (3:1 molar ratio). Then samples were dried by evaporating the solvent under a stream of nitrogen and then with high vacuum for 2 h in a speed vac. The samples were hydrated in a desired volume of buffer (25 mM HEPES, 150 mM NaCl, pH 7.4) with stirring to facilitate dispersion. Multilayered vesicles were sonicated in an FB-15049 sonicator bath (Fisher Scientific Inc., Waltham, MA, USA) at 30 °C for 1 h to obtain SUVs for AFM and TEM experiments.

Transmission electron microscopy (TEM) imaging An excess of SUVs was incubated with 2.1 µM PmPV2 for 30 min at 37 °C. After treatment, 5 µL of the liposome suspension was placed onto a 300-square-mesh copper grid covered with a Formvar carbon support film (Micro to Nano VOF, Netherlands) and fixed for 1 min. Samples were then negative stained with 50 µL of a 1% (w/v) phosphotungstic acid solution for 30 s. Images at different amplifications were taken using a TEM/STEM FEI Talos F200X microscope (Thermo Scientific) at 200 kV.

Patch-clamp recordings

Caco-2 cells were allowed to settle onto the cover glass bottom of a 3 ml experimental chamber. The cells were observed with a mechanically stabilized, inverted microscope (Telaval 3, Carl Zeiss, Jena, Germany) equipped with a 40x objective lens. The chamber was perfused for 15 min, at 1 ml.min-1 by gravity, with extracellular saline solution before the patch-clamp experiment was started. Application of test solutions was performed through a multi barreled pipette positioned close to the cell being investigated. All experiments were performed at 22 °C. The standard tight-seal whole-cell configuration of the patch-clamp technique was used

[63] following two different protocols. First, the cells were clamped using voltage ramps from -50 mV to +60 mV and the macroscopic evoked currents were measured before and after adding PmPV2 to a final concentration of 0.05 mg/mL and after washing cells with the extracellular solution. Secondly, cells were clamped at a holding potential of -50 mV, hence evoking a macroscopic holding current, which we measured before and after adding PmPV2 to a final concentration of 0.005 mg/mL to the bath solution. Glass pipettes were drawn from WPI PG52165-4 glass on a two-stage vertical micropipette puller (PP-83, Narishige Scientific Instrument Laboratories, Tokyo, Japan) and pipette resistance ranged from 2 to 4 MOhms. Ionic currents were measured with an Axopatch 200A amplifier (Axon Instruments, Foster City, CA) filtered at 2 kHz, and digitized (Digidata 1440 Axon Instruments, Foster City, CA) at a sample frequency of 20 kHz. The extracellular saline solution used for recording whole cell ionic currents had a composition similar to the physiological extracellular solution containing 130 mM NaCl, 4.7 mM KCl, 2.5 mM CaCl2, 6 mM glucose, and 5 mM HEPES; the intracellular solution had 130 mM KCl, 5 mM Na2ATP, 1 mM MgCl2, 0.1 mM EGTA, and 5 mM HEPES. The pH of both solutions was adjusted to 7.4 and 7.2, respectively, with NaOH

Accession numbers

The NS-EM density map has been deposited in the EMBD under accession code EMD-21097. The SAXS data has been deposited in the SASBDB under accession code SASDEN3.

Availability of data and material

The datasets generated during the SAXS experiments are available in the SASBDB repository (https://www.sasbdb.org/data/SASDEN3/ohuzme8q9a/). The NS-EM datasets are available in the wwPDB repository (https://deposit.wwpdb.org/deposition/) with accession code EMD-21097. All other data generated or analysed during this study are included in this published article and its supplementary information files.

Acknowledgements

SI, LHO, JC, VM and HH are members of CONICET, Argentina. MSD is member of CIC.BA, Argentina. TB and MLG are Doctoral and Postdoctoral students, respectively with scholarships from CONICET. We thank L. Bauzá for her technical assistance. We thank M. Ramos for his help in figure designs. We are grateful to Single Particle Cryo-EM staff from Brazilian Nanotechnology National Laboratory (LNNano, Campinas) for providing access to their facilities and for the technical assistance during EM experiments. We thank LNLS - Brazilian Synchrotron Light Laboratory for access to their facilities and for the technical assistance during SAXS experiments (project SAXS1-17746).

Competing interests

We have no competing interests.

