



OPEN ACCESS

EDITED BY
Salvatore Santamaria,
University of Surrey, United Kingdom

REVIEWED BY
Ana Crnkovic,
National Institute of Chemistry, Slovenia
Rajan Sankaranarayanan,
Centre for Cellular & Molecular Biology
(CCMB), India

*CORRESPONDENCE
F. Xavier Gomis-Rüth,
✉ fxgr@ibmb.csic.es
Walter Stöcker,
✉ stoecker@uni-mainz.de

SPECIALTY SECTION
This article was submitted to Protein
Biochemistry for Basic and Applied
Sciences,
a section of the journal
Frontiers in Molecular Biosciences

RECEIVED 26 October 2022
ACCEPTED 30 November 2022
PUBLISHED 04 January 2023

CITATION
Gomis-Rüth FX and Stöcker W (2023),
Structural and evolutionary insights into
astacin metallopeptidases.
Front. Mol. Biosci. 9:1080836.
doi: 10.3389/fmolb.2022.1080836

COPYRIGHT
© 2023 Gomis-Rüth and Stöcker. This is
an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

Structural and evolutionary insights into astacin metallopeptidases

F. Xavier Gomis-Rüth^{1*} and Walter Stöcker^{2*}

¹Proteolysis Laboratory, Department of Structural Biology, Molecular Biology Institute of Barcelona (IBMB), Higher Scientific Research Council (CSIC), Barcelona, Catalonia, Spain, ²Institute of Molecular Physiology (IMP), Johannes Gutenberg-University Mainz (JGU), Mainz, Germany

The astacins are a family of metallopeptidases (MPs) that has been extensively described from animals. They are multidomain extracellular proteins, which have a conserved core architecture encompassing a signal peptide for secretion, a prodomain or prosegment and a zinc-dependent catalytic domain (CD). This constellation is found in the archetypal name-giving digestive enzyme astacin from the European crayfish *Astacus astacus*. Astacin catalytic domains span ~200 residues and consist of two subdomains that flank an extended active-site cleft. They share several structural elements including a long zinc-binding consensus sequence (HEXXHXXGXXH) immediately followed by an EXRXDRD motif, which features a family-specific glutamate. In addition, a downstream SIMHY-motif encompasses a “Met-turn” methionine and a zinc-binding tyrosine. The overall architecture and some structural features of astacin catalytic domains match those of other more distantly related MPs, which together constitute the metzincin clan of metallopeptidases. We further analysed the structures of PRO-, MAM, TRAF, CUB and EGF-like domains, and described their essential molecular determinants. In addition, we investigated the distribution of astacins across kingdoms and their phylogenetic origin. Through extensive sequence searches we found astacin CDs in > 25,000 sequences down the tree of life from humans beyond Metazoa, including Choanoflagellata, Filasterea and Ichtyospora. We also found < 400 sequences scattered across non-holozoan eukaryotes including some fungi and one virus, as well as in selected taxa of archaea and bacteria that are pathogens or colonizers of animal hosts, but not in plants. Overall, we propose that astacins originate in the root of Holozoa consistent with Darwinian descent and that the latter genes might be the result of horizontal gene transfer from holozoan donors.

KEYWORDS

evolution of metallopeptidases, catalytic domain (CD), darwinian descent, horizontal gene transfer (HGT), phylogeny of enzymes

1 Introduction

The astacins are a family of extracellular zinc-dependent metalloproteinases (MPs) named after a digestive enzyme discovered in the 1960s in the European crayfish *Astacus astacus* L., as named by Linnaeus (Linnaeus, 1758), which was also referred to as *Astacus fluviatilis* F. by Fabricius (Fabricius, 1796). The enzyme was first named “*Astacus* protease” or “low-molecular-weight protease” (Pfleiderer et al., 1967), and the designation “astacin” was coined after related proteins had been found in other organisms [for reviews, see (Dumermuth et al., 1991; Jiang and Bond, 1992; Stöcker et al., 1993; Bond and Beynon, 1995; Zwilling and Stöcker, 1997; Gomis-Rüth et al., 2012a; Stöcker et al., 2013a; Stöcker et al., 2013b; Bond, 2019)]. Moreover, the astacins were the first identified members of the metzincin clan of MPs together with the matrix metalloproteinases, serralysins and adamalysins/a-disintegrin- and-metalloproteinases (ADAMs), which share common topologies and zinc-binding environments as inferred from structural studies (Bode et al., 1993; Stöcker and Bode, 1995; Gomis-Rüth, 2003; Sterchi, 2008; Gomis-Rüth, 2009; Cerdà-Costa and Gomis-Rüth, 2014; Arolas et al., 2018). Astacins function as protein degraders during digestion, developmental tissue turnover and differentiation, and embryonic hatching, but also as sophisticated shedders of membrane-bound substrates (Jiang and Bond, 1992; Gomis-Rüth et al., 2012a; Stöcker et al., 2013b; Bond, 2019). They are subdivided into the bone morphogenetic protein 1 (BMP1)/“Tolloid”-like proteinases (BTPs), meprin, hatching enzymes, and other astacins (Sterchi et al., 2008). Recent genomes have unravelled a plethora of genes encoding proteins annotated as astacins in various organisms within metazoans, which date back to 760 million years ago (Berman 2019).

In this article, we both dissected reported molecular structures and calculated new high-confidence computational models to analyse the molecular determinants of the most relevant astacin domains. Based on structural and molecular specifications of the prototypic astacin catalytic domain (CD), we further performed comprehensive sequence similarity searches to identify potential family members outside vertebrates to locate the origin of astacins according to Darwinian descent. Finally, we screened and reviewed the literature available for functional and evolutionary implications of the distinct astacin subfamilies outside vertebrates.

2 Results and discussion

2.1 Architecture and function of relevant astacin-family domains

Astacins across all phyla minimally comprise a zinc-binding CD, which is preceded by an upstream propeptide or prodomain

(PRO) for latency and a signal peptide (S) for targeting to the plasmalemma or the extracellular space in animals [(Gomis-Rüth et al., 2012a); Figure 1]. However, most astacins are multidomain proteins, which have acquired a diverse set of additional domains (Figure 1). We retrieved reported experimental crystal structures of a CD, a “MAM” domain [first identified in meprin, A5 protein and receptor protein tyrosine phosphatase μ ; (Cismasiu et al., 2004)], a “TRAF” domain [reminiscent of tumour-necrosis-factor receptor-associated factor; (Park, 2018)] and PRO domains, and computed high-confidence computational models of the “CUB” domain [first identified at the sequence level in complement subcomponents C1r/C1s, Uegf and BMP1; (Bork and Beckmann, 1993)] and epidermal growth factor (EGF)-like domains (see the Methods section) for their molecular analysis.

The CD is ascribed to protein family Pfam-01400 (Figure 1) and spans ~200 residues. It contains two or three disulfide bonds at variable positions (Gomis-Rüth et al., 2012a) and is divided into an upper N-terminal subdomain and a lower C-terminal subdomain by an extended active-site cleft, as first revealed by the crystal structure of archetypal crayfish astacin (Bode et al., 1992; Gomis-Rüth et al., 1993) (Figure 2A). The N-terminal subdomain is rich in regular secondary structure and contains a hallmark five-stranded β -sheet (β 1– β 5; Figure 2A), whose lowermost strand β 4 frames the upper rim of the active-site cleft when viewed in the standard orientation of MPs (Gomis-Rüth et al., 2012b) (Figure 2A, left), and two helices: the “backing helix” and the “active-site helix”. The latter encompasses most of a characteristic zinc-binding motif (HEXXHXXGXXH; amino-acid one letter code; X stands for any residue), which is found across astacins and other metzincin families (Bode et al., 1993; Stöcker et al., 1993; Yiallourous et al., 2000; Gomis-Rüth et al., 2012a; Cerdà-Costa and Gomis-Rüth, 2014; Arolas et al., 2018). The helix includes the first two zinc-liganding histidines and the general base/acid glutamate required for catalysis (Arolas et al., 2018). After the glycine of the motif, the chain undergoes a sharp turn and enters the C-terminal subdomain, which is more irregular and just encompasses a short β -ribbon (β 6– β 7) and a “C-terminal helix” as regular secondary structure (Figure 2A). The C-terminal subdomain provides two more zinc ligands, viz., the third histidine of the motif and a downstream tyrosine, which is swung out in a “tyrosine-switch” motion upon substrate binding to stabilize the reaction intermediate during catalysis (Grams et al., 1996; Yiallourous et al., 2000). This tyrosine is found two positions after another conserved element within metzincins, the “Met-turn” methionine (Bode et al., 1993; Tallant et al., 2010), which creates a hydrophobic base for the metal-binding site (Tallant et al., 2010). The tyrosine and the methionine are embedded in a characteristic SIMHY-motif in astacins (Stöcker et al., 1993).

Finally, another structural characteristic of astacin CDs is an unaccessible N-terminus (Gomis-Rüth et al., 2012a; Guevara et al., 2022). Maturation cleavage occurs at a bond that is

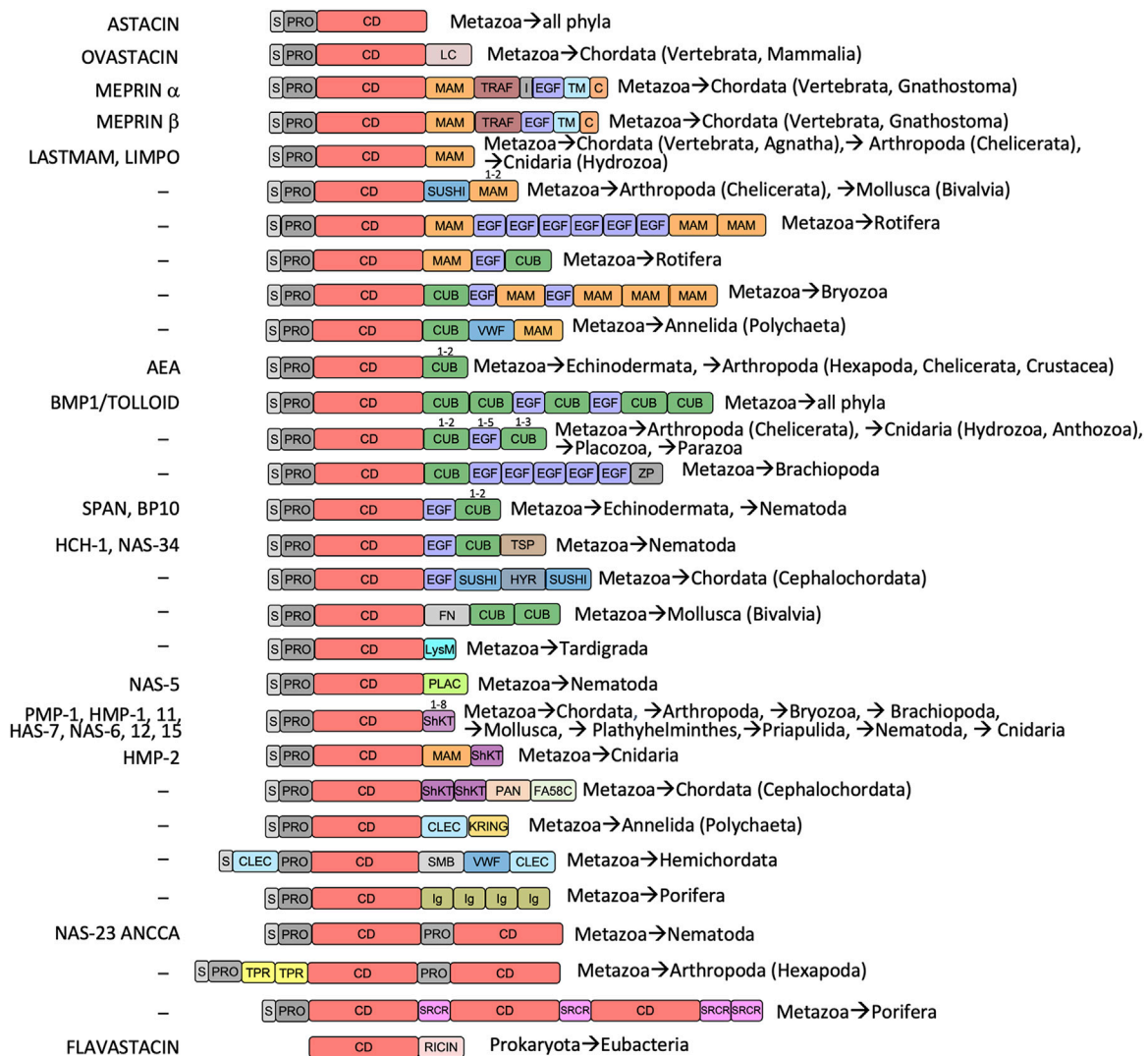


FIGURE 1

Astacin domain combinations and phylogenetic occurrence. Metazoan astacins minimally comprise an N-terminal signal-peptide for extracellular secretion (S), a propeptide or prodomain conferring latency (PRO) and a catalytic zinc-dependent metalloprotease domain (CD). Moreover, most astacins evince additional domains, listed with *Prosite* database codes (PS; <https://prosite.expasy.org>): ABC (ABC-transporter; PS00211), CD (catalytic protease domain; PS51864), C (cytoplasmic tail), CLEC (C-type lectin; PS50041), CUB (found in complement subcomponents C1r/C1s, Uegf and BMP1; PS01180), EGF (epidermal growth factor-like; PS00022), FA58C (factor-5/8 type-C domain; PS500229), FN3 (fibronectin type-III domain; PS508539), HYR (hyalin repeat protein; PS50825), I (intervening domain in meprin α , contains a furin cleavage site), IG (immunoglobulin-like; PS508359), KRING (kringle; PS00021), LC (low complexity domain, disordered), LCCL (*Limulus* C-domain; PS50820), LYSM (extracellular receptor domain; PS51782), MAM (found in meprin, A5 protein and receptor protein tyrosine phosphatase μ ; PS00740), MATH (meprin and traf homology domain; PS50144), PAN (also dubbed APPLE; found in plasma kallikrein and factor XI; PS50948), PLAC (polycystin-1, lipoxigenase and α -toxin; PS50095), PTX (pentraxin; PS51828), RICIN (ricin-type lectin; PS50231), SH2 (SARC homology domain; PS50001), ShKt (K⁺-channel-blocking *Stichodactyla helianthus* toxin; PS51670), SMB (somatomedin; PS50958), SRCR (cysteine-rich scavenger receptor; PS50287), SUSHI (sushi adhesion domain; PS50923), TPR (tetratricopeptide repeat; PS50005), TSP (thrombospondin-like domain; PS50092), VWF (von-Willebrand-factor domain; PS50234), ZF-UBR (UBR-type zinc-finger; PS51157) and ZP2 (zona pellucida protein 2 domain; PS51034). On the left, typical astacin family members are listed, for which physiological functions are documented (see [Supplementary Table S1](#) for complete protein and gene names, and *UniProt* access codes). Phylogenetic occurrences are indicated on the right.

occluded in the zymogen, which entails that partial unfolding of the segment flanking the activation site and/or preliminary cleavages are required for activation (Guevara et al., 2010; Arolas et al., 2012; Guevara et al., 2022). Upon final cleavage,

the first six or seven residues of the mature enzyme are amply repositioned and penetrate the mature enzyme moiety, so the first two or three residues are completely buried in the molecular structure (Gomis-Rüth et al., 1993; Arolas et al., 2012; Ran et al.,

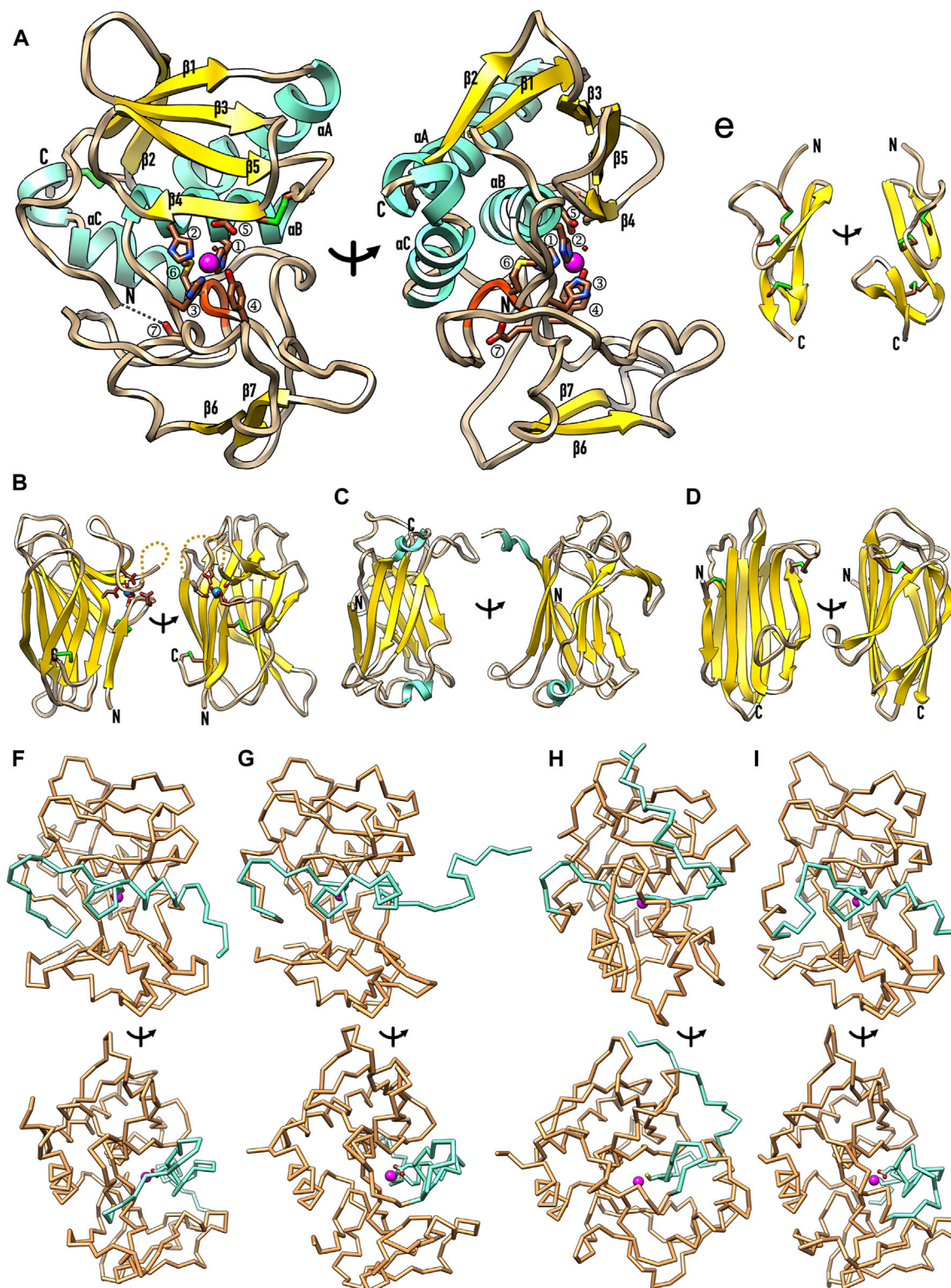


FIGURE 2

Representative structures of the most relevant astacin domains. (A) Ribbon-type plot of the mature *Astacus astacus* crayfish astacin catalytic domain [PDB 1AST; residues 50–251, see UniProt P07584; (Bode et al., 1992; Gomis-Rüth et al., 1993)], which is shown in the standard orientation of (Continued)

FIGURE 2 (Continued)

MPs [left; (Gomis-Rüth et al., 2012b)] and vertically rotated by 90 degrees (right). Regular secondary structure elements are shown as yellow β -strands ($\beta 1$ – $\beta 7$) and aquamarine α -helices (αA – αC). The first five strands constitute the typical five-stranded β -sheet of astacins (Gomis-Rüth et al., 2012a) and the helices are dubbed “backing helix” (αA), “active-site helix” (αB) and “C-terminal helix” (αC). The latter is split in two by a kink. Unbound mature astacin has its catalytic zinc cation (magenta sphere) bound in trigonal-bipyramidal coordination by the three histidines ($\textcircled{1}$ – $\textcircled{3}$) of a characteristic zinc-binding motif [HEXXHXXGXXH; (Bode et al., 1993)] plus a more distal downstream tyrosine ($\textcircled{4}$) and the catalytic solvent molecule [small red sphere; (Arolas et al., 2018)]. The glutamate within the motif ($\textcircled{5}$) is the general base/acid for catalysis (Arolas et al., 2018). The “Met-turn” with the conserved methionine ($\textcircled{6}$; (Bode et al., 1993; Tallant et al., 2010)) is shown as an orange ribbon. The mature N-terminal residue (labelled N) is bound to the family-specific glutamate (E^{103}) [$\textcircled{7}$; (Gomis-Rüth, 2003)] after the third zinc-binding histidine. The C-terminus is also labelled (C) and the two disulfide bonds of the structure (C^{42} – C^{198} and C^{64} – C^{84}) are further displayed with sulphur atoms in green. (B) The structure of the unique EGF-like domain of human BMP1 predicted with *AlphaFold* (Jumper et al., 2021) shows two β -ribbons and three disulfide bonds. Two orthogonal orientations are displayed. (C) Experimental structure of the MAM domain of meprin β [PDB 4GWM; (Arolas et al., 2012)] in two orthogonal orientations. The β -sandwich domain (residues 259–427, see *UniProt* Q16820) features two disulfide bonds and a structural sodium cation (blue sphere) octahedrally coordinated by six protein oxygens. (D) Structure of the first CUB domain of human BMP predicted with *AlphaFold* in two orthogonal orientations, which show a β -sandwich architecture with two disulfide bonds. (E) Experimental structure of the TRAF domain of meprin β [PDB 4GWM; (Arolas et al., 2012)] in two orthogonal orientations. The β -sandwich domain (residues 428–597, see *UniProt* Q16820) has two short helices and a β -ribbon grafted into strand-connecting loops. (F–I) Experimental zymogen structures as Ca-traces in standard orientation (top panels) and after a vertical 90-degree rotation (bottom panels) of (F) crayfish astacin [PDB 3LQ0; (Guevara et al., 2010)], (G) human meprin β [PDB 4GWM; (Arolas et al., 2012)], (H) myroilysin from the bacterium *Myroides* sp. [PDB 5GWD; (Xu et al., 2017)] and (I) astacin from the horseshoe crab *Limulus polyphemus* [PDB 8A28; (Guevara et al., 2022)]. Only the PROs (aquamarine) and CDs (sandy brown) are displayed for clarity, together with the catalytic zinc ions (magenta spheres) and the side chains of the respective aspartate/cysteine-switch residue.

2020; Guevara et al., 2022). Moreover, the new N-terminus binds the “family-specific” glutamate immediately after the third zinc-binding histidine (Bode et al., 1993; Gomis-Rüth, 2003), either directly through its side chain or through the α -amino group *via* a solvent molecule (Figure 2A). This feature is unique among MPs and reminiscent of trypsin-like serine endopeptidases, which dedicate an aspartate next to the catalytic serine to bind the likewise buried mature N-terminus (Bode et al., 1986). The astacin glutamate is immediately followed by an XXXRDRD motif (Gomis-Rüth, 2003) whose charged residues establish interactions relevant for domain stability.

A MAM domain is found after the CD in meprins α and β , *Limulus* and *Hydra* astacins and other (potential) family members (Figure 1) (Arolas et al., 2012; Eckhard et al., 2021; Guevara et al., 2022). The crystal structure of human meprin β (Arolas et al., 2012; Eckhard et al., 2021) reveals that its MAM domain is a β -sandwich consisting of a four- and a five-stranded antiparallel β -sheet, which are twisted and rotated ~ 25 degrees relative to each other (Figure 2B). The domain conforms to a jelly-roll architecture featuring two four-stranded Greek key motifs and is connected by two disulfide bonds. Furthermore, the domain has a sodium-binding site, at which the cation is octahedrally coordinated by six oxygens from side chains and the main chain of the protein (Figure 2B). The overall architecture of the domain conforms to the structural criteria defined for the MAM protein family (Pfam-00629), which was identified *in silico* in meprin α and β , A5 protein, and receptor protein tyrosine phosphatase μ (Beckmann and Bork, 1993). Comparison with other MAM domains reveals that the central β -sandwich is conserved but the loops responsible for functionality deviate, as well as the metal-binding capacity and arrangement (Aricescu et al., 2006; Yelland and Djordjevic, 2016). This domain appears to have adhesive functions and, in meprin β , it contributes to

dimerization by bringing the CD and TRAF domains together (Arolas et al., 2012; Eckhard et al., 2021).

Uniquely for astacins, meprins α and β exhibit a TRAF domain downstream of the MAM domain (Figure 1) (Arolas et al., 2012; Eckhard et al., 2021). The crystal structure of human meprin β (Arolas et al., 2012; Eckhard et al., 2021) shows that this moiety features two twisted four-stranded antiparallel β -sheets, which are rotated ~ 40 degrees relative to each other (Figure 2C, left) and give rise to a flatter sandwich than in MAM (compare Figure 2B, right and Figure 2C, right). The strands are connected by loops of variable length, which include two short helical segments plus a short β -ribbon and give rise to a double Greek key architecture. The second Greek key is inserted into the first one but does not form a jelly roll. The only cysteine of this domain (C^{492}) is buried and unbound, the N- and the C-terminus are on contiguous β -strands of the front β -sheet (Figure 2C, left), the C-terminus protrudes from the top surface of the domain (Figure 2C, right). In general, the TRAF domain of meprin β resembles tumor-necrosis-factor receptor-associated factors, which are mediators of cell activation engaged in homo- and heterodimerization and originated the TRAF protein family (Pfam-00917) (Zapata et al., 2001).

Further relevant for astacins are CUB domains (Figure 1), which were first identified in complement subcomponents C1r/C1s, Uegf and BMP1 (Bork and Beckmann, 1993) and form protein family Pfam-00431. They occur in BTP-subfamily astacins including BMP1, as well as in echinoderm astacins, a paralogue within *A. astacus* and several other orthologues in up to five copies (Figure 1). According to a highly reliable *AlphaFold* computational model (see Figure 2D and the Methods section), the first CUB domain of human BMP1 would be a β -sandwich made of an antiparallel four-stranded β -sheet and a mixed parallel/antiparallel five-stranded β -sheet, which would be both partially twisted. Their strands would be nearly parallel

(Figure 2D, left), in contrast to MAM (Figure 2B, left) and TRAF (Figure 2C, left), and connected by mostly short loops. Two disulfide bonds would crosslink the domain. CUB domains were apparently present in the last common ancestor of eumetazoans and are currently found in synaptic proteins (González-Calvo et al., 2022). Remarkably, combinations of CUB and MAM domains are found in neuropilins, which are receptors for axon guidance cues and play synaptic roles (González-Calvo et al., 2022). Moreover, a CUB domain is engaged in the “Venus-flytrap” mechanism of inhibition of endopeptidases by the human pan-peptidase tetrameric inhibitor α_2 -macroglobulin. It participates in major structural rearrangement of the C-terminal half of the protomer, which further includes three more domains (Marrero et al., 2012; Goulas et al., 2017; Luque et al., 2022). A CUB domain also participates in the “snap-trap” mechanism of monomeric α_2 -macroglobulin-related inhibitors from commensal and pathogenic bacteria such as *Escherichia coli* and *Salmonella enterica* (Wong and Dessen, 2014; Garcia-Ferrer et al., 2015; Goulas et al., 2017).

Next, EGF domains (Pfam-00008) are widely present in up to six copies in several astacins, including meprins, BTPs and proteins from nematodes and echinoderms (Figure 1). Generally, they are found in many animal proteins in the extracellular part of membrane-bound or secreted proteins (Bork et al., 1996). In meprin β , the EGF-like domain is considered a hinge domain, which moves the dimer from a membrane-proximal position for cleavage of transmembrane substrates, such as the amyloid precursor protein, to a membrane-distal position upon binding to its endogenous inhibitor fetuin B (Karmilin et al., 2019; Eckhard et al., 2021). We obtained a generally reliable *AlphaFold* computational model (see the Methods section) for the EGF-like domain of human BMP1 (see Figure 2E). It revealed a ~40-residue structure cross-connected by three disulfide bonds for structural integrity and two β -hairpins, which overall conform to the standard architecture of these domains (Wouters et al., 2005).

Finally, large diversity is found across astacin PROs, which range between 34 and 486 residues and just share the motif FXGDI among animal orthologues (Guevara et al., 2010; Gomis-Rüth et al., 2012a). The PROs of crayfish astacin (Figure 2F), human meprin β (Figure 2G), bacterial myroilysin (Figure 2H) and horseshoe crab astacin (Figure 2I) have been structurally characterized. They revealed essentially unstructured peptides running along the cleft of the CD in the opposite direction of a true substrate, which precludes their intramolecular cleavage (Guevara et al., 2010; Arolas et al., 2012; Xu et al., 2017; Guevara et al., 2022). The catalytic solvent molecule bound to the catalytic zinc ion in the mature CD (Figure 2A) is replaced by either the aspartate of the motif in the three animal zymogens (Guevara et al., 2010; Arolas et al., 2012; Guevara et al., 2022) or a cysteine in the bacterial enzyme (Xu et al., 2017), which lacks the motif. These residues operate according to an “aspartate-switch” or “cysteine-switch” mechanism of latency, respectively. Such

mechanisms have been also reported, among others, for the MPs fragilysin-3 from *Bacteroides fragilis* (Goulas et al., 2011) and matrix metalloproteinases (Springman et al., 1990; Rosenblum et al., 2007), respectively.

2.2 Astacins possibly originate in Holozoa

Multicellularity presumably originated several times in unicellular opisthokont holozoans, which have been suggested as precursors of metazoans [(Sebé-Pedrós et al., 2017; Berman 2019)]. To better understand the hierarchical clustering of the distinct phyla within Holozoa, which originate 1.3 billion years ago (Berman 2019), and to put our phylogenetic studies into context, we tentatively assembled a consensus dendrogram based on current literature (Figure 3) given the apparent disparity in the available models (see Section 3.2). This hypothesis entails that Holozoa would split into Teretosporea, themselves consisting of Corallochytra/Pluriformea (alias Opisthokonta incertae sedis) and Ichthyosporea, and Filozoa. These, in turn, would divide into Filasterea and Choanozoa. The latter would consist of Choanoflagellata, which are unicellular flagellates, and Metazoa, which encompass the multicellular animals and date back to about 760 million years ago (Berman 2019). Up the tree, Bilateria would englobe animals with a plane of symmetry (including Xenacoelomorpha), except echinoderms, which evince post-larval (secondary) pentaradial symmetry. They sequentially would team up with Cnidaria, Placozoa, Porifera and Ctenophora to eventually form Metazoa (Figure 3).

We performed searches for astacins in several protein and gene databases (see Section 3.1), which revealed > 25,000 entries for potential peptidases of the M12A family. This is how astacins are defined in the MEROPS database of peptidases and their inhibitors [www.ebi.ac.uk/merops; (Rawlings and Bateman, 2021)]. In addition, > 12,000 sequences from > 1,000 species of identified and putative family members were found within family PF01400 within the PFAM database (Mistry et al., 2021). At this point, high-confidence manually curated sequence searches were performed with the sequence of the mature CD of crayfish astacin. The resulting hit sequences were verified to span the entire CD and contain the intact zinc-binding motif, as well as the family-specific glutamate followed by the XXRXDRD and SIMHY motifs, with just minimal conservative substitutions (Figure 4 reproduces selected aligned example sequences). They were further checked to contain a PRO with the zinc-blocking aspartate. A subgroup of sequences was chosen for alignments, phylogenetic tree construction and physiological considerations (Supplementary Table S1; Sections 2.4–2.7). In addition, Table 1 presents a selection of described and potential non-vertebrate metazoan astacins.

We consistently found astacin sequences from humans down the tree of life until the root of subphylum Vertebrata (a selection

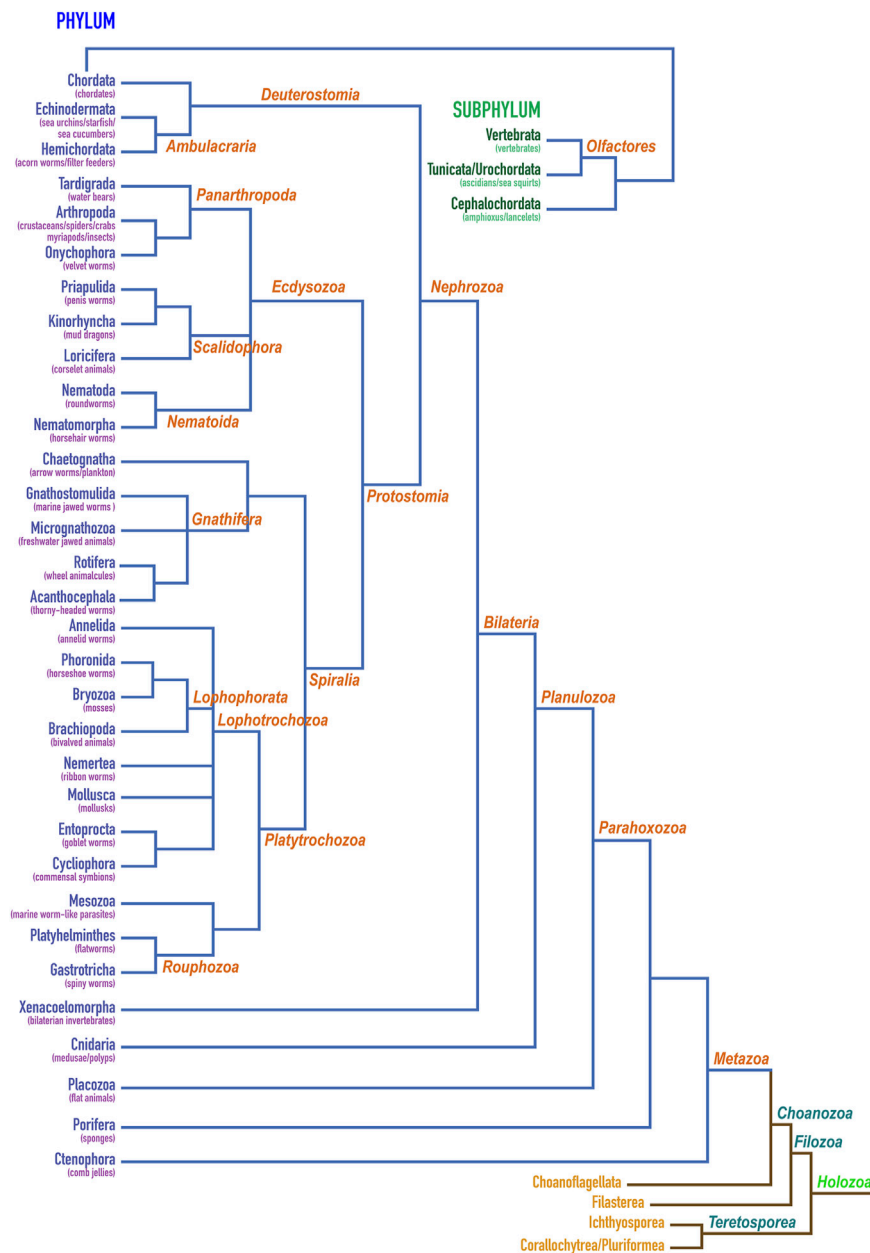
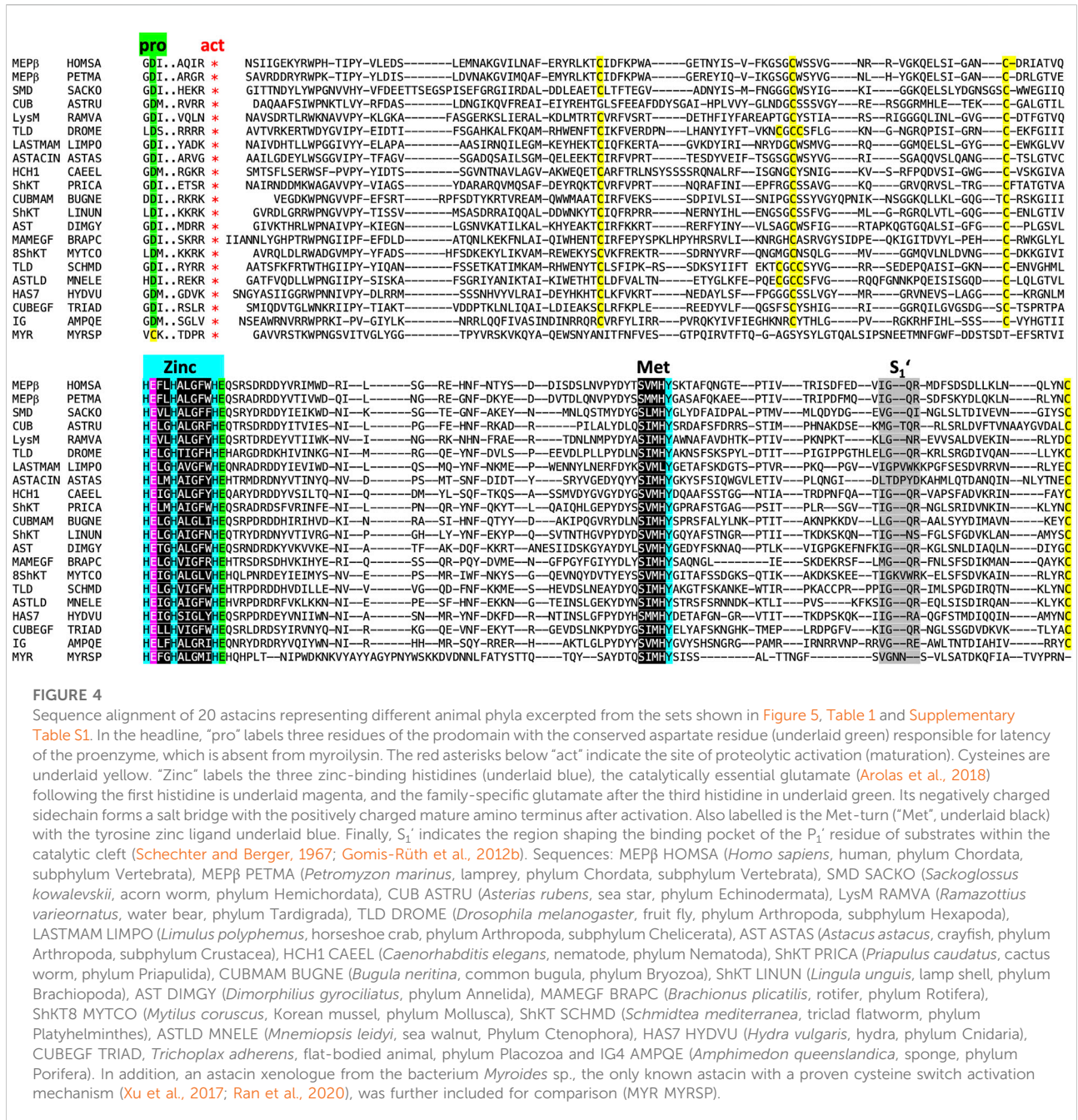


FIGURE 3
 Classification of holozoans. Dendrogram depicting the herein proposed hierarchical clustering of phyla within holozoans assembled based on current literature (Ryan et al., 2010; Ruggiero et al., 2015; Torruella et al., 2015; Cannon et al., 2016; Lu et al., 2017; Sebé-Pedrós et al., 2017; Whelan et al., 2017; Adl et al., 2019; Giribet et al., 2019; Laumer et al., 2019; Marlétaz et al., 2019; Sogabe et al., 2019; Hickman et al., 2020; Schoch et al., 2020; Schulze and Kawauchi, 2021). Phylum Chordata is further shown for its constituting subphyla Vertebrata, Tunicata/Urochordata and Cephalochordata. The first two give rise to Olfactores.

is provided by [Supplementary Table S1](#)), among which the jawless fishes (Agnatha) are most basal, with two extant genera: lampreys and hagfishes ([Supplementary Table S1](#)). Vertebrata associate with the subphyla Tunicata/Urochordata, which includes sea squirts and the base tunicate *Ciona intestinalis*, and Cephalochordata, which features the lancelet

(amphioxus), to form the phylum Chordata ([Figure 3](#)). These taxa also evinced abundant astacins. In addition, we could find sequences for the phyla Echinodermata and Hemichordata within Ambulacraria, which together with Chordata form Deuterostomia. Within Ecdysozoa, we could retrieve sequences from Tardigrada and Arthropoda but not



Onychophora within Panarthropoda; Priapulida but not Kinorhyncha or Loricifera within Scaldiphora; and Nematoda but not Nematomorpha within Nematoida. Out of Gnatifera plus Chaetognatha, we could only get hits for Rotifera. As to Lophotrochozoa, we found astacins in Annelida, Bryozoa, Brachiopoda and Mollusca. Next, we identified potential orthologues within Platyhelminthes but not Mesozoa and Gastrotricha. All these phyla constitute Protostomia, which together with Deuterostomia form Nephrozoa (Figure 3). The latter give rise to Bilateria together with Xenacoelomorpha, for

which we could find sequences in two organisms (Table 1). Finally, we could identify astacins in all the remaining more primitive metazoan phyla: Cnidaria, Placozoa, Porifera and Ctenomorpha. All these results (Figure 3; Table 1) suggest that astacins are systematically present at least down to the root of Metazoa.

Choanoflagellata have been discussed as animal precursors for a long time, since they are similar to choanocytes of sponges (Hickman et al., 2020). However, a search for astacins within *Monosiga brevicolis* (King et al., 2008), *Salpingoeca rosetta*

TABLE 1 Representative sequences of studied and potential non-vertebrate metazoan astacins.

Phylum Chordata
Subphylum Tunicata/Urochordata
Sea squirt <i>Halocynthia roretzi</i> Cho et al. (2008); base tunicate <i>Ciona intestinalis</i> Davis and Smith (2002)
Subphylum Cephalochordata
Florida lancelet <i>Branchiostoma floridae</i> UP C3Y5H0; lancelet/amphioxus <i>Branchiostoma belcheri</i> UP A0A6P5AEI0, UP A9JR45, UP A0A6P4Z1P4, UP A0A6P5ALI5
Phylum Echinodermata
Sea urchin <i>Paracentrotus lividus</i> Lepage et al. (1992); sea urchin <i>Strongylocentrotus purpuratus</i> Reynolds et al. (1992); sea cucumber <i>Holothuria glaberrima</i> Mashanov et al. (2012); sea star <i>Asterias rubens</i> GB OP067654.1
Phylum Hemichordata
Acorn worm <i>Saccoglossus kowalevskii</i> Freeman et al. (2012)
Phylum Tardigrada
Water bear <i>Ramazzotius varieornatus</i> UP A0A1D1VZ89, UP A0A1D1VI43; moss piglet <i>Hypsibius exemplaris</i> GB OWA53846; water bear <i>Hypsibius dujardini</i> UP A0A1W0WK06
Phylum Arthropoda
Subphylum Crustacea
Crayfish <i>Astacus astacus</i> Stöcker et al. (1990); Kamchatka crab <i>Paralithodes camtschatica</i> Semenova et al. (2006); common fish louse <i>Argulus foliaceus</i> AmbuAli et al. (2020); water fleas <i>Daphnia pulex</i> Schwerin et al. (2009) and <i>Daphnia magna</i> UP A0A0P6DSE7, UP A0A0P5US34, UP A0A0P6HCB0; shrimp <i>Penaeus vannamei</i> UP A0A423TAW0, UP A0A3R7NMP2, UP A0A3R7N8Y5
Subphylum Chelicerata
Brazilian brown recluse spider <i>Loxosceles intermedia</i> da Silveira et al. (2007); Chilean/Peruvian recluse spider <i>Loxosceles laeta</i> (Medina-Santos et al., 2019); yellow garden spider <i>Argiope aurantia</i> Foradori et al. (2006); Adanson's house jumper <i>Hasarius adansonii</i> Diab et al. (2021); mygalomorph spider <i>Actinopus</i> spp. Prosdocimi et al. (2011); orbweaver spider <i>Gasteracantha cancriformis</i> Prosdocimi et al. (2011); Brazilian tarantula <i>Grammostola iheringi</i> Borges et al. (2016); spitting spider <i>Scytodes thoracica</i> Zobel-Thropp et al. (2014a); brown widow spider <i>Latrodectus geometricus</i> Khamtorn et al. (2020); black widow spider <i>Latrodectus hesperus</i> UP E7D164; cellar spider <i>Physocyclus mexicanus</i> Zobel-Thropp et al. (2019); wolf spider <i>Pardosa pseudoannulata</i> Huang et al. (2018); Brazilian white-knee tarantula <i>Acanthoscurria geniculata</i> and African social velvet spider <i>Stegodyphus mimosarum</i> Walter et al. (2017); African hermit spider <i>Nephilengys cruentata</i> Fuzita et al. (2015a); primitive hunting spider <i>Plectreurys tristis</i> Zobel-Thropp et al. (2014b); yellow scorpion <i>Tityus serrulatus</i> Fuzita et al. (2015b); horseshoe crab <i>Limulus polyphemus</i> (Becker-Pauly et al., 2009); wheat curl mite <i>Aceria tosicella</i> UP A0A8B8R4B3
Subphylum Hexapoda
Fruitfly <i>Drosophila melanogaster</i> Shimell et al. (1991); sand fly <i>Lutzomyia longipalpis</i> Jochim et al. (2008); maybeetle <i>Melolontha melolontha</i> Wagner et al. (2002); tse-tse fly <i>Glossina morsitans</i> Yan et al. (2002); fall armyworm <i>Spodoptera frugiperda</i> Ferreira et al. (2007); Bertha armyworm <i>Mamestra configurata</i> Toprak et al. (2016); silkworm <i>Bombyx mori</i> Lu et al. (2010); Chinese oak silkworm <i>Antheraea pernyi</i> (Tang et al. 2011a); Chinese wild silkworm <i>Bombyx mandarina</i> Tang et al. (2011b); Australian sheep blowfly <i>Lucilia cuprina</i> Young et al. (2000); ixodid cattle tick <i>Rhipicephalus microplus</i> Barnard et al. (2012); American dog tick <i>Dermacentor variabilis</i> Sonenshine et al. (2011); spinybacked orbweaver <i>Gasteracantha cancriformis</i> (Prosdocimi et al. 2011); African malaria mosquito <i>Anopheles gambiae</i> Riehle et al. (2002); yellow-fever mosquito <i>Aedes aegypti</i> UP Q17KW5, UP Q16JR6; sandfly <i>Phlebotomus kandelakii</i> A0A6B2EJJ9; biting mite <i>Culicoides sonorensis</i> UP A0A336L7R2
Phylum Priapulida
Cactus worm <i>Priapulid caudatus</i> GB XP_014678976
Phylum Nematoda
Nematode <i>Caenorhabditis elegans</i> Möhrle et al. (2003), Park et al. (2010); parasite nematode <i>Trichinella spiralis</i> Lun et al. (2003); roundworm <i>Onchocerca volvulus</i> Borchert et al. (2007); dog hookworms <i>Ancylostoma caninum</i> Williamson et al. (2006) and <i>Ancylostoma ceylanicum</i> Baska et al. (2013); brown stomach worm <i>Teladorsagia circumcincta</i> Stepek et al. (2015); New World hookworm <i>Necator americanus</i> Ranjit et al. (2006); brown stomach worm <i>Ostertagia ostertagi</i> de Maere et al. (2005); rat small-intestine nematode <i>Strongyloides ratti</i> Soblik et al. (2011); human small-intestine nematode <i>Strongyloides stercoralis</i> Gómez Gallego et al. (2005), Varatharajulu et al. (2011); interstine nematode <i>Strongyloides venezuelensis</i> Yoshida et al. (2011); threadworm <i>Strongyloides papillosus</i> Hunt et al. (2016); possum roundworm <i>Parastrongyloides trichosuri</i> Hunt et al. (2016); free-living nematode <i>Rhabditophanes</i> sp. Hunt et al. (2016); rat hookworm <i>Nippostrongylus brasiliensis</i> Sotillo et al. (2014); entomoparasitic nematode <i>Steinernema carpocapsae</i> Jing et al. (2010); barber's pole worm <i>Haemonchus contortus</i> Stepek et al. (2010); Stepek et al. (2011); roundworm <i>Brugia malayi</i> Stepek et al. (2010); Stepek et al. (2011); roundworm <i>Angiostrongylus cantonensis</i> ; roundworms <i>Pristionchus pacificus</i> , <i>Meloidogyne hapla</i> , and <i>Meloidogyne incognita</i> Park et al. (2010); rodent roundworm <i>Heligmosomoides polygyrus</i> Hewitson et al. (2011). See also Martin-Galiano and Sotillo (2022).
Phylum Rotifera
(Bdelloid) rotifers <i>Brachionus plicatilis</i> GB RNA28629, GB RNA07860, GB RNA15157; <i>Brachionus ibericus</i> GB ADR79275; and <i>Brachionus calyciflorus</i> GB CAF0958500, GB CAF0895916, GB CAF0832385; <i>Adineta ricciae</i> GB CAF0946005, GB CAF0977041, GB CAF0768045; <i>Adineta vaga</i> GB UJR10667, GB UJR30378, GB UJR12793; <i>Adineta steineri</i> GB CAF1145707, GB CAF0769233, GB CAF1137664; <i>Rotaria</i> sp. Silwood-1/2 GB CAF3391332, GB CAF2377320, GB CAF4963388; <i>Rotaria sordida</i> GB CAF0972930, GB CAF1460355, GB CAF1002947; <i>Rotaria magnacalcarata</i> GB CAF1378535, GB CAF4188303, GB CAF5207147; <i>Rotaria socialis</i> GB CAF3283866, GB CAF3682864, GB CAF4691708; <i>Didymodactylo carnosus</i> GB CAF1098425, GB CAF0882139, GB CAF1230341

(Continued on following page)

TABLE 1 (Continued) Representative sequences of studied and potential non-vertebrate metazoan astacins.

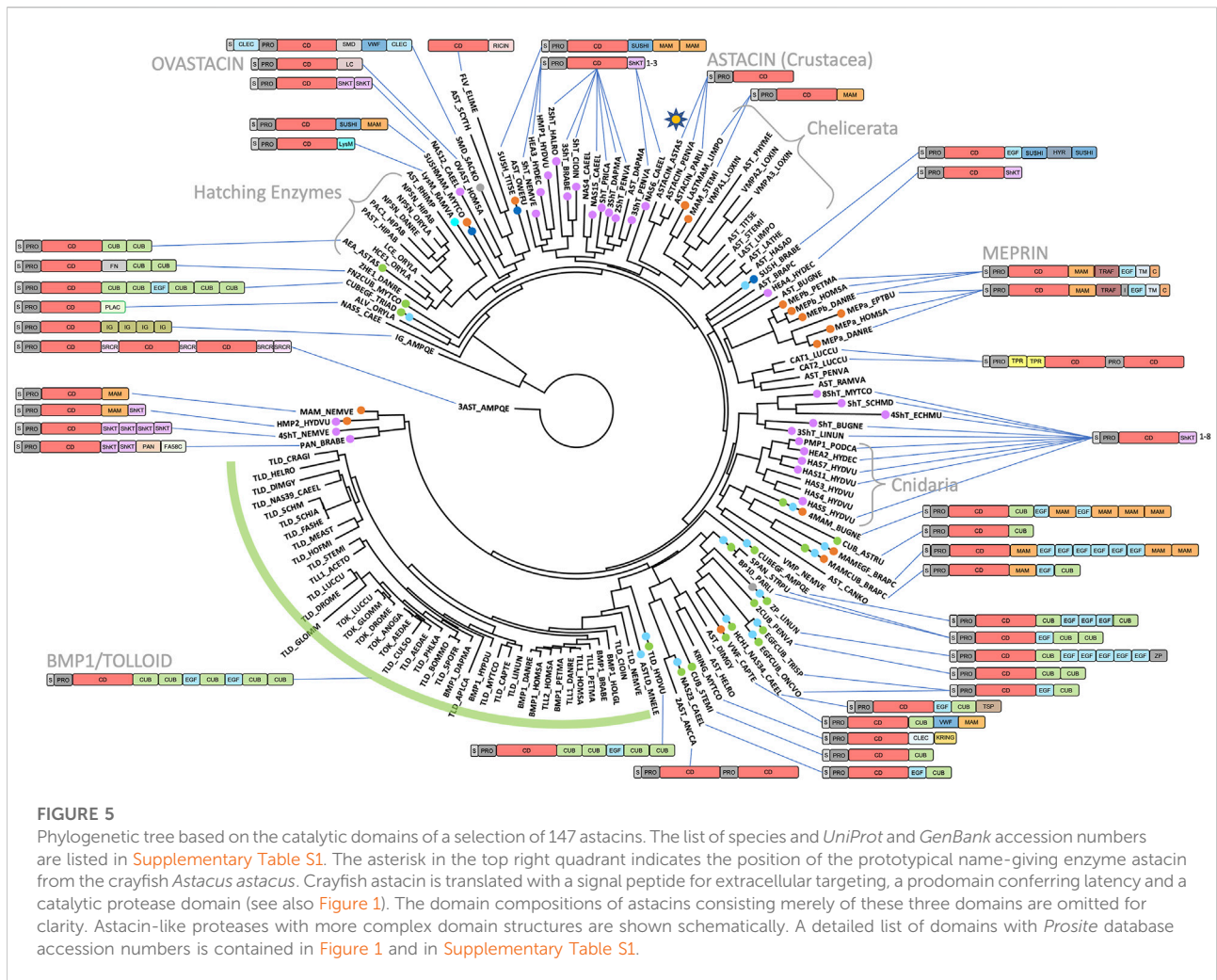
<i>Phylum Annelida</i>
Tubeworm <i>Owenia fusiformis</i> GB CAC9583006, GB CAH1790161, GB CAH1785993; meiofaunal worm <i>Dimorphilus gyrociliatus</i> GB CAD5123995, GB CAD5123793, GB CAD5125857; leech <i>Helobdella robusta</i> GB XP_009017824, GB XP_009026841, GB XP_009025904; segmented worm <i>Capitella teleta</i> GB ELT87968, GB ELU00391, GB ELU00287; Satsuma tubeworm <i>Lamellibrachia satsuma</i> GB KAI0241173, GB KAI0224812, GB KAI0229203
<i>Phylum Bryozoa</i>
Common bugula <i>Bugula neritina</i> GB KAF6036083, GB KAF6021779, GB KAF6027899
<i>Phylum Brachiopoda</i>
Common oriental lamp shell <i>Lingula anatina</i> GB XP_013381380, GB XP_013399408, GB XP_013417809; lamp shell <i>Lingula unguis</i> UP A0A1S3HHT9, UP A0A1S3KFN6, UP A0A1S3IEM6
<i>Phylum Mollusca</i>
Sea hare <i>Aplysia californica</i> Liu et al. (1997); spear squid <i>Loligo bleekeri</i> Yokozawa et al. (2002); Akoya pearl oyster <i>Pinctada fucata</i> Xiong et al. (2006); Pacific oyster <i>Crassostrea gigas</i> (Roberts et al., 2009); Suminoe oyster <i>Crassostrea ariakensis</i> (Yang and Wu, (2009); myosinase from the Japanese flying squid <i>Todarodes pacificus</i> Yokozawa et al. (2002); vampire snail <i>Cumia reticulata</i> Gerdol et al. (2019); golden cuttlefish <i>Sepia esculenta</i> Kumar et al. (2018); bigfin reef squid <i>Sepioteuthis lessoniana</i> Kanzawa et al. (2008); spear squid <i>Loligo bleekeri</i> Kanzawa et al. (2005); Mediterranean mussel <i>Mytilus galloprovincialis</i> Yu et al. (2016); <i>Mytilus coruscus</i> , Korean mussel UP A0A6J8APQ9, UP A0A6J8DI28, UP A0A6J8BR94, UP A0A6J8CIB2, UP A0A6J7ZT88, UP A0A060Q6V2
<i>Phylum Platyhelminthes</i>
Triclad flatworms <i>Schmidtea mediterranea</i> Isolani et al. (2013), UP A0A4Z2DK01 and <i>Dugesia japonica</i> Isolani et al. (2013); blood flukes <i>Schistosoma mansoni</i> and <i>Schistosoma japonicum</i> Park et al. (2010); liver fluke <i>Fasciola hepatica</i> UP A0A4E0RFD2; tapeworm <i>Echinococcus multilocularis</i> UP A0A068Y1M8. See also Martin-Galiano and Sotillo (2022).
<i>Phylum Xenacoelomorpha</i>
Parasitic aquatic worm <i>Meara stichopi</i> GB AVK72361; panter worm <i>Hofstenia miamia</i> GB AID23683
<i>Phylum Cnidaria</i>
Fresh-water polyp <i>Hydra vulgaris</i> Sarras et al. (2002); hydroid <i>Hydractinia echinata</i> Möhrle et al. (2006); jellyfish <i>Podocoryna carnea</i> Pan et al. (1998); carp endoparasite <i>Sphaerospora molnari</i> Hartigan et al. (2020); blade fire coral <i>Millepora complanata</i> Hernández-Elizárraga et al. (2019); lion's mane jellyfish <i>Cyanea capillata</i> Liu et al. (2015); Nomura's jellyfish <i>Nemopilema nomurai</i> Kang et al. (2014); sea anemone <i>Nematostella vectensis</i> Möhrle et al. (2006)
<i>Phylum Placozoa</i>
Flat-bodied animal <i>Trichoplax adhaerens</i> GB XP_002110676, GB XP_002110677, GB XP_002113592 and <i>Trichoplax</i> sp. H2 GB RDD40966, GB RDD38391, GB RDD44281
<i>Phylum Porifera</i>
Sponge <i>Amphimedon queenslandica</i> GB XP_019850969, GB XP_019859271, GB XP_019850968
<i>Phylum Ctenophora</i>
Sea walnut <i>Mnemiopsis leidy</i> GB AEP16401
<i>Bacterial astacins^a</i>
<i>Flavobacterium meningosepticum</i> flavastacin [UP Q47899; Tarentino et al. (1995)]; myroilysin from <i>Myroides profundus</i> (UP B5B0E6) and <i>Myroides</i> sp. CSLB8 (UP A0A0P0DZ84) Xu et al. (2017), Ran et al. (2020)

GB, GenBank code (www.ncbi.nlm.nih.gov/genbank); UP, UniProt code (www.uniprot.org). Phyla are consistent with Figure 3.

^aDespite not being from metazoans, these proteins have been listed as they are the only biochemically and structurally characterized astacins outside Holozoa. They appear in Figures 1, 5.

[formerly *Proterospongia* sp.; (Fairclough et al., 2013)] and *Monosiga ovata* of this phylum did not reveal significantly similar proteins. In addition, organisms from the next related taxa, Filasterea (*Capsospora owczarzaki*), Ichthyosporea (*Sphaeroforma arctica*) and Corallochytrium/Pluriformea (*Corallochytrium limacisporum* and *Syssomonas multiformis*), which together with Metazoa and Choanoflagellata constitute the Holozoa (Figure 3), did not contain potential orthologs either in the generally available databases. At this point, we got access to

unpublished genome sequences from a series of organisms at the root of Holozoa, viz., the Choanoflagellata *Acanthoeca spectabilis*, *Choanoeca perplexa*, *Diaphanoeca grandis*, *Salpingoeca dolichothecata*, *Salpingoeca infusorum*, *Salpingoeca kvevrii*, *Salpingoeca macrocollata*, *Salpingoeca punica*, *Salpingoeca urceolata* and *Stephanoeca diplocostata*; the Filastereum *Ministeria vibrans*; and the Ichthyosporea *Amoebidium parasiticum*, *Abeoforma whisleri* and *Pirum gemmata* [I. Ruiz-Trillo & M. Leger, personal communication,



and (Torruella et al., 2015; Carr et al., 2017; Grau-Bové et al., 2017)]. We identified several potential astacin orthologues in *Diaphanoeca grandis* (e.g., comp29462_c0_seq1:1552770, comp29716_c0_seq2:1462332 and comp33115_c0_seq4: 851467), *Salpingoeca dolichothecata* (e.g., comp24492_c1_seq1:7011852, comp14956_c0_seq2:1981400 and comp26781_c0_seq6:13261), *Amoebidium parasiticum* (*Amoebidium parasiticum*_Apar_comp21710_c1_seq1m30654/1239) and *Ministeria vibrans* (*Ministeria vibrans*_Mvib_g2222t1/16202), whose sequences are provided in Supplementary Figure S1. Altogether, these findings would suggest that astacins probably originate in the root of Holozoa according to Darwinian descent.

A selection of invertebrate and vertebrate sequences (Supplementary Table S1) enabled us to construct a phylogenetic tree for metazoan astacins (Figure 5), which was solely based on a sequence alignment of the respective CDs. Three sequences encompassed multiple CDs, which originate in the most basal sponge *Amphimedon queenslandica* (UniProt code

[UP] A0A1X7U9V1), the nematode hookworm *Ancylostoma caninum* (UP A0A368GTC8) and the blow fly *Lucilia cuprina* (UP A0A0L0BYD0). This tree does not mirror the phylogenesis of organisms presented in Figure 3 since it contains both orthologous and paralogous astacins. Indeed, members of the distinct subfamilies evinced separate clustering (Figure 5). The tree is not comprehensive, since only a selection of astacins comprising the minimal setup of typical sequential and structural motifs (as outlined above) were included. Nevertheless, certain clusters of astacins can be recognized. These are, starting clockwise in the upper left quadrant of the circular tree (Figure 5), the 1) hatching enzymes, which originate from the same root as ovastacin; 2) ShKT-carrying astacins from chordates, nematodes, cnidarians, priapulids and arthropods (chelicerates and crustaceans), which include the prototypal crayfish astacin; 3) a clade containing the meprins; 4) a second cluster of ShKT-astacins, mostly from cnidarians (right centre); 5) proteins rich in MAM, EGF and CUB domains (right bottom); and 6) the BTPs (left bottom).

2.3 Scattered presence of astacins outside Holozoa

Searches in non-holozoan Eukaryota including plants, fungi and the fungus-like Oomycota revealed merely < 400 sequences scattered across Alveolata, Stramenopila, Rhizaria, Archaeplastida (Haptista), Excavata (Discoba) and some Amoebozoa clades. By contrast, no astacins were detected in any of the other eukaryotic taxa except for the silver mallet wood *Rhodamnia argentea* [GenBank code (GB) XP_030553468], the only plant orthologue retrieved. However, this entry was debunked as a contamination with a tolloid-like chelicerate astacin from the wheat curl mite *Aceria tosichella* (UP A0A8B8R4B3).

The observation of astacin-like proteins within Stramenopila is remarkable since this taxon alone already accounts for ~200 of the hits. These are heterokonts that were formerly grouped into fungi (Holomycota) but currently are considered to be closer to brown algae than to fungi (Sebé-Pedrós et al., 2017). They further include the Oomycetes (“egg fungi”), a clade containing many parasitic/saprophytic organisms, which in turn encompass most of the hits within Stramenopila. Among them is *Aphanomyces astaci*, a well-known parasite of the North American crayfish *Cambarus clarkii*, which developed resistance against this pest. However, when American crayfish were brought to Europe in the late 19th century, *A. astaci* infection caused the “crayfish plague” in the endogenous crayfish population (*A. astacus*), which almost caused its extinction (Diéguez-Urbeondo and Söderhäll, 1993). Oomycetes are known champions of horizontal gene transfer (HGT) (Koonin et al., 2001; Keeling and Palmer, 2008), thereby gathering enzymes useful to target their prey (Judelson, 2017). This genetic transfer route could thus explain the generally scattered but locally focused presence of astacin CDs, which is inconsistent with Darwinian descent, in eukaryotes outside Holozoa.

HGT could also account for the sporadic occurrence of astacin-like peptidases in archaea, bacteria and viruses, which would thus also be xenologues (Koonin et al., 2001). Examples of archaeal sequences were found in *Candidatus korarcheota* (UP A0A662SFB1), *Nitrosopumilus* sp. (GB MCA9827382), *Methanotrichaceae* archaeon (GB MBN1323470) and *Halobacteriales* archaeon QH_6_64_20 (GB PSP40402). Viral sequences were restricted to Lutzomyia reovirus 2 (UP A0A0H4M9A8). The more populous bacterial examples were from *Bacillus cereus* (GB WP_235610182), *Acinetobacter baumannii* (GB WP_207273295), *Klebsiella pneumoniae* (GB NAU77905), *Bacillus thuringiensis* (GB WP_228528809), *Legionella pneumophila* (GB WP_061484376), *Bacillus mycoides* (GB WP_186320991), among others. Overall, the vast majority of bacterial astacin hosts live in intimate contact with animals, which would facilitate HGT of genes from eukaryotes to prokaryotes. Among them are those of the biochemically

studied proteins flavastacin from *Flavobacterium meningosepticum* [also known as *Elizabethkingia meningoseptica*; UP Q47899; (Tarentino et al., 1995)] and myroilysins from *Myroides profundus* (UP B5B0E6) and *Myroides* sp. CSLB8 (UP A0A0P0DZ84) (Xu et al., 2017; Ran et al., 2020). In the latter case, zymogenic latency was shown to follow a different mechanism from the animal forms [see Section 2.1; (Guevara et al., 2022)], which would further support an HGT event as the origin of its presence in the bacterium. This is reminiscent of the aforementioned fragilysin-3, which originates in a member of the human colon microbiota. Its CD was proposed to be an adamalysin/ADAM xenologue acquired by HGT from the host that separately evolved to derive a distinct mechanism of latency (Goulas et al., 2011; Goulas et al., 2013).

2.4 Functional and evolutionary aspects of BMP1/tolloid-like peptidases

In the phylum Chordata, subphylum Vertebrata, order Mammalia, six genes encode astacin peptidases [*bmp1*, *tll1*, *tll2*, *mepa*, *mepb* and *astl*; <http://degradome.uniovi.es/met.html>; (Quesada et al., 2009)]. The first three genes code for the BMP1, TLL1 and TLL2 proteins, which belong to the BTPs (Wozney et al., 1988; Nguyen et al., 1994). They are characterized by an arrangement of five CUB domains and two EGF-like domains C-terminal of the CD and are present in virtually all metazoan phyla (Figures 1, 5). BTPs are important for dorsoventral patterning in metazoan embryogenesis (Sato and Sargent, 1990; Holley et al., 1996; de Robertis, 2008; de Robertis and Tejada-Muñoz, 2022). In deuterostomes, they are also crucial for extracellular matrix assembly through the processing of precursors of matrix components, growth factors and their receptors (Kessler et al., 1996; Ge and Greenspan, 2006). The CUB and EGF-like domains of these astacins bear important exosites for substrate recognition and targeting toward the extracellular matrix (Sieron et al., 2000; Garrigue-Antar et al., 2004; Ge and Greenspan, 2006; Hintze et al., 2006; Hopkins et al., 2007; Wermter et al., 2007). Although BTPs are closely linked with bilateral morphogenesis, which supposedly originated in “Urbilateria” (de Robertis, 2008; de Robertis and Tejada-Muñoz, 2022), they are not restricted to Bilateria and are also present in radially symmetric metazoans, Cnidaria (Saina et al., 2009) and Ctenophora (Pang et al., 2011). Related proteins with slightly deviating CUB/EGF arrangements were also found in the basal placozoan species *Trichoplax adhaerens* (CUBEGF-TRIAD) and in several other astacins (Figure 5). Taken together, BTPs are pan-metazoan astacins with essential functions in embryogenesis, which exert further additional functions in several metazoan phyla.

2.5 Functional and evolutionary aspects of meprin metallopeptidases

Two other human astacin genes, *mepa* and *mepb*, encode meprin α and meprin β for which orthologs have only been detected among vertebrates. Both meprins are membrane bound but meprin α is released already in the trans-Golgi network by furin cleavage and stays membrane bound only in association with meprin β . The latter is a “shedase”, which releases cell-surface proteins such as growth factors, cytokines, receptors, as well as amyloid precursor protein through cleavage at its β -secretase site. Deregulation of meprins leads to neurodegenerative diseases, changes in barrier function (such as in the blood brain barrier), inflammatory bowel disease, fibrosis, nephritis and cancer (Sterchi et al., 2008; Arolas et al., 2012; Becker-Pauly and Pietrzik, 2016; Arnold et al., 2017; Eckhard et al., 2021; Gindorf et al., 2021; Bayly-Jones et al., 2022; Werny et al., 2022). The unique domain composition of meprins includes MAM, TRAF and EGF-like domains (Figures 1, 5), and although MAM- and EGF-containing astacins have been identified in other metazoan phyla (see Figures 1, 5; Supplementary Table S1), none of these are apparently membrane bound or exhibit comparable physiological potential to vertebrate meprins. Finally, our database searches unravelled meprin-like astacin-CDs also in basal vertebrates such as lamprey (MEP β _PETMA) and hagfish (MEP α _EPTBU), which just encompass the S, PRO, CD and MAM moieties (see Figures 1, 5; Supplementary Table S1), similarly to a reported horseshoe crab enzyme [LASTMAM_LIMPO; (Becker-Pauly et al., 2009; Guevara et al., 2022)].

2.6 Astacins in animal reproduction

The sixth human astacin is ovastacin, which is encoded by the *astl* gene and is expressed only in oocytes among mammals (Burkart et al., 2012). Absence of ovastacin results in subfertility, since it is released during the cortical reaction after intrusion of a sperm cell and causes hardening of the zona pellucida of the extracellular matrix surrounding the egg. This provides rigidity and robustness to the resulting embryo until its implantation in the uterus (Stöcker et al., 2014; Körschgen et al., 2017). Ovastacin has the basic domain composition S-PRO-CD, which is followed by a disordered domain of unknown function. This domain stays connected with the oolemma after the release of the enzyme into the perivitelline space during the cortical reaction, which suggests a function in membrane anchoring and shedding of ovastacin (Körschgen et al., 2017). A similar function to ovastacin was reported for alveolin from the medaka fish *Oryzias latipes*, which likewise hardens the envelope of the fertilized egg (zygote) in bony fishes (Shibata et al., 2000).

In egg-laying vertebrates like fishes, amphibians, reptiles and birds, a specialized group of astacin MPs termed hatching

enzymes has evolved (Nagasawa et al., 2022). They are absent from mammals and involved in the cleavage of the eggshell. They optionally contain an additional pair of cysteine residues in the N-terminal subdomain of their mature CDs compared to crayfish astacin, as well as extra C-terminal CUB domains (Figures 1, 5). Hatching enzymes are also present in egg-laying invertebrates, such as the crayfish *A. astacus*, which in addition to the prototypic digestive astacin archetype possesses the “*Astacus* embryonic astacin” (AEA_ASTAS in Figure 5; Supplementary Table S1; Geier and Zwilling, 1998). Finally, a reproductive astacin was also described from *Drosophila* seminal plasma (SEMP1_DROME; Figure 5; Supplementary Table S1). This MP is involved in a proteolytic cascade that triggers sperm capacitation and thus regulates fertility in the fruitfly (LaFlamme et al., 2012; LaFlamme and Wolfner, 2013; LaFlamme et al., 2014).

2.7 Astacins of helminths and cnidarians

The genome analysis of the roundworm *Caenorhabditis elegans* uncovered 40 astacins termed “nematode astacins” (Möhrlen et al., 2003; Park et al., 2010). Moreover, in a comprehensive analysis of 154 helminth species of the phyla Nematoda and Platyhelminthes, many of them parasitic, an enormous radiation of astacins was also observed (Martín-Galiano and Sotillo, 2022). Most remarkable are the > 100 different additional domains that occur downstream of the CD in variable combinations, thus yielding an enormous functional versatility for these proteins. These domains do not only include protein-protein and protein-carbohydrate interacting domains, but also additional enzymatic functions, such as trypsin-like serine peptidases and hydroxylases.

Particularly striking are astacins containing “ShKT” domains, which mimic a toxin from the sea anemone *Styrodactyla helianthus* that blocks potassium channels (Castañeda et al., 1995). We found such ShKT-astacins in Nematoda (CAEEL, ONCVO, ANCCA and TRISP; see Figure 5; Table 1; Supplementary Table S1), Platyhelminthes (SCHMD, ECHMU and SCHJA), Arthropoda (PENVA and DAPMA), Priapulida (PRICA), Bryozoa (BUGNE and LINUN), Mollusca (MYTCO) and Cnidaria (HYDVU, HYDEC, PODCA, NEMVE), as well as in the lower chordates sea squirt (HALRO and CIOIN) and lancelet (BRABE). In Figure 5, ShKT-astacins are labelled with pink branch tips (see Supplementary Table S1). These astacins are mostly expressed in epithelia forming barriers to the environment or in the digestive tract, which suggests functions in protection and defence, as well as preservation of the epithelial integrity (Park et al., 2010; Isolani et al., 2018). However, considering the parasitic lifestyle of many of the organisms harbouring ShKT-astacins, the combination of proteolytic and toxin domains may also challenge the respective host (Moran et al., 2013;

Martín-Galiano and Sotillo, 2022). Similarly, toxicity has also been reported for astacins lacking ShKT domains from spider venoms (Trevisan-Silva et al., 2010), in which other proteins may take over the role of the latter domains in linking proteolytic activity with specific toxicity. This is reminiscent of the snake venom MPs from the adamalysin/ADAM family, for which forms spanning only the CD are not haemorrhagic while those encompassing further C-terminal disintegrin-like and cysteine-rich domains may be haemorrhagic (Fox and Serrano, 2009; Gomis-Rüth et al., 2013; Herrera et al., 2018). Finally, many ShKT-astacins carry additional CUB, EGF, MAM, etc. domains that may serve to modulate activity. Examples are the morphogenetically active *Hydra vulgaris* ShKT-proteins HMP1, HMP2 and HAS7 (Figure 5; Supplementary Table S1). While HMP1 is involved in head formation and head regeneration (Yan et al., 1995), HMP2 has a function in foot morphogenesis (Yan et al., 2000) and HAS7 specifically cleaves *Hydra* WNT3 during head morphogenesis and thereby restricts head organizer formation (Ziegler et al., 2021; Holstein, 2022).

2.8 New functions of astacins

A surprising function for an astacin peptidase was very recently reported for the sea star *Asterias rubens*, which underpins the enormous versatility of nature's toolbox. In zoology textbooks [e.g., (Hickman et al., 2020)], the locomotion of Asterozoa on a surface is usually explained as the concerted action of a multitude of tiny sucking pods in the podia of the tube feet lining the oral face of the animal's arms. However, pod attachment is apparently not based on suction but on a secreted glue consisting of adhesive matrix proteins that are left on the surface after detachment. The latter is mediated by an astacin MP spanning a CD and a CUB domain (CUB_ASTRU; see Figure 5; Supplementary Table S1), which is specifically secreted by de-adhesive gland cells and releases the adhesive material from the surface of the tube feet (Algrain et al., 2022).

3 Materials and methods

3.1 Database searches

Sequence searches were performed with the *Psi-Blast* program (Altschul and Koonin, 1998) within *UniProt* (<https://www.uniprot.org>) and *GenBank* (<https://www.ncbi.nlm.nih.gov/genbank>) using standard parameters, as well as with the *Ensembl* genome browser (<http://www.ensembl.org>), within the *Choanobase* database (King, 2005) and the *Merops* database (<https://www.ebi.ac.uk/merops>; (Rawlings and Bateman,

2021)]. Literature searches for described astacins were performed with the keyword "astacin*" within *PubMed* (<https://pubmed.ncbi.nlm.nih.gov>), which retrieved 342 papers. Hits were manually curated.

3.2 Compilation of a dendrogram for holozoans

The currently available trees for holozoans in zoology textbooks such as (Hickman et al., 2020) present several discrepancies with models proposed by recent research publications (Cannon et al., 2016; Giribet et al., 2019; Laumer et al., 2019). These are based on massive sequence data generated in the last decade by increasingly affordable sequencing methods, first through the Illumina MiSeq platform and more recently through MinION sequencing (Santos et al., 2020), which are also partially applicable to natural history collections (Folk et al., 2021). As an example, phylum Chaetognatha was considered a separate clade at the same level as Spiralia and Ecdysozoa, while it is currently envisaged as a sister clade of Gnathifera within Spiralia (Kocot et al., 2017). Also setting phylum Porifera at the root of Metazoa contradicts recent models, which choose Ctenophora for this place (Giribet et al., 2019; Laumer et al., 2019). Finally, recent work proposed new models at the base of Holozoa, with Ichtyosporea and Corallochytra/Pluriformea forming Teretosporea, which together with Filozoa give rise to Holozoa (Torruella et al., 2015; Arroyo et al., 2020). Accordingly, we assembled a dendrogram for holozoans based on consensus information extracted from these and other recent publications (Ryan et al., 2010; Ruggiero et al., 2015; Torruella et al., 2015; Cannon et al., 2016; Kocot et al., 2017; Lu et al., 2017; Sebé-Pedrós et al., 2017; Whelan et al., 2017; Adl et al., 2019; Giribet et al., 2019; Laumer et al., 2019; Marlétaz et al., 2019; Sogabe et al., 2019; Schoch et al., 2020; Schulze and Kawachi, 2021).

3.3 Computation of alignments and phylogenetic trees

Amino-acid sequence alignments and phylogenetic trees were computed using the *Seaview* program (<http://doua.prabi.fr/software/seaview>) (Galtier et al., 1996; Gouy et al., 2010). Alignments were performed with *Clustal Omega* within *Seaview* (Sievers et al., 2011) using default parameters. Manual adjustment of the S₁' regions in Figure 4 was performed based on overlays of the X-ray crystal structures of crayfish astacin [Protein Data Bank (PDB) access codes 1AST, 1QJI] and zebrafish hatching enzyme 1 (PDB 3LQB). Phylogenetic trees were calculated with *PhyML* using the maximum likelihood approach as implemented in *Seaview* (Guindon et al., 2010). The BLOSUM62 scoring matrix was used and 100 bootstrap

replications were computed in each case. Trees were represented with *Figtree* (<http://tree.bio.ed.ac.uk>).

3.4 Computation of three-dimensional structural models

Precise and accurate computational models of specific astacin domains were obtained with the *AlphaFold* program (Jumper et al., 2021; Tunyasuvunakool et al., 2021). To this aim, the program was locally installed in a high-performance computing cluster operated with linux and the amino acid sequences were fed into the program, which was run employing standard parameters. Quality of the predicted models was monitored through the average predicted local distance difference test (pLDD1) value. Values exceeding 90% are considered to originate in high-accuracy models, while those above 80% correspond to generally correct models for the backbone (Tunyasuvunakool et al., 2021). The unique EGF-like domain of human BMP1 (residues 547–590, see UP P13497) was predicted as a generally correct model for the backbone given an average pLDD1 value of 82.5%. Moreover, the first CUB domain of human BMP1 (residues 321–434, see UP P13497) was predicted highly accurately (average pLDD1 = 91.5%). Three-dimensional structure figures were prepared using *Chimera* (Goddard et al., 2018).

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors upon reasonable request.

Author contributions

FXG-R and WS conceived and supervised the project, performed calculations and sequence searches, and wrote the manuscript.

References

- Adl, S. M., Bass, D., Lane, C. E., Lukes, J., Schoch, C. L., Smirnov, A., et al. (2019). Revisions to the classification, nomenclature, and diversity of eukaryotes. *J. Eukaryot. Microbiol.* 66 (1), 4–119. doi:10.1111/jeu.12691
- Algrain, M., Hennebert, E., Bertemes, P., Wattiez, R., Flammang, P., and Lengerer, B. (2022). In the footsteps of sea stars: Deciphering the catalogue of proteins involved in underwater temporary adhesion. *Open Biol.* 12 (8), 220103. doi:10.1098/rsob.220103
- Altschul, S. F., and Koonin, E. V. (1998). Iterated profile searches with PSI-BLAST—a tool for discovery in protein databases. *Trends biochem. Sci.* 23 (11), 444–447. doi:10.1016/s0968-0004(98)01298-5
- AmbuAli, A., Monaghan, S. J., McLean, K., Inglis, N. F., Bekaert, M., Wehner, S., et al. (2020). Identification of proteins from the secretory/excretory products (SEPs) of the branchiuran ectoparasite *Argulus foliaceus* (Linnaeus, 1758) reveals unique

Funding

This study was supported in part by grants from Spanish and Catalan public and private bodies (grant/fellowship references PID2019-107725RG-I00 and PDC2022-133344-I00 by MICIN/AEI/10.13039/501100011033, 2017SGR3 and Fundació La Marató de TV3 201815).

Acknowledgments

Special thanks to I. Ruiz-Trillo and M. Leger from the Institute of Evolutionary Biology (CSIC-UPF) for their assistance in identifying astacin sequences at the root of Holozoa.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.1080836/full#supplementary-material>

secreted proteins amongst haematophagous ecdysozoa. *Parasit. Vectors* 13 (1), 88. doi:10.1186/s13071-020-3964-z

Aricescu, A. R., Hon, W. C., Siebold, C., Lu, W., van der Merwe, P. A., and Jones, E. Y. (2006). Molecular analysis of receptor protein tyrosine phosphatase μ -mediated cell adhesion. *EMBO J.* 25 (4), 701–712. doi:10.1038/sj.emboj.7600974

Arnold, P., Boll, I., Rothaug, M., Schumacher, N., Schmidt, F., Wichert, R., et al. (2017). Meprin metalloproteases generate biologically active soluble interleukin-6 receptor to induce trans-signaling. *Sci. Rep.* 7, 44053. doi:10.1038/srep44053

Arolas, J. L., Broder, C., Jefferson, T., Guevara, T., Sterchi, E. E., Bode, W., et al. (2012). Structural basis for the sheddase function of human meprin β metalloproteinase at the plasma membrane. *Proc. Natl. Acad. Sci. U. S. A.* 109 (40), 16131–16136. doi:10.1073/pnas.1211076109

- Arolas, J. L., Goulas, T., Cuppari, A., and Gomis-Rüth, F. X. (2018). Multiple architectures and mechanisms of latency in metallopeptidase zymogens. *Chem. Rev.* 118 (11), 5581–5597. doi:10.1021/acs.chemrev.8b00030
- Arroyo, A. S., Iannes, R., Baptiste, E., and Ruiz-Trillo, I. (2020). Gene similarity networks unveil a potential novel unicellular group closely related to animals from the Tara Oceans Expedition. *Genome Biol. Evol.* 12 (9), 1664–1678. doi:10.1093/gbe/evaa117
- Barnard, A.-C., Nijhof, A. M., Gaspar, A. R. M., Neitz, A. W. H., Jongejan, F., and Maritz-Olivier, C. (2012). Expression profiling, gene silencing and transcriptional networking of metzincin metalloproteases in the cattle tick, *Rhipicephalus (Boophilus) microplus*. *Vet. Parasitol.* 186 (3–4), 403–414. doi:10.1016/j.vetpar.2011.11.026
- Baska, P., Wisniewski, M., Krzyzowska, M., Dlugosz, E., Zygnier, W., Gorski, P., et al. (2013). Molecular cloning and characterisation of *in vitro* immune response against astacin-like metalloprotease Ace-MTP-2 from. *Exp. Parasitol.* 133 (4), 472–482. doi:10.1016/j.exppara.2013.01.006
- Bayly-Jones, C., Lupton, C. J., Fritz, C., Venugopal, H., Ramsbeck, D., Wermann, M., et al. (2022). Helical ultrastructure of the metalloprotease meprin α in complex with a small molecule inhibitor. *Nat. Commun.* 13 (1), 6178. doi:10.1038/s41467-022-33893-7
- Becker-Pauly, C., Bruns, B. C., Damm, O., Schütte, A., Hammouti, K., Burmester, T., et al. (2009). News from an ancient world: Two novel astacin metalloproteases from the horseshoe crab. *J. Mol. Biol.* 385 (1), 236–248. doi:10.1016/j.jmb.2008.10.062
- Becker-Pauly, C., and Pietrzik, C. U. (2016). The metalloprotease meprin β is an alternative β -secretase of APP. *Front. Mol. Neurosci.* 9, 159. doi:10.3389/fnmol.2016.00159
- Beckmann, G., and Bork, P. (1993). An adhesive domain detected in functionally diverse receptors. *Trends biochem. Sci.* 18 (2), 40–41. doi:10.1016/0968-0004(93)90049-s
- Berman, J. J. (2019). “Chapter 5 – phylogeny: Eukaryotes to chordates,” in *Evolution’s clinical guidebook – translating ancient genes into precision medicine*. Editor J. J. Berman (London (UK): Academic Press Elsevier). doi:10.1016/B978-0-12-817126-4.00005-9
- Bode, W., Gomis-Rüth, F. X., Huber, R., Zwilling, R., and Stöcker, W. (1992). Structure of astacin and implications for activation of astacins and zinc-ligation of collagenases. *Nature* 358, 164–167. doi:10.1038/358164a0
- Bode, W., Gomis-Rüth, F. X., and Stöcker, W. (1993). Astacins, serralytins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the ‘metzincins’. *FEBS Lett.*, 331, 134–140. doi:10.1016/0014-5793(93)80312-i
- Bode, W., and Huber, R. (1986). “Crystal structure of pancreatic serine endopeptidases,” in *Molecular and cellular basis of digestion*. Editors P. Desnuelle and H. Sjöström&O. Norén (Amsterdam: Elsevier).
- Bond, J. S., and Beynon, R. J. (1995). The astacin family of metalloendopeptidases. *Protein Sci.* 4, 1247–1261. doi:10.1002/pro.5560040701
- Bond, J. S. (2019). Proteases: History, discovery, and roles in health and disease. *J. Biol. Chem.* 294 (5), 1643–1651. doi:10.1074/jbc.TM118.004156
- Borchert, N., Becker-Pauly, C., Wagner, A., Fischer, P., Stöcker, W., and Brattig, N. W. (2007). Identification and characterization of onchoastacin, an astacin-like metalloproteinase from the filaria *Onchocerca volvulus*. *Microbes Infect.* 9 (4), 498–506. doi:10.1016/j.micinf.2007.01.007
- Borges, M. H., Figueiredo, S. G., Leprevost, F. V., de Lima, M. E., Cordeiro, M. do N., Diniz, M. R., et al. (2016). Venomous extract protein profile of Brazilian tarantula *Grammostola iheringi*: Searching for potential biotechnological applications. *J. Proteomics* 136, 35–47. doi:10.1016/j.jprot.2016.01.013
- Bork, P., and Beckmann, G. (1993). The CUB domain. A widespread module in developmentally regulated proteins. *J. Mol. Biol.* 231 (2), 539–545. doi:10.1006/jmbi.1993.1305
- Bork, P., Downing, A. K., Kieffer, B., and Campbell, I. D. (1996). Structure and distribution of modules in extracellular proteins. *Q. Rev. Biophys.* 29 (2), 119–167. doi:10.1017/s0033583500005783
- Burkart, A. D., Xiong, B., Baibakov, B., Jiménez-Movilla, M., and Dean, J. (2012). Ovastacin, a cortical granule protease, cleaves ZP2 in the zona pellucida to prevent polyspermy. *J. Cell Biol.* 197 (1), 37–44. doi:10.1083/jcb.201112094
- Cannon, J. T., Vellutini, B. C., Smith, J., 3rd, Ronquist, F., Jondelius, U., and Hejnol, A. (2016). Xenacoelomorpha is the sister group to Nephrozoa. *Nature* 530 (7588), 89–93. doi:10.1038/nature16520
- Carr, M., Richter, D. J., Fozouni, P., Smith, T. J., Jeuck, A., Leadbeater, B. S. C., et al. (2017). A six-gene phylogeny provides new insights into choanoflagellate evolution. *Mol. Phylogenet. Evol.* 107, 166–178. doi:10.1016/j.ympev.2016.10.011
- Castañeda, O., Sotolongo, V., Amor, A. M., Stöcklin, R., Anderson, A. J., Harvey, A. L., et al. (1995). Characterization of a potassium channel toxin from the Caribbean Sea anemone *Stichodactyla helianthus*. *Toxicon* 33 (5), 603–613. doi:10.1016/0041-0101(95)00013-c
- Cerdà-Costa, N., and Gomis-Rüth, F. X. (2014). Architecture and function of metallopeptidase catalytic domains. *Protein Sci.* 23 (2), 123–144. doi:10.1002/pro.2400
- Cho, H. K., Nam, B. H., Kong, H. J., Han, H. S., Hur, Y. B., Choi, T. J., et al. (2008). Identification of softness syndrome-associated candidate genes and DNA sequence variation in the sea squirt, *Halocynthia roretzi*. *Mar. Biotechnol.* 10 (4), 447–456. doi:10.1007/s10126-008-9084-y
- Cismasiu, V. B., Denes, S. A., Reilander, H., Michel, H., and Szedlaczek, S. E. (2004). The MAM (meprin/A5-protein/PTPmu) domain is a homophilic binding site promoting the lateral dimerization of receptor-like protein-tyrosine phosphatase mu. *J. Biol. Chem.* 279 (26), 26922–26931. doi:10.1074/jbc.M313115200
- da Silveira, R. B., Wille, A. C., Chaim, O. M., Appel, M. H., Silva, D. T., Franco, C. R., et al. (2007). Identification, cloning, expression and functional characterization of an astacin-like metalloprotease toxin from *Loxosceles intermedia* (Brown spider) venom. *Biochem. J.* 406 (2), 355–363. doi:10.1042/BJ20070363
- Davis, S. W., and Smith, W. C. (2002). Expression cloning in ascidians: Isolation of a novel member of the astacin protease family. *Dev. Genes Evol.* 212 (2), 81–86. doi:10.1007/s00427-002-0211-x
- de Maere, V., Vercauteren, I., Geldhof, P., Gevaert, K., Vercruyse, J., and Claerebout, E. (2005). Molecular analysis of astacin-like metalloproteases of *Ostertagia ostertagi*. *Parasitology* 130, 89–98. doi:10.1017/s0031182004006274
- de Robertis, E. M. (2008). Evo-devo: Variations on ancestral themes. *Cell* 132 (2), 185–195. doi:10.1016/j.cell.2008.01.003
- de Robertis, E. M., and Tejada-Muñoz, N. (2022). Evo-Devo of Urbilateria and its larval forms. *Dev. Biol.* 487, 10–20. doi:10.1016/j.ydbio.2022.04.003
- Diab, M. R., Ahmed, M. M., Hussein, E. H., and Mohammed, A. M. A. (2021). Isolation of the Astacin-like metalloprotease coding gene (astl) and assessment of its insecticidal activity towards *Spodoptera littoralis* and *Sitophilus oryzae*. *Jordan J. Biol. Sci.* 14 (4), 677–685. doi:10.54319/jjbs/140407
- Diéguez-Uribeondo, J., and Söderhäll, K. (1993). *Procambarus clarkii* Girard as a vector for the crayfish plague fungus, *Aphanomyces astaci* Schikora. *Aquac. Res.* 24 (6), 761–765. doi:10.1111/j.1365-2109.1993.tb00655.x
- Dumermuth, E., Sterchi, E. E., Jiang, W. P., Wolz, R. L., Bond, J. S., Flannery, A. V., et al. (1991). The astacin family of metalloendopeptidases. *J. Biol. Chem.* 266, 21381–21385. doi:10.1016/s0021-9258(18)54648-2
- Eckhard, U., Körschgen, H., von Wiegen, N., Stöcker, W., and Gomis-Rüth, F. X. (2021). The crystal structure of a 250-kDa heterotetrameric particle explains inhibition of sheddase meprin β by endogenous fetuin-B. *Proc. Natl. Acad. Sci. U. S. A.* 118 (14), e2023839118. doi:10.1073/pnas.2023839118
- Fabricius, J. C. (1796). *Nomenclator entomologicus enumerans insecta omnia in J. C. Fabricii Entomologia Systematica emendata et aucta 1792*. Manchester (UK): Impensis Nicholson.
- Fairclough, S. R., Chen, Z., Kramer, E., Zeng, Q., Young, S., Robertson, H. M., et al. (2013). Premetazoan genome evolution and the regulation of cell differentiation in the choanoflagellate *Salpingoeca rosetta*. *Genome Biol.* 14 (2), R15. doi:10.1186/gb-2013-14-2-r15
- Ferreira, A. H., Cristofolletti, P. T., Lorenzini, D. M., Guerra, L. O., Paiva, P. B., Briones, M. R., et al. (2007). Identification of midgut microvillar proteins from *Tenebrio molitor* and *Spodoptera frugiperda* by cDNA library screenings with antibodies. *J. Insect Physiol.* 53 (11), 1112–1124. doi:10.1016/j.jinsphys.2007.06.007
- Folk, R. A., Kates, H. R., LaFrance, R., Soltis, D. E., Soltis, P. S., and Guralnick, R. P. (2021). High-throughput methods for efficiently building massive phylogenies from natural history collections. *Appl. Plant Sci.* 9 (2), e11410. doi:10.1002/aps3.11410
- Foradori, M. J., Tillinghast, E. K., Smith, J. S., Townley, M. A., and Mooney, R. E. (2006). Astacin family metallopeptidases and serine peptidase inhibitors in spider digestive fluid. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 143 (3), 257–268. doi:10.1016/j.cbpb.2005.08.012
- Fox, J. W., and Serrano, S. M. (2009). Timeline of key events in snake venom metalloproteinase research. *J. Proteomics* 72 (2), 200–209. doi:10.1016/j.jprot.2009.01.015
- Freeman, R., Ikuta, T., Wu, M., Koyanagi, R., Kawashima, T., Tagawa, K., et al. (2012). Identical genomic organization of two hemichordate hox clusters. *Curr. Biol.* 22 (21), 2053–2058. doi:10.1016/j.cub.2012.08.052
- Fuzita, F. J., Pinkse, M. W. H., Patane, J. S. L., Juliano, M. A., Verhaert, P. D. E. M., and Lopes, A. R. (2015). Biochemical, transcriptomic and proteomic analyses of digestion in the scorpion *Tityus serrulatus*: Insights into function and evolution of

- digestion in an ancient arthropod. *PLoS One* 10 (4), e0123841. doi:10.1371/journal.pone.0123841
- Fuzita, F. J., Pinkse, M. W. H., Verhaert, P. D. E. M., and Lopes, A. R. (2015). Cysteine cathepsins as digestive enzymes in the spider *Nephilengys cruentata*. *Insect biochem. Mol. Biol.* 60, 47–58. doi:10.1016/j.ibmb.2015.03.005
- Galtier, N., Gouy, M., and Gautier, C. (1996). SEAVIEW and PHYLO_WIN: Two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.* 12 (6), 543–548. doi:10.1093/bioinformatics/12.6.543
- García-Ferrer, I., Arède, P., Gómez-Blanco, J., Luque, D., Duquerroy, S., Castón, J. R., et al. (2015). Structural and functional insights into *Escherichia coli* α_2 -macroglobulin endopeptidase snap-trap inhibition. *Proc. Natl. Acad. Sci. U. S. A.* 112, 8290–8295. doi:10.1073/pnas.1506538112
- Garrigue-Antar, L., Francois, V., and Kadler, K. E. (2004). Deletion of epidermal growth factor-like domains converts mammalian tolloid into a chordinase and effective procollagen C-proteinase. *J. Biol. Chem.* 279 (48), 49835–49841. doi:10.1074/jbc.M408134200
- Ge, G., and Greenspan, D. S. (2006). Developmental roles of the BMP1/TLD metalloproteinases. *Birth Defects Res. C Embryo Today* 78 (1), 47–68. doi:10.1002/bdrc.20060
- Geier, G., and Zwilling, R. (1998). Cloning and characterization of a cDNA coding for *Astacus embryonic astacin*, a member of the astacin family of metalloproteases from the crayfish *Astacus astacus*. *Eur. J. Biochem.* 253 (3), 796–803. doi:10.1046/j.1432-1327.1998.2530796.x
- Gerdol, M., Cervelli, M., Mariottini, P., Oliverio, M., Dutertre, S., and Modica, M. V. (2019). A recurrent motif: Diversity and evolution of ShKT domain containing proteins in the vampire snail *Cumia reticulata*. *Toxins (Basel)* 11 (2), 106. doi:10.3390/toxins11020106
- Gindorf, M., Storck, S. E., Ohler, A., Scharfenberg, F., Becker-Pauly, C., and Pietrzik, C. U. (2021). Meprin β : A novel regulator of blood-brain barrier integrity. *J. Cereb. Blood Flow. Metab.* 41 (1), 31–44. doi:10.1177/0271678X20905206
- Giribet, G., and Edgecombe, G. D. (2019). “Perspectives in animal phylogeny and evolution,” in *Perspectives on evolutionary and developmental biology essays for Alessandro Minelli*. Editor G. Fusco (Padova (I: Padova University Press).
- Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H., et al. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* 27 (1), 14–25. doi:10.1002/pro.3235
- Gómez Gallego, S., Loukas, A., Slade, R. W., Neva, F. A., Varatharajulu, R., Nutman, T. B., et al. (2005). Identification of an astacin-like metallo-proteinase transcript from the infective larvae of *Strongyloides stercoralis*. *Parasitol. Int.* 54 (2), 123–133. doi:10.1016/j.parint.2005.02.002
- Gomis-Rüth, F. X., Botelho, T. O., and Bode, W. (2012). A standard orientation for metallopeptidases. *Biochim. Biophys. Acta* 1824, 157–163. doi:10.1016/j.bbapap.2011.04.014
- Gomis-Rüth, F. X. (2009). Catalytic domain architecture of metzincin metalloproteases. *J. Biol. Chem.* 284, 15353–15357. doi:10.1074/jbc.R800069200
- Gomis-Rüth, F. X., Stöcker, W., Huber, R., Zwilling, R., and Bode, W. (1993). Refined 1.8 Å X-ray crystal structure of astacin, a zinc-endopeptidase from the crayfish *Astacus astacus* L. Structure determination, refinement, molecular structure and comparison with thermolysin. *J. Mol. Biol.* 229 (4), 945–968. doi:10.1006/jmbi.1993.1098
- Gomis-Rüth, F. X. (2003). Structural aspects of the metzincin clan of metalloendopeptidases. *Mol. Biotechnol.* 24, 157–202. doi:10.1385/MB:24:2:157
- Gomis-Rüth, F. X., Trillo-Muyo, S., and Stöcker, W. (2012). Functional and structural insights into astacin metallopeptidases. *Biol. Chem.* 393, 1027–1041. doi:10.1515/hsz-2012-0149
- Gomis-Rüth, F. X. (2013). “Zinc adamalysins,” in *Encyclopedia of metalloproteins*. Editors V. N. Uversky, R. H. Kretsinger&E, and A. Permyakov (Heidelberg (Germany): Springer-Verlag).
- González-Calvo, I., Cizeron, M., Bessereau, J. L., and Selimi, F. (2022). Synapse formation and function across species: Ancient roles for CCP, CUB, and TSP-1 structural domains. *Front. Neurosci.* 16, 866444. doi:10.3389/fnins.2022.866444
- Goulas, T., Arolas, J. L., and Gomis-Rüth, F. X. (2011). Structure, function and latency regulation of a bacterial enterotoxin potentially derived from a mammalian adamalysin/ADAM xenolog. *Proc. Natl. Acad. Sci. U. S. A.* 108 (5), 1856–1861. doi:10.1073/pnas.1012173108
- Goulas, T., García-Ferrer, I., Marrero, A., Marino-Puertas, L., Duquerroy, S., and Gomis-Rüth, F. X. (2017). Structural and functional insight into pan-endopeptidase inhibition by α_2 -macroglobulins. *Biol. Chem.* 398, 975–994. doi:10.1515/hsz-2016-0329
- Goulas, T., and Gomis-Rüth, F. X. (2013). “Fragilysin,” in *Handbook of proteolytic enzymes*. Editors N. D. Rawlings&G and S. Salvesen (Oxford, Great Britain: Academic Press).
- Gouy, M., Guindon, S., and Gascuel, O. (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* 27 (2), 221–224. doi:10.1093/molbev/msp259
- Grams, F., Dive, V., Yiotakis, A., Yiallourous, I., Vassiliou, S., Zwilling, R., et al. (1996). Structure of astacin with a transition-state analogue inhibitor. *Nat. Struct. Biol.* 3, 671–675. doi:10.1038/nsb0896-671
- Grau-Bové, X., Torruella, G., Donachie, S., Suga, H., Leonard, G., Richards, T. A., et al. (2017). Dynamics of genomic innovation in the unicellular ancestry of animals. *eLife* 6, e26036. doi:10.7554/eLife.26036
- Guevara, T., Rodríguez-Banqueri, A., Stöcker, W., Becker-Pauly, C., and Gomis-Rüth, F. X. (2022). Zymogenic latency in a ~250-million-year-old astacin metallopeptidase. *Acta Crystallogr. Sect. D.* 78, 1347–1357. doi:10.1107/S2059798322009688
- Guevara, T., Yiallourous, I., Kappelhoff, R., Bissdorf, S., Stöcker, W., and Gomis-Rüth, F. X. (2010). Proenzyme structure and activation of astacin metallopeptidase. *J. Biol. Chem.* 285 (18), 13958–13965. doi:10.1074/jbc.M109.097436
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* 59 (3), 307–321. doi:10.1093/sysbio/syq010
- Hartigan, A., Kosakyan, A., Peckova, H., Eszterbauer, E., and Holzer, A. S. (2020). Transcriptome of *Sphaerospora molnari* (Cnidaria, Myxosporae) blood stages provides proteolytic arsenal as potential therapeutic targets against sphaerosporosis in common carp. *BMC Genomics* 21 (1), 404. doi:10.1186/s12864-020-6705-y
- Hernández-Elizárraga, V. H., Olguín-López, N., Hernández-Matehuala, R., Ocharán-Mercado, A., Cruz-Hernández, A., Guevara-González, R. G., et al. (2019). Comparative analysis of the soluble proteome and the cytolytic activity of unbleached and bleached *Millepora complanata* (“Fire Coral”) from the Mexican Caribbean. *Mar. Drugs* 17 (7), 393. doi:10.3390/md17070393
- Herrera, C., Escalante, T., Rucavado, A., Fox, J. W., and Gutiérrez, J. M. (2018). Metalloproteinases in disease: Identification of biomarkers of tissue damage through proteomics. *Expert Rev. Proteomics* 15 (12), 967–982. doi:10.1080/14789450.2018.1538800
- Hewitson, J. P., Harcus, Y., Murray, J., van Agtmaal, M., Filbey, K. J., Grainger, J. R., et al. (2011). Proteomic analysis of secretory products from the model gastrointestinal nematode *Heligmosomoides polygyrus* reveals dominance of venom allergen-like (VAL) proteins. *J. Proteomics* 74 (9), 1573–1594. doi:10.1016/j.jprot.2011.06.002
- Hickman, C., Keen, S., Eisenhour, D., Larson, A., and f'Anson, H. (2020). *Integrated principles of zoology*. New York (USA): McGraw-Hill Education.
- Hintze, V., Höwel, M., Wermter, C., Grosse Berkhoff, E., Becker-Pauly, C., Beermann, B., et al. (2006). The interaction of recombinant subdomains of the procollagen C-proteinase with procollagen I provides a quantitative explanation for functional differences between the two splice variants, mammalian tolloid and bone morphogenetic protein 1. *Biochemistry* 45 (21), 6741–6748. doi:10.1021/bi060228k
- Holley, S. A., Neul, J. L., Attisano, L., Wrana, J. L., Sasai, Y., O'Connor, M. B., et al. (1996). The *Xenopus* dorsalizing factor noggin ventralizes *Drosophila* embryos by preventing DPP from activating its receptor. *Cell* 86 (4), 607–617. doi:10.1016/s0092-8674(00)80134-8
- Holstein, T. W. (2022). The role of cnidarian developmental biology in unraveling axis formation and Wnt signaling. *Dev. Biol.* 487, 74–98. doi:10.1016/j.ydbio.2022.04.005
- Hopkins, D. R., Keles, S., and Greenspan, D. S. (2007). The bone morphogenetic protein 1/Tolloid-like metalloproteinases. *Matrix Biol.* 26 (7), 508–523. doi:10.1016/j.matbio.2007.05.004
- Huang, L., Wang, Z., Yu, N., Li, J., and Liu, Z. (2018). Toxin diversity revealed by the venom gland transcriptome of *Pardosa pseudoannulata*, a natural enemy of several insect pests. *Comp. Biochem. Physiol. Part D. Genomics Proteomics* 28, 172–182. doi:10.1016/j.cbd.2018.09.002
- Hunt, V. L., Tsai, I. J., Coghlan, A., Reid, A. J., Holroyd, N., Foth, B. J., et al. (2016). The genomic basis of parasitism in the *Strongyloides* clade of nematodes. *Nat. Genet.* 48 (3), 299–307. doi:10.1038/ng.3495
- Isolani, M. E., Abril, J. F., Salo, E., Deri, P., Bianucci, A. M., and Batistoni, R. (2013). Planarians as a model to assess *in vivo* the role of matrix metalloproteinase genes during homeostasis and regeneration. *PLoS ONE* 8 (2), e55649. doi:10.1371/journal.pone.0055649
- Isolani, M. E., Batistoni, R., Ippolito, C., Bianucci, A. M., Marracci, S., and Rossi, L. (2018). Astacin gene family of metalloproteinases in planarians: Structural organization and tissue distribution. *Gene Expr. Patterns* 28, 77–86. doi:10.1016/j.gep.2018.03.003
- Jiang, W., and Bond, J. S. (1992). Families of metalloendopeptidases and their relationships. *FEBS Lett.* 312 (2-3), 110–114. doi:10.1016/0014-5793(92)80916-5

- Jing, Y., Toubarro, D., Hao, Y., and Simoes, N. (2010). Cloning, characterisation and heterologous expression of an astacin metalloprotease, Sc-AST, from the entomoparasitic nematode *Steinernema carpocapsae*. *Mol. Biochem. Parasitol.* 174 (2), 101–108. doi:10.1016/j.molbiopara.2010.07.004
- Jochim, R. C., Teixeira, C. R., Laughinghouse, A., Mu, J., Oliveira, F., Gomes, R. B., et al. (2008). The midgut transcriptome of *Lutzomyia longipalpis*: Comparative analysis of cDNA libraries from sugar-fed, blood-fed, post-digested and *Leishmania infantum* chagasi-infected sand flies. *BMC Genomics* 9, 15. doi:10.1186/1471-2164-9-15
- Judelson, H. S. (2017). Metabolic diversity and novelties in the oomycetes. *Annu. Rev. Microbiol.* 71, 21–39. doi:10.1146/annurev-micro-090816-093609
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596 (7873), 583–589. doi:10.1038/s41586-021-03819-2
- Kang, C., Han, D. Y., Park, K. I., Pyo, M. J., Heo, Y., Lee, H., et al. (2014). Characterization and neutralization of *Nemopilema nomurai* (Scyphozoa: Rhizostomeae) jellyfish venom using polyclonal antibody. *Toxicon* 86, 116–125. doi:10.1016/j.toxicon.2014.04.005
- Kanzawa, N., Ogawa, T., Asakura, M., Okiyama, K., Honda, M., and Tsuchiya, T. (2008). Comparative expression and tissue distribution analyses of astacin-like squid metalloprotease in squid and cuttlefish. *Zool. Sci.* 25 (1), 14–21. doi:10.2108/zsj.25.14
- Kanzawa, N., Tatewaki, S., Watanabe, R., Kuniyama, I., Iwashita, H., Nakamura, K., et al. (2005). Expression and tissue distribution of astacin-like squid metalloprotease (ALSM). *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 142 (2), 153–163. doi:10.1016/j.cbpc.2005.05.018
- Karmilin, K., Schmitz, C., Kuske, M., Körschgen, H., Olf, M., Meyer, K., et al. (2019). Mammalian plasma fetuin-B is a selective inhibitor of ovastacin and meprin metalloproteinases. *Sci. Rep.* 9, 546. doi:10.1038/s41598-018-37024-5
- Keeling, P. J., and Palmer, J. D. (2008). Horizontal gene transfer in eukaryotic evolution. *Nat. Rev. Genet.* 9 (8), 605–618. doi:10.1038/nrg2386
- Kessler, E., Takahara, K., Biniaminov, L., Brusel, M., and Greenspan, D. S. (1996). Bone morphogenetic protein-1: the type I procollagen C-proteinase. *Science* 271, 360–362. doi:10.1126/science.271.5247.360
- Khamtorn, P., Rungsa, P., Jangpromma, N., Klaynongsruang, S., Daduang, J., Tessiri, T., et al. (2020). Partial proteomic analysis of Brown widow spider (*Latrodectus geometricus*) venom to determine the biological activities. *Toxicon* 8, 100062. doi:10.1016/j.toxcon.2020.100062
- King, N. (2005). Choanoflagellates. *Curr. Biol.* 15 (4), R113–R114. doi:10.1016/j.cub.2005.02.004
- King, N., Westbrook, M. J., Young, S. L., Kuo, A., Abedin, M., Chapman, J., et al. (2008). The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. *Nature* 451 (7180), 783–788. doi:10.1038/nature06617
- Kocot, K. M., Struck, T. H., Merkel, J., Waits, D. S., Todt, C., Brannock, P. M., et al. (2017). Phylogenomics of Lophotrochozoa with consideration of systematic error. *Syst. Biol.* 66 (2), 256–282. doi:10.1093/sysbio/syw079
- Koonin, E. V., Makarova, K. S., and Aravind, L. (2001). Horizontal gene transfer in prokaryotes: Quantification and classification. *Annu. Rev. Microbiol.* 55, 709–742. doi:10.1146/annurev.micro.55.1.709
- Körschgen, H., Kuske, M., Karmilin, K., Yiallourous, I., Balbach, M., Floehr, J., et al. (2017). Intracellular activation of ovastacin mediates pre-fertilization hardening of the zona pellucida. *Mol. Hum. Reprod.* 23 (9), 607–616. doi:10.1093/molehr/gax040
- Kumar, P., Kannan, M., ArunPrasanna, V., Vaseeharan, B., and Vijayakumar, S. (2018). Proteomics analysis of crude squid ink isolated from *Sepia esculenta* for their antimicrobial, antibiofilm and cytotoxic properties. *Microb. Pathog.* 116, 345–350. doi:10.1016/j.micpath.2018.01.039
- LaFlamme, B. A., Avila, F. W., Michalski, K., and Wolfner, M. F. (2014). A *Drosophila* protease cascade member, seminal metalloprotease-1, is activated stepwise by male factors and requires female factors for full activity. *Genetics* 196 (4), 1117–1129. doi:10.1534/genetics.113.160101
- LaFlamme, B. A., Ram, K. R., and Wolfner, M. F. (2012). The *Drosophila melanogaster* seminal fluid protease "seminase" regulates proteolytic and post-mating reproductive processes. *PLoS Genet.* 8 (1), e1002435. doi:10.1371/journal.pgen.1002435
- LaFlamme, B. A., and Wolfner, M. F. (2013). Identification and function of proteolysis regulators in seminal fluid. *Mol. Reprod. Dev.* 80 (2), 80–101. doi:10.1002/mrd.22130
- Laumer, C. E., Fernández, R., Lemer, S., Combosch, D., Kocot, K. M., Riesgo, A., et al. (2019). Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proc. Biol. Sci.* 286, 20190831. doi:10.1098/rspb.2019.0831
- Lepage, T., Ghiglione, C., and Gache, C. (1992). Spatial and temporal expression pattern during sea urchin embryogenesis of a gene coding for a protease homologous to the human protein BMP-1 and to the product of the *Drosophila* dorsal-ventral patterning gene tolloid. *Development* 114 (1), 147–163. doi:10.1242/dev.114.1.147
- Linnæus, C. (1758). *Systema nature per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis synonymis, locis*. Stockholm, Sweden: Laurentius Salvius, Holmiae.
- Liu, G., Zhou, Y., Liu, D., Wang, Q., Ruan, Z., He, Q., et al. (2015). Global transcriptome analysis of the tentacle of the jellyfish *Cyanea capillata* using deep sequencing and expressed sequence tags: Insight into the toxin- and degenerative disease-related transcripts. *PLoS One* 10 (11), e0142680. doi:10.1371/journal.pone.0142680
- Liu, Q. R., Hattar, S., Endo, S., MacPhee, K., Zhang, H., Cleary, L. J., et al. (1997). A developmental gene (Tolloid/BMP-1) is regulated in *Aplysia* neurons by treatments that induce long-term sensitization. *J. Neurosci.* 17 (2), 755–764. doi:10.1523/JNEUROSCI.17-02-00755.1997
- Lu, F. H., Tang, S. M., Shen, X. J., Wang, N., Zhao, Q. L., Zhang, G. Z., et al. (2010). Molecular cloning and characterization of hatching enzyme-like gene in the silkworm, *Bombyx mori*. *Mol. Biol. Rep.* 37 (3), 1175–1182. doi:10.1007/s11033-009-9483-9
- Lu, T.-M., Kanda, M., Satoh, N., and Furuya, H. (2017). The phylogenetic position of dicyemid mesozoans offers insights into spiralian evolution. *Zool. Lett.* 3, 6. doi:10.1186/s40851-017-0068-5
- Lun, H. M., Mak, C. H., and Ko, R. C. (2003). Characterization and cloning of metallo-proteinase in the excretory/secretory products of the infective-stage larva of *Trichinella spiralis*. *Parasitol. Res.* 90 (1), 27–37. doi:10.1007/s00436-002-0815-0
- Luque, D., Goulas, T., Mata, C. P., Mendes, S. R., Gomis-Rüth, F. X., and Castón, J. R. (2022). Cryo-EM structures show the mechanistic basis of pan-peptidase inhibition by human α_2 -macroglobulin. *Proc. Natl. Acad. Sci. U. S. A.* 119 (19), e2200102119. doi:10.1073/pnas.2200102119
- Marlétaz, F., Peijnenburg, K. T. C. A., Goto, T., Satoh, N., and Rokhsar, D. S. (2019). A new spiralian phylogeny places the enigmatic arrow worms among Gnathiferans. *Curr. Biol.* 29 (2), 312e3–318. doi:10.1016/j.cub.2018.11.042
- Marrero, A., Duquerroy, S., Trapani, S., Goulas, T., Guevara, T., Andersen, G. R., et al. (2012). The crystal structure of human α_2 -macroglobulin reveals a unique molecular cage. *Angew. Chem. Int. Ed. Engl.* 51, 3340–3344. doi:10.1002/anie.201108015
- Martin-Galiano, A. J., and Sotillo, J. (2022). Insights into the functional expansion of the astacin peptidase family in parasitic helminths. *Int. J. Parasitol.* 52 (4), 243–251. doi:10.1016/j.ijpara.2021.09.001
- Mashanov, V. S., Zueva, O. R., and García-Arriaras, J. E. (2012). Expression of Wnt9, TCTP, and Bmp1/Tll in sea cucumber visceral regeneration. *Gene Expr. Patterns* 12 (1–2), 24–35. doi:10.1016/j.gep.2011.10.003
- Medina-Santos, R., Guerra-Duarte, C., de Almeida Lima, S., Costal-Oliveira, F., Alves de Aquino, P., Oliveira do Carmo, A., et al. (2019). Diversity of astacin-like metalloproteases identified by transcriptomic analysis in Peruvian *Loxosceles laeta* spider venom and *in vitro* activity characterization. *Biochimie* 167, 81–92. doi:10.1016/j.biochi.2019.08.017
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., et al. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Res.* 49 (D1), D412–D419. doi:10.1093/nar/gkaa913
- Möhrle, F., Hutter, H., and Zwilling, R. (2003). The astacin protein family in *Caenorhabditis elegans*. *Eur. J. Biochem.* 270 (24), 49093891–49094920. doi:10.1046/j.1432-1033.2003.03891.x
- Möhrle, F., Maniura, M., Plickert, G., Frohne, M., and Frank, U. (2006). Evolution of astacin-like metalloproteases in animals and their function in development. *Evol. Dev.* 8 (2), 223–231. doi:10.1111/j.1525-142X.2006.00092.x
- Moran, Y., Praher, D., Schlesinger, A., Ayalon, A., Tal, Y., and Technau, U. (2013). Analysis of soluble protein contents from the nematocysts of a model sea anemone sheds light on venom evolution. *Mar. Biotechnol.* 15 (3), 329–339. doi:10.1007/s10126-012-9491-y
- Nagasawa, T., Kawaguchi, M., Nishi, K., and Yasumasu, S. (2022). Molecular evolution of hatching enzymes and their paralogous genes in vertebrates. *BMC Ecol. Evol.* 22 (1), 9. doi:10.1186/s12862-022-01966-2
- Nguyen, T., Jamal, J., Shimell, M. J., Arora, K., and O'Connor, M. B. (1994). Characterization of tolloid-related-1: A BMP-1-like product that is required during larval and pupal stages of *Drosophila* development. *Dev. Biol.* 166 (2), 569–586. doi:10.1006/dbio.1994.1338
- Pan, T.-L., Gröger, H., Schmid, V., and Spring, J. (1998). A toxin homology domain in an astacin-like metalloproteinase of the jellyfish *Podocoryne carnea* with a dual role in digestion and development. *Dev. Genes Evol.* 208, 259–266. doi:10.1007/s004270050180
- Pang, K., Ryan, J. F., Baxevanis, A. D., and Martindale, M. Q. (2011). Evolution of the TGF- β signaling pathway and its potential role in the ctenophore, *Mnemiopsis leidyi*. *PLoS One* 6 (9), e24152. doi:10.1371/journal.pone.0024152

- Park, H. H. (2018). Structure of TRAF family: Current understanding of receptor recognition. *Front. Immunol.* 9, 1999. doi:10.3389/fimmu.2018.01999
- Park, J.-O., Pan, J., Möhrlein, F., Schupp, M.-O., Johnsen, R., Baillie, D. L., et al. (2010). Characterization of the astacin family of metalloproteases in *C. elegans*. *BMC Dev. Biol.* 10, 14. doi:10.1186/1471-213X-10-14
- Pfleiderer, G., Zwiling, R., and Sonneborn, H. H. (1967). On the evolution of endopeptidases. 3. A protease of molecular weight 11, 000 and a trypsin-like fraction from *Astacus fluviatilis* Fabr. *Hoppe. Seylers. Z. Physiol. Chem.* 348 (10), 1319–1331.
- Prodocimi, F., Bittencourt, D., da Silva, F. R., Kirst, M., Motta, P. C., and Rech, E. L. (2011). Spinning gland transcriptomes from two main clades of spiders (order: Araneae)—insights on their molecular, anatomical and behavioral evolution. *PLoS One* 6 (6), e21634. doi:10.1371/journal.pone.0021634
- Quesada, V., Ordoñez, G. R., Sánchez, L. M., Puente, X. S., and López-Otin, C. (2009). The degradome database: Mammalian proteases and diseases of proteolysis. *Nucleic Acids Res.* 37, D239–D243. doi:10.1093/nar/gkn570
- Ran, T., Li, W., Sun, B., Xu, M., Qiu, S., Xu, D. Q., et al. (2020). Crystal structure of mature myrolysin and implication for its activation mechanism. *Int. J. Biol. Macromol.* 156, 1556–1564. doi:10.1016/j.ijbiomac.2019.11.205
- Ranjit, N., Jones, M. K., Stenzel, D. J., Gasser, R. B., and Loukas, A. (2006). A survey of the intestinal transcriptomes of the hookworms, *Necator americanus* and *Ancylostoma caninum*, using tissues isolated by laser microdissection microscopy. *Int. J. Parasitol.* 36 (6), 701–710. doi:10.1016/j.ijpara.2006.01.015
- Rawlings, N. D., and Bateman, A. (2021). How to use the MEROPS database and website to help understand peptidase specificity. *Protein Sci.* 30 (1), 83–92. doi:10.1002/pro.3948
- Reynolds, S. D., Angerer, L. M., Palis, J., Nasir, A., and Angerer, R. C. (1992). Early mRNAs, spatially restricted along the animal-vegetal axis of sea urchin embryos, include one encoding a protein related to tolloid and BMP-1. *Development* 114 (3), 769–786. doi:10.1242/dev.114.3.769
- Riehle, M. A., Garczynski, S. F., Crim, J. W., Hill, C. A., and Brown, M. R. (2002). Neuropeptides and peptide hormones in *Anopheles gambiae*. *Science* 298 (5591), 172–175. doi:10.1126/science.1076827
- Roberts, S., Goetz, G., White, S., and Goetz, F. (2009). Analysis of genes isolated from plated hemocytes of the Pacific oyster, *Crassostrea gigas*. *Mar. Biotechnol.* 11 (1), 24–44. doi:10.1007/s10126-008-9117-6
- Rosenblum, G., Meroueh, S., Toth, M., Fisher, J. F., Fridman, R., Mobashery, S., et al. (2007). Molecular structures and dynamics of the stepwise activation mechanism of a matrix metalloproteinase zymogen: Challenging the cysteine switch dogma. *J. Am. Chem. Soc.* 129 (44), 13566–13574. doi:10.1021/ja0739411
- Ruggiero, M. A., Gordon, D. P., Orrell, T. M., Bailly, N., Bourgoin, T., Brusca, R. C., et al. (2015). A higher level classification of all living organisms. *PLoS One* 10 (4), e0119248. doi:10.1371/journal.pone.0119248
- Ryan, J. F., Pang, K., Program, N. C. S., Mullikin, J. C., Martindale, M. Q., and Baxevanis, A. D. (2010). The homeodomain complement of the ctenophore *Mnemiopsis leidyi* suggests that Ctenophora and Porifera diverged prior to the ParaHoxozoa. *EvoDevo* 1 (1), 9. doi:10.1186/2041-9139-1-9
- Saina, M., Genikhovich, G., Renfer, E., and Technau, U. (2009). BMPs and chordin regulate patterning of the directive axis in a sea anemone. *Proc. Natl. Acad. Sci. U. S. A.* 106 (44), 18592–18597. doi:10.1073/pnas.0900151106
- Santos, A., van Aarle, R., Barrientos, L., and Martínez-Urtaza, J. (2020). Computational methods for 16S metabarcoding studies using Nanopore sequencing data. *Comput. Struct. Biotechnol. J.* 18, 296–305. doi:10.1016/j.csbj.2020.01.005
- Sarras, M. P., Jr., Yan, L., Leontovich, A., and Zhang, J. S. (2002). Structure, expression, and developmental function of early divergent forms of metalloproteinases in hydra. *Cell Res.* 12 (3–4), 163–176. doi:10.1038/sj.cr.7290123
- Sato, S. M., and Sargent, T. D. (1990). Molecular approach to dorsoanterior development in *Xenopus laevis*. *Dev. Biol.* 137, 135–141. doi:10.1016/0012-1606(90)90014-a
- Schechter, I., and Berger, A. (1967). On the size of the active site in proteases. I. Papain. 1967. *Biochem. Biophys. Res. Commun.* 27, 497–502. doi:10.1016/j.bbrc.2012.08.015
- Schoch, C. L., Ciuffo, S., Domrachev, M., Hotton, C. L., Kannan, S., Khovanskaya, R., et al. (2020). NCBI taxonomy: A comprehensive update on curation, resources and tools. *Database* 2020, baaa062–21. doi:10.1093/database/baaa062
- Schulze, A. J., and Kawauchi, G. Y. (2021). How many sipunculan species are hiding in our oceans? *Diversity* 13 (2), 43. doi:10.3390/d13020043
- Schwerin, S., Zeis, B., Lamkemeyer, T., Paul, R. J., Koch, M., Madlung, J., et al. (2009). Acclimatory responses of the *Daphnia pulex* proteome to environmental changes. II. Chronic exposure to different temperatures (10 and 20°C) mainly affects protein metabolism. *BMC Physiol.* 9, 8. doi:10.1186/1472-6793-9-8
- Sebé-Pedrós, A., Degnan, B. M., and Ruiz-Trillo, I. (2017). The origin of Metazoa: A unicellular perspective. *Nat. Rev. Genet.* 18 (8), 498–512. doi:10.1038/nrg.2017.21
- Semenova, S. A., Rudenskaya, G. N., Rebrikov, D. V., and Isaev, V. A. (2006). cDNA cloning, purification and properties of *Paralithodes camtschatica* metalloprotease. *Protein Pept. Lett.* 13 (6), 571–575. doi:10.2174/092986606777145887
- Shibata, Y., Iwamatsu, T., Oba, Y., Kobayashi, D., Tanaka, M., Nagahama, Y., et al. (2000). Identification and cDNA cloning of alveolin, an extracellular metalloproteinase, which induces chorion hardening of medaka (*Oryzias latipes*) eggs upon fertilization. *J. Biol. Chem.* 275 (12), 8349–8354. doi:10.1074/jbc.275.12.8349
- Shimell, M. J., Ferguson, E. L., Childs, S. R., and O'Connor, M. B. (1991). The *Drosophila* dorsal-ventral patterning gene *tolloid* is related to human bone morphogenetic protein 1. *Cell* 67, 469–481. doi:10.1016/0092-8674(91)90522-Z
- Sieron, A. L., Tretiakova, A., Jameson, B. A., Segall, M. L., Lund-Katz, S., Khan, M. T., et al. (2000). Structure and function of procollagen C-proteinase (mTolloid) domains determined by protease digestion, circular dichroism, binding to procollagen type I, and computer modeling. *Biochemistry* 39 (12), 3231–3239. doi:10.1021/bi992312o
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539. doi:10.1038/msb.2011.75
- Soblik, H., Younis, A. E., Mitreva, M., Renard, B. Y., Kirchner, M., Geisinger, F., et al. (2011). Life cycle stage-resolved proteomic analysis of the excretome/secretome from *Strongyloides ratti*—identification of stage-specific proteases. *Mol. Cell. Proteomics* 10 (12), M111.010157010157. doi:10.1074/mcp.M111.010157
- Sogabe, S., Hatleberg, W. L., Kocot, K. M., Say, T. E., Stoupin, D., Roper, K. E., et al. (2019). Pluripotency and the origin of animal multicellularity. *Nature* 570 (7762), 519–522. doi:10.1038/s41586-019-1290-4
- Sonenshine, D. E., Bissinger, B. W., Egekwo, N., Donohue, K. V., Khalil, S. M., and Roe, R. M. (2011). First transcriptome of the testis-vas deferens-male accessory gland and proteome of the spermatophore from *Dermacentor variabilis* (Acari: Ixodidae). *PLoS One* 6 (9), e24711. doi:10.1371/journal.pone.0024711
- Sotillo, J., Sánchez-Flores, A., Cantacessi, C., Harcus, Y., Pickering, D., Bouchery, T., et al. (2014). Secreted proteomes of different developmental stages of the gastrointestinal nematode *Nippostrongylus brasiliensis*. *Mol. Cell. Proteomics* 13 (10), 2736–2751. doi:10.1074/mcp.M114.038950
- Springman, E. B., Angleton, E. L., Birkedal-Hansen, H., and Van Wart, H. E. (1990). Multiple modes of activation of latent human fibroblast collagenase: Evidence for the role of a Cys73 active-site zinc complex in latency and a "cysteine switch" mechanism for activation. *Proc. Natl. Acad. Sci. U. S. A.* 87, 364–368. doi:10.1073/pnas.87.1.364
- Stepek, G., McCormack, G., Birnie, A. J., and Page, A. P. (2011). The astacin metalloprotease moulting enzyme NAS-36 is required for normal cuticle ecdysis in free-living and parasitic nematodes. *Parasitology* 138 (2), 237–248. doi:10.1017/S0031182010001113
- Stepek, G., McCormack, G., and Page, A. P. (2010). Collagen processing and cuticle formation is catalysed by the astacin metalloprotease DPY-31 in free-living and parasitic nematodes. *Int. J. Parasitol.* 40 (5), 533–542. doi:10.1016/j.ijpara.2009.10.007
- Stepek, G., McCormack, G., Winter, A. D., and Page, A. P. (2015). A highly conserved, inhibitable astacin metalloprotease from *Teladorsagia circumcincta* is required for cuticle formation and nematode development. *Int. J. Parasitol.* 45 (5), 345–355. doi:10.1016/j.ijpara.2015.01.004
- Sterchi, E. E. (2008). Special issue: Metzincin metalloproteinases. *Mol. Asp. Med.* 29 (5), 255–257. doi:10.1016/j.mam.2008.08.007
- Sterchi, E. E., Stöcker, W., and Bond, J. S. (2008). Meprens, membrane-bound and secreted astacin metalloproteinases. *Mol. Asp. Med.* 29, 309–328. doi:10.1016/j.mam.2008.08.002
- Stöcker, W., and Bode, W. (1995). Structural features of a superfamily of zinc-endopeptidases: The metzincins. *Curr. Opin. Struct. Biol.* 5 (3), 383–390. doi:10.1016/0959-440x(95)80101-4
- Stöcker, W., and Gomis-Rüth, F. X. (2013). "Astacins: Proteases in development and tissue differentiation," in *Proteases: Structure and function*. Editor K. Brix&W. Stöcker (Vienna, Austria: Springer-Verlag). doi:10.1007/978-3-7091-0885-7
- Stöcker, W., Gomis-Rüth, F. X., Bode, W., and Zwilling, R. (1993). Implications of the three-dimensional structure of astacin for the structure and function of the astacin-family of zinc-endopeptidases. *Eur. J. Biochem.* 214, 215–231. doi:10.1111/j.1432-1033.1993.tb17915.x

- Stöcker, W., Karmilin, K., Hildebrand, A., Westphal, H., Yiallourous, I., Weiskirchen, R., et al. (2014). Mammalian gamete fusion depends on the inhibition of ovastacin by fetuin-B. *Biol. Chem.* 395 (10), 1195–1199. doi:10.1515/hsz-2014-0189
- Stöcker, W., Ng, M., and Auld, D. S. (1990). Fluorescent oligopeptide substrates for kinetic characterization of the specificity of *Astacus* protease. *Biochemistry* 29, 10418–10425. doi:10.1021/bi00497a018
- Stöcker, W., and Yiallourous, I. (2013). "Chapter 188 - astacin," in *Handbook of proteolytic enzymes*. Editors N. D. Rawlings & G. and S. Salvesen (Oxford: Academic Press). doi:10.1016/B978-0-12-382219-2.00188-5
- Tallant, C., García-Castellanos, R., Baumann, U., and Gomis-Rüth, F. X. (2010). On the relevance of the Met-turn methionine in metzincins. *J. Biol. Chem.* 285, 13951–13957. doi:10.1074/jbc.M109.083378
- Tang, S. M., Ding, L. P., Shen, X. J., Wu, J., Lu, F. H., Zhou, Q. M., et al. (2011). Full-length cDNA cloning and expression characteristic analysis of hatching enzyme-like gene in the Chinese oak silkworm. *Antheraea Pernyi. Sci. Seric.* 27, 1000–1007.
- Tang, S. M., Zhao, X. H., Wu, J., Qiu, Z. Y., and Shen, X. J. (2011). Cloning and bioinformatics analysis of hatching enzyme gene in Chinese wild silkworm (*Bombyx mandarina*). *Acta agri. Jiangxi* 23, 1–5.
- Tarentino, A. L., Quinones, G., Grimwood, B. G., Hauer, C. R., and Plummer, T. H., Jr. (1995). Molecular cloning and sequence analysis of flavastacin: An O-glycosylated prokaryotic zinc metalloendopeptidase. *Arch. Biochem. Biophys.* 319 (1), 281–285. doi:10.1006/abbi.1995.1293
- Toprak, U., Erlanson, M., Baldwin, D., Karcz, S., Wan, L., Coutu, C., et al. (2016). Identification of the *Mamestra configurata* (Lepidoptera: Noctuidae) peritrophic matrix proteins and enzymes involved in peritrophic matrix chitin metabolism. *Insect Sci.* 23 (5), 656–674. doi:10.1111/1744-7917.12225
- Torruella, G., de Mendoza, A., Grau-Bové, X., Antó, M., Chaplin, M. A., del Campo, J., et al. (2015). Phylogenomics reveals convergent evolution of lifestyles in close relatives of animals and fungi. *Curr. Biol.* 25 (18), 2404–2410. doi:10.1016/j.cub.2015.07.053
- Trevisan-Silva, D., Gremski, L. H., Chaim, O. M., da Silveira, R. B., Meissner, G. O., Mangili, O. C., et al. (2010). Astacin-like metalloproteases are a gene family of toxins present in the venom of different species of the Brown spider (genus *Loxosceles*). *Biochimie* 92 (1), 21–32. doi:10.1016/j.biochi.2009.10.003
- Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Zidek, A., et al. (2021). Highly accurate protein structure prediction for the human proteome. *Nature* 596 (7873), 590–596. doi:10.1038/s41586-021-03828-1
- Varatharajulu, R., Parandaman, V., Ndao, M., Andersen, J. F., and Neva, F. A. (2011). *Strongyloides stercoralis* excretory/secretory protein strongylastacin specifically recognized by IgE antibodies in infected human sera. *Microbiol. Immunol.* 55 (2), 115–122. doi:10.1111/j.1348-0421.2010.00289.x
- Wagner, W., Möhrlen, F., and Schnetter, W. (2002). Characterization of the proteolytic enzymes in the midgut of the European Cockchafer, *Melolontha melolontha* (Coleoptera: Scarabaeidae). *Insect biochem. Mol. Biol.* 32 (7), 803–814. doi:10.1016/s0965-1748(01)00167-9
- Walter, A., Bechsgaard, J., Scavenius, C., Dyrland, T. S., Sanggaard, K. W., Enghild, J. J., et al. (2017). Characterisation of protein families in spider digestive fluids and their role in extra-oral digestion. *BMC Genomics* 18 (1), 600. doi:10.1186/s12864-017-3987-9
- Wermter, C., Höwel, M., Hintze, V., Bombosch, B., Aufenvenne, K., Yiallourous, I., et al. (2007). The protease domain of procollagen C-proteinase (BMP1) lacks substrate selectivity, which is conferred by non-proteolytic domains. *Biol. Chem.* 388 (5), 513–521. doi:10.1515/BC.2007.054
- Werny, L., Colmorgen, C., and Becker-Pauly, C. (2022). Regulation of meprin metalloproteases in mucosal homeostasis. *Biochim. Biophys. Acta. Mol. Cell Res.* 1869 (1), 119158. doi:10.1016/j.bbamcr.2021.119158
- Whelan, N. V., Kocot, K. M., Moroz, T. P., Mukherjee, K., Williams, P., Paulay, G., et al. (2017). Ctenophore relationships and their placement as the sister group to all other animals. *Nat. Ecol. Evol.* 1 (11), 1737–1746. doi:10.1038/s41559-017-0331-3
- Williamson, A. L., Lustigman, S., Oksov, Y., Deumic, V., Plieskatt, J., Mendez, S., et al. (2006). *Ancylostoma caninum* MTP-1, an astacin-like metalloprotease secreted by infective hookworm larvae, is involved in tissue migration. *Infect. Immun.* 74 (2), 961–967. doi:10.1128/IAI.74.2.961-967.2006
- Wong, S. G., and Dessen, A. (2014). Structure of a bacterial α_2 -macroglobulin reveals mimicry of eukaryotic innate immunity. *Nat. Commun.* 5, 4917. doi:10.1038/ncomms5917
- Wouters, M. A., Rigoutsos, I., Chu, C. K., Feng, L. L., Sparrow, D. B., and Dunwoodie, S. L. (2005). Evolution of distinct EGF domains with specific functions. *Protein Sci.* 14 (4), 1091–1103. doi:10.1110/ps.041207005
- Wozney, J. M., Rosen, V., Celeste, A. J., Mitscock, L. M., Whitters, M. J., Kriz, R. W., et al. (1988). Novel regulators of bone formation: Molecular clones and activities. *Science* 242, 1528–1534. doi:10.1126/science.3201241
- Xiong, X., Chen, L., Li, Y., Xie, L., and Zhang, R. (2006). Pf-ALMP, a novel astacin-like metalloproteinase with cysteine arrays, is abundant in hemocytes of pearl oyster *Pinctada fucata*. *Biochim. Biophys. Acta* 1759 (11–12), 526–534. doi:10.1016/j.bbaxp.2006.09.006
- Xu, D., Zhou, J., Lou, X., He, J., Ran, T., and Wang, W. (2017). Myroilysin Is a new bacterial member of the M12A family of metzincin metalloproteinases and is activated by a cysteine switch mechanism. *J. Biol. Chem.* 292 (13), 5195–5206. doi:10.1074/jbc.M116.758110
- Yan, J., Cheng, Q., Li, C. B., and Aksoy, S. (2002). Molecular characterization of three gut genes from *Glossina morsitans morsitans*: Cathepsin B, zinc-metalloprotease and zinc-carboxypeptidase. *Insect Mol. Biol.* 11 (1), 57–65. doi:10.1046/j.0962-1075.2001.00308.x
- Yan, L., Fei, K., Zhang, J., Dexter, S., and Sarras, M. P., Jr. (2000). Identification and characterization of hydra metalloproteinase 2 (HMP2): A meprin-like astacin metalloproteinase that functions in foot morphogenesis. *Development* 127 (1), 129–141. doi:10.1242/dev.127.1.129
- Yan, L., Pollock, G. H., Nagase, H., and Sarras, M. P., Jr. (1995). A 25.7 x 10³ Mr hydra metalloproteinase (HMP1), a member of the astacin family, localizes to the extracellular matrix of *Hydra vulgaris* in a head-specific manner and has a developmental function. *Development* 121, 1591–1602. doi:10.1242/dev.121.6.1591
- Yang, S., and Wu, X. (2009). Tollid-like gene in *Crassostrea ariakensis*: Molecular cloning, structural characterization and expression by RLO stimulation. *Fish. Shellfish Immunol.* 27 (2), 130–135. doi:10.1016/j.fsi.2008.11.020
- Yelland, T., and Djordjevic, S. (2016). Crystal structure of the neuropilin-1 MAM domain: Completing the neuropilin-1 ectodomain picture. *Structure* 24 (11), 2008–2015. doi:10.1016/j.str.2016.08.017
- Yiallourous, I., Grosse-Berkhoff, E., and Stöcker, W. (2000). The roles of Glu93 and Tyr149 in astacin-like zinc peptidases. *FEBS Lett.* 484, 224–228. doi:10.1016/S0014-5793(00)02163-3
- Yokozawa, Y., Tamai, H., Tatewaki, S., Tajima, T., Tsuchiya, T., and Kanzawa, N. (2002). Cloning and biochemical characterization of astacin-like squid metalloprotease. *J. Biochem.* 132 (5), 751–758. doi:10.1093/oxfordjournals.jbchem.a003283
- Yoshida, A., Nagayasu, E., Nishimaki, A., Sawaguchi, A., Yanagawa, S., and Maruyama, H. (2011). Transcripts analysis of infective larvae of an intestinal nematode, *Strongyloides venezuelensis*. *Parasitol. Int.* 60 (1), 75–83. doi:10.1016/j.parint.2010.10.007
- Young, A. R., Mancuso, N., Meeusen, E. N. T., and Bowles, V. M. (2000). Characterisation of proteases involved in egg hatching of the sheep blowfly, *Lucilia cuprina*. *Int. J. Parasitol.* 30 (8), 925–932. doi:10.1016/s0020-7519(00)00073-4
- Yu, D., Ji, C., Zhao, J., and Wu, H. (2016). Proteomic and metabolomic analysis on the toxicological effects of as (III) and as (V) in juvenile mussel *Mytilus galloprovincialis*. *Chemosphere* 150, 194–201. doi:10.1016/j.chemosphere.2016.01.113
- Zapata, J. M., Pawlowski, K., Haas, E., Ware, C. F., Godzik, A., and Reed, J. C. (2001). A diverse family of proteins containing tumor necrosis factor receptor-associated factor domains. *J. Biol. Chem.* 276 (26), 24242–24252. doi:10.1074/jbc.M100354200
- Ziegler, B., Yiallourous, I., Trageser, B., Kumar, S., Mercker, M., Kling, S., et al. (2021). The Wnt-specific astacin proteinase HAS-7 restricts head organizer formation in *Hydra*. *BMC Biol.* 19 (1), 120. doi:10.1186/s12915-021-01046-9
- Zobel-Thropp, P. A., Correa, S. M., Garb, J. E., and Binford, G. J. (2014). Spit and venom from scytodes spiders: A diverse and distinct cocktail. *J. Proteome Res.* 13 (2), 817–835. doi:10.1021/pr400875s
- Zobel-Thropp, P. A., Mullins, J., Kristensen, C., Kronmiller, B. A., David, C. L., Brecci, L. A., et al. (2019). Not so dangerous after all? Venom composition and potency of the pholcid (daddy long-leg) spider *Physocyclus mexicanus*. *Front. Ecol. Evol.* 7, 256. doi:10.3389/fevo.2019.00256
- Zobel-Thropp, P. A., Thomas, E. Z., David, C. L., Brecci, L. A., and Binford, G. J. (2014). *Plectreurys tristis* venom: A proteomic and transcriptomic analysis. *J. Venom. Res.* 5, 33–47.
- Zwilling, R., and Stöcker, W. (1997). "The astacins — structure and function of a new protein family," in *Naturwissenschaftliche Forschungsergebnisse* (Hamburg (Germany): Verlag Dr. Kovač).