Expansion and Neofunctionalization of Actinoporin-like Genes in Mediterranean Mussel (*Mytilus galloprovincialis*)

Neža Koritnik[1,†], Marco Gerdol[2,†], Gašper Šolinc[1], Tomaž Švigelj[1], Simon Caserman[1], Franci Merzel[3], Ellie Holden[4,5], Justin LP Benesch[4,5], Francesco Trenti[6], Graziano Guella[6], Alberto Pallavicini [ID][2,7], Maria Vittoria Modica[8], Marjetka Podobnik[1], and Gregor Anderluh [ID][1,*]

[1]Department of Molecular Biology and Nanobiotechnology, National Institute of Chemistry, Ljubljana, Slovenia

[2]Department of Life Sciences, University of Trieste, Trieste, Italy

[3]Theory Department, National Institute of Chemistry, Ljubljana, Slovenia

[4]Department of Chemistry, Oxford University, Oxford, UK

[5]Kavli Institute for Nanoscience Discovery, University of Oxford, Oxford, UK

[6]Bioorganic Chemistry Lab, Department of Physics, University of Trento, Via Sommarive 14, 38123 Povo, Italy

[7]Division of Oceanography, National Institute of Oceanography and Applied Geophysics, Trieste, Italy

[8]Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, Napoli, Italy

[†]Contributed equally to the work presented in the paper.

[*]Corresponding author: E-mail: gregor.anderluh@ki.si.

Abstract

Pore-forming toxins are an important component of the venom of many animals. Actinoporins are potent cytolysins that were first detected in the venom of sea anemones; however, they are occasionally found in animals other than cnidarians and are expanded in a few predatory gastropods. Here, we report the presence of 27 unique actinoporin-like genes with monophyletic origin in *Mytilus galloprovincialis*, which we have termed mytiporins. These mytiporins exhibited a remarkable level of molecular diversity and gene presence–absence variation, which warranted further studies aimed at elucidating their functional role. We structurally and functionally characterized mytiporin-1 and found significant differences from the archetypal actinoporin fragaceatoxin C. Mytiporin-1 showed weaker permeabilization activity, no specificity towards sphingomyelin, and weak activity in model lipid systems with negatively charged lipids. In contrast to fragaceatoxin C, which forms octameric pores, functional mytiporin-1 pores on negatively charged lipid membranes were hexameric. Similar hexameric pores were observed for coluporin-26 from *Cumia reticulata* and a conoporin from *Conus andremenezi*. This indicates that also other molluscan actinoporin-like proteins differ from fragaceatoxin C. Although the functional role of mytiporins in the context of molluscan physiology remains to be elucidated, the lineage-specific gene family expansion event that characterizes mytiporins indicates that strong selective forces acted on their molecular diversification. Given the tissue distribution of mytiporins, this process may have broadened the taxonomic breadth of their biological targets, which would have important implications for digestive processes or mucosal immunity.

Key words: pore-forming toxins, actinoporins, actinoporin-like proteins, *Mytilus galloprovincialis*, mytiporins, protein pores.

Significance

The evolution of toxin-related proteins is not well understood. We discover an expansion of the actinoporin-related protein family with monophyletic origin in *Mytilus galloprovincialis*. The representative protein has altered functional properties compared to archetypal toxins. We show that the conserved structural toxin fold has independently acquired different functions during animal evolution. Such new structures may have contributed to the venom potency in some phyla and to feeding or immune responses in other phyla.

Introduction

The plasma membrane of cells represents one of the first barriers to external attack, and therefore it is not surprising that many toxins act at the level of the plasma membrane (Menestrina et al. 2003). Pore-forming toxins (PFTs) are proteins that are often used for predation or defense. They can form pores in the lipid membranes of cells and thus damage the integrity of cells. They are ubiquitously distributed in nature and are an important component of the venom of many animals (Peraro and van der Goot 2016; Podobnik and Anderluh 2017; Verma et al. 2021).

Actinoporins are PFTs originally discovered in sea anemones (Kem 1988; Maček 1992; Anderluh and Maček 2002). They represent one of the best-studied PFT families, and thus much is known about their structure and mechanism of action (Rojko et al. 2016). The most studied actinoporins are from sea anemones and include fragaceatoxin C (FraC) from *Actinia fragacea*, equinatoxin II (EqtII) from *Actinia equina*, and sticholysins from *Stichodactyla helianthus* (Rojko et al. 2016; Rivera-de Torre et al. 2020; Caaveiro and Tsumoto 2021). Actinoporins can recognize the membrane lipid sphingomyelin (SM) and can thus efficiently bind to SM-containing lipid membranes (Bakrač et al. 2008; Bakrač and Anderluh 2010). SM is recognized by a binding site located on one side of the β-sandwich, which is the central structural element of actinoporins (Bakrač et al. 2008; Tanaka et al. 2015). After binding, the N-terminus is transferred into the lipid membrane in an α-helix arrangement (Hong et al. 2002; Drechsler et al. 2006; Tanaka et al. 2015). Several monomers then oligomerize at the membrane plane, and finally, the N-terminal α-helices are transferred across the membrane to form a functional channel lined by the cluster of α-helices (Kristan et al. 2007; Tanaka et al. 2015). Actinoporins thus represent α-helical PFTs (Gouaux 1997; Peraro and van der Goot 2016). Crystallographic studies have revealed an octameric arrangement of FraC (Tanaka et al. 2015), and different stoichiometries have been proposed for other actinoporins but have not been characterized at the structural level yet (Belmonte et al. 1993; Baker et al. 2014; Rojko et al. 2016).

Recent studies have suggested that many proteins that are similar to actinoporins, that is actinoporin-like proteins (ALPs), are widely distributed (Gutierrez-Aguirre et al. 2006; Macrander and Daly 2016). ALPs have been found in other cnidarian species, such as *Hydra* (Sher et al. 2005) and the reef-building coral *Stylophora pistilata* (Ben-Ari et al. 2018), but also in fish, plants (Gutierrez-Aguirre et al. 2006), sponges (Sandoval and McCormack 2022), and multiple species of molluscs. Three ALPs, named echotoxins, were purified from the salivary gland of the marine snail *Monoplex echo* and were found to be hemolytic and lethal to mice (Shiomi et al. 2002). They were also effectively inhibited by gangliosides, which were proposed to be their receptors in erythrocyte membranes (Shiomi et al. 2002). Transcriptomic and proteomic approaches revealed the presence of ALPs in the salivary gland of the closely related giant triton *Charonia tritonis* (Bose et al. 2017). ALPs have been detected in the venom glands and in some cases also in the salivary glands of a large number of conoidean snails, such as various species of *Conus* (Leonardi et al. 2012; Safavi-Hemami et al. 2014; Pardos-Blas et al. 2019, 2021, 2022; Sudewi et al. 2019; Fouda et al. 2021), the drilliids *Clavus canalicularis* and *Clavus davidgilmouri* (Lu et al. 2020), and the terebrids *Triplostephanus anilis* and *Terebra subulata* (Gorson et al. 2015). Although multiple ALP isoforms were reported in some *Conus* species (e.g., nine isoforms in *Conus geographus* (Safavi-Hemami et al. 2014)), it is in the vampire snail *Cumia reticulata* that a remarkable ALP family expansion was subsequently discovered (Modica et al. 2015; Gerdol et al. 2018). *Cumia reticulata* ALPs, coluporins, represent a well-diversified protein family that is present in many copies and with a tissue-specific expression pattern (Gerdol et al. 2018). In bivalves, an ALP, clamlysin, was purified from the freshwater clam *Corbicula japonica* and partially functionally characterized, revealing its SM-specificity and other functional features reminiscent of actinoporins (Takara et al. 2011).

The biological role of actinoporins is not yet fully understood. Actinoporins constitute a significant component of the venom of sea anemones and, along with sodium and potassium channel inhibitors (Norton 2009; Jouiaei et al. 2015), were proposed to defend against predators, enable these sessile organisms to prey, prevent adhesion of other organisms, and participate in intraspecies aggression (Maček 1992; Anderluh and Maček 2002). In *Hydra*, it has been suggested that PFTs facilitate the digestion of prey, destroying cells after pore formation due to osmotic imbalance (Sher et al. 2008). A similar role was ascribed to coluporins, which may have a profound role in the

feeding behavior of vampire snails, possibly acting in synergy with other venom components and enhancing envenomation efficacy (Gerdol et al. 2018). PFTs have also a well-established role as components of the arsenal of toxic molecules that may facilitate the elimination of invading pathogens (including bacterial and eukaryotic parasites) in the context of the innate immune response (McCormack et al. 2013; Verma et al. 2021). A nice example is the aerolysin-like PFT lysenin, which is expressed by the earthworm (Podobnik, Kisovec et al. 2017) and plays an important role in eliminating parasitic cells from the coelomic fluid (Shogomori and Kobayashi 2008; Opper et al. 2013). A completely different function has been suggested for plant ALPs. For example, bryoporin may play a role in water stress management in moss (Hoang et al. 2009).

In this article, we report the discovery of an ALP from the Mediterranean mussel *Mytilus galloprovincialis* with unusual extensions at both the N- and C-termini. Usually, the C-terminus of ALPs is well defined by the length of the last β-strand in the β-sandwich, whereas extensions at the N-terminus can prevent the permeabilizing activity of some actinoporins, for example EqtII (Hong et al. 2002; Kristan et al. 2007). A comprehensive genomic analysis aimed at identifying homologous sequences in the reference pan-genome of *M. galloprovincialis* identified 26 additional sequences belonging to the mytiporins gene family, likely originating from a relatively recent lineage-specific gene family expansion event. A phylogenetic analysis of metazoan ALPs was conducted to identify major evolutionary patterns for this family. Furthermore, a detailed functional and structural analysis of the first member of this large mussel ALP family, mytiporin-1, revealed altered functional properties compared with anthozoan actinoporins, which were shared with some other ALPs from venomous snails. Collectively, our results suggest that the conserved structural fold of actinoporins, associated with several previously described PFTs, may have independently acquired different functions during the course of animal evolution. Such new structures may have contributed to venom potency in some phyla and to feeding or the immune response in other phyla.

Results

The Diversity of Mytiporins in *Mytilus galloprovincialis*

The original identification of an ALP from *M. galloprovincialis* derives from a screening of the database of expressed sequence tags dbEST (Boguski et al. 1993), which was conducted with the algorithm tblastn (Altschul et al. 1997) using the echotoxin 2 sequence as a query (Uniprot ID Q76CA2) (Kawashima et al. 2003). Assembled expressed sequence tags from *M. galloprovincialis* yielded a putative protein consisting of 230 amino acids (supplementary fig. 1, Supplementary Material online). The first 19 amino acids were predicted to be a signal peptide (Teufel et al. 2022).

We named the putative mature protein mytiporin-1 (MYTP1). MYTP1 is clearly a homolog of actinoporins from sea anemones, such as EqtII and FraC. It shares 21% identical and 51% similar amino acids with FraC (supplementary fig. 2, Supplementary Material online). The release of the complete genome sequence of *M. galloprovincialis* in 2020 (Gerdol et al. 2020) further corroborated the correctness of the MYTP1 sequence, even though a few nonsynonymous substitutions were identified that likely derive from allelic variants, as will be discussed below.

We next exploited the availability of this genomic resource to obtain a comprehensive overview of the distribution of actinoporin-like genes in this species. Overall, a total of 27 different mytiporin genes were identified in *M. galloprovincialis*, of which 19 were present in the reference genome (i.e., LOLA; (Gerdol et al. 2020)), whereas the remaining 8 were found in multiple resequenced individuals. Just ten out of 27 mytiporin genes, i.e. mytiporin-1, -3, -4, -5, -9, -17, -19, -21, -25, and -27, belonged to the *core* gene set, and the remaining 17 genes were *dispensable* and therefore predicted to be subjected to presence–absence variation (fig. 1A). Of these, some were only observed in a few individuals (e.g., mytiporin-18 and -23 were only found in 3 out of 16 mussels), whereas others were widespread in the analyzed dataset (e.g., mytiporin-2 was found in 15 out of 16 mussels). The number of mytiporin genes found in each individual ranged from 15 to 19, averaging 17. As previously reported (Gerdol et al. 2020), mussel dispensable genes are associated with genomic regions that can be found in a homozygous, hemizygous or nullizygous state in any given individual, depending on the combination between parental haplotypes, and should therefore not be interpreted as highly divergent allelic variants belonging to the same locus.

The 27 *M. galloprovincialis* mytiporin sequences displayed a relatively high diversity, with pairwise amino acid sequence similarity as low as 15% within the porin domain region and variable length and domain architecture (supplementary table S1 and fig. 3, Supplementary Material online). The size of the mytiporin precursors ranged from 176 amino acid residues (in mytiporin-22) to 462 amino acid residues (in mytiporin-9). Seven mytiporins displayed a signal peptide for secretion into the extracellular space (fig. 1A). While its absence would suggest a cytosolic location for the remaining sequences, SecretomeP (Bendtsen et al. 2004) revealed that leaderless secretion was very likely for 12 mytiporins and doubtful for 7 mytiporins due to borderline NN-scores. In summary, most mytiporins are probably targeted to the extracellular environment, either through classical or leaderless secretion. None of the 27 *M. galloprovincialis* mytiporins displayed transmembrane regions or additional domains in the N-terminus of variable length. The actinoporin domain was, in all cases, positioned at the C-terminus of the protein.
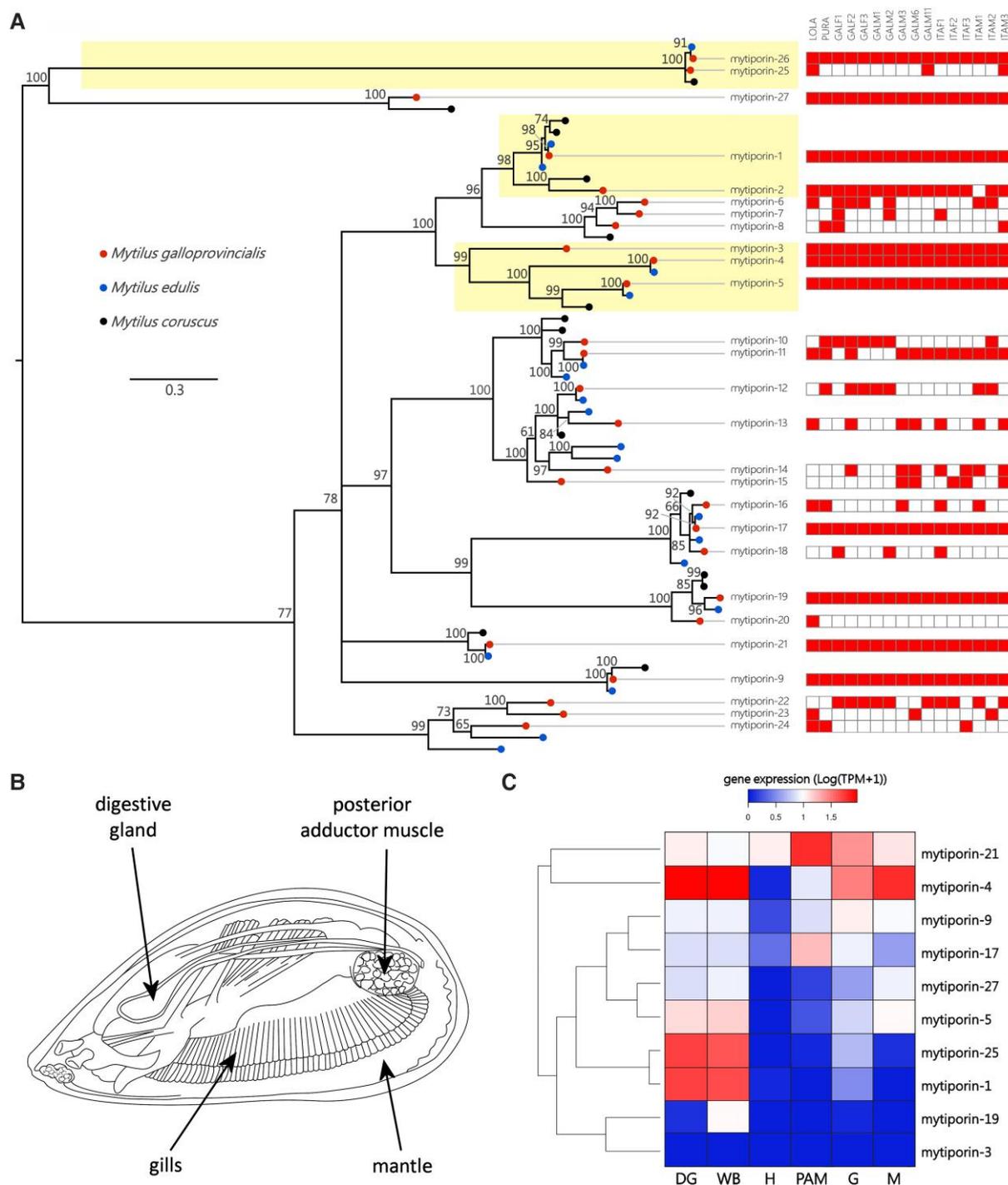
Fig. 1—Actinoporin-like genes in *Mytilus* spp. (*A*) Bayesian phylogeny of *Mytilus* spp. mytiporins, based on the multiple sequence alignment of the porin domain. The names of the sequences refer to those identified in *M. galloprovincialis* and discussed in the text. The sequences that display a classical signal peptide for secretion are highlighted. The numbers shown near each node represent posterior probability support values. Nodes with low posterior probability support (i.e., < 0.5) were collapsed. The tree is represented as midpoint-rooted phylogram. The presence or the absence (full and empty squares, respectively; right) of any given mytiporin gene in the reference genome (LOLA) and in 15 additional individuals subjected to whole-genome resequencing (PURA indicates mussels studied by Murgarella et al. (2016); the prefixes "GAL" and "ITA" indicate mussels from two distinct populations, as reported in the main text). (*B*) A schematic representation of the internal anatomy of a *M. galloprovincialis* adult individual. (*C*) Gene expression profile of the ten core mytiporin genes in six tissues. The cladogram clustered mytiporins with similar tissue distribution. Gene expression levels are reported as transcripts per million (TPM; see Materials and Methods for details). DG, digestive gland; G, gills; H, hemolymph; M, mantle; PAM, posterior adductor muscle, WB, whole body.

Table 1

The Number of ALP-encoding Genes Identified in Fully Sequenced Molluscan Genomes

Bivalvia		Gastropoda		Polyplacophora	
Archivesica marissinica	0	*Gigantopelta aegis*	0	*Acanthopleura granulata*	0
Argopecten irradians irradians	7	*Elysia chlorotica*	0		
Argopecten purpuratus	8	*Lottia gigantea*	0	**Cephalopoda**	
Bathymodiolus platifrons	14	*Marisa cornuarietis*	0	*Octopus sinensis*	0
Crassostrea gigas	1	*Pomacea maculata*	0	*Haplochaena maculata*	0
Crassostrea virginica	1	*Pomacea canaliculata*	0	*Watasenia scintillans*	0
Cyclina sinensis	6	*Lanistes nyassanus*	0	*Octopus bimaculoides*	0
Dreissena rostriformis	12	*Achatina fulica*	0	*Sepia pharaonis*	0
Lutraria rhynchaena	0	*Achatina immaculata*	0		
Magallana hongkongensis	1	*Conus consor*	3		
Megalonaias nervosa	2	*Lautoconus ventricosus*	2		
Mizuhopecten yessoensis	4	*Haliotis rufescens*	1		
Modiolus philippinarum	2	*Hliotis discus*	0		
Mytilus coruscus	14	*Haliotis laevigata*	0		
Mytilus edulis	19	*Aplysia californica*	0		
Panopaea generosa	2	*Chrysomallon squamiferum*	3		
Pecten maximus	3	*Dracogyra subfuscus*	0		
Perna viridis	8	*Biomphalaria glabrata*	0		
Pinctada fucata	2				
Ruditapes philippinarum	1				
Saccostrea glomerata	1				
Scapharca broughtonii	1				
Sinonovacula constricta	0				

In parallel with *M. galloprovincialis*, the recently released genomes of the congeneric species *M. edulis* and *M. coruscus* were also analyzed, revealing a similar number of ALPs, that is 19 and 14, respectively. Of note, additional *dispensable* mytiporin genes not represented in the reference genome assemblies might be present in both species. Phylogenetic inference showed that most *M. galloprovincialis* genes had orthologous genes in both sister species, even though the orthology ratio was often not 1:1, particularly in the case of *dispensable* genes (fig. 1*A*).

It is also worth noting that several dispensable mytiporin genes, such as mytiporin-22, -23, and -24, or mytiporin-12, -13, -14, and -15, were placed with high posterior probability support within the same branches, indicating a possible recent origin by gene duplication, followed by rapid sequence diversification. This observation is consistent with the data collected from the analysis of the mussel pangenome, which have previously revealed a correlation between gene PAV and gene families subjected to recent lineage-specific expansion events (Gerdol et al. 2020).

The Gene Expression Profile of Mytiporins

The ten core mytiporin genes displayed broadly different expression patterns, displaying significant variability both in terms of tissue specificity and peak expression values (fig. 1*B* and *C*). Mytiporin-3 and -19 were not expressed at biologically significant levels under physiological conditions, that is they did not exceed 1 transcript per million

(TPM), which is usually considered the threshold level to define a gene as expressed. Three mytiporins with a canonical signal peptide were strongly expressed in the digestive gland: mytiporin-1, -4, and -25. Mytiporin-4 was also expressed at significant levels (>10 TPM) in gills and mantle. Mytiporin-21 was expressed in multiple tissues, but mostly in the posterior adductor muscle. The other mytiporins did not show strong tissue specificity, being expressed at moderate levels (usually < 10 TPM) in the digestive gland, whole body, gills, and mantle. Curiously, no mytiporin (except mytiporin-21) was expressed at significant levels in the hemolymph.

Distribution of Actinoporin-like Genes in Mollusca

Our comparative genomics analysis highlighted a highly asymmetric distribution of genes encoding ALPs in Mollusca. In general, Bivalvia showed a higher number of such genes (mean = 5.25) than Gastropoda (mean = 0.5), in which ALPs were often completely missing (in 12 out of the 16 species analyzed). No ALP-encoding genes could be detected in Cephalopoda and Polyplacophora. Nevertheless, the low number of species with fully sequenced genomes available (five and one, respectively) is currently not sufficient to rule out the possibility that ALP genes are present in other members belonging to these two molluscan classes (table 1). The observations reported for Gastropoda are consistent with the data collected in a previous study (Gerdol et al. 2018), which was primarily

based on transcriptomic data. ALP-encoding genes were rarely found, with a few exceptions linked with lineage-specific expansions that occurred in species with peculiar feeding habits, such as coluporins in *C. reticulata* (Gerdol et al. 2018). Conversely, the number of ALP-encoding genes in Bivalvia was highly variable, ranging from zero (in three species) to 20 (in *M. edulis*), with most bivalve genomes containing only one or two genes. Few exceptions were represented by scallops from the family Pectinidae (three to eight genes), the Chinese venus *Cyclina sinensis* (six genes), the freshwater zebra mussel *Dreissena rostriformis* (12 genes), and marine mussels from the family Mytilidae. A lineage-specific gene family expansion event was evident in Mytilidae because of the identification of eight ALP-encoding genes in *Perna viridis*, 14 ALP-encoding genes in *Bathymodiolus platifrons,* and > 10 ALP-encoding genes in three *Mytilus* species, as anticipated in the previous sections (fig. 1A). Specifically, 14 and 19 ALP-encoding genes were identified in *M. coruscus* and *M. edulis*, respectively. *Modiolus philippinarum* was an exception, as its genome encoded only two ALP-encoding genes.

From an evolutionary perspective, the monophyletic origins of all mytiporins, except mytiporins-25 and -26 (see the Discussion), was highly supported (posterior probability = 1), implying a shared evolutionary origin from a single ancestral mytiporin gene. Moreover, as previously revealed by another study (Gerdol et al. 2018), all three major groups of molluscan ALPs (i.e., mytiporins, conoporins, and coluporins), were grouped in a single highly supported (posterior probability = 1) monophyletic cluster. This cluster was independent from cnidarian ALPs (actinoporins and ALPs from *Hydra*), suggesting a shared evolutionary origin from a prototypical ALP-encoding gene present in the latest common ancestor of all Mollusca (fig. 2).

Lipid Membrane Binding of Mytiporin-1

Next, we investigated the functional and structural properties of a typical mytiporin. We chose MYTP1 because of its high expression, unusual tissue distribution (fig. 1) and extensions at the N- and C-termini, which may affect functional and structural properties of the pore. MYTP1 is expressed in the digestive gland (fig. 1C), suggesting its important role in digestion. We compared the properties of MYTP1 with those of FraC (which was used in this study as a control representing a typical actinoporin) and several other molluscan ALPs. We chose coluporin-16 (Col16) and -26 (Col26) from *C. reticulata* (Gerdol et al. 2018) and a conoporin from *Conus andremenezi* (Cand) to represent each of the major groups of molluscan ALPs (supplementary fig. 2, Supplementary Material online, fig. 2). We expressed all proteins in *E. coli* and purified them to homogeneity (supplementary fig. 4, Supplementary Material online).

First, we investigated the ability of MYTP1 to bind to lipid membranes. Interestingly, MYTP1 significantly bound only to multilamellar vesicles (MLVs) composed of negatively charged 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoglycerol (POPG). Conversely, it did not bind to MLVs composed of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC; a phosphatidylcholine i.e. regularly used as a negative control lipid for membrane interactions of actinoporins, e.g., EqtII (Hong et al. 2002; Rojko et al. 2013)), POPC:cholesterol (chol) 1:1 (all stated lipid ratios are molar), or POPC:SM 1:1 (fig. 3). By contrast, FraC bound particularly well to POPC:SM 1:1 membranes, consistent with its preference for SM-containing membranes (Bellomio et al. 2009). Of all the molluscan proteins tested, Cand bound almost indiscriminately to all membranes (fig. 3). Col16 and Col26 behaved similarly to MYTP1 and showed the strongest binding to POPG membranes (fig. 3). Furthermore, MYTP1 binding was dependent on salt concentration ([NaCl]). MYTP1 did not bind to POPG membranes in the presence of 500 mM NaCl (fig. 3).

Permeabilizing Activity of Mytiporin-1

Next, we examined the release of the fluorescent probe calcein from large unilamellar vesicles (LUVs). FraC and MYTP1 did not release calcein from control, POPC-containing LUVs. FraC (at > 1 nM) showed almost complete calcein release from LUVs composed of POPC:SM 1:1. By contrast, MYTP1 was much less efficient (supplementary fig. 5a, b, Supplementary Material online). However, it was very efficient in releasing calcein from POPG-containing LUVs, whereas FraC permeabilized these LUVs much less efficiently (supplementary fig. 5a, b, Supplementary Material online). For MYTP1, calcein release was dependent on [NaCl] (supplementary fig. 5c, Supplementary Material online). All these results are consistent with the results regarding lipid membrane binding (fig. 3).

Pore Formation by Mytiporin-1 in Planar Lipid Bilayers

We then investigated the pore-forming activity of MYTP1 in planar lipid bilayers, an approach that allows direct visualization of pore formation in membranes by measuring electrical currents (Marchioretto et al. 2013). We observed stepwise increases in electrical currents after the addition of MYTP1 monomers (at 8.4 µM) to 1,2-diphytanoyl-*sn*-glycero-3-phospho-choline (DPhPC):POPG 1:1 lipid membranes in a buffer with a low salt concentration (100 mM NaCl) and at an applied potential of −50 mV. We interpret these increases, which were usually observed less than 5 min after the measurements began, as the formation and insertion of MYTP1 pores into the membranes. After pore insertion, the current traces exhibited large noise, indicating the dynamic nature of the pores (fig. 4A). Although most of the inserted pores had a well-defined size (conductance of 395 ± 6 pS; $n = 58$,
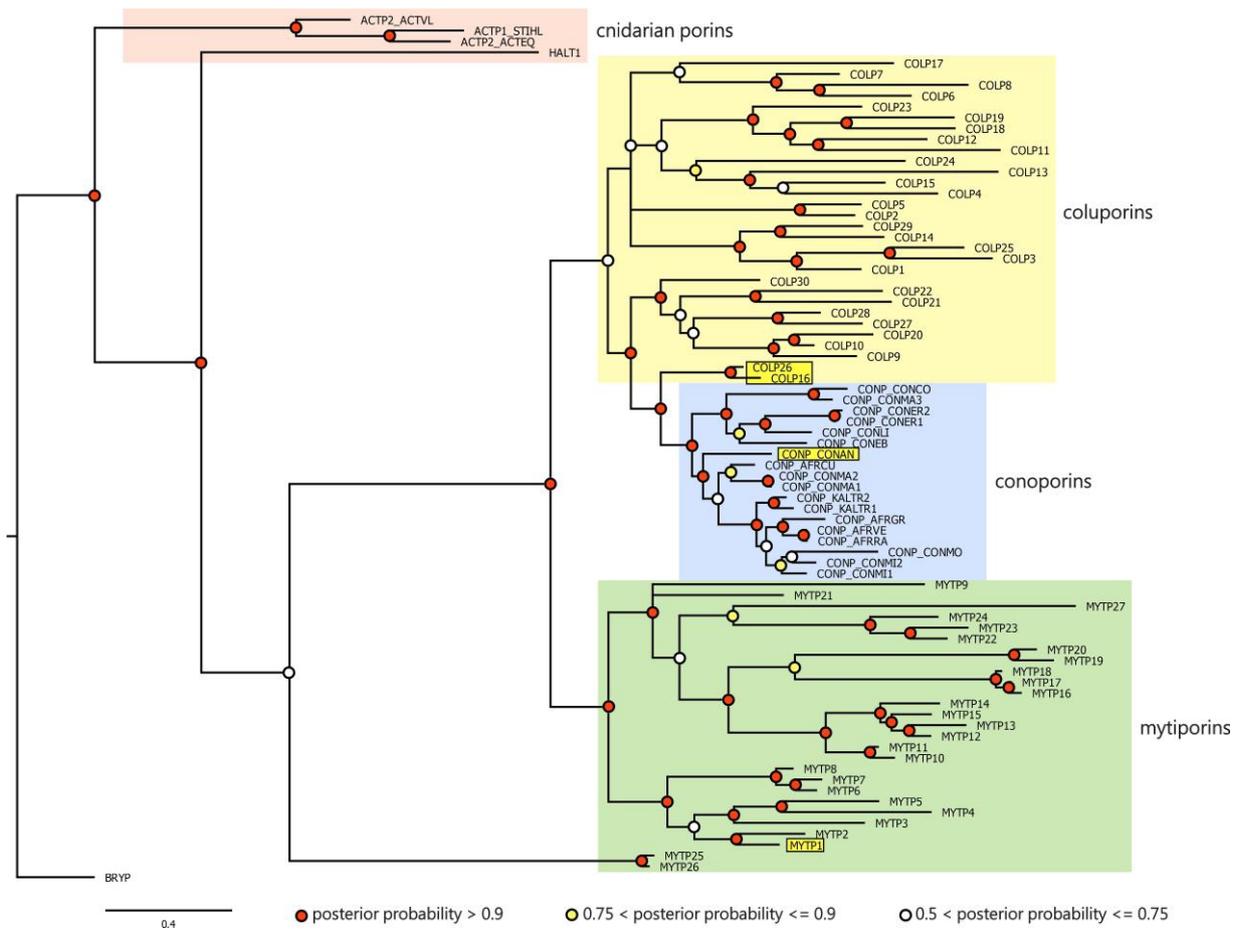
Fig. 2—Bayesian phylogeny of selected metazoan porins characterized by the presence of the Pfam domain PF06369. Molluscan mytiporins, coluporins, and conoporins and cnidarian actinoporins and ALPs were selected. The *Physcomitrella patens* bryoporin, a plant actinoporin-like protein, was used as an outgroup for tree-rooting purposes. ACTEQ, *Actinia equina*; ACTVL, *Actineria villosa*; AFRCU, *Africonus cuneolus*; AFRGR, *Africonus graham*; AFRRA, *Africonus raulsilvai*; AFRVE, *Africonus verdensis*; BRYP, bryoporin; COLP, coluporin; CONAN, *Conus andremenezi*; CONCO, *Conus consors*; CONEB, *Conus ebreaus*; CONER, *Conus ermineus*; CONLI, *Conus lividus*; CONMA, *Conus magus*; CONMI, *Conus miliaris*; CONMO, *Conus monile*; CONP, conoporin; HALT1: *Hydra magnipapillata* actinoporin-like toxin 1; KALTR, *Kalloconus trochulus*; MYTP, mytiporin; STIHL, *Stichodactyla helianthus*. The posterior probability support for the node is denoted with colored circles. The ALPs characterized in this study are highlighted in yellow boxes. Cnidarian actinoporins used in this analysis represent some of the most studied representatives, for example EqtII (ACTP2_ACTEQ), sticholysin I (ACTP1_STIHL), and HALT-1 (HALT1).

mean ± standard error of fit with a Gaussian distribution) (fig. 4B), some inserted pores were much larger (approximately 5% of the pores exhibited conductance greater than 800 pS). The voltage–current relationship shows an approximately linear relationship (fig. 4C). Pore insertions in DPhPC:POPG 1:1 membranes at the higher salt concentration (500 mM NaCl) were observed only when a higher negative potential was used (−150 mV) and showed unstable insertions and high noise (fig. 4A). No insertions were observed at 1000 mM NaCl (fig. 4D). Similarly, no insertions into DPhPC membranes were observed at 100 mM NaCl (fig. 4D). These results are consistent with the finding that MYTP1 activity requires negatively charged lipid membranes and is salt-dependent, as observed in the binding (fig. 3) and permeabilization assays (supplementary fig. 5, Supplementary Material online).

Activity of Mytiporin-1 on Cells

We next examined the activity of MYTP1 on some model cell lines and cell isolates. We checked the permeabilizing activity induced by different PFTs on bovine and fish red blood cells, the mouse macrophage cell line Raw 264.7, the fish skin cell line EPC, and *M. galloprovincialis* hemocytes. MYTP1-induced hemolysis of bovine erythrocytes at a concentration approximately 10-fold higher than that of FraC. Coluporins were found to be even less effective (fig. 5A and B). The hemolytic activity of FraC on bovine and fish erythrocytes was comparable, and MYTP1 activity was not detected on fish erythrocytes (fig. 5B). PFT-induced permeabilization of plasma membranes in Raw 264.7 and EPC cell lines was detected as increased fluorescence after entry of the membrane-impermeable fluorescent nuclear
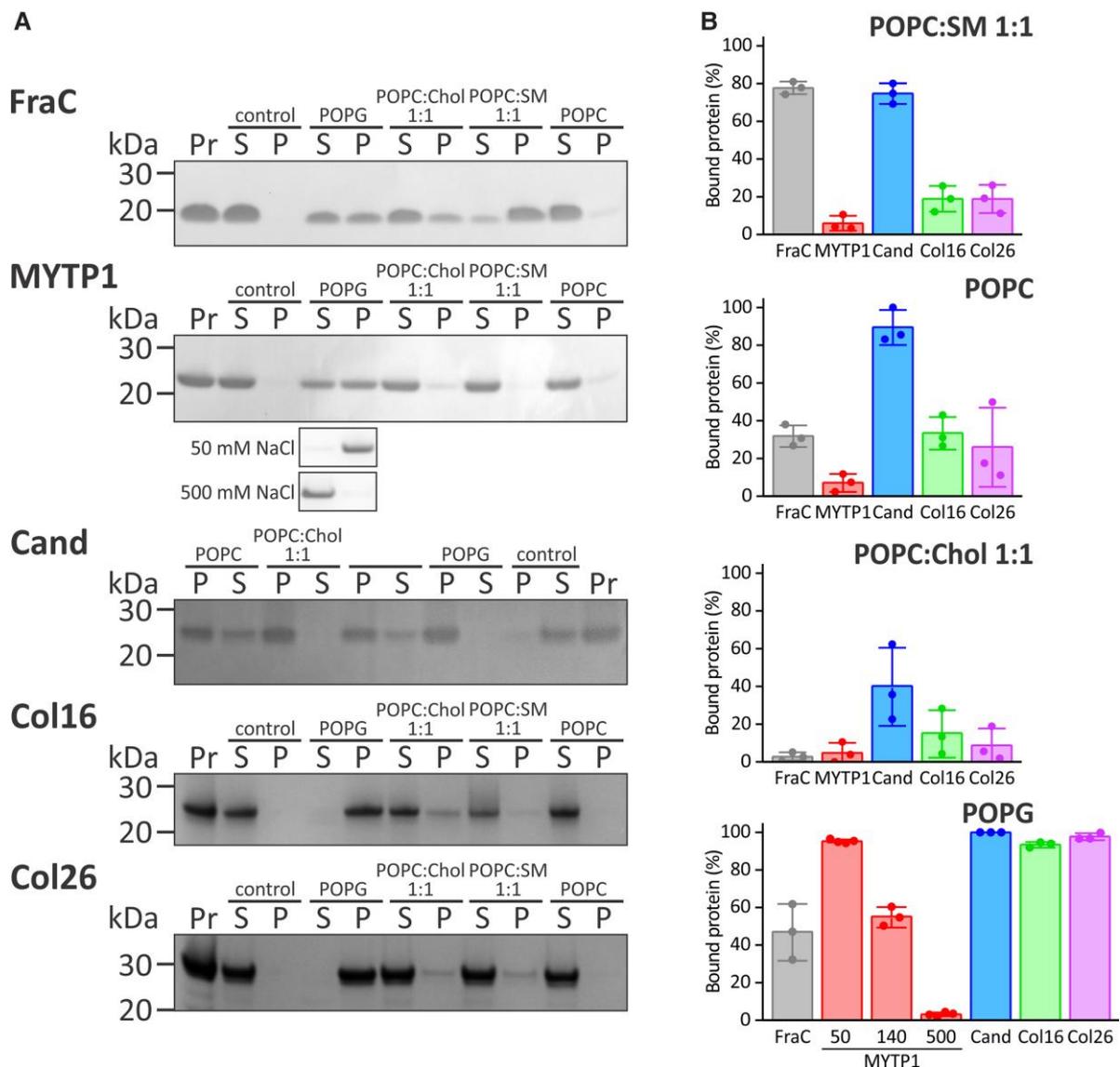
Fig. 3—Membrane binding by mytiporin-1 (MYTP1) and ALPs. (*A*) Binding of fragaceatoxin C (FraC) and other molluscan ALPs. Pr, total amount of protein used in the assay; S, supernatant; *P*, pellet; control, the sample in the absence of multilamellar vesicles; POPC, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine; POPG, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoglycerol; Chol, cholesterol; SM, sphingomyelin. The [NaCl] was 140 mM in phosphate-buffered saline, except when indicated differently (in the MYTP1 sample). (*B*) Quantification of bands from *A* for each composition of lipid membranes. Mean ± standard deviation and individual data points are shown for three independent experiments. For MLVs composed of POPG, three different salt concentrations (in mM) are shown for MYTP1.

stain SYTOX™ Green (fig. 5*C* and *D*). In these cells, the difference in activity between FraC and MYTP1 and the delay in signal onset in the case of MYTP1 was even more pronounced. While MYTP1 did not permeabilize Raw 264.7 cells, it weakly permeabilized EPC cells (fig. 5*C* and *D*). By contrast, FraC and MYTP1 exhibited no permeabilization activity in mussel hemocytes (fig. 5*E*). In their presence, hemocyte nuclei remained unstained but were stained in the presence of 0.5% Triton X-100 (a positive control) or the PFT perfringolysin O from *Clostridium perfringens* (which forms pores approximately 30 nm in diameter (Johnson

and Heuck 2014)). Accordingly, the metabolic activity of hemocytes remained unchanged in the presence of FraC and MYTP1, suggesting no cytotoxic activity. This metabolic activity was abolished upon treatment with the cholesterol-dependent cytolysin perfringolysin O or 0.5% Triton X-100 (fig. 5*F*). We also checked the binding of MYTP1 to model membranes composed of a mixture of the main ceramide aminoethylphosphonates (CAEP) of *M. galoprovincialis*, a major sphingolipid of marine invertebrates (Balbi et al. 2022). CAEP is a lipid class composed of a ceramide-phosphoethanolamine analog bearing a phosphono group
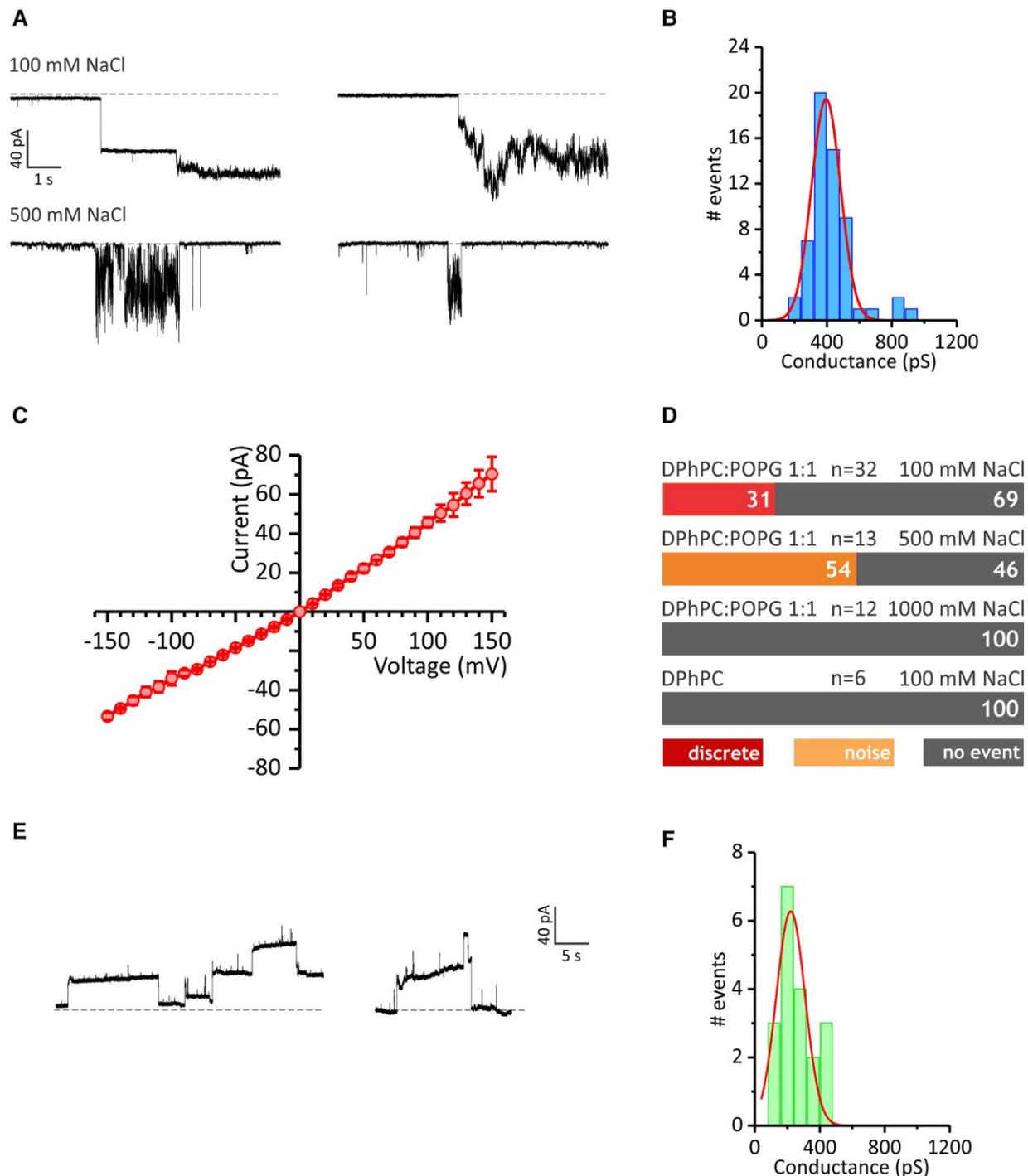
FIG. 4—Permeabilizing activity of mytiporin-1 (MYTP1) on planar lipid bilayers. (A) Typical current traces of planar lipid bilayers composed of 1,2-diphytanoyl-*sn*-glycero-3-phospho-choline (DPhPC):1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoglycerol (POPG) 1:1 exposed to 8.4 μM MYTP1 monomer. The salt concentration in the buffer (50 mM Tris/HCl, pH 7.4) is indicated. The applied potential was −50 mV and −150 mV at 100 and 500 mM NaCl, respectively. (B) A histogram of the conductance data obtained from traces at 100 mM NaCl. The curve shows a fit with a Gaussian distribution. (C) The voltage–current relationships; each point represents the mean ± standard deviation of four independent experiments. (D) Quantification of the experiments shown in A. The numbers denote the percentages of each observed event. *n* represents the number of independent membrane preparations for each lipid composition. (E) Insertion of soluble MYTP1 pores (at 1.3 μM MYTP1) at 100 mM NaCl and +150 mV applied potential. (F) Histogram of the conductance data obtained from experiments presented in E (at 100 mM NaCl). The curve shows a fit with a Gaussian distribution.
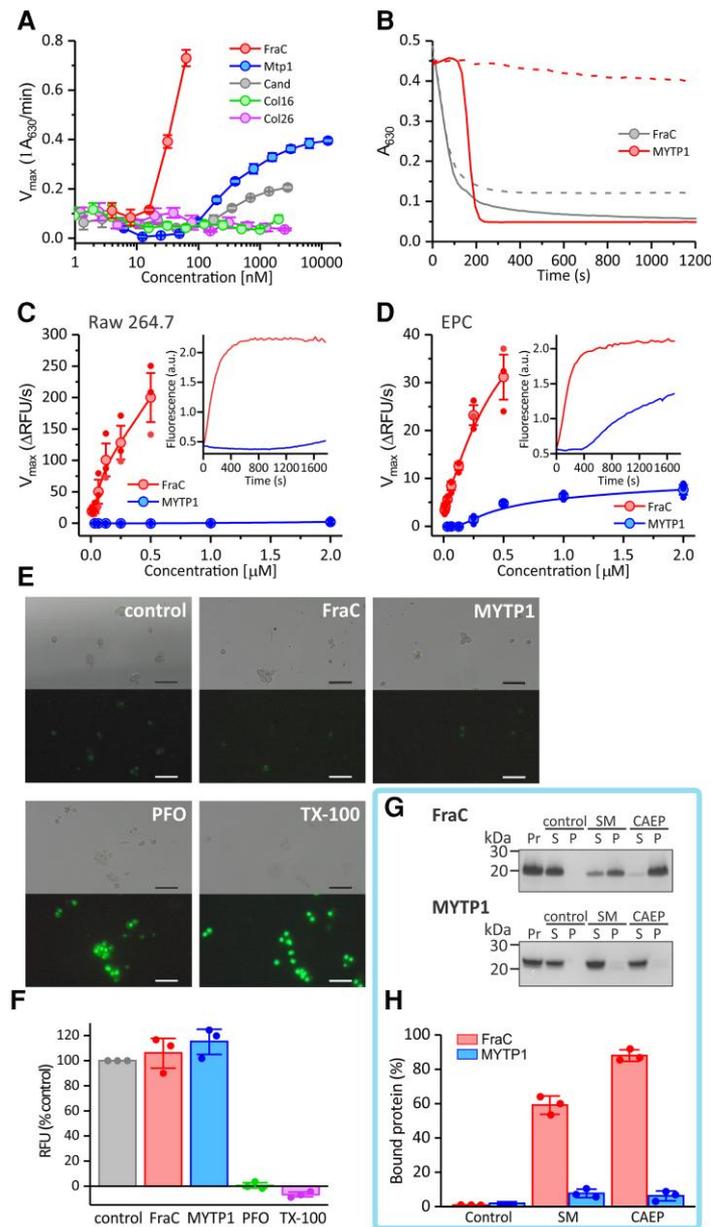
FIG. 5—Activity of mytiporin-1 (MYTP1) in cells. (A) The maximal rate of hemolysis of bovine red blood cells induced by FraC and ALPs. (B) Kinetics of hemolysis induced by 30 nM FraC (gray) and 6300 nM MYTP1 (red). Solid line, bovine red blood cells; dashed line, fish red blood cells. (C, D) Maximum rates of fluorescence increases (expressed in relative fluorescent units [RFU] per second) after permeabilization of Raw 264.7 (C) and EPC (D) cells with different concentrations of FraC (red) and MYTP1 (blue) in the presence of SYTOX™ Green Nuclear Stain. The inserts show the kinetics of fluorescence increases after treatment with 0.25 µM FraC (red) and 2 µM MYTP1 (blue). (E) Bright-field and fluorescence images of hemocytes adherent to µ-slide channels in the presence of pore-forming proteins (at 5 µM), 0.5% Triton X-100 (TX-100), or phosphate-buffered saline (control). PFO, perfringolysin from *Clostridium perfringens*. All treatment solutions contained SYTOX™ Green nuclear stain. Scale bars: 50 µm. (F) Viability of *Mytilus galloprovincialis* hemocytes in the presence of pore-forming proteins (at 4 µM), 0.5% TX-100, or phosphate-buffered saline (control). Viability was determined according to the conversion rate of PrestoBlue™ cell viability reagent and normalized to the control. (G) Binding of FraC and MYTP1 to 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine: sphingomyelin:cholesterol (chol) 5:3:2 (SM) or POPC:ceramide aminoethylphosphonate:chol 5:3:2 (CAEP). Pr, total amount of protein used in the assay; S, supernatant; P, pellet; control, the sample in the absence of MLVs. (H) Quantification of bands from G for each composition of lipid membranes. Mean ± standard deviation of three independent experiments are shown in all panels. Individual data points are also shown in panels C, D, F and H.

in its polar head instead of a phosphate group (Balbi et al. 2022). We produced MLVs composed of POPC:SM:chol 5:3:2 or POPC:CAEP:chol 5:3:2 that closely mimicked the outer leaflet of the hemocyte plasma membrane from which CAEP were isolated (Balbi et al. 2022). FraC showed strong binding, as indicated by its specificity for SM (fig. 5G and H). On the other hand, MYTP1 did not bind to any of the MLVs, in agreement with absence of permeabilizing activity in any of the model systems tested, including mussel hemocytes. In summary, MYTP1 shows significantly less permeabilizing and cytotoxic activity in various model cell systems in comparison to FraC.

Model of the Mytiporin-1 Monomer

We created a model of monomeric MYTP1 using AlphaFold (Jumper et al. 2021). The overall structure of the MYTP1 model resembles an actinoporin fold with a β-sandwich core consisting of two sheets, consisting of five and six β-strands. The β-sandwich core is flanked on both sides by α-helices (supplementary fig. 6a, Supplementary Material online). Compared with the typical actinoporins FraC and EqtII, the central β-sandwich remains almost unchanged except for two additional β-strands one on each sheet. A short β4 is located on the outside of the β-sandwich close to the start of the N-terminal α-helices. The longer β9 is following the α-helix (α 3) in the C-terminal part of the protein (supplementary figs. 2 and 6b, Supplementary Material online) (Athanasiadis et al. 2001; Mechaly et al. 2011). As expected from the sequence, both the N- and C-termini diverge the most from the actinoporin fold. The N-terminus, which in actinoporins forms an α-helix that undergoes conformational change and forms the transmembrane region of the pore, is longer in MYTP1 and consists of two helices. A short strand β1 is positioned between these two helices (supplementary fig. 6a, Supplementary Material online). According to the model, neither helix occupies the same position as the α1 helix of FraC (supplementary fig. 6a and b, Supplementary Material online). The C-terminus in MYTP1 is also longer, and the extension is structurally organized as the α-helix α4. This helix is located at the top of the β-sandwich with respect to the membrane-binding region of actinoporins and most likely does not play a role in membrane binding. Another difference between the two structures is the overall surface electrostatic potential: MYTP1 is much less positively charged (supplementary fig. 6c, Supplementary Material online).

The similarities in the core also extend to the membrane-binding loops, for which no significant differences were observed. The C-terminal helix α2 is involved in membrane binding of FraC (Tanaka et al. 2015) (supplementary fig. 6b, Supplementary Material online). In MYTP1, the corresponding structural element α3 is similarly positioned, the most notable difference being the longer loop between α3 and β9, and the β9 strand that is absent in FraC (supplementary fig. 6b, Supplementary Material online). Perhaps the largest differences between the two proteins can be found in the lipid-binding region (fig. 6A). X-ray structures of FraC contain at least three bound SM molecules (Tanaka et al. 2015). SM recognition in the initial membrane-binding step is mediated by side chains of residues in loops between strands β5-β6 and β7-β8 and helix α2, according to mutagenesis data and structural models (Anderluh et al. 1998; Hong et al. 2002; Bakrač et al. 2008; Tanaka et al. 2015) (fig. 6B). This region of the molecule allows binding of two SM molecules (L2 and L3, based on structural work on FraC ((Tanaka et al. 2015); fig. 6A and B). In FraC, SM-binding sites are formed by side chains of 13 and 7 residues for L2 and L3 SM molecules, respectively (fig. 6C) (Tanaka et al. 2015). These residues are almost completely conserved in actinoporins, whereas conservation in mytiporins is limited to only one glycine and proline residue (Gly88 and Pro111 in MYTP1, corresponding to Gly85 and Pro107 in FraC) (fig. 6B and D, supplementary fig. 3b, Supplementary Material online). In all other positions of the putative lipid-binding site in MYTP1, there is almost no conservation. In particular, in MYTP1, amino acids are present that have side chains with different physicochemical properties, that is charged and polar amino acid residues instead of aromatic ones (see, e.g., the binding site for L3 in fig. 6D). Furthermore, Glu120 in MYTP1 significantly affects the charge in the L3 binding site (fig. 6A). In addition, loop β7-β8 is shorter by three to four residues in mytiporins-11–20. Consequently, some of the most important residues for binding and recognizing SM (i.e., Trp112 and Tyr113 in EqtII (Bakrač et al. 2008)), are missing from this loop in mytiporins (fig. 6D). Finally, according to the structural model, the L3 SM-binding site in MYTP1 could be further occluded by the side chain of Arg153 from the β9 strand (fig. 6A), which would further prevent efficient recognition and binding of an L3 SM molecule. In summary, according to the predicted 3D structure of MYTP1, the shape and charge of the binding site for the two SM molecules are significantly different in MYTP1 and other mytiporins compared with actinoporins (fig. 6), thereby most likely preventing efficient SM binding. This is consistent with the functional data showing that MYTP1 is much less specific for SM compared with FraC (figs. 3 and 4).

Stoichiometry of Mytiporin-1 Pores

We next prepared soluble MYTP1 pores to provide further structural information. We incubated MYTP1 with LUVs comprising POPG (the most optimal membrane composition; fig. 3), extracted resultant pores with detergent Tween-80, and further purified the solubilized pores by cation exchange chromatography (supplementary fig.7,
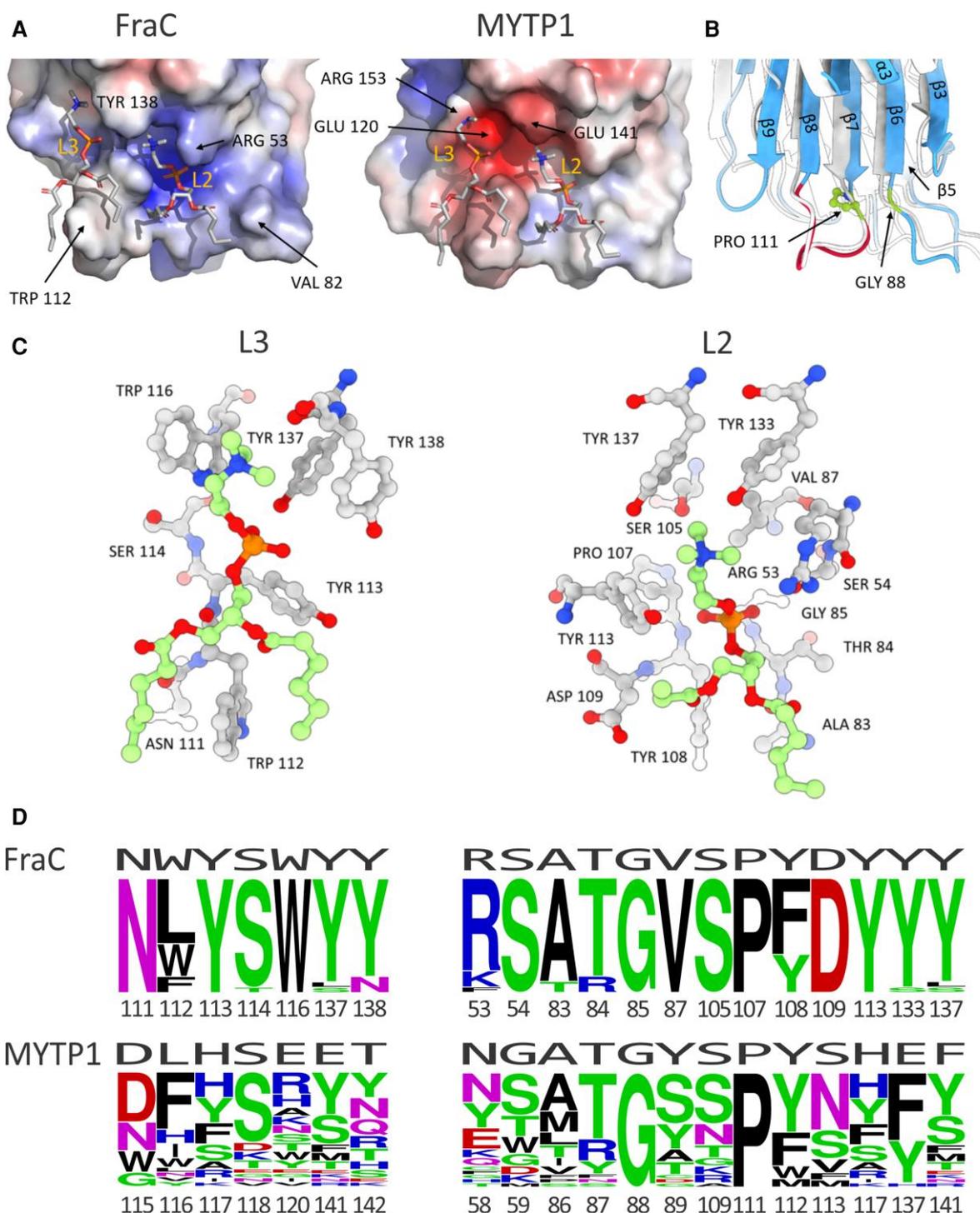
Fig. 6—Lipid-binding sites of actinoporins and mytiporins. (A) Surface of fragaceatoxin C (FraC) and the AlphaFold model for mytiporin-1 (MYTP1), colored according to electrostatic potential (red, negative potential; blue, positive potential). Sphingomyelin (SM) molecules L2 and L3 from FraC structures with bound lipid analogs are shown with sticks (white, carbon; blue, nitrogen; red, oxygen). Some of the residues are labeled. (B) Superimposed models of FraC (white) (PDB-ID:4TSO) and the AlphaFold model of MYTP1 (blue) in the same orientation as in A. Secondary structural elements of MYTP1 are labeled red (β7-β8 loop) and green (conserved glycine (Gly88) and proline (Pro111) residues; according to the MYTP1 numbering). (C) Binding sites of FraC for L2 and L3 SM. FraC numbering is used (PDB:4TSO). Green, SM. The coloring of side chains is as in A for SM. (D) Frequency analysis of amino acid residues that participate in L2 and L3 SM binding. The frequency distributions were prepared by Sequence logos (Crooks et al. 2004) for all full-length actinoporins in the UniProt database (21 sequences) and all mytiporins from this study. The residues in FraC and MYTP1 are indicated above the logos. The numbering of residues is shown below for each of the proteins. Green, YSTG; black, ALVMPWF; red, DE; blue, HRK; magenta, NQ.

Supplementary Material online). We added these purified pores to DPhPC:POPG 1:1 planar lipid bilayers at 100 mM NaCl and observed well-defined stepwise increases in the current (fig. 4E). Once inserted, these pores had much lower noise than pores formed on the membranes by oligomerization of MYTP1 monomers (i.e., compare the noise of open pores in figs. 4A and E). However, the pores were not stably inserted, as frequent decreases of current were also observed (fig. 5E). Furthermore, most of these inserted pores had a well-defined size, although smaller (with a conductance of 221 ± 14 pS; $n = 19$, mean ± standard error of fit with a Gaussian distribution) than pores formed from monomers (fig. 4B and F).

To provide an accurate mass and therefore stoichiometry of MYTP1 pores, native mass spectrometry was performed. The data analysis shows a charged series for a single species with its dominant peak at 5707 m/z at a charge of +25 (fig. 7A). This translates to a mass of 142,667 Da, corresponding to the hexameric form of the MYTP1 pore (fig. 7B). Adding additional energy to the complex by ramping activation potentials from 75 V (where only hexamer is present) to 120 V dissociates the hexamer into a monomer (23,775 Da) and pentamer (118,888 Da), validating that the native species is comprised of six non-covalently assembled mytiporin monomers (supplementary fig. 8, Supplementary Material online). Furthermore, upon close inspection, the peaks of the hexamer charge series include additional peaks that correspond to the addition of 1–4 lipids attached to the protein complex, each with a mass of approximately 750 Da (fig. 7B). Mass photometry was also performed to provide a solution-based mass confirmation in an alternate buffer system (phosphate-buffered saline), and the results align with the mass spectrometry-determined masses (supplementary fig. 9, Supplementary Material online).

We also examined the stoichiometry of isolated MYTP1 pores using cryo-electron microscopy (cryo-EM) (supplementary fig. 10, Supplementary Material online). After reference-free 2D classification, only hexameric pores were observed (supplementary fig. 10b, Supplementary Material online), suggesting that this is, if not the only, at least the predominant stoichiometry (fig. 7C and D). We also imaged LUVs with bound ALPs from molluscs (Col16, Col26, and Cand). For LUVs, we used the lipid composition that exhibited the strongest binding in the MLV binding experiments (fig. 3). On the vesicles, Col26 and Cand formed visible circular oligomers (fig. 7C) with a hexameric stoichiometry (fig. 7D). However, ordered oligomers were not observed for Col16 (fig. 7C).

The Structure of Mytiporin-1 Pores

To determine the structure of the isolated MYTP1 pore, we collected data on this sample and performed a single-particle analysis with cryo-EM (supplementary fig. 10,

Supplementary Material online, fig. 7E and F). C6 symmetry was used for 3D refinement because cryo-EM 2D classification of MYTP1 particles revealed hexamers (supplementary fig. 10b, Supplementary Material online). The overall resolution of the final map was approximately 4.5 Å (supplementary fig. 10e, Supplementary Material online), which only allowed a low-resolution structural model of the soluble MYTP1 pore to be built. The map is well defined only in the pore cap region built of a ring of six β-sandwiches, but not in the transmembrane channel. A truncated AlphaFold model of monomeric MYTP1, encompassing residues Asn26-Leu193, was used as a starting point, fitted to the map, and modified by hand. The model of MYTP1 fits well into the map within the β-sandwich region including the membrane-binding area. Due to the poor density, the C- and N-termini could not be modeled, and only the beginning of the potential transmembrane N-terminus could be traced (fig. 7E, supplementary fig. 10c, d, Supplementary Material online). The outer diameter of MYTP1 pore (9.3 nm) is smaller compared with that of the FraC pore (11.6 nm, PDB:4TSY), due to two fewer protomers in the oligomer (fig. 7F). The position of α3 of MYTP1 and the corresponding α2 of FraC is similar as both are located outside the pore. A side view of the pore shows that the protomers in MYTP1 are more perpendicular to the membrane plane than in FraC (fig. 7E, supplementary fig. 11a, Supplementary Material online). The different number of protomers and the angles at which each protomer interacts with the membrane result in a different interaction area between protomers in the upper part of the pore of MYTP1 in contrast to the FraC pore.

To obtain information about the dynamics of different parts of the MYTP1 pore, we performed molecular dynamics simulations. As shown by three equivalent 300 ns long atomistic molecular dynamics simulations of MYTP1 in the POPC:SM 1:1 membrane, the orientations of the N-terminal helices vary considerably, whereas the parts of the protein that forms the upper part of the pore remain almost unchanged (supplementary fig. 12a and b, Supplementary Material online). The superimposed structures of all protomers are shown in (supplementary fig. 12c, Supplementary Material online). The helical motion occurs mainly on the transmembrane part of the pore and can be characterized by twisted rotation or phase-shifted tangential displacement of the triple helices. Therefore, the calculated root-mean-square fluctuation per residue (supplementary fig. 12d, Supplementary Material online) clearly shows an increasing divergence of the helices, which increases with distance from the β-sandwich. The high mobility of helices agrees with the high noise observed in the planar lipid bilayer experiments (fig. 4) and the fact that the helical region is not resolved in the pore structure (fig. 7E and F).
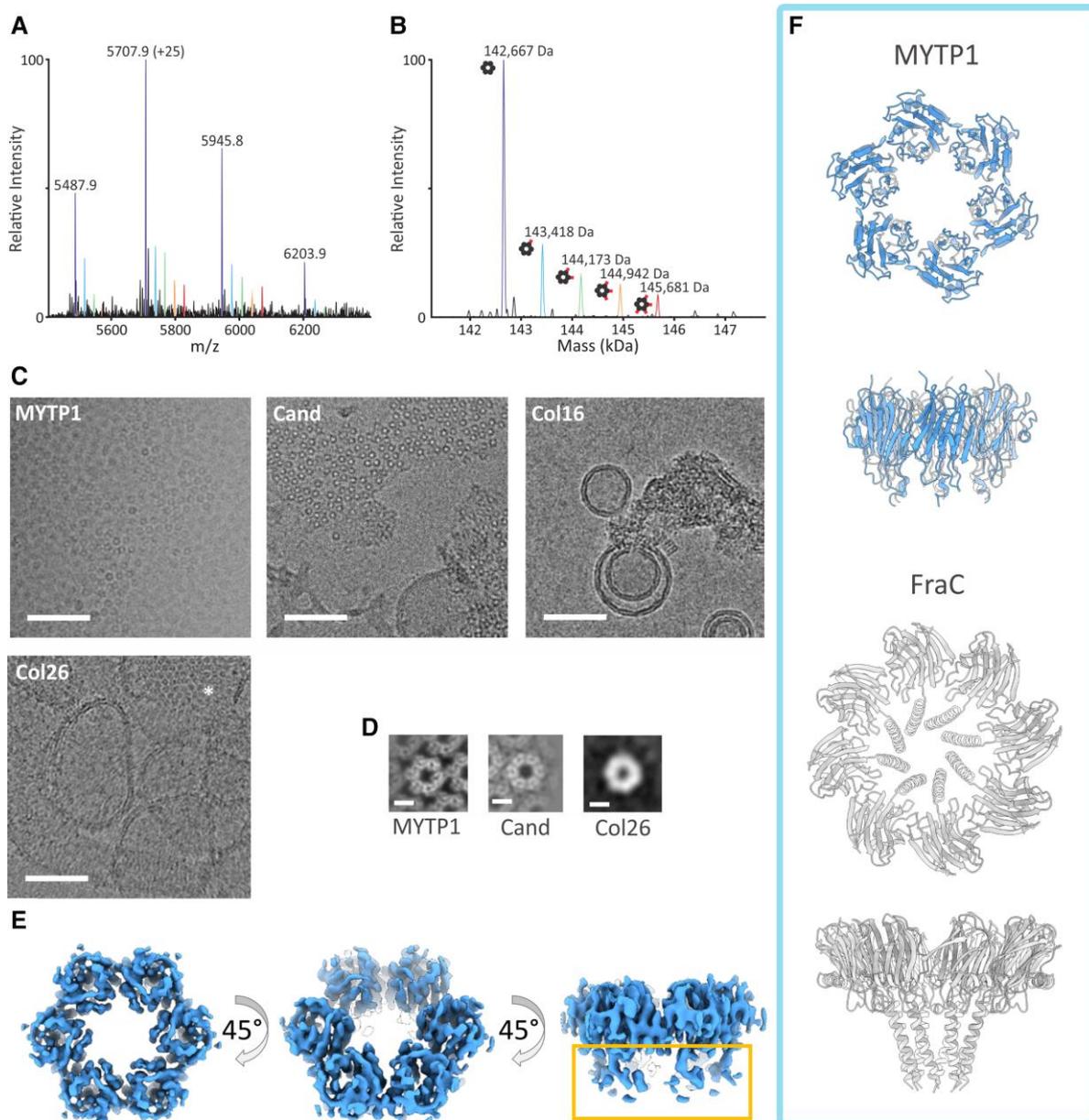
FIG. 7—Stoichiometry and structure of mytiporin-1 (MYTP1) pores. (A) Charge series of the MYTP1 hexamer with its dominant peak residing at 5707.9 m/z with a charge of +25 (purple). Additional charge series can be seen for the addition of 1–4 lipid adducts (approximately 750 Da) to the complex (blue, green, orange, red). (B) Resultant masses deconvolved from the charge series as presented in A; hexamer (purple) and the addition of 1–4 lipid adducts (blue, green, orange, red). (C) Transmission cryo-electron micrographs of porin pores. Isolated MYTP1 pores, POPC:SM 1:1 MLVs with bound Cand pores, POPG MLVs with bound Col16, and POPG MLVs with bound Col26 (* indicates visible pores). Scale bars: 50 nm. (D) 2D classes of ALPs pores. Scale bars: 5 nm. (E) Cryo-EM density map of the MYTP1 pore at different orientations. The poorly defined transmembrane region of the pore is framed. (F) Side-by-side comparison of the MYTP1 pore (blue) and fragaceatoxin C (FraC) pore (gray; PDB: 4TSY) at two orientations.

Polymorphisms Associated with the Mytiporin-1 Sequence

A total of 35 polymorphic amino acid sites resulting from non-synonymous single-nucleotide polymorphisms (SNPs) were annotated in MYTP1 mature proteins of 16 resequenced *M. galloprovincialis* individuals (supplementary table S2, Supplementary Material online). While some of these variants were observed with low frequency in the resequenced individuals, others were widespread and sometimes found in all individuals, possibly due to an association with one of the two paralogous genes predicted to be present. Of all SNPs resulting in amino acid changes, only two (Trp4* and Leu29*, both recorded in a single individual) were linked with a premature truncation of the polypeptide

chain, thereby indicating a null allele. Among the most common sequence variants identified (supplementary table S2 and fig. 13, Supplementary Material online), those observed in more than 50% of the resequenced individuals were Ser13Thr, Ala37Val, Ala68Val, Asn101Gly, Thr142Ala/Ser, Phe151Leu, and Lys202Val. Several polymorphisms were involved in conservative substitutions among hydrophobic residues, with minor implications on the structural properties of MYTP1. Ser13Thr is located on the N-terminal helix and could play a role in pore formation or affect channel properties. Ala68Val is positioned on the membrane-binding loop; however, substituting one hydrophobic residue for another probably does not affect binding. Thr142Ala/Ser lies in the area that interacts with neighboring subunits in the pore; however, substitution to a similarly sized side chain probably does not affect activity (supplementary fig. 13, Supplementary Material online).

Discussion

Similar to several other large gene families previously described in marine mussels, mytiporins are characterized by a seemingly complex evolutionary history, strongly affected by gene presence–absence variation. Indeed, while a total of 27 unique gene sequences were identified in the 16 resequenced *M. galloprovincialis* individuals, only 10 of these were *core* genes shared by all mussels (fig. 1A). The other mytiporin genes were *dispensable*, that is characterized by complex presence–absence variation patterns, which resulted in 15–19 unique genes per individual. The complexity of the mytiporin gene family was further complicated by the presence of nearly identical gene copies, such as in the case of MYTP1, whose coverage and associated sequence polymorphisms were consistent with the presence of two paralogs (supplementary table S2, Supplementary Material online). The presence of two nearly identical MYTP1 paralogous gene copies in *M. galloprovincialis* is further supported by the detection of two orthologous genes in both *M. edulis* and *M. coruscus* (fig. 1A). In addition, our genomic screening revealed the presence of several mytiporin-like pseudogenes (which were not analyzed in detail in this study), occasionally lacking one exon or containing nonsense or frameshift mutations. These pseudogenes usually shared high similarity with *dispensable* mytiporin genes, suggesting an origin from gene duplication, followed by acquisition of loss-of-function mutations.

Altogether, these observations suggest that the molecular diversity of mytiporins is the product of recent tandem gene duplication events, which might have occurred in the *Mytilus* spp. lineage. This idea is supported by multiple lines of evidence: (i) the presence of a very small number of porin-like genes in most other bivalve species, in stark contrast with > 10 mytiporins genes in the three *Mytilus* species (table 1); (ii) the strongly supported monophyly of nearly all

mytiporins (except mytiporins-25 and -26) (fig. 2); (iii) the evident presence of orthologs for most mytiporin genes between *M. galloprovincialis* and the closely related species *M. edulis* and *M. coruscus*, which strongly suggests a shared ancestry pre-dating the speciation events (fig. 1A).

Curiously, the mytiporins with a canonical signal peptide were not grouped in a monophyletic clade but were instead classified into three distinct groups (fig. 1A), indicating that this feature was likely acquired (or lost) multiple times along the evolution of *Mytilus*. Furthermore, dispensable mytiporins most often belonged to branches characterized by multiple sequences with close similarity, suggesting a link between recent gene duplication and presence–absence variation. It remains to be established whether mytiporins have undergone functional specialization in parallel with their primary sequence diversification and acquisition of different regulatory regions and gene expression profiles.

Overall, our Bayesian phylogenetic analysis suggests that the actinoporin-like gene family underwent multiple independent expansion and contraction events during the evolution of molluscs (fig. 2). These events have resulted in relatively large collections of diversified porins in some taxa, including *Mytilus* spp. and *C. reticulata*. Conversely, cone snails, despite possessing a few actinoporin-like genes (table 1), did not show any evidence of expansion in any species with a fully sequenced genome or assembled transcriptome available (Leonardi et al. 2012; Safavi-Hemami et al. 2014; Pardos-Blas et al. 2019, 2021, 2022; Sudewi et al. 2019; Fouda et al. 2021). Importantly, the phylogenetic tree of molluscan ALPs reveals that the intraspecific sequence diversity of both mytiporins in *M. galloprovincialis* and coluporins in *C. reticulata* is higher than the interspecific diversity of conoporins in cone snails, suggesting that strong selective pressures might have led, through accelerated evolution, to the acquisition of relatively large repertoires of diversified porin genes in mussels and vampire snails. The occurrence of similar processes in several gene families encoding toxins and venoms components in predatory animals has often been interpreted as the product of prey-driven natural selection (Chang and Duda 2012; Vonk et al. 2013; Sunagar and Moran 2015; Shibata et al. 2018). However, while this interpretation may apply to coluporins, it is inconsistent with the filter-feeding habits of *M. galloprovincialis*. Interestingly, selective processes similar to those described for venom proteins have been previously described in a number of bivalve gene families encoding both receptors (Zhang et al. 2015) and effectors (Rey-Campos et al. 2020, 2021) of the innate immune response. Such evolutionary dynamics are thought to be mostly driven by host–pathogen interactions.

Although all the sequences included in the phylogenetic tree of ALPs, including bryoporin (which is produced by an organism belonging to the kingdom of Plantae (Šolinc et al. 2022)), share a very similar 3D structure and the presence of

a recognizable protein domain, it is unclear whether they share a common evolutionary origin or are the product of convergent evolution. Consequently, while cnidarian and molluscan ALPs may derive from an ancestral metazoan actinoporin-like gene, the possibility that they have been acquired independently, either through *de novo* processes or horizontal gene transfer, cannot be excluded. Furthermore, the highly divergent sequences of mytiporins-25 and -26 were placed in a basal, intermediate position between cnidarian and molluscan porins in the phylogenetic tree (fig. 2). These sequences may derive from an ancient gene duplication event, whose descendants were lost in gastropods and only retained in mussels. Conversely, these sequences could be the product of an independent acquisition event, which may have included horizontal gene transfer from an unknown donor.

The structural and functional data of MYTP1, together with some other examples of molluscan ALPs, indicate that the molluscan branch of ALPs acquired different functional and structural features compared with anthozoan actinoporins during molluscan evolution and diversification. The three main and profound differences are described below.

(i) The SM-preference, a hallmark of actinoporins, is not observed in molluscan ALPs. This is most likely due to substantial changes in amino acids and the size of the crucial loop at SM-binding site(s) (fig. 6). As such, the SM-binding site exhibits profoundly changed electrostatic and morphological properties, which prevent effective SM recognition (figs. 3 and supplementary fig. 5, Supplementary Material online). MYTP1 and other molluscan ALPs may instead rely on the electrostatic attraction of the negatively charged membrane in the initial binding step. This initial electrostatic interaction is sufficient to retain MYTP1 on the membrane, triggering its oligomerization into hexamers. Surprisingly, we could purify functional hexamers formed on POPG membranes (fig. 4E and supplementary fig. 7, Supplementary Material online) and observed functional pores only on POPG-containing membranes at low salt concentration (fig. 4). Accordingly, the specificity of echotoxins from *Monoplex echo* for gangliosides (Shiomi et al. 2002) may thus be explained by the high amount of negatively charged functional groups in the headgroup of these lipids. Alternatively, the lipid-binding surface of MYTP1 could be shaped to specifically engage with a different, as yet unidentified molecular target. (ii) Molluscan ALPs (for which we could obtain data regarding pores) form hexameric pores that distinctively differ from the octameric arrangement of FraC (Tanaka et al. 2015). (iii) The permeabilizing activity of MYTP1 is significantly lower than that of FraC. In all the examined cell models, MYTP1 exhibited significantly lower activity than that of FraC. Actinoporins are extremely potent toxins (Anderluh and Maček 2002; Šuput 2009). For a variety of cells, they are highly toxic and lytic already at nanomolar concentrations (Maček et al. 1994; Meunier et al. 2000) due to their high affinity to SM-containing membranes. They also show cardiotoxic effects (Bunc et al. 1999) and are highly lethal in mammals, with an i.v. LD_{50} of 20–100 µg/kg in mice (Maček and Lebez 1988). By contrast, MYTP1 exhibited very poor activity in various cells (fig. 5) and only exhibited modest permeabilizing activity in negatively charged model membranes (i.e., POPG or POPG:phosphatidylcholine mixtures). This, at least in part, can be explained by the extension at its N-terminus. All the studied ALPs, and also echotoxin II (supplementary fig. 2, Supplementary Material online), possess a polar extension at the N-terminus containing 5–15 amino acids. It was shown for EqtII that any extension at its N-terminus inhibits its permeabilizing activity (Hong et al. 2002; Kristan et al. 2007), possibly due to difficulties in transferring the N-terminus across the lipid bilayer. This may be the underlying reason for the poor activity of MYTP1 and other molluscan ALPs. Among all the mussel tissues and cell types, we could reliably assess the activity of MYTP1 only against hemocytes, determining a lack of activity (fig. 5E and F), which is due to absence of binding to CAEP, a major sphingolipid class of molluscs (fig. 5G and H). Altogether, these data imply that MYTP1 may target the membranes of foreign cells, either in the context of the immune response against invading microorganisms or during the digestive process.

In the absence of clear functional data, the reliability of these two alternative hypotheses can only be evaluated based on the pattern of mytiporin expression in different tissues. Unfortunately, this analysis was complicated by the *dispensable* nature of most mytiporin genes, which might have led to significant noise in available RNA-seq datasets. The expression trends of the 17 *dispensable* mytiporins should be investigated in detail in individuals with matched genomic and transcriptomic data. Nevertheless, we collected some preliminary information concerning the tissue specificity of the ten mytiporin *core* genes (fig. 1C). Both mytiporins-3 and -19 exhibited very low expression levels in all tissues. This may either suggest limited biological function (which would be consistent with the detection of a mytiporin-3 pseudogene with in-frame stop codons in *M. edulis*) or that their expression is tightly regulated by unidentified stimuli or linked with tissue-specific microbiomes (Musella et al. 2020). Most mytiporins did not show strong tissue specificity, whereas mytiporins-1, -4 and -25 were preferentially expressed in the digestive gland. Conversely, no mytiporin (except mytiporin-21), was expressed at significant levels in hemocytes, circulating cells with phagocytic activity that are key players in pathogen recognition and elimination in molluscs.

Although the lack of expression of mytiporins in hemocytes should rule out the involvement of mytiporins in hemocyte-mediated immune processes, it does not

necessarily preclude a possible role in the context of mucosal immunity. Mucosal immunity acts at the main sites of contact between bivalve tissues and the external environment and is particularly important in filter-feeding organisms, such as bivalves (Allam and Pales Espinosa 2016). Pore-forming proteins play a key role as immune effectors in all metazoans, contributing to the killing of invading microbes. For example, in all vertebrates, C9 acts as the final effector molecule of the lytic pathway of the complement system, forming pores in bacterial cell envelopes and leading to their death (Doorduijn et al. 2021). Although no sequence clearly homologous to C9 exists in bivalves, some authors have previously hypothesized that other proteins belonging to the same structural family (MACPF) or showing homology with other PFT families might be activated by the proto-complement system in *M. galloprovincialis* in response to lectin-mediated microbe recognition (Mitta et al. 1999; Estevez-Calvar et al. 2011). Additionally, multiple multi-gene families of antimicrobial peptides, whose mode of action usually involves the permeabilization or disruption of microbial membranes, have been extensively studied in mussels (Mitta et al. 1999, 2000), not all of which exhibit a strictly hemocyte-specific expression pattern (Leoni et al. 2017).

Altogether, the specific expression of some mytiporins (including MYTP1) in the digestive gland paired with the lack of mytiporin expression in hemocytes suggests that this gene family is unlikely to play a role in mussel humoral immunity and more likely to play a role in food digestion. PFTs and several other enzymes are also involved in disrupting cells of unicellular planktonic organisms, such as microalgae, which are ingested by filter-feeders as food (Sher et al. 2008). It has been observed that mussels and other filter-feeding bivalves can rely on bacteria, in addition to phytoplankton, as a relevant food source under certain conditions (Langdon and Newell 1990; Ezgeta-Balić et al. 2012; Lüskow and Riisgård 2018). Most marine bacteria are Gram-negative and have a significant negative charge both in the inner and outer membrane, which may facilitate survival under conditions of high pressure and salinity (Murzyn et al. 2005; Leone et al. 2007). The finding that MYTP1 acts on negatively charged model membranes, albeit weakly, could support the involvement of PFTs in digestive processes adapted to bacteria. Their putative involvement in digestion could also explain the high intraspecific diversification of these proteins that might confer a greater trophic flexibility to an opportunistic filter-feeder such as *Mytilus*. Indeed, similar adaptations, linked to the advantages of broadening the spectrum of potential hosts/preys, have been associated with the diversification of other feeding-related protein families in molluscs, including coluporins from the slow-moving opportunistic parasite *C. reticulata* (Gerdol et al. 2018). Nevertheless, until the specificity of action of any mytiporin towards cell

membranes of potential food sources will be unequivocally established, the possibility that these proteins may be involved in immune defense at mucosal interfaces (including those of the digestive tract) should not be discarded.

Regardless of the involvement of mytiporins in digestion, the innate immune response, or both, their remarkable primary sequence divergence and lineage-specific expansion in mussels indicate that the newly acquired paralogous gene copies might have undergone neofunctionalization during evolution. This may have broadened the spectrum of their molecular partners and, consequently, the taxonomic range of membranes that they can recognize. In summary, we have shown that ALPs underwent a significant expansion and diversification in *M. galloprovincialis*, mirroring the situation observed in other molluscan species with a predatory feeding behavior. The characterization of MYTP1 and other molluscan ALPs revealed the acquisition of previously unreported structural and functional features that significantly differ from those of archetypal anthozoan actinoporins. These are most likely linked with neofunctionalization; however, the novel biological functions remain to be clarified.

Materials and Methods

Materials

The following was used: KOD Hot Start DNA-Polymerase (Novagen), restriction enzymes *Nde*I and *Hind*III (New England BioLabs), T4 DNA ligase (Thermo Fisher Scientific), various lipids (Avanti Polar Lipids), and tobacco etch virus (TEV) protease and fragaceatoxin C (FraC) (Bellomio et al. 2009; Tanaka et al. 2015) were recombinant proteins prepared by the standard protocols of the National Institute of Chemistry (Slovenia). Cloning, expression, and purification of MYTP1, Cand, Col16, and Col26 is described in the Supplementary Materials Online. CAEP were isolated from mussel hemocytes as described elsewhere (Balbi et al. 2022).

Identification and Characterization of Mytiporin Sequences in *Mytilus galloprovincialis*

The proteins encoded by the gene models annotated in the *M. galloprovincialis* genome v.mg10 (Gerdol et al. 2020) were initially screened with HMMER (Finn et al. 2011) to recover the sequences characterized by the presence of the conserved sea anemone cytotoxic protein Pfam domain PF06369, detected with a threshold e-value of 0.05. However, several previously reported molluscan porins are characterized by marked primary sequence differences and significant evolutionary divergence from those (mainly of cnidarian origin) used for the generation of the aforementioned domain in the Pfam database (Finn et al. 2010). Hence, an updated molluscan-specific Hidden

Markov Model was generated based on the multiple sequence alignment of all available molluscan porin protein sequences (Gerdol et al. 2018). Both the multiple sequence alignment and Hidden Markov Model were recursively updated with newly retrieved sequences, and searches were repeated until no additional match could be detected.

The genome of *M. galloprovincialis*, and several other molluscs, is characterized by massive gene presence–absence variation, which results in the presence of *core* genes, found in all individuals, and *dispensable* genes, associated with hemizygous genomic regions and only found in some individuals (Calcino et al. 2021). To identify possible additional mytiporin *dispensable* genes not present in the reference genome, several previously published transcriptomes obtained from multiple *M. galloprovincialis* tissues (see the supplementary material from Gerdol et al. (2020)) were screened using the same strategy described above, following their translation to proteins with Transdecoder v.5.5.0 (https://github.com/TransDecoder). Only transcripts including full-length coding sequences (i.e., including an initial ATG codon and a translation termination signal) were considered for further analysis. The encoded protein sequences were compared with those obtained from annotated gene models and only those that did not display nearly identical matches (i.e., pairwise identity at the amino acid level < 97%) were added to the mytiporin dataset. Those showing a level of similarity higher than the threshold were considered likely to derive from allelic variants of the same locus and were therefore discarded.

The position of the porin domain was recorded for all complete mytiporin sequences, which were further subjected to analysis with Phobius (Kall et al. 2007) to identify signal peptide regions for secretion and the transmembrane domain and with Interproscan to detect the presence of additional conserved domains (Finn et al. 2017). The mytiporin sequences lacking a signal peptide were further screened with SecretomeP v.2.0 to investigate the possibility of leaderless secretion (Bendtsen et al. 2004).

Determination of Presence–Absence Variation

Raw paired-end Illumina whole-genome resequencing data were available for nine *M. galloprovincialis* individuals collected in Galicia, Spain (PURA, GALF1, GALF2, GALF3, GALM1, GALM2, GALM3, GALM6, and GALM11) and six individuals collected in Italy (ITAF1, ITAF2, ITAF3, ITAM1, ITAM2, and ITAM3) (Murgarella et al. 2016; Gerdol et al. 2020). Following the trimming procedure described elsewhere (Gerdol et al. 2020), these were mapped against the coding nucleotide sequence of mussel mytiporins with the *map reads to contigs* tool included in the CLC genomics Workbench v.21 (Qiagen, Hilden, Germany), using the following parameters: length fraction = 0.5, similarity fraction = 0.9. The presence or the absence of any given

gene was assessed based on the observed read coverage. A mytiporin gene was identified as "present" if the observed coverage was in line or higher than the expectations for a *dispensable* gene present in a hemizygous state (further details about this threshold are provided in (Gerdol et al. 2020)) and "absent" if these criteria were not met. The visual inspection of a uniform read mapping along the full-length of the coding sequence was further employed to confirm the results obtained *in silico*.

Evaluation of Gene Expression Levels

Considering that most mytiporin genes are dispensable, the determination of their expression levels may be significantly affected by the genomic background of the individuals selected for RNA extraction. Hence, only core mytiporin genes (i.e., mytiporin-1, -3, -4, -5, -9, -17, -19, -21, -25, and -27) were selected for the *in silico* computation of gene expression values. This analysis was based on several previously published RNA-seq datasets (as previously reported in Gerdol et al. (2020)), derived from several independent experiments and from mussels collected from different geographical locations. Hence, this analysis was intended to provide a broad overview of tissue specificity, and the gene expression levels were computed as the average level of all samples belonging to the following tissues: hemocytes, digestive gland, gills, mantle, and posterior adductor muscle. Gene expression levels were calculated based on the number of uniquely mapped reads obtained with stringent mapping performed with the CLC Genomics Workbench v.21 (length fraction = 0.75, similarity fraction = 0.98) and reported as TPM (Wagner et al. 2012).

Phylogenetic Analyses

The porin domain sequence was extracted from all available porins, discarding the unalignable and highly divergent N-termini. Two different sequence datasets were then aligned with MUSCLE (Edgar 2004) to generate two multiple sequence alignments: dataset A included only mussel (*M. galloprovincialis*, *M. edulis*, and *M. coruscus*) sequences, whereas dataset B included a selected group of ALPs. Dataset B included all *M. galloprovincialis* mytiporins, all *C. reticulata* coluporins (GenBank accession ID: MH194204-MH194233), all available Conidae ALPs (GenBank accession ID: MF576552.1, MN517481.1, MN517479.1, KY963310.1, MF443899.1, KU881667.1, MN517478.1, MH360667.1, and MH360453.1; UniProKB accession IDs: P0DKQ8.1; the sequence from *Conus miliaris* was retrieved from the supplementary materials of Weese and Duda (2019)); the sequences from *Africonus* spp. and *Kalloconus* spp. were retrieved from the supplementary materials of Abalde et al. (2020), four previously functionally characterized cnidarian actinoporins (GenBank accession ID: XM_012707306.2 and XM_031694881.1;

UniProtKB accession ID: P81662.1 and D2YZQ3.1), and a bryoporin (GenBank accession ID: XM_024506210.1). This dataset was processed with CD-HIT (Li and Godzik 2006) to reduce redundancy (based on a 75% pairwise sequence identity threshold).

The multiple sequence alignments were subjected to analysis with Modeltest-NG (Darriba et al. 2019) to detect the best-fitting model of molecular evolution, which was found to be LG + G + I in both cases. Bayesian phylogenetic inference analysis was run with MrBayes v.3.2.7a (Huelsenbeck and Ronquist 2001) using two MCMC analyses run in parallel, with 200,000 (dataset A) and 500,000 (dataset B) generations each. The first 25% of the generated trees was discarded during the *burnin* process, and run convergence was checked with Tracer v.1.7.1 (Rambaut et al. 2018). The resulting trees were graphically represented with FigTree v.1.1.4 (http://tree.bio.ed.ac.uk/software/figtree) and rooted with a midpoint strategy (dataset A) or by arbitrarily using bryoporin as an outgroup (dataset B). Nodes supported by posterior probability values < 0.5 were collapsed.

Sedimentation Assay

MLVs were prepared in PBS (pH 7.4) and consisted of the following lipids: (1) 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1′-rac-glycerol) (POPG), (2) 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC), (3) POPC and cholesterol (Chol) at a 1:1 molar ratio, and (4) POPC and sphingomyelin (SM) at a 1:1 molar ratio. In order to assess binding to hemocyte-derived sphingolipid CAEP, we used MLVs composed of POPC:CAEP:chol at 5:3:2 molar ratio and used POPC:SM:chol 5:3:2 MLVs as a control. The lipid mixtures were dissolved in chloroform, which was then evaporated using Rotavapor R-300 (Büchi 15 Labortechnik AG) to produce a lipid film. The lipid film was dissolved in PBS, flash-frozen in liquid nitrogen, and thawed five times. Next, 2 μg of FraC, MYTP1, Cand, Col16, and Col26 were separately incubated with MLVs in PBS at a molar ratio of 1:2000 for 1 h at 25 °C and 250 rpm. The reaction mixtures were centrifuged at 16,100 × *g* for 10 min to generate supernatant and pellet, both of which were then analyzed by SDS-PAGE. Protein bands were analyzed by ImageJ to quantify the amount of proteins in the supernatant and pellet. In addition, POPG-containing MLVs were also prepared in PBS containing 50 or 500 mM NaCl (instead of 140 mM NaCl) and assessed for binding of MYTP1. All experiments were performed three times.

Hemolysis Assay

Hemolytic activity of proteins was measured using bovine and fish erythrocytes. Erythrocytes were washed in erythrocyte buffer (50 mM Tris/HCl, 140 mM NaCl, pH 7). Two-fold serial dilutions of protein in erythrocyte buffer (100 μl) were prepared in 96-well microplates (Costar). Bovine or fish erythrocytes in erythrocyte buffer (100 μl) were added to the protein dilutions to a final absorbance at 630 nm of 0.5. The final concentration of protein in the first well was 15.3 μM for FraC, 12.6 μM for MYTP1, 14.9 μM for Cand, and 18.8 μM for Col16 and Col26. Absorbance at 630 nm was measured for 20 min (at 20 s intervals) at 25 °C using the SynergyTM MX microplate reader (BioTek). The experimental curves were processed by linear regression with three data points using Gen5 software (BioTek) to determine the maximum hemolysis rate (V_{max}). All measurements were performed three times.

Electrical Measurements in Planar Lipid Bilayers

Monomeric MYTP1 and soluble MYTP1 pores were analyzed by electrical measurements in planar lipid bilayers. An integrated chip-based recording setup, Orbit mini, and EDR2 software (Nanion Technologies) were used for the recordings. Measurements were performed in parallel with multi-electrode cavity array chips (Ionera Technologies). Prior to the experiments, an electrolyte solution (50 mM Tris/HCl, and 100, 500, or 1000 mM NaCl, pH 7.4) was added to the multi-electrode cavity array chips, and lipids dissolved in octane (10 mg/ml) were used to deposit planar lipid bilayers over the electrodes. POPG and 1,2-diphytanoyl-*sn*-glycero-3-phospho-choline (DPhPC) were used to prepare the planar lipid bilayer. Monomeric MYTP1 or MYTP1 pores were added to the *cis* side of the bilayer at final concentrations of 8.4 μM and approximately 1.3 μM, respectively. A starting voltage (−100 to 150 mV) was applied to promote pore insertion. The sampling rate of the measurements was 20 kHz. Current measurements were analyzed using Clampfit software (11.2).

Mass Spectrometry and Photometry

The buffer of the purified MYTP1 pore sample was exchanged (Bio-Spin 6 Columns, Bio Rad) into 200 mM ammonium acetate to a final monomeric concentration of 2 μM. Native MS data were acquired by nanoelectrospraying the sample onto a QExactive UHMR mass spectrometer (ThermoFisher) using gold-plated 1.2 μm capillaries prepared in-house, as previously described (Kondrat et al. 2015). Resultant spectra were deconvolved and analyzed using UniDec (Marty et al. 2015). To acquire the mass photometry data (supplementary fig. 8, Supplementary Material online), the sample was further diluted to 50 nM and measured on a Refeyn OneMP mass photometer and analyzed using DiscoverMP (Refyn Ltd) as described elsewhere (Young et al. 2018).

Cryo-Electron Microscopy and Data Processing

For MYTP1, isolated pore samples were used. The other three porins, Col16, Col26, and Cand, were imaged on

MLVs. Col16 and Col26 (8 µM) were incubated with MLVs (2 mM) composed of POPG, and Cand (10 µM) was incubated with MLVs (2 mM) composed of POPC:SM at a lipid molar ratio of 1:1. In all cases, 3 µl of suspension was applied to a glow-discharged (GloQube® Plus, Quorum, UK) Quantifoil R1.2/1.3 300-mesh copper holey carbon grid (Quantifoil, Germany), blotted under 100% humidity at 4 °C for 6–7 s, and plunged into liquid ethane using a Mark IV Vitrobot (Thermo Fisher Scientific, USA). Micrographs were collected on a cryo-transmission electron microscope (Glacios, Thermo Fisher Scientific, USA) with a Falcon 3EC direct electron detector (Thermo Fisher Scientific, USA) and operated at 200 kV using the EPU software (Thermo Fisher Scientific, USA). Images were recorded in counting mode with a pixel size of 0.98 Å for MYTP1 and Cand and 1.6 Å for Col16 and Col26. All micrographs were dose-fractioned into 40 frames with a total dose of 30 e−/Å^2.

All steps of processing were performed in cryoSPARC 2.4 with built-in algorithms (Punjani et al. 2017). Cryo-EM data were analyzed following the typical steps of a single-particle analysis protocol, stopping at 2D classification for all porins except MYTP1, for which the 3D volume was reconstructed. Micrographs were dose-weighted and motion-corrected. After contrast transfer function estimation, particles were handpicked and underwent 2D classifications to create 2D class averages. MYTP1 data was further processed by 2D class-based template picking. Additional 2D and 3D classification steps and homogeneous refinement with 278610 particles were performed to create the final 3D volume map. The MYTP1 monomer was modelled using the AlphaFold plugin (Jumper et al. 2021) in ChimeraX version 1.3 (Pettersen et al. 2004). The pore model was built-in Coot (Emsley and Cowtan 2004) based on the predicted monomer model. Due to insufficient density quality, the N-terminal helices and C-termini were not modelled in the case of the pore.

Cell Cultures

The fish skin cell line EPC (American Type Culture Collection No. CRL-2872) was cultured in Eagle's Minimum Essential Medium supplemented with 10% heat-inactivated fetal bovine serum (both from Sigma) in a humidified incubator at 25 °C and 5% CO_2. The murine macrophage cell line Raw 264.7 (American Type Culture Collection No. TIB-71) was grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum (both from Sigma) in a humidified incubator at 37 °C and 5% CO_2. For cell membrane permeability assays, EPC and Raw 264.7 cell suspensions were rinsed twice in basic medium without fetal bovine serum and adjusted to a concentration of 10^6 cells/ml in DMEM containing SYTOX™ Green dye (5 µM; Invitrogen). Next, 25 µl of the cell suspension was added to each well of a white 384-well clear-bottomed microtiter plate. Toxin dilutions (25 µl) were added, and fluorescence was measured repeatedly for 30 min. The response was plotted as a function of fluorescence intensity versus time. The highest temporal increase in fluorescence intensity was determined at the inflection point of the logistic fit, as determined by Origin laboratory software, and was used as a relative measure of cell membrane permeabilization. A linear fit was used instead for negative controls and treatments with no response.

Isolation of Hemolymph from *Mytilus galloprovincialis*

The surface of *M. galloprovincialis* was first briefly cleaned. Mussels were held ventral side up, and scissors were slowly inserted between the shells at the exit site of the byssus fibers and rotated to separate the shells. The interior was rinsed with sterile PBS. The salinity of PBS for all procedures was increased by adding 11.1 g/l of NaCl. Then, hemolymph was collected from the posterior adductor muscle using a 1 ml syringe with a 26-G needle filled with 100 µl sterile PBS containing 10 µM EDTA to prevent clogging of the hemolymph. The hemolymph was examined under a bright field microscope and immediately used for the experiments.

Hemocyte Permeabilization and Viability Assays

Isolated hemolymph (30 µl) was seeded into the channels of a µ-slide (µ-Slide VI 0.4 Ibidi) and left in a humidified incubator at 15 °C for 3 h to allow cells to adhere. The attached cells were rinsed and treated in a single step with toxins (60 µl, 4 µM in PBS). Triton X-100 in PBS was used at a concentration of 0.5%. The toxin and Triton X-100 solutions contained SYTOX™ Green stain (at a final concentration of 12.5 µM; Invitrogen) for fluorescent labelling of the nuclei of the permeabilized cells. Bright field and fluorescence images were taken 30 min after treatment near the inlet of the µ-channels.

The isolated hemolymph was centrifuged, and the hemocytes were rinsed in PBS containing 10 µM EDTA. Finally, hemocytes were suspended in L-15 medium with increased salinity by adding 11.1 g/L of NaCl and 10% PrestoBlue™ (Invitrogen) cell viability reagent. The suspension (25 µl) was added to each well of a white 384-well clear-bottom microtiter plate. Toxins or Triton X-100 were then added, and repeated fluorescence measurements were performed for 2.5 h. The slope of the linear increase in fluorescence intensity was used as a relative measure of cell viability. Three independent experiments, that is independent preparations of hemocytes, were performed.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgements

Data Availability

All sequences have been deposited in GenBank; the accession codes are listed in supplementary table S1, Supplementary Material online.

Literature Cited

Abalde S, Tenorio MJ, Afonso CML, Zardoya R. 2020. Comparative transcriptomics of the venoms of continental and insular radiations of West African cones. Proc Biol Sci 287:20200794.

Allam B, Pales Espinosa E. 2016. Bivalve immunity and response to infections: are we looking at the right place? Fish Shellfish Immunol 53:4–12.

Altschul SF, et al. 1997. Gapped blast and psi-blast: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402.

Anderluh G, et al. 1998. Avidin-fitc topological studies with three cysteine mutants of equinatoxin II, a sea anemone pore-forming protein. Biochem Biophys Res Commun 242:187–190.

Anderluh G, Maček P. 2002. Cytolytic peptide and protein toxins from sea anemones (Anthozoa: Actiniaria). Toxicon 40:111–124.

Athanasiadis A, Anderluh G, Maček P, Turk D. 2001. Crystal structure of the soluble form of equinatoxin II, a pore-forming toxin from the sea anemone *Actinia equina*. Structure 9:341–346.

Baker MA, Rojko N, Cronin B, Anderluh G, Wallace MI. 2014. Photobleaching reveals heterogeneous stoichiometry for equinatoxin ii oligomers. ChemBioChem 15:2139–2145.

Bakrač B, et al. 2008. Molecular determinants of sphingomyelin specificity of a eukaryotic pore-forming toxin. J Biol Chem 283: 18665–18677.

Bakrač B, Anderluh G. 2010. Molecular mechanism of sphingomyelin-specific membrane binding and pore formation by actinoporins. Adv Exp Med Biol 677:106–115.

Balbi T, et al. 2022. Ceramide aminoethylphosphonate as a new molecular target for pore-forming aegerolysin-based protein complexes. Front Mol Biosci 9:902706.

Bellomio A, et al. 2009. Purification, cloning and characterization of fragaceatoxin C, a novel actinoporin from the sea anemone *Actinia fragacea*. Toxicon 54:869–880.

Belmonte G, Pederzolli C, Maček P, Menestrina G. 1993. Pore formation by the sea anemone cytolysin equinatoxin II in red blood cells and model lipid membranes. J Membr Biol 131:11–22.

Ben-Ari H, Paz M, Sher D. 2018. The chemical armament of reef-building corals: inter- and intra-specific variation and the identification of an unusual actinoporin in *Stylophora pistilata*. Sci Rep 8: 251.

Bendtsen JD, Jensen LJ, Blom N, Von Heijne G, Brunak S. 2004. Feature-based prediction of non-classical and leaderless protein secretion. Protein Eng Des Sel 17:349–356.

Boguski MS, Lowe TM, Tolstoshev CM. 1993. Dbest-database for "expressed sequence tags". Nat Genet 4:332–333.

Bose U, et al. 2017. Multiomics analysis of the giant triton snail salivary gland, a crown-of-thorns starfish predator. Sci Rep 7:6000.

Bunc M, Drevenšek G, Budihna M, Šuput D. 1999. Effects of equinatoxin II from *Actinia equina* (L.) on isolated rat heart: the role of direct cardiotoxic effects in equinatoxin II lethality. Toxicon 37: 109–123.

Caaveiro JMM, Tsumoto K. 2021. Molecular basis for the activation of actinoporins by lipids. Methods Enzymol 649:277–306.

Calcino AD, Kenny NJ, Gerdol M. 2021. Single individual structural variant detection uncovers widespread hemizygosity in molluscs. Philos Trans R Soc Lond B Biol Sci 376:20200153.

Chang D, Duda TF Jr. 2012. Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. Mol Biol Evol 29:2019–2029.

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. Weblogo: a sequence logo generator. Genome Res 14:1188–1190.

Darriba D, et al. 2019. Modeltest-ng: a new and scalable tool for the selection of DNA and protein evolutionary models. Mol Biol Evol 37:291–294.

Doorduijn DJ, et al. 2021. Polymerization of C9 enhances bacterial cell envelope damage and killing by membrane attack complex pores. PLoS Pathog 17:e1010051.

Drechsler A, et al. 2006. Structure and activity of the N-terminal region of the eukaryotic cytolysin equinatoxin II. Biochemistry 45:1818–1828.

Edgar RC. 2004. Muscle: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797.

Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D 60:2126–2132.

Estevez-Calvar N, Romero A, Figueras A, Novoa B. 2011. Involvement of pore-forming molecules in immune defense and development of the Mediterranean mussel (*Mytilus galloprovincialis*). Dev Comp Immunol 35:1017–1031.

Ezgeta-Balić D, Najdek M, Peharda M, Blažina M. 2012. Seasonal fatty acid profile analysis to trace origin of food sources of four commercially important bivalves. Aquaculture 334–337:89–100.

Finn RD, et al. 2010. The Pfam protein families database. Nucleic Acids Res 38:D211–D222.

Finn RD, et al. 2017. Interpro in 2017-beyond protein family and domain annotations. Nucleic Acids Res 45:D190–D199.

Finn RD, Clements J, Eddy SR. 2011. HMMER Web server: interactive sequence similarity searching. Nucleic Acids Res 39:W29–W37.

Fouda MMA, Abdel-Wahab M, Mohammadien A, Germoush MO, Sarhan M. 2021. Proteomic analysis of red sea *Conus taeniatus* venom reveals potential biological applications. J Venom Anim Toxins Incl Trop Dis 27:e20210023.

Gerdol M, et al. 2020. Massive gene presence-absence variation shapes an open pan-genome in the Mediterranean mussel. Genome Biol 21:275.

Gerdol M, Cervelli M, Oliverio M, Modica MV. 2018. Piercing fishes: porin expansion and adaptation to hematophagy in the vampire snail *Cumia reticulata*. Mol Biol Evol 35:2654–2668.

Gorson J, et al. 2015. Molecular diversity and gene evolution of the venom arsenal of terebridae predatory marine snails. Genome Biol Evol 7:1761–1778.

Gouaux E. 1997. Channel-forming toxins: tales of transformation. Curr Opin Struct Biol 7:566–573.

Gutierrez-Aguirre I, Trontelj P, Maček P, Lakey JH, Anderluh G. 2006. Membrane binding of zebrafish actinoporin-like protein: aF

domains, a novel superfamily of cell membrane binding domains. Biochem J 398:381–392.

Hoang QT, et al. 2009. An actinoporin plays a key role in water stress in the moss *Physcomitrella patens*. New Phytol 184:502–510.

Hong Q, et al. 2002. Two-step membrane binding by equinatoxin II, a pore-forming toxin from the sea anemone, involves an exposed aromatic cluster and a flexible helix. J Biol Chem 277: 41916–41924.

Huelsenbeck JP, Ronquist F. 2001. Mrbayes: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Johnson BB, Heuck AP. 2014. Perfringolysin o structure and mechanism of pore formation as a paradigm for cholesterol-dependent cytolysins. In: Anderluh G and Gilbert R, editors. MACPF/CDC proteins - agents of defence, attack and invasion. In: Netherlands: Springer. p. 63–81.

Jouiaei M, et al. 2015. Evolution of an ancient venom: recognition of a novel family of cnidarian toxins and the common evolutionary origin of sodium and potassium neurotoxins in sea anemone. Mol Biol Evol 32:1598–1610.

Jumper J, et al. 2021. Highly accurate protein structure prediction with Alphafold. Nature 596:583–589.

Kall L, Krogh A, Sonnhammer EL. 2007. Advantages of combined transmembrane topology and signal peptide prediction-the Phobius web server. Nucleic Acids Res 35:W429–W432.

Kawashima Y, Nagai H, Ishida M, Nagashima Y, Shiomi K. 2003. Primary structure of echotoxin 2, an actinoporin-like hemolytic toxin from the salivary gland of the marine gastropod *Monoplex echo*. Toxicon 42:491–497.

Kem WR. 1988. Sea anemone toxins: structure and action. In: Hessinger DA and Lenhoff HM, editors. The biology of nematocysts. San Diego: Academic Press. p. 375–405.

Kondrat FDL, Struwe WB, Benesch JLP. 2015. Native mass spectrometry: towards high-throughput structural proteomics. In: Owens R, editors. Structural proteomics. Methods in molecular biology. New York: Humana Press. p. 349–371.

Kristan K, Viero G, Maček P, Dalla Serra M, Anderluh G. 2007. The equinatoxin N-terminus is transferred across planar lipid membranes and helps to stabilize the transmembrane pore. FEBS J 274:539–550.

Langdon CJ, Newell RIE. 1990. Utilization of detritus and bacteria as food sources by two bivalve suspension-feeders, the oyster *Crassostrea virginica* and the mussel *Geukensia demissa*. Marine Ecol Prog Ser 58:299–310.

Leonardi A, et al. 2012. *Conus consors* snail venom proteomics proposes functions, pathways, and novel families involved in its venomic system. J Proteome Res 11:5046–5058.

Leone S, Silipo A ELN, Lanzetta R, Parrilli M, Molinaro A. 2007. Molecular structure of endotoxins from gram-negative marine bacteria: an update. Mar Drugs 5:85–112.

Leoni G, et al. 2017. Myticalins: a novel multigenic family of linear, cationic antimicrobial peptides from marine mussels (*Mytilus* spp.). Mar Drugs 15:261.

Li W, Godzik A. 2006. CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659.

Lu A, et al. 2020. Transcriptomic profiling reveals extraordinary diversity of venom peptides in unexplored predatory gastropods of the genus *Clavus*. Genome Biol Evol 12:684–700.

Lüskow F, Riisgård HU. 2018. *In situ* filtration rates of blue mussels (*Mytilus edulis*) measured by an open-top chamber method. Open J Mar Sci 8:395–406.

Maček P. 1992. Polypeptide cytolytic toxins from sea anemones (Actiniaria). FEMS Microbiol Immunol 5:121–129.

Maček P, Belmonte G, Pederzolli C, Menestrina G. 1994. Mechanism of action of equinatoxin II, a cytolysin from the sea anemone *Actinia equina* L. Belonging to the family of actinoporins. Toxicology 87:205–227.

Maček P, Lebez D. 1988. Isolation and characterization of three lethal and hemolytic toxins from the sea anemone *Actinia equina* L. Toxicon 26:441–451.

Macrander J, Daly M. 2016. Evolution of the cytolytic pore-forming proteins (actinoporins) in sea anemones. Toxins (Basel) 8:368.

Marchioretto M, Podobnik M, Dalla Serra M, Anderluh G. 2013. What planar lipid membranes tell us about the pore-forming activity of cholesterol-dependent cytolysins. Biophys Chem 182:64–70.

Marty MT, et al. 2015. Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. Anal Chem 87:4370–4376.

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.

Mechaly AE, et al. 2011. Structural insights into the oligomerization and architecture of eukaryotic membrane pore-forming toxins. Structure 19:181–191.

Menestrina G, Dalla Serra M, Lazarovici P. 2003. Pore-forming peptides and protein toxins. London: CRC Press.

Meunier FA, et al. 2000. Ca(2+) and Na(+) contribute to the swelling of differentiated neuroblastoma cells induced by equinatoxin-II. Toxicon 38:1547–1560.

Mitta G, Hubert F, Noel T, Roch P. 1999. Myticin, a novel cysteine-rich antimicrobial peptide isolated from haemocytes and plasma of the mussel *Mytilus galloprovincialis*. Eur J Biochem 265:71–78.

Mitta G, Vandenbulcke F, Hubert F, Salzet M, Roch P. 2000. Involvement of mytilins in mussel antimicrobial defense. J Biol Chem 275:12954–12962.

Modica MV, Lombardo F, Franchini P, Oliverio M. 2015. The venomous cocktail of the vampire snail *Colubraria reticulata* (Mollusca, Gastropoda). BMC Genomics 16:441.

Murgarella M, et al. 2016. A first insight into the genome of the filter-feeder mussel *Mytilus galloprovincialis*. PLoS One 11:e0151561.

Murzyn K, Róg T, Pasenkiewicz-Gierula M. 2005. Phosphatidylethanolamine-phosphatidylglycerol bilayer as a model of the inner bacterial membrane. Biophys J 88:1091–1103.

Musella M, et al. 2020. Tissue-scale microbiota of the Mediterranean mussel (*Mytilus galloprovincialis*) and its relationship with the environment. Sci Total Environ 717:137209.

Norton RS. 2009. Structures of sea anemone toxins. Toxicon 54: 1075–1088.

Opper B, Bognár A, Heidt D, Németh P, Engelmann P. 2013. Revising lysenin expression of earthworm coelomocytes. Dev Comp Immunol 39:214–218.

Pardos-Blas JR, et al. 2021. The genome of the venomous snail *Lautoconus ventricosus* sheds light on the origin of conotoxin diversity. Gigascience 10:1–15.

Pardos-Blas JR, Irisarri I, Abalde S, Tenorio MJ, Zardoya R. 2019. Conotoxin diversity in the venom gland transcriptome of the magician's Cone, *Pionoconus magus*. Mar Drugs 17:553.

Pardos-Blas JR, Tenorio MJ, Galindo JCG, Zardoya R. 2022. Comparative venomics of the cryptic cone snail species *Virroconus ebraeus* and *Virroconus judaeus*. Mar Drugs 20:149.

Peraro MD, van der Goot FG. 2016. Pore-forming toxins: ancient, but never really out of fashion. Nat Rev Microbiol 14:77–92.

Pettersen EF, et al. 2004. UCSF Chimera-a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612.

Podobnik M, Anderluh G. 2017. Pore-forming toxins in Cnidaria. Semin Cell Dev Biol 72:133–141.

Podobnik M, Kisovec M, Anderluh G. 2017. Molecular mechanism of pore formation by aerolysin-like proteins. Philos Trans Royal Society B 372:20160209.

Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. Cryosparc: algorithms for rapid unsupervised cryo-em structure determination. Nat Methods 14:290–296.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using tracer 1.7. System Biol 67:901–904.

Rey-Campos M, Novoa B, Pallavicini A, Gerdol M, Figueras A. 2020. Comparative genomics reveals a significant sequence variability of myticin genes in *Mytilus galloprovincialis*. Biomolecules 10:943.

Rey-Campos M, Novoa B, Pallavicini A, Gerdol M, Figueras A. 2021. Comparative genomics reveals 13 different isoforms of mytimycins (a-m) in *Mytilus galloprovincialis*. Int J Mol Sci 22:3235.

Rivera de Torre E, et al. 2020. Structural and functional characterization of sticholysin III: a newly discovered actinoporin within the venom of the sea anemone *Stichodactyla helianthus*. Arch Biochem Biophys 689:108435.

Rojko N, et al. 2013. Membrane damage by an alpha-helical pore-forming protein, equinatoxin II, proceeds through a succession of ordered steps. J Biol Chem 288:23704–23715.

Rojko N, Dalla Serra M, Maček P, Anderluh G. 2016. Pore formation by actinoporins, cytolysins from sea anemones. Biochim Biophys Acta 1858:446–456.

Safavi-Hemami H, et al. 2014. Combined proteomic and transcriptomic interrogation of the venom gland of *Conus geographus* uncovers novel components and functional compartmentalization. Mol Cell Proteomics 13:938–953.

Sandoval K, McCormack GP. 2022. Actinoporin-like proteins are widely distributed in the phylum Porifera. Mar Drugs 20:74.

Sher D, et al. 2005. Toxic polypeptides of the hydra-a bioinformatic approach to cnidarian allomones. Toxicon 45:865–879.

Sher D, Fishman Y, Melamed-Book N, Zhang M, Zlotkin E. 2008. Osmotically driven prey disintegration in the gastrovascular cavity of the green hydra by a pore-forming protein. FASEB J 22:207–214.

Shibata H, et al. 2018. The habu genome reveals accelerated evolution of venom protein genes. Sci Rep 8:11300.

Shiomi K, Kawashima Y, Mizukami M, Nagashima Y. 2002. Properties of proteinaceous toxins in the salivary gland of the marine gastropod (*Monoplex echo*). Toxicon 40:563–571.

Shogomori H, Kobayashi T. 2008. Lysenin: a sphingomyelin specific pore-forming toxin. Biochim Biophys Acta 1780:612–618.

Šolinc G, et al. 2022. Pore-forming moss protein bryoporin is structurally and functionally related to actinoporins from evolutionarily distant cnidarians. J Biol Chem. 298:102455

Sudewi AAR, et al. 2019. Selecting potential neuronal drug leads from conotoxins of various venomous marine cone snails in Bali, Indonesia. ACS Omega 4:19483–19490.

Sunagar K, Moran Y. 2015. The rise and fall of an evolutionary innovation: contrasting strategies of venom evolution in ancient and young animals. PLoS Genet 11:e1005596.

Šuput D. 2009. *In vivo* effects of cnidarian toxins and venoms. Toxicon 54:1190–1200.

Takara T, et al. 2011. Purification, molecular cloning, and application of a novel sphingomyelin-binding protein (clamlysin) from the brackishwater clam, *Corbicula japonica*. Biochim Biophys Acta 1811:323–332.

Tanaka K, Caaveiro JM, Morante K, Gonzalez-Manas JM, Tsumoto K. 2015. Structural basis for self-assembly of a cytolytic pore lined by protein and lipid. Nat Commun 6:6337.

Teufel F, et al. 2022. Signalp 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol. 40:1023–1025.

Verma P, Gandhi S, Lata K, Chattopadhyay K. 2021. Pore-forming toxins in infection and immunity. Biochem Soc Trans 49:455–465.

Vonk FJ, et al. 2013. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. Proc Natl Acad Sci USA 110:20651–20656.

Wagner GP, Kin K, Lynch VJ. 2012. Measurement of mRNA abundance using RNA-seq data: rpkm measure is inconsistent among samples. Theory Biosci 131:281–285.

Weese DA, Duda TF Jr. 2019. Effects of predator-prey interactions on predator traits: differentiation of diets and venoms of a marine snail. Toxins (Basel) 11:299.

Young G, et al. 2018. Quantitative mass imaging of single biological macromolecules. Science 360:423–427.

Zhang L, et al. 2015. Massive expansion and functional divergence of innate immune genes in a protostome. Sci Rep 5:8693.

Associate editor: Dr. Adam Eyre-Walker