References

[1] Anderluh G, Kisovec M, Krasevec N, Gilbert RJ. Distribution of MACPF/CDC proteins. Subcell Biochem. 2014;80:7-30.
[2] Peraro MD, van der Goot FG. Pore-forming toxins: ancient, but never really out of fashion. Nat Rev Microbiol. 2016;14:77-92.
[3] Ellisdon AM, Reboul CF, Panjikar S, Huynh K, Oellig CA, Winter KL, et al. Stonefish toxin defines an ancient branch of the perforin-like superfamily. Proc Natl Acad Sci U S A. 2015;112:15360-5.

[4] Nagai H, Oshiro N, Takuwa-Kuroda K, Iwanaga S, Nozaki M, Nakajima T. A new polypeptide toxin from the nematocyst venom of an Okinawan sea anemone Phyllodiscus semoni (Japanese name "unbachi-isoginchaku"). Biosci Biotechnol Biochem. 2002;66:2621-5.
[5] Oshiro N, Kobayashi C, Iwanaga S, Nozaki M, Namikoshi M, Spring J, et al. A new membrane-attack complex/perforin (MACPF) domain lethal toxin from the nematocyst venom of the Okinawan sea anemone *Actineria villosa*. Toxicon. 2004;43:225-8.
[6] Dreon MS, Frassa MV, Ceolin M, Ituarte S, Qiu JW, Sun J, et al. Novel animal defenses against predation: A snail egg neurotoxin combining lectin and pore-forming chains that resembles plant defense and bacteria attack toxins. PLoS One. 2013;8:e63782.
[7] Heras H, Frassa MV, Fernández PE, Galosi CM, Gimeno EJ, Dreon MS. First egg protein with a neurotoxic effect on mice. Toxicon. 2008;52:481-8.
[8] Giglio ML, Ituarte S, Ibañez AE, Dreon MS, Prieto E, Fernández PE, et al. Novel Role for Animal Innate Immune Molecules: Enterotoxic Activity of a Snail Egg MACPF-Toxin. Frontiers in Immunology. 2020;11.
[9] Bishnoi R, Khatri I, Subramanian S, Ramya TN. Prevalence of the F-type lectin domain. Glycobiology. 2015;25:888-901.
[10] Mu H, Sun J, Heras H, Chu KH, Qiu JW. An integrated proteomic and transcriptomic analysis of perivitelline fluid proteins in a freshwater gastropod laying aerial eggs. J Proteomics. 2017;155:22-30.
[11] Odumosu O, Nicholas D, Yano H, Langridge W. AB Toxins: A Paradigm Switch from Deadly to Desirable. Toxins (Basel). 2010;2:1612-45.
[12] Giglio ML, Ituarte S, Pasquevich MY, Heras H. The eggs of the apple snail *Pomacea maculata* are defended by indigestible polysaccharides and toxic proteins. Can J Zool. 2016;94:777-85.
[13] Yusa Y, Sugiura N, Ichinose K. Predation on the apple snail, *Pomacea canaliculata* (Ampullariidae), by the Norway rat, *Rattus norvegicus*, in the field. Veliger. 2000;43:349-53.
[14] Sun J, Mu H, Ip JCH, Li R, Xu T, Accorsi A, et al. Signatures of Divergence, Invasiveness and Terrestralization Revealed by Four Apple Snail Genomes. Mol Biol Evol. 2019;36:1507-20.
[15] Rosado CJ, Buckle AM, Law RH, Butcher RE, Kan WT, Bird CH, et al. A common fold mediates vertebrate defense and bacterial attack. Science. 2007;317:1548-51.
[16] Slade DJ, Lovelace LL, Chruszcz M, Minor W, Lebioda L, Sodetz JM. Crystal structure of the MACPF domain of human complement protein C8 alpha in complex with the C8 gamma subunit. J Mol Biol. 2008;379:331-42.
[17] Kopec KO, Lupas AN. beta-Propeller blades as ancestral peptides in protein evolution. PLoS One. 2013;8:e77074.
[18] Chaudhuri I, Soding J, Lupas AN. Evolution of the beta-propeller fold. Proteins. 2008;71:795-803.
[19] Kikhney AG, Svergun DI. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. FEBS Letters. 2015;589:2570-7.
[20] Xue Z, Xu D, Wang Y, Zhang Y. ThreaDom: extracting protein domain boundary information from multiple threading alignments. Bioinformatics. 2013;29:i247-56.
[21] Holford M, Daly M, King GF, Norton RS. Venoms to the rescue. Science. 2018;361:842-4.
[22] Luna-Ramirez KS, Aguilar MB, Falcon A, Heimer de la Cotera EP, Olivera BM, Maillo M. An O-conotoxin from the vermivorous *Conus spurius* active on mice and mollusks. Peptides. 2007;28:24-30.
[23] Gawade SP. Snake venom neurotoxins: Pharmacological classification. J toxicol Toxin reviews. 2004;23:37-96.
[24] Frassa MV, Ceolín M, Dreon MS, Heras H. Structure and stability of the neurotoxin PV2 from the eggs of the Apple Snail *Pomacea canaliculata*. Biochim Biophys Acta. 2010;1804:1492-9.
[25] Rudd PM, Elliot T, Cresswell P, Wilson IA, Dwek RA. Glycosylation and the immune system. Science. 2001;291:2370-6.
[26] Hoffmann JA, Kafatos FC, Janeway J, C.A, Ezenowitz RAB. Phylogenetic perspectives in innate immunity. Science. 1999;284:1313-8.

[27] Falnes PØ, Sandvig K. Penetration of protein toxins into cells. Current Opinion in Cell Biology. 2000;12:407-13.
[28] Ros U, Garcia-Saez AJ. More Than a Pore: The Interplay of Pore-Forming Proteins and Lipid Membranes. J Membr Biol. 2015;248:545-61.
[29] Parker CL, Sodetz JM. Role of the human C8 subunits in complement-mediated bacterial killing: evidence that C8γ is not essential. Molecular Immunology. 2002;39:453-8.
[30] Bonnardel F, Kumar A, Wimmerova M, Lahmann M, Perez S, Varrot A, et al. Architecture and Evolution of Blade Assembly in beta-propeller Lectins. Structure. 2019;27:764-75 e3.
[31] Jawad Z, Paoli M. Novel Sequences Propel Familiar Folds. Structure. 2002;10:447-54.
[32] Chen CK, Chan NL, Wang AH. The many blades of the beta-propeller proteins: conserved but versatile. Trends Biochem Sci. 2011;36:553-61.
[33] Fulop V, Jones DT. Beta propellers: structural rigidity and functional diversity. Curr Opin Struct Biol. 1999;9:715-21.
[34] Reboul CF, Whisstock JC, Dunstone MA. Giant MACPF/CDC pore forming toxins: A class of their own. Biochim Biophys Acta. 2016;1858:475-86.
[35] Mu H, Sun J, Heras H, Chu KH, Qiu J-W. Dataset for the proteomic and transcriptomic analyses of perivitelline fluid proteins in Pomacea snail eggs. Data in Brief. 2017;15:203-7.
[36] Marchioretto M, Podobnik M, Dalla Serra M, Anderluh G. What planar lipid membranes tell us about the pore-forming activity of cholesterol-dependent cytolysins. Biophys Chem. 2013;182:64-70.
[37] Podack ER, Ding-E Young J, Cohn A. Isolation and biochemical and functional characterization of perforin 1 from cytolytic T-cell granules. Proc Natl Acad Sci U S A. 1986;83:3050-.
[38] Boraston AB, Lammerts van Bueren A, Ficko-Blean E, Abbott DW. Chapter 3.29. Carbohydrate-Protein Interactions: Carbohydrate-Binding Modules. In: Kamerling JP, editor. Comprehensive Glycoscience From Chemistry to Systems Biology Vol 3. Utrecht: Elsevier Science; 2007. p. 661-96.
[39] Pasquevich MY, Dreon MS, Heras H. The major egg reserve protein from the invasive apple snail *Pomacea maculata* is a complex carotenoprotein related to those of *Pomacea canaliculata* and *Pomacea scalaris*. Comp Biochem Physiol. 2014;169 B:63-71.
[40] Lowry OH, Rosenbrough NJ, Farr AL, Randall R. Protein measurement with the Folin phenol reagent. J Biol Chem. 1951;193:265-75.
[41] Ip JCH, Mu H, Chen Q, Sun J, Ituarte S, Heras H, et al. AmpuBase: a transcriptome database for eight species of apple snails (Gastropoda: Ampullariidae). BMC Genomics. 2018;19:179.
[42] Lu S, Cao Y, Fan SB, Chen ZL, Fang RQ, He SM, et al. Mapping disulfide bonds from sub-micrograms of purified proteins or micrograms of complex protein mixtures. Biophys Rep. 2018;4:68-81.
[43] Lu S, Fan S-B, Yang B, Li Y-X, Meng J-M, Wu L, et al. Mapping native disulfide bonds at a proteome scale. Nature Methods. 2015;12:329.
[44] Yang B, Wu Y-J, Zhu M, Fan S-B, Lin J, Zhang K, et al. Identification of cross-linked peptides from complex samples. Nature Methods. 2012;9:904.
[45] Council NR. Guide for the Care and Use of Laboratory Animals. 8th ed. Washington, DC: National Academies Press; 2011.
[46] Finney DJ. Probit Analysis. 3rd ed. New York: Cambridge University Press; 1971.
[47] McGuckin WF, McKenzie BF. Clin Chim. 1958;4:476-9.
[48] Streitz JM, Madden MT, Salo W, Bernadino KP, Deutsch JL, Deutsch JC. Differentiation of mucinous from non-mucinous pancreatic cyst fluid using dual-stained, 1 dimensional polyacrylamide gel electrophoresis. Clin Proteomics. 2014;11:42.
[49] Butler W, Smith JCP, Schenilder H. Biochim Biophys Acta. 1970;219:514-.
[50] Whitmore L, Wallace BA. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers. 2008;89:392-400.

[51] Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. Journal of Applied Crystallography. 2003;36:1277-82.
[52] Franke D, Svergun DI. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. Journal of Applied Crystallography. 2009;42:342-6.
[53] Svergun DI. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys J. 1999;76:2879-86.
[54] Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols. 2015;10:845.
[55] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605-12.
[56] Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 2019;47:W636-W41.
[57] Holm L. Benchmarking fold detection by DaliLite v.5. Bioinformatics. 2019;35:5326-7.
[58] Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol. 1999;292:195-202.
[59] Grant T, Rohou A, Grigorieff N. cisTEM, user-friendly software for single-particle image processing. Elife. 2018;7.
[60] Tang G, Peng L, Baldwin PR, Mann DS, Jiang W, Rees I, Ludtke SJ. EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol. 2007,157: 38-46.
[61] Terwilliger TC. Finding non-crystallographic symmetry in density maps of macromolecular structures. J Struct Funct Genomics 2013;14: 91-5.
[62] Nicholls RA, Tykac M, Kovalevskiy O, Murshudov GN. Current approaches for the fitting and refinement of atomic models into cryo-EM maps using *CCP-EM*. Acta Crystallogr D Struct Biol. 2018,74:492-505.
[63] Hamill OP, Marty A, Neher E, Sakmann B, Sigworth FJ. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. Pflugers Arch. 1981;391:85-100.
[64] Xu Q, Abdubek P, Astakhova T, Axelrod HL, Bakolitsa C, Cai X, et al. Structure of a membrane-attack complex/perforin (MACPF) family protein from the human gut symbiont Bacteroides thetaiotaomicron. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2010;66:1297-305.
[65] Chatzidaki-Livanis M, Coyne MJ, Comstock LE. An antimicrobial protein of the gut symbiont Bacteroides fragilis with a MACPF domain of host immune proteins. Mol Microbiol. 2014;94:1361-74.

Figure legends

Figure 1. Identification of PV2 toxin from the poisonous eggs of *P. maculata*. (A) Egg clutch of the apple snail *P. maculata*. (B) Egg fluid from apple snail eggs was subjected to ultracentrifugation in NaBr gradient and isolation of native PmPV2 by ionic exchange and exclusion columns. *Insets*: ultracentrifugation tube showing PVF fractions and native-PAGE of purified PmPV2 (red arrowheads). (C) Molecular mass determination PmPV2 by SLS. (D) PmPV2 subunit composition analyzed by SDS-PAGE demonstrates the sample is dimeric with a single band corresponding to dimeric PmPV2 shown in lane SDS. Lane SDS+ βME shows a sample that has been deliberately monomerized following incubation with β-mercaptoethanol (βME) as reducing agent.

Figure 2. Tertiary and quaternary structure of PmPV2. (A) Schematic architecture of PmPV2. Domains are shown in orange (MACPF) purple (IMAD) and blue (Lectin) boxes. The cysteine residues involved in the interchain disulfide bond are highlighted in red. (B) 3D homology modeling of PmPV2 subunits highlighting characteristic regions of 6-blade β-propeller lectin domain in PmPV2-31 subunit (left), and MACPF domain in PmPV2-67 (right). (C) Gallery of representative 2D class averages showing the most populated views of the protein. (D) 3D TEM map of PmPV2 obtained from reference-free 2D class averages: left, side view; right, top view. Scale bars are shown. (E) PmPV2 *ab-initio* volume obtained by SAXS (yellow): left, side view; right, top view. Please refer to figure S10 to see a more detailed comparison between SAXS and TEM 3D maps. (F) Rigid-body fitting of MACPF (orange) and Lectin (blue) domains into the two possible hands of the negative-stain TEM density map (transparent gray): top, right-handed; bottom, left-handed. Different orientation are shown to illustrate the fitted domains across the dimer-of-heterodimers.

Figure 3. PmPV2 form pores and perforates membranes. (A) Lectin activity of PmPV2 on erythrocytes (upper panel) and hemagglutinating activity of PmPV2 preincubated with monosaccharides (lower panel). D-mannose (Man), D-galactose (Gal), D-galactosamine (GalNH2), N-acetyl-D-galactosamine (GalNac), D-glucose (Glc), D-glucosamine (GlcNH2), N-

acetyl-D-glucosamine (GlcNac), L-Fucose (Fuc). (B) Patch clamp experiments: Typical whole cell holding current obtained from a Caco-2 cell continuously clamped at -50 mV, before (upper panel) and after extracellular perfusion of PmPV2 (middle panel) or perfused with PmPV2 preincubated with GlnNH2 (lower panel). (C) TEM imaging of PmPV2 pore formed on liposomes. (a) POPC:Cho liposomes without PmPV2 treatment 120kx amplification. Bar 100 nm. (b) Liposomes carrying PmPV2 pore-like structures (arrow). 120kx amplification. Bar 100 nm. (c-d) Top view of ring-like structure form by PmPV2 on the liposome surface at 225kx (c) and 640kx amplification. Bar 20 nm.

Figure 4. Phylogeny and occurrence of IMAD domain. (A) Unrooted phylogenetic tree of homologous sequences of Ct-PmPV2-67. Homologues were retrieved from BLASTp analysis, sequence aligned by MUSCLE and phylogeny reconstructed using MrBayes. Node numbers represent Bayesian posterior probabilities (in percentage) of finding a given clade. (B) Domain architecture of Ct-PmPV2-67 homologue sequences. The boxes highlight the relative position and architecture of the Nt-MACPF and Ct-IMAD domains, as found by ThreaDom and Pfam analysis. AEK10751.1: MACPF domain containing protein (*Mytilus galloprovincialis*); P0C8G6.2: Perivitelline-2 67 kDa subunit (*Pomacea canaliculata*); AMZ02450.1: Perivitellin-2 67 kDa subunit-like (*Littorina littorea*); XP_013094890.1, XP_013093733.1, XP_013094888.1, XP_013095051.1, and XP_013094774.1: Perivitellin-2 67 kDa subunit-like (*Biomphalaria glabrata*); RUS72664.1: Hypothetical protein (*Elysia chlorotica*); XP_005112978.1: Perivitellin-2 67 kDa subunit-like (*Aplysia californica*); RWR99060.1, and RWS10239.1: Perivitellin-2 67 kDa subunit-like (*Dinothrombium tinctorium*); XP_013787111.1: Perivitellin-2 67 kDa subunit-like (*Limulus polyphemus*); XP_006817276.1: Perivitellin-2 67 kDa subunit-like (*Saccoglossus kowalevskii*); XP_019620290.1: Uncharacterized protein (*Branchiostoma belcheri*); XP_014669680.1: Uncharacterized protein (*Priapulus caudatus*); XP_028408379.1: Perivitellin-2 67 kDa subunit-like (*Dendronephthya gigantea*).

Fig. 5. Schematic representation of MACPF proteins highlighting their domain assembly. Immune proteins includes C6, C7, C8α, C8β and C9 MAC, perforin-1 and perforin-2

(macrophage expressed gene-1, MPEG-1) from animals and CAD1 from plants [1]; Developmental MACPFs includes a MPEG-1 orthologous from trophoblast (EPCS50), astrotactins (ASN1 and ASN2), apextrin, bone morphogenic protein/retinoic acid-inducible neural-specific proteins (BRINP1, BRINP2 and BRINP3) and torso-like from animals and the conidium-specific SpoC1-C1C protein and pleurotolysin (Ply) from fungi [1]; Venoms or poisons includes several cnidarian toxins (PsTX-60A, PsTX-60B and AvTx-60A) [4,5], the stone fish sturnatoxin (SNTX) [3] and perivitellin-2 (PV2) from *Pomacea* snails [6,7,8 this study]. Bacteria exotoxins include *Photorabdus luminicese-MACPF* (Plu-MACPF), and the CT153, BSAP-1 and BT_3439 proteins [1,64,65]. Protist MACPF are found in apicomplexan as *Plasmodium falciparum* perforin-like protein (PfPLP), the sporozoite micronemal protein essential for cell traversal-1 (SPECT2) and the membrane-attack ookinete protein (MAOP) [1]. ANX: annexin-like domain; ApeC: apextrin C-terminal domain; C2: calcium-dependent plasma membrane-binding domain; CCP: sushi (CCP/CCR) domain; EGF-like: epidermal Growth Factor domain; FAT: focal adhesion-targeting domain; FIMAC: factor I module MAC domain; FN3: fibronectin type III domain; IMAD: invertebrate MACPF accessory domain; LDLRA: low density lipoprotein receptor class A; MABP: MVB12-associated β-prism domain; MIR: mannosyltransferase, inositol triphosphate receptor domain; THX: thioredoxin; TM: transmembrane region; TSP: thrombospondin domain; PRYSPRY: PRY spIa and ryanodine receptor domain.

Declarations

Ethics approval and consent to participate

The experiment with mice was approved by the "Comité Institucional para el Cuidado y Uso de Animales de Laboratorio" (CICUAL) of the School of Medicine, Universidad Nacional de La Plata (UNLP) (Assurance No. P 01012016) and were carried out in accordance with the Guide for the Care and Use of Laboratory Animals (Guide for care and use of laboratory animals. Washington: Academic Press; 2011).

Authors' contributions

SI MG MSD SM VM HH conceived and designed the experiments. MG SI JC SM TB VM MSD JI LHO performed the experiments. MG JI JWQ MSD SI VM SB EP JC LHO HH analysed the data. HH JWQ JC VM LHO contributed reagents/materials/analysis tools. MG SI MSD JC SM TB LHO VM JWQ JI HH wrote the paper.

Funding

This work was supported by funding from Ministry of Science and Technology of Argentina Grant (Agencia Nacional de Promoción Científica y Tecnológica, PICT 2014-0850 to HH and PICT 2013-0122 to SI.), Consejo Nacional de Investigaciones Científicas y Técnicas, CONICET (PIP 0051 to HH), General Research Fund of Hong Kong (HKBU 12301415 to JW Q), and by partial financial support from LNLS–Brazilian Synchrotron Light Laboratory/MCT (Project SAXS1-17746), and LNNano (Project TEM 24346).

CrediT Author Statement

SI MG MSD SM VM HH conceived and designed the experiments.

MG SI JC SM TB VM MSD JI LHO performed the Investigation, experiments.

MG JI JWQ MSD SI VM SB EP JC LHO HH analysed the data.

HH JWQ JC VM LHO contributed with funding acquisition for reagents/materials/analysis tools.

MG SI MSD JC SM TB LHO VM JWQ JI HH wrote the original draft paper, review and editing.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

M.L. Giglio, S. Ituarte, V. Milesi, M.S. Dreon, T.R. Brola, J. Caramelo, J.C.H. Ip, S. Maté, J.W. Qiu, L.H. Otero and H. Heras

Highlights

- MACPF pore-forming proteins with toxic function are very rare in animals.
- PmPV2 is a dimer in solution and protomers combine immune proteins into an AB toxin.

- PmPV2 tachylectin binds to membrane glycans acting as a delivery subunit.
- PmPV2 MACPF subunit forms pores and alters membrane conductance.
- The structure of PmPV2 from snail eggs differs from all known MACPFs